

**НЕЛИНЕЙНЫЕ
ВОЛНЫ' 2002**



Российская академия наук
Институт прикладной физики

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2002

Нижний Новгород
2003

Издано по решению редакционно-издательского совета
Института прикладной физики РАН

УДК 534.222.2

ББК 22.312я4

Н49

Нелинейные волны' 2002 / Отв. ред. А. В. Гапонов-Грехов,
В. И. Некоркин. — Нижний Новгород: ИПФ РАН, 2003. — 448 с.

ISBN 5-8048-0036-1

В сборник включены статьи, написанные по материалам лекций, прочитанных на XI школе по нелинейным волнам (Нижний Новгород, март 2002 г.). Обсуждаются проблемы коллективной динамики сложных систем, обработки, преобразования и передачи информации, волновой динамики, нелинейной оптики, а также нелинейные явления в плазме и др.

Книга рассчитана на специалистов, занимающихся исследованием нелинейных явлений, а также аспирантов и студентов соответствующих специальностей.

Ответственные редакторы

*академик А. В. ГАПОНОВ-ГРЕХОВ,
доктор физико-математических наук В. И. НЕКОРКИН*

Рецензенты

чл.-корр. РАН М. И. Рабинович, д. ф.-м. н. В. Г. Гавриленко

*Проведение школы и издание сборника осуществлены при поддержке
Министерства промышленности, науки и технологий Российской Федерации,
Российской академии наук,
Российского фонда фундаментальных исследований,
Минатома Российской Федерации,
Администрации Нижегородской области.*

ISBN 5-8048-0036-1

© Институт прикладной физики РАН, 2003 г.

СОДЕРЖАНИЕ

<i>Предисловие</i>	5
--------------------------	---

Информационные аспекты нелинейной динамики

<i>В. Б. Казанцев, В. И. Некоркин.</i> Динамика колебательных нейронов. Информационные аспекты	9
<i>Я. И. Мольков, А. М. Фейгин.</i> Прогноз качественного поведения динамической системы по хаотическому временному ряду	34
<i>А. С. Дмитриев.</i> Динамический хаос и информация	53
<i>В. Д. Шалфеев, В. В. Матросов.</i> Хаотически модулированные колебания в связанных фазовых системах	77
<i>В. Г. Яхно.</i> Модели нейрноподобных систем. Динамические режимы преобразования информации	90

Общие вопросы нелинейной динамики

<i>Г. С. Голицын.</i> Белый шум как основа объяснения многих статистических закономерностей в природе	117
<i>А. М. Фридман, А. В. Клименко.</i> Две компоненты сейсмической активности Земли и их связь с особенностями суточного вращения	133
<i>Е. А. Кузнецов.</i> Коллапс в гидродинамике	156
<i>Ю. И. Троицкая.</i> Нелинейные волны и структуры в стратифицированных сдвиговых потоках	179
<i>А. Б. Езерский.</i> Связанные состояния и пространственно-временной хаос топологических дефектов.....	189
<i>Е. Н. Пелиновский.</i> Нелинейные модели генерации цунами движущимися источниками	199
<i>Е. И. Якубович.</i> Новый подход к гидродинамике несжимаемой жидкости	211
<i>С. Флах, А. Е. Мирошниченко.</i> Дискретные бризеры	223

Нелинейные явления в плазме и оптике

<i>М. Ю. Рябикин, А. М. Сергеев.</i> Динамика атома в сверхсильном поле и генерация аттосекундных импульсов	235
<i>М. Д. Токман, А. Г. Литвак, М. А. Ерухимова, А. Ю. Крячко.</i> Стимулированное излучение без инверсии и электромагнитно-индуцированная прозрачность в плазме и электронных потоках	250
<i>Г. М. Фрайман, А. А. Балакин.</i> Столкновения электронов с ионами в сильном лазерном поле	268
<i>В. Ю. Трахтенгерц, Д. И. Иудин, А. Н. Григорьев.</i> О фрактальной динамике активных сред	287
<i>Л. С. Мхитарьян.</i> Нелинейная стадия развития возмущений при ускорении оболочек	303
<i>Г. Г. Кочемасов.</i> Некоторые нелинейные задачи лазерного термоядерного синтеза	314
<i>А. С. Кингсен.</i> Нелинейные волны в электронной магнитной гидродинамике	329

Нелинейные явления в медицине и биологии

<i>В. Л. Воейков.</i> Нелинейные колебания на молекулярном и клеточном уровнях в биологических системах: возможные источники происхождения и функциональная роль в процессах жизнедеятельности	345
<i>Ю. М. Романовский, А. В. Нетребко, О. А. Чичигина.</i> Проблема затухания субглобулярных колебаний белковых молекул в воде	359
<i>Ю. И. Неймарк.</i> Математическая модель иммунного ответа организма на вторжение инфекции	372

Математические проблемы нелинейной динамики

<i>Н. Г. Макаренко.</i> Фракталы, мультифрактальные меры и аттракторы	381
<i>А. П. Кузнецов, С. П. Кузнецов, И. Р. Сатаев.</i> Критическое поведение при переходе к хаосу через удвоения периода. Модельные отображения и ренормгрупповой анализ	395
<i>А. И. Саичев, С. Г. Уткин.</i> Законы аномальной диффузии	416
<i>С. Н. Гурбатов, А. И. Саичев.</i> Простейшие модели потоков частиц	432

Предисловие

Одно из ведущих направлений современной физико-математической науки составляют исследования в области нелинейной динамики, нелинейных волн и нелинейного поведения сложных пространственно-временных систем. И на этом направлении отечественная наука традиционно занимала и занимает, несомненно, достойное положение. Прогресс в понимании таких фундаментальных нелинейных явлений, как автоколебания, автоволны, синхронизация, солитоны, когерентные структуры, хаос, турбулентность – имеет ключевое значение не только для развития физики, химии, биологии и других наук, но и для развития прикладных исследований и создания новых технологий.

Школа «Нелинейные волны-2002» стала одиннадцатой по счету. Предыдущие десять школ (1972–1995), на которых активно обсуждались фундаментальные проблемы нелинейной динамики, как сейчас очевидно, послужили одним из реальных факторов, способствовавших поддержанию позиций отечественной науки не только в области нелинейной динамики, но и в ряде современных направлений физики и математики, связанных с нелинейной динамикой.

За последние пять – десять лет и у нас в стране, и за рубежом основные направления в нелинейной динамике были сконцентрированы на изучении коллективной динамики сложных систем и, в особенности, на выяснении возможностей и путей управления пространственно-временной динамикой вплоть до подавления хаоса и образования структур, что имеет важное значение для ряда задач теории неравновесных сред, физики плазмы, гидро- и аэродинамики, нейродинамики, задач обработки, преобразования и передачи информации. В качестве примеров можно указать на идею использования динамического хаоса для передачи информации, открывающую интересные перспективы создания новых средств связи, на идею построения нейрокомпьютерных систем обработки информации на базе принципов нейродинамики и др. Существенный интерес наблюдается к «информационной динамике сложных систем» – новому междисциплинарному направлению в нелинейной науке, основанному на исследовании свойств живых и искусственных нейронных систем и моделировании интеллектуальных функций мозга. Такие систе-

мы способны не только передавать и обрабатывать поступившую информацию, но и генерировать новую. Именно этой проблематике, которую условно можно обозначить как «информационные аспекты нелинейной динамики», на школе «Нелинейные волны-2002» была посвящена значительная часть лекций и докладов.

Таким образом, сложившаяся на предыдущих школах традиция широко представлять сравнительно новые для теории колебаний и волн проблемы была продолжена и на школе «Нелинейные волны-2002». Помимо «информационных аспектов нелинейной динамики», на школе обсуждались традиционные проблемы, связанные с общими вопросами волновой динамики, нелинейными явлениями в океане и атмосфере. Другой цикл лекций был посвящен нелинейным явлениям в плазме, нелинейной оптике и явлениям взаимодействия сверхсильных полей с веществом. Большой интерес вызвали лекции по проблемам нелинейной динамики в медицине и биологии. На школе присутствовал также традиционный цикл лекций по математическим проблемам теории колебаний и волн.

Наряду с вошедшими в этот сборник работами, на школе были также прочитаны лекции О. В. Руденко «Нелинейная акустика и медицина», Л. В. Келдыша «Многофотонная ионизация атомов сверхкороткими лазерными импульсами», А. А. Андронova «Проблемы блоховского генератора в сверхрешетках», В. Я. Демиховского «Квантовый компьютер: варианты реализации», В. Б. Гильденбурга «Нелинейная динамика волновых пучков и пакетов в ионизируемой среде», Г. Т. Гурия «Макроскопическое структурообразование в динамике крови», В. Н. Белых «Гомоклинический хаос», С. А. Кашенко «Асимптотические методы анализа сложной динамики в системах с запаздыванием».

Конечно, представленные на школе лекции и доклады не исчерпывают всего многообразия исследований в области современной нелинейной динамики. Однако мы надеемся, что ключевые, наиболее актуальные направления этих исследований на школе 2002 года в той или иной степени были представлены.

В заключение мы выражаем признательность авторам, представившим материалы в данный сборник, и надеемся, что они примут участие в работе следующих Нижегородских школ по нелинейным волнам.

*А. В. Гапонов-Грехов,
В. И. Некоркин*

Информационные аспекты нелинейной динамики

ДИНАМИКА КОЛЕБАТЕЛЬНЫХ НЕЙРОНОВ. ИНФОРМАЦИОННЫЕ АСПЕКТЫ

В. Б. Казанцев, В. И. Некоркин

1. Введение

Одним из наиболее интересных направлений современной нелинейной физики является исследование явлений и процессов в так называемых нейродинамических системах. Нейродинамические системы – это системы, моделирующие структуру и свойства живых клеток, нейронов и нейронных ансамблей. Построение и исследование физических моделей таких систем является очень сложной, однако весьма привлекательной задачей. Трудности вызваны двумя основными причинами. С одной стороны, сам объект исследования устроен чрезвычайно сложно [1–8]. Так, например, по последним оценкам нейрофизиологов головной мозг состоит более чем из 10^{11} нервных клеток, взаимодействующих между собой. Кроме того, существует 10^3 различных типов этих клеток. С другой стороны, получение экспериментальных данных имеет очевидные трудности, связанные, во-первых, с малыми размерами нервных клеток (~0,2 мкм), во-вторых – со сложностью самих измерений – объект должен оставаться живым. Поэтому здесь важен *физический*, модельный подход. Исследование динамики даже сильно упрощенных нейродинамических моделей позволяет объяснить основные механизмы функционирования реальных систем, а также предсказать их поведение в тех или иных условиях. Чрезвычайно важен и прикладной аспект. Во-первых, такие исследования могут сыграть определяющую роль в диагностике различных динамических заболеваний (*dynamic deceases*) и указать способы их лечения. Яркий пример – возникновение сердечных фибрилляций и аритмий, вызванных образованием в сердечной ткани спиральных волн, нарушающих нормальную работу сердца [1]. Во-вторых, на основе моделей нейродинамики разрабатываются информационно-вычислительные системы (нейрокомпьютеры) нового поколения. В этой области, наряду с уже имеющимися успехами (например, различные системы распознавания образов), имеются очень большие перспективы. В частности, многообещающим выглядят моделирование процессов ассоциативной памяти, пространственно-временной селекции, построение систем управления, контроля и координации движений автономных машин-роботов с использованием принципов нейродинамики [9–11]. Эффективность таких систем будет на несколько порядков превосходить мощности современных компьютеров. Для сравнения, простейшее хватательное движение пальца руки, управляемое нейронами головного мозга, требует одновременного сокращения около 50 мышц [7]. Для управления таким движением с по-

мощью компьютера, т. е. выбора оптимальной комбинации мышц в течение 1 мс, требуется процессор с частотой 10^6 ГГц!

Нейродинамические системы подобно реальным нейронным ансамблям представляют собой сети взаимосвязанных элементов – «нейронов», локализованных в узлах пространственной решетки с различной архитектурой. В отличие от известных «формальных» нейронных сетей (персептрон), где элемент является либо логической ячейкой («0» или «1»), либо функцией связи входа и выхода, в *нейродинамических* моделях, о которых пойдет речь, элемент обладает собственной, в некоторых случаях нетривиальной динамикой [12–16]. Таким образом, поведение таких систем характеризуется развитием процессов как во времени, так и в пространстве. Другими словами, нейродинамические системы можно трактовать как некоторые дискретные неравновесные среды. Несмотря на большое число элементов, за счет введения межэлементных связей они обладают удивительным свойством – способностью к *самоорганизации*. В процессе автономной эволюции они способны производить когерентные пространственные или пространственно-временные динамические структуры без специфического воздействия извне [1, 17–19]. Это свойство является характерной чертой реальных нейронных ансамблей. В частности, структуры электрической активности коры головного мозга свидетельствуют о том, что именно с помощью такого рода организации осуществляются различные функции контроля жизнедеятельности организма.

В силу чрезвычайной сложности объекта исследования получить точную модель нейронного ансамбля, и даже отдельно взятого нейрона, с использованием уравнений клеточной мембраны (проводимость мембраны, ионные токи через каналы и т. д.) достаточно сложно. Примером такой модели являются известные уравнения Ходжкина – Хаксли, описывающие возбудимые свойства мембран с помощью уравнений баланса ионных токов через мембрану. В самом простейшем выражении такая модель вовлекает большое число переменных (до 6) и параметров (до 15). Ясно, что изучение свойств даже сравнительно малых ансамблей таких элементов является затруднительным. Для понимания коллективной динамики нейронных сетей используются более простые модели, лишь *качественно* отражающие основные характеристики реальных нейронов. Это так называемые функциональные модели (*behavior-based models*). Их преимуществом, кроме простоты, является еще и то, что они сравнительно легко реализуются в виде электрических аналоговых схем.

2. Динамика нейронов с подпороговыми колебаниями

С функциональной точки зрения можно выделить три основных состояния нейронов – относительного покоя, когда величины внешнего стимула или спонтанных возмущений недостаточно для преодоления некоторого порога, при превышении которого нейрон переходит в состоя-

ние возбуждения (активности) и генерирует импульс (потенциал действия) или последовательность импульсов, и, наконец, состояние рефрактерности, когда нейрон не реагирует на внешние стимулы. Среди большого разнообразия видов нейронов важное место занимают нейроны, демонстрирующие колебательную активность ниже порога возбуждения (далее колебательные нейроны). Например, результаты исследований нейрофизиологов [5–8] показывают, что в различных частях головного мозга существует колебательная активность на определенных частотах (таламо-кортикальные колебания ~ 40 Гц, оливо-церебеллярные колебания ~ 10 Гц). Динамические механизмы такой активности связаны, как правило, с колебательными свойствами системы на уровне отдельных нейронов.

2.1. Модель нейрона с «нелинейным восстановлением»

Одной из наиболее простых моделей нейронов, отражающих их основные характеристики, является известная модель ФитцХью – Нагумо [4]. Рассмотрим модификацию этой модели, позволяющей учесть колебательную активность нейронов ниже порога возбуждения:

$$\begin{cases} \dot{u} = f(u) - v; \\ \dot{v} = \varepsilon(g(u) - v - I). \end{cases} \quad (1)$$

Здесь переменная u определяет (качественно) мембранный потенциал клетки, v – «восстанавливающая» переменная (определяет динамику ионных токов, отвечающих за восстановление потенциала покоя нейрона). Нелинейность $f(u)$ – кубического типа (для определенности $f(u) = u - u^3/3$); ε – параметр, определяющий характерные временные масштабы импульса возбуждения; I – постоянный внешний стимул (например, ток стимуляции). В отличие от классической модели во второе уравнение системы (1) введена нелинейная функция $g(u) = \alpha_1 u$ при $u < 0$ при $\alpha_2 u$, $u \geq 0$. Такая модификация, фактически, означает учет нелинейных свойств динамики ионных токов зависимости от мембранного потенциала клетки.

Рассмотрим динамику системы (1) на фазовой плоскости (рис. 1).

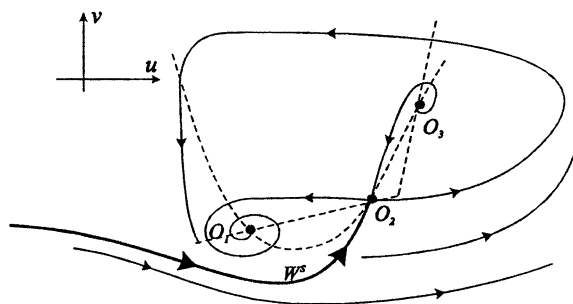


Рис. 1. Фазовая плоскость модели нейрона с нелинейным восстановлением

При выбранных параметрах она имеет три состояния равновесия O_1 ($u^{(1)}, v^{(1)}$) (устойчивый фокус с собственными значениями $\lambda_{1,2} = -h_1 \pm i\omega_1$), O_2 ($u^{(2)}, v^{(2)}$) (седло, $\lambda_1 > 0, \lambda_2 < 0$) и O_3 ($u^{(3)}, v^{(3)}$) (неустойчивый фокус, $\lambda_{1,2} = h_2 \pm i\omega_2$). Устойчивое состояние равновесия O_1 отвечает состоянию покоя нейрона, входящая сепаратриса седла O_2 определяет порог возбуждения. Отметим, что затухающие колебания в окрестности состояния покоя приводят к появлению у нейрона характерных собственных временных масштабов.

Отклик нейрона на внешний сигнал. При появлении внешнего стимула, превышающего порог, нейрон генерирует импульс возбуждения – отклик. Внешним стимулом является последовательность импульсов, получаемая нейроном от соседей через синаптическую связь. Действие синаптической связи заключается в следующем. Каждый приходящий импульс за счет различных биохимических механизмов создает так называемый постсинаптический потенциал (ПСП) [2, 3]. Если его величина сравнительно мала, то нейрон возвращается к состоянию покоя. При достижении порога возбуждения нейрон генерирует отклик. Синаптическая связь может быть возбуждающей и тормозной. В первом случае постсинаптический потенциал (ВПСП) имеет положительный знак и деполяризует мембрану, во втором случае (ТПСП) – отрицательный и гиперполяризует мембрану (рис. 2).



Рис. 2. Возбуждающий и тормозной постсинаптический потенциал

Применительно к модели (1) аппроксимируем действие синаптической связи следующим образом. При получении входного импульса в момент времени t_0 величина переменной u (мембранный потенциал) изменяется на величину u_p ($u_p > 0$ для ВПСП и $u_p < 0$ для ТПСП). Далее, до прихода следующего импульса в момент времени $t_0 + \tau_p$ динамика нейрона определяется автономной системой (1). Таким образом,

$$\begin{aligned} u(t) &= u(t, u_{i_0}, v_{i_0}); \\ v(t) &= v(t, u_{i_0}, v_{i_0}); \\ u_{i_0} &= u(t_0) + u_p; \quad v_{i_0} = v(t_0). \end{aligned} \quad (2)$$

На фазовой плоскости (см. рис. 1) условия (2) определяют мгновенный скачок переменной u в момент времени прихода импульса t_0 на величину u_p .

Предположим, что нейрон получает входной сигнал в форме периодической последовательности импульсов с межимпульсным (межспайковым) интервалом τ_p . Обозначим состояние системы после прихода n -го импульса (u_n, v_n). Тогда для следующего импульса

$$u_{n+1} = u(\tau_p, u_n, v_n) + u_p; \quad (3)$$

$$v_{n+1} = v(\tau_p, u_n, v_n).$$

Следовательно, траектории неавтономной системы (1), (2) определяют в трехмерном фазовом пространстве нелинейное двумерное отображение Пуанкаре (3). Пусть в начальный момент времени $t = 0$ нейрон находится в состоянии покоя O_1 . Тогда состояние нейрона на момент прихода n -го импульса будет определяться траекторией отображения (3) с начальными условиями $u_1 = u^{(1)} + u_p$, $v_1 = v^{(1)}$. Порог возбуждения в модели задается входящей сепаратрисой седла W^s : $v = W(u)$ (см. рис. 1). Тогда условие генерации импульса отклика в ответ на n -й входной импульс можно записать в виде

$$v_n - W(u_n) < 0. \quad (4)$$

Интегрирующий отклик. Рассмотрим сначала случай, когда подпороговые колебания отсутствуют (состояние равновесия O_1 – устойчивый узел). При $\varepsilon \ll 1$ отображение (3) можно получить аналитически [20]:

$$u_{n+1} = u^{(1)} + \frac{(u_n - u^{(1)})[u_p - (u_p + A)\exp(-A\tau_p)] - u_p A}{(u_n - u^{(1)})[1 - \exp(-A\tau_p)] - A} + u_p; \quad (5)$$

$$v_{n+1} = v_n,$$

где $A = \sqrt{\alpha_1^2 - 4(\alpha_1 + I - 2/3) - \alpha_1}$. Условие генерации импульса отклика (4) в этом случае переписывается в виде $u_n > A$. Динамика отображения (5) представлена на рис. 3, а.

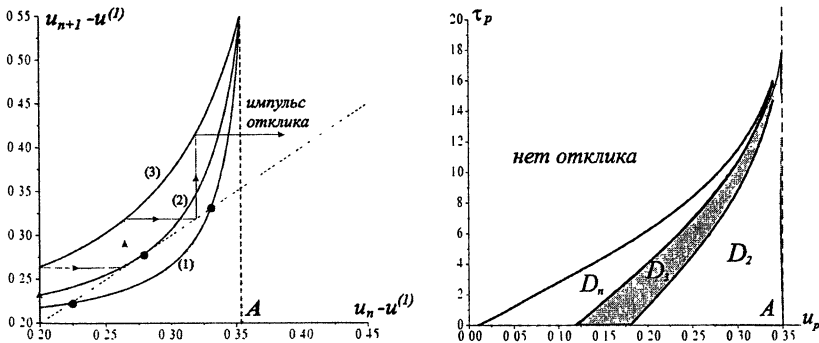


Рис. 3. Интегрирующий отклик (а) и точечное отображение (5) при уменьшении межимпульсного интервала ($\alpha_1 = 2$, $I = 0,4$, $u_p = 0,2$, $\varepsilon = 0,001$) Кривая 1 – $\tau_p = 9$, 2 – $\tau_p = 7,24$; 3 – $\tau_p = 5$. Области отклика (б): D_n – отклик на n импульсов.

В случае, когда оно имеет две неподвижные точки, все траектории стремятся к устойчивой из них, расположенной ниже порога возбуждения, и отклик невозможен. При исчезновении неподвижных точек все траектории отображения (5) монотонно возрастают, преодолевая порог при определенном числе входных импульсов n . Другими словами, система «суммирует» («интегрирует») входные импульсы в зависимости от интервала следования τ_p и величины синаптической связи u_p . На рис. 3, б представлены области отклика на дуплет и триплет входных импульсов.

Резонансный отклик. Рассмотрим теперь отклик при наличии подпороговых колебаний (см. рис. 1). В этом случае отображение (3) можно получить численным интегрированием системы (1) с соответствующими начальными условиями. Области отклика системы на определенное число импульсов (дуплет и триплет), полученные согласно формуле (4), представлены на рис. 4, а. Их расположение на плоскости (u_p, τ_p) носит резонансный характер. Отметим, что подпороговые колебания являются существенно неизохронными и области отклика имеют достаточно сложную структуру. Таким образом, нейрон с подпороговыми колебаниями обладает свойством селективного отклика. Более того, кроме частотной селективности, отклик сохраняет и интегрирующие свойства. Система генерирует импульс отклика в ответ на определенное число входных импульсов, следующих с заданным межспайковым интервалом. На рис. 4, б представлена эволюция переменной u (мембранного потенциала) при отклике на $n = 8$ импульсов во входной последовательности.

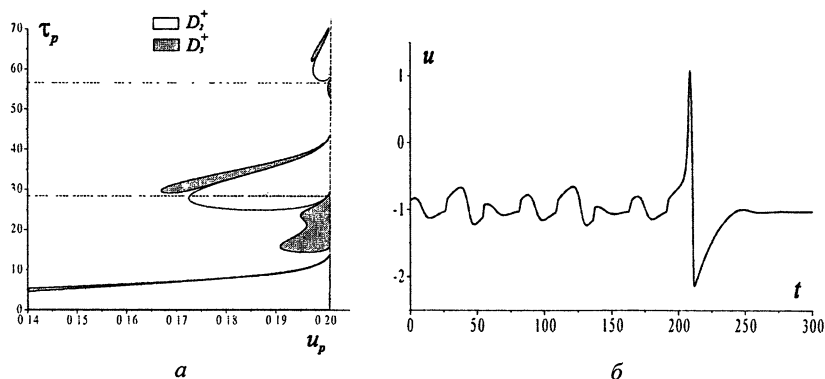


Рис. 4. Области резонансного отклика D_n^+ на ВПСП (а) и отклик на последовательность $n = 8$ входных импульсов (б): $\varepsilon = 0,1$; $\alpha_1 = 0,5$; $\alpha_2 = 10$; $I = 0,15$; $u_p = 0,172$; $\tau_p = 27,5$.

Отметим здесь несколько важных информационных аспектов. Во-первых, наличие частотно-селективных свойств определяет возможность мультиплексирования сигналов и использования одного канала для пере-

дачи больших объемов информации. Во-вторых, наличие вычислительных (интегрирующих) свойств на выделенных частотах позволяет колебательным нейронам осуществлять обработку информации раздельно или параллельно, что является ключом к увеличению производительности любой вычислительной системы.

Резонансный отклик на ТПСП. Некоторые нейроны способны генерировать отклик и в случае тормозной синаптической связи ($u_p < 0$). Этот эффект также может быть описан в рамках модели (1), (2). Порог возбуждения, определяемый сепаратрисой W^s , является существенно нелинейной функцией (см. рис. 1). Более того, эта функция возрастает при уменьшении u . Тогда при получении последовательности постсинаптических импульсов отрицательной полярности существует возможность преодолеть порог возбуждения (4).

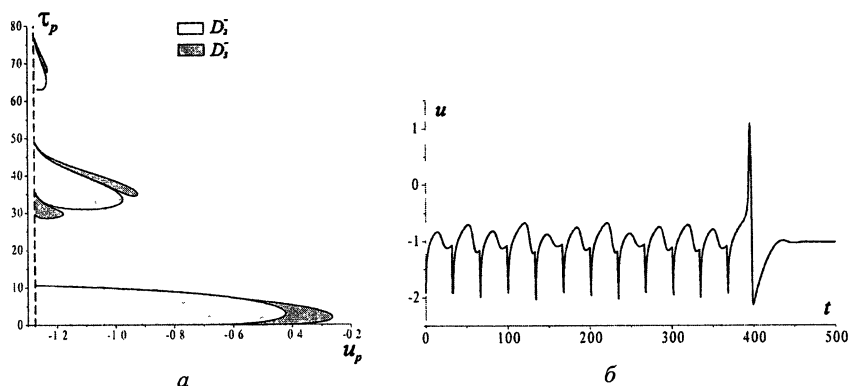


Рис. 5. Области резонансного отклика D_n^- на ТПСП (а) и отклик на последовательность $n = 12$ входных импульсов (б): $\varepsilon = 0,1$; $\alpha_1 = 0,5$; $\alpha_2 = 10$; $l = 0,15$; $u_p = -0,96$; $\tau_p = 33,5$.

Области резонансно-интегрирующего отклика на ТПСП представлены на рис. 5, а, рисунок б иллюстрирует отклик модели на $n = 12$ входных импульсов.

2.2. Однородная модель колебательного нейрона

В формировании колебаний в упомянутой выше оливо-церебеллярной системы основную роль играют нейроны нижних олив (нижние оливы – две группы нейронов, имеющие форму оливок и расположенные в стволе головного мозга) [5–8]. Эта система осуществляет контроль и координацию движений, определяя универсальный ритм (10 Гц) мышечных сокращений. Нейроны олив обладают очень специфическими свойствами. В невозбужденном состоянии они производят спонтанные, близкие к гармоническим колебания ниже порога возбуждения (подпороговые

колебания) с частотой $\Omega \sim 10$ Гц и амплитудой $5 \div 10$ мВ. Импульсы возбуждения возникают на пике подпороговых колебаний. Кроме того, существует два различных порога возбуждения. При достижении первого (низкого) порога на пике подпороговых колебаний нейрон генерирует широкий импульс ($T_1 \sim 0,1/\Omega$) относительно малой амплитуды (~ 20 мВ). Появление таких импульсов определяется в основном токами ионов Ca^{2+} , поэтому будем называть их Ca^{2+} -импульсы. При достижении второго порога нейрон генерирует мощные ($60 \div 100$ мВ) узкие ($T_2 \sim 0,01/\Omega$) импульсы (Na^+ -импульсы), появление которых также соотносится с максимумами подпороговых колебаний. Экспериментальные данные на рис. 6, полученные с нейронов нижних олив при различных видах стимуляции, иллюстрируют основные особенности динамики этих нейронов.

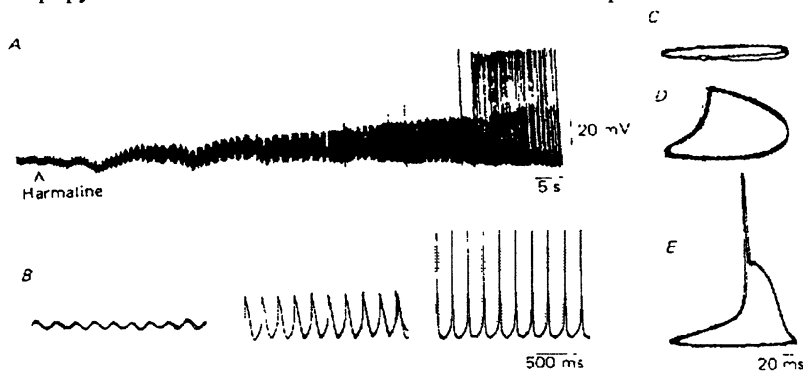


Рис. 6. Экспериментальные данные, полученные с нейронов нижних олив морской свинки на срезе в условиях *in vitro* (рисунок взят из работы Llinas and Yarom, 1986)

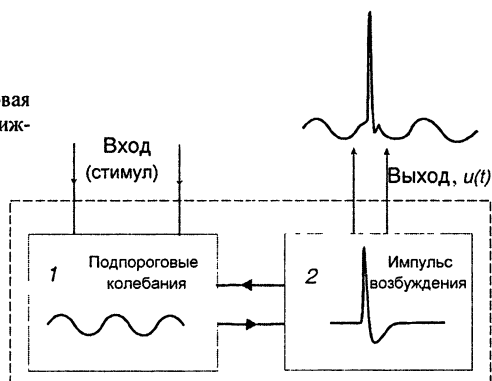
Рассмотрим сначала модель нейрона, имеющую лишь один порог возбуждения (по Na^+). Модель состоит из двух взаимодействующих блоков, отражающих различные функции нейрона (рис. 7) [21–24]. Первый блок описывает подпороговые колебания и может быть реализован в виде генератора Ван дер Поля в мягком режиме возбуждения. Второй блок отвечает за формирование импульса и реализуется в виде возбудимого элемента ФитцХью – Нагумо. При введении нелинейной связи между блоками получаем, что динамика модели может быть описана следующей системой 4-го порядка:

$$\begin{cases} \varepsilon_1 \dot{u} = f(u) - v - y, \\ \dot{v} = \varepsilon_2 (u + I), \\ \dot{x} = y, \\ \dot{y} = [\gamma(1 + \alpha I + \beta u) - lx^2]y - \omega_0^2 x, \end{cases} \quad (6)$$

где переменные x и y описывают динамику первого блока; u и v – второго; $f(u)$ – нелинейная функция кубической формы; $\varepsilon_1, \varepsilon_2$ – малые положитель-

ные параметры; I – постоянный внешний стимул; $\gamma, \beta, I > 0, \alpha < 0$ – параметры, характеризующие динамику переменных Ван дер Поля и связь между блоками.

Рис. 7. Однопороговая модель нейрона нижних олив



Заметим, что динамика переменной u в модели (6) качественно отражает эволюцию мембранного потенциала нейрона, переменные x, y, v – динамику ионных токов, параметр I определяет уровень деполяризации нейрона. Введение в систему (6) двух малых параметров ϵ_1, ϵ_2 необходимо для согласования характерных временных масштабов блоков (длительность импульса и период подпороговых колебаний) в соответствии с экспериментальными данными. Связь между блоками осуществляется следующим образом. Близкие к гармоническим колебания с первого блока (x, y) изменяют состояние второго по переменной u . В свою очередь, изменение амплитуды подпороговых колебаний происходит в зависимости от мембранного потенциала u и внешнего стимула I .

Рассмотрим динамику системы (6). Она имеет одну быструю и три медленные переменные. С помощью теории релаксационных колебаний установлено, что аттракторы системы (6) либо лежат в окрестности устойчивого многообразия медленных движений, либо формируются за счет движений между компонентами этого многообразия [22]. При этом на многообразии медленных движений траектории системы (6) совершают движения либо на конической поверхности, либо вблизи нее (рис. 8, $z \equiv u + y$).

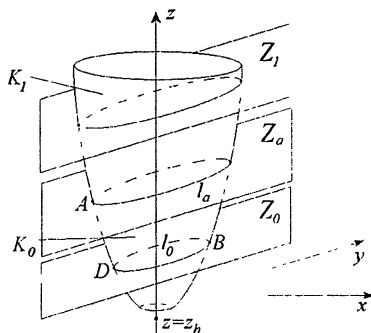


Рис. 8. Качественный вид фазового пространства системы (6). K_0, K_1 – устойчивые компоненты поверхности медленных движений, Z_0, Z_1 – плоскости переключения между K_0 и K_1 .

Эта коническая поверхность близка к поверхности, образованной предельным циклом изолированной системы Ван дер Поля (блок 1), в которой переменная $u = u_0 = \text{const}$ играет роль управляющего параметра. Поскольку движения системы (6) происходят вблизи (на) конической поверхности, траектории системы (6) порождают одномерное точечное отображение Пуанкаре Π . В случае, когда «обмена» траекториями между окрестностями устойчивых компонент (поверхности K_0 и K_1) многообразия медленных движений не происходит, отображение Π имеет единственную устойчивую неподвижную точку, соответствующую устойчивому предельному циклу C_0 системы (6). Цикл C_0 определяет квазигармонические подпороговые колебания модели (6) (рис. 9).

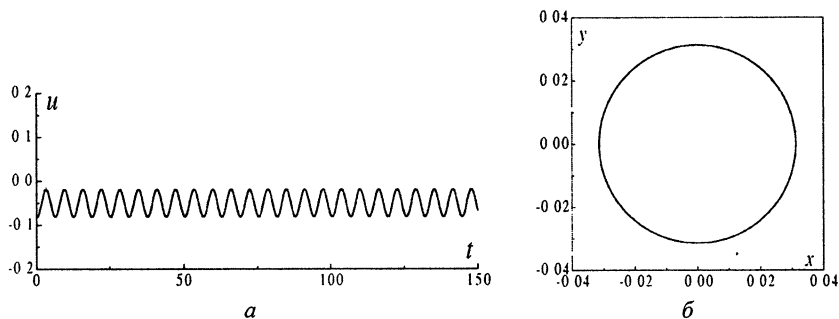


Рис. 9. Квазигармонические колебания, отвечающие предельному циклу C_0 эволюция мембранного потенциала (а) и фазовая траектория в проекции на плоскость (x, u) (б); $\alpha = -5$, $\beta = 5$, $\gamma = 0,001$, $l = 2$, $m_2 = 0,1$, $I = 0,05$.

При уменьшении параметра l цикл C_0 поднимается по компоненте K_0 и при некотором $l = l_0$ касается плоскости Z_a (см. рис. 8), что приводит к разрушению цикла C_0 . При $l < l_0$ возникает «обмен» траекториями между компонентами K_0 и K_1 , т. е. в системе (6) реализуются релаксационные колебания. Вид отображения Π в этом случае представлен на рис. 10.

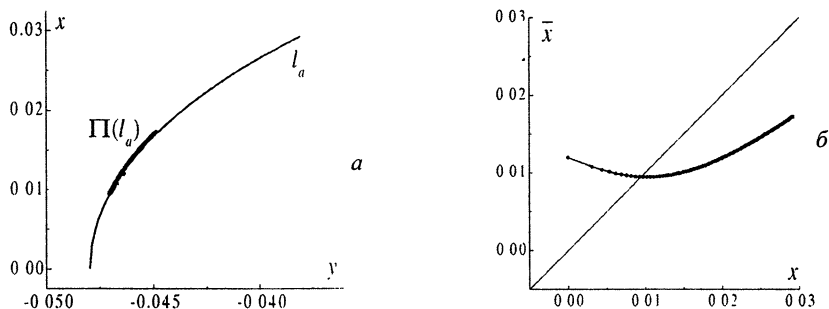


Рис. 10. Отображение Π для $l = -0,02$. а – преобразование кривой l_a в плоскости (x, y) , б – функция последования по переменной x .

Отображение Π имеет устойчивую неподвижную точку, ей отвечает устойчивый релаксационный предельный цикл системы (6) (рис. 11). В этом случае в модели реализуются подпороговые квазисинусоидальные колебания, на их фоне происходят резкие импульсные выбросы переменной u (напомним, u моделирует изменение мембранного потенциала нейрона).

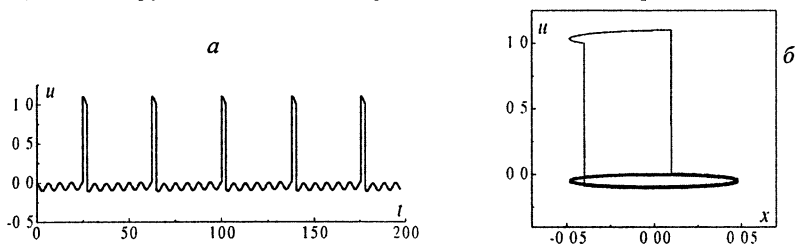


Рис. 11. Релаксационный предельный цикл для $I = -0,02$: эволюция мембранного потенциала (а) и фазовая траектория в проекции на плоскость (x, u) (б).

Представленный на рис. 11 предельный цикл является простейшим релаксационным циклом системы (6). Существуют и более сложные длиннопериодические релаксационные циклы, задающие другие соотношения между числом подпороговых колебаний и импульсов возбуждения. В этом случае отображение Π является разрывным, состоящим из нескольких непрерывных ветвей (рис. 12).

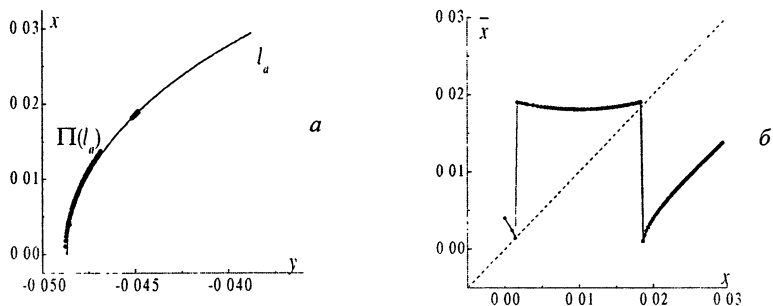


Рис. 12. Разрывное отображение Π для $I = -0,027$: преобразование кривой l_a в плоскости (x, y) (а) и функция последования по переменной x (б).

На рис. 13 представлен предельный цикл, соответствующий появлению импульсов на каждом 5-м и 6-м максимумах подпороговых колебаний в определенной последовательности. Заметим, что изменением параметров модели можно добиться различных интервалов следования импульсов (межспайкового интервала), кратных периоду подпороговых колебаний.

Таким образом, однопороговая колебательная модель нейрона (6) описывает генерацию импульсов возбуждения на фоне подпороговых колебаний и демонстрирует достаточно хорошее соответствие с данными

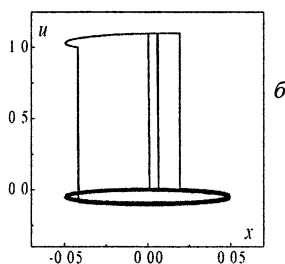
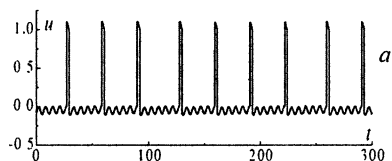
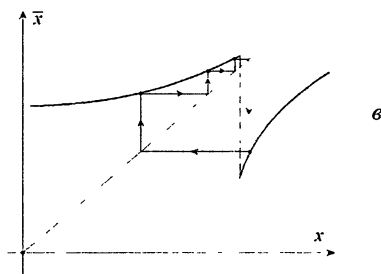


Рис. 13. Длиннопериодический релаксационный предельный цикл для $I = -0,027$: *а* – эволюция мембранного потенциала; *б* – фазовая траектория в проекции на плоскость (x, u) ; *в* – качественный вид траектории отображения, отвечающей предельному циклу.



экспериментов (см. рис. 8). Более того, структура модели в виде взаимодействующих блоков с различными свойствами позволяет моделировать и другие колебания, присущие различным типам нейронов. В частности, при определенном соотношении характерных временных масштабов блоков модели колебания имеют форму известных спайк-берст-колебаний и могут быть как регулярными, так и хаотическими. Пример хаотической временной реализации и соответствующее хаотическому аттрактору отображение Пуанкаре представлены на рис. 14.

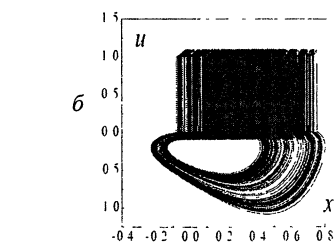
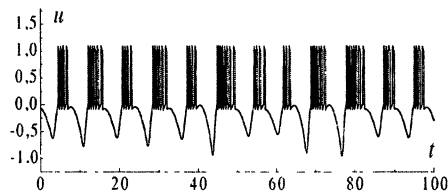
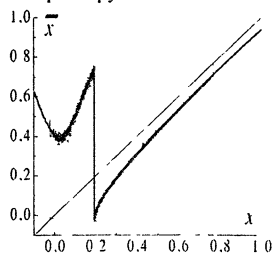


Рис. 14. Хаотические спайк-берст-колебания в модели (6) $I = -0,027$: *а* – эволюция мембранного потенциала, *б* – фазовая траектория в проекции на плоскость (x, u) ; *в* – отображение Пуанкаре, отвечающее хаотическому аттрактору



Генератор импульсов с подпороговыми колебаниями. Блочная структура модели (см. рис. 7) позволяет сравнительно легко реализовать ее в виде аналоговой радиотехнической схемы – генератора импульсов с подпороговыми колебаниями (рис. 15). За подпороговые колебания здесь

отвечает классический генератор синусоидальных колебаний (генератор ВдП) с полевым транзистором T_1 (блок I). В схему включены также два эмиттерных повторителя на транзисторах T_2 – необходим для управления режимом работы генератора ВдП – и T_3 – обеспечивает одностороннюю связь генератора ВдП с генератором импульсов, функцию которого выполняет работающий в жестком режиме возбуждения блокинг-генератор (блок II), запуск которого осуществляется напряжением синусоидальных колебаний генератора ВдП, поступающих на базу транзистора T_4 . С другой стороны, короткие импульсы блокинг-генератора, изменяя через эмиттерный повторитель T_2 напряжение питания активного элемента T_1 , могут влиять на параметры синусоидальных колебаний генератора ВдП.

На рис. 16 представлены лабораторная установка генератора импульсов с подпороговыми колебаниями и типичные временные реализации. Колебания генератора демонстрируют хорошее соответствие с экспериментальными данными (см. рис. 6) и с результатами теоретических исследований (см. рис. 11, 14).

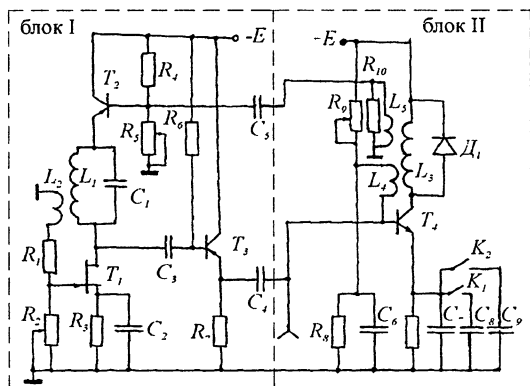


Рис. 15. Электрическая схема генератора импульсов с подпороговыми колебаниями (из работы V.V. Парко е. а., 2002 [24]).

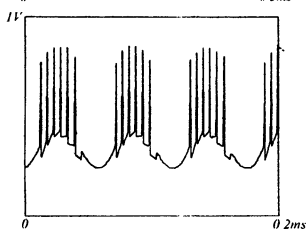
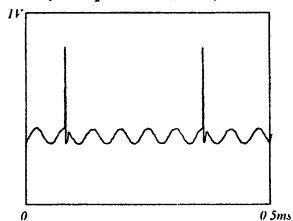
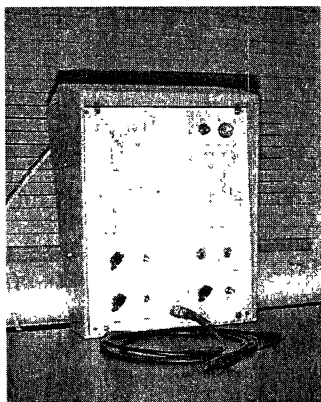


Рис. 16. Лабораторная установка генератора импульсов (слева) с подпороговыми колебаниями. Импульсы на фоне подпороговых колебаний (вверху), спайк-берст-колебания (внизу)

2.3. Двухпороговая модель колебательного нейрона

Для получения более точного описания динамики нейрона нижних олив введем колебательную модель нейрона с двумя порогами возбуждения. Она состоит из трех взаимодействующих блоков, отвечающих за различные функции (рис. 17). В отличие от однопороговой модели (см. рис. 7)

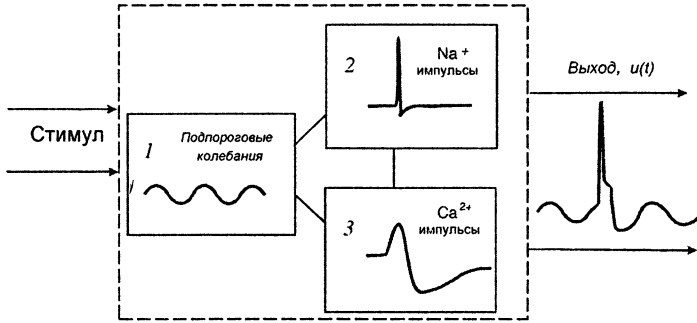


Рис. 17. Блок-диаграмма двухпороговой модели нижних олив

в схему включен дополнительный блок 3, отвечающий за генерацию Ca^{2+} -импульсов с низкого порога. Этот блок, как и блок 2, реализуется с помощью модели ФитцХью – Нагумо с другими значениями параметров для получения требуемой формы Ca^{2+} -импульсов. Динамика модели описывается системой дифференциальных уравнений 6-го порядка [23]:

$$\begin{cases} \dot{x} = y, & \text{(а)} \\ \dot{y} = [\gamma(1 + \alpha u)x^2]y - \omega^2(1 + \beta u)x, & \text{(б)} \\ \varepsilon_2 \dot{w} = g(w) - z - x, & \text{(в)} \\ \dot{z} = 0,5(w - I_2)(w^2 + 0,1), & \text{(г)} \\ \varepsilon_1 \dot{u} = f(u) - v + hw, & \text{(д)} \\ \dot{v} = 0,05(u - I_1)(u^2 + 0,5). & \text{(е)} \end{cases} \quad (7)$$

Здесь переменные x , y описывают динамику генератора Ван дер Поля (блок (а)), переменные w , z и u , v описывают блоки (в) и (б), отвечающие за генерацию импульсов с низкого и высокого порогов соответственно. Параметры I_1 и I_2 изменяют уровни деполяризации блоков (в) и (б) относительно двух различных порогов. Нелинейные функции $f(u)$, $g(u)$ выбираются кусочно-линейными для обеспечения требуемой формы импульса и взаимного расположения порогов возбуждения.

$$f(u) = \begin{cases} -1,5u, & u < a, \\ 0,2u - 1,7a, & a \leq u \leq 4, \\ -1,6u - 1,7a + 9,6, & a > 4. \end{cases} \quad g(w) = \begin{cases} -2w, & w < b, \\ 3w - 5b, & b \leq w \leq 1, \\ -5w - 5b + 8, & w > 1. \end{cases}$$

Остальные параметры и константы выбираются для обеспечения требуемого соотношения между характерными временными масштабами импульсов и подпороговых колебаний. Численное моделирование системы (7) проводилось при фиксированных $\varepsilon_1 = 0,005$, $\varepsilon_2 = 0,01$, $a = 1,8$, $b = 0,5$, $\omega^2 = 0,63$, $h = 0,5$, $\alpha = -0,5$, $\beta = 0,3$, $\gamma = 0,21$, а параметры I_1 и I_2 выбирались в качестве контрольных.

Основные динамические режимы. Для сравнительно малых значений I_1 и I_2 в модели реализуются квазигармонические подпороговые колебания. При увеличении одного из контрольных параметров (I_1 или I_2) одна из подсистем (7) – (б) или (в) – на пике квазигармонических колебаний достигает порога возбуждения, и модель генерирует импульс возбуждения. Возможны три различных типа импульсов.

Импульсы с низкого порога (Ca^{2+} -импульсы) возникают при возбуждении блока (в) – см. систему (7), в то время как блок (б) остается невозбужденным (рис. 18). Появление импульсов с высокого порога (Na^+ -импульсы) происходит аналогично, только в этом случае блок (б) возбужден, а блок (в) находится ниже порога возбуждения ($I_1 = 1,69$, $I_2 = -0,65$).

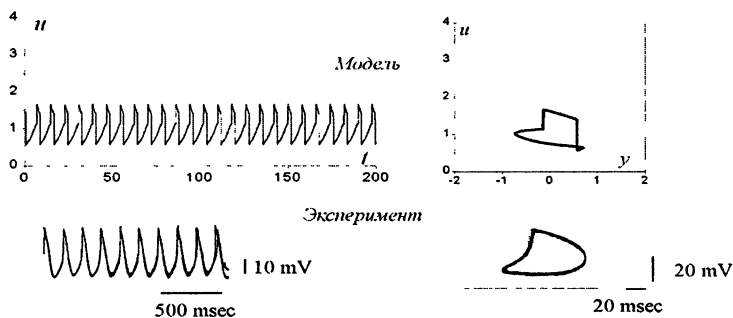


Рис. 18. Динамика модели в сравнении с экспериментальными данными (из статьи *Llinas and Yarom, 1986*). Импульсы с низкого порога возбуждения. $I_1 = 1,14$; $I_2 = -0,63$.

Эволюция мембранного потенциала качественно подобна однопороговой модели (см. рис. 11, 12). Наиболее интересен случай двухпороговой генерации, при котором импульс с низкого порога (блок (в)) запускает генератор импульсов с высоким порогом (блок (б)). Таким образом, импульс возбуждения появляется в результате последовательного преодоления системой (7) двух порогов ($\text{Ca}^{2+} + \text{Na}^+$ -импульс) (рис. 19).

Отметим, что изменением контрольных параметров I_1 и I_2 , меняющих, фактически, расположение подпороговых колебаний относительно двух порогов возбуждения, можно добиться любого интервала следования импульсов, кратного периоду подпороговых колебаний. Наличие второго порога возбуждения по Ca^{2+} играет важную функциональную роль в динамике нейронов нижних олив и всей оливо-цереbellарной петли.

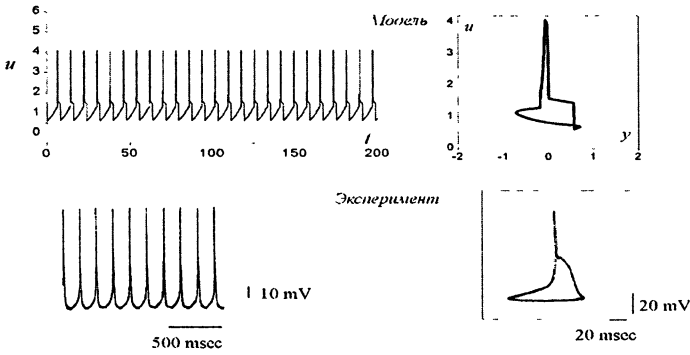


Рис. 19 Двухпороговая генерация импульсов. Динамика модели в сравнении с экспериментальными данными (из статьи *Linas and Yarom, 1986*). $I_1 = 1,97$, $I_2 = -0,27$.

Импульсы Ca^{2+} фактически являются опорными для основных Na^+ -импульсов, вызывающих распространение потенциала действия по аксонам нейронов нижних олив. Таким образом, строго фиксируется момент появления основных импульсов, так как пик опорного импульса является достаточно острым (по сравнению с максимумом подпороговых колебаний в однопороговой модели).

Двухпороговая генерация при наличии внешнего стимула. Для иллюстрации двухпороговой генерации рассмотрим динамику системы (7) при наличии внешнего стимула. Положим в уравнениях (7) $I_1 = I_1^0 + I_1^s(t)$, $I_2 = I_2^0 + I_2^s(t)$, где I_1^0 , I_2^0 – некоторые константы, а $I_1^s(t)$, $I_2^s(t)$ – медленно меняющиеся функции времени. Для исключения разных резонансных явлений предположим, что подпороговые колебания подавлены ($\gamma < 0$). Динамика модели при воздействии пилообразного сигнала показана на рис. 20. При различных уровнях деполяризации появляются составные ($\text{Ca}^{2+} + \text{Na}^+$) импульсы и Na^+ -импульсы с высокого порога возбуждения.

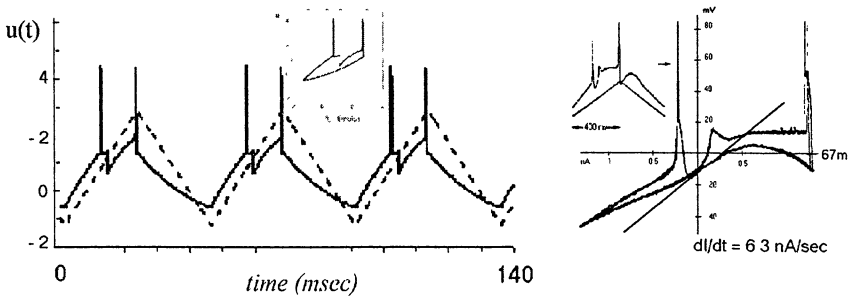


Рис. 20. Отклик двухпороговой модели на пилообразный внешний сигнал. Экспериментальные данные (по статье *Linas, Yarom, 1986*).

Отметим, что модель демонстрирует очень хорошее качественное соответствие с экспериментальными данными (см. рис. 6, 20).

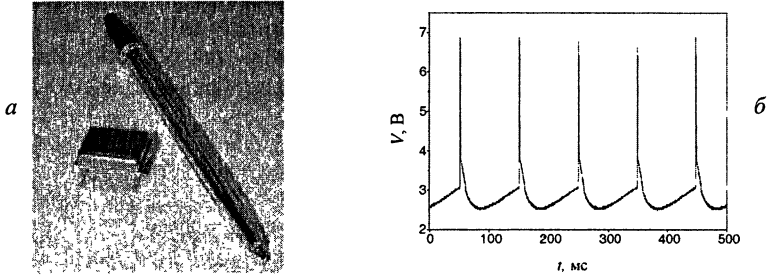


Рис. 21. Реализация двухпороговой модели колебательного нейрона в виде микрочипа (а) и двухпороговая генерация импульсов (б).

На основе блочной структуры (см. рис. 17) двухпороговая модель нейрона нижних олив была реализована в виде микрочипа (гибридной интегральной микросхемы) (рис. 21).

3. Коллективная динамика ансамбля колебательных нейронов

Рассмотрим поведение ансамбля колебательных нейронов, используя введенную в разд. 2.1 модель с нелинейным восстановлением. Предположим, что нейроны связаны между собой электрически, т. е. через электрический синапс [2, 3]. Действие синаптической связи электрического типа основано на существовании между соседними нейронами так называемых гэп-контактов. Такой контакт приводит к обобществлению клеточной мембраны между нейронами в некоторой узкой области. Вследствие этого ток синаптической связи определяется сопротивлением контакта и разностью потенциалов между нейронами. Таким образом, синаптическая связь электрического типа достаточно хорошо аппроксимируется линейной резистивной (диффузионной) связью. Коллективное поведение ансамбля таких нейронов описывается динамической системой:

$$\begin{cases} \dot{u}_j = f(u_j) - v_j + d(u_{j-1} - 2u_j + u_{j+1}) \\ \dot{v}_j = \epsilon[g(u_j) - v_j - I] \end{cases} \quad j = 1, 2, \dots, N, \quad (8)$$

где j – номер элемента-нейрона, N – число элементов в ансамбле, а d характеризует силу электрической связи между элементами. Одной из основных проблем динамики нейронных ансамблей является установление формы и механизмов распространения активности, локально возникающей в определенной точке пространства. Рассмотрим эту задачу, считая, что элементы-нейроны ансамбля локализованы в узлах одномерной пространственной решетки (цепочки). В этом случае индекс j играет роль

пространственной координаты, и динамика системы (8) является пространственно-временной. Рассмотрим распространение активности в системе (8) в форме бегущих волн – импульсов и волновых фронтов [25]. Будем предполагать, что характерный пространственный масштаб волн достаточно велик, т. е. на профиле волны укладывается много элементов цепочки (длинноволновая аппроксимация). Тогда профили возможных решений описываются следующей системой:

$$\begin{cases} \dot{u} = y, \\ \dot{y} = \frac{c^2}{d}(y - f(u) + v), \\ \dot{v} = \varepsilon[g(u) - v - I], \end{cases} \quad (9)$$

где точкой обозначено дифференцирование по $\xi = t - j/c$ – бегущей координате. При параметрах, отвечающих колебательной активности элемента-нейрона ниже порога возбуждения (см. рис. 1), система (8) имеет три состояния равновесия типа седло-фокус. Два из этих состояний равновесия O_1 и O_3 имеют неустойчивое одномерное и устойчивое двумерное многообразия, а состояние равновесия O_2 – наоборот, устойчивое одномерное и неустойчивое двумерное многообразия (рис. 22, б). Бегущие импульсы и фронты отвечают соответственно гомоклиническим и гетероклиническим траекториям системы (9). На рис. 22, а представлены результаты численного изучения бифуркационных множеств, отвечающих гомоклиническим (множество Γ) и гетероклиническим (множество H) траекториям системы (9). Каждое из этих множеств имеет форму спирали с общим центром. Точкам множества Γ соответствуют гомоклинические траектории, образованные одномерной сепаратрисой точки O_1 , которая возвращается в O_1 , «охватывая» точку O_3 . При этом по мере движения по Γ к центру спирали сепаратриса совершает постоянно увеличивающееся число осцилляций в окрестности точки O_2 .

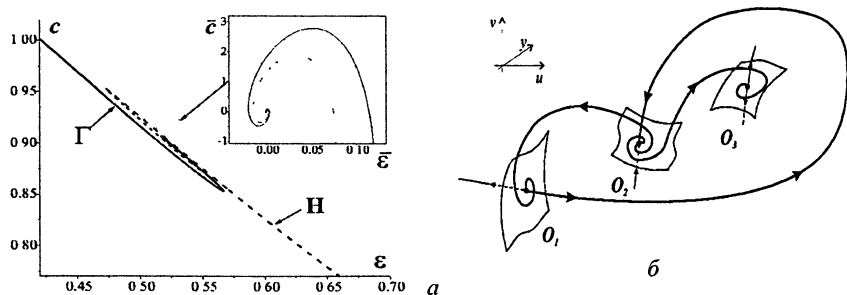


Рис. 22. Зависимость скорости распространения импульсов (кривая Γ) и волновых фронтов (H) в системе (9) (а) и гетероклинический контур для параметров центра спирали (б), отвечающий бифуркации коразмерности 2. $\alpha_1 = 0,5$, $\alpha_2 = 2$, $I = 0,2$

Это свойство гомоклинических траекторий системы (9) иллюстрирует рис. 23, на котором показаны соответствующие временные реализации.

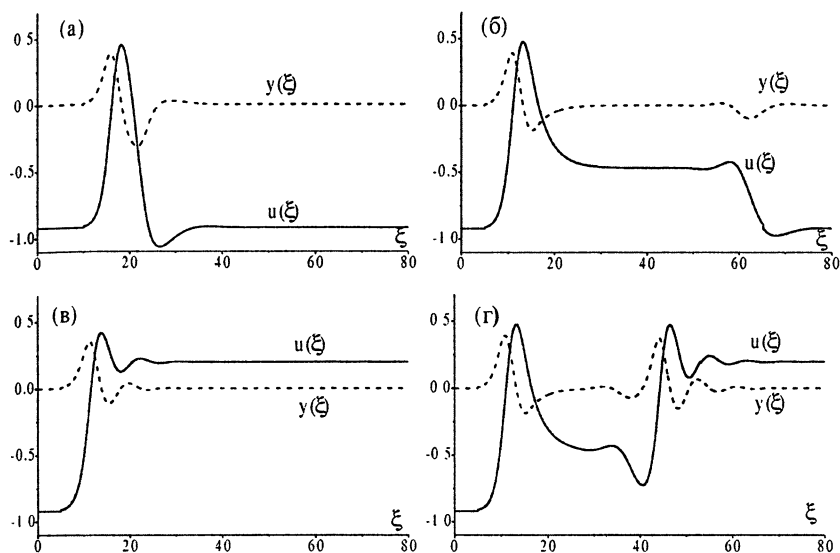


Рис. 23. Профили гомо- и гетероклинических траекторий в системе (9). $\alpha_1 = 0,5$; $\alpha_2 = 2$, $I = 0,2$; $d = 1$; $\varepsilon = 0,52276$. $c = 0,893$ (а); $c = 0,90041$ (б); $c = 0,902$ (в); $c = 0,90042$ (г).

Точкам множества H соответствуют гетероклинические траектории, образованные сепаратрисой седло-фокуса O_1 , асимптотически приближающейся к точке O_3 . Здесь также при изменении параметров по H к центру спирали сепаратриса осциллирует в окрестности точки O_3 . Это приводит к увеличивающемуся числу осцилляций на профиле волновых фронтов. Центр спирали отвечает бифуркации коразмерности 2 и существованию в фазовом пространстве гетероклинического контура, образованного траекториями всех трех седло-фокусов системы (9). Таким образом, для значений параметра ε вблизи центра спиралей система (9) имеет одновременно счетное множество разнообразных волновых движений. Поскольку ветви спиралей Γ и H тесно прижаты друг к другу, скорости волн мало отличаются. Существование большого числа различных волновых движений в этой области параметров ведет к формированию сложных волновых структур в системе (8). Было установлено, что устойчивыми могут быть лишь импульсы, имеющие наименьшую скорость (см. рис. 23, а, нижняя ветвь кривой Γ на рис. 22, а). Все другие бегущие волны, в том числе импульсы сложной формы и все волновые фронты (рис. 23, б-г), оказались неустойчивыми. Это вполне ожидаемый результат, так как в рассматриваемой области параметров пространственно-однородные состояния O_2 и

O_3 неустойчивы. На рис. 24 представлены пространственно-временные диаграммы, показывающие два наиболее типичных сценария распространения активности в ансамбле (9).

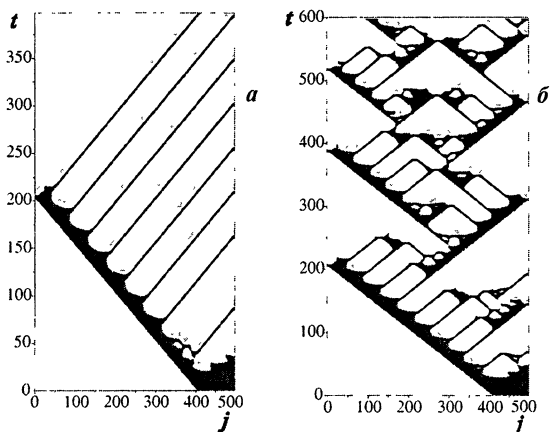


Рис. 24. Пространственно-временные диаграммы активности ансамбля (8). *a* – образование последовательности импульсов вследствие распада волнового фронта ($\alpha_1 = 0,5$, $\alpha_2 = 2$, $l = 0,2$, $d = 1$, $\varepsilon = 0,53$); *б* – пространственно-временной хаос ($\alpha_1 = 0,5$, $\alpha_2 = 2$, $l = 0,2$, $d = 1$, $\varepsilon = 0,57$)

На рис. 24, *a* начальное возмущение группы элементов начинает распространяться в виде волнового фронта, который в силу неустойчивости распадается на серию устойчивых импульсов возбуждения, распространяющихся в обратном направлении. Более интересный сценарий имеет место при увеличении параметра ε (вблизи первой точки поворота кривой Γ). Здесь сами импульсы возбуждения теряют устойчивость, распадаясь на пару волновых фронтов или импульсов, которые, в свою очередь, по мере развития неустойчивости распадаются на очередную пару импульсов и т. д. (рис. 24, *б*). Результатом таких волновых неустойчивостей является пространственно-временной хаос, развивающийся как следствие существования счетного числа неустойчивых бегущих волн. Другим словами, система переходит из окрестности одной неустойчивой волны к другой, формируя пространственно-временную структуру.

Заметим, что описанная выше пространственно-временная динамика характерна не только для нейронных ансамблей. Так, «идейно» близкие сценарии возникновения сложных пространственно-временных паттернов были обнаружены в модели химической реакции CO окисления поверхности Pt (110) [26–27].

4. Нейродинамические системы контроля и координации движений

Как уже отмечалось выше, колебательные нейроны играют важную роль в информационно-вычислительных функциях нервной системы. Моделирование динамики таких нейронов открывает возможности для перспективных приложений создания систем контроля, управления и вычислений, использующих нейродинамические принципы [9, 10]. Рассмотрим

один из примеров таких приложений – нейродинамическую систему контроля и координации движений автономных машин – роботов.

Из работ нейрофизиологов [7, 8] следует, что любое движение происходит не непрерывно, а в некотором смысле (кинематическом) последовательно – имеет характерный временной масштаб, определяемый частотой порядка 10 Гц. Этот масштаб означает, что нервная система осуществляет контроль над сокращением мышц не чаще, чем через временной интервал около 100 мс, что позволяет, во-первых, сократить «вычислительные ресурсы», необходимые для контроля над движением; во-вторых – придать движению определенную «эластичность» – возможность адаптации к внешним условиям, точной настройки и т. д. На основе этих представлений была предложена нейродинамическая схема дискретного контроля и координации движений автономных машин – роботов (рис. 25).

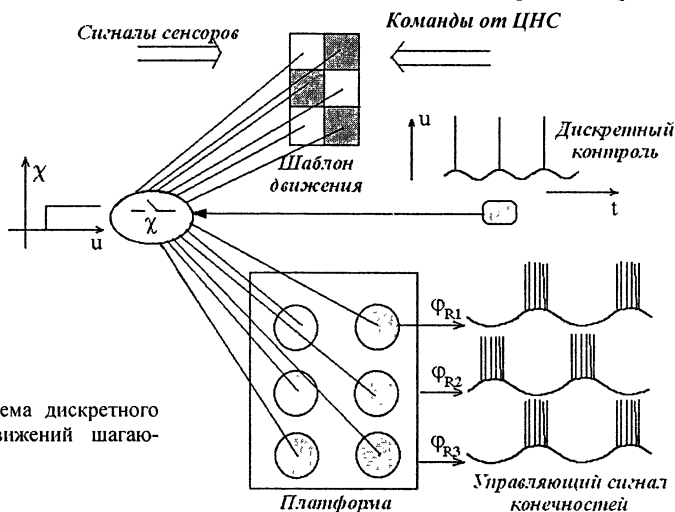


Рис. 25 Схема дискретного контроля движений шагающего робота

Основной задачей системы контроля шагающего робота (например, имеющего 6 конечностей) является обеспечение стабильности машины при движении по неровной поверхности. Идея предложенной здесь системы управления заключается во введении блока дискретного контроля, отражающего функции нейронов нижних олив. В данном случае блок был реализован с помощью однопороговой модели одного нейрона (6). Функция его состоит в следующем. Он генерирует контрольные импульсы с определенной частотой (10 Гц). В промежутках между импульсами конечности адаптируются к условиям поверхности самостоятельно за счет различных локальных механизмов автоподстройки, которые мы здесь не рассматриваем. В моменты появления импульсов осуществляется контроль над положением конечностей и восстановление фазовых соотношений («шаблона движения»), обеспечивающих стабильный режим ходьбы.

В частности, для устойчивости шестиконечного робота в каждый момент времени три его конечности должны находиться на поверхности. Предположим, что каждая из конечностей управляется некоторым квазипериодическим сигналом, и фаза этого сигнала однозначно определяет положение ноги робота в каждый момент времени. Тогда действие системы контроля сводится к взаимной координации шести фазовых переменных. Для этих целей введем следующую динамическую систему:

$$\dot{\varphi}_i = h\Phi(t, t_k, \Delta t)(\varphi_i^* - \varphi_i) + \xi_i, \quad I = 1, 2, \dots, 6, \quad (10)$$

где φ_i – фаза соответствующей конечности, φ_i^* – «опорная» фаза, задаваемая в соответствии с выбранным шаблоном движения. Функция Φ моделирует действие блока дискретного контроля и имеет вид

$$\Phi(t, t_k, \Delta t) = \begin{cases} 0, & t < t_k, \\ 1, & t_k < t < t_k + \Delta t, \end{cases} \quad (11)$$

где t_k – момент прихода k -го контрольного импульса, генерируемого системой (6) на максимумах подпороговых колебаний с периодом T_c . Длительность действия системы контроля после каждого импульса, Δt , достаточно мала, $\Delta t \sim 0,1T_c$. Величины ξ_i – случайные величины, не коррелированные между собой и равномерно распределенные в некотором интервале $\{-\xi_0, \xi_0\}$. Эти величины моделируют расстройки фаз конечностей, возникающие при адаптации системы к локальным неровностям поверхности за счет механизмов автоподстройки. В целом система работает следующим образом. При отсутствии контрольного сигнала фазы конечностей эволюционируют случайным образом. После прихода контрольного

сигнала в течение времени Δt система восстанавливает (корректирует) фазовые соотношения согласно шаблону движения (платформа соединяется с блоком шаблона движения на рис. 26). Скорость коррекции определяется параметром h . Отметим, что и остальные блоки системы контроля и координации движений (генератор ритма («шаблона») движения, системы автоподстройки) могут быть реализованы с помощью различных динамических систем. Например, различные шаблоны движения (фазовые соотношения движения конечностей) можно задавать с помощью бистабильных нейродинамических систем

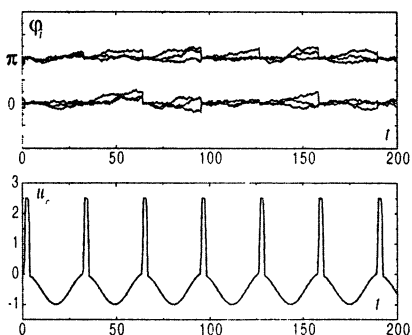


Рис. 26. Дискретный контроль фаз конечностей шестиногого робота с помощью (6) (контрольный сигнал – нижняя картинка) и (10). Фазы расположены в двух группах, отличающихся на π , что соответствует быстрому шаблону ходьбы (см рис 25)

в виде устойчивых пространственных структур бистабильных решеток [28].

5. Заключение

В работе исследованы нейродинамические системы, моделирующие динамику нейронов с колебательной активностью ниже порога возбуждения. Показано, что учет колебательных свойств позволяет описать ряд интересных как с фундаментальной, так и с прикладной точек зрения нелинейных эффектов. (i) Изучена динамика модифицированной модели ФитцХью – Нагумо с нелинейным восстановлением. Эта модель имеет строго определенный порог возбуждения (сепаратриса седла) и затухающие подпороговые колебания. При получении входного сигнала в виде последовательности импульсов с различным межспайковым интервалом такая система обладает селективным, интегрорезонансным откликом. Она способна выделять из последовательности определенное число импульсов с заданным интервалом следования. Это определяет возможности эффективного распознавания информации и параллельной обработки сигналов в ансамблях таких нейронов. С точки зрения перспективных приложений, такая модель может быть использована как минимальная единица в информационно-вычислительных системах нейродинамического типа – нейрокомпьютерах. (ii) Предложена модель нейронов, обладающих спонтанной колебательной активностью ниже порога возбуждения. Примером таких нейронов являются нейроны нижних олив, играющие ключевую роль в осуществлении контроля и координации движений в центральной нервной системе. Модель имеет достаточно простую, блочную структуру, где каждый из блоков отвечает за определенную функцию нейрона (подпороговые колебания, генерация импульса с одного или нескольких порогов возбуждения) и реализуется в виде сравнительно простых динамических систем (генератор Ван дер Поля, элемент ФитцХью – Нагумо). Динамика модели изучена с помощью анализа фазового пространства и построения отображений Пуанкаре. Показано, что модель воспроизводит практически все режимы колебаний нейрона (спонтанных и в присутствии внешнего стимула), полученные в экспериментах нейрофизиологов. Специфической особенностью модели является возможность двухпороговой генерации импульсов возбуждения. Более того, блочная структура придает модели определенную универсальность. В частности, при изменении характерных временных масштабов блоков она способна воспроизводить хаотические спайк-берст-колебания, типичные для других нейронов нервной системы. С прикладной точки зрения, модель сравнительно проста для реализации в виде радиозлектронных устройств, так как каждый из блоков фактически представляет собой известные радиотехнические схемы (генератор синусоидальных колебаний, блокинг-генератор в ждущем режиме). (iii) Наличие колебательной активности у отдельного

нейрона приводит к нетривиальным эффектам пространственно-временной динамики ансамблей таких единиц. Представленное в работе исследование цепочки электрически связанных элементов-нейронов с нелинейным восстановлением показало, что такая система способна формировать различные (в том числе и хаотические) пространственно-временные структуры активности. Эти структуры представляют собой импульсы и фронты возбуждения, неустойчивость которых приводит к установлению в системе самоподдерживающихся колебаний с определенной пространственной конфигурацией. При этом каждый из элементов генерирует последовательность импульсов возбуждения с различным интервалом следования. Такие структуры, с одной стороны, показывают возможность появления колебаний в ансамблях, элементы которых не являются автоколебательными (затухающие колебания ниже порога возбуждения, см. рис. 1). С другой стороны, колебания элементов ансамбля оказываются «сфазированными» между собой определенным образом согласно пространственно-временному паттерну и фактически формируют некоторые информационные последовательности импульсов, где информация кодируется в изменяющемся межспайковом интервале. (iv) На основе разработанной модели нейрона нижних олив предложена система контроля и координации движений автономных машин-роботов. Основной идеей здесь является введение в систему управления блока дискретного контроля, корректирующего ошибки, возникающие за счет действия механизмов автоподстройки при движении, например, по неровной поверхности.

Проведенное исследование колебательных нейродинамических систем показывает ряд перспективных приложений создания информационно-вычислительных систем, где такие модели будут играть ключевую роль (нейрокомпьютеры, системы распознавания, системы управления движением и др.).

Эти исследования проводились по госконтракту № 40.020.1.1.1168 Минпромнауки РФ и при частичной финансовой поддержке грантов РФФИ (00-02-16400, MAC 02-02-06287), российско-французского гранта 04610PA, гранта INTAS – YSF01/02-24.

Литература

1. *Murray J. D.* *Mathematical Biology*. Second Corrected Edition Springer-Verlag, Berlin, 1993.
2. *Kandel E. R., Schwartz J. H., Jessell T. M. (Eds.)* *Principles of Neural Science*, Third Edition. Prentice-Hall Intern. Inc., 1991.
3. *Рубин А. Б.* Биофизика. М.: Книжный дом Университет, 1999. Т. 1-2.
4. *FitzHugh R.* *Mathematical models of excitation and propagation in nerve* // *Biological Engineering* / Ed. H. P. Schwan, 1969. P. 3-85.
5. *Llinas R., Yarom Y.* *Oscillatory properties of guinea-pig inferior olivary neurones and their pharmacological modulation: An in vitro study* // *J. Physiol., Lond.* 1986. V. 376. P. 163.
6. *Llinas R.* *The intrinsic electrophysiological properties of Mammalian neurons: insight into central nervous system function* // *Science*. 1988. V. 242. P. 1654.

7. *Welsh J. P., Llinas R.* Some organizing principles for the control of movement based on olivocerebellar physiology / Eds. de Zeeuw C. I., Strata P. and Voodg // *Progress in Brain Research*. 1997. V. 114.
8. *Llinas R.* I of vortex. From neurons to self. The MIT Press, Massachusetts, 2002.
9. *Grillner S.* Neural networks for vertebrate locomotion // *Scientific American*. 1996. V. 64.
10. *Ekeberg O.* A combined neuronal and mechanical model of fish swimming // *Biol. Cybern.* 1993. V. 69. P. 363.
11. *Казанцев В. Б., Некоркин В. И.* Принципы контроля и координации движений на основе динамики нейронов головного мозга // *Изв. Вузов. Прикладная нелинейная динамика*. 2001 Т. 9, № 1. С. 38-48.
12. *Абарбанель Г. Д. И., Рабинович М. И., Сельверстон А., Баженов М. В., Хуэрта Р., Суцник М. М., Рубчинский Л. Л.* Синхронизация в нейронных ансамблях // *УФН*. 1996. Т. 166, № 4. С. 363-390.
13. *Борисюк Г. М., Борисюк Р. М., Казанович Я. Б., Лузянина Т. Б., Турова Т. С., Цымбалюк Т. С.* Осцилляторные нейронные сети. Математические результаты и приложения // *Математическое моделирование*. 1992. Т. 4, № 12. С. 3.
14. *Izhikevich E. M.* Neural excitability, spiking and bursting // *Int. J. Bifurc. Chaos*. 2000. V. 10, № 6. P. 1171-1266.
15. *Eguia M. C., Rabinovich M. I., Abarbanel H. D. I.* Information transmission and recovery in neuron communication channels // *Phys. Rev. E*. 2000. V. 62. P. 7111-7122.
16. *Brenner N., Strong S. P., Koberle R., Bialek W., de Ruyter van Steveninck R. R.* Synergy in a neural code // *Neural Computation*. 2000. V. 12. P. 1531-1552.
17. *Haken H.* Principles of Brain functioning. A synergetic approach to Brain activity, behavior and cognition. Springer-Verlag, Berlin, 1996.
18. *Nekorkin V. I., Velarde M. G.* Synergetic of active lattice systems. Berlin, Springer-Verlag, 2002.
19. *Babloyantz A., Lourenco C.* Computation with chaos. A paradigm for cortical activity // *Proc. Nat. Acad. Sci. USA*. 1994. V. 91. P. 9027-9031.
20. *Kazantsev V. B.* Selective communication and information processing by excitable systems // *Phys. Rev. E*. 2001. V. 64. P. 056210.
21. *Казанцев В. Б., Некоркин В. И., Велардэ М. Г.* Модель нейрона с осцилляторной активностью ниже порога возбуждения // *Изв. вузов Радиофизика* 1998. Т. 41, № 12. С. 1593-1603.
22. *Nekorkin V. I., Kazantsev V. B., Velarde M. G.* Spike-burst and other oscillations in a system composed of two coupled, drastically different elements // *Eur. Phys. J. B*. 2000. V. 16. P. 147-155.
23. *Velarde M. G., Nekorkin V. I., Kazantsev V. B., Makarenko V. A., Llinas R.* Modeling Inferior Olive Dynamics // *Neural Networks*. 2001. V. 15. P. 5-10.
24. *Papko V. V., Kazantsev V. B., Nekorkin V. I., Prestunov V. A., Makarenko V. I., Llinas R.* The electronic circuit modeling inferior olive neuron // *Phys. Rev. E*. 2002 (submitted).
25. *Kazantsev V. B., Nekorkin V. I., Binczak S., Bilbault J. M.* Spiral bifurcation set for nonlinear waves in excitable medium // *Phys. Rev. Lett*. 2002 (submitted).
26. *Zimmerman M. G., Firlé S. O., Natiello M. A., Hilderbrand M. E., Eiswirth M., Bar M., Bangia A. K., Kevrekidis I. G.* Pulse bifurcation and transition to spatiotemporal chaos in an excitable reaction-diffusion model // *Physica D*. 1997. V. 100. P. 92-104.
27. *Or-Guil M., Krishnan J., Kevrekidis I. G., Bar M.* Pulse bifurcation and instabilities in an excitable medium // *Phys. Rev. E*. 2001. V. 64. P. 046212.
28. *Nekorkin V. I., Makarov V. A., Kazantsev V. B., Velarde M. G.* Spatial disorder and pattern formation in lattices of coupled bistable systems // *Physica D*. 1997. V. 100. P. 330-342.

ПРОГНОЗ КАЧЕСТВЕННОГО ПОВЕДЕНИЯ ДИНАМИЧЕСКОЙ СИСТЕМЫ ПО ХАОТИЧЕСКОМУ ВРЕМЕННОМУ РЯДУ

Я. И. Мольков, А. М. Фейгин

Введение

Мы обсудим общий подход к прогнозу качественного поведения динамической системы (ДС), основанный на нелинейно-динамическом анализе порожденного системой временного ряда (ВР). Возможность построения такого прогноза *без* привлечения априорной информации о нелинейно-динамических свойствах системы, в том числе – без построения модели «из первых принципов», была осознана в самые последние годы [1]. Предложенные к настоящему времени алгоритмы, реализующие общий подход, предназначены для анализа «высококачественных» (достаточно протяженных и слабо зашумленных) ВР и, в силу этого, испытаны только с помощью различных компьютерных моделей динамических систем [1–3]. Тем не менее, у нас нет сомнения в возможности усовершенствования алгоритмов до уровня, необходимого для прогнозирования поведения ДС по сильно зашумленным ВР, что позволяет надеяться на построение прогностических моделей реальных систем. «Идеологические» основания такой уверенности содержатся в изложенном ниже общем подходе. Мы опишем также наиболее «продвинутой» на сегодняшний день «универсальный» алгоритм построения прогноза (на основе моделей в виде искусственных нейронных сетей), не накладывающий принципиальных ограничений на размерность моделируемой системы. Возможности данного алгоритма будут продемонстрированы на примере построения прогноза качественного поведения по хаотическому временному ряду, сгенерированному системой Макея – Гласса.

1. Байесов подход к построению математических моделей по данным измерений

Предлагаемый подход к прогнозу качественного поведения ДС состоит в построении и исследовании особых *прогностических* математических моделей изучаемой системы. Наряду со специфической для реконструкции *динамических* систем техникой, которая основана на знаменитых «теоремах о вложениях», доказанных около двадцати лет назад Такенсом [4], подход содержит в качестве обязательной компоненты отыскание распределения вероятностей параметров прогностической модели. Мы начнем с обсуждения именно этой («статистической») компоненты. Такая последовательность изложения позволит поставить прогноз качественно-го поведения ДС в контекст общей задачи математического моделирова-

ния и описать корректный подход к отысканию значений параметров модели, «оптимальных» с точки зрения ее адекватности изучаемой системе.

При решении общей задачи моделирования подразумевается, что модель снабжена некоторым фиксированным набором параметров, варьированием которых она возможно лучше совмещается с данными эксперимента. В одних случаях модель представляет собой ряд функций определенного класса (например, полиномов), а параметры являются коэффициентами этого ряда. В других вид модели может диктоваться лежащими в основе теоретическими представлениями («первыми принципами»), которым экспериментальные данные предположительно удовлетворяют. Количественную характеристику близости данных эксперимента и модели называют ценовой функцией (ЦФ), маленькие значения которой отвечают хорошему приближению, а большие – плохому. Задача моделирования тем самым сводится к минимизации ценовой функции по параметрам модели.

Хорошо известной и наиболее широко используемой ЦФ является среднеквадратичная невязка χ^2 :

$$\chi^2 = \sum_{k=1}^K (y_k - f(x_k, \mu))^2. \quad (1)$$

Здесь модель $f(x_k, \mu)$ предположительно связывает K пар результатов измерений $\{x_k, y_k\}_{k=1}^K$. Подгонка модели к экспериментальным данным состоит в отыскании значений параметров $\mu = \{\mu_m\}_{m=1}^M$, минимизирующих ЦФ (1). Данный способ поиска оптимальных значений параметров модели носит название метода наименьших квадратов (МНК).

Однако возникает принципиальный вопрос: является ли среднеквадратичная невязка «универсальной» ЦФ, минимизация которой обеспечивает наибольшую близость модели к экспериментальным данным во всех возможных ситуациях? Мы попытаемся дать конструктивный ответ на этот вопрос, т. е. предложить универсальный способ отыскания адекватной ЦФ для любой возникающей ситуации.

При анализе случая общего положения мы должны учесть, прежде всего, что реальные экспериментальные данные всегда содержат шумовую составляющую (вследствие, например, конечной точности измерений). Следовательно, измеренные значения отличаются от тех, которые были сгенерированы изучаемой системой, и эти отличия являются случайными величинами. Поэтому вопрос, на который нужно ответить для отыскания адекватной ЦФ, может быть сформулирован следующим образом: *Какова вероятность того, что наблюдаемые данные эксперимента были произведены системой, описываемой предлагаемой моделью при определенных значениях ее (модели) параметров?*

Наиболее близким по постановке задачи является подход, предложенный Байесом (Bayes) [5, 6]. В кратком изложении он выглядит следующим образом. Допустим, что система, над которой производится эксперимент, обладает набором свойств (параметров), прямое измерение которых невозможно. Теорема, сформулированная Байесом, говорит, что апостериорное условное распределение вероятности этих *ненаблюдаемых* параметров (часто называемое правдоподобием) пропорционально произведению их априорного распределения вероятности и условного распределения вероятности полученных экспериментальных результатов.

Казалось бы, теорема Байеса отвечает прямо на поставленный вопрос. Однако необходимо отметить факт, часто обделяемый вниманием исследователей. А именно: Байес предполагает, что ненаблюдаемые являются свойствами самой *системы*, тогда как в задачах моделирования параметры характеризуют *модель*, вид которой выбирается произвольно. Другими словами, в канонических условиях применимости теоремы Байеса функционал $y_k = f(x_k, \mu)$ описывает *систему* и известен *a priori*, а при построении модели по экспериментальным данным реконструкция данного функционала является одной из *целей* исследования. Очевидно, что расхождение между моделью и системой («дефект модели» [7]) будет присутствовать всегда вследствие конечности набора параметров. Вопрос о влиянии данного эффекта не прост и довольно плохо изучен; способ, позволяющий оценить его влияние, мы обсудим ниже. Пока же будем считать, что дефектом модели можно пренебречь.

Таким образом, сделав некоторые предположения и используя теорему Байеса, можно получить оценку условного распределения вероятности параметров модели, которое может быть, в частности, использовано для отыскания «правильной» ЦФ для выбора оптимальных значений параметров. Кроме того, могут быть вычислены любые характеристики системы и их распределения, что дает возможность ответить на существенно более широкий круг вопросов. В случае (а это, как правило, так), когда аналитически получить результат невозможно, мощным численным средством построения распределений вероятности является метод Монте-Карло.

Формальная запись вышесказанного выглядит следующим образом. Пусть имеется некоторое свойство системы $\kappa = F[f(\cdot)]$, являющееся произвольным функционалом модели $f: A \rightarrow B$, предположительно связывающей результаты некоторых измерений $y \in B$ и $x \in A$. Допустим, что в нашем распоряжении имеется K таких пар $\{x_k, y_k\}_{k=1}^K$, измеренных с конечной точностью. Отметим, что характер распределений вероятности ошибок измерений может быть различным и должен определяться независимо. Пусть известно, что ошибки ξ и η (будем далее называть их

шумами) аддитивны и распределены в соответствии с плотностями вероятностей $w_x(\xi)$ и $w_y(\eta)$ соответственно. Тогда введем незашумленные значения переменных u и v (часто называемые латентными переменными), определяемые как $x_k = u_k + \xi_k$, $y_k = v_k + \eta_k$, которые, естественно, функционально связаны моделью $v_k = f(u_k)$.

Вычислим теперь плотность вероятности, с которой можно наблюдать для таких x именно такой набор y при условии, что система описывается моделью f и латентными переменными u :

$$p(x, y | f, u) = \prod_{k=1}^K w_x(x_k - u_k) w_y(y_k - f(u_k)). \quad (2)$$

Следует отметить, что если оценка параметров распределений шумов в свою очередь произведена с конечной точностью, то эти параметры также должны быть внесены в список ненаблюдаемых наряду с f и u .

Остается сделать последний шаг, применив теорему Байеса, в соответствии с которой правдоподобие или плотность вероятности ненаблюдаемых величин принимать конкретные значения f и u при условии таких результатов измерений есть не что иное, как

$$p(f, u | x, y) \propto p(f, u) p(x, y | f, u). \quad (3)$$

Данное выражение приобретает смысл, если под зависимостью от модели f понимается зависимость от конечного набора параметров μ , ее характеризующих, поэтому на данном этапе необходимо конкретизировать модель. Как уже отмечалось, шаг этот является произвольным, и вопрос о его влиянии на конечный результат мы пока оставим без комментариев. Априорное распределение вероятности $p(f, u) := p(\mu, u)$ является константой в отсутствие априорных предположений или сведений о параметрах модели и латентных переменных. Отметим еще раз, что многомерное распределение (3) должно включать в себя распределения оценок параметров распределений шумов, если таковые делались. Вспомним, что прикладное значение задачи состоит в оценке распределения вероятности для величины $k = F[f(\cdot)]$, которая может быть произведена на основании (3) методом Монте-Карло.

2. Некоторые канонические частные случаи

Пусть с ошибкой η определяются только значения величины y , и пусть случайная величина η является гауссовой и не зависимой между отсчетами, тогда в (2) $w_x(\xi) = \delta(\xi)$, $w_y(\eta) = (2\pi\sigma^2)^{-1/2} \exp(-\eta^2 / 2\sigma^2)$.

Правдоподобие (3) приобретет следующий вид:

$$\begin{aligned}
 p(\mu | x, y) &\propto \prod_{k=1}^K \exp \left\{ -\frac{(y_k - f(x_k, \mu))^2}{2\sigma^2} \right\} = \\
 &= \exp \left\{ -\sum_{k=1}^K \frac{(y_k - f(x_k, \mu))^2}{2\sigma^2} \right\}.
 \end{aligned} \tag{4}$$

Введем ЦФ $C(\mu) = \sum_{k=1}^K \frac{(y_k - f(x_k, \mu))^2}{2\sigma^2}$ и увидим, что $\min_{\mu} C(\mu)$ точно соответствует максимуму (4). Поскольку $C(\mu)$ с точностью до постоянного множителя совпадает со среднеквадратичной невязкой χ^2 (см. (1)), в данной ситуации адекватным способом поиска оптимальных значений параметров является МНК.

Пусть теперь ошибки измерения величин x не нулевые, но также гауссовы и независимы. Тогда получим правдоподобие в форме

$$p(\mu, u | x, y) \propto \prod_{k=1}^K \exp \left\{ -\frac{(y_k - f(u_k, \mu))^2}{2\sigma_y^2} \right\} \exp \left\{ -\frac{(x_k - u_k)^2}{2\sigma_x^2} \right\}, \tag{5}$$

максимум которого будет соответствовать минимуму ЦФ

$$C(\mu, u) = \sum_{k=1}^K \left(\frac{(y_k - f(u_k, \mu))^2}{2\sigma_y^2} + \frac{(x_k - u_k)^2}{2\sigma_x^2} \right). \tag{6}$$

Следует отметить, что в данном случае минимизация производится не только по параметрам модели, но и по латентным переменным. Этот подход называется обобщенным МНК [8].

И, наконец, последняя ремарка, касающаяся простейшего способа верификации моделей. Предположим, что $f(\cdot)$ действительно описывает моделируемую систему, а шумы действительно гауссовы и независимы как между собой, так и между отсчетами. Тогда ценовые функции, приведенные выше, являются суммами квадратов нормально распределенных независимых случайных величин с единичной дисперсией. Это означает, что невязка $\chi^2 \propto C$ является случайной величиной и распределена в соответствии с χ^2 -распределением. Это распределение выражается через неполную гамма-функцию $Q = \Gamma(v/2, \chi^2/2)$ и дает вероятность данной величине быть больше своего аргумента. Число степеней свободы « v » χ^2 -распределения составляет количество членов суммы за вычетом числа переменных, по которым производится оптимизация. Для хороших моделей данная характеристика Q не должна быть ни слишком маленькой, ни слишком большой. Отрицательный результат в проверке такой статистической гипотезы означает, что либо неудачно выбран класс моделей, либо не верны представления о природе шумов. Отметим, что данный резуль-

тат является точным только для ценовых функций, допускающих сведение к квадратичной форме по переменным, по которым производится минимизация. Это, в частности, означает, что модель должна линейно зависеть от своих параметров.

3. Особенности моделирования динамических систем

Динамической называют систему, будущее поведение которой однозначно определяется текущим состоянием, задаваемым конечным набором величин (часто называемых начальными условиями). Пространство состояний системы называют фазовым. Координаты вектора состояния в фазовом пространстве называют динамическими переменными системы. Размерность фазового пространства есть размерность системы. Эволюция системы определяется эволюционным оператором, который позволяет по текущему состоянию системы вычислить ее состояние в будущем.

С точки зрения построения математической модели по данным измерений, динамическая система (ДС) имеет две принципиальные особенности.

Во-первых, в канонической постановке в качестве экспериментальных данных для моделирования динамических систем имеются, как правило, результаты *пассивного* наблюдения в виде временного ряда *единственной* динамической переменной (скалярного временного ряда). То, что по таким данным возможно восстановить эволюцию системы в целом (реконструировать траекторию в фазовом пространстве системы), впервые продемонстрировано в работе [9], в которой было предложено использовать координаты с задержками для восстановления вектора состояния системы. Идея состояла в том, что набор достаточного числа значений имеющейся единственной динамической переменной с некоторым шагом по времени можно интерпретировать как вектор состояния системы. После того, как Такенс [4] доказал это утверждение (а также несколько других полезных теорем), и дал верхнюю оценку на количество необходимых фазовых координат, стало ясно, что задача моделирования ДС в приведенной постановке не является фантастикой. Данный подход к восстановлению фазового пространства системы получил название метода координат с задержками. Количество необходимых координат является размерностью вложения.

Во-вторых, целью моделирования в случае ДС является построение либо модели оператора эволюции, воспроизводящей восстановленную фазовую траекторию, либо модели самой этой траектории.

Существенное ограничение на класс систем, которые целесообразно моделировать таким образом, накладывает требование о пассивности наблюдения. Мы не можем при измерении отклика системы «предлагать» оператору эволюции начальное состояние по своему выбору, а должны довольствоваться наблюдением ее собственной динамики. Интуитивно

ясно, что объем информации о системе, содержащейся в исходном временном ряде, тем больше, чем больший объем в фазовом пространстве занимает аттрактор, являющийся предельным множеством для восстановленной фазовой траектории. По этой причине наибольший интерес представляют ДС, демонстрирующие динамический хаос, поскольку именно в этом режиме фазовое пространство ДС содержит «наибольшее» предельное множество, которое называют странным аттрактором.

Напомним, что хаотический тип поведения ДС впервые был обнаружен Лоренцем [10] при анализе простых моделей конвективной циркуляции в атмосфере. Это открытие породило большое число спекуляций о природе тех или иных физических процессов. Дело в том, что хаотические процессы, на первый взгляд, ничем не отличаются от случайных (стохастических). Для того чтобы доказать динамическую природу наблюдаемого процесса, необходимо применение методов, весьма чувствительных к точности данных и их объему. Для многих «природных» систем точность измерений их характеристик и протяженность ВР, достигнутые на сегодняшний день, недостаточны для однозначного вывода о хаотической или стохастической природе системы. Тем не менее, число утверждений о выявлении, на основе данных наблюдений, хаотического динамического поведения, имеет тенденцию к расширению по мере приложения современных методов нелинейно-динамического анализа к новым базам данных различной природы. Обнаружение такого поведения продекларировано для различных систем, определяющих протекание важнейших процессов в атмосфере и гидросфере Земли (эволюцию озонового слоя [11], поведение концентраций химических составляющих атмосферы в приземном слое воздуха [12], крупномасштабных вариаций поверхностной температуры тропических вод Тихого океана (явление Эль-Ниньо) [13, 14]), тектоническую активность [15], а также для различных систем в живых организмах [16, 17].

4. Алгоритм построения моделей ДС: общее описание

Как уже говорилось, объектом моделирования ДС является фазовая траектория или воспроизводящий ее оператор эволюции. Поэтому первый шаг алгоритма – восстановление по хаотическому ВР фазового пространства моделируемой ДС, для чего прежде всего необходимо определить размерность вложения. Для этого существует набор методов, на обсуждении которых мы не будем здесь подробно останавливаться. Упомянем лишь об одном из прямых методов, называемом методом фальшивых соседей [18]. Идея его состоит в том, что оператор эволюции должен быть однозначным, т. е. одному прообразу должен соответствовать единственный образ, или, как это часто интерпретируют, фазовая траектория не должна иметь самопересечений. Для этого, последовательно увеличивая размерность вложения при восстановлении фазового пространства, сле-

дят за тем, чтобы сегменты исходного ряда, дающие близкие образы в пространстве размерности D , давали близкие образы и в пространстве размерности $D + 1$. Если этого не происходит, значит, размерность недостаточна. Преимущество данного метода в том, что он прямо отвечает на вопрос о достаточности некоторой размерности вложения для построения модели, тогда как косвенные методы (например, вычисление корреляционной размерности с последующим применением оценки Такенса), как правило, приводят к ее завышению. Знание размерности фазового пространства (наименьшей размерности вложения) позволяет восстановить наблюдаемую фазовую траекторию ДС по временному ряду единственной динамической переменной – например, поясненным в разд. 3 методом фазовых координат с задержкой [19].

Рассмотрим теперь процедуру реконструкции оператора эволюции, воспроизводящего восстановленную по хаотическому ВР фазовую траекторию. Мы ограничимся обсуждением специфики построения моделей эволюционных операторов в виде точечных отображений (ТО) в фазовом пространстве системы (такие операторы описывают процессы с дискретным временем), опустим изложение процедуры извлечения из исходного ВР функции последования, которую должна воспроизводить модель в виде ТО. Для рассматриваемого в разд. 5, 6 случая слабо зашумленного ВР данная процедура (после восстановления фазовой траектории) вполне очевидна. (Вопрос о построении функции последования по сильно зашумленному ВР гораздо менее тривиален; его обсуждение выходит за рамки данной статьи). Существенные (специфические для ДС) сложности возникают при применении Байесова подхода для отыскания (по данным измерений) плотности вероятности параметров модели.

Поясним характер возникающей проблемы на простейшем примере одномерных отображений. Пусть известно, что скалярный ряд $\{x_k\}_{k=1}^K$ произведен некоторым одномерным ТО и содержит аддитивную стохастическую добавку, имеющую плотность вероятности $w(\xi)$. Тогда, введя латентные переменные u_k , получим

$$x_k = u_k + \xi_k, \quad (7)$$

$$u_{k+1} = f(u_k, \mu) \quad (8)$$

и, определив формально в (2) $y_k = x_{k+1}$, для априорной плотности вероятности можем записать:

$$p(x | \mu, u) = \prod_{k=1}^{K-1} w(x_k - u_k) w(x_{k+1} - f(u_k, \mu)).$$

К сожалению, данный результат является неверным: очевидно, что в случае, когда результаты эксперимента суть ВР, сгенерированный ДС, латентные переменные $\{u_k\}_{k=1}^{K-1}$ не являются независимыми (как это пола-

галось при записи выражения (2)); напротив, $u_k = f^k(u_0)$. Поэтому, как недавно указали авторы работы [20], вместо (2) априорную плотность вероятности наблюдать значения $\{x_k\}_{k=1}^K$, описывающие состояния ДС в моменты времени $\{t_k\}_{k=1}^K$, следует определить выражением

$$p(x | \mu, u) = \prod_{k=1}^K w(x_k - f^k(u_0, \mu)). \quad (9)$$

Выше u_0 – значение латентной переменной («не зашумленное» состояние системы) в начальный момент времени, $f^k(\cdot)$ – k -кратное (последовательно примененное k раз) ТО. Другими словами, «динамичность» системы уменьшает количество латентных переменных до одной. Казалось бы, данное обстоятельство кардинально облегчает (по крайней мере, технически) использование требуемой нам апостериорной плотности вероятности

$$p(\mu, u_0 | x) \propto p(\mu, u_0) \times p(x | \mu, u_0) \quad (10)$$

параметров модели и латентных переменных, в том числе существенно упрощает отыскание (с помощью следующей из (9), (10) ЦФ) их наиболее вероятных значений. Можно, однако, понять, что для хаотической реализации выражение (10) для достаточно больших K (с учетом нормировки) будет отлично от нуля на множестве значений параметров меры нуль. Поясним данное утверждение, предположив, что погрешность измерения ξ_k представляет собой белый гауссов шум с дисперсией ε^2 и нулевым средним. Тогда из (9), (10) получаем

$$p(\mu, u_0 | x) \propto \exp\left(-\sum_{k=1}^K (x_k - f^k(u_0; \mu))^2 / 2\varepsilon^2\right). \quad (11)$$

Очевидно, что при $K \rightarrow \infty$ правдоподобие (11) стремится к δ -функции: лишь при $\mu = \mu_0$, отвечающем *точному* совпадению модели и моделируемой ДС, и начальном (латентном) состоянии u_0 , расположенном *точно* на «истинной» фазовой траектории, сумма в показателе экспоненты будет ограниченной по величине. Все остальные значения μ и u_0 зададут фазовую траекторию, не совпадающую с «истинной». Из-за экспоненциально быстрого «разбегания» траекторий на хаотическом аттракторе, который соответствует исходному ВР, в этих случаях плотность вероятности (11) будет с ростом K стремиться к нулю. Таким образом, воспользоваться корректным выражением (11) для плотности вероятности в интересующем нас случае хаотического ВР технически невозможно.

Мы можем предложить два способа обойти указанную трудность. Первый из них развивает идею авторов [21] о модификации обобщенного МНК (6) с целью отыскания ЦФ, приводящей к более точной оценке па-

раметров модели, построенной по зашумленному хаотическому ВР. Идея модификации состоит в использовании хаотичности ВР: из-за экспоненциально быстрого «разбегания» фазовых траекторий отсчеты, разделенные временным интервалом порядка λ^{-1} (обратной величины старшего показателя Ляпунова), можно считать независимыми. Модифицированная ЦФ получается разбиением ВР на последовательные фрагменты длины l , $l \leq \lambda^{-1}$, и предварительной минимизации среднеквадратичной невязки вида (11) по латентной переменной в пределах каждого отдельного фрагмента:

$$C(\mu) = \sum_{s=0}^{(K/l)-1} \min_{u_s} \sum_{j=1}^l (x_{s/l+j} - f^j(u_s; \mu))^2. \quad (12)$$

Применение такого способа в рамках Байесова подхода встречается, однако, принципиально обусловленные технические сложности и требует отдельного обсуждения.

Второй способ, впервые использованный в работе [20], предлагает учесть *стохастичность* ДС, т. е. присутствие шумовой компоненты η в *операторе эволюции* системы. Соответствующий шум часто называют динамическим (или собственным, или интерактивным). Полагая его аддитивным, мы должны вместо (7), (8) записать

$$x_k = u_k + \xi_k, \quad (13)$$

$$u_{k+1} = f(u_k, \mu) + \eta_k. \quad (14)$$

После этого, считая шум измерений ξ и динамический шум η независимыми, для априорной плотности вероятности получаем выражение

$$p(x | \mu, u) = \prod_{k=1}^K w_{\xi}(x_k - u_k) w_{\eta}(u_k - f(u_{k-1}, \mu)), \quad (15)$$

откуда с помощью теоремы Байеса находим требуемую апостериорную плотность вероятности

$$p(\mu, u | x) \propto p(\mu, u) p(x | \mu, u) \quad (16)$$

того, что наблюдаемые x являются зашумленными значениями латентных переменных u , сгенерированными при значениях параметров μ .

Отметим, что знание плотностей вероятности шумов (динамического и измерений) дает возможность выделить все режимы поведения наблюдаемой ДС, которые могли породить исходный ВР, и рассчитать вероятность того, что ВР порожден каждым из выявленных режимов. Для этого нужно сгенерировать с помощью модели $f(u, \mu)$ достаточно большой («статистический») ансамбль ВР, отвечающих различным наборам параметров μ и латентных переменных u , распределенных в соответствии с плотностью вероятности (16). Анализ каждого ВР из данного ансамбля позволяет выявить все «подходящие» режимы поведения системы и рассчитать их вероятности.

В заключение данного раздела обсудим вопрос о погрешностях, вносимых самим алгоритмом построения модели ДС по хаотическому ВР. Очевидно, что все три входящие в алгоритм операции: восстановление фазовой траектории, извлечение из фазовой траектории функции последования и, наконец, собственно построение модели (аппроксимация функции последования аналитически записанным ТО) – влекут за собой погрешности. Погрешности первых двух операций обусловлены дискретностью и конечной протяженностью исходного ВР; нетрудно понять, что они играют точно такую же роль, как шум измерений ξ (см. (13)). Ясно также, что корректной является трактовка данных погрешностей как случайных величин. Погрешность третьей операции мы уже упоминали (в разд. 1), назвав ее, вслед за авторами [7], дефектом модели, характеризующим неизбежную неадекватность модели моделируемой системе. Очевидно, что дефект модели может быть описан как аддитивная поправка к «незашумленному» оператору эволюции системы $f(u, \mu)$ и учтен, аналогично динамическому шуму, в виде слагаемого в правой части (14). Ясно, однако, что дефект модели, вообще говоря, *не* является случайной величиной; какие-либо попытки оценить его влияние на качество моделирования нам неизвестны. Можно, тем не менее, ожидать, что если дефект модели существенно меньше других обсуждавшихся погрешностей, то его влиянием можно пренебречь. Понятно также, что дефект модели определяет предельный уровень фильтрации шумов. Основываясь на приведенных качественных соображениях, мы будем далее оценивать дефект модели, трактуя его как динамический шум [3].

5. Прогностические модели динамических систем на основе искусственных нейронных сетей [3]

Обсудим теперь постановку задачи прогноза качественного поведения ДС по скалярному хаотическому ВР.

Прежде всего, мы ограничимся анализом ситуации, когда изменение поведения системы является следствием регулярных изменений внешних по отношению к системе условий, приводящих к регулярной зависимости от времени (трендам) ее управляющих параметров. Кроме того, мы рассмотрим случай слабо нестационарных ВР, порожденных слабо неавтономными ДС. В этом случае характерный временной масштаб трендов параметров τ_p намного превышает как характерное время τ_v эволюции динамических переменных, так и «длину» ВР T , т. е. выполняются неравенства

$$\tau_p > (\gg) T \gg \tau_v. \quad (17)$$

Наш интерес к ситуации (17) объясняется тремя обстоятельствами. Во-первых, существенное различие временных масштабов позволяет дос-

таточно надежно установить факт нестационарности ВР. Во-вторых, существующие сегодня методы определения размерности ДС (наименьшей размерности вложения) по ВР применимы только в случае слабо нестационарных ВР. В-третьих, слабая неавтономность характерна для многих «природных» ДС.

В условиях (17) мы можем представить зависимость параметров модели от времени $\mu(t)$ в виде ряда Тейлора и ограничиться первыми двумя членами разложения. Очевидно, что после этого оператор эволюции $f(\cdot)$ можно записать в виде

$$f(u_{k-1}, \dots, u_{k-N}; \mu(t)) = f_0(u_{k-1}, \dots, u_{k-N}; \mu_0) + \beta \cdot t_{k-1} \cdot f_1(u_{k-1}, \dots, u_{k-N}; \mu_1). \quad (18)$$

Здесь, в отличие от (14), рассматривается общий случай ТО N -го порядка; параметры модели $\{\mu_0, \mu_1, \beta\} \in \mathcal{R}^{\{M_0+M_1+\beta\}}$ не зависят от времени.

Линейная аппроксимация зависимости оператора эволюции от времени, использованная в (18), является адекватной при анализе почти всех слабо нестационарных ВР. Исключением является случай, когда количественные характеристики наблюдаемого аттрактора, изменяющиеся вследствие неавтономности системы, достигают своих экстремальных значений *внутри* исходного ВР. В этом случае в разложении зависимости $\mu(t)$ в ряд необходимо учесть члены второго порядка малости. Такая ситуация имеет, например, место при построении прогностической модели мезосферной фотохимической системы по ВР, отвечающему так называемому «внешнему» хаотическому аттрактору данной системы [1–3, 22].

Как уже говорилось, предложенные к настоящему времени алгоритмы позволяют строить прогностические модели на основе достаточно протяженных, измеренных с малым дискретом и слабо зашумленных в процессе измерений ВР. В соответствии с введенной выше терминологией сказанное означает, что в этом случае из всех обсуждавшихся погрешностей при построении модели следует учитывать только динамический шум. Другими словами, мы должны положить в (13) $\xi = 0$ (точнее, считать плотность вероятности шума измерений δ -функцией: $w(\xi) = \delta(\xi)$), так что из (15) с учетом (16) получаем для условной плотности вероятности параметров модели выражение

$$p(\mu | x) = \prod_{k=N}^K w_\eta(x_k - f(x_{k-1}, \dots, x_{k-N}, \mu(t))), \quad (19)$$

где w_η есть плотность вероятности динамического шума, в который включен, как уже было сказано, и дефект модели. Поэтому выражение (19) полностью адекватно случаю, когда ВР сгенерирован компьютерной моделью ДС и основная погрешность обусловлена дефектом модели. Всюду далее мы будем считать, что случайная величина η является гаус-

совой и независимой между отсчетами, тогда $w_{\eta} = (2\pi\sigma^2)^{-1/2} \exp(-\eta^2/2\sigma^2)$ и (19) переписывается в виде

$$p(\mu | x) \propto \exp \left\{ - \sum_{k=N}^K [x_k - f(x_{k-1}, \dots, x_{k-N}, \mu)]^2 / 2\sigma^2 \right\}. \quad (20)$$

Выше при записи (19), (20) мы учли отсутствие шума измерений, положив $u_k = x_k$.

Таким образом, в случае аддитивного динамического шума мы приходим к каноническому случаю, когда адекватной является ЦФ в виде среднеквадратичной невязки, а оптимальные значения параметров следует искать с помощью МНК (см. разд. 2).

Следующий очень существенный вопрос заключается в выборе вида модели $f(\cdot)$. Для того, чтобы алгоритм позволял строить прогностические модели различных (в том числе – высокоразмерных) ДС, функции $f(\cdot)$ и $f(\cdot)$ в (18) должны давать возможность аппроксимировать практически любую однозначную функцию произвольного числа переменных с любой наперед заданной точностью. Такой универсальностью обладают искусственные нейронные сети [23], позволяющие повышать точность аппроксимации простым увеличением числа нейронов, без изменения вида функции. Еще одним полезным свойством нейронной сети (НС) является ее слабая чувствительность к изменению аргументов вне «области обучения» (области изменения аргументов функции в пределах исходного ВР), что обеспечивает глобальную устойчивость получаемых моделей. В работах [3, 22] мы использовали в качестве функций $f_0(\cdot)$ и $f_1(\cdot)$ персептрон – простейшую трехслойную НС [24], линейно входящие параметры которого искались, в соответствии с (18), в виде линейных функций времени:

$$f(x_{k-1}, \dots, x_{k-N}; t_{k-1}; \mu(t)) = \sum_{i=1}^m (\alpha_i + t_{k-1}\beta_i) \tanh \left(\sum_{j=1}^N w_{ki} x_{k-j} + \gamma_i \right). \quad (21)$$

Здесь m – число нейронов во внутреннем слое сети, $\{\alpha, \beta, w, \gamma\} \in \mathfrak{R}^{(N+3)m}$ – полный набор параметров модели, и считалось $u_k = x_k$ (учтено отсутствие шумов измерений).

Предложенная в [3] процедура обучения (отыскания оптимальных значений параметров) нейронной сети (21) также явно использует медленность зависимости оператора эволюции от времени. Процесс обучения проводится в два этапа. На первом ищутся значения параметров $\tilde{v}, \tilde{w}, \tilde{\gamma}$, обеспечивающие минимум «автономной» среднеквадратичной невязки

$$\chi_a^2 = \left\langle \left[x_k - \sum_{i=1}^m v_i \cdot \tanh \left(\sum_{j=1}^N w_{ji} x_{k-j} + \gamma_i \right) \right]^2 \right\rangle_k. \quad (22)$$

где $\langle \cdot \rangle_k$ имеет смысл среднего по соответствующему индексу. На втором этапе фиксируются значения параметров $w = \tilde{w}$, $\gamma = \tilde{\gamma}$ и находятся значения параметров $\bar{\alpha}$, $\bar{\beta}$, соответствующие минимуму «неавтономной» среднеквадратичной невязки

$$\chi_n^2 = \left\langle \left[x_k - \sum_{i=1}^m (\alpha_i + t_{k-1} \cdot \beta_i) \cdot \tanh \left(\sum_{j=1}^N \tilde{w}_{ji} x_{k-j} + \tilde{\gamma}_i \right) \right]^2 \right\rangle_k. \quad (23)$$

Поведение полученной в результате обучения модели

$$x_k = \sum_{i=1}^m \bar{v}_i(t) \cdot \tanh \left(\sum_{j=1}^N \tilde{w}_{ji} x_{k-j} + \tilde{\gamma}_i \right) + \eta_k. \quad (24)$$

$$\bar{v}_i(t) = \bar{\alpha}_i + \bar{\beta}_i \cdot t \quad (25)$$

определяется *единственным* управляющим параметром – временем t ; при этом (25) описывает наиболее вероятный тренд «неавтономных» параметров системы.

Обсудим теперь вопрос об использовании построенной модели для прогноза качественного поведения ДС. Нетрудно понять, что «естественный» путь – экстраполяция наиболее вероятного тренда (25) за пределы исходного ВР и анализ поведения модели (24) при «будущих» значениях параметров – недостаточен. Наибольшая вероятность тренда не означает, вообще говоря, что отвечающий ему прогноз наиболее вероятен. Более того, ответ на вопрос о вероятностях предсказываемых режимов поведения ДС путем анализа *единственного* тренда вообще невозможен. Чтобы ответить на данный вопрос, необходимо исследовать *структуру пространства параметров* модели в окрестности экстраполированного наиболее вероятного тренда. Для этого следует применить Байесов подход в полном объеме. Мы реализуем его с помощью метода Монте-Карло: построим ансамбль трендов «неавтономных» параметров $v_i(t) = \alpha_i + \beta_i \cdot t$, распределенных в каждый момент времени в «будущем» в соответствии с функцией распределения (20), и исследуем качественное поведение прогностической модели (20), (21), отвечающее каждому из входящих в ансамбль трендов. Такой анализ позволяет получить статистические ансамбли моментов времени предсказываемых бифуркаций, после чего нетрудно рассчитать вероятностные характеристики прогноза: зависимости от времени вероятности обнаружить у ДС в заданный момент времени определенный тип поведения, плотности вероятности и интегральной вероятности момента конкретной предсказываемой бифуркации и пр.

6. Пример: прогноз качественного поведения системы Макея–Гласса

В этом разделе мы кратко опишем результаты построения прогностической модели и прогноза бифуркаций по слабо нестационарному хаотическому ВР, сгенерированному одной из так называемых «систем с задержкой», позволяющих моделировать с помощью систем обыкновенных дифференциальных уравнений процессы, протекающие в ряде распределенных динамических систем. Мы рассмотрим систему Макея–Гласса, моделирующую процесс регенерации лейкоцитов у больных хронической лейкемией [25]:

$$\frac{dx(t)}{dt} = -b \cdot x(t) + f(x(t-\tau)), \quad (26)$$

$$f(x) = a \cdot x / (1 + x^n). \quad (27)$$

Поведение данной системы определяется параметрами a , b , n и τ . Всюду ниже для первых трех из них были выбраны следующие значения: $a = 0,2$; $b = 0,1$; $n = 10$. В качестве управляющего мы использовали задержку τ , фигурирующую в правой части уравнения (26). Увеличение данного параметра в пределах от $\tau = 4,53$ до $\tau = 100$ последовательно приводит систему от периодических осцилляций с периодом $T = 3\tau$ (в интервале $4,53 < \tau < 13,3$) через каскад удвоения периода в режим хаотических осцилляций, возникающий при $\tau = 16,8$. При дальнейшем росте задержки система демонстрирует хаотические осцилляции, прерываемые окнами регулярного поведения, причем корреляционная размерность хаотического аттрактора d в среднем растет: от $d = 1,95$ при $\tau = 17$ до $d = 7,50$ при $\tau = 100$ [26].

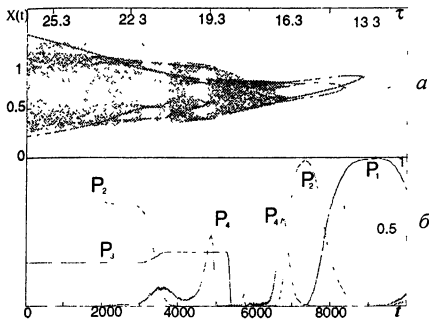
Для построения прогностической модели использован ВР $x(t)$, на протяжении которого восстановленная фазовая траектория 1000 раз пересекала секущую Пуанкаре. В качестве последней выбирался ноль второй фазовой координаты $x^{(2)}(t) = x(t + \Delta t)$, величина Δt была равна 9. В пределах ВР параметр τ менялся (по линейному закону) от $\tau = 18,4$ до $\tau = 17,2$. На всем протяжении исходного ВР система демонстрировала хаотические осцилляции; корреляционная размерность соответствующего аттрактора была близка к 2, а наименьшая размерность вложения $d_E = 3$.

Построение прогноза по данному ВР было произведено с помощью алгоритма на основе нейронных сетей, описанного в разд. 5. Для реконструкции оператора эволюции использовалась функция последования третьего порядка

$$x_k = f(x_{k-1}, x_{k-2}, x_{k-3}; \mu(t)) + \eta_k, \quad (28)$$

аппроксимация которой была произведена нейронной сетью (21) с тремя входами ($N = 3$) и пятью нейронам во внутреннем слое ($m = 5$).

Результаты прогноза качественного поведения системы (26), (27) вне пределов исходного ВР – в «прошлом» при $t \in [0; 5000]$ и «будущем» при $t \in [6000; 10000]$ – приведены на рисунке (исходный ВР принадлежит временному отрезку $t \in [5000; 6000]$). На рисунке *а* серыми точками показана бифуркационная диаграмма, полученная при экстраполяции за пределы исходного ВР наиболее вероятного тренда (25). На рисунке *б* приведены зависимости от времени $P_k(t)$ вероятностей обнаружения предсказанных режимов поведения, найденных в результате обработки статистического ансамбля предсказанных режимов.



Система Макея–Гласса. результаты прогноза *а* – бифуркационная диаграмма, соответствующая среднему тренду (серые точки) и «правильная» бифуркационная диаграмма (черные точки); *б* – зависимости от времени $P_k(t)$ вероятностей обнаружения предсказанных режимов поведения

Для построения такого ансамбля было рассчитано и проанализировано 1000 трендов, параметры которых в каждый момент времени были распределены в соответствии с плотностью вероятности (20). Входящая в (20) величина σ^2 выбиралась равной нормированному (на разность числа отсчетов в исходном ВР и числа параметров модели) минимальному значению неавтономной невязки (23):

$$\sigma^2 = (\chi_n^2)_{\min} / [K - (N + 3)m]. \quad (29)$$

Кроме того, на рисунке *а* черными точками построена «правильная» бифуркационная диаграмма, полученная прямым расчетом эволюции системы (26), (27) во всем анализируемом временном интервале $t \in [0; 10000]$; управляющий параметр τ изменялся в этом интервале (см. верхнюю горизонтальную ось на рисунке *а*) по тому же закону, что и в пределах исходного ВР.

Мы видим, что предложенный алгоритм позволил правильно предсказать «будущий» выход из хаоса через обратный каскад удвоения вплоть до возникновения осцилляций на основной частоте. Примечательно, что максимальная вероятность обнаружить предсказанные режимы увеличивается при удалении в «будущее» (см. рисунок *б*). Объяснение данного «парадокса» почти тривиально: максимальная вероятность обна-

ружения режима определяется, во-первых, максимальным значением интегральной вероятности входа в режим и, во-вторых, степенью перекрытия во времени плотностей вероятности бифуркаций входа в режим и выхода из него. В рассматриваемом случае максимальные интегральные вероятности входа во все три «будущих» режима оказываются равными единице, а степень перекрытия плотностей вероятностей входа и выхода тем больше, чем короче «продолжительность жизни» режима. Самым короткоживущим является ближайший к границе исходного ВР 4-периодический режим, самым долгоживущим – наиболее удаленный в будущее режим осцилляций на основной частоте, что и объясняет обнаруженный парадокс.

Что касается прогноза в «прошлое», мы видим правильно предсказанный выход из хаоса через перемежаемость в 4-периодический режим с последующим переходом (через каскад удвоения периода) в режим хаотических осцилляций. Отметим при этом, что предсказанная топологическая структура возникающего в «прошлом» хаотического аттрактора отличается от структуры «правильного» хаотического аттрактора системы Макея–Гласса, отвечающего соответствующей области в ее пространстве параметров. Дальнейший прогноз в «прошлое» полностью ошибочен: вместо чередования узких областей регулярного и хаотического поведения, завершающегося переходом в широкую область хаотических осцилляций, модель предсказывает режимы двух и трех периодических осцилляций. Причина данной ошибки – рост размерности аттрактора при увеличении задержки τ , происходящий при $\tau \geq 21$. Очевидно, что имеется фундаментальный запрет на прогноз более сложного (более высокоразмерного) поведения по более простому.

Заключение

(об ограничениях и дальнейших возможностях)

1. Принципиальным для применения изложенного подхода к прогнозу качественного поведения ДС является *предсказание бифуркаций более сложного поведения в более простое*: размерность аттракторов, возникающих в фазовом пространстве ДС в результате бифуркаций, не должна превышать размерности аттрактора, отвечающего исходному ВР. Достаточно очевидно, например, что описанная общая процедура не позволяет предсказать каскад бифуркаций удвоения периода в случае, когда исходный временной ряд содержит информацию о простейшем для такого каскада типе поведения. Другой пример подобного рода обсуждался нами в разд. 6. Подчеркнем, что предложенный алгоритм позволяет преодолевать окна регулярного поведения между хаотическими аттракторами, однако использование для построения прогностической модели наиболее сложного типа поведения имеет решающее значение.

2. Предложенный алгоритм на основе искусственных нейронных сетей, обладая такими важными достоинствами, как универсальность по отношению к реконструкции оператора эволюции системы произвольной размерности и глобальная устойчивость, может ограничивать возможности прогноза. Это происходит в случае, когда характерный размер хаотического аттрактора увеличивается в направлении течения времени, интересующем нас при построении прогноза. В такой ситуации качество прогноза с помощью «нейронно-сетевой» модели ухудшается из-за малой чувствительности сети к изменениям аргументов *вне* области ее обучения. В этом случае, как отмечается в [3, 22], более адекватными могут оказаться модели, сочетающие нейронную сеть в качестве автономной компоненты прогностической модели (18), и систему других независимых функций (например, многомерных полиномов) в качестве неавтономной компоненты.

3. Предложенный подход к построению прогностических моделей дает принципиальную возможность верифицировать существующие модели из первых принципов, на основании которых производятся современные оценки будущей эволюции различных динамических систем. Несовпадение нелинейно-динамических свойств прогностической модели и модели из первых принципов будет означать необходимость корректирования последней. Последующее сравнение верифицированной модели из первых принципов с прогностической моделью позволит сопоставить динамическим переменным и неавтономным параметрам последней реальные физические, химические и прочие характеристики моделируемой системы. В результате можно рассчитывать на (а) выделение «главных» динамических переменных, которые определяют эволюцию наблюдаемой системы; (б) выяснение, тренды каких именно параметров реальной системы могут привести к предсказанным бифуркациям; (в) оценку моментов бифуркаций и их количественных последствий.

Работа выполнена при поддержке РФФИ (проект 02-02-17080), Комплексной программы научных исследований Президиума РАН «Нелинейная динамика и солитоны» (проект 3.3) и Министерства промышленности, науки и технологий РФ (контракт № 40.020.1.1.1168).

Литература

- 1 *Фейгин А М, Мольков Я И, Мухин Д Н, Лоскутов Е М.* Прогноз бифуркаций слабеавтономных динамических систем на основе наблюдаемых временных рядов *Препринт ИПФ РАН, № 508. 1999 53 с*
- 2 *Фейгин А М, Мольков Я И., Мухин Д Н, Лоскутов Е М.* Прогноз качественного поведения динамической системы по хаотическому временному ряду // *Изв. вузов Радиофизика. 2001. Т 44, № 5–6 С 376–399.*
- 3 *Feigin A M, Molkov Y I, Mukhin D. N and Loskutov E. M* Investigation of nonlinear dynamical properties by the observed complex behaviour as a basis for construction of the dynamical models of atmospheric photochemical systems // *Faraday discussions 2002 V 120 P 105–123*

4. *Takens F* Detecting strange attractor in turbulence In D A Rand and L-S Young (Eds.) // *Dynamical Systems and Turbulence*. Warwick, 1980. Lecture Notes in Mathematics. Berlin: Springer, 1981 V 898. P 366
5. *Корн Г и Корн Т* Справочник по математике М. Наука, 1970 484 с
6. *Боев Г П* Теория вероятностей М Гостехтеориздат, 1950. 368 с
7. *Бутковский О Я, Кравцов Ю А, Логунов М Ю.* Анализ погрешностей восстановления параметров нелинейного отображения по зашумленным хаотическим временным рядам // *Изв вузов Радиофизика* 2002 Т 45, № 1 С 55–66.
8. *Press W H, Teukolsky S A, Vetterling W T and Flannery B P* Numerical Recipes in C (The Art of Scientific Computing) Second Edition Cambridge University Press, 1988 874 p.
9. *Packard N H, Crutchfield J. P., Farmer J D and Shaw R S* Geometry from a time series // *Phys Rev Lett.* 1980 V 45 P 712
10. *Lorenz E M* Deterministic nonperiodic flow // *Journal of the Atmospheric Science* 1963. V 20 P 130
11. *Yang P, Brasseur G P, Gille J C, et al* Dimensionalities of ozone attractors and their global distribution // *Physica D* 1994 V 76 P 331
12. *Biswas P, Li I-F. and Islam S* Estimation of dominant degree of freedom for air pollutant concentration data application to ozone measurements // *Atmospheric Environment* 1994 V 28 P. 1707
13. *Wang B, Barcilon A and Fang Z* Stochastic dynamics of El Niño-Southern oscillation // *Journal of the Atmospheric Science* 1999 V 56 P 5.
14. *Saunders A, Ghil M.* A Boolean delay equation model of ENSO variability // *Physica D* 2001. V 160 P 54.
15. *Srivastava H N, Bhattacharya S N and Sinha Ray K C* Strange attractor characteristics of earthquakes in Shillong plateau and adjoining regions // *Geophys Res Lett* 1996 V 23 P 3519
16. *Abarbanel H D I, Huerta R., Rabinovich M. I, et al.* Synchronized action of synaptically coupled chaotic single neurons. I. Simulations using realistic model neurons // *Neural Comput* 1996. V 8 P 1567
17. *Frank G W, Lookman T, Nerenberg M A. H, et al* Chaotic time series analyses of epileptic seizures // *Physica D.* 1990. V 46 P 427
18. *Abarbanel H D I* Analysis of Observed Chaotic Data New York Springer-Verlag, 1997 272 p
19. *Анищенко В С, Вадивасова Т Е, Астахов В В* Нелинейная динамика хаотических и стохастических систем Саратов Изд. СГУ, 1999 368 с
20. *Meyer R and Christesen N* Bayesian reconstruction of chaotic dynamical systems // *Phys Rev E* 2000 V 62 P 3535
21. *Kantz H, Jaeger L* Improved cost function for modeling of noisy chaotic time series // *Physica D* 1997 V 109. P 59
22. *Фейгин А М* Нелинейно-динамические модели атмосферных фотохимических систем Дис на соискание уч ст д ф-м н // ИФФ РАН Нижний Новгород, 2002
23. *Hornik K., Stinchcombe M and White H.* Multilayer feedforward networks are universal approximators // *Neural Networks* 1989 V 2 P 359
24. *Arbib M A (Edit)* The Handbook of Brain Theory and Neural Networks The MIT Press, 1995
25. *Mackey M C and Glass L* Oscillation and chaos in physiological control systems // *Science* 1977 V 197 P 287
26. *Неймарк Ю И, Ланда П С* Стохастические и хаотические колебания М Наука, 1987 424 с

ДИНАМИЧЕСКИЙ ХАОС И ИНФОРМАЦИЯ

А. С. Дмитриев

Введение

Открытие и изучение динамического хаоса заставили пересмотреть многие привычные представления о процессах в окружающем нас мире. Одним из классов процессов, где роль хаоса и нелинейной динамики в целом чрезвычайно важна, являются информационные процессы в живых системах (мозге, сердечно-сосудистой системе, отдельных нейронных ансамблях). Исследование этих процессов, а также интерес к применению хаоса для решения таких классических задач теории информации [1–5], как обработка, передача сигналов, привлекают внимание к информационным свойствам самого динамического хаоса и порождающих его динамических систем (см., например, [6–9]).

Ниже хаотические колебания (хаос) рассматриваются как носитель информации, а системы, их генерирующие, как специфические источники информации. Важным свойством хаотических колебаний как информационных сигналов является то, что средний объем информации в единицу времени у них строго определен. Отсюда следует, например, что они могут быть переданы без искажений по каналу с ограниченной пропускной способностью. Для других видов аналоговых сигналов, содержащих информацию, например для речевых сигналов или белого шума, это невозможно.

Непосредственно информация, содержащаяся в хаотических сигналах, не является «полезной». В этом смысле образ хаотических колебаний в фазовом пространстве – хаотический аттрактор – может рассматриваться в качестве резервуара, наполненного некоторой неконтролируемой информацией. Типичными хранимыми «информационными блоками» в резервуаре являются неустойчивые предельные циклы, число которых в хаотическом аттракторе счетно. Теоретически объем информации, хранящийся в хаотическом аттракторе, неограничен. Можно ли сделать так, чтобы хотя бы часть этого хранилища соответствовала содержательной информации? Оказывается, это возможно, и более того, разработана теория, обеспечивающая построение таких «хранилищ» полезной информации с практически неограниченной емкостью.

Следующий вопрос, связанный с информационными свойствами хаотических траекторий звучит так: можно ли динамике произвольной динамической системы, не нарушая ее структуры сопоставить требуемую информационную последовательность? Сформулируем этот вопрос несколько иначе: нельзя ли хаотическую систему использовать в качестве кодера, преобразующего информационный сигнал в хаотический? В статье приводятся простые примеры, показывающие, что построение такого кодера вместе с соответствующим ему декодером возможно.

Простейшим типом взаимодействия хаотических сигналов с информационными и шумовыми сигналами является линейное взаимодействие. Различные варианты взаимодействия этих сигналов приводят к содержательным задачам, анализ которых показывает, что как методы, так и результаты решения этих задач с учетом информационных свойств имеют фундаментальные отличия от аналогичных задач, решаемых для нехаотических сигналов. В частности, это относится к задачам очистки хаотических сигналов от шума (фильтрации), синхронизации хаотических систем и разделению хаотических сигналов. Оказывается, что при малых уровнях шума задачи с хаотическими сигналами решаются значительно эффективнее, чем с обычными сигналами. Однако в этих задачах существует порог на уровень шума, при превышении которого качество решения задачи резко падает.

Примером нелинейного взаимодействия хаотических и информационных сигналов является модуляция информационным сигналом хаотического сигнала, или, в более широкой постановке, ввод информационного сигнала в хаотический сигнал и последующее извлечение информационного сигнала.

Как только речь заходит об обработке информации, возникает вопрос о связи нелинейной динамики и вычислительных процессов. Здесь полезно вспомнить о работах в области предельной эффективности вычислений и передачи информации (принцип Ландауэра и обратимые вычисления Беннета). Во-первых, в этой проблеме, как и в процессах обработки хаотических сигналов, возникают принципиальные ограничения (подобные второму закону термодинамики). Во-вторых, появляется возможность проанализировать общность явления генерации информации в вычислительных процессах с одной стороны и в динамических системах с хаосом с другой стороны. В третьих, заявленное в упомянутых работах единство процессов вычислений, измерений и передачи информации дает основание рассматривать вычислительный процесс (и его реализацию на компьютере) как «канал связи» во временной области. Необычным свойством такого типа «каналов связи» является то, что поток информации на его выходе может быть больше, чем поток информации на входе. В-четвертых, анализ проблемы предельной эффективности и предельной эффективности передачи информации проводился Ландауэром на примере двухъямного потенциала – вырожденного варианта нелинейного осциллятора, который представляет собой базовый элемент при исследовании явлений нелинейной динамики. Последнее обстоятельство наводит на мысль о едином корне обсуждающихся проблем. И тут мы подходим к вопросу о взаимодействии систем, производящих информацию с окружающей средой. Эти системы можно разделить на два класса: информационно-замкнутые (не обменивающиеся информацией с внешней средой) и информационно-открытые. Анализируются условия информационной

замкнутости и информационной открытости, обсуждается вопрос о связи информационной и термодинамической открытости систем.

1. Синтез хаотических систем с заданной информацией на периодических траекториях

Итак, каждую траекторию динамической системы можно рассматривать как некоторый информационный сигнал. Тем самым совокупность траекторий отображения представляет собой своеобразное «хранилище» информации в виде множества траекторий системы. Это «хранилище» обладает рядом интересных свойств, проявление которых зависит от того, имеются ли аттракторы в динамической системе и какого они типа.

Рассмотрим некоторые из этих свойств. Если в момент времени $t = t_0$ задать начальные условия, то формально траектория отображения известна и однозначно определена для всех $t > t_0$ (аналог теоремы Коши для отображений). Если траектории вдоль ее длины сопоставлен информационный сигнал, он может быть воспроизведен путем итерирования отображения с заданными начальными условиями, т. е. информация извлекается из памяти путем решения эволюционного уравнения.

«Хранилище» информации как динамическая память обладает естественными ассоциативными свойствами. Действительно, задавая любой фрагмент сигнала, можно воспроизвести все его последующие значения, а в случае циклического сигнала – и весь сигнал целиком.

Пусть в отображении имеется единственный аттрактор – неподвижная точка. Тогда траектория, задаваемая любыми начальными условиями из области определения отображения, сходится к неподвижной точке, а соответствующий ей информационный сигнал – к постоянному значению. Текущее информационное содержание такого сигнала по мере сходимости уменьшается. «Информационное содержание» самих устойчивых неподвижных точек в случае одномерного отображения – это значение единственной константы. Увеличение объема записанной информации может происходить только за счет точности задания координаты – точки равновесия, т. е. за счет улучшения разрешающей способности.

Если в отображении имеется аттрактор – устойчивый предельный цикл, то все траектории со временем стягиваются к нему. Периодическую траекторию, соответствующую предельному циклу, можно рассматривать как периодически воспроизводимый информационный сигнал. Объем информации в таком сигнале определяется длиной цикла и степенью его изрезанности (информационной насыщенностью).

Если в системе имеется единственный хаотический аттрактор, то траектория от итерации к итерации теряет информацию о начальных условиях и отображение производит собственную информацию. При небольшой скорости производства информации аттрактор остается хорошо локализованным в фазовом пространстве. Такие хаотические аттракторы

в принципе могут хранить достаточно большие объемы информации. Если же производство информации велико, то хаотический аттрактор как целое становится объектом с большой неопределенностью и как хранилище информации теряет свою привлекательность. Однако можно использовать в качестве хранилища информации его систему неустойчивых циклов. Действительно, хаотические аттракторы содержат счетное число периодических орбит [10–13]. Это вполне определенные орбиты, которые могут рассматриваться как траектории, соответствующие информационным сигналам. Однако информация, соответствующая таким сигналам, неконтролируема. Для того чтобы сделать часть циклических траекторий соответствующими содержательной информации, в работе [14] был предложен метод синтеза одномерных отображений, содержащих систему циклических траекторий с заданной структурой. На основе этого метода была продемонстрирована возможность записи информации на циклических траекториях, реализация ассоциативной памяти и других базовых операций [15–25].

Сжатие информации. В процессе записи производится специальное кодирование (ортогонализации) информационных последовательностей. Побочным результатом ортогонализации является обратимое (без потерь) сжатие информации – уменьшение общего объема данных, необходимых для полного и точного описания информационных объектов, например текстов или изображений. Если взглянуть на этот метод кодирования как на метод сжатия информации, то оказывается, что коэффициент сжатия, достигаемый при его применении, примерно совпадает с тем, что дают другие известные методы сжатия информации без потерь – арифметический, Лемпеля–Зива (ЛЗ), Хаффмана [26].

2. Организация хаотических последовательностей, содержащих требуемую информацию

Хаотическая система генерирует информацию. Эта информация неконтролируема по содержанию. Можно ли заставить систему генерировать требуемую полезную информацию? Полезную в том смысле, чтобы это была информация, которую мы должны передать (закодировать в виде хаотического сигнала). Рассмотрим отображение сдвига Бернулли:

$$x_{n+1} = 2x_n \pmod{1}. \quad (1)$$

Свойства траекторий этого отображения удобно анализировать, используя двоичное представление значений переменной x_n . Пусть траектория отображения стартует с начального значения, лежащего внутри отрезка $[0, 1]$ и представленного в двоичной систем счисления

$$x_0 = \sum_{v=1}^{\infty} a_v 2^{-v}, \quad a_v \in \{0, 1\}, \quad (2)$$

т. е. $x_0 = 0.a_1a_2..$ Действие отображения (1) на начальное условие (2) заключается в сдвиге мантииссы на одну позицию влево и отбрасывании первого члена α_1 . Таким образом,

$$x_1 = 0.a_2a_3.. \quad (3)$$

То же самое происходит на последующих итерациях. Поэтому

$$x_n = 0.a_{n+1}a_{n+2}.. \quad (4)$$

Представление справедливо при бесконечной точности вычислений. Пусть теперь точность вычислений конечна, например, числа описываются M битами. Тогда

$$x_0 = \sum_{v=1}^M a_v 2^{-v}, \quad a_v \in \{0, 1\}, \quad (5)$$

т. е. $x_0 = 0.a_1a_2...a_M$. Применяя отображение сдвига Бернулли к x_0 , получаем x_1 :

$$x_1 = 0.a_2...a_M^* , \quad (6)$$

где $*$ означает новую открытую позицию. Как она заполняется? В зависимости от устройства компьютера на последней позиции равновероятно появляется 0 или 1. Мы можем на каждой итерации искусственно вставлять в последнюю позицию нужный информационный бит

$$x_1 = 0.a_2...a_M a_{M+1} . \quad (7)$$

Извлечение этого информационного бита осуществляется путем M -кратного итерирования отображения (7).

Приведенный пример – тривиальный, но он показывает принципиальную возможность подобного кодирования. Сама по себе такая кодировка интереса не представляет. В качестве менее тривиального примера рассмотрим симметричное «tent»-отображение, описываемое уравнением

$$x_{n+1} = \begin{cases} 2x_n, & \text{если } 0 \leq x_0 \leq 1/2 \\ 2(1-x_n), & \text{если } 1/2 < x_0 \leq 1 \end{cases} \quad (8)$$

Эта система, как и сдвиг Бернулли, обладает полной символической динамикой. Снова рассмотрим двоичное представление переменной (5). Преобразование переменной при применении к ней отображения сдвига Бернулли в этом случае не так тривиально. Действительно,

$$x_1 = \begin{cases} 0.a_2a_3...a_M S_1, & \text{если } 0 \leq x_0 \leq 1/2 \\ 0.b_2b_3...b_M S_2, & \text{если } 1/2 < x_0 \leq 1 \end{cases} \quad (9)$$

где S_1 и S_2 – появившиеся после итерации открытые позиции. Как связаны a_i и b_i ?

$$1.000\dots0 - 0.a_1a_2\dots a_M = 0.b_1b_2\dots b_M. \quad (10)$$

На открытую позицию, как и в случае отображения сдвига Бернулли, вводится информационный бит. Получаемая при итерациях «хаотическая» траектория содержит полезную информацию, но «запутана» по сравнению с траекторией получаемой при итерациях отображения сдвига Бернулли. В каждом «хаотическом» отсчете будет содержаться ровно 1 бит полезной информации. Итак, мы построили хаотический кодер. Как организовать декодер? Приемник получает последовательность зашумленных «хаотических» отсчетов. Наша задача – извлечь заложенную в них информацию. Будем действовать следующим образом. Введем отображение обратное «tent»-отображению. В декодер (обратное отображение) поступают, вообще говоря, искаженные хаотические отсчеты y_n :

$$y_n = x_n + w_n, \quad (11)$$

где w_n – шум в канале. Пусть уровень шума мал по сравнению с уровнем хаотического отсчета, но достаточно велик по сравнению с амплитудой младшего разряда отсчета. При этих условиях мы не можем установить – непосредственно какой бит вводится в младший разряд. С другой стороны у нас имеется информация о том, какую из ветвей выбрать при итерировании двузначного отображения, обратного «tent»-отображению [33, 34]. Для извлечения заложенной в n -м отсчете информации итерируем обратное отображение M раз, взяв в качестве начального условия y_n . Для устранения неоднозначности при итерировании используются принятые ранее отсчеты y_{n-M}, \dots, y_{n-1} . При этом элементы последовательности восстанавливаются с предельно возможной точностью, и переданная информация извлекается из последнего двоичного знака.

3. Линейное взаимодействие хаоса, информационных сигналов и шума

Классическая теория информации имеет дело с двумя типами источников, порождающих информацию (энтропию): *источниками сообщений и источниками шума*.

Источник сообщений может порождать в единицу времени конечное количество информации, как это имеет место в случае использования конечного алфавита. Он также может порождать в единицу количество информации от нуля до бесконечности, если рассматривать непрерывные значения сигналов, и иметь возможность измерять их с неограниченной точностью. Однако введение конечной точности измерений приводит к необходимости рассматривать только состояния, различимые при данной точности измерений. Их число конечно и информация, которая может быть воспринята от источника, также становится ограниченной. Причиной ограниченности объема воспринимаемой информации могут также служить шумы в канале связи.

Стандартной моделью *источника шума* является источник белого шума. Белый шум в источниках с дискретным временем представляет собой последовательность нормально распределенных независимых отсчетов. Формально такой шум обладает бесконечной энтропией.

Генераторы хаоса также можно рассматривать как источники информации, причем как источники с весьма специфическими свойствами:

- сигналы этих источников принимают непрерывные значения;
- средняя скорость генерации информации ограничена, и в простейшем случае систем с единственным положительным ляпуновским показателем λ определяется выражением

$$I = \lambda \log_2 e \quad (12)$$

Задачей теории информации является анализ взаимодействия источников информации (энтропии) между собой и с приемниками. Фактически при наличии источников сообщений и источников шума речь может идти только о передаче информации в отсутствие шума и передаче информации в присутствии шума, т. е. либо о взаимодействии нескольких источников сообщений, либо о взаимодействии источников сообщений и источников шума. Дополнение этой пары источниками хаоса существенно расширяет и усложняет картину. Теперь возможны следующие новые комбинации взаимодействия источников энтропии (информации): источник хаоса – источник шума; источник хаоса – источник хаоса; источник сообщений – источник хаоса и, наконец, источник сообщений – источник хаоса – источник шума. Покажем, что перечисленным типам взаимодействия отвечают содержательные информационные и коммуникационные задачи.

Источник хаоса – источник шума

Хаотическая самосинхронизация в присутствии шума. Хаотическую синхронизацию [27–32] можно рассматривать как процесс передачи информации через канал с шумом от хаотического источника к приемнику, который должен воспроизвести переданный сигнал либо точно, либо с допустимым уровнем искажений [33–36]. Здесь очень важно то обстоятельство, что независимо от «физического содержания» процесса синхронизации он может осуществляться и поддерживаться при наличии канала связи между «передатчиком» и «приемником», обладающим достаточной пропускной способностью, причем количественная величина необходимой пропускной способности канала определяется через степень хаотичности сигнала «передатчика».

Фактически в этом случае мы имеем дело с некоторым обобщением понятия синхронизации, поскольку она понимается не как навязывание поведения одной системы другой, а как получение точной или приближенной копии сигнала на приемной стороне (копия включает в себя и «синхронизм» во времени с учетом времени распространения сигнала).

Важным обстоятельством представляется также тот факт, что содержательная сторона в рассмотрении синхронизации как процесса передачи информации имеется только в случае хаотических сигналов, поскольку количество информации в периодических, в любых других регулярных, а также в нерегулярных, предсказуемых на большие времена сигналах, равно нулю.

Наконец, анализ хаотической синхронизации с точки зрения теории информации дает повод рассуждать о принципах согласованного поведения и «синхронизации» систем, в которых заведомо производится обмен информацией, но которые формально нельзя отнести к числу хаотических. Примерами таких систем могут служить биологические сообщества, большие экономические системы, большие информационные системы (телевидение, Интернет), сообщества государств и т. д. Можно предположить, несмотря на материальные потоки, физическое взаимодействие подсистем: а) фундаментальной причиной координированного поведения являются процессы обмена информацией между подсистемами; б) в системах с «синхронизацией» производство информации способно в значительной степени компенсировать изменения материальных потоков; в) навязывание согласованного поведения проще в тех системах, где скорость производства информации в подсистемах минимальна; г) информационные потоки могут играть как интегрирующую, так и дезинтегрирующую роль для целостности систем.

Другими примерами задач взаимодействия хаотических сигналов (хаотических систем) и шума являются: радио, акустическая и оптическая локация на основе хаотических сигналов; различение хаотических и шумовых сигналов [37, 38]; выделение хаотических сигналов на фоне помех [30].

Источник хаоса – источник информационных сообщений

Передача информации с использованием хаоса. Начиная с 1992 года был предложен ряд способов передачи сигналов, использующих хаотическую динамику: хаотическая маскировка (*chaotic masking*); переключение хаотических режимов; нелинейное подмешивание (*nonlinear mixing*); дуальное нелинейное преобразование (*inverse systems*); опережающее управление сечением Пуанкаре (*Predictive Poincare control modulation*); частотная модуляция хаотическим сигналом и др.

Хаотическая маскировка. Под хаотической маскировкой можно понимать несколько разных задач, каждая из которых имеет свою специфику и требует выполнения различного набора условий для своего решения.

З а д а ч а 1. Пусть имеется источник «паразитных» излучений, которые содержат информацию, нежелательную для приема посторонним наблюдателем. С другой стороны нет необходимости в приеме сообщений и «своими» пользователями. В этом случае источник хаоса может применяться для маскировки «паразитного» сигнала.

Задача 2. Имеется источник сообщений, информация от которого должна приниматься «своим» пользователем (пользователями) и нежелательно, чтобы она принималась посторонними наблюдателями. В этом случае источник хаоса может применяться для маскировки информационного сигнала. Свои пользователи обеспечиваются компенсаторами маскирующего сигнала.

Задача 3. Имеется n своих и m чужих источников сообщений, работающих в одном и том же диапазоне частот. Задача заключается в том, чтобы «накрыть» этот диапазон маскирующим «одеялом» так, чтобы во-первых, обеспечить скрытность работы своих пользователей, с одновременной возможностью для них компенсировать маскирующий эффект хаотического сигнала; во-вторых, максимально затруднить работу чужих радиоэлектронных средств.

Источник хаоса – источник хаоса

Разделение хаотических сигналов. Пусть имеются две, вообще говоря, неодинаковые («ведущие») динамические системы, генерирующие хаотические колебания. Эти колебания складываются, и суммарный сигнал передается к паре других («ведомых») динамических систем, которые могут быть связаны между собой. Возможна ли в этих условиях обратная операция разделения двух колебательных процессов в паре ведомых динамических систем? Это и есть задача разделения хаотических сигналов [40–42].

Источник информации – источник хаоса - источник шума

В качестве примеров такого типа взаимодействий можно привести системы передачи информации с использованием хаоса (о которых шла речь выше), работающие в реальных условиях с помехами. Далее эти многопользовательские системы связи, использующие хаотические несущие с выделением своих сообщений на «шумовом» фоне «чужих» сигналов, и, наконец, радиосреда со многими потребителями, использующими как обычные, так и хаотические сигналы и создающие помехи друг другу.

4. Модуляция и демодуляция хаотического носителя информационным сигналом

В большинстве современных систем связи в качестве носителя информации используются гармонические колебания. Информационный сигнал в передатчике модулирует эти колебания по амплитуде, частоте или фазе, а в приемнике информация выделяется с помощью обратной операции – демодуляции. Модуляция носителя может осуществляться либо за счет модуляции уже сформированных гармонических колебаний, либо путем управления параметрами генератора в процессе формирования колебаний.

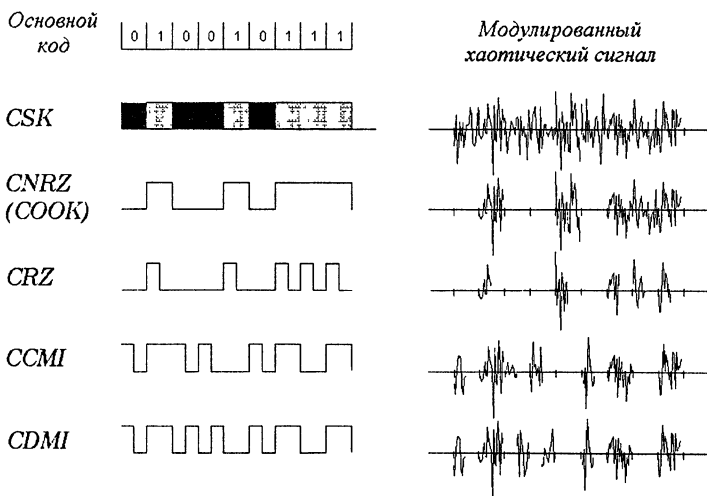
Аналогичным образом можно производить модуляцию хаотического сигнала информационным сигналом. Однако возможности здесь значительно шире. Действительно, если в случае гармонических сигналов управляемых характеристик всего три (амплитуда, фаза и частота), то в случае хаотических колебаний даже небольшие изменения значения параметра одного из элементов источника хаоса приводит к изменению характера колебаний, которое может быть надежно зафиксировано. Это означает, что у источников хаоса с изменяемыми параметрами элементов потенциально имеется широкий набор схем ввода информационного сигнала в хаотический сигнал (модуляции хаотического сигнала информационным сигналом).

В системах связи традиционен интерес к сигналам с широкой полосой частот, которая используется как для увеличения скорости передачи информации, так и для повышения устойчивости работы систем, при наличии возмущающих факторов. Хаотические сигналы являются принципиально широкополосными, и именно это обстоятельство привлекает внимание к ним как к носителям для передачи информации.

Рассмотрим некоторые методы модуляции хаотического сигнала информационным сигналом и соответствующие методы демодуляции.

Пусть требуется передавать двоичные данные в виде последовательности нулей и единиц. Как можно промодулировать хаотический сигнал, чтобы его фрагменты соответствовали нулям и единицам, а на приемном конце эти нули и единицы можно было выделить?

Несколько возможных решений приведено на рисунке.



Методы модуляции хаотического сигнала для передачи цифровой информации

1. Нули и единицы кодируются последовательно расположенными фрагментами хаотического сигнала двух видов. Первый сигнал кодирует нули, второй – единицы. Приемник должен содержать устройство, способное фиксировать соответствующие временные интервалы и идентифицировать два различных хаотических режима. Это метод модуляции, основанный на переключении хаотических режимов [43–45] (CSK – *chaotic shift keying* – в англоязычной терминологии).

2. Как и в предыдущем случае, нули и единицы кодируются последовательно расположенными фрагментами двух сигналов. Но теперь один из них, кодирующий нули, – нулевой сигнал, а второй, кодирующий единицы, – хаотический. Этот метод модуляции в случае нехаотических сигналов носит название «метод без возвращения к нулю» (NRZ – *Not Return to Zero*). В случае хаотического сигнала его логично назвать «методом без возвращения к нулю для хаотических сигналов» (CNRZ). В англоязычной литературе используется также термин OOK (*On-Off Keying*) для нехаотических сигналов и соответственно COOK (*Chaotic On-Off Keying*) для хаотических.

3. В методе с возвращением к нулю (RZ – *Return to Zero*) через время, соответствующее положению периода после прохождения сигнала «1», производится возврат к нулевому состоянию. Нули кодируются отсутствием на соответствующих позициях сигнала. В случае хаотического сигнала – это «метод с возвращением к нулю для хаотических сигналов» (CRZ – *Chaotic Return to Zero*).

4. Для модуляции хаотических сигналов можно также использовать метод кодирования, в котором период делится на две части и производится попеременное инверсное преобразование «0» в «10», а «1» в «11» или в «00». В случае регулярных сигналов этот метод называется CMI. В случае хаотического сигнала – CSMI.

5. Применим и метод модуляции, называемый в случае нехаотических сигналов DMI, в котором при инвертировании нулей и единиц период изменяется в два раза. В случае хаотического сигнала его логично называть CDMI.

6. Еще одной возможностью является случайное или хаотическое размещение позиций, на которых располагаются приходящие сигналы, причем закон расположения позиций известен как отправителю, так и получателю.

Перечисленные типы модуляции (кодирования) предназначены для канала с одним пользователем. Однако часть из них может быть обобщена на многопользовательские системы. Например, если в методе CRZ увеличить интервалы времени, на которые производится возврат к нулю (т. е. интервалы времени, в которые сигнал не передается), то в них можно разместить сигналы от других пользователей, обеспечивая тем самым временное разделение сигналов для многопользовательского доступа.

Приведенные примеры методов модуляции хаотических сигналов показывают, что здесь применимы как традиционные для техники связи, так и относительно новые подходы. А как обстоит дело с демодуляцией?

Для демодуляции хаотических сигналов могут применяться как когерентные, так и некогерентные методы. Среди когерентных методов наибольший интерес представляют методы, использующие хаотическую синхронизацию. Основную идею применения хаотической синхронизации для когерентного приема поясняет рисунок. Здесь хаотическая синхронизация применяется для демодуляции сигналов при использовании модуляции с переключением хаотических режимов. Методы когерентной демодуляции на основе хаотической синхронизации имеют ряд очень привлекательных черт, однако им свойствен такой существенный недостаток, как слабая устойчивость к рассогласованию параметров передатчика и приемника, шума в канале и другим возмущающим факторам.

Поэтому наряду с усилиями, направленными на поиск более устойчивых схем когерентного приема, уделяется большое внимание некогерентным методам приема. В простейшем случае это могут быть методы прямого детектирования хаотического сигнала с выделением огибающей, соответствующей информационному сигналу. Расчеты показывают, что при всей своей простоте этот метод демодуляции является достаточно эффективным, а в сложных условиях распространения сигнала даже более эффективным, чем когерентные методы.

5. Информация и вычисления

До недавнего времени информация рассматривалась как нефизическая, простая запись состояния осязаемой материальной Вселенной, существующая вне физики и никак не связанная с областью, управляемой законами физики. Однако изучение физики, вовлекаемой в обработку информации (особенно с применением компьютеров) привело к твердому выводу, что информация больше не метафора. Она не является чем-то абстрактным и субъективным, но реальна как атомы, энергия или горы. Информация может быть количественно определена. Она всегда воплощена в некотором физическом представлении, таком, как чернила на бумаге, отверстия в перфокарте, магнитные структуры на флоппи-дисках или расположение атомов в ДНК. Одно из самых глубоких и наиболее важных открытий современной науки: информация физична. В своем гипертрофированном варианте принцип физичности информации привел к идее о Вселенной-компьютере, которая, в свою очередь, базируется на менее содержательном принципе, утверждающем, что Вселенная может быть фактически описана в терминах обработки информации.

Принцип Ландауэра. В 50-е годы в ИВМ проводились исследования по выяснению основных физических ограничений на вычислительный

процесс. Изучением проблемы, которая в то время имела слабое отношение к технике разработки компьютеров, занимался Р. Ландауэр. Реально даже в наши дни компьютеры расходуют энергии значительно больше, чем принципиальный фундаментальный нижний предел.

«Когда реальность находится так далеко от фундаментальных пределов, пределы не используются в качестве ориентиров для технологов,— писал Ландауэр.— Тогда почему мы этим занимаемся? Потому, что это находится в самом центре науки. Наука, и физика в особенности, выражена в терминах математики, т. е. в терминах правил для оперирования с числами. Информация, численная или другая, не является абстракцией, но неизбежно связана с физическими представлениями... Поэтому обработка информации неизбежно связана с физической величиной, ее содержанием и ее законами» [46].

В своих поисках (конечных) пределов обработки информации Ландауэр руководствовался примерами Шенноновского канала наиболее эффективной системы связи и вопросом Карно о предельной эффективности тепловой машины. Задача была найти путь понимания фундаментальных пределов для эффективности вычислений, которые (пределы) не зависят от конкретной вычислительной машины.

Общий знаменатель между теорией Шеннона и теорией паровой машины заключается в том, что в обоих случаях были найдены пределы, по которым то, что вы пытаетесь сделать, не зависит от деталей выбранной технологии. Точно так же должен существовать путь обсуждения вычислительного процесса и его энергетических пределов, безотносительно к тому, будем ли мы говорить о релейной логике или логике на основе интегральных микросхем, должен существовать путь делать это на более фундаментальной основе.

Идеальная эффективность. Насколько в принципе может быть эффективен компьютер? Наряду с упоминавшимися выше теориями Карно и Шеннона, при анализе этого вопроса на Ландауэра оказали влияние взгляды Леона Бриллюэна. В середине 50-х годов вышла ставшая знаменитой книга Бриллюэна «Наука и теория информации» [47], в которой рассматривалась связь между наукой (в основном физикой) и информацией. Тот вид, в котором это было сделано, Ландауэр считал неудовлетворительным. Он посчитал, что должен существовать более интересный путь размышления об этом.

В частности, в своей книге Бриллюэн анализировал парадокс демона Максвелла с позиций допущения, что демон не может видеть в темноте. Следовательно, требовалось бы подсветить приближающуюся молекулу для того, чтобы увидеть ее и измерить ее скорость. Энергия, требуемая для подсветки, превышала бы энергию, спасенную демоном, оставляя справедливым второй закон термодинамики. Другие ученые того времени объясняли парадокс демона в такой же манере.

В конечном итоге этот подход был основан на вычислениях Шеннона, касающихся того, какое требуется количество энергии, чтобы послать сообщение. *Измерение чего-то означает приобретение информации, что в основе своей то же самое, что и передача информации.* И Шеннон показал, что посылка бита информации по телефонной линии требует минимального количества энергии, которой обладает типичная молекула, находящаяся в окрестности при комнатной температуре.

Ландауэр при анализе парадокса демона Максвелла осознал, *что если измерения подобны передаче информации* (что становится ясно из анализа ситуации с демоном Максвелла с помощью теории Шеннона), *то они также подобны вычислениям* [48–50] – все три процесса относятся к обмену информацией. Если Бриллюэн был прав, то вычисления требовали бы неизбежной потери энергии на каждом шаге вычислений, тех же самых, что каждое измерение для молекулы демона Максвелла. Однако Ландауэр предположил, что существует не только этот путь для передачи информации. Анализ Шеннона был проведен для специального случая. Существуют пути передачи сообщений, отличные от колебаний в проводах.

«Я не обязан передавать информацию именно этим путем (с помощью колебаний в проводах), – говорит Ландауэр. – Я могу передать вам флоппи-диск или использовать для этой цели обыкновенную почту». Обоснование этой идеи было изложено в статье «Энергетические требования для передачи информации» [51]. Однако эта статья не вызвала такого резонанса, как другая – главная статья Ландауэра.

Главная статья Ландауэра (в соавторстве с Дж. Свенсоном) опубликована в 1961 г. в *IBM Research Journal* [52]. Там было объяснено, что вычисления сами по себе не требуют никакого, даже минимального расхода энергии. *Соединение двух систем вместе для обеспечения информационного обмена может быть реализовано таким образом, что не будет никакой потери энергии.* Идеальная вычислительная система может в принципе манипулировать информацией, не производя тепловых отходов вообще. Ключ объяснения состоял в понимании того, что вычисления могли быть все более и более медленными для уменьшения трения, вызывающего тепловые потери. Стирание же информации всегда производит тепло, которое рассеивается в окружающей среде.

Необходимые потери (затраты) энергии не зависят от того, как вы стираете информацию и какого она вида. Тот факт, что стирание требует (некоторой) минимальной затраты энергии, известен теперь среди компьютерных физиков как *принцип Ландауэра*.

Пример с баскетбольными мячами. Простой пример с баскетбольными мячами демонстрирует, что стирание информации производит тепло, которое рассеивается в окружающей среде. Рассмотрим систему для представления информации, состоящую из двух мячей. Положите один из

них около своей левой ноги, а другой держите в правой руке. Мяч на полу соответствует состоянию «0», мяч в руке – состоянию «1». Вы играете роль человека – флоппи-диска с емкостью информационной памяти (записи) в два бита. Теперь ваше задание стереть бит «1» в вашей правой руке. Решение звучит довольно просто – бросить мяч. Но мяч после броска не займет просто состояние «0» на полу. Вместо этого он отскочит. Фактически, если вы имеете совершенный эластичный мяч и хороший жесткий пол, он отскочит прямо обратно вам в руку, в положение, соответствующее состоянию «1». Чтобы перейти в состояние «0» – стереть бит «1» – мяч должен испытать трение о молекулы воздуха и пола. После каждого отскока мяч будет достигать все меньшей и меньшей высоты. В конечном счете трение оставляет мяч на полу и «1» стирается. Но почему? Только потому, что энергия от отскочившего мяча была передана полу и воздуху. В вакууме и с полом без потерь мяч отскакивал бы вам в руки постоянно (неопределенно долго). Информация может быть стерта, только если в процессе расходуется энергия.

Обратимые вычисления. Итак, энергия нужна для вычислений, поскольку необходимо стирать информацию по пути. Удаление информации требует затрат энергии.

Спустя несколько лет, коллега Ландауэра по IBM Беннетт переформулировал вопрос следующим образом. Допустим, вы нашли способ вычислений без стирания какой-либо информации. Тогда вы могли бы не затрачивать никакой энергии.

Предположим, вы не пытаетесь спасти всю промежуточную информацию, рассуждал Беннетт. Если все шаги вычислительного процесса могут быть обращены, заключил он, тогда вся предыдущая информация может быть реконструирована. Предположим, что вы разработали вычислительные шаги так, что они могут быть восстановлены без размещения каждого шага в памяти. Вы программируете компьютер не запоминать все, что он сделал, а фиксировать место, где это было сделано. Другими словами, в любой точке в потоке вычислений (вычислительного потока) вы можете переключить устройство «на обратный ход» и послать компьютер к его исходной точке. Такой «обратимый компьютер» способен восстановить все логические шаги, включенные в формирование его вычислений. Запускаем компьютер вперед для получения вашего ответа, затем запускаем его назад для получения всей информации, которая была вовлечена для получения этого ответа.

Работа Беннетта [53] показала, что компьютер в принципе не обязан в действительности хранить любую информацию. Поэтому вычисления могут быть организованы так, чтобы тратить энергию настолько мало, насколько это потребуется и без необходимости в больших объемах памяти. «Если вы делаете все верно, это не замедляет все очень сильно и не заставляет использовать очень большую память», – говорил Беннетт. По-

думайте о переходе реки с использованием нескольких специальных камней. Вы можете собирать камни и размещать их обратно в воде, согласно определенным правилам. Применяя эти правила, вы сможете пересечь реку, используя лишь несколько камней. А еще использование правил позволяет вам точно восстановить свои шаги, и таким образом информация о том, как вы переходите реку, будет сохранена. Точно также, обратимые компьютеры, используя соответствующие правила, могут осуществлять сложные вычисления, избегая совершать действия по сохранению информации о каждом шаге на пути.

Когда Беннетт впервые предложил обратимые вычисления, сам Ландауэр сомневался. В течение месяцев он предполагал, что при анализе что-то было упущено. Ричард Фейнман также сначала был настроен скептически. Однако оба в конце концов осознали, что идея Беннетта работает [54]. Часто говорят, что «существует некоторое, абсолютно минимальное количество энергии, необходимое для обработки одного бита информации». «Это просто неправда», – возражал Ландауэр много раз. «В действительности не существует никаких пределов на обработку информации, – настаивал он, – измерение не требует диссипации энергии, точно так же как *передача информации*» [51].

В течение многих лет схема обратимых вычислений, предложенная Беннеттом, представляла интерес только для фундаментальной физики, а не для создания реальных чипов. Но постепенно обратимые вычисления входят в жизнь. Так, Ральф Меркль из исследовательского центра фирмы Хегох в Калифорнии считает, что для практики обратимые вычисления не обязаны быть совершенными. Энергию можно экономить, даже если только часть вычислений будет обратимой.

Схема обратимых вычислений, предложенная Беннеттом, обходит пределы, налагаемые принципом Ландауэра. Поэтому цель конструкторов, занимающихся снижением потребления энергии компьютерами, заключается в том, чтобы использовать типы обратимых шагов, которые сохраняют исходную информацию настолько широко, насколько это возможно. Некоторые инструкции в компьютере обратимы, а другие – нет. Вы не обязаны иметь полную обратимость для всего компьютера. Грамотное проектирование аппаратуры может сделать большинство шагов обратимыми, спасая большое количество информации и, таким образом, спасая большое количество энергии. Обратимые вычисления иллюстрируют, как важно принимать во внимание при конструировании компьютеров тот фундаментальный факт, что информация реальна и физична.

6. Обобщенный канал связи и вычисления

Теория Шеннона имеет удивительную особенность. Для выяснения возможности передачи информации из одной точки в другую она использует понятие канала связи, в котором не фигурирует в качестве перемен-

ной расстояние. Вместе с тем понятие канала связи позволяет рассматривать передачу информации не только через пространство, но и через время. По существу, имеется, как минимум, два типа коммуникационных каналов:

- 1) канал для передачи информации между двумя пространственными точками, расстояние между которыми D ,
- 2) канал для передачи информации через временной интервал T .

Примерами передачи информации через канал второго типа являются записи информации на бумаге, жестком диске, CD с последующим ее прочтением. В ряде случаев к каналам второго типа (передача информации через интервалы времени) применимо понятие пропускной способности канала. Появляется возможность осознать взаимосвязь между вычислениями и передачей информации, обозначенную в принципе Ландауэра.

Рассмотрим следующую структуру коммуникационного канала второго типа. Имеется источник информационного сообщения в виде начальных условий для решения некоторой задачи на компьютере. Обозначим эти начальные условия через $x \in X$, где X – множество всех начальных условий для данной программы P , реализованной на компьютере. Будем рассматривать компьютер как коммуникационный канал, на вход которого поступает информация x , содержащая некоторое количество бит, и на выходе которого через время T появляется результат работы программы (алгоритма, компьютера) в виде решения $y \in Y$, принадлежащих множеству возможных для этой задачи решений Y .

О передаче какой информации идет речь? Перед решением задачи была полная неопределенность в отношении того, что из себя будет представлять результат работы программы. Пусть Y состоит из N возможных равновероятных решений. В результате работы компьютера мы получаем знания о том, какое именно решение выбрано. Таким образом, неопределенность полностью исчезла, и на выходе появляется информация в объеме $\log_2 N$ бит. Эта информация и была передана через канал за время T . Отсюда пропускная способность канала второго типа равна

$$C = \frac{\log_2 N}{T} \quad (14)$$

Значение C определяется рядом факторов, в частности производительностью компьютера и сложностью задачи.

Таким образом, вычислениям можно сопоставить некоторый коммуникационный канал, что подтверждает общность процесса передачи информации и процесса вычислений.

Справедливо ли обратное, т. е. может ли произвольный коммуникационной канал быть представлен в виде вычислительного процесса? Какими свойствами обладают каналы типа вычислительного процесса?

Рассмотрим пример с синхронизацией двух хаотических систем через передачу информации от ведущей системы к ведомой [33–36].

Пусть необходимо обеспечить точность синхронизации в n бит. Пренебрежем тем фактом, что из-за разбегания траекторий ведомая система действует как усилитель шумов (т. е. понижает пропускную способность канала «связи»). Тогда требуемая точность передачи отсчетов равна n битам, и, следовательно, необходимая пропускная способность канала связи равна n бит на отсчет хаотической системы.

Пусть для определенности динамическая система представляет собой одномерное отображение и генерирует один бит информации на отсчет. Тогда, применяя преобразователь хаотических отсчетов в двоичный поток, мы получаем поток данных в канале первого типа, равный 1 биту на отсчет динамической системы.

Восстановление хаотических отсчетов на входе ведомой хаотической системы в схеме синхронизации через передачу информации осуществляется обратный преобразователь. Он производит на каждом временном интервале, соответствующем длине такта системы, ровно n бит информации, путем обработки входного сигнала. В каком-то смысле он в n раз усиливает (увеличивает) информационный поток. Тем самым преобразователь увеличивает общую пропускную способность «комбинированного» канала связи, состоящего из последовательно соединенных пространственного и временного каналов. Эта дополнительная информация производится самим обратным преобразователем.

Вывод 1. Устройства обработки принимаемого сигнала могут повышать пропускную способность канала связи.

Этот вывод не противоречит обычным представлениям о пропускной способности канала связи. Действительно, фильтрацией шумов мы также повышаем пропускную способность канала, поскольку уменьшаем уровень шумов в приемнике. Однако здесь имеется и существенная разница.

Вывод 2. Линейные устройства обработки информации не позволяют увеличить поток выходной информации по отношению к входному (в канал). Применение же схемы обработки вычислительного (вообще говоря, необратимого) процесса может приводить к дополнительному производству информации и, соответственно, к увеличению ее потока по сравнению с входным потоком.

Замечание. Представим себе, что на передающей и приемной сторонах имеются две одинаковые машины Тьюринга. Каждая из машин при своем функционировании в силу логической необратимости производит в процессе своей работы информацию. Мы можем по шагам передавать информацию от машины передатчика к машине приемнику. Для того чтобы передавать всю эту информацию, необходим достаточно высокоскоростной канал. Однако можно поступить и по-другому. Задать одинаковые начальные состояния машин (для чего передать в машину прием-

ника необходимые начальные условия) и запустить машину в приемнике. При этом получатель будет принимать в точности ту же информацию, которую он получал при непосредственной передаче данных с выхода машины передатчика.

Возникает вопрос: а не банальность ли это? *Может быть, информация фактически уже имеется на приемной стороне и «зарыта» в конструкции машины Тьюринга?* Ответ заключается в том, что множество возможных начальных состояний настолько велико, что случайно воспроизвести нужный результат не представляется возможным. По существу, здесь идет обмен информацией, передаваемой по каналу, на информацию, порождаемую машиной Тьюринга. Но для внешнего наблюдателя это одна и та же информация.

7. Информационно-открытые системы

При синхронизации двух систем связи через передачу информации уменьшение объема данных, передаваемых через пространственный канал связи, по сравнению с синхронизацией на основе передачи самого хаотического сигнала компенсируется дополнительными вычислениями, которые трактуются как преобразование и «усиление» информации во временном канале связи. «Экономия» в пропускной способности канала связи весьма значительна: примерно в 10 раз для синхронизации с относительной точностью порядка 10^{-3} .

Имеются ли пути дальнейшего уменьшения объема передаваемой информации за счет дополнительных вычислений на приемной стороне? Чтобы ответить на этот вопрос, необходимо ввести понятия информационно-открытых и информационно-замкнутых физических систем.

Пусть имеется динамическая система Σ , описываемая некоторыми эволюционными уравнениями для переменной $x \in R^N$, где N – размерность фазового пространства системы. Поведение Σ можно изучать, квантуя ее по времени и переменной состояния. Основная идея заключается в делении множества возможных состояний на конечное число ячеек и в нахождении маршрута, определяющего, к какой ячейке относится состояние системы при каждом такте часов. Каждая ячейка описывает *информационное состояние системы* и ассоциируется с некоторым «символом» и, таким образом, информационная эволюция системы описывается бесконечной последовательностью символов.

Будем называть систему Σ информационно-открытой системой, если она связана с внешней по отношению к ней средой таким образом, что эта связь приводит к изменению информационных состояний системы в процессе ее эволюции относительно такой же системы, не связанной с внешней средой. В противном случае будем называть систему Σ информационно-замкнутой системой.

Термодинамическая и информационная открытость системы – вещи, вообще говоря, разные. В частности, термодинамическая открытость системы еще не означает информационной открытости (хотя такое и может быть, см. пример с баскетбольными мячами).

Каждая динамическая система, на которую воздействует внешний шум, может рассматриваться по отношению к среде, порождающей этот шум, как кандидат на информационно-открытую для этой среды систему. Она будет информационно открытой в том случае, если уровень внешнего шума таков, что шум способен изменить информационную эволюцию системы.

Рассмотрим примеры, которые поясняют понятие информационной открытости.

Пример 1. Пусть имеется диссипативный осциллятор с двухъямным потенциалом и внешним гауссовским шумом. Динамика осциллятора описывается уравнениями

$$\ddot{x} + \alpha \dot{x} + V'_x(x) = \xi(t). \quad (15)$$

Связь с внешней средой осуществляется через шум $\xi(t)$. Система имеет два информационных состояния: около левого и около правого положений равновесия. Обозначим их соответственно через состояние «0» и состояние «1». Если $\alpha \gg 1$, то это случай передемпфированного осциллятора, динамика которого в нормированном виде описывается как

$$\dot{x} = -V'_x + \xi(\tau), \quad (16)$$

где $\tau = \alpha t$; $\langle \eta(\tau)\eta(0) \rangle = 2D\delta(\tau) = \sigma^2$ и для определенности потенциал $V(x)$ имеет вид

$$V(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4. \quad (17)$$

При малых значениях σ система является информационно-замкнутой, поскольку связь с внешней средой, осуществляемая за счет воздействия гауссовского шума не меняет информационного состояния системы. Однако при увеличении σ вероятность перехода в другое состояние начинает существенно отличаться от нуля. При больших амплитудах шума траектория время от времени перескакивает через барьер, система «воспринимает» внешнюю информацию, ее информационное состояние меняется при каждом перескоке, и она становится информационно открытой. Количественно вероятность перехода r_k от одного состояния к другому определяет соотношение Крамерса:

$$r_k = \frac{1}{2} \exp\left(-\frac{1}{2\sigma^2}\right). \quad (18)$$

В рассмотренном примере существенно наличие *порогового значения* внешнего шума, ниже которого система является информационно-замкнутой, а выше – информационно-открытой.

Пример 2. Физически реализованная хаотическая гиперболическая система описывается одномерным отображением

$$x_n = f(x_n) + \xi_n, \quad \langle \xi_n \xi_m \rangle = \delta_{nm} \sigma^2, \quad (19)$$

где δ_{nm} – символ Кронекера. Гиперболические хаотические системы обладают высокой чувствительностью к возмущениям. В них сколь угодно малое возмущение траектории порождает траекторию, экспоненциально быстро расходящуюся с невозмущенной и, следовательно, с другим информационным содержанием. Эволюция информационных состояний самой возмущенной системы также отличается от информационной эволюции невозмущенной системы. Поэтому гиперболическая система является информационно-открытой при сколь угодно малом значении σ , т. е. для нее *порог внешнего шума, при превышении которого информационно-замкнутая система становится информационно-открытой, отсутствует (равен нулю)*.

Таким образом, информационный обмен между гиперболической системой и внешней средой существует при любом ненулевом уровне взаимодействия. Это подтверждает идею Ландауэра о том, что передача информации, вообще говоря, не требует затрат энергии.

Пример 3. В отсутствие внешнего воздействия информационно-замкнутыми системами являются системы с конечным числом состояний. Важным подмножеством (подклассом) таких систем являются типичные компьютеры.

Вернемся к синхронизации.

При обычном подходе к хаотической синхронизации мы должны постоянно передавать от одной системы к другой, как минимум, одну из компонент сигнала (содержащую всю информацию о нем).

Мы можем сократить объем передаваемой информации до некоторого минимального объема, достаточного для восстановления сигнала. Предельное значение этого объема равно производству информации самой системой.

Наконец, если ведущая и ведомая системы являются информационно-замкнутыми и реализуются на типичных процессорах, то по наблюдаемому сигналу от ведущей системы мы можем идентифицировать ее информационное состояние, после чего всю необходимую для синхронизации информацию (расчет траектории) производить в приемнике, уменьшив поток информации через коммуникационный канал до нуля.

При случайной потере синхронизации ее можно будет восстановить путем дополнительных краткосрочных наблюдений за приходящим сигналом.

Заключение

Интерес к взаимосвязи детерминированного хаоса и информации наблюдается уже в течение длительного времени [55–61]. Однако в последние годы в этом интересе появились принципиально новые элементы. Они связаны, прежде всего, с применением хаотических сигналов для передачи сообщений. Оказалось, например, что такое явление, как синхронизация, обобщенное на хаотические системы, может рассматриваться через процесс передачи информации от одной системы к другой. Хаотические колебания могут выступать как в роли носителей информации, так и в роли процессов, маскирующих информационные сигналы. Они могут применяться при активном зондировании (в частности, радиолокации) и служить информативными индикаторами характера естественных и искусственных процессов. Применение динамического хаоса для передачи информации порождает новые возможности для мультиплексирования и демultipлексирования сигналов.

Перечисленные вопросы составляют только часть задач, в которых проявляется необходимость анализа информационных аспектов взаимодействия хаотических сигналов между собой, с другими типами сигналов и их воздействия на соответствующие приемные устройства. Характер этих взаимодействий существенно отличается от соответствующих взаимодействий для типичных сигналов, рассматриваемых в классической теории информации, и может стать мощным источником повышения эффективности коммуникационных систем.

Литература

1. *Котельников В А* Теория потенциальной помехоустойчивости. М. Госэнергоиздат, 1956.
2. *Шеннон К* Работы по теории информации и кибернетике. М.: ИЛ, 1963
3. *Харкевич А А* Борьба с помехами. М.: Наука, 1965
4. *Петрович Н Т, Размахнин М К.* Системы связи с шумоподобными сигналами. М.: Сов радио, 1969
5. *Варакин Л Е* Системы связи с шумоподобными сигналами. М.: Радио и связь, 1985.
6. *Дмитриев А С, Панас А И., Старков С О* Динамический хаос как парадигма современных систем связи // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 1997 № 10 С. 4-26
7. *Шалфеев В Д, Осипов Г В, Козлов А К, Волковский А Р* Хаотические колебания – генерация, синхронизация, управление // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 1997. № 10 С. 27-49.
8. *Дмитриев А. С., Старков С О.* Передача сообщений с использованием динамического хаоса и классическая теория информации // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 1998 № 11. С 4-32.
9. *Андреев Ю В., Дмитриев А. С., Емец С В., Панас А. И., Старков С О., Балабин А М., Дмитриев А. А, Кишик В В., Кузьмин Л В, Борисенко А Г* Стратегии использования динамического хаоса в коммуникационных системах и компьютерных сетях. Разделение кодера источника и кодера канала // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 2000 № 11 С. 4-26.

10. Шарковский А. Н. Сосуществование циклов непрерывного преобразования прямой в себя // Украинский математический журнал. 1964. № 1. С. 61-71.
11. Gunaratne G. N., Procaccia I. Organization of chaos // Phys. Rev. Lett 1987. V. 59, № 13. P. 1377-1380.
12. Cvitanovich P. Invariant measurement of strange sets in terms of cycles // Phys. Rev. Lett 1988. V. 61, № 24. P. 2729-2732.
13. Дмитриев А. С., Старков С. О., Широков М. Е. Структура периодических орбит хаотической автоколебательной системы, описываемой разностными уравнениями 2-го порядка // РЭ 1994. Т. 39, № 9. С. 1392-1399.
14. Дмитриев А. С. Запись и восстановление информации в одномерных динамических системах // РЭ. 1991. Т. 36, № 1. С. 101-108.
15. Dmitriev A. S., Panas A. I., Starkov S. O. Storing and recognizing information based on stable cycles of one-dimensional map // Phys. Lett. A. 1991. V. 155. P. 494-499.
16. Andreyev Yu. V., Dmitriev A. S., Starkov S. O. information processing in 1-D systems with chaos // IEEE Transactions on Circuits and Systems 1997. V. 44, № 1. P. 21-28.
17. Andreyev Yu. V., Dmitriev A. S., Chua L. O., Wu C.W. Associative and random access memory using one-dimensional maps // Int. J. of Bifurcation and Chaos. 1992. V. 3. P. 483-504.
18. Дмитриев А. С. Хаос и обработка информации в нелинейных динамических системах // РЭ 1993. Т. 38, № 1. С. 1-24.
19. Андреев Ю. В., Дмитриев А. С. Запись и восстановление изображений в одномерных динамических системах // РЭ. 1994. Т. 39, № 1. С. 104-113.
20. Андреев Ю. В., Бельский Ю. Л., Дмитриев А. С. Запись и восстановление информации с использованием устойчивых циклов двумерных и многомерных отображений // РЭ 1994. Т. 39, № 1. С. 114-123.
21. Андреев Ю. В., Бельский Ю. Л., Дмитриев А. С., Куминов Д. А. Динамические системы с хаосом как среда для записи, хранения и обработки информации // Изв. вузов Радиофизика. 1994. Т. 37. С. 1003-1019.
22. Andreyev Yu.V., Dmitriev A. S., Kuminov D. A., Chua L. O., Wu C.W. 1-d maps, chaos and neural networks for information processing // Int. J. of Bifurcation and Chaos. 1996. V. 6, № 4. P. 627-646.
23. Andreyev Yu. V., Belsky Yu. L., Dmitriev A. S., Kuminov D. A. Information processing using dynamical chaos // IEEE Transactions on Neural Networks. 1996. V. 7. P. 290-299
24. Andreyev Yu. V., Dmitriev A. S., Starkov S. O. Information Processing in 1-D Systems with Chaos // IEEE Trans. on Circuits and Systems. 1997. V. 44, № 1. P. 21-28.
25. Андреев Ю. В., Дмитриев А. С., Куминов Д. А. Хаотические процессоры // Успехи современной радиоэлектроники. Зарубежная радиоэлектроника. 1997. № 10. С. 50-79.
26. Haykin S. Communication Systems. NY: John Wiley & Sons, 1994.
27. Fujisaka H., Yamada T. Stability Theory of synchronized motion in coupled-oscillator systems // Prog. Theor. Phys. 1983. V. 69, № 1. P. 32-46.
28. Афраймович В. С., Веричев Н. Н., Рабинович М. И. Стохастическая синхронизация колебаний в диссипативных системах // Изв. вузов Радиофизика 1986. Т. 28, № 9. С. 1050-1060.
29. Pikovsky A. S. On the interaction of strange attractors // Zeitschrift fur Physik B. 1984. V. 55. P. 149-154.
30. Pecora L. M., Carroll T. L. Synchronization in chaotic systems // Phys. Rev. Lett. 1990. V. 64, № 8. P. 821-824.
31. Rulkov N. F., Sushchic M. M., Tsimring L. S., Abarbanel H. D. I. Generalized synchronization of chaos in directionally coupled chaotic systems // Phys. Rev. E. 1995. V. 51. P. 980-995.
32. Rosenblum M. G., Pikovsky A. S. Kurths J. Phase synchronization of chaotic oscillators // Phys. Rev. Lett 1996. V. 76. P. 1804-1807.
33. Дмитриев А. С. Хаотическая синхронизация как информационный процесс // Изв. вузов. Радиофизика. 1998. Т. 41, № 12. С. 1497-1509.
34. Dmitriev A. S., Kassin G. A., Khilinski A. D., Chaotic. Synchronization. Information viewpoint // Int. J. of Bifurcation and Chaos 2000. V. 10, № 4. P. 749-761.

35. *Дмитриев А. С., Касьян Г. А., Хаслер М., Хилинский А. Д.* Хаотическая синхронизация двумерных динамических систем на основе передачи информации об их состояниях // РЭ. 2001. Т. 46, № 5. С. 566-575.
36. *Дмитриев А. С., Касьян Г. А., Хилинский А. Д.* Хаотическая синхронизация отображений Хенона. Информационный подход // ПЖТФ. 2002. Т. 28, № 9. С. 36-41.
37. *Eckman J. P., Ruelle D.* Ergodic theory of chaos and strange attractors // Review of Modern Physics. 1985. V. 57, № 3. Part 1 P. 617-656.
38. *Афраймович В. С., Рейман А. М.* Размерность и энтропия в многомерных системах. // Нелинейные волны. Динамика и эволюция / Под ред. А. В. Гапонова-Грехова, М. И. Рабиновича М.: Наука, 1989. С. 238-262.
39. *Abarbanel H. D. I.* Analysis of observed chaotic data. NY.: Springer, 1995.
40. *Tsinring L., Sushchik M.* Multiplexing chaotic signals using synchronization // Phys. Lett. A. V. 213. № 3-4 P. 155-160.
41. *Андреев Ю. В., Дмитриев А. С., Ефремова Е. В.* Разделение хаотических сигналов при наличии шума // РЭ. 2001. Т. 46, № 12. С. 1346-1355.
42. *Andreyev Yu. V., Dmitriev A. S., Efremova E. V.* Dynamic separation of chaotic signals in the presence of noise // Phys. Rev. E. 2002. V. 65. P. 046220-046224.
43. *Parlitz U., Chua L., Kocarev L., Halle K., Shang A.* Transmission of digital signals by chaotic synchronization // Int. J. of Bifurcation and Chaos. 1992. V. 2 № 4. P. 973-977
44. *Бельский Ю. Л., Дмитриев А. С.* Передача информации с использованием динамического хаоса // РЭ. 1993. Т. 38 № 7 С. 1310-1315.
45. *Dedieu H., Kennedy M., Hasler M.* Chaos shift keying modulation and demodulation of a chaotic carrier using self synchronizing Chua's circuits // IEEE Trans. Circuits and Systems. V. CAS-40. № 10 P. 634-642.
46. *Landauer R.* Computation: A Fundamental Physical View // Physica Scripta. 1987. V. 35. P. 88-95. Reprinted in Harray S. Left and Andrew F. Rex eds., Maxwell's Demon (Princeton, NY: Princeton University Press, 1990), P. 260.
47. *Бриллюэн Л.* Наука и теория информации М.: ГИФМЛ, 1960.
48. *Landauer R.* Information is inevitably physical // Feynman and computation / Ed. A.J.G. Hey. Reading, Mass.: Persue, 1999.
49. *Landauer R.* Information is physical // Phys. Today. May. 1991. P. 23-29.
50. *Landauer R.* The physical nature of information // Phys. Lett., A217. P. 188.
51. *Landauer R.* Energy requirements to communications // Appl Phys. Lett. 1987.V. 51. № 24. P. 2056-2058.
52. *Ландауэр Р.* Необратимость и выделение тепла в процессе вычислений // Квантовый компьютер и квантовые вычисления Т. 2. С. 9-32. (Из IBM J. Research Develop, 1961. V. 5. P. 183).
53. *Беннетт Ч.* Логическая обратимость вычислений // Квантовый компьютер и квантовые вычисления Т. 2. С. 33-52. (Из IBM J. Research Develop. 1973. V. 17. P. 525).
54. *Фейнман Р.* Feynman lectures on computation. Reading, Massachusetts: Addison-Wesley Pub. Com., 1996.
55. *Shaw R.* Strange attractors, chaotic behavior and information flow // Naturforsch. 1981. V. 36A. P. 80-112.
56. *Николис Г., Пригожин И.* Познание сложного. М. Мир, 1991
57. *Хакен Г.* Информация и самоорганизация. Макроскопический подход к сложным системам М.: Мир, 1991.
58. *Николис Дж.* Динамика иерархических систем. Эволюционное представление М.: Мир, 1989.
59. *Nicolis J. S.* Chaos and information processing. Singapore: World Scientific, 1991.
60. *Кадомцев Б. В.* Динамика и информация // Нелинейные волны. Структуры и бифуркации М. Наука, 1987 С. 45-53.
61. *Deco G., Schurmann B.* Information dynamics. Foundations and applications. NY.: Springer, 2001.

ХАОТИЧЕСКИ МОДУЛИРОВАННЫЕ КОЛЕБАНИЯ В СВЯЗАННЫХ ФАЗОВЫХ СИСТЕМАХ

В. Д. Шалфеев, В. В. Матросов

Введение

Целью настоящей работы является обсуждение вопросов возбуждения и генерации хаотически модулированных колебаний в ансамблях связанных автогенераторов с фазовым управлением, синхронизации таких колебаний и их использования для передачи информации.

Идея использования динамического (детерминированного) хаоса для передачи информации очень привлекательна, поскольку в этом случае возможно сочетание как традиционных статистических методов обработки информации (например, корреляционных методов), так и специфических динамических методов, основанных на свойствах нелинейной динамики. Высказанная в [1] идея использования в системах связи хаотических автоколебаний в качестве несущих в целях обеспечения скрытности передачи информации нашла достаточно широкий отклик в литературе [2–21]. И хотя последующие исследования не обнаружили высокой криптостойкости рассмотренных вариантов систем связи с хаотической несущей, тем не менее очевидные преимущества таких систем, связанные со свойствами широкополосности и сверхширокополосности хаотических сигналов, а также со сравнительной простотой реализации генераторов хаотических колебаний, наличием разных вариантов управления хаотическими колебаниями, и следовательно богатых возможностей модуляции хаотической несущей, ставят исследования как в области связи на базе хаотических автоколебаний, так и в области других новых информационных технологий на базе динамического хаоса в ряд весьма актуальных и перспективных [11].

В большинстве публикаций по передаче информации с использованием динамического хаоса рассматриваются варианты схем для генерации и синхронизации хаотических колебаний, не использующие системы фазовой автоподстройки частоты (ФАП) [2–14], за очень редким исключением [12, 14–21]. В то же время, известно, что традиционные методы связи, использующие регулярные колебания в качестве несущих, в подавляющем большинстве основаны именно на широком применении систем ФАП для управления частотой модуляции, демодуляции, оптимальности, устойчивости, управляемости и технологичности систем ФАП (стандартные микросхемы ФАП разработаны и выпускаются в широком ассортименте), вызывает интерес исследование возможности эффективной генерации хаотических колебаний на базе систем ФАП, возможности осуществления процесса синхронизации (подстройки) таких хаотических

колебаний и построения на этой основе систем связи с использованием хаотических колебаний. Ниже проводится обсуждение этих вопросов.

1. Возбуждение хаотических колебаний

1.1. Хаотически модулированные колебания в изолированной ФАП

Система ФАП является типовым кольцом автоматического управления частотой генератора Γ (рис. 1). Периодические колебания с выхода генератора Γ с текущей фазой φ_r сравниваются на фазовом дискриминаторе (ФД) с колебаниями опорного сигнала с текущей фазой φ_c , в результате чего на выходе ФД формируется сигнал, зависящий от разности фаз $\varphi = \varphi_c - \varphi_r$, который после прохождения фильтра Φ подается на управляющий элемент УЭ, непосредственно изменяющий частоту генератора Γ . Математическая модель такой системы ФАП может быть записана в операторной форме:

$$\frac{p\varphi}{\Omega} + K(p)\sin\varphi = \gamma. \quad (1)$$

Здесь $p \equiv d/dt$, Ω – максимальное отклонение частоты генератора Γ , которое может компенсировать цепь управления, $\gamma = \Delta\omega/\Omega$, $\Delta\omega$ – отклонение частоты свободного генератора Γ от опорной частоты ω_c .

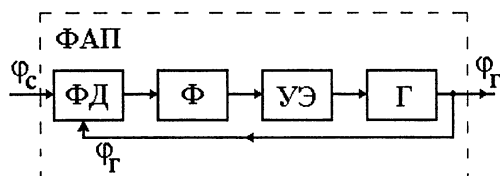


Рис. 1. Система ФАП.

В простейшем случае фильтра первого порядка, когда $K(p) = (1 + ap)^{-1}$, уравнение (1) можно записать в виде

$$\frac{d\varphi}{d\tau} = y, \quad \varepsilon \frac{dy}{d\tau} = \gamma - y - \sin\varphi, \quad (2)$$

где $\tau = \Omega t$, $\varepsilon = a\Omega$. В результате изучения этих уравнений качественными методами теории колебаний установлено, что система ФАП с фильтром первого порядка в цепи управления имеет два стационарных режима:

- *Режим синхронизации генератора опорным сигналом* – режим, при котором частоты управляемого генератора и опорного сигнала равны, а разность фаз принимает некоторое постоянное значение. На фазовом цилиндре $V_0 = \{\varphi \pmod{2\pi}, y\}$ этому режиму отвечает устойчивое состояние равновесия с координатами $\varphi = \arcsin \gamma$, $y = 0$.

- *Режим биений* – режим, при котором разность фаз постоянно растет, а средняя разность частот имеет некоторое постоянное значение. Этому режиму соответствует устойчивый предельный цикл, охватывающий фазовый цилиндр V_0 .

Более сложные динамические режимы в системе ФАП, в том числе хаотические, реализуются при использовании в цепи управления фильтров более высокого порядка. Например, если в цепь управления ФАП поставить фильтр второго порядка с коэффициентом передачи $K(p) = (1 + ap + bp^2)^{-1}$, то уравнение (1) принимает вид

$$\frac{d\varphi}{d\tau} = y, \quad \frac{dy}{d\tau} = z, \quad \mu \frac{dz}{d\tau} = \gamma - y - \sin \varphi - \varepsilon z, \quad (3)$$

где $\mu = b\Omega^2$, $\varepsilon = a\Omega$.

На рис. 2 даны [22] примеры проекций характерных колебательных аттракторов на плоскость (φ, y) , для которых отсутствуют вращательные движения по координате φ .

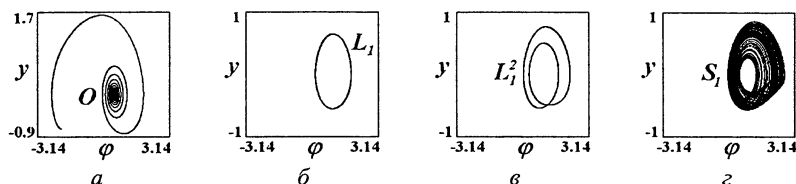


Рис. 2. Примеры колебательных аттракторов системы (3)

Рисунок 2, *а* соответствует режиму синхронизации генератора Γ опорным сигналом, рис. 2, *б, в* – режиму квазисинхронизации генератора опорным сигналом, при котором имеет место периодическая угловая модуляция колебания на выходе генератора Γ , при которой средняя частота генератора совпадает с опорной частотой, а рис. 2, *з* соответствует режиму квазисинхронизации, при котором имеет место хаотическая угловая модуляция колебаний со средней частотой, стабилизированной по опорной частоте. Будем далее называть такой режим режимом хаотически модулированных колебаний (ХМК). Ниже остановимся на режиме возбуждения ХМК как наиболее интересном с точки зрения его синхронизации и использования для передачи информации. Заметим, что в системе (3) возможны и другие режимы, в том числе режимы биений как периодические, так и хаотические, на которых мы далее не будем останавливаться.

Можно сделать вывод, что введение фильтра второго порядка в цепь обратной связи ФАП дает возможность возбудить режим ХМК. Однако детальное исследование этого режима показало, что область существования его в пространстве параметров системы (3) сравнительно невелика [17, 27], что может представить определенные трудности при практиче-

ском использовании таких систем в качестве генераторов ХМК. Попытки добиться увеличения областей генерации ХМК путем изменения нелинейности фазового дискриминатора и параметров фильтра оказались малоэффективными [21]. В [17, 18, 27] предложен другой способ генерации ХМК, основанный на использовании свойств коллективной динамики малых ансамблей связанных систем ФАП с простой (нехаотической) индивидуальной динамикой [23–26, 28]. Остановимся подробнее на этом варианте генерации ХМК. Для примера рассмотрим ансамбль из трех последовательно (каскадно) связанных систем ФАП.

1.2. Хаотически модулированные колебания в ансамбле систем ФАП

Структурная схема ансамбля, состоящего из трех каскадно (последовательно) связанных систем фазовой автоподстройки (КФАП), представлена на рис. 3. Наряду с основными связями в КФАП применяются дополнительные связи «назад», которые реализуются путем передачи сигналов фазового рассогласования с ФД₂ и ФД₃ через некоторые преобразующие устройства K_1 и K_2 в цепь управления первого и второго генераторов соответственно.

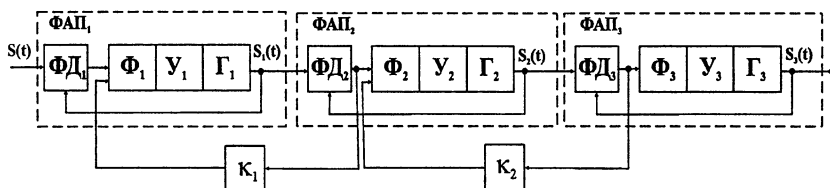


Рис. 3. Каскадное соединение ФАП

Уравнения динамики такой системы в операторной форме записываются в виде [29]:

$$\begin{aligned} \frac{p\varphi_1}{\Omega_1} &= \gamma_1 - K_1(p)[\sin\varphi_1 - \kappa_1 \sin(\varphi_2 - \varphi_1)], \\ \frac{p\varphi_2}{\Omega_2} &= \gamma_2 - K_2(p)[\sin(\varphi_2 - \varphi_1) - \kappa_2 \sin(\varphi_3 - \varphi_2)], \\ \frac{p\varphi_3}{\Omega_3} &= \gamma_3 - K_3(p)\sin(\varphi_3 - \varphi_2), \end{aligned} \quad (4)$$

где $\varphi_i = \varphi_c - \varphi_{\Gamma i}$ – текущие фазовые рассогласования.

Рассмотрим ансамбль, состоящий из одинаковых элементов с наиболее простыми фильтрами в цепях управления, т. е. случай $\Omega_1 = \Omega_2 = \Omega_3 = \Omega$ и $K_i(p) = (1 + ap)^{-1}$. Тогда из (4) получается следующая математическая модель КФАП:

$$\begin{aligned}
\frac{d\varphi_1}{d\tau} &= y_1, & \varepsilon_1 \frac{dy_1}{d\tau} &= \gamma_1 - y_1 - \sin \varphi_1 - \kappa_1 \sin(\varphi_2 - \varphi_1), \\
\frac{d\varphi_2}{d\tau} &= y_2, & \varepsilon_2 \frac{dy_2}{d\tau} &= \gamma_2 - y_2 - \sin(\varphi_2 - \varphi_1) - \kappa_2 \sin(\varphi_3 - \varphi_2), \\
\frac{d\varphi_3}{d\tau} &= y_3, & \varepsilon_3 \frac{dy_3}{d\tau} &= \gamma_3 - y_3 - \sin(\varphi_3 - \varphi_2),
\end{aligned} \tag{5}$$

где $\varepsilon_i = \Omega a_i$. Система (5) определена на шестимерном фазовом цилиндре $V_6 = \{\varphi_1(\text{mod}2\pi), y_1, \varphi_2(\text{mod}2\pi), y_2, \varphi_3(\text{mod}2\pi), y_3\}$ в восьмимерном пространстве параметров $\Lambda_1 = \{\gamma_1, \varepsilon_1, \gamma_2, \varepsilon_2, \gamma_3, \varepsilon_3, \kappa_1, \kappa_2\}$. Она обладает большим разнообразием стационарных движений, которые и описывают различные динамические режимы КФАП. Остановимся на хаотических квазисинхронных режимах модели (5).

Хаотическим режимам КФАП отвечают странные хаотические аттракторы математической модели (5). Эти аттракторы, аналогично регулярным, могут охватывать фазовый цилиндр V_6 по всем циклическим координатам или только по некоторым, а могут и не охватывать ни по одной из координат. Тип аттрактора определяет вид колебаний на выходе управляемых генераторов. Для определения типа аттрактора будем использовать индексы вращения $[J_1, J_2, J_3]$, где J_i принимает значение 1, если по φ_i -координате происходит вращение, и 0 – в противном случае. Отсутствие вращения по φ_i -координате означает, что i -й генератор работает в квазисинхронном режиме, когда остальные генераторы находятся в режиме биений. Выход фазовых траекторий на хаотический колебательный аттрактор с индексом вращения $[0, 0, 0]$ приводит к генерации хаотически модулированных колебаний (ХМК) на выходе всех генераторов ансамбля. Это *глобальный ХМК-режим* КФАП.

Генерации хаотически модулированных колебаний на выходе отдельных генераторов соответствуют колебательно-вращательные хаотические аттракторы, у которых индексы вращения содержат как нули, так и единицы. Эти аттракторы определяют режим частичной генерации хаотически модулированных колебаний ансамблем ФАП или *частичный ХМК-режим* КФАП. Наконец, наличие в фазовом пространстве вращательного хаотического аттрактора с индексом вращения $[1, 1, 1]$ отвечает *глобальному хаотическому режиму биений*.

Хаотические аттракторы в фазовых пространствах моделей КФАП появляются в результате бифуркаций регулярных аттракторов. Ниже рассмотрены некоторые сценарии возникновения и эволюции хаотических колебаний в моделях КФАП, обнаруженные в процессе численных экспериментов.

Переход к хаосу через каскад удвоения периода предельного цикла. Это один из основных переходов к хаотическим колебаниям в моделях

КФАП. Причем он справедлив для любых типов движений, т. е. не зависит от индекса вращения.

Рисунок 4 иллюстрирует эволюцию колебательного предельного цикла при увеличении параметра κ_2 от 0,63 до 0,95. На рис. 4, а представлена однопараметрическая бифуркационная диаграмма отображения Пуанкаре, характеризующая переход от глобального регулярного квазисинхронного режима к регулярному режиму частичной квазисинхронизации через глобальный и частичный ХМК-режимы.

Стартовым состоянием бифуркационной диаграммы служит колебательный предельный цикл. При увеличении параметра κ_2 этот предельный цикл проходит через серию бифуркаций удвоения периода: при

$$\kappa_2^{(1)} = 0,638407, \kappa_2^{(2)} = 0,709835, \kappa_2^{(3)} = 0,724888, \kappa_2^{(4)} = 0,728068$$

и т. д. В результате этих бифуркаций в фазовом пространстве модели (5) рождается хаотический аттрактор с индексом вращения $[0, 0, 0]$. Он определяет глобальный ХМК-режим КФАП. Пример хаотического аттрактора, отвечающего за генерацию хаотически модулированных колебаний всеми генераторами ансамбля, с его спектром и автокорреляционной функцией представлен на рис. 4, б. Хаотичность этого аттрактора подтверждается значением максимального ляпуновского показателя:

$$\lambda_{\max} = 0,125.$$

Дальнейшее усиление связи κ_2 приводит к срыву квазисинхронного режима третьего генератора на хаотический режим биений, при этом первый и второй генераторы продолжают работать в режиме генерации ХМК. Таким образом, в ансамбле устанавливается частичный ХМК-режим. На бифуркационной диаграмме интервал значений параметра κ_2 , при которых аттрактор, порождающий отображение Пуанкаре, имеет вращение по координате φ_3 , отмечен темной линией над диаграммой. Колебательно-вращательный хаотический аттрактор с индексом вращения $[0, 0, 1]$, отвечающий за генерацию хаотически модулированных колебаний на выходе первого и второго генераторов, со спектром и автокорреляционной функцией представлены на рис. 4, в, д. Максимальный ляпуновский показатель этого аттрактора положителен и равен $\lambda_{\max} = 0,138$.

Сравнительный анализ характеристик (ляпуновских показателей, спектров и автокорреляционных функций) колебательного и колебательно-вращательного аттракторов свидетельствует, что свойства хаотически модулированных колебаний на выходе первого и второго генераторов могут быть существенно модифицированы путем изменения динамического режима третьего генератора за счет силы второй дополнительной связи.

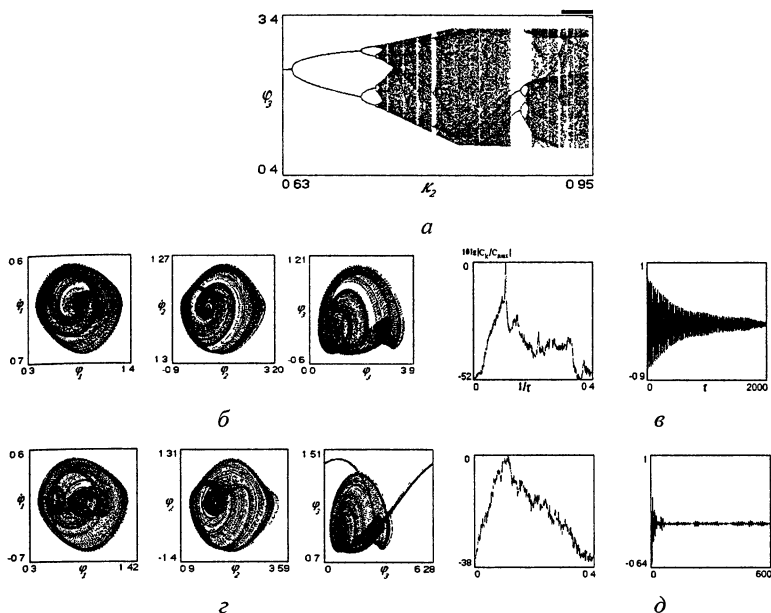


Рис. 4. Однопараметрическая бифуркационная диаграмма отображения Пуанкаре модели (5) при $\gamma_1 = \gamma_2 = \gamma_3 = 0,5$, $\kappa_1 = 0,5$, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1$ (а); проекции, спектры и автокорреляционные функции хаотических аттракторов: колебательного при $\kappa_2 = 0,8$ (б), (в), колебательно-вращательных типа $[0, 0, 1]$ при $\kappa_2 = 0,93$ (е), (д).

Переход к хаосу через удвоение инвариантных торов. На рис. 5 представлена бифуркационная диаграмма (а), зависимость ляпуновских показателей (б) и проекции отображения Пуанкаре (в–з), иллюстрирующие переход к хаотическим колебаниям через удвоение торов. На бифуркационной диаграмме по вертикали отложены локальные и глобальный минимумы отображения Пуанкаре по координате y_2 . Начальным состоянием для построения этой бифуркационной диаграммы служил инвариантный тор системы (5) при $\gamma_1 = 1,7$, $\gamma_2 = 0,3$, $\gamma_3 = 0,5$, $\kappa_1 = 1,4$, $\kappa_2 = 0,35$, $\varepsilon_1 = 0$, $\varepsilon_2 = 50$, $\varepsilon_3 = 65$. Этот аттрактор имеет индекс вращения $[1, 0, 0]$, проекция отображения Пуанкаре дана на рис. 5, в. Увеличение параметра ε_3 приводит к бифуркациям рождения торов периодов два (рис. 5, з, при $\varepsilon_3 = 69$), четыре (рис. 5, д, при $\varepsilon_3 = 70$), восемь (рис. 5, е, при $\varepsilon_3 = 70,2$) и далее возникновению хаотических аттракторов (рис. 5, ж при $\varepsilon_3 = 70,4$; рис. 5, з при $\varepsilon_3 = 70,7$). На рис. 5, б приведены зависимости двух максимальных по модулю ляпуновских показателей от параметра ε_3 , при этом еще два показателя практически (с точностью до параметров расчета) равны нулю, а еще один отрицательный. Прохождение инвариантного то-

ра через бифуркацию удвоения сопровождается практическим обнулением максимального ляпуновского показателя, а переход к хаотическому аттрактору выходом максимального показателя в область положительных значений. При $\varepsilon_3 \approx 71,84$ хаотический аттрактор разрушается, и фазовые траектории из его окрестности переходят на инвариантный тор типа $[1, 1, 1]$.

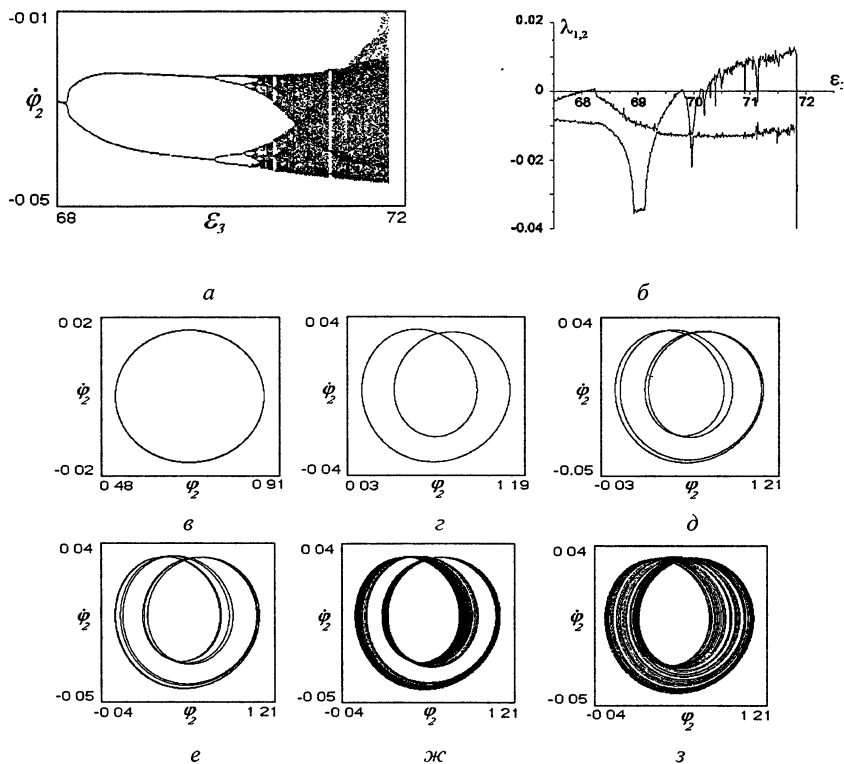


Рис. 5. Переход к хаотическим колебаниям через бифуркации удвоения двумерного тора: однопараметрическая бифуркационная диаграмма (а), зависимость двух максимальных по модулю ляпуновских показателей от параметра ε_3 (б), двумерные (ϕ_2, ψ_2) проекции отображения Пуанкаре (в-з).

В процессе численного моделирования системы (5) были обнаружены и другие сценарии перехода к хаотическим колебаниям, в частности переход к хаосу через перемежаемость первого и второго рода, через бифуркации периодических движений на торе, ведущих к его разрушению, а также бифуркации возникновения хаотического аттрактора в результате разрушения инвариантного тора как целого [29].

Заметим, что переход от синхронного режима к квазисинхронным колебаниям у генераторов, объединенных в ансамбль, может быть осуществлен путем организации дополнительных связей. При этом процесс возбуждения протекает мягко, не сопровождаясь бифуркационными явлениями в фазовом пространстве математической модели, а структура возникающих колебаний определяется колебаниями цепи управления генератора, откуда поступает дополнительное управление.

2. Синхронизация хаотически модулированных колебаний

Рассмотрим теперь задачу синхронизации хаотических колебаний двух однонаправленно связанных генераторов хаотических колебаний – задающего и синхронизируемого. Представим ФАП1 в качестве задающего генератора ХМК и однокольцевую ФАП2 в качестве синхронизируемого генератора ХМК, причем в ФАП1 и ФАП2 выбираются параметры близкие, но не одинаковые. Под синхронизацией двух ХМК будем понимать «полную синхронизацию», при которой выходные координаты двух хаотических генераторов, оставаясь хаотическими, полностью совпадают. Для достижения синхронизации введем однонаправленные связи ФАП1 с ФАП2 по координатам φ , y , z разностного типа. На рис. 6 представлен вариант соединения ФАП1 и ФАП2 однонаправленной связью через дополнительный фазовый детектор ФД₁₂. На выходе ФД₁₂ образуется сигнал $\delta_\varphi \sin(\varphi_2 - \varphi_1)$, который используется для управления частотой генератора Г2. Для достижения «полной хаотической синхронизации» можно попытаться ввести связи по другим координатам. Например, можно сделать вариант соединения ФАП1 и ФАП2 через дополнительный частотный дискриминатор ЧД₁₂, в этом случае на его выходе формируется сигнал $\delta_y \Phi(y_2 - y_1)$, поступающий в цепь управления генератора Г2. В первом приближении характеристику частотного дискриминатора можно считать линейной и принять $\delta_y \Phi(y_2 - y_1) = \delta_y(y_2 - y_1)$. Наконец, можно осуществить однонаправленную связь и по координате z , так, чтобы в цепь управления Г2 поступал сигнал $\delta_z(z_2 - z_1)$.

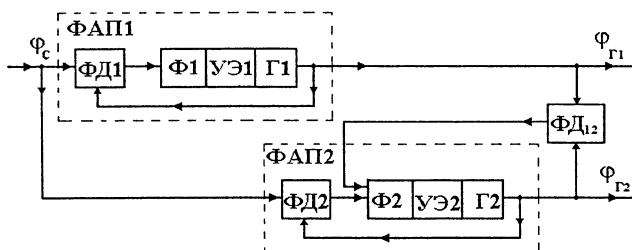


Рис. 6. Синхронизация двух однокольцевых ФАП

На рис. 7 даны результаты моделирования процесса синхронизации хаотических колебаний ФАП1 и ФАП2 [27]. Начальные условия: $y_1^0 = \gamma_1$, $y_2^0 = \gamma_2$, $z_1^0 = 0$, $z_2^0 = 0$, φ_1^0 , φ_2^0 – произвольные. Параметры γ , ε , μ выбирались в ФАП1 и ФАП2 не одинаковыми, но достаточно близкими (отличия в значениях параметров составляли 1 – 5 %). При отсутствии связей (или введении только δ_φ) синхронизации колебаний в ФАП1 (рис. 7, а) и в ФАП2 (рис. 7, б) нет (рис. 7, в). С введением связей колебания в ФАП2 подстраиваются под колебания в ФАП1, причем точность подстройки (синхронизации) увеличивается с ростом коэффициентом связей (рис. 7, г – линейная связь по z , рис. 7, д – линейная связь по y , рис. 7, е, ж, з – смешанное линейное управление по z и y , рис. 7, и – нелинейное управление по y). Однако, следует иметь в виду, что введение достаточно больших значений коэффициентов связей может привести к срыву синхронизации и переходу системы ФАП2 на другой тип аттрактора. По-видимому, это связано с малостью области параметров, соответствующих генерации рассматриваемых хаотических колебаний.

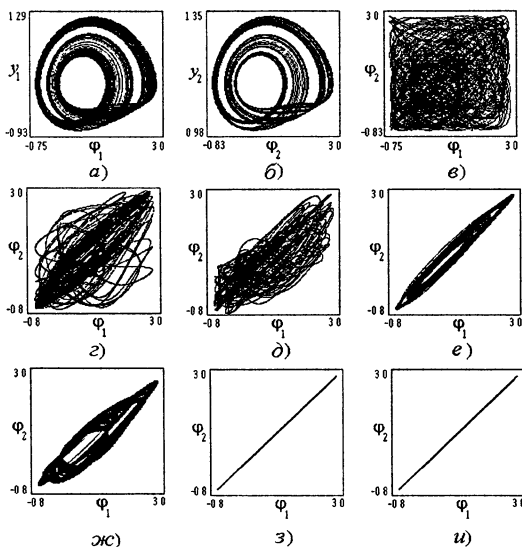


Рис. 7. Синхронизация хаотических колебаний в схеме на рис. 6 в зависимости от связей

Таким образом, в результате компьютерного моделирования показана возможность осуществления режима полной синхронизации ХМК в двух неидентичных ФАП с близкими значениями параметров. Аналогичное рассмотрение возможности подстройки ХМК проведено и для случаев, когда генераторами ХМК являются ансамбли связанных ФАП [27].

3. Передача информации с использованием динамического хаоса

Опираясь на полученные результаты по генерации и синхронизации ХМК, можно предложить различные варианты передачи информации с использованием ХМК в качестве несущих колебаний. В [17, 27] представлен вариант моделирования процесса передачи аналогового сигнала. Рассмотрим вариант построения системы передачи бинарного сигнала. Используя метод переключения хаотических сигналов [8], можно построить передатчик на основе двух генераторов ХМК1 и ХМК2, переключаемых в соответствии с заданным информационным бинарным сигналом, а в приемнике использовать генератор ХМК3, параметры которого выбраны так, что он может быть синхронизован одним из генераторов передатчика, например ХМК1, и не может быть синхронизован другим – ХМК2. Результаты компьютерного моделирования такого способа передачи информации показаны на рис. 8 (проекции аттракторов, соответствующих ХМК1 и ХМК2, входной информационный сигнал в передатчике и выходной сигнал в приемнике).

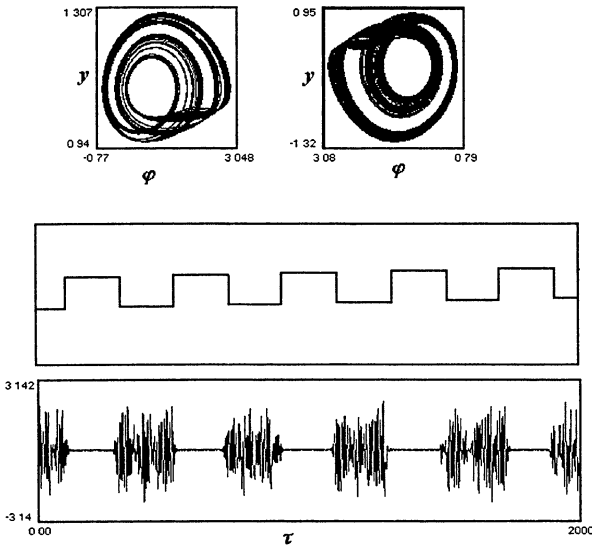


Рис. 8. Передача бинарного сигнала

В результате проведенных компьютерных экспериментов можно сделать вывод о возможности использовании малых ансамблей связанных систем ФАП для создания генераторов хаотически модулированных колебаний, их синхронизации и построения на этой основе систем связи с использованием хаотической несущей.

Работа выполнена при финансовой поддержке РФФИ (грант 02-02-17573) и Программ «Ведущие научные школы России» (грант 00-15-96582), «Университеты России» (проект УР.03.01.027).

Литература

1. Pecora L. M., Carroll T. L. Synchronization in chaotic system // Phys. Rev. Lett. 1990. V. 64, № 8. P. 821-824.
2. Oppenheim A. L., Wornell G. W., Isabell S. H., Cuomo K. M. Signal processing in the context of chaotic systems // Proc. IEEE ICASSP. 1992. V. 6. P. 117-120.
3. Kocarev L., Halle K. S., Eckert K., Chua L. O., Parlitz U. Experimental demonstration of secure communication via chaotic synchronization // Int. J. Bifurcation and Chaos. 1992. V. 2. № 3. P. 709-713.
4. Hayes S., Grebogi G., Ott E. Communication with Chaos // Phys. Rev. Lett. 1993. V. 70. P. 3031-3034.
5. Волковский А. П., Рульков Н. Ф. Синхронный хаотический отклик нелинейной колебательной системы как принцип детектирования информационной компоненты хаоса // Письма в ЖТФ. 1993. Т. 9, № 3. С. 71-75
6. Козлов А. К., Шалфеев В. Д. Избирательное подавление детерминированных хаотических сигналов // Письма в ЖТФ. 1993. Т. 19, № 3. С. 83-86.
7. Dmitriev A. S., Panas A. I., Starkov S. O. Experiments on speech and music signals transmission using chaos // Int. J. Bifurcation and Chaos. 1995. V. 5. P. 1249-1254.
8. Dedieu H., Kennedy M. P., Hasler M. Chaos shift keying: modulation and demodulation of a chaotic carrier using self-synchronizing Chua's Circuits // IEEE Trans CAS. 1993. V. 40, № 10. P. 634-642.
9. Hasler M. Engineering chaos for secure communication systems // Phyl. Trans R Soc. Lond A. 1995. V. 353. P. 115-126.
10. Kennedy M. P. Communication with chaos: state of the art and engineering challenges // Proc. NDES96 Seville. 1996. P. 1-8.
11. Дмитриев А. С., Панас А. И., Старков С. О. Динамический хаос как парадигма современных систем связи // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 1997 № 10 С. 4-26.
12. Шалфеев В. Д., Осипов Г. В., Козлов А. К., Волковский А. П. Хаотические колебания – генерация, синхронизация, управление // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники 1997 № 10. С. 27-49.
13. Boccaletti S., Farini A., Arecchi F. N. Adaptive synchronization of chaos for secure communication // Phys. Rev E 1997. V. 55, № 5. P. 4979-4981.
14. Капранов М. В., Кулешов В. Н., Ларионова М. В., Морозов А. Г., Удалов Н. Н. Свойства передачи информации с манипуляцией параметрами и начальными условиями генераторов хаотических колебаний // Зарубежная электроника. Успехи современной радиоэлектроники 2000 № 11. С. 48-60.
15. Sato A., Endo T. Experiments of secure communications of phase-locked loops // Proc. NDES'94. 1994. P. 117-122
16. Smyth N., Crowley C., Kennedy M. P. Improved receiver for CSK spread spectrum communications using analog phase locked loop chaos // Proc. NDES'96 Seville. 1996. P. 27-32.

17. *Korzinova M. V., Matrosov V. V., Shalfeev V. D.* Communications using cascade coupled phase-locked loops chaos // *Int. J. Bifurcation and Chaos*. V. 9, № 5. P. 963-973.
18. *Шалфеев В. Д., Матросов В. В., Корзинова М. В.* Каскадное соединение фазовых систем – генерация и синхронизация хаотических колебаний // *Вестник ННГУ. Нелинейная динамика – синхронизация и хаос* III. Н. Новгород, 1998. С. 35-44.
19. *Kolumban G., Vizvari B.* Nonlinear dynamics and chaotic behaviour of the analog phase-locked loop // *Proc. NDES'95*. Dublin. 1995. P. 99-102.
20. *Shalfeev V. D., Matrosov V. V., Korzinova M. V.* Chaos in Phase Systems: Generation and Synchronization. // *Controlling Chaos and Bifurcations in Engineering Systems* / Ed. G. Chen. CRC Press. Boca - Raton - London – New -York - Washington, D.C. 2000. P. 529-558.
21. *Shalfeev V. D., Matrosov V. V.* Dynamical chaos in phase-locked loops // *Chaos in circuits and systems* / Eds. G. Chen and T. Ueta. World Scientific Publishing Company. Singapore. 2002. P. 130-150.
22. *Матросов В. В.* Регулярные и хаотические колебания в фазовой системе // *Письма в ЖТФ*. 1996 Т 22, № 23. С.4-8.
23. *Матросов В. В., Корзинова М. В.* Коллективная динамика каскадного соединения фазовых систем // *Изв. вузов. Прикладная нелинейная динамика*. 1994. № 2. С. 10-16.
24. *Матросов В. В., Шалфеев В. Д.* О влиянии связей на нелинейную динамику двух каскадно-связанных управляемых генераторов // *Изв. вузов. Радиофизика* 1995. Т. 38, № 3-4. С. 275-279.
25. *Матросов В. В., Корзинова М. В.* Синхронные и автоколебательные режимы каскадного соединения фазовых систем // *Вестник ННГУ. Нелинейная динамика – синхронизация и хаос* Н Новгород 1996 С. 77-92.
26. *Матросов В. В.* Некоторые особенности динамического поведения каскадного соединения двух фазовых систем // *Изв. вузов Прикладная нелинейная динамика*. 1997. Т. 5, № 6 С. 52-61.
27. *Шалфеев В. Д., Матросов В. В., Корзинова М. В.* Динамический хаос в ансамблях связанных фазовых систем // *Зарубежная радиоэлектроника Успехи современной радиоэлектроники*. 1998 № 11 С. 44-56
28. *Афраймович В. С., Неколкин В. И., Осипов Г. В., Шалфеев В. Д.* Устойчивость, структуры и хаос в нелинейных сетях синхронизации Горький: ИПФ АН СССР, 1989
29. *Matrosov V. V., Kasatkin D. V.* The Dynamics of three cascade coupled phase systems // *Proc of the Int. Conf. dedicated to the 100th Anniversary of A. A. Andronov*. Nizhny Novgorod, Russia, 2002. P. 225-227.

МОДЕЛИ НЕЙРОПОДОБНЫХ СИСТЕМ. ДИНАМИЧЕСКИЕ РЕЖИМЫ ПРЕОБРАЗОВАНИЯ ИНФОРМАЦИИ

В. Г. Яхно

1. Введение

Сложное поведение многих природных систем определяется различными вариантами адаптивных реакций на разнообразные внешние сигналы. Такие системы обычно состоят из взаимосвязанных однотипных активных элементов. К наиболее известным системам преобразования информационных сигналов относятся биомембраны, нейроны, нейронные ансамбли, нервные ткани, искусственные аналоги нейронных систем, и многие другие активные распределенные системы.

Обычно распределенные системы называют *нейроподобными*, если они состоят из активных элементов с *несколькими устойчивыми (или «квазистойчивыми») состояниями* и между такими неравновесными элементами существуют *нелокальные пространственные связи* (см., например, [1–15]).

Среди различных моделей можно выделить несколько групп моделей для нейроподобных систем. Сразу же отметим, что любое предлагаемое разбиение на «базовые» модели может быть выполнено неоднозначным образом. Варианты в сильной степени зависят от представлений исследователя о важности и необходимой степени «огрубления» тех или иных деталей изучаемых процессов.

К первой группе относятся модели однородных одно-, двух-, трех- и более компонентных неравновесных, нейроподобных систем. Каждая из компонент в таких моделях может, в частности, характеризоваться своими схемами активных механизмов и значениями релаксационных временных параметров.

Такие модели описывают параллельное преобразование входных воздействий на основе возможных комбинаций из «базовых» структур коллективной активности – элементарных нелинейных структур. Процессы формирования и взаимодействия таких структур используются для описания динамики однородных нейроподобных систем.

Ко второй группе относятся модели «элементарных» классификаторов или ячеек принятия решений с фиксированными алгоритмами, в которых возможно управление алгоритмами обработки изображений с целью оптимизации режимов распознавания как под особенности вида сигнала, так и под задачи, решаемые адаптивной распознающей системой.

Операции обработки фрагментов сложного изображения в блоках моделей адаптивных распознавателей могут осуществляться в параллель-

ном режиме. Для этого используются модели неравновесных распределенных систем первой группы. Точность принимаемых решений и оптимальность функционирования адаптивных распознавателей зависит от алгоритмов, адекватно соответствующих задаче, и оптимальных параметров работы этих алгоритмов.

На основе рассматриваемых в этой работе двух групп базовых моделей возможно конструирование распознающих систем с более точными операциями кодирования, например, за счет подстройки архитектуры системы, автоматического выбора алгоритмов под временные и ситуационные особенности обрабатываемого сигнала, а также за счет формирования ассоциаций между потоками сигналов различных модальностей (видео, акустические, тактильные, химические сигналы и т. п.).

Важно отметить, что изучение режимов реагирования базовых моделей нейроноподобных систем относится к классической области интересов нелинейной динамики, связанной как с классификацией моделей, так и с классификацией динамических режимов для каждой группы базовых моделей.

2. Модели однородных нейроноподобных систем

Нейроноподобные системы представляют собой слои из взаимодействующих активных элементов (нейронных ансамблей) с пространственно распределенными активирующими и угнетающими (тормозными) воздействиями. Особенности вывода уравнений для усредненной активности нейроноподобных элементов с *классическим* типом связей рассматривались в работах [1–11].

Известно существование однородных участков в нервной системе головного мозга. Таким образом, например, организованы периферические отделы анализаторных систем. Представления об однородности могут быть с успехом распространены на стабильные по функциональной организации участки нервной системы, предназначенные для параллельной обработки потока сенсорной информации.

При создании однородной модели эти структуры представляются состоящими из плоских слоев, образованных одинаковыми по своим свойствам элементами, которые связаны определенным образом с соседями. Опираясь на результаты работы [8], рассмотрим простейший, одномерный вариант модели, описывающий изменение уровней активационных процессов в аксонных, дендритных деревьях и в телах нейронов с помощью усредненных переменных. Основные параметры такой распределенной системы будем считать однородными и оценивать усредненными по пространству величинами. Схема одномерной структуры взаимосвязей между элементами в классической нейронной сети из возбуждающих и тормозных нейронов показана на рис. 1.

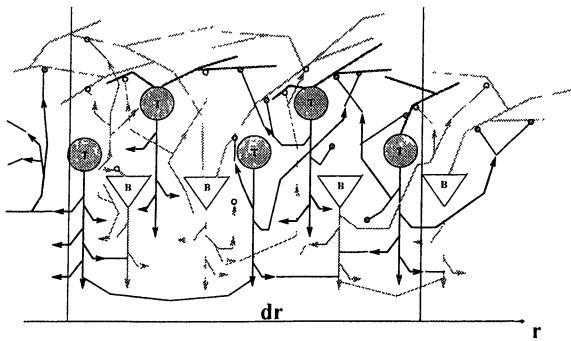


Рис. 1. Схема связей между дендритными и аксонными деревьями в ансамбле нейроноподобных элементов с возбуждающими (белые) и тормозными (темные) влияниями

Нейроны связаны между собой через густую сеть аксонных и дендритных нервных волокон. Импульсы, генерируемые телами нейронов, по древовидной сети аксонных волокон распространяются до синапсов, которые выделяют нейромедиаторы (возбуждающие или тормозные, в зависимости от вида нейрона). Нейромедиаторы действуют на мембрану дендритных волокон, и связанные с этими воздействиями изменения электрических потенциалов передаются на тело нейрона. В результате на мембране тела нейрона формируется так называемый постсинаптический потенциал (ПСП).

Для возбуждающих нейронов постсинаптический потенциал обозначают как ВПСП, а для тормозных нейронов — ТПСП. При превышении ВПСП или ТПСП соответствующей пороговой величины t_{0B} или t_{0T} , мембрана тела нейрона начинает генерировать импульсы, которые в свою очередь снова распространяются по аксонному дереву и передают свои сигналы на дендриты других нейронов через соответствующие нейромедиаторы.

Характерные изменения $V(t)$ — напряжения на мембране пейсмекерных клеток — показаны на рис. 2.

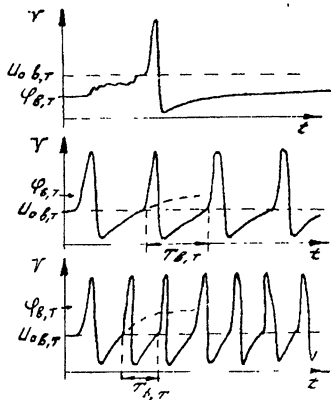


Рис. 2. Потенциалы действия, генерируемые мембраной тела нейрона

Длительность импульса (потенциала действия) порядка нескольких миллисекунд, амплитуда порядка 0,1 В. Потенциалы действия разделены участками, где напряжение меняется медленно. За возбужденным состоянием следует состояние рефрактерности, в течение которого нейрон не способен реагировать на приходящие к нему сигналы возбуждения. Через некоторое время способность к возбуждению восстанавливается, и нейрон возвращается к состоянию покоя. Волна потенциала действия распространяется вдоль аксона со скоростью от нескольких метров до десятков метров в секунду. Интенсивность сти-

мула, действующего на нейрон, определяет частоту генерируемых импульсов. Отметим, что в таких клетках частота генерации импульсов зависит от соотношения между порогом $t_{0B,T}$ и величиной действующего ВПСП или ТПСП. Фазы импульсов, генерируемых различными нейронами в ансамблях, считаются распределенными случайно.

Итоговое изменение потенциала на теле нейрона, получившееся в результате временного и пространственного суммирования синаптических сигналов, зависит от частот нервных импульсов, приходящих на дендритное дерево данного нейрона по аксонным волокнам от других нейронов.

В рассматриваемой модели параметры нейронного ансамбля и входящих в него нейронов считаются однородными и оцениваются усредненными по пространству величинами. Усреднение проводится на длинах порядка 0,1 мм, много меньших, чем характерные размеры нейронного ансамбля (до нескольких сантиметров), но значительно превышающих размеры тела отдельного нейрона ($\sim 0,01$ мм).

Введен следующий набор усредненных переменных:

$u_B(t, \bar{r})$ – средняя доля активных («зажженных») волокон в единице объема в аксонных деревьях возбуждающих нейронов;

$u_T(t, \bar{r})$ – средняя доля активных волокон в единице объема в аксонных деревьях тормозных нейронов;

ϕ_B – величина среднего ВПСП для нейронов с возбуждающим медиатором;

ϕ_T – величина среднего ТПСП для нейронов с тормозным медиатором;

$E_{B,T} []$ – средние числа возбуждающих или тормозных нейронов в единице объема, у которых постсинаптический потенциал (ПСП) превосходит пороговое значение активации нейронов;

t_{0B} и t_{0T} – соответственно средние пороговые значения активации возбуждающих и тормозных нейронов;

Для этих переменных составляются два набора уравнений, описывающих прохождение сигналов по аксонным и дендритным деревьям нейронного ансамбля. Первый набор уравнений выводится из баланса численности «зажженных» волокон в аксонных деревьях, причем уравнения баланса составляются как для импульсов, бегущих в волокнах слева направо ($\bar{\rho}_{B,T}$), так и для импульсов, бегущих справа налево ($\bar{\rho}_{B,T}$). При этом изменение доли волокон с импульсами, бегущими слева направо, определяется балансным уравнением (см. схему для выделенного отрезка r до $r + dr$ на рис. 1):

$$d\bar{\rho}_{B,T}/dt = -\Gamma_1 \bar{\rho}_{B,T} - \Gamma_2 \bar{\rho}_{B,T} + \Gamma_3 \bar{\rho}_{B,T} + \Gamma_4 \bar{\rho}_{B,T} + v_{B,T} \cdot (d\bar{\rho}_{B,T}/dr) + \frac{E_{B,T}[\phi_{B,T}(r, t)]}{T_{B,T}} \quad (1)$$

Слагаемые в правой части уравнения означают:

$-\Gamma_1 \bar{\rho}_{B,T}$ – убыль импульсов за счет достижения ими синапсов;

$-\Gamma_2 \bar{\rho}_{B,T}$ – убыль импульсов за счет поворота волокон аксонных деревьев и перехода их в импульсы, бегущие справа налево;

$+\Gamma_3 \bar{\rho}_{B,T}$ – увеличение количества волокон с импульсами за счет ветвления аксонных волокон с прежним направлением;

$+\Gamma_4 \bar{\rho}_{B,T}$ – увеличение доли волокон с импульсами за счет поворота волокон аксонных деревьев и перехода импульсов, бежавших справа налево в импульсы бегущие слева направо;

$v_{B,T} \cdot (d\bar{\rho}_{B,T}/dr)$ – разница между количеством «зажженных» волокон с левой и правой сторон выделенного отрезка от r до $r + dr$; $v_{B,T}$ – средние величины проекций скоростей импульсов возбуждения в аксонных волокнах на ось r ($v_{B,T} \sim 5$ м/с [5–8, 16–17]);

$E_{B,T} [\phi_{B,T}(r, t)]/T_{B,T}$ – увеличение количества волокон с импульсами внутри отрезка от r до $r + dr$ за счет генерации импульсов нейронами с ПСП превышающим порог возбуждения; предполагается, что средние длины аксонных деревьев значительно превышают длину отрезка усреднения; $T_{B,T}$ – средний период, с которым импульсы генерируются нейронами, если их ПСП превышает порог возбуждения (см. рис. 2) ($T_B \sim 10$ мс; $T_T \sim 5$ мс [5–8, 16–17]).

Аналогичное уравнение баланса записывается для изменяющейся доли волокон с импульсами, бегущими справа налево:

$$d\bar{\rho}_{B,T}/dt = -\Gamma_1 \bar{\rho}_{B,T} - \Gamma_4 \bar{\rho}_{B,T} + \Gamma_3 \bar{\rho}_{B,T} + \Gamma_2 \bar{\rho}_{B,T} + v_{B,T} \cdot (d\bar{\rho}_{B,T}/dr) + \frac{E_{B,T} [\phi_{B,T}(r, t)]}{T_{B,T}} \quad (2)$$

Зависимости $E_{B,T}[\]$, описывающие генерацию импульсов нейронами, можно переписать в виде

$$E_{B,T}[\] = P_{B,T} F_{B,T}[-t_{0B,T} + \phi_{B,T}],$$

где $P_{B,T}$ – плотность нейронов ($P_B \sim 2 \cdot 10^{-5}$ 1/мкм³; $P_T \sim 10^{-7}$ 1/мкм³ [59, 63, 72]), а нелинейные, нормированные на единицу функции $F_B[-t_{0B} + \phi_B]$ и $F_T[-t_{0T} + \phi_T]$ имеют ступенчатый вид около значений $(t_{0B} - \phi_B) = 0$ и $(t_{0T} - \phi_T) = 0$ (рис. 3).

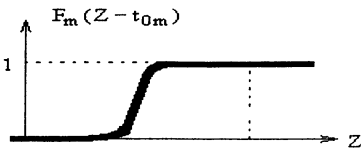


Рис. 3. Характерный вид нелинейной функции, описывающей процесс генерации импульсов

Второй набор уравнений определяется процессами распространения сигналов в дендритных волокнах.

Если пренебречь временем задержки передачи сигнала по дендритному дереву, то усредненные величины ПСП ϕ_B и ϕ_T в зависимости от числа «зажженных» волокон различного типа, окружающих «усредненную клетку», могут быть записаны с помощью уравнений

$$\phi_B = \int_{-\infty}^{+\infty} \Phi_{BB}(\bar{\xi} - \bar{r}) u_B(t, \bar{\xi}) d\bar{\xi} - \int_{-\infty}^{+\infty} \Phi_{BT}(\bar{\xi} - \bar{r}) u_T(t, \bar{\xi}) d\bar{\xi}, \quad (3)$$

$$\phi_T = \int_{-\infty}^{+\infty} \Phi_{TB}(\bar{\xi} - \bar{r}) u_B(t, \bar{\xi}) d\bar{\xi} - \int_{-\infty}^{+\infty} \Phi_{TT}(\bar{\xi} - \bar{r}) u_T(t, \bar{\xi}) d\bar{\xi}. \quad (4)$$

Функции $\Phi_{mp}(\xi - \bar{r})$ задают величины пространственных воздействий на потенциал мембраны тела «усредненного нейрона» ($m = B, T$) в точке \bar{r} , со стороны аксонных окончаний ($p = B, T$), от «усредненных нейронов», расположенных в точке ξ . Поскольку сигнал на дендриты одинаково передается синапсами от аксонов, приходящих с разных направлений, то при определении суммарной величины ПСП используются переменные $u_{B,T}(t, \bar{r}) = (\bar{\rho}_{B,T} + \bar{\rho}_{B,T})$. При введении дополнительных переменных $u'_{B,T}(t, \bar{r}) = (\bar{\rho}_{B,T} - \bar{\rho}_{B,T})$ и предположении равноправности правого и левого направлений в структуре сети (т. е. изотропности архитектуры связей в одномерном ансамбле, когда $\Gamma_2 = \Gamma_4$, $\Gamma_3 = \Gamma_3$) уравнения (1) и (2) можно переписать (складывая и вычитая их) в следующем виде:

$$du'_{B,T}(t, r)/dt = -\frac{1}{\tau_{2B,T}} u'_{B,T}(t, \bar{r}) + v_{B,T} (du'_{B,T}(t, r)/dr) + 2P_{B,T} \frac{E_{B,T} [\phi_{B,T}(r, t)]}{T_{B,T}}, \quad (5)$$

$$du_{B,T}(t, r)/dt = -\frac{1}{\tau_{2B,T}} u_{B,T}(t, \bar{r}) + v_{B,T} (du_{B,T}(t, r)/dr), \quad (6)$$

где $\tau_{1B,T} = 1/(\Gamma_1 - \Gamma_3)$ – характерное время последствия, т. е. среднее время между генерацией импульса нейроном и реакцией дендритов, на котором оканчиваются аксонные окончания этого нейрона. Это время включает в себя время задержки импульсов на синапсах и время, в течение которого волокна в аксонном дереве могут находиться в «зажженном» состоянии в случае одновременного прекращения активности всех нейронов ($\tau_{1B} \sim 20$ мс; $\tau_{1T} \sim 200$ мс [5–8, 16–17]); $\tau_{2B,T} = 1/(\Gamma_1 + 2\Gamma_2 - \Gamma_3)$ – среднее время, за которое фронт активации в аксоне проходит расстояние между точками ответвления коллатералей, поворачивающих в обратную сторону ($\tau_{2B} \sim 10^{-2}$ мс; $\tau_{2T} \sim 2 \cdot 10^{-2}$ мс [5–8, 16–17]).

При этом для изотропного случая, когда количество импульсов, бегущих в волокнах слева направо, может быстро сравняться с количеством импульсов, бегущих справа налево, получают уравнения следующего вида:

$$\partial u_B(t, \vec{r}) / \partial t = -\frac{1}{\tau_{1B}} u_B(t, \vec{r}) + D_B \cdot \partial^2 u_B(t, \vec{r}) / \partial x^2 + \beta_{I/B} F_B[-t_{0B} + \phi_B], \quad (7)$$

$$\partial u_T(t, \vec{r}) / \partial t = -\frac{1}{\tau_{1T}} u_T(t, \vec{r}) + D_T \cdot \partial^2 u_T(t, \vec{r}) / \partial x^2 + \beta_{I/T} F_T[-t_{0T} + \phi_T], \quad (8)$$

где $D_i = v_i^2 \tau_{2i}$, $i = B$ или T для соответственно возбуждающих или тормозных нейронов, а $\beta_{I/B,T} = 2P_{B,T} / T_{B,T}$.

Решение системы уравнений (3)–(4) и (7)–(8) при заданном начальном условии $u_{B,T}(t=0, r) = u_{0B,T}(r)$ описывает пространственно-временную динамику в такой однородной нейроноподобной системе. Такая модельная система представляет собой упрощенный вариант уравнений, приведенных в работах Сбитнева [5–7]. Основное отличие от модели Вильсона и Коузена [2] заключается в том, что в последней пространственные связи между нейронами предполагаются неизменными во времени, тогда как на самом деле среднее число связей (волокон в «зажженном» состоянии) меняется и время их изменений порядка $\tau_{2B,T}$.

Рассмотрим ряд возможных вариантов моделей при различном виде основных параметров системы.

В первом случае предположим, что размер дендритных деревьев много меньше $v_{B,T} \sqrt{\tau_{2B,T} \cdot \tau_{1B,T}}$ – диффузионной длины за счет аксонных связей. Тогда от модели в интегродифференциальной форме можно перейти к уравнениям типа «реакция–диффузия»:

$$\partial u(t, \vec{r}) / \partial t = D_B \cdot \partial^2 u(t, \vec{r}) / \partial x^2 + \Psi(u, g), \quad (9)$$

$$\partial g(t, \vec{r}) / \partial t = D_T \cdot \partial^2 g(t, \vec{r}) / \partial x^2 + \varphi(u, g), \quad (10)$$

в которых уровень активности тормозных нейронов переобозначен через переменную $g(t, \vec{r})$, а функции $\Psi(u, g)$ и $\varphi(u, g)$ определяются особенностями механизмов неравновесности в рассматриваемой системе. Для нейроноподобной системы с набором M популяций возбуждающих нейронов со своими порогами и характеристиками дендритных деревьев и I популяций для тормозных нейронов нелинейные функции записываются в виде

$$\Psi(u, g) = -k_{11} u(t, \vec{r}) + \sum_{m=1}^M k_{21m} F_{1m}[-t_{0m} + \gamma_{11m} u(t, \vec{\xi}) - \gamma_{12m} g(t, \vec{\xi})], \quad (11)$$

$$\varphi(u, g) = -k_{12} g(t, \vec{r}) + \sum_{m=1}^I k_{22m} F_{2m}[-t_{0m} + \gamma_{21m} u(t, \vec{\xi}) - \gamma_{22m} g(t, \vec{\xi})]. \quad (12)$$

Варианты таких зависимостей для наиболее интересных динамических реакций в случае одного и двух видов возбуждающих нейронов и

одного вида тормозных нейронов показаны на рис. 4. Видно, что пороговые характеристики различных популяций нейронов позволяют формировать многоуровневую активную систему. При соответствующем подборе параметров в такой системе можно сформировать различные варианты триггерных, ждущих либо автоколебательных режимов отдельных нейроноподобных элементов.

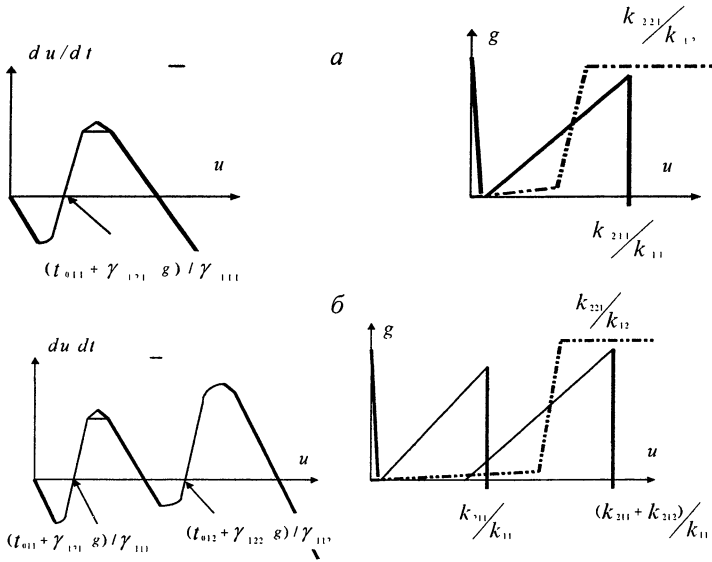


Рис. 4. Варианты вида нелинейной функции $\Psi(u, g)$ при заданных значениях переменной $g(t, \bar{r})$ и нуль-изоклин $\Psi(u, g) = 0, \phi(u, g) = 0$ в уравнениях (9)–(12), используемых для описания сосредоточенного нейроноподобного элемента *а* – случай взаимодействия одного вида возбуждающих нейронов и одного вида тормозных нейронов, *б* – случай совместной работы двух ансамблей возбуждающих нейронов и одного ансамбля тормозных нейронов

Разнообразные варианты моделей могут быть сформированы за счет изменения вида функций пространственных связей, задаваемых видом дендритных деревьев и характером расположения на них синаптических окончаний от аксонных деревьев. Для случая, когда активность популяции тормозных нейронов описывается линейной функцией и влияние тормозных нейронов на себя мало ($\Phi_{T,T}(\xi - \bar{r}) \sim 0$), а также размер дендритных деревьев тормозных нейронов много больше $v_T \sqrt{\tau_{2T} \cdot \tau_{1T}}$ – диффузионной длины – за счет аксонных связей, системы уравнений (3)–(4) и (7)–(8) можно свести к упрощенному варианту однокомпонентной «базовой» модели с латеральными (боковыми в пределах слоя) связями:

$$g / \tau_2 = \beta_{F2} \left[-t_{0T} + \int_{-\infty}^{\infty} \Phi_{T,B}(\xi - r) u(\xi, t) d\xi \right], \quad (13)$$

$$\frac{du}{dt} = -\frac{u}{\tau_1} + \beta_{F2} F_B \left[-t_{0B} + \int_{-\infty}^{\infty} \Phi_{B,B}(\xi - r) u(\xi, t) d\xi - \int_{-\infty}^{\infty} \Phi_{B,T}(\xi - r) g(\xi, t) d\xi \right] + D_1 \frac{\partial^2 u(t, \vec{r})}{\partial r^2}, \quad (14)$$

или в виде одного уравнения:

$$\frac{du}{dt} = -\frac{u}{\tau_1} + \beta_{F1} F_B \left[-t_{01} + \int_{-\infty}^{\infty} \Phi_1(\xi - r) u(\xi, t) d\xi + D_1 \frac{\partial^2 u(t, \vec{r})}{\partial r^2} \right], \quad (15)$$

где

$$\Phi_1(\xi - \vec{r}) = \Phi_{BB}(\xi - \vec{r}) - \tau_2 \beta_{F2} \int_{-\infty}^{+\infty} \Phi_{T,B}(\xi - \eta) \Phi_{B,T}(\eta - \vec{r}) d\xi, \quad (16)$$

$$t_{01} = t_{0B} - t_{0T} \tau_2 \beta_{F2} \int_{-\infty}^{+\infty} \Phi_{B,T}(\xi) d\xi.$$

Выполненные преобразования из (3)–(4), (7)–(8) в (15) показывают, что сеть взаимосвязанных возбуждающих и тормозных нейронов может работать как единый ансамбль из активных элементов, обладающий достаточно сложной, немонотонной зависимостью (16) вида функции пространственных связей. Примеры вариантов формирования немонотонных функций пространственных связей показаны на рис. 5.

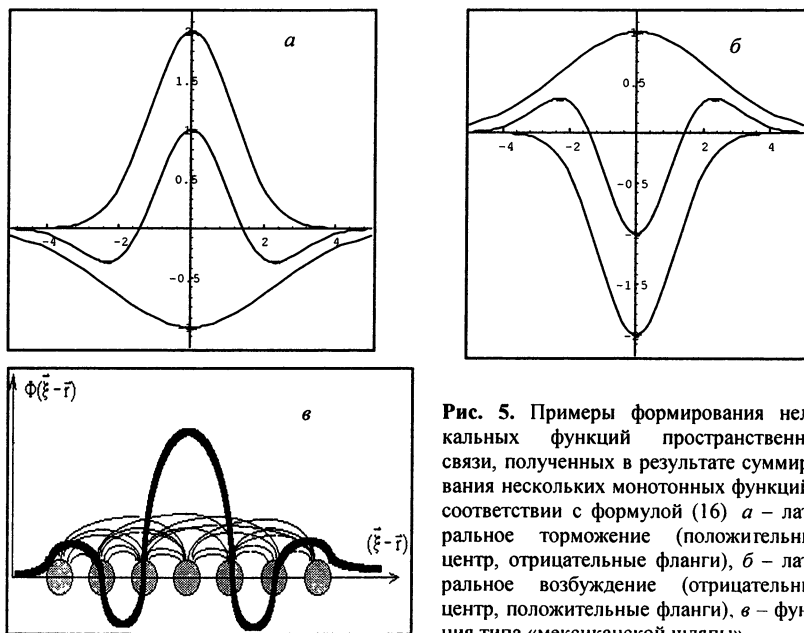


Рис. 5. Примеры формирования нелокальных функций пространственной связи, полученных в результате суммирования нескольких монотонных функций в соответствии с формулой (16) *а* – латеральное торможение (положительный центр, отрицательные фланги), *б* – латеральное возбуждение (отрицательный центр, положительные фланги), *в* – функция типа «мексиканской шляпы»

Другой вариант модели, соответствующий случаю двухкомпонентной системы:

$$\frac{du}{dt} = -\frac{u}{\tau_1} + \beta_{F1}(g) F_1 \left[-t_1(g) + \int_{-\infty}^{\infty} \Phi_1(\xi - r, g) u(\xi, t) d\xi \right] + D_1 \frac{\partial^2 u(t, \vec{r})}{\partial r^2}, \quad (17)$$

$$\frac{dg}{dt} = -\frac{g}{\tau_2} + \beta_{F2}(g) F_2 \left[-t_2(g) + \int_{-\infty}^{\infty} \Phi_2(\xi - r, g) u(\xi, t) d\xi \right] + D_2 \frac{\partial^2 g(t, \vec{r})}{\partial r^2} \quad (18)$$

может быть получен в предположении совместной работы многих ансамблей возбуждающих и тормозных нейронов. Одни ансамбли обеспечивают формирование и изменения функций пространственной связи, другие – изменения величин порогов $t_{1,2}$, или $\tau_{1,2}^{-1}$; $\beta_{F1,2}(g)$ – плотность активированных нейронных клеток (это возможно, если аксон-аксонные синапсы блокируют распространение импульсов от возбужденных нейронов). Результат работы многих ансамблей может быть описан с помощью двух наиболее важных (или можно назвать эффективных) переменных $u(t, \vec{r})$ и $g(t, \vec{r})$.

В таблице приведены примеры возможного изменения основных характеристик двухкомпонентных уравнений при формировании упрощенного вида «базовых» моделей.

$\frac{\tau_u}{\tau_g}$	$\tau_{F_{1,2}}(g)$	$u = \frac{\tau_u}{\tau_g} F_1(u, g)$ $g = \frac{\tau_g}{\tau_u} F_2(u, g)$	$F_{1,2}[z]$	$t h_1(g)$	$\Phi_1(\xi)$	$\Phi_2(\xi)$	$\alpha = \alpha(g)$	ДИСКРЕТНОСТЬ ПО		
								(t, r)	(t)	(r)
$\frac{\tau_u}{\tau_g} \ll 1$	$\tau_{F_{1,2}} = \text{const}$									
$\frac{\tau_u}{\tau_g} = 1$										

Учитывая предположения, сделанные при выводе данного набора базовых моделей, нетрудно составить перечень возможных вариантов для их дальнейших модификаций. Более реалистичные модели сетевых систем из взаимосвязанных нейроноподобных элементов потребуют включить в рассмотрение и проводить описание:

а) совместной работы трех, четырех, и может быть даже большего числа переменных (так называемые многокомпонентные системы);

б) одновременной работы нескольких механизмов неравновесности, приводящих к трем и большему числу устойчивых стационарных состояний активных элементов;

в) влияний, связанных с разнообразными неоднородностями, известными для реальных нейроноподобных систем;

г) влияний возможных анизотропных свойств реальных сетевых систем;

д) изменений динамики поведения, вызванных внутренними или внешними шумовыми воздействиями и т. п.

Для каждого вида модифицированных моделей потребуются дальнейшее изучение особенностей автоволновых или более сложных пространственно-временных структур коллективной активности в сети из взаимосвязанных нейроноподобных элементов.

В зависимости от различных сочетаний характерных параметров нейроноподобной среды можно получить разнообразные варианты «базовой» модели (7)–(18).

Из наиболее характерных моделей можно выделить следующие группы.

1. Функции $\Phi_1(\bar{\xi} - \bar{r})$ и $\Phi_2(\bar{\xi} - \bar{r})$ представлены монотонно спадающими зависимостями, что качественно соответствует диффузионным пространственным связям в уравнениях (9) или (10). Тогда при различных видах $F_{B,T}[-t_{0B,T} + \phi_{B,T}]$ реализуются свойства триггерного переключателя, ждущего автогенератора, генератора периодических сигналов или многоуровневого устройства (см. рис. 4). Такого вида «диффузионные» модели используются для описания процессов структурообразования в «плазменных» системах [18–27]). Также они используются при рассмотрении процессов распространения импульсов возбуждения в сердечной, мышечной тканях [18–20, 28–29]), в клеточных популяциях при морфологических трансформациях [30–31]), при описании других клеточных систем [30–32]. Диффузионные модели применяют для описания различных химических реакций [24, 33–35], их используют в алгоритмах обработки информации [36–47].

2. Функции $\Phi_1(\bar{\xi} - \bar{r})$ имеют вид так называемого «латерального» торможения, $\Phi(r) = (1 - b(r)^2)e^{-\alpha(r)^2}$, а $F_m[\]$ имеет ступенчатый вид, то такая система построена в основном для описания процессов в однородных нейронных сетях сетчатки глаза живых объектов, коры некоторых отделов головного мозга и т. п. [2, 4, 8–11, 14, 48–49]), и получила сейчас популяционную «социологическую» интерпретацию. К такому типу могут быть отнесены системы, в которых «нормальная» реакция активных

элементов связана с положительным воздействием на «ближайшее» окружение («родственники»), отрицательным, угнетающим воздействием на ближних соседей, способных конкурировать с «родственниками», и слабым воздействием или нейтральным отношением к дальним элементам популяции. Если вместо одной ступеньки будет $\Pi[Z]$, определенная через две функции Хевисайда: $F[Z - T_1] - F[Z - T_2]$, то приходим к моделям простейших «социально-экономических» процессов [50–52].

3. Функции $\Phi_i(\xi - \bar{r})$ могут иметь несколько переходов от положительного к отрицательному воздействию и наоборот. Либо себя и «родственников» система «угнетает», а стимулирующее воздействие оказывает ближайшим соседям – своеобразная «филантропическая» реакция (см. рис. 5, б) и многие другие варианты.

4. Свою группу составляют также сети из дискретно расположенных в пространстве нейроноподобных активных элементов, известных как *Cellular Neural/Nonlinear Network*, которые рассматривались, например, в работах [53–57].

Нетрудно оценить число разнообразных сочетаний характерных параметров «базовой» модели (17)–(18), которое оказывается порядка нескольких сотен вариантов. Универсальность выбранной модели нейроноподобной среды заключается в возможности широкого использования моделей типа (17)–(18). Они применимы для описания динамики поведения многих природных активных распределенных систем. Их также используют для параллельных режимов преобразования во всех известных сейчас основных операциях предварительной обработки информации.

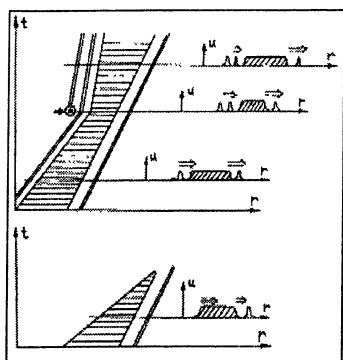
Приведенный набор уравнений (7)–(18) и варианты функций (см. таблицу), представляют собой один из вариантов классификации двухкомпонентных систем нейроноподобного типа.

3. Примеры динамических режимов преобразования сигналов в нейроноподобных системах первой группы

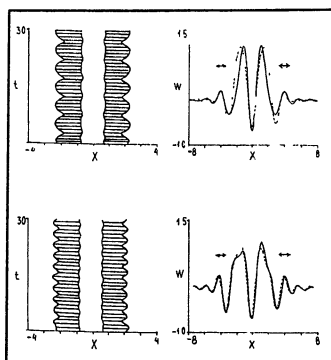
Динамические процессы в моделях однородных нейроноподобных систем с нелокальными пространственными связями (первая группа моделей) были рассмотрены вначале с помощью аналитических методов, а затем изучены в компьютерных экспериментах. Основная проблема в таком исследовании связана с нахождением набора возможных стационарных пространственных структур (аттракторов) в таких неравновесных системах (см. уравнения (7) – (18)), а также с анализом переходных процессов между такими структурами. При рассмотрении процессов в нейроноподобных моделях первой группы оказалось возможным использовать основные элементы качественного анализа, применявшегося ранее при рассмотрении автоволновых процессов в однородных неравновесных

системах с диффузионными пространственными связями [8, 18–20]. Затем, дополнительно, были выведены также новые характеристики для описания особенностей динамики структур коллективной активности в системах с нелокальными пространственными связями [12–14].

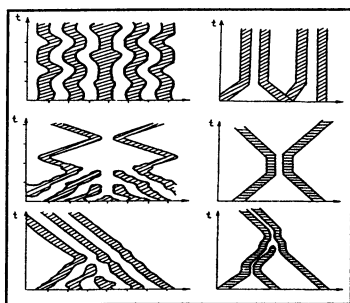
На рис. 6. приведены наиболее характерные примеры динамических структур в однородных двухкомпонентных нейроноподобных системах [12–14, 58–60].



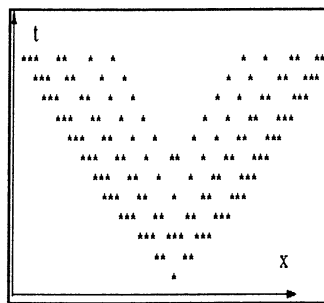
a



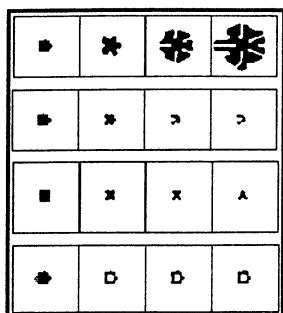
б



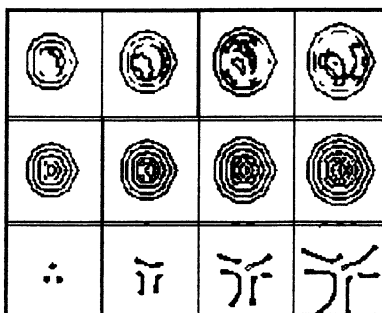
в



г



д



е

Как видно из примеров на рис. 6, нелокальность пространственной связи обеспечивает существование новых структур [12–14, 56–58], связанных с дополнительными устойчивыми импульсами, пульсирующими режимами и режимами отражения импульсов при взаимодействии, «мигающими» режимами распространения активных областей (рис. 6, *a*), неустойчивостями при распространении плоского фронта, а также рядом других режимов, вызванных нелокальностью пространственных связей (рис. 6, *б*).

Для исследования более полного набора автоволновых структур в однородных нейроноподобных системах были разработаны специализированные варианты интерактивных программных продуктов. Анализ показал, что основные особенности характерных решений (в том числе и показанных на рис. 6) управляются тремя типами параметров:

а) параметры, определяющие временную динамику отдельного элемента: τ_1^{-1} , τ_2^{-1} и $\beta_{F1}(g)$, $\beta_{F2}(g)$, t_{0m} , $F_{im}[\]$, непрерывность или дискретность изменения переменных во времени;

б) параметры, определяющие пространственные связи между активными элементами в пространстве: D_i , $\Phi_{mp}(\vec{\xi} - \vec{\eta})$; дискретность расположения активных элементов в пространстве;

в) параметры, связанные с особенностями вида начального условия и формируемой пространственной структуры.

Цель исследований, проводимых с помощью такой интерактивной системы, состоит в поиске для каждого вида моделей (например, полученных из приведенной таблицы) наиболее полного набора возможных решений как в одномерном, так и в двумерном случаях.

Рис. 6. Примеры структур в однородных нейроноподобных системах. *a*–*z* – 8 режимов характерной автоволновой динамики структур в одномерных системах; *x* – пространственная ось, *y* – временная ось; активные состояния заштрихованы; *a* – варианты переходных режимов при формировании областей активности с дополнительными импульсами у фронта перехода за счет нелокальных связей; чем больше дополнительных импульсов, тем меньше скорость распространения такого фронта, снижение скорости заднего фронта (верхний рисунок) вызвано стимуляцией дополнительного импульса внешним воздействием; *б* – два взаимодействующих «пульсирующих» импульса в случае синфазных (верхний рисунок) и противофазных (нижний рисунок) движений их фронтов, *в* – формирование «пачки» остановившихся импульсов в случае «пульсирующего» и обычного режимов их существования (сверху); примеры процессов отражения импульсов в случае «пульсирующего» и обычного режимов их существования (средний рисунок); подавление импульсов и формирование «пачки» распространяющихся импульсов в случае «пульсирующего» и обычного режимов их существования (внизу), *г* – формирование распространяющихся областей «мигающей» активности за счет собственного подавления и латерального возбуждения, *д* – пример формирующейся структуры, вызванной неустойчивостью фронта в расширяющейся области активности (верхний рисунок); два варианта формирования «скелетона» из-за неустойчивости фронта в сжимающейся области активности (средние рисунки); выделение контура (нижний рисунок), *e* – примеры формирования структур в системе с двумя порогоми активации и разным видом нелокальных пространственных связей, *д-е* – в каждом квадрате – двумерная структура активности, вдоль каждой из горизонтальных линий приведены по 4 примера эволюции структур во времени

Понятно, что чем подробнее будет набор таких решений, тем более точным будет язык описания динамических режимов в природных нейроноподобных системах.

Полученные решения для описания автоволновых процессов использовались, например:

а) при интерпретации динамики взаимодействия волн распространяющейся депрессии [59];

б) для описания режимов нормального восприятия и патологических нарушений при прохождении сенсорных сигналов в корковых структурах мозга [60];

в) при оценке параметров автоволновых процессов в сердечной мышце [29, 61] и сравнении модельных расчетов автоволн выброса кальция в мышечных клетках с экспериментальными данными [62–65].

Другая интересная область применения полученных решений связана с определением условий, при которых начальное изображение может трансформироваться в набор упрощенных изображений, которые, в свою очередь, используются для вычисления необходимых кодовых признаков. В частности, было продемонстрировано, что в одних и тех же однородных нейроноподобных системах за счет изменения параметров системы можно осуществить широкий набор преобразований начального изображения. К таким операциям, в частности, относятся: выделение контуров и скелетон для изучаемых объектов на изображении; выделение линий заданного направления и объектов заданного размера; определение областей с заданной текстурой и границ между разными текстурами; локализация точек пересечения линий и многие другие операции (см. рис.7) [66–70].



Рис. 7. Пример схемы, используемой для параллельного преобразования входного изображения в наборы требуемых кодовых описаний отдельных фрагментов

Для выполнения всех этих операций в лаборатории автоволновых процессов ИПФ РАН использовались одно- или двухкомпонентные модели нейроноподобных систем. Такие модели исходно ориентированы на режимы параллельного преобразования входных изображений. До настоящего времени не встречалось ограничений для использования алгоритмов нейроноподобного кодирования при выделении любых фрагментов из анализируемого входного изображения. Аппаратная реализация таких алгоритмов в параллельной архитектуре позволит резко сократить время предобработки сложных изображений большого масштаба.

4. Схемные модели адаптивных распознающих систем

Основной предмет теории распознавания образов составляет изучение различных алгоритмов преобразования потока сигналов какого-либо одного заданного типа, а также рассмотрение операций сравнения соответствующих «кодовых» величин, получаемых при описании видео, акустических, тактильных и других информационных сигналов (см., например, [71–73]). В теории параллельных вычислений на нейроподобных средах модели такого типа относятся к системам с «крупнозернистой» структурой, потому что каждый блок такой системы выполняет свою заданную операцию и расположен в определенной «ветке», пути (последовательности алгоритмов) преобразования сигналов (в отличие от однородного «мелкозернистого» уровня) [74]. Такая «крупнозернистая» структура позволяет использовать рабочий язык огрубленного описания динамических процессов при трансформации потоков информационных сигналов в системах с различными параметрами. Фактически, такие схемы описывают основные наборы последовательных или параллельных вычислений при обработке информационных сигналов.

Модели для элементарных систем с адаптивными свойствами состоят из блоков преобразования информационных сигналов, блоков сохранения кодового описания этих сигналов и используемых для их обработки алгоритмов (моделей), блоков оценки близости сигналов и принятия решений. Если требуется использование режимов параллельного преобразования сигналов, то работа всех этих блоков может осуществляться на основе моделей однородных нейроподобных систем.

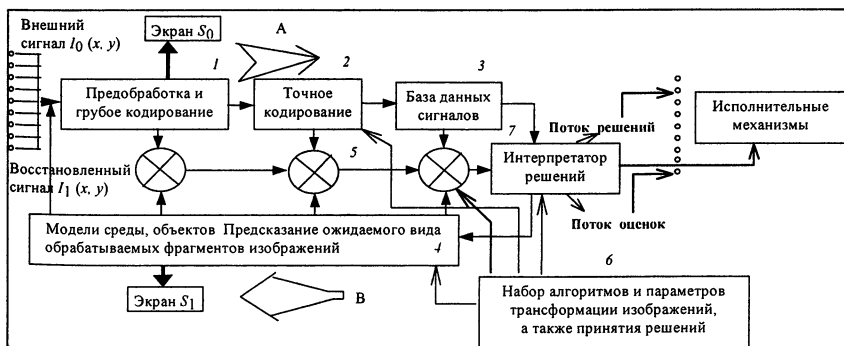


Рис. 8. Этапы трансформации потоков информационных данных и взаимодействие между различными обрабатывающими блоками в адаптивной системе принятия решений с фиксированным набором алгоритмов (нейроподобные системы второй группы)

Архитектура нейроподобных адаптивных систем распознавания ориентирована на выполнение операций оптимальной настройки системы на вид обрабатываемого изображения при фиксированном и заранее заданном наборе используемых алгоритмов. В модельном описании функциональных

взаимосвязей желательно выбрать минимальный набор операций, требуемый для реализации поставленных перед системой целей. На рис. 8 представлен набор последовательностей операций (по мнению автора, необходимый и достаточный) для автоматического управления адаптивно-классификационными процессами и связанными с ними процессами принятия решений [75–80]. Схема построена на взаимодействии следующих основных стадий или последовательностей операций с информационными сигналами:

а) последовательность операций «кодирования» (стрелка А) – исходное изображение на входе (которое может контролироваться на экране S_0) преобразуется в дерево кодовых значений (блоки 1–3), которые представляют собой описание особенностей исходного изображения в терминах эксперта-учителя;

б) последовательность операций обратного восстановления, «генерации» предполагаемого входного изображения по имеющимся в архиве кодовым значениям (стрелка Б, блоки 3–4), т. е. формирование представления данного классификатора о том первоначальном сигнале, по которому было создано такое дерево кодовых значений (это изображение может контролироваться на экране S_1);

в) последовательность операций для формирования оценок точности соответствия или «невязок» (блоки 5) кодовых значений на соответствующих друг другу уровнях в путях «кодирования» и восстановления изображений, смысл этих оценок носит мотивационный характер, потому что по их величине формируется решение в блоке 7;

г) алгоритмы формирования сигнала «принятие решения» (блок 7), на основе полученных оценок сравнения. С помощью эксперта-учителя или алгоритмов, определяющих стратегию поведения, возможно управление вариантами алгоритмов кодирования и восстановления потоков обрабатываемых изображений (блок 6).

Аналогичные архитектуры взаимодействия сигналов при поиске оптимальных алгоритмов для их преобразования рассматривались также, например, в работах [81–83]. Практически все известные системы кодирования сложных сигналов также могут быть интерпретированы на основе такой схемы. Стандартные процедуры при обучении людей также соответствуют представленной на схеме последовательности операций (см. рис. 8). Практически любой процесс обучения представляет собой уточнение и согласование операций кодового описания с операциями восстановления исходного сигнала из кодов и сопоставления полученных результатов.

Основная особенность этой базовой модельной системы заключается в формальном представлении необходимых и достаточных управляющих воздействий между информационными потоками и алгоритмами обработки этих потоков. В частности, вычисленные параметры изображений информационных сигналов (входные и промежуточные изображения, разнообразные кодовые описания, оценки точности соответствия, или

«невязки» – мотивационные оценки) влияют на величины управляющих параметров для каждой из операций (кодирования-декодирования, вычисления мотивационных оценок (невязок), принятия решений), которые в свою очередь изменяют параметры информационного потока. По виду динамики этого процесса вычисляются оценки точности для используемых алгоритмов кодирования-декодирования и формируются мотивационные сигналы для реакций системы. По ним, например, определяются условия, при которых необходимо корректировать параметры действующих алгоритмов или проводить замену «старых» алгоритмов на «новые» алгоритмы кодирования.

5. Примеры динамических процессов при настройке нейроподобных распознающих систем

Рассмотрим некоторые режимы работы такой динамической системы.

Точность работы алгоритмов распознавания, даже при небольшом наборе обучающих примеров, может быть проверена по результатам циклического процесса, когда полученное кодовое описание входного изображения восстанавливается в интерпретацию этого входного изображения, затем восстановленное изображение снова запускается на кодирование и так далее. Скорость сходимости значений в этом цикле «кодирование → восстановление → кодирование → ...» и величины отличий от первоначально полученных значений характеризуют адекватность и точность выбранных алгоритмов. Такой режим, как известно, используется при обучении живых систем. Он позволяет на малом количестве примеров адаптировать алгоритмы на выделение наиболее важных признаков изображения. В этом процессе можно увидеть также аналогию с тем, что называется процессом сознания, т. е., процессом, когда изучаемый образ находится в цикле периодической обработки, анализа и проверки. Если же система не использует процесс восстановления с последующим сравнением (в частности, если предыдущий опыт работы уже убедил ее в корректности выбранных алгоритмов), то решение принимается сразу же по кодовым признакам, и передается на запуск исполнительных механизмов. Этот режим, по-видимому, аналогичен инстинктивным, неосозанным, реакциям в живых системах.

В качестве наглядного подтверждения существования у животных функциональной системы восстановления зрительных образов, можно привести известные всем факты четкого восприятия изображений во сне. В условиях, когда сигнал на внешних сенсорах отсутствует, внутренние процессы «активации кодов в базе данных» запускают функционирование нейронных ансамблей, связанных с восстановлением зрительных образов. В бодрствующем состоянии у животного внешние сигналы от сенсоров, по-видимому, превалируют над сигналами внутренней интерпретации кодов.

Из обыденной жизни и медицинской практики хорошо известно также явление галлюцинаций, при котором некоторые люди могут четко воспринимать (ощущать) отсутствующе на самом деле сенсорные сигналы.

Все живые системы, несомненно, проявляют свойства иерархически организованных адаптивных распознающих систем. Поэтому важно использовать известные для них данные при разработке технических систем распознавания. Архитектура моделей адаптивных распознающих систем позволяет формализовать процессы настройки алгоритмов и использовать некоторые режимы живых систем для использования их в технических системах распознавания с повышенной точностью. В частности, при разработке алгоритмов биометрической идентификации человека была использована архитектура адаптивной настройки.

Обычно эффективные кодирования проверяют в процедуре тестирования большого набора входных сигналов с подсчетом статистических оценок – FAR, FRR (ошибок первого и второго рода). Эта стандартная операция проверки обычно занимает много времени, а также с ее помощью трудно определить, какие операции в алгоритме кодирования приводят к потере точности.

Модель адаптивной распознающей системы позволяет выполнить настройку многих операций даже на одном или малом количестве примеров входных сигналов (см., например, рис. 9). При этом подбираются наиболее оптимальные параметры выбранного алгоритма для выполнения поставленной цели. Проверяется также устойчивость алгоритмов кодирования в процедуре замкнутого, итеративного процесса: кодирование – восстановление, затем повторное кодирование восстановленного изображения. В результате были выбраны наиболее адекватные режимы распознавания человека по изображению его руки, что привело к повышению точности принятия решений в несколько раз.

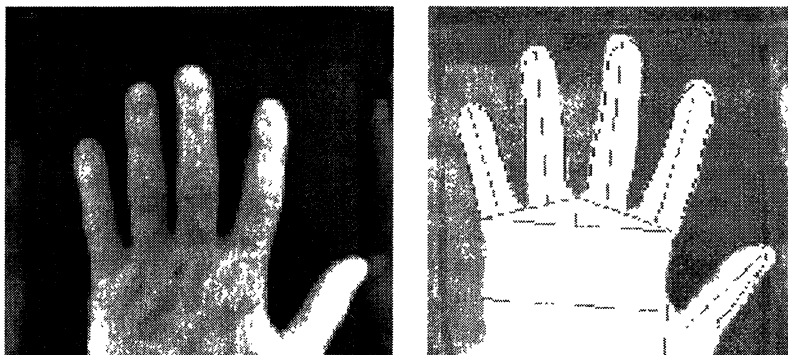


Рис. 9. Пример сопоставления результатов преобразования исходного изображения в контурное и обратное восстановление из кодовой таблицы контура, показанного на исходном полутоновом изображении

Аналогичный подход настройки алгоритмов позволил оптимизировать работу «*Music Recognition System*» – системы распознавания искомых музыкальных фрагментов при анализе непрерывного звукового потока.

Сейчас такую архитектуру автоматической настройки биометрических систем предполагается использовать во всех новых разработках лаборатории автоволновых процессов ИПФ РАН.

Использование формализованных режимов настройки моделей адаптивной распознающей системы представляет интерес также для диагностики и «терапии» различных психологических особенностей взаимодействия живых систем.

Рассмотрим, например, некоторые, весьма понятные с житейской точки зрения, режимы непонимания при общении между адаптивными распознающими системами.

а) Во взаимодействующих системах используются для распознавания разные модели ситуаций и связанные с ними алгоритмы кодирования – декодирования (это может быть связано с разными целями, мотивациями разных систем или просто с отсутствием подходящих алгоритмов), в результате наблюдается ситуация полного непонимания (нарушения в блоках 1, 2). В частном случае, алгоритм одной из систем может быть ориентирован на «образное» кодирование, ассоциативные связи в семантической сети «образов», в то время как другая система строит восприятие на логической модели из цепочки выводов и опирается на оценки обоснованности выводов. Это пример взаимодействия людей с образным и логическим мышлением. Известно, что такие люди очень часто не понимают друг друга [84].

б) Входные сигналы, передаваемые между адаптивными системами, представлены в некотором заранее закодированном «жаргонном» виде (несоответствия алгоритмов в блоках 1, 2, 4). Взаимодействующие распознающие системы либо имеют возможность (мотивации) произвести перекодировку к уже имеющимся у них кодовым представлениям (адаптация в блоке 7), и тогда проявляются элементы понимания, либо не имеют такой возможности (например, потому что перестройка используемых моделей требует дополнительных затрат энергии), и тогда – непонимание (отсутствие настроек в блоке 7).

в) Во входных сигналах распознающая система в первую очередь выделяет такие признаки ситуации, которые препятствуют включению тех адекватных алгоритмов обработки, которые необходимы для распознавания принимаемого потока информации (нарушение работы в блоке 7 ведет к нарушениям в блоках 1, 2, 4). Например, распознающий автомат для понимания потока информации должен затрачивать определенный уровень энергии, в то же самое время он получает сигналы от соседей о малой важности этой информации. В этом случае, если приоритеты сигналов от соседей (или экономии своей энергии) оказываются выше, чем коэффициент

мотивации к восприятию входного информационного потока, адаптации алгоритмов распознавания не происходит, и адаптивная система отключается от процесса анализа входных информационных сигналов и т. д.

Приведенные примеры объединяются одним известным афоризмом: *«Вы не воспринимаете нас не потому, что наши представления неверны или сложны, а потому, что наши представления не входят в ваши понятия»*.

Список возможных жизненных и экспериментальных примеров, подтверждающих эффективность схемы (см. рис. 8) для адаптации используемых алгоритмов, можно увеличивать и дальше. В тоже время, более перспективным для разработки такого вида адаптивных систем было бы обнаружение условий и фактов (например, по предложению от вдумчивых читателей), когда описанная схема динамической системы становится неэффективной.

Заключение

Приведенные в данной работе наборы моделей для однородных нейроноподобных систем и модельных схем для процессов адаптивного распознавания сложных потоков информации ориентированы на описание характерной динамики как биологических систем, так и соответствующих им искусственных аналогов. Из экспериментов известно, что временное изменение пространственных структур активности в живых системах встречается на всех уровнях иерархической организации, включая молекулярный, мембранный, клеточный, популяционный и другие уровни. Почти во всех случаях такие «нелинейные динамические структуры» связаны с различными вариантами функциональных состояний этих объектов.

Собранные к настоящему времени экспериментальные и теоретические данные позволяют утверждать, что различные виды динамических структур активности в нейроноподобных системах, а также особенности преобразования информационных потоков в модельных распознающих системах могут быть основой для описания поведения биологических систем, а также сложных реакций в физических и химических системах, вовлеченных в обработку информации. Следовательно, совместное использование набора базовых моделей и данных о динамике структур самоорганизационной активности (получивших названия автоволновых процессов, динамических структур, и т. п.) формирует язык для описания режимов поведения многих природных нейроноподобных систем.

Приведенные здесь базовые модели и их решения обсуждались с сотрудниками лаборатории автоволновых процессов ИПФ РАН. Всем сотрудникам, решившим заниматься этой тематикой, я искренне благодарен. Особую признательность выражаю И. В. Нуйдель, Н. С. Беллюстину, А. А. Тельных, П. Г. Хурлапову, А. Л. Иванову и А. В. Чайкину за их активное желание реализовать обсуждаемые идеи в конкретных разработках.

Работа выполнялась при частичной поддержке гранта РФФИ 01-01-00388, российско-голландского гранта NWO-RFBR №005-RUS 99/2, а также частичной поддержке гранта АФГИР №РМО-10214-BNL №36943.

Литература

- 1 Позин Н В Моделирование нейронных структур М. Наука, 1970
- 2 Wilson H R, Cowan J D A mathematical theory of the functional dynamics of cortical and thalamic neuron tissue // Kibernetik 1973 V 13 P. 55-80, Biophys J 1972 V 12. P 1-24
- 3 Соколов Е Н, Шмелев Л А Нейробионика М Наука, 1983, Соколов Е Н, Вайт-квичюс Г Г Нейроинтеллект от нейрона к нейрокомпьютеру М Наука, 1989
- 4 Свечников С В, Шквар А М Нейротехнические системы обработки информации Киев Наукова думка, 1983
- 5 Сбитнев В И // Биофизика 1975 Т. 20 С 699-702.
- 6 Сбитнев В И // Биофизика 1976 Т 21. С 1072-1076
- 7 Сбитнев В И // Биофизика 1977 Т 22 С 523-528
- 8 Кудряшов А. В., Яхно В Г Распространение областей повышенной импульсной активности в нейронной сети // Динамика биологических систем 1978 Вып. 2 С. 45-59
- 9 Tuckwell H C, Miura R M A mathematical model for spreading cortical depression // Biophys. J. 1978 23. P 257.
- 10 Голубев В. Н., Рабинович М И, Таланов В. И., Шкловер В., Яхно В. Г. Критические явления в неоднородных возбудимых средах Моделирование на «ТВ-аналоге» // Письма в ЖЭТФ. 1985 Т 42, вып 3. С 84-87
11. Фролов А. А., Муравьев И П. Нейронные модели ассоциативной памяти М : Наука, 1987.
12. Мастеров А. В., Рабинович М. И., Толков В Н., Яхно В. Г Исследование режимов взаимодействия автоволн и автоструктур в нейроподобных средах // Коллективная динамика возбудителей и структурообразование в биологических тканях ИПФ АН СССР. 1988 С 89-104
- 13 Мастеров А. В., Яхно В Г Анализ стационарных волн в однокомпонентной возбудимой среде с нелокальными связями // ИПФ АН СССР. Коллективная динамика возбудителей и структурообразование в биологических тканях. 1988 С 198-208.
14. Yakhno V G. Basic models of hierarchy neuron-like systems and ways to analyse some of their complex reactions // Optical Memory & Neural Network 1995 V. 4, № 2 P. 141-155
- 15 Горбань А. Н. Обучение в нейронных сетях. М · Параграф, 1990
- 16 Экклз Д Тормозные пути в центральной нервной системе М. Мир, 1971.
- 17 Прибрам К. Языки мозга. М. Прогресс, 1975
18. Васильев В. А., Романовский Ю М., Яхно В Г Автоволновые процессы // М : Наука, 1987.
19. Васильев В А, Романовский Ю. М., Яхно В Г Автоволновые процессы в распределенных кинетических системах // УФН. 1979 Т. 128, вып 4 С 625-666.
- 20 Vasiliev V A, Romanovskii Y. M, Chernavskii D C, Yakhno V G. Autowave processes in kinetic systems. Spatial and temporal self-organization in physics, chemistry, biology, and medicine // With 113 figures, D. Reidel Publishing Company, 1987
- 21 Кернер Б. С., Осипов В В Автосолитоны М Наука, 1991.
22. Кернер Б С, Осипов В. В // ЖЭТФ. 1978. Т 74. С 1675-1697; 1980. Т 79, С 2218-2238, 1982. Т 83 С 2201-2214, 1985 Т 89 С 589-607, Микроэлектроника 1981 Т 10 С 407-432; 1983. Т 12 С. 512-529, 1985 Т 14 С 389-407,
- 23 Поляков В П, Поляков С В Автоволновые процессы в нижней ионосфере, модифицированной мощным радиоизлучением // Автоволновые процессы в системах с диффузией / Под ред М Т Греховой Горький ИПФ АН СССР, 1981 С 166-175
- 24 Франк-Каменецкий Д А Диффузия и теплопередача в химической кинетике М Наука, 1967

- 25 Поляков С В., Яхно В Г. О термодиффузионном механизме генерации неоднородностей электронной концентрации в F-слое ионосферы // Физика плазмы 1980 Т 6, № 2 С 383-387
- 26 Семенов В Е Волна пробоя в самосогласованном поле электромагнитного волнового пучка // Физика плазмы 1982 Т 8, № 3 С 613-618
- 27 Райзер Ю П Лазерная искра и распространение разрядов М Наука, 1974
28. Романовский Ю М, Степанова Н В, Чернавский Д С. Математическое моделирование в биофизике М. Наука, 1975, Математическая биофизика. М. Наука, 1984
- 29 Иваницкий Г Р, Кринский В И, Сельков Е Е Математическая биофизика клетки М Наука, 1978.
- 30 Рабинович М. И., Езерский А. Б. Динамическая теория формообразования М. Янус-К, 1998 192 с
- 31 Белинцев Б. Н // УФН. 1983 Т 141 С 55-101
- 32 Романовский Ю М, Теплов В А // УФН 1995 Т 165, № 5. С 555-578
33. Абарбанель Г Д И, Рабинович М И и др Синхронизация в нейронных ансамблях // УФН 1993 Т 166, № 4. С 363-390.
- 34 Колебательные процессы в биологических и химических системах М· Наука, 1967
35. Жаботинский А. М Концентрационные автоколебания М. Наука, 1974
36. Рамбиди Н Г., Чернавский Д. С., Кринский В. И. и др Недискретные биомолекулярные информационно-логические устройства: физические принципы и возможные пути построения // Итоги Науки и Техники ВИНТИ. Сер. «Физические и математические модели нейронных сетей». 1992 Т 5 Ч. 1, С 49-178
- 37 Rambidi N. G, Maksimychev A. V, Usatov A V Implementation of image processing operations using light-sensitive chemical dynamic media // Optical Memory & Neural Networks 1995 V 4, № 2, P 99-109.
- 38 Conrad M Molecular computing: the lock-key paradigm // Computer 1992 V 25, № 11 P. 6-9.
- 39 Conrad M. Emergent computation through self-assembly // Nanobiology. 1993. V 2, P 5-30
- 40 Balkarey Yu. I, Cohen A S., Evtikhov M G, Elinson M I. Autowave collective excitations in neural networks with feedbacks // Optical Memory & Neural Networks. 1995 V 4, № 2 P. 119-126.
- 41 Балкарей Ю. И., Елинсон М И Нейроголография основанная на автоволновых средах // Микроэлектроника 1993. Т 22, № 5 С. 14-19
- 42 Krinsky V I., Viktashev V. N, Efimov I R Autowave principles for parallel image processing // Physica D. 1991. V 49, № 1-2. P 247-253.
- 43 Kuhnert L. A new optical photocemical memory-device in a light-sensitive chemical active medium // Nature 1986. V. 319. P 393.
- 44 Коваленко А С, Яцимирский К. Б., Лучина Л Н, Андреев Е А, Тихонова Л П О влиянии низкоинтенсивного электромагнитного излучения на концентрационные волны Белоусова-Жаботинского // ДАН 1989 Т 304, № 5 P 1143-1147.
- 45 Kuhnert L, Agladze K I, Krinsky V I Image processing using light-sensitive chemical wave // Nature 1989 V 337 P 244
- 46 Королева В И., Кузнецова Г Д Оценка возбуждающих и тормозных факторов в динамике циклического возбуждения коры большого мозга // Электрофизиологические исследования стационарной активности в головном мозге М Наука, 1983 С 173
- 47 Королева В И, Горелова Н А Свойства распространяющейся депрессии во время разных фаз циклического возбуждения коры мозга крысы // Нейрофизиология 1983 Т 15, № 3 С 226
- 48 Bellustin N S, Zemskov S P, Yakhno V G Dynamics of neuron-like media consisting of adaptive elements // Proc of SPIE: Optical Neural Networks / Ed A Mikhaelian 1994 V 2430 P 155-164

49 *Bellustin N S, Zemskov S. P., Yakhno V G* Neuron-like elements for modeling of social and economic problems // Proc of The Second International Symposium on Neuroinformatics and Neurocomputers Rostov-on-Don, Russia, September, 20-23, 1995 P 311-315

50 *Беллустин Н С, Яхно В Г* Пространственно-временная динамика в простейших экономических системах // Нелинейные волны Синхронизация и структуры Ч 2 / Под ред М. И. Рабиновича, М. М. Сушика, В. Д. Шалфеева Нижний Новгород. ННГУ, 1995 С 99-104

51 *Chua L O and Yang L* Cellular neural networks: theory Applications // IEEE Trans On Circuit and Systems 1988 V 35 P 1257-1290

52 *Chua L O, Roska T* The CNN Paradigm // IEEE Trans. on Circuits and Systems I 1993 V 40, № 3 P 147-156.

53 *Rekeczky Cs, Roska T and Ushida A* CNN-based difference-controlled adaptive nonlinear image filters // Int J of Circuit Theory and Applications 1998 V 26 P 375-423

54 *Rekeczky Cs and Chua L. O.* Computing with front propagation active contour and skeleton models in continuous-time CNN // J of VLSI Signal Processing Systems 1999. V 23, № 2/3. P 373-402

55. *Rekeczky Cs* Active contour and skeleton models in continuous-time CNN // Special Session of the 14th IEEE European Conference on Circuit Theory and Design ECCTD'99 in Stresa, August, 1999

56 *Яхно В Г* Процессы самоорганизации в распределенных нейроноподобных системах Примеры возможных применений // Нейроинформатика 2001 Лекции по нейроинформатике М. МИФИ, 2001 С. 103-141

57 *Yakhno V G* Dynamics of autowave processes in neuron-like systems and CNN technology // Proc of the 6th IEEE Int. Workshop on Cellular Neural Networks and their Applications, Catania, Italy, 2000 P 33-38.

58 *Yakhno V G, Bellustin N S, Krasil'nikova I G., Kuznetsov S. O, Nuidel I V, Panfilov A I, Perminov A. O, Shadrin A V, Shevyrev A A* Research decision-making system operating with composite image fragments using neuron-like algorithms // Radiophysics. 1994 V. 37, № 8. P 961-986

59. *Кузнецова Г Д., Пелиновский Д. Е., Яхно В Г* Математические модели динамики волн распространяющейся депрессии в коре головного мозга // Изв вузов Прикладная Нелинейная Динамика 1994 Т 2, № 3-4. С 86-99

60. *Kuznetsova G. D., Nuidel I V, Khurlapov P G, Yakhno V G* Modeling of normal and pathological sensor activity image transformation versions in an animal's cortex // Optical Memory and Neural Networks Optical Information Science and Technology / ed Andrei L Micaelian Proc of SPIE, 1998 V 3402 P. 486-499

61 *Кринский В. И., Яхно В Г* Спиральные волны возбуждения в сердечной мышце // Нелинейные волны / Отв ред. М. И Рабинович Горький. ИПФ АН СССР 1980 С 200-214

62. *Канышева С А, Ключков Б Н, Толков В Н, Яхно В Г* Моделирование динамических структур в мышечных тканях // Медицинская биомеханика Рига, 1986 С 145-150

63 *Coleman A W, Coleman J R, Griffin J D, Weltman J K, Chapman K M* Methylxanthine-induced escalation a propagated wave phenomenon observed in skeletal muscle developing in culture // Proc Nat Acad Sci. USA 1972 V 69, № 3 P. 613-616

64 *Kort A A, Lakatta e g* Propagation velocity and frequency of spontaneous microscopic waves in intact rat papillary muscle are Ca²⁺ - dependent // Biophys J 1984 V 45, № 2, p 2 P 94a

65 *Lea T J* Free calcium measurement in cells // Nature 1977 V 269, № 5624 P 108

66 *Bellustin N S, Kuznetsov S O, Nuidel I V, Yakhno V G* Neural networks with close nonlocal coupling for analysing composite images // Neurocomputing 1991 V 3 P 231-246

67 *Yakhno V G, Nuidel I V* Modeling of sensory information transformations // Neuro-computer as the basis of thinking computers М Наука, 1993 P 207- 223

68 *Yakhno V G, Bellustin N S, Krasilnikova I G, Kuznetsov S O, Nuidel I V, Panfilov A I, Perminov A O, Shadrin A V, Shevyrev A A* research decision-making system operating

with composite image fragments using neuron-like algorithms // Radiophysics 1994 V 37, №8 P 961-986.

69 *Nuidel I V, Kuznetsov S O* Using of Neural Networks for Image Processing // Radiophysics 1994 V 37, № 8 P. 1053-1061

70 *Kuznetsov S O, Nuidel I V, Panfilov A. I, Yakhno V G* Image preprocessing by neuron-like algorithms // Optical Memory and Neural Networks Optical Information Science and Technology / ed Andrei L Micaelian Proc of SPIE 1998 V 3402 P 479-485

71. *Браверман Э М, Мучник И Б* Структурные методы обработки эмпирических данных М. Наука, 1983

72. *Дуда Р, Харп П* Распознавание образов и анализ сцен М. МИР, 1976.

73 *Патрик Э* Основы теории распознавания образов М Советское Радио, 1980.

74 *Arbib M A* Schemes and neural networks for sixth generation computing // J of Parallel and Distributed Computing. 1989. V. 6, № 2 P 185 -216

75 *Telnykh A A., Yakhno V G* Neuron-like models of the second and third levels - adaptive recognition system // Proc. of XII International conference on Neurocybernetics. Neurocybernetics Problems. 1999 P 164-168

76 *Яхно В Г., Нуйдель И В, Тельных А. А., Бондаренко Б Н, Сборщиков И Ф., Хилько А И* Метод адаптивного распознавания информационных образов и система для его осуществления Российский Патент № 2160467, 1999

77 *Бондаренко Б Н., Нуйдель И В, Сборщиков И. Ф, Тельных А. А., Хилько А И, Яхно В. Г* Адаптивное построение изображений нейроноподобной исследовательской системой // Препринт ИПФ РАН Нижний Новгород, 2001 № 559 29 С.

78. *Бондаренко Б Н, Коган А. Н., Нуйдель И В., Сборщиков А. В., Тельных А. А., Хилько А И., Хурлапов П. Г, Яхно В Г.* Разработка адаптивных алгоритмов распознавания сложных изображений в нейроноподобных средах // Труды четвертой конференции по радиофизике 5 мая 2000 Н. Новгород, 2000. С 207-208

79. *Яхно В Г* Процессы самоорганизации в распределенных нейроноподобных системах. Примеры возможных применений // Нейроинформатика 2001 Лекции по нейроинформатике М : МИФИ, 2001 С. 103-141

80. *Yakhno V. G.* Dynamics of autowave processes in Neuron-like systems and CNN technology // Cellular Neural Networks and their Applications Proc. of the 6th IEEE International Workshop, Catania, Italy, 2000 P 33-38

81. *Fukushima K.* Neural network model of selective attention in visual pattern recognition and associative recall // Applied Optics. 1983. V 26, № 23 P 4985-4992; Neural network for visual pattern recognition Computer. 1988 P 65-67

82 *Fukushima K* Cognitron a self-organizing multilayered neural network model // NHK Technical Monograph. 1981 № 30, Neocognitron: a new algorithm for pattern recognition tolerant of deformation and shifts in position NHK Technical Labs Tokyo. Pattern Recognition, 1982. V. 15. P. 455-169,

83. *Зверев В А* Физические основы формирования изображений волновыми полями // ИПФ РАН 1998 252 с

84. *Уолтер Грэй.* Живой мозг. М : Мир, 1966 295 с

Общие вопросы нелинейной динамики

БЕЛЫЙ ШУМ КАК ОСНОВА ОБЪЯСНЕНИЯ МНОГИХ СТАТИСТИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ В ПРИРОДЕ

Г. С. Голицын

Введение

Несмотря на сложность и многообразие природных явлений в их пространственном и временном развитии, носящем часто случайный характер, продолжают обнаруживаться простые закономерности в их распределениях вероятностей. Особенно это относится к моментам этих распределений, прежде всего, ко вторым моментам: спектрам, структурным функциям и др. Эти закономерности, в случае спектров – степенные, обычно наблюдаются на ограниченных интервалах частот (временных и/или пространственных), определяемых внешними по отношению к рассматриваемому явлению причинами и собственными масштабами реакции рассматриваемой системы на те или иные воздействия.

Во многих случаях эти воздействия (силы) имеют случайный характер и их корреляции существенно меньше, чем время реакции системы. Тогда в определенном интервале масштабов корреляционную функцию случайных сил можно аппроксимировать δ -функцией – прием, часто употребляемый в современной физике [1, 2]. Сразу подчеркнем это принципиальное отличие от классической статистической физики, где случайным является распределение скоростей, вводится понятие термодинамического равновесия, а здесь случайным принимается распределение ускорений, т. е. сил на единицу массы изучаемой системы. Это приводит к новой физике, когда в одном интервале масштабов имеется ввод энергии, а ее реализация происходит в других масштабах пространства или времени. Ее начало было положено в 1941 г. классическими работами А. Н. Колмогорова [3, 4] и его аспиранта А. М. Обухова [5, 6] по теории турбулентности.

Здесь, следуя А. М. Яглому и Е. А. Новикову [7–9], на примере локально однородной и локально изотропной турбулентности мы изложим основные моменты описания таких систем. Известные законы турбулентности будут получены путем рассмотрения поведения лагранжевых жидких частиц. Будет показано, что их вторые моменты (временные спектры) тождественны универсальной части спектра морского волнения – [10]. Известная формула для структурной функции лагранжевой частицы (см. ниже (1)), приведенная впервые в 1944 г. в первом издании книги [11] (см. историю в [12]), может быть проинтерпретирована так же, как распределение (кумулятивное) частот событий, вызываемых накоплением энергии в системе под действием случайных сил и имеющих пороговую природу (см. [13, 14]). Таким образом объясняется закон Гутенберга–

Рихтера повторяемости землетрясений в зависимости от их интенсивности. Аналогичные законы известны для частот цунами, оползней и т. п. На этой же основе определена форма энергетического спектра космических лучей [13, 14].

Отметим, что математические основы для рассмотрения реакции системы, интегрирующей воздействия как стохастический стационарный процесс, были развиты А. Н. Колмогоровым в двух заметках в 1940 г. [15, 16] (см. также [17]). Реакция на случайные воздействия также является случайным процессом, но со стационарными приращениями первого порядка. Общую теорию случайных процессов со случайными приращениями n -го порядка развил в 1955 г. А. М. Яглом [18]. Здесь она находит свое применение для описания перемещений лагранжевой частицы в пространстве (процесс со вторыми стационарными приращениями), откуда следует закон Ричардсона–Обухова [19, 6] относительной диффузии в турбулентном потоке.

В заключение будет сформулировано правило (принцип) скорейшей реакции системы на внешние воздействия, во многом эквивалентное подходу, обсуждаемому выше, и дан ряд примеров, его поясняющих.

Турбулентность Колмогорова–Обухова

А. М. Яглом [7] еще в 1949 г. обнаружил, что в инерционном интервале турбулентности, т. е. для масштабов r , меньших внешнего масштаба L_0 и больших внутреннего масштаба $l_v = (v^3/\varepsilon)^{1/4}$, ускорение лагранжевых жидких частиц (ЛЖЧ) является стационарным случайным процессом с постоянным спектром интенсивности ε , т. е. является белым шумом. Здесь ν – кинематическая вязкость среды, ε – скорость диссипации (генерации, в статистически и динамически равновесном случае – *steady-state*) кинетической энергии турбулентности. При этом временная структурная функция скорости ЛЖЧ будет

$$D_v(\tau) = \langle [u(t + \tau) - u(t)]^2 \rangle = \langle [v^2(\tau)] \rangle = \varepsilon \tau. \quad (1)$$

Постоянному спектру ускорений a_ε соответствует корреляционная функция (для временных масштабов $\tau \gg \tau_v = (\nu/\varepsilon)^{1/2}$)

$$B_a(\tau) = \varepsilon \delta(\tau) = \int_0^\infty \varepsilon \cos \omega \tau \, d\omega. \quad (2)$$

Е. А. Новиков [9] рассмотрел эту проблему для уравнения типа Ланжевена, получил экспоненциальную затухающую корреляционную функцию, и все следующие ниже результаты оказываются справедливыми у него для времен $\tau \gg \tau_v$. Временные спектры связаны со структурными функциями соотношением [15–17]

$$D(\tau) = 2 \int_0^\infty (1 - \cos \omega \tau) E(\omega) d\omega. \quad (3)$$

Согласно (3) из соотношения(1) получаем (для инерционного интервала) формулы (П.7) – см. приложение:

$$E_v(\omega)=\varepsilon\omega^{-2}, E_r=\varepsilon\omega^{-4}.$$

Очевидно, что они могут быть сразу получены из соображений размерности только. Из формулы (П.8) $\langle r^2 \rangle = r^2 = \varepsilon\tau^3$ (см. Приложение) выражаем масштаб времени

$$\tau = (r^2/\varepsilon)^{1/3}. \quad (4)$$

Подставляя его в (1) и (3), получаем результаты Колмогорова [3] и Обухова [5, 6]:

$$D_v(r) = (\varepsilon r)^{2/3}, \quad (5)$$

$$E(k) = \varepsilon^{2/3} k^{-5/3}. \quad (6)$$

Коэффициент относительной диффузии получаем, дифференцируя (П.8) по τ и с учетом

$$K = \frac{dr^2}{dt} \sim \varepsilon\tau^2 \sim \varepsilon^{1/3} r^{4/3}. \quad (7)$$

Последнее соотношение и есть знаменитый закон Ричардсона–Обухова [19, 6]. А. М. Обухов получил его из соображений подобия и размерности [5, 6]. Здесь он получен из гипотезы о «белом шуме» воздействий в определенном интервале времен, соответствующем инерционному интервалу. Другими словами, результаты Колмогорова–Обухова являются пространственным отображением «красного шума» поля скоростей лагранжевых жидких частиц, возникающих из «белого шума» их ускорений.

Специально проведенные численные расчеты [20] средних квадратов относительных скоростей и смещений в ансамбле частиц, на каждую из которых действует единичная сила со спектром, имеющим вид белого шума, но с разными распределениями вероятностей, показали, что, во-первых, равенство (1) и (П.8) выполняются с высокой точностью для ансамбля, состоящего из 30 (и более) частиц. Даже для 10 частиц они выполняются вполне удовлетворительно, особенно (П.8). Во-вторых, что можно было предвидеть и заранее, характер распределения вероятностей для сил не влияет на вычисляемые вторые моменты. В частности, гауссово распределение или распределение с разными знаками асимметрии (положительной для двумерной [21] и отрицательной для трехмерной турбулентности [4]) дают искомые зависимости с одними и теми же численными коэффициентами [20].

Морское волнение и закон Ричардсона

О том, как и почему закон Ричардсона (7) выполняется в природе, см. [10, 22]. Особо следует остановиться на относительной диффузии на по-

верхности океана [23–25]. О том, что этот закон выполняется на поверхности моря, Ричардсон и Стоммел заявили в 1948 г. в заметке [24]. Результаты обширных экспериментов многих авторов были обобщены А. Окубо и Р. В. Озмидовым в 1970 г. [24] (см. также книгу [25]). В [24] было показано, что существуют два интервала масштабов – примерно от 10 м до 1 км и от 10 до 3000 км, где в каждом из них действует закон Ричардсона–Обухова, но со значениями скорости диссипации кинетической энергии, различающимися примерно на полтора порядка (если считать численные коэффициенты в (7) одинаковыми в обоих интервалах).

В интервале 1–10 км существует разброс данных, но каких-либо систематических изменений с масштабом r нет. Автор в [22] обратил внимание на тот факт, что в частотном спектре возвышений морского волнения имеется универсальная часть, простирающаяся примерно на декаду почти сразу за частотой максимума спектра, где спектр пропорционален ω^{-4} . Впервые этот результат был получен в 1966 г. В. Е. Захаровым [26] и объяснен из соображений подобия и размерности в 1973 г. Тоба [27] на своих данных измерений как

$$E_h(\omega) = \alpha u_* g \omega^{-4}, \quad (8)$$

где α – численный коэффициент, g – ускорение силы тяжести. Отметим, что u_* – динамическая скорость (скорость трения) в воде. Размерность величины $u_* g$ есть мощность на единицу массы, т. е. совпадает с размерностью скорости диссипации кинетической энергии ϵ в (1). Очевидно тождество формул (8) и (П.8). Из условия несжимаемости воды следует, что зависимость $\propto \omega^{-4}$ должна выполняться и для горизонтальных смещений ЛЖЧ в поле волнения, откуда вытекает закон для среднего квадрата горизонтальных смещений жидких частиц $\propto \tau^3$ и, очевидно, закон Ричардсона. Даже во втором издании 1996 г. книги [12] отмечалась неясность причин выполнения закона Ричардсона в экспериментах по диффузии на поверхности океана. Мы видим, что для интервала 10–10³ м это связано с универсальностью частотного спектра морских волн, тождественного временному спектру ЛЖЧ в турбулентности, и несжимаемостью воды [10, 22]. Накачка энергии осуществляется путем обрушения достаточно длинных волн, в процессе которого лишь очень небольшая часть $\alpha u_* g$ тратится на генерацию горизонтальной турбулентности вблизи водной поверхности (см. [28]).

По поводу выполнения закона Ричардсона в интервале 10–3000 км в [22] высказана гипотеза, что здесь действует обратный каскад энергии двумерной турбулентности. Вертикальный масштаб движения в верхнем квазиоднородном слое океана толщиной порядка 50 м при этом на 2,5–5 порядков меньше горизонтального масштаба движения, что обеспечивает практическую двумерность для масштабов более 10 км. Специально поставленные лабораторные эксперименты [29] подтвердили, что в услови-

ях двумерной турбулентности хорошо выполняется зависимость $\propto \tau^3$ для среднего квадрата относительных смещений жидких частиц, т. е. наблюдаются все условия для выполнения закона Ричардсона (7). При этом обратный каскад может черпать свою энергию из того же процесса обрушения волн, поток которой, как сказано выше, на полтора порядка меньше потока в сторону больших волновых чисел (малых масштабов). Свой вклад в диффузию в больших масштабах могут давать внутренние волны, выходящие на поверхность океана, мезомасштабные неоднородности поля ветра и т. п. Весь вопрос требует, конечно, тщательного рассмотрения.

Сравним закон относительной диффузии Ричардсона–Обухова с законом молекулярной диффузии или диффузии броуновской частицы. В обоих последних случаях средний квадрат смещения частиц равен [30]

$$\langle \Delta x^2(t) \rangle = 2nDt, \quad (9)$$

где n – размерность пространства, D – коэффициент диффузии. В трехмерном случае $2n = 6$. Ричардсон в 1929 г. [30] описал целую серию своих изобретательных опытов по определению коэффициента относительной диффузии ЛЖЧ. Это описание начинается с визуальной оценки коэффициента молекулярной диффузии, очень поучительной для целей образования. Солнечный луч в темной комнате высвечивает пылинки: «Как пляшут пылинки в полдневных лучах...», – писал Ф. И. Тютчев. За секунду пылинка перемещается на расстояние порядка 1 см. Тогда из (9) $D = \langle \Delta x^2(t) \rangle / 6t \approx 0,16 \text{ см}^2/\text{с}$. Это значение практически совпадает со справочным значением $0,15 \text{ см}^2/\text{с}$.

А. М. Обухов [31] обратил внимание на формальное сходство формул (9) и (1): $\langle \Delta v^2(t) \rangle \approx \epsilon t$. Поэтому он назвал величину ϵ коэффициентом диффузии в пространстве скоростей. Еще раз отметим, что в турбулентности случайными являются ускорения, а в молекулярной теории – импульсы молекул. В этом и состоит коренное отличие статистической физики от теории процессов со случайными воздействиями, т. е. силами.

Явления, родственные турбулентности

Из других явлений, близких по статистическим свойствам турбулентности, отметим относительные движения рек на поверхности горных ледников при их движениях со средней скоростью нескольких метров в сутки. А. Б. Казанский [32] провел такие измерения на Памире на двух ледниках. Относительные перемещения рек, расставленных в разных частях поверхности ледника, измерялись в течение 8 часов, вычислялись скорости, которые затем осреднялись. С не очень большим разбросом найденные структурные функции скоростей для рек оказались подчиняющимися «закону 2/3». Казанский объясняет наблюдаемый закон тем, что движущийся ледник состоит из кусков различных размеров, взаимодействующих друг с другом. Можно, очевидно, предположить, что

время корреляции ускорений (сил), действующих на отдельные куски, было много меньше 8 часов, времени осреднения данных об изменении положения реек. При этом наша схема случайных воздействий должна срабатывать, что и подтверждается наблюдениями.

Из других явлений, родственных турбулентности с точки зрения развитаго здесь подхода, упомянем термическую и плотностную конвекцию, турбулентные флуктуации пассивной примеси, о которых пойдет речь в предпоследнем разделе, также низкочастотные флуктуации ряда климатических процессов. Подход к описанию последних был предложен К. Хасельманом [33] на основе «белого шума» воздействий. Например, для температуры ВКС (верхнего квазиоднородного слоя) океана это флуктуации радиационного баланса на морской поверхности, вызванные, в частности, переменной облачностью (подробнее об этом см. [34] и литературу там).

Статистика событий

Рассмотрим статистику событий различной интенсивности (энергии), регистрируемых с определенной частотой во времени и пространстве. События – такие, как землетрясения, цунами, оползни – являются пороговыми явлениями, когда система набирает в течение определенного времени некоторый запас энергии (упругой в случае землетрясений) и затем сбрасывает ее тем или иным способом. В случае землетрясений – путем образования разрывов в земной коре, в случае оползня потенциальная энергия породы (разжиженной) сбрасывается в процессе быстрого схода оползня и т. п.

Более детально, помимо землетрясений, рассмотрим природу космических лучей. Частицы их образуются в процессах взрыва сверхновых звезд и последующего ускорения в ударных волнах этих взрывов [35]. Это так называемый механизм ускорения Ферми, качественно им предложенный в 1950 г. (см. [35]). В предположении случайности процесса ускорения можно объяснить наблюдаемую форму энергетического спектра космических лучей, регистрируемого на земной поверхности. Впервые это было сделано на основе соображений подобия и размерности в [36] и более полно в [14].

Начнем рассмотрение с очевидной формулы

$$A \approx \frac{dA}{dt} \tau = \dot{A} \tau \quad (10)$$

для приращения величины A за время τ . Очевидна формальная аналогия с формулами (1) для структурной функции скорости или (9) для среднего квадрата смещения броуновской частицы. Если A – случайная (квадратичная) величина, то при $\varepsilon_A = dA/dt = \text{const}$ аналогия ясна. Отличия при рассмотрении статистики событий, вроде землетрясений, возникают из-за того, что при этом время ожидания τ должно быть кумулятивным време-

нем $\tau(\geq A)$ наблюдения события с интенсивностью равной или большей A . С этим временем связана частота

$$N(\geq A) = \frac{1}{\tau(\geq A)} = \int_A^{\infty} \frac{dA}{\tau(A)}, \quad (11)$$

лишь эта частота имеет размерность обратного времени, а дифференциальная частота $\tau(A)$ имеет размерность $[A \cdot T]$. Предполагается, что интеграл (11) сходится благодаря тому, что существует, по тем или иным причинам, верхний предел роста величины A , т. е. для этих событий время ожидания бесконечно.

С помощью (11) из (10) получаем

$$N(\geq A) = A^{-1} \dot{A}, \quad (12)$$

что и является основой рассмотрения статистики событий, зависимостей типа интенсивность (энергия) – частота.

Статистика землетрясений – закон Гутенберга–Рихтера (ГР)

Этот эмпирический закон обычно приводится в виде [37]

$$\lg N(\geq m) = a - bm, \quad (13)$$

где m – магнитуда землетрясений, a, b – численные коэффициенты. Величина $b \approx 1$ для $0 < m < 7$. Для больших землетрясений и землетрясений вблизи срединно-океанических хребтов $b \approx 15$. В [14] показано, что магнитуда связана с площадью разрыва S в земной коре соотношением

$$m = \lg(S/S_0), \quad S \approx 10^2 \text{ м}^2 = 1 \text{ ар (сотка)}. \quad (14)$$

Конечно, эта связь статистическая, но коэффициент корреляции величины m из (14) с магнитудой сейсмических каталогов $r = 0,91$. Потенцируя (13) и учтя (14), мы сразу видим, что $b = 1$ в полном соответствии с (12). Связь природы изменений величины b от 1 до 1,5 выяснили в 1988 г. С. С. Григорян [38] и в 1989 г. Рандл [39], связав эти изменения с толщиной коры: тонкая кора рвется вся и процесс распространения трещины в ней одномерный, а в толстой коре при не слишком сильных землетрясениях разрыв распространяется двумерным образом.

В последние десятилетия в теоретической сейсмологии используется величина сейсмического момента M , связанная с магнитудой соотношением

$$m = \frac{2}{3} \lg M - 6, \quad (15)$$

где M выражено в системе СИ – (ньютон на метр) $\text{Н} \cdot \text{м}$; μ – модуль сдвига, величина, как правило, близкая к $3 \cdot 10^{10}$ Па; S – площадь разрыва, u – среднее смещение блоков коры вдоль разрыва. Приведем формулу Рихтера для регистрируемой энергии сейсмических волн, излучаемых в процессе землетрясений:

$$\lg E = 1,5 m + 4,8,$$

откуда и из (15) следует

$$E = 6,3 \cdot 10^5 M. \quad (16)$$

Автор в работе [40] развил теорию подобия для землетрясений. Известно [37, 14], что величина сброса напряжений $\Delta\sigma$ не имеет прямой зависимости от m и разброс ее для разных землетрясений невелик (см. рис. 1 в [14]), составляя в среднем примерно $4 \text{ Мпа} = 4 \cdot 10^6 \text{ бар}$. С ее помощью можно составить масштаб длины

$$L_0 = (M/\Delta\sigma)^{1/3}. \quad (17)$$

Этот масштаб статистически хорошо описывает [14, 40] длину и площадь разрыва, смещение блоков $u = (\Delta\sigma/\mu)L_0$. Его отношение к толщине коры h определяет естественный параметр подобия для процесса землетрясений:

$$\Pi = L_0/h. \quad (18)$$

Для вывода закона ГР в терминах сейсмического момента воспользуемся общей формулой (12), которую запишем с учетом появившегося параметра подобия (18) в виде

$$N(\geq M) = M^{-1} \dot{M} f(\Pi), \quad (19)$$

что, вследствие (15), соответствует $b = 1,5$ в законе ГР, если $f(\Pi) = \text{const}$.

Безразмерная функция $f(\Pi)$ должна определяться из наблюдений. Сравнение с данными глобального каталога землетрясений показывает [40], что величина b возрастает от 1 до 1,5 при $\Pi \approx 1$. Сравнивая (15), (13) и (19), отмечаем, что при $\Pi \geq 1$ функция $f(\Pi) = \text{const}$. Величину \dot{M} Я. Я. Каган [42] определил как $36 \cdot 10^{20} \text{ Н} \cdot \text{м/год} = 1,2 \cdot 10^{14} \text{ Вт}$, что формально примерно втрое больше мощности глобального потока геотермического тепла F – основы всей геодинамики. С учетом этого, сравнивая (19) с глобальным каталогом землетрясений, получаем, что $\text{const} = 0,13$. При $\Pi \ll 1$ можно разложить функцию $f(\Pi)$ в ряд Тэйлора, ограничившись лишь линейным по аргументу членом. Тогда с учетом (17) и (18) получим

$$N(\geq M) = aM^{-2/3}, \quad a = 0,13 \dot{M} (\Delta\sigma^{1/3} h)^{-1} = 0,4 F (\Delta\sigma^{1/3} h)^{-1}. \quad (20)$$

Таким образом, в [40] была раскрыта природа эмпирической величины a в законе ГР. Она оказывается связанной с потоком тепла $F = 4 \cdot 10^{13} \text{ Вт}$ (или региональным значением \dot{M} (см. [42]), толщиной коры, величиной сброса напряжения).

Приведенные здесь цифры дают возможность оценить характеристики цикла энергии в процессах геодинамики. Зная величину \dot{M} , можно оценить с помощью (16) мощность сейсмических волн. Это будет величина порядка $2,3 \cdot 10^{11} \text{ Вт}$. Согласно [37] в эти волны идет менее 0,1 доли полной энергии, высвобождаемой в процессе землетрясения. Тогда доля геотермического потока, идущая на землетрясение, будет порядка $(2 \cdot 10^{10}) / (0,1 \cdot 4 \cdot 10^{13}) = 5 \cdot 10^{-3}$, т. е. порядка половины процента.

Для сравнения отметим, что мощность развития тропического урагана тоже порядка 10^{13} Вт [42]. Но их бывает около 80 в год, и они приносят населению Земли существенно больше вреда, чем землетрясения.

Тот факт, что соотношение (12), составляющее основу закона ГР, можно рассматривать как процесс стохастической накачки упругой энергии в коре, подтверждают спектры деформаций, полученные по данным измерений лазерным интерферометром на глубине 2 км на Баксанской нейтринной станции [43]. Спектры деформаций в интервале частот 10^5 –1 Гц близки к f^{-2} , где f – частота. Это значит, что спектр производных деформаций по времени близок к спектру белого шума. Упругая (свободная) энергия квадратично связана с тензором деформаций. Поэтому она, в соответствии с (1), будет расти пропорционально времени.

Эта энергия, накапливаясь годами и десятилетиями, сбрасывается, когда достигает некоторого порогового значения. Такого рода процессы называются самоорганизующимися критическими (self-organized criticality [2, 14]).

Зависимости типа закона ГР обнаружены для распределений цунами по высотам (см. [14]), для индуцируемых частот (например откачкой газа) землетрясений [44], всплесков гамма-излучения на некоторых нейтронных звездах [45], оползней [14] и т. д.

Спектр энергии космических лучей

Рассмотрим здесь лишь частицы КЛ с энергией $E \geq 10^{10}$ эВ = 10 ГэВ. Известно (все данные из [35]), что скорость их генерации $G \approx 3 \cdot 10^{33}$ Вт (примерно 3 сверхновых в нашей галактике за 100 лет), плотность энергии $\omega = 0,5$ эВ/см³ для частиц в интервале 1 – 10^3 ГэВ. Кумулятивный спектр космических лучей

$$I(\geq E) = \begin{cases} 1E^{-1/7}, & 10 \leq E \leq 3 \cdot 10^6 \text{ ГэВ}, \\ 10^{-7} E^{-2.1}, & 3 \cdot 10^6 \leq E \leq 10^9 \text{ ГэВ}, \end{cases} \quad (21)$$

Стохастический механизм Ферми предполагает, что набор энергии частиц пропорционален времени

$$E \sim G\tau(\geq E) \quad (22)$$

в полной аналогии с законом ГР. Отличие спектра космических лучей в том, что измерение идет на единицу площади (а по закону ГР – на территорию региона или всей поверхности Земли). Если бы мы еще знали пространственную плотность КЛ величину $n(\geq E)$, то смогли бы оценить искомый спектр как

$$I(\geq E) = \frac{v}{4\pi} n(\geq E), \quad n(\geq E) = \int_E^{\infty} n(E) dE, \quad (23)$$

где v – скорость движения облаков космических лучей в нашей Галактике. К счастью, последнюю величину в (23) мы можем оценить сверху как

$$n(\geq E) = \int_E^{\infty} \frac{E}{E} n(E) dE \leq \frac{1}{E} \int_E^{\infty} E n(E) dE = \frac{\omega(\geq E)}{E} \leq \frac{\omega}{E} \quad (24)$$

через известную, по крайней мере для указанного выше интервала от 1 до 10^3 ГэВ, величину ω объемной плотности энергии космических лучей. Величина $n(\geq E)$ есть объемная концентрация числа частиц, и тогда можно оценить расстояние между частицами:

$$l = [n(\geq E)]^{-1/3} \leq (E/\omega)^{1/3}. \quad (25)$$

Отсюда следует, что частицы движутся в пространственных трубках площадью сечения

$$S(\geq E) \leq [n(\geq E)]^{-2/3} \geq (E/\omega)^{2/3}. \quad (26)$$

Таким образом, имеется естественный масштаб частоты из (22) как $N(\geq E) = [\tau(\geq E)]^{-1} \sim G/E$ и оценка (снизу) величины $S^{-1} \geq E$. Тогда для кумулятивного спектра можно записать:

$$I(\geq E) \leq \frac{c}{\tau(\geq E)} \frac{1}{S(\geq E)} = c \frac{G}{E} \left(\frac{\omega}{E} \right)^{2/3} = c G \omega^{2/3} E^{-5/3}. \quad (27)$$

Показатель степени в нашей оценке сверху оказывается практически совпадающим с показателем 1,7 для измеренных спектров. Численный коэффициент $c = 3 \cdot 10^{-38}$, что получается из сравнения формулы (27) с первой строкой системы (21). Такая малая величина получается потому, что мы измеряем площадь, например в квадратных метрах, а площадь сечения диска Галактики порядка 10^{39} м^2 .

Так называемое «колени», с которого начинается укручение в спектре при $E \geq 3 \cdot 10^6$ ГэВ, объясняется тем, что ларморовский радиус частицы с этой энергией в магнитном поле с плотностью энергии порядка ω начинает уже чувствовать конечную толщину радиуса диска. Постоянство плотности энергии для частиц с $E \geq 3 \cdot 10^6$ ГэВ становится слишком грубым приближением, так как эта плотность должна падать с ростом E . Однако можно оценить (снова сверху) поведение величины $\omega(\geq E)$ с помощью (27), (23)–(26), откуда

$$\frac{\omega(\geq E)}{E} > n(\geq E) = \frac{4\pi}{v} I(\geq E) \sim \frac{4\pi}{v} G \omega^{2/3} E^{-5/3}.$$

Тогда единичная площадь

$$S^{-1}(\geq E) \geq \left(\frac{4\pi}{v} G \omega^{2/3} E^{-5/3} \right)^{2/3} = \left(\frac{4\pi G}{v} \right)^{2/3} \omega^{4/9} E^{-10/9}.$$

Время, т. е. частоту, возьмем опять из (22) и получим

$$I(\geq E) \sim \left(\frac{4\pi}{v}\right)^{2/3} G^{5/3} \omega^{4/9} E^{-n}, \quad n = \frac{19}{9} = 2 + \frac{1}{9}, \quad (28)$$

что удивительным образом согласуется с эмпирическим показателем 2.1 в (21).

Во всех случаях, рассмотренных в этом разделе, решающим является стохастическая природа набора энергии системой, ведущая к ее линейной зависимости от времени. Это роднит статистику событий с турбулентностью и родственными ей явлениями.

Обратная времени величина есть частота, т. е. число событий за единицу времени. Отсюда и возникает закон ГР, осложненный в случае космических лучей необходимостью учета и пространственной частоты частиц космических лучей.

Правило скорейшей реакции

Это правило, сформулированное впервые автором в 1980 г. [46], гласит, что если мы знаем воздействие (мощности) на систему, то ее результирующая энергия оценивается как произведение мощности на наименьшее время реакции, присущее этой системе. В [46] в качестве примеров приводились кинетическая энергия планетных атмосфер (не слишком быстро вращающихся), колмогоровская турбулентность, термическая конвекция. В 1997 г. была подготовлена обширная работа [13] и опубликована заметка [47], где количество примеров было существенно расширено. Например, в гидродинамике таким образом были получены (с точностью до численных коэффициентов) формулы для скоростей в трубах под влиянием градиента давления в вязком и турбулентном режимах, колмогоровские законы турбулентности, в том числе в вязком интервале масштабов, скорости конвекции для турбулентных и вязких режимов, в том числе с учетом вращения, законы ГР и для космических лучей (верхняя строка (21)). Наименьшее время в случаях гидродинамики находилось из параметров подобия, которые представлялись в виде отношений двух характерных времен. Например, число Рейнольдса представлялось как

$$Re = \frac{ur}{v} = \frac{u}{r} \frac{r^2}{v} = \frac{\tau_v}{\tau_a}. \quad (29)$$

Если почти все описанные выше результаты могут быть получены и из соображений размерности, то это не так с последним результатом, например, если порознь рассматривать величины ε , v , r . Наше правило объединяет две первые величины в одну с размерностью времени, что и обеспечивает успех. При $Re \leq 1$ действует $\tau_v = r^2/v$, и мы получаем структурную функцию в интервале диссипации, при $Re \gg 1$ $\tau_d = r/v$, и мы получаем закономерности инерционного интервала турбулентности. В случае землетрясений и космических лучей, где нет параметров подобия,

представимых в виде отношений двух времен, существует лишь одно время $\tau(\geq A) = A / \dot{A} = A(dA/dt)^{-1}$ и именно оно использовалось для вывода распределений частота – интенсивность, т. е. энергетических спектров.

На основе этого правила можно вывести формулу для кинетической энергии общей циркуляции атмосферы для планеты типа Земли в виде произведения мощности солнечной энергии и времени установления локального статистического равновесия в глобальном масштабе $\tau = r/c$, где r – радиус планеты, c – скорость звука, вывести формулы типа Стокса для падения частички в поле силы тяжести, определить масштаб скорости при конвекции вращающейся жидкости [48] и т. д. и т. п. [14]. Вкратце упомянем поведение скалярной примеси в турбулентном потоке с интенсивностью генерации турбулентности ε и скоростью генерации неоднородности в поле примеси N . Средний квадрат изменений примеси на расстоянии r (структурная функция) в инерционном интервале будет $D_0(r) \approx N\tau_\alpha \approx Nr/u$. Так как $u \approx (\varepsilon r)^{1/3}$, то $D_0(r) = N\varepsilon^{-1/3}r^{2/3}$, закон, полученный в 1949 г. А. М. Обузовым (см. [12]).

Заключение

Все процессы, описанные здесь, есть реакция некоторой системы на очень мелкомасштабные по времени случайные воздействия, их в некотором интервале масштабов можно описать белым шумом. При этом, например, результаты Колмогорова – Обухова понимаются как пространственное (эйлерово) отображение красного шума поля скоростей, вызванного белым шумом ускорений, действующих на лагранжеву жидкую частицу. Во всех разобранных случаях приращение энергии системы пропорционально времени воздействия на нее. В статистике событий задача перевертывается: мы измеряем интенсивность или энергию события, знаем время ожидания события (т. е. частоту), хотим оценить мощность воздействия и понять, почему распределения энергия – частота (землетрясений и родственные явления), как правило, обратно пропорциональны энергии. Ответ прост и заключается в виде формулы типа (10), т. е. что произведение мощности на время реакции есть энергия, набираемая системой в процессе стохастического (или регулярного) воздействия на нее. Отсюда белый шум воздействия и правило скоростей реакции. В случае космических лучей необходимо оценивать их объемную концентрацию в зависимости от энергии E , что удастся, по крайней мере, сверху. В случае землетрясений и подобных им процессов работает принцип самоорганизованной критичности [42]: реакция системы начинается, когда накапливаемая энергия превосходит некоторый порог. Если времена корреляции воздействий становятся как-то сравнимыми с временами реакции системы, то спектр воздействий будет уже подниматься в сторону низких частот и система будет набирать энер-

гию как τ^n , где уже $n > 1$. Тогда временной спектр реакции системы будет падать как ω^{-n-1} , т. е. быстрее, чем ω^{-2} .

В заключение дадим упорядоченный список рассмотренных здесь процессов и событий (вопросы, рассматривавшиеся автором, помечены звездочкой).

1. Непрерывные процессы.
 - 1.1 Локально-однородная и изотропная турбулентность.
 - 1.2 Спектр морского волнения (и диффузия Ричардсона*).
 - 1.3 Пассивный скаляр в турбулентном потоке.
 - 1.4 Правило скорейшей реакции* (взаимозаменяемое, с некоторыми оговорками, с δ -коррелированными воздействиями).
 - 1.4.1 Термическая и плотностная конвекция в вязком* и турбулентном режимах (мантия Земли*, термики и т. д.).
 - 1.4.2 Конвекция во вращающихся жидкостях* (жидкое ядро Земли, ураганы, глубокая конвекция в океане и т. п.).
 - 1.4.3 Циркуляция планетных атмосфер* (Земля, Марс, Венера, Титан).
 - 1.4.4 Течение в трубах при вязком и турбулентном режимах.
 - 1.4.5 Течение под влиянием силы тяжести (реки, наносы и т. п.)
2. Статистика событий.
 - 2.1 Самоорганизованная критичность (землетрясения, оползни, гамма-вспышки нейтронных звезд и т. п.*).
 - 2.2 Цунами.
 - 2.3 Космические лучи*.

Приложение

Здесь мы привели необходимые математические сведения, строгие выводы которых можно найти в [12] ((24.2) и др.). Для простоты пренебрежем векторной природой сил, ускорений, скорости, смещений, поскольку все равно численные коэффициенты надо определять из экспериментов, реальных или точных численных. Уравнение движения для ЛЖЧ есть

$$x = a(t). \quad (\text{П.1})$$

Для приращения скорости за время τ имеем

$$\dot{x}(\tau) - \dot{x}(0) = \int_0^\tau a(t) dt = V(\tau) - V(0), \quad (\text{П.2})$$

а для структурной функции, т. е. среднего квадрата приращения скорости,

$$\langle [V(\tau) - V(0)]^2 \rangle = D_v^4(\tau) = \int_0^\tau \int_0^\tau B_v(t' - t'') dt' dt'' = 2 \int_0^\tau (\tau - s) B_v(s) ds. \quad (\text{П.3})$$

Рассмотрим интервал времен, где корреляционная функция ускорений может быть представлена как

$$B_a(\tau) = a^2 \tau_0 \delta(\tau) = \varepsilon \delta(\tau), \quad (\text{П.4})$$

где $a^2 = \varepsilon^{3/2} v^{-1/2}$, $\tau_0 = (v/\varepsilon)^{1/2}$ (см. [12, 7] и др.). Тогда спектр ускорений ЛЖЧ будет белым шумом, а временная структурная функция скорости частицы из (П.3) с учётом (П.4) запишется как

$$D_v^L(\tau) = 2\varepsilon\tau \approx \varepsilon\tau. \quad (\text{П.5})$$

Для случайных процессов со стационарными приращениями первого порядка предложена в [15–17] (см. также [12] и др.) связь между спектральной плотностью и структурной функцией:

$$D(\tau) = 2 \int_0^{\infty} (1 - \cos \omega\tau) E(\omega) d\omega, \quad (\text{П.6})$$

которую с помощью известных тригонометрических соотношений можно представить так же, как

$$D(\tau) = 4 \int_0^{\infty} \sin^2\left(\frac{\omega\tau}{2}\right) E(\omega) d\omega. \quad (\text{П.7})$$

Средний квадрат смещений ЛЖЧ получается интегрированием (П.2). С учетом (П.5), опять с точностью до численного коэффициента,

$$\langle x^2(\tau) \rangle = r^2(\tau) \approx \varepsilon\tau^3. \quad (\text{П.8})$$

Поле смещений будет уже случайным процессом со вторыми стационарными приращениями [18]. В инерционном интервале, если спектр ускорений $E_a(\omega)$ – белый шум, то спектр скоростей будет $E_v(\omega) = \varepsilon\omega^{-2}$, а спектр смещений

$$E_r(\omega) = \omega^{-2} E_v(\omega) = \omega^{-4} E_a(\omega) = \varepsilon\omega^{-4}. \quad (\text{П.9})$$

Структурная функция смещений будет связана со спектральной плотностью (П.9) соотношением (П.7). Для существования интеграла в (П.7) функция $E(\omega)$ должна в нуле расти медленнее, чем ω^{-3} , либо надо переопределить структурную функцию как вторую разность: $D(\tau) = D(0) - 2D(\tau) + D(2\tau)$, которая уже будет стационарным процессом.

Литература

1. *Кляцкин В И* Стохастические уравнения глазами физика М. Физматлит, 2001. С 528.
2. *Sornette D.* Critical phenomena in natural sciences. Springer, 2000. P. 434.
3. *Колмогоров А Н* Локальная структура турбулентности в несжимаемой жидкости при очень больших числах Рейнольдса // ДАН СССР. 1941 Т. 30. № 2 С 299–303.

4. Колмогоров А. Н. Рассеяние энергии при локально изотропной турбулентности // ДАН СССР. 1941. Т. 32, № 1. С. 19–21.
5. Обухов А. М. О распределении энергии в спектре турбулентного потока // ДАН СССР. 1941. Т. 32, № 1. С. 22–24.
6. Обухов А. М. О распределении энергии в спектре турбулентного потока // Изв. АН СССР. Сер. геогр. геофиз. 1941. Т. 5, № 4–5. С. 453–466.
7. Яглом А. М. О поле ускорений в турбулентном потоке // ДАН СССР. 1949. Т. 61, № 5. С. 195–198.
8. Обухов А. М., Яглом А. М. Микроструктура турбулентного потока // ПММ. 1951. Т. 15, № 1. С. 3–26.
9. Новиков Е. А. Метод случайных сил в теории турбулентности // ЖЭТФ. 1963. Т. 44, № 6. С. 2159–2168.
10. Голицын Г. С. Методические основы теории турбулентности и спектра морского волнения // Изв. АН СССР. ФАО. 2001. Т. 37, № 4. С. 438–445.
11. Ландау Л. Д., Лифшиц Е. М. Механика сплошных сред. М.: ГИТТЛ, 1953. С. 734.
12. Монин А. С., Яглом А. М. Статистическая гидромеханика Т. 2. 1-е изд. М.: Физматгиз, 1967. 720 с. 2-е изд. СПб.: Гидрометеоздат, 1996. 742 с.
13. Golitsyn G. S. Convection in viscous and rotating fluids from the viewpoint of the forced flow theory // Buoyant convection in geophysical flows / Eds. E. Plate et. al. Kluwer Ac. Publ., 1998, P. 129–155.
14. Голицын Г. С. Место закона Гутенберга–Рихтера среди других статистических законов природы // Вычисл. сейсмология 2001, № 32. С. 138–161.
15. Колмогоров А. Н. Кривые в гильбертовском пространстве, инвариантные по отношению к однопараметрической группе движений // ДАН СССР. 1940. Т. 26, № 1. С. 6–9.
16. Колмогоров А. Н. Спираль Винера и некоторые другие интересные кривые в гильбертовском пространстве // ДАН СССР. 1940. Т. 26, № 1. С. 115–118.
17. Neumann, J., von, Schoenberg I. J. Fourier integrals and metric geometry // Trans. Amer. Math. Soc. 1941. HV. 50, № 2. P. 226–251.
18. Яглом А. М. Корреляционная теория процессов со случайными стационарными n -ми приращениями // Матем. сб. 1955. Т. 37, № 1. С. 141–196.
19. Richardson L. F. Atmospheric diffusion shown on a distance-neighbour graph // Proc. Roy. Soc. London. 1926. V. A 110, № 156. P. 709–737.
20. Голицын Г. С., Пономарев В. М. // Средние квадраты относительных скоростей и смещений в ансамбле частиц при белом шуме действующих на них сил // Письма в ЖЭТФ. 2002 (в печати).
21. Lindborg E. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? // J. Fluid Mech. 1999. V. 388. P. 259–288.
22. Голицын Г. С. Объяснение закона относительной турбулентной диффузии в атмосфере и на поверхности океана // ДАН. 2001. Т. 380, № 5. С. 676–679.
23. Richardson L. F., Stommel H. Note on eddy-diffusion in the sea // J. Meteorol. 1948. V. 5, № 5. P. 238–240.
24. Окубо А., Озмидов Р. В. Эмпирическое соотношение между коэффициентом горизонтальной диффузии и масштабом явления // Изв. РАН. ФАО. 1970. Т. 6, № 5. С. 534–536.
25. Озмидов Р. В. Турбулентная диффузия в океане Л. Гидрометеоздат, 1985. 284 с.
26. Захаров В. Е., Филоненко Н. Н. Спектр энергии для стохастических колебаний жидкой поверхности // ДАН СССР. 1966. Т. 170, № 6. С. 1292–1295.
27. Toba Y. Local balance in the air-sea boundary processes. III. On the spectrum of wind waves // J. Oceanogr. Soc. Japan. 1973. V. 29. P. 209–220.
28. Phillips O. M. Spectral and statistical properties of the equilibrium range in wind generated gravity waves // J. Fluid Mech. 1985. V. 156. P. 505–531.
29. Jullien M.-C. Paret J., Tabeling P. Richardson pair dispersion in two-dimensional turbulence // Phys. Rev. Lett. 1999. V. 82 (14). P. 2872–2875.
30. Richardson L.F. A search for the law of atmospheric diffusion // Beitr. Phys. Freien Atmos. 1929. B. 1S. Nr. 1.5.24–29.

31. *Обухов А. М.* Description of turbulence in terms of lagrangian variables // *Adv. Geophys.* 1959. V. 6. P. 113–115.
32. *Kazansky A. B.* Studing the motion energy dissipation process in highly fractured glaciers using remote sensing technique // *Ann. Glaciologu* 1987. V. 9. P. 239–248.
33. *Hasselmann K.* Stochastic climate models. Part I. Theory // *Tellus*. 1976. V. 28, № 6. P. 473–485.
34. *Голицын Г. С., Демченко П. Ф.* Статистические свойства простой энергобалансовой модели климата // *Изв. АН СССР. ФАО.* 1980 Т 16, № 12 С 1235–1242.
35. *Гинзбург В. Л.* (ред) *Астрофизика космических лучей* М · Физматгиз 1990 534 с.
36. *Голицын Г. С.* Спектр космических лучей с точки зрения теории подобия // *Письма в Астрон. журн.* 1997 Т 23 С 127–132.
37. *Kanamori H.* Mechanics of earthquakes // *Annu. Rev. Earth Planet. Sci.* 1994. V. 22. P. 207–237.
38. *Григорян С. С.* О механизме генерации землетрясений и содержании эмпирических соотношений в сейсмологии // *ДАН СССР.* 1988 Т. 299, № 6 С. 1094–1101.
39. *Rundle J. B.* Derivation of the complete Gutenberg – Richter magnituge – frequency relation using the principle of scale invaviance // *J. Geophys. Res* 1989 V 96 (6). P. 12337–12342.
40. *Голицын Г. С.* Землетрясения с точки зрения теории подобия // *ДАН* 1996 Т. 346, № 4. С. 536–539.
41. *Kagan Y. Y.* Seismic moment-frequency relation for shallow earthquakes: regional comparison // *J. Geophys. Res.* 1997 V. 102 P. 2835–2852.
42. *Голицын Г. С.* Статистика и энергетика тропических циклонов // *ДАН* 1997 Т 354, № 4 С 535–538.
43. *Руденко В. Н., Милюков В. К., Нестеров В. В., Иванов И. П.* Состояние установок Московского университета по приему гравитационных волн // *Вестник МГУ Астрон астрофиз* 1994 Т. 5 С 93–99
44. *Голицын Г. С.* Некоторые закономерности наведенной сейсмичности // *ДАН* 1998 Т. 361, № 6 С 822–824
45. *Голицын Г. С.* Звездотрясения на SGR 1806-20 и других нейтронных звездах // *Письма в Астрон журн* Т 24, № 4 С 716–721.
46. *Голицын Г. С.* Исследование конвекции с геофизическими приложениями и аналогиями Л *Гидрометеоздат*, 1980 С. 56 (см также: *G. S. Golitsyn* // *J Fluid Mech.* 1979. V. 95 P. 567–608).
47. *Голицын Г. С.* Принцип скорейшей реакции в гидродинамике, геофизике, астрофизике // *ДАН* 1997 № 3 С 321–324.
48. *Голицын Г. С.* Конвективные режимы на различных вращающихся геофизических и астрофизических объектах // *Изв АН СССР. ФАО* 1991 Т 27, № 1 С 13–21.

ДВЕ КОМПОНЕНТЫ СЕЙСМИЧЕСКОЙ АКТИВНОСТИ ЗЕМЛИ И ИХ СВЯЗЬ С ОСОБЕННОСТЯМИ СУТОЧНОГО ВРАЩЕНИЯ

А. М. Фридман, А. В. Клименко

«Сейсмология сложилась позднее большинства физических наук. Сейчас также трудно представить себе сейсмолога без его основного прибора – сейсмографа, как и астронома без телескопа. Между тем телескоп был построен около 1600 г., а первые эффективные сейсмографы – между 1879 и 1890 гг.»

Чарльз Рихтер,
«Элементарная сейсмология»

1. Предварительные замечания

Предсказание землетрясений требует формулировки достаточного признака их возникновения, который пока неизвестен. В качестве одной из кажущихся сейчас актуальными проблем в сейсмологии можно назвать, например, следующую: установление корреляций между временной характеристикой сейсмической активности, $N(t)$, являющейся годовым числом землетрясений с магнитудой выше некоторой, и основными временными функциями, характеризующими вращение Земли: ее угловой скорости $\Omega(t)$ и угловым ускорением $|d\Omega/dt|$.

Нам представляется, что ответ на поставленный выше вопрос окажется полезным при решении одной из наиболее трудных проблем естествознания – построении динамической модели сейсмической активности Земли. При этом главная трудность сейсмологических задач, по сравнению, например, с астрономическими, состоит не только в том, что экспериментальная сейсмология была развита значительно позже наблюдательной астрономии, но и в сложности внутреннего строения Земли, делающей построение сейсмической динамики необычайно трудным. Именно поэтому нам представляется привлекательным построение прежде всего феноменологической модели сейсмической активности (аналогично тому, как термодинамика была построена существенно раньше статистической физики). Для этого и потребуется, в частности, знание тех корреляционных соотношений, о которых упоминалось выше. Однако прежде чем приступить к описанию этих достаточно важных корреляций между особенностями вращения Земли и функцией, характеризующую ее

сейсмическую активность $N(t)$, следует, видимо, хотя бы очень кратко пояснить, как определяются величины $N(t)$, т. е. как детектируются землетрясения и их основные характеристики.

Из очага землетрясения (размер которого зависит от энергии землетрясения, примем для определенности, что его порядок составляет несколько километров) испускаются сейсмические волны. Последние представляют собой низкочастотные звуковые волны в твердой упругой Земле. Они делятся на поверхностные и объемные. Первые распространяются вдоль земной поверхности, вторые, в свою очередь, делятся на продольные (упругие волны сжатия) и поперечные (упругие волны сдвига). Объемные волны в буквальном смысле слова «просвечивают» нашу планету и подобно рентгеновскому анализу позволяют выявлять внутреннее строение Земли без непосредственного проникновения в ее недра (рис. 1). Скорость продольных волн в 1,7 раза больше скорости поперечных, поэтому продольные волны первыми регистрируются на сейсмографах и называются P -волнами (primary). Поперечные волны называются S -волнами (secondary).

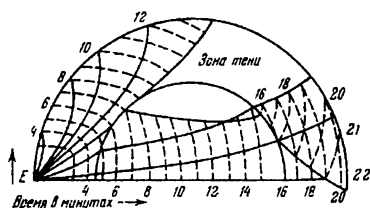


Рис. 1. Распространение сейсмических P -волн в Земле. Сечение Земли показывает пути сейсмических P -волн, излучаемых из очага землетрясения, расположенного непосредственно под эпицентром (точка E) Прерывистые линии (изохроны) указывают время прихода в минутах P -волн в различные точки земной поверхности P -волны не регистрируются в обширной зоне тени, что обусловлено преломлением этих волн на границе мантия – ядро [1]

Скорости объемных волн выражаются через модули упругости (k – модуль сжатия, μ – модуль сдвига) и плотность ρ среды известными выражениями:

$$v_p = \sqrt{\left(k + \frac{4}{3}\mu\right) / \rho}, \quad v_s = \sqrt{\mu / \rho}. \quad (1)$$

Согласно полученным сейсмологией данным Землю можно условно считать состоящей из трех основных областей: коры, мантии и ядра (рис. 2). Кора отделена от мантии резкой сейсмической границей, на которой свойства среды скачкообразно изменяются (возрастают скорости v_p и v_s и плотность ρ). Толщина земной коры нерегулярна, она изменя-

ется примерно от 10 км (с учетом толщи воды) в океанах до 100 км и более в горных районах континентальных областей. Вклад земной коры в полную массу Земли и ее момент инерции мал, поэтому обычно при рассмотрении Земли в целом земную кору представляют в виде однородного слоя эффективной толщины порядка 35 км. Ниже коры в интервале глубин 35–2885 км расположена силикатная оболочка, или мантия Земли.

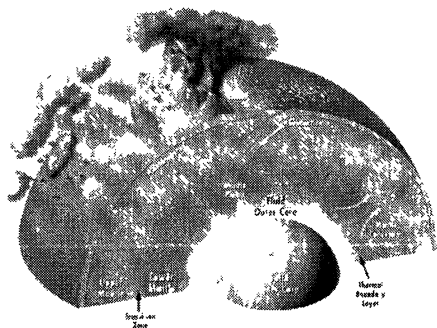


Рис. 2. Внутреннее строение Земли, состоящее из трех основных областей: коры, мантии и ядра.

Наконец, центральная часть Земли расположена в интервале глубин 2885–6371 км, образуя ядро. Так как средняя плотность Земли $\rho_{\text{ср}} \approx 5,5 \text{ г/см}^3$, а плотности гранита и базальта, образующих поверхностный слой коры, равны соответственно 2,8 и 3,0 г/см^3 , то ядро должно быть достаточно тяжелым. Заметим, что плотность железных метеоритов составляет $\approx 7,85 \text{ г/см}^3$.

На границе мантия – ядро скорость P -волн скачком падает от 13,6 до 8,1 км/с, скорость S -волн падает от 7,3 км/с до 0, плотность возрастает от 5,5 до 10 г/см^3 . Так как в ядре $v_S = 0$, следовательно, согласно (1) $\mu = 0$, т. е. ядро является жидким.

Чарльз Рихтер ввел в сейсмологию понятие магнитуды землетрясения M_S , которая представляет собой количественную характеристику последнего, связанную с параметрами поверхностных волн, возбуждаемыми этим землетрясением (шкала Рихтера):

$$M_S = \lg \frac{(a/1 \text{ мкм})}{(T/1 \text{ с})} + f(\Delta, h) + C, \quad (2)$$

где a – амплитуда смещения почвы в микронах в поверхностных волнах с периодом T в секундах ($T \sim 20 \text{ с}$), Δ – эпицентральный расстояние, h – глубина очага землетрясения (рис. 3). Эмпирическая формула $f(\Delta, h)$ позволяет приводить все наблюдения к стационарному эпицентральному расстоянию $\Delta = 100 \text{ км}$, а C – поправка на отклонение грунта от «стандартного».

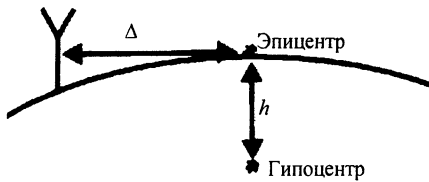


Рис. 3. Гипоцентр землетрясения, находящийся на глубине h , его проекция на поверхность земли, называемая эпицентром, и расстояние от эпицентра до сейсмической станции – Δ

Логарифмическая шкала позволяет одной формулой охватить значительный интервал в интенсивности землетрясений. Сильнейшие землетрясения характеризуются магнитудами, большими восьми баллов по шкале Рихтера. Так, катастрофическое чилийское землетрясение 22 мая 1960 г. имело магнитуду $M_S = 8,3$. Это землетрясение вошло в историю еще и потому, что после него впервые был зарегистрирован весь спектр собственных колебаний Земли. Магнитуды слабых толчков, которые регистрируются, доходят до отрицательных величин (вплоть до -3).

Глубокие землетрясения излучают объемные волны. Характеризующая их магнитуда обозначается m и определяется формулой аналогичной (2), где период $T = 1$ с. Между обеими магнитудами существует связь (шкала Гутенберга – Рихтера):

$$m = 2,5 + 0,63M_S. \quad (3)$$

Определение магнитуды было выбрано таким образом, чтобы величины отражали энергетическую характеристику землетрясений. Представим себе, что сейсмическое излучение очага землетрясения состояло бы только из монохроматической волны с периодом $T = 20$ с. Тогда энергия такого излучения $E_S \sim v^2$, где $v \sim a/T$ – скорость вертикальных колебаний почвы в поверхностной волне, тогда согласно (2) и (3)

$$\lg E_S \sim 2 \lg v \sim 2 \lg \frac{a}{T} \approx BM_S + A, \quad (4)$$

где $B = 2$, $A = -\lg f(\Delta, h) - C$. (5)

Из-за более сложного спектра излучения и учета некоторых дополнительных факторов на практике имеем

$$B \approx 1,5, \quad A \approx 11,8. \quad (6)$$

Окончательно можно записать

$$\lg E_S \approx 11,8 + 1,5M_S. \quad (7)$$

Отсюда легко оценить энергию чилийского землетрясения (в эргах), подставив в (7) магнитуду $M_S = 8,3$. Получим $E_S \sim 10^{24}$ эрг.

По глубине гипоцентров землетрясения классифицируются на неглубокие (литосферные), $h < 70$ км, промежуточные (астеносферные), $h \sim 70-300$ км, и глубокие, $h > 300$ км. Глубже 720 км землетрясения не наблюдались.

2. Две компоненты сейсмической активности Земли

На рис. 4 изображены временные кривые ежегодного числа землетрясений с $m > 4,0$; $4,5$; $5,0$; $5,5$ с 1964 по 2000 гг. Здесь и ниже используются данные каталога Национального информационного центра по землетрясениям Геологической службы США [2], начиная лишь с 1964 г., когда определения землетрясений с магнитудой $m \geq 4,0$ проводились наиболее надежно для всей Земли. На всех кривых были видны два основных пика сейсмической активности (1965, 1983–1986 и 1997 гг.) и менее выраженный пик 1976–1977 гг. Увеличение сейсмической активности в середине 60-х, 80-х и 90-х годов и более спокойный промежуточный интервал (с медленным повышением активности) можно считать глобальной закономерностью сейсмического режима Земли в рассматриваемый период, поскольку отмеченные выше особенности наблюдаются не только для различных магнитуд, но и в каждом из полушарий: Северном, Южном, Западном и Восточном. Действительно на рис. 5, а, в, где показано изменение числа землетрясений с $m > 4,0$ и $4,5$ в 1964–1988 гг. в Северном и Южном полушариях, мы вновь видим общие максимумы сейсмической активности в середине 60-х, 80-х и 90-х годов. Что же касается интервала так называемой «спокойной» сейсмичности, то здесь хорошо прослеживается зеркальная симметрия графиков числа землетрясений в Северном и Южном полушариях.

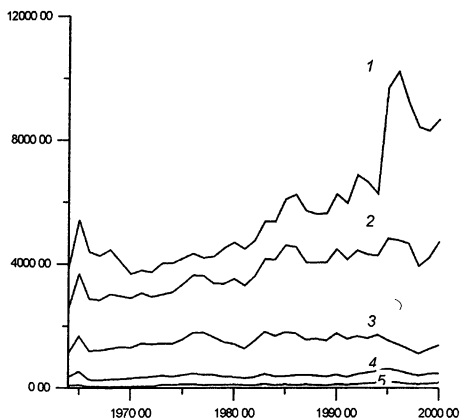


Рис. 4. Вариации глобальной сейсмической активности Земли (1964–1988 гг.). Цифрами обозначены временные кривые ежегодного числа землетрясений $N(t)$ с соответствующей магнитудой m ; кривая 1 ($m \geq 4,0$), 2 ($m \geq 4,5$), 3 ($m \geq 5,0$), 4 ($m \geq 5,5$), 5 ($m \geq 6,0$) Данные 1964–2000 гг

Определим, из каких компонент состоит сейсмическая активность различных полушарий. На рис. 5, б, г показаны сглаженные данные по каждому полушарию (скользящее среднее по 5 годам) и разность между годовыми данными и сглаженной кривой. Если коэффициент корреляции между кривыми, изображенными на рис. 5, а, составляет $Q = 0,91 \pm 0,03$, то между сглаженными кривыми (см. рис. 5, б) $Q_T = 0,98 \pm 0,01$ (используемые формулы для вычисления коэффициентов корреляции приведены в Приложении). В последнем случае мы ввели индекс T (*Total* – глобальный), указывающий на то, что сглаженные кривые отражают временное

изменение сейсмической активности всего земного шара. У кривых на рис. 5, а корреляция $Q < Q_T$, следовательно, Q содержит кроме крупномасштабной Q_T некую мелкомасштабную составляющую. Как следует из вида нижних кривых (см. рис. 5, б), в «спокойный» период мелкомасштабная компонента проявляет отчетливую зеркальную симметрию с отрицательными коэффициентами корреляции, в то время как во время «активного» периода корреляция между M -компонентами сейсмической активности сильно положительная: $Q_M = 0,41 \pm 0,14$ для всего временного интервала и $Q_M = -0,82 \pm 0,09$ для спокойного интервала 1966–1983 гг. В обозначении коэффициента корреляции, характеризующего зеркальную симметрию, используется индекс M (*Mirror* – зеркальный). Для сглаженных кривых (см. рис. 5, з) коэффициент корреляции $Q_T = 0,95 \pm 0,02$, для мелкомасштабной M -компоненты – $Q_M = -0,63 \pm 0,13$ для 1966–1983 гг.

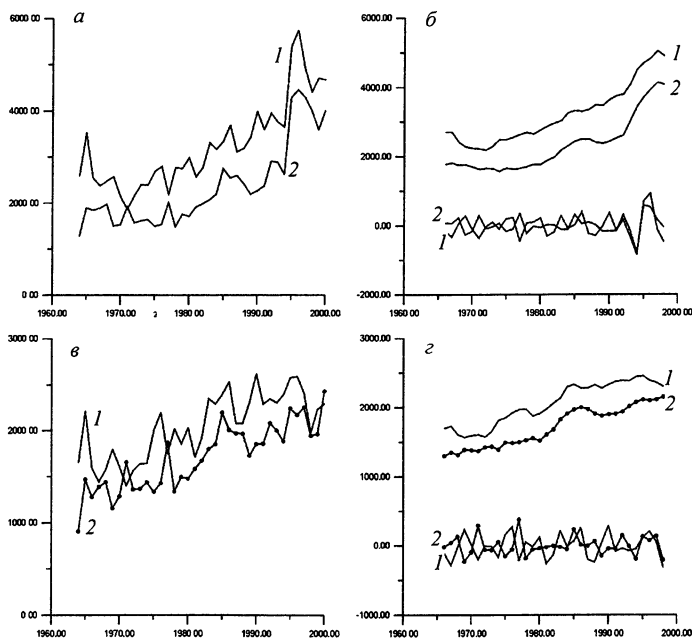


Рис. 5. Изменение числа землетрясений с $m > 4,0$ (а, б) и с $m > 4,5$ (в, з) в Северном (1) и Южном (2) полушариях Земли:

а – годовые данные, коэффициент линейной корреляции между двумя кривыми $Q = 0,91 \pm 0,03$,

б – скользящее среднее по 5 годам, коэффициент линейной корреляции $Q_T = 0,98 \pm 0,01$, разность между годовыми и усредненными данными для каждого полушария, коэффициент линейной корреляции $Q_M = 0,41 \pm 0,14$ для всего временного интервала, $Q_M = -0,82 \pm 0,09$ для «спокойного» периода 1966—1983 гг,

в – годовые числа событий для Северного и Южного полушарий,

з – сглаженные кривые и разности. Для двух верхних кривых $Q_T = 0,95 \pm 0,02$, для двух нижних $Q_M = -0,22 \pm 0,17$ за весь период и $Q_M = -0,63 \pm 0,15$ для 1966–1983 гг

Аналогичные результаты были получены как для Западного и Восточного полушарий [3], так и для северных и южных частей Восточного и Западного полушарий. В двух последних случаях приведем коэффициенты корреляции: $Q_T = 0,82 \pm 0,06$, $Q_M = -0,63 \pm 0,13$ и $Q_I = 0,89 \pm 0,03$, $Q_M = -0,80 \pm 0,21$. Величина Q_M вычислялась для «спокойного» периода 1966–1983 гг.

Изложенное выше демонстрирует наличие двух компонент сейсмической активности Земли (T - и M -компонент) для наиболее крупных регионов (минимальный размер – 1/4 поверхности Земли). В работе [3] аналогичные результаты были получены для регионов, размер которых составлял 1/16 часть земной поверхности, причем показано, что зеркальная симметрия M -компоненты наблюдается только в сейсмически «спокойные» годы, тогда как в годы высокой сейсмической активности корреляция M -компонент существенно положительна.

3. Корреляция кривой $N(t)$, описывающей временное изменение годовой сейсмической активности Земли, с кривыми $\Omega(t)$ и $|d\Omega/dt|$, описывающими временное изменение угловой скорости вращения Земли и ее углового ускорения

На рис. 6 показаны графики временной зависимости угловой скорости вращения Земли $\Omega(t)$ и числа слабых землетрясений $N(t)$ с магнитудой $m \geq 4,0$. (Усредненные по годам данные о неравномерности вращения Земли взяты из работы Моррисона [4] и затем продолжены до 2000 г. усреднением по годам данных из [5]). Коэффициент корреляции $Q = 0,40 \pm 0,14$ для $m \geq 4,0$, и он систематически падает с ростом магнитуды:

$$m \geq 4,5$$

$$Q = 0,25 \pm 0,15$$

$$m \geq 5,0$$

$$Q = -0,20 \pm 0,16$$

$$m \geq 5,5$$

$$Q = 0,00 \pm 0,16$$

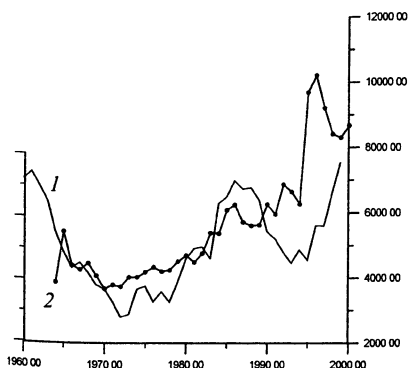


Рис. 6. Графики угловой скорости вращения Земли $\Omega(t)$ (1) и ежегодного числа землетрясений всей Земли с магнитудой $m \geq 4,0$ (2). $Q = 0,50 \pm 0,13$. При усреднении графиков по пяти годам (скользящее среднее) коэффициент корреляции становится $Q = 0,56 \pm 0,13$

Естественно ожидать, что с функцией $N(t)$, описывающей изменение со временем ежегодного числа сильнейших землетрясений, лучше будет коррелировать та характерная функция вращения Земли, которая лучше

описывает изменения $\Omega(t)$, в частности, учитывает ее большие временные производные. Очевидно, что в качестве такой функции следует выбрать $|d\Omega/dt|$, поскольку рост напряжений в земной коре должен наблюдаться как при ускорениях, так и при торможениях. Именно это изложено в работе [6].

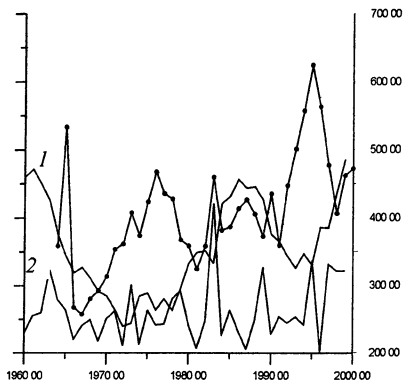


Рис. 7. Графики числа землетрясений $N(t)$ всей Земли с $m \geq 5,5$ (кривая с точками), угловой скорости вращения Земли $\Omega(t)$ и модуля ее производной $|d\Omega/dt|$. Корреляция между $N(t)$ и $\Omega(t)$ $Q = 0,14 \pm 0,17$. Корреляция между $N(t)$ и $|d\Omega/dt|$ $Q = 0,28 \pm 0,16$. При усреднении графиков $N(t)$ и $\Omega(t)$ по пяти годам получим корреляция между $N(t)$ и $\Omega(t)$ $Q = 0,23 \pm 0,17$, корреляция между $N(t)$ и $|d\Omega/dt|$ $Q = 0,51 \pm 0,13$.

На рис. 7 изображены графики числа сильных землетрясений $N(t)$ для $m \geq 5,5$ и модуля производной $|d\Omega/dt|$. Обращает на себя внимание систематическое повышение коэффициента корреляции по мере роста магнитуды m . Если при малых магнитудах корреляция вообще отсутствует, то при больших m корреляция становится статистически значимой.

Для всей Земли		Для северного полушария (см. рис. 8)	
$m \geq 4,0$	$Q = 0,09 \pm 0,20$	$m \geq 4,0$	$Q = 0,20 \pm 0,20$
$m \geq 4,5$	$Q = 0,16 \pm 0,20$	$m \geq 4,5$	$Q = 0,25 \pm 0,19$
$m \geq 5,0$	$Q = 0,33 \pm 0,18$	$m \geq 5,0$	$Q = 0,33 \pm 0,12$
$m \geq 5,5$	$Q = 0,28 \pm 0,17$	$m \geq 5,5$	$Q = 0,30 \pm 0,12$

(при $m \geq 5,5$ $Q = 0,58 \pm 0,16$ – за период 1972–1988 гг.)

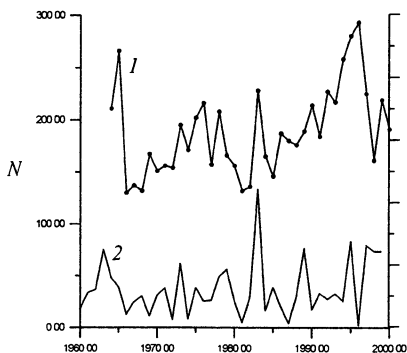


Рис. 8. Графики числа землетрясений $N(t)$ северного полушария Земли с $m \geq 5,5$ (1) и модуля производной $|d\Omega/dt|$ (2), $Q = 0,30 \pm 0,15$

4. Основная гипотеза

Наблюдаемые вариации угловой скорости вращения Земли, $\Omega(t)$, и ее временной производной, $|d\Omega/dt|$, должны неизбежно вызывать напряжения (растяжения–сжатия) во всех трех основных областях Земли (коре, мантии и ядре). В поверхностном слое Земли это может привести, например, к возмущению движения литосферных плит. Следствием последнего могут являться разрушения границы соприкасающихся плит, воспринимаемые как повышенная сейсмическая активность в пограничных зонах плит (землетрясения в областях субдукции и спрединга).

Ниже представим некоторые аргументы в пользу этой гипотезы, высказанной ранее в работах [6–9].

Оценим вариацию энергии вращения поверхностного слоя Земли глубиной L и сравним с энергией, выделяемой при землетрясениях за тот же промежуток времени. Энергия вращения Земли равна (см., например, Allen, 1958):

$$E = \frac{1}{2} I \Omega^2 \approx \frac{1}{5} M R^2 \Omega^2 = \frac{4}{5} \pi R^5 \rho \Omega^2 \approx 2 \times 10^{36} [\text{эрг}], \quad (8)$$

где R , ρ – средний радиус и плотности Земли. Энергия вращения поверхностного слоя Земли толщиной L (км) есть

$$E_L \approx E \cdot 5 \frac{L}{R} \approx 0,4 \left(\frac{L}{300} \right) \cdot 10^{36} [\text{эрг}] = 0,2 \left(\frac{L}{300} \right) E, \quad (9)$$

При оценке мы полагали $L/R \ll 1$.

Как следует из астрономических наблюдений, характерное изменение периода вращения Земли в течение 10 лет составляет $\Delta T \approx (1 \div 5) \cdot 10^{-3}$ с [7, 8].

Оценим вариацию энергии вращения поверхностного шарового слоя Земли глубиной L км за 10 лет, используя упомянутую выше оценку изменения периода вращения ΔT за это время

$$\delta E_L \approx E_L \cdot 2 \frac{\Delta \Omega}{\Omega} \approx (1 \div 5) \cdot \left(\frac{L}{300} \right) \cdot 10^{28} [\text{эрг}], \quad (10)$$

где использовано условие $\Delta \Omega / \Omega \ll 1$.

Согласно с [9], максимальная энергии землетрясений за 10 лет составляет

$$\delta E_{eq} \approx 5 \cdot 10^{25} [\text{эрг}]. \quad (11)$$

Следовательно, величина энергетического «резервуара», возникающая вследствие изменения угловой скорости вращения даже части сейсмически активного слоя Земли глубиной 300 км, на 3 порядка превосходит энергию землетрясений за тот же период времени. Поскольку локализация гипоцентров землетрясений отнюдь не ограничена 300-кило-

метровой глубиной (как мы видим, в этом случае мы пренебрегаем глубокими, часто являющимися наиболее сильными, землетрясениями), то сейсмическая энергия δE_{EQ} вполне может черпаться из наблюдаемого изменения энергии вращения Земли δE_L .

Показанный в предыдущем параграфе факт отсутствия корреляции между числом слабых землетрясений $N(t)$ и $|d\Omega/dt|$ не позволяет сделать вывод об отсутствии связи землетрясений малых магнитуд с ускорением вращения Земли.

Действительно, проведенное авторами [10] усреднение двух графиков $\bar{N}(t)$ и $|d\Omega/dt|$ по нескольким годам и смещение назад по времени графика $\bar{N}(t)$ привели к высокой корреляции между ними. Последний факт свидетельствует о наличии причинно–следственной связи сейсмической активности Земли с угловым ускорением (замедлением) вращения Земли: причина – ускорение (замедление) вращения, следствие – рост глобального числа землетрясений [10].

Ниже, следуя изложенному в работе [11]*, приведем результаты корреляционного анализа связи глобальной сейсмичности с неравномерностью вращения Земли в зависимости от глубины очагов землетрясений, типа границ литосферных плит (зон спрединга и субдукции) и других особенностей регионов землетрясений.

Затем продемонстрируем явление отрицательной корреляции между M -компонентами, обнаруженными в сейсмической активности на границах тектонических плит: в зонах субдукции, в системе срединно–океанических хребтов и континентальных горных систем. И наконец, покажем, что сейсмическая активность Земли зависит от широты, что является дополнительным аргументом в пользу важной роли неоднородностей вращения Земли в инициировании землетрясений.

5. Сейсмичность зон спрединга и неравномерность вращения Земли

На карту мира нанесены все землетрясения за 1964–2000 гг. с магнитудой $m \geq 4,0$ (рис. 9). Самые широкие сейсмоактивные области по краям Тихого океана принадлежат зонам субдукции, а более узкие дорожки сейсмоактивности по срединно–океаническим хребтам – областям спрединга (зоны субдукции отличаются также и по значительно большей глубине землетрясений).

* При этом каталог землетрясений [2] дополнен данными последнего десятилетия (1991–2000)

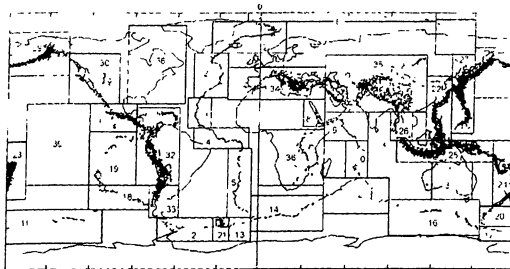


Рис. 9. Карта землетрясений с магнитудой $m \geq 4,0$ (отмечено точками) за период 1964–2000 гг. Указаны границы и нумерация рассматриваемых в статье регионов. Участки 1–20 соответствуют областям спрединга, 21–39 – зонам субдукции

Проанализируем суммарную сейсмичность мировой системы срединно-океанических хребтов (Южно- и Северо-Атлантические хребты и их арктическое продолжение, Восточно-Тихоокеанское поднятие, разлом Красного моря и Аравийско-Индийский хребет, антарктическое кольцо поднятий и хребтов – зоны 1–20 по рис. 9).

В табл. 1 показаны коэффициенты корреляции (в зависимости от величины магнитуды) между годовыми значениями $|d\Omega/dt|$ и ежегодным числом землетрясений во всех зонах спрединга. Все землетрясения в этих зонах, как правило, сосредоточены в слое 0–50 км. Корреляция достаточно заметна особенно после 1969 г. Заметим, что для наиболее сильных землетрясений ($m \geq 6,0$) корреляция отсутствует.

Таблица 1

Коэффициент корреляции Q между годовыми значениями $|d\Omega/dt|$ и ежегодным числом землетрясений в области спрединга

Магнитуда	Средняя величина Q за 1964–2000 гг.	Средняя величина Q за периоды
<i>А. Во всех зонах спрединга</i>		(1969–1980)
> 4,0	$0,27 \pm 0,16$	$0,45 \pm 0,17$
> 4,5	$0,22 \pm 0,16$	$0,46 \pm 0,17$
> 5,0	$0,13 \pm 0,17$	$0,53 \pm 0,15$
> 5,1	$0,22 \pm 0,16$	$0,55 \pm 0,15$
> 5,5	$0,30 \pm 0,15$	$0,30 \pm 0,19$
> 6,0	$0,19 \pm 0,16$	$-0,04 \pm 0,21$
<i>Б. На стыке Африканской и Индийской плит (зоны 8–10)</i>		(1972–1990)
> 4,0	$0,43 \pm 0,14$	$0,67 \pm 0,14$
> 4,5	$0,47 \pm 0,13$	$0,76 \pm 0,10$
> 5,0	$0,40 \pm 0,14$	$0,73 \pm 0,11$
> 5,5	$0,23 \pm 0,16$	$0,49 \pm 0,17$
<i>В. Вдоль Аравийско-Индийского океанического хребта (зона 10)</i>		(1969–1990)
> 4,0	$0,56 \pm 0,12$	$0,69 \pm 0,11$
> 4,5	$0,55 \pm 0,12$	$0,69 \pm 0,11$
> 5,0	$0,50 \pm 0,13$	$0,66 \pm 0,12$
> 5,5	$0,32 \pm 0,15$	$0,07 \pm 0,21$

Рассмотрим активность границы спрединга между Африканской и Индийской плитами (Красное море–Аравийско-Индийский срединно-океанический хребет или зоны 8–10 согласно рис. 9). Результаты исследования этого региона для различных магнитуд приведены в табл. 1 (пп. Б и В). Корреляции велики для 1972–1990 гг., зон 8–10 и для 1969–1990 гг., зоны 10. В последнем случае для наиболее сильных землетрясений ($m \geq 5,5$) корреляция вновь отсутствует.

6. Связь сейсмичности отдельных зон субдукции с неравномерностью вращения Земли

Исследуем связь сейсмичности зон субдукции $N(t)$ с изменениями длительности суток $|d\Omega/dt|$ (табл. 2). Зоны субдукции включают Тихоокеанское кольцо: от Новой Зеландии через Аляску до Чили, а также ветви Филиппинского и Зондского желобов и область Южно-Сандвичева желоба (зоны 21–39 согласно рис. 9). Как следует из табл. 2, немалая корреляция между годовыми значениями $|d\Omega/dt|$ и ежегодным числом землетрясений в зонах субдукции $N(t)$ имеется лишь при значительных магнитудах ($m \geq 5,5-6,0$), т. е. подтверждается вывод статей [10, 12]. В них же, в частности, показано, что при усреднении двух кривых, $N(t)$ и $|d\Omega/dt|$, по интервалу 3–5 лет, когда тем самым доминирующей в сейсмическом процессе оказывается T -компонента, коэффициенты корреляции возрастают при всех значениях m и оказываются значимыми при $m \geq 4,0$. Таким образом, с модулем производной скорости вращения Земли коррелирует суммарная сейсмическая активность как зон спрединга, так и зон субдукции.

Таблица 2

Коэффициент корреляции Q между годовыми значениями $|d\Omega/dt|$ и ежегодным числом землетрясений во всех областях субдукции

Магнитуда	Средняя величина Q за период 1964–2000 гг.	Средняя величина Q за период 1969–1988 гг.
> 4,0	0,24 ± 0,16	0,04 ± 0,22
> 4,5	0,18 ± 0,16	0,11 ± 0,22
> 5,0	0,01 ± 0,17	0,33 ± 0,20
> 5,5	0,23 ± 0,16	0,51 ± 0,17
> 6,0	0,27 ± 0,16	0,57 ± 0,15
> 6,1	0,27 ± 0,16	0,41 ± 0,19

Связь глубины землетрясений с долготой в интервале 135–145° в. д. и 35–40° с. ш. (о. Хонсю) показана на рис. 10. Виден горизонтальный слой землетрясений от 0 до 70 км в толще Евразийской литосферной плиты и наклонный слой землетрясений вдоль опускающейся Тихоокеанской плиты вплоть до 500 км. Особенно сейсмически активна область соприкосновения Тихоокеанской и Евроазиатской плит. Именно эти землетрясения наиболее сильно коррелируют с ускорением вращения Земли.

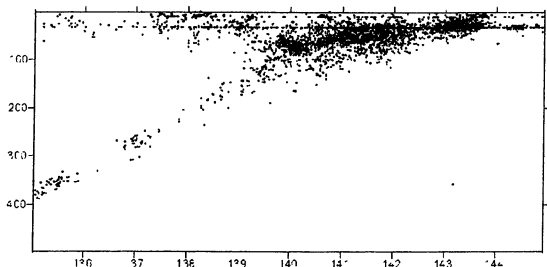


Рис. 10 Связь глубины землетрясений с $m \geq 4,0$ с долготой в интервале $135-145^\circ$ в д и $35-40^\circ$ с ш. (о Хонсю) Виден горизонтальный слой землетрясений в толще Евразийской литосферной плиты от 0 до 70 км и наклонный слой землетрясений вдоль опускающейся Тихоокеанской плиты.

Результаты описанных выше исследований, представленные в табл. 3 (см. более подробно [11], где установлена также связь внутриплитовых землетрясений с неравномерностью вращения Земли), дают возможность сделать уверенный вывод о связи ускорения (замедления) суточного вращения Земли и глобальной сейсмической активности. Эта корреляция прослеживается по всем типам сейсмоактивных регионов: зонам спрединга, субдукции, касательного сдвига плит и внутриплитовых землетрясений.

Таблица 3

Сводная таблица результатов исследования корреляции между значениями $|d\Omega/dt|$ и ежегодным числом землетрясений (скользящее среднее по пяти годам для $N(t)$ и $\Omega(t)$)

Регион	Магнитуда	Коэффициент корреляции за весь период 1964–2000 гг.
Вся планета	$\geq 4,0$	$0,47 \pm 0,15$
	$\geq 5,0$	$0,17 \pm 0,18$
	$\geq 6,0$	$0,54 \pm 0,13$
Все зоны спрединга	$\geq 4,0$	$0,49 \pm 0,14$
	$\geq 5,0$	$0,49 \pm 0,14$
	$\geq 6,0$	$0,68 \pm 0,10$
Все зоны субдукции	$\geq 4,0$	$0,51 \pm 0,13$
	$\geq 5,0$	$0,30 \pm 0,16$
	$\geq 6,0$	$0,61 \pm 0,11$

7. *M*- и *T*-компоненты

1. *M*- и *T*-компоненты в наиболее сейсмически активном районе Земли на стыке Евразийской, Тихоокеанской и Индо-австралийской плит. В этой части статьи попытаемся на некоторых частных и, на наш взгляд, ярких примерах проиллюстрировать характер *M*-компоненты

сейсмической активности – пока только на феноменологическом уровне. Где могут лучше проявиться черты M -компоненты как не в наиболее сейсмически активном районе Земли?

Ранее мы привели ряд примеров проявления зеркальной симметрии M -компоненты сейсмичности, в частности между северной и южной частями Восточного полушария коэффициент корреляции $Q_M = -0,63 \pm 0,13$ (при усреднении по 5 годам событий с $m \geq 4,0$ за 1964–1988 гг.). Разбивая основные сейсмические зоны на отдельные регионы и варьируя их границы, можно определить, что коэффициент отрицательной корреляции наиболее высок между двумя районами: Гималайско-Фиджийским и Филиппинско-Камчатским – рис. 11, на котором указаны усредненные положения землетрясений относительно разлома в диапазоне глубин 70–120 км, всех толчков в интервале 120–400 км и глубже 400 км. Направления движения плит – от мелких землетрясений к более глубоким. Острова не нанесены (за исключением о. Калимантан), чтобы не затенять значки землетрясений.

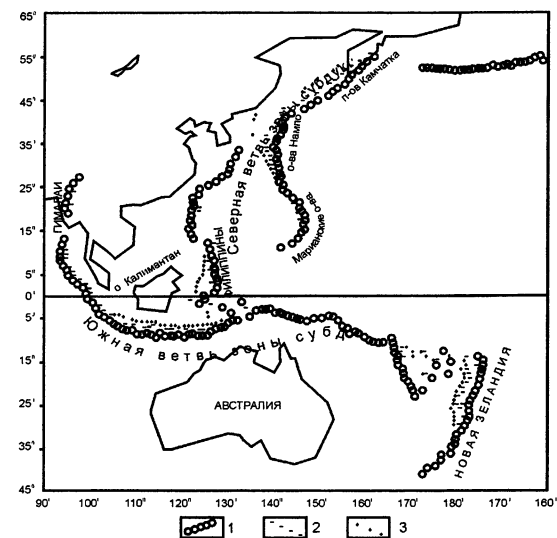


Рис. 11. Самый активный сейсмический район Земли на стыке Евразийской, Тихоокеанской и Индо-Австралийской плит. Нанесены землетрясения с гипоцентрами 1 – 70–120 км, 2 – 120–400 км; 3 – > 400 км

В данных ветвях, где происходит примерно половина всех землетрясений мира, $Q_M = -0,74 \pm 0,09$ (при усреднении по 5 годам событий с $m \geq 4,0$). Северной ветвью зоны субдукции будем называть область, расположенную в Северном полушарии (до 65° с. ш.) между 115 и 162° в. д. В нее же включена зона Южного полушария с координатами 0–5° ю. ш. и 115–130° в. д. (южное окончание Филиппинской зоны). Южной ветвью

зоны субдукции будем называть район, который начинается в Северном полушарии – в Гималаях, южнее 30° с. ш. и идет на юг до экватора. Далее (за исключением упомянутой южной области Филиппин) эта ветвь включает в себя все землетрясения от экватора до 22° ю. ш. Область, лежащую восточнее Южной ветви почти перпендикулярно ее основной части, назовем Новозеландской зоной.

На рис. 12, а изображены графики количества землетрясений с $m \geq 4,0$ в Южной и Северной ветвях зон субдукции. Отчетливо видны две особенности графика – плавный рост активности в обеих ветвях (обусловленный, видимо, ростом сети сейсмостанций) и зеркально-симметричное расположение экстремумов на обеих кривых.

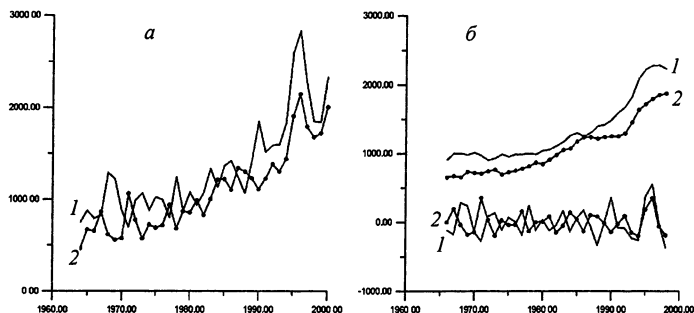


Рис. 12. Количество ежегодных землетрясений за 1964–2000 гг. в Северной (1) и Южной (2) ветвях зон субдукции. Магнитуда толчков $m \geq 4,0$ а – годовые данные $Q = 0,79 \pm 0,06$; б – две составляющие сейсмической активности Северной и Южной ветвей: усредненные по пяти годам T -компоненты (верхние кривые) и быстропеременные M -компоненты (нижние). Коэффициенты корреляции за весь период времени $Q_T = 0,96 \pm 0,01$ и $Q_M = -0,12 \pm 0,17$, за «спокойный» период 1964–1990 гг. $Q_T = 0,89 \pm 0,04$ и $Q_M = -0,60 \pm 0,13$

Приведенные на рис. 12, б верхние кривые суть кривые рис. 12, а, усредненные по 5 годам. Нижние кривые на рис. 12, б являются результатом вычитания значений верхних (усредненных) кривых из значений неусредненных кривых рис. 12, а. Число и магнитуда сейсмических событий такие же, как и на рис. 12, а. Факт зеркальной симметрии нижних кривых на протяжении всего времени отчетливо виден. Коэффициент корреляции между усредненными кривыми (T -компонентами): $Q_T = 0,89 \pm 0,04$ и между M -компонентами: $Q_M = -0,60 \pm 0,13$ за 1964–1990 гг.

Для феноменологического исследования данного явления попробуем понять основной признак отличия Южной и Северной ветвей друг от друга. Вызвана ли данная сейсмическая зеркальная симметрия их взаимно перпендикулярной ориентацией или главным фактором является то, что Южная ветвь – зона контакта между Индо-Австралийской и другими пли-

плитами (Тихоокеанской и Евразийской), а Северная – зона контакта как раз между двумя последними плитами?

На рис. 11 отчетливо видны направления движения субдуцирующих Тихоокеанской и Индо-Австралийской плит – от мелких землетрясений к более глубоким. Видно, что зоной субдукции Тихоокеанской плиты, практически перпендикулярной к направлению ее движения, являются районы: Камчатка–Япония–о-ва Нампо и Марианские и зона Филиппин, а также район Новой Зеландии. Зона субдукции Индо-Австралийской плиты идет по островам: Суматра, Ява, Новая Гвинея, Соломоновы о-ва и Новые Гебриды. Рассмотрим Новозеландский регион, который географически принадлежит к продолжению Южной ветви, но ориентирован практически поперек основной части Южной ветви так же, как Северная ветвь. При этом Новозеландская зона является вместе с Северной ветвью зоной субдукции Тихоокеанской плиты, в то время как Южная ветвь – зона субдукции Индо-Австралийской. Прибавление новозеландских землетрясений (что увеличивает общее число событий с $m \geq 4,0$ на 10%) к землетрясениям Южной ветви на 1/6 уменьшает по модулю отрицательную корреляцию: с $Q_M = -0,68 \pm 0,11$ до $Q_M = -0,57 \pm 0,14$. Прибавление этих событий к Северной ветви практически не меняет коэффициента корреляции: $Q_M = -0,69 \pm 0,11$ (здесь значение M -компоненты вычисляется для спокойного периода 1964–1990 гг.). Следовательно, динамика землетрясений в Новозеландском регионе близка к динамике сейсмичности в Северной ветви.

2. M -компонента срединно-океанических хребтов. Тихоокеанская и Индо-Австралийская зоны субдукции образуют T -образное перекрестье с центром вблизи экватора. Если в данном регионе действует фактор, перераспределяющий напряжение в точке T -образного стыка тектонических плит, то, может быть, аналогичный фактор действует и в других подобных точках пересечений разломов?

В Индийском океане, в точке с координатами 25° ю. ш. и 70° в. д. сходятся три срединно-океанических хребта (на границе зон 10 и 15, см. рис. 9): с севера идет Аравийско-Индийский хребет (мы начинаем рассмотрение этого разлома с Аденского залива, с 50° в. д. и на юге заканчиваем на перекрестье, на 25° ю. ш.). С юго-востока и юго-запада к пересечению подходят цепи антарктических хребтов. Анализ показывает, что сейсмичность Аравийско-Индийского хребта коррелирует (Q велико по модулю, но отрицательно) с сейсмичностью указанных антарктических хребтов (рис. 13, a , где кривая без точек есть m -компонента вдоль Аравийско-Индийского хребта, кривая с точками – M -компонента вдоль цепи антарктических хребтов; коэффициент корреляции $Q = -0,71 \pm 0,10$).

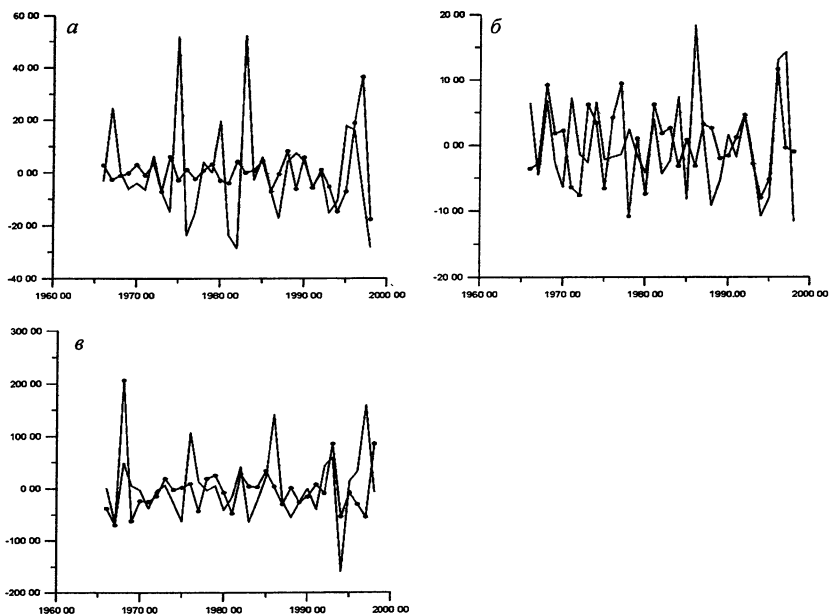


Рис. 13. Отрицательная корреляция сейсмической активности на перекрестье трех систем срединно-океанических хребтов:

a – в Индийском океане (координаты этой точки – 25° ю. ш. и 70° в. д.), $Q_M = -0,71 \pm 0,10$; *б* – на юге Атлантического океана (координаты этой точки – 53° ю.ш. и 5° з.д.), $Q_M = -0,79 \pm 0,08$, *в* – на юго-востоке Тихого океана (координаты этой точки – 30° ю. ш и 110° з. д.), $Q_M = -0,76 \pm 0,08$. Кривая с точками есть *M*-компонента сейсмичности вдоль Чилийского поднятия и центральной части Восточно-Тихоокеанского поднятия. Кривая без точек – *M*-компонента сейсмичности Южно-Тихоокеанского поднятия и южной части Восточно-Тихоокеанского поднятия. Здесь коэффициенты корреляции вычислены для «спокойного» периода 1964–1988 гг

Последующий анализ еще двух *T*-образных перекрестьев в Тихом и Атлантическом океанах (см. рис. 13, *б, в*) показал аналогичную картину.

3. Анализ сейсмичности Альпийско-Гималайского пояса. В сейсмичности Альпийско-Гималайского пояса выделим четыре характерные зоны: *Гималайскую* – от 76 до 95° в. д. и от 26 до 33° с. ш.; *Тибетскую*, которая является расширением Гималайской до 50° с. ш. и включает Гималаи, Тибет и Тянь-Шань; *Сычуаньскую* – от 95 до 106° в. д. и от 20 до 35° с. ш.; *Индокитайскую*, которая является расширением Сычуаньской до 15° ю. ш. Первые две зоны характеризуются широтными направлениями хребтов (складчатости), т. е. сжатие происходит по направлению север–юг. Для третьей и четвертой областей типичны меридиональные хребты и разломы: сжатие происходит по линии запад–восток. Как мы видели ранее, при перпендикулярном действии силы возникает

M -компонента сейсмического процесса. Из вышеуказанных четырех зон можно составить четыре парные комбинации с поперечным действием сил. Вычисление коэффициентов корреляции в каждой такой паре между усредненными (по 3 годам) кривыми активности (Q_T) и разница между ежегодными землетрясениями (Q_M) при $m \geq 4$ дают параметры, приведенные в табл. 4.

Таблица 4

Коэффициенты корреляции
между усредненными (по 3 годам) кривыми активности (Q_T)
и разницей между ежегодными землетрясениями (Q_M) при $m \geq 4$

Зона	Q_T	Q_M	Q_M за «спокойный» период 1964–1988 гг.
Гималайская–Сычуаньская	$0,80 \pm 0,06$	$0,05 \pm 0,17$	$-0,62 \pm 0,12$
Тибетская–Сычуаньская	$0,78 \pm 0,07$	$-0,12 \pm 0,17$	$-0,64 \pm 0,12$
Тибетская–Индокитайская	$0,91 \pm 0,03$	$-0,26 \pm 0,16$	$-0,73 \pm 0,09$
Гималайская–Индокитайская	$0,90 \pm 0,03$	$-0,25 \pm 0,16$	$-0,84 \pm 0,06$

8. О связи сейсмической активности Земли с широтой в зависимости от глубины эпицентров

Исследования зависимости плотности землетрясений от широты проводились ранее [14]. В качестве базы данных был использован тот же каталог Национального географического информационного центра по землетрясениям Геологической службы США VX DAT за 1962–2000 гг. [2]. Были построены графики зависимости плотности землетрясений от широты их эпицентров $\rho_{\theta_n} = N_{\theta_n} / S_{\theta_n}$, где N_{θ_n} – количество землетрясений, произошедших в интервале широт $(\theta_n - 5^\circ; \theta_n + 5^\circ)$; S_{θ_n} – соответствующая площадь, выраженная в единицах $2\pi R^2$, R – радиус Земли. Величина θ_n , которой соответствует определенное число землетрясений N_{θ_n} , удовлетворяет условию $\theta_n = 10n$, где $n = 0, \pm 1, \pm 2, \dots, \pm 9$. Поэтому все экстремумы на кривых $N(\theta)$, либо на использованных нами кривых $\rho(\theta)$, находятся лишь в точках $\theta = \theta_n$.

На рис. 14, а приведены графики распределения полной плотности землетрясений в зависимости от широты эпицентров. Кривые 1–5 соответствуют землетрясениям с магнитудой $M_b \geq 4,0; 4,5; 5,0; 5,5; 6,0$ соответственно. Видны четкая зависимость плотности землетрясений от широты для землетрясений со всеми магнитудами и два максимума – первый на 40° с. ш. и второй, меньший, на 10° ю. ш. Как мы видим, для всех значений магнитуд положение максимумов оказываются совпадающими.

На рис. 14, б представлены те же графики, но для поверхностных землетрясений с глубиной гипоцентров менее 30 км. Видно, что по сравнению с Северным полушарием в Южном полушарии поверхностных землетрясений зарегистрировано существенно меньше.

График распределения землетрясений с глубиной гипоцентров 30–100 км (см. рис. 14, в) практически повторяет таковой на рис. 14, а, с той лишь разницей, что максимальная плотность составляет 450000 вместо 800000.

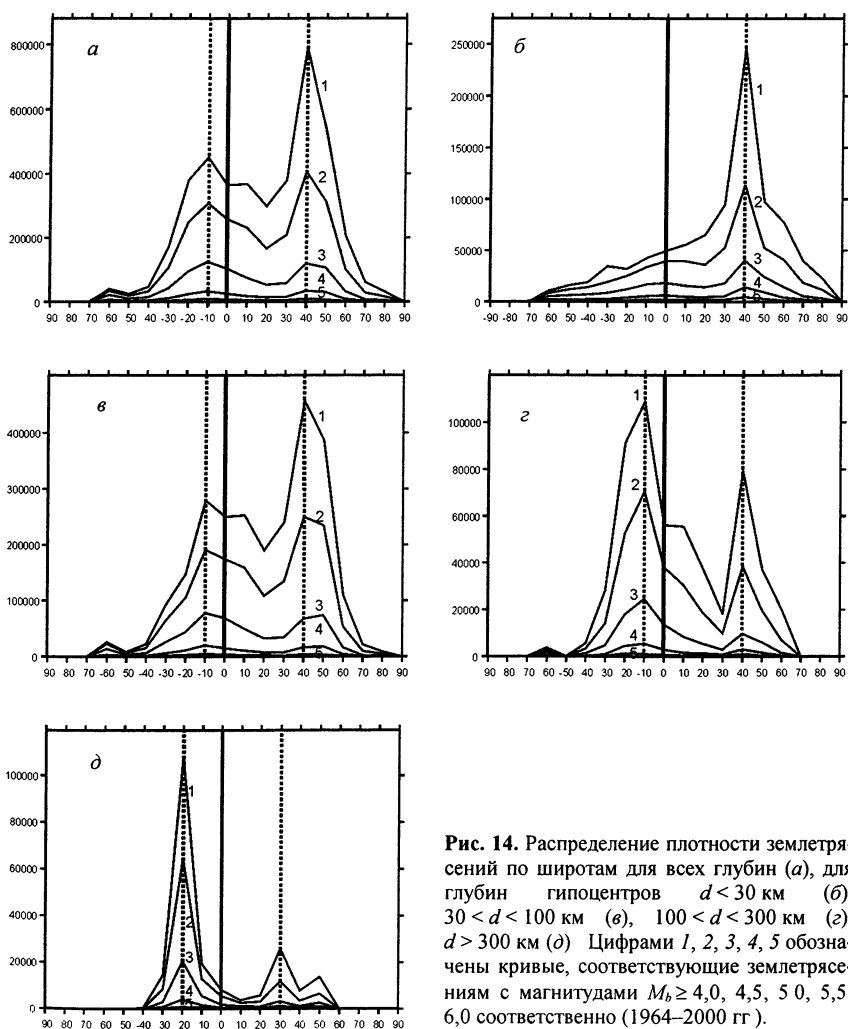


Рис. 14. Распределение плотности землетрясений по широтам для всех глубин (а), для глубин гипоцентров $d < 30$ км (б), $30 < d < 100$ км (в), $100 < d < 300$ км (г), $d > 300$ км (д) Цифрами 1, 2, 3, 4, 5 обозначены кривые, соответствующие землетрясениям с магнитудами $M_b \geq 4,0, 4,5, 5,0, 5,5, 6,0$ соответственно (1964–2000 гг.).

В интервале глубин 100–300 км (рис. 14, *з*) максимальная плотность землетрясений уже приходится на Южное полушарие. С дальнейшим ростом глубины количество землетрясений, приходящихся на Северное полушарие, резко падает, в отличие от Южного, где землетрясения происходят вплоть до глубины в 600–700 км (рис. 14, *д*).

9. Заключение

1. Как было показано в более ранних работах [3, 6, 10–13], сейсмическую активность Земли можно представить в виде двух составляющих – глобальной, медленно меняющейся *T*-компоненты и быстро меняющейся *M*-компоненты. Этот факт был подтвержден на основе более полных данных за 1964–2000 гг. Также было обнаружено принципиальное отличие в поведении *M*-компоненты в «спокойные» периоды и периоды максимальной сейсмической активности. В «спокойные» периоды *M*-компонента сейсмической активности достаточно крупных регионов демонстрирует отрицательную корреляцию, в то время как в активные периоды корреляция становится положительной, что хорошо заметно даже «на глаз», например, на графиках 5, б, в. Поскольку величина Q_M характеризует коэффициент корреляции графиков сейсмической активности в соседних регионах (плитах), т. е. является численной характеристикой *M*-компоненты, то причина разных знаков Q_M в «спокойные» и «активные» периоды может состоять в следующем: когда в одной плите происходит землетрясение, это приводит к уменьшению внутриплитовых напряжений, что, соответственно, приводит к уменьшению вероятности возникновения землетрясения в соседней плите. Но в периоды высокой сейсмической активности внутриплитовые напряжения настолько велики, что уменьшения напряжений, вызванные землетрясениями в одной области, слабо влияют на соседнюю, и поэтому землетрясения происходят в обеих областях.

2. В областях наибольшей сейсмической активности – вдоль границ тектонических плит, в системе срединно-океанических хребтов и континентальных горных системах – обнаружен широко распространенный феномен отрицательной корреляции – *M*-компонента сейсмичности (табл. 5). Характерные времена полного цикла напряжений в *M*-компоненте сейсмичности в данных точках составляют 2–3 года. Этот феномен не связан со случайными флуктуациями, так как в каждой рассмотренной системе отрицательная корреляция сейсмичности между различными ветвями сохраняется на протяжении всего «спокойного» периода (~27 лет). Обнаруженный эффект противофазы сейсмичности происходит без сдвигов по времени.

Коэффициент корреляции между M -компонентами
(разность между годовыми и усредненными по 3 годам данными)

Регион	Тип сравниваемых разломов	Для M -компоненты	Магнитуда
Запад Тихого океана	Зоны субдукции	$-0,75 \pm 0,09$	$> 4,0$
Индийский океан	Перекресток срединных хребтов	$-0,71 \pm 0,10$	$> 4,0$
Атлантический океан	—"	$-0,79 \pm 0,08$	$5,4 > m > 4,3$
Юго-восток Тихого океана	—"	$-0,76 \pm 0,08$	$6,1 > m > 4,5$
Восток Тихого океана	—"	$-0,57 \pm 0,14$	$> 4,1$
Антарктика	Противолежащие срединные хребты	$-0,77 \pm 0,08$	$5,4 > m > 4,3$
Тибет – Индокитай	Континентальные хребты с разной ориентацией	$-0,77 \pm 0,08$	$5,0 > m > 4,0$

3. Для тех же регионов (i) усредненные кривые сейсмоактивности $\bar{N}_i(t)$ сильных ($> 5,4$) землетрясений почти всегда демонстрируют высокую степень положительной корреляции (T -компонента сейсмической активности). Типичные времена вариаций сейсмичности в этом случае 5–6 лет.

Таким образом, существует, по крайней мере, два класса динамических факторов, определяющих изменение сейсмичности в крупных тектонических структурах:

а) Глобальный механизм (T -компонента), определяющий изменение во времени общего количества сильных землетрясений. Возможно, этот механизм тесно связан с неравномерностью вращения Земли;

б) Механизм (M -компоненты), вызывающий переменное напряжение на стыках тектонических плит (особенно заметное на перекрестьях), влияет только на слабые землетрясения. Возможно, этот механизм связан со своеобразными колебаниями в движении тектонических плит.

Подтвердить или опровергнуть предположение о колебательном характере тектонических движений с периодом 2–3 года, по-видимому, можно будет только пользуясь данными измерений с помощью техники космической геодезии [15].

4. Исследована зависимость сейсмической активности от двух параметров: широты и глубины. В частности, особенностью Северного полушария является неизменность локализации максимумов плотности землетрясений с различными магнитудами на широте 40° с. ш. В то же время в Южном полушарии с ростом глубины происходит смещение максимума

плотности землетрясений: от экватора – для гипоцентров с глубиной менее 30 км, до 20° ю. ш. – для гипоцентров с глубиной свыше 300 км.

Таким образом, очевидна сильная зависимость сейсмической активности Земли от широты, причем эта зависимость одинакова для землетрясений со всеми магнитудами. Это указывает на связь сейсмической активности с вращением Земли.

Интересно, что для землетрясений с гипоцентрами глубже 100 км сейсмическая активность Южного полушария превосходит таковую для Северного и для глубоких землетрясений она полностью доминирует.

Работа выполнена при поддержке РФФИ (грант № 98–05–65005).

Приложение. Нахождение коэффициента корреляции:

Для вычисления коэффициента корреляции использовалась формула для линейной корреляции (корреляция Пирсона):

$$Q = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N \cdot \sigma_x \cdot \sigma_y}.$$

Здесь используются обозначения:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \text{ – среднее значение } x, \sigma_x = \sqrt{D_x} \text{ – стандартное отклонение,}$$

$$D_x = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \text{ – дисперсия, аналогично для } y\text{-компоненты.}$$

Ошибка определения коэффициента корреляции:

$$\Delta Q = \frac{1 - Q^2}{\sqrt{N}}.$$

Литература

- 1 Жарков В. И. Внутреннее строение Земли и планет. М : Наука, 1978. 192 с
- 2 AT (1928–1990) US Geological Survey, National Earthquake Information Center.
- 3 Горькавый Н. Н., Левицкий Л. С., Тайдакова Т. А. и др. О выявлении трех компонент и сейсмической активности Земли // Физика Земли 1994, № 10. С. 23–32.
- 4 Morrison L. V. ВИН Annual Reports. 1984 P. B15–B16
- 5 IERS Annual Report Observatoire de Paris, 1992. P. 127–128.
- 6 Горькавый Н. Н., Минин В. А., Тайдакова Т. А., Фридман А. М. Существуют ли астрономические причины сильнейших землетрясений // Астрономический циркуляр. 1989. № 1540 С 35–36.
- 7 Munk W H, Macdonald G T. H. The rotating of the Earth. Cambridge: The Uneversity Press, 1960.
- 8 Moritz H., Muller I I Earth rotating. Theory and observation. Vugar Publ Comp New York. 1987
- 9 Kanamari H // Proc. Int School of Phys «Enrico Fermi». Course 85, 1983 P 596.
- 10 Горькавый Н. Н., Левицкий Л. С., Тайдакова Т. А. и др. О корреляции графиков угловой скорости вращения Земли и модуля временной производной с частотой землетрясений и зависимости от их магнитуды // Физика Земли 1994. № 10 С 33–38

- 11 Горькавый Н Н., Левицкий Л. С., Тайдакова Т. А и др О зависимости корреляции между региональной сейсмичностью Земли и неравномерностью ее вращения от глубины очагов землетрясений // Физика Земли. 1999. № 10. С 52–66
12. Горькавый Н Н., Трапезников Ю А., Фридман А. М О глобальной составляющей сейсмического процесса и ее связи с наблюдательными особенностями вращения Земли // Докл. РАН 1994. Т. 338, № 4 С. 525–527.
- 13 Горькавый Н Н., Левицкий Л С., Тайдакова Т А и др Об отрицательной корреляции сейсмической активности между зонами субдукции Тихоокеанской и Индо-Австралийской плит и на перекрестье срединно-океанических хребтов // Физика Земли 1999 № 11 С. 28–39
- 14 Левин Б В, Чирков Е Б Особенности широтного распределения сейсмичности и вращение Земли // Вулканология и сейсмология 1999 № 6 С. 65–69.
- 15 Кротошкин П Н, Трапезников Ю А. Вариации угловой скорости вращения Земли, колебаний полюса и скорости дрейфа геомагнитного поля и их возможная связь с геотектоническими процессами // Изв АН СССР Сер. геол. 1963. № 11. С. 32–50
- 16 Шейдеггер А Основы геодинамики М Недра, 1987.
17. Татевян С К. Роль спутниковых локационных измерений в изучении современной геодинамики // Исследование Земли из космоса 1999 № 1 С 87–96

Е. А. Кузнецов

1. Введение

В начале статьи предложим новое описание вихревой динамики идеальной жидкости как движение сжимаемой заряженной жидкости, двигающейся в эффективных электрическом и магнитном полях. Отсюда следует представление вихревых линий, построенное в работе [13]. Покажем, что коллапс завихренности в гидродинамике следует рассматривать как процесс опрокидывания вихревых линий.

Коллапс, как процесс образования особенности за конечное время из начально гладкого распределения, играет принципиально важную роль в физике, представляя собой один из наиболее эффективных механизмов диссипации энергии. В частности, поэтому коллапс различных плазменных волн, например коллапс ленгмюровских волн [1], является определяющим фактором при нагреве плазмы высокочастотным полем или электронными пучками. В гидродинамике коллапс играет также важную роль, например для турбулентности морского волнения. Оценки времен диссипации поверхностной гравитационной волны за счет вязкости дают времена, составляющие десять часов и более, хотя, как всем хорошо известно, после шторма море успокаивается за десятки минут, но не часов. Объяснение наблюдаемого времени затухания турбулентности морского волнения связано с образованием барашков на гребне волн, сопровождающееся брызгами, что и обеспечивает эффективную диссипацию морского волнения. Формирование особенностей в виде заострений поверхности жидкости представляет собой процесс коллапса для поверхностных волн.

Однако, не только этим интересен коллапс. Он весьма важен с точки зрения теории колмогоровских спектров турбулентности. Хорошо известно, что всякая особенность в физическом пространстве волновых векторов \mathbf{k} порождает степенное поведение амплитуд фурье-гармоник в области больших волновых чисел. Если особенности распределены по пространству и по времени случайным образом с некоторой мерой, то от особенностей в спектре турбулентности возникнут соответствующие степенные хвосты. Впервые эта идея была использована Филлипсом в 1958 году [2] для нахождения спектра турбулентности поверхностных гравитационных волн. Пусть поверхность жидкости $z = \eta(x, t)$ (z — вертикальная координата) испытывает излом в точке $x = x_0$, т. е. функция η , будучи непрерывной, имеет в этой

точке скачок производных, соответственно ее вторая производная в этой точке оказывается пропорциональной $\delta(x - x_0)$. Отсюда фурье-образ $\eta(k)$ этой функции приобретает при больших k степенной хвост, пропорциональный $1/k^2$. Для спектра турбулентности $\varepsilon(\omega)$ это дает степенное поведение — спектр Филлипса [2]:

$$\varepsilon(\omega) \sim \omega^{-5}. \quad (1)$$

Здесь $\varepsilon(\omega)$ нормирован так, что $\int \varepsilon(\omega) d\omega$ есть плотность энергии волн (на единицу площади), ω и k связаны между собой дисперсионным соотношением $\omega = \sqrt{gk}$, g — ускорение поля тяжести.

Интересно сравнить спектр Филлипса (1) с колмогоровским спектром Захарова — Филоненко [5]:

$$\varepsilon(\omega) \sim \omega^{-4}, \quad (2)$$

который может быть получен из решения кинетических уравнений для волн. В отличие от (1), спектр Захарова — Филоненко соответствует каскадному процессу передачи энергии из больших масштабов в малые в предположении малости углов наклона поверхности от горизонта. Последнее означает, что дисперсионные (волновые) эффекты являются доминирующими по сравнению с нелинейными, в то время как при формировании спектра (1) ситуация противоположна. По этой причине спектр Филлипса (1) называют сильнотурбулентным.

Другой пример — акустическая турбулентность. Согласно Б. Б. Кадомцеву и В. И. Петвиашвили [5] акустическая турбулентность — это набор ударных волн. На каждой ударной волне плотность ρ испытывает скачок. По этой причине производная в этих точках есть δ -функция, а фурье-амплитуда от скачка плотности даст зависимость $\sim 1/k$. Для спектра турбулентности отсюда соответственно имеем [5]

$$\varepsilon_\omega \sim \omega^{-2}, \quad (3)$$

где $\omega = kc_s$, c_s — скорость звука.

Спектр слабой звуковой турбулентности — спектр Захарова — Сагдеева [6],

$$\varepsilon_\omega \sim \omega^{-3/2}, \quad (4)$$

также отличается от сильнотурбулентного спектра (3). Различие в спектрах как для акустики, так и для турбулентности поверхностных волн связано с наличием линейного закона дисперсии волн. Спектры (1) и (3) описывают сильнонелинейные режимы, в то время как спектры (2) и (4) соответствуют пределу слабой нелинейности.

Обратимся теперь к развитой гидродинамической турбулентности при больших числах Рейнольдса, $Re \gg 1$. В этом случае в инерционном интервале масштабов — промежуточной области между энерго-содержащей и вязкой областями — спектр турбулентности, согласно А. Н. Колмогорову [7] и А. М. Обухова [8], определяется единственной величиной P — постоянным потоком энергии по масштабам (совпадающим с темпом диссипации энергии в единице объема):

$$\varepsilon_k = CP^{2/3}k^{-5/3}. \quad (5)$$

В этом случае для флуктуации скорости v с заданным масштабом r из инерционного интервала колмогоровскому спектру (5) соответствует распределение

$$v \sim P^{1/3}r^{1/3}, \quad (6)$$

которое может быть получено из соображений размерности (см., например, [9]). Из этих же соображений для завихренности скорости, $\Omega = \text{rot } \mathbf{v}$, мы можем записать

$$\Omega \sim P^{1/2}r^{-2/3}, \quad (7)$$

так что распределение завихренности имеет особенность при $r \rightarrow 0$.

Что это за особенность и каковы физические причины ее возникновения? На эти вопросы возможны несколько вариантов ответов. Один из них — особенность (7) может быть «отголоском» реальных коллапсов, в результате которых в поле завихренности за конечное время происходит формирование особенностей. Если реализуется такой сценарий, то спектр (7) следует трактовать как аналог спектра Филлипса.

В данной работе мы приведем ряд аргументов в пользу того, что коллапс завихренности — образование особенности завихренности за конечное время — может объяснять сингулярное распределение (7). Согласно работам [10, 11, 12] коллапс в гидродинамике может рассматриваться как процесс опрокидывания вихревых линий, т. е. как процесс, аналогичный опрокидыванию в газовой динамике. В отличие от опрокидывания в газодинамике, когда одна траектория лагранжевой частицы пересекается с траекторией другой лагранжевой частицы, опрокидывание вихревых линий означает, что одна вихревая линия догоняет другую вихревую линию. Для гладкого непрерывного распределения завихренности опрокидывание в первый момент наступает при касании вихревых линий. При этом в точке касания завихренность обращается в бесконечность. И это возможно несмотря

на несжимаемость векторных полей завихренности и скорости жидкости. Существование такого типа коллапса была продемонстрирована впервые в [10] для трехмерной интегрируемой гидродинамики с гамильтонианом $H = \int |\Omega| dr$, введенной в работе Рубана и автора [13]. Такая модель может быть получена из гидродинамики Эйлера в приближении локальной индукции.

В основе описания опрокидывания вихревых линий лежит так называемое представление вихревых линий, предложенное в работах [1–14]. Это представление является смешанным — лагранжево-эйлеровым — описанием, когда каждая вихревая линия нумеруется двумерным лагранжевым маркером, а другой параметр задает саму вихревую линию. Иными словами, представление вихревых линий есть не что иное как переход в движущуюся криволинейную систему координат, связанную с вихревыми линиями.

План статьи следующий. Во втором разделе приведена необходимая информация о представлении вихревых линий (ПВЛ). В частности, показано, что при локальном описании течений переменные Клебша могут быть использованы в качестве лагранжевых маркеров вихревых линий. В ситуации общего положения переход к представлению вихревых линий эквивалентен рассмотрению новой сжимаемой гидродинамики заряженной жидкости, двигающейся под действием эффективных электрического и магнитного полей. Здесь же рассмотрен вопрос о применении ПВЛ к течениям вязкой жидкости. Получено уравнение диффузионного типа, описывающее динамику инварианта Коши в вязком случае с «тензором диффузии», определяемым метрикой ПВЛ. По своей форме это уравнение совпало с уравнением, выведенным в [15]. При этом уравнения движения вихревых линий — в своей прежней форме — понимаются как уравнения, задающие переход в криволинейную систему координат. Полученные точные уравнения для описания вязких течений можно рассматривать как результат точного разделения двух временных масштабов — инерционного (нелинейного по сути) и вязкого.

В следующем разделе рассматриваются общие свойства отображений вблизи точки опрокидывания. Показано, в частности, что для газодинамики пыли поведение завихренности вблизи точки коллапса такое же, как и при опрокидывании вихревых линий в трехмерной интегрируемой гидродинамике. Предполагая, что коллапс в гидродинамике представляет собой опрокидывание вихревых линий, в четвертом разделе исследуется структура особенности завихренности. Показано, что в момент опрокидывания особенность завихренности сильно анизотропна, а ее главная зависимость совпадает с *колмогоровской*

зависимостью (7). В последнем разделе обсуждаются полученные результаты, приведено их сравнение с существующими численными работами, а также со строгими математическими результатами.

2. Основные уравнения

Как известно (см., например, обзоры [17, 18]), уравнения Эйлера для несжимаемой жидкости

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \nabla) \mathbf{v} = -\nabla p, \quad \operatorname{div} \mathbf{v} = 0 \quad (8)$$

как для двумерных, так и для трехмерных течений обладают бесконечным числом интегралов движения. Это так называемые лагранжевы инварианты Коши. Наиболее просто выражение для инварианта Коши можно получить исходя из теоремы Кельвина о сохранении циркуляции скорости:

$$\Gamma = \oint (\mathbf{v} \cdot d\mathbf{l}), \quad (9)$$

где контур интегрирования $C[\mathbf{r}(t)]$ движется вместе с жидкостью. Если в этом выражении перейти от эйлеровых координат \mathbf{r} к лагранжевым \mathbf{a} , то (9) переписется в виде

$$\Gamma = \oint \dot{x}_i \cdot \frac{\partial x_i}{\partial a_k} da_k,$$

где контур $C[\mathbf{a}]$ уже будет неподвижным.

В силу произвольности контура $C[\mathbf{a}]$ и благодаря формуле Стокса, отсюда немедленно следует, что величина

$$\mathbf{I} = \operatorname{rot}_{\mathbf{a}} \left(\dot{x}_i \frac{\partial x_i}{\partial \mathbf{a}} \right) \quad (10)$$

сохраняется в каждой точке \mathbf{a} . Это и есть лагранжев инвариант Коши.

Сохранение этих инвариантов, как это впервые было показано Салмоном [17], обязано специальной симметрии — бесконечной симметрии относительно переобозначения лагранжевых маркеров (relabeling symmetry), оставляющей инвариантным действие. Если лагранжевы координаты \mathbf{a} в (10) совпадают с начальными положениями жидких частиц, то инвариант \mathbf{I} совпадает с начальной завихренностью $\Omega_0(\mathbf{a})$. Можно показать, что инварианты Коши характеризуют замороженность вихревых линий в жидкости. Это очень важное свойство, согласно которому жидкие (лагранжевы) частицы не могут покинуть

собственную вихревую линию, на которой они находились в начальный момент времени. Для лагранжевых частиц, таким образом, имеется одна незамороженная степень свободы — движение вдоль вихревой линии. Однако такое движение, как это видно из уравнений движения для завихренности

$$\frac{\partial \Omega}{\partial t} = \text{rot} [\mathbf{v} \times \Omega], \quad (11)$$

не изменяет ее значения. С этой точки зрения вихревая линия является инвариантным объектом и естественно поэтому перейти к такому описанию, где эта инвариантность видна с самого начала. Такое описание — представление вихревых линий — было предложено в работах Рубана и автора [13, 14].

2.1. Локальное представление

Рассмотрим вихревую трубку с непрерывно распределенной завихренностью. Будем каждую вихревую линию нумеровать двумерным маркером ξ и введем параметр s , задающий данную вихревую линию. И пусть функция

$$\mathbf{r} = \mathbf{R}(\xi, s) \quad (12)$$

определяет соответствующий переход в криволинейную, связанную с вихревыми линиями систему координат. В этом случае завихренность в точке \mathbf{r} будет задаваться интегралом [13]:

$$\Omega(\mathbf{r}) = \int d^2 \xi \int \frac{\partial \mathbf{R}}{\partial s} \delta(\mathbf{r} - \mathbf{R}(\xi, s)) ds. \quad (13)$$

Здесь вклад в завихренность от вихревой линии с маркером ξ дается интегралом по s , в котором производная $\partial \mathbf{R} / \partial s$ есть касательный вектор к вихревой линии. Суммирование (интегрирование по ξ) дает значение Ω в точке \mathbf{r} . Наличие δ -функции в (13) позволяет полностью снять интегрирование, так что

$$\Omega(\mathbf{r}, t) = \frac{1}{J} \cdot \frac{\partial \mathbf{R}}{\partial s}, \quad (14)$$

где

$$J = \frac{\partial(x, y, z)}{\partial(\xi_2, \xi_2, s)} \quad (15)$$

— якобиан преобразования (12).

Формула (14) может быть получена исходя из представления завихренности через переменные Клебша λ и μ :

$$\mathbf{\Omega} = [\nabla\lambda \times \nabla\mu]. \quad (16)$$

Геометрический смысл этих переменных хорошо известен: пересечение поверхностей $\lambda = \text{const}$ и $\mu = \text{const}$ задает вихревую линию. Известно также, что в несжимаемом случае переменные Клебша являются лагранжевыми инвариантами, неизменными вдоль траектории жидких частиц:

$$\frac{\partial\lambda}{\partial t} + (\mathbf{v}\nabla)\lambda = 0; \quad \frac{\partial\mu}{\partial t} + (\mathbf{v}\nabla)\mu = 0.$$

И тогда эти переменные можно взять в качестве маркеров вихревых линий. Легко устанавливается, что переход в (16) к новым переменным

$$\lambda = \lambda(x, y, z), \quad \mu = \mu(x, y, z), \quad s = s(x, y, z),$$

где s — параметр, задающий данную вихревую линию, приводит к выражению (14) с якобианом

$$J = \frac{\partial(x, y, z)}{\partial(\lambda, \mu, s)}.$$

Таким образом, представления (14) и (16) эквивалентны друг другу.

Согласно теореме Дарбу переменные Клебша могут быть введены локально всегда, но не глобально. Если переменные Клебша λ и μ могут быть введены всюду, так что скорость

$$\mathbf{v} = \lambda\nabla\mu + \nabla\phi, \quad \text{div } \mathbf{v} = \Delta\phi + \text{div } \lambda\nabla\mu = 0,$$

то топологический инвариант — инвариант спиральности $\int (\mathbf{v} \cdot \text{rot } \mathbf{v}) d\mathbf{r}$, характеризующий степень зацепления вихревых линий, — будет тождественно равен нулю. Поэтому, чтобы ввести представление вихревых линий для течений с нетривиальной топологией, необходимо обратиться непосредственно к изначальным уравнениям движения — уравнениям (8), (11) для скорости и ее завихренности.

2.2. ПВЛ — ситуация общего положения

Согласно уравнению (11) тангенциальная к вектору $\mathbf{\Omega}$ компонента скорости \mathbf{v}_τ не оказывает влияния (прямого) на изменение завихренности, т. е. в (11) вместо скорости \mathbf{v} нужно поставить ее поперечную компоненту \mathbf{v}_n .

Уравнение движения для поперечной скорости \mathbf{v}_n следует непосредственно из уравнения (8). Оно имеет вид уравнения движения частицы в электромагнитном поле:

$$\frac{\partial \mathbf{v}_n}{\partial t} + (\mathbf{v}_n \nabla) \mathbf{v}_n = \mathbf{E} + [\mathbf{v}_n \times \mathbf{H}], \quad (17)$$

где эффективные поля — электрическое и магнитное — даются выражениями:

$$\mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial \mathbf{v}_\tau}{\partial t}, \quad (18)$$

$$\mathbf{H} = \text{rot } \mathbf{A}, \quad (19)$$

где скалярный и векторный потенциалы (φ и \mathbf{A} соответственно) выражаются посредством формул

$$\varphi = p + \frac{\mathbf{v}_\tau^2}{2}, \quad \mathbf{A} = \mathbf{v}_\tau, \quad (20)$$

так что два уравнения Максвелла

$$\text{div } \mathbf{H} = 0, \quad \frac{\partial \mathbf{H}}{\partial t} = -\text{rot } \mathbf{E}$$

автоматически удовлетворяются. При этом на векторный потенциал \mathbf{A} наложена калибровка

$$\text{div } \mathbf{A} = -\text{div } \mathbf{v}_n,$$

которая эквивалентна условию $\text{div } \mathbf{v} = 0$.

Два других уравнения Максвелла могут быть также записаны, но не несут никакой дополнительной нагрузки, поскольку плотность заряда ρ и ток \mathbf{j} формально определяются из соотношений (18), (19). Базовым здесь является само уравнение движения (17) для нормальной компоненты скорости, которое представляет собой уравнение движения для нерелятивистской частицы с зарядом и массой, равными единице, скорость света при этом также равна единице.

Уравнение движения (17) записано в эйлеровом представлении. Чтобы перейти к его лагранжевой формулировке, нужно рассмотреть уравнения для «траекторий», которые определяются скоростью \mathbf{v}_n :

$$\frac{d\mathbf{R}}{dt} = \mathbf{v}_n(\mathbf{R}, t) \quad (21)$$

с начальными условиями

$$\mathbf{R}|_{t=0} = \mathbf{a}.$$

Решение уравнения (21) задает отображение

$$\mathbf{r} = \mathbf{R}(\mathbf{a}, t), \quad (22)$$

определяющее переход от эйлерового описания к новому лагранжево-му описанию.

Уравнения движения в новых переменных суть уравнения Гамильтона:

$$\dot{\mathbf{P}} = -\frac{\partial h}{\partial \mathbf{R}}, \quad \dot{\mathbf{R}} = \frac{\partial h}{\partial \mathbf{P}}, \quad (23)$$

где точка означает дифференцирование по времени при фиксированном значении \mathbf{a} , $\mathbf{P} = \mathbf{v}_n + \mathbf{A} \equiv \mathbf{v}$ есть обобщенный импульс, а гамильтониан частицы h , являясь функцией импульса \mathbf{P} и координаты \mathbf{R} , дается стандартным выражением:

$$h = \frac{1}{2}(\mathbf{P} - \mathbf{A})^2 + \varphi \equiv p + \frac{\mathbf{v}^2}{2}.$$

Первое уравнение системы (23) — это уравнение движения (17), записанное в переменных \mathbf{a} , t , а второе уравнение совпадает с (21). Для «новой» гидродинамики (17) или ее гамильтоновской формулировки (23) будет справедлива теорема Кельвина (она же теорема Лиувилля):

$$\Gamma = \oint (\mathbf{P} \cdot d\mathbf{R}), \quad (24)$$

где интегрирование ведется по замкнутому, двигающемуся вместе с «жидкостью» контуру. Отсюда точно так же, как это было сделано выше при выводе (10), следует выражение для «нового» инварианта Коши:

$$\mathbf{I} = \text{rot}_{\mathbf{a}} \left(P_i \frac{\partial x_i}{\partial \mathbf{a}} \right). \quad (25)$$

Его отличие от оригинального инварианта Коши (10) состоит в том, что в уравнении движения (21) вместо скорости \mathbf{v} стоит ее нормальная компонента \mathbf{v}_n . Как следствие, «новая» гидродинамика является сжимаемой: $\text{div } \mathbf{v}_n \neq 0$. Поэтому на якобиан J преобразования (22) не накладывается никаких ограничений.

Из формулы (25) легко может быть получено выражение для завихренности $\boldsymbol{\Omega}$ (ср. с [13, 14]):

$$\boldsymbol{\Omega}(\mathbf{r}, t) = \frac{(\boldsymbol{\Omega}(\mathbf{a}) \cdot \nabla_{\mathbf{a}}) \mathbf{R}(\mathbf{a}, t)}{J}, \quad (26)$$

где J — якобиан преобразования (22), равный

$$J = \frac{\partial(x_1, x_2, x_3)}{\partial(a_1, a_2, a_3)}.$$

Здесь мы учли, что обобщенный импульс \mathbf{P} совпадает со скоростью \mathbf{v} , включая момент времени $t = 0$: $\mathbf{P}_0(\mathbf{a}) \equiv \mathbf{v}_0(\mathbf{a})$. В этом соотношении $\Omega_0(\mathbf{a})$ есть новый инвариант Коши (совпадающий с начальной завихренностью), имеющий нулевую дивергенцию: $\operatorname{div}_a \Omega_0(a) = 0$.

Представление (26) обобщает соотношение (16) на произвольную топологию вихревых линий. Переменные \mathbf{a} в этом выражении локально можно понимать как набор ξ и s .

Таким образом, в случае общего положения уравнение движения вихревых линий имеет вид (21), которое дополняется соотношением (26) и уравнением

$$\Omega(\mathbf{r}, t) = \operatorname{rot}_r \mathbf{v}(\mathbf{r}, t) \quad (27)$$

с условием $\operatorname{div}_r \mathbf{v}(\mathbf{r}, t) = 0$.

Уравнения движения (21), (27) вместе с соотношением (26) можно рассматривать как результат частичного интегрирования уравнения Эйлера (8). Эти уравнения разрешены относительно инвариантов Коши — бесконечного числа интегралов движения, что принципиально важно при численном интегрировании (см. [11, 12]). Для этой системы инварианты Коши сохраняются автоматически, в то время как при прямом интегрировании уравнений Эйлера нужно следить, в какой степени инварианты Коши являются сохраняющимися величинами. По-видимому, этот факт есть одно из главных ограничений, определяющих точность дискретных численных схем при прямом интегрировании уравнений Эйлера.

Другим важным свойством представления вихревых линий является отсутствие каких-либо ограничений на якобиан J , какие, например, имеют место при переходе от эйлерова описания к лагранжевому, когда якобиан равен единице. При этом $1/J$ имеет смысл плотности n вихревых линий. Эта величина в силу уравнения (21) как функция \mathbf{r} и t подчиняется уравнению непрерывности:

$$\frac{\partial n}{\partial t} + \operatorname{div}_r (n \mathbf{v}_n) = 0. \quad (28)$$

В этом уравнении $\operatorname{div}_r \mathbf{v}_n \neq 0$, поскольку только полная скорость \mathbf{v} имеет нулевую дивергенцию.

В заключение этого раздела рассмотрим вопрос о применении ПВЛ к течениям вязкой жидкости.

Запишем уравнение Навье-Стокса для завихренности Ω :

$$\frac{\partial \Omega}{\partial t} = \text{rot}[\mathbf{v} \times \Omega] - \nu \text{rot rot } \Omega. \quad (29)$$

В уравнении (29) совершим переход к переменным \mathbf{a} и t посредством замен, определяемых уравнением (21), дополняемым соотношением Коши (26):

$$\Omega(\mathbf{r}, t) = \frac{(\Omega_0 \cdot \nabla_a) \mathbf{R}(\mathbf{a}, t)}{J}, \quad (30)$$

в котором Ω_0 есть функция не только координат \mathbf{a} , но и времени t : $\Omega_0 = \Omega_0(\mathbf{a}, t)$.

При подстановке (30) в (29) первое слагаемое в правой части сократится в силу (21). В результате уравнение (29) запишется в виде

$$\frac{1}{J} \left(\frac{\partial \Omega_0}{\partial t} \cdot \nabla_a \right) \mathbf{R} = -\nu \text{rot rot} \left\{ \frac{1}{J} (\Omega_0 \cdot \nabla_a) \mathbf{R} \right\}. \quad (31)$$

На следующем этапе необходимо дифференцирование по \mathbf{r} в правой части (31) заменить на дифференцирование по переменным \mathbf{a} . В результате простых, но достаточно громоздких вычислений уравнение (31) превращается в уравнение для $\Omega_0(\mathbf{a}, t)$:

$$\frac{\partial \Omega_0}{\partial t} = -\nu \text{rot}_a \left(\hat{g} \text{rot}_a \left(\frac{\hat{g}}{J} \Omega_0 \right) \right). \quad (32)$$

Это уравнение линейно относительно Ω_0 , в нем \hat{g} — метрический тензор, равный

$$g_{\alpha\beta} = \frac{\partial x_i}{\partial a_\alpha} \cdot \frac{\partial x_i}{\partial a_\beta}.$$

Уравнение (32) для инварианта Коши по своей форме совпадает с уравнением, полученным Зеньковичем и Якубовичем [15] для несжимаемой гидродинамики, когда переменные \mathbf{a} имеют смысл лагранжеских маркеров жидких частиц. В уравнении Зеньковича — Якубовича якобиан J предполагается независимым от времени, в простейшем варианте он равен единице. В этом и состоит принципиальное отличие уравнения Зеньковича — Якубовича от уравнения (32).

Примечательной особенностью полученной системы является *точное разделение* двух различных временных масштабов, ответственных за инерционные (по сути — нелинейные) и вязкие процессы. Первые описываются уравнением (21), а вязкие процессы — уравнением

диффузионного типа (32), в котором «коэффициент» диффузии, пропорциональный коэффициенту вязкости ν , имеет сложную тензорную структуру, определяемую метрикой отображения $\mathbf{r} = \mathbf{R}(\mathbf{a}, t)$.

3. Опрокидывание вихревых линий

Якобиан J , таким образом, может принимать произвольные значения, в частности, обращаться в нуль в некоторой точке. Согласно (25) это означает появление у завихренности особенности. Для гидродинамики Эйлера строгое доказательство этого факта пока отсутствует. Строго доказана возможность появления такого типа особенности за конечное время (т. е. коллапса) за счет обращения якобиана в нуль для трехмерной интегрируемой гидродинамики [10], когда в уравнении Эйлера для завихренности (11) связь между скоростью \mathbf{v} и завихренностью Ω (обе — обобщенные) определяется соотношением

$$\mathbf{v} = \text{rot} \frac{\delta \mathcal{H}}{\delta \Omega}, \quad \mathcal{H} = \int |\Omega| dr. \quad (33)$$

Гамильтониан \mathcal{H} такого вида был впервые предложен И. М. Халатниковым [18] для сверхтекучести, однако модифицированное таким образом уравнение Эйлера также имеет смысл для идеальной гидродинамики. Как было показано в [13, 14], модель (33) может быть получена из уравнений Эйлера в приближении локальной индукции.

Появление особенности в точке, где впервые якобиан обращается в нуль ($J = 0$) означает, что одна вихревая линия касается в этой точке другой вихревой линии. Этот процесс в [10] был назван опрокидыванием вихревых линий — он аналогичен процессу опрокидывания в газодинамике, когда образование сингулярности в плотности и градиенте скорости возникает в случае касания лагранжевых частиц.

В работе Рубана и автора [10] было показано, что появление особенности в трехмерной интегрируемой гидродинамике обязано отсутствию в этой модели всякого взаимодействия между вихревыми линиями. Это есть гидродинамика свободных, непрерывно распределенных по пространству вихревых линий, имеющих свою внутреннюю нелинейную динамику. В этом состоит основная причина опрокидывания вихревых линий для модели (33). Важно, что этот процесс оказывается возможным несмотря на бездивергентность как самого поля скорости, так и поля завихренности. Пожалуй, это есть один из основных результатов работы [10]. Следует отметить также, что для трехмерной

интегрируемой гидродинамики, в ситуации общего положения, завихренность вблизи точки опрокидывания ведет себя взрывным образом: $\Omega \sim (t_0 - t)^{-1}$, где t_0 — время коллапса. Сингулярность возникает за счет обращения якобиана J в нуль.

3.1. Газовая динамика пыли

Покажем, что таким же образом, как и для трехмерной интегрируемой гидродинамики, ведет себя завихренность вблизи точки опрокидывания для газодинамики пыли:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \rho \mathbf{v} = 0, \quad \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \nabla) \mathbf{v} = 0, \quad (34)$$

где ρ — плотность пыли, а \mathbf{v} — ее скорость.

Для этого удобно ввести матрицу U с матричными элементами

$$U_{ij} = \frac{\partial v_j}{\partial x_i}.$$

Уравнение движения для U следует непосредственно из второго уравнения (34) после его дифференцирования по координатам:

$$\frac{dU}{dt} = -U^2 \left[\frac{d}{dt} \equiv \frac{\partial}{\partial t} + (\mathbf{v} \nabla) \right]. \quad (35)$$

Отсюда уравнение просто интегрируется:

$$U(\mathbf{r}, t) = U_0(\mathbf{a}) [1 + U_0(\mathbf{a})t]^{-1}, \quad (36)$$

где $U_0(\mathbf{a})$ — начальное значение матрицы U , зависящее от начальной координаты \mathbf{a} лагранжевых частиц, связанных с текущей координатой \mathbf{r} соотношением

$$\mathbf{r} = \mathbf{a} + \mathbf{v}_0(\mathbf{a})t. \quad (37)$$

Предположим теперь, что все собственные значения матрицы $U_0(\mathbf{a})$ чисто действительные, равные $\lambda_1, \lambda_2, \lambda_3$. Такое предположение означает, что симметричная часть матрицы U , так называемый тензор напряжений, $S = \frac{1}{2}(U + U^T)$ превышает ее антисимметричную часть $\hat{\Omega} = \frac{1}{2}(U - U^T)$ — тензор завихренности (T означает здесь транспонирование). Это условие выполняется, если

$$2(\operatorname{tr} S)^2 - \frac{2}{3}\operatorname{tr}(S^2) \geq \Omega^2.$$

Тогда матрица U (36) может быть разложена по проекторам матрицы $U_0(a)$, соответствующим каждому собственному значению:

$$U = \sum_m \frac{\lambda_m(a)}{1 + \lambda_m(a)t} \cdot P^{(m)}(a). \quad (38)$$

Здесь $P^{(m)}$ — одномерные проекторы ($P^2 = P$), которые строятся стандартным образом через собственные векторы прямой ($U_0\psi = \lambda\psi$) и сопряженной ($\varphi U_0 = \lambda\varphi$) спектральных задач для матрицы U_0 :

$$P_{\alpha\beta}^{(n)}(a) = \varphi_\alpha^{(n)}(a) \cdot \psi_\beta^{(n)}(a),$$

при этом векторы $\varphi^{(n)}$ и $\psi^{(m)}$ взаимно ортогональны:

$$\psi_\alpha^{(n)} \cdot \varphi_\alpha^{(m)} = \delta_{nm}.$$

Согласно (38) определитель матрицы U и якобиан отображения (37) будут представлять собой произведения:

$$\det U = \prod_k \frac{\lambda_k}{1 - \lambda_k t}, \quad J = \prod_k (1 - \lambda_k t).$$

Из (38) видно, что особенности в U возникают за счет обращения знаменателей $1 - \lambda_k t$ в нуль. Впервые особенность у U появляется при $t = t_0$, определяемом из условия (ср. с [19])

$$t_0 = \min_{m,a} [-1/\lambda_m(\mathbf{a})] > 0. \quad (39)$$

Отсюда следует, что вблизи сингулярной точки в сумме (38) выживает только один член:

$$U \approx - \frac{P^{(n)}}{\tau + \gamma_{\alpha\beta} \Delta a_\alpha \Delta a_\beta}. \quad (40)$$

Здесь проектор $P^{(n)}$ берется в точке $\mathbf{a} = \mathbf{a}_0$ при $m = n$, соответствующих минимуму (39), $\tau = t_0 - t$, $\Delta \mathbf{a} = \mathbf{a} - \mathbf{a}_0$, а

$$2\gamma_{\alpha\beta} = - \left. \frac{\partial^2 \lambda_{0n}^{-1}}{\partial a_\alpha \partial a_\beta} \right|_{\mathbf{a}=\mathbf{a}_0, t=t_0}$$

является положительно определенной матрицей.

Замечательная формула (40) демонстрирует, во-первых, что матрица U стремится к вырожденной при $t \rightarrow t_0$, во-вторых, что обе части

матрицы U — ее симметричная часть S и антисимметричная часть — тензор завихренности Ω — одновременно становятся бесконечными при $t = t_0$ в точке $\mathbf{a} = \mathbf{a}_0$. Интересно также отметить, что вблизи сингулярной точки отношение между этими частями фиксировано и контролируется двумя соотношениями, следующими из определения проектора P :

$$P_S = P_S^2 + P_A^2, \quad P_A = P_S P_A + P_A P_S,$$

где P_S и P_A есть соответственно симметричная («потенциальная») и антисимметричная (вихревая) части проектора P . В частности, второе соотношение обеспечивает коллапсирующее решение уравнения Эйлера для завихренности

$$\frac{\partial \Omega}{\partial t} = \text{rot} [\mathbf{v} \times \Omega],$$

которое по форме совпадает с (11).

Таким образом, завихренность Ω , как и для трехмерной интегрируемой гидродинамики (ср. с [10]), имеет в точке коллапса $\mathbf{a} = \mathbf{a}_0$ особенность типа $1/\tau$.

3.2. Структура сингулярности

Обратимся вновь к несжимаемой идеальной гидродинамике, описываемой посредством уравнений Эйлера (8). Будем предполагать, что коллапс происходит благодаря опрокидыванию вихревых линий. Обозначим через $\tilde{\tau}(a) > 0$ решение уравнения

$$J(a, t) = 0,$$

и пусть

$$t_0 = \min_a \tilde{t}(a),$$

где минимум достигается в точке $\mathbf{a} = \mathbf{a}_0$. Около точки минимума (t_0, \mathbf{a}_0) якобиан может быть разложен в ряд (ср. с [10]):

$$J = \alpha(t_0 - t) + \gamma_{ij} \Delta a_i \Delta a_j + \dots, \quad (41)$$

где $\alpha > 0$, γ_{ij} — положительно определенная матрица.

Разложение (41) имеет простой геометрический смысл. Поверхность $J = J(\mathbf{a}, t) > 0$ при своей деформации в момент $t = t_0$ впервые касается плоскости $J = 0$. В ситуации общего положения такое касание имеет место в отдельной точке и оно квадратично.

Разложение (41) предполагает, что одно из собственных значений матрицы Якоби $\hat{J}_{\alpha\beta} = \partial x_\beta / \partial a_\alpha$, скажем, λ_3 , обращается в нуль согласно (41), а два других собственных значения $\lambda_{1,2}$ (действительны они или комплексны) стремятся при $t \rightarrow t_0$ в точке коллапса $\mathbf{a} = \mathbf{a}_0$ к постоянным значениям. В этом смысле данная ситуация полностью подобна рассмотренной выше для газовой динамики пыли. Если в разложении (41) по какой-либо причине коэффициенты α или γ_{ij} будут равны нулю, то такое разложение уже не будет соответствовать ситуации общего положения.

Для режима (41) при $t \rightarrow t_0$

$$\Omega = \frac{\mathbf{R}_s(\mathbf{a}_0 t_0)}{\tau(\alpha + \gamma_{ij}\eta_i\eta_j)}, \quad (42)$$

где $\eta = \Delta \mathbf{a} \cdot \tau^{-1/2}$ — автомодельная переменная, т. е. $\max |\Omega| \sim 1/\tau$. Здесь мы считаем, что числитель в (42) — тангенциальный вектор к вихревой линии $\partial \mathbf{R} / \partial s$ — в точке коллапса не обращается в нуль, поскольку обращение якобиана в нуль в ситуации общего положения означает, что три вектора $\partial \mathbf{R} / \partial s$, $\partial \mathbf{R} / \partial \xi_2$ и $\partial \mathbf{R} / \partial \xi_3$ в этой точке при $t = t_0$ лежат в одной плоскости и ни один из них не обращается в нуль. В вырожденной ситуации последнее, вообще говоря, не обязательно.

Формула (42) демонстрирует сингулярное поведение завихренности при опрокидывании вихревых линий во вспомогательном пространстве переменных \mathbf{a} . В физическом пространстве структура особенности будет выглядеть иначе. Чтобы почувствовать эту разницу, достаточно обратиться к одномерной ситуации, или, говоря на языке теории катастроф, к одномерной складке, когда

$$J \equiv \frac{\partial x}{\partial a} = \alpha\tau + \gamma a^2. \quad (43)$$

Здесь для простоты мы поместили начало координат в точку $a = a_0$.

Соотношение (43) просто интегрируется:

$$x = \alpha\tau a + \frac{\gamma a^3}{3},$$

откуда следует, что в физическом пространстве возникает другой тип автомодельности $x \sim \tau^{3/2}$, т. е.

$$J = \tau f(x\tau^{-3/2}).$$

В момент опрокидывания $\tau = 0$ якобиан $J \sim x^{2/3}$. Для газовой динамики пыли, описываемой системой (34), это означает появление

степенных сингулярностей для плотности и градиента скорости (см., например, [20]):

$$\rho \sim \frac{\partial v}{\partial x} \sim \frac{1}{x^{2/3}}.$$

Отсюда становится ясным, что в трехмерном случае в r -пространстве должна возникать сильная анизотропия, связанная с различным поведением собственных значений матрицы Якоби.

Наша задача теперь — найти для трехмерного случая аналоги уравнения (43). Учитывая, что нас интересует малая окрестность вблизи точки опрокидывания, мы можем ограничиться локальным представлением (14):

$$\Omega(\mathbf{r}, t) = \frac{1}{J} \cdot \frac{\partial \mathbf{R}}{\partial s}, \quad (44)$$

где

$$J = \left(\left[\frac{\partial \mathbf{R}}{\partial \xi_1} \times \frac{\partial \mathbf{R}}{\partial \xi_2} \right] \cdot \frac{\partial \mathbf{R}}{\partial s} \right) \quad (45)$$

— якобиан преобразования от переменных (x, y, z) к переменным (ξ_1, ξ_2, s) .

Выражение (44) инвариантно относительно перепараметризации, т. е. замен $s \rightarrow \bar{s}(s, \xi_1, \xi_2, t)$. Мы выберем такую параметризацию, когда вектор $\partial \mathbf{R} / \partial s$ при $t = t_0$ в точке коллапса равен единичному вектору \mathbf{e}_1 .

Поскольку якобиан J (45) при $t = t_0$ в точке $\mathbf{a} = \mathbf{a}_0$ обращается в нуль ($J(\mathbf{a}_0, t_0) = 0$), то три вектора $\partial \mathbf{R} / \partial s$, $\partial \mathbf{R} / \partial \xi_1$ и $\partial \mathbf{R} / \partial \xi_2$ в этой точке компланарны — лежат в одной и той же плоскости, т. е. в точке касания

$$\frac{\partial \mathbf{R}}{\partial \xi_1} = c_1 \cdot \mathbf{e}_1, \quad \frac{\partial \mathbf{R}}{\partial \xi_2} = c_2 \cdot \mathbf{e}_2, \quad \frac{\partial \mathbf{R}}{\partial s} = \mathbf{e}_1. \quad (46)$$

Здесь \mathbf{e}_α — единичные векторы ($\alpha = 1, 2$), c_α — константы. При этом без ограничения общности векторы \mathbf{e}_1 и \mathbf{e}_2 можно считать ортогональными друг другу.

В малой окрестности точки $\mathbf{a} = \mathbf{a}_0$ наиболее важной добавкой к касательному вектору будет составляющая вектора, перпендикулярная плоскости с образующими $(\mathbf{e}_1, \mathbf{e}_2)$:

$$\frac{\partial \mathbf{R}}{\partial s} = \mathbf{e}_z + U[\mathbf{e}_1 \times \mathbf{e}_2], \quad |U| \ll 1. \quad (47)$$

В силу условий совместности

$$\frac{\partial^2 \mathbf{R}}{\partial s \partial \xi_\alpha} = \frac{\partial^2 \mathbf{R}}{\partial \xi_\alpha \partial s}, \quad \frac{\partial^2 \mathbf{R}}{\partial \xi_2 \partial \xi_1} = \frac{\partial^2 \mathbf{R}}{\partial \xi_1 \partial \xi_2} \quad (48)$$

два других вектора $\partial \mathbf{R} / \partial \xi_1$, $\partial \mathbf{R} / \partial \xi_2$ будут также иметь перпендикулярные добавки:

$$\frac{\partial \mathbf{R}}{\partial \xi_1} = c_1 \mathbf{e}_1 + V[\mathbf{e}_1 \times \mathbf{e}_2], \quad (49)$$

$$\frac{\partial \mathbf{R}}{\partial \xi_2} = c_2 \mathbf{e}_2 + W[\mathbf{e}_1 \times \mathbf{e}_2]. \quad (50)$$

Благодаря (47) функции U, V, W связаны между собой соотношениями

$$\frac{\partial U}{\partial \xi_1} = \frac{\partial V}{\partial s}, \quad \frac{\partial U}{\partial \xi_2} = \frac{\partial W}{\partial s}, \quad \frac{\partial V}{\partial \xi_2} = \frac{\partial U}{\partial \xi_1},$$

которые показывают, что вектор с компонентами U, V, W может быть представлен в виде градиента некоторой скалярной функции:

$$U = \frac{\partial \chi}{\partial s}, \quad V = \frac{\partial \chi}{\partial \xi_1}, \quad W = \frac{\partial \chi}{\partial \xi_2}. \quad (51)$$

В главном порядке собственные значения матрицы Якоби \hat{J} равны:

$$\lambda_1 = c_1, \quad \lambda_2 = c_2, \quad (52)$$

$$\lambda_3 = \frac{\partial \chi}{\partial s} - \frac{1}{c_1} \frac{\partial \chi}{\partial \xi_1}. \quad (53)$$

Выпуклость J в окрестности точки опрокидывания означает, что для соответствующего значения λ_3 мы можем записать:

$$\lambda_3 = (c_1 c_2)^{-1} (\alpha \tau + \hat{\gamma} a^2). \quad (54)$$

Здесь начало координат помещено в точку \mathbf{a}_0 , так что вектор $a = (\xi_1, \xi_2, s)$ в (54) предполагается малым, $\hat{\gamma} a^2 \equiv \gamma_{ij} a_i a_j$.

Совершая преобразование

$$\tilde{s} = s, \quad \tilde{\xi}_1 = \xi_1 + \frac{s}{c_1},$$

собственное значение λ_3 (53) в новых переменных запишем как производную от \tilde{s} :

$$\lambda_3 = \frac{\partial \chi}{\partial \tilde{s}}. \quad (55)$$

(Здесь и всюду ниже знак тильды опущен.) В результате интегрирования соотношений (47), (49), (50) для искомого отображения получим выражение

$$\mathbf{R} = c_1 \xi_1 \mathbf{e}_1 + c_2 \xi_2 \mathbf{e}_2 + \chi \mathbf{e}_3, \quad (56)$$

где функция χ определяется из интегрирования уравнения (55):

$$c_1 c_2 \frac{\partial \chi}{\partial s} = \alpha \tau + \gamma_{\alpha\beta} \xi_\alpha \xi_\beta + 2\gamma_{3\beta} \xi_\beta s + \gamma s^2. \quad (57)$$

Здесь и всюду ниже греческие индексы α, β пробегает значения 1, 2, $\gamma \equiv \gamma_{33}$. В этом уравнении линейное по s слагаемое вносит анизотропию в плоскости ξ_α . В дальнейшем будем рассматривать симметричную ситуацию, когда коэффициенты $\gamma_{3\beta} = 0$. Обобщение на случай произвольной анизотропии представляется достаточно простым.

В ситуации, когда $\gamma_{3\beta} = 0$, из (57) следует:

$$x_3 = (c_1 c_2)^{-1} \left[(\alpha \tau + \gamma_{\alpha\beta} \xi_\alpha \xi_\beta) s + \frac{\gamma s^3}{3} \right]. \quad (58)$$

Выражение $\gamma_{\alpha\beta} \xi_\alpha \xi_\beta$ в (57), (58) можно рассматривать как перенормировку времени τ . До момента опрокидывания $\tau = 0$ выражение в скобках в (58) всегда положительно. При $\tau < 0$ круглая скобка обращается в нуль на эллипсе:

$$\gamma_{\alpha\beta} \xi_\alpha \xi_\beta = \alpha |\tau|.$$

Это уравнение при $t > t_0$ задает расширение области опрокидывания — каустики. В силу (56) расширение размеров этого эллипса будет пропорционально $\sqrt{|\tau|}$.

При $\tau > 0$, т. е. до момента опрокидывания, искомое отображение задается неявно (56) совместно с (58), а завихренность в главном порядке имеет вид

$$\Omega = \frac{\mathbf{e}_1}{\alpha \tau + \gamma s^2 + \tilde{\gamma}_{\alpha\beta} x_\alpha x_\beta (c_1 c_2)^{-1}}, \quad (59)$$

где параметр s подлежит исключению с помощью уравнения (56), $\tilde{\gamma}_{\alpha\beta} = \gamma_{\alpha\beta} (c_\alpha c_\beta)^{-1}$.

Отсюда видно, что в \mathbf{g} -пространстве автомодельности по переменным x_3 и x_α различны: вдоль третьей оси $x_3 \sim \tau^{3/2}$, а по двум другим направлениям $x_\alpha \sim \tau^{1/2}$. Таким образом, распределение завихренности Ω оказывается сильно анизотропным:

$$\Omega = \frac{1}{\tau} \mathbf{G} \left(\frac{x_\alpha}{\tau^{1/2}}, \frac{x_3}{\tau^{3/2}} \right), \quad (60)$$

имеющим при $\tau \rightarrow 0$ блинообразную форму с сильно различающимися масштабами.

Из (59) следует также, что при $\tau \rightarrow 0$, т. е. когда

$$\eta_3 = \frac{x_3}{\tau^{3/2}} \rightarrow \infty \quad \text{и} \quad \eta_\alpha = \frac{x_\alpha}{\tau^{1/2}} \rightarrow \infty,$$

завихренность приобретает особенность с сильной анизотропией. В области $|x_3| \gg C|x_\perp|^3$, где $C = \tilde{\gamma}_\perp^{3/2} \gamma^{-1/2} (c_1 c_2)^{-1}$, лидирующим членом в знаменателе (59) при $\tau = 0$ является слагаемое, пропорциональное s^2 , где $s = (3c_1 c_2 x_3 / \gamma)^{1/3}$. Соответственно в этой области Ω имеет степенную сингулярность относительно x_3 :

$$\Omega = \frac{\mathbf{e}_1}{\gamma^{1/3} (3c_1 c_2 x_3)^{2/3}}. \quad (61)$$

В очень узкой области $x_\perp \gg C^{-1/3} x_3^{1/3}$, т. е. практически в поперечном направлении к $\mathbf{e}_3 = [\mathbf{e}_1 \times \mathbf{e}_2]$, у Ω имеется особенность типа $1/x_\perp^2$:

$$\Omega = \frac{\mathbf{e}_1}{\tilde{\gamma}_{\alpha\beta} x_\alpha x_\beta}. \quad (62)$$

Из (61) и (62) для Ω можно составить интерполяционную формулу:

$$\Omega = \frac{\mathbf{e}_1}{\gamma^{1/3} (3c_1 c_2 x_3)^{2/3} + \tilde{\gamma}_{\alpha\beta} x_\alpha x_\beta}. \quad (63)$$

Таким образом, завихренность, определяемая посредством (61), демонстрирует сингулярное поведение, аналогичное колмогоровской зависимости (7). В дальнейшем зависимость (61) мы будем называть колмогоровской.

Как нестационарная (60), так и предельная (61) зависимости $\Omega(r, t)$ при опрокидывании справедливы в области $L < \gamma^{-1/2}$. При $L \sim \gamma^{-1/2}$ значение Ω должно сравниваться со средним значением, которое порядка начального Ω_0 . Поэтому константа пропорциональности между Ω и $x_3^{-2/3}$ оценивается как $P^{1/3}$, где $P \sim \Omega_0^3 L^2$ по порядку величины совпадает с колмогоровским значением для потока энергии.

3.3. О диссипации на начальной стадии опрокидывания

Известно, что диссипация энергии определяется интегралом $E_\nu = \nu \int_V (\text{rot } \mathbf{v})^2 d\mathbf{r}$, где ν — коэффициент вязкости. Подставляя в это выражение (42) и переходя затем в этом интеграле от переменных \mathbf{r} к переменным \mathbf{a} , получим

$$\nu \int_V \frac{d\mathbf{a}}{\alpha\tau + \gamma_{ij} a_i a_j}. \quad (64)$$

Здесь интегрирование ведется по объему примерно с тем же размером, что и L , поскольку коллапсирующее решение сшивается с более медленным основным течением как раз на масштабах порядка L . Из (64) следует, во-первых, что этот интеграл сходится при $a \rightarrow 0$ и $\tau \rightarrow 0$, т. е. вклад от особенности в этот интеграл мал — по этой причине в [11] этот коллапс был назван сверхслабым. Во-вторых, переходя в (64) к автомодельным переменным $\eta = a/\tau^{1/2}$, легко видеть, что интеграл набирает свою величину на больших масштабах $\sim L\tau^{-1/2}$, где интеграл линейно растет с $L/\tau^{1/2}$. Этот рост компенсируется множителем $\tau^{1/2}$, возникшим после перехода к переменным η . В результате E_ν представляет собой независящую от времени величину порядка

$$E_\nu \sim \nu \Omega_0^2 L^3,$$

т.е. малую величину по сравнению с потоком $Q \sim PL^3$ в эту область. Таким образом, коллапс на начальной стадии — вплоть до формирования *точечной* особенности — не обеспечивает необходимой диссипации энергии. Согласование должно возникнуть на следующей стадии — стадии возникновения каустик и их расширения. Этот вопрос, однако, выходит за рамки данной статьи и будет обсужден в последующих работах. Здесь важным является то, что вязкость не влияет на колмогоровскую асимптотику завихренности (61).

4. Заключительные замечания

Таким образом, мы показали, что вклад в энергию от особенности оказывается малым — пропорциональным $\tau^{3/2}$. Однако интеграл $\int (\nabla \Omega)^2 d\mathbf{r}$ при $t \rightarrow t_0$ уже расходится. Поэтому при опрокидывании решение $\mathbf{v} = \mathbf{v}(\mathbf{r}, t)$ при $t \geq t_0$ не продолжимо в пространстве Соболева $H^2(\mathcal{R}^3)$ с нормой $\left(\sum_{q \leq 2} \int (\nabla^q f)^2 d\mathbf{r} \right)^{1/2}$, где q — мультииндекс. Согласно теореме, доказанной в [22], этого достаточно, чтобы расходился интеграл

$$\int_0^{t_0} \sup_{\mathbf{r}} |\Omega| dt = \infty. \quad (65)$$

Этот критерий, являясь необходимым и достаточным для коллапса, выполняется для (59). По этому критерию асимптотическое выражение для завихренности (59) относится к пограничному случаю, при котором интеграл (65) расходится логарифмически. Важно отметить, что многие существующие численные эксперименты, прежде всего Керра [23], Пелца [24], Пелца и Боратава [25], Грауэра, Марлиани

и Гермашевского [26], показывают поведение завихренности вблизи точки коллапса как $(t_0 - t)^{-1}$, что согласуется с разложением (59). Однако из этих численных экспериментов, основанных на прямом интегрировании уравнений Эйлера, невозможно сделать вывод о том, является ли данный процесс взрывного роста опрокидыванием вихревых линий. Этот пробел был восполнен в работах [12, 13], в которых численное решение задачи о коллапсе было выполнено в рамках частично проинтегрированных уравнений Эйлера, т. е. разрешенных относительно инвариантов Коши. Несмотря на недостаточно высокое пространственное разрешение (число точек — 128^3), было показано, что коллапс имеет место за счет обращения якобиана в нуль. Якобиан в точке коллапса по мере приближения к $t = t_0$ с высокой точностью обращался в нуль по *линейному* закону, при этом коэффициенты разложения $\gamma_{\alpha\beta}$ в точке минимума J практически не менялись во времени. Все это свидетельствовало в пользу того, что коллапс в гидродинамике есть процесс опрокидывания вихревых линий. Если это так, то в соответствии с результатами, представленными в данной статье, в окрестности точки опрокидывания при $t \rightarrow t_0$ должна формироваться особенность завихренности с колмогоровским поведением.

Автор благодарит Е. И. Якубовича за возможность ознакомиться с работой [15] до ее опубликования. Данная работа была поддержана РФФИ (грант 00-01-00929), Программой ведущих научных школ РФ (грант 00-15-96007) и INTAS (грант 00-00292).

Литература

1. Захаров В. Е. Коллапс ленгмюровских волн // ЖЭТФ. 1972. **62**. 1745
2. Phillips O. H. The equilibrium range in the spectrum of wind-generated water waves // J. Fluid Mech. 1958. **4**. 426.
3. Захаров В. Е., Филоненко Н. Н. Спектр энергий стохастических колебаний поверхности жидкости // ДАН СССР. 1967. **170**. 1292.
4. Zakharov V. E. Weakly nonlinear waves on the surface of an ideal finite depth fluid // Amer. Math. Soc. Transl. (2) 1998. **182**. 167–197.
5. Кадомцев В. Б., Петвиашвили В. И. О звуковой турбулентности // ДАН СССР. 1973. **208**. 794.
6. Захаров В. Е., Сагдеев Р. З. Спектр акустической турбулентности // ДАН СССР [Sov. Phys. Dokl. **15**. 439 (1970)].
7. Колмогоров А. Н. Локальная структура турбулентности в несжимаемой жидкости при больших числах Рейнольдса // ДАН СССР. 1941. **30**. 9.
8. Обузов А. М. О распределении энергии в спектре турбулентного потока // ДАН СССР 1941. **32** 22.

9. Ландау Л. Д., Лифшиц Е. М. Гидродинамика. М.: Наука, 1988.
10. Кузнецов Е. А., Рубан В. П. Коллапс вихревых линий в гидродинамике // ЖЭТФ. 2000. **118**. 893.
11. Желиговский В. А., Кузнецов Е. А., Подвигина О. Н. Численное моделирование коллапса в идеальной несжимаемой гидродинамике // Письма в ЖЭТФ. 2001. **74**. 402.
12. Kuznetsov E. A., Podvigina O. N., Zheligovsky V. A. Numerical evidence of breaking of vortex lines in an ideal fluid // Proc. of IUTAH Symp. «Tubes, Sheets and Singularities in Fluid Dynamics». Zakopane, Kluwer, 2002 (in press).
13. Кузнецов Е. А., Рубан В. П. Гамильтоновская динамика вихревых линий в системах гидродинамического типа // Письма в ЖЭТФ. 1998. **67**. 1015.
14. Kuznetsov E. A., Ruban V. P. Hamiltonian dynamics of vortex and magnetic lines in the hydrodynamic type models // Phys. Rev. E. 2000. **61**. 831.
15. Yakubovich E. I., Zenkovich D. A. Matrix Fluid Dynamics // Proc. of Int. Conf. «Progress in Nonl. Sci.» (July 2001, N. Novgorod, Russia), vol. II «Frontiers of Nonlinear Physics»/ Ed. by A. G. Litvak, N. Novgorod, 2002. P. 282–287; physics/0110004.
16. Захаров В. Е., Кузнецов Е. А. Гамильтоновский формализм для нелинейных волн // УФН 1977. **137**. 1137.
17. Salmon R. Hamiltonian Fluid Mechanics // Ann. Rev. Fluid Mech. 1988. **20**. 225.
18. Халатников И. М. Теория сверхтекучести. М.: Наука, 1971.
19. Frisch U. Turbulence. The legacy of A.N.Kolmogorov. Cambridge Univ. Press (1995).
20. Shandarin S. F., Zeldovich Ya. B. // Rev. Mod. Phys. 1989. **61**. 185.
21. Захаров В. Е., Кузнецов Е. А. Квазиклассическая теория трехмерного волнового коллапса // ЖЭТФ. 1986 **91**. 1310.
22. Beale J. T., Kato T., Majda A. J. Remarks on the breakdown of smooth solutions for the 3-D Euler equations // Comm Math. Phys 1984. **94**. 61.
23. Kerr R. M. Evidence for a singularity of the 3-dimensional, incompressible Euler equations // Phys Fluids A. 1993. **5**. 1725.
24. Pelz R. B. Locally self-similar, finite-time collapse in a high-symmetry vortex filament model // Phys. Rev. E. 1997. **55**. 1617.
25. Boratav O. N., Pelz R. B. Direct numerical simulation of transition to turbulence from high-symmetry initial condition // Phys Fluids. 1994. **6**. 2757.
26. Grauer R., Mariani C., Germaschewski K. Adaptive mesh refinement for singular solutions of the incompressible Euler equations // Phys. Rev. Lett. 1998. **80**. 4177.

НЕЛИНЕЙНЫЕ ВОЛНЫ И СТРУКТУРЫ В СТРАТИФИЦИРОВАННЫХ СДВИГОВЫХ ПОТОКАХ

Ю. И. Троицкая

Введение

Для природных потоков водных и воздушных масс характерно наличие стратификации по плотности и скорости потока. Моделью таких течений является плоскопараллельный стратифицированный сдвиговый поток с профилями горизонтальной скорости $\vec{u} = U_0(z)\vec{x}_0$ и плотности $\rho = \rho_0(z)$. По отношению к возмущениям такого потока стратификация,

характеризуемая частотой плавучести $N = \left(-\frac{g}{\rho_0} \frac{d\rho_0}{dz} \right)^{1/2}$, играет роль

стабилизирующего фактора, а сдвиг скорости, определяемый величиной $\frac{dU}{dz}$, – дестабилизирующего. Их баланс определяет условие устойчиво-

сти потока и характеризуется безразмерным параметром $Ri = N^2 / U_z^2$ – числом Ричардсона. Достаточное условие устойчивости плоскопараллельного стратифицированного сдвигового потока – относительно бесконечно малых возмущений – критерий Майлса–Ховарда [1, 2]: $Ri > 1/4$ в любой точке потока. Если в какой-либо области потока $Ri < 1/4$, то могут существовать экспоненциально нарастающие собственные волны. При больших надкритичностях волны становятся сильно нелинейными и динамика потока может стать сложной, приводя к развитию турбулентности

Однако даже в слабонадкритических течениях, а также для волн в устойчиво стратифицированных потоках существуют области сильной нелинейности волновых полей – критические слои (КС), в которых волновые возмущения гидродинамических полей порядка средних. Такие слои образуются вблизи критических уровней (КУ), на которых скорость потока равна фазовой скорости волновых возмущений [3, 4]. При этом возникает сильное взаимодействие волн с потоком, называемое резонансным.

В настоящей работе приведен обзор результатов исследований нелинейного резонансного взаимодействия волн со стратифицированными сдвиговыми потоками. Статья имеет три раздела: в 1-м вводится понятие КС в стратифицированном сдвиговом потоке, во 2-м обсуждается взаимодействие волн с динамически устойчивым потоком, а в 3-м перечисляются основные результаты, полученные при исследовании взаимодействия волн с динамически неустойчивым потоком.

1. Критический слой в стратифицированном сдвиговом потоке

Представления о КС возникают при асимптотическом анализе гармонических волновых возмущений малой амплитуды. При больших числах Рейнольдса для описания таких волн применимо линейное невязкое приближение, и комплексная амплитуда возмущений функции тока ψ_1 удовлетворяет уравнению Тейлора–Гольдштейна (Т–Г):

$$\frac{d^2\psi_1}{dz^2} + \frac{U_{0zz}}{U_0 - \omega/k} \psi_1 + \left[\frac{N^2}{(U_0 - \omega/k)^2} - k^2 \right] \psi_1 = 0. \quad (1)$$

Если при $z = z_c$ скорость потока совпадает с фазовой скоростью волны $U_0(z_c) = \omega/k$, то z_c является КУ, на котором уравнение Т–Г имеет полюс второго порядка. Решение уравнения Т–Г вблизи КУ можно найти методом Фробениуса. Оно имеет точку ветвления:

$$\psi_1 = A(z - z_c)^{1/2+\mu} \varphi_a(z) + B(z - z_c)^{1/2-\mu} \varphi_b(z), \quad (2)$$

где $\mu = \sqrt{4^{-1} - \text{Ri}_c}$, $\text{Ri}_c = \text{Ri}(z_c)$.

При описании взаимодействия волн с потоком важную роль играет радиационное напряжение $T = \langle u_w \rangle = -\langle \psi_{1y} \psi_{1x} \rangle$. Производная от него по z определяет силу, действующую на поток в направлении x . Интегрирование уравнения Т–Г показывает, что эта величина постоянна вне КУ, а на КУ испытывает скачок [3]. Это означает, что в нем на поток действует радиационная сила, которая может вызвать существенную деформацию среднего потока.

Запишем в окрестности КУ выражение для полной функции тока:

$$\Psi = \frac{U_z (z - z_c)^2}{2} + \varepsilon \text{Re} \left[\left(A(z - z_c)^{1/2+\mu} + B(z - z_c)^{1/2-\mu} \right) e^{-i(\omega t - kx)} \right].$$

Из него видно, что с приближением к КУ амплитуда волновых возмущений растет быстрее, чем невозмущенное течение, т. е. даже возмущение малой амплитуды в окрестности КУ может стать сильно нелинейным. Масштаб, на котором возмущения становятся порядка средних полей, называется масштабом нелинейного КС, δ_{NL} . Его величина зависит от стратификации [5]: при $\text{Ri}_c > 1/4$ $\delta_{NL} = \varepsilon^{2/3}$, при $\text{Ri}_c < 1/4$ $\delta_{NL} = \varepsilon^{2/(3+\sqrt{1-4\text{Ri}_c})}$.

В КС, т. е. в окрестности КУ порядка δ_{NL} , уравнение Т–Г не применимо даже для волны малой амплитуды, а нужно решать полную систему уравнений гидродинамики. Заметим, что поскольку в КС возмущения функции тока порядка средних, то в нем линии тока замкнуты. Возникающая при этом структура течения носит название «кошачьих глаз» [6].

Формальная математическая процедура описания взаимодействия волны с течением в КС – это метод сращиваемых асимптотических разложений. Главная задача при этом – найти связь волновых полей и средних величин по разные стороны от КС – правила обхода. Заметим, что такая постановка задачи аналогична задаче о нелинейном взаимодействии частиц плазмы с монохроматической волной [7].

Во внутренних переменных нелинейного КС

$$Y = (z - z_c) / \delta_{NL}; \quad \Phi = \psi / \delta_{NL}^2; \quad B = b / \delta_{NL}; \quad \tau = t \delta_{NL},$$

система уравнений гидродинамики имеет вид:

$$\begin{aligned} \frac{\partial X}{\partial \tau} + \frac{\partial \Phi}{\partial Y} \frac{\partial X}{\partial x} - \frac{\partial \Phi}{\partial x} \frac{\partial X}{\partial Y} - \text{Ri} \frac{\partial B}{\partial x} &= \lambda \frac{\partial^2 X}{\partial Y^2}, \\ \frac{\partial B}{\partial \tau} + \frac{\partial \Phi}{\partial Y} \frac{\partial B}{\partial x} - \frac{\partial \Phi}{\partial x} \frac{\partial B}{\partial Y} &= \frac{\lambda}{\text{Pr}} \frac{\partial^2 B}{\partial Y^2}, \\ X &= \frac{\partial^2 \Phi}{\partial Y^2}. \end{aligned} \quad (3)$$

Из 1-го уравнения системы (3) очевидно, что $\lambda = 1 / \text{Re} \delta_{NL}^3$ имеет смысл обратного числа Рейнольдса течения в КС.

2. Взаимодействие волн с динамически устойчивым стратифицированным сдвиговым потоком

2.1. Нелинейный вязко-диффузионный критический слой в стратифицированном сдвиговом потоке

Рассмотрим особенности КС в стратифицированной жидкости с $\text{Ri} > 1/4$, подробно исследованного в [8]. В этом случае в жидкости могут распространяться внутренние волны, модифицированные за счет присутствия сдвига скорости, это соответствует тому, что в решении Фробениуса (2) $\mu = i\sqrt{\text{Ri} - 4}^{-1}$ – чисто мнимое число. При этом в (2) слагаемое со знаком (-) соответствует волне с положительной групповой скоростью, а (+) – с отрицательной. Тогда задачу о правиле обхода можно сформулировать в терминах коэффициентов отражения и прохождения.

Итак, пусть на стратифицированный сдвиговый поток падает гармоническая волна с заданной амплитудой A , имеющая на нем КУ. Требуется найти коэффициенты отражения R и прохождения T . В нелинейном КС возможны эффекты детектирования и генерации гармоник. Это означает, что в задачу нахождения правил обхода входит определение средних полей скорости и плотности по разные стороны от КС и амплитуд излучаемых высших гармоник.

Правила обхода (т. е. связь амплитуд волн и средних полей) определяются параметрами λ , Ri и Pr , входящими в систему (3). Нелинейность

течения внутри КС характеризуется параметром λ (обратное число Рейнольдса). При большом λ течение внутри КС вязкое, и нелинейными эффектами можно пренебречь. В этом случае получается известное правило обхода: скачка среднего течения нет ($\Gamma_+ = \Gamma_-$), высшие гармоники не излучаются ($R_n = T_n = 0$, при $n > 1$), отраженная волна отсутствует ($R = 0$), прошедшая волна сильно ослабляется [6]. Другая картина взаимодействия волны с КС возникает в начальной задаче о падении волны конечной амплитуды на КС. При этом на достаточно малых временах можно не учитывать влияния вязкости. В такой постановке задача была рассмотрена в [9]. Было показано, что при этом возникают среднее течение и отраженная волна, нарастающие во времени.

На больших временах от начала взаимодействия волны с потоком необходимо учитывать вязкость, что должно приводить к формированию стационарного течения в КС. При этом возникает скачок завихренности (или излом скорости) через КС. Причем в случае $Ri > 1/4$ он возникает в нулевом порядке теории возмущений по амплитуде волны, т. е. имеет порядок невозмущенного течения. Происхождение его связано со следующим. В КС имеет место скачок радиационных напряжений, т. е. действует радиационная сила, которая в стационарном течении должна быть уравновешена вязкой силой. Это означает, что через КС должен быть скачок вязких напряжений и пропорциональный ему скачок средней завихренности. В отличие от линейного вязкого КС возникает отраженная волна с коэффициентом отражения R .

Скачок завихренности и коэффициенты отражения и прохождения определяются параметрами течения в КС: λ , Ri и Pr . Зависимости от λ при различных Ri для $Pr = 0,71$, соответствующего температурной стратификации воздуха, были получены в [8]. Для температурной стратификации воды $Pr = 7$. На рис. 1 приведены зависимости при $Pr = 0,71$ и $Pr = 7$. Видно, что при $Pr = 7$ существенно больше коэффициент отражения, что является признаком более сильной нелинейности.

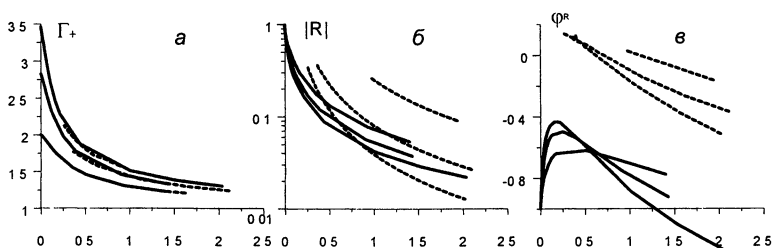


Рис. 1. Зависимости параметров течения от λ при различных Ri и Pr *a* – скачок завихренности, *б* и *в* – модуль и фаза коэффициента отражения. Сплошные линии – $Pr = 0,71$, пунктирные – $Pr = 7$

Решение во внешней области определяется сильно нелинейной областью, КС. Рассмотрим, какова в нем структура течения при различных числах Прандтля. На рис. 2 показаны картины линий тока и линий равной плотности течения в окрестности КС при $Pr = 0,71$ и $Pr = 7$ при близких значениях параметра нелинейности и числа Ричардсона. Видно, что при большем значении числа Прандтля структура поля плотности внутри КС существенно более «нелинейна», а структуры линий тока практически одинаковы. Качественно формирование такой структуры поля плотности можно пояснить следующим образом. Вблизи седловых точек «кошачьих глаз» деформирующее поле скорости приводит к сильному растягиванию линий равной плотности, а диффузия приводит к «расплыванию» образующейся при этом тонкой «ленты». В результате формируется характерная структура поля плотности, изображенная на рис. 2. Подобная структура поля плотности внутри КС наблюдалась в лабораторных экспериментах [10].

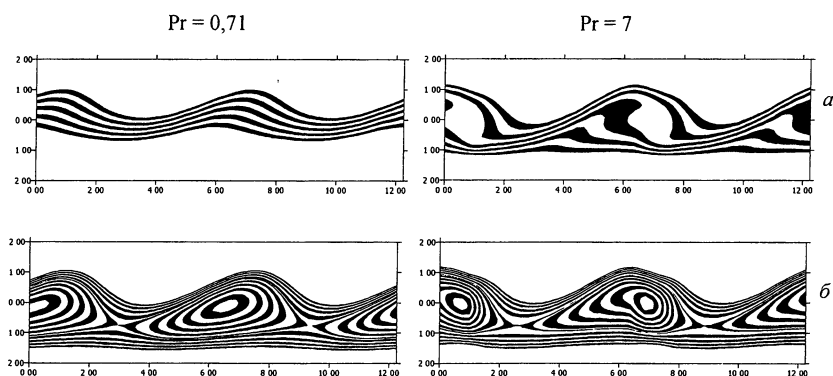


Рис. 2. Структура течения в критическом слое при различных числах Прандтля a – изолинии плотности, b – линии тока

Итак, в рамках стационарной задачи получается, что возмущения средней скорости растут при удалении от КС. Возникает вопрос о реализуемости такого течения, особенно важный для случая $Ri > 1/4$, когда деформация профиля средней скорости имеет порядок невозмущенного течения. Ответить на него можно, если учесть нестационарность течения во внешней области [11]. В этом случае средняя горизонтальная скорость удовлетворяет диффузионному уравнению с источником:

$$\frac{\partial U}{\partial \tau} - \lambda \frac{\partial^2 U}{\partial Y^2} = F_{rad}, \quad (4)$$

где источник – это радиационная сила, отличная от нуля только внутри КС. На бесконечности профиль скорости не возмущен.

При постоянстве скачка завихренности через КС можно найти автомодельное решение уравнения (4), справедливое при $\tau \gg \delta_{NL}^2 / \lambda$. Характерный вид профиля скорости, соответствующий этому решению, представлен на рис. 3. Это течение имеет два масштаба: толщину КС и ширину переходной области $\sqrt{\lambda t} \gg \delta_{NL}$, на которой происходит согласование возмущенного течения около критического слоя и невозмущенного течения на бесконечности. Следует обратить внимание на то, что при этом происходит смещение КС навстречу падающей волне пропорционально \sqrt{t} .

Течения, подобные тем, которые были получены в рамках асимптотической процедуры, наблюдались в численных и лабораторных экспериментах [12–14].

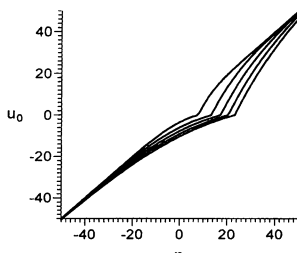


Рис. 3. Профили скорости в последовательные моменты времени. Смещение КС происходит слева направо

2.2. Квазилинейная теория взаимодействия волновых пакетов со стратифицированным сдвиговым потоком

До сих пор речь шла о взаимодействии с потоком гармонической волны. Пусть теперь на поток падает пакет волн с конечной шириной спектра. Тогда, если размер КС мал по сравнению с шириной области взаимодействия, применимо квазилинейное приближение, аналогичное используемому в физике плазмы. В рамках этого подхода в [15] была рассмотрена задача о падении пакета внутренних волн с конечной шириной спектра на стратифицированный сдвиговый поток. Изучалась необратимая деформация средних гидродинамических полей (в данной задаче полей скорости и плотности), аналогичная деформации функции распределения в физике плазмы [5]. Оказалось, что при $Ri > 1/4$ деформация средней скорости описывается уравнением Римана, а не нелинейным диффузионным уравнением (ср. в плазме [5]), а необратимой деформации поля средней плотности не происходит.

В физике плазмы хорошо известно, что коэффициент диффузии, описывающий квазилинейную деформацию функции распределения, содержит компоненту, описывающую процесс «кажущейся» или «адиабатической» диффузии. Аналог такого процесса для стратифицированного сдвигового потока с $Ri > 1/4$ был рассмотрен в [16].

Одним из основных источников внутренних волн в природе (атмосфере и океане) является обтекание препятствий потоками. Однако в этом случае рассмотрение деформации сдвигового потока за счет взаимодействия с двумерными возмущениями в квазилинейном приближении невозможно, поскольку положение КС всех волн совпадает и находится на уровне, где скорость потока равна нулю. Это вырождение можно снять, если рассмотреть поток, имеющий две горизонтальные компоненты скорости $\vec{U}(z) = (U(z), V(z))$, на который падают волны с горизонтальным волновым вектором $\vec{k} = (k_x, k_y)$, излучаемые при обтекании неровной поверхности. Такая задача была рассмотрена в [17]. При этом координата КС определяется из условия перпендикулярности векторов $\vec{U}(z_c)$ и \vec{k} , т. е. $(\vec{k}\vec{U}(z_c)) = 0$. В [17] изучалась необратимая квазилинейная деформация исходного течения. При этом оказалось, что необратимой деформации поля плотности не происходит, как и в [15]. Модуль средней скорости также не меняется. Это обусловлено тем, что в каждом КС радиационная сила, сонправленная с волновым вектором, действует перпендикулярно вектору скорости, а значит, не совершает работы. При этом вектор скорости поворачивается, причем угол поворота удовлетворяет уравнению простой волны.

3. Резонансное взаимодействие волн с динамически неустойчивым стратифицированным сдвиговым потоком ($Ri < 1/4$)

3.1. Асимптотические модели

Основные результаты в рамках асимптотической теории для этого случая были получены в 70-е годы. При этом прежде всего были исследованы случаи $Ri \ll 1$ [18, 19], и зависимость скачка средней завихренности от λ была найдена только для однородной жидкости [19]. Также были проведены исследования асимптотическими методами эволюции слабонадкритического стратифицированного сдвигового потока с числом Ri близким к $1/4$ (см., например, [20–21]). После этого асимптотические методы для КС в потоке с $Ri < 1/4$ не развивались, а предпочтение было отдано численным методам.

3.2. Численные модели динамики неустойчивого стратифицированного сдвигового потока

В большинстве работ [22–27] рассматривалось нарастание возмущений в слое смещения. Если минимальное число Ричардсона такого течения меньше $1/4$, то такое течение является неустойчивым.

В первых численных работах [22–24] было исследовано установление периодических по x двумерных возмущений, имеющих максимальный линейный инкремент (их волновое число k_{\max}), численное решение проводилось на одном периоде волны. При этом устанавливалось течение, в котором структура линий равной завихренности и плотности, а также механизм их формирования аналогичны обсуждаемым выше для случая $Ri > 1/4$. Такое течение может стать неустойчивым относительно вторичных возмущений. При этом численно исследованы два принципиально разных типа неустойчивости: трехмерная мелкомасштабная [22–26] и двумерная субгармоническая [27, 28].

Остановимся подробнее на субгармонической неустойчивости. При ее моделировании интегрирование уравнений гидродинамики проводилось на области, занимавшей две длины волны наиболее неустойчивой моды. Вначале нарастало наиболее неустойчивое возмущение, а затем происходило слияние вихрей. Такое поведение объясняется в [28] на языке трехволнового резонансного взаимодействия волн с волновыми числами $k_2 = k_{\max}$ (волновым числом наиболее неустойчивой моды) и двух волн с $k_1 = k_{\max}/2$. Поскольку для волн в слое смешения дисперсия слабая, то для них выполняется условие резонанса $\omega_2 = \omega_{\max}$, $\omega_1 = \omega_{\max}/2$. Тогда можно записать уравнения для амплитуд основной моды A_2 и ее субгармоники A_1 :

$$\frac{dA_1}{dt} = \gamma_1 A_1^* A_2, \quad \frac{dA_2}{dt} = \gamma_2 A_1^2.$$

Причем для наиболее неустойчивой моды $\gamma_1, \gamma_2 > 0$, наблюдается взрывной рост амплитуд гармоник.

3.3. Лабораторные эксперименты

Структуры, которые образуются в результате развития неустойчивости в стратифицированных сдвиговых потоках, наблюдались экспериментально. На рис. 4 показано развитие неустойчивости Кельвина-Гельмгольца, зарегистрированное в [29]. Замечательно, что в условиях, когда нет выделенного масштаба, развивается наиболее неустойчивая мода. Формирование такой простой структуры, по-видимому, является следствием конкуренции мод.



Рис. 4. Развитие неустойчивости Кельвина-Гельмгольца

Дальнейшее развитие неустойчивости было прослежено в экспериментах [30], где было продемонстрировано формирование характерной двумерной структуры, слияние, а затем разрушение образовавшейся структуры за счет мелкомасштабных трехмерных возмущений.

Заключение

Основным современным методом исследования нелинейных волн и структур в стратифицированных сдвиговых потоках является прямое численное моделирование. Главная его цель – воспроизведение основных свойств течений, наблюдаемых в экспериментах. По существу, прямое численное моделирование представляет собой рафинированный компьютерный аналог реального эксперимента. Для качественного понимания динамики происходящих процессов требуется построение простых физических моделей, в основе которых лежат методы теории нелинейных колебаний и волн.

Работа подготовлена при финансовой поддержке Президентской программы по поддержке ведущих научных школ России (проекты 01-15-96772 и 01-15-99399).

Литература

- 1 Miles J W On the stability of heterogeneous shear flow // J Fluid Mech 1961 **10** Pt 4 P 496–509
- 2 Hovard L N Note to a paper of John Miles // J Fluid Mech 1961 **10**. Pt 4 P. 509–514
- 3 Маслюк С А Неустойчивости и переход в сдвиговых течениях // Гидродинамические неустойчивости и переход к турбулентности / Под ред. Х Суинни, Дж Голлаба М Мир, 1981 С 218–270.
- 4 Maslowe S A Critical layers in shear flows // Ann Rev Fluid Mech 1986 **18** P 405–432
- 5 Maslowe S A The generation of clear air turbulence by nonlinear wave // Stud Appl Math 1972 **51**. № 1 P 1–16
- 6 Тернер Дж Эффекты плавучести в жидкостях М Мир, 1977 431 с
- 7 Вопросы теории плазмы Вып 7 М Энергоатомиздат, 1973
- 8 Troitskaya Yu I Viscous diffusion nonlinear critical layer in a stratified shear flow // J. Fluid Mech 1991 **233** P 25–48
- 9 Brown S N, Stewartson K On the nonlinear reflection of a gravity wave at a critical layer Part 1 // J Fluid Mech 1980 V 100 P 577–595, Part 2 1982 **115** P 217–230, Part 3 1982 V 115 P 231–250
- 10 Thorp S A An experimental study of critical layers // J Fluid Mech 1981 **103** P 321–344
- 11 Troitskaya Yu I, Reznik S N Quasi-steady dissipative nonlinear critical layer in a stratified shear flow // Phys Fluids 1996 **8**, № 12 P 3313–3328
- 12 Fritts D The nonlinear gravity wave - critical layer interaction // J Atm Sci 1978 **35** P 397–413
- 13 Koop C G, McGee B Measurements of internal gravity waves in a continuously stratified shear flow // J Fluid Mech 1986 **172** P 453–480

14. *Coop C G* A preliminary investigation of the interaction of internal gravity waves with a steady shearing motion // *J Fluid Mech* 1981 **113** P 347–386.
15. *Цимринг Л Ш* Дестабилизация первоначально устойчивого возмущения в стратифицированном сдвиговом потоке // *Океанология* 1982 **22**, № 4 С 540–544
16. *Ерухимова Т Л, Токман М Д, Трахтенгерц В Ю* К квазилинейной теории взаимодействия внутренних гравитационных волн со сдвиговыми течениями // *Изв АН ФАО* 1998 **34(6)** С 827–834
17. *Резник С Н, Троицкая Ю И* Квазилинейная модель деформации стратифицированного ветра, меняющего направление над случайно-неоднородной поверхностью // *Изв вузов Радиофизика* 1999 **42** (3) С 255–265
18. *Haberman R* Wave induced distortions of slightly stratified shear flow // *J Fluid Mech* 1973 **58** P 727–735
19. *Haberman R* Critical layers in parallel flows // *Stud Appl Math* 1972 **51**, № 2 P 139–161.
20. *Stewartson K* Marginally stable inviscid flows with critical layers // *J. Appl Math* 1981 **27** P 133–175
21. *Churilov S M, Shukhman I G* Nonlinear stability of a stratified shear flow in the regime with an unsteady critical layer // *J Fluid Mech* 1988 **194** P 187–216
22. *Lin S J, Corcos G M* The mixing layer deterministic models of turbulent flow Part 3 The effect of plane strain on the dynamic of streamwise vortices // *J. Fluid Mach* 1982 **141** P 139–178
23. *Klaassen G P, Peltier W R* The onset of turbulence in finite amplitude Kelvin-Helmholtz billows // *J Fluid Mech* 1985 **155** P 1–35
24. *Klaassen G P, Peltier W R* The influence of stratification on secondary instability in free shear layers // *J. Fluid Mech* 1991 **227** P 71–106
25. *Smyth W D, Peltier W R* Instability and transition in finite amplitude Kelvin-Helmholtz and Holmboe waves // *J Fluid Mech* 1991 **228** P 387–415
26. *Caulfield C P, Peltier W R* The anatomy of the mixing transition in homogeneous and stratified free shear layers // *J Fluid Mech* 2000 **413** P 1–47
27. *Klaassen G P, Peltier W R* The role of transverse secondary instabilities in the evolution of free shear layers // *J Fluid Mech* 1989 **202** P 367–402
28. *Collins D A, Maslowe S A* Vortex pairing and resonant wave interactions in a stratified free shear layer // *J Fluid Mech* 1988 **191** P 465–480
29. *Lawrence G A, Browand F K, Redekopp L G* The stability of a sheared density interface // *Phys Fluids A* 1990 **3(10)** P 2360–2370
30. *Coop C G, Browand F K* Instability and turbulence in a stratified fluid with shear // *J Fluid Mech* 1979 **93** P 135–159

СВЯЗАННЫЕ СОСТОЯНИЯ И ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ХАОС ТОПОЛОГИЧЕСКИХ ДЕФЕКТОВ

А. Б. Езерский

Введение

Нарушения порядка и переход к хаосу в пространственно-периодических структурах, возникающих в самых различных по своей физической природе системах, могут быть связаны с возникновением и взаимодействием топологических дефектов. В простейшем случае, когда поле параметра представляет собой роликовую структуру, о топологическом дефекте говорят, если в одной части пространства на один период больше, чем в другой (рис. 1, *a*). Если описывать такое поле u гармонической функцией с изменяющейся амплитудой и фазой: $u(x, y) = R(x, y) \cos[kx + \Phi(x, y)]$ (x, y – декартовы координаты, k – волновое число), то поля R и Φ будут выглядеть так, как показано на рис. 1, *b* и 1, *в*. Амплитуда уменьшается до нуля в начале координат и стремится к постоянной величине на большом удалении, а фаза монотонно увеличивается

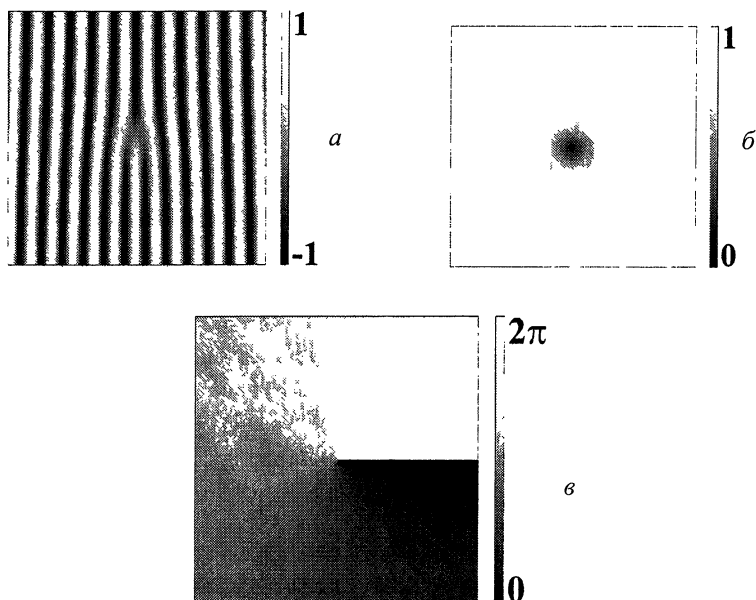


Рис. 1. Топологический дефект в пространственно-периодической структуре
a – поле параметра порядка, *b* – поле амплитуды, *в* – поле фазы

от 0 до 2π при обходе вокруг точки с координатами (0, 0). Такие топологические дефекты могут быть обнаружены, когда пространственная периодичность возникает в диссипативной среде из-за неустойчивости однородного распределения, например, при термоконвекции в слое жидкости, подогреваемого снизу, или в лазерном пучке, распространяющемся в нелинейной среде, или на шкуре зебры, периодическая окраска которой обусловлена неоднородным распределением реагентов при химических реакциях, происходящих в тканях животного. За внешней похожестью зрительных образов лежат достаточно глубокие физические причины формирования похожих нарушений порядка в перечисленных выше системах. В этих системах, кроме того, что существует определенный масштаб неустойчивости, есть одинаковые нелинейные механизмы ее ограничения, а линейная диссипация энергии связана с диффузией поля u . Поэтому топологические дефекты, возникающие в различных периодических полях обладают похожими свойствами – они могут образовывать связанные состояния, доменные стенки, в ансамбле взаимодействующих дефектов возникает пространственно-временной хаос. Подобные эффекты изучаются в данной работе на примере нескольких гидродинамических систем, где возникают пространственно периодические структуры. Статья организована следующим образом. Вначале будет представлена математическая модель топологического дефекта (разд. 1), затем – экспериментальные данные по взаимодействию дефектов в трех системах – в конвекции Марангони – Бенара в слое жидкости, подогреваемого снизу (разд. 2), в ряби Фарадея (разд. 3) и в вихревой дорожке Кармана за нагретым цилиндром (разд. 4). В заключение сформулированы выводы, касающиеся динамики дефектов в периодических структурах.

1. Математическая модель топологического дефекта

Для описания динамики дефектов очень часто используются уравнения, которые получаются феноменологически, а не последовательным выводом из фундаментальных уравнений (уравнений Навье – Стокса, Максвелла или кинетических уравнений, описывающих химические реакции). Одно из таких уравнений – диссипативный предел уравнения Гинзбурга – Ландау для комплексной амплитуды $A(x, y, t)$ роликовой структуры $u = A(x, y, t)e^{ikx} + k.c.$ – имеет следующий вид:

$$\frac{\partial A}{\partial t} = \epsilon A - A |A|^2 + \left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \right) A. \quad (1)$$

Первое слагаемое в правой части (1) при $\epsilon > 0$ описывает неустойчивость возмущений с длиной волны $\lambda = 2\pi/k$, второе слагаемое – нелинейные потери, приводящие к ограничению амплитуды, последнее – линейные

потери, связанные с диффузией. Если представить комплексную амплитуду в виде $A = R(x, y) \exp[i\Phi(x, y)]$, то уравнение (1) примет вид

$$\begin{aligned} \frac{\partial R}{\partial t} &= \varepsilon R - R(\nabla\Phi)^2 - R^3 + \Delta R, \\ \frac{\partial \Phi}{\partial t} &= \frac{2}{R} \nabla\Phi \nabla R + \Delta\Phi, \end{aligned} \tag{2}$$

где $\Delta = \frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2}$, $\nabla = \bar{x}_0 \frac{\partial}{\partial x} + \bar{y}_0 \frac{\partial}{\partial y}$, \bar{x}_0 , \bar{y}_0 – единичные векторы.

Найдем стационарное решение системы (2), соответствующее топологическому дефекту. Перейдем к полярным координатам $x = r \cos \phi$, $y = r \sin \phi$ и будем считать, что амплитуда R зависит только от r , а фаза Φ – от угла ϕ . В этом случае произведение градиентов амплитуды и фазы во втором уравнении системы (2) $\nabla\Phi(\phi)\nabla R(r) = 0$, и можно найти решение для фазы $\Phi = \pm n\phi$, где $n = 1, 2, 3 \dots$. Уравнение для амплитуды аналитически не решается, однако можно найти асимптотики при $r \rightarrow 0$ и $r \rightarrow \infty$. Для $n = 1$ они имеют вид

$$R(r) \sim \begin{cases} r, & r \rightarrow 0, \\ (1 - \frac{1}{2\varepsilon r^2}), & r \rightarrow \infty. \end{cases}$$

Если теперь рассчитать поле u , для комплексной амплитуды A , фаза которой равна азимутальному углу $\Phi = \phi$, а амплитуда R равна нулю в начале координат и выходит на постоянное значение на больших расстояниях, то получится как раз структура, представленная на рис. 1, а. Следует отметить, что топологический дефект в пространственно периодической структуре можно рассматривать как двумерный аналог краевой дислокации в кристаллических решетках. Роль дополнительного периода в твердых телах играет слой атомов, введенный в совершенный кристалл. Такие краевые дислокации постулировал в 20-х годах прошлого века Бюргерс для объяснения пластической деформации кристаллов. После изобретения электронного микроскопа краевые дислокации были действительно обнаружены в твердых телах.

2. Динамика топологических дефектов в гексагональной структуре, возникающей при термоконвекции

При превышении критической разности температур границ слоя жидкости, однородно подогреваемого снизу, режим, температуропроводности сменится режимом конвекции. Если верхняя граница свободная, а толщина слоя достаточно мала, то возникновение конвекции обусловле-

но зависимость коэффициента поверхностного натяжения от температуры (конвекция Марангони – Бенара). Типичным для конвекции такого рода является возникновение шестигранных ячеек, которые обусловлены синхронизацией трех мод:

$$u = R_1 \exp(i\vec{k}_1\vec{r} + i\Phi_1) + R_2 \exp(i\vec{k}_2\vec{r} + i\Phi_2) + R_3 \exp(i\vec{k}_3\vec{r} + i\Phi_3) + k.c.$$

Здесь $\vec{k}_1, \vec{k}_2, \vec{k}_3$ – волновые векторы, угол между которыми составляет 120° . Шестигранная решетка, которая возникает как результат развития неустойчивости из случайных возмущений, редко бывает совершенной. Нарушения порядка, которые наиболее часто наблюдаются в такой структуре, заключаются в том, что одна из ячеек имеет 5 соседей, а другая 7. Такое нарушение порядка называется пента-гепта-дефектом. Как показал эксперимент [1], пента-гепта-дефект при термоконвекции возникает из-за притяжения двух топологических зарядов, принадлежащих разным модам. Для того чтобы отобразить движение топологических зарядов при обработке изображений, подсчитывался параметр синхронизации p : $p = -\cos(\Phi_1 + \Phi_2 + \Phi_3)$. Эволюция структуры и поля параметра порядка при взаимодействии двух топологических дефектов показана на рис. 2.

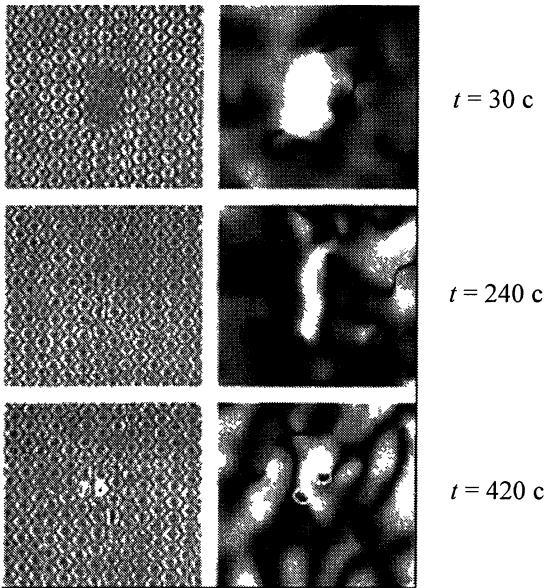


Рис. 2. Возникновение коридора при взаимодействии двух топологических дефектов в гексагональной структуре. Левая колонка – эволюция поля яркости изображения, правая – поля параметра синхронизации p . Цифрами 5 и 7 обозначены ячейки, имеющие 5 и 7 соседей, кружками – положение дефектов на плоскости параметра синхронизации

На начальном этапе взаимодействия зарядов формировалась область, где наблюдался большой градиент параметра p , так называемый коридор (ср. рис. 2, $t = 240$ с), а потом происходило сближение зарядов с возрастающей скоростью вдоль коридора и формировался пенто-гепто-дефект. Коридоры при различных начальных значениях разности фаз мод пред-

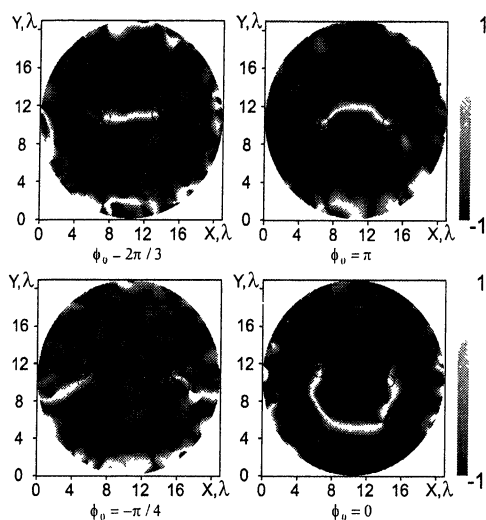


Рис. 3. Форма коридоров в поле параметра синхронизации p при различном сдвиге фаз ϕ_0 между модами, образующими гексагональную структуру

ставлены на рис. 3. Дефекты могли сближаться либо уходить к боковым границам. Образование пента-гепта-дефекта может быть изучено на основе модельных уравнений, подобных уравнению (1), описывающих взаимодействие трех резонансно связанных мод [2]. В численных расчетах было обнаружено существование коридора и движение топологических зарядов вдоль коридора. Так же, как в эксперименте, в численном расчете наблюдалось как сближение топологических дефектов с возрастающей скоростью и возникновение пента-гепта-дефекта, так и уход топологических дефектов к границам области интегрирования. В физическом эксперименте, тем не менее, наблюдались процессы взаимодействия дефектов, которые не описывались модельными уравнениями. Так, в численном моделировании количество топологических дефектов в модах, образующих шестигранную решетку, сохранялось. В эксперименте [1] вдоль коридора могли возникать новые пары дефектов с противоположными зарядами. Движение топологического дефекта в этом случае происходило как цепь рождений и аннигиляций дефектов, располагающихся в окрестности коридора [1]. Такой процесс модельными уравнениями не описывался.

Существование одного или нескольких топологических дефектов характерно для небольшой надкритичности ϵ . С увеличением ϵ в шестигранной решетке количество дефектов увеличивается. Простейшая характеристика степени нарушения порядка – это зависимость среднего количества дефектов от надкритичности. Такая характеристика конвекции Марангони – Бенара была получена в работе [3]. Процент ячеек с дефек-

тами N в [3] аппроксимировался формулой Аррениуса $N \sim \exp(-\frac{B}{\varepsilon})$, где B – эмпирическая константа. При малых надкритичностях происходит быстрый рост числа дефектов, а при больших надкритичностях, когда число дефектов становится сравнимым с числом ячеек в системе – наблюдается насыщение.

3. Связанные состояния дефектов

— в системе параметрически возбуждаемых капиллярных волн

Связанные состояния топологических дефектов могут наблюдаться и в решетках, образованных параметрически возбуждаемыми волнами на поверхности жидкости. При пространственно однородных вертикальных колебаниях вязкой жидкости с частотой ω_p на ее поверхности возникает тетрагональная структура, образованная двумя стоячими волнами с частотами $\omega = \omega_p/2$ и ортогональными волновыми векторами. Возникновение тетрагональной структуры характерно для параметрического возбуждения волн на поверхности вязкой жидкости в резонаторах с произвольной формой боковых границ, размеры которых существенно больше, чем длина волны возмущений [4]. Возникновение каждой из стоячих волн можно представлять как распад поля с частотой ω_p на две встречные волны с частотами $\omega_p/2$ и волновыми векторами $\pm k$:

$$\omega_p = \frac{\omega_p}{2} + \frac{\omega_p}{2}; \quad 0 = (k) + (-k).$$

Тетрагональная решетка с нарушением порядка показана на рис. 4, а. Если представить решетку как сумму двух стоячих волн (см. рисунки б и в), то можно понять, что такое нарушение порядка – связанное состояние, представляющее собой суперпозицию двух топологических зарядов, – существует лишь в одной из стоячих волн. Как показано в работе [4], это связанное состояние содержит топологические заряды, разнесенные вдоль направления распространения волн. Представить стоячую волну с дефектом можно в виде [5]

$$A_+ \cos(\frac{\omega_p}{2} t - kx + \Phi_+) + A_- \cos(\frac{\omega_p}{2} t + kx + \Phi_-),$$

где фазы Φ_{\pm} и амплитуды A_{\pm} имеют вид, аналогичный решению уравнения (1). Пояснить возникновение связанного состояния можно следующим образом. Если в волне с частотой $\omega_p/2$ и волновым числом k , распространяющейся вправо, возникло возмущения фазы, то благодаря связи через пространственно однородную накачку с частотой ω_p возмущения фазы возникают и в волне, распространяющейся влево с частотой $\omega_p/2$.

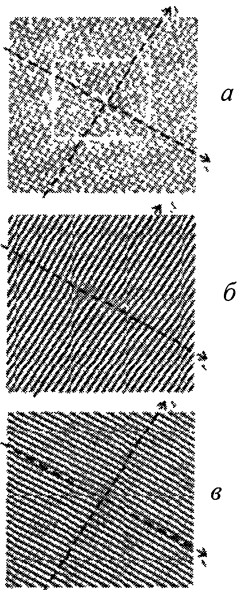


Рис. 4. Связанное состояние из двух топологических дефектов, принадлежащих одной стоячей волне *a* – поле поверхностных капиллярных волн, *б* – стоячая волна со связанными топологическими дефектами, *в* – стоячая волна без дефектов

Благодаря взаимодействию волн с накачкой такое связанное состояние обладает большим запасом устойчивости. Оно может двигаться как целое, причем расстояние между топологическими зарядами в таком связанном состоянии зависит от параметров возбуждаемых волн и надкритичности. Связанные состояния с противоположными топологическими зарядами могут аннигилировать, что приводит к возникновению совершенной структуры. Если связанные состояния имеют один и тот же заряд и принадлежат одной стоячей волне, то они могут образовывать долгоживущие доменные стенки, разделяющие области с различной ориентацией тетрагональных решеток [4]. Различные типы взаимодействия связанных состояний в параметрически возбуждаемых структурах изучались в работе [4].

При увеличении амплитуды колебаний жидкого слоя количество топологических дефектов в пространственно периодической структуре возрастает (на рис. 5 показано фазовое поле одной из стоячих волн для структуры с дефектами). Взаимодействие дефектов приводит к тому, что невозможно выделить отдельные пары дефектов, образующие связанное состояние, как это можно сделать для уединенного дефекта. Тем не менее, так же как и для термокапиллярной конвекции, пространственно-временной хаос поверхностных волн можно характеризовать зависимостью числа дефектов от надкритичности. Такая зависимость приведена на рис. 6. Наблюдается быстрое увеличение числа дефектов, которое может быть аппроксимировано функцией Аррениуса.



Рис. 5. Поле фазы одной из стоячих волн для режима пространственно-временного хаоса Кружочками отмечены топологические дефекты

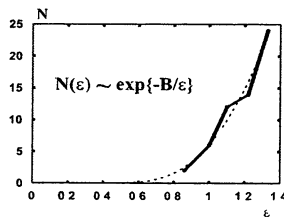


Рис. 6. Зависимость количества дефектов от надкритичности в параметрически возбуждаемых волнах

4. Дефекты в дорожке Кармана за нагретым цилиндром

В периодической дорожке вихрей, возникающей за нагретым цилиндром, обтекаемым воздушным потоком, возможно возникновение нарушений порядка [6], которое может быть сконструировано из топологических дефектов. В отсутствие нагрева за цилиндром возникает периодическая дорожка, свойства которой подробно исследуются на протяжении длительного времени. Нагрев цилиндра приводит к тому, что горячий воздух из пограничного слоя вовлекается в движение и приводит к возникновению изгибных колебаний вихрей, которые нарастают по амплитуде вниз по потоку. Увеличение амплитуды таких колебаний приводит к перезамыканию вихрей и возникновению дефектов, показанных на рис. 7. На временной реализации пульсаций температуры появлению дефекта соответствует уменьшение амплитуды колебаний до нуля и уменьшению

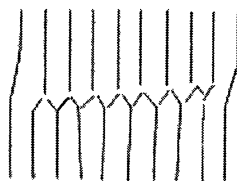
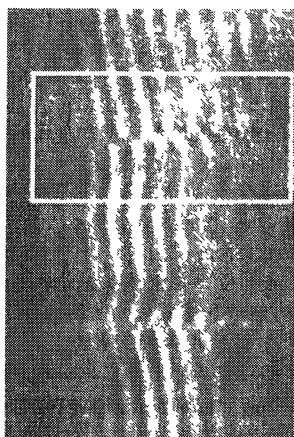


Рис. 7. Визуализация вихревой дорожки за нагретым цилиндром [6] Слева область дорожки выделена белым прямоугольником Эксперименты проводились при $Re = 61$ и разности температур цилиндра и воздушного потока $\Delta T = 150^\circ\text{C}$

фазы сигнала на 2π (рис. 8). Один пространственный период теряется, если в периодической решетке вихрей возникает два топологических дефекта с противоположными зарядами, разнесенными в продольном и поперечном направлениях. При увеличении мощности нагрева количество топологических дефектов увеличивается и, так же как для пента-гептадефектов при термоконвекции и дефектов в параметрически возбуждаемой ряби, пространственно-временной хаос в дорожке за нагретым цилиндром можно характеризовать зависимостью их числа от управляющего параметра. В данном случае – от мощности нагрева. Такая зависимость приведена на рис. 9.

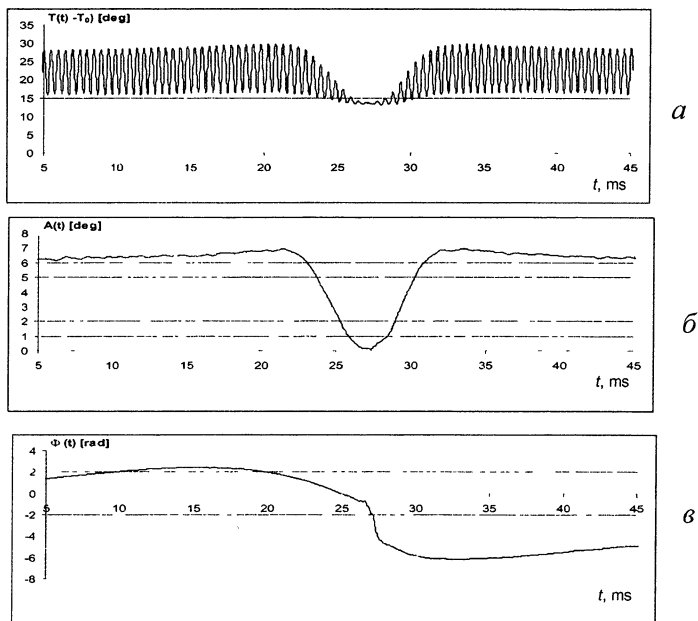


Рис. 8. Пульсации температуры в следе за цилиндром для режима возникновения топологических дефектов. *а* – временной ряд, *б* – амплитуда, *в* – фаза сигнала

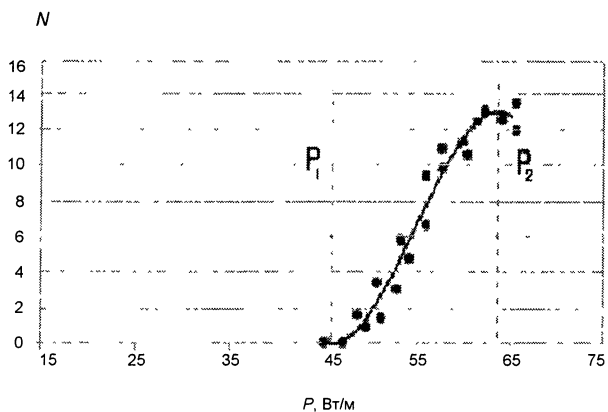


Рис. 9. Зависимость количества дефектов N , возникающих в дорожке за 1 секунду при $Re = 60$, от мощности нагрева P , рассчитанная по реализациям, содержащим 10^4 периодов P_1 – мощность, при которой наблюдается переход от регулярного срыва к дорожке с топологическими дефектами, P_2 – мощность, при которой происходит насыщение числа возникающих в дорожке дефектов

Заключение

Итак, показано, что топологические дефекты можно использовать как элементарные объекты для конструирования сложных пространственных паттернов. Динамика топологических дефектов в пространственно периодических структурах имеет много общего вне зависимости от того, в какой именно физической системе реализуется тот или иной паттерн. На этапе, когда из-за неустойчивости уже возникли пространственно периодические решетки, динамика отдельных вихрей или ячеек, а не динамика мод, из которых эта решетка образовалась, определяет поведение системы. При этом аналогия «топологический дефект – краевая дислокация» является, на наш взгляд, достаточно продуктивной. Так, возможно, например, проведение аналогии между возникновением дефектов в изогнутом кристалле и генерацией топологических дефектов при искривлении ее фронта [7].

Эти исследования проводились при частичной поддержке грантов РФФИ (проекты 99-02-16493, 02-02-17314), а также Президентской программы поддержки ведущих научных школ России.

Литература

- 1 *Afenchenko V O, Ezersky A B, Nazarovskiy A V, Velarde M G* Experimental evidence on the structure and evolution of penta-hepta-defect in hexagonal lattices due to Benard – Marangoni convection // *Int J Bif Chaos* 2001 **60**, № 5. P 1261–1273
- 2 *Rabinovich M I, Tzimring L S* Dynamics of dislocations in hexagonal patterns // *Phys Rev. E.* 1994 **49**, № 3. R35–R40
- 3 *Perez-Garcia C, Cerisier P., Ocellly R* Pattern selection in the Bernard-Marangoni instability Propagation in systems far from equilibrium / Ed by J. E. Wesfreid, H R Brand, P Manneville, N Boccacia. Berlin. Springer – Verlag, 1988
- 4 *Ezersky A B, Ermoshin D A, Kijshko S V* Dynamics of defects in parametrically excited capillary ripples // *Phys Rev. E.* 1995 **51**, № 4 P. 4411–4418
- 5 *Ezersky A B, Ermoshin D A, Kijshko S V* Bound states of topological defects in parametrically excited capillary ripples // *Physica D* 2001 **152-153** P 310–324
- 6 *Езерский А Б, Гариб М, Хаммаши М* Пространственно-временная структура следа за нагретым цилиндром // ПМТФ 1994, вып 5 С 74–81
- 7 *Ezersky A B, Kijshko S V, Nazarovskiy A V* The dynamics of bound states of topological defects in extended spatially periodic structures // *Proceedings of 6th Experimental Chaos Conference* Potsdam, 2001

НЕЛИНЕЙНЫЕ МОДЕЛИ ГЕНЕРАЦИИ ЦУНАМИ ДВИЖУЩИМИСЯ ИСТОЧНИКАМИ

Е. Н. Пелиновский

Введение

Под волнами цунами принято понимать поверхностные гравитационные волны, возникающие в море вследствие крупномасштабных непродолжительных воздействий (подводные землетрясения, извержения подводных вулканов, подводные оползни, сход лавин, падение в воду обломков скал и метеороидов, резкое изменение метеорологических условий). Характерная длительность цунами составляет 5–60 мин, длина 1–1000 км, высота может достигать десятков метров. Разнообразные сведения о цунами можно найти, например, в [1–5], а также на сайтах [6]. Последнее десятилетие XX века характеризовалось повышенной частотой появления цунами (рис. 1), в том числе и в России. Наиболее известное из них – цунами 5 октября 1994 г. высотой около 10 м, возникшее после землетрясения на Южных Курилах [7]. Многие из катастрофических цунами последних лет были обследованы международными экспедициями, что позволило унифицировать получаемые данные, необходимые для проверки моделей генерации, распространения и наката на берег волн цунами.

Теоретические модели волн цунами уже обсуждались на нелинейных школах [8] (см. также [1–3]). Здесь мы бы хотели обсудить новый взгляд на генерацию волн цунами, сформировавшийся, по существу, за последние 5 лет. Нелинейность в этих процессах является существенной, хотя во многих моделях она не так «выперта», спрятанная в численных моделях, учитывающих реальный рельеф дна и сложную структуру подвижного морского грунта. В традиционных моделях генерации цунами при подводных землетрясениях считается, что волны возбуждаются при мгновенном вертикальном смещении участка дна (не обязательно как целого, отдельные участки дна могут двигаться в разных направлениях) – так называемая поршневая модель генерации волн цунами. Сами смещения дна определяются как остаточные смещения при решении сейсмологической (упругой) модели развития землетрясения. Такая модель предсказывает параметры волн цунами как функции магнитуды (энергии) землетрясения, глубины фокуса и геометрических характеристик разлома (длины и ори-

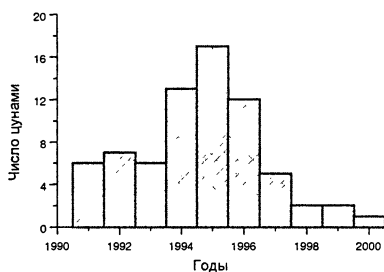


Рис. 1. Количество цунами в последнем десятилетии XX века

ентации). Однако уже давно было замечено, что иногда слабые землетрясения, которые согласно поршневой модели не должны вызывать заметных цунами, на самом деле их вызывают. Уже тогда была высказана гипотеза, что землетрясение играет роль триггера, приводящего в движение осадочные слои на наклонном морском дне (а большинство землетрясений происходит именно на границе континентального шельфа, на границе тектонических плит). Возбуждаемый при этом огромный оползень движется с большими скоростями и генерирует волны, в принципе по тому же сценарию, что и обычные корабельные волны.

Эта гипотеза так часто обсуждалась на совещаниях специалистов на протяжении 20 лет, что я даже не могу назвать имя автора этой гипотезы. Но только в последнее время появились прямые доказательства в пользу этого механизма, и он сразу стал очень популярен [9]. В частности, землетрясение, случившееся в Папуа – Новая Гвинея 17 июля 1998 года, привело к возникновению цунами высотой около 15 м, которое унесло более 2200 жизней. С энергетической точки зрения землетрясение было сравнительно слабым, его магнитуда (логарифм энергии) составила 7,1. Это значение близко к пороговому значению магнитуды цунамигенного землетрясения (7,0), при котором служба цунами должна объявлять тревогу. Сейсмическое и гидроакустическое зондирование, выполненное после землетрясения, вкупе с имеющимися данными до землетрясения, показали, что вблизи побережья в зоне очага произошел массивный подводный оползень, который и послужил причиной столь катастрофического цунами [10]. Через год, 17 августа 1999 года, подобное катастрофическое событие произошло в Турции, в Измитовском заливе Мраморного моря (магнитуда 7,4). Здесь часть берега (достаточно низкого) шириной около 200 м сошла в воду вместе со строениями, вызвав цунами высотой около 3 м [11]. Сходы оползней в воду и генерация при этом цунами были хорошо известны специалистам; достаточно напомнить, что в бухте Литуя (Аляска) в 1958 году сход лавины привел к заплеску воды на противоположный берег в 524 м и высоте волны цунами на выходе из бухты около 20 м. Однако число цунами, связанных с оползнями «в чистом виде», невелико в общей статистике цунами (не более 10 %). Теперь же, возможно, придется сделать вывод, что оползни могут быть главной причиной цунами и землетрясения (рассматриваемые как основной источник цунами) генерируют цунами через инициирование оползневых движений. Это подтвердил и статистический анализ связи сильных землетрясений с зонами осадконакопления в океане, показавший, что примерно 30 % цунами в той или иной степени связаны с оползнями и обвалами [12]. Дополнительным аргументом является обнаруженная корреляция цунами с приливом, который влияет на стабильность берегового склона [13]. Оползневой механизм генерации волн цунами сейчас является «горячей» точкой в пробле-

ме цунами; ему был посвящен специальный симпозиум, организованный НАТО в 2001 году [14].

В настоящей статье будут обсуждаться только нелинейные аспекты генерации волн цунами оползнями, которые могут быть полезны и для других разделов нелинейной физики.

1. Движение оползня

Прежде всего обсудим механику самого оползня, чтобы оценить его характерные параметры. Рассмотрим здесь простейшую ситуацию твердого недеформирующегося оползня, движущегося по наклонному дну (рис. 2).

Координату x направим вдоль наклонного дна (угол α к горизонту). В проекции на эту ось действуют следующие силы: скатывающая компонента силы тяжести (за минусом силы Архимеда), $F_g = (\rho - \rho_0)gV \sin \alpha$, сила трения $F_f = (\rho - \rho_0)\mu gV \cos \alpha$, гидравлическое сопротивление в турбулентном

потоке $F_h = \rho_0 C_D S U^2 / 2$ и волновое сопротивление F_w (вклад этой

компоненты оказался мал [15]). Здесь ρ и ρ_0 – плотности оползня и воды, $V = LS$ – объем оползня, L и S – его длина и поперечное сечение, μ – коэффициент кулоновского трения (приблизительно 0,005) и C_D – коэффициент лобового сопротивления (его величина порядка 0,002).

Динамика оползня в такой упрощенной постановке описывается уравнением Ньютона:

$$\rho V \frac{dU}{dt} = (\rho - \rho_0)gV(\sin \alpha - \mu \cos \alpha) - \frac{\rho_0}{2} C_D S U^2. \quad (1)$$

Решения этого уравнения тривиальны; если наклон дна мал, то независимо от начальной скорости оползень будет тормозиться. Такая ситуация может реализоваться при падении скал, имеющих большую начальную скорость при вхождении в воду. Если же начальная скорость отсутствует, то для инициации движения необходимо иметь либо увеличение наклона дна (как на континентальном склоне), либо разрушение стабильности оползня (приводящего к уменьшению кулоновского трения), либо и то и другое. Землетрясение как раз и играет роль триггера,

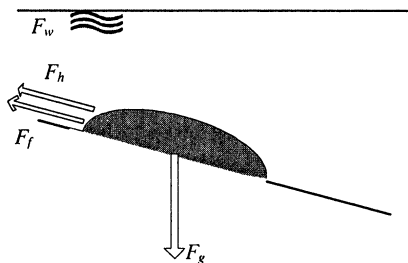


Рис. 2. Простейшая геометрия движущегося оползня и действующие на него силы

вызывая уменьшение сцепления между оползневой массой и твердым грунтом. Скорость движения оползня, как и его положение в пространстве, в этом случае находится в явном виде [15, 16]:

$$U(t) = U_{\infty} \tanh\left(\frac{\rho_0 C_D U_{\infty} t}{2\rho L}\right), \quad (2)$$

$$x(t) = x_0 + \frac{2\rho L}{\rho_0 C_D} \ln \cosh\left(\frac{\rho_0 C_D U_{\infty} t}{2\rho L}\right), \quad (3)$$

где установившаяся скорость есть

$$U_{\infty} = \sqrt{\frac{\rho - \rho_0}{\rho_0} \frac{2gL}{C_D} |\sin \alpha - \mu \cdot \cos \alpha|}. \quad (4)$$

Принимая горизонтальный размер оползня в 10 км и уклон дна в 10^{-2} , скорость оползня может достигать значений в сотни метров в секунду. Такая скорость сопоставима с максимальной скоростью распространения волн в океане (около 200 м/с), и, следовательно, возможно резонансное возбуждение значительных волн цунами. Так, оползень, образовавшийся при проведении строительных работ при расширении территории аэропорта г. Ниццы (Франция) 16 октября 1976 года достиг скорости 10 м/с (измерено по времени разрушения телефонных кабелей, проложенных по дну Средиземного моря) и привел к генерации цунами высотой 3–4 м, погубив 6 человек [17]. Примерно 8000 лет назад гигантский оползень в Норвегии, известный в литературе как Storrega Slide, достиг скорости в 50 м/с, вызвав цунами в Гренландии, Исландии и Шотландии до 5 м (это событие сейчас хорошо документировано геологическими данными) [16]. И наконец, нижегородский пример речного цунами: «в 1597 году, 18 июня, разразился над Нижним Новгородом гнев Божий: оторвавшись от матерой земли, оползла гора в Печерском монастыре и с растущим на ней лесом двинулась в Волгу сажен на 50, а местами и более, вследствие чего на реке поднялись бугры великие: стоявшие на воде под монастырем суда выбросило на берег сажен в 20 от воды и более» [18]. Причиной оползня стали дожди, нарушившие устойчивость волжского откоса.

2. Вынужденное уравнение Кортевега – де Вриза

Простейшей физической моделью резонансного возбуждения волн цунами может служить так называемое вынужденное уравнение Кортевега – де Вриза:

$$\frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} = \frac{\partial f}{\partial x}, \quad (5)$$

где η – смещение водной поверхности; $c = (gh)^{1/2}$ – скорость длинных волн на воде; $\alpha = 3c/2h$ – нелинейный и $\beta = ch^2/6$ – дисперсионный коэффициенты; $f = -cz_b/2$ – внешняя сила, связанная с движением оползня по дну с высотой $z_b(x, t)$; h – глубина океана, предполагаемая постоянной и g – ускорение силы тяжести. Отметим, что уравнение (5) может быть получено для волн цунами, генерируемых крупномасштабными изменениями атмосферного давления, так называемых метеоцунами [3, 19, 20]. Хотя уравнение (5) может быть выведено строго, с использованием асимптотической процедуры при малости нелинейности, дисперсии и возбуждения [3, 21, 22], его можно написать и сразу в силу аддитивности малых параметров в первом приближении. Мы будем предполагать, что оползень движется с постоянной скоростью V , которая близка к максимальной скорости длинных волн c (именно в этом случае и может быть получено одноволновое эволюционное уравнение), при этом движение оползня может быть как сверхкритическим ($V > c$), так и докритическим ($V < c$). Рассмотрим сначала несколько простейших решений, отвечающих различным предельным случаям.

2.1. Линейное возбуждение длинных волн

Если нелинейностью и дисперсией можно пренебречь, уравнение (5) решается тривиально, и с учетом нулевых начальных условий оно имеет вид

$$\eta(x, t) = \frac{f(x - Vt) - f(x - ct)}{c - V}. \quad (6)$$

В отсутствие резонанса ($V \neq c$) возбуждаются две волны: одна вынужденная, бегущая со скоростью источника, и одна свободная (в рамках исходных уравнений имеется также свободная волна, бегущая в противоположную сторону, однако в окрестности резонанса ее амплитуда достаточно мала [23]). Форма обеих волн неизменна (в отсутствие дисперсии), и их амплитуда (противоположная по знаку) значительна в окрестности резонанса (существенно превышая высоту оползня). Волна излучается назад при сверхкритическом движении источника и вперед, если источник докритический. При резонансе ($V = c$) волновое поле не является стационарным и возрастает во времени:

$$\eta(x, t) = t \frac{\partial f(x - ct)}{\partial x}. \quad (7)$$

Уже эти простейшие решения демонстрируют важность эффектов нелинейности, дисперсии и нестационарности при генерации волн.

2.2. Эффекты дисперсии при линейном возбуждении волн

Учет дисперсии в уравнении (5) в рамках линейного приближения приводит к решению в интегральной форме:

$$\eta(x, t) = \int \frac{A(k)}{c_{ph}(k) - V} \left(e^{ik(x-Vt)} - e^{ik(x-c_{ph}t)} \right) dk, \quad (8)$$

где $c_{ph} = c - \beta k^2$ – фазовая скорость волн на воде в окрестности длинноволнового предела, $A(k)$ – комплексная спектральная амплитуда. Если движение оползня сверхкритическое ($V > c$), то излучение назад (свободное решение), сформированное в начальный момент, будет расплываться со временем, трансформируясь в затухающую функцию Эйри на больших расстояниях от источника. Над оползнем будет повышение уровня воды (вынужденное решение) гладкой формы (в недисперсионном приближении оно повторяет форму смещения дна), оно и «выживает» на больших временах. Если же источник движется с докритической скоростью, то возможен резонанс с какой-нибудь спектральной компонентой и излучение позади области источника представляет собой незатухающую монохроматическую волну вида

$$\eta(x, t) = \frac{A(k_*)}{\partial c_{ph} / \partial k(k_*)} e^{ik_*(x-Vt)}, \quad (9)$$

где k_* находится из резонансного условия: $c_{ph}(k_*) = V$. Если оползень движется почти в резонансе с максимальной скоростью поверхностных волн ($V \sim c$), то знаменатель в (9) мал (он равен $-2\beta k_*$) и амплитуда излучения достаточно велика, так что необходимо принимать во внимание нелинейные эффекты.

2.3. Нелинейная генерация длинных волн

Пренебрежем теперь дисперсионным слагаемым в (5) и рассмотрим стационарную генерацию нелинейных волн (гидравлическое приближение). В этом случае (5) интегрируется и превращается в алгебраическое уравнение

$$(c - V)\eta + (\alpha\eta^2)/2 = f, \quad (10)$$

причем возникающая при этом константа может быть сразу положена нулю в силу условия Зоммерфельда. В безразмерных переменных ($z = z_b/h$, $\zeta = \eta/h$, $Fr = V/c$) соотношение (10) имеет вид

$$(Fr - 1)\zeta - 3\zeta^2/4 = z/2 \quad (11)$$

и представляет собой параболу (рис. 3), ограниченную по величине вынуждающей силы (высоте оползня) сверху значением

$$z_{\max} = 2(Fr - 1)^2/3. \quad (12)$$

Характер волнового поля зависит от высоты оползня. Если она достаточно мала (меньше z_{\max} при фиксированной скорости или числе Фруда), то волновое поле всюду остается гладким и его профиль находится из квадратного уравнения (11). Другая ситуация будет реализовываться для сильных возмущений. На краю оползня волновое поле опять удовлетворяет стационарному уравнению (11) и находится в явном виде (при этом мы движемся по нижней ветке параболы на рис. 3). В точках, где высота оползня начинает превышать максимальное значение (12), нет решений алгебраического уравнения (11). В этой области волновое поле перестает быть стационарным. На противоположном конце оползня, где его высота опять становится меньше максимального значения (12), волновое поле может быть стационарным, но его амплитуда может стремиться как к нулю (нижняя ветка параболы на рис. 3), так и к фиксированному значению $\zeta_{\max} = 4(\text{Fr} - 1)/3$ (верхняя ветка параболы). Выбор между этими решениями в рамках стационарного приближения не может быть сделан.

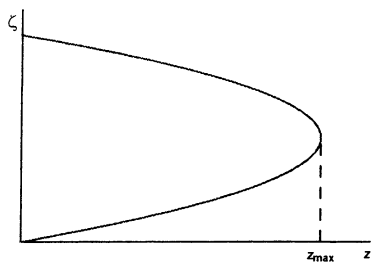


Рис. 3. Связь между смещением водной поверхности и высотой оползня

В центральной части достаточно большого оползня, когда $z > z_{\max}$, волновое поле должно содержать мелкомасштабные компоненты или быть нестационарным. Нестационарность, конечно же, проявляется и для более малых высот оползня. Она приводит к формированию ударных волн и, следовательно, опять к появлению областей с мелкомасштабными неоднородностями волнового поля.

2.4. Стационарные решения
вынужденного уравнения Кортевега – де Вриза

Рассмотрим теперь стационарные решения уравнения (5), отвечающие генерации вынужденных волн. Они находятся из обыкновенного дифференциального уравнения

$$\beta \frac{d^2 \eta}{dx^2} - (V - c)\eta + \frac{\alpha \eta^2}{2} = f, \tag{13}$$

причем опять константу интегрирования можно положить равной нулю. Особенно простые решения [22, 24] получаются в случае короткой внешней силы, когда возможна ее аппроксимация дельта-функцией:

$$f = B\delta(x/l), \tag{14}$$

где B и l – амплитуда и ширина вынуждающей силы (для оползня $B < 0$). Вне зоны источника решение находится из «свободного» уравнения Кортевега – де Вриза, где, однако, скорость распространения фиксирована и равна скорости движения источника. В случае сверхкритического источника таким решением является солитон единственной амплитуды:

$$\eta(x, t) = A \operatorname{sech}^2 \left[\sqrt{\frac{\alpha A}{12\beta}} (x - X - Vt) \right], \quad \frac{\alpha A}{3} = V - c. \quad (15)$$

Положение солитона, координата X , может быть любым относительно источника.

Эти два солитонных решения одинаковой амплитуды, расположенные по обе стороны от источника, должны быть сшиты с помощью граничных условий, получаемых из (13):

$$\eta|_- = \eta|_+, \quad \left. \frac{d\eta}{dx} \right|_+ - \left. \frac{d\eta}{dx} \right|_- = \frac{Bl}{\beta}. \quad (16)$$

Для оползня $B < 0$, поэтому единственной квазисолитонной формой волны, привязанной к источнику, является односторонняя симметричная кривая, изображенная на рис. 4. Если высота оползня достаточно мала, то

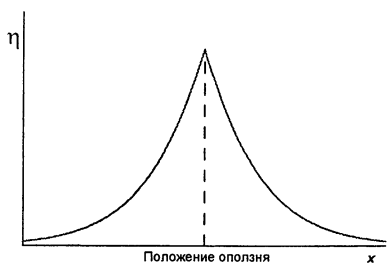


Рис. 4. Стационарное квазисолитонное возмущение над подводным оползнем

профиль волны близок к точному солитону. Однако это решение не является единственным. В принципе, после прохождения источника волновое поле может не возвращаться к нулевому значению, а стремиться к некоторому пьедесталу. Пьедестал эквивалентен другому значению глубины бассейна, изменяющему величину скорости распространения длинных волн, $c = (gh)^{1/2}$. Если глубина увеличилась, то за источником

может быть реализован докритический режим (в отличие от области перед источником). В этом случае решение (13) соответствует уже не единственной, а периодической волне, так называемой кноидальной (не будем приводить здесь громоздких выражений для этой волны). Таким образом, движение оползня может приводить к генерации солитона на его фронте и кноидальной волны за ним.

Разнообразные примеры стационарных волновых полей приводятся в [22]. Однако проблема их устойчивости и возникновения из нестационарных возмущений остается открытой в рамках стационарного приближения. Ниже мы дадим один из примеров такой генерации.

3. Взаимодействие солитона с движущимся источником

Предположим, что солитон на воде уже сформировался и начинает взаимодействовать с движущимся источником, амплитуда которого считается малой. В этом случае естественно предположить, что солитон локально остается солитоном, а его параметры (амплитуда и положение) меняются из-за взаимодействия. Аналитически солитон описывается формулой (15) уже без ограничения на его амплитуду и скорость. Это типичная задача на метод возмущений для солитонов, и в первом приближении параметры солитона находятся из уравнения энергетического баланса и скорости распространения (в системе отсчета, связанной с источником) [25–28]:

$$\frac{d}{dt} \int_{-\infty}^{\infty} \frac{\eta^2}{2} dx = \int_{-\infty}^{\infty} \eta \frac{\partial f}{\partial x} dx, \quad \frac{dX}{dt} = c - V + \frac{\alpha A}{3}. \quad (17)$$

Для дельтаобразного источника (14) интеграл (17) может быть вычислен точно, и в переменных

$$\gamma = \sqrt{(\alpha A)/12\beta}, \quad \theta = \gamma X, \quad (18)$$

уравнения (17) принимают вид

$$\frac{d\gamma}{dt} = -\frac{\alpha B l}{12\beta} \gamma \frac{\tanh \theta}{\cosh^2 \theta}, \quad \frac{d\theta}{dt} = \gamma(c - V + 4\beta\gamma^2), \quad (19)$$

где мы пренебрегли малым членом, пропорциональным амплитуде источника в правой части второго уравнения. Интегральные кривые этой динамической системы находятся явно:

$$\frac{\alpha B l}{24\beta} \operatorname{sech}^2 \theta = (c - V)\gamma + \frac{4\beta}{3}\gamma^3 + \text{const}. \quad (20)$$

Если ввести безразмерные параметры

$$4\beta\gamma_0^2 = |c - V|, \quad G = \gamma/\gamma_0, \quad Q = (\alpha B l)/(96\beta^2\gamma_0^3), \quad (21)$$

кривые (20) будут содержать только один параметр (не считая константы интегрирования):

$$Q \operatorname{sech}^2 \theta = -G \operatorname{sign}(V - c) + \frac{G^3}{3} + \text{const}, \quad (22)$$

при этом амплитуда солитона есть $A \sim G^2$.

Фазовые портреты динамической системы (19) в этих переменных показаны на рис. 5 как для докритического, так и сверхкритического режимов (принято, что $Q < 0$, что соответствует движению оползня). Если источник движется медленно (докритический режим), то солитон не может находиться в резонансе с ним (солитон всегда «быстрый») и обгоняет его (в этом случае отсутствуют состояния равновесия на фазовой плоскости, рис. 5, а). Взаимодействие приводит только к обратимому изменению

амплитуды достаточно энергетического солитона. Ниже гомоклинической траектории на рис. 5, *a* (она касается оси абсцисс) возможна как генерация слабых солитонов (и они уходят вперед от источника), так и их поглощение (догоняющих источник).

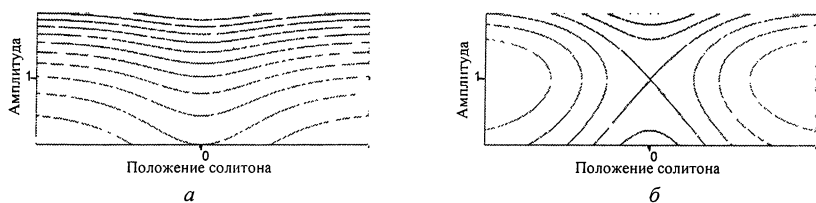


Рис. 5. Фазовая плоскость взаимодействия солитона с движущимся источником в докритическом (*a*) и сверхкритическом (*b*) режиме

В случае сверхкритического движения источника он может находиться в резонансе с солитоном определенной амплитуды и на фазовой плоскости (рис. 5, *б*) появляется состояние равновесия. Оно оказывается седлом, и, следовательно, резонансный солитон, по существу найденный в рамках стационарного приближения (рис. 4), является неустойчивым. В результате солитон убегает (или отстает) от источника. Взаимодействие приводит к необратимому изменению амплитуды солитона. Более быстрый солитон, догоняя источник, теряет энергию (а следовательно, и скорость) и отстает от него. Более медленный солитон, напротив, получает энергию от движущегося источника и ускоряется, убегая от источника. Максимальное изменение амплитуды солитона может достигать $3A_0$, где A_0 – амплитуда резонансного солитона – легко находится из (18) и (21). Солитоны большой амплитуды (верхние кривые на рис. 5, *б*) в силу большой скорости обгоняют источник, и в этом случае взаимодействие носит обратимый характер. Слабые солитоны, генерируемые на фронте источника, не успевают развиться и опять поглощаются источником (нижняя кривая вблизи положения источника). Более полная картина взаимодействия солитона с движущимся источником с учетом его нестационарности и диссипации волновой энергии описана в [25–28].

Итак, уже эта простейшая задача о взаимодействии солитона с движущимся источником показывает все особенности генерации нелинейных волн и, что особенно важно, нестационарность волнового поля. Применительно к проблеме цунами, как это видно из приведенных модельных задач, оползень может вызывать сильные волны как впереди его, так и позади, приводя к наводнениям на ближайших берегах.

Заключение

Мы привели ряд простейших теоретических задач в проблеме генерации волн цунами движущимися оползнями, демонстрирующих важ-

ность нелинейных эффектов. На практике, в связи с сильно возросшими вычислительными возможностями, такие упрощенные модели уже не используются, а применяются максимально полные модели волн на воде (теория мелкой воды, уравнения Буссинеска, полные по нелинейности, или исходные уравнения) и оползневого движения с учетом его реологии [29–36]. Результаты расчетов в целом находятся в удовлетворительном согласии с наблюдаемыми данными, и сейчас по существу определяется, какие же вычислительные модели являются наиболее адекватными и что они не могут описать.

Исследования проводились при частичной финансовой поддержке ИНТАС (99-1068 и 01-2156), РФФИ (02-05-65107) и программы РАН по нелинейным волнам.

Литература

- 1 Мурти Т Сейсмические морские волны – цунами Л Гидрометеиздат 1981 448 с
- 2 Пелиновский Е Н Нелинейная динамика волн цунами Горький ИПФ РАН, 1982 228 с
- 3 Пелиновский Е Н Гидродинамика волн цунами Нижний Новгород ИПФ РАН, 1996 276 с
- 4 Заякин Ю А Цунами на Дальнем Востоке России Петропавловск-Камчатский Камшат, 1996 88 с
- 5 Bryant E Tsunami Cambridge Univ Press. 2001 320 p
- 6 Лаборатория цунами ИВМиМГ СО РАН <http://tsun.sccc.ru> и Тихоокеанская морская исследовательская лаборатория НОАА <http://pml.noaa.gov>
- 7 Иващенко А И, Гусяков В К, Джумагалиев В А и др Шикотанское цунами 5 октября 1994 г // Докл РАН 1996. Т 348 С. 532–538
- 8 Пелиновский Е Н Волны цунами // Нелинейные волны М Наука, 1981 С 187–203
- 9 Keating B H, Waythomas C F, Dawson A G Landslides and tsunamis (Pageoph Topical Volumes, V 157) Berlin, Birkhauser Verlag, 2000. 1313 p
- 10 Tappin D R, Watts P, McMurty G. M., Lofoy Y, Matsumoto T The Sissano, Papua New Guinea tsunami of July 1998 – offshore evidence on the source mechanism // Marine Geology 2001 V. 175 P 1–23
- 11 Altinik Y, Tinti T, Alpar B, Yalciner A C, Ersoy S, Bortolucci E, Armigliato A The tsunami of August 17, 1999 in Izmit Bay, Turkey // Natural Hazards 2001 V 24 P 133–146
- 12 Gusiakov V K «Red», «green» and «blue» Pacific tsunamigenic earthquakes and their relation with conditions of oceanic sedimentation // Tsunamis at the end of a critical decade (Ed G Hebenstreit) Kluwer, Dordrecht, 2001 С 17–32
- 13 Куликов Е А, Рабинович А Б, Файн И В, Борнхолд Б Д, Томсон Р Е Генерация цунами оползнями на тихоокеанском побережье Северной Америки и роль приливов // Океанология 1998 Т 38 С 323–328
- 14 Yalciner A, Pelinovsky E, Synolakis K, Okal E NATO Advanced Research Workshop «Underwater ground failures on tsunami generation, modeling, risk and mitigation» Istanbul, 2001 264 p
- 15 Pelinovsky E, Poplavsky A Simplified model of tsunami generation by submarine landslides // Phys Chem Earth 1997 V 21 P 13–17
- 16 Harbitz C Model simulations of tsunami generated by the Storegga Slides // Marine Geology 1992 V 105 P 1–21

- 17 *Assier-Rzadkiewicz S, Heinrich P, Sabatier P C, Savoye B, Bourillet J F* Numerical modeling of a landslide-generated tsunami the 1979 Nice event // PAGEOPH 2000. V 157 P 1707–1728
- 18 *Гауцский А* Нижегородский летописец Нижегородская ярмарка 2001 717 с.
- 19 *Akylas T R* On excitation of long nonlinear water waves by a moving pressure distribution // J Fluid Mech 1984. V 141 P 455–466
- 20 *Pelinovsky E, Talipova T, Kurkin A, Kharif C* Nonlinear mechanism of tsunami wave generation by atmospheric disturbances // Natural Hazards and Earth System Sciences 2001 V 1 P 243–250
- 21 *Cole S L* Transient waves produced by flow past a bump // Wave Motion. 1985 V. 7. P 579–587
- 22 *Shen S S* A course on nonlinear waves. Dordrecht, Kluwer, 1993 188 p
- 23 *Tinti S, Bortolucci E, Chiavettieri C* Tsunami excitation by submarine slides in shallow-water approximation // PAGEOPH 2001 V 158 P 759–797
- 24 *Malomed B A* Interaction of a moving dipole with a soliton in the KdV equation // Physica D 1988 V 32 P 393–408
- 25 *Grimshaw R, Pelinovsky E, Tian X* Interaction of solitary wave with an external force // Physica D 1994 V 77 P 405–433
- 26 *Гримишоу Р, Пелиновский Е Н* Взаимодействие уединенных поверхностных и внутренних волн с движущимся возмущением // Докл РАН 1995 Т 344 С 394–396.
- 27 *Grimshaw R, Pelinovsky E, Sakov P* Interaction of a solitary wave with an external force moving with variable speed // Stud Applied Math 1996 V 97 P 235–276
- 28 *Grimshaw R, Pelinovsky E, Bezen A* Hysteresis phenomena in the interaction of a damped solitary wave with an external force // Wave Motion 1997 V 26 P 253–274
- 29 *Assier-Rzadkiewicz S, Mariotti C., Heinrich P* Numerical simulation of submarine landslides and their hydraulic effects // J Waterway, Port, Coastal, and Ocean Eng 1997 V 124. P 149–157
- 30 *Imamura F, Gica E C* Numerical model for tsunami generation due to subaqueous landslide along a coast // Science Tsunami Hazards 1996 V 14 P 13–28
- 31 *Jiang L, LeBlond P H* The coupling of a submarine slide and the surface waves which it generates // J Geophys Research 1992 V. 97(C8) P 12,731–12,744
- 32 *Johnsgard H, Pedersen G* Slide-generated waves in near-shore regions A Lagrangian description // Phys. Chem Earth 1996 V 21 P 45–49
- 33 *Madsen P A, Schaffer H A* High-order Boussinesq-type equations for surface gravity waves – derivation and analysis // Phil Trans Royal Soc London 1998. V. 356 P 1–59
- 34 *Tanioka Y, Satake K* Tsunami generation by horizontal displacement of ocean bottom // Geophys Res Letters 1996 V 23 P. 861–864.
- 35 *Watts P* Wavemaker curves for tsunami generated by underwater landslides // J. Waterway, Port, Coastal, and Ocean Eng 1998 V 124 P 127–137
- 36 *Wei G, Kirby J. T, Grilli S, Subramanya R* A fully nonlinear Boussinesq model for surface waves 1 Highly nonlinear unsteady waves // J. Fluid Mech 1995. V 294 P. 71–92.

НОВЫЙ ПОДХОД К ГИДРОДИНАМИКЕ НЕСЖИМАЕМОЙ ЖИДКОСТИ

Е. И. Якубович

1. Идеи, которые будут изложены в настоящей работе, рассредоточены по журнальным статьям, а некоторые из них еще нигде не опубликованы.

План изложения будет таков. После микроскопического ликбеза по лагранжеву описанию динамики жидкости расскажем о нашем взгляде на описание течения жидкости, затем о двух формах уравнений гидродинамики, которые из него следуют. Одна из этих форм теперь уже лишь относительно новая, а другая – совсем новая. Далее будут обозначены некоторые пути исследований. Последний раздел будет посвящен одной из центральных проблем гидродинамики: как учесть вязкость при лагранжевом описании течения жидкости; точнее, как наш подход к описанию течения жидкости позволяет учесть ее вязкость.

2. Как хорошо известно, существуют два способа описания течения жидкости – Эйлера и Лагранжа. Оба способа сформулировал Эйлер в 1755 году (за 33 года до Лагранжа) в трактате «Общие принципы движения жидкостей». Применение первого способа до сих пор по частоте намного опережает второй. Нам представляется, что единственной существенной причиной для этого может быть только трудность учета вязкости при лагранжевом подходе.

Сущность обоих подходов, как известно, такова: при эйлеровом подходе жидкость рассматривают как поле скоростей в пространстве переменных X, Y, Z, T ; а при лагранжевом следят за каждой частицей жидкости, предварительно их «пометив» (перенумеровав). Таким образом, неизвестными функциями будут X, Y, Z , а независимыми переменными – «номера» частиц и время. Выбор этих меток в значительной степени произволен, необходимо лишь следить за однозначностью и чтобы не было частиц без имени. За такие метки принимают три произвольные постоянные a, b, c из решения уравнения Ньютона для частицы. Можно, например, выбрать за лагранжевы переменные a, b, c начальные координаты частиц жидкости. Тогда множество всех частиц, имеющих координаты $b = 0, c = 0$ будет расположено вдоль оси изменения координаты a (при $t = 0$ она совпадает с осью X). С течением времени эта жидкая ось будет искривляться и перемещаться в пространстве. То же самое можно сказать об осях b и c . Эти оси будем считать осями «жидкой» системы координат. Эта жидкая система координат, вообще говоря, криволинейна, неортогональна и нестационарна.

Чтобы получить уравнения Лагранжа для частиц жидкости, надо прервать уравнения Ньютона для частицы, куда входят производные по неизвестным функциям X, Y, Z , в уравнения, куда вместо них войдут про-

изводные по переменным a, b, c, t . Для этого уравнение Ньютона $\mathbf{r}_{tt} = -\nabla p$ умножим на вектор касательный к одной из жидких координатных кривых \mathbf{r}_a : $(\mathbf{r}_{tt}\mathbf{r}_a) = -(\nabla p\mathbf{r}_a) = -p_a$. Это и есть одно из трех уравнений Лагранжа. В развернутой форме они имеют вид [1]

$$\begin{aligned} X_{tt}X_a + Y_{tt}Y_a + Z_{tt}Z_a &= -p_a, \\ X_{tt}X_b + Y_{tt}Y_b + Z_{tt}Z_b &= -p_b, \\ X_{tt}X_c + Y_{tt}Y_c + Z_{tt}Z_c &= -p_c. \end{aligned} \quad (1)$$

Кроме них, есть еще уравнение непрерывности (сохранение объема):

$$\frac{\partial}{\partial t} \det \begin{pmatrix} X_a & X_b & X_c \\ Y_a & Y_b & Y_c \\ Z_a & Z_b & Z_c \end{pmatrix} = 0. \quad (2)$$

Для переменных Лагранжа, совпадающих с начальными координатами частиц, этот детерминант тождественно равен единице.

Можно получить другую форму этих уравнений, которая давно известна, но редко приводится в учебниках. Это аналог уравнения Гельмгольца для завихренности. Возьмем перекрестные производные по лагранжевым переменным от пары уравнений Лагранжа (сравните взятие ротора от уравнений Эйлера при получении уравнений Гельмгольца) и вычтем одно из другого. Правые части с давлениями пропадут, а левые станут производными по времени. После интегрирования по t получим три уравнения [1]:

$$\begin{aligned} X_{tb}X_c - X_{tc}X_b + Y_{tb}Y_c - Y_{tc}Y_b + Z_{tb}Z_c - Z_{tc}Z_b &= S_1(a, b, c), \\ X_{tc}X_a - X_{ta}X_c + Y_{tc}Y_a - Y_{ta}Y_c + Z_{tc}Z_a - Z_{ta}Z_c &= S_2(a, b, c), \\ X_{ta}X_b - X_{tb}X_a + Y_{ta}Y_b - Y_{tb}Y_a + Z_{ta}Z_b - Z_{tb}Z_a &= S_3(a, b, c). \end{aligned} \quad (3)$$

Можно показать, что три постоянные, в правых частях этих соотношений – это компоненты завихренности в «жидкой» системе координат [2, 3]:

$$\boldsymbol{\Omega} = S_1\mathbf{r}_a + S_2\mathbf{r}_b + S_3\mathbf{r}_c. \quad (4)$$

Замечательно, что удалось избавиться от давления без повышения порядка уравнений.

Представленная система уравнений предлагалась для исследования течений жидкости в переменных Лагранжа [1]. Она не лучше и не хуже уравнений Эйлера. Просто, как обычно, одни проблемы выглядят проще при одном подходе, а другие – при другом. Кроме того, возможно появление совсем новых проблем. В этом мы видим смысл поиска различных подходов.

3. В чем сущность нашего подхода? Давайте наблюдать за относительным перемещением частиц-соседей. Пусть имеется бесконечно малый материальный элемент жидкости $d\mathbf{r} = \{dX, dY, dZ\}$, соединяющий две

частицы с бесконечно малым приращением лагранжевых переменных $da = \{da, db, dc\}$. Будем рассматривать течение жидкости как непрерывную деформацию этого жидкого элемента. Она наглядно описывается матрицей Якоби с элементами – производными текущих координат частиц по лагранжевым переменным

$$\hat{R} = \begin{pmatrix} X_a & X_b & X_c \\ Y_a & Y_b & Y_c \\ Z_a & Z_b & Z_c \end{pmatrix} \quad (5)$$

и связывает приращения

$$d\mathbf{X} = \hat{R}da. \quad (6)$$

По способу введения матрицы \hat{R} она похожа на тензор деформации в теории упругости. Однако есть существенное отличие: в теории упругости тензор деформации описывает изменение *длины* бесконечно малого элемента [4], а у нас матрица описывает изменение *векторного* бесконечно малого элемента, его деформацию и, что очень важно, вращение. Конечно, зная \hat{R} , легко получить координаты частиц жидкости.

Наша цель – показать, что матрица Якоби может выступать как фундаментальная величина, которой вполне достаточно для описания движения жидкости – подобно скорости при эйлеровом описании или координатам частицы X, Y, Z при лагранжевом. Прежде всего, \hat{R} непосредственно входит в уравнение непрерывности:

$$\det \hat{R} = \det \hat{R}_0, \quad \hat{R}_0 = \hat{R}|_{t=0}. \quad (7)$$

Удивительное дело, но система уравнений (3) сворачивается в одно матричное уравнение для \hat{R} [5]

$$\hat{R}_i^T \hat{R} - \hat{R}^T \hat{R}_i = \hat{S}, \quad \hat{S} = \begin{pmatrix} 0 & S_3 & -S_2 \\ -S_3 & 0 & S_1 \\ S_2 & -S_1 & 0 \end{pmatrix}. \quad (8)$$

Здесь верхний индекс T означает транспонирование матрицы. К этим уравнениям следует добавить условие того, что матрица \hat{R} не произвольна, а связана с координатами частиц согласно своему определению (5). Из последнего, очевидно, следует, что перекрестные производные столбцов матрицы \hat{R} должны равняться друг другу:

$$\partial R_{ik} / \partial a_n = \partial R_{in} / \partial a_k. \quad (9)$$

(на это ограничение области решений уравнения (8) указал Зенькович [5]). Иногда мы называем его условием градиентности, иногда – условием совместности. Надо сказать, что в теории упругости тоже есть похожее условие, оно там характеризует отсутствие дислокаций. Заметим, что полу-

ченные уравнения не содержат других неизвестных функций, кроме матрицы Якоби. Отметим ряд важных свойств полученных уравнений.

1) Они однородны, т. е. \hat{R} всюду входит в одинаковых степенях. Это позволяет надеяться на получение в будущем автомодельных решений.

2) Уравнение для завихренности содержит только производные по времени, уравнение совместности – производные только по лагранжевым переменным, а уравнение непрерывности вообще алгебраическое.

3) Можно привести тот замечательный факт, что несмотря на нелинейность, система не меняет своего вида при введении комплексных величин:

$$\begin{aligned} W_1 &= \frac{1}{\sqrt{2}}(X + iY), & W_2 &= \frac{1}{\sqrt{2}}(X - iY), & W_3 &= Z, \\ \xi_1 &= \frac{1}{\sqrt{2}}(a + ib), & \xi_2 &= \frac{1}{\sqrt{2}}(a - ib), & \xi_3 &= c. \end{aligned} \quad (10)$$

Вместо \hat{R} теперь возникнет матрица

$$L_{ik} = \partial W_i / \partial \xi_k. \quad (11)$$

Необходимо только вместо операции транспонирования применять эрмитово сопряжение и антисимметричную матрицу \hat{S} заменить на комплексную матрицу:

$$\hat{M} = \begin{pmatrix} -iS_3 & 0 & \frac{1}{\sqrt{2}}(-S_2 + iS_1) \\ 0 & iS_3 & \frac{1}{\sqrt{2}}(-S_2 - iS_1) \\ \frac{1}{\sqrt{2}}(S_2 + iS_1) & \frac{1}{\sqrt{2}}(S_2 - iS_1) & 0 \end{pmatrix}. \quad (12)$$

В этих обозначениях уравнение движения будет выглядеть как

$$\hat{L}_i^* \hat{L} - \hat{L}^* \hat{L}_i = \hat{M}, \quad (13)$$

где верхний индекс звездочка означает эрмитово сопряжение. Уравнение непрерывности и условие градиентности также будут иметь прежний вид:

$$\begin{aligned} \det \hat{L} &= \det \hat{L}_0, \\ \partial L_{ik} / \partial \xi_n &= \partial L_{in} / \partial \xi_k. \end{aligned} \quad (14)$$

4. Одним из основных достоинств матричного описания, на наш взгляд, является его компактный, «блочный» характер. Вспомним, что связка двух переменных – комплексное описание – привела в физике ко многим результатам. Связка трех переменных – векторное описание – также облегчило решение многих задач. Преимущество связки девяти переменных – матричного описания течения жидкости – проиллюстрируем следующим примером.

Найдем решение для матрицы \hat{L} в виде произведения трех матриц:

$$\hat{L} = \hat{U}(t) \hat{L}_0 \hat{V}(t), \quad (15)$$

где крайние матрицы зависят только от времени, а средняя – только от координат. Пример сильно упростится, если матрицы \hat{U} и \hat{V} выбрать диагональными вида

$$\hat{U} = \begin{pmatrix} e^{-i\gamma t} & 0 & 0 \\ 0 & e^{i\gamma t} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{V} = \begin{pmatrix} e^{-i\omega t} & 0 & 0 \\ 0 & e^{i\omega t} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (16)$$

Уравнение непрерывности (14) при этом удовлетворится тождественно. Условие градиентности выполнится автоматически, если предположить следующий частный вид средней матрицы:

$$\hat{L}_0 = \begin{pmatrix} L_{11}(\xi_1) & L_{12}(\xi_2) & L_{13}(c) \\ L_{21}(\xi_1) & L_{22}(\xi_2) & L_{23}(c) \\ L_{31}(\xi_1) & L_{32}(\xi_2) & L_{33}(c) \end{pmatrix}. \quad (17)$$

Теперь нам осталось лишь конкретизировать вид этих функций $L_{ik}(\xi_r)$, чтобы удовлетворить уравнению движения (13). Тут довольно большой выбор возможностей. Простой вариант – выбрать среднюю матрицу вида

$$\hat{L}_0 = \begin{pmatrix} L_{11}(\xi_1) & \bar{L}_{21}(\xi_1) & 0 \\ L_{21}(\xi_1) & \bar{L}_{11}(\xi_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (18)$$

Здесь элементы L_{11} и L_{12} – две произвольные аналитические функции. Черта сверху над функцией означает ее комплексное сопряжение. Результирующая матрица L удовлетворит уравнению движения и всем остальным условиям. Она описывает известный класс двумерных течений – так называемых птолемеевских течений [6]. Траектории частиц жидкости для такого течения описываются комплексной функцией $W_1 = \frac{1}{\sqrt{2}}(X + iY) = \int L_{11}(\xi_1) d\xi_1 e^{-i(\omega+\gamma)t} + \int \bar{L}_{21}(\xi_1) d\bar{\xi}_1 e^{i(\omega-\gamma)t}$, представляют собой сумму двух круговых вращений с различными частотами и имеют вид эпи- или гипоциклоид. Оказывается, ряд известных точных решений лагранжевой гидродинамики, таких, как волны Герстнера или течение внутри вихря Кирхгофа [1], относятся к частным случаям этого семейства решений. С помощью решений этого типа удалось найти аналитическое описание широкого класса локализованных нестационарных двумерных вихрей в потенциальном потоке. С помощью других известных способов описания течений решать подобные задачи не представлялось возможным. В то же время простота получения этих решений при нашем подходе демонстрирует его перспективность.

Трансформация двумерных решений птолемеевского типа в трехмерные получится добавлением в матрице \hat{L}_0 вместо нулей специально по-

добранных функций. Например, одно из таких обобщений получается, если нули в третьем столбце и третьей строке заменить:

$$\hat{L}_0 = \begin{pmatrix} L_{11}(\xi_1) & \bar{L}_{21}(\xi_1) & -\omega \\ L_{21}(\xi_1) & \bar{L}_{11}(\xi_1) & -\omega \\ (\omega + 2\gamma)L_{11} + (\omega - 2\gamma)L_{21} & (\omega + 2\gamma)\bar{L}_{11} + (\omega - 2\gamma)\bar{L}_{21} & 1 \end{pmatrix}. \quad (19)$$

Результирующая матрица \hat{L} (15) удовлетворит всем уравнениям гидродинамики и опишет трехмерное вихревое нестационарное течение жидкости. Траектории частиц будут четырехчастотными по времени с частотами ω , γ , $\omega - \gamma$, $\omega + \gamma$. Они представят собой намотки на тороидальные поверхности, образованные путем вращения различным образом ориентированных в пространстве эллипсов вокруг оси Z . Для несоизмеримых частот ω и γ эти намотки квазипериодичны и каждая траектория заполняет всю поверхность соответствующей тороидальной фигуры. Аналитическое выражение для траекторий имеет вид

$$\begin{aligned} X + iY &= \left(\int L_{11}(\xi_1) d\xi_1 e^{-i\omega t} + \int \bar{L}_{21}(\xi_1) d\bar{\xi}_1 e^{i\omega t} - \omega c \right) e^{-i\gamma t}, \\ Z &= 2 \operatorname{Re} \left\{ \left[(\omega + 2\gamma) \int L_{11} d\xi_1 + (\omega + 2\gamma) \int L_{12} d\xi_1 \right] e^{-i\omega t} \right\} + c. \end{aligned} \quad (20)$$

Любопытно заметить, что выражение для каждой координаты частицы X , Y , Z в этом решении состоит из 9 слагаемых. Это означает, что подстановка их (координат) в одно из уравнений движения Лагранжа, скажем для проверки, приведет к многочлену ($9 \times 9 \times 6 = 486$ членов!). Вспомним, что таких уравнений три. Еще больше членов будет содержать детерминант из уравнения непрерывности. Поэтому нахождение такого решения непосредственно из уравнений Лагранжа представляет значительные трудности. Отметим, что в общем случае в переменных Эйлера эти решения (как и толемеевские) не могут быть представлены в принципе. Наш подход позволяет получать такие решения в компактном виде и довольно просто (приведенное решение (19) получено Д. А. Зеньковичем [5]).

5. Перейдем теперь к второй форме матричных уравнений гидродинамики, еще нигде не опубликованной. Если первая форма имела своим аналогом уравнение Гельмгольца (в переменных Эйлера), или уравнение Коши (в переменных Лагранжа), то вторая форма, о которой сейчас пойдет речь, вообще не имеет аналога.

Идея основана на физическом смысле математической операции умножения матрицы \hat{R} на векторную величину da . Деформацию жидкого элемента и его поворот (с сохранением длины) можно отдельно описать, представив \hat{R} в виде произведения ортогональной матрицы на симметричную: $\hat{R} = \hat{\theta} \hat{D}$, где первая матрица ортогональна, а вторая симметрич-

на. При этом ортогональная матрица опишет вращение, а симметричная – деформацию бесконечно малого жидкого элемента. Такое представление в теории матриц называется полярным представлением матрицы [7]. Оно аналогично представлению комплексного числа в полярной форме. Мы получим уравнения для этих матриц, т. е. уравнения для деформации и вращения бесконечно малого жидкого элемента, и посмотрим, можно ли извлечь из этого какую-нибудь выгоду.

Матричная форма уравнений гидродинамики $(\hat{R}^T)_t \hat{R} - \hat{R}^T \hat{R}_t = \hat{S}$ дает ограничение лишь на антисимметричную часть матрицы $(\hat{R}^T)_t \hat{R}$. Симметрическая часть этой матрицы может быть произвольной функцией времени, она не должна лишь противоречить условию непрерывности. Легко видеть, что симметрическая часть от $(\hat{R}^T)_t \hat{R}$ равна производной по времени от $\hat{R}^T \hat{R} = \hat{D}^2$. Таким образом, достаточно, чтобы временная зависимость матрицы \hat{D} не противоречила уравнению непрерывности:

$$\frac{\partial}{\partial t}(\det \hat{D}) = 0. \quad (21)$$

Найдем уравнение для ортогональной матрицы $\hat{\theta}$. Подставив выражение для \hat{R} в матричное уравнение гидродинамики, после некоторых выкладок получим уравнение для $\hat{\theta}$:

$$\hat{\theta}_t = \hat{\theta} \hat{A}, \quad (22)$$

где матрица \hat{A} зависит только от \hat{D} и не зависит от $\hat{\theta}$: $\hat{A}(a, t) = -2^{-1}(\hat{D}^{-1} \hat{S} \hat{D}^{-1} + \hat{D}^{-1} \hat{D}_t - \hat{D}_t \hat{D}^{-1})$ – антисимметричная матрица.

Полученное уравнение вместе с уравнением непрерывности и уравнением градиентности

$$\frac{\partial}{\partial a_j}(\theta_{nk} D_{km}) = \frac{\partial}{\partial a_m}(\theta_{nk} D_{kj}) \quad (23)$$

являются новой полярной формой уравнений гидродинамики. Интересным свойством этой системы является независимость уравнения непрерывности от $\hat{\theta}$. Другим важным свойством полученных уравнений является их линейность по отношению к $\hat{\theta}$.

6. Приведем пример использования полученных уравнений. Рассмотрим течение с матрицей R , зависящей только от времени. Так бывает при линейной зависимости скоростей от координат X, Y, Z в переменных Эйлера, или (что то же самое) от a, b, c в переменных Лагранжа. В этом случае уравнение градиентности становится лишним, так как удовлетворяется автоматически. Даже с учетом всех указанных упрощений система

уравнений гидродинамики (21) и (22) еще достаточно сложна. Действительно, уравнение для $\hat{\theta}$ – система трех уравнений с переменными коэффициентами. Однако в отличие от первой формы матричных уравнений гидродинамики структура уравнений гидродинамики в полярной форме позволяет развить теорию возмущений для течений указанного типа.

Действительно, пусть мы знаем решение $\hat{R}_0 = \hat{\theta}_0 \hat{D}_0$ и возмущение для \hat{D} , равное \hat{D}_1 . Ищем решение для $\hat{\theta}$, мало отличающееся от $\hat{\theta}_0$. Итак,

$$\hat{A} = \hat{A}_0 + \varepsilon \hat{A}_1, \quad \varepsilon \ll 1. \quad (23)$$

Возмущение \hat{A}_1 выражается через возмущения \hat{D}_1 . Уравнение для возмущения $\hat{\theta}$ будет с переменными коэффициентами, и, казалось бы, все очень сложно. Однако будем искать возмущение для $\hat{\theta}$ в виде (в этом вся идея)

$$\hat{\theta} = \hat{\theta}_0 + \varepsilon \hat{\Gamma} \hat{\theta}_0, \quad (24)$$

где $\hat{\Gamma}$ – неизвестная (антисимметричная) матрица. Подставив это выражение в уравнение (22) для $\hat{\theta}$, получим дифференциальное уравнение для $\hat{\Gamma}$:

$$\hat{\Gamma}_t = \hat{\theta}_0 \hat{A}_1 \hat{\theta}_0^{-1}, \quad (25)$$

где неизвестная матрица $\hat{\Gamma}$ вообще не входит в правую часть, которая целиком известна. Мы избавились таким способом от переменных коэффициентов и получаем выражение для возмущения $\hat{\theta}$ в квадратурах. Интересно, что аналогичным приемом может быть получено в квадратурах возмущение любого порядка.

На этом примере мы закончим рассмотрение второй формы матричных уравнений и перейдем к учету вязкости в переменных Лагранжа.

7. Учет вязкости в уравнении Эйлера приводит к уравнению Навье – Стокса. Оно выводилось 18 лет (1827–1845). Первое, что приходит в голову для вывода аналогичного уравнения в переменных Лагранжа, – это прямая замена переменных X, Y, Z в уравнении Навье – Стокса на лагранжевы переменные. Однако при такой прямолинейной процедуре возникают трудно интерпретируемые, на наш взгляд, образования типа сумм двойных якобианов от одних переменных по другим [8]. Полученное таким образом сорок лет назад выражение никем до сих пор не употреблялось.

Основная наша идея заключается в следующем. В отсутствие вязкости проекции завихренности на кривые подвижные, неортогональные оси жидкой системы координат постоянны (см. первую часть статьи). Как покажет результат, в присутствии вязкости они уже таковыми не будут. Для нашей цели выберем их за новые неизвестные функции. Окажется, что при наличии вязкости они будут описываться уравнением диффузии в жидкой системе координат. Это уравнение и будет искомым. По существу

ву, оно будет аналогом уравнения Гельмгольца для вязкой жидкости в переменных Лагранжа. Вывод уравнений проведем следующим образом. Уравнение Ньютона, в котором учтена вязкость (коэффициент вязкости $-\eta$),

$$\mathbf{r}_{tt} = -\nabla_x p - \eta \operatorname{rot}_x (\operatorname{rot}_x \mathbf{r}_t) \quad (26)$$

умножим слева на матрицу \hat{R}^T . Тогда получим

$$\hat{R}^T \mathbf{r}_{tt} = -\nabla_a p - \eta \hat{R}^T \operatorname{rot}_x (\operatorname{rot}_x \mathbf{r}_t). \quad (27)$$

Индексы x и a внизу у знака векторной дифференциальной операции означают, что дифференцирование ведется по эйлеровым или лагранжевым переменным соответственно. Возьмем от этого уравнения операцию ротора по лагранжевым переменным:

$$\operatorname{rot}_a (\hat{R}^T \mathbf{r}_{tt}) = -\eta \operatorname{rot}_a (\hat{R}^T \operatorname{rot}_x (\operatorname{rot}_x \mathbf{r}_t)). \quad (28)$$

Теперь преобразуем в этом уравнении роторы по переменным X в роторы по лагранжевым переменным. Для этого воспользуемся тождеством, выражающим операцию ротора от одних переменных через другие переменные [9]:

$$\operatorname{rot}_x \mathbf{A} = (\hat{R}/|D|) \operatorname{rot}_a (\hat{R}^T \mathbf{A}) \quad (29)$$

(здесь \mathbf{A} – произвольный вектор, через $|D|$ обозначен якобиан преобразования, равный $\det \hat{R}$), после чего уравнение (27) примет вид

$$\operatorname{rot}_a (\hat{R}^T \mathbf{r}_{tt}) = -\eta \operatorname{rot}_a (|D|^{-1} \hat{R}^T \hat{R} \operatorname{rot}_a (|D|^{-1} \hat{R}^T \hat{R} \operatorname{rot}_a (\hat{R}^T \mathbf{r}_t))). \quad (30)$$

Совокупность инвариантов Коши S_i можно записать в векторном виде:

$$\mathbf{S} = \operatorname{rot}_a (\hat{R}^T \mathbf{r}_t). \quad (31)$$

Учтем, что метрический тензор, характеризующий жидкую систему координат, равен

$$\hat{g} = \hat{R}^T \hat{R} \quad (32)$$

и подставим все это в уравнение (29). В итоге получим

$$\mathbf{S}_t = -\eta \operatorname{rot}_a \left(\frac{\hat{g}}{|D|} \operatorname{rot}_a \left(\frac{\hat{g}}{|D|} \mathbf{S} \right) \right). \quad (33)$$

Уравнения (31) и (33) и есть искомые уравнения гидродинамики для вязкой жидкости в лагранжевых переменных. Важно отметить, что дифференциальное уравнение для \hat{R} (уравнение (31)) в этой системе уравнений совпадает с уравнением (8) для \hat{R} в гидродинамике идеальной жидкости.

В случае двумерного течения уравнение (33) несколько упростится. Вектор Коши будет иметь всего одну компоненту, направленную, скажем, по оси c . Кроме того, метрический тензор под знаком внутреннего ротора будет вообще отсутствовать. Уравнение (33) примет вид

$$\begin{aligned}
S_t \mathbf{c}_0 &= -\eta \operatorname{rot}_a \left(\hat{\mathbf{g}} \left(\mathbf{a}_0 \frac{\partial S/|D|}{\partial b} - \mathbf{b}_0 \frac{\partial S/|D|}{\partial a} \right) \right) = \\
&= -\eta \left[\frac{\partial}{\partial b} \left(\frac{g_{11}}{|D|} \frac{\partial S/|D|}{\partial b} - \frac{g_{12}}{|D|} \frac{\partial S/|D|}{\partial a} \right) - \frac{\partial}{\partial a} \left(\frac{g_{12}}{|D|} \frac{\partial S/|D|}{\partial b} - \frac{g_{22}}{|D|} \frac{\partial S/|D|}{\partial a} \right) \right]. \quad (34)
\end{aligned}$$

8. В качестве простой иллюстрации рассмотрим пример описания этими уравнениями стационарного сдвигового течения. В переменных Эйлера имеем поле скорости $\mathbf{V} = \{u(y), 0, 0\}$. Как легко видеть, при нашем описании матрица R равна

$$\hat{R} = \begin{pmatrix} 1 & u'(b)t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (35)$$

а матрица \hat{S} , как это следует из уравнения (8), будет

$$\hat{S} = \begin{pmatrix} 0 & -u'(b) & 0 \\ u'(b) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (36)$$

Отсюда вектор $\mathbf{S} = \{0, -u'(b), 0\}$, т. е. имеет только одну компоненту и, естественно, не зависит от времени. Метрический тензор, напротив, будет зависеть от времени (напомним, $\hat{\mathbf{g}} = \hat{R}^T \hat{R}$), и будет равен

$$\hat{\mathbf{g}} = \begin{pmatrix} 1 & u't & 0 \\ u't & u'^2 t^2 + 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (37)$$

Подставив все это в уравнение гидродинамики для S (34), получим в итоге дифференциальное уравнение для $u(b)$ третьего порядка:

$$u''' = 0, \quad (38)$$

из общего решения которого $u = \alpha b^2 + \beta b + \gamma$ следует, в частности, течение Пуазейля (граничные условия $-u = 0$ при $b = 0$ и $b = h$).

9. Казалось бы, полученные уравнения (31), (33) течений вязкой жидкости слишком сложны для использования. Однако малость коэффициента вязкости позволяет существенно их упростить.

Пусть характерное время изменения X равно T , a – характерный масштаб изменения S , тогда, если $T^{-1} \gg \eta/a^2$ (большое число Рейнольдса), то в уравнении для S вместо точного значения метрического тензора

можно поставить его среднее по времени T значение. Более того, его выражение под знаком усреднения можно брать из решения для течения идеальной жидкости. Траектории частиц будут выглядеть как при отсутствии вязкости, но с плавно меняющимися параметрами.

В указанном приближении возможны такие режимы течений, когда все пространство можно разбить на области двух типов движений жидкости: «плавные», уже отмеченные, и «резкие», представляющие собой относительно тонкие слои между «плавными» течениями с существенно разными S .

Сформулируем граничные условия для бесконечно тонкого пограничного слоя указанного типа между двумя типами течений с существенно разными завихренностями – условия на скачке завихренности. Пусть переменные a и b меняются вдоль границы скачка, а уравнение $c = 0$ определяет эту граничную поверхность. Из условия непрерывности следует, что нормальные компоненты скоростей на границе обеих областей равны друг другу, а касательные равны из-за конечности S на границе. Из равенства касательных компонент скоростей следует равенство циркуляций по произвольному контуру, лежащему на граничной поверхности. Последнее означает равенство нормальных компонент завихренностей с обеих сторон границы. Покажем, что из непрерывности нормальных компонент завихренности следует непрерывность S_3 на границе. Действительно, как это видно из выбранных обозначений, векторы $\partial\mathbf{X}/\partial a$ и $\partial\mathbf{X}/\partial b$ касательны к граничной поверхности. Отсюда следует, что их векторное произведение параллельно нормали к этой поверхности. Ясно поэтому, что смешанное произведение $(\boldsymbol{\Omega}[\frac{\partial\mathbf{X}}{\partial a} \times \frac{\partial\mathbf{X}}{\partial b}])$ непрерывно на границе. С другой стороны, как следует из равенства (4), это смешанное произведение равно S_3 . Следовательно, S_3 непрерывно на поверхности разрыва.

Второе граничное условие для S может быть найдено из уравнения (33). Выберем пересекающий границу прямоугольный контур ABCD, две параллельных стороны которого (AB и CD) идут вдоль границы ($c = 0$), а две другие стороны (BC и DA) – ей перпендикулярны. Вычислим поток вектора S , через контур ABCD, применим теорему Стокса и устремим стороны контура BC и DA к нулю; после этих трех операций получим условие – на скачке завихренности касательные к границе разрыва компоненты вектора

$$\left. \frac{\hat{g}}{|D|} \operatorname{rot}_a \left(\frac{\hat{g}}{|D|} S \right) \right|_{c=0} \quad (39)$$

непрерывны. В частности, для двумерного течения на резкой границе между вихревой областью и потенциальной (пусть граница в этом случае будет $b = 0$) будем иметь условие

$$g_{11} \frac{\partial S/|D|}{\partial b} - g_{12} \frac{\partial S/|D|}{\partial a} \Big|_{c=0} = 0. \quad (40)$$

Другой предельный случай – это большой коэффициент вязкости ($\eta/a^2 \gg T^{-1}$). Тогда в уравнении (33) можно пренебречь членом $\partial S/\partial t$ и время будет входить в уравнение для S лишь параметром. При этом уравнения для координат (31), конечно, будут содержать производные по времени. Все это будет выполняться до тех пор, пока производные в правой части (33) не станут настолько малыми, что правая часть сравняется с отброшенной левой.

Работа по «освоению» уравнений (31), (33) для вязкой жидкости в переменных Лагранжа еще только началась. Предстоит много сделать, чтобы достичь такого же уровня их понимания, как и уравнений Навье–Стокса. Это относится, конечно, и к уравнениям (7)–(9) и (21)–(23) матричной гидродинамики.

Литература

- 1 Ламб Г Гидродинамика М Гостехиздат, 1947 928 с
- 2 Saffman P G Vortex Dynamics Cambridge University Press, 1993
- 3 Бэтчелор Дж Введение в динамику жидкости. М., Мир, 1973 758 с
- 4 Ландау Л Д, Лифшиц Е М Теоретическая физика Т VII Теория упругости М Наука, 1987
- 5 Абрашкин А А, Зенькович Д А, Якубович Е И Матричная формулировка гидродинамики и трехмерное обобщение птолемеевских течений // Радиофизика 1996 Т 39, № 6 С 619–622
- 6 Абрашкин А А, Якубович Е И. О плоских вихревых течениях идеальной жидкости // Докл АН СССР 1984 Т 276, № 1 С 76–78.
- 7 Гантмахер Ф Р Теория матриц. М., Наука, 1967 317 с
- 8 Кочин Н Е, Кибель И А, Розе Н В Теоретическая гидромеханика Ч II М Физматгиз, 1963 396 с
- 9 Овсянников Л В Задача о неустановившемся движении жидкости со свободной границей Новосибирск Наука, Сиб отделение, 1967 5–75 с

ДИСКРЕТНЫЕ БРИЗЕРЫ

С. Флах, А. Е. Мирошниченко

1. Введение

О существовании локальных возбуждений в сложных системах известно уже давно. Это так называемые *солитонные* решения уравнения Кортевега – де Вриза (КдВ), sine-Гордон (СГ), ϕ^4 и др. Среди решений СГ известны такие, как *кинк*, *анти-кинк*, и в особенности *бризер* – локализованное в пространстве и периодическое во времени состояние (рис. 1).

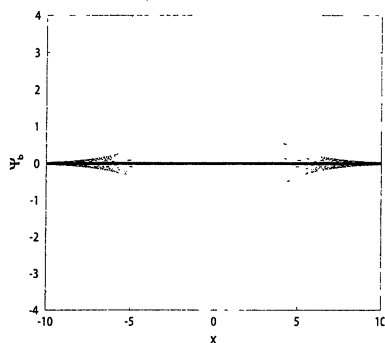


Рис. 1. СГ-бризер в различные моменты времени.

Однако это состояние структурно неустойчиво из-за существования неограниченного спектра малых колебаний, что характерно для подавляющего большинства континуальных уравнений (см. обсуждения в [1]).

В пространственно дискретных системах дело обстоит иначе. Рассмотрим d -мерную решетку, состоящую из упорядоченных в пространстве узлов. Расстояние между узлами фиксировано, что приводит к наличию дискретной трансляционной инвариантности. Состояние каждого

узла с индексом l описывается в простейшем случае канонической парой – координатой X_l и импульсом P_l , которые являются функцией времени t . Гамильтониан H всей системы есть положительно определённая функция всех координат и импульсов, с дополнительным требованием, что он обладает всеми симметриями решетки. Эволюция системы описывается гамильтоновой системой уравнений $dX_l/dt = \partial H/\partial P_l$ и $dP_l/dt = -\partial H/\partial X_l$. Без потери общности можно предположить, что $H = 0$ при $X_l = P_l = 0$. Это так называемое основное состояние решетки.

Линеаризуя уравнения движения вокруг основного состояния $H = 0$, мы получаем задачу на собственные значения. В силу трансляционной инвариантности собственные вектора будут делокализованными плоскими волнами, а собственные значения Ω_q (частоты) принадлежат фоновому спектру. В силу дискретной трансляционной инвариантности Ω_q есть периодическая функция волнового вектора q . Более того, фоновый спектр ограничен $|\Omega_q| \leq \Omega_{\max}$. Как будет показано позднее, это необходимое, но не достаточное условие существования бризеров. В зависимо-

сти от наличия или отсутствия голдстоуновской моды спектр Ω_q может быть акустическим (безщелевым) или оптическим (с наличием щели $|\Omega_q| \geq \Omega_{\min}$). С увеличением числа степеней свободы на узел увеличивается число фононных зон Ω_q^j с возможными щелями между ними.

Нас будут интересовать пространственно локализованные и периодические во времени решения полной нелинейной системы уравнений движений, а именно

$$X_l(t) = X_l(t + T_b) + \lambda m_l, \quad P_l(t) = P_l(t + T_b), \quad X_{l \rightarrow \infty} \rightarrow 0, \quad P_{l \rightarrow \infty} \rightarrow 0, \quad (1)$$

где T_b – период, m_l – целые числа и λ – вещественное число (для этого уравнения должны быть инвариантны под сдвигами по X_l на λ). Такое состояние называется *дискретным бризером*. В большинстве рассмотренных примеров $\lambda = 0$. Если $\lambda \neq 0$ и $m_l \neq 0$ для конечного числа узлов, то состояние называется *ротобризером* [1–3].

Если такое состояние существует, то его можно разложить в ряд Фурье по времени: $X_l(t) = \sum_k A_{kl} e^{ik\omega_b t} + \lambda m_l(t/T_b)$ $\omega_b = 2\pi/T_b$. Пространственная локализация подразумевает, что $A_{k,l \rightarrow \infty} \rightarrow 0$ и $m_{|l| > |l_0|} = 0$. Подставляя это разложение в уравнения движения, получаем систему связанных алгебраических уравнений для фурье-коэффициентов. Вдали от центра бризера все фурье-коэффициенты малы и должны стремиться к нулю. В силу малости амплитуд уравнения могут быть линеаризованы. Эта линеаризация приводит к расщеплению в k -пространстве, и мы получаем для каждого k линейное уравнение для A_{kl} (связанных по l). Это уравнение содержит $k\omega_b$ как параметр. Если $k\omega_b = \Omega_q$, то соответствующие амплитуды не будут спадать в пространстве, а будут периодически изменяться вдоль решётки. Тем самым мы получаем необходимое условие существования бризера:

$$k\omega_b \neq \Omega_q \quad (2)$$

– условие отсутствия резонансов гармоник бризера с фононным спектром [4]. Для оптического спектра всегда можно найти области частот, где существует бризер. Для случая акустического спектра нулевая гармоника бризера ($k = 0$) всегда находится в резонансе с $\Omega_{q=0} = 0$. Мы обсудим эту ситуацию более подробно позднее.

Как было сказано выше, отсутствие резонанса (2) – это необходимое условие существования бризера. Более детальный анализ показывает, что бризеры, как периодические орбиты, бифуркируют (отщепляются) от границ фононного спектра. Для этого необходимо присутствие нелинейных членов в уравнениях движения [1, 5].

Дискретные бризеры (периодические орбиты) появляются в основном как однопараметрические семейства замкнутых траекторий. В качестве параметра семейства может выступать частота, энергия, действие и т. д. Следует заметить, что мы не требуем существования дополнительных топологических свойств гамильтониана (таких, как наличие барьера, топологических инвариантов или интегрируемости). Для некоторых типов бризеров существует такой предел по параметру, когда бризер локализуется и все амплитуды стремятся к нулю [1].

С помощью условия отсутствия резонанса (2) мы можем исключить существование локализованных в пространстве квази-периодических во времени состояний. В самом деле, в наипростейшем случае мы должны удовлетворить следующему условию для несоизмерных частот ω_1, ω_2 и всевозможных пар целых чисел $((k_1, k_2): k_1\omega_1 + k_2\omega_2 \neq \Omega_q$. Доказательства того, что это не так, представлены в [4].

Доказательства существования дискретных бризеров приведены в [6] и [7].

Экспериментальные подтверждения существования дискретных бризеров были получены в столь разных областях, как локализация решеточных колебаний кристаллов PtCl [8], локализация света в слабо связанных волноводах с нелинейной средой [9], локализация спиновых флуктуаций в антиферромагнетиках [10], а также в системах связанных джоузефсоновских переходов, что будет обсуждено подробнее ниже.

Заметим, что бризеры могут существовать и в автономной диссипативной системе с внешней накачкой. В отличие от Гамильтонова случая периодические орбиты бризера не принадлежат однопараметрическим семействам решений, а соответствуют устойчивым предельным циклам, которые изолированы в фазовом пространстве [11].

2. Пространственная локализация бризеров

Рассмотрим гамильтониан

$$H = \sum_l \left[2^{-1} P_l^2 + V(X_l) + \sum_{l' \neq l} W_{l-l'}(X_l - X_{l'}) \right], \quad (3)$$

где $V(z), W_l(z)$ – неотрицательные функции и $V(0) = W_l(0) = 0$. Если $\partial^2 V / \partial z^2$ отлично от нуля при $z = 0$, тогда мы имеем оптический фононный спектр Ω_q , в противном случае спектр будет акустическим. На рис. 2 схематично представлены бризерные состояния для различных типов потенциалов.

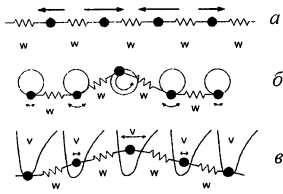


Рис. 2. Схематическое распределение амплитуд бризеров для конкретных потенциалов (3)

а $V(z) \equiv 0, W(z) = 2^{-1}z^2 + 4^{-1}z^4$,

б $V(z) \equiv 0, W(z) = 1 - \cos(z)$,

в $V(z) = 2^{-1}z^2 + 4^{-1}z^4, V'(z) = 2^{-1}z^2$ Здесь W описывает взаимодействие между ближайшими соседями

Если условие $W_{l>l_c} = 0$ выполняется для конечной области взаимодействия (короткодействующее взаимодействие) или W_l экспоненциально убывает с l , то Ω_q^2 – аналитическая функция переменной q . Для того чтобы вычислить спадание амплитуд бризера, мы используем уравнение для фурье-коэффициентов A_{kl} . С помощью метода функции Грина получаем

$$A_{kl} \propto \int_{\Xi} \frac{\cos(ql)}{(k\omega_b)^2 - \Omega_q^2} d^d q, \tag{4}$$

где интегрирование берется по первой зоне Бриллюэна Ξ . В силу сходимости ряда Фурье мы заключаем, что для короткодействующего взаимодействия A_{kl} экспоненциально спадают с l (рис. 3), а показатель экспоненты зависит от k [1].

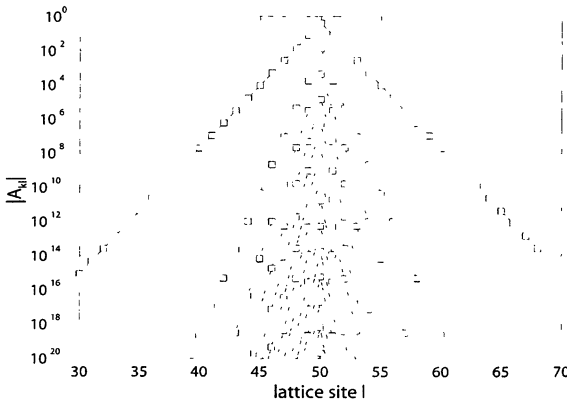


Рис. 3. Амплитуды A_{kl} как функции номера узла l Амплитуды с одинаковыми k соединены линиями (см [1])

Показатель экспоненты спадаания A_{kl} (4) стремится к нулю, как только $k\omega_b$ достигает границы спектра Ω_q . Заметим, что в этом приближении линеаризация алгебраических уравнений на A_{kl} вдали от центра бризера несправедлива для конечного числа $k' \neq k$ и нелинейные поправки должны быть учтены, что приводит к изменению соответствующих показателей [12].

2.1. Алгебраически спадающее взаимодействие

Рассмотрим одномерную решетку с алгебраически спадающим взаимодействием $W_l(z) \propto 1/l^s$ и $\partial^2 V / \partial^2 z|_{z=0} \neq 0$. Так как Ω_q^2 уже не будет аналитична по q , выражение (4) подразумевает, что на больших расстояниях пространственное спадание фурье-коэффициентов будет также алгебраическим $A_{kl} \propto 1/l^s$. Однако при $s \rightarrow \infty$ взаимодействие становится короткодействующим (взаимодействие ближайших соседей). Для понимания перехода к экспоненциальному спаду в этом пределе рассмотрим случай, когда $k\omega_b$ близко к границе зоны Ω_q , которая характеризуется некоторым волновым вектором q_c . Так как подинтегральная функция (4) расходится вблизи q_c , мы можем использовать приближение стационарной фазы и разложить Ω_q^2 около q_c , учитывая только первый порядок разложения. Для случая $s > 3$ ведущий порядок зависимости Ω_q^2 от q будет пропорционален $(q - q_c)^2$. Неаналитичность скрыта в членах высшего порядка по $(q - q_c)$ и не проявляется в этом приближении. Тогда мы получаем экспоненциальное спадание A_{kl} в пространстве. Однако мы знаем, что асимптотическая зависимость A_{kl} по l алгебраическая и заключаем, что для частот бризера, близких к границе спектра Ω_q , пространственная локализация будет экспоненциальна до некоторого характерного расстояния l_c , но становится алгебраической на больших расстояниях [13] (рис. 4).

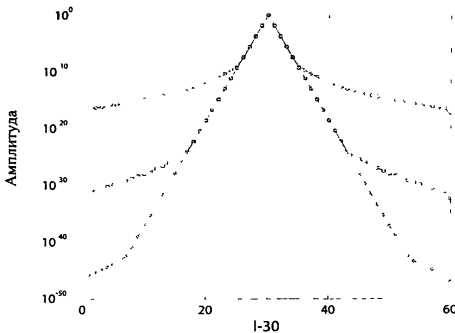


Рис. 4. Амплитуды дискретного бризера для $s = 10, 20, 30$. Четко виден переход от экспоненциального к степенному спаду [13]

Численные расчеты подтверждают это предположение. Переходное критическое расстояние может быть оценено как

$$\ln l_c / l_c \approx \nu / s, \quad (5)$$

где ν – экспонента пространственного спада, полученного в рамках приближения стационарной фазы. Из результата (5) следует, что $l_c \rightarrow \infty$ при $s \rightarrow \infty$, как и ожидалось для случая короткодействующего взаимодействия. Удивительным является то, что при $\nu \rightarrow 0$ (т. е. $k\omega_b \rightarrow \Omega_{q_c}$) мы получаем $l_c \rightarrow \infty$. Это означает, что как только частота бризера (или его гармоники) подходят к границе фононной зоны Ω_q , решение становится бесконечно слабо экспоненциально локализованным во всем пространстве [13].

2.2. Наличие голдстоуновской моды

Когда фононный спектр Ω_q содержит нуль $\Omega_{q=0} = 0$, линеаризованные уравнения вокруг основного состояния $H = 0$ содержат голдстоуновскую моду как решение. В этом случае нулевая гармоника бризера $k = 0$ заслуживает особого внимания. Все высшие гармоники ($k \neq 0$) могут быть проанализированы как и для случая оптического спектра. Если гамильтониан инвариантен под пространственным отражением $X_l \rightarrow -X_l$, то периодическое во времени решение может быть также инвариантно под этим преобразованием и $A_{kl} \equiv 0$ для четных k , включая $k = 0$. Однако если данная симметрия нарушена, все коэффициенты Фурье будут ненулевыми, включая $k = 0$.

Предположим, что Ω_q^2 аналитична по q . Так как нулевая гармоника $k = 0$ не может спадать экспоненциально на больших расстояниях от центра бризера, основная часть решения на больших расстояниях будет представлена нулевой гармоникой, медленно убывающей функцией, являющейся статической деформацией решетки. Соответствующее линеаризованное уравнение будет подобно уравнению для напряжения в механике деформированного твердого тела под действием локальной деформации (ядра бризера). Напряжение в этом случае будет спадать алгебраически в пространстве. Результирующий алгебраический спад $A_{0l} \propto 1/|l|^{d-1}$ был подтвержден численно для случая $d = 2$ [14] (рис. 5).

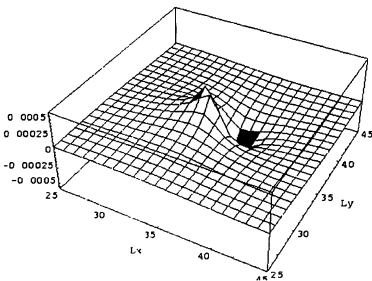


Рис. 5. Статическая деформация акустического бризера [14]

3. Порог по энергии

Прямое следствие пространственного спада бризера – это возможное наличие ненулевого порога по энергии. Напомним, что бризер представляет собой однопараметрическое семейство периодических во времени траекторий в фазовом пространстве. Наличие или отсутствие порога по энергии является физически важным. Энергия бризера может исчезнуть только в пределе нулевой амплитуды, т. е. когда ω_b достигает границ зоны Ω_q . Энергия бризера E_b оценивается как

$$E_b \propto \int_1^{\infty} r^{d-1} F_d^2(\delta r) dr, \quad (6)$$

где плотность энергии пропорциональна $A_r^2 \propto F_d^2(\delta r)$. Так как в рассматриваемом пределе пространственная локализация слабо экспоненциальна (вне зависимости от аналитических свойств Ω_q^2), функция $F_d(\delta r)$ ограничена экспоненциальной функцией с экспонентой δ . Предполагая, что дисперсия вблизи границ зоны Ω_q в ведущем порядке квадратична по $(q - q_c)$, находим, что $\delta \propto |\omega_b - \Omega_{q_c}|$. В то же время, используя теорию возмущения для слабонелинейных плоских волн с амплитудой A и частотой ω_b , можно оценить $|\omega_b - \Omega_{q_c}| \propto A^2$. Так как бризер в данном пределе представляет собой слегка деформированную (локализованную) плоскую волну, мы приходим к следующему выражению [15]:

$$E_b \propto |\omega_b - \Omega_{q_c}|^{1-d/2}. \quad (7)$$

Этот результат подразумевает, что энергия бризера не может принимать сколь угодно малые значения для $d \geq 2$.

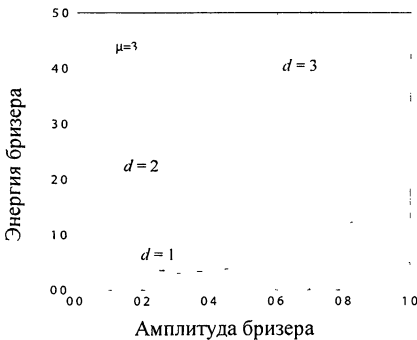


Рис. 6. Зависимость энергии бризера от амплитуды в решетках с размерностями $d = 1, 2, 3$ [15]

Следовательно, в этом случае у бризера есть ненулевое наименьшее значение энергии соответственно (рис. 6). В некоторых случаях ненуле-

вой порог по энергии существует и в одномерных системах [15]. Неаналитичность дисперсии Ω_q^2 тоже может приводить к данному эффекту (понижению критической размерности решетки) [13].

4. Ротобризеры в системе джозефсоновских контактов

Система джозефсоновских контактов является универсальным экспериментальным объектом для изучения нелинейных эффектов. Анизотропная решетка под действием постоянного тока, представленная на рис. 7, – одна из простейших систем, в которой можно возбудить дискретные бризеры.

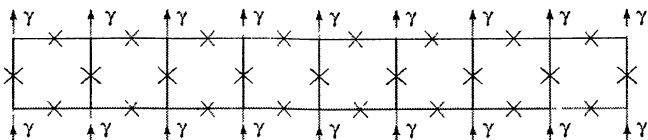


Рис. 7. Система Джозефсоновских контактов. Каждый контакт обозначен как \times , γ – внешний ток, протекающий через вертикальные контакты

Динамика одиночного джозефсоновского контакта описывается временной эволюцией разности фаз волновых функций двух сверхпроводников, соединённых слабой связью. Такой контакт может находиться в двух устойчивых состояниях сверхпроводящем и резистивном.

Когда через контакт протекает ток I_l , разность фаз $\phi_l(t)$ описывается уравнением

$$\ddot{\phi}_l + \alpha \dot{\phi}_l + \sin \phi_l = I_l. \quad (8)$$

Время нормализовано к $t_0 = (C\Phi_0 / (2\pi I_c))^{1/2}$, где Φ_0 – квант магнитного потока, C – емкость и I_c – критический ток контакта, $\alpha = (\Phi_0 / (2\pi I_c C R_N^2))^{1/2}$ – безразмерное сопротивление контакта. Обозначим через ϕ_l^v , ϕ_l^h и $\tilde{\phi}_l^h$ разности фаз вертикальных, правых верхних и правых нижних контактов в l -й ячейке. Используя закон Кирхгофа и учитывая условие квантования потока, получаем следующую систему уравнений [16]:

$$\begin{aligned} \ddot{\phi}_l^v + \alpha \dot{\phi}_l^v + \sin \phi_l^v &= \gamma + (\Delta \phi_l^v - \nabla \tilde{\phi}_{l-1}^h + \nabla \phi_{l-1}^h) / \beta_L, \\ \ddot{\phi}_l^h + \alpha \dot{\phi}_l^h + \sin \phi_l^h &= -(\phi_l^h - \tilde{\phi}_l^h + \nabla \phi_l^v) / (\eta \beta_L), \\ \ddot{\tilde{\phi}}_l^h + \alpha \dot{\tilde{\phi}}_l^h + \sin \tilde{\phi}_l^h &= (\phi_l^h - \tilde{\phi}_l^h + \nabla \phi_l^v) / (\eta \beta_L), \end{aligned} \quad (9)$$

где γ – нормализованный постоянный ток в единицах I_{cV} , $\beta_L = 2\pi L I_{cV} / \Phi_0$ – параметр дискретности, L – самоиндукция элементар-

ной ячейки и $\eta = I_{cH} / I_{cV}$ – параметр анизотропии, равный отношению величин критических токов на горизонтальном и вертикальном контактах. Разностные операторы определяются как $\nabla\phi_l = \phi_{l+1} - \phi_l$, $\Delta\phi_l = \phi_{l+1} - 2\phi_l + \phi_{l-1}$.

Бризер (в данном случае ротобризер) соответствует такому состоянию, когда конечное число контактов находится в резистивном состоянии $m_l \neq 0$, тогда как остальные контакты осциллируют около сверхпроводящего состояния ($m_l = 0$). Экспериментально были обнаружены бризеры различных типов с а) симметрией верх-вниз, б) симметрией право-лево, в) инверсией и г) отсутствием симметрии [17, 18, 19]. Каждая группа может содержать любое число вертикальных резистивных контактов n_r (рис. 8).

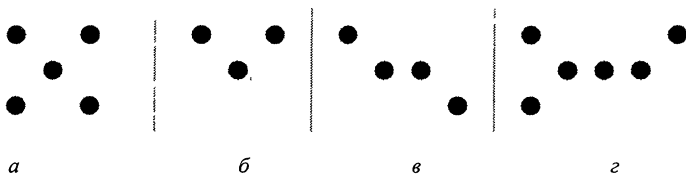


Рис. 8. Классификация бризеров по симметриям в системе контактов джозефсона. Черные кружочки обозначают контакты, находящиеся в резистивном состоянии а – симметрия *верх-вниз*, б – симметрия *право-лево*, в – *инверсия* и г – отсутствие симметрии. Число вертикальных резистивных контактов может быть любым $n_r \geq 1$

С экспериментальной точки зрения каждый бризер характеризуется своей вольт-амперной характеристикой. Среднее падение напряжения через вертикальный контакт равняется $V = T_b^{-1} \int_0^{T_b} \phi_l^v dt$. Из законов Кирхгофа следует что для бризеров типа а и г $V = 2\omega_b$, для бризеров типа б и з $V = \omega_b$.

Усреднив уравнение движения на период бризера, можно получить следующие приближённые значения для частот различных типов бризеров [20]:

$$\omega_b^a = \frac{n_r \gamma}{2\alpha(n_r + \eta)}, \quad \omega_b^{c,b} = \frac{n_r \gamma}{\alpha(n_r + 2\eta)}, \quad \omega_b^d = \frac{n_r \gamma}{\alpha(2n_r + 3\eta)}, \quad (10)$$

Авторы благодарят М. В. Фистуль за полезную дискуссию, а также Н. Б. Мирошниченко за помощь при оформлении работы.

Литература

- 1 *Flach S, Willis C R*. Discrete Breathers // *Phys Rep* 1998 V 295 P 181.
- 2 *Aubry S* Breathers in nonlinear lattices. existence, linear stability and quantization // *Physica D* 1997 V 103 P 201
- 3 *Sievers A J, Page J B*. Dynamical properties of solids VII phonon physics the cutting edge // Elsevier, Amsterdam, 1995
- 4 *Flach S* Conditions on the existence of localized excitations in nonlinear discrete systems // *Phys Rev E* 1994 V 50 P 3134
- 5 *Flach S* Tangent bifurcation of band edge plane waves, dynamical symmetry breaking and vibrational localization // *Physica D*. 1996 V 91 P 22
- 6 *MacKay R S and Aubry S* Proof of existence of breathers for time-reversible or hamiltonian networks of weakly coupled oscillators // *Nonlinearity* 1994. V 7 P 1623
- 7 *Flach S* Existence of localized excitations in nonlinear discrete systems // *Phys Rev E* 1994 V 51 P 1503
- 8 *Swanson B I, Brozik J A., Love S. P., Strouse G F, Shreve A P, Bishop A R and Wang W-Z* Observation of intrinsically localized modes in a discrete low-dimensional material // *Phys Rev Lett* 1999 V 82 P 3288
- 9 *Eisenberg H S, Silberberg Y, Morandotti R, Boyd A. R and Aitchison J S* Discrete spatial optical solitons in waveguide arrays // *Phys Rev Lett* 1998 V 81 P 3383
- 10 *Schwarz U T, English L Q. and Sievers A J* Experimental generation and observation of intrinsic localized spin wave modes in an antiferromagnet // *Phys Rev Lett* 1999 V 83 P 223
- 11 *MacKay R S, Sepulchre J A* Stability of discrete breathers // *Physica D* 1998 V 119 P 148
- 12 *Flach S* Obtaining breathers in nonlinear hamiltonian Lattices // *Phys Rev E* 1995 V 51 P 3579
- 13 *Flach S* Breathers on lattices with long-range interactions // *Phys Rev E* 1998 V 58 P R 4116
- 14 *Flach S, Kladko K, Takeno S* Acoustic breathers in two-dimensional lattices // *Phys Rev Lett*. 1997 V 79 P 4838
- 15 *Flach S, Kladko K, MacKay R. S* Energy thresholds of discrete breathers in one-, Two- and three-dimensional lattices // *Phys. Rev Lett* 1997 V 78 P 1207.
- 16 *Flach S and Spicci M*. Rotobreather dynamics in underdamped Josephson junction ladders // *J Phys Condens Matter* 1999 V 11 P. 321
- 17 *Trías E, Mazo J J, Orlando T P* Discrete breathers in nonlinear lattices experimental detection in a josephson array // *Phys Rev Lett* 2000 V 84 P 741
- 18 *Binder P, Abraimov D, Ustinov A V, Flach S, Zolotaryuk Y* Observation of breathers in Josephson ladders // *Phys. Rev. Lett.* 2000 V 84 P 745
- 19 *Binder P, Abraimov D, Ustinov A V* Diversity of discrete breathers observed in a Josephson ladder // *Phys Rev E* 2000 V 62 P 2858
- 20 *Miroshnichenko A E, Flach S, Fistul M V, Zolotaryuk Y, Page J B* Breather in Josephson junction ladders resonances and electromagnetic waves spectroscopy // *Phys Rev E* 2001 V 64 P 066601

Нелинейные явления в плазме и оптике

ДИНАМИКА АТОМА В СВЕРХСИЛЬНОМ ПОЛЕ И ГЕНЕРАЦИЯ АТТОСЕКУНДНЫХ ИМПУЛЬСОВ

М. Ю. Рябикин, А. М. Сергеев

Введение

Физика сверхбыстрых процессов вплотную подошла к вопросу о создании когерентных источников субфемтосекундного излучения. Важно отметить, что диапазон длительностей импульсов в единицы фемтосекунд, соответствующий одному периоду оптического поля, является естественным пределом для используемых методов синхронизации мод в квантовых генераторах света, поэтому для преодоления аттосекундного барьера требуются новые физические подходы. Очевидно, они должны основываться на эффектах взаимодействия мощных фемтосекундных импульсов с нелинейными средами, обладающими сверхширокополосным нелинейным откликом с масштабом, превышающим частоту оптического излучения. К настоящему времени рассматриваются три основные концепции продвижения к субфемтосекундному диапазону длительностей. В первой из них предлагается использовать сверхширокополосный континуум, возбуждаемый при самомодуляции мощного фемтосекундного оптического импульса в керровской нелинейной среде (например, газ в капилляре [1]), для сжатия на оптической системе с адаптируемыми к фазовым соотношениям в континууме дисперсионными характеристиками. Вторая концепция основана на использовании эффекта самомодуляции фемтосекундного импульса при возбуждении рамановского молекулярного отклика в газе [2, 3]. Наибольший интерес, на наш взгляд, представляет третья концепция, в которой аттосекундный импульс возбуждается за счет эффекта генерации высоких гармоник лазерного излучения в условиях быстрой ионизации атомов разреженного газа на переднем фронте мощного фемтосекундного лазерного импульса [4–7] или с использованием модуляции эллиптичности поляризации лазерного поля [8]. В данной статье изложены теоретические основы этой концепции и современное состояние ее развития.

1. Возбуждение всплесков высоких гармоник излучения при ионизации газа мощными оптическими импульсами

Напряженности лазерных электрических полей с интенсивностями порядка 10^{16} Вт/см² сравнимы или превосходят соответствующие атомные значения для наружных электронных оболочек. Это означает, что доминирующим фактором в формировании нелинейного отклика атома в сверхсильном оптическом поле является ионизация, которая происходит в течение нескольких периодов электромагнитных колебаний на переднем

фронте лазерного импульса. Именно в этот сравнительно небольшой промежуток времени нелинейная поляризация атома становится наиболее сильной, и особенности микроскопической динамики должны быть наиболее отчетливо представлены в спектрах возбуждаемого атомом излучения. Это излучение является по природе тормозным и возникает при освобождении электрона из внутриатомного потенциала или при соударениях ускоренного лазерным полем электрона с атомами и ионами.

В случае разреженного газа, когда взаимодействием освободившихся электронов с соседними атомами за время прохождения сверхкороткого импульса можно пренебречь, особую роль приобретают так называемые возвратные соударения с родительскими ионами. Как показывает классический анализ движения возникающих свободных электронов в линейно поляризованном оптическом поле, около половины этих частиц возвращаются назад к родительским ионам и взаимодействуют (соударяются) с ними хотя бы еще один раз. Если средняя осцилляторная энергия $U_p = e^2 E_0^2 / 4m\omega_0^2$, которую набирает электрон, двигаясь свободно в лазерном поле с амплитудой E_0 и частотой ω_0 , существенно превосходит потенциал ионизации атома U_i , возвратные соударения сопровождаются излучением высокоэнергетичных квантов и являются причиной возникновения своеобразного высокочастотного плато в спектре нелинейного поляриационного отклика атома [9]. Природа этого плато находит объяснение в известных свойствах тормозного излучения заряженной частицы при рассеянии на притягивающем кулоновском центре. Максимальная энергия испускаемого кванта, очевидно, должна соответствовать максимально возможной энергии электронов при возвратных соударениях. Эта простая закономерность выражается в формуле для частоты обрезания плато: $\hbar\omega_{\max} \approx 3,2U_p + U_i$. В случае линейно поляризованного лазерного импульса, благодаря повторяемости с полупериодом оптического поля актов ионизации, происходящих с различными атомами в каждом физически малом объеме, излучение атомной системы, соответствующее плато, сфазировано во времени и представляет собой набор высоких нечетных гармоник частоты лазерного поля. Ясно, что спектры гармоник служат эффективным инструментом для изучения особенностей микродинамики атома посредством макроскопических измерений характеристик оптического излучения.

Для детального анализа трансформации спектра ионизирующего лазерного импульса удобно использовать самосогласованную модель [4], состоящую из уравнений Максвелла для электрического поля волны $E(z, t)$, записанных в простейшей форме редуцированного (в пренебрежении отраженной волной) одномерного волнового уравнения, и уравнения Шредингера с внутриатомным потенциалом $V(x, y)$ (x, y – координаты в плоскости поляризации электрического поля) для одноэлектронной волновой функции:

$$\frac{\partial^2 \mathbf{E}}{\partial z \partial t} = -\frac{1}{2}(\mathbf{E} + \mathbf{R}), \quad (1)$$

$$i \frac{\partial \Psi}{\partial t} = -\frac{1}{2} \frac{\partial^2 \Psi}{\partial x^2} - \frac{1}{2} \frac{\partial^2 \Psi}{\partial y^2} + V(x, y) \Psi + \mathbf{rE}(z, t) \Psi. \quad (2)$$

Компонента $\mathbf{R}(t) = \int |\Psi|^2 (\partial V / \partial \mathbf{r}) dx dy$ поляризационного отклика среды – правая часть уравнения (1) – является источником интересующих нас коротковолновых всплесков излучения; z – координата вдоль направления распространения лазерного излучения, измеряемая в единицах $me^4 c / \omega_p \hbar^3$; $t \rightarrow t - z/c$ – время, отсчитываемое от передней границы импульса; $\omega_p^2 = 4\pi e^2 N/m$ – характерная плазменная частота, соответствующая плотности газа. Время, внутриатомные координата и потенциал, электрическое поле в импульсе нормированы на соответствующие атомные величины $t_a = \hbar^3 / me^4$, $x_a = \hbar^2 / me^2$, $V_a = me^4 / \hbar^2$, $E_a = m^2 e^5 / \hbar^4$. Несмотря на кажущуюся краткость формулировки, эта система уравнений имеет весьма универсальный характер. Все многообразие нелинейной оптики скрыто во внутриатомном потенциале $V(x, y)$. Выбирая его в соответствующем виде, можно моделировать нелинейный отклик практически любой квантовой системы, от двухуровневой до пространственно-периодической с зонной структурой уровней.

Впервые система (1)–(2) была использована в работах [4, 5] для исследования взаимодействия сверхкороткого линейно поляризованного лазерного импульса с газом «одномерных» квантовых частиц, имеющих модельный атомный потенциал $V(x) = -(1 + x^2)^{-0.5}$ [10] с энергией связи основного электронного состояния 0,67 а. е. Обнаружено, что при прохождении амплитудой поля A критического диапазона значений порядка 0,11–0,13 а. е. происходит быстрое обогащение спектра гармоник и уширение их линий. Анализ функции \bar{R} показывает, что это связано с уменьшением жизни атомов. Чем больше амплитуда лазерного импульса при фиксированной длительности, тем меньше отдельные гармоники различимы в спектре в результате перекрытия их спектров и образования в коротковолновой части спектра континуума. Другой важной особенностью спектра гармоник является их синий сдвиг, обусловленный повышением частоты возбуждающего оптического поля при ионизации газа. Проведенный анализ показал, что синий сдвиг несущей является также основным фактором насыщения роста амплитуды n -й гармоники из-за возникновения фазового рассогласования на длине $z = \lambda \omega^2 / n \omega_p^2$. Полученная в расчетах максимально достижимая амплитуда третьей гармоники составила $\approx 0,1 A$ при длительности ее всплеска менее трех периодов возбуждающего поля. Длительность всплесков гармоник на начальной

стадии нарастания их амплитуды вдоль трассы взаимодействия уменьшалась с номером гармоники. Это обстоятельство позволило впервые предположить [5], что на фронте ионизирующего фемтосекундного лазерного импульса возможно возбуждение всплеска когерентного тормозного излучения аттосекундной длительности.

2. Особенности ионизации атома мощными оптическими импульсами, содержащими малое число периодов поля, и генерация аттосекундного континуума

Для импульсов с плавной огибающей, ввиду резкой зависимости скорости ионизации от амплитуды волны, эффективная величина энергии осцилляторного движения освободившегося электрона ограничена значением поля, при котором наступает быстрая ионизация атома ($\approx 0,1E_a$). Данное ограничение может быть преодолено, как было показано в работе [11], при использовании лазерных импульсов с быстро нарастающей амплитудой (или импульсов с малым числом периодов поля). При этом оказывается возможным существенное повышение эффективности описанного выше процесса преобразования частот оптического излучения и дальнейшего продвижение в сторону уменьшения длины волны и укорочения импульсов генерируемого излучения.

Преимущества использования лазерных импульсов с малым числом периодов поля могут быть поняты из рассмотрения классических траекторий электронов, ответственных за генерацию высокоэнергичных фотонов при возвратных соударениях. В случае поля с постоянной амплитудой возвратными являются только траектории электронов, дрейфовая скорость которых направлена к родительскому иону, что имеет место для электронов, высвобождающихся из атомов в интервалы времени, когда напряженность электрического поля падает по абсолютному значению от максимума до нуля. Максимальную кинетическую энергию в момент возврата к родительскому иону имеют электроны, вылетевшие из атомов при фазе поля $\varphi \approx 18^\circ$ ($E = E_0 \cos \omega_0 t$) [9]. В случае, если амплитуда поля растет со временем, траектории оказываются возвратными и для части электронов с $\varphi < 0^\circ$. При этом в спектре тормозного излучения появляется синий сдвиг гармоник, связанный с уменьшением от одного периода поля к другому интервала между возвратными соударениями электронов с одинаковой энергией. С увеличением темпа нарастания β (в единицах частоты лазерного поля) амплитуды лазерного импульса в процесс генерации тормозного излучения при возвратных соударениях вовлекается все большая доля электронов, причем их энергия в момент столкновения с родительскими ионами может быть существенно выше, чем при $\beta = 0$. При $\beta \geq 0,28$ возвратными становятся траектории всех высвобожденных из атомов электронов, а максимальная энергия соударения приблизительно в 7 раз больше, чем в поле со стационарной амплитудой.

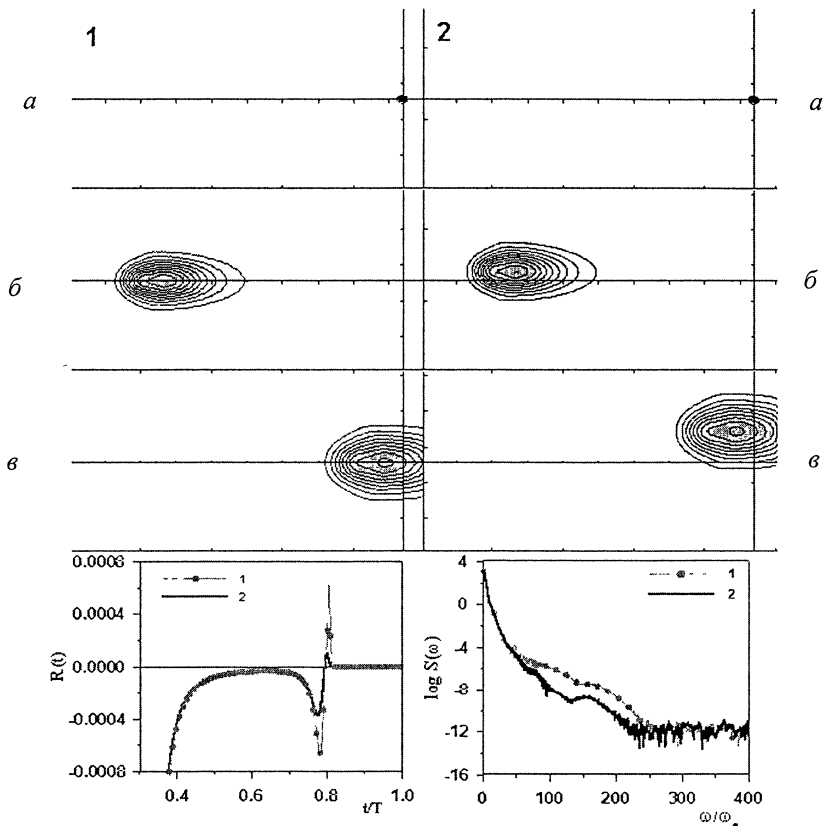


Рис. 1. Электронный волновой пакет в точке поворота (б) и в момент соударения с родительским ионом (в) Центр исходного распределения (а) совпадает с местом расположения родительского иона. Нижний рисунок – фрагмент временной зависимости атомного поляризационного отклика (слева) и его спектр (справа) Индексы 1 и 2 – результаты, полученные без учета магнитного поля волны (1) и с его учетом (2)

Если вероятность ионизации атома до смены знака поля близка к единице, то в результате образуется локализованный электронный волновой пакет, ускоренно движущийся под действием поля как целое и налетающий после смены его знака на родительский ион с высокой скоростью, что порождает мощный всплеск излучения высокоэнергичных фотонов длительностью в доли периода лазерного поля (см. рис. 1) [12]. Это и есть механизм генерации аттосекундного континуума. В сравнении с упомянутыми во введении концепциями генерации субфемтосекундных импульсов при использовании эффектов фазовой самомодуляции или самосжатия импульсов, в нашем случае потенциально достижимы существенно более короткие длительности всплесков излучения при должном

решении проблемы временной и пространственной когерентности при их возбуждении.

Так, например, очевидно, что процесс генерации всплесков довольно чувствителен к конкретному временному профилю импульса (или абсолютной фазе поля в возбуждающем импульсе) [13], и поэтому можно представить себе противоположную ситуацию, когда эффект невозвращения основной доли освобожденных электронов приведет к резкому понижению эффективности тормозного излучения. Случиться это может, например, в случае, если поле проходит диапазон критических для атома напряженностей на растущем участке, но близко к центру импульса, когда амплитуда осцилляций перестает нарастать. При этом рождение сгустка свободных электронов происходит в основном при $\varphi < 0^\circ$, а значение β близко к 0. Таким образом, для определенного класса профилей лазерных импульсов с малым числом периодов поля существует эффект запрета генерации всплесков аттосекундных импульсов.

Для изучения пространственно-временной динамики возбуждаемых всплесков в работе [14] использовалась техника вейвлет-анализа, в котором в качестве базиса был выбран вейвлет Морле $W(t, t_0) = \exp(it) \times \exp[-(t - t_0)/\sigma^2]$ с фиксированным параметром $\sigma > 2\pi$. Было установлено, что источник всплесков локализован в узком временном интервале с длительностью, много меньшей оптического периода, что позволяет использовать его для генерации аттосекундных импульсов. Для доказательства сохранения временной когерентности был произведен расчет процесса возбуждения и распространения в плазме ультракороткого импульса длительностью менее 150 ас, соответствующего спектральному диапазону 13–14,2 нм при возбуждении лазерной накачкой на длине волны 800 нм.

Идея повышения эффективности трансформации электромагнитной энергии в рентгеновский диапазон при укорочении длительности возбуждающего излучения была реализована в экспериментальных работах [15, 16], где при воздействии на атомы гелия импульсов длительностью около 25 фс, генерируемых Ti:Sa-лазерной системой с $\lambda = 800$ нм, была впервые достигнута генерация ультракоротких импульсов когерентного рентгеновского излучения с длиной волны 2,7 нм, попадающей в диапазон "водяного окна". Теоретически обнаруженный эффект синего сдвига частот гармоник в отклике отдельного атома в отсутствие сдвига частоты возбуждающего поля при возрастании его амплитуды [6] был также подтвержден в последующих экспериментальных работах [17].

3. Физические ограничения на эффективность генерации и длительность аттосекундных импульсов

Кроме упомянутой выше зависимости эффективности возбуждения субфемтосекундных всплесков от фазового рассинхронизма волн в обра-

зующейся плазме и абсолютной фазы поля в оптическом импульсе, существуют важные факторы ограничения генерации, обусловленные различными особенностями эволюции пакета электронных состояний при движении в сильном поле вне атома.

Одной из особенностей атомной нелинейности в сверхсильных полях является ее резкая зависимость от поляризации оптического поля [18]. Ясно, что если основной вклад в генерацию гармоник вносит тормозное излучение при возвратных соударениях, эффективность этого процесса должна драматически падать даже в присутствии малой циркулярно поляризованной компоненты на фоне сильного линейно поляризованного поля. Причина этого заключается в отклонении вновь родившихся электронов при движении вне атома от прямолинейной траектории за счет действия поперечной компоненты поля, а также в исчезновении возвратных соударений. В работе [19] исследованы эволюция Ψ -функции волнового пакета электронов и спектры поляризационного отклика при численном моделировании квантово-механической задачи о динамике одноэлектронного атома с двумерным потенциалом $V = -\frac{1}{\sqrt{x^2+y^2+0.01}}$ в зависимости от

поляризации оптического поля в лазерном импульсе. В случае линейно поляризованного поля с амплитудой в несколько десятых долей E_a разрушение атома происходит за 2–3 периода поля с частотой $\omega_0 = 0, 2\omega_a$, причем значительная часть волнового пакета электронов, освобождаемых из внутриатомного потенциала, возвращается назад и вновь рассеивается на нем. В результате в спектре отклика наблюдается медленно убывающий высокоэнергичный хвост вплоть до частот $\omega \approx 30\omega_0$. В случае лазерного импульса с той же энергией, но циркулярной поляризацией поля Ψ -функция волнового пакета освободившихся электронов имеет форму спирали, раскручивающейся с течением времени от атомного центра (рис. 2), т. е. возвратные соударения отсутствуют. В спектре отклика, соответственно, исчезает высокочастотная часть и остаются только черты, обусловленные тормозным излучением электронов при вылете из атома. Таким образом, применение света различной поляризации позволяет подчеркнуть или ослабить отдельные черты нелинейного отклика атома и в результате селективно диагностировать микропроцессы в атоме в присутствии сверхсильного лазерного поля. С другой стороны, можно сделать вывод, что для эффективного возбуждения субфемтосекундных импульсов необходимо тщательно контролировать поляризацию оптического импульса.

Другим ограничением эффективности генерации аттосекундных всплесков является влияние силы Лоренца со стороны магнитного поля лазерной волны [12], отклоняющей траектории электронов от прямолинейных, что особенно важно для световых полей с интенсивностями $I \geq 3 \cdot 10^{16}$ Вт/см², в которых электроны разгоняются до скоростей

$v \geq 0,1$ с. Степень влияния магнитного поля волны при возвратных соударениях в значительной мере зависит от ширины электронного волнового пакета при его прохождении вблизи родительского иона. Это влияние может оказаться существенным, если средняя скорость дрейфа электронного волнового пакета в направлении распространения лазерной волны больше или порядка скорости его расплывания в этом направлении. Данное сравнение дает оценку критического значения напряженности электрического поля: $E_0^{cr} \sim 10\omega_0$, из которой следует, что влияние магнитного поля возрастает с уменьшением его частоты и для излучения видимого диапазона становится существенным при полях E_0 порядка атомных.

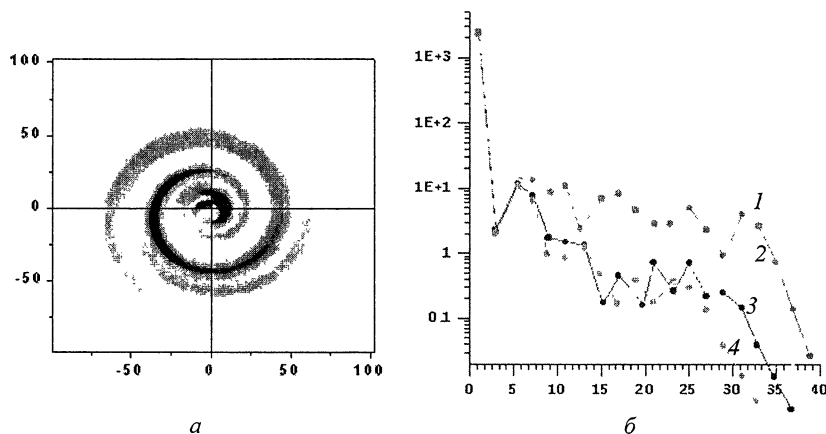


Рис. 2. Электронный волновой пакет в поле с циркулярной поляризацией (слева) и спектры гармоник, генерируемых атомом в поле с $\omega_0=0,2$ а. е и $E_x^2 + E_y^2 = 0,25$ а. е, при различных эллиптичностях поля $\varepsilon = E_x/E_y$: $\varepsilon = 0$ (1); 0,2 (2); 0,3 (3); 0,4 (4) (справа)

Для детальных расчетов возбуждения аттосекундных импульсов при учете силы Лоренца можно использовать уравнение

$$i \frac{\partial \Psi}{\partial t} = V(r)\Psi + \frac{1}{2} \left(\vec{p} + \frac{\mathbf{A}}{c} \right)^2 \Psi \quad (3)$$

с разложением векторного потенциала по координате z : $\mathbf{A}(z,t) = e_x A(z,t) \approx e_x [A_0(t) + B(t)z]$ и двумерную модель атома со сглаженным кулоновским потенциалом $V(r) = -V_0(a^2 + r^2)^{-1/2}$.

Представленный на рис. 1 (с индексом 2) случай $a = 0,3$ и $V_0 = 0,712$ моделирует атом водорода с энергией основного состояния $U_0 = -0,5$ а. е (13,6 эВ). Для линейно поляризованного поля вида $E(\tau) = E_0 f(\tau) \sin \omega_0 \tau$,

где $\tau = t - z/c$, $f(\tau) = \exp(2\omega_0\tau/\pi) - 1$, $E_0 = 0,36$ и $\omega_0 = 0,114$ (что соответствует второй гармонике излучения Ti:Sa-лазера с длиной волны 400 нм), под действием магнитного поля центр волнового пакета к моменту столкновения смещается от начала координат в направлении распространения лазерной волны на расстояние более 20 боровских радиусов, что почти вдвое превышает его полуширину в этом направлении. В результате всплеск, ответственный за генерацию высокоэнергичных фотонов с длительностью порядка 0,1 Т (~100 ас) имеет амплитуду на порядок меньше в сравнении со случаем без силы Лоренца, что находит свое отражение в спектре генерируемого излучения.

Двойкой, с точки зрения генерации аттосекундных всплесков, оказывается роль квантовой диффузии. С одной стороны, она приводит к уменьшению электронной плотности пакета, с другой – частично компенсирует действие факторов (эллиптичность поляризации, сила Лоренца), приводящих к искривлению электронных траекторий. Именно благодаря последнему обстоятельству даже при отклонении траектории центра волнового пакета от прямолинейной на расстояние, существенно превышающее атомные размеры, эффективность генерации высокоэнергичных фотонов сохраняется на достаточно высоком уровне.

4. Возбуждение аттосекундных импульсов при быстрой ионизации молекул оптическим импульсом

Дальнейшее продвижение по пути улучшения эффективности генерации аттосекундных импульсов и высоких гармоник может быть связано с использованием вместо атомов молекул или молекулярных кластеров как частиц с потенциально большим нелинейным откликом для преобразования энергии оптического излучения в коротковолновый диапазон. Молекулы и кластеры обладают разнообразием размеров, форм, ориентаций по отношению к возбуждающему лазерному полю – т. е. новых по сравнению с атомными системами факторов, которые могут быть использованы для оптимизации нелинейного отклика. В выполненных в последние годы теоретических исследованиях сформулирован ряд возможных преимуществ молекулярных систем. Среди них – расширение области коротковолнового плато в спектрах генерации гармоник за счет соударения ускоренного лазерным полем электрона с «чужим» ионом в молекуле [20]; возбуждение, наряду с традиционными нечетными, также и четных гармоник со сравнимой по величине амплитудой при ионизации молекул, состоящих из разных атомов [21]; возрастание эффективности генерации гармоник при ориентации возбуждающего поля поперек оси молекулы [22]. Однако в проведенных к настоящему времени экспериментальных исследованиях с молекулярными газами (см., например, [23]) пока не удалось обнаружить режимы с заметным увеличением эффективности генерации гармоник по сравнению с инертными газами. Возможным объясне-

нием является то, что, как показано ниже, режим оптимального нелинейного отклика соответствует существенно неравновесному расположению ядер в молекуле с расстоянием между ними, превышающим значение, определяемое минимумом кривой потенциальной энергии молекулы. Другими словами, для эффективной трансформации оптического излучения в коротковолновый диапазон молекула должна находиться в высоковозбужденном колебательном состоянии либо в распадном терме в процессе начавшейся диссоциации. Такие условия, очевидно, требуют специальной подготовки молекулярной среды в эксперименте.

Рассмотрим широко принятую модель [20, 22, 24] простейшей молекулярной системы – двумерный аналог иона X_2^+ со сглаженным кулоновским потенциалом $V(x,y)$, образованным двумя одинаковыми однозарядными центрами, удерживающими единственный электрон:

$$V(x, y) = -\frac{1}{\sqrt{(x - R/2)^2 + y^2 + a^2}} - \frac{1}{\sqrt{(x + R/2)^2 + y^2 + a^2}}. \quad (4)$$

Здесь R – расстояние между ядрами; a – параметр сглаживания, выбираемый обычно из условия совпадения энергии ионизации модельного иона из нижнего электронного терма с его величиной для реального молекулярного иона. Стационарные состояния электрона в потенциале (4) могут быть получены, например, методом численного интегрирования уравнения Шредингера (2) с нулевым внешним полем в мнимом времени. Найденная таким образом кривая потенциальной энергии молекулы $U(R)$ имеет минимум, соответствующий, в частности, для параметра сглаживания $a = (0,5)^{1/2}$ равновесному расстоянию между ядрами $R = 2,4$ а. е., потенциалу ионизации $U_i = 28,5$ эВ и энергии диссоциации $U_D = 2,9$ эВ, что близко к экспериментально измеренным значениям для реального молекулярного иона H_2^+ .

Как уже отмечалось выше, основным фактором, ограничивающим величину нелинейного отклика ионизируемой частицы в линейно поляризованном оптическом поле с нерелятивистской интенсивностью, является диффузия волнового пакета электронов при их ускорении вне атома (молекулы). Время ускорения, составляющее часть оптического периода, значительно превосходит внутриатомные времена изменения волновой функции электронов, в результате чего к моменту возвратного соударения волновой пакет существенно расплывается, и в генерации тормозного излучения принимает участие лишь малая часть ускоренных электронов. На языке классического описания – поперечная (по отношению к ионизирующему полю) компонента импульса у большей части освобожденных из атомов электронов отклоняет траекторию их движения далеко в сторону от рассеивающего центра.

Очевидно, данная ситуация может быть изменена, если в исходном распределении Ψ -функции электронов до включения лазерного поля бу-

дет обогащена компонента, характеризующая малым значением поперечного импульса. Расплывание такого пакета при ускорении оптическим полем будет замедлено, и, следовательно, возрастет количество электронов, участвующих в возвратных соударениях. Именно такая ситуация характерна, в отличие от атома, для молекулярных систем, само существование которых обусловлено наличием делокализованной компоненты в распределении Ψ -функции электронов, расположенной в межъядерной области и ответственной за возникновение притяжения ядер друг к другу.

Уменьшение распыливания электронного волнового пакета при увеличении межъядерного расстояния и, как следствие, значительное возрастание нелинейного отклика в коротковолновой части спектра наиболее наглядно прослеживаются в режиме генерации аттосекундных импульсов при полной ионизации молекулы в течение одного периода быстро нарастающего лазерного поля. Соответствующая картина свободна от множественной интерференции волновых пакетов, которая неизбежно возникает при ионизации, растянутой на несколько периодов поля, и затрудняет четкую идентификацию эффекта. На рис. 3 представлена полученная в результате численного решения уравнения Шредингера структура волнового пакета в момент перед его возвратным соударением с молекулярным остовом с размерами $R = 2,4; 5; 7,2$ и 18 а. е. при наложении перпендикулярного молекулярной оси электрического поля $E(\tau) = E_0 f(\tau) \sin \omega_0 \tau$,

$E_0 = 0,36$, $\omega_0 = 0,114$. Здесь же показаны распределения исходного состояния электронов по поперечному импульсу $P(k_x) = \int_{-\infty}^{\infty} |\Psi(k_x, k_y)|^2 dk_y$.

Заштрихованная область соответствует электронам, которые за время свободного движения смещаются в поперечном направлении на расстояние, не превышающее R .

Временная зависимость второй производной дипольного момента молекулярного иона в направлении вектора поляризации лазерного поля, демонстрирующая возбуждение всплеска с длительностью менее 200 ас, и соответствующие спектры отклика приведены на рис. 4. Заметим, что эффективность генерации коротковолновых квантов с частотами $\omega > 30\omega_0$ для межъядерного расстояния $R = 7,2$ а. е., по меньшей мере, на порядок превышает соответствующее значение для равновесной конфигурации молекулы с $R = 2,4$ а. е.

Принципиальным вопросом для сравнения откликов атомной и молекулярной систем является поведение эффективности генерации при больших R . Очевидно, что исследуемый нами молекулярный ион превращается в пределе $R \rightarrow \infty$ в пару невзаимодействующих однозарядных ионов с суммарным электронным зарядом, равным единице. Возвратное тормозное излучение в такой системе в направлении, перпендикулярном ее оси, тождественно излучению одного атома с единичным электронным зарядом. Таким образом, увеличивая межъядерное расстояние, мы имеем воз-

возможность в рамках данной модели прямо ответить на вопрос, насколько молекулярная система является более эффективным генератором коротковолновых квантов, чем аналогичная атомная система, при одинаковом числе «работающих» на тормозное излучение электронов. На рис. 4, б жирной линией представлен спектр аттосекундного всплеска при ионизации атома с потенциалом $V = -1/(x^2+y^2+0,5)^{1/2}$. Видно, что при оптимальном межъядерном расстоянии ($R \approx 7,2$ а. е.) молекулярная система дает выигрыш почти на порядок во всей коротковолновой части спектра. Однако, вообще говоря, отклик атомной системы не всегда мал по сравнению с молекулярным. Более того, он превосходит, по крайней мере в рамках рассматриваемой модели, отклик молекулы, находящейся в равновесном состоянии ($R = 2,4$ а. е.).

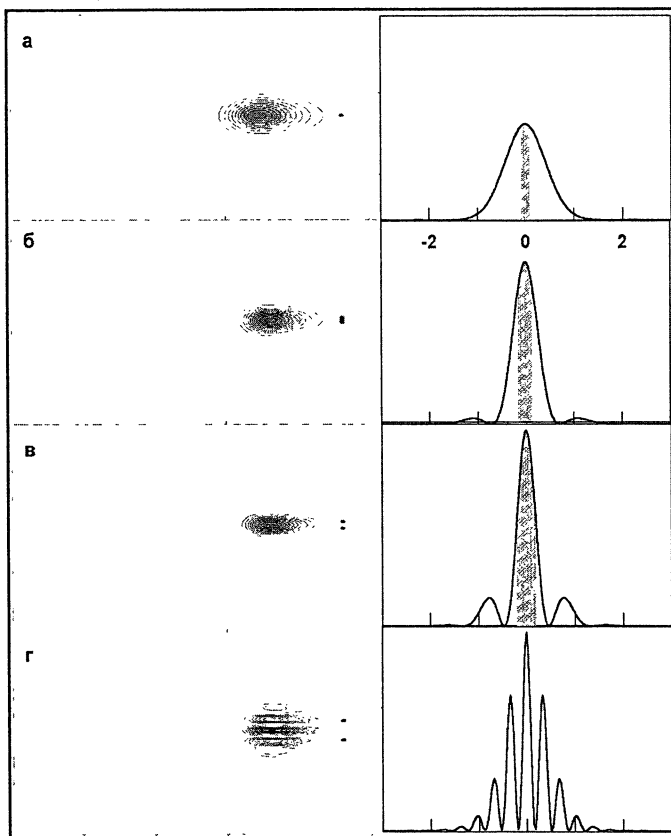


Рис. 3. Электронный волновой пакет перед возвратным соударением с ионным остовом (также изображен на рисунке) (слева) и распределение электронов по поперечному импульсу в исходном состоянии (справа) Результаты приведены для $R = 2,4$ (а), 5 (б), 7,2 (в) и 18 (г)

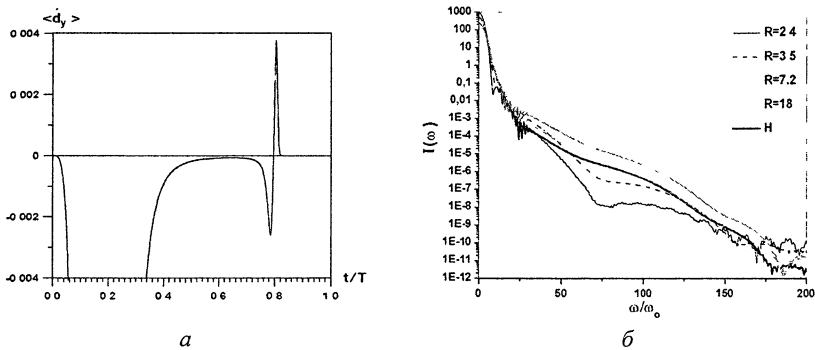


Рис. 4. Временная зависимость поляризационного отклика молекулярного иона в направлении лазерного поля при $R = 7,2$ (а) и спектры молекулярного отклика при $R = 2,4, 3,5, 7,2$ и 18 (б). На рисунке б приведен также спектр отклика для двумерной модели атома с $V(x, y) = -(x^2 + y^2 + 0,5)^{-1/2}$.

Отметим, что постепенный переход от молекулярного отклика к атомному сопровождается проявлением новой, «перекрестной» компоненты возвратного тормозного излучения. Она обусловлена электронами, вырывающимися из окрестности одного иона, ускоряемыми лазерным полем и соударяющимися с другим ионом. Присутствие таких электронов и их преобладание над делокализованными электронами при $R > 20$ а. е. отражается в интерференционной модуляции структуры волнового пакета при его возвращении к молекулярному остову (см. рис. 3, з). При еще большем раздвижении ядер части волнового пакета, соответствующие исходно локализованным вблизи притягивающих центров электронам, при расплывании в процессе ускорения не успевают существенно перекрыться, в результате чего отклик системы превращается в чисто атомный.

Аналогичный вывод о существовании оптимального межъядерного расстояния, существенно превышающего равновесное, для трансформации энергии лазерного импульса в вакуумный ультра-фиолетовый (ВУФ) и рентгеновский диапазон, подтверждается для процесса генерации высоких гармоник излучения при ионизации молекул в течение нескольких периодов оптического поля.

В заключение представляется интересным привести экспериментальную оценку возможностей для создания сверхкороткоимпульсных источников когерентного ВУФ-излучения, которые открываются при использовании оптимального режима генерации аттосекундных импульсов и высоких гармоник при ионизации молекул. Достигнутый к настоящему времени коэффициент конверсии энергии мощного фемтосекундного оптического импульса в импульс высоких гармоник излучения в ВУФ-диапазоне (около 50 нм) при выполнении условия фазового синхронизма в инертном газе приближается к 10^{-4} [25]. При увеличении этого коэффициента на

порядок за счет перехода к молекулярным газам и при использовании компактного Ti:Sa-лазера с длительностью импульсов 20 фс, частотой повторения 1 кГц и энергией 5 мДж средняя мощность когерентного ВУФ-излучения составит 5 мВт. Этот уровень соответствует рекордным значениям, достигнутым в настоящее время только на крупномасштабных ускорительных комплексах с лазерами на свободных электронах.

Данная работа выполнена при поддержке Российского фонда фундаментальных исследований (грант N 01-02-18006). Авторы выражают признательность М. Д. Чернобровцевой за помощь при проведении численных расчетов и подготовке материалов для выступления на Всероссийской школе по нелинейным волнам.

Литература

1. *Nisoli M., De Silvestri S., Svelto O. et al.* Compression of high-energy laser pulses below 5 fs // *Opt. Lett.* 1997. **22**, № 8 P 522-524.
2. *Kaplan A. E.* Subfemtosecond pulses in mode-locked 2π solitons of the cascade stimulated Raman scattering // *Phys. Rev. Lett.* 1994. **73**, № 9 P 1243-1246.
3. *Harris S. E. and Sokolov A. V.* Subfemtosecond pulse generation by molecular modulation // *Phys. Rev. Lett.* 1998. **81**, № 14. P. 2894-2897.
4. *Ванин Е В, Даунер М С., Ким А В, Сергеев А М* Возбуждение сверхкоротких всплесков гармоник излучения при ионизации газа мощным лазерным импульсом // *Письма в ЖЭТФ* 1993 **58**, № 12 С 964-969.
5. *Sergeev A M, Kim A V, Vanin E. V., Downer M. C.* Rapidly ionized atoms as a source of ultrashort coherent XUV radiation // *High-field interactions and short-wave-length generation. OSA Technical Digest Series (Optical Society of America, Washington DC), 1994.* **16**. P. 130-132.
6. *Sergeev A. M., Kim A V., Vanin E. V., Farina D., Lontano M., Downer M. C.* Atom in superstrong atom field towards subfemtosecond XUV sources // *Proc. SPIE* 1996. **2770**. P. 36-45.
7. *Christov I. P., Murnane M. M., Kapteyn H. C.* High harmonic generation of attosecond pulses in the "single-cycle" regime // *Phys. Rev. Lett.* 1997. **78**, № 7. P. 1251-1254
8. *Corkum P. B., Burnett N. H., Ivanov M. Yu.* Subfemtosecond pulses // *Opt Lett.* 1994, **19**, № 22 P 1870-1872.
9. *Corkum P. B.* Plasma perspective on strong-field multiphoton ionization // *Phys. Rev. Lett* 1993. **71**, № 13 P 1994-1997.
10. *Javanainen J., Eberly J. H., Su Q.* Numerical simulations of multiphoton ionization and above-threshold electron spectra // *Phys Rev. A* 1988. **38**, № 7 P 3430-3446.
11. *Sergeev A M, Kim A V., Vanin E. V.* Few-optical-cycle pulse interactions with matter. models and nonlinear effects // *Proc. SPIE* 1996. **2701**. P. 235-246
12. *Ким А В, Рябикин М Ю, Сергеев А М* От фемтосекундных к аттосекундным импульсам // *УФН* 1999. **169**, № 1. С. 58-66
13. *Sergeev A M, Gildenburg V. B., Kim A. V., Lontano M, Quiroga-Teixeiro M. L.* Few-optical-cycle pulse interactoin with plasmas: Models and nonlinear effects // *Superstrong fields in plasmas /Eds M Lontano, G. Mourou, F Pegoraro, E Sindoni American Institute of Physics, 1998.* P 15-31
14. *Kim A. V, Chernobrovtsseva M D, Kartashov D V, Sergeev A M.* Generation and stimulated amplification of high energy photon bursts at gas ionization by few-optical-cycle laser pulses // *Applications of High Field and Short Wavelength Sources VII OSA Technical Digest Series (Optical Society of America, Wasington DC, 1997), 1997* **7** P 223-225.
15. *Zhou J, Peatross J, Murnane M. M et al.* Enhanced high-harmonic generation using 25 fs laser pulses // *Phys. Rev Lett.* 1996 **76**, № 5 P.752-755.

16. Chang Z., Rundquist A., Wang H. et al. Generation of coherent soft X-rays at 2.7 nm using high harmonics // Phys. Rev. Lett. 1997. **79**, № 16. P. 2967-2970.
17. Shin H. J., Lee D. G., Cha Y. H. et al. Generation of nonadiabatic blueshift of high harmonics in an intense femtosecond laser field // Phys. Rev. Lett. 1999. **83**, № 13 P 2544-2547.
18. Budil K. S., Salieres P., L'Huillier A. et al. Influence of ellipticity on harmonic generation // Phys. Rev. A. 1993. **48**, № 5 P 3437-3440.
19. Бабин А. А., Ким А. В., Куселев А. М., Сергеев А. М., Степанов А. Н. Взаимодействие сверхсильных лазерных полей с веществом гипотезы, эффекты, приложения // Изв. Вузов Радиофизика 1996. **39**, № 6 С. 713-734.
20. Bandrauk A. D., and Yu H. High-order harmonic generation by one- and two-electron molecular ions with intense laser pulses // Phys. Rev. A. 1999. **59**, № 1 P 539-548.
21. Kreibich T., Lein M., Engel V, Gross E. K. U. Even-harmonic generation due to beyond-Born-Oppenheimer dynamics // Phys. Rev. Lett. 2001. **87**, № 10. Art. no. 103901.
22. Lappas D. G, and Marangos J. P. Orientation dependence of high-order harmonic generation in hydrogen molecular ions // J. Phys. B. At. Mol. Opt. Phys. 2000 **33**, № 21 P 4679-4689
23. Velotta R., Hay N., Mason M. B. et al High-order harmonic generation in aligned molecules // Phys. Rev. Lett. 2001. **87**, № 18 Art. no. 183901.
24. Lein M, Hay N., Velotta R et al Role of the intramolecular phase in high-harmonic generation // Phys. Rev. Lett 2002. **88**, № 18 Art no. 183903.
25. Hergott J.-F., Kovacev M, Merdji H et al Extreme-ultraviolet high-order harmonic pulses in the microjoule range // Phys. Rev. A. 2002. **62**, № 2. Art. no 021801.

СТИМУЛИРОВАННОЕ ИЗЛУЧЕНИЕ БЕЗ ИНВЕРСИИ И ЭЛЕКТРОМАГНИТНО-ИНДУЦИРОВАННАЯ ПРОЗРАЧНОСТЬ В ПЛАЗМЕ И ЭЛЕКТРОННЫХ ПОТОКАХ

М. Д. Токман, А. Г. Литвак, М. А. Ерухимова, А. Ю. Крячко

Введение

В последние годы интенсивно исследуется взаимодействие электромагнитных волн с многоуровневыми квантовыми системами, находящимися в так называемых когерентных состояниях. Наиболее широко известными примерами таких процессов являются стимулированное излучение без инверсии (СБИ) [1] и электромагнитно-индуцированная прозрачность (ЭИП)* [2], а также тесно связанные с ЭИП «замедление» [3] и «остановка» света [4, 5]. Интерес к этим эффектам обусловлен не только их возможными приложениями (см. в этой связи [1, 2, 6]), но и их связью с принципиальными вопросами физики взаимодействия излучения с веществом. В процессе их реализации имеет место кардинальное изменение активных и реактивных свойств резонансной среды при ее взаимодействии с относительно слабым излучением (т. е. не способным существенно изменить распределение населенностей по уровням). Таким образом, реализация эффектов СБИ и ЭИП связана с тем, что взаимодействие немонохроматического излучения с резонансными квантовыми переходами отнюдь не всегда однозначно определяется распределением населенностей. В силу известной общности радиационных процессов в ансамблях квантовых и классических осцилляторов-электронов [7, 8, 9] можно надеяться, что и система классических частиц со спадающим энергетическим спектром в области резонанса может находиться в состояниях, в которых возможно стимулированное излучение или распространение электромагнитных волн без затухания. Именно этот аспект электродинамики плазмы и электронных потоков является основным предметом данного обзора, хотя далее мы кратко остановимся и на соответствующих особенностях квантовых прототипов рассматриваемых классических систем.

При всей кажущейся парадоксальности эффектов СБИ и ЭИП речь идет, в сущности, о параметрических радиационных процессах, которые интенсивно исследуются уже много десятилетий. В частности, еще в 1957 году появились работы [10, 11, 12], из результатов которых, в принципе, следовала возможность реализации эффектов СБИ и ЭИП в трехуровне-

* Стандартные аббревиатуры в англоязычной литературе. LWI (lasing without inversion) и EIT (electromagnetically induced transparency)

вой квантовой системе. Эти работы, однако, не получили немедленного развития, так как в то время динамикой многоуровневых систем в электромагнитном поле интересовались в основном в связи с проблемой создания инверсии населенностей на одном из резонансных переходов. Активные исследования собственно особенностей распространения многочастотных полей в многоуровневых когерентных средах, которые и привели к теоретическому и экспериментальному открытию эффектов СБИ и ЭИП, ведутся примерно с середины 90-х (см. обзоры [1, 2] и приведенные там ссылки).

Базовой теоретической моделью в этих исследованиях является так называемая трехуровневая Λ -схема (рис. 1). В рамках Λ -схемы обычно рассматриваются две высокочастотные (ВЧ) электромагнитные волны с частотами, близкими к резонансным частотам переходов $|3\rangle \rightarrow |1\rangle$ и $|3\rangle \rightarrow |2\rangle$, параметрически связанные за счет колебаний (квантовой когерентности) на низкочастотном (НЧ) переходе $|2\rangle \rightarrow |1\rangle$.

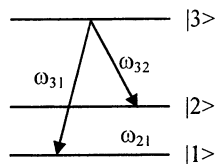


Рис. 1. Трехуровневая Λ -схема.

При этом состояние НЧ-перехода может определяться как действием сторонних источников, так и самими ВЧ-полями; во втором случае обычно рассматривается ситуация типа резонансного рассеяния (подробнее в этой связи см. [1]), когда одна из ВЧ-волн является управляющей (накачкой), а другая – пробной (сигнальной). В режиме СБИ бихроматическое излучение или высокочастотная пробная волна (в случае резонансного рассеяния) могут усиливаться в отсутствие инверсии на любом из переходов Λ -схемы, при этом электроны отдают свою энергию электромагнитному полю. Эффект ЭИП, строго говоря, является просто порогом СБИ, однако для ЭИП в режиме резонансного рассеяния характерны настолько специфические режимы распространения пробной волны, что этот случай представляет особый интерес (см. [2–6]).

Поиски классических аналогов квантовых когерентных сред впервые были предприняты в работах [13, 14] (система типа оптического клистрона) и [15] (параметрическое взаимодействие волн в изотропной плазме) и далее развиты в [16] (СБИ) и [17] (ЭИП). Однако рассмотренные в этих работах примеры лишь с известной натяжкой могут считаться классическими аналогами соответствующих эффектов в трехуровневых системах, так как в них отсутствует либо характерное для квантовой системы параметрическое взаимодействие разночастотных электромагнитных мод, либо столь же характерное резонансное взаимодействие электромагнитных мод со средой.

Прямым классическим аналогом (т. е. полностью эквивалентным исходным квантовым системам с точки зрения поведения электромагнитного поля) оказалось параметрическое взаимодействие волн в ансамбле

электронов, совершающих циклотронные колебания в магнитном поле. Классическим аналогом простейшего варианта СБИ (Λ -схема с предварительно «приготовленной» низкочастотной когерентностью) оказалась циклотронная параметрическая неустойчивость двух гармоник гирочастоты, связанных за счет предварительной модуляции функции распределения электронов по фазам циклотронного вращения на «разностной» гармонике [18] (при этом фактически эквивалентный Λ -схеме трехуровневый «блок» оказывается включенным в систему уровней Ландау, описывающих движение квантового электрона в магнитном поле). Соответствующие эффекту ЭИП особенности распространения волн были теоретически обнаружены в [19, 20] для электронно-циклотронной волны в холодной плазме, параметрически связанной с электростатической модой за счет электромагнитной накачки.

Таким образом, можно считать несомненным фактом то, что эффекты СБИ и ЭИП имеют общезначимую природу (точно так же, как и стандартное стимулированное излучение, в равной мере характерное для квантовых и классических систем; см., например, [9]). В частности, в работах [18–20] показано, что с точки зрения общей теории колебаний и волн эффект СБИ в простейшем варианте представляет собой рассеяние в среде с модулированной проводимостью, а эффект ЭИП является «параметрическим» вариантом известного в теории колебаний эффекта динамического демпфирования [21]. Эти исследования получили развитие в работах [22–24] (СБИ) и [25] (ЭИП). На основе данного общего подхода в дальнейшем удалось также предсказать эффекты электромагнитно и акустически индуцированной прозрачности в ансамблях двухуровневых систем, вовлеченных в коллективные колебания. В таких системах резонансное поглощение электромагнитного излучения на квантовом переходе подавлено за счет параметрической связи с коллективными колебаниями (оптическими [26] или акустическими [27] фононами).

В разделе 1 данного обзора на примере «безынервного» циклотронного мазера рассмотрены особенности эффекта СБИ в классической системе. В разделе 2 приведена гидродинамическая теория эффекта ЭИП для электронно-циклотронных волн в высокотемпературной плазме.

1. «Безынервная» генерация циклотронного излучения

1.1. Феноменологическая трактовка эффекта СБИ

В простейших квантовых СБИ-системах (т. е. построенных по принципу Λ -схемы с предварительно «приготовленной» низкочастотной когерентностью) и в их классических аналогах эффект «безынервного» усиления по сути является процессом параметрического взаимодействия когерентных ВЧ-мод в модулированной среде электронов, которое в опре-

деленных условиях приводит не к перекачке энергии из одной моды в другую, а к усилению обеих мод. При этом электронный ансамбль является «безынверсным» в том смысле, что он является устойчивым по отношению к генерации каждой из ВЧ-мод по отдельности (например, в случае, когда распространение второй ВЧ-моды «запрещено» внешними электродинамическими условиями).

Обычно параметрическое взаимодействие двух ВЧ-мод в среде с модулированными электромагнитными характеристиками в простейшем случае может быть описано укороченными уравнениями [28]:

$$\begin{cases} \beta_1 \dot{E}_1 + \gamma_1 E_1 = \delta \varepsilon E_2, \\ \beta_2 \dot{E}_2 + \gamma_2 E_2 = -\delta \varepsilon^* E_1, \end{cases} \quad (1)$$

где $E_{1,2}$ – комплексные амплитуды волн с частотами и волновыми векторами $\omega_{1,2}$ и $k_{1,2}$; $\delta \varepsilon$ – комплексная амплитуда возмущения некоторого электродинамического параметра среды с частотой $\Omega = \omega_1 - \omega_2$ и волновым вектором $\kappa = k_1 - k_2$; константы β_1, β_2 определяются линейной дисперсией волн; величины $\gamma_{1,2}$ описывают линейную диссипацию. Одновременное усиление двух ВЧ-мод в рамках системы (1) возможно, только если коэффициенты β_1 и β_2 имеют разные знаки. Этот хорошо известный случай соответствует волнам с разным знаком энергии. В силу этого обстоятельства в [29] эффект СБИ также интерпретировался как взаимодействие мод с разным знаком энергии. Однако обычно эффект СБИ рассматривается для пары волн, ни одна из которых не способна усиливаться в системе с линейной диссипацией в отсутствие параметрической связи, т. е. не обладает основным свойством волн с отрицательной энергией [30]. Для волн с положительной энергией одновременное усиление двух параметрически связанных ВЧ-мод возможно при смене знака в правой части одного из уравнений (1) (или при домножении правых частей (1) на мнимую единицу; заменой $\delta \varepsilon \rightarrow i \delta \varepsilon$ оба варианта сводятся друг к другу). В работе [18] было отмечено, что в квантовой Λ -схеме механизм «безынверсного» усиления сводится к параметрическому взаимодействию мод в среде с промодулированной проводимостью, которое как раз и описывается соответствующими уравнениями. При этом усредненная (немодулированная) компонента проводимости может быть положительной, что обеспечивает поглощение мод в отсутствие параметрической связи между ними – в этом смысле система «безынверсна».

Данная общефизическая трактовка механизма «безынверсного» усиления позволила сформулировать условия реализации этого эффекта в ансамбле классических электронов. Поскольку активный отклик на внешнее ВЧ-воздействие связан с частицами, находящимися в условиях синхронизма с волнами, для реализации данного механизма стимулированно-

го бихроматического излучения необходимо промодулировать функцию распределения частиц в области фазового пространства, где резонансное условие Доплера [8] выполнено для обеих ВЧ-волн (а не только для их биений, как при стандартном рассеянии [31]). Важно, что при этом усредненный по периоду НЧ-модуляции энергетический спектр может быть спадающим, т. е. «безынверсным».

Как показали исследования этого эффекта в различных схемах «безынверсного» циклотронного мазера [22, 32], в классических системах (в отличие от известных квантовых) возможна реализация «безынверсного» усиления также и в отсутствие частиц, резонансных парциальным волнам. При этом необходимый для усиления характер параметрической связи мод обеспечивается не за счет модуляции проводимости среды, а за счет «противофазной» модуляции парциальных восприимчивостей электронного ансамбля для двух разных мод. Заметим, что в этом случае ансамбль электронов в отсутствие модуляции является «чисто реактивной» (бездиссипативной) средой, однако в среде с пространственной или временной дисперсией модуляция некоторого параметра с необходимостью приводит к комплексному отклику на монохроматическое воздействие [33].

В данном разделе мы рассмотрим один из вариантов «безынверсного» циклотронного стимулированного излучения, на примере которого можно проиллюстрировать все особенности различных режимов «безынверсного» усиления в классической системе.

1.2. Основные уравнения

Рассмотрим две плоские волны, распространяющиеся под углом к постоянному магнитному полю $\mathbf{B} = Bz_0$ (рис. 2), имеющие одинаковое поперечное волновое число и разные частоты и продольные волновые числа:

$$\mathbf{E} = y_0 \sum_{j=1}^2 \text{Re } E_j \exp(ik_{\perp}x + ik_{\parallel j}z - i\omega_j t).$$

Пусть условие доплеровского резонанса на первой гармонике гирочастоты с обеими волнами выполнено для электронов с компонентами импульса $p_{\parallel R} = mc\beta_{\parallel R}$ и $p_{\perp R} = mc\beta_{\perp R}$:

$$\omega_j = \omega_B / \gamma_R + ck_{\parallel j} \beta_{\parallel R},$$

где $\omega_B = eB/mc$, m – масса покоя частицы, c – скорость света, $\gamma_R = (1 + \beta_{\parallel R}^2 + \beta_{\perp R}^2)^{1/2}$ и $\beta_{\parallel R}$ – соответственно релятивистский гамма-фактор и нормированная на скорость света продольная скорость резонансных частиц. Рассмотрим взаимодействие этого поля с ансамблем электронов, импульсы которых близки к резонансному значению.

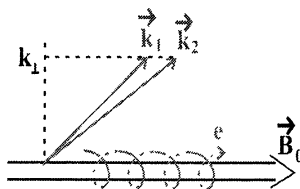


Рис. 2. «Безынверсный» циклотронный мазер на первой гармонике гирочастоты

Этот ансамбль описываем функцией распределения по «медленным» переменным $f(r, \chi, \theta, z, X, t)$, где $r = \rho_{\parallel} - \rho_{\parallel R}$, $\chi = \rho_{\perp}^2/2 - \rho_{\perp R}^2/2$, θ – фаза циклотронного вращения, X – поперечная координата центра «ларморовского кружка». Для функции распределения используем следующую форму уравнения Лиувилля [18, 22]:

$$\left(\frac{\partial}{\partial t} + \omega_H \frac{\partial}{\partial \theta} + c\beta_{\parallel} \frac{\partial}{\partial z} \right) f + \sum_{j=1}^2 F_j \left(\frac{\omega_B}{\omega_j} \frac{\partial}{\partial \chi} + \eta_{\parallel j} \frac{\partial}{\partial r} \right) f = 0. \quad (2)$$

Здесь:

$$\omega_H = \omega_B \left(1 - \chi/\gamma_R^2 - \beta_{\parallel R} r/\gamma_R \right) / \gamma_R, \quad \beta_{\parallel} = \beta_{\parallel R} + r \left(1 - \beta_{\parallel R}^2 \right) / \gamma_R - \beta_{\parallel R} r / \gamma_R^2,$$

$$F_j = G \operatorname{Re} \alpha_j \exp \left(i\theta + ik_{\perp} X + ik_{\parallel j} z - i\omega_j t \right), \quad G = (\rho_{\perp R} / \gamma_R) J_1'(k_{\perp} r_H),$$

$$\eta_{\parallel j} = ck_{\parallel j} / \omega_j, \quad r = c\rho_{\perp R} / \omega_B, \quad J_1' - \text{производная функции Бесселя,}$$

$\alpha_j = eE_j / mc$ – нормированная амплитуда волны. Для амплитуд волн также пользуемся укороченными уравнениями:

$$\partial \alpha_j / \partial t = - (2\pi e / mc) I_j. \quad (3)$$

При этом предполагаем, что волновые векторы и частоты волн связаны «вакуумным» дисперсионным соотношением. Для простоты рассмотрим начальную задачу (краевая задача обсуждалась в [22]). В уравнении (3) амплитуды резонансных гармоник тока выражаются через функцию распределения следующим образом [18, 22, 34]:

$$I_j = ecG \left\langle \int d\chi dr d\theta f(\chi, r, \theta, z, X, t) \exp(-i\theta - ik_{\perp} X - ik_{\parallel j} z + i\omega_j t) \right\rangle_{X, z, t}. \quad (4)$$

Параметрическая связь рассматриваемых волн обеспечивается предварительной модуляцией функции распределения, которую в начальный момент времени задаем в виде

$$f_{in}(t=0) = f_0(\chi, r) + f_M(\chi, r) \cos(kz + \varphi_M), \quad (5)$$

где $k = k_{\parallel 1} - k_{\parallel 2}$. При этом невозмущенная функция распределения в последующие моменты времени отличается от (5) фазой модулированной компоненты, принимающей вид: $\varphi_M + kz - ck\beta_{\parallel} t = \varphi_M + kz - \Omega t + \Delta_M t$, где $\Omega = \omega_1 - \omega_2$,

$$\Delta_M = \Omega - ck\beta_{\parallel} = \Omega\chi/\gamma_R^2 - ckr \left(1 - \beta_{\parallel R}^2 \right) / \gamma_R - \quad (6)$$

отстройка синхронизма частицы с биениями волн. Расстройки парциального циклотронного синхронизма с ВЧ волнами определяются выражениями

$$\Delta_j = \omega_j - \omega_H - ck_{\parallel j} \beta_{\parallel} = \left(\omega_j / \gamma_R \right) \left[\chi / \gamma_R + r \left(\beta_{\parallel R} - \eta_{\parallel j} \right) \right] = \Delta_{\chi} + \Delta_{\eta}. \quad (7)$$

Решим кинетическое уравнение (2) в линейном приближении по амплитуде волн. Рассмотрим асимптотическое решение, отвечающее правилу обхода Ландау и справедливое при условии:

$$\langle \delta \Delta_{\parallel} \rangle t \gg 1 \quad (8)$$

(здесь $\langle \delta \Delta_{\parallel} \rangle$ – определяемый функцией распределения характерный разброс расстройк синхронизма в пучке). При этом ограничимся временами, на которых можно пренебречь «баллистической релаксацией» промодулированной составляющей функции распределения:

$$\Delta_M t \ll 1. \quad (9)$$

Для одновременного выполнения условий (8) и (9) необходима, во-первых, спектральная близость мод: $\omega_1 \approx \omega_2 \approx \omega$ и, во-вторых, условие $r |\beta_{\parallel R} - n_{\parallel}| \ll \langle \delta \chi \rangle / \gamma_R$, означающее, что разброс расстройк синхронизма в ансамбле определяется главным образом разбросом релятивистских расстройк из-за их зависимости от поперечного импульса. В итоге с учетом (3) и (4) получим следующие уравнения параметрической связи мод:

$$\begin{cases} \dot{\alpha}_1 + (\Gamma^0 + iD^0) \alpha_1 = -\exp(i\phi_M t) \left[(\Gamma^s + iD^s) + (\Gamma^a + iD^a) \right] \alpha_2, \\ \dot{\alpha}_2 + (\Gamma^0 + iD^0) \alpha_2 = -\exp(-i\phi_M t) \left[(\Gamma^s + iD^s) - (\Gamma^a + iD^a) \right] \alpha_1. \end{cases} \quad (10)$$

Здесь в левой части представлены стандартные коэффициенты, определяющие линейную восприимчивость ансамбля электронов:

$$\Gamma^0 = \left(2\pi^2 e^2 G^2 \omega / m\gamma_R \right) (1 - n_{\parallel}^2) \int d\chi dr f_0(\chi, r) \partial \delta(\Delta_{\chi}) / \partial \Delta_{\chi},$$

$$D^0 = \left(2\pi e^2 G^2 \omega / m\gamma_R \right) (1 - n_{\parallel}^2) \int d\chi dr f_0(\chi, r) \wp / \Delta_{\chi}^2,$$

а правые части уравнений определяют обусловленную модуляцией параметрическую связь мод. Коэффициенты с индексом s определяются симметричной частью функции $f_M(r) = f_M^s(r) + f_M^a(r)$ (т. е. удовлетворяющей условию $f_M^s(\chi, r) = f_M^s(\chi, -r)$):

$$\Gamma^s = \pi^2 e^2 G^2 \omega / m\gamma_R (1 - n_{\parallel 1} n_{\parallel 2}) \int d\chi dr f_M^s(\chi, r) \partial \delta(\Delta_{\chi}) / \partial \Delta_{\chi},$$

$$D^s = \left(\pi e^2 G^2 \omega / m\gamma_R \right) (1 - n_{\parallel 1} n_{\parallel 2}) \int d\chi dr f_M^s(\chi, r) \wp / \Delta_{\chi}^2,$$

а коэффициенты с индексом a – антисимметричной частью модулированной компоненты функции распределения, удовлетворяющей условию $f_M^a(\chi, r) = -f_M^a(\chi, -r)$:

$$\Gamma^a = \left(\pi e^2 G^2 \omega / m\gamma_R^2 \right) (1 - n_{\parallel 1} n_{\parallel 2}) \times$$

$$\int d\chi dr f_M^a(\chi, r) r \left\{ -c\kappa (1 - \beta_{\parallel R}) \quad t \wp / \Delta_{\chi}^2 + \pi \left[\partial^2 \delta(\Delta_{\chi}) / \partial \Delta_{\chi}^2 \right] \omega (n_{\parallel 1} - \beta_{\parallel R}) \right\},$$

$$D^a = \left(\pi e^2 G^2 \omega / m\gamma_R^2 \right) (1 - n_{\parallel 1} n_{\parallel 2}) \times$$

$$\int d\chi dr f_M^a(\chi, r) r \left\{ -c\kappa (1 - \beta_{\parallel R}) \quad t\pi \left[\partial \delta(\Delta_{\chi}) / \partial \Delta_{\chi} \right] + \left(2\wp / \Delta_{\chi}^3 \right) \omega (n_{\parallel 1} - \beta_{\parallel R}) \right\}.$$

1.3. Различные режимы «безынервной» параметрической генерации

Как следует из полученной системы уравнений (10), вид параметрической связи, и соответственно условие и механизм усиления рассматриваемых ВЧ-мод, существенным образом зависит от типа модуляции. Можно выделить две характерные ситуации.

а) Функция распределения промодулирована *по поперечным импульсам* (при этом считаем, что $f_M^a(r) = 0$, в то время как антисимметричная часть функции $f_M(\chi)$ отлична от нуля). Существенно, что при такой модуляции отстройки синхронизма (7) Δ_1 и Δ_2 осциллируют в фазе, за счет чего в промодулированной среде формируются сходные отклики на оба ВЧ-поля. При этом получаем, что в системе усиливается бихроматическая экспоненциальная мода $\alpha_j \propto \exp(\mu t)$. Условие усиления имеет вид

$$|\Gamma^a| > \Gamma^0.$$

Это и есть режим «безынервного» усиления за счет модуляции функции распределения резонансных частиц; при этом немодулированная компонента функции «неинвертирована», так что $\Gamma^0 > 0$. Усиление двух мод достигается в данном случае за счет модуляции проводимости, оптимальным образом сфазированной с биениями ВЧ-мод. Это прямой классический аналог «безынервного» усиления в квантовой Λ -схеме. В работе [18] данный режим генерации был продемонстрирован на примере несколько иной схемы взаимодействия, когда ВЧ моды распространяются поперек постоянного магнитного поля, а их частоты близки к частотам разных гармоник циклотронной частоты; при этом электронный ансамбль с разбросом по энергиям промодулирован на разностной гармонике.

б) Функция распределения промодулирована *по продольным импульсам* (при этом считаем, что $f_M^s(r) = 0$). При такой модуляции отстройки Δ_1 и Δ_2 (7) осциллируют в противофазе, поскольку для волноводных мод параметры $\beta_{||R} - n_{||j}$ имеют разный знак. При этом отклики на первое и второе ВЧ-поля также оказываются промодулированными в противофазе. В результате, в системе может усиливаться бихроматическая экспоненциальная мода $\alpha_j \propto \exp(\mu t)$. Условие усиления имеет вид

$$|D^a| > \Gamma^0.$$

Последнее условие может быть выполнено и при отсутствии в системе частиц, резонансных ВЧ-волнам (когда $\Gamma^0 = 0$). Это и есть режим усиления в «реактивной» среде, который оказывается возможным в силу специфичной зависимости отстроек синхронизма от компонент импульса электронов, одновременно учитывающей как релятивистскую отстройку гирочастоты, так и доплеровский сдвиг.

Отметим, что необходимым для такого режима взаимодействия волн и частиц модулированный электронный ансамбль может быть сформирован, например, в рамках следующей схемы. Пусть вначале имеется моноэнергетический пучок с широким питч-угловым разбросом, соответствующий широко распространенной магнетронно-инжекторной пушке: при наклонном распространении волн относительно магнитного поля питч-угловой разброс играет ту же роль, что и широкий энергетический разброс при поперечном распространении. Модуляция функции распределения по продольному импульсу осуществляется путем кратковременного взаимодействия такого пучка с продольным полем «накачки» $E = \text{Re}[z_0 E_0 \exp(ikz - i\Omega t)]$ в предварительной секции. При этом необходимая для возникновения неустойчивости модуляция электронного пучка формируется уже на линейной стадии взаимодействия электронного ансамбля с полем «накачки» (см. [22]).

Механизм энергообмена между бихроматическим ВЧ-полем и промодулированной средой нерезонансных частиц проанализирован в [35]. Для отбора энергии у электронного ансамбля необходимо согласование движения частиц с пространственной структурой волны биений, что и происходит при развитии данной неустойчивости в изначально промодулированном по продольной скорости электронном ансамбле. Важно отметить, что для реализации такого режима принципиально наличие магнитного поля. При его отсутствии возможно лишь стандартное рассеяние, так как в этом случае в «сопровождающей» волну модуляции системе отсчета сохраняется сумма средней энергии осцилляторного движения и энергии медленного движения ведущего центра [31].

1.4. О насыщении «безынверсной» неустойчивости

Одним из исследованных (см. [23]) механизмов стабилизации «безынверсной» неустойчивости является «баллистическая перегруппировка» модулированной фракции электронов, проявляющаяся на временах $\langle \delta \Delta_M \rangle t \geq 1$, где $\langle \delta \Delta_M \rangle$ – характерный разброс по отстройкам параметрического синхронизма (6) в ансамбле электронов.

В работе [24] исследован так называемый «квазилинейный» механизм насыщения «безынверсной» неустойчивости на резонансных частицах, обусловленный самосогласованной эволюцией энергетического спектра и модулированной компоненты функции распределения. В ходе такого усложненного варианта «квазилинейной диффузии» система электронный ансамбль + электромагнитное поле релаксирует к стационарному состоянию, при котором образуется не стандартное «плато» [30], а спадающий энергетический спектр. Важно отметить, что в процессе генерации уменьшается энергия электронного ансамбля: в области резонансных частиц формируется более резкий спад по энергии, чем в начальном состоянии.

Помимо указанных механизмов насыщения возможны, в принципе, «захват» частиц в мощном поле бихроматической волны конечной амплитуды и нелинейный сдвиг частот гармоник, однако эти процессы применительно к «безынервному» стимулированному излучению еще не исследованы.

1.5. Об энергетической эффективности «безынервной» генерации в классической системе

Как и в процессе обычного стимулированного излучения, в случае «безынервной» генерации происходит передача энергии от неравновесной среды к электромагнитному полю. Кардинальное различие заключается в характере неравновесности. Если для стандартного стимулированного излучения необходим инвертированный энергетический спектр, то для «безынервного» режима нужна пространственно-временная модуляция функции распределения. Это обстоятельство и определяет возможность использования СБИ для конверсии монохроматического низкочастотного излучения в немонохроматическое высокочастотное. Ключевым моментом при оценке эффективности возможных «безынервных» электронных приборов является вопрос о соотношении между энерговыкладом электронов в усиливаемое ВЧ-поле и энергозатратами на создание предварительной модуляции пучка. Как показано в [22], для квазимоноэнергетических пучков с широким питч-угловым разбросом (типичных для обычных магнетронно-инжекторных пушек) реализация эффекта СБИ возможна в рамках «линейного» режима НЧ-модуляции, когда энергия электронного ансамбля увеличивается лишь во втором порядке по амплитуде модулирующего поля. В принципе, аналогичная ситуация может иметь место и для пучков с конечной шириной энергетического спектра электронов [18]. В этих случаях энерговыклад электронов может превосходить потери энергии на модуляцию. Принципиально иная ситуация имеет место в случае использования монотонно спадающего энергетического спектра электронов. При этом для достижения порога «безынервной» генерации необходима НЧ-модуляция в нелинейном режиме «захвата» частиц, когда неизбежно существенное искажение усредненной компоненты функции распределения (см. [18]). Для этого режима на основании результатов развитой в [24] квазилинейной теории можно сделать вывод, что энерговыклад НЧ-поля в модулирующей секции оказывается в несколько раз больше энерговыклада электронов в ВЧ-поле.

Интересной с точки зрения экспериментальных приложений представляется реализация рассмотренных здесь механизмов «безынервной» циклотронной параметрической генерации в лазерах на свободных электронах, действующих на основе ондуляторного (а не циклотронного) механизма взаимодействия частиц с ВЧ-полем.

2. Электромагнитно индуцированная прозрачность для электронно-циклотронных волн в плазме

2.1. Простейшая модель ЭИП

Построение простейшей классической модели ЭИП существенно опирается на то, что система уравнений для компонент матрицы плотности квантовой системы соответствует, вообще говоря, уравнениям, описывающим возбуждение системы классических связанных осцилляторов внешним резонансным воздействием (см. [19, 20]). С точки зрения реализации режима ЭИП важнейшей особенностью уравнений для матрицы плотности является то обстоятельство, что в них внешние источники (т. е. поля) не только непосредственно возбуждают осцилляторы, но и обеспечивают параметрические связи между ними. Поэтому аналогичная классическая колебательная система представляет собой два колебательных контура, связанных через гармонически меняющуюся взаимную индуктивность (рис. 3). В данном случае гармоническая связь между осцилляторами выступает в роли управляющей волны в распределенной системе, а внешняя ЭДС – в роли пробной волны.

Можно показать [19, 20], что частотная зависимость импеданса для первого контура на рис. 3 полностью эквивалентна выражению для восприимчивости квантовой системы по отношению к сигнальному полю в режиме ЭИП. При условии эффективного резонансного возбуждения контура с частотой ω_{21} резонансная раскачка колебаний в первом контуре может быть подавлена. Отметим,

что сам по себе эффект подавления резонансной раскачки осциллятора за счет связи с другой колебательной системой хорошо известен: это так называемое динамическое демпфирование (см., например, [21]). Однако в своем стандартном варианте этот эффект обеспечивается обычной линейной связью двух осцилляторов, что приводит к возбуждению демпфирующей системы на частоте вынуждающей обобщенной силы. В рассматриваемом нами случае «параметрическая» связь осцилляторов приводит к раскачке демпфирующей системы на комбинационной частоте. В этом основная особенность рассматриваемого эффекта.

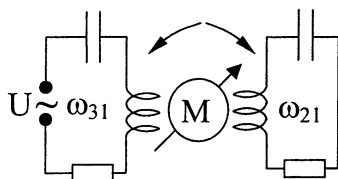


Рис. 3. Эквивалентная колебательная система с сосредоточенными параметрами

2.2. ЭИП для циклотронных волн в плазме

Для волн в плазме существует аналогичная возможность подавления резонансного взаимодействия волн с частицами за счет параметрического возбуждения коллективных степеней свободы, например электростатиче-

ских колебаний. В качестве примера рассмотрим гидродинамическую теорию ЭИП для электронно-циклотронного резонанса в магнитоактивной плазме. При всей условности гидродинамической модели в ее рамках удастся понять основные особенности, связанные с волновой природой «демпфирующей подсистемы». Далее будем придерживаться той же постановки задачи, что и в работе [25].

Пусть в магнитоактивной плазме вдоль постоянного внешнего магнитного поля $\mathbf{B} = Bz_0$ распространяются две циркулярно поляризованные волны с электрическим полем, вращающимся по направлению циклотронного вращения электронов (необыкновенные волны):

$$\mathbf{E}_\perp(z, t) = \text{Re} \left\{ \mathbf{e}_+ \left[E_1 \exp(-i\omega_1 t + ik_1 z) + E_2 \exp(-i\omega_2 t + ik_2 z) \right] \right\}, \quad (11)$$

где $\mathbf{e}_\pm = 2^{-1/2}(\mathbf{x}_0 + i\mathbf{y}_0)$ – вектор поляризации волн; $\mathbf{x}_0, \mathbf{y}_0, \mathbf{z}_0$ – орты декартовых осей.

Колебания поперечной и продольной (относительно постоянного магнитного поля) компонент скорости электронов определяются уравнениями Эйлера с учетом воздействия силы Лоренца со стороны волновых полей:

$$\begin{cases} \frac{\partial \mathbf{V}_\perp}{\partial t} + \omega_B [\mathbf{V}_\perp, \mathbf{z}_0] + \gamma \mathbf{V}_\perp + V_\parallel \frac{\partial \mathbf{V}_\perp}{\partial z} = -\frac{e}{m} \left(\mathbf{E}_\perp + V_\parallel \frac{\partial}{\partial z} \int_{-\infty}^t \mathbf{E}_\perp dt \right), \\ \frac{\partial V_\parallel}{\partial t} + \gamma V_\parallel + V_\parallel \frac{\partial V_\parallel}{\partial z} = -\frac{1}{mN_e} \frac{\partial p}{\partial z} - \frac{e}{m} E_p + \frac{e}{m} \mathbf{V}_\perp \frac{\partial}{\partial z} \int_{-\infty}^t \mathbf{E}_\perp dt. \end{cases} \quad (12)$$

Здесь γ – эффективная частота соударений, E_p – электрическое поле в плазменной волне, p – газокINETическое давление, N_e – концентрация электронов. Систему (12) необходимо дополнить уравнением непрерывности

$$\partial N_e / \partial t + \partial(N_e V_\parallel) / \partial z = 0 \quad (13)$$

и уравнением, описывающим возбуждение электрического поля в плазменной волне:

$$\partial E_p / \partial t = -4\pi j_z = 4\pi e N_e V_\parallel. \quad (14)$$

Здесь j_z – продольная компонента электронного тока. Предполагая постоянство концентрации ионов N_i и квазинейтральность

$$N_i = N_0 = \text{const}, \quad |N_e - N_i| \ll N_e, N_i, \quad (15)$$

из (12)–(14) можно получить систему уравнений, описывающих возбуждение продольных коллективных колебаний пондеромоторной силой со стороны ВЧ-полей (рассмотрим для определенности адиабатический процесс):

$$\begin{cases} \frac{\partial^2 n}{\partial t^2} + \gamma \frac{\partial n}{\partial t} + \omega_p^2 n - 3V_T^2 \frac{\partial^2 n}{\partial z^2} = -\frac{eN_0}{m} \frac{\partial}{\partial z} \mathbf{V}_\perp \int_{-\infty}^t \mathbf{E}_\perp dt, \\ V_{\parallel} = -N_0^{-1} \int_{-\infty}^z (\partial n / \partial t) dz. \end{cases} \quad (16)$$

Здесь $n = N_e - N_0$ – возмущение электронной концентрации, $\omega_p = (4\pi e N_0 / m)^{1/2}$ – электронная плазменная частота, $V_T = (T/m)^{1/2}$ и T – соответственно тепловая скорость и температура электронов.

Предположим выполненными следующие условия синхронизма:

$$|\omega_1 - \omega_B|, \quad \gamma \ll \omega_1, \omega_B, \quad (17a)$$

$$|\omega_L - \omega_p|, \quad \gamma \ll \omega_L, \omega_p, \quad (17б)$$

$$k_{1,2,L} V_T \ll \omega_{1,2,L}. \quad (17в)$$

Здесь $\omega_L = \omega_1 - \omega_2$ и $k_L = k_1 - k_2$ – соответственно частота и волновой вектор биений сигнальной волны и волны накачки. Неравенство (17a) означает выполнение условий ЭЦР для сигнальной волны и, следовательно, существование циклотронного поглощения данной волны в отсутствие волны накачки. Неравенство (17б) определяет условие возбуждения плазменных волн на комбинационной частоте ω_L . Наконец, третье неравенство представляет собой условие слабой пространственной дисперсии.

С учетом условий синхронизма (17a), (17б) в уравнениях (12)–(14) можно оставить только слагаемые с «резонансными» частотами, т. е. воспользоваться методом укороченных уравнений. Важно отметить, что при выводе данной системы уравнений необходимо учитывать нелинейную компоненту поперечной скорости с поляризацией, соответствующей обыкновенной волне (электрическое поле в которой вращается противоположно направлению циклотронного вращения электронов). Дело в том, что за счет раскачки коллективных степеней свободы резонансные компоненты скорости оказываются подавленными, поэтому учет данной компоненты приводит к заметным, хотя и не принципиальным поправкам. Из полученной системы укороченных уравнений можно найти выражение для эффективного показателя преломления сигнальной волны [25]:

$$N^2 = 1 - \frac{\omega_p^2 Z_p + \xi_{EC} \omega_L \omega_2 (2k_1/k_2 - 1 - \omega_1/2\omega_2)}{\omega_1 (\omega_1 - \omega_B + i\gamma) Z_p - \omega_L^2 \omega_1 \xi_{EC}}, \quad (18)$$

где $\xi_{EC} = |eE_2 / 2m(\omega_2 - \omega_B + i\gamma)|^2 = |V_2|^2 / (2V_{ph})^2$ – безразмерный параметр нелинейности, представляющий собой отношение квадратов осцилляторной и фазовой скоростей для поля накачки, $Z_p = \omega_L^2 - \omega_p^2 - 3k_L^2 V_T^2 + i\gamma \omega_L$, причем соотношение $Z_p = 0$ представляет собой стандартное

«гидродинамическое» дисперсионное уравнение для плазменных волн. Выражение (18) отличается от полученного в работах [19, 20] для холодной плазмы поправками, обусловленными учетом теплового движения.

Обсудим теперь особенности закона дисперсии сигнальной волны в условиях ЭИП для «теплой» плазмы. Для этого представим соотношение (18) в виде

$$Z_p(N^2 - N_0^2) = -\xi_{EC} \frac{\omega_2 \omega_L^2}{\omega_1 - \omega_B + i\gamma} \left(1 - N^2 + \frac{\omega_p^2}{\omega_1 \omega_L} \left(\frac{2k_1}{k_2} - 1 - \frac{\omega_1}{2\omega_2} \right) \right), \quad (19)$$

где $N_0^2 = 1 - \omega_p^2 / [\omega_1(\omega_1 - \omega_B + i\gamma)]$ – «линейный» показатель преломления сигнальной волны в холодной магнитоактивной плазме (см., например, [7]). Величина ξ_{EC} является малым параметром задачи: для разумных значений интенсивности излучения накачки (порядка $10 \div 100$ кВт/см²) значение ξ_{EC} составляет $10^{-6} \div 10^{-5}$. Вследствие этого определяемые (19) дисперсионные кривые сигнальной волны проходят либо вблизи линейной дисперсионной кривой сигнальной волны, определяемой соотношением $N^2 = N_0^2$, либо вблизи дисперсионной кривой плазменных волн $Z_p = 0$ (смещенной, разумеется, на частоту и волновой вектор накачки ω_2 и k_2). Указанное поведение дисперсионных кривых нарушается лишь в окрестности точек пересечения зависимостей $N^2 = N_0^2$ и $Z_p = 0$, где дисперсионные кривые сигнальной волны «переходят» с одной кривой на другую. Все вышесказанное подтверждается численными расчетами для соотношения (18), проиллюстрированными на рис. 4, из которого видно, что поведение дисперсионных кривых в области ЭИП действительно определяется законом дисперсии плазменных волн. В частности, на дисперсионных кривых существуют участки, соответствующие сильному групповому замедлению сигнальной волны (что характерно и для трехуровневой квантовой системы), вплоть до обращения групповой скорости в нуль (что невозможно в трехуровневой квантовой системе).

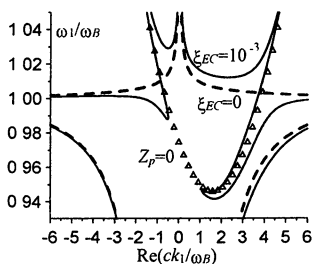


Рис. 4. Закон дисперсии ЭЦ волны в условиях ЭИП $\omega_p/\omega_B = 0,75$, $\omega_2/\omega_B = 0,2$, $\gamma/\omega_B = 2,5 \cdot 10^{-3}$, $V_T/c = 7,5 \cdot 10^{-2}$

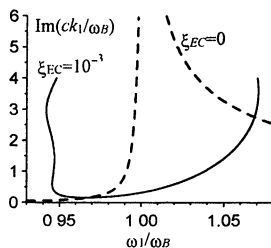


Рис. 5. Формирование окна прозрачности ($\text{Re}(k_1) > 0$) Все параметры как для рис 4

Полученный результат соответствует сформулированной ранее теоретико-колебательной трактовке эффекта ЭИП. Для реализации режима ЭИП принципиальным является параметрическое возбуждение «демпфирующих» колебаний, которые в данном случае удовлетворяют дисперсионному уравнению $Z_p = 0$.

На рис. 5 показан профиль линии поглощения сигнальной волны в условиях ЭИП. Видно, что существует диапазон частот, в котором резонансное поглощение волны подавлено, причем минимум поглощения приблизительно соответствует точке на дисперсионной кривой с нулевой групповой скоростью.

Заметим, что пороговое для реализации ЭИП значение амплитуды накачки ξ_{EC}^* определяется диссипативными эффектами (см. также [19, 20]): $\xi_{EC}^* = (\gamma^2 / \omega_p \omega_B)$.

2.3. О возможности наблюдения эффекта ЭИП в плазме

В рамках гидродинамической теории для «теплой» плазмы из (18) следует, что ширина полосы прозрачности сигнальной волны $\Delta\omega$ в условиях ЭИП определяется соотношением $|(\omega_1 - \omega_B) \text{Re} Z_p| \approx \omega_p^2 \omega_B \xi_{EC}$, что приводит к следующему выражению для $\Delta\omega$:

$$\Delta\omega \sim \max \left\{ (\omega_p \omega_B \xi_{EC})^{1/2}, k_L^2 V_T^2 / \omega_p \right\}. \quad (20)$$

Из данного соотношения, в частности, следует, что ширина полосы прозрачности в «теплой» плазме может быть существенно больше, чем в холодной, если температура плазмы достаточно велика:

$$T / (mc^2) > \omega_p^2 (ck_L)^{-2} (\omega_B \xi_{EC} / \omega_p)^{1/2}. \quad (21)$$

Важно отметить, что ширина полосы может и вообще не зависеть от ξ_{EC} . Так, например, для интенсивности накачки 100 кВт/см^2 температура плазмы должна превышать 150 эВ .

Приведенные выше оценки основных параметров «окна» ЭИП позволяют сформулировать, по крайней мере, наиболее пессимистические требования к условиям реализации данного эффекта в плазме. Ограничение на допустимую неоднородность магнитного поля B и концентрацию N_0 определяется соотношением (20):

$$\delta B, \delta N_0 \sim \max \left\{ (\omega_B \xi_{EC} / \omega_p)^{1/2}, k_L^2 V_T^2 / \omega_p^2 \right\}. \quad (22)$$

Здесь δB и δN_0 – относительная «расстройка» B и N_0 соответственно. Для интенсивности накачки порядка 10 кВт/см^2 и температур порядка 1 кэВ и 10 кэВ значения δB и δN_0 должны быть не более 1% и 5% соответственно.

При этом характерный размер рабочей области должен быть порядка одного метра (см. [25]). Не исключено, однако, что более детальная теория ЭИП для неоднородного слоя приведет к более мягким ограничениям.

Таким образом, особенности эффекта ЭИП в плазме определяются особенностями дисперсии плазменных волн. Поведение дисперсионных кривых в «окне» ЭИП в основном определяется законом дисперсии «демпфирующей системы» (плазменных волн) и практически не зависит от интенсивности волны накачки (последняя, однако, должна превышать некоторое пороговое значение); эффект ЭИП имеет место в неизменном диапазоне значений частоты и волнового вектора сигнальной волны при достаточно произвольных значениях интенсивности накачки, в то время как в приближении холодной плазмы (и в квантовой трехуровневой системе) соответствующий интервал частот сужается при уменьшении интенсивности накачки. Необходимо отметить, что важнейшие особенности поведения дисперсионных кривых и линии поглощения сигнальной волны, а также оценки для ширины полосы прозрачности и допустимой неоднородности плазмы сохраняются и при более строгом исследовании данного эффекта в рамках кинетической теории (см. [25]).

Заключение

В заключение обсудим возможные приложения рассмотренных выше систем. Эффект СБИ в ансамблях классических электронов может быть, в принципе, использован с целью конверсии низкочастотного монохроматического излучения в высокочастотное немонахроматическое циклотронное или ондуляторное излучение. При этом, в зависимости от свойств электронного потока, энерговклад ансамбля электронов в ВЧ-излучение может быть как меньше поглощаемой электронами энергии низкочастотного поля, так и больше ее. В любом случае, однако, этот способ конверсии должен быть не очень критичен к качеству электронного пучка – в силу «безынверсной» природы рассматриваемого варианта стимулированного излучения.

Что касается применения эффекта ЭИП в физике плазмы и электронике, то здесь возможны спектроскопические приложения. Однако для разработки конкретных схем необходимо развить теорию ЭИП с учетом неоднородности типичных плазменных объектов.

В [36] сформулировано также предложение об использовании ЭИП-режима в компактных плазменных ускорителях, при этом характерное для ЭИП сильное групповое замедление может быть использовано для сжатия электромагнитных импульсов. Вместо ВЧ-волны накачки в [36] предложено использовать магнитный ондулятор. Такая схема позволяет повысить эффективность энерговклада сигнальной волны в электростатические

колебания, хотя и существенно ограничивает область допустимых параметров плазмы (необходимо обеспечить близость плазменной и циклотронной частот).

Работа поддержана грантами РФФИ (проекты № 01-02-17388 и № 01-02-17779).

Литература

- 1 *Kocharovskaya O* Amplification and lasing without inversion // *Phys Rep* 1992 V 219, № 3–6 P 175–190
- 2 *Harris S E* Electromagnetically induced transparency // *Phys Today* 1997 V 50 P 36–42
- 3 *Marangos J* Slow light in cool atoms // *Nature* 1999. V 397 P 559–560
- 4 *Liu C, Dutton Z, Behroozi C. H, Hau L V* Observation of coherent optical information storage in an atomic medium using halted light pulses // *Nature* 2001 V 409 P 490–493
- 5 *Phillips D F, Fleischhauer A, Mair A., et al* Storage of light in atomic vapor // *Phys Rev Lett* 2001 V 86, № 5 P 783–786.
- 6 *Harris S E* Nonlinear optics at low light levels // *Phys Rev. Lett* 1999 V 82, № 23 P 4611–4614
- 7 *Гинзбург В Л* Распространение электромагнитных волн в плазме М Наука, 1967 684 с
- 8 *Железняков В В* Излучение в астрофизической плазме М Наука, 1997 528 с
- 9 *Гапонов А В, Петелин М И, Юлпатов В К* Индуцированное излучение возбужденных классических осцилляторов и его использование в высокочастотной электронике // Изв ВУЗов, сер Радиофизика 1967 Т. 10, № 9–10 С 1414–1453
- 10 *Scovil H E D, Feher G, Seidel H* Operation of a solid state maser // *Phys. Rev* 1957 V 105, № 2. P 762–763
- 11 *Конторович В М, Прохоров А М* О нелинейных эффектах взаимодействия резонансных полей в молекулярном генераторе и усилителе // ЖЭТФ. 1957 Т. 33 С 1428–1430
- 12 *Javan A* Theory of three-level maser // *Phys Rev* 1957 V 107, № 6. P 1579–1589
- 13 *Sherman B., Kurizki G., Nikonov D E, Scully M O* Universal classical mechanism of free-electron lasing without inversion // *Phys. Rev Lett* 1995 V 75, № 25–18. P 4602–4605.
- 14 *Nikonov D E, Sherman B, Kurizki G, Scully M O* Lasing without inversion in Cherenkov free-electron lasers // *Opt Commun* 1996 V 123, № 1–3 P 363–371.
- 15 *Harris S E* Electromagnetically induced transparency in an ideal plasma // *Phys Rev Lett.* 1996 V 77, № 27 P 5357–5360
- 16 *Artemiev A I, Fedorov M V, Rostovtsev Y V et al* Free-electron Laser without inversion gain optimization and implementation scheme // *Phys Rev Lett* 2000 V 85, № 21 P 4510–4513.
- 17 *Gordon D F, Mori W B, Joshi C* On the possibility of electromagnetically induced transparency in a plasma // *Phys of Plasmas* 2000 V 7, № 8. P 3145–3166
- 18 *Гапонов-Грехов А В, Токман М Д* О «безынверсионном» стимулированном излучении неравновесных коллективов классических осцилляторов // ЖЭТФ. 1997 Т 112, № 4(10) С 1176–1196
- 19 *Litvak A G, Tokman M D* Electromagnetically induced transparency in ensembles of classical oscillators // *Phys Rev Lett* 2002 V 88, № 9 P 095003
- 20 *Лутвак А Г, Токман М Д* Об эффекте электромагнитной индуцированной прозрачности в классических системах // Изв ВУЗов, сер Радиофизика 2001 Т 44, № 5–6 С 407–418
- 21 *Pippard A B* The Physics of vibration Cambridge University Press, Cambridge, U K ,

1989 656 p, Рабинович М И, Трубецков Д И Введение в теорию колебаний и волн. М Наука, 1984 432 с

22 Ерухимова М А, Токман М Д К теории параметрической генерации циклотронного излучения // ЖЭТФ 2000 Т 118, № 2(8) С 291–301

23 Ерухимова М А, Токман М Д К теории «безынервного» мазера на свободных электронах // Изв ВУЗов, сер Радиофизика 2001. Т 44, № 1–2 С 190–198

24 Ерухимова М А, Токман М Д Квазилинейная теория «безынервной» циклотронной генерации // Физика плазмы 2001 Т 27, № 10. С. 921–927

25 Крячко А Ю, Литвак А Г, Токман М Д Электромагнитно индуцированная прозрачность в высокотемпературной магнитоактивной плазме // ЖЭТФ 2002 Т 122, № 4(10) С 805–813

26 Litvak A G, Mironov V A., Tokman M D, Kryachko A Yu Collective and nonlinear effects in the conditions of electromagnetically induced transparency // Proceedings of the International conference «Progress in nonlinear science» dedicated to the 100th anniversary of A. A. Andronov 2002 V 2 P. 445–455

27 Radeonychev Y V, Tokman M D, Litvak A G, Kocharovskaya O Vibrationally induced transparency in optically dense resonance medium // XVII International Conference on Coherent and Nonlinear Optics (ICONO'2001), Minsk, Belarus, June 26–July 1, Technical digest 2001 P 84, Radeonychev Y. V, Tokman M D, Litvak A G, Kocharovskaya O. Transparency induced by modulation of atomic transition // Proceedings of the International conference «Progress in nonlinear science» dedicated to the 100th anniversary of A. A. Andronov 2002. V 2 P 497

28 Бломберген Н Нелинейная оптика М Мир, 1966 424 с

29. Kocharovskaya O., Mandel P Amplification without inversion the double-lambda scheme // Phys Rev A 1990 V 42, № 1 P 523–535

30 Кадомцев Б Б Коллективные явления в плазме М Наука, 1988 303 с

31 Литвак А Г Динамические нелинейные электромагнитные явления в плазме // Вопросы теории плазмы / Под ред М А Леонтовича М.: Энергоатомиздат, 1980. С 164

32 Erukhimova M A, Tokman M D. About the explosive parametric instability of modes with positive energy in the process of electromagnetic field and electron ensembles interaction // Proceedings of the international workshop «Strong microwaves in plasma» 2000. V. 2 P 827–832.

33 Ландау Л Д, Лифшиц Е М Электродинамика сплошных сред М. Наука, 1982. 623 с

34 Братман В. Л., Гинзбург Н С., Нусинович Г С. и др Циклотронные и синхротронные мазеры // Релятивистская высокочастотная электроника / Под ред А В Гапонова-Грехова Горький ИПФ АН СССР, 1979 С 157–216

35. Erukhimova M A, Tokman M D To the theory of parametric instability of the cyclotron radiation in reactive electron medium // Proceedings of the 29th EPS Conference on Plasma Phys and Contr Fusion, Montreux, 17–21 June 2002 ECA 2002 V 26B (готовится к публикации, доступно по адресу http://elise.epfl.ch/pdf/P4_001.pdf)

36 Shvets G, Wurtele J S Transparency of magnetized plasma at the cyclotron frequency // Phys Rev Lett 2002. V 89, № 11 P 115003

СТОЛКНОВЕНИЯ ЭЛЕКТРОНОВ С ИОНАМИ В СИЛЬНОМ ЛАЗЕРНОМ ПОЛЕ

Г. М. Фрайман, А. А. Балакин

1. Введение

Интерес к проблеме парных электрон-ионных соударений в сильных электромагнитных полях наблюдается уже в течение 40 лет [1–5], что само по себе свидетельствует о ее фундаментальности и важности. Особенно актуальной стала эта проблема в последние годы в связи с разработкой и созданием мощных лазеров с плотностями энергии до 10^{18} – 10^{21} Вт/см², соответствующими ультрарелятивистским осцилляторным энергиям электронов. Возникающие при этом физические явления при исследовании взаимодействия подобных полей с обычной и кластерной плазмой и возможные приложения (от лазерного УТС до преобразования оптического излучения в излучение рентгеновского диапазона длин волн) привели к тому, что проблема парных соударений в плазме из чисто академической превратилась в важную прикладную проблему. Именно изложение современного состояния проблемы парных электрон-ионных соударений в плазме в присутствии сильных полей электромагнитного излучения и является основной целью предлагаемой лекции.

Традиционно теоретическое исследование электрон-ионных соударений в сильных электромагнитных полях проводят на основе трех моделей. Все эти модели базируются на приближении парных соударений, т. е. предполагается, что вероятность одновременного столкновения трех частиц в одной точке пространства пренебрежимо мала. Это приводит к тому, что все характеристики интеграла столкновений для одночастичной функции распределения могут быть найдены из решения задачи рассеяния пучка невзаимодействующих (тестовых) электронов на одном ионе [6].

Следует отметить, что во многих вторичных работах изложение существенно усложнилось по сравнению с оригинальными работами. В то же время такое усложнение привело лишь к чисто иллюзорной строгости и не добавило каких-либо новых идей по сравнению с оригинальными работами. Поэтому мы в основном ссылаемся только на те публикации, в которых достигался определенный прогресс в понимании, либо на работы имеющие обзорный характер.

Наибольшее распространение получила модель «малоуглового» рассеяния [5–8], когда в качестве невозмущенной дрейфовой траек-

тории электрона выбирается прямолинейная и все эффекты оцениваются в рамках теории возмущений вдоль этой траектории. Очевидно, что в рамках этого приближения столкновения электронов с ионами происходят в различные, некоррелированные между собой моменты времени. То есть полагается, что в случае пучка с однородным начальным распределением частиц, в том числе и по фазам поля, моменты столкновения¹ также будут равномерно распределены по периоду поля. Кроме того, в рамках этого приближения не учитывается возможность притяжения электрона к иону в процессе рассеяния. Другими словами, в рамках приближения прямолинейных траекторий предполагается, что электрон не может сильно искривить свою траекторию как в течение всего процесса рассеяния, так и до момента последнего удара. Фактически, это то же самое приближение, которое было использовано в [6].

Другая модель — низкочастотное приближение [8,9] — описывает столкновения в том числе и с большими углами рассеяния. При этом предполагается, что достаточно сильное внешнее электрическое поле ускоряет электрон до и после столкновения (кулоновское поле иона на этих стадиях считается несущественным), а в процессе мгновенного рассеяния важно только статическое поле ближайшего иона. Как и в модели прямолинейных траекторий, считается, что столкновения происходят в случайные моменты времени. Поскольку в такой постановке задачи вклад от рассеяния на большие углы мал, то результат с логарифмической точностью совпадает с результатом малоуглового приближения.

Квантовая модель (борновское приближение по кулоновскому потенциалу [10]) приводит к тем же результатам, что и предыдущие два приближения, в силу учета эффектов первого порядка в квазиклассическом разложении.

Таким образом, во всех перечисленных выше приближениях получались результаты, различающиеся только логарифмическими множителями. Основной причиной подобного совпадения, по-видимому, были общие для всех предположения о некоррелированности моментов столкновений и невозможности электрону искривить свою дрейфовую траекторию (притянуться к иону) в процессе многократных осцилляций около иона. Особенно наглядно это продемонстрировано в недавних работах [7], где автор непосредственно из кинетическо-

¹Ниже под моментом столкновения будет пониматься момент наиболее близкого подхода электрона к иону при рассеянии. Само это столкновение будем называть последним столкновением, или последним ударом.

го уравнения с интегралом столкновений в форме Ландау² получает, опять-таки с логарифмической точностью, те же результаты (для эффективной частоты столкновений, эффективности генерации гармоник и т. д.), что и в цитированных работах. По-видимому, совпадение результатов, даваемых тремя различными, на первый взгляд, приближениями, стало причиной угасания интереса к этой тематике более чем на тридцать лет.

2. Динамика частиц в сильных полях

Какие же эффекты опускаются в традиционных моделях? Рассмотрим, как описывается столкновение традиционно. Напомним, что в отсутствие кулоновского поля иона движение частицы представляет собой суперпозицию из осцилляций в поле электромагнитной волны с осцилляторной скоростью $v_{osc} = eE/m\omega_0$ (ω_0 и E – частота и амплитуда поля электромагнитной волны) и дрейфа с постоянной дрейфовой скоростью v . Во всех перечисленных моделях полагается, что дрейф происходит по прямой линии и осуществляет лишь «доставку» частиц в область взаимодействия, а затем происходит собственно столкновение частицы с ионом. В разных моделях это столкновение описывается с разной степенью точности (с учетом или без учета близких столкновений, когда рассеяние происходит на большие углы, квантово-механически или в рамках классической динамики). Поскольку частицы считаются равномерно распределенными в пространстве перед столкновением, то вклад частиц, рассеивающихся на большие углы, мал. Действительно, из задачи Резерфорда известно, что из всех частиц, имевших начальную скорость v , на большой угол рассеиваются только частицы, пролетающие мимо иона на расстоянии меньшем резерфордовского радиуса $b_v = e^2 Z/mv^2$. Поскольку в сильных полях считается, что рассеяние происходит с осцилляторной скоростью $v_{osc} \gg v$, а соответствующий резерфордовский радиус $b_{osc} = e^2 Z/mv_{osc}^2$ мал, то вклад рассеяния на большие углы мал. После столкновения частица считается покинувшей область рассеяния.

Численное моделирование показывает другую динамику рассеяния частицы. Это демонстрирует рис. 1, на котором приведена дрейфовая траектория электрона. В сильных полях из-за большого радиуса осцилляций $r_{osc} = eE/m\omega_0^2 \gg b_{osc}$ частица многократно возвращается к иону и испытывает много далеких (иногда их называют

²Напомним [6], что вывод интеграла столкновений Ландау базируется на предположениях о равномерности моментов столкновений и определяющем вкладе дальних (почти прямолинейных) рассеяний.

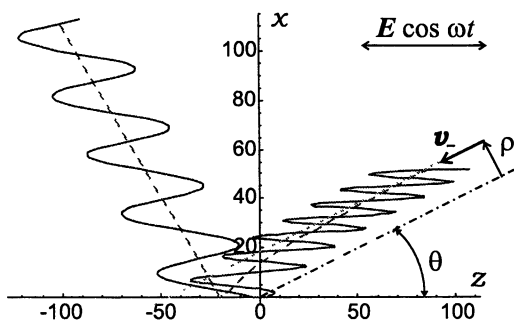


Рис. 1. Типичная траектория электрона (сплошная линия), дрейфовая траектория (пунктир) и параметры рассеиваемого электрона (v – скорость, ρ – прицельный параметр, θ – угол между скоростью и полем E). Точками изображена дрейфовая траектория в «прямолинейном» приближении.

«мягкими») столкновений за время рассеяния. Дрейфовая энергия частицы при этом практически не меняется, но происходит существенное искривление *дрейфовой траектории*³. В результате перед последним («жестким») ударом, фактически и изменяющим энергию частиц, электроны оказываются значительно ближе к иону, чем при прямолинейном дрейфовом движении. Вследствие этого изменение энергии частиц, находящихся вначале далеко от иона, при рассеянии существенно возрастает и появляется множество других новых эффектов. Именно анализу этих эффектов при электрон-ионных столкновениях в сильных электромагнитных полях с учетом искривления дрейфовых траекторий частиц при рассеянии посвящена эта работа.

Здесь в основном рассматриваются процессы столкновений в нерелятивистских полях. Такое ограничение оправдано существенными упрощениями в аналитике. Исключение составляют разделы, связанные с обсуждением результатов экспериментов (п. 7), в которых реализуются и релятивистские уровни интенсивностей. Кроме этого, мы ограничимся классическим описанием. Краткий анализ квантовых поправок вынесен в заключительную часть.

³ Дрейфовой траекторией частицы назовем траекторию, получающуюся вычитанием осцилляций частицы из ее траектории в лабораторной системе координат.

3. Приближение парных соударений

В этом разделе мы обсудим, как в сильных полях задача об отыскании интеграла электрон-ионных столкновений в уравнении для одночастичной функции распределения в разреженной плазме сводится к задаче о рассеянии пучков тестовых электронов на одном ионе. Поскольку этот результат представляется вполне очевидным, мы приведем лишь краткий вывод с целью продемонстрировать, что так называемое приближение парных столкновений отнюдь не эквивалентно приближению малоуглового рассеяния

Наиболее простым способом получения уравнения для одночастичной функции распределения электронов является метод «цепочки» Боголюбова [6]. Как известно, идея метода состоит в последовательном переходе от многочастичной функции распределения для всех частиц плазмы к одночастичной, двухчастичной и т. д. При этом в разреженной плазме достаточно ограничиться двухчастичной функцией распределения, т. е. оборвать цепочку уже на втором шаге. Эта процедура хорошо известна и подробно описана во многих учебниках. Наше замечание состоит в том, что аналогичные расщепления корреляций удается сделать и для плазмы в сильном поле электромагнитной волны. Наиболее легко эту процедуру выполнить в дрейфовых координатах. Их удобство состоит в том, что при переходе к дрейфовым координатам левая часть кинетического уравнения становится не зависящей от внешнего поля и, соответственно, времени. Вся информация о внешнем сильном поле содержится в интеграле столкновений, который становится явно зависящим от времени:

$$St[f] = \frac{\partial}{\partial p_i} \int \frac{\partial U}{\partial r_i} G(\mathbf{r}, \mathbf{p}, t; \mathbf{r}', \mathbf{p}', t') \frac{\partial U}{\partial r'_j} \frac{\partial f}{\partial p'_j} d^3 r' d^3 p' dt' d^3 r. \quad (1)$$

Здесь $U(\mathbf{r}, t) = e^2 Z / |\mathbf{r} - \mathbf{r}_{osc}(t)|$ — осциллирующий кулоновский потенциал иона.

При выводе этого выражения использованы обычные условия. А именно, как и в задаче без полей, считается, что влияние других ионов на процесс рассеяния пренебрежимо мало, т. е. что плазма достаточно редкая. Другими словами, в объеме взаимодействия v_{int} должно быть мало частиц:

$$nV_{int} \ll 1. \quad (2)$$

В плазме со слабыми полями это условие сводится к условию того, что в дебаевской сфере было много частиц.

Физический смысл интеграла столкновений (1) обычный. Он отвечает за изменение дрейфовых импульсов и кинетической энергии

электронов при рассеянии на ионах в поле электромагнитной волны. Само это поле предполагается плавным в пространстве на масштабе столкновения и медленно изменяющимся во времени. Это обычно всегда выполнено, поскольку всюду речь идет о квазимонохроматических полях. Следует также отметить, что так же, как и в интеграле столкновений Больцмана, стоящая в ядре функция $G(\dots)$ — не что иное, как функция Грина задачи рассеяния пучка электронов с фиксированным на бесконечности начальным дрейфовым импульсом на одном ионе. Подробнее эта задача будет рассмотрена ниже. Здесь же заметим, что результаты предшественников наиболее просто классифицируются в терминах данной функции Грина. В частности, приближение прямолинейных траекторий (иногда не совсем точно называемое малоугловым) соответствует тому, что в функции Грина начальный и конечный импульс считаются равными, а дрейфовая траектория прямолинейной. Соответственно, в приближении мгновенного удара [9] при подсчете среднего за период энергообмена с полем волны также считается, что частицы до удара движутся равномерно и прямолинейно. Различие состоит лишь в более точном учете данных рассеяния для малых прицельных параметров. Поскольку в этих моделях основной вклад в энергообмен вносят дальние столкновения (напомним, что в обеих моделях прицельный параметр при рассеянии не изменяется), то вклад близких столкновений с малыми прицельными параметрами не существен и результаты этих моделей совпадают.

Теперь нам ясно, чем могут быть вызваны новые результаты. Поскольку электроны осциллируют вблизи иона много периодов, может оказаться существенным притяжение и, как следствие, изменение плотности тестовых электронов вблизи иона. Второй, менее очевидный результат состоит в том, что сами фазы столкновений перестают быть равновероятными. То есть процесс притяжения сопровождается группировкой электронов.

4. Сильное поле

Обсудим теперь, при каких интенсивностях полей указанные эффекты оказываются существенными. Или, что же следует понимать под сильными полями?

Для ответа на эти вопросы надо проанализировать функцию Грина G в (1), которая определяется траекториями тестовых частиц. Последние описываются уравнением Ньютона (нерелятивистский случай):

$$\ddot{r} = -\frac{\partial}{\partial r} \frac{1}{|r - r_{osc}|}, \quad r_{osc} = -\frac{1}{\Omega^2} \cos \Omega t. \quad (3)$$

Это уравнение записано в безразмерном виде после введения характерных масштабов

$$r_E = \sqrt{\frac{eZ}{E}}, \quad \omega_E = \sqrt[4]{\frac{eE^3}{m^2Z}}, \quad v_E = \sqrt[4]{\frac{Ze^3E}{m^2}}. \quad (4)$$

Единственным безразмерным параметром⁴, входящим в уравнение (3) и характеризующим поле накачки, является безразмерная частота:

$$\Omega = \frac{\omega_0}{\omega_E} = \left(\frac{\omega_0^4 m^2 Z}{e E^3} \right)^{\frac{1}{4}}. \quad (5)$$

В параметр Ω амплитуда E и частота ω_0 входят в комбинации ω_0^4/E^3 . Это означает, что столкновения в сильных высокочастотных полях (например, лазерного диапазона) подобны столкновениям в более слабых низкочастотных полях (например, микроволнового диапазона)⁵. Это означает также, что структура фазового пространства определяется только этим параметром и принципиально различна в областях высокочастотного (слабого, $\Omega \gg 1$) и низкочастотного (сильного, $\Omega \ll 1$) поля. Как изменяется эта структура при изменении параметра?

При больших значения Ω (слабых полях) заметные изменения в структуре фазового пространства будут только в малой окрестности сепаратрисы, отделяющей пролетные гиперболические траектории от периодических. По мере роста поля (уменьшении параметра Ω) объем стохастического слоя увеличивается и при $\Omega \ll 1$ становится порядка $1/\Omega \gg 1$. В размерных переменных области стохастичности это соответствует осцилляторному радиусу в координатном пространстве и осцилляторному импульсу в импульсном пространстве. Качественно эта эволюция изображена на рис. 2.

С точки зрения задачи рассеяния это означает, что в слабых полях ($\Omega \gg 1$) или пока дрейфовая (тепловая) скорость электрона велика по сравнению с осцилляторной, динамика частиц регулярная, и энергообмен с полем может быть рассчитан по модели Ландау. Напротив, в случае сильного поля ($\Omega \ll 1$) в условиях, когда осцилляторная скорость существенно больше дрейфовой, область нерегулярной динамики в фазовом пространстве системы становится определяющей.

⁴Это верно, конечно же, только в рамках применимости уравнения (3), т. е. для нерелятивистского случая.

⁵Очевидно, что соответствующим образом должны быть пересчитаны и другие параметры плазмы: температура и концентрация.

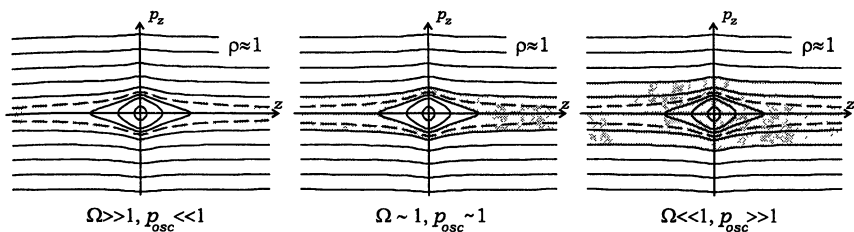


Рис. 2. Изменение области стохастичности (качественно) в фазовом пространстве при изменении параметра Ω . Пунктиром изображена сепаратричная кривая.

Уже эти простые соображения показывают, что в ранних работах 60—70-х годов эти эффекты не могли быть еще учтены адекватно, поскольку само понятие стохастичности еще только зарождалось. Кроме этого, и для численного решения эта задача была еще весьма трудной.

5. Численные результаты

Что же дают численные расчеты? На рис. 3 приведены дрейфовые траектории электронов для продольного (рис. 3, *a*, начальная скорость вдоль поля) и поперечного рассеяния (рис. 3, *b*). Видно, что эти траектории существенно зависят от фазы влета (фазы поля в момент старта). Особенно сильно эта зависимость проявляется для малых прицельных параметров. Так, если далекие частицы движутся почти по одной линии и хорошо подходят для описания в рамках малоуглового приближения (см. рис. 3, *a*), то уже на границе области притяжения (области начальных координат, из которой частицы попадают в малую окрестность иона $\sim b_{osc}$) наблюдается сильный разброс траекторий. Внутри области притяжения траектории ведут себя стохастическим образом.

Из рис. 3 видно, что наряду с «лобовыми» столкновениями (частицы, летящие точно вдоль оси Oz на расстояниях порядка $b_{osc} = \Omega^2$) изменение энергии того же порядка испытывают и частицы с прицельными параметрами, много большими b_{osc} . В традиционных моделях мы бы увидели заметное рассеяние только при лобовых столкновениях. Из рис. 3 видно и второе следствие притяжения: зависимость изменения энергии от прицельных параметров не является взаимно однозначной функцией. В сильных полях при $v < v_{osc}$ реализуется ситуация, когда в одну и ту же точку пространства (в частности, в область b_{osc} вблизи иона) приходят частицы с существенно разными

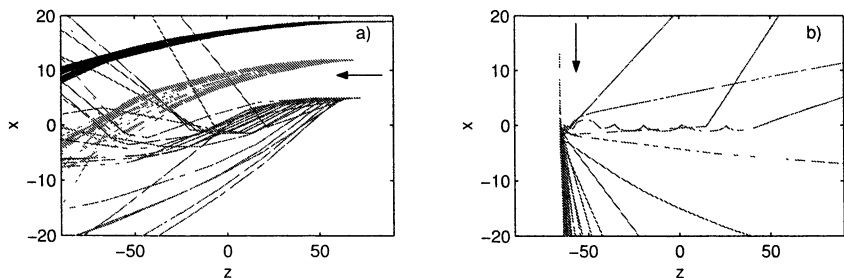


Рис. 3. Дрейфовые траектории частиц при одинаковых начальных прицельных параметрах, но разных фазах поля для $\Omega = 0, 125$ и $v = 0,1v_{osc}$. Стрелка указывает направление начальной скорости частиц. Разному цвету соответствуют траектории с разными прицельными параметрами. Видно появление сложной (стохастической) динамики при малых прицельных параметрах.

прицельными параметрами. То есть реализуется многопотоковость, отсутствующая в слабых полях. Наряду с этим возникает стохастичность — малое изменение начальных параметров приводит к сильному разбеганию траекторий и различию конечных скоростей частиц.

Отметим, что граница области притяжения одновременно является и границей сильного изменения энергии частиц (рис. 4)[11]. Действительно, для того чтобы сильно изменить свою энергию частице надо приблизиться к иону, но частицы вне области притяжения к иону близко подлететь не могут, следовательно, они не могут сильно изменить и свою энергию. Микрохарактеристикой влияния притяжения является изменение функции распределения тестовых частиц от координат. Напомним, что в традиционных моделях эта функция предполагалась константой. В расчетах же эта функция оказалась пропорциональной $\sim 1/\rho$ (ρ — поперечная координата перед жестким столкновением). Это интегрируемая сингулярная зависимость, приводящая к усилению роли близких столкновений и, как следствие, увеличению всех энергетических процессов (джоулевого нагрева, генерации гармоник излучения и быстрых частиц).

Еще одна важная для нас информация приведена на рис. 4, *с*. Речь идет о зависимости фазы жесткого (последнего) столкновения от фазы влета и прицельного параметра. Достаточно неожиданный результат заключается в том, что процесс притяжения сопровождается заметной группировкой электронов [12]. Для количественной оценки этого эффекта мы нашли фазу поля в момент, когда электрон ближе всего находится около иона в течение всего процесса рассеяния (фаза

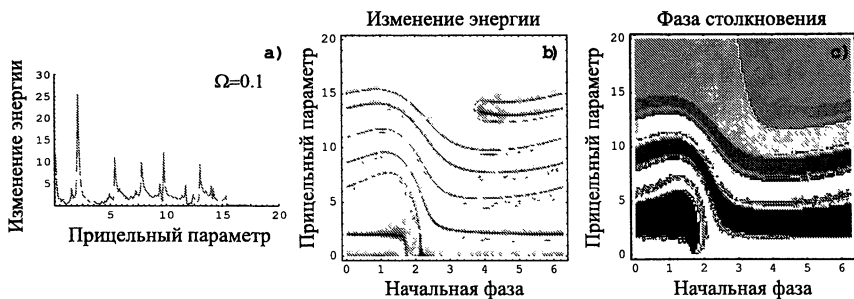


Рис. 4. Зависимость изменения энергии от прицельного параметра (a), усредненная по фазе влета; изменение энергии (b) и фаза жесткого (последнего) столкновения (c) в зависимости от прицельного параметра и начальной фазы для продольного падения. Более темный цвет соответствует большему значению величины.

жесткого столкновения). Замечательный результат состоит в том, что все электроны, испытывающие заметный энергообмен (сравни с рис. 4, b) имеют близкие фазы жесткого удара. Более внимательный анализ этих фаз приводит к выводу, что они соответствуют максимумам осцилляторной скорости.

В заключение этого раздела отметим, что даже для $\Omega = 0,1$ проведение численных расчетов полной задачи рассеяния (для всех углов между полем и пучком) — занятие весьма длительное и требующее предельных возможностей современных персональных компьютеров. В то же время, скажем, для потоков 10^{18} Вт/см² при длине волны 1 мк, интересующий нас малый параметр $\Omega = 0,01$. Поэтому особую ценность приобретают аналитические результаты.

6. Аналитика

Посмотрим, что удастся сделать в этом направлении. Прежде всего заметим, что задача рассеяния (3) описывается гамильтоновскими уравнениями с периодическим во времени потенциалом. Это означает, что можно перейти к точечным отображениям для дрейфовых координат и импульсов частицы через период. Следовательно, в силу гамильтоновости, проблема сводится к отысканию одной скалярной функции — производящей функции отображения. Из теоретической механики хорошо известно, что для отыскания производящей функции достаточно вычислить действие на одном периоде и выразить его

как функцию, например, начальных координат и конечных импульсов $S = S(\mathbf{r}_{ini}, \mathbf{p}_{out})$. При этом оставшаяся пара канонических координат, соответственно, равна:

$$\mathbf{p}_{ini} = \frac{\partial S}{\partial \mathbf{r}_{ini}}, \quad \mathbf{r}_{out} = \frac{\partial S}{\partial \mathbf{p}_{out}}. \quad (6)$$

Следует отметить, что на одном периоде мы можем использовать различные приближения.

Ряд важных свойств отображения (6) можно обнаружить в важном частном случае квазипродольного распространения в пренебрежении смещением частицы поперек электрического поля из-за дрейфовой скорости $v_{\perp} \ll 2\pi/\rho v_{osc}$ ⁶. При этом отображение распадается на сдвиговую часть для продольных координаты и импульса и отображения для поперечной координаты $\rho = \sqrt{x^2 + y^2}$:

$$\bar{\rho} = \left| \rho - \frac{2\pi}{\rho} \right|. \quad (7)$$

Графически отображение (7) представлено на рис. 5. Отображение (7) не описывает рассеяние частиц на большие углы и их уход из области рассеяния. Для учета таких частиц можно ввести область потерь с шириной порядка $b_{osc} = \Omega^2 \xrightarrow{\Omega \rightarrow 0} 0$ (заштрихованная область на рис. 5), при попадании в которую частицы считаются ушедшими из области рассеяния.

Рассмотрим свойства отображения (7). Во-первых, необходимо отметить появление многопоточности, состоящей в том, что в одну и ту же точку пространства из-за притяжения одновременно приходят частицы, стартовавшие с разных прицельных параметров. Простейший пример — наряду с частицами, имеющими малые начальные параметры $\rho_{ini} \leq b_{osc}$, в эту же окрестность попадают частицы, пролетевшие мимо иона один раз, рассеявшиеся на малый угол и попавшие в ион при возврате (рис. 5). Прицельный параметр ρ_1 , соответствующий этим частицам легко найти: $\rho_1 = \sqrt{2}a\pi$. Аналогично, существуют частицы, совершившие две, три и более осцилляций с прицельными параметрами ρ_1, ρ_2, \dots . Всего таких групп частиц будет $N \approx r_{osc}/vT \sim v_{osc}/v$. При этом все частицы с $\rho < \rho_N$ в конечном итоге оказываются в области $\bar{\rho} \leq \rho_1$.

Вторым важным свойством отображения (7) является появление сингулярности в функции корреляции. Как уже отмечалось выше, все

⁶При достаточно малой дрейфовой скорости все качественные эффекты, обсуждаемые в этом параграфе, будут иметь место и для ЦП.

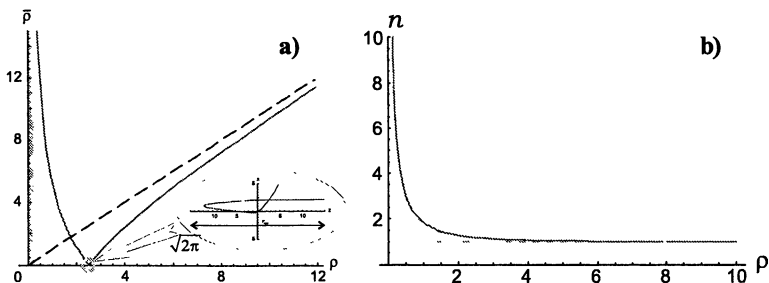


Рис. 5. Зависимость прицельного параметра $\bar{\rho}$ через итерацию отображения (7) от прицельного параметра ρ (а). Заштрихована область потерь частиц при рассеянии на большие углы. Особо выделена точка, из которой частицы попадают точно в ион через одну итерацию отображения. Справа (б) изображена зависимость концентрации частиц от прицельного параметра ρ для: невозмущенного распределения (пунктир) и предельного значения через бесконечное количество итераций (сплошная линия), являющегося инвариантной мерой отображения (7).

частицы с прицельными параметрами $\rho \leq \rho_{att}$ стягиваются в процессе рассеяния в область $\rho \leq \rho_1$. Каково же распределение частиц перед «последним» столкновением, т. е. в диапазоне $[0, \rho_1]$? Для ответа на этот вопрос найдем изменение концентрации (или, что то же самое, пространственной плотности функции корреляции) через итерацию отображения (7) из закона сохранения числа частиц:

$$\bar{n} \bar{\rho} d\bar{\rho} d\bar{z} = n \rho d\rho dz.$$

Выражая из отображения (7) старую координату ρ через новую $\bar{\rho}$ и используя неравенство $\rho > \rho_1$ (рассматриваются частицы, еще не попавшие в область $[0, \rho_1]$), находим (см. рис. 5, б)⁷:

$$\bar{n} = n(\rho) \frac{\rho^3}{\bar{\rho}(\rho^2 + 2\alpha\pi)}, \quad \rho(\bar{\rho}) = \frac{1}{2} \left(\bar{\rho} + \sqrt{\bar{\rho}^2 + 8\alpha\pi} \right). \quad (8)$$

В частности, при больших прицельных параметрах $\bar{\rho} \gg \sqrt{8\alpha\pi} = 2\rho_1$, как и следует ожидать, концентрация практически не меняется. Од-

⁷В действительности норму отображения (7) легко найти, и она равна $1/\rho$. Однако, как уже говорилось ранее, это отображение с потерями: частицы, подлетевшие близко к нулю, уходят из области рассеяния. Таким образом, в таком отображении можно ввести только норму по отношению к пролетным частицам. Именно такая ситуация изображена на рисунке 5, б — распределение частиц, еще не попавших в область $(0, \rho_1)$.

нако при малых прицельных параметрах в концентрации появляется особенность (ср. рис. 4):

$$\bar{n} \underset{\bar{\rho} \ll \rho_1}{\simeq} n \frac{\rho_1}{2\bar{\rho}}. \quad (9)$$

Учет $N \gg 1$ пролетов частиц приводит к распределению концентрации перед последним ударом вида

$$n \approx n_0 \left(1 + \frac{a}{\rho} \right), \quad (10)$$

где коэффициент $a \simeq b_v$ должен определяться из полного отображения (в том числе и для поперечного падения пучка).

Третье свойство отображения (7) – появление стохастической динамики. В предыдущем пункте уже было отмечено, что динамика рассеяния частиц в сильном поле накачки оказывается очень сложной, стохастической. Этот же факт демонстрирует отображение (7). Действительно, с одной стороны, оно является глобально притягивающим⁸. С другой стороны, центральная часть отображения является локально неустойчивой. В результате происходит перемешивание частиц. Это и есть один из критериев стохастической динамики. Кулоновский центр при этом играет роль «организатора» сложного стохастического аттрактора.

Необычность кулоновского аттрактора состоит в постоянной потере частиц благодаря рассеянию на большие углы. Частицы, попавшие вследствие притяжения к иону при многократных осцилляциях (при регулярной динамике) или вследствие диффузии Арнольда (стохастического «блуждания» частицы при стохастической динамике) в малую окрестность иона, испытывают сильноугловое рассеяние и набирают скорость, сравнимую с осцилляторной скоростью. В результате они, с одной стороны, вылетают из области аттрактора, а с другой стороны, дают вклад в столкновительный ток и, соответственно, в когерентное излучение. Поскольку при рассеянии на большие углы частицы получают скорость, направленную против мгновенного направления осцилляторной скорости, следует ожидать появления интенсивной генерации столкновительного тока и, соответственно, когерентного излучения, растущего с ростом области стохастичности, т. е. с ростом интенсивности поля накачки.

⁸Необходимо отметить, что попадание частицы в область $\rho \leq \sqrt{b_{osc} r_E}$ сопровождается сильным изменением ее скорости и возможным уходом из области рассеяния, так что отображение (7) перестает быть применимым.

7. Экспериментальные следствия

Обсудим результаты некоторых экспериментов. Ниже будет показано, что в режиме сильного поля эффективности всех процессов в плазме — джоулева нагрева, генерации гармоник, образования «хвостов» быстрых частиц — не убывают и слабо зависят от поляризации поля. Мы приведем также результаты известных экспериментов по взаимодействию интенсивного лазерного излучения с плазмой, которые сравнительно просто объясняются развитой выше теорией.

Начнем с анализа эффективности джоулева нагрева и генерации быстрых электронов при столкновениях. В недавних экспериментах с петаваттным лазером в США [13] с энергией импульса до 50 Дж при длительностях 0,5–10 пс и интенсивностях излучения 10^{18} – 10^{21} Вт/см² при взаимодействии с заранее приготовленной плазмой с концентрацией 10^{17} – 17^{19} см⁻³ — было обнаружено, в частности, что при таких интенсивностях нагрев основной массы электронов слабо зависит от поля и в основном определяется только длительностью импульса излучения. Кроме того, было обнаружено, что процесс взаимодействия сопровождается образованием сравнительно большого числа быстрых электронов с суммарной энергией до десятков процентов от энергии лазерного импульса. То есть было обнаружено, что излучение хорошо поглощается, причем энергия быстрых частиц сравнима с тепловой энергией основной массы. Именно на этом, фактически, основаны все дальнейшие результаты этой группы и, по видимому, результаты японских исследователей [14] по обнаружению быстрых протонов и нейтронов при таких взаимодействиях.

Откуда же берутся быстрые электроны, и почему поглощение столь заметно? Авторы либо произносят общие слова о сильной турбулентности, либо осторожно говорят о непонятности этих эффектов. Посмотрим, что предсказывает наша теория. На рис. 6 приведены экспериментальные результаты нескольких работ [13, 14] и сплошной линией показан результат наших оценок для нормированного на квадрат концентрации плазмы числа быстрых электронов от нормированного импульса электрона pc . Мы видим хорошее совпадение в диапазоне энергий быстрых частиц от десятков килоэлектронвольт до 30 МэВ.

Следует отметить, что авторы [14] в своих работах говорят о зависимости от энергии электронов. Мы же построили зависимость от произведения импульса на скорость света. Это одно и то же для релятивистской области энергий (при интенсивностях больших 10^{18} Вт/см²), но принципиально различно при малых энергиях. Поэтому включение результатов [14] в график является определенным произволом с нашей стороны. Мы сделали это, поскольку только в этой работе

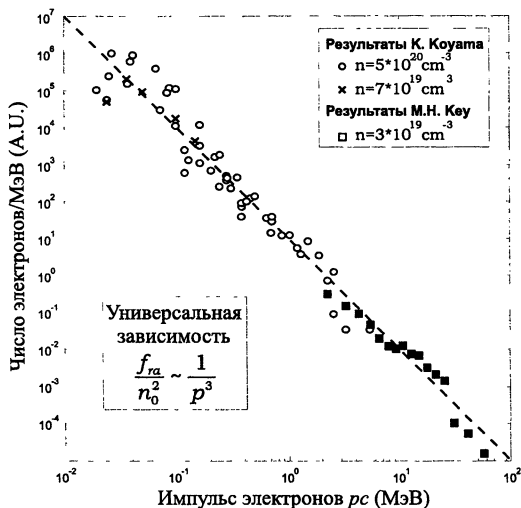


Рис. 6. Сравнение экспериментальных [13, 14] и теоретической (пунктир) зависимостей распределения горячих электронов по импульсу.

прослеживается зависимость от концентрации плазмы. Кроме того, не исключена простая описка. Дело в том, что авторы используют схему с многоканальным сцинтиллятором, в которой число фотонов в каждом канале при преобразовании в видимое излучение действительно определяется энергией каждого электрона. Но само число быстрых электронов, попавших в канал, определяется только его продольным импульсом (и в релятивистском, и в нерелятивистском случаях). Дело в том, что для попадания электронов в канал авторы используют малоугловое отклонение электронов в промежутке с магнитным полем. Но малый угол отклонения (номер канала) в этом случае зависит только от продольного импульса.

Нетрудно оценить и эффективность джоулевого нагрева. С учетом эффекта притяжения джоулевый нагрев, обусловленный столкновениями электронов с ионами, монотонно растет с ростом поля и в сильных полях определяется простым соотношением:

$$Q \approx 4\pi n^2 m v_{osc}^2 \left(2v_{osc} b_{osc}^2 \ln \frac{r_a}{b_{osc}} + v b_v b_{osc} \right). \quad (11)$$

То есть в сильных полях он действительно почти не зависит от интенсивности (ею определяется только размер области, занятой сильным

полем, и степень ионизации), и, следовательно, температура определяется только длительностью импульса (рис. 7, а). В частности, для экспериментальных условий [13] мы действительно получаем температуру порядка нескольких сотен электронвольт.

Еще одна группа экспериментов состоит в наблюдении высоких гармоник в разреженной плазме с сильным лазерным полем (см. обзор [3]). Авторы наблюдают непонятно большие эффективности излучения на гармониках (вплоть до 11–13) основной частоты при интенсивностях порядка 10^{18} Вт/см². Если вначале они произносили слова о сильной турбулентности, то в последнее время стали полагать, что излучение связано с гармониками в движении отдельных электронов в сильных полях (так называемое нелинейное излучение свободных электронов, нелинейное томсоновское рассеяние). Эта точка зрения представляется нам неоправданной.

Остановимся на этом вопросе подробнее. Действительно, в столь сильных полях магнитное поле волны становится существенным, и каждый электрон в поле линейно-поляризованной электромагнитной волны движется по так называемой фигуре в форме восьмерки [15]. То есть релятивизм существен и гармоники в движении отдельного электрона велики. Однако в плазме, когда на длине волны λ много электронов (в условиях экспериментов $n = 10^{19}$ см³ и $\lambda^3 \simeq 10^{-6}$ см⁻³), при отыскании плотности тока этот эффект подавляется из-за модуляции плотности электронов (релятивистского преобразования масштабов). Этот факт известен [15], но, по-видимому, забыт к настоящему времени. Проще всего его напомнить для нашего случая холодной плазмы в рамках релятивистской гидродинамики. Действительно, полагая, что плотность дрейфовых центров n_0 осциллирующих электронов неизменна, для плотности поперечного тока в поле волны имеем

$$j_{\perp} = env_{\perp}. \quad (12)$$

Здесь $v_{\perp} = \frac{e}{c} A \cdot \exp i(\omega t - kx) \cdot (1 - v_x/c)$. Видно, что из-за релятивизма спектр скорости сильно обогащается. А как обстоят дела со спектром тока? Из уравнения непрерывности

$$\frac{\partial n}{\partial t} + \operatorname{div} j = 0 \quad (13)$$

имеем $n = n_0 / (1 - \frac{v_x}{c})$. То есть концентрация также оказывается промодулированной. Окончательно для плотности тока имеем

$$j_{\perp} = \frac{e}{c} n_0 A \cdot \exp i(\omega t - kx) ! \quad (14)$$

Таким образом, гармоники могут появиться лишь в порядках n/n_{cr} . То есть интенсивность излучения свободных электронов на гармониках очень мала.

А что дают столкновения? Для понимания очень важен не только эффект притяжения, но и эффект группировки (п. 5). Поскольку мы убедились, что электроны не просто притягиваются к электрону, а еще и группируются по фазе поля, ясно, что они могут хорошо излучать на гармониках. На рис. 7, *b* приведена зависимость сечения когерентного излучения χ_{coh} , характеризующая интенсивность когерентного излучения (подробнее см. [12]), от номера гармоники. Это сечение является основной характеристикой когерентной компоненты столкновительного тока.



Рис. 7. Качественная зависимость темпа нагрева от интенсивности накачки (*a*); зависимости сечения когерентного излучения χ_{coh} от номера излучаемой гармоники (*b*) для разных частот поля накачки Ω .

Отметим, что в случае изотропной дрейфовой функции в распределении электронов наблюдаются только нечетные гармоники. Из рис. 7 видно, что когерентное излучение должно пропадать в слабых полях ($\Omega \ll 1$) и при больших температурах ($v_{osc} < v$). Особенно интересен факт, что излучение должно наблюдаться с не меньшей интенсивностью и при циркулярной поляризации накачки.

8. Заключение

Таким образом, в сильных электромагнитных полях в плазме основные энергетические процессы (джоулев нагрев, генерация быстрых электронов, когерентное излучение на гармониках накачки) ста-

новятся все более интенсивными. Основным параметром, определяющим понятие сильного поля, является

$$1/\Omega = \sqrt[4]{\frac{eE^3}{m^2 Z\omega^4}} \simeq 10 \frac{(P/10^{16})^{3/8}}{Z^{1/4}} (\lambda) \gg 1.$$

Здесь в размерной оценке мощность измеряется в вольтах на сантиметр квадратный, а длина волны в микронах. Видно, что если в лазерном диапазоне поле является сильным только, как минимум, для тераваттных лазеров, то в микроволновом диапазоне длин волн мощность может быть существенно меньше.

Важными условиями применимости модели парных соударений в сильных полях являются условия на длительность импульса и концентрацию плазмы. Условие на длительность импульса получается из требования, чтобы электрон успел проосциллировать вблизи иона нужное число раз. Поскольку основной вклад во все интегральные характеристики рассеяния дают электроны с дрейфовыми скоростями, направленными перпендикулярно к полю волны накачки, то достаточно, чтобы импульс был всего несколько периодов. Аналогично, условие на плотность (точнее на разреженность) плазмы состоит в том, чтобы объем взаимодействия при столкновениях был мал по сравнению с межчастичным объемом ($nV_{int} \ll 1$). Учитывая, что основной вклад в столкновения дают поперечные рассеяния, и следовательно $V_{int} \approx b_j^2 \cdot r_{osc}$, для концентрации имеем следующую оценку: $n < 10^{19} T^2 \omega^2 / Z^2 \sqrt{P}$. Здесь, как и всюду в оценках, температура в электронвольтах, частота в 10^{15} рад/с, P в 10^{16} Вт/см². Видно, что в микроволновом диапазоне длин волн это условие достаточно трудно выполнимо. Более слабая оценка получается для границы реализации корреляционных эффектов ($r_E \ll r_{de}$). Для этого достаточно $n_e \ll 9 \cdot 10^{22} T \sqrt{P} / Z^2 \text{см}^{-3}$.

Еще один важный вопрос, который мы собирались обсудить, почему в сильных полях можно пользоваться классическим описанием? Дело в том [16], что в силу исключительности кулоновской задачи рассеяния, в которой дифференциальное сечение рассеяния и в квантовой, и в классической задачах полностью совпадают, для применимости классического описания достаточно, чтобы дебройлевская длина волны электрона, оцененная по осцилляторной скорости, была мала по сравнению r_E -размером области, из которой и происходит рассеяние. Нетрудно видеть, что с ростом поля это условие становится все более слабым: $P \gg 2 \cdot 10^{-6} \omega^4 / Z^2$.

В заключение отметим, что за неимением места мы обсудили только первоначальные следствия эффектов столкновений в сильных

полях. Прежде всего за кадром остались очень важные вопросы об электродинамических явлениях в плазме с сильными столкновениями. Дело в том, что сама по себе плазма в сильном поле является существенно неравновесной средой и наличие когерентных столкновительных эффектов должно и приводит к специфическим диссипативным неустойчивостям в такой среде.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты 01-02-16575, 02-02-06005).

Литература

1. Marcuse D. // Bell. Syst. Tech. J. 1962. **41**. 1557.
2. Corkum P. B., Burnett M. H., Ivanov M. Y. // Opt. Lett. 1994. **19**. 1.
3. Chen S.-Y., Maksimchuk A., Umstadter D. // Phys. Rev. Lett. 2000. **84**. 5528.
4. Yudin G. L., Ivanov M. Y. // Phys. Rev. A. 2001. **63**. 033404.
5. Dawson J., Oberman C. // Phys. Fluids. 1962. **5**. 517; Силин В. П. // ЖЭТФ. 1964. **47**. 2254; Shvets G., Fisch N. J. // Phys. Plasmas. 1997. **4**. 428.
6. Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. М.: Наука, 1979.
7. Силин В. П. // ЖЭТФ. 1998. **114**. 864.
8. Mittleman M. H. Introduction to the theory of laser-atom interactions. Plenum Press, New York — London, 1993.
9. Kroll N. M., Watson K. M. // Phys. Rev. A. 1973. **8**. 804.
10. Бункин Ф. В., Казаков А. Е., Федоров М. В. // УФН. 1973. **107**, 559; Pert G. J. // Phys. Rev. E. 1995. **51**. 4778.
11. Fraiman G. M., Mironov V. A., Balakin A. A. // Phys. Rev. Lett. 1999. **82**. 319; Фрайман Г. М., Миронов В. А., Балакин А. А. // ЖЭТФ. 1999. **115**. 463; Fraiman G. M., Mironov V. A., Balakin A. A. // Phys. of plasma. 2001. **8**. 2502.
12. Фрайман Г. М., Миронов В. А., Балакин А. А. // ЖЭТФ. 2001. **120**. 797.
13. Key M. H., Cable M. D., et. all // Phys. Plasmas. 1998. **5**. 1966; Hatchet S.P., Brown C.G., et. all // Phys. Plasmas. 2000. **7**. 2076.
14. Koyama K., Saito N., Tanimoto M. // ICPP 2000. Quebec. Canada, 2000, V. ICPP 4051. P. MP1.067.
15. Ландау Л. Д., Лифшиц Е. М. // Теория поля. М.: Наука, 1988.
16. Ландау Л. Д., Лифшиц Е. М. // Квантовая механика. М.: Наука, 1988.

О ФРАКТАЛЬНОЙ ДИНАМИКЕ АКТИВНЫХ СРЕД

В. Ю. Трахтенгерц, Д. И. Иудин, А. Н. Григорьев

1. Введение

Фрактальная динамика активных сред привлекает в последнее время всё большее внимание, и это не случайно. Фрактальные объекты обнаруживаются в природе повсеместно. Контуры береговых линий и русла рек, растительный покров и лесные пожары, структура облачности и грозовые электрические разряды, просачивание жидкости сквозь грунты и явления агрегации, эволюция популяций и многое другое — всё это примеры фрактального поведения. В природе фрактальные структуры, как правило, являются стохастическими, т. е. присущее фракталам самоподобие в их внутреннем устройстве проявляется в виде универсальных вероятностных распределений и спектров. Именно такие «случайные» фракталы нас будут интересовать в дальнейшем. Простой и вместе с тем содержательный пример стохастической фрактальной конструкции связан с явлением протекания, или перколяции (от английского — *percolate*) [1, 2].

Перколяция характеризует широкий класс явлений в двухфазных системах с контрастными физическими свойствами компонент: протекание жидкостей через пористые среды, прохождение электрического тока в композитах металл — диэлектрик, распространение фронта возбуждения в сильно неоднородных активных средах, передачу информации через случайные каналы связи и т. п.

Возьмем в качестве примера пористую среду, представляющую собой случайно однородное распределение полостей с характерным размером a много меньшим размера системы L (представьте себе приличный кусок швейцарского сыра). Пусть p характеризует долю объема, занятого порами, тогда $1 - p$ соответствует доле объема, занятого твердой матрицей. При $p \ll 1$ отдельные поры с большой вероятностью изолированы друг от друга или, сливаясь, образуют небольшие кластеры пустот, при этом рассматриваемый нами пористый материал непроницаем. По мере роста концентрации пор p средний размер кластеров увеличивается. При концентрации, близкой к некоторой критической $p = p_c$, появляется кластер, называемый перколяционным, связывающий противоположные стороны системы. Материал становится прозрачным для фильтрационных течений. С дальнейшим ростом пористости доля объема, принадлежащая перколяционному кластеру пустот (плотность перколяционного кластера), возрастает. Со-

ответственно, средний размер конечных кластеров, не принадлежащих перколяционному кластеру, уменьшается, и при $p \sim 1$ все поры сливаются, очевидно, в единый кластер. Линейные размеры конечных кластеров ниже и выше p_c характеризуются корреляционной длиной R_c , которая определена как среднее расстояние между узлами одного и того же конечного кластера. С приближением к p_c R_c возрастает как

$$R_c \approx a|p - p_c|^{-\nu}, \quad (1)$$

с одинаковым показателем ν – критическим индексом радиуса-корреляции – ниже и выше порога.

Можно привести множество других примеров, где в определенном образом приготовленных системах перколяция по одной из контрастных по физическим свойствам компонент приводит к содержательным эффектам. Однако более интересная физическая ситуация возникает тогда, когда развитие и взаимопревращение компонент в двухфазной системе обусловлено динамикой самой системы. Естественно, что такие превращения возникают только в активных средах и сопровождаются подводом к системе свободной энергии и диссипацией. При этом ключевую роль начинает играть явление так называемой самоорганизованной критичности (СОК), когда система автоматически эволюционирует к состоянию динамического равновесия, близкому к порогу перколяции [3].

Можно заметить, что описанная выше динамическая система с взаимным превращением компонент близка к классической модели активной бистабильной среды с восстановлением (ингибитором), которая широко используется при моделировании автоволновых процессов. Ниже эта аналогия обсуждается более детально. Наша цель — определить область параметров, в которой указанная активная среда обнаруживает фрактальную, либо регулярную динамику. В качестве иллюстрации рассматривается фрактальная динамика электрических зарядов в грозовом облаке.

2. Бистабильная среда с восстановлением в представлении клеточных автоматов

Простейшей базовой моделью активных бистабильных сред с восстановлением (ингибитором) является система двух уравнений [4, 5]:

$$\begin{aligned} \frac{\partial u}{\partial t} &= f(u, v) + D\Delta u, \\ \frac{\partial v}{\partial t} &= -\epsilon(v - v_0(u)), \end{aligned} \quad (2)$$

где в зависимости от задачи u может служить локальной разностью потенциала в нервном волокне или грозовом облаке, температурой при распространении фронтов горения, концентрацией химического реагента и т. д. Ингибитор v характеризует скорость восстановления активной компоненты, причем член $v_0(u)$ определяет переход системы в исходное (пассивное) состояние по мере «выгорания» активного вещества. В большинстве реальных систем ϵ – малый параметр. Для бистабильных сред соответствующая (2) точечная система ($D = 0$) представляет собой релаксационную систему с N -образной характеристикой (рис. 1), которая имеет три состояния равновесия: два устойчивых (пассивное u_1 и критическое u_c) и одно неустойчивое (активное u_a). Качественно, действие ингибитора v приводит к перемещению кривой $f(u, v)$ вдоль вертикальной оси, как это показано на рис. 1. При этом соответствующая (2) точечная система может быть как ждущим генератором одиночных импульсов (если минимум кривой $f(u, v)$ остается в области отрицательных значений), так и автогенератором (если в процессе уменьшения ингибитора v реализуется случай, соответствующий верхней кривой на рис. 1). Учет диффузионного слагаемого в первом уравнении системы (2) приводит к генерации уединенных бегущих импульсов возбуждения в случае ждущего режима, либо незаходящей серии импульсов в случае автоколебаний. Такое решение показано на рис. 2. Ширина и скорость импульса определяется параметрами ϵ и D ($|f| \sim 1$) [4, 5].

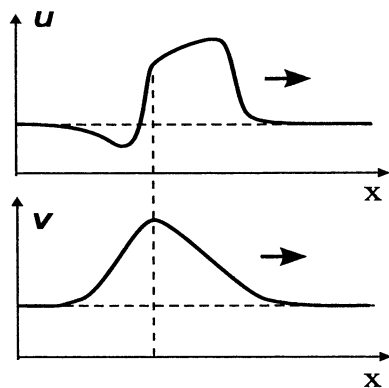
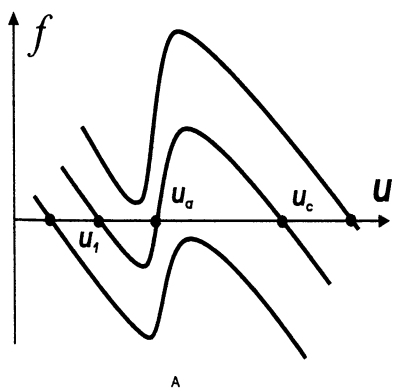


Рис. 1. Три возможные ситуации для сосредоточенного активного элемента: нижняя кривая — пассивное состояние, средняя — ждущий или бистабильный режим, верхняя — самовозбуждение и возникновение автоколебаний.

Рис. 2. Уединенный бегущий импульс в возбудимой среде. На верхнем графике показан профиль температуры, на нижнем — профиль концентрации ингибитора.

Многочисленные исследования базовой модели (2) и ее обобщений, проведенные в основном для двухмерного случая, обнаружили разнообразные типы регулярной и нерегулярной динамики в виде плоских, круговых и спиральных волн возбуждений и их модификаций, возникающих при наличии дислокаций, фазовых корректоров и т. п. Возникает вопрос, есть ли место в этой совокупности взаимодействий для фрактальной динамики в том смысле, как это сформулировано во введении? Забегая вперед, можно дать однозначный ответ, что такая фрактальная динамика возникает в определенной области управляющих параметров. Таких параметров в нашей задаче три: это параметр ϵ , характеризующий, по- существу, скорость подвода свободной энергии к системе, отношение $x_c = u_c/u_a$, определяющее относительное расстояние между устойчивым (критическим) и неустойчивым (точка активации) состояниями равновесия (см. рис. 1), и частота ξ появления внешнего (запускающего) сигнала на входе системы. Параметры ϵ и ξ безразмерные и сравниваются с обратным временем формирования импульса возбуждения. Важную роль играют и характерные пространственные масштабы системы. И наконец, мы должны учесть пространственно-временные флуктуации при заселении системы активной фазой. Роль этих флуктуаций убывает по закону больших чисел по мере заселения среды, однако они могут оказаться принципиальными при $\epsilon \rightarrow 0$.

Исследовать роль указанных факторов в рамках системы (2) чрезвычайно сложно, особенно в интересующем нас трёхмерном случае, поэтому мы перейдем к более упрощенному описанию динамики (2) на языке клеточных автоматов. Представим нашу возбудимую среду в виде трехмерной решетки клеточных автоматов. Имея в виду дальнейшие приложения развиваемой модели к грозовому электричеству, поставим каждому узлу решетки в соответствие зависящую от времени скалярную величину u , характеризующую значение электрического потенциала [6]. Эволюция потенциального рельефа на решетке определяется в рамках модели случайного роста. Именно на каждом шаге модельного времени происходит распределение по узлам решетки нормально распределенных случайных добавок к величине скалярного поля. При этом каждый узел совершает независимое от соседей броуновское движение в пространстве значений скалярного потенциала. Рост потенциального рельефа ограничен некоторым критическим значением разности потенциалов u_c между соседними узлами, по достижении которого происходит пробой — выравнивание потенциала между соседями. Предполагается, что на следующем шаге модельного времени возникший пробой может инициировать пробой соседних

связей (инфицировать соседей), если приложенная к ним разность потенциалов превышает некоторый фиксированный уровень - уровень активации u_a , значение которого меньше критического. Указанная схема взаимодействия соседних автоматов в решетке отражает, по существу, все основные свойства системы (2) в режиме автоволн. Значения $u = 0$, $u = u_a$ и $u = u_c$ соответствуют состояниям равновесия точечной системы. Роль ингибитора в представлении клеточных автоматов заключена в характере эволюции разностей потенциалов между ячейками: медленном росте с последующим резким переходом в нуль. Активация соседних ячеек соответствует передаче возбуждения посредством диффузии. При этом шаг счёта по времени и размер ячейки в системе клеточных автоматов можно связать с масштабом и скоростью движения импульса возбуждения в исходной системе (2). Следует заметить, что постановка проблемы в рамках клеточных автоматов может относиться и к более широкому кругу задач, в которых размер ячейки и механизм передачи возбуждения соседям обусловлены причинами, отличными от обсуждавшихся при рассмотрении (2).

3. Результаты компьютерного моделирования

В рассматриваемом режиме автоволн на решетке клеточных автоматов управляющими параметрами являются скорость роста ϵ интенсивности электрического поля в ячейках, определяемая соотношением

$$\langle \Delta u^2 \rangle = \epsilon t, \quad (3)$$

где $\langle \Delta u^2 \rangle$ — среднеквадратичная разность потенциалов между соседними ячейками, а также отношение $x_c = u_c/u_a$. Проведенные нами компьютерные эксперименты [6] показали, что при достаточно большом значении параметра ϵ в рассматриваемой системе клеточных автоматов реализуются хорошо известные регулярные режимы автоволн. Пример такого режима, реализованного на двумерной решетке, показан на рис. 3. Расчеты проводились на квадратной решетке 90×90 . Параметр ϵ превышал обратное время пробега волной активации линейного размера решетки. Заметим, что характерный пространственный масштаб спиралей на рис. 3 и типичное расстояние между ними обратно пропорциональны скорости роста ϵ .

Поведение системы меняется кардинально, когда ϵ становится много меньше единицы. Фрактальная динамика обнаруживается в области параметров $\epsilon \ll 1$ и $x_c \equiv u_c/u_a > 4$ как в двумерном, так и в трёхмерном случаях. После некоторого переходного процесса система

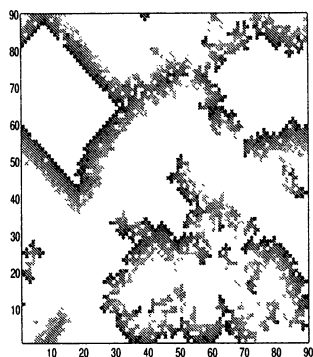


Рис. 3. Автоволны на квадратной решетке 90×90

приходит в состояние динамического равновесия вблизи порога перколяции. Этот критический уровень определяется равновесием между

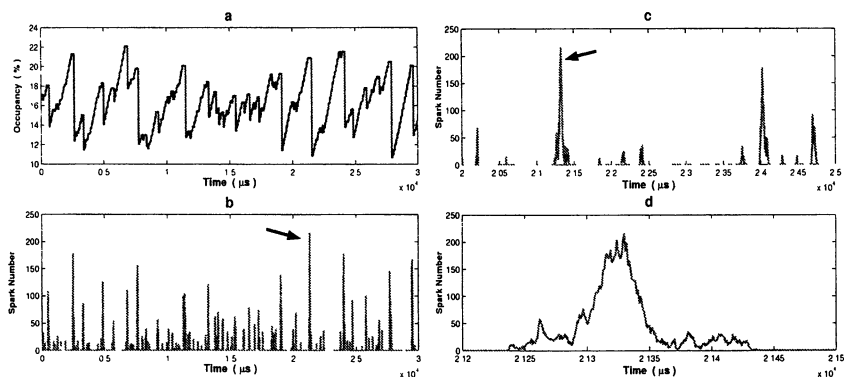


Рис. 4. Динамика в трехмерном случае: *a* — эволюция числа активированных пар, *b* — эволюция числа пробоев, *c* — *d* то же, что и *b* при последовательном увеличении разрешения по времени для эпизодов, помеченных стрелками.

процессами роста мелкомасштабного потенциального рельефа и процессами выравнивания потенциалов в соседних ячейках при возникновении пробоя между ними (рис. 4, *a*). Появление новых активированных связей компенсируется их исчезновением в процессе металлизации. Можно сказать, что происходит заселение среды активированными элементами. На больших временах при грубом временном разрешении события представляют собой дельтаобразные вспышки,

частота появления которых обратно пропорциональна их амплитуде (рис. 4, *b*). Напротив, при более высоком разрешении, на коротких временах выявляется тонкая структура каждого события, определяемая процессом выгорания отдельного кластера активированных связей (рис. 4, *c — d*). Принципиально важным оказывается тот факт, что кластеры активированных связей оказываются фрактальными структурами.

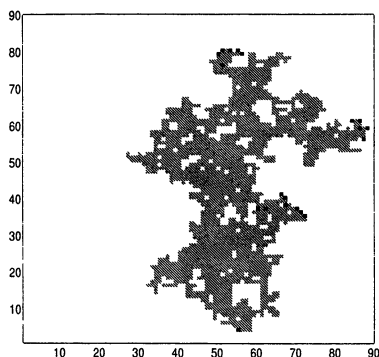


Рис. 5. Кластер активированных элементов на квадратной решетке 90×90

Мгновенная фотография процесса выгорания кластера активированных связей приведена на рис. 5. Черные пиксели на рисунке соответствуют фронту ионизации — элементам структуры, пробитым на данном шаге модельного времени. Вычисленная по результатам моделирования фрактальная размерность кластеров практически совпала с размерностью перколяционного кластера.

4. Приближение среднего поля

В этом разделе мы хотим выяснить общие критерии, обеспечивающие появление критической динамики. Численный эксперимент, например, обнаружил существование критического уровня заселенности, флуктуации вблизи которого обладают скейлинговыми свойствами. Кроме того, в эксперименте была выявлена зависимость этого уровня от числа ближайших соседей для заданной решетки автоматов, т. е. от ее координационного числа z . Чем большее число элементов находится в контакте с зараженным, тем выше вероятность размножения инфекции и тем, следовательно, ниже критический уровень

заселенности. Координационное число возрастает с увеличением размерности решетки. В пределе высоких координационных чисел для расчета критического уровня заселенности может быть использовано приближение среднего поля [7].

Введем функцию распределения связей по абсолютной величине перепада потенциала на них $f(u)$. Величина $f(u)du$ определяет долю связей с абсолютной величиной разности потенциалов в интервале $[u, u + du]$. В любой фиксированный момент времени все связи в системе можно отнести к трем типам: 1) активные, их удельное число $p = \int_{u_a}^{u_c} f(u)du$; 2) пассивные, с удельным числом $q = \int_0^{u_a} f(u)du$; и 3) «горячие», т. е. те, что возвращаются за время τ из активного состояния в пассивное с нулевой разностью потенциалов.

Эволюция функции распределения связей по величине разности потенциалов на их концах описывается следующим уравнением:

$$\frac{\partial f}{\partial t} = \epsilon \frac{\partial^2 f}{\partial u^2} + I - \Omega f, \quad (4)$$

где коэффициент ϵ играет роль коэффициента диффузии связей по величине разности потенциалов, а декремент Ω определяется соотношением:

$$\Omega = \begin{cases} 0 & 0 \leq u \leq u_a, \\ \Omega_0 & u_a \leq u \leq u_c. \end{cases} \quad (5)$$

Последний диссипативный член характеризует обусловленный микропробоями механизм ограничения роста. Понятно, что все связи, вырванные из процесса роста в интервале между уровнем активации и критическим уровнем, возвращаются через время τ в точку с нулевой разностью потенциалов:

$$I = \begin{cases} m\delta(0) & 0 \leq u \leq u_a, \\ 0 & u_a \leq u \leq u_c, \end{cases} \quad (6)$$

где $\delta(u)$ – дельта-функция Дирака,

$$m = -\epsilon \left. \frac{\partial f}{\partial u} \right|_{u=u_c} + \Omega p = s + \Omega_0 p. \quad (7)$$

Здесь величина $s = -\epsilon \frac{\partial f}{\partial u}$ характеризует плотность потока связей в пространстве разностей потенциалов. Первое слагаемое в правой части (7) соответствует удельному числу связей, достигших в единицу времени критического значения u_c разности потенциалов. Второе слагаемое соответствует доле связей выводимых из процесса роста благодаря механизму активации. Оба слагаемых в сумме характеризуют скорость появления связей в точке $u = 0$.

При отсутствии последних двух членов это уравнение представляет собой уравнение Фоккера – Планка, которое описывает диффузионное расширение функции распределения в пространстве разностей потенциалов в процессе роста.

Справедливо условие сохранения числа связей в системе

$$p + q + m\tau = 1. \quad (8)$$

При этом функция распределения учитывает только активные и пассивные связи и, следовательно,

$$\int_0^{u_c} f(u)du = 1 - m\tau. \quad (9)$$

Введем новые переменные и обозначения:

$$T = u_a^2/\epsilon; \quad \tilde{t} = t/T; \quad x = u/u_a; \quad x_c = u_c/u_a; \quad (10)$$

тогда (4) переписывается в виде

$$\frac{\partial f}{\partial \tilde{t}} = \epsilon \frac{\partial^2 f}{\partial x^2} + \tilde{I} - \alpha f, \quad (11)$$

где

$$\alpha = \begin{cases} 0 & 0 \leq x < 1, \\ \Omega_0 T & 1 \leq x \leq x_c, \end{cases} \quad \tilde{I} = \begin{cases} mT\delta(0) & 0 \leq x < 1, \\ 0 & 1 \leq x \leq x_c. \end{cases} \quad (12)$$

Стационарное решение уравнения (11) имеет вид

$$f = \begin{cases} C_2 - C_1 x, & 0 \leq x < 1, \\ (C_2 - C_1) \exp \sqrt{\alpha}(1 - x), & 1 \leq x \leq x_c. \end{cases} \quad (13)$$

Используя непрерывность потока в точке $x = 1$ и условие нормировки (8), получаем

$$C_2 = C_1 \left(1 + \frac{1}{\sqrt{\alpha}} \right), \quad C_1 = \left(\frac{1 - \exp \sqrt{\alpha}(1 - x_c)}{\alpha} + \frac{1}{2} + \frac{1}{\sqrt{\alpha}} + \frac{\tau}{T} \right)^{-1}. \quad (14)$$

Компактные выражения для стационарных значений p и m имеют вид

$$p_0 = C_1 \left\{ 1 - \exp [\sqrt{\alpha}(1 - x_c)] \right\} / \alpha, \quad m_0 = C_1/T. \quad (15)$$

Характерный вид функции распределения связей по величине перепада потенциала на них представлен на рис. 6. Там же приведена функция распределения, полученная из численного эксперимента. Видно

хорошее соответствие среднеполевого подхода и численного эксперимента для $z \geq 4$.

Возвращаясь к дискретной модели, заметим, что величина коэффициента диффузии ϵ пропорциональна дисперсии $\tilde{\delta}$ нормально распределенных порций потенциала, добавляемых к изменяющемуся рельефу на каждом шаге модельного времени: $\epsilon = \tilde{\delta}^2/\tau$. Таким образом, $T = u_a^2/\tilde{\delta}\tau$. В счете мы выбирали время релаксации τ равным шагу модельного времени: $\tau = 1$.

Коэффициент Ω , характеризующий диссипацию, можно связать с координационным числом дискретной модели z . Действительно, интегрируя уравнение (4) по u в интервале от u_a до u_c , получаем

$$\frac{dp}{dt} = s(u_c) - s(u_a) - \Omega_0 p, \quad (16)$$

где величина s определена в соответствии с (7). Уравнение (16) описывает изменение числа активированных связей. Последний член в уравнении соответствует индуцированному выгоранию связей и равен произведению удельного числа выгорающих в единицу времени связей m на число их ближайших активированных соседей pz . Следовательно,

$$\Omega_0 = mz \quad (17)$$

или

$$z = \alpha/C_1. \quad (18)$$

С увеличением числа ближайших соседей число элементов «инфицированных» на одном шаге модельного времени увеличивается. При этом экспоненциальный член в (15) убывает, и, как видно из рис. 7, значение p_0 асимптотически стремится к $1/z$. Ситуация здесь полностью эквивалентна перколяционному распространению инфекции на решетках с большим числом измерений, где величина $1/z$ соответствует порогу протекания. В пространствах с размерностью два и три сильные флуктуации нарушают среднеполевую асимптотику. Тем не менее вид решения (13) сохраняется и для малых размерностей пространства. Другим важным параметром нашей задачи является отношение x_c . На рис. 7 показана зависимость критического числа $p_0(z)$ при разных значениях этого параметра.

5. Роль размеров системы

Если в начальный момент времени рельеф потенциала отсутствовал: $f(u, 0) = \delta(0)$, то до появления первой критической пары в системе распределение связей в пространстве разностей потенциалов будет

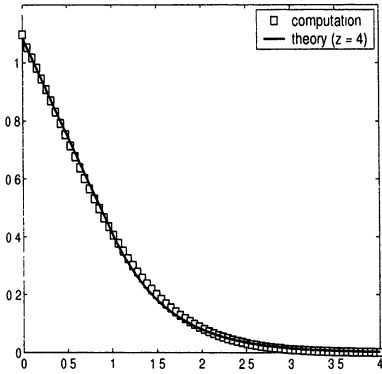


Рис. 6. Функция распределения связей по величине перепада потенциала для $x_c = 4$. Сплошная линия соответствует теоретической зависимости с $z = 4$.

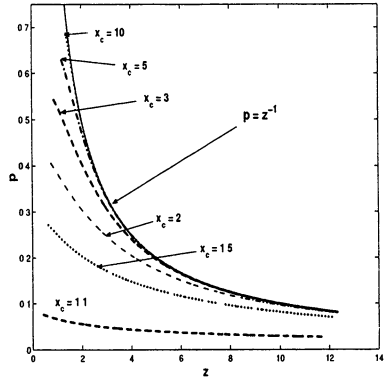


Рис. 7. Зависимость критического параметра p от величины z при различных соотношениях критического уровня и уровня активации.

гауссовским:

$$f(u, t) = \sqrt{\frac{2}{\pi \tilde{\delta} t / \tau}} \exp \left\{ -\frac{u^2}{2 \tilde{\delta} t / \tau} \right\}, \quad t < T_c. \quad (19)$$

Здесь характерный интервал времени T_c до появления первой критической пары зависит от размеров системы и может быть рассчитан из следующего соотношения:

$$\int_{u_c}^{\infty} f(u, T_c) du = 1/M, \quad (20)$$

где M – число связей в системе (для простой кубической решетки, имеющей N^3 узлов, $M = 3(N^3 + 2N - 1)$). Соотношение (20) подразумевает появление хотя бы одной критической связи в ограниченной системе. Заметим, что в термодинамическом пределе, когда мы имеем дело с безграничной системой, интервал времени T_c стремится к нулю: критические пары мгновенно появляются с началом процесса роста. Представим интеграл в левой части (20) через интеграл вероятности:

$$\int_{u_c}^{\infty} f(u, T_c) du = 1 - \int_0^{u_c} f(u, T_c) du = 1 - \text{erf}(y)/2, \quad (21)$$

где

$$y = \frac{u_c}{\sqrt{\tilde{\delta} T_c / \tau}}, \quad (22)$$

и используя известную асимптотику для интеграла вероятности при больших аргументах:

$$1 - \operatorname{erf}(y)/2 \approx \frac{\exp(-y)}{\sqrt{\pi y}} \left(1 - \frac{1}{2y^2}\right). \quad (23)$$

получаем выражение для оценки времени T_c :

$$\frac{1}{M} \approx \frac{\exp(-y)}{\sqrt{\pi y}} \left(1 - \frac{1}{2y^2}\right). \quad (24)$$

Для кубической решетки, имеющей миллион узлов, $M = 300597$. Выбирая параметр $u_c^2/\bar{\delta} = 2,5 \cdot 10^5$ и используя (22) и (24), мы находим $T_c \approx 100\tau$. Влияние размеров системы на характер динамики оказывается несущественным при условии, что к моменту первого пробоя удельное число активированных связей в системе не превышает порога протекания для задачи узлов на соответствующей решетке:

$$\int_{u_c}^{\infty} f(u, T_c) du = \frac{1}{2} \left\{ \operatorname{erf}(y) - \operatorname{erf}(yu_a/u_c) \right\} \lesssim p_c. \quad (25)$$

Из последнего выражения мы получаем минимально допустимую величину отношения u_a/u_c . Для выбранных выше значений параметров минимально допустимая величина отношения $u_a/u_c \simeq 0,25$. Если же условие (25) не выполнено, в системе появляется ряд новых динамических режимов: возникновение периодических и стохастических автоколебаний, распространение незагущающих фронтов горения при периодических граничных условиях и т. д. Упомянутые эффекты связаны исключительно с ограниченностью системы и, следовательно, отсутствуют в термодинамическом пределе.

6. Фрактальная динамика электрических разрядов в грозовом облаке

Грозное облако (ГО) может служить яркой иллюстрацией к рассмотренным выше эффектам фрактальной динамики [6, 8]. ГО безусловно является активной средой со свободной энергией в виде восходящих воздушных потоков, сильной турбулентности и фазовых превращений, сопровождаемых интенсивной зарядкой облачных частиц и генерацией электрического поля разных масштабов [9–12]. Радиополучение от микрозарядов, возникающих на предварительной стадии молниевых разрядов, свидетельствует о развитой мелкомасштабной структуре электрического поля с масштабами $l \sim 10 \div 10^2$ м [8, 12].

Существуют и физические механизмы, которые могут приводить к генерации такой ячеистой структуры электрического поля. Указанные ячейки служат основой для моделирования ГО трехмерной решеткой клеточных автоматов. Существование предварительной фазы грозового разряда в виде многочисленных микроразрядов, длящейся несколько десятых долей секунды без заметного изменения макроскопического электрического поля в облаке, указывает на то, что электрический пробой, который определяется критической разностью потенциалов u_c между соседними ячейками, возникает в первую очередь в мелкомасштабных структурах. Пробой в отдельно взятой ячейке может инициировать электрические разряды в соседних ячейках. К этому приводит ряд причин, главной из которых в условиях ГО служит, по-видимому, электрический пробой на убегающих электронах (ПУЭ) [13]. Суть ПУЭ состоит в том, что в процессе первого, по существу, холодного разряда с высоким порогом u_c рождаются энергичные электроны МэВ-ных энергий, которые могут запустить лавину ионизации в соседних ячейках при гораздо меньшей разности потенциалов $u_a \ll u_c$. Оценки показывают, что путь торможения МэВ-ных электронов в воздухе на высоте $h \sim 5$ км составляет $a \sim 50$ м. При $l \lesssim a$ эти электроны могут служить эффективным механизмом «размножения» микроразрядов, даже если в соседних ячейках разность потенциалов $u \sim u_a \ll u_c$. Реально отношение критического потенциала u_c к потенциалу активации u_a составляет $x_c = u_c/u_a \sim 5$. Время жизни возникающей цепочки разрядов составляет $\tau_c \sim 10^2$ мкс, и на это время она является, по существу, металлическим проводником электрического тока. В процессе разряда разность потенциалов на активных ячейках обращается в нуль, после чего мелкомасштабный электрический рельеф нарастает вновь за счет свободной энергии ГО. Этот рост, однако, на много порядков медленнее процесса «выгорания» активных ячеек. Согласно теоретическим оценкам [14–16], время развития самой быстрой из известных для ГО пучково-плазменной неустойчивости составляет $T \sim 1 \div 10$ с, т. е. в случае ГО $\epsilon = \tau_c/T \sim 10^{-4} \div 10^{-5} \ll 1$.

Таким образом, динамика микроразрядов в ГО в точности описывается рассмотренной в предыдущих разделах моделью, причем значения управляющих параметров $\epsilon \ll 1$ и $x_c \gtrsim 4$ лежат в области фрактальной динамики. Отсюда, кроме эффектов самоподобия и соответствующих универсальных спектров, следует принципиальный вывод о том, что ГО по своим электрическим свойствам находится вблизи порога перколяции, когда относительная доля активированных ячеек не превышает критической величины $p_c = 0,34$ для задачи

узлов на кубической решетке. В условиях динамического равновесия обновление активированных ячеек происходит за время T , т. е. число происходящих в облаке разрядов в единицу времени равно $F = \frac{p_c V}{T l^3}$, где V – объем активной части облака, а l – характерный масштаб ячейки. Для $V^{\frac{1}{3}} \sim 3$ км, $T \sim 3$ с и $l \sim 30$ м имеем $F \sim 10^5$ с $^{-1}$, что неплохо согласуется с экспериментом. Пребывание вблизи порога перколяции означает, что в облаке присутствуют проводящие кластеры с широким спектром размеров, достигающих диаметра облака. В такой ситуации облако критически реагирует даже на слабое крупномасштабное поле. Проводящий кластер поляризуется во внешнем поле, и на его концах возникают условия для самоподдержания электрического разряда и за пределами активной части облака [17]. Такой самоподдерживающийся разряд и формирует, по-видимому, лидерный канал молнии. Проведенное компьютерное моделирование фрактальной динамики ГО в рамках 3D-решетки клеточных автоматов в присутствии крупномасштабного поля продемонстрировало принципиальное изменение динамики системы: при включении поля происходило практически полное выгорание активных ячеек. Пример описанной стадии активности ГО в присутствии внешнего поля показан на рис. 8. Видно, что число разрядов в кластере, инициированном слабым крупномасштабным электрическим полем, более чем на порядок превосходит уровень фоновой активности (сравните с рис. 5). Характерная конфигурация возникающего проводящего кластера показана на рис. 9.

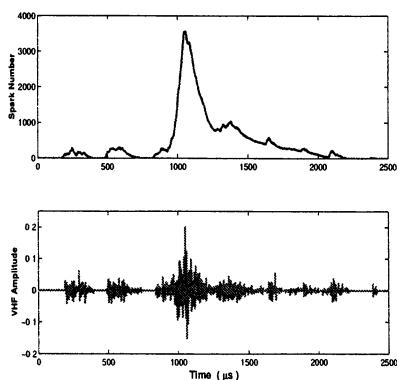


Рис. 8. Динамика трехмерной системы в присутствии внешнего поля.

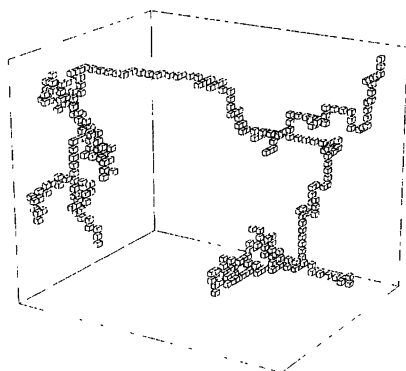


Рис. 9. Характерный вид проводящего кластера.

7. Заключение

Подводя итог проведенному выше анализу, можно утверждать, что фрактальная динамика присуща широкому классу нелинейных диссипативных сред. Переход к фрактальной динамике определяется значением управляющих параметров, главными из которых являются отношение скорости диссипации к скорости поступления свободной энергии в систему (параметр ϵ) и отношение характерного масштаба l элементарных возбуждений к линейному размеру системы L . Для природных систем, как правило, реализуется случай, когда $\epsilon \ll 1$ и $l/L \ll 1$, и имеет место фрактальная динамика. Принципиальную роль при этом играют эффекты динамической перколяции, когда число активных элементов среды балансирует вблизи порога перколяции. Сами элементы организуются в самоподобные кластеры разных размеров со степенным распределением по масштабам, причем длина максимального кластера сравнима с размерами системы. Эффективным методом описания вероятностных характеристик фрактальной динамики может служить среднеполевой подход, основанный на введении функции распределения по состояниям активности ячеек.

Фрактальная динамика среды меняется кардинально при появлении в системе крупномасштабного «силового» поля. В среде нервных волокон или в грозовом облаке такое поле возникает, например, под воздействием разности потенциалов, приложенной к границам системы. Численные эксперименты показывают, что возникает своего рода синхронизация¹ роста кластеров, и этот рост лавинно нарастает, пока не выгорят практически все активные ячейки. Этот процесс требует дальнейшего, более глубокого исследования, но уже сейчас можно высказать гипотезу, что природа изобрела удивительно надежный и экономный ($\epsilon \rightarrow 0$) канал передачи информации, основанный на фрактальной динамике активных сред.

Литература

1. *Stauffer D., Aharony A.* Introduction to Percolation Theory, Taylor & Francis. 1992. 181 p.
2. *Bunde A., Havlin S.* Percolation, in fractals and disordered systems / Ed. A. Bunde, S. Havlin. Springer, 1966. P. 59–177.
3. *Bak P.* How nature works. Oxford Univ. Press, 1997. 212 p.
4. *Лоскутов А. Ю., Михайлов А. С.* Введение в синергетику. М.: Наука, 1990. 272 с.

¹Этот термин применительно к росту кластеров предложен Л. В. Келдышем

5. Трубецков Д. И. След вдохновений и трудов упорных. Саратов: изд-во гос. УНЦ «Колледж», 2001. С. 13.
6. Иудин Д. И., Трахтенгерц В. Ю. Фрактальная структура нелинейной динамики электрического заряда в грозовом облаке // Изв. вузов. Радиофизика. 2001. № 5—6. С. 419—437.
7. *Vespignani A., Zapperi S.* How self-organized criticality works: A unified mean-field picture // *Phys. Rev. E.* 1998. V. 57, № 6. P. 6345—6362.
8. Иудин Д. И., Трахтенгерц В. Ю. Фрактальная динамика электрического заряда в грозовом облаке // Физика атм. и океана. 2000. № 5. С. 650—662.
9. *MacGorwan D. R., Rust W. D.* The electrical nature of storms. Oxford Un. Press, 1998.
10. *Uman M. A.* The Lightning Discharge // *Internat. Geophys. Series.* 1987. 39. 377.
11. *Rakov V. A., Uman M. A.* K and M changes in close lightning ground flashes in Florida // *J. Geophys. Res.* 1990. V. D 11. P. 18,631—18,647.
12. *Proctor D. E., Uytendogaardt R., Meredith B. M.* VHF Radio Pictures of Lightning Flashes to Ground // *J. Geophys. Res.* 1988. V. 93, D10. P. 12,683—12,727.
13. Гуревич А. В., Зыбин К. П. Пробой на убегающих электронах электрические разряды во время грозы // УФН. 2000. Т. 173, № 11. С. 1177—1199.
14. Трахтенгерц В. Ю. О природе электрических ячеек в грозовом облаке // ДАН СССР. 1989. Т. 308. С. 584.
15. *Trakhtengerts V. Y.* Electrical impedance of a thunderstorm cloud // *Proc. of ICAE-92. St.-Petersburg. Russia,* 1992. V. 2. P. 416—417.
16. Мареев Е. А., Сорокин А. Е., Трахтенгерц В. Ю. Эффекты коллективной зарядки облака в многотоковой аэрозольной плазме // Физика плазмы. 1999. № 3. С. 289—300.
17. Базелян Э. М., Райзер Ю. П. Механизм притяжения молнии и проблема лазерного управления молнией // УФН. 2000. Т. 170, № 7. С. 753—769.

НЕЛИНЕЙНАЯ СТАДИЯ РАЗВИТИЯ ВОЗМУЩЕНИЙ ПРИ УСКОРЕНИИ ОБОЛОЧЕК

Л. С. Мхитарьян

Подавляющее большинство встречающихся течений плазмы неустойчиво. Имеем ли мы дело с плазменным шнуром в токамаках, или с лайнером, сжимающимся текущим по нему током, или со сферической мишенью, облучаемой лазерным светом и под действием возникающего давления сходящейся к центру – во всех этих случаях движение плазмы неустойчиво.

Почти всегда есть возможность проанализировать развитие неустойчивостей лишь в линейном приближении. Расскажем здесь об одном из немногих случаев, когда эволюция возмущений может быть прослежена и в нелинейной стадии – это известная задача Э. Отта [1] о разгоне бесконечно тонкой пластины постоянным давлением. Расскажем также о приложении решения этой задачи к анализу влияния возмущений на воспламенение термоядерного топлива в мишенях для лазерного синтеза с инерционным удержанием.

1. Задача Отта

1.1. Постановка задачи и общее решение

Пусть бесконечно тонкая оболочка без прочности (высокие скорости или жидкие и газообразные (плазма) среды) разгоняется постоянным во времени и пространстве давлением.

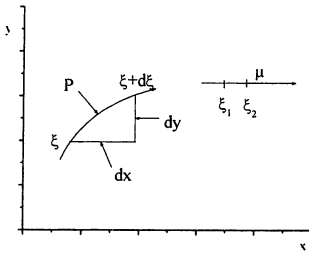


Рис. 1. Лагранжевы координаты на пластине

На идеально плоской и однородной пластинке с поверхностной массой μ введём лагранжевую координату ξ (рис. 1). Если в начальный момент форма и скорость пластинки заданы, то как они будут меняться под действием приложенного давления? Другими словами, каковы функции $x(\xi, t)$ и $y(\xi, t)$? Уравнения движения гласят ($g = P / \mu$):

$$\frac{\partial^2 x}{\partial t^2} = g \frac{\partial y}{\partial \xi}, \quad \frac{\partial^2 y}{\partial t^2} = -g \frac{\partial x}{\partial \xi}. \quad (1)$$

Вводя теперь комплексную функцию $z = x + iy$, будем иметь

$$\frac{\partial^2 z}{\partial t^2} = -ig \frac{\partial z}{\partial \xi} \quad (2)$$

У этого линейного уравнения помимо очевидного решения, описывающего полёт невозмущённой пластины, есть также решение в виде суммы бегущих волн. Общее решение (2) имеет вид

$$z = \xi - i \frac{gt^2}{2} + \sum_{k=-\infty}^{k=\infty} z_k \exp(i\omega t - ik\xi). \quad (3)$$

Подставляя это в (2) получим дисперсионное уравнение

$$\omega^2 = gk. \quad (4)$$

Видно, что имеется два существенно различных типа решений. В одном $\omega = \pm\sqrt{gk}$ (для $k > 0$), и амплитуды различных гармоник колеблются со временем. В другом $\omega = \pm i\sqrt{gk}$ (для $k < 0$), и амплитуды экспоненциально нарастают.

Обозначив $\sqrt{gk} = \gamma$, перепишем (3) в виде

$$z = \xi - i \frac{gt^2}{2} + \sum_{k>0} \left[z_k^{(1)} \sin \gamma t + z_k^{(2)} \cos \gamma t \right] e^{-ik\xi} + \left[z_k^{(3)} \operatorname{sh} \gamma t + z_k^{(4)} \operatorname{ch} \gamma t \right] e^{ik\xi}.$$

Восемь констант, сюда входящих, определяются начальными условиями: начальной формой пластины $x(\xi, 0)$, $y(\xi, 0)$ и начальной её скоростью $\dot{x}(\xi, 0)$, $\dot{y}(\xi, 0)$. При разложении в ряд Фурье этих функций каждое условие определит по две константы. Не будем заниматься общим исследованием решения, а рассмотрим ряд частных примеров.

1.2. Колеблющиеся решения

Выбирая подходящим образом начало отсчёта координат и времени, запишем колеблющуюся моду в виде

$$x = \xi + a \cdot \sin(\gamma t) \cdot \cos(k\xi), \quad y = -\frac{gt^2}{2} - a \cdot \sin(\gamma t) \cdot \sin(k\xi).$$

Мы выбрали случай, когда в начальный момент пластинка плоская, но ей придана начальная скорость. Видно, что в ходе колебаний вещество перетекает вдоль пластинки, так что её поверхностная масса всё время меняется:

$$\mu \frac{d\xi}{dl} = \frac{\mu}{\sqrt{\left(\frac{\partial x}{\partial \xi}\right)^2 + \left(\frac{\partial y}{\partial \xi}\right)^2}} = \frac{\mu}{\sqrt{1 - 2ak \sin(k\xi) \sin(\gamma t) + [ak \sin(\gamma t)]^2}}.$$

Эти изменения происходят так, что отставшие участки делаются легче, а опередившие – тяжелее. Ясно, что малейшее нарушение распределения массы вдоль пластинки разрушит режим колебаний.

1.3. Нарастающие моды

Пусть пластинка разгоняется из состояния покоя, имея вначале не плоскую форму

$$x = \xi + a \cdot \text{ch}(\gamma t) \cdot \cos(k\xi), \quad y = -\frac{gt^2}{2} + a \cdot \text{ch}(\gamma t) \cdot \sin(k\xi). \quad (5)$$

Амплитуда возмущения $A(t) = a \cdot \text{ch}(\gamma t)$ экспоненциально нарастает со временем (при $\gamma t \gg 1$). Области струй (т. е. есть выпуклостей в ту сторону, откуда действует давление) расположены при $2\pi n \leq k\xi \leq \pi(2n+1)$, $n = 0, \pm 1, \pm 2, \dots$. Ширина каждой области (см. (5)) есть

$$\Delta x = \frac{\pi(2n+1) - 2\pi n}{k} + A \cdot [\cos \pi(2n+1) - \cos 2\pi n] = \frac{\pi}{k} - 2A.$$

С увеличением амплитуды ширина области струй уменьшается, а ширина области пузырей увеличивается. Масса на единицу поверхности с течением времени меняется:

$$\frac{\mu d\xi}{\sqrt{dx^2 + dy^2}} = \frac{\mu}{\sqrt{1 - 2Ak \sin(k\xi) + A^2 k^2}}. \quad (6)$$

В области пузырей ($\sin(k\xi) < 0$) она уменьшается, а в области струй ($\sin(k\xi) > 0$) – растёт. Поэтому при увеличении амплитуды радиус кривизны отставших участков пластины делается всё меньше и меньше, пока, наконец, на пластине не возникают острые углы. В самом деле

$$\frac{dy}{dx} = \frac{\partial y}{\partial \xi} \cdot \frac{\partial \xi}{\partial x} = \frac{A(t) \cdot k \cdot \cos(k\xi)}{1 - A(t) \cdot k \cdot \sin(k\xi)}.$$

И как только будет $A(t) \cdot k = 1$, пластина в точках $k\xi = \frac{\pi}{2} + 2\pi n$ (се-

редина струй) примет вертикальное положение. Поверхностная масса (6) в этих точках обратится в бесконечность. В середине пузыря ($\sin(k\xi) = -1$) поверхностная масса к этому времени уменьшится вдвое.

При дальнейшем увеличении амплитуды решение Отта теряет силу: оно перестаёт быть однозначным. На пластине возникают самопересечения, показанные на рис. 2.

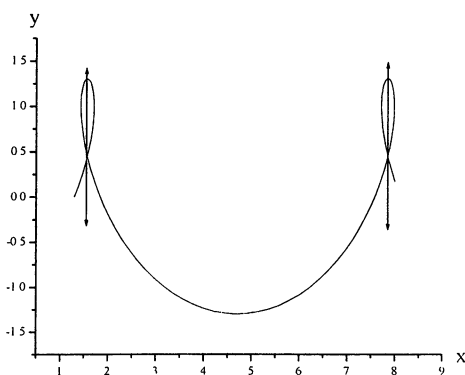


Рис. 2. Образование струй при $Ak > 1$

В том месте, где в решении Отта возникают петли, на самом деле образуются кумулятивные струи. При этом пест движется в ту сторону, откуда действует давление (и несёт основную массу), а быстрая струя движется в противоположном направлении.

После образования кумулятивной струи давление в пластинке в области струй существенно меняется, и дальнейшую эволюцию формы пластины решение Отта, конечно, не в состоянии передать. Однако, и это важно в практических приложениях, возмущения давления из области струи не передаются на расстояния, превышающие толщину пластины. Поэтому, если интересоваться эволюцией массы в пузыре, уравнение (6) остаётся справедливым и тогда, когда Ak уже заметно превосходит единицу.

1.4. Переход к нелинейности. Высшие гармоники

Характерный признак нелинейной стадии – появление гармоник, отсутствовавших в начальном возмущении. Из рис. 2 видно, что к моменту образования струй возбуждены все гармоники вплоть до самых высоких. Как это происходило?

Чтобы проследить за возникновением высших гармоник, достаточно разложить $y(x)$ из (5) в ряд Фурье:

$$y(x) = \sum_{n=1}^{\infty} y_n \sin(nkx), \quad y_n = \frac{k}{\pi} \int_{-\pi/k}^{\pi/k} y(x) \cdot \sin(nkx) \cdot dx.$$

Используя (5), перейдём здесь от интегрирования по x к интегрированию по ξ ($dx \rightarrow \frac{\partial x}{\partial \xi} d\xi$). Тогда

$$y_n = \frac{kA}{\pi} \int_{-\pi/k}^{\pi/k} \sin(k\xi) \cdot \sin[nk(\xi + A \cdot \cos(k\xi))] \cdot (1 - Ak \cdot \sin(k\xi)) \cdot d\xi,$$

$$A = a \cdot \text{ch}(\gamma t).$$

Преобразовав синус суммы двух углов и раскрыв скобки, мы получим известные из теории цилиндрических функций интегралы. Так что результат может быть выражен аналитически через функции Бесселя:

$$y_n = \frac{A}{n} \cdot [J_{n-1}(nAk) - J_{n+1}(nAk)]. \quad (7)$$

Эволюция первых пяти гармоник и приведена на рис. 3, где изображено отношение амплитуды гармоники с номером n к амплитуде первой гармоники в зависимости от Ak .

Используя асимптотическое разложение функций Бесселя при больших значениях индекса, легко убедиться, что

$$\begin{aligned} \text{а) для } Ak \ll 1 \text{ и } n \gg 1 & \quad y_n / y_1 \approx \frac{1}{n} \left(\frac{e}{2} \cdot Ak \right)^{n-1}, \\ \text{б) для } Ak = 1 \text{ и } n \gg 1 & \quad y_n / y_1 \approx (1,54/n) \cdot \sqrt{2/\pi} \cdot (2n)^{-1/4}. \end{aligned}$$

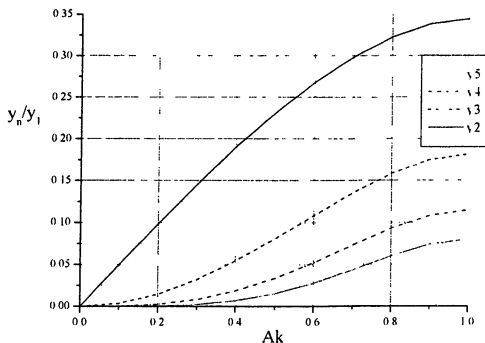


Рис. 3. Эволюция гармоник при ускорении пластины

Видно, что при малых амплитудах возмущения высшие гармоники очень медленно развиваются. В момент образования струи амплитуда гармоник слабо убывает с возрастанием их номера. Уже при $Ak \sim 0,3$ вклад второй гармоники становится заметным. Можно считать что нелинейная стадия (при всей условности этого определения), начинается именно с этого момента.

1.5. Поверхности с евклидовой геометрией. Цилиндр

Уравнения Отта применимы к любой поверхности, геометрия на которой евклидова. Ограничимся рассмотрением простейшего случая в качестве примера.

Пусть мы имеем цилиндр с массой M на единицу длины, который, имея начальную скорость u_0 и начальный радиус R_0 , тормозится изнутри постоянным давлением. Введя в качестве лагранжевой переменной угол φ , будем иметь вместо (1)

$$\begin{aligned} \frac{\partial^2 x}{\partial t^2} &= \gamma^2 \frac{\partial y}{\partial \varphi}, \quad 0 \leq \varphi \leq 2\pi, \\ \frac{\partial^2 y}{\partial t^2} &= -\gamma^2 \frac{\partial x}{\partial \varphi}, \quad \mu d\varphi = \frac{M}{2\pi} d\varphi, \quad \gamma^2 = \frac{2\pi P}{M}. \end{aligned} \quad (8)$$

Представим решение этой системы в виде суммы цилиндрически симметричного движения и возмущения

$$\begin{aligned} x(\varphi, t) &= R(t) \cdot \cos(\varphi) + \tilde{x}(\varphi, t), \\ y(\varphi, t) &= R(t) \cdot \sin(\varphi) + \tilde{y}(\varphi, t). \end{aligned} \quad (9)$$

Невозмущённое решение удовлетворяет уравнению $\dot{R} = \gamma^2 R$, решение которого есть $R(t) = R_0 \text{ch}(\gamma t) - (u_0 / \gamma) \text{sh}(\gamma t)$. При $u_0 < R_0 \gamma$ цилиндр не достигает оси, останавливаясь на радиусе $R_{\min} = R_0 \left(1 - \frac{u_0^2}{\gamma^2 R_0^2} \right)^{1/2}$.

Пусть теперь цилиндр имеет вначале только возмущения формы:

$$R(\varphi, 0) = R_0 \cdot (1 + a \cdot \cos(m\varphi)), \quad a \ll 1.$$

Это означает, что

$$x(\varphi, 0) = R_0 \cos \varphi + \frac{R_0 a}{2} [\cos(m+1)\varphi + \cos(m-1)\varphi],$$

$$y(\varphi, 0) = R_0 \sin \varphi + \frac{R_0 a}{2} [\sin(m+1)\varphi - \sin(m-1)\varphi].$$

Поэтому интересующая нас мода будет

$$\hat{x}(\varphi, t) = \frac{R_0 a}{2} [\cos(m+1)\varphi \cdot \operatorname{ch} \sqrt{m+1} \gamma t + \cos(m-1)\varphi \cdot \cos \sqrt{m-1} \gamma t],$$

$$\hat{y}(\varphi, t) = \frac{R_0 a}{2} [\sin(m+1)\varphi \cdot \operatorname{ch} \sqrt{m+1} \gamma t - \sin(m-1)\varphi \cdot \cos \sqrt{m-1} \gamma t].$$

Почему для пластины инкремент был \sqrt{gk} , а здесь $\gamma \sqrt{m \pm 1}$? Уско-

рение цилиндра есть $\frac{2\pi R(t)P}{M}$. А волновой вектор моды с номером m есть $m/R(t)$ (в самом деле, $k = 2\pi/\lambda$, а длина волны есть, конечно, $\lambda = 2\pi R(t)/m$). Поэтому и для цилиндра инкремент равен \sqrt{gk} !

Растущая мода быстро обгоняет колеблющуюся, так что при $\operatorname{ch} \sqrt{m} \gamma t \gg 1$ форма цилиндра определяется уравнениями

$$\tilde{x} = \cos \varphi + A \cos(m+1)\varphi, \quad \tilde{y} = \sin \varphi + A \sin(m+1)\varphi, \quad (10)$$

где $\tilde{x} = \frac{x}{R(t)}$, $\tilde{y} = \frac{y}{R(t)}$, $A(t) = \frac{R_0 a}{2R(t)} \operatorname{ch} \sqrt{m+1} \gamma t$.

И здесь при $A \cdot (m+1) > 1$ возникают самопересечения. В самом деле, самопересечение означает, что кривая дважды (по крайней мере) проходит через одну и ту же точку. Когда кривая задана параметрически, это равносильно тому, что существует два значения параметра, при которых

$$x(\varphi_1) = x(\varphi_2), \quad y(\varphi_1) = y(\varphi_2). \quad (11)$$

Но всегда можно положить $\varphi_1 = \varphi_0 + \varepsilon$, $\varphi_2 = \varphi_0 - \varepsilon$. Самопересечение возникает впервые, когда $\varepsilon \rightarrow 0$. Раскрывая тригонометрические функции суммы двух углов, запишем (11) для кривой (10) в виде

$$\sin \varphi_0 \cdot \sin \varepsilon + A \sin(m+1)\varphi_0 \cdot \sin(m+1)\varepsilon = 0,$$

$$\cos \varphi_0 \cdot \sin \varepsilon + A \cos(m+1)\varphi_0 \cdot \sin(m+1)\varepsilon = 0.$$

При $\varepsilon \rightarrow 0$ отсюда получим уравнение для φ_0 , решение которого есть

$$-\cos m\varphi_0 = \frac{1 + A^2(m+1)^2}{2A(m+1)}.$$

Функция, стоящая справа, имеет минимум, равный единице, при $A \cdot (m+1) = 1$. И, значит, $\varphi_0 = \frac{\pi (2k+1)}{m}$, $k = 0, 1, 2 \dots$ Струи возникают там, где начальный радиус оболочки был минимальным (в области провалов).

Для примера на рис. 4 приведена форма цилиндрической оболочки (10), возмущённой третьей гармоникой, на момент образования струй ($A = 0,25$, $m = 3$).

Видно, что в дальнейшем область внутри оболочки будет разделена струями на три подобласти. Если струи пересекутся в центре, возникнет многосвязная конфигурация с развитой поверхностью.

Если рассматривать возмущения в виде наложения нескольких гармоник, мы наблюдали бы весьма прихотливое их взаимодействие, зависящее от соотношения фаз и амплитуд этих гармоник. Весьма показательным случаем, когда длины волн двух гармоник сильно отличаются. На рис. 5 показан результат взаимодействия гармоник с $m = 3$ и $m = 29$ ($A_3 = 1/4$, $A_{29} = 1/15$).

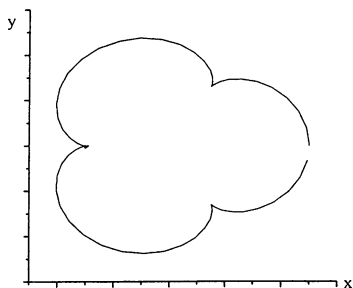


Рис. 4. Форма цилиндра на момент начала образования струй

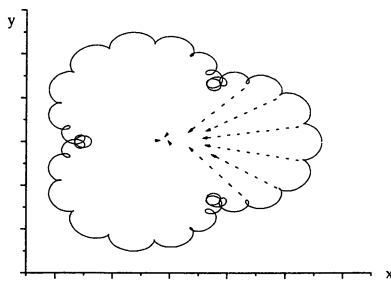


Рис. 5. Взаимодействие гармоник с $m = 3$ и $m = 29$ при сжатии цилиндра

Видно, что струи, за которые ответственна третья гармоника, возникли раньше, чем стало $A \cdot (m+1) = 1$ (три области с самопересечениями). А на фоне крупномасштабной гармоники появилось множество струй от мелкомасштабной. При дальнейшем развитии возмущений каждая из трёх областей, которые мы видели на рис. 4, будет разбита на множество подобластей струями от мелкомасштабной гармоники. На рис. 5 это изображено пунктирными стрелками. И еще одно замечание.

Мы рассматривали случаи, когда ускоряющее (или тормозящее) давление не зависело от времени. На практике это обычно не так. Но если время, за которое произведение gk заметно меняется, мало в сравнении

со временем нарастания возмущения $\frac{1}{gk} \frac{\partial gk}{\partial t} \ll \sqrt{gk}$, можно утверждать,

что возмущение будет успевать подстраиваться под текущее значение инкремента. Применяя метод, аналогичный ВКБ-приближению, легко убедиться, что возмущение будет расти в этом случае по закону

$$A(t) = A_0 \exp\left(\int \sqrt{gk} dt\right).$$

2. Приложение решения Отта

2.1. Мишень для зажигания на американской лазерной установке

Во ВНИИЭФ мы с коллегами исследовали влияние роста возмущений на зажигание термоядерного топлива в мишени, которая планируется для исследований на строящейся американской лазерной установке NIF (рис. 6).

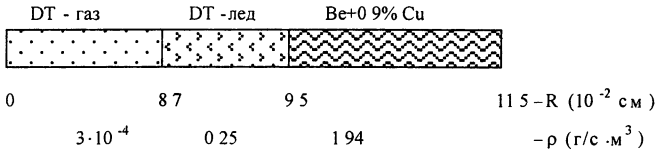


Рис. 6. Геометрия мишени для установки NIF.

Предполагается, что лазерная энергия на этой установке будет конвертироваться в рентгеновское излучение со спектром, близким к планковскому, с температурой, зависящей от времени так, как показано на рис. 7.

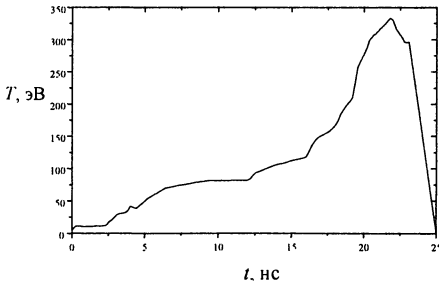


Рис. 7. Временной ход температуры рентгеновского излучения, падающего на мишень

Бериллиевая оболочка мишени под действием рентгеновского излучения испаряется, и возникшая малоплотная плазма разгоняет неиспаренную (плотную) часть оболочки и DT-топливо к центру. Ситуация на фронте испарения (абляции) с очевидностью неустойчива.

Ударная волна, возникшая на фронте абляции, достигнув внутренней поверхности DT-льда, генерирует волну разгрузки, в которой из-за возмущений появляются тангенциальные (вдоль оболочки) компоненты скорости. Они переносятся волной разгрузки на фронт абляции и там нарастают, вызывая перетекание массы внутри оболочки из DT-льда. Если неоднородности оболочки, вызванные этим, становятся большими, DT-топливо не воспламеняется.

2.2. Результаты двумерных расчетов

Для анализа проблемы мы использовали программу МИМОЗА-НД [2], способную рассчитывать двумерные газодинамические течения вместе с радиационным переносом энергии. В начальный момент времени на внутренней поверхности ДТ-льда задавалось возмущение формы в виде сферической гармоники:

$$R(\theta) = R_0 + A_n P_n(\cos \theta).$$

Для $n = 4, 10, 30$ проводилась серия расчётов, в которых начальная амплитуда гармоник последовательно увеличивалась до тех пор, пока не наступал отказ в зажигании мишени. Для иллюстрации приведем форму ДТ-льда в одном из расчётов на момент времени, близкий к зажиганию топлива (рис. 8).

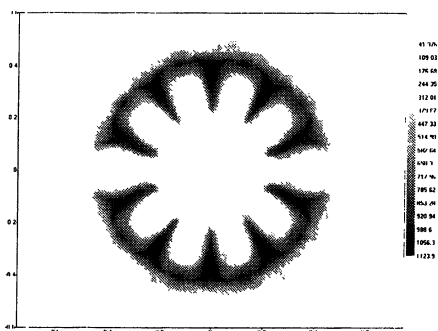


Рис. 8. Форма ДТ-льда, полученная в расчете по программе МИМОЗА-НД на момент времени $t = 22,8$ нс для возмущения с $n = 10$

Видно, как из начального возмущения сформировались пузыри и струи – совсем так же, как мы уже видели, исследуя решение Отта для цилиндра. Масса ДТ-льда перетекает из области пузырей в область струй. Оказывается (рис. 9), что для любой гармоники отказ в зажигании мишени наступает, когда масса в области пузырей снизится примерно на 30%.

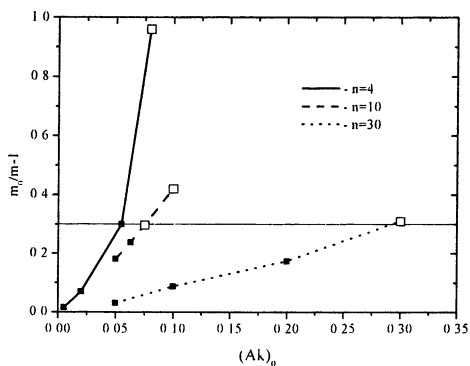


Рис. 9. Потеря массы в «пузыре» в зависимости от начальной амплитуды возмущения для различных гармоник. Расчеты, в которых мишень загорелась (■) и не загорелась (□)

Мы использовали решение Отта для построения феноменологической модели явления. Но чтобы применить его к обсуждаемой проблеме, потребовалась небольшая модернизация:

$$\frac{1}{m_b} \frac{dm_b}{dt} = \frac{D}{\Delta_0} (Ak)_0 - \frac{\eta}{k_0 \Delta_0} \cdot \frac{(ak) \frac{df_-}{dt}}{(1 + (ak) f_-(t))}, \quad (12)$$

$$f_-(t) = \sinh \left(\int_{t_0}^t \gamma dt \right) - \sin \left(\int_{t_0}^t \gamma dt \right),$$

$$a = \alpha \frac{(Ak)_0 R_0}{n} \exp \left(-\frac{n}{R_0} \Delta_0 \right) = \alpha A_0 \exp(-k_0 \Delta_0), \quad (13)$$

$$\gamma = \sqrt{g_{ab}(t)k(t) - \beta V_{ab}(t)k(t)}, \quad k(t) = \frac{n}{R_{ab}(t)}. \quad (14)$$

Здесь m_b – масса в середине пузыря, D – скорость ударной волны, Δ_0 – толщина оболочки из ДТ-льда, t_0 – момент выхода ударной волны на внутреннюю поверхность ДТ-льда, k – волновой вектор возмущения в этот момент, R_{ab} – радиус фронта абляции (R_0 – его значение в момент t_0), g_{ab} – ускорение фронта абляции, V_{ab} – скорость абляции: скорость испарения массы с единицы поверхности, делённая на среднюю плотность ДТ-льда, α и β – параметры модели.

Первый член в уравнении (12) описывает перетекание массы в пузыре из-за тангенциальной скорости, вызванной ударной волной при её выходе на внутреннюю поверхность ДТ-льда. Второй член описывает перетекание массы из-за развития возмущений скорости, перенесённых на фронт абляции волной разгрузки. Он отличается от (6) тем, что в (12) произведено усреднение тангенциальной скорости по толщине оболочки из ДТ-льда. Так мы учитывали влияние конечности толщины оболочки в изучаемой мишени. Выражение для амплитуды (13) передаёт затухание тангенциальных скоростей при движении волны разгрузки от внутренней поверхности ДТ-льда к фронту абляции. И, наконец, второй член в выражении (14) для инкремента учитывает неравномерный унос массы (а значит, уменьшение амплитуды возмущений) с фронта абляции из-за теплового потока.

Сравнение эволюции массы в середине пузыря, предсказываемое моделью, с результатами двумерных расчётов приведено на рис. 10; видно вполне удовлетворительное согласие.

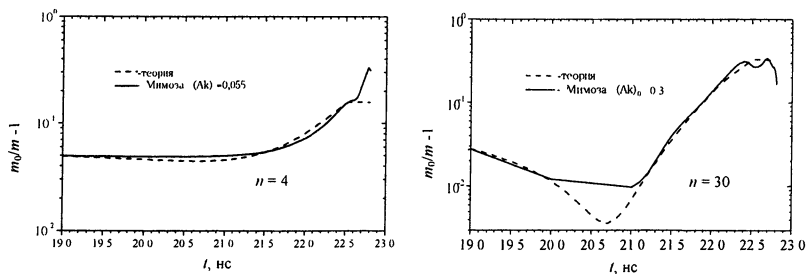


Рис. 10. Сравнение эволюции массы в пузырье, предсказываемой моделью, с полученной в расчетах по программе МИМОЗА-НД для различных номеров гармоник

Зная эволюцию массы в пузырье, легко найти значения начальных амплитуд возмущений на внутренней поверхности ДТ-льда, при которых $m_0/m-1$ достигает заданного значения для любой гармоники. Плоскость «начальная амплитуда возмущения – номер гармоники» окажется, таким образом, разделённой на две области, в одной из которых мишень вспыхивает, а в другой – нет. Результат такого построения приведён на рис. 11.

Видно, что положение кривой, ниже которой мишень горит, а выше – отказывает, модель передаёт достаточно аккуратно.

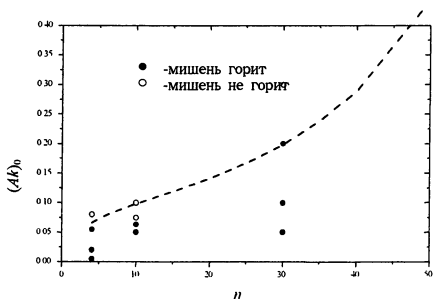


Рис. 11. Предсказываемая моделью зависимость начальной амплитуды возмущения, при которой масса в пузырье снижается на 30%, от номера гармоники (пунктирная кривая). Кружками обозначены результаты расчетов по программе МИМОЗА-НД

Итак, решение Отта может служить надёжным путеводителем при анализе нелинейной стадии развития возмущений даже в столь сложной проблеме, как мишени для лазерного управляемого термоядерного синтеза.

Литература

- 1 Ott Edward Nonlinear evolution of the Rayleigh – Taylor instability of a thin layer // Phys Rev Lett 1972 V 29, № 21 P 1429–1432
- 2 Софронов И Д, Бельков С А, Винокуров О А, Мхитарьян Л С, Рябикина Н А Методика расчета спектрального переноса излучения в двумерном комплексе МИМОЗА-НД // Вопросы атомной науки и техники Сер Математическое моделирование физических процессов, 2000 Вып 1 С 8–15

НЕКОТОРЫЕ НЕЛИНЕЙНЫЕ ЗАДАЧИ ЛАЗЕРНОГО ТЕРМОЯДЕРНОГО СИНТЕЗА

Г. Г. Кочемасов

Введение

Лазерный термоядерный синтез (ЛТС) – это направление исследований, ставящее своей целью использование лазеров для зажигания и поддержания горения малых масс термоядерного топлива. Первые предложения по ЛТС были даны в начале 60-х годов, сразу после изобретения первого лазера (А. Д. Сахаров, неопубликованное сообщение [1], Н. Г. Басов, О. Н. Крохин [2]). Основной задачей современного этапа исследований является осуществление зажигания термоядерного топлива (смеси дейтерия D и трития T), т. е. получение энергии, выделившейся в ходе термоядерных реакций, существенно превышающей энергию, вложенную в DT-топливо. В последнее десятилетие весьма актуальным стало также исследование более широкого круга явлений, происходящих в горячей и плотной плазме и в определенной степени подобных тем, что происходят в астрофизических масштабах. Более отдаленной является проблема создания энергетически выгодной термоядерной энергетики.

ЛТС – комплексная научно-техническая проблема, включающая разработку мощных (≥ 500 ТВт) и энергетичных (≥ 2 МДж) лазеров, конструирование и изготовление мишеней с размерами $\leq 1\div 10$ мкм, проведение экспериментов по сжатию и нагреву DT-топлива, диагностирование нестационарной плазмы с высоким временным (≤ 10 пс) и пространственным (≤ 10 мкм) разрешением, теоретическую и расчетную разработку постановок экспериментов и анализ экспериментальных данных.

В 70-х годах основным подходом к реализации ЛТС был подход, предложенный А. Д. Сахаровым: интенсивное лазерное излучение облучает со всех сторон сферическую оболочку, наполненную DT-топливом, вызывая испарение (абляцию) внешнего слоя оболочки. Реактивный эффект разлетающейся плазмы вызывает ускоренное движение оставшейся части оболочки к центру, что в свою очередь приводит к сжатию и нагреву DT-топлива до температур (~ 10 кэВ $\approx 10^8$ К), достаточных для инициирования термоядерных реакций. В последние 15 лет наибольшие успехи были достигнуты при использовании метода непрямого облучения, когда лазерное излучение предварительно преобразовывается в рентгеновское, которое и облучает оболочку с DT-топливом.

Теоретическое и расчетное описание работы лазерных микромишеней базируется в основном на электродинамике, физике плазмы, газовой динамике, атомной и ядерной физике. Большинство из протекающих процессов и весь процесс в целом являются нелинейными. Достаточно упомянуть не-

линейное взаимодействие лазерного света (интенсивность $\geq 10^{14}$ Вт/см²) с разлетающейся плазмой, нелинейный процесс переноса тепла из горячей плазмы в холодные слои оболочки путем электронной и лучистой теплопроводности, формирование и распространение ударной волны, нелинейное изэнтропическое сжатие и нагрев ДТ-топлива и пр. Сколько-нибудь подробное описание хотя бы одного из упомянутых процессов потребовало бы отдельной работы. Поэтому ниже приведем вводный курс физики ЛТС и краткий обзор исследований по физике горячей и плотной плазмы, выполненных на лазерной установке Российского федерального ядерного центра – ВНИИЭФ «Искра-5». Расскажем также о планах создания установки следующего поколения «Искра-6» и строительстве модуля этой установки «Луч»¹.

1. Основы физики ЛТС

1.1. Зажигание ДТ-топлива. Для осуществления реакций синтеза реагирующие ядра должны быть сближены на расстояние порядка 10^{-13} см, после чего процесс их слияния происходит с заметной вероятностью за счет туннельного эффекта. Из всех возможных реакций синтеза предпочтительной оказывается реакция ядер изотопов водорода-дейтерия и трития, происходящая с выделением энергии 17,6 МэВ. При этом около 14 МэВ уносит нейтрон, а остальная энергия приходится на ядро He⁴ (α -частицу). Для преодоления потенциального барьера ядер, в состав которых входят протоны, должна быть сообщена энергия ~ 10 кэВ, что соответствует температуре $\sim 10^8$ К. Константа скорости данной реакции составляет около 10^{-17} см³·с⁻¹ при $T \cong 10$ кэВ [5].

При температурах $T \cong 10$ кэВ плазма полностью ионизована и состоит из ядер дейтерия D и трития T и электронов. Поэтому для нагрева ДТ-плазмы до температуры $T \cong 10$ кэВ массой $m = \frac{4}{3}\pi R^3 \rho = \rho V$ требуется энергия

$$E_{\text{ДТ}} = \frac{3}{2}T(n_e + n_D + n_T)V = \frac{3}{A}N_A mT = 115 mT \text{ [МДж]}, \quad (1)$$

где n_e, n_D, n_T – концентрации электронов, дейтронов и тритонов; A – средний атомный номер ДТ-смеси ($A = 2,5$); N_A – число Авогадро; масса измеряется в граммах, температура – в килоэлектронвольтах.

Какую энергию $E_{\text{ТЯ}}$ можно при этом получить из топлива? Она дается выражением

¹ Обзор исследований по физике горячей и плотной плазмы основывается на докладе автора на IV Харитоновских научных чтениях [3] (см также [4] и цитированную там литературу)

$$E_{\text{ТЯ}} = \varepsilon(\overline{\sigma V}) n_{\text{D}} n_{\text{T}} V t = \varepsilon(\overline{\sigma V}) t \frac{m \rho N_A^2}{2A^2}, \quad (2)$$

где ε – энергия, выделяющаяся в одном акте реакции; $(\overline{\sigma V})$ – константа скорости реакции. Если t достаточно велико, то $E_{\text{ТЯ}} > E_{\text{ДТ}}$. В инерционном ТС, в том числе и ЛТС, время t ограничивается разгрузкой плазмы, характерное время которой

$$\tau \cong \frac{R}{3C_S}, \quad (3)$$

где $C_S = \left(\frac{2}{3} C_V T\right)^{1/2}$ – изотермическая скорость звука, C_V – теплоемкость плазмы – см. (1). Условие $E_{\text{ТЯ}} > E_{\text{ДТ}}$ трансформируется в

$$\rho R > \frac{48 A C_S T}{N_A (\overline{\sigma V})} = f(T). \quad (4)$$

Функция $f(T)$ показана на рис. 1, где она приведена в уточненном по сравнению с (4) виде, взятом из [6]. Видно, что значение «оптической» толщины составляет $\rho R = 0,3$ г/см² при $T = 10$ кэВ. Выражение для энергии, требуемой для нагрева ДТ-топлива, можно переписать в виде

$$E_{\text{ДТ}} = 115 T \frac{4}{3} \pi \frac{(\rho R)^3}{\rho^2} \text{ [МДж]}. \quad (5)$$

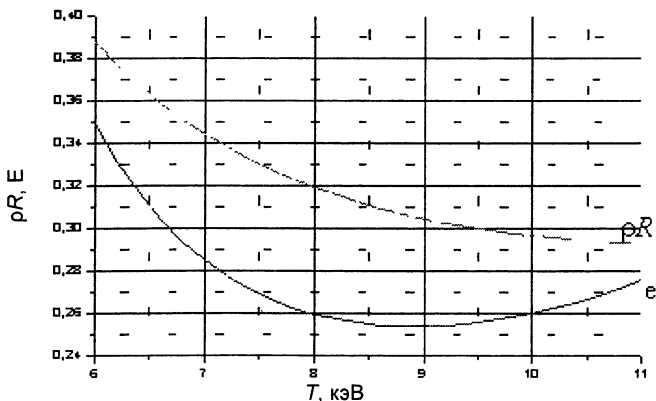


Рис. 1. Оптическая толщина ρR топлива (в г/см²) и энергия мишени (в произвольных единицах) в зависимости от температуры

Если теперь для ρR взять минимально возможное значение, следующее из (4), то получим зависимость $E_{DT}(T)$, также изображенную на рис. 1. Эта функция имеет минимум при $T = 9$ кэВ:

$$E_{DT}(\min) = 115 \cdot 10 \cdot \frac{4}{3} \pi \frac{(0.3)^3}{\rho^2} \approx \frac{130}{\rho^2} \text{ [МДж]}. \quad (6)$$

Таким образом, для уменьшения затрат энергии топливо нужно сжимать. Если, например, удастся сжать до $\rho = 100 \text{ г/см}^3$, E_{DT} составит 13 кДж. Цена вопроса о сжатии высока, если учесть, что не вся энергия лазерного излучения передается топливу.

Действительно, в соответствии с предложением А. Д. Сахарова в ЛТС для сжатия и нагрева DT-топлива его помещают в сферическую оболочку (из СН, Ве, SiO₂ и т. п.) и облучают внешнюю поверхность лазерными лучами (рис. 2, а).

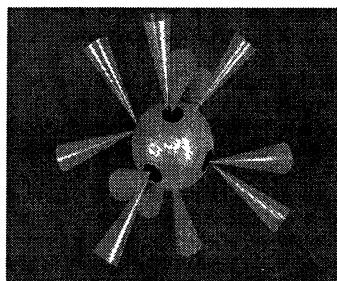
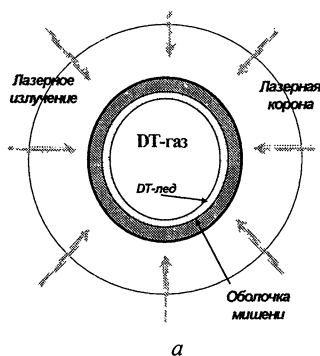


Рис. 2. Схемы лазерных мишеней а – прямого облучения и б – рентгеновская

Оболочка является сферическим вариантом ракеты: внешний слой оболочки испаряется и толкает ее к центру, сжимая и нагревая (сначала при прохождении ударной волны и затем при изэнтропическом сжатии) DT-топливо. Так как воздействующее лазерное излучение является высокоинтенсивным ($I \geq 10^{14} \text{ Вт/см}^2$), испаряемый слой плазмы является высокотемпературной плазмой. Его называют лазерной короной. Температура короны не менее 1 кэВ, поэтому ∇T внутри короны можно с хорошей точностью пренебречь. Отсюда следует, что лазерная корона описывается изотермической волной разгрузки [7]:

$$\rho = \rho_f \exp\left(-\frac{x - x_f}{C_T t}\right), \quad V = C_l + \frac{x}{t}, \quad (7)$$

где ρ_f – плотность на фронте волны, C_T – изотермическая скорость звука. Для изотермической короны энергия (приходящаяся на единицу площади поверхности оболочки), как следует из (7), выражается как

$$E = 4\rho_f C_T^3 t. \quad (8)$$

При записи (8) мы предположили, что показатель адиабаты $\gamma = \frac{C_p}{C_v} = \frac{5}{3}$.

Испаряемая в единицу времени масса (также с единицы площади) есть

$$\dot{m} = \rho_f C_T, \quad (9)$$

а давление короны, передаваемое оболочке через тонкий переходной слой, есть

$$P_a \cong \rho_f C_T^2. \quad (10)$$

Теперь, воспользовавшись уравнениями Циолковского

$$m = m_0 - \dot{m}t, \quad m\ddot{r} = P_a, \quad (11)$$

можно определить скорость оболочки

$$V(t) = \frac{dr}{dt} = \frac{P_a}{\dot{m}} \ln \frac{m_0}{m(t)} = C_T \ln \frac{m_0}{m(t)} \quad (12)$$

и КПД сферической ракеты:

$$\eta = \frac{E_{\text{кин}}}{E} = \frac{4\rho_f C_T^3 t}{\frac{mV^2}{2}} = \frac{1}{8} \frac{m(t)}{m_0 - m(t)} \left[\ln \frac{m_0}{m(t)} \right]^2. \quad (13)$$

Максимальная передача энергии от короны к оболочке происходит при $\frac{m}{m_0} \cong 5$, что приводит к $\eta \cong 0,08$. Более аккуратные численные расчеты

дают обычно значения

$$\eta \cong 0,05 - 0,12. \quad (14)$$

Зная КПД ракеты и энергию нагрева ДТ, можно оценить характерный масштаб энергии лазера, требуемый для зажигания:

$$E_{\text{ДТ}} = E_{\text{лаз}} \eta \eta_1 \Rightarrow E_{\text{лаз}} = E_{\text{ДТ}} (\eta \eta_1)^{-1}. \quad (15)$$

Здесь дополнительно учтено, что энергия, передаваемая от оболочки к топливу в момент торможения, составляет $\eta_1 \cong 0,3$. Принимая $\rho_{\text{ДТ}} = 100 \text{ г/см}^2$, $E_{\text{ДТ}} = 13 \text{ кДж}$, $\eta = 0,08$, $\eta_1 = 0,3$, получим $E_{\text{лаз}} \cong 500 \text{ кДж}$.

В процессе сжатия топлива его объем уменьшается, а плотность нарастает. Ясно, что существует ограничение сверху на сжатие, определяе-

мое отклонением реального процесса от идеальной сферической симметрии. Кроме того, сжатие не может быть меньше некоторого значения, если поставлена задача достичь некоторое конечное значение плотности. Действительно, давление лазерной короны совершает работу $P_a V_0$ (V_0 – начальный объем оболочки), которая с коэффициентом η_1 трансформируется во внутреннюю энергию топлива $P_a V_0 = C_V \rho T V \eta_1^{-1}$, т. е.

$$P_a \delta_V \cong P_a \frac{V_0}{V} = C_V \rho T \eta_1^{-1}. \quad (16)$$

При $T = 10$ кэВ, $C_V = 115 \frac{\text{МДж}}{\text{г} \cdot \text{кэВ}}$, $\eta_1 = 0,5$

$$P_a \delta_V = 25 \cdot 10^6 \left(\frac{\rho}{100 \text{ г/см}^3} \right). \quad (17)$$

В лазерной короне достижимы давления не более $P_a \cong 10^2$ Мбар, что соответствует интенсивности лазерного излучения $I \lambda_\mu^2 \leq 10^{15}$ Вт/см², где λ_μ – длина волны излучения в микронах. При больших интенсивностях неприемлемо большая доля поглощенной плазмой энергии переходит в энергию быстрых электронов. Отсюда следуют неравенства

$$\delta_V \geq 2,5 \cdot 10^4 \left(\frac{\rho}{100 \text{ г/см}^3} \right), \quad \frac{R_0}{R} \geq 30 \left(\frac{\rho}{100 \text{ г/см}^3} \right).$$

Принимая $\rho = 100$ г/см³, $\rho R = 0,3$ г/см², получим $R = 30$ мкм. Следовательно, начальный радиус оболочки должен быть около $R_0 \cong 1$ мкм.

За какое время надо выделить лазерную энергию? Воспользуемся балансом энергии:

$$E_L = 4\pi R_0^2 \tau_L I_a.$$

При $E_L = 500$ кДж, $R_0 = 1$ мм, $I_a = 10^{15}$ г/см², $\lambda = 0,35$ мкм получим $\tau_L \cong 4$ нс.

Набранную оболочкой скорость можно оценить, полагая, что она разгоняется до конечной скорости V при сжатии от R_0 до $\frac{R_0}{2}$:

$$V_{\text{ср}} \tau = \frac{V}{2} \tau = \frac{R_0}{2},$$

откуда следует, что $V \cong 3 \cdot 10$ см/с. Глубокие сжатия требуют высокой симметрии облучения. Если давление на фронте абляции отличается на

некотором участке поверхности оболочки на δP_A , то скорость оболочки в этом месте изменится на $\delta V = V \frac{\delta P_a}{P_a}$. При схождении к центру положение

этого участка δR будет отличаться от среднего на $\delta R = \delta V \frac{R_0}{V} = \frac{\delta P_a}{P_a} R_0$.

Отклонения от сферической симметрии можно считать сравнительно малыми, если $\delta R \leq \frac{1}{3} R$, т. е.

$$\frac{\delta P_a}{P_a} \leq \frac{1}{3} \frac{R}{R_0}. \quad (18)$$

При $\frac{R}{R_0} = 30$ получаем, что неоднородность давления, а значит и лазерной интенсивности, должна быть не выше одного процента.

Проблема получения столь высокой однородности облучения при прямом лазерном воздействии на поверхность оболочки до сих пор не решена. Вдобавок в мишенях прямого облучения необходимы большие запасы по энергии (порядка тройки) для преодоления разрушительного эффекта потери симметрии при развитии неустойчивости Рэлея – Тейлора. Последняя развивается в процессе разгона оболочки, так как легкая лазерная корона давит на тяжелую оболочку, а также на стадии торможения, когда легкий DT-газ тормозит оболочку. В связи с этим последние 15 лет интенсивно разрабатывается альтернативный подход, основанный на конверсии лазерного излучения в рентгеновское.

1.2. Рентгеновские мишени. В мишенях непрямого воздействия² лазерное излучение трансформируется в рентгеновское внутри бокса-конвертора, который имеет обычно цилиндрическую либо сферическую форму (рис. 2, б). Внутренняя поверхность бокса покрывается материалом с большим Z , обычно золотом. Лазерное излучение вводится внутрь бокса через отверстия ввода, относительная площадь которых невелика ($\leq 0,1$). Попадая на внутреннюю стенку бокса, лазерное излучение вызывает образование лазерной короны с температурой $1 \div 4$ кэВ ($I_L \cong 10^{14} - 10^{15}$ Вт/см²) и характерной плотностью порядка 10^{-2} г/см³ (для $\lambda = 0,35$ мкм). Плазма из золота в таких условиях является эффективным излучателем: около 50–60 % поглощенной энергии высвечивается в виде неравновесного излучения в диапазоне $h\nu \leq 3$ кэВ. Половина энергии, излученной из

² Во ВНИИЭФ первое рассмотрение подобного подхода было выполнено в 1968 г Н Б Бабичевым, С Б Кормером, Н А Поповым (см [1]) С Б Кормер внес также определяющий вклад в становление исследований по ЛТС во ВНИИЭФ и создание мощных лазерных установок «Искра-4» [15] и «Искра-5»

данного элемента лазерной энергии, распространяется от стенки (рис. 3). Другая часть, суммируясь с излучением, пришедшим от участков с другими угловыми координатами, приводит к образованию новой волны разгрузки – рентгеновской короны с температурой 0,2–0,3 кэВ и характерной плотностью 0,3–0,7 г/см³. Из-за большого различия в плотности лазерная корона практически не влияет на рентгеновскую. Эффективную энергетическую температуру T рентгеновского излучения, генерируемого внутри бокса можно оценить исходя из баланса мощности [8, 9]:

$$\frac{E_L \xi_x}{S_{\text{бокса}} \tau_L} = \sigma T^4 [\beta + (1-\beta)(1-\alpha)] = I_{\text{cp}} \xi_x, \quad (19)$$

где $S_{\text{бокса}}$ – площадь внутренней поверхности бокса, E_L и τ_L – энергия и длительность лазерного импульса, ξ_x – коэффициент первичной конверсии лазерного излучения в рентгеновское в лазерной короне, β – относительная площадь отверстий, I_{cp} – средняя интенсивность облучения стенок лазерным излучением, α – альbedo рентгеновской короны. Для определения температуры T по формуле (19) нужно знать ξ_x и β . Как показывают расчеты и аналитическая модель [8], величина ξ_x в интересующих нас условиях меняется сравнительно слабо и составляет $\xi_x \cong 0,5-0,6$. Что касается альbedo, то эта величина является возрастающей функцией температуры и времени: $\alpha = \alpha(T, \tau_L)$ [8, 9] и составляет величину $\alpha = 0,7-0,9$ при $\tau_L = 0,3-3$ нс, $I_{\text{cp}} = 10^{14}-10^{15}$ Вт/см². Для $I_{\text{cp}} = 3 \cdot 10^{14}$ Вт/см², $\alpha = 0,8$, $\xi_x = 0,6$, $\beta = 0,1$ получаем $T \cong 290$ эВ.

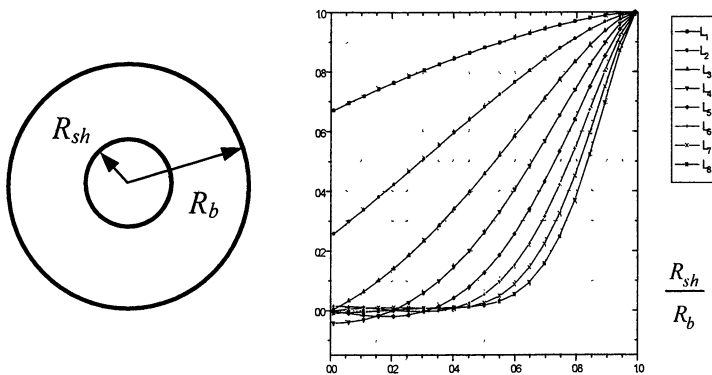


Рис. 3. Зависимость степени сглаживания от отношения радиуса оболочки к радиусу бокса

$\frac{R_{sh}}{R_b}$ для сферических гармоник с номерами l_i . При $\frac{R_{sh}}{R_b} \leq 0,25$ сглаживаются все гармоники кроме первой, второй и третьей

Если оболочка с ДТ-топливом покрыта пластиком, то давление на ее поверхность может достигать значения близкого к $P = \frac{2}{3} C_V \rho T \cong 200$ Мбар, соответствующего идеальному газу с плотностью, равной начальной плотности пластика ρ и нагретому до температуры T . Степень близости зависит от соотношения между внутренней и кинетической энергиями, которое при высоких T и малых τ_L может быть малым.

Основное достоинство непрямого метода состоит в возможности получения сравнительно однородного облучения, что связано с двумя причинами: 1) некогерентным характером рентгеновского поля (складываются интенсивности); 2) мелкомасштабные неоднородности рентгеновского излучения, исходящего с поверхности бокса, сглаживаются по мере распространения от стенок к оболочке. Количественно последнюю тенденцию удобно изобразить в виде коэффициентов ослабления различных сферических гармоник как функций отношения радиуса оболочки r к радиусу бокса R (рис. 4). Из рисунка следует, что при $\frac{r}{R} \cong 0,15$ «выживают» только гармоники с номером $l \geq 4$.

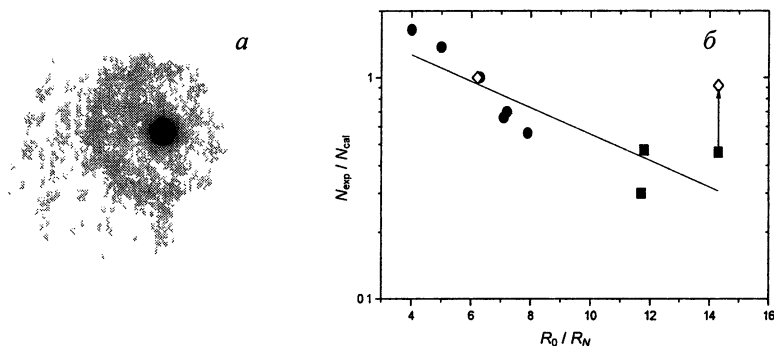


Рис. 4. Обскурограмма центральной мишени (а) и зависимость отношения экспериментального и расчетного нейтронного выхода от расчетной степени радиального схождения центральной капсулы на момент генерации нейтронов (б): ● – оболочка без покрытия, ■ – оболочка с полипарасилиленовым покрытием, ◇ – расчеты с учетом турбулентного перемешивания

Недостатком непрямого метода облучения является его сравнительно невысокая энергетическая эффективность. Действительно, оболочке с ДТ-топливом передается доля энергии рентгеновского излучения

$$\eta_{\Pi} = \frac{S_{06}}{S_6} [\beta + (1-\beta)(1+\beta)] + S_{06} = \frac{1}{1 + \frac{S_6}{S_{06}} [\beta + (1-\beta)(1+\beta)]}$$

При $\frac{R_{об}}{R_6} = 0,2$ $\frac{S_{об}}{S_6} = 0,04$. Принимая, как и раньше, $\beta = 0,1$, $\alpha = 0,8$, получаем $\eta_{\Pi} \cong 0,18$. Учитывая, что коэффициент первичной конверсии лазерного излучения в рентгеновское $\xi_v \cong 0,6$, получаем $E_{об} \cong 0,18 \cdot 0,6 E_L \cong 0,1 E_L$. В итоге характерная энергия лазера с энергией, достаточной для зажигания в режиме непрямого облучения, составляет $\cong 1,5 \cdot 10^6$ Дж.

Определенные надежды на снижение требуемой энергии несет схема двухэтапной работы, когда процессы сжатия и нагрева ДТ-топлива происходят раздельно. Для пояснения обратимся к формуле (5) и предположим, что значения $\rho R = 0,3$ г/см² и $\rho = 100$ г/см³ прежние, а температура T заметно меньше 10 кэВ, например, $T = 0,5$ кэВ. В этом случае требуемая энергия лазера, затрачиваемая на сжатие, составит вместо 500 кДж около $E_L \cong 25$ кДж. В результате первого этапа мы будем иметь сферически симметричное сжатое топливо с необходимым значением параметра ρR , но с недостаточно высокой температурой. Время существования такого сгустка по-прежнему определяется формулой (3). Если $R = 30$ мкм, $T = 0,5$ кэВ, $C_S \cong 2,6 \cdot 10^7$ см/с, следовательно, $\tau \cong 40$ нс. Поэтому, если удастся за время меньшее τ , например за 10 нс, нагреть топливо до температуры $T \cong 10$ кэВ, то оно вспыхнет. Соответствующий режим иногда называют быстрым зажиганием (русская версия английского термина «fast ignition» [10]). Для второй фазы при этом требуется около 15–30 кДж. Данное направление в настоящее время также исследуется в ряде лабораторий мира, в частности в университете г. Осака.

2. Обзор экспериментов с мишенями непрямого облучения на установке «Искра-5»

2.1. Исследование работы рентгеновской мишени в однородном рентгеновском поле.

Основным направлением экспериментальных исследований на установке «Искра-5» были исследования работы рентгеновской мишени, состоящей из сферического медного бокса диаметром 1,3–4 мм, в центре которого помещалась сферически симметричная стеклянная капсула, содержащая ДТ-газ (см. рис. 2, б) [11]. Номинальным значением диаметра бокса было 2 мм. На внутреннюю поверхность бокса наносился слой золота толщиной около 1 мкм. Этот слой служил конвертером лазерного излучения установки «Искра-5», которое вводилось внутрь бокса через 6 отверстий и облучало его внутреннюю поверхность. В типичных экспериментах энергия лазерного излучения составляла около 10 кДж при длительности импульса $\sim 0,35$ нс.

Рентгеновское излучение, генерируемое в боксе с температурой около 150 кэВ, поглощалось центральной капсулой, что приводило к сжатию оболочки и нагреву ДТ-газа. При отношении радиуса бокса к радиусу

центральной капсулы в диапазоне 5–10 расчетная неоднородность рентгеновского потока на поверхности капсулы составляет не хуже 2–3 %.

Кроме рентгеновского излучения, на капсулу воздействует плазма разлетающейся лазерной короны. Поскольку скорость движения вещества ($V \cong 10^7\text{--}10^8$ см/с) существенно меньше скорости света, то существует интервал времени $\tau \approx \varnothing_{\text{бок}} / 2V \cong 1$ нс (при $\varnothing_{\text{бок}} \cong 2$ мм), в течение которого оболочка не подвергается воздействию потока вещества стенок бокса. При длительности лазерного импульса порядка 0,3 нс и времени схлопывания 0,5 нс воздействием потока вещества можно пренебречь. Однако сравнительно большая длина волны лазерного излучения установки «Искра-5» при интенсивностях не менее $3 \cdot 10^{14}$ Вт/см² приводит к генерации горячих электронов с температурой $T_h \cong 20\text{--}50$ кэВ, которые, в свою очередь, ускоряют некоторую часть ионов (Au) до скоростей $(3\text{--}5) \cdot 10^8$ см/с. В соответствии с этим время τ уменьшается до 0,3 нс, поэтому для оболочек со сравнительно большим временем схождения ($\geq 0,5$ нс) необходимо иметь в виду возможность дополнительного энергоклада в оболочку мишени путем переноса энергии быстрыми ионами.

Основные вопросы, которые ставились в экспериментах с такого типа мишенями, заключались в исследовании характеристик рентгеновского поля, его спектра, эффективной температуры, степени симметрии. Кроме того, для понимания процессов, происходящих в мишени, нужно знать такие характеристики, как скорость полета оболочки, степень сжатия DT-газа и его температуру. С целью исследования этих вопросов в экспериментах варьировались такие параметры центральной капсулы, как ее диаметр (от 270 мкм до 0,9 мм), толщина (от 1 до 40 мкм), начальное давление DT-газа (от 3 до 50 атм). Для увеличения толщины на некоторые мишени наносился слой аблятора из полипараксилилена или материала, содержащего кремний.

Проведен большой цикл экспериментов, в которых было показано, что выбранная геометрия облучения обеспечивает высокую степень симметрии рентгеновского поля на поверхности центральной капсулы. Фотография рентгеновского свечения центральной мишени, полученная с помощью камеры-обскуры (см. рис. 4), показывает высокую степень симметрии сжатой области, что подтверждает расчетные оценки.

Регистрация спектра рентгеновского излучения, генерируемого стенками бокса, показала, что спектр является неравновесным. Абсолютные измерения энергии рентгеновского излучения показали, что рентгеновский поток, облучающий поверхность стеклянной капсулы, соответствует эффективной температуре 160–170 эВ.

Измерения времени схождения оболочки $\tau_{\text{гр}}$, проведенные с помощью щелевого рентгеновского фотохронографа, позволяют оценить типичную скорость ее движения при толщине оболочки примерно 5–7 мкм величиной, равной около $3 \cdot 10^7$ см/с.

Интегральный нейтронный выход измерялся с помощью комплекса методик, основанных на активации меди и индия, а также методе затяннутой регистрации. Спектр нейтронов, а затем и температура ионов DT определялись по времяпролетной методике, имеющей два канала с пролетными базами $L_1 = 1249$ см и $L_2 = 1676$ см. Нейтронный выход изменялся, в зависимости от параметров оболочки и условий эксперимента, в диапазоне $10^7 - 10^{10}$. Для мишени диаметром 280 мкм и толщиной оболочки 5 мкм нейтронный выход был на уровне $3 \cdot 10^9$, что отвечает температуре DT-газа, измеренной с помощью времяпролетной методики, 2,5 кэВ.

Анализ экспериментальных результатов проводился на основе расчетов по программе СНДП [12], представляющей собой одномерный код расчета радиационной газовой динамики. Анализ показал, что генерация нейтронного выхода удовлетворительно описывается в рамках сферически симметричных расчетов (рис. 5) вплоть до объемных степеней сжатия газа порядка 10^3 (это соответствует 10-кратному радиальному сжатию оболочки на рис. 4). При еще больших степенях сжатия снижение нейтронного выхода по отношению к сферически-симметричному расчету может быть объяснено влиянием турбулентного перемешивания.



Рис. 5. Схема эксперимента по созданию на поверхности сферически-симметричной капсулы асимметричного рентгеновского поля (а) и зависимость нейтронного выхода и степени неоднородности рентгеновского поля на поверхности капсулы σ_{rms} от величины сдвига мишени относительно центра бокса (б)

2.2. Эксперименты с преднамеренно асимметричным облучением.

В другой серии экспериментов осуществлялось нарушение однородности рентгеновского поля на поверхности сферически симметричной стеклянной капсулы. Такую неоднородность можно создавать различными способами. Мы выбрали метод, базирующийся на введении дополнительного отверстия в боксе и сдвиге сферической капсулы в направлении этого дополнительного отверстия (рис. 5). Измеренный нейтронный выход в этой серии менялся на 2–3 порядка величины.

Проведенное сравнение полученных экспериментальных результатов с результатами газодинамических расчетов сжатия центральных капсул по программе МИМОЗА-НД [12], с параметрами мишени и рентгеновского импульса, соответствующими эксперименту (см. рис. 5), позволяет констатировать качественное и количественное согласие между экспериментальными и расчетными данными в широком диапазоне изменения асимметрии рентгеновского поля. Наблюдается удовлетворительное согласие расчетных и экспериментальных значений нейтронного выхода во всем исследованном диапазоне сдвигов, за исключением эксперимента со сдвигом $\Delta = R_{\text{box}}$.

Наиболее быстро увеличение асимметрии происходит при приближении мишени к отверстию для ввода лазерного излучения на расстояние, сравнимое с ее диаметром. В этом случае по DT-газу распространяется почти плоская ударная волна, на фронте которой и происходит образование нейтронов. Полное их количество при этом на 2÷3 порядка меньше, чем при центральном расположении мишени.

3. Установка «Луч» – модуль мощной лазерной установки «Искра-6»

В РФЯЦ–ВНИИЭФ совместно с кооперацией ведущих институтов страны по лазерной науке и технике проводятся работы по созданию мощной лазерной установки на неодимовом стекле «Искра-6» с энергией излучения в импульсе порядка 300 кДж, длительностью импульса 3 нс, длиной волны излучения 0,35 мкм [4]. Она предназначена для изучения свойств высокотемпературной плотной плазмы и решения ряда задач термоядерного синтеза. Установка «Искра-6» будет состоять из 16 модулей, каждый из которых должен содержать восемь лазерных каналов, всего в установке «Искра-6» будет 128 лазерных каналов, или лазерных лучей.

В качестве первого шага разрабатывается модуль установки «Искра-6» установка «Луч». Установка «Луч» содержит четыре лазерных канала с полной энергией излучения порядка 16 кДж на основной частоте (~12 кДж на третьей гармонике). Основные параметры установки «Луч»: число каналов – 4; энергия излучения на $1\omega_0$ – 16 кДж; энергия излучения на $3\omega_0$ – 12 кДж; форма импульса излучения – профилированная; длительность основного импульса – регулируемая, $\tau_{0,5} = 3 \div 5$ нс; длительность предимпульса – регулируемая, $\tau_{0,5} = 5 \div 15$ нс; размер пучка – 200 × 200 мм; расходимость излучения – $\Theta_{0,8} \leq 10^{-4}$ рад; энергия конденсаторной батареи – 5 МДж.

Основная цель создания установки «Луч» – отработать ключевые элементы лазерного канала, проверить правильность выбора конструкции модуля установки «Искра-6», реализовать заданные параметры излучения канала, разработать конструкторскую документацию на основные узлы

установки. На установках «Искра-6» и «Луч» так же, как и на установке NIF (ЛЛНЛ, США), используются многопроходные лазерные усилители, имеющие некоторые преимущества по сравнению с системой прямого усиления. Основное из них – высокая унификация усилительных элементов, что приводит к повышению компактности установки и снижению ее стоимости.

В настоящее время работы по созданию установки «Луч» ведутся достаточно активно. Разработана конструкторская документация на основные узлы установки, и они в основном изготовлены. Создан стенд для отработки системы формирования опорного излучения. Временное профилирование импульса излучения осуществляется двумя затворами Поккельса, управляемыми формирующими длинными линиями. Длительность пьедестала регулируется в пределах 5 – 10 нс, амплитуда – от 10^{-3} до 10^{-1} от интенсивности основного импульса. Передний фронт основного импульса профилируется по заданному закону. Зарегистрированы пространственные параметры лазерного луча после прохождения излучения задающего генератора двухпроходного преусилителя. Качество пучка (ближняя и дальняя зоны) близки к дифракционному. Стабильность длины волны излучения задающего генератора составляет примерно $0,1 \text{ \AA}$. Отрабатывается адаптивное зеркало, предназначенное для корректировки волнового фронта излучения при вводе его в усилительные каналы. Создан энергетический модуль установки «Луч» с энергозапасом около 300 кДж, на котором будут отрабатываться лампы накачки, их надежность, ресурс работы и полупроводниковый твердотельный разрядник (РВД-25), который, по предварительным данным, выдерживает 10^5 – 10^6 включений при токе порядка 150 кА.

Физический пуск установки «Луч» планируется осуществить в конце 2002 года.

* * *

Сорок лет исследований по ЛТС привели к значительному прогрессу в понимании физики процессов и развитию технологии, но проблема зажигания ДТ-топлива все еще не решена. Ближайшие 10 лет пройдут под знаком строительства лазеров мегаджоульного уровня и в ожидании первых попыток зажечь термоядерное топливо.

Литература

- 1 *Попов Н. А.* АДС и ЛТС // Вопросы атомной науки и техники. Сер. Математическое моделирование физических процессов 1992 Вып 4 С 53–54
- 2 *Басов Н. Г., Крохин О Н* Условия разогрева плазмы излучением оптического генератора // ЖЭТФ. 1964. Т 46(1) С 171
- 3 *Кириллов Г. А., Кочемасов Г Г* и др Обзор исследований по физике горячей и плотной плазмы, выполненных на лазерной установке «Искра-5» Доклад на IV Харитоновских научных чтениях, Саров, 2002

4 Кириллов Г А, Кочемасов Г Г и др Обзор программы ЛТС в РФЯЦ-ВНИИЭФ // Труды РФЯЦ-ВНИИЭФ Саров, 2001. Вып 1 С. 242–253

5 Арцимович Л А Управляемые термоядерные реакции 2-е издание. М, 1963

6 Мхитарьян Л С Физика работы мишеней для ЛТС // Научно-популярный журнал «Атом» 1999 С 16–19

7 Ландау Л Д, Лифшиц Е М Гидродинамика, М Наука, 1986

8 Mkhitarian L S, Kochemasov G G Laser-X-ray conversion, In Laser Interactions with Matter // Proceedings of the 23 rd ECLIM Oxford. September, 1994 P 351–354

9 Lindl J D Development of the indirect-drive approach to ICF and the target physics basis for ignition and gain // Phys Plasmas 1995 V. 2 P 3933–4073

10 Tabak M, Hammer J, Glinsky M E et al Ignition and high gain with ultrapowerful lasers // Phys. Plasmas 1994 V. 1 P 1626–1634

11 Абзаев Ф М, Бельков С А, Бессараб А В и др Сжатие и нагрев сферических термоядерных мишеней при непрямом (рентгеновском) облучении на установке «Искра-5» // ЖЭТФ. 1998 Т. 114, вып 1 С 155–170

12 Бельков С А, Долголева Г В Модель среднего иона для расчета кинетики ионизации населенностей возбужденных уровней и спектральных коэффициентов переноса излучения в программе СНДП // Вопросы атомной науки и техники Сер Математическое моделирование физических процессов 1992. Вып. 1 С. 59

13 Mkhitarian L S, Bel'kov S A, Bessarab A V et al Asymmetry influence on indirect-drive Iskra-5 implosions // Inertial Fusion Science and Applications 99 / Editors C Labanne, W J Hogan, K A Tanaka Elsevier, 2000. P. 142.

14 Софронов И Д, Бельков С А, Винокуров О А и др. Методика расчета спектрального переноса излучения в двумерном комплексе «Мимоза-НД» // Вопросы атомной науки и техники Сер. Математическое моделирование физических процессов 2000. Вып. 1 С 8

15. Корнер С Б Фотодиссоциационные лазеры для управляемого термоядерного синтеза // Изв. АН СССР. Сер физическая 1980 Т 44, вып 10 С 2002–2020

НЕЛИНЕЙНЫЕ ВОЛНЫ В ЭЛЕКТРОННОЙ МАГНИТНОЙ ГИДРОДИНАМИКЕ

А. С. Кингсен

Электронная магнитная гидродинамика (ЭМГ) – сравнительно молодая область теоретической физики плазмы, развивающаяся преимущественно в связи с развитием программ импульсного управляемого термоядерного синтеза, а в последнее время – также и астрофизики. Проявляться ЭМГ-эффекты могут не только в плазме, но и в металлах – твердых и жидких, в полупроводниках, тонких пленках и т. д. Преимущественно наше изложение будет следовать обзорам [1] и отчасти [2] с надлежащей адаптацией применительно к задачам данной статьи. Последние два десятилетия ознаменовались «взрывным» развитием ЭМГ. Иногда возникает невольный вопрос: почему столь очевидные и наглядные эффекты не были обнаружены ранее? Возможная причина – в том, что становление данного направления происходило в общем контексте нелинейной физики плазмы. Существенным стимулом прогресса послужили успехи теории турбулентности и нелинейных структур.

Основные уравнения

В основе модели ЭМГ лежит известная иерархия времен релаксации:

$$\tau_{ee} : \tau_{ii} : \tau_{ei} = 1 : \sqrt{M/m} : (M/m), \quad (1)$$

которой обусловлена применимость двухкомпонентной (в общем случае – многокомпонентной) магнитной гидродинамики. В самом деле, из соотношений (1) следует, что равновесие внутри каждого из компонентов устанавливается значительно раньше, чем полное равновесие. Это дает нам право в широком диапазоне параметров пользоваться парой уравнений Эйлера, связанных лишь силой трения между компонентами:

$$\begin{aligned} n \frac{d(m\mathbf{v}_e)}{dt} &= -ne \left(\mathbf{E} + \frac{1}{c} [\mathbf{v}_e, \mathbf{B}] \right) - \nabla P_e - R_{ie} n_i n (\mathbf{v}_e - \mathbf{v}_i), \\ n \frac{d(M\mathbf{v}_i)}{dt} &= Zn_i e \left(\mathbf{E} + \frac{1}{c} [\mathbf{v}_i, \mathbf{B}] \right) - \nabla P_i - R_{ie} n_i n (\mathbf{v}_i - \mathbf{v}_e), \end{aligned} \quad (2)$$

вместе с уравнениями переноса и уравнениями непрерывности для каждого компонента и уравнениями Максвелла. Однако даже упрощенная таким образом система уравнений оказывается намного сложнее и труднее для количественного исследования, чем, например, та, что описывает магнитную гидродинамику (МГД). Но в определенном «окне параметров», как будет ясно из дальнейшего, она может быть редуцирована к виду не более сложному, чем МГД. Прежде всего, здесь и далее мы будем

обсуждать лишь такие процессы и явления, в которых не нарушается квазинейтральность. Это, в частности, означает, что токовая скорость не превосходит электронной тепловой. (В обратном случае в сценарий вмешивается неустойчивость Бунемана, приводящая к образованию двойных слоев.) Наложим еще одно ограничение: рассмотрим такой класс движений, в которых смещение ионов пренебрежимо мало из-за большой инерции. Тогда можно отбросить второе из уравнений (2), но количественно это означает, что скорость течения электронов (токовая скорость) должна быть много больше, чем характерная массовая, в качестве таковой естественно взять альфвеновскую или ионно-звуковую. Эти оценки совпадают, если учесть, что в результате самоорганизации динамической системы естественно ожидать установления одинаковых по порядку величин газокINETического и магнитного давления: $B^2 \sim nT$. Итак, мы считаем выполненной следующую цепочку неравенств:

$$c_A, v_A \ll j/ne \ll v_{Te}, v_{Ae}. \quad (3)$$

Обозначим характерный пространственный масштаб задачи через a . По условию квазинейтральности $a \gg c/\omega_{pe}$. Тем самым можно пренебречь и током смещения; соответственно, теорема о циркуляции

$$\text{rot } \mathbf{B} = \frac{4\pi}{c} \mathbf{j} \quad (4)$$

трансформируется в оценку $B \approx (4\pi/c)ja$. Как следствие, (3) можно переписать в виде $ja\sqrt{4\pi/Mnc^2} \approx B/\sqrt{4\pi nM} \ll j/ne \ll B/\sqrt{4\pi nM} \approx ja\sqrt{4\pi/mnc^2}$, а из него, в свою очередь, следует

$$c/\omega_{pi} \gg a \gg c/\omega_{pe}. \quad (5)$$

Условия (3) и (5) определяют пределы применимости ЭМГ. В качестве основных уравнений ЭМГ удобно использовать уравнения диффузии и замороженности магнитного поля. Напомним сначала, как они выглядят в МГД-приближении. Достаточно очевидная цепочка равенств

$$\mathbf{E} + \frac{1}{c}[\mathbf{v}, \mathbf{B}] = \frac{1}{\sigma} \mathbf{j} = \frac{c}{4\pi} \text{rot } \frac{1}{\sigma} \mathbf{B}$$

в сочетании с уравнением электромагнитной индукции

$$c \text{rot } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

дает в предположении $\sigma = \text{const}$

$$\frac{\partial \mathbf{B}}{\partial t} = \text{rot}[\mathbf{v}, \mathbf{B}] + \frac{c^2}{4\pi\sigma} \nabla^2 \mathbf{B}. \quad (6)$$

Уравнение (6) содержит два эффекта: диффузию поля ($\nabla^2 \mathbf{B}$) и замороженность поля в проводящую среду ($\text{rot}[\mathbf{v}, \mathbf{B}]$). Хорошо известны многие проявления последнего эффекта – ограничимся указанием на гигантские магнитные поля нейтронных звезд как следствие замороженности поля в вещество при гравитационном коллапсе. При рассмотрении двухкомпонентной МГД (см. (2)) возникает вопрос о том, что же и куда в действительности заморожено, или, точнее, какой более общий закон вырождается в пределе МГД в уравнение (6). Чтобы ответ выглядел как можно более ясным, упростим насколько возможно нашу задачу, положив $\sigma = \text{const}$ и $P_\alpha = P_\alpha(n_\alpha)$; $\alpha = i, e$ – это закон типа политропы. Взяв ротор от обоих уравнений (2) и комбинируя импульсы $\mathbf{p}_{e,i}$ с вектор-потенциалом \mathbf{A} (напоминаем: $\mathbf{B} \equiv \text{rot}\mathbf{A}$) в обобщенном импульсе компонента $\alpha = i, e$

$$\mathbf{P}_\alpha = \mathbf{p}_\alpha + \frac{e_\alpha}{c} \mathbf{A}, \quad (7)$$

получим систему двух уравнений:

$$\frac{\partial}{\partial t} \text{rot}\mathbf{P}_e = \text{rot}[\mathbf{v}_e, \text{rot}\mathbf{P}_e] - \frac{e}{\sigma} \text{rot}\mathbf{j}, \quad (8)$$

$$\frac{\partial}{\partial t} \text{rot}\mathbf{P}_i = \text{rot}[\mathbf{v}_i, \text{rot}\mathbf{P}_i] + \frac{Ze}{\sigma} \text{rot}\mathbf{j}. \quad (9)$$

В пределе однокомпонентной МГД ($\mathbf{p}_e \rightarrow 0$, $\mathbf{v}_e \approx \mathbf{v}_i$) в случае, например, идеальной проводимости $\sigma \rightarrow \infty$ система уравнений (8), (9) и в самом деле может быть сведена к одному уравнению типа (6). Если при этом в обобщенном импульсе доминирует его механическая компонента \mathbf{p}_i , то $\text{rot}\mathbf{v}$ оказывается заморожен в вещество – это идеальная гидродинамика невязкой несжимаемой жидкости. В обратном случае в вещество заморожено поле, а в общем случае, как можно видеть из системы уравнений (8)–(9), в каждый компонент системы заморожен ротор обобщенного импульса данного компонента.

Перейдем к случаю ЭМГ. Тогда существенным остается лишь одно уравнение (8); более того, как правило, само оно может быть упрощено при использовании двух малых параметров:

$$\partial/\partial t \ll \omega_{Be}; \quad c/\omega_{pe} \ll a.$$

В результате получаем простой ответ:

$$\frac{\partial \mathbf{B}}{\partial t} + \text{rot}[\mathbf{j}/ne, \mathbf{B}] = -\frac{c}{\sigma} \text{rot}\mathbf{j}, \quad (10)$$

т. е. поле оказывается замороженным в электронное течение. Уравнение (10) можно использовать вместе с уравнением $\text{div}\mathbf{j} = 0$, но оно превращается в тождество, как следствие (4), и оба они вместе справедливы при

$\partial/\partial t \ll \omega_{pe}$, т. е. когда ток смещения несуществен, а электронное течение не нарушает квазинейтральности. В принципе уравнение (10) может быть получено непосредственно из обобщенного закона Ома в пренебрежении электронной инерцией, но с учетом эффекта Холла:

$$0 = -ne\mathbf{E} + \frac{1}{c}[\mathbf{j}, \mathbf{B}] + \frac{ne}{\sigma} \mathbf{j}. \quad (11)$$

Взяв ротор от уравнения (11), снова получаем (10). Этот вывод, конечно, короче, но не столь определенны границы применимости. Во всяком случае, очень часто можно вместо ЭМГ рассматривать МГД с доминирующим эффектом Холла.

Что касается движения ионов (см. (9)), то, вообще говоря, оно может быть нетривиальным, и его описание не сведется к уравнению в замороженности поля, ибо не может быть поле заморожено сразу в два компонента. Как правило, в границах применимости ЭМГ ионное движение оказывается близким к идеальной гидродинамике с замороженным ротором скорости. При этом характерные электронные и ионные времена существенно различны: можно рассматривать динамику электронных течений на фоне медленной вариации внешних условий, вызванной ионным движением.

Нелинейный транспорт поля. Конвективная волна переноса

Пусть плазма занимает полупространство $z > 0$ и в момент времени $t = 0$ поле в плазме $B = 0$, и в то же время в любой момент поле на границе равно постоянной величине \mathbf{B}_0 . Для определенности мы вначале рассмотрим плоскую геометрию с $\mathbf{B} \parallel \mathbf{e}_y$ и $\partial/\partial y = 0$ (рис. 1).

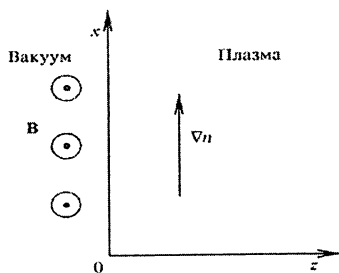


Рис. 1. Геометрия задачи о конвективном скин-эффекте

В такой плоской геометрии уравнение (10) может выродиться в самое ординарное МГД-уравнение диффузии магнитного поля. В самом деле,

$$[\mathbf{j}, \mathbf{B}] = -\frac{c}{8\pi} \nabla B^2 + \frac{c}{4\pi} (\mathbf{B} \nabla) \mathbf{B}. \quad (12)$$

Второй член в правой части (12), который один и может дать нелинейный вклад в rot $[\mathbf{j}, \mathbf{B}]$, в случае прямых силовых линий тождественно равен нулю, поскольку напрямую связан с вектором кривизны магнитных силовых линий с:

$$(\mathbf{B}\nabla)\mathbf{B} = B^2\mathbf{c}, \quad |\mathbf{c}| \equiv R_B^{-1}, \quad \mathbf{c} \parallel -\mathbf{R}_B,$$

где \mathbf{R}_B – текущий радиус кривизны. Однако конвективный член в уравнении (10), т. е. второй в левой части, может и не быть равным нулю, если существует отличный от нуля градиент концентрации. Чтобы «очистить» эффект от всего лишнего, положим опять $\sigma = \text{const}$, и тогда

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{c}{8\pi e} [\nabla B^2, \nabla(n^{-1})] = \frac{c^2}{4\pi\sigma} \nabla^2 \mathbf{B}. \quad (13)$$

По-прежнему стремясь к максимально простому ответу, рассмотрим такую геометрию задачи, когда вектор ∇n параллелен поверхности плазмы и ортогонален \mathbf{B} , как это изображено на рис. 1. Тогда нелинейный конвективный член обуславливает трансляцию поля в направлении оси z . Пока глубина проникновения не слишком велика, что определяется неравенством

$$\frac{\partial^2}{\partial z^2} \gg \frac{\partial^2}{\partial x^2}, \quad (14)$$

обычным для хрестоматийной задачи о скин-эффекте, уравнение (13) может быть представлено в классическом виде уравнения Бюргерса:

$$\frac{\partial B}{\partial t} + \kappa B \frac{\partial B}{\partial z} = D \frac{\partial^2 B}{\partial z^2}, \quad (15)$$

где $\kappa = (c/4\pi e)(\partial n^{-1}/\partial x)$ и $D = c^2/4\pi\sigma$.

Итак, даже в самой простой геометрии и в наиболее упрощающих задачу предположениях она в рамках ЭМГ (с учетом первого ненулевого ЭМГ-эффекта) уже оказывается нелинейной. Но нелинейность эта – особого рода. Дело в том, что уравнение Бюргерса относится к так называемым скрытолинейным уравнениям. Несложной подстановкой

$$B = (2D/\kappa)(\partial \ln |\varphi| / \partial z) \quad (16)$$

оно переводится в линейное уравнение диффузии

$$\frac{\partial \varphi}{\partial t} = D \frac{\partial^2 \varphi}{\partial z^2}, \quad (17)$$

и притом множества решений обоих уравнений совпадают. В частности, из наших начально-граничных условий

$$\begin{aligned} t = 0: \quad B = 0, \quad z > 0, \\ t \geq 0: \quad B = B_0, \quad z = 0 \end{aligned} \quad (18)$$

следует точное решение:

$$\varphi = -\operatorname{erf}\left[z/2\sqrt{Dt}\right] + \exp\left[(\kappa B_0/2D)(z + \kappa B_0 t/2)\right] \times \operatorname{erfc}\left[(z + \kappa B_0 t)/2\sqrt{Dt}\right]. \quad (19)$$

Единственная польза, которую из него можно извлечь, состоит в демонстрации точной интегрируемости уравнения (15); будучи нелинейным, оно остается столь же полностью интегрируемым, как и уравнение диффузии, описывающее стандартный скин-эффект. Для понимания физического сценария лучше все же исходить непосредственно из уравнения (13). Ясно, что на ранней стадии процесса переноса поля (точнее, при $t < D/\kappa^2 B_0^2$), пока велик градиент B , проникновение поля в плазму обусловлено обычной диффузией. Но когда профиль B в достаточной мере сгладится, на первый план выходит перенос поля токовыми электронами (см. уравнение (10)), и характер решения меняется.

Рассмотрим вначале тот случай, когда векторы \mathbf{B} , ∇n и направление вглубь плазмы \mathbf{e}_z (или $-\nabla B^2$) образуют левую тройку, как показано на рис. 1. Для такой геометрии $\kappa < 0$ в уравнении (15), а это значит, что нелинейный вынос поля конкурирует с диффузионным вносом, и в результате при $t \rightarrow \infty$ устанавливается стационарное состояние

$$B = B_0 / (1 - \kappa B_0 z / 2D) \quad (20)$$

(во избежание недоразумений напомним: $\kappa < 0$) – это асимптотика решения (19):

$$\varphi_{t \rightarrow \infty} \rightarrow (\pi Dt)^{-1/2} (2D/\kappa B_0 - z).$$

Уравнение (20) представляет собой пример «запертого» на границе поля, хотя в одном отношении пример этот не безупречен: полный магнитный поток, проникший в проводящую среду, $\Phi = \int_0^\infty B dz$, оказывается в данном решении бесконечным. Но, во всяком случае, можно видеть, что стационарный профиль во временной задаче – нечто сугубо отличное от линейной теории скин-эффекта. В следующем разделе мы продемонстрируем более корректный пример такого рода, а сейчас обратимся непосредственно к более интересному случаю.

Предположим, что теперь \mathbf{B} , ∇n и \mathbf{e}_z образуют правую тройку. В сравнении с рис. 1, один из векторов, \mathbf{B} или ∇n , должен быть ориентирован противоположным образом – тогда $\kappa > 0$. Теперь уже конвективный член не соревнуется с диффузионным, но перехватывает на себя весь транспорт поля вглубь плазмы, который происходит даже и в случае иде-

альной проводимости $\sigma \rightarrow \infty$ ($D \rightarrow 0$). Мы будем искать решение уравнения (15) в виде бегущей волны, как обычно поступают, например, в теории УВ:

$$B \equiv B(\xi), \quad \xi = z - ut. \quad (21)$$

Это дает возможность редуцировать уравнение (15) к следующему виду:

$$-uB'_\xi + (\kappa/2)(B^2)'_\xi = DB''_{\xi\xi}, \quad (22)$$

для которого, в свою очередь, мгновенно находится первый интеграл:

$$\text{const} - uB + (\kappa/2)B^2 = DB'_\xi. \quad (23)$$

Пусть профиль тока уже достаточно сгладился у левой границы $z = 0$: $B'_\xi \cong 0$ при $B = B_0$, тогда

$$u = \kappa B_0 / 2. \quad (24)$$

Для получения окончательного ответа перепишем (23) в виде

$$u(b^2 - b) = D \frac{db}{d\xi}, \quad \xi \rightarrow \infty \Rightarrow b \rightarrow 0, \quad b'_\xi \rightarrow 0, \quad (25)$$

где $b = B / B_0$. Уравнение (25) без труда интегрируется:

$$B = \left(\frac{B_0}{2} \right) \left\{ 1 - \text{th} \left[\frac{\kappa B_0}{4D} (z - ut) \right] \right\}. \quad (26)$$

Это, быть может, наиболее важный результат, касающийся конвективного транспорта поля в модели ЭМГ (Кингсеп, Мохов, Чукбар, 1984). Мы будем называть его ЭМГ-волной. Она похожа на УВ, но не является таковой, поскольку в ней не происходит движения вещества. Но она также распространяется с постоянной скоростью $u \sim V_A \frac{c}{\alpha \omega_{pi}} \gg V_A$, где

$\alpha \sim \left(\frac{\partial}{\partial x} \right)^{-1}$; ее профиль остается постоянным во времени и лишь движется, как показано на рис. 2; при этом ее передний фронт формируется как следствие конкуренции между нелинейностью и диффузией. Образование ЭМГ-волны можно проследить аналитически с помощью подстановки (16). В сравнении со стационарным профилем (20) этот результат требует более мягкого условия $t \gg D / (\kappa B_0)^2$, т. е. длина распространения должна быть много больше ширины фронта волны.

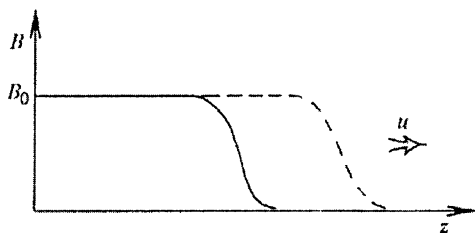


Рис. 2. ЭМГ-волна проникновения поля / тока

Если градиент плотности достаточно мал, точнее, $ut < \sqrt{Dt}$, где τ – характерный масштаб времени в задаче, ЭМГ-волна не может сформироваться таким образом. Но в геометрии более общего вида, когда силовые линии уже не прямые, уравнение эволюции поля опять становится существенно нелинейным

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{c}{4\pi ne} \text{rot}(\mathbf{B}\nabla)\mathbf{B} = \frac{c^2}{4\pi\sigma} \nabla^2 \mathbf{B}, \quad (27)$$

даже и в однородной плазме. Скажем, для случая цилиндрической симметрии с $\partial/\partial\varphi \equiv 0$ опять получается уравнение Бюргера (15).

Экспериментальное исследование ЭМГ-волны

Уже вскоре после предсказания ЭМГ-волн в [3] последовал ряд экспериментальных результатов, подтверждающих эффект проникновения поля в проводящую среду в виде сильно нелинейной волны [4–7]. Мы приведем в качестве примера результаты, полученные в работе Долгачева и др. [8], где использовался сильноточный генератор электронного пучка пулеметного типа, работавший в режиме одиночных выстрелов.

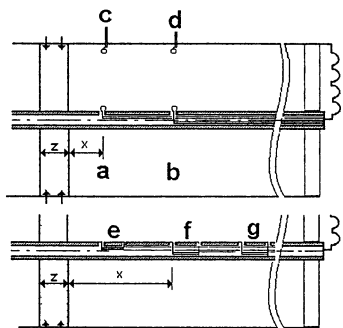


Рис. 3. Схема расположения датчиков в области ППТ (слева) – нагрузка (справа): a, b, c, d – петли для измерения dB/dt , размещенные на центральном (катод) и наружном электродах на расстоянии 15 и 30 см от плазменных пушек; e, f, g – шунты, врезанные в центральный электрод на расстоянии 15, 30, 45 см от плоскости расположения пушек

Выходной узел ускорителя был устроен на основе плазменного прерывателя тока (ППТ – рис. 3). Параметры медленного каскада: емкость батареи генератора Маркса – 0,4 мкФ, амплитуда напряжения, тока и по-

лупериод выходного импульса – соответственно 170 кВ, 120 кА и 1,2 мкс. Диаметры электродов ППТ – соответственно 14 и 160 мм. Для заполнения зазора ППТ использовался набор из 24 плазменных пушек, расположенных по окружности внешнего электрода вакуумной коаксиальной линии. На выходе ППТ получался импульс длительностью 100–150 нс с амплитудой напряжения до 600 кВ. Мы не будем рассматривать подробно сценарий работы ППТ, ограничимся лишь фазой проводимости, в течение которой проводились измерения поля (тока), проникающего в разреженную плазму, заполняющую область коаксиала между зазором ППТ и нагрузкой. Электротехнические датчики – шунты и магнитные зонды – показаны на рис. 3.

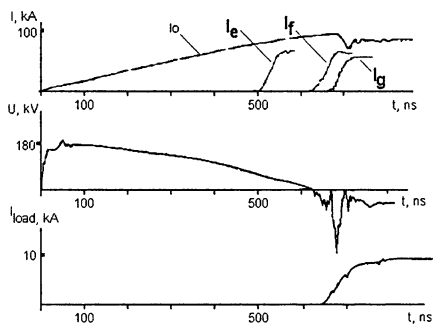


Рис. 4. Осциллограммы тока генератора (I_0), выходного напряжения ППТ (U) и тока на индуктивную нагрузку (I_{load}). Прохождение ЭМГ-волны по разреженной плазме, заполняющей коаксиальную линию, отображается сигналами шунтов I_e , I_f , I_g .

ЭМГ-волна была зарегистрирована в области ППТ – нагрузка, причем по мере распространения в сторону нагрузки она ускорялась и слегка уширялась, как можно видеть из рис. 4 (см. сигналы с шунтов I_e , I_f и I_g). Эффект ускорения можно объяснить на основании выражения (24) для

скорости ЭМГ-волны; в самом деле, из нее следует $u \sim V_A \frac{c}{\alpha \omega_{pi}} \propto n^{-1}$, так

что при движении фронта в направлении падающей плотности плазмы скорость и должна возрастать. Эффект уширения фронта не столь очевиден: на первый взгляд, из (26) следует, что в случае спитцеровской проводимости ширина фронта $\Delta \propto n$, т. е. следует ожидать, напротив, его укручения. Примем, однако, во внимание, что неравенства (3) определяют не только применимость ЭМГ, но и условия для эффективного возбуждения ионно-звуковой неустойчивости тока и, как следствие, возможный режим аномальной сопротивляемости. А посему воспользуемся вместо проводимости Спитцера формулами Сагдеева (см., например, [9]):

$$\sigma \cong \omega_{pe} \frac{neV_{Te}}{j}; \quad j \cong nec_e \sqrt{\frac{MT_i}{mT_e}} \left(\frac{E^2}{8\pi nT_e} \right)^{1/4}. \quad (28)$$

Подставим эту зависимость в выражение для Δ и примем во внимание, что плотность тока на фронте волны обратно пропорциональна ширине

фронта, $j \propto I/\Delta$, где I – величина полного тока в плазме. В случае аксиальной симметрии вместо характерного масштаба плоской задачи $a \equiv |\nabla \ln n|^{-1}$, нужно использовать другой: $\delta \equiv |\nabla \ln(nr^2)|^{-1}$. Получаем

$$\Delta \approx \frac{ne\delta c^2}{\sigma} \propto \frac{nI}{n^{3/2}\Delta} \Rightarrow \Delta \propto n^{-1/4}. \quad (29)$$

Тем самым получает объяснение как сам эффект уширения, так и его более слабая по сравнению с ускорением зависимость от плотности плазменного облака.

Разумеется, ЭМГ-волна конвективного переноса магнитного поля многократно идентифицировалась и в результате численного моделирования. Экспериментальные доказательства существования такого эффекта представляются намного более ценными – в особенности это важно именно в нелинейной физике.

Модификация классической скин-волновой задачи

Представляет определенный интерес – хотя бы с точки зрения принципа соответствия – рассмотрение с учетом эффектов ЭМГ классической задачи о скин-эффекте, хотя она и является несколько искусственной для тех параметров проводящей среды, при которых работает ЭМГ. Эта задача была решена в работе [10].

Граничное условие (18) наиболее типично в условиях, характерных для импульсных плазменных систем, где обычно и работает лучше всего приближение ЭМГ. Обратимся к хрестоматийному граничному условию

$$B_{(z=0)} = B_0 \cos \omega t \quad (30)$$

или

$$B_{(z=0)} = B_0 \theta(t) \cos \omega t. \quad (31)$$

Эта задача важна еще и потому, что нелинейные эффекты, обусловленные ∇n или ∇T (эффект Нернста–Эттингаузена), могут реально вмешаться в электродинамику даже и при вполне классической постановке, хотя бы на уровне поправок. Пусть все приближения останутся теми же, что и в уравнениях (15) и (13), и геометрия задачи пусть будет той же, что на рис. 1. Нелинейность конвективного переноса поля удобно характеризовать безразмерным параметром ξ , который мы определим как отношение нелинейного пространственного масштаба $\delta_{ef} = D/\kappa B_0$ к линейной глубине скин-слоя $\delta_{sk} = \sqrt{2D/\omega}$:

$$\xi \equiv \sqrt{2} \frac{\delta_{ef}}{\delta_{sk}} = \frac{\sqrt{v_{ef} \omega}}{V_{Ae} |\partial(\ln n)/\partial x|},$$

где v_{ef} — эффективная частота столкновений, а V_{Ae} — электронная альфвеновская скорость.

Очевидно, что линейная теория скин-эффекта, базирующаяся на уравнении диффузии поля, инвариантна относительно перемены знака поля. Но уравнению Бюргерса, а следовательно задаче о нелинейном конвективном скин-эффекте, такая инвариантность уже несвойственна. (В этом и заключена причина, по которой так важна топология тройки векторов $\nabla n, \mathbf{B}, \mathbf{e}_z$). В плоской геометрии, в пределах применимости уравнения Бюргерса (15), поле определенной ориентации (скажем, $B > 0$) эффективно проникает в плазму (см., например, (26)), тогда как поле противоположной ориентации (соответственно, $B < 0$) обнаруживает тенденцию к локализации вблизи поверхности плазмы — в духе решения (20). Значит, есть основания ожидать, что у этой задачи стационарного решения не существует. Импульсы положительной полярности должны периодически уходить вглубь плазмы, и таким образом поддерживаются постоянные потоки энергии и магнитного поля в плазменный объем. Поэтому при квазистационарном граничном условии (30) нужно дополнить периодическое решение ненулевым, положительным и постоянным во времени полем при $z \rightarrow \infty$. Это вполне подтверждаются точным решением.

Усредняя (15) по периоду ($2\pi/\omega$), дважды интегрируя по z , получим

$$B_{(z=\infty)} = \frac{K}{2D} \int_0^{\infty} (\langle B^2 \rangle - B^2_{(z=\infty)}) dz > 0. \quad (32)$$

Положим $B_{(z=\infty)} = \alpha B_0$, $0 < \alpha < 1$. Будем искать функцию $\varphi(z)$ — см. уравнение (16)). Заметим, что периодическое поведение во времени величины B не означает того же для φ . Решая диффузионное уравнение для φ при $z \rightarrow \infty$, нетрудно получить $\varphi \propto \exp(\alpha^2 \omega t / 4\xi^2)$. Для произвольного z естественно искать решение в виде

$$\varphi = \tilde{\varphi} \exp(\alpha^2 \omega t / 4\xi^2), \quad (33)$$

где $\tilde{\varphi}$ предполагается периодичной во времени. Эта функция может быть получена в форме фурье-разложения. Ограниченность B накладывает условие положительности $\tilde{\varphi}$, а отсюда, в свою очередь, следует закон спадаания фурье-гармоник с номером. В итоге имеем

$$\tilde{\varphi} = \sum_{n=-\infty}^{+\infty} \tilde{\varphi}_n \exp\left[-(A_n / 4\xi)(\omega / D)^{1/2} z - in\omega t\right], \quad (34)$$

$$A_n = 2\sqrt{\alpha^2 - 4\xi^2 ni}, \quad \text{Re } A_n > 0, \quad \tilde{\varphi}_{-n} = \tilde{\varphi}_n^*.$$

Используя (34), можно переписать (30) в виде

$$A_n = \psi_{n+1} + \psi_n^{-1}, \quad \psi_n \equiv \tilde{\varphi}_n / \tilde{\varphi}_{n-1} < 1. \quad (35)$$

Принимая во внимание, что $\tilde{\varphi}$ – вещественная функция, определенная независимо от произвольного числового коэффициента, можно положить $\tilde{\varphi}_n = 1$. Из (35) находим

$$\psi_n = 1 / \left(A_n - 1 / \left(A_{n+1} - 1 / \left(A_{n+2} - \dots \right) \right) \right). \quad (36)$$

Теперь можно получить и общий член ряда (34):

$$\tilde{\varphi}_n = \prod_{m=1}^n 1 / \left(A_m - 1 / \left(A_{m+1} - \dots \right) \right). \quad (37)$$

Из (35) и (37) следует важный результат:

$$\alpha = \text{Re} 1 / \left(A_1 - 1 / \left(A_2 - 1 / \left(A_3 - \dots \right) \right) \right). \quad (38)$$

Таким образом, решение и в самом деле, помимо элемента периодичности, обязательно включает ненулевое поле при $z \rightarrow \infty$. Уравнение (38) позволяет найти функцию $\alpha(\xi)$, а в конечном счете – и точное решение

$$B = - \left(\frac{2D}{\kappa} \right) \frac{\partial}{\partial z} \ln \left\{ \exp \left(- \frac{\alpha \kappa B_0 z}{2D} \right) + 2 \text{Re} \sum_{n=-\infty}^{+\infty} \tilde{\varphi}_n \exp \left[- \left(\frac{A_n}{4\xi} \right) \sqrt{\frac{\omega}{D}} z - i n \omega t \right] \right\}, \quad (39)$$

согласованное с формулами (37), (38). В общем случае при произвольной величине ξ результат может быть представлен только рядом Фурье и цепными дробями. Но он оказывается весьма простым и наглядным в обоих предельных случаях: сильной и слабой нелинейности. Если нелинейность слаба, т. е. $\xi \gg 1$, то, как можно вычислить, $\alpha = 2^{-5/2} \xi^{-1}$, так что

$$B = B_0 \exp(-z / \delta_{sk}) \cos(z / \delta_{sk} - \omega t) \Leftrightarrow z \ll \delta_{sk} \ln \xi; \quad (40)$$

$$B = B_0 / \left(2^{5/2} \xi \right) \Leftrightarrow z \gg \delta_{sk} \ln \xi.$$

Итак, в этом предельном случае, т. е. в универсальной по сути ситуации линейного скин-эффекта с малой конвективной поправкой, классическое решение дополняется расширяющимся вглубь плазмы плато малой амплитуды и постоянного знака. И конечно, он может реализоваться не только в плазме, но и в металлах, только природа нелинейности там будет, скорее всего, иная – не ∇_n , а ∇T .

Столь же наглядно решение в случае сильной нелинейности при $\xi \gg 1$. Здесь удобнее не искать предельный переход в точном решении, а исходить непосредственно из уравнения (15). Начнем с рассмотрения импульса отрицательной полярности. Характерное время формирования стационарного профиля вида (20) (определенное не из полного потока, а

из наклона кривой) по порядку величины равно $\tau \approx \xi^2 / \omega \ll \omega^{-1}$. Это означает, что проникающее отрицательное поле «следит» за граничным условием, образуя вблизи поверхности осциллирующий профиль

$$B \equiv B_0 \cos \omega t / \left[1 + (\kappa z / 2D) B_0 \cos \omega t \right]$$

с характерным масштабом $\delta_{ef} = \delta_{sk} \xi$ и глубиной проникновения $\delta_{ef} \approx \delta_{sk}$.

Далее мы повторяем вывод уравнения (32), но интегрируя по $z > \delta_{sk}$. В этой области вклад могут давать только положительные импульсы; здесь, к тому же, можно пренебречь диффузионным членом в (15). Получаем $B_{(z=\infty)}^2 \equiv \langle B^2 \theta(B) \rangle_{(z=0)}$, а окончательный результат при граничном условии (30) оказывается следующим:

$$B_{(z=\infty)} = B_0 / 2. \quad (41)$$

Таким образом, вместо скин-эффекта, мы для случая сильной нелинейности получили другой, не менее интересный: детектирование (выпрямление) сигнала магнитного поля плоским плазменным слоем.

Примерно так же можно рассмотреть и более реалистичную задачу с граничным условием (31). Для краткости ограничимся формулировкой результатов, достаточно очевидных, если опираться на рассуждения и точные решения, приведенные выше. В пределе сильной нелинейности $\xi \ll 1$ решение совпадает с квазистационарным при z меньших некоторого $z_0(t)$. При $z > z_0(t)$ имеем $B(z, t) = 0$. Величину $z_0(t)$ можно оценить, взяв интеграл от всего уравнения (15) по z и по t :

$$z_0 \approx 2^{-5/2} \xi^{-1} \kappa B_0 t. \quad (42)$$

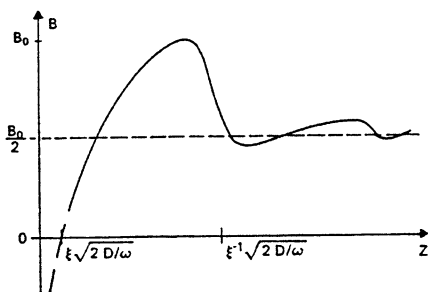


Рис. 5. Детектирование магнитного сигнала в случае сильной нелинейности

Уравнение (42) дает оценку ширины расширяющегося плато, амплитуда которого близка к $B_0 / 2$. Более детальный вид уходящего в глубину плазмы сигнала положительной полярности представлен на рис. 5.

Полезно заметить, что хотя основные уравнения и приведены у нас к одномерной форме, в действительности задачи электронной магнитной гидродинамики всегда, как минимум, двумерны. Конвекция поля в направлении оси z обусловлена не столько током j_z (которого может вообще не быть), сколько токами j_x и j_r , нормальными к e_z . Можно видеть, что в реальной геометрии, когда плазма так или иначе ограничена в направлении, перпендикулярном e_z , токовый перенос магнитной энергии должен происходить вдоль границ проводника из области, где магнитное поле заперто, в область, где имеет место ускоренное проникновение. Нетрудно убедиться и в том, что при условии большого параметра Холла вектор Пойнтинга лежит в плоскости фронта нелинейной волны.

Отметим особо, что в рамках ЭМГ (исключая тривиальный вырожденный случай) классический скин-эффект вообще невозможен. Поле в проводящей среде уходит на бесконечность, имеет при этом определенный знак, распространяется в виде нелинейной волны, амплитуда которой может достигать половины амплитуды осциллирующего поля на границе.

Хотя ЭМГ-волны ранее не были известны, известны их в некотором роде линейные аналоги, причем как в плазме, так и в твердом теле. Это дрейфовые волны и геликоны (вистлеры, свистящие атмосферерики), которые могут распространяться на фоне достаточно сильного постоянного поля. Предельный переход от нашего случая к этим хорошо известным линейным волнам не вполне тривиален: конечная величина поля B одновременно играет роль амплитуды линейной волны и внешнего поля, в котором эта волна распространяется.

Литература

1. Кингсен А. С., Чукбар К. В., Яньков В. В., Электронная магнитная гидродинамика // Вопросы теории плазмы / Под ред. Б. Б. Кадомцева Вып 16 М : Атомиздат, 1987. С. 243.
2. Gordeev A. V., Kingsep A. S., Rudakov L. I. // Phys. Reports, 1994. V. 243. P. 216.
3. Кингсен А. С., Мохов Ю. В., Чукбар К. В. // Физика плазмы, 1984 Т. 10. С. 854.
4. Weber B. V., Colombath D. G., Commisso R. J. e. a. // Proc. VI Int. Conf. BEAMS'86, Kobe, 1986. P. 882; see also: Weber B. V., Commisso R. J., Meger R. A. e. a. // Appl. Phys. Letters, 1984. V. 45. P. 1043.
5. Maron Y., Sarid E., and Zahavi O. // Phys. Rev. A. 1989. V. 39. P. 5842.
6. Bystritskii V. M., Krasik Ya. E., Lisitsin I. V. e. a. // IEEE Trans. Plasma Sci. 1991. V. 19. P. 607.
7. Commisso R. J., Goodrich P. J., Grossmann J. M. e. a. // Phys. Fluids B. 1992, V. 4(7). P. 2368.
8. Долгачев Г. И., Кингсен А. С., Ушаков А. Г. // Физика плазмы 2001. Т. 27. С. 64.
9. Кингсен А. С. // Физика плазмы. 1991. Т. 17. С. 582.
10. Калда Я. Л., Кингсен А. С. // Физика плазмы. 1988 Т. 14 С. 1519.

Нелинейные явления в медицине и биологии

НЕЛИНЕЙНЫЕ КОЛЕБАНИЯ НА МОЛЕКУЛЯРНОМ И КЛЕТОЧНОМ УРОВНЯХ В БИОЛОГИЧЕСКИХ СИСТЕМАХ: ВОЗМОЖНЫЕ ИСТОЧНИКИ ПРОИСХОЖДЕНИЯ И ФУНКЦИОНАЛЬНАЯ РОЛЬ В ПРОЦЕССАХ ЖИЗНЕДЕЯТЕЛЬНОСТИ

В. Л. Воейков

Введение

До конца прошлого века изучение колебательных процессов не относилось к приоритетным направлениям исследований в молекулярной и клеточной биологии. Традиционно эти разделы биологии базируются на представлениях классической термодинамики и органической химии, энзимологии и фармакологии. Так, считается, что скорость биохимических процессов определяется концентрациями участвующих реагентов (закон действующих масс) и соотношением констант элементарных реакций. Чрезвычайно высокая скорость протекания реакций в клетке объясняется каталитической активностью соответствующих ферментов, а высочайшая избирательность этих процессов, практически полное отсутствие побочных реакций – молекулярной комплементарностью участников химических превращений (принцип «ключ-замок»). Протекание в клетке энергопотребляющих процессов объясняют тем, что они сопряжены с термодинамически выгодными реакциями, в ходе которых энергия освобождается. А увеличение степени упорядоченности продуктов при синтезе специфических биополимеров из мономеров, явное понижение энтропии в ходе развития живых организмов объясняют тем, что как структура полимеров, так и осуществляемые с их участием процессы жизнедеятельности, включая весь ход биологического развития, закодированы в последовательностях эволюционно отобранных генов [1].

Благодаря развитию новых методов исследования, позволяющих наблюдать в реальном времени и с высокой степенью разрешения за ходом протекания внутриклеточных процессов, установлено множество фактов, которые не согласуются со сложившимися в классической биохимии взглядами, и в самые последние годы положение стало существенно меняться. В клеточную и молекулярную биологию стали проникать представления неравновесной термодинамики, синергетики, теории диссипативных динамических систем, теории нелинейных колебаний. Хотя экспериментальные исследования колебательных режимов в биохимических реакциях, протекающих как в живых клетках, так и *in vitro*, стали развиваться только недавно и затрагивают пока весьма ограниченный круг ре-

акций, уже ясно, что колебательные режимы играют фундаментальную роль в нормальных процессах жизнедеятельности.

Ниже мы кратко опишем ряд наиболее ярких и широко распространенных биохимических процессов, для которых характерны нелинейные колебания во времени и в пространстве живых систем, и рассмотрим роль колебательных режимов в регуляции процессов жизнедеятельности. Далее мы обсудим вопрос о наиболее общих эндогенных причинах возникновения осцилляций в ходе биохимических процессов. Мы полагаем, что в основе биоритмов лежат окислительные реакции с участием активных форм кислорода, протекающие в водных системах («дыхание»), в ходе которых в условиях, характерных для существования жизни, могут спонтанно возникать осцилляции, которые служат ритмоводителями процессов более высоких уровней организации.

1. Кальциевые осцилляции и осцилляции в ходе окислительно-восстановительных процессов в клетках

То, что в нервных и мышечных клетках в ответ на факторы внешней среды или по внутренним причинам могут быть индуцированы периодические процессы (нервный импульс, мышечное сокращение), известно давно. Менее хорошо известно, что и в невозбудимых клетках происходят колебания мембранного потенциала, синтеза РНК и белков, что осциллируют различные показатели при гликолизе, фотосинтезе, митохондриальном дыхании [2]. Особое внимание в последнее время привлекают кальциевые осцилляции. Известно, что в ответ на взаимодействие многих гормонов и нейромедиаторов с находящимися на клеточной поверхности рецепторами, ионы Ca^{2+} освобождаются из особых содержащих их пузырьков в цитоплазму. Там они активируют Ca^{2+} -связывающие белки и ферменты, запуская каскад реакций, приводящих к характерному ответу клетки на гормон. Долго считалось, что Ca^{2+} действует по принципу «доза-эффект», т. е. интенсивность ответа монотонно растет с увеличением содержания Ca^{2+} в цитоплазме. Появление методов, позволяющих осуществлять непрерывную регистрацию внутриклеточной активности Ca^{2+} , позволило показать, что кальциевый ответ носит колебательный характер, причем не только во времени, но и в пространстве как индивидуальной клетки, так и многоклеточных комплексов [3]. Кальциевые волны представляют собой распространение кальциевых осцилляций по клеточному пластику после стимуляции хотя бы одной клетки, входящей в его состав. Нередко в зависимости от концентрации (дозы) внешнего регуляторного сигнала меняется частота кальциевых спайков при их постоянной амплитуде. На ряде примеров установлено, что в зависимости от частоты осцилляций меняется спектр работающих в данный момент в клетке генов, т. е. режим кальциевых осцилляций обеспечивает тонкую настройку рабо-

ты генома [4]. Изменение частоты кальциевых осцилляций в ответ на изменение концентрации (амплитуды) внешнего регуляторного сигнала говорит о возможности аналого-цифрового преобразования, происходящего на клеточном уровне. Такой характер физиологической реакции может существенно повысить точность восприятия клеткой сигналов из окружающей среды за счет снижения шумовой компоненты [5]. Кроме того, при частотном режиме чувствительность регулируемых кальцием систем к изменениям его среднестатистической концентрации во много раз выше, чем при его отсутствии [6], т. е. частотный режим регуляции клеточной активности намного экономичнее, чем амплитудный.

Другой недавно открытый колебательный процесс на клеточном уровне – это осцилляции в клетках уровня важного метаболита – NADPH/NADP (никотинамид адениндинуклеотид фосфат восстановленный/окисленный). Обычно это соединение рассматривают только как кофактор многих ферментов, участвующих в реакциях окисления-восстановления, и обеспечивающих организм энергией в реакциях гликолиза и окислительного фосфорилирования. Однако оно может использоваться и как донор электронов для непосредственного одноэлектронного восстановления кислорода, например, в реакции:



Продукт этой реакции – супероксид анион радикал ($\text{O}_2^{\bullet -}$) – представитель семейства так называемых активных форм кислорода (АФК). В модельных реакциях *in vitro* эти соединения обладают очень высокой и мало контролируемой химической активностью. Поэтому АФК считаются исключительно опасными соединениями, вызывающими разнообразные повреждения клеток (более подробно о свойствах АФК см. ниже). Оказалось, однако, что в организме есть ферменты, в частности NADPH-оксидазы, которые катализируют перенос электронов с этого субстрата прямо на кислород. Вначале такие ферменты были обнаружены у клеток иммунной системы – нейтрофилов, играющих роль первой линии обороны против проникших в организм бактерий и вирусов. Поэтому продуцирование АФК клетками иммунной системы считали экзотикой – платой за необходимость защищаться от еще более опасных патогенов. Сейчас стало ясно, что NADPH-оксидазы и аналогичные ферменты присутствуют практически во всех клетках, и не имеющих отношения к иммунной системе [7]. Генерация АФК происходит непрерывно и находится под контролем различных регуляторных факторов.

Совсем недавно было убедительно показано, что продуцирование АФК нейтрофилами и другими клетками осуществляется в осцилляторном режиме, причем периоды колебаний, лежащие в диапазоне от десятков секунд до минут, определяются функциональной активностью клеток [8] Осцилляции продукции АФК тесно связаны с осцилляциями внутри-

клеточной концентрации NADPH. Гормоны, регулирующие продуцирование АФК нейтрофилами, действуют на параметры осцилляций: одни увеличивают их амплитуду, другие – частоту, например, удваивая ее, причем, что наиболее важно, вторые не оказывают своего действия, если амплитуда осцилляций не превышает некоторого порога. При воспалительных состояниях у нейтрофилов нарушается регулярный режим колебаний – из строго периодического он становится хаотичным. Нормальный колебательный режим можно восстановить, например, сначала охладив, а затем повысив температуру среды клеток до физиологической (возможно, этим объясняется польза холодных компрессов при воспалениях). [9]. Содержание NADPH непрерывно осциллирует в клетках самого разного типа, причем частота и амплитуда осцилляций коррелируют с клеточной активностью. Так, у клеток почки в культуре частота осцилляций в покое составляет 0,005 Гц, а при миграции – 0,05 Гц с более чем двукратным ростом амплитуды [10].

2. Клеточное дыхание

В приведенных выше примерах речь шла об осцилляциях Ca^{2+} и NADPH. Есть ли между ними какая-нибудь связь? Осцилляции NADPH непосредственно связаны с образованием АФК, т. е. с прямым восстановлением O_2 электронами, которые отдает ему NADPH. Следует напомнить, что живые организмы используют кислород для окисления пищевых продуктов, извлекая при этом энергию, обеспечивающую их жизнедеятельность. В биохимии рассматривается только один путь кислородного дыхания – окислительное фосфорилирование в митохондриях. Этот процесс начинается с аналога NADPH – NADH (никотинамид адениндинуклеотида), который при окислении отдает электроны в электрон-транспортную цепь митохондрий. Энергия разности редокс-потенциалов между NADH и O_2 разбивается при этом на дробные порции, позволяя на пути движения электронов от одного фермента к другому запасать ее в молекулах АТФ. При расщеплении молекулы АТФ освобождается энергия, эквивалентная 0,4–0,5 эВ. Кислород при митохондриальном дыхании выступает как конечный акцептор потерявших свою энергию электронов. В биохимии и клеточной биологии, собственно, только этот процесс и считается дыханием.

Но если электрон(ы) с NADPH или NADH напрямую сбрасываются на кислород, то на этом более коротком пути полного одноэлектронного восстановления молекулы кислорода до воды четырьмя электронами не только образуются АФК, но и освобождаются гораздо более “тяжелые” кванты энергии – порядка 1–2 эВ. Вклад этих процессов в биоэнергетику вообще не рассматривается, хотя на них расходуется до трети, а иногда и более потребляемого организмом кислорода [11]. Есть данные и о том,

что энергия, освобождающаяся при реакциях, протекающих с участием АФК, не диссипирует полностью и немедленно, а используется клеткой для реализации многих функций. Поэтому, если считать «дыханием» любые окислительно-восстановительные процессы, в которых используется кислород и энергия которых идет на поддержание функциональной активности клетки, то осцилляции NADPH в отдельных клетках отражают колебательный режим их дыхания. В покое дыхание клетки редкое и «поверхностное»; при действии на нее одних биорегуляторов оно становится более «глубоким» (рост амплитуды), а если в этом состоянии на клетку действуют агенты второго типа, ее дыхание учащается.

Прямых данных о связи генерации АФК с кальциевыми осцилляциями пока мало. Но известно, что дыхание клетки, измеренное по суммарному потреблению кислорода, кальциевые колебания и функциональная активность клетки тесно связаны. Такие данные получены, например, на изолированных бета-клетках поджелудочной железы, которые в ответ на добавление глюкозы секретируют инсулин. При этом в ответ на повышение в среде бета-клеток содержания глюкозы не просто увеличивается скорость секреции ими инсулина, а его выброс осуществляется в колебательном режиме. При диабете часто нарушается не столько общее содержание инсулина в крови, сколько нормальный ритм его поступления в кровь [12]. Между стимулирующим фактором – глюкозой – и повышением секреторной активности бета-клеток лежит сложная цепь биохимических процессов, и одно из ранних событий в этой цепи – это возникновение осцилляций внутриклеточного кальция. Недавно было установлено, что самая первая реакция бета-клетки в ответ на добавление в среду глюкозы – это возникновение выраженных колебаний в потреблении ею кислорода [13]. Глюкоза стимулирует «углубление» дыхания клетки даже если осцилляции Ca^{2+} искусственно подавлены (в последнем случае клетка в ответ на глюкозу не секретирует инсулин). Если же осцилляции Ca^{2+} не заблокированы, они развиваются точно в фазе с осцилляциями дыхания (период осцилляций составляет $3,1 \pm 0,1$ мин), причем продолжаться без угасания они могут в отдельной клетке в течение нескольких часов.

Авторы работы [13] рассматривали «углубление» клеточного дыхания только с позиции митохондриального потребления кислорода, тогда как становится очевидным, что при любом потреблении кислорода клетками возрастает и скорость его одноэлектронного восстановления с образованием АФК. Именно этот процесс, по нашему мнению, автоматически порождает колебательные режимы сначала на молекулярном, а затем и на клеточном уровне, поскольку, как будет показано ниже, окислительные реакции с участием O_2 , протекающие даже в совершенно неорганизованных, казалось бы, системах, например водных растворах простых органических соединений, имеют тенденцию к самоорганизации во времени и

пространстве, выражающейся, в частности, в возникновении колебательных режимов потребления кислорода. Необходимым условием для возникновения колебательных режимов в таких системах является образование АФК, и ниже мы более подробно остановимся на свойствах химических реакций с их участием.

3. Энергетика процессов с участием активных форм кислорода

АФК – это продукты одноэлектронного восстановления молекулы кислорода. К ним относится супероксид-анион радикал ($O_2^{\bullet-}$), гидропероксильные радикалы ($HO\bullet$ и $ROO\bullet$, где R – остаток органической молекулы). Когда эти частицы получают еще один электрон (вместе с H^+), они превращаются в перекиси (H_2O_2 или $ROOH$). Перекиси – это не свободные радикалы, а молекулы, но они легко восстанавливаются в присутствии доноров электронов, превращаясь в чрезвычайно энергичные радикалы – гидроксил- или алкил(алкоксил)-радикалы ($HO\bullet$ или $RO\bullet$). Последние отнимают электрон (вместе с водородом) у любого подходящего донора, превращаясь в воду или спирт (ROH). Само существование АФК обусловлено уникальными свойствами молекулярного кислорода – только он в своем основном состоянии триплетен, т. е. является бирадикалом – частицей с двумя неспаренными электронами [$M\uparrow\uparrow$], а значит, обладает избыточным запасом энергии по сравнению с молекулами в невозбужденном синглетном состоянии [$M\uparrow\downarrow$], в котором все их электроны спарены. Но такого состояния у кислорода нет – его синглетное по энергии еще выше, чем триплетное. Таким образом, как триплетное, так и синглетное состояния кислорода – это возбужденные, богатые энергией состояния.

До последнего времени АФК рассматривались лишь с точки зрения их химической агрессивности: будучи сильными окислителями, они могут повреждать биологически важные молекулы. Те из них, что являются свободными радикалами, инициируют цепные процессы, при которых повреждения могут лавинообразно нарастать. Однако, как отмечалось выше, на одноэлектронное восстановление кислорода в организме идет существенная доля от общего его потребления. Но возникающие при этом АФК тут же устраняются специализированными ферментами или антиоксидантами, например, аскорбатом или токоферолом. Какой же смысл в достаточно интенсивном производстве АФК, если они немедленно устраняются?

Мы полагаем, что смысл этот заключается в энергетических особенностях реакций с участием АФК. При рекомбинации радикалов, приводящей к их устранению, или при распаде и полном восстановлении перекисей освобождаются порции энергии, эквивалентные энергии квантов видимого и даже УФ-света. Так, при рекомбинации двух гидропероксиль-

ных радикалов ($2 \text{HO}\cdot \rightarrow \text{H}_2\text{O}_2 + {}^1\text{O}_2$) кислород появляется в синглетном состоянии, энергия которого, по меньшей мере, на 1 эВ выше энергии триплетного кислорода. При восстановлении H_2O_2 до воды освобождается еще 2 эВ. Эти реакции катализируют ферменты супероксиддисмутазы и каталазы, которые могут работать с частотой, превышающей 1 МГц. Освобождение квантов энергии со столь высокой частотой затрудняет быстрое рассеяние энергии электронного возбуждения в тепло. Всего при полном восстановлении молекулы O_2 четырьмя электронами освобождается 8 эВ (для сравнения укажем, что энергия УФ-фотона с $\lambda=250$ нм равна 5 эВ). Беспольное рассеяние столь ценной энергии маловероятно еще и потому, что ее генерация происходит в организованной клеточной и внеклеточной среде. В ряде работ, действительно, показано, что энергия, освобождающаяся в реакциях с участием АФК, может использоваться в качестве энергии активации или для модуляции ферментативной активности [14, 15]. Если же эта энергия освобождается ритмично, то реакции с участием АФК могут выступать в роли ритмоводителей зависимых от них процессов.

4. Колебательные режимы реакций с участием АФК

В водных растворах карбонильных соединений (например, глюкозы или метилглиоксаля) и аминокислот протекает amino-карбонильная реакция или так называемая реакция Мейера (РМ). В ходе ее образуются соединения, способные восстанавливать O_2 . При этом в реакционной системе появляются АФК, и их реакции сопровождаются излучением фотонов. Мы регистрировали это излучение с использованием детекторов одиночных фотонов. Оказалось, что если наблюдение вести в течение достаточно длительного времени, при этом реакцию систему не “беспокоить”, то при развитии РМ даже в простейших по исходному химическому составу растворах можно наблюдать возникновение колебаний излучения, не угасающих в течение многих часов и даже дней (рис. 1) [16, 17].

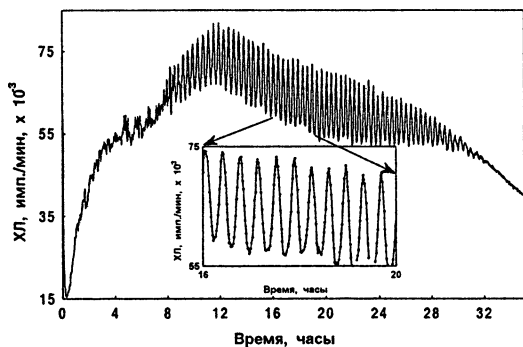


Рис. 1. Кинетика развития излучения в растворе глюкозаглицин (60 мМ, 60 мМ; 20 мл; рН 11,0).

В развитии колебательного режима чрезвычайно важную роль играет кислород. Если реакционную систему продувать азотом и устранить контакт с воздухом, то излучение из нее практически отсутствует, однако как только контакт с воздухом восстанавливается, происходит резкий всплеск интенсивности излучения, причем сразу в колебательном режиме (рис. 2).

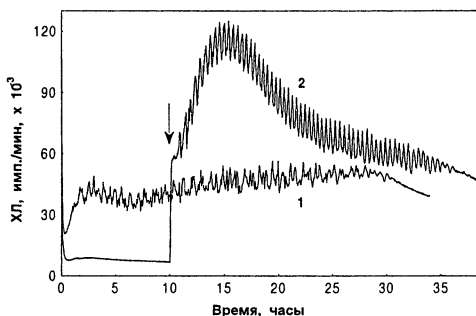


Рис. 2. Излучение в системе глюкоза – глицин (60 мМ, 60 мМ, 20 мл; рН 11,0) в аэробных (1) и анаэробных (2) условиях. Стрелка указывает момент восстановления аэробных условий.

В реакционных системах, в которых в качестве карбонильного компонента использовали весьма активный реагент – метилглиоксаль (МГ) развиваются 2 волны излучения, причем только вторая модулирована колебаниями. При одновременном измерении содержания кислорода и излучения из таких систем выяснилось, что вторая волна и осцилляторный режим излучения появляются, когда растворенного кислорода в реакционной системе не остается (рис. 3).

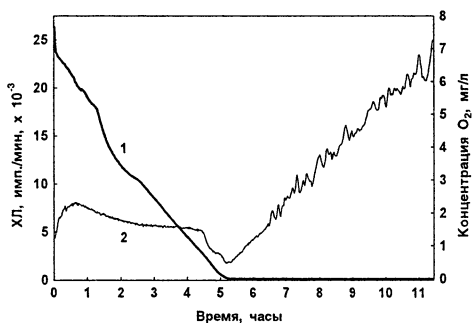


Рис. 3. Одновременное измерение концентрации кислорода в толще раствора (1) и излучения (2) в системе МГ – этаноламин (10 мМ, 30 мМ, 20 мл; рН 10,3)

Таким образом, развитие колебательного режима происходит в условиях, когда кислород поступает в раствор из газовой фазы, а его восстановители – из толщи раствора. Поэтому основная доля эмиссии приходится на самые верхние слои, находящиеся в контакте с воздухом. Исходя из этого, можно было бы предположить, что с увеличением соприкасающейся с воздухом площади поверхности раствора интенсивность излучения будет расти. В действительности оказалось, что с увеличением площади поверхности она уменьшается (рис. 4), причем при постоянном объ-

еме с уменьшением площади свободной поверхности максимальная и интегральная интенсивность излучения второй волны, а также вероятность развития колебательного режима возрастают по закону, близкому к гиперболическому: $\Sigma I = k/S^{2,3}$, где ΣI – интегральная интенсивность излучения в процентах от минимальной, S – площадь поверхности в процентах от минимальной. Следовательно, чем сильнее затруднена диффузия кислорода в раствор, тем интенсивнее идут в нем сопровождаемые излучением реакции.

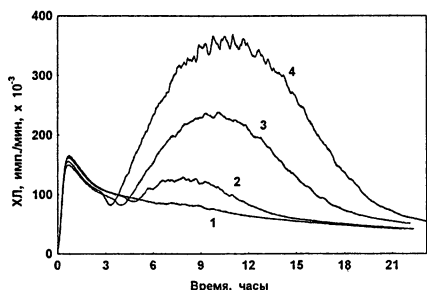


Рис. 4. Зависимость кинетики излучения из системы МГ – этаноламин (10 мМ, 30 мМ, 6 мл; рН 10,3 (карб Б 50 мМ)) от площади свободной поверхности:

- 1 – 420 мм²;
- 2 – 380 мм²;
- 3 – 340 мм²;
- 4 – 300 мм².

С другой стороны, при постоянной площади поверхности интенсивность и продолжительность сопровождаемых излучением процессов тем больше, чем больше объем реакционной системы. Таким образом, когда реакционная система контактирует с воздухом, развитие второй волны излучения, модулированной его колебаниями, возможно при прочих равных условиях лишь если параметр P , пропорциональный объему системы и обратно пропорциональный контактирующей с воздухом площади ее поверхности, превышает некоторую критическую величину ($P > V/S$).

Для развития сопровождаемой колебаниями волны излучения характерны и другие критические (предельные) условия – температура, рН, концентрация реагентов, наличие или отсутствие катализаторов окислительно-восстановительных процессов. Например, в субкритических условиях, когда вторая волна развивается практически без колебаний, а интенсивность ее очень низка, внесение в реакционную смесь классического антиоксиданта аскорбата в очень низких концентрациях способствует как увеличению ее амплитуды и продолжительности, так и развитию колебательного режима (рис. 5). С другой стороны, при превышении концентрации аскорбата выше некоторой величины (в данном случае – 1 мМ) он выступает в роли классического антиоксиданта – ингибитора разветвленно-цепных процессов, значительно тормозя развитие второй волны излучения.

Реакции с участием АФК относятся к окислительно-восстановительным, поэтому можно было предположить, что колебания излучения должны коррелировать с колебаниями окислительно-восстановительного

потенциала (ОВП). ОВП измеряли точечным платиновым электродом по отношению к потенциалу хлорсеребряного электрода. В самом начале реакции ОВП характеризуется достаточно высоким значением. В момент прогорания первого пика излучения (когда кислород в растворе исчерпывается), потенциал значительно снижается. При поверхностном расположении электрода (платиновый шарик находится на глубине 3 мм) в дальнейшем с появлением второй волны излучения, модулированной колебаниями, ОВП начинает постепенно возрастать, при этом развиваются выраженные колебания его значения, которые, как видно на рис. 6, имеют те же частотные характеристики, что и колебания излучения.

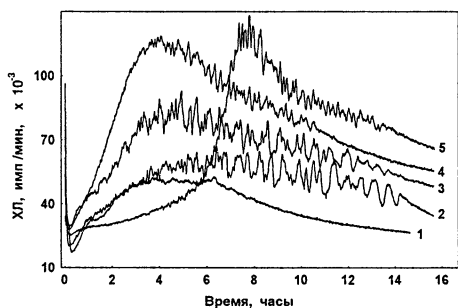


Рис. 5. Влияние аскорбата на излучение из системы глюкоза-глицин (60 мМ, 60 мМ, 10 мл, рН 11,0):

- 1 – контроль,
- 2 – 1,0 мкМ,
- 3 – 10 мкМ,
- 4 – 100 мкМ,
- 5 – 1 мМ аскорбата

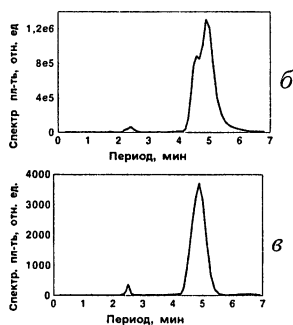
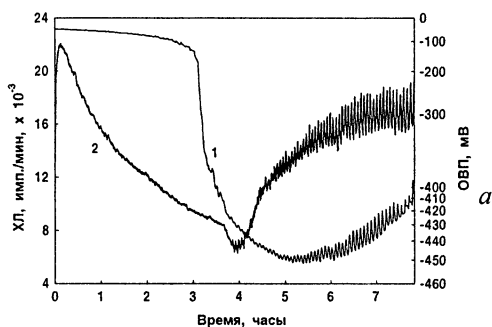


Рис. 6. Одновременное измерение ОВП (1) и излучения (2) в приповерхностном слое в системе метилглиоксаль – этаноламин (10 мМ, 30 мМ, 40 мл, 440 мм²; рН 10,3) (а). Периодограммы, полученные путем фурье-анализа временных рядов ОВП (б) и излучения (в), начиная с 4-го часа протекания реакции

5. Возможные причины развития колебательного режима излучения в ходе amino-карбонильной реакции

Итак, необходимыми условиями для развития колебательного режима излучения и ОВП при amino-карбонильной реакции служат: 1) нали-

чие резкого градиента кислорода между реакционной системой и газовой фазой, с которой раствор контактирует, 2) ограничение скорости поступления кислорода в реакционную систему путем, например, уменьшения площади поверхности или увеличения ее объема и отсутствием перемешивания, 3) создание других условий, способствующих преодолению определенного соотношения между содержанием в растворе легко окисляемых соединений и поступающего туда кислорода (например, внесение в среду доноров электронов (аскорбата), повышение значения pH, повышение концентрации реагентов выше некоторого критического уровня).

При лимитированной скорости диффузии кислорода в раствор в реакционной системе начинают накапливаться восстановительные эквиваленты, что доказывается снижением значения ОВП раствора. Когда концентрация кислорода, продолжающего диффундировать в раствор, превышает некоторую критическую величину, инициируется цепной процесс окисления им восстановленных соединений, при котором концентрация активных центров (АФК) растет в геометрической прогрессии. Ускорение (разгорание) процесса обусловлено, по нашему мнению, тем, что значительная часть восстановленных соединений, накопившихся в анаэробных или гипоксических условиях, представлена метастабильными молекулами перекисного типа. Их расщепление, при котором происходит размножение активных центров, дополнительно активирующих кислород, требует энергии активации. Мы полагаем, что эту роль играет энергия электронного возбуждения, генерируемая в реакциях с участием АФК. Ускорение потребления кислорода за счет появления новых активных центров не компенсируется его диффузией в раствор, в котором снова начинается накопление недоокисленных (восстанавливающих) соединений. Излучение из системы затухает, с тем чтобы позднее развилась его новая волна.

6. Участие электронно-возбужденных состояний в развитии колебательных процессов

Согласно предложенной выше схеме механизма развития колебательного режима излучения и ОВП, важным фактором для его возникновения является разветвление цепей, способствующее резкому ускорению потребления кислорода. При изотермических условиях, в которых протекают подобного рода процессы, энергия активации для расщепления недоокисленных метастабильных продуктов, в частности пероксидов, может поставляться за счет реакций рекомбинации радикалов, при которых образуются продукты в электронно-возбужденном состоянии (ЭВС). Кванты энергии, которые освобождаются при их релаксации, вполне достаточны для разрыва пероксидных связей. То, что генерация ЭВС является существенным фактором в развитии колебательных режимов окислительно-

восстановительных процессов, протекающих в водных средах, следует из представленных выше данных и других полученных нами результатов.

В частности, как можно видеть на многих графиках, сопровождаемая колебаниями волна излучения постепенно спадает, и колебания исчезают, однако даже через много дней система продолжает служить стабильным источником излучения. На этой стадии его интенсивность не меняется даже при сильном перемешивании реакционной смеси. Но оказалось, что развитие новых волн излучения из таких систем может быть индуцировано гораздо более «мягкими» воздействиями. Такое действие оказывает внесение в раствор солей меди или железа в микромолярных концентрациях, облучение систем очень слабыми источниками света (даже при оптическом контакте «угасшей» системы с системой с более высоким уровнем излучения) и, что на первый взгляд особенно удивительно, создание анокисических условий.

Индукцию в «угасшей» системе процессов, сопровождающихся фотонной эмиссией, под действием столь разных факторов можно, по видимому, объяснить с единых позиций. Из рис. 2 видно, что при развитии второй волны излучения среднее значение ОВП реакционного раствора постепенно растет, следовательно, содержание в нем восстановительных эквивалентов уменьшается, снижается способность восстанавливать кислород, а значит, падает и средний уровень АФК, и плотность энергии, обеспечивающей возникновение новых активных центров. При этом образуются все более сложные соединения, которые поглощают и рассеивают энергию электронно-возбужденных состояний и/или обрывают реакционные цепи, превращаясь в достаточно стабильные, а значит, малоактивные радикалы. Их можно рассматривать как сильные антиоксиданты, поскольку они не могут активировать кислород и тем самым тормозят разветвление цепей. Кислород, который продолжает поступать в раствор, захватывается соединениями с двойными связями, что приводит к образованию еще более сложных продуктов. Поскольку энергетический выход такого рода реакций значительно ниже, чем реакций рекомбинаций «легких» кислородных радикалов, средняя интенсивность излучения снижается. Если же в такую систему внести «слабый» антиоксидант, то он может реинициировать разветвление цепей. Этот эффект был давно обнаружен на примере вырожденно-разветвленных цепных реакций окисления полимеров, затухающих задолго до полного окисления субстрата [18]. «Слабый» антиоксидант, являющийся донором электронов, способствует расщеплению метастабильных соединений перекисного типа, что приводит к «размножению» активных центров, увеличению числа реакционных цепей, возобновлению одноэлектронного восстановления O_2 и разгору реакции. Ионы меди и железа катализируют распад перекисей, и, вероятно, именно по этой причине при их внесении в реакционную смесь

происходит разгорание «затухшего» процесса. Разгорание процесса при подсвечивании системы даже таким слабым источником, как другая реакционная система, можно объяснить тем, что многие метастабильные перекисные соединения окрашены и при поглощении фотонов с достаточным уровнем энергии они распадаются, порождая новые реакционные цепи.

Наконец, обнаруженное нами «разгорание» процесса после продувания реакционной смеси инертными газами и изоляции ее от контакта с воздухом также свидетельствует, что роль кислорода в данном процессе двоякая – на ранней стадии он абсолютно необходим для его разгорания, а на поздней выступает в роли тушителя по описанным выше причинам. Разгорание процесса после удаления из системы кислорода объясняется, вероятнее всего, тем, что сохраняющийся уровень эндогенной генерации электронно-возбужденных состояний в отсутствие кислорода оказывается достаточным для распада пероксидных соединений и размножения активных центров. При этом ОВП системы снижается, т.е. растет ее восстановительный потенциал. После возобновления контакта с воздухом система вновь приобретает способность к бурной генерации АФК, сопровождающейся резкой и длительно не затухающей вспышкой излучения, вновь модулированной колебаниями, подобно тому, как это происходит в системе, исходно изолированной от кислорода (см. рис. 2, кривая 2).

Заключение

Важнейшая роль колебательных режимов биохимических реакций в осуществлении процессов жизнедеятельности уже не может подвергаться сомнению. По поводу источника происхождения колебаний высказываются различные предположения. До последнего времени, однако, в этой связи не уделялось достаточного внимания тому факту, что любые живые системы, начиная от клеточного уровня и выше, могут существовать только за счет той энергии, которую они сами генерируют в ходе окислительно-восстановительных реакций. Наиболее эффективным источником энергии является дыхание – потребление кислорода, окисляющего пищевые субстраты. Описанная выше реакционная система является, с нашей точки зрения, простейшей моделью того, как зарождается и осуществляется дыхательный процесс. Окислительно-восстановительные реакции, которые начинаются в исходно бесструктурной среде, имеют тенденцию к самоорганизации во времени и пространстве, проявляющейся, в частности, в возникновении устойчивых колебательных режимов. Как показано выше, условия для их возникновения отнюдь не являются экзотическими, более того, они характерны для любой живой системы, будь то многоклеточный организм, орган или ткань, отдельная клетка. При этом существенно, что речь в данном случае идет об осцилляциях электромагнитной природы – осцилляциях излучения, отражающего освобождение квантов

энергии высокой плотности и ОВП, отражающего токи электрически заряженных частиц. Такие осцилляции – не только свидетельство того, что в подобных системах протекают коллективные динамические процессы, но именно они могут выступать, с одной стороны, в роли наиболее фундаментальных ритмоводителей для вышестоящих уровней биологической организации, а с другой – рецепторами внешних электромагнитных полей, способных при определенных условиях модулировать их характеристики.

Литература

1. Альбертс Б. Брей Д., Льюис Дж и др Молекулярная биология клетки В 3-х т М.: Мир, 1994 517+539+504 с
2. Rapp P. E. An atlas of cellular oscillator // J. Exp Biol. 1979. V. 81. P. 281-306.
3. Berridge M. J., Cobbold P. H., Cuthbertson K. S. R Spatial and temporal aspects of cell signaling // Phil. Trans. R Soc. Lond B. 1988 V. 320. P. 325-343.
4. Li W-H., Llopis J., Whitney M et al Cell-permeant caged InsP3 ester shows that Ca²⁺ spike frequency can optimize gene expression. // Nature. 1998. V. 392. P. 936-941.
5. Trewavas A J., Malho R. Signal perception and transduction: the origin of the phenotype. // Plant Cell. 1997. V. 9. P. 1181-1195.
6. Dupont G., Swillens S., Clair C., et al. Hierarchical organization of calcium signals in hepatocytes: from experiments to models // Biochim. Biophys. Acta. 2000. V. 1498. P. 134-152.
7. Sauer H., Wartenberg M., Hescheler J. Reactive oxygen species as intracellular messengers during cell growth and differentiation // Cell Physiol. Biochem. 2001. V. 11. P. 173-186.
8. Rosenspire A. J., Kindzelskii A L., Petty, H. R. Interferon- and sinusoidal electric fields signal by modulating NAD(P)H oscillations in polarized neutrophils // Biophys. J. 2000. V. 79. P. 3001-3008
9. Kindzelskii, A. L., Zhou, M. J., Haugland, R. P. et al. Oscillatory pericellular proteolysis and oxidant deposition during neutrophil locomotion // Biophys. J. 1998 V. 74. P. 90-97.
10. Rosenspire A J., Kindzelskii A L., Petty H. R. Pulsed DC electric fields couple to natural NAD(P)H oscillations in HT-1080 fibrosarcoma cells // J. Cell Sci. 2001. V. 114. P. 1515-1520.
11. Voeikov V. Reactive Oxygen Species, Water, Photons, and Life // Rivista di Biologia/Biology Forum. 2001. V. 94. P. 193-214
12. O'Rahilly S., Turner R. C., Matthews D R. Impaired pulsatile secretion of insulin in relatives of patients with non-insulin-dependent diabetes // N. Engl J. Med. 1988. V. 318. P. 1225-1230.
13. Porterfield D., Corkey R., Sanger R. et al. Oxygen consumption oscillates in single clonal pancreatic beta-cells (HIT) // Diabetes. 2000. V. 49. P. 1511–1516.
14. Cilento G., Adam W. From free radicals to electronically excited species. // Free Radic. Biol. Med. 1995. V. 19. P. 103-114.
15. Баскаков И В, Воейков В Л Роль электронно-возбужденных состояний в биохимических процессах // Биохимия 1996 Т 61 С. 837-844.
16. Воейков В. Л., Колдунов В В, Кононов Д С Новый колебательный процесс в водных растворах соединений, содержащих карбонильные и аминогруппы // Кинетика и катализ 2001. Т. 42. С 670-672.
17. Воейков В Л, Колдунов В. В., Кононов Д. С Длительные колебания хемилюминесценции в ходе амино-карбонильной реакции в водных растворах // Журн. физ. химии 2001. Т. 75. С. 1579-1585.
18. Нейман М В О противоположных тенденциях в механизме окислительной деструкции полимеров // Химическая кинетика и цепные реакции М. Наука, 1966. С. 408-430.

ПРОБЛЕМА ЗАТУХАНИЯ СУБГЛОБУЛЯРНЫХ КОЛЕБАНИЙ БЕЛКОВЫХ МОЛЕКУЛ В ВОДЕ

Ю. М. Романовский, А. В. Нетребко, О. А. Чичигина

Введение

Молекулы белков-ферментов, или, как их называют в настоящее время, молекулярные машины, осуществляют катализ многочисленных и взаимосвязанных химических превращений в живых организмах. При этом каждый фермент осуществляет вполне определенный набор операций над теми или иными молекулами-субстратами, превращая их в продукты реакции. Например, ферменты-гидролазы, к которым относятся α -химотрипсин (АХТ), ацеталхолинэстераза и др. являются «молекулярными ножницами», эффективно разрывающими химические связи. В дальнейшем рассмотрении в качестве примера мы будем использовать молекулу АХТ (см., например, [1, 2]).

Молекулярные машины работают в водном окружении. Они способствуют тем химическим превращениям, которые, в принципе, могут осуществиться в воде и без их участия, но со скоростями на много порядков меньшими. Например, АХТ разрывает пептидные связи ($-N = C-$), объединяющие аминокислоты в пептидную цепь – простейшую белковую структуру. Он «режет» белки на отдельные аминокислоты, которые уже могут усваиваться организмом в процессе пищеварения. Разрыв пептидной связи может произойти и в воде в результате тепловых флуктуаций, но с очень малой вероятностью. Даже и в этом случае действует катализатор – ион гидроксила. В процессе диффузии он может «случайно» удачно встретиться с атомной группой ($-N = C-$), разрыхлить электронное облако двойной связи и тем самым способствовать действию случайных напряжений, возникающих в результате тепловых флуктуаций (рис. 1).

Молекула белка-фермента представляет собой сложную конструкцию, состоящую из сотен и даже тысяч аминокислот.

Альфа-спирали и бета-структуры аминокислот объединяются в субглобулы и кластеры с помощью ионных, химических, водородных и ван-дер-ваальсовых связей. В целом такая макромолекула представляет собой сложную колебательную систему с тысячами степеней свободы. При этом отдельные субглобулы и кластеры являются более жесткими образованиями и объединяются «мягкими» связями. Трехмерная структура белков в кристаллическом состоянии хорошо расшифровывается современными рентгеновскими методами, а их динамика, как в кристаллическом виде, так и в растворе, исследуется различными спектроскопическими методами. Белковая молекула может находиться в различных конформациях,

которые отличаются друг от друга расположением кластеров и отдельных атомных групп. Когда молекула субстрата или ее часть (обрабатываемая «деталь») специфически связывается в кармане активного центра (АЦ) молекулярной машины, комплекс «субстрат – белок» образует конформацию, при которой разрываемая связь располагается вблизи активных заряженных атомных групп.

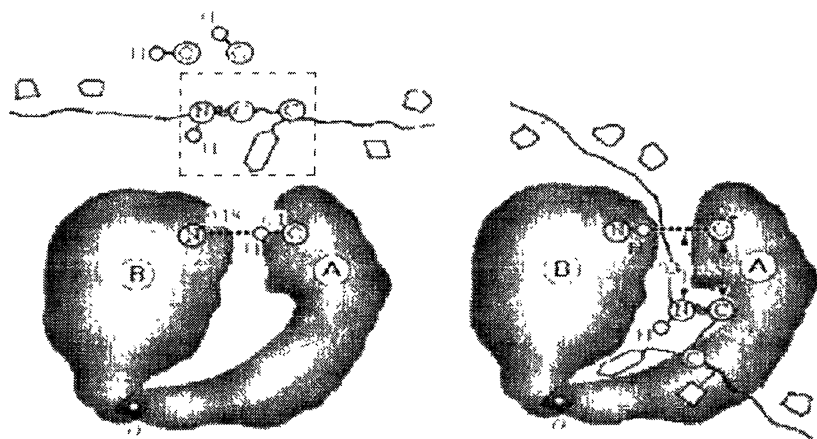


Рис. 1. Схема молекулы α -химотрипсина (слева), состоящая из двух субглобул А и В, пептидная цепочка вне АЦ, состоящая из аминокислот, связанных последовательно пептидными связями, и ион гидроксила OH^- Между субглобулами находится АЦ, существенным элементом которого является комплекс с водородной связью $\text{N} \cdots \text{HO}$. Атомы O и N расположены, соответственно, на субглобулах А и В (расстояния даны в нанометрах) Пунктиром обведены атакуемое звено пептидной цепи (пептидная связь показана волнистыми линиями) и неизменяемая часть субстрата Справа – неизменяемая часть субстрата закреплена в кармане АЦ, водородная связь и пептидная связь зафиксированы в определенном положении относительно иона O^- , при этом протон перешел к атому N Точка O – «шарнир», соединяющий субглобулы А и В [3, 4, 5, 6]

По меткому выражению М. В. Волькенштейна [7], молекула субстрата попадает в «структурированный раствор». Иначе можно сказать, что она располагается в трехмерном силовом поле, испытывая при этом статические и случайные напряжения. В структурированном растворе роль диффузии играют флуктуационные движения, роль гидроксила в АЦ АХТ играет ион кислорода одной из аминокислот, располагающийся на удачной позиции вблизи пептидной связи (см. рис. 1) на субглобуле А. Этот ион кислорода образуется при переходе H к атому N на субглобуле В. Переход H от O к N становится возможным благодаря конформациям в АЦ, возникающим при посадке субстрата.

Вся совокупность движений способствует также эффективному проникновению субстрата в АЦ и выходу продуктов реакции из него. Таким образом, для понимания механизмов работы ферментов очень важно знать характеристики случайных движений всех элементов конструкции, и в частности субглобул. АЦ АХТ располагается между субглобулами, поэтому весьма важно знать характерные амплитуды, частоты и добротности колебаний субглобул, возникающих в момент сорбции субстрата в АЦ. Именно они и случайные воздействия молекул воды во многом определяют флуктуации параметров силовых полей в АЦ со связанным субстратом.

Данные КР-спектроскопии (рамановской) свидетельствуют, что многие белки, состоящие из двух субглобул в кристаллическом состоянии, и в частности АХТ, имеют резонансы на частотах порядка 10^{12} Гц. Это подтверждается также численными оценками. Однако в водном растворе оптическими методами подобные резонансы наблюдать не удается. Использование формулы Стокса для оценки добротности к тому же показывает, что субглобулярные колебания передемпфированы. Этот факт охладил пыл экспериментаторов в поисках резонансов на таких частотах; кроме того, КР-спектры в воде для этого диапазона получить очень трудно [8, 9].

Целью цикла исследований, представленных в настоящей работе, является оценка характеристик затухания таких субглобулярных колебаний с помощью методов статистической физики и молекулярной динамики. Так, результаты, полученные методами механики сплошной среды, представляются не всегда неправомерными [8, 9, 10, 11]. Сначала кратко остановимся на последних.

1. Оценка добротности методом Стокса – Ламба

Вместо модели из двух почти сферических субглобул АХТ, которые имеют некий локальный шарнир, а роль пружин играют водородные связи, введем упрощенную модель одной колеблющейся субглобулы с половинной массой этой молекулы. Ее масса $M = 12,5$ тыс. ат. ед. $\approx 4 \cdot 10^{-20}$ г и радиус $R = 20$ Å, характерная частота $\omega = 2\pi \cdot 10^{12}$ с $^{-1}$, а плотность белка мало отличается от плотности воды ρ . Пусть такая масса совершает колебания с амплитудой a и скоростью $V(t)$. Тогда согласно теории Стокса – Ламба (см. [12]) сила, действующая на глобулу, будет равна

$$F = 6\pi\eta R \left(1 + \frac{R}{\delta}\right) V + 3\pi R^2 \sqrt{\frac{2\eta\rho}{\omega}} \left(1 + \frac{2R}{9\delta}\right) \frac{dV}{dt} = F_{fr} + F_{in}, \quad (1)$$

где первый член – сила трения $F_{fr} = hV$ (h – коэффициент трения), второй – сила инерции, обусловленная наличием присоединенной массы, η – коэффициент вязкости воды, а

$$\delta = \sqrt{\frac{2\eta}{\rho\omega}} \quad (2)$$

– толщина пограничного слоя. Эти выражения справедливы, когда $\delta \ll 2R$, $(2R)^2\omega \gg \eta / \rho$ и $a \ll 2R$. В реальности $a \approx 0,1-0,3 \text{ \AA}$, и эти условия выполняются. При этом $\delta \approx 5 \text{ \AA}$ и лишь немного превосходит радиус молекулы воды, а амплитуда колебаний много меньше ее диаметра. В этих условиях представление о воде как о вязкой сплошной среде становится не вполне адекватным. Добротность колебаний

$$Q = \frac{\omega}{2\gamma} = \omega \frac{M}{h}. \quad (3)$$

Здесь γ – декремент затухания, полная масса M должна включать присоединенную массу, которая близка к массе глобулы и массу водяной «шубы» (порядка нескольких сотен масс молекул воды), которая всегда окружает глобулу. При этом $Q \approx 2$. Помимо тех замечаний, которые мы уже сделали, отметим, что на микроуровне молекулы воды «нелинейно» взаимодействуют друг с другом и с субглобулой. Это взаимодействие можно описать приближенно с помощью потенциала Леннарда – Джонса. Поэтому можно ожидать зависимости Q не только от частоты, но и от амплитуды колебаний.

Приведем также оценку добротности валентных колебаний малой молекулы этана ($\text{H}_3\text{C} - \text{CH}_3$) в воде. Как уже говорилось, валентные колебания в малых атомных группах в АЦ фермента могут играть важную функциональную роль. Согласно формулам (1) – (3) для таких колебаний $Q < 1$ [8, 9, 11].

2. Затухание колебаний субглобул при взаимодействии с молекулами воды. Аналитическое рассмотрение

2.1 Случай быстрых взаимодействий с частицами среды

Рассмотрим одномерный линейный осциллятор-поршень со скоростью $V(t) = V_0 \cos \omega t$ массой M в окружении легких частиц среды $m = M/10^3$, движущихся со скоростью v . Найдем, как кинетическая энергия частицы среды («молекулы воды») ΔE изменится в результате взаимодействия с таким осциллятором. На такую же величину, но с обратным знаком изменится энергия осциллятора. Будем считать, что изменениями в движении осциллятора в результате единичного взаимодействия можно пренебречь. Для начала рассмотрим случай абсолютно упругого удара, что соответствует удару с малым временем взаимодействия τ_{int} по сравнению с периодом колебания поршня $T = 2\pi / \omega$ ($\tau_{\text{int}} \ll T$), изменение энергии для такого случая будем записывать с индексом «0» (ΔE_0).

Из закона отражения от движущегося поршня скорость частицы после удара $v' = 2V_0 \cos \phi - v$, где $\phi = \omega t$ – фаза колебаний поршня в момент соударения.

На основе всех представленных здесь рассуждений получена зависимость изменения энергии от фазы колебания поршня:

$$\Delta E_0 = -2mV_0^2(y - \cos \omega t) \cos \omega t,$$

где $y = v / V_0$. Это хорошо согласуется с данными компьютерного эксперимента. Полученное выражение будет ниже использовано для усреднения изменения энергии по времени и по скоростям молекул воды.

2.2. Случай медленных взаимодействий с частицами среды

Теперь рассмотрим противоположный предел – случай медленных взаимодействий субглобулы и частиц среды, когда $\tau_{\text{int}} \gg T$. Перейдем в неинерциальную систему отсчета, движущуюся с поршнем. Тогда частица среды движется в потенциале, созданном суперпозицией потенциала взаимодействия с поршнем и быстро периодически меняющегося потенциала силы инерции. Движение в такой потенциальной яме будет обратимым, следовательно, изменения кинетической энергии частицы после прохождения такой потенциальной ямы туда и обратно не произойдет.

Полуфеноменологический переход от одного предельного случая к другому можно совершить с помощью формулы

$$\Delta E = \frac{T}{T + \tau_{\text{int}}} \Delta E_0.$$

Применимость этой формулы проверялась нами с помощью компьютерных экспериментов. При малых значениях τ_{int} зависимость ΔE от него прослеживается слабо. Так же мало меняется ΔE при малой глубине потенциала взаимодействия. Это связано с тем, что время взаимодействия в этом случае хоть и велико, но слабо зависит от глубины ямы, будучи определяемым по известной формуле

$$\tau_{\text{int}} = \sqrt{\frac{m}{2}} \int_{-R}^R \frac{dr}{\sqrt{E - U(r)}},$$

где E – полная энергия частицы (воды), она же – кинетическая энергия частицы до начала взаимодействия, R – радиус взаимодействия.

2.3. Среднее за период изменение энергии

Для усреднения изменения энергии в результате столкновения частицы с колеблющимся поршнем необходимо учесть линейную зависимость вероятности столкновения от относительной скорости движения частицы. Распределение вероятностей запишется в виде

$$w(\phi) = \frac{1}{2\pi v} (v + V_0 \cos \phi),$$

при условии, что $y > \cos \phi$, иначе столкновение невозможно, $w(\phi) = 0$. Мы считаем, что частица подлетает справа, ее скорость $-v$.

Усредняя изменение энергии по фазам с данным распределением, получаем для $y \geq 1$: $\langle \Delta E_0 \rangle = 2mV_0^2$; а для $y < 1$ ситуация гораздо сложнее. Столкновения будут происходить только при определенных значениях фазы, так как частица не догонит поршень при значениях фазы вблизи π . Таким образом, фаза меняется от некоторого значения ψ до ϕ_y . Последнее ограничение определяется тем, что частица догоняет поршень в тот момент, когда их скорости становятся равными ($V = -v$) и удара не происходит, или, что то же самое, он происходит, но скорости не изменяются в результате него, т. е. частица «касается» поршня. Очевидно, что $\phi_y = \arccos(-y)$ и $\pi/2 < \phi_y < \pi$. Тогда значение ψ будет определяться условием того, что в предыдущий период колебания при значении фазы $\phi_y - 2\pi$ частица коснулась стенки и продолжала двигаться с той же скоростью, пока не столкнулась с поршнем. Тогда изменение фазы поршня $\psi - \phi_y + 2\pi$, изменение координаты частицы $(\sin \psi - \sin \phi_y)V_0/\omega$. Окончательно получаем уравнение для фазы ψ , начиная с которой возможно столкновение:

$$y(\psi - \phi_y + 2\pi) = \sin \phi_y - \sin \psi.$$

Отсюда видно, что $-\pi < \psi < \pi/2$.

Усредняя изменение энергии в пределах от ψ до ϕ_y , получаем

$$\langle \Delta E_0 \rangle = \frac{mV_0^2}{\pi y} \left[2\pi y - y^2 \sin \psi - \frac{y}{2} \sin 2\psi - \frac{1}{3} (\sin^3 \phi - \sin^3 \psi) \right],$$

$$\langle \Delta E_0 \rangle = 2mV_0^2 F(y).$$

Это выражение соответствует $y < 1$. Сложную функцию $F(y)$ можно приближенно представить в виде $\tilde{F}(y) = Ay + By^2 + Cy^3$ на интервале $0 < y < 0,4$, где $A = 5,7$; $B = 5,7$; $C = 9,8$. Разложение $F(y)$ в ряд вблизи 0 дает выражение $F(y) = 2\pi y$. А на оставшемся интервале $0,4 < y$ имеем $F(y) = 1$. Такое упрощенное выражение может быть использовано в дальнейшем при усреднении по v , это усреднение мы будем обозначать чертой сверху. Таким образом, мы получили убывание энергии, отдаваемой поршнем при увеличении скорости его колебаний по сравнению со скоростью движения частиц среды. При неизменной амплитуде это означает убывание затухания при увеличении частоты колебаний.

Такой аномальный рост эффективной добротности становится заметным при отношении скоростей $y_0 = 0,4$. Это связано с тем, что при быстрых колебаниях столкновения с частицами среды происходят только при максимальном отклонении и нулевой скорости стенки, т. е. они не влияют на ее движение.

Уравнение изменения энергии выглядит следующим образом:

$$\frac{d\bar{E}}{dt} = 2\langle\Delta E_0\rangle v, \quad (4)$$

где $v = nSv/6$ – частота ударов частиц среды, n – концентрация, S – площадь осциллятора, множитель 2 появился из-за того, что «молекулы воды» налетают с двух сторон по отношению к поршню-субглобуле, а 6 – от различия направлений движения частиц среды.

После усреднения по скоростям частиц среды, предполагая их нормальное распределение, получаем зависимость для производной энергии от обезразмеренной скорости движения осциллятора, представленную на рис. 2:

$$z = V_0 y_0 \sqrt{\frac{m}{2kT}}$$

– параметр, характеризующий отношение скорости субглобулы к средней скорости теплового движения частиц среды. Здесь выбран характерный масштаб времени

$$\tau = \frac{3}{Sn} \left(\frac{m}{kT} \right)^{1/2}.$$

Мы видим наличие минимума затухания при $z \approx 1,5$, что соответствует скоростям и амплитудам, приводимым выше, и частотам колебаний «субглобул» для различных молекул. Так, для субглобулы АХТ, для которой $\tau \sim 10^{-14}$ с и амплитуда колебаний a порядка 1 \AA , $z \sim 1$; для этана $\tau \sim 10^{-13}$ с, $a \sim 0,01 \text{ \AA}$, z – также порядка единицы. Здесь $V_0 = a\omega$. Далее, решая уравнение (4), получаем среднюю энергию как функцию времени (рис. 3). Видим ее сильное отличие от экспоненциальной зависимости.

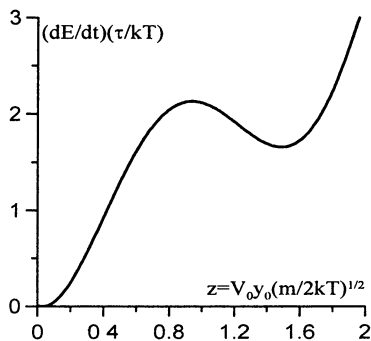


Рис. 2. Зависимость средней безразмерной производной энергии колебаний субглобулы от отношения ее скорости к характерной скорости движения частиц среды

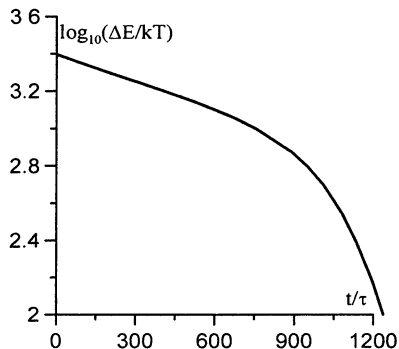


Рис. 3. Зависимость средней энергии (ΔE) колебания субглобулы от безразмерного времени

Мы получили, что при больших амплитудах колебания эффективная добротность много больше, чем добротность, вычисляемая по Стоксу. Причем в рассмотренном нами диапазоне амплитуд она меняется от 10^2 до 10^1 для АХТ и от 10^4 до 10^2 для этана. Для малых амплитуд ($< 0,1 \text{ \AA}$) добротность на порядок меньше, что согласуется с выводами п. 3.

Важно отметить, что аналитические вычисления проведены только для случая абсолютно упругого удара. Учет вида потенциала взаимодействия заметным образом усложняет вычисления и делает необходимым проведение компьютерного эксперимента.

3. Оценка затухания методом молекулярной динамики

Если аналитические оценки делались для простоты в случае 1D, то здесь мы рассмотрим более общий случай 2D-системы. Кластер (или глобула) в виде диска массой M и радиусом R помещается в центре квадратного «ящика» размером $L \times L$ (например, «в кармане АЦ») и прикрепляется к стенкам четырьмя пружинами с жесткостями k (см. [8, 9, 11]). Потенциальная энергия данной системы в случае линейных пружин записывается следующим образом:

$$U(x, y) = \frac{k}{2} \left(\sqrt{y^2 + (l-x)^2} - l \right)^2 + \frac{k}{2} \left(\sqrt{x^2 + (l-y)^2} - l \right)^2 + \frac{k}{2} \left(\sqrt{y^2 + (l+x)^2} - l \right)^2 + \frac{k}{2} \left(\sqrt{x^2 + (l+y)^2} - l \right)^2,$$

где l – длины пружин; x, y – декартовы координаты положения центра масс диска (предполагается, что положение равновесия находится в начале координат).

В пространстве между диском и отражающими стенками помещались n малых дисков массой m и радиусом r (молекулы воды). Взаимодействия между всеми дисками, а также между дисками воды и стенками «ящика» задавались в виде потенциала Леннарда – Джонса или его упрощенной модификации:

$$U_{LJ} = \begin{cases} 4U_0 \left[\left(\frac{d_0}{d} \right)^{12} - \left(\frac{d_0}{d} \right)^6 \right] & \text{при } d < d_0 \sqrt[6]{2}; \\ -U_0 & \text{при } d > d_0 \sqrt[6]{2}. \end{cases}$$

Здесь d – расстояние между центрами дисков или расстояние от центра диска воды до стенки; d_0 полагается равным $R + r$ для случая «большой – маленький» диски, $r + r$ для случая «маленький – маленький» диски, r для случая «маленький диск – стенка»; U_0 – коэффициент, характеризующий «жесткость» потенциала взаимодействия (от его величины зависит степень сближения дисков между собой).

Уравнения движения центра масс диска-глобулы запишем в виде

$$\begin{cases} M \frac{d^2 x_g}{dt^2} = -\frac{\partial U(x_g, y_g)}{\partial x} - \sum_{i=1}^n \frac{\partial}{\partial x} U_{i,j}(r_{gw}), \\ M \frac{d^2 y_g}{dt^2} = -\frac{\partial U(x_g, y_g)}{\partial y} - \sum_{i=1}^n \frac{\partial}{\partial y} U_{i,j}(r_{gw}). \end{cases}$$

Здесь r_{gw} – расстояние между центрами масс диска-глобулы и дисков-воды.

Уравнения движения центра масс j -го диска молекулы воды:

$$\begin{cases} m \frac{d^2 x_w}{dt^2} = -\frac{\partial U_{i,j}(r_{gw})}{\partial x} - \frac{\partial U_{i,j}(r_{ws})}{\partial x} - \sum_{\substack{i=1 \\ i \neq j}}^n \frac{\partial}{\partial x} U_{i,j}(r_{ww}), \\ M \frac{d^2 y_w}{dt^2} = -\frac{\partial U_{i,j}(r_{gw})}{\partial y} - \frac{\partial U_{i,j}(r_{ws})}{\partial y} - \sum_{\substack{i=1 \\ i \neq j}}^n \frac{\partial}{\partial y} U_{i,j}(r_{ww}). \end{cases}$$

Здесь r_{ww} – расстояние между центрами дисков воды, r_{ws} – расстояние от центра малого диска до стенки.

Начальное положение дисков показано на рис. 4 ($U_0 = 0,25$) и 5 ($U_0 = 25$), модуль начальных скоростей у всех молекул воды был одинаков и соответствовал средней скорости воды при обычной температуре – 300 м/с ($3 \text{ \AA}/\text{пс}$). Направления же скоростей выбирались из условия равномерного распределения их по углам. По прошествии 200 пс, за которые распределение скоростей становилось близким к максвелловскому, глобула получала начальный толчок в горизонтальном направлении. После этого наблюдался процесс релаксации ее колебаний.

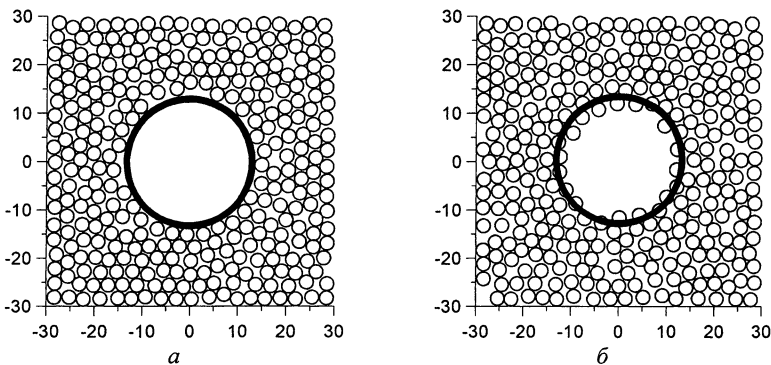


Рис. 4. Первоначальное расположение глобулы и молекул воды для случаев «жестких» ($U_0 = 25$) (а) и «мягких» ($U_0 = 0,25$) (б) потенциалов взаимодействия

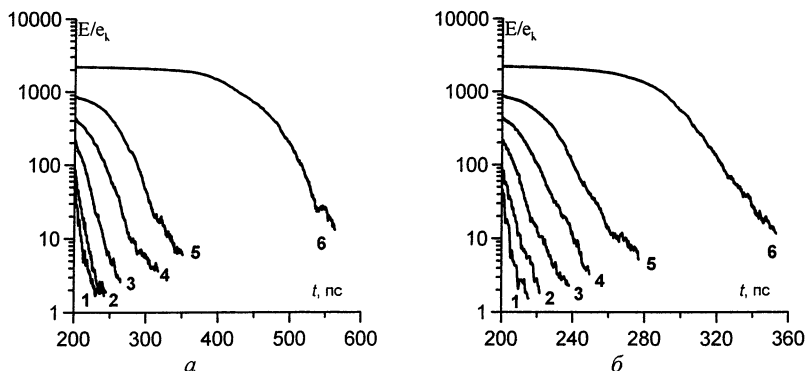


Рис. 5. Релаксация полной энергии глобулы в зависимости от времени для различных вариантов «жесткости» потенциалов (а и б) частот свободных колебаний.

В расчетах $n = 356$; $M/m = 100$; $r = 1,5 \text{ \AA}$; $R = 15 \text{ \AA}$; $L = 60 \text{ \AA}$; $l = 50 \text{ \AA}$. Жесткости пружин подбирались из условия, что свободные колебания глобулы происходят с частотами порядка 10^{-12} Гц (1/пс). В расчетах

$\omega = \sqrt{\frac{2k}{M}}$ задавалось равным $\sqrt{4}$ (кривая 1 на рис. 5); $\sqrt{10}$ (кривая 2);

$\sqrt{20}$ (кривая 3); $\sqrt{40}$ (кривая 4); $\sqrt{100}$ (кривая 5); $\sqrt{200}$ 1/пс (кривая 6).

Амплитуда свободных колебаний глобулы (без учета взаимодействия с молекулами воды) полагалась равной 1 \AA . Соответственно скорость первоначального толчка для глобулы $V_0 = \sqrt{\frac{2k}{M}}$ Å/пс. Таким образом, были

рассмотрены варианты, когда скорости движения глобулы сравнимы со скоростями хаотического движения молекул воды, а также превосходят их на порядок. Были рассмотрены три варианта жесткости потенциалов взаимодействия частиц: «мягкий» – $U_0 = 0,25$, «средний» – $U_0 = 2,5$ и «жесткий» – $U_0 = 25$. Все данные получены путем усреднения по десяти реализациям.

На рис. 5 представлены зависимости полной энергии E (кинетическая плюс потенциальная) глобулы (без учета энергии взаимодействия с молекулами воды), отнесенной к первоначальной кинетической энергии одной молекулы воды $e_k = kT$ от времени. Эти расчеты обобщают данные работ [8, 9, 11].

Из представленных результатов можно сделать следующие выводы. Время релаксации существенным образом зависит от скорости движения глобулы. При больших значениях происходят почти незатухающие колебания. Визуализация происходящих процессов во время счета показывает, что глобула, благодаря своей относительно большой массе и энергии,

«пробивает» для себя дорожку, в результате чего взаимодействия с окружающими молекулами воды происходят только в моменты, когда пружины растянуты и, следовательно, собственная скорость (а значит, и кинетическая энергия) глобулы практически равна нулю. В результате передачи энергии не происходит. В свою очередь, молекулы воды, имея незначительную скорость, не успевают проникать в зону больших скоростей глобулы. Единственная возможность получить значительный прирост энергии имеется у молекул воды, попадающих в область высоких скоростей глобулы «сверху» и «снизу». Именно они в основном и приводят к релаксации на первом этапе. По мере развития этого процесса скорость движения глобулы постепенно уменьшается, а у молекул воды наоборот увеличивается, и в тот момент, когда они примерно сравниваются, начинается практически лавинообразный процесс рассеяния энергии.

На скорость релаксации большое влияние оказывает «жесткость» потенциалов взаимодействия молекул воды. Чем ниже эта «жесткость», тем легче глобуле «пробить» себе дорожку. При высоких «жесткостях» молекулы воды просто физически не могут быть удалены из области больших скоростей глобулы, а следовательно, происходит весьма интенсивный обмен энергиями.

В наших предыдущих работах [8, 9, 11] при малых начальных энергиях и амплитудах затухание в начале процесса релаксации больше, чем на его «хвосте», когда средняя амплитуда глобулы близка к амплитуде теплового равновесия. При этом средняя эффективная добротность при мягком взаимодействии молекул воды, соответствующая энергии водородных связей, больше, чем оценки по Стоксу – Ламбу для субглобул АХТ, в несколько раз. Релаксация энергии колебаний субглобулы для этого случая соответствует кривой 1 на рис. 5. Добротность же валентных колебаний молекулы этана, полученная на нашей модели, превосходит оценку добротности, полученную на основе теории сплошной среды, на три порядка. Этот результат совпадает с данными, полученными методом молекулярной динамики, сделанными в 3D-случае с учетом более точных потенциалов взаимодействия молекул воды между собой и с молекулой этана (см. [13]).

Заключение

Нами была исследована скорость затухания колебаний субглобулы в среде как аналитическими методами, так и с помощью компьютерного моделирования. Получено явное отклонение от обычной линейной зависимости силы трения от скорости. А именно, обнаружено гораздо более медленное затухание колебаний. Исследованы различные причины такого эффекта. Вот эти причины:

1. Зависимость скорости затухания от типа взаимодействия субглобулы с частицами среды. В случае мягких взаимодействий мы получили

более медленное затухание. Этот вопрос исследовался главным образом методами молекулярной динамики и представлен графически. Теоретические оценки позволяют предположить, что время затухания в случае медленных взаимодействий увеличивается на величину порядка времени взаимодействия.

2. Принципиальное отличие колебательного движения, происходящего в замкнутой области пространства, от поступательного. И компьютерное, и аналитическое исследования показывают сильное, но не полное подавление затухания при больших скоростях колебаний. Это происходит при отношении средней скорости движения частиц среды к амплитуде скорости колебания равном 0,4. Тогда начинает сказываться зависимость вероятности соударения субглобулы с частицей среды от фазы ее колебания. Большинство соударений происходят в моменты, когда субглобула отклонена в крайнее положение и ее скорость близка к нулю. При таких соударениях потери энергии субглобулой минимальны. Можно сказать, что при быстром движении она «успевает распахивать молекулы среды» и движется в пустом пространстве.

3. Компьютерное моделирование показало нетривиальные эффекты, связанные с нагреванием среды частицей вблизи области ее движения, что означает увеличение скорости частиц среды, окружающих субглобулу, и несколько уменьшает эффект, описанный в п. 2.

4. Мы предполагаем, что заметное влияние на процесс затухания колебаний в воде могут оказывать свойства самой воды. Компьютерное моделирование показало, что при мягких взаимодействиях частиц среды друг с другом затухание происходит медленнее, однако это может быть связано с особенностями используемой модели.

5. Отметим, что релаксация энергии колебаний «глобулы» во времени при высоких начальных энергиях (кривая б на рис. 5) совпадает с кривой на рис. 3, рассчитанной аналитически. Однако физические условия в обоих случаях отличаются. В 2D-модели система замкнута, число молекул ограничено, молекулы воды взаимодействуют между собой. Наиболее близки результаты, полученные в п. п. 2 и 3 для «жестких взаимодействий».

6. Более точные 3D-модели могут быть построены с использованием «трехточечных» моделей молекул воды [14] и с учетом водной шубы на поверхности молекулы фермента.

7. Полученный эффект увеличения добротности в зависимости от соотношения скоростей поршня и частиц среды можно рассматривать как элемент самоорганизации системы. При оптимальной температуре термостата (когда $z = 1,5$) система дольше проявляет свои внутренние колебательные свойства.

Авторы выражают глубокую признательность А. Ю. Чикишеву за плодотворные дискуссии.

Работа была поддержана грантом «Научные школы России» (00-15-97843), грантом РФФИ (01-03-33163), грантом INTAS (01-0450) и грантом по Междисциплинарному проекту МГУ «Молекулярная динамика ферментов».

Литература

- 1 *Ферит Э.* Структура и механизм действия ферментов М Мир, 1980
- 2 *Чернавский Д С, Чернавская Н М* Белок-машина Биологические макромолекулярные конструкции М Изд-во Москов ун-та, 1999
- 3 *Нетребко А В, Нетребко Н В, Романовский Ю М, Хургин Ю. И, Шидловская Е Г* Сложные модуляционные режимы и стохастизация колебаний в кластерных динамических моделях макромолекул // Изв вузов Прикладная нелинейная динамика 1994 Т 2, № 3 С 26–43
- 4 *Нетребко А В, Нетребко Н В, Романовский Ю М, Хургин Ю. И, Эбелинг В* Стохастическая кластерная динамика фермент-субстратного комплекса // Изв. вузов. Прикладная нелинейная динамика 1996 Т 4, № 3 С 53–66
- 5 *Романовский Ю М, Тихомирова Н К, Хургин Ю И* Электромеханическая модель фермент-субстратного комплекса // Биофизика. 1979 Т 24 С 442
- 6 *Шувалова Е В, Кубасов А А, Романовский Ю М, Чикишев А. Ю* Динамика переноса протона в активном центре белка-фермента α -химотрипсина // Изв вузов Прикладная нелинейная динамика. 2000. Т 8, № 5. С 23–35
- 7 *Волькенштейн М В, Голованов И Б, Соболев В М* Молекулярные орбитали в энзимологии. М.. Наука, 1982
- 8 *Romanovsky Yu. M., Netrebko A. V., Chikishev A Yu* Are the subglobular oscillations of protein molecules in water overdamped? In Saratov Fall Meeting 2001 Optical Technologies in Biophysics and Medicine. III / Ed Valery V Tuchin // Proc Of SPIE. V 4707. P 16–29
- 9 *Ebeling W., Romanovsky Yu., Schimansky-Geier L* (eds.) Stochastics Dynamics of Reacting Biomolecules. Singapore World Scientific 2002
- 10 *Романовский Ю. М, Эбелинг В* (ред) Молекулярная динамика ферментов М Изд-во Москов. ун-та, 2000
- 11 *Romanovsky Yu M., Chikishev A Yu., Kroo S. V, Netrebko A. V.* Computer simulation of the decay of nonlinear subglobular oscillations in aqueous solutions of protein molecules in Optical Diagnostics and Sensing of Biological Fluids and Glucose and Cholesterol Monitoring II / Eds Alexander V Priezzhev, Gerald L Cote // Proc. Of SPIE 2002. V. 4624. P. 139–149
- 12 *Ландау Л. Д., Лифшиц Е. М* Гидродинамика. М Наука, 1986
- 13 *Shaitan K V, Saraikin S S* On the effect of the amplitude of fluctuations on the friction coefficient of the Brownian oscillator in water // Biofizika, 2000 45 P 407–413.
- 14 *Wood D W* Computer simulation of water and aqueous solutions // Water. A comprehensive treatise / Ed F Franks 6. Recent advances. Plenum Press NY – London, 1979 P 279–409

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ИММУННОГО ОТВЕТА ОРГАНИЗМА НА ВТОРЖЕНИЕ ИНФЕКЦИИ

Ю. И. Неймарк

Иммунные системы животного и человека очень сложны. Сложно и многообразно их устройство, сложны и многовариантны их реакции на инфекцию, сложны и многолики средства противодействия. Сегодня описание иммунной системы занимает пару томов, и при этом еще многое остается незатронутым и неизвестным. Вместе с тем иммунная система, сколь бы сложна она ни была, предназначена оказывать противодействие и уничтожать вторгшуюся в организм чужеродную и вредную для него инфекцию, и в общих понятиях это противодействие инфекции и ее вредное влияние на организм могут быть описаны, опуская детали того, как это фактически происходит.

Есть инфекция, которая стремится размножиться и воздействует на организм, подавляя его жизнеспособность. Есть организм, который пытается ее истребить, организуя кроме изначального противодействия еще и дополнительное, как только ее обнаружит. Быстрота и эффективность ответа организма зависят от наследственных факторов и от приобретенного «жизненного опыта», а также от состояния организма в данный момент и резервов, которыми он располагает.

Таким образом, несмотря на всю сложность иммунного ответа организма на инфекцию, на первый план выступают три фактора: инфекция, ее размножение и влияние на организм, противодействие организма и возможности его реализации, зависящие от его состояния. Эти факторы можно количественно охарактеризовать тремя величинами: x , y и z – численностью инфекции, величиной противодействия организма и потенциалом организма. Инфекция, помимо своей численности x , характеризуется быстротой размножения в среде организма и подавляющим действием на организм. Противодействие характеризуется своим количеством y и его эффективностью, быстротой и мощностью пополнения w , зависящими от организма и его потенциала z . Все сказанное позволяет описать динамику иммунного ответа – изменение величин x , y , w и z – следующими четырьмя дифференциальными уравнениями:

$$\dot{x} = \lambda x - \frac{ax y}{1 + \alpha x} - \bar{\epsilon} x^2,$$

$$\begin{aligned}
 \dot{y} &= \begin{cases} -\frac{\bar{b}xy}{1+\alpha x} + w = K & \text{при } y > 0 \text{ или } y = 0 \text{ и } K \geq 0, \\ 0 & \text{при } y = 0 \text{ и } K < 0, \end{cases} \\
 \tau \dot{w} + w &= \begin{cases} 0 & \text{при } x \leq \bar{x}_0, \\ \bar{B}z(2z_0 - z)(x + \bar{\beta}x^2) & \text{при } x > \bar{x}_0, \end{cases} \\
 \dot{z} &= \begin{cases} \frac{c(z_0 - z)}{1 + \gamma x} - dy - e = F & \text{при } z \geq 0 \text{ или } z = 0 \text{ и } F \geq 0, \\ 0 & \text{при } z = 0 \text{ и } F < 0. \end{cases} \quad (1)
 \end{aligned}$$

В этих уравнениях x_0 – порог чувствительности организма к инфекции, τ – временная задержка дополнительного иммунного ответа, величина которого зависит от объема инфекции x и потенциала организма z . При $\bar{\beta} = 0$ зависимость от x линейная, при $\bar{\beta} > 0$ она форсирована, а при $\bar{\beta} < 0$, напротив, уменьшена. Последнее уравнение системы (1) описывает изменение потенциала z , при этом $0 \leq z \leq z_0$. Первый член последнего уравнения (1) – это скорость пополнения потенциала z до его предельного значения z_0 , второй и третий – его расход на иммунный ответ организма y и на потребности самого организма. Обращение z в нуль трактуется как полное истощение организма, ведущее к его гибели.

У четырех уравнений (1) 13 параметров. Параметр λ может иметь любой знак: $\lambda < 0$ означает, что инфекция, попав в среду организма, гибнет; при $\lambda > 0$ она все же размножается, но ей мешает в этом противодействие организма y , которого может и не быть. В дальнейшем $\lambda > 0$. Все остальные параметры, кроме $\bar{\beta}$, положительны. Процесс нарастания и динамики противодействия y инфекции x описывается последними тремя уравнениями.

Хотя сказанного достаточно для понимания предлагаемой модели (1), все же дадим более подробное описание каждого из ее уравнений. В первом уравнении член λx соответствует размножению инфекции, второй – ее уничтожению фактором y противодействия организма. Этот последний член, естественно, равен нулю при $x = 0$, максимален при $x = \infty$ и линейно зависит от y . Последний член отражает возможность конкуренции за условия размножения. Второе уравнение описывает изменение противодействия инфекции y . Первый член во втором уравнении при $y > 0$ соответствует тому, что на каждую единицу уничтоженной инфекции x происходит определенное уменьшение y . Второй член w в этом же уравнении – это скорость пополнения организмом противодействия y . Третье уравнение указывает, что иммунный ответ на инфекцию имеет место только

при $x > x_0$, происходит с временной задержкой и определяется величиной, зависящей от потенциала организма z , причем пропорционально x или с некоторым форсированием при $\bar{\beta} > 0$ или, наоборот, уменьшением с ростом x при $\bar{\beta} < 0$. Последнее уравнение описывает изменение потенциала z организма. В этом уравнении $0 \leq z < z_0$. Первый член правой части увеличивает z при $z < z_0$. Скорость этого роста потенциала убывает с увеличением вредного влияния инфекции x . Второй и третий члены – затраты на поддержание противодействия y и всего организма.

Модель (1) для исследования достаточно сложна, но ее можно значительно упростить, ограничившись переменными x и y . Это тем более возможно, что роль остальных уравнений довольно очевидна и может быть мысленно учтена и понята. Упрощенная модель имеет вид

$$\dot{x} = \lambda x - \frac{\alpha xy}{1 + \alpha x} - \bar{\epsilon} x^2,$$

$$\dot{y} = \begin{cases} -\frac{bxy}{1 + \alpha x} & \text{при } x \leq \bar{x}_0, \\ -\frac{bxy}{1 + \alpha x} + \bar{B}(x + \bar{\beta} x^2) = K & \text{при } x > \bar{x}_0 \text{ и } y > 0, \\ & \text{или } y = 0 \text{ и } K \geq 0, \\ 0 & \text{при } x > \bar{x}_0 \text{ и } y > 0 \text{ и } K < 0. \end{cases} \quad (2)$$

Отличие этих уравнений (2) от (1) в том, что величина \bar{B} теперь постоянна, в то время как в уравнениях (1) она меняется согласно двум последним уравнениям. Формально перейти от системы (1) к системе (2) можно, положив $\tau = 0$ и приняв z постоянным. Величина $\tau > 0$ определяет быстроту возрастания w от нуля до предельного значения $\bar{\bar{B}}z(2z_0 - z)(x + \bar{\beta} x^2)$, а \bar{B} в уравнениях (2) равно $\bar{\bar{B}}z(2z_0 - z)$ при принимаемом постоянном z . Так что роль последних двух уравнений в (1) может быть учтена как изменяемость \bar{B} в уравнениях (2).

Число параметров в упрощенных уравнениях (2) может быть уменьшено путем изменения масштабов по времени t и переменным x и y :

$$\dot{x} = x - \frac{xy}{1 + x} - \epsilon x^2,$$

$$y = \begin{cases} -\frac{bxy}{1+x} & \text{при } x \leq x_0, \\ -\frac{bxy}{1+x} + B(x + \beta x^2) = K & \text{при } x > x_0 \text{ и } y > 0, \\ & \text{или } y = 0 \text{ и } K \geq 0, \\ 0 & \text{при } x > x_0 \text{ и } K < 0. \end{cases} \quad (3)$$

Поймем, что следует из упрощенной модели (3), сначала приняв $\varepsilon = \beta = 0$. При этом остается три параметра: x_0 , b и B . Как оказывается, возможно два существенно разных случая: $b > B$ и $b < B$. Второй случай, $b < B$, распадается еще на два: $b < 1$ и $b > 1$. Так что имеют место четыре разных вида фазовых портретов, представленных на рис. 1.

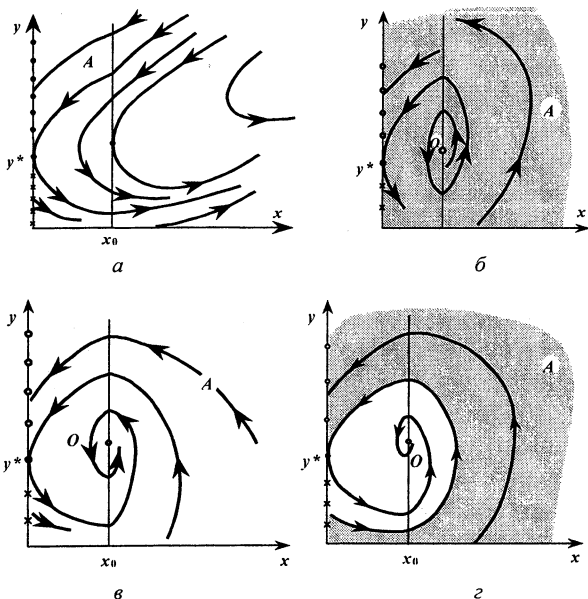


Рис. 1. Фазовые портреты упрощенной модели (3) при $\beta = \varepsilon = 0$, соответственно при $b > B$, $b < B$ и $b < 1$, $b < B$ и $b > 1$. A – область начальных условий, для которых имеет место выздоровление. Вне области A инфекция сохраняется на уровне чувствительности x_0 и имеет место бациллоносительство.

Как видно из рис. 1, в случае $b > B$ организм избегает гибели только при начальных условиях в области A , т. е. при достаточно высоком начальном уровне y . При недостаточном начальном противодействии y

организм гибнет (происходит неограниченное возрастание x , влекущее, согласно последнему уравнению (1), к занулению потенциала z). Это происходит даже при самой незначительной начальной инфицируемости. Напротив, при $b < B$ организм никогда не гибнет. Однако исход встречи с инфекцией может кончиться либо полным ее уничтожением (область A и приход к одному из равновесий, $x = 0, y > y^*$), либо сохранением инфекции на пороге чувствительности (весь положительный октант, кроме области A) благодаря приходу к устойчивому равновесию O . Полное выздоровление при начальных условиях в области A сопровождается одним или, возможно, несколькими обострениями (рис. 1, б). То, насколько ярко выражены эти обострения, зависит от массивности инфекции x и начального значения величины иммунного противодействия y . Носительство инфекции и равновесное сожительство с ней происходит путем колебательного прихода к устойчивому равновесию O (рис. 1, в-г).

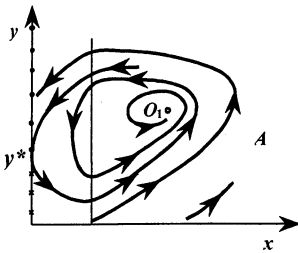
По рис. 1 можно проследить, каковы последствия уменьшения количества инфекции x , например, с помощью антибиотиков или других лекарств, в предположении, что это не оказывает существенного влияния на иммунный ответ, т. е. на величины b и B в уравнениях (3) и значение y . В случае $b > B$ и $y > y^*$ достаточное уменьшение инфекции x ведет к более быстрому выздоровлению или даже спасению организма от гибели. При $y < y^*$ выздоровление требует полного уничтожения инфекции ($x = 0$). Это выздоровление носит неустойчивый характер: малейшее повторное внесение инфекции или остаточная инфекция влекут после временного улучшения повторное заболевание, оканчивающееся гибелью организма. В случае $b < B$ роль такого лекарственного воздействия скромнее и не всегда положительна, так как может привести к сожительству с инфекцией, хотя и облегчает течение заболевания. Вместе с тем необходимо учитывать, что значительное и длительное обострение, через которое может произойти, согласно рис. 1, б-г выздоровление, может реально не состояться в силу не учитываемого нами уменьшения потенциала z , влекущего уменьшение величины B в уравнениях (3). И еще необходимо отметить, что на самом деле, согласно полной модели (1), вначале, в момент заражения, $B = 0$, и лишь затем достигает своего, рассматриваемого нами значения, причем рост B в течение времени порядка τ происходит только после того, как x превзойдет x_0 .

Приведенные рисунки фазовых портретов, как правило, указывают на желательность привнесенного извне увеличения y , стимуляции иммунной системы, что убыстряет выздоровление и позволяет избежать перехода в хроническую инфицированность. Это будет видно и из дальнейшего, когда выявится роль β . Но, пожалуй, самый существенный вывод, кроме

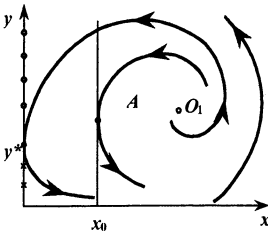
крайней желательности обеспечения $b < B$, это то, что целесообразно поддерживать изначальное значение y больше, чем y^* (см. рис. 1).

Действительно, если в начальный момент, до заражения, иммунная система находится в равновесии, $x = 0, y \geq 0$, то при $y < y^*$ в силу неустойчивости равновесия даже незначительное увеличение x приводит к заболеванию, в то время как при $y > y^*$ происходит уничтожение инфекции и наступление нового близкого равновесного состояния. Уменьшение y может произойти в силу отсутствия асимптотической устойчивости состояний равновесия $x = 0, y > y^*$. Напротив, при сожительстве с инфекцией состояние равновесия организма O асимптотически устойчиво, и в этом смысле оно оберегает организм от вторичного заражения. Конечно, это только при стабильности параметров b и B , при их изменении эта инфекция, наоборот, может быть причиной повторного заболевания.

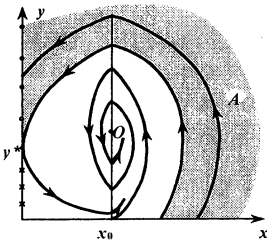
Теперь посмотрим, что нового вносят параметры β и ϵ , не учитываемые ранее ($\epsilon = \beta = 0$). Ограничимся иллюстрацией ряда фазовых портретов, срисованных с дисплея ЭВМ. Фазовые портреты (рис. 2)



a $b = 1,8, B = 1,0, \epsilon = 0, \beta = 0,2$ Состояние равновесия O_1 устойчиво и соответствует хроническому заболеванию В области A – выздоровление



б $b = 0,8, B = 0,5, \epsilon = 0, \beta = 0,2$ Состояние равновесия O_1 неустойчиво и при любых начальных условиях – выздоровление



в $b = 1, B = 0,9, \epsilon = 0,01, \beta = 0,05$ В области A – выздоровление, вне A – бациллоносительство

Рис. 2. Вид фазового портрета упрощенной модели (3) при $x_0 = 1$

демонстрируют роль форсированного иммунного ответа $\beta > 0$ на весьма неблагоприятном случае $b > B$. При этом все кончается хроническим заболеванием, бактерионосительством или выздоровлением. Конечно, длительное хроническое заболевание может навлечь существенное уменьшение потенциала z и следующее за ним уменьшение величины B , что может привести к фатальным последствиям. На рис. 3 представлен другой вариант форсировки иммунного ответа ($\beta > 0$), при котором возможно течение болезни с периодически повторяющимися рецидивами (автоколебание Γ).

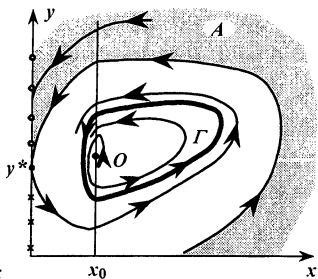


Рис. 3. Вид фазового портрета упрощенной модели (3) при $x_0 = 1, b = 0,9, B = 0,9, \epsilon = 0,01$ и $\beta = 0,05$. В области A – выздоровление, вне A – периодическое течение болезни (автоколебание Γ)

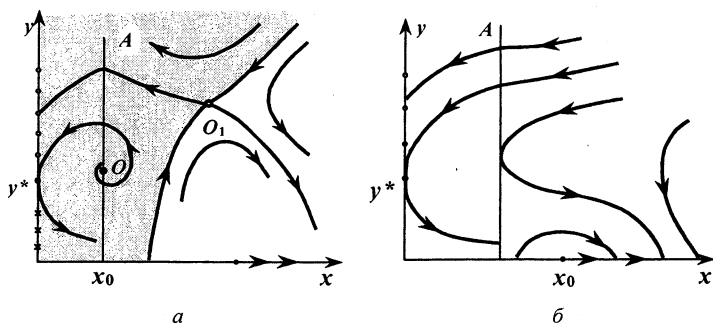


Рис. 4. Вид фазового портрета упрощенной модели (3) при $x_0 = 1, \epsilon = 0$, $a - b = 0,5, B = 2, \beta = -0,1$, $b - b = 0,8; B = 1, \beta = -0,2$. В области A – выздоровление, вне A – гибель

Что может произойти при ослаблении иммунного ответа ($\beta < 0$), демонстрируют рис. 4: несмотря на то, что $B > b$, возможен летальный исход. К ослаблению иммунного ответа может привести не только $\beta < 0$, но и уменьшение потенциала z согласно последнему, не учитываемому упрощенной системой (3), уравнению системы (1).

Математические проблемы нелинейной динамики

ФРАКТАЛЫ, МУЛЬТИФРАКТАЛЬНЫЕ МЕРЫ И АТТРАКТОРЫ

Н. Г. Макаренко

Эта статья является расширенным вариантом лекции, прочитанной на школе «Нелинейные волны» 2002 г. Она дополнена фрагментами из интернетовских *Заметок о Фракталах* [16] и работы [17]. Приведенное ниже не претендует ни на строгость изложения, ни на какую-либо завершенность и, конечно, не заменяет систематических курсов по геометрии фракталов [1–6]. Это лишь введение в основы интересной и популярной своими приложениями области.

1. Размерности и меры

В эпоху *Анри Пуанкаре* произвольное гладкое многообразие называли просто *непрерывностью*; под размерностью (параметрической) понималось минимальное число координат, необходимых для идентификации точки многообразия. Однако Кантор показал, что в рамках такого определения невозможно различить плоскость и линию. Действительно, пусть $x = (x_1, x_2)$ – точка квадрата $I = [0, 1] \times [0, 1]$. Ее координаты $0 \leq x_1, x_2 \leq 1$ можно представить в виде бинарных дробей: $x_1 = 0, \alpha_1, \alpha_2, \dots$; $x_2 = 0, \beta_1, \beta_2, \dots$, где α_i, β_i – нули либо единицы. Сопоставим паре разложений x_1, x_2 последовательность $z = 0, \alpha_1, \beta_1, \alpha_2, \beta_2, \dots$, наследующую *лексикографический порядок* исходных оригиналов. Мы получили отображение $z: I \rightarrow [0, 1]$. Следовательно, существует способ параметризации точек плоскости только одной координатой – извилистой линией! Разумеется, это возможно лишь в том случае, если мы не связываем себя условием непрерывности обратного отображения. *Анри Пуанкаре* заметил, что отказ от непрерывности трудно примирить с интуитивными логическими принципами. То, что было предложено им взамен, основано на весьма остроумной концепции *физической непрерывности* – цепочки, в которой каждый элемент неотличим от соседнего, но оба конца цепи легко различимы [7]. Размерность *непрерывности* вводится с помощью понятия *купюры*. Последняя образуется совокупностью элементов, изъятых из непрерывности. Если в результате такой операции непрерывность разделяется на две части, а сама купюра состоит лишь из конечного числа элементов, не образующих связного множества, то размерность непрерывности равна единице. Эти понятия легко переносятся на евклидовы геометрические объекты. Замкнутая кривая делится на две части двумя отмеченными точками. Сфера распадается на две половины сече-

нием, сделанным по замкнутой кривой. Пространство нельзя поделить точками или линиями, но можно сделать это замкнутыми поверхностями. Таким образом, математическая непрерывность имеет n измерений, если ее можно разбить на части, произведя в ней одно или несколько сечений, которые сами являются непрерывностями $n-1$ измерения. Это рекуррентное определение размерности, которое предполагает, что объемы – части пространства, поверхности – границы объемов, линии – границы поверхностей, а точки – границы линий. Формализация этих идей привела Брауэра, Урысона и Менгера к индуктивному определению *топологической размерности* [8]. Размерность любого конечного или счетного множества точек есть нуль. Размерность любого связного множества точек есть $d_i + 1$, если его можно разрезать на 2 несвязных куска, исключив как минимум d_i -мерное множество точек, т. е. сделав d_i -мерный разрез. Ясно, что топологическая размерность всегда есть целое число.

Размерность можно определить, опираясь на понятие покрытия и практические способы вычисления евклидовых мер – длин и площадей. Однако первым был, по-видимому, итальянский математик Никколо Фонтана Тарталья (1500–1557). Он заметил, что на три неколлинеарные точки плоскости можно натянуть треугольник, а четырьмя некопланарными точкам R^3 соответствует тетраэдр. Этим выпуклым телам можно приписать ненулевую меру – площадь и объем соответственно. Тарталья заключил, что пространство является n -мерным, если для его $n + 1$ точек, находящихся в общем положении, существует полиэдр ненулевого объема. В подходе Борхарда и Минковского [9] мера геометрического объекта некоторой размерности определяется как *скорость изменения меры* другого объекта большей размерности, покрывающего исходный, при *линейном* убывании некоторого характерного масштаба. Заметим, что это определение не зависит от выбора координат. Рассмотрим плоскую спрямляемую кривую C . Перемещение нормали длины 2ϵ так, чтобы ее середина двигалась вдоль кривой, образует *шарф Минковского* площадью $-S(\epsilon)$. Можно показать [9], что существует $\lim S(\epsilon)/2\epsilon$, при $\epsilon \rightarrow 0$, который называется длиной C . Аналогичным образом можно определить площадь как $\lim V(\epsilon)/2\epsilon$, $\epsilon \rightarrow 0$, где $V(\epsilon)$ – объем тела (*сосиски Минковского*), заметаемого отрезками 2ϵ , нормальными к поверхности.

Для определения размерности используется следующее обобщение. Регулярный многомерный объект можно представить в форме прямого произведения объектов низшей размерности. Тогда его мера выражается в виде степени характерного размера ϵ , элемента покрытия (разбиения) – ϵ^{\dim} . Степенная зависимость меры от ϵ называется *скейлингом*. Теперь размерность (\dim) можно определить как скорость изменения меры при измельчении разбиения, т. е. уменьшения ϵ по *степенному* закону. Поскольку при этом нас интересует только функциональная связь, вместо суммарного объема элементов разбиения (покрытия) часто используют

число таких элементов. Но покрытия (разбиения) можно реализовать многими способами. Поэтому для получения корректных численных оценок берется минимальное число элементов покрытия или разбиения. В последнем случае подсчитываются только непустые кубы или шары, содержащие, по меньшей мере, хотя бы одну точку множества.

Эти соображения лежат в основе размерности по Колмогорову, которую чаще называют емкостью, и размерности Хаусдорфа – Безиковича. Пусть (X, d) – полное метрическое пространство с метрикой d и $\dim X = d$. Рассмотрим непустое компактное подмножество $F \in X$. Для $\varepsilon > 0$ построим замкнутый d -мерный шар $B(x, \varepsilon)$ с центром в точке $x \in X$. Подсчитаем наименьшее число $N(\varepsilon)$ замкнутых шаров радиуса ε , необходимых для покрытия множества F :

$$N(\varepsilon) = \text{наименьшему целому } t, \text{ такому, что } F \subset \bigcup_m B(x_n, \varepsilon) \text{ для } \{x_n | n = 1, 2, \dots, t\} \in X.$$

Можно доказать, что такое число действительно существует [6]. Тогда, если существует скейлинг $N(\varepsilon) \propto \varepsilon^{d_c}$, при $\varepsilon \rightarrow 0$, то d_c называют колмогоровской емкостью множества. Более точно, для двух функций f и g определим отношение $f(\varepsilon) \propto g(\varepsilon)$, если $\lim \{ \ln f(\varepsilon) / \ln g(\varepsilon) \} = 1$ при $\varepsilon \rightarrow 0$. Пусть $V(1)$ – объем d -мерной единичной сферы. Тогда емкость можно определить как такое число d_c , при котором для $\varepsilon \rightarrow 0$ существует отличный от нуля предел $\lim \{ V(1)N(\varepsilon)\varepsilon^d \}$. В этом случае $N(\varepsilon) \propto V(1)\varepsilon^{d_c}$. Поэтому емкость определяется как предел, если он вообще существует, отношения

$$\{ \ln N(\varepsilon) - \ln V(1) \} / \{ \ln(1/\varepsilon) \}.$$

Очевидно, что $\ln V(1) / \ln(1/\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$, поэтому этот член обычно опускают. Заметим, что $d_c \leq d$ и не обязательно целое число! Существуют две теоремы [6], которые позволяют перейти к дискретному варианту для ε . Первая из них утверждает, что емкость компактного подмножества F совпадает с пределом

$$\lim_{n \rightarrow \infty} \{ \ln N(\varepsilon_n) \} / \{ \ln(1/\varepsilon_n) \},$$

где $\varepsilon_n = cr^n$, $0 < r < 1$, $c > 0$, $n = 1, 2, \dots$, так что радиусы шаров можно менять дискретно. Вторая теорема позволяет использовать боксы размером $1/2^n$, и тогда исчезает необходимость нормировки $V(1)$. Если $N_n(F)$ – число таких боксов, пересекающих F , то емкость определяется как предел при $n \rightarrow \infty$:

$$\lim \left\{ \ln N_n(F) \right\} / \left\{ \ln 2^n \right\}.$$

Заметим, что понятие меры, вообще говоря, не связано с метрикой или топологией. Однако, хаусдорфова мера строится исходя из метрики [8, 10]. Тем не менее она пригодна для пространств, которые даже локально не являются евклидовыми. Эта мера принадлежит классу *полуаддитивных* мер, т. е.

$$\mu(A) \leq \sum \mu(A_i) \quad (\cup A_i = A),$$

поэтому, например, построить для нее операцию интегрирования не так просто. Мы опять используем идею покрытия и одно чрезвычайно простое соображение: емкость или объем n -мерного тела легко оценить, взяв его характерный размер l в степени n . Например, площадь квадрата $\propto l^2$, объем куба $\propto l^3$ и т. д. Если n заранее не известно, мы просто перебираем различные степени. «Правильное» значение – это то, для которого выполняется условие полуаддитивности, дополненное требованием «устойчивости» истинного n : произвольные разбиения тела должны давать близкие величины суммарного объема. Характерный размер произвольного множества A в метрическом пространстве (X, d) – это *диаметр* множества: $\text{diam } A = \sup d(x, y)$. Здесь верхняя грань берется по всем $x, y \in A$. Выберем произвольное $\varepsilon > 0$ и назовем ε -покрытием множества A счетное покрытие этого множества замкнутыми шарами: $\{B_i\}$ ($\text{diam } B_i < \varepsilon$). Образуем *крупнозернистую α -мерную меру* для произвольного, не обязательно целого числа $\alpha > 0$:

$$l_\alpha(A, \varepsilon) = \inf \sum (\text{diam } B_i)^\alpha,$$

где нижняя грань берется по всем ε -покрытиям. Когда ε убывает, \inf возрастает, или во всяком случае не убывает. Поэтому предел при $\varepsilon \rightarrow 0$, конечный или бесконечный, существует и называется *α -мерной мерой Хаусдорфа* для A :

$$l_\alpha(A) = \lim_{\varepsilon \rightarrow 0} l_\alpha(A, \varepsilon).$$

Нетрудно убедиться, что функция l_α *монотонна*: $l_\alpha(A) \leq l_\alpha(F)$, если $A \subset F$ и *полуаддитивна*:

$$l_\alpha(\cup_n A_n) \leq \sum l_\alpha(A_n).$$

Хаусдорфова размерность множества A характеризует поведение $l_\alpha(A)$ как функцию от α такую, что если $l_\alpha(A)$ *конечна*, то $l_\beta(A) = 0$ для любого $\beta > \alpha$. Иными словами, если при пробном подсчете «объемов» мы превысили истинное значение размерности A , мера l_α сразу дает нуль. Если $\beta < \alpha$, то при $\varepsilon \rightarrow 0$ и $l_\beta(A) = \infty$. Поэтому существует «точка перехода», соответствующая α^* конечному значению меры Хаусдорфа:

$$l_\alpha(A) = \infty, \forall \alpha < \alpha^* \text{ и } l_\alpha(A) = 0 \forall \alpha > \alpha^* .$$

Это значение α^* и называется *размерностью Хаусдорфа*: $\alpha^* = \dim_H A$. Она однозначно определяется условиями:

$$\dim_H A = \sup \{ \alpha \mid l_\alpha(A) = \infty \} = \inf \{ \alpha \mid l_\alpha(A) = 0 \} .$$

Рассмотрим простой пример – кривую C длины L , покрытие которой заменено равномерным разбиением множества на боксы одинакового размера ε . Очевидно, что необходимое число боксов составляет L/ε и аналогом меры Хаусдорфа служит величина $l_\alpha(C, \varepsilon) = L\varepsilon^{\alpha-1}$. Ясно, что $l_\alpha(C, \varepsilon) \rightarrow \infty$ при $\varepsilon \rightarrow 0$, если $\alpha < 1$, и $l_\alpha(C, \varepsilon) \rightarrow 0$, если $\alpha > 1$. «Правильная» величина размерности $\alpha = 1$ соответствует конечной мере.

Разумеется, компьютерные оценки размерности Хаусдорфа невозможны: нельзя перебрать все покрытия. Именно эта причина вынуждает нас при численных расчетах переходить к емкости, когда все шары данного покрытия одинаковы. В этом случае оптимальным будет конечное покрытие, которое содержит минимальное число слагаемых в выражении для меры. Следует заметить, однако, что d_H для некоторых относительно простых множеств может существенно отличаться от d_c .

Остается еще один неприятный момент, связанный с существованием предела. Если $N(\varepsilon)$ – минимальное число шаров диаметром ε , покрывающих множество, то емкость, по определению, – это предел $\ln N(\varepsilon)/(-\ln \varepsilon)$ при $\varepsilon \rightarrow 0$. Но предел никогда не может быть достигнут в численных экспериментах. Самое большее, на что мы можем рассчитывать – это ограниченность указанной дроби. Здесь можно прибегнуть к следующей уловке. Рассмотрим числовую последовательность диаметров шаров в порядке их убывания: $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k, \dots$. Можно доказать что если максимум отношения $\ln N(\varepsilon)/-\ln \varepsilon$, как функция k , перестает расти задолго до предела, *которого можно достичь при вычислении*, то необходимая оценка емкости существует. В этой связи уместно упомянуть о попытках обобщения определения емкости на бесконечномерные пространства. Разумеется, в этом случае указанный предел просто равен бесконечности. Однако, в ряде случаев число $N(\varepsilon)$ будет расти как функция $1/\exp(\varepsilon^d)$, где d – некоторая константа, которую условно можно считать размерностью [11]. Таким образом, для бесконечномерных пространств под размерностью можно понимать предел при $\varepsilon \rightarrow 0$:

$$\underline{\lim} [-\log \log N(\varepsilon) / \log \varepsilon] .$$

Вместо $N(\varepsilon)$ в этом определении можно использовать число точек ε сети – наименьшее число таких точек, для которых каждая точка пространства удалена хотя бы от одной из них на расстояние не больше ε .

2. Дробные размерности

История дробных размерностей началась с попыток английского физика *Ричардсона* измерить длину побережья Британии [1, 12]. Располагая подробной картой, он аппроксимировал линию побережья ломаной L_b , составленной из b отдельных хорд длиной b , все вершины которой лежали на берегу. *Ричардсон* полагал, что в этом случае существует предел $L_b \rightarrow L$ при $b \rightarrow 0$, как бывает для гладких кривых. Однако оказалось, что с уменьшением b суммарная длина ломаной растет до бесконечности по закону

$$L_b = \lambda b^{1-D}, \quad D < 2.$$

Если построить на каждом звене b квадрат, то суммарная площадь квадратов равна $b^2 N = b^2 L_b / b = \lambda b^{2-D} \rightarrow 0$, при $b \rightarrow 0 \quad D < 2$. Таким образом,

береговая линия имеет бесконечную длину и порождает нулевую площадь! Первое, что приходит в голову, это сомнения в корректности такой аппроксимации. В математике уже давно известны случаи, когда метод хорд не работал. Классический пример привел еще *Лебег*: возьмем равносторонний треугольник ABC и соединим середины трех его сторон (рис. 1). Очевидно,

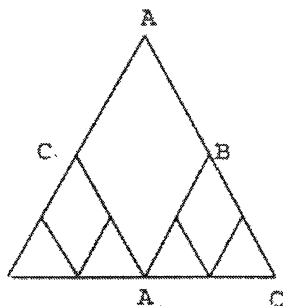


Рис. 1

Продолжая процесс разбиения, мы приближаем ломаную к стороне BC . Если считать последнюю пределом ломаной, получим $AB + AC = BC$. Разумеется, сторона BC не является пределом ломаной – она приближается к BC лишь по положению, но не по направлению. Все дело здесь в степени гладкости. Длина кривой выражается интегралом, под знаком которого стоят производные от уравнения кривой, следовательно, гладкость кривой – по меньшей мере C^1 .

Зубец пилы – это функция $f(x) = |x|$, для которой отношение $[f(h) - f(0)]/h$ при $h \rightarrow 0$ стремится к двум разным пределам: к 1, когда $h > 0$, или к -1 при $h < 0$. Пила не имеет производных в бесконечном числе вершин! Более общий случай, когда точки, в которых отсутствует производная, составляют почти всю область определения функции, приводит к ситуации, которая кажется неприемлемой в рамках гладкого анализа. «С омерзением и ужасом я отворачиваюсь от этой зловердной язвы – непрерывных функций, нигде не имеющих производных», – писал *Эрмит Стильтесу* в 1893 г. Существование таких функций не является тривиальным и с интуитивной точки зрения. Как заметил *А. Пуанкаре*, «надо

быть очень сведущим, чтобы не считать очевидным, что каждая кривая имеет касательную и в самом деле, если представить себе эту кривую и некоторую прямую как две узкие полосы, то всегда можно расположить их так, что они могут иметь общую часть, не пересекаясь» [7]. Для описания этой ситуации формулу линейных приращений $\Delta u = \mu \Delta x + O(\Delta x^2)$ следует заменить на $\Delta u = \mu_H (\Delta x)^H$, где H – показатель Гельдера, а μ и μ_H – обычная и гильдеровская производные. Для гладких функций $H = 1$; в случае $H < 1$ вместо касательной имеется криволинейный конус $\Delta u \approx \Delta x^H$ [13]. Объекты, для которых $H < 1$, а показатель D в формуле Ричардсона строго больше единицы, обычно называют фракталами, хотя это определение не является строгим. Рассмотрим несколько примеров так называемых конструктивных [2] фракталов.

Пример 1. Множество Кантора.

Пусть $F_0 = [0, 1]$. Выбросим из F_0 интервал $(1/3, 2/3)$, а то, что останется, обозначим F_1 . Затем выбросим из F_1 интервалы $(1/9, 2/9)$ и $(7/9, 8/9)$ и получим F_2 (рис. 2).

Продолжая этот процесс, мы придем к убывающей последовательности замкнутых интервалов $\{F_n\}$.

Множество $F = \bigcap F_n$ называют канторовым множеством. Пусть S – множество «выброшенных» кусков отрезка $[0, 1]$ при построении F , т. е. $S = (1/3, 2/3) \cup (1/9, 2/9) \cup (7/9, 8/9) \cup \dots$

Тогда лебегова мера S равна сумме: $1/3 + 2 \times 1/9 + 4 \times 1/27 + \dots = 1$. Таким

образом, при построении F мы выбросили всю длину отрезка! Запишем каждое из чисел $x \in [0, 1]$ в троичной системе:

$$x = a_1/3 + a_2/3^2 + \dots + a_n/3^n + \dots,$$

где $a_n = 0, 1$ или 2 . Легко убедиться, что F принадлежат только те x , которые могут быть записаны хотя бы одним способом так, что в последовательности числителей $a_1, a_2, \dots, a_n, \dots$ ни разу не встречается единица. Совокупность таких последовательностей имеет мощность континуума.

Вычислим емкость канторова множества. При k -й итерации нам необходимо $N = 2^k$ боксов размером $\varepsilon = 3^{-k}$ для покрытия всех фрагментов. По-

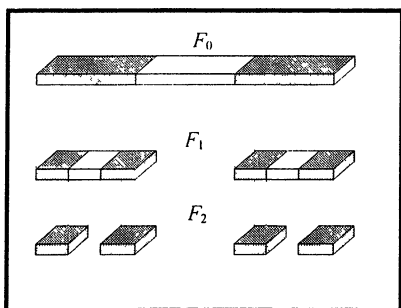


Рис. 2

этому $d_c = \ln 2 / \ln 3 \approx 0,63$, т. е. меньше, чем размерность исходного отрезка. Существует функциональный аналог множества *Кантора* – функция *Кантора*. Распределим равномерно на F единичную массу (меру) с плотностью μ . Тогда функция

$$F(x) = \int_0^x d\mu(x)$$

описывает распределение меры на канторовом носителе. Она является непрерывной возрастающей функцией, которая, тем не менее, почти всюду горизонтальна! Ее называют иногда «дьявольской лестницей» (рис. 3).

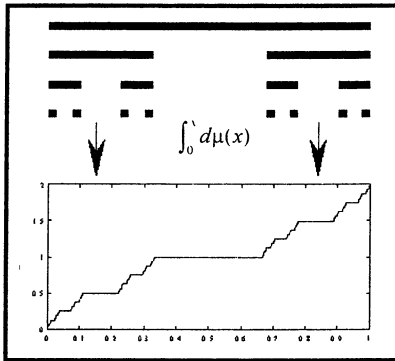


Рис. 3

Пример 2. Кривая Кох.

Возьмем равносторонний треугольник и сделаем следующую элементарную операцию: каждая сторона делится на 3 части, после чего средний сегмент заменяется на равносторонний треугольник. Операция повторяется n раз. То, что получится при $n \rightarrow \infty$, и есть триада Кох (рис. 4).

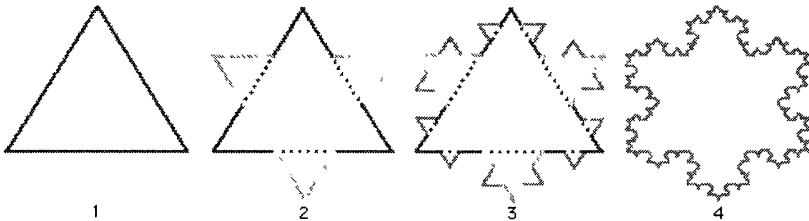


Рис. 4

Это множество самоподобно, точно так же как и сам процесс его построения. На любом шаге его линейный элемент длины l заменяется на

$N = 4$ элемента длиной $lr^{-1} = x_1, x_2$ каждый. Поэтому емкость $d_c = \ln N / \ln r = \ln 4 / \ln 3 \approx 1,2618$. Триада Кох является математической моделью кривой побережья, с которой работал Ричардсон.

Пример 3. Функция Вейерштрасса.

Рассмотрим непрерывную функцию $f(x)$, которая удовлетворяет функциональному уравнению $f(x) = af(bx) + \cos x$ при $0 < a < 1$; $ab \geq 1$. С помощью итераций легко найти:

$$f_0(x) = \cos x$$

$$f_1(x) = a \cos(bx) + \cos x$$

$$f_2(x) = a^2 \cos b^2 x + a \cos bx + \cos x$$

.....

$$f_{n+1}(x) = af_n(bx) + \cos x.$$

Поэтому

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n x)$$

– это и есть функция Вейерштрасса [5, 14]. Ряд равномерно сходится в любом интервале, так что $f(x)$ всюду непрерывна. Если $ab > 1$, то ряд, полученный почленным дифференцированием суммы, расходится. Можно доказать, что при $ab > 1 + 3\pi/2$ для всех значений x функция $f(x)$ не имеет производной [14].

Мандельброт указал, что функция Вейерштрасса фрактальна: ее график – кривая, имеющая размерность больше единицы. Фрактальность означает отсутствие *наименьшего* масштаба [1], но функция Вейерштрасса имеет *наибольший* масштаб $ab \geq 1$, а это не всегда удобно для моделей. Поэтому в [14] было предложено следующее обобщение:

$$W(t) = \sum \left[1 - \exp(i\gamma^n t) \right] \exp(i\phi_n t) \gamma^{(D-2)n}, \quad 1 < D < 2, \quad \gamma > 1.$$

Частоты $\{\gamma^n\}$ образуют геометрическую прогрессию в диапазоне $[0, \infty]$, а выбор фазы ϕ_n определяет детерминированное или стохастическое поведение $W(t)$. Фрактальная природа $W(t)$ проявляется в том, что при увеличении масштаба структура графика в точности повторяется. Пусть $\phi_n = kn$. Легко проверить, что $W(t)$ аффинноподобна [1]:

$$W(\gamma t) = \exp(-ik) \gamma^{2-D} W(t).$$

В приложениях рассматривают обычно ряд косинусов $\text{Re} W(t)$ при $k = 0$ или альтернированный ряд синусов $\text{Im} W(t)$ при $k = \pi$. Если $\{\phi_n\}$ – мно-

жество случайных чисел, $W(t)$ ведет себя как стохастическая нестационарная функция. Тем не менее свойства самоподобия (скейлинг) сохраняются, но уже не для самой $W(t)$, а для ее среднеквадратических приращений, которые не зависят от t . Для всех

$$D \langle |W(t)|^2 \rangle = V(t)$$

удовлетворяет скейлингу:

$$V(\gamma\tau) = \gamma^{4-2D} V(\tau).$$

Идея первого доказательства фрактальности $W(t)$ принадлежит *Ори* и основана на потенциальном определении фрактальной размерности. Он рассмотрел положительные электрические заряды, распределенные в соответствии с графиком $\text{Re}W(t)$. Тогда значение энергии на единицу t определяется модифицированным потенциалом Кулона $\propto r^{-\alpha}$, где α не обязательно равно единице. Энергия E рассматривается как функция $E = E(\alpha)$, а фрактальная размерность определяется как наибольшее α , для которого $E(\alpha) \neq \infty$ [14]. Наибольший интерес представляют три значения D . Для $D = 1,5$ (рис. 5; 1000 точек, $\phi_n = \text{const}$) функция $W(t)$ моделирует броуновское блуждание частицы вдоль оси W .

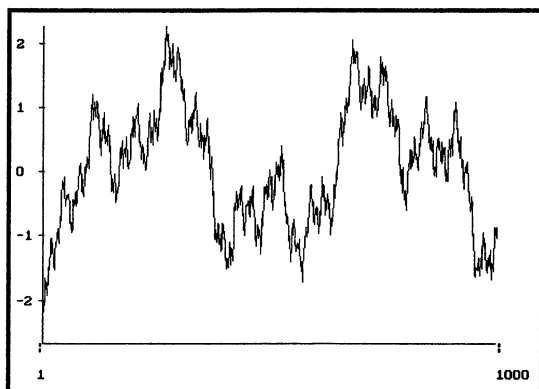


Рис. 5

В пределе $D \rightarrow 2$ график $W(t)$ почти замечает плоскость, демонстрируя модель $1/f$ -шума. Наконец, при $D = 1$ это модель квазикристаллической поверхности.

3. Фракталы, аттракторы и мультифрактальные меры

Описанный выше способ построения функции *Вейерштрасса* является частным случаем результатов *де Рама* [15], который в 1957 году исследовал решения функционального уравнения:

$$f(x) = F\{f(x)\} + g(x),$$

где $g(x)$ – заданная функция, а F – сжимающее отображение, определенное в пространстве ограниченных функций на интервале $[0, 1]$ и обладающее свойством $F(0) = 0$. Сжатие означает выполнение неравенства $\|Ff_1 - Ff_2\| \leq \lambda \|f_1 - f_2\|$ для некоторого $0 < \lambda < 1$ любых двух функций f_1, f_2 и супремум-нормы $\|f\| \equiv \sup_{[0,1]} |f(x)|$. Существование единственного решения уравнения *де Рама* следует из теоремы *Банаха* о сжимающем отображении и того факта, что $f \rightarrow Ff + g$ – сжатие. Если положить [13]

$$F_{\alpha,\beta}f(x) \equiv \begin{cases} \alpha f(2x), & x \in [0, 1/2], \\ \beta f(2x-1), & x \in (1/2, 1], \end{cases}$$

то функция *Вейерштрасса* получается при $\alpha = \beta = a$, $g(x) = \cos \pi x$. Идея построения произвольного фрактала с помощью системы сжимающих итераций была предложена *Хатчинсоном* [4, 6]. *Системой итеративных функций* (СИФ) $(X, \{f_i\})$, $i = 1, 2, \dots, k$, называют набор сжимающих отображений

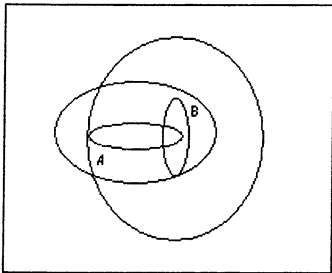


Рис. 6

ний $\{f_i\} : d(f_i(x), f_i(y)) \leq s_i d(x, y)$ в полном метрическом пространстве (X, d) . Коэффициентом сжатия СИФ называют $s = \max\{s_i : i = \overline{1, k}\}$. Определим для (X, d) полное метрическое пространство «фракталов» $(H(X), h)$, где $H(X)$ – множество непустых компактных подмножеств X и h – метрика Хаусдорфа. Точнее, пусть $B(x, r)$ – замкнутый шар радиуса r с центром в точке x . Для произвольного множества

$A \in X$ дилатацией A_r радиуса r множества A называется $\bigcup_{x \in A} B(x, r)$. Таким образом, дилатация множества A – это добавление к A всех точек, лежащих на расстоянии не более r от его границы. Пусть A, B – непустые компактные подмножества из X . Тогда расстояние (метрика) Хаусдорфа определяется как (рис. 6)

$$h(A, B) = \min \{r > 0 \mid A \subset B_r; B \subset A_r\}.$$

Таким образом, это минимальное из двух чисел: первое из них получается расширением множества A до тех пор, пока его образ не поглотит B , второе – дилатацией B , пока она не поглотит A .

Для любого $B \in H(X)$ определим оператор Хатчинсона $T: H(X) \rightarrow H(X)$

как $T(B) = \bigcup_{i=1}^k f_i(B)$, где $T(B) = \{T(x) \mid x \in B\}$, и его композицию $T^{\circ n}$ порядка

n : $T^{\circ 0}(B) = B$, $T^{\circ(i+1)}(B) = T^{\circ 1}(T(B))$. Последовательность множеств, полученная в результате итерирования $T(B)$, т. е. $\{B, T(B), T^{\circ 2}(B), \dots, T^{\circ n}(B), \dots\}$,

называется *орбитой* B для $(H(X), T)$. Пару $(H(X), T)$ можно рассматривать как *детерминированную дискретную динамическую систему* с пространством состояний $H(X)$ и преобразованием T .

Согласно *теореме Банаха* о неподвижной точке, для любой начальной точки $B_0 \in H$ последовательность $B_0, B_1 = T(B_0), B_2 = T(B_1) \equiv T^{\circ 2}(B_0), \dots$ сходится к единственной неподвижной точке $A \in H$, которая является решением уравнения $T(A) = A$. Очевидно, что $\lim_{n \rightarrow \infty} T^{\circ n}(B) = A$. Неподвижная точка A называется *аттрактором* СИФ,

или *фракталом*.

Рассмотрим простой пример. Пусть $X = [0, 1]$, $f_1 = 1/3x$, $f_2 = 1/3x + 2/3$.

Первая из функций сжимает единичный отрезок в 3 раза, вторая делает то же самое и сдвигает полученный образ на 2/3. Рассмотрим их объединение: $T = f_1 \cup f_2$. Последовательность итераций $T(X)$ при $n \rightarrow \infty$ очевидно приводит к канторову множеству $F = \lim_{n \rightarrow \infty} T^{\circ n}[0, 1]$. Таким образом, фрактал всегда является предельным образом динамики сжатия. Существуют глубокие аналогии между динамикой СИФ и диссипативными динамическими системами и их аттракторами [4]. Однако эта тема выходит за рамки статьи. Если в последнем примере выбирать f_1 с вероятностью p_1 , а f_2 с вероятностью p_2 , то получим канторово множество в «черно-белосерых» тонах, где более темным фрагментам соответствует область применения отображения с большей вероятностью. Цветовая гамма повторяется при переходе к более тонкому разрешению. Таким образом, раскраска фрактала становится самоподобной – ее называют обычно мультифрактальной мерой. Обычное канторово множество также можно рассматривать как «одноцветную» меру – «краску», нанесенную на отрезок в точках, принадлежащих F . Мультифрактал соответствует множеству одноцветных фракталов, окрашенных с разной, но масштабно согласован-

ной интенсивностью. Для описания таких мер используют мультифрактальный формализм, существующий в нескольких вариантах [1, 3, 5]. Ограничимся здесь лишь одним из них.

Рассмотрим один грамм краски (меру), нанесенный на отрезок (носитель меры) произвольным образом. Разделим носитель меры на равные интервалы (боксы) C_δ размером δ . Пусть $C = \{C_\delta\}$ – множество всех таких боксов. Предположим, что мера внутри бокса аппроксимируется степенным законом:

$$\mu(C_\delta) \approx \delta^{\alpha(C_\delta)},$$

где величины

$$\alpha(C_\delta) = \frac{\log \mu(C_\delta)}{\log \delta}$$

называются *крупнозернистыми гельдеровскими показателями сингулярности меры*.

Пусть $N_\delta(\alpha, \varepsilon) = \#\{C_\delta : \alpha(C_\delta) \in (\alpha - \varepsilon, \alpha + \varepsilon)\}$, где $\#$ – число непустых боксов, содержащих меру с показателем $\alpha \pm \varepsilon$. Тогда *крупнозернистый мультифрактальный спектр больших отклонений* [18, 19] определяется выражением

$$f_g(\alpha) = \lim_{\varepsilon \rightarrow 0} \limsup_{\delta \rightarrow 0} \frac{\log N_\delta(\alpha, \varepsilon)}{\log(1/\delta)}.$$

Последнее выражение *по форме* совпадает с определением емкости множества, образованного боксами с выделенным значением показателя сингулярности меры. Термин *крупнозернистый* объясняется конечной точностью оценки α . Форма определения $f_g(\alpha)$ связана с теоремой о больших отклонениях, позволяющей дать вероятностную интерпретацию спектра [18]. А именно, вероятность найти $\alpha(C_\delta) \approx \alpha$ ведет себя приблизительно как $N_\delta(\alpha, \varepsilon)/N_\delta = P_\delta[\alpha(C_\delta) \approx \alpha] \approx \delta^{d_T - f_g(\alpha)}$, где N_δ – общее число боксов, содержащих меру, а d_T – топологическая размерность носителя. Таким образом, пара $\{\alpha, f_g(\alpha)\}$ описывает структуру произвольной меры. Если мера обладает самоподобными свойствами (мультифрактальна), то ее спектр $f_g(\alpha)$ представляет собой выпуклую кривую.

Литература

- 1 Федер Е Фракталы М Мир, 1991
- 2 Морозов А Д Введение в теорию фракталов Н Новгород Изд-во ННГУ, 1999
- 3 Божокин С В, Паршин Д А Фракталы и мультифракталы Москва–Ижевск РХД, 2001

- 4 *Кроновер Р М* Фракталы и хаос в динамических системах М Постмаркет, 2000
- 5 *Falconer К* Fractal Geometry Mathematical Foundations and Applications John Wiley & Sons, 1990
- 6 *Barnsly М* Fractals Everywhere Orlando Academic Press, 1988
- 7 *Пуанкаре А* О Науке М Наука, 1983 559 с
- 8 *Горелик Г Е* Размерность пространства. М МГУ, 1983
- 9 *Лебег А* Об измерении величин М ГУПИМП, 1960.
- 10 *Билленгслей П* Эргодическая теория и информация М Мир, 1969
- 11 *Арнольд В И* О представлении функций нескольких переменных в виде суперпозиции функций меньшего числа // Математическое просвещение 1958 Вып 3 С 41–61
- 12 *Баренблатт Г И* Подобие, автомодельность, промежуточная асимптотика Л Гидрометеиздат, 1982
- 13 *Зельдович Я Б, Соколов Д Д* Фракталы, подобие, промежуточная асимптотика // УФН 1985 Т 146, № 3 С 493–506
- 14 *Berry М V, Lewis Z V* On the Weierstrass-Mandelbrot fractal function // Proc Roy Soc London Ser A 1980 370, № 1743 P 459–484
- 15 *Tasaki S, Gilbert T, Dorfman J R* An Analytical Construction of the SRB Measures for Baker-type Maps – available [http // babbage sissa it/chaos-dyn/9801031](http://babbage.sissa.it/chaos-dyn/9801031)
- 16 *Макаренко Н Г* Заметки о фракталах – available [http //www keldysh ru/dpt_17/works/mak/index htm](http://www.keldysh.ru/dpt_17/works/mak/index.htm)
- 17 *Макаренко Н Г* Фракталы, аттракторы, нейронные сети и все такое // Лекции по нейроинформатике Ч 2 // Нейроинформатика–2002 IV Всерос науч-техн конф М МИФИ, 2002 С 121–169
- 18 *Riedi R, Scheuring I* Conditional and relative multifractal spectra // Fractals 1997 5, № 1 P 153–168
- 19 *Canus Ch, Vehel J, Tricot C* Continuous large deviation multifractal spectrum definition and estimation – available [http //www-rocq inria fr/fractales](http://www-rocq.inria.fr/fractales)

КРИТИЧЕСКОЕ ПОВЕДЕНИЕ ПРИ ПЕРЕХОДЕ К ХАОСУ ЧЕРЕЗ УДВОЕНИЯ ПЕРИОДА. МОДЕЛЬНЫЕ ОТОБРАЖЕНИЯ И РЕНОРМГРУППОВОЙ АНАЛИЗ

А. П. Кузнецов, С. П. Кузнецов, И. Р. Сатаев

1. Введение

Возможность описания турбулентности как динамического процесса в пространственно распределенных системах давно привлекает большое внимание исследователей. Один из важных вопросов — как возникает пространственно-временной хаос из простого регулярного режима при изменении управляющих параметров?

Для систем с малой размерностью фазового пространства прорыв в понимании перехода к хаосу в нелинейных системах связан с открытием Фейгенбаумом универсальности удвоений периода и развитием ренормгруппового (РГ) подхода. Простейший класс систем, демонстрирующих фейгенбаумовский тип поведения, — это одномерные (1D) необратимые отображения, однако тот же самый тип перехода к хаосу встречается во многих диссипативных нелинейных системах разной физической природы [1, 2, 3, 4].

Предположим, что имеется некоторая нелинейная система, поведением которой можно управлять путем изменения нескольких параметров. Фейгенбаумовское поведение на пороге хаоса является феноменом коразмерности один. Например, если мы имеем систему с тремя управляющими параметрами, бифуркации удвоения периода будут происходить на некоторых двумерных поверхностях в трехмерном пространстве параметров, и их последовательность будет накапливаться к границе хаоса — фейгенбаумовской критической поверхности.

Можно предположить, что при движении вдоль границы хаоса в пространстве параметров мы можем встретить нефейгенбаумовское поведение на некоторых критических линиях (феномены коразмерности два) или в некоторых критических точках (феномены коразмерности три). Естественно ожидать, что выявленные таким образом типы поведения также должны допускать РГ-анализ и проявлять специфические свойства количественной универсальности и скейлинга. В этом случае мы говорим об определенном типе критического поведения (или типе критичности).

Одна возможность возникновения нефейгенбаумовского критического поведения, ассоциирующегося с удвоениями периода, состоит в том, что динамика на пороге хаоса остается по существу одномерной, но соответствующее 1D-отображение возмущено настолько, что покидает фейгенбаумовский класс универсальности. Другая возможность заключается в выходе некоторой новой моды на порог устойчивости, что ведет к увеличению эффективной размерности динамики.

В статье представлена коллекция типов критического поведения, которые могут иметь место при многопараметрическом анализе перехода к хаосу. Для каждого типа дается сводка основных универсальных констант, а также простой пример — одно- или двумерное отображение, являющееся представителем данного класса универсальности¹.

2. Критическое поведение в одномерных отображениях

Рассмотрим сначала типы критичности, ассоциирующиеся с одномерной динамикой на пороге хаоса. Каждый из них будет отвечать некоторому решению РГ-уравнения Фейгенбаума — Цвитановича [1]:

$$g(x) = \alpha g[g(x/\alpha)], \quad (1)$$

где α — универсальная постоянная (фактор скейлинга), специфическая для данного типа критичности. Решить уравнение (1) означает найти одновременно и функцию $g(x)$ (обычно численно, в виде полиномиального разложения), и константу α . О функции $g(x)$ говорят как о неподвижной точке РГ-уравнения.

Фактически, функция $g(x)$ представляет асимптотическую форму надлежащим образом нормированного оператора эволюции для 2^k -итераций исходного отображения точно в критической точке. Существование предела при $k \rightarrow \infty$ обеспечивается перенормировкой динамической переменной $x \sim \alpha^{-k}$.

Следующий шаг состоит в анализе малых возмущений неподвижной точки РГ-уравнения. Предположим, что в исходном отображении мы немного изменяем какой-либо из параметров и сдвигаемся тем самым из критической точки. Теперь оператор эволюции, определенный на 2^k -итерациях, будет содержать некоторое возмущение функции $g(x)$, и его поведение можно изучать с помощью линеаризованно-

¹Полученные численно решения РГ-уравнений в виде достаточно громоздких полиномиальных аппроксимаций полностью в тексте не приводятся; их можно найти на нашем веб-сайте www.sgtnd.tserv.ru [5] и в работе [6].

го уравнения РГ. Это приводит к следующей задаче на собственные значения [1]:

$$\nu u(x) = \alpha [g'(g(x/a))u(x/a) + u(g(x/a))]. \quad (2)$$

Решения уравнения (2) представляют собой моды возмущения неподвижной точки $g(x)$. Существенными будут моды с собственными значениями $|\nu| > 1$, определяющие асимптотику при больших k , причем из их числа следует исключить те, которые ассоциируются с инфинитезимальными заменами переменных. Число n существенных собственных векторов определяет *коразмерность* типа критичности: это минимальное количество управляющих параметров, при котором данная критическая ситуация может встретиться в некоторой точке пространства параметров. (В самом деле, требуя, чтобы коэффициенты при n существенных модах обратились в нуль, мы задаем точно n условий на параметры исходного отображения.)

В выбранной надлежащим образом локальной системе координат («скейлинговые координаты») n -мерное пространство параметров будет иметь универсальную топографию, специфическую для каждого типа критичности.

Скейлинговые координаты определяются так, чтобы сдвиг из критической точки вдоль каждой координатной оси в пространстве параметров приводил к появлению возмущения, содержащего одну существенную собственную моду линеаризованного уравнения РГ.

Существенные собственные числа $\nu = \delta_1, \dots, \delta_n$ являются тогда *факторами скейлинга*: если использовать их как факторы увеличения вдоль соответствующих координатных осей, то будет наблюдаться повторение структуры топографии пространства параметров во все меньшей и меньшей окрестности критической точки.

Для коразмерностей выше единицы задача нахождения в явном виде скейлинговых координат нетривиальна, причем форма соответствующей замены зависит от конкретных соотношений между существенными собственными числами [6].

Предположим, мы нашли критическую точку коразмерности два, и $\delta_1 > \delta_2^k > 1$ при $k = 1, \dots, K$, но $\delta_2^{K+1} > \delta_1$. Тогда выражения для двух управляющих параметров через скейлинговые координаты (C_1, C_2) должны содержать члены $C_1, C_2, C_2^2, \dots, C_2^K$. Только при $K = 1$ достаточно линейной замены координат. В случае коразмерности три следует учесть члены вида $C_2^{k_1} C_3^{k_2}$, если выполняется неравенство $\delta_2^{k_1} \delta_3^{k_2} < \delta_1$. (Все сказанное выше остается справедливым и для 2D-отображений.)

Важной количественной характеристикой типа критичности является *универсальный мультипликатор*. Он может быть получен как

производная от функции $g(x)$ в неподвижной точке $x_* = g(x_*)$: $\mu_* = g'(x_*)$. Для конкретных систем это значение мультипликатора реализуется в критической точке для циклов периода 2^k в асимптотике больших k . Данное свойство полезно для точного численного нахождения положения критической точки в пространстве параметров. Для этого отыскиваем такие значения управляющих параметров, при которых мультипликаторы равны μ_* для соответствующего числа орбит с достаточно долгими периодами 2^k .

Фейгенбаумовская критичность (тип F) является наиболее известным типом поведения при удвоениях периода. Он связан с неподвижной точкой РГ-уравнения (1): $g(x) = 1 - 1,5276x^2 + 0,1048x^4 + \dots$, причем $\alpha = -2,5029078750959$ [1]. Линеаризованный оператор РГ (2) имеет в спектре одно существенное собственное число $\nu = \delta = 4,669201609103$, т. е. коразмерность равна единице: $\text{CoDim}_F = 1$.

Классическим примером фейгенбаумовской критичности является поведение логистического отображения

$$x_{n+1} = 1 - \lambda x_n^2 \quad (3)$$

в критической точке $\lambda_c = 1,40115518909205$ [2]. Здесь отображение имеет бесконечное счетное множество неустойчивых циклов периода 2^k . Асимптотически, при $k \rightarrow \infty$, эти циклы характеризуются одним и тем же универсальным мультипликатором $\mu = -1,6011913280736$. Аттрактор в критической точке представляет собой фрактальное множество, размерность Хаусдорфа которого равна $D_0 = 0,5380451435805$ [7]. Локально, в окрестности экстремума $x = 0$, структура аттрактора воспроизводится при изменении масштаба переменной x в α раз.

Топография пространства параметров вблизи фейгенбаумовской критической точки содержит при $\lambda < \lambda_c$ последовательность областей устойчивой регулярной динамики с периодами 2^k и сложно устроенное множество периодических «окон» и хаотических режимов в за-критической области $\lambda > \lambda_c$. Эта структура демонстрирует свойства скейлинга: в малых масштабах топография пространства параметров воспроизводит себя при увеличении в δ раз [1–4].

Трикритическое поведение (тип T) ассоциируется с решением РГ-уравнения, которое представляется в виде разложения по степеням x^4 : $g(x) = 1 - 1,8341x^4 + 0,01296x^8 + 0,3119x^{12} + \dots$, при этом $\alpha = -1,6903029714$ [8, 9]. Линеаризованный оператор РГ (2) имеет в спектре три существенных собственных числа $\delta = 7,284686217$, α^2 , α^3 [8, 10, 11]. Это означает, что коразмерность равна трем: $\text{CoDim}_T = 3$.

В самом деле, сколько условий должно быть удовлетворено, чтобы найти трикритическую точку в одномерном отображении? Во-первых,

два условия необходимы, чтобы иметь экстремум четвертого порядка: вторая и третья производные должны обращаться в нуль в этой точке. Еще один параметр нужен, чтобы прийти в критическую точку через каскад удвоений периода при сохранении двух предыдущих условий.

В качестве конкретного примера возьмем произвольное гладкое одномерное отображение, содержащее в разложении Тейлора члены до четвертого порядка включительно. Устраним кубический член сдвигом начала координат и пересчитаем масштаб динамической переменной так, чтобы нормировать постоянный член на единицу. В результате получим

$$x_{n+1} = 1 - Ax_n^2 - Bx_n^4 - Cx_n, \quad (4)$$

где A, B, C — параметры.

При $A = 0$ и $C = 0$ это отображение демонстрирует с увеличением B каскад удвоений периода, накапливающийся к трикритической точке $B_c = 1,594901356229$ [8]. В ней существует бесконечное множество неустойчивых 2^k -циклов, универсальное асимптотическое значение мультипликатора равно $\mu = -2,05094049$. Размерность Хаусдорфа трикритического аттрактора $D_0 = 0,642575065$.

Вопрос о коразмерности трикритичности в одномерных отображениях не столь тривиален, как может показаться из приведенного обсуждения. Дело в том, что экстремум четвертого порядка может появиться не в исходном отображении, а в его итерациях, а трикритичность возникает как феномен коразмерности два [12]. Пусть имеем гладкое 1D-отображение с двумя квадратичными экстремумами и двумя управляющими параметрами. В общем случае в пространстве параметров может существовать кривая, на которой выполнено условие: один экстремум отображается точно во второй. Если мы находимся на этой кривой, то дважды итерированное отображение будет иметь экстремум четвертого порядка. Если вдоль этой кривой имеет место каскад удвоений периода, то пределом его будет трикритическая точка. Можно показать, что в этом случае собственное число α^3 исключается (своего рода *скрытая симметрия*, см. [10]), и только $\delta_1 = \delta$ и $\delta_2 = \alpha^2$ являются существенными. Например, отображение

$$x_{n+1} = A - Bx_n + x_n^3 \quad (5)$$

имеет трикритическую точку при $A_c = 0,2426987573$, а также при $B_c = 1,9513857778$.

Топография пространства параметров в окрестности трикритической точки коразмерности два приведена на рис. 1 для модельного отображения (5). Видно, что в скейлинговых координатах картинки выглядят подобными друг другу при пересчете масштабов.

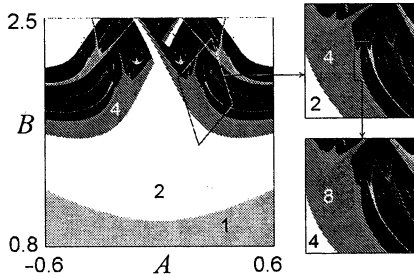


Рис. 1. Топография плоскости параметров для модельного отображения (5). На этом и следующих рисунках области различных периодических режимов обозначены тонами серого цвета, черным — области неперiodической динамики, белым — расходимость итераций модельного отображения. На нижних диаграммах трикритическая точка расположена в центре, и используются скейлинговые координаты: $A - A_c = -0,59986C_1 + 0,21928C_2$, $B - B_c = C_1 + C_2$.

В диссипативных системах, не сводящихся к 1D-отображению точно, трикритичность коразмерности два не встречается: в игру неизбежно вступает третья собственная мода линеаризованного РГ-уравнения. Таким образом, чтобы наблюдать трикритичность, необходимо все же иметь три параметра [10].

Критичность S-типа («Six-power») ассоциируется с решением РГ-уравнения, представленным разложением по степеням x^6 : $g(x) = 1 - 1,9077x^6 - 0.3329x^{12} + \dots$, причем $\alpha = -1,4677424503$ [6, 9]. Линеаризованный оператор РГ (2) имеет в спектре три существенных собственных числа $\delta = 9,296246833$, α^2 , α^3 , α^4 , α^5 , т. е. $\text{CoDim}_S = 5$. Действительно, чтобы иметь экстремум шестого порядка необходимо выполнить пять дополнительных условий: 2-я, 3-я, 4-я и 5-я производные должны быть равны нулю в одной и той же точке, а пятый параметр нужен, чтобы реализовать накопление бифуркаций удвоения периода.

Однако в 1D-отображениях критичность S-типа может появляться как феномен коразмерности три [13]. Предположим, что отображение имеет квадратичный экстремум и кубическую точку перегиба, и одна из этих точек отображается точно во вторую. Тогда двукратно итерированное отображение будет иметь экстремум шестого порядка. В трехмерном пространстве параметров может существовать кривая, на которой выполняется указанное условие. Если вдоль этой кривой происходит каскад удвоений периода, точка накопления будет критической точкой S-типа.

В модельном отображении (7) при условии отображения экстремума в точку перегиба находим критическую точку S_1 : $A_c = 1,8724481923$, $B_c = -1,6252052847$, $C_c = 1,0940161015$. Если, наоборот, точка перегиба отображается в точку экстремума, получаем критическую точку S_2 : $A_c = 1,3799094808$, $B_c = -0,5574097012$, $C_c = 1,1818211223$. В обоих случаях существенными являются три собственных числа $\delta_1 = \delta$, $\delta_2 = \alpha^4$, $\delta_3 = \alpha^2$ для S_1 , и $\delta_1 = \delta$, $\delta_2 = \alpha^3$, $\delta_3 = \alpha^2$ для S_2 .

В критической точке S существуют неустойчивые циклы всех периодов 2^k , универсальное асимптотическое значение мультипликатора равно $\mu = -2,32150547$. Размерность Хаусдорфа критического аттрактора $D_0 = 0,683433256$.

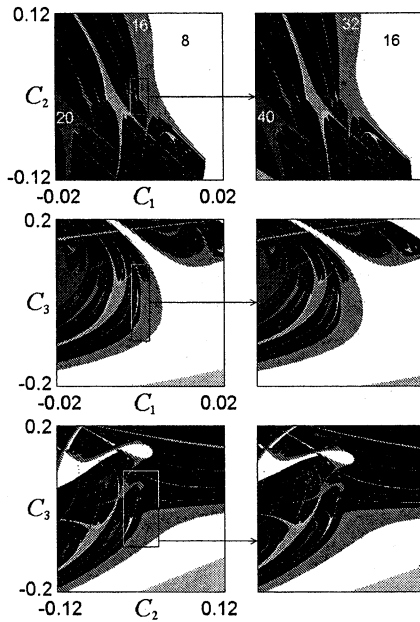


Рис. 2. Топография сечений пространства параметров вблизи критической точки S_1 для отображения (7) в скейлинговых координатах: $A - A_c = -0,5285(C_1 + 0,7803C_3^2) - 0,8097C_2 + C_3$, $B - B_c = (C_1 + 0,7803C_3^2) + C_2 - 0,8488C_3$, $C - C_c = -0,1267(C_1 + 0,7803C_3^2) + 0,0822C_2 + 0,5907C_3$. Критическая точка расположена в центре каждого рисунка.

Для того чтобы выписать соотношения для скейлинговых координат в пространстве параметров в окрестности критической точки S , следует учесть некоторые нелинейные члены. На рис. 2 показаны

карты режимов в сечении пространства параметров координатными плоскостями (C_i, C_j) и проиллюстрированы свойства скейлинга вблизи критической точки S_1 .

Критичность E -типа («eight-power») ассоциируется с решением РГ-уравнения, представляющимся в виде разложения по степеням x^8 : $g(x) = 1 - 1,8973x^8 - 0,7388x^{16} + \dots$, при этом $\alpha = -1,358017279$ [9, 13].

Линеаризованный оператор РГ (2) имеет в спектре семь существенных собственных чисел $\delta = 10,94862427, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6, \alpha^7$, т. е. $\text{CoDim}_E=7$. Происхождение этого числа может быть объяснено так же, как это было сделано для типов T и S . Тем не менее поведение E -типа встречается в 1D-отображениях как феномен коразмерности три [13]. В самом деле, если 1D-отображение имеет три квадратичных экстремума $x_1, 2, 3$, причем x_1 отображается в x_2 , и $x_2 \rightarrow x_3$, то проитерированное отображение будет иметь экстремум восьмого порядка. В трехмерном пространстве параметров это условие выполняется в общем случае на некоторой кривой. Если вдоль этой кривой происходит каскад удвоений периода, точка накопления будет критической точкой E -типа. Только три из семи собственных чисел являются существенными в этом случае: $\delta_1 = \delta, \delta_2 = \alpha^4, \delta_3 = \alpha^2$. Для модельного отображения (7) такая критическая точка расположена в $A_c = 2,4493669341, B_c = -1,2604157306, C_c = 0,7009546250$.

Универсальный мультипликатор для циклов большого периода 2^k равен $\mu = -2,51408835$. Размерность Хаусдорфа критического аттрактора $D_0 = 0,707102082$.

На рис. 3 показаны сечения пространства параметров координатными плоскостями (C_i, C_j) и продемонстрированы их скейлинговые свойства.

Итак, из всех обсуждаемых выше критических ситуаций только классический фейгенбаумовский тип имеет всегда одну и ту же коразмерность один – и для 1D-отображений, и в общем случае. Для других типов критичности вопрос оказывается более тонким. Тип T для 1D-отображений может быть обнаружен в двухпараметрических семействах, а типы S и E – в трехпараметрических, хотя коразмерность, вычисленная с помощью РГ-анализа, выше. Для отображений большей размерности или для систем дифференциальных уравнений, типы T, S , и E будут появляться как феномены коразмерности 3, 5, и 7 соответственно.

Тем не менее было обнаружено, что при двухпараметрическом анализе реалистичных систем трикритические скейлинговые свойства могут оставаться справедливыми вплоть до удивительно глубоких

уровней удвоения периода («псевдо-трикритическое поведение») [11, 14]. В некотором смысле трикритическое поведение выступает как промежуточная асимптотика, и может оказаться очень трудно, если вообще возможно, наблюдать заметные отличия от истинной трикритичности в эксперименте или при непосредственном численном моделировании. По-видимому, в том же самом смысле, т. е. в качестве промежуточной асимптотики, при трехпараметрическом анализе реальных нелинейных диссипативных систем могут обнаруживаться также типы скейлингового поведения E и S .

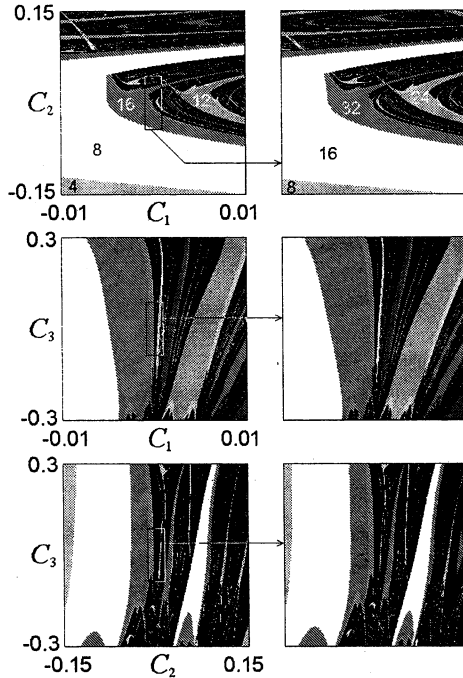


Рис. 3. Топография сечений пространства параметров вблизи критической точки E -типа для отображения (7) в скейлинговых координатах $A - A_c = (C_1 - 0, 2006C_3^2 + 0, 3443C_2C_3 - 0, 1384C_3^3) + C_2 - 0, 9456C_3, B - B_c = -0, 5597 \times (C_1 - 0, 2006C_3^2 + 0, 3443C_2C_3 - 0, 1384C_3^3) - 0, 4571C_2 + C_3, C - C_c = -0, 3388 \times (C_1 - 0, 2006C_3^2 + 0, 3443C_2C_3 - 0, 1384C_3^3) + 0, 3176C_2 - 0, 0065C_3$. Критическая точка расположена в центре каждого рисунка.

Перейдем теперь к типам критического поведения, которые возникают вследствие наличия дополнительного измерения в фазовом пространстве и не встречаются в 1D-отображениях.

3. Системы с однонаправленной связью

Обратимся сначала к специальному классу систем. Предположим, что система допускает разбиение на две подсистемы с однонаправленной связью, так что модельное отображение представляется в виде $x_n + 1 = F_1(x_n)$, $y_n + 1 = F_2(x_n, y_n)$. Поскольку однонаправленная связь оказывается физически реализуемой [11, 15], это оправдывает специальный интерес к исследованию такого рода систем.

РГ-анализ приводит к следующей системе уравнений для пары функций g и f [16, 17]:

$$g(x) = \alpha g(g(x/\alpha)), \quad (6)$$

$$f(x, y) = \beta f(g(x/\alpha), f(x/\alpha, y/\beta)). \quad (7)$$

Первое соотношение совпадает с уравнением Фейгенбаума — Цвитановича, так что проблема состоит в анализе второго уравнения. Функция $f(x, y)$ и орбитальный фактор скейлинга β должны быть получены из этого уравнения, в то время как функция $g(x)$ известна.

Вследствие двумерности системы критическое поведение характеризуется двумя универсальными мультипликаторами, каждый из которых, в силу однонаправленности связи, ассоциируется с одной из подсистем. Они равны производным от функций $g(x)$ и $f(x, y)$ в неподвижной точке $x_* = g(x_*)$, $y_* = f(x_*, y_*)$: $\mu_1 = g'(x_*)$, $\mu_2 = f'_y(x_*, y_*)$.

Для неподвижной точки РГ-уравнений (6) и (7) мы можем рассмотреть класс возмущений, которые не нарушают однонаправленность связи. Соответствующая задача на собственные числа может быть решена раздельно для двух подпространств: в присутствии и в отсутствие возмущений первой подсистемы. Собственные числа для возмущений первого типа получаются из уравнения (2). Во втором случае справедливо уравнение [16]:

$$\nu v(x, y) = \beta [f'(g(x/\alpha), f(x/\alpha, y/\beta))v(x/\alpha, y/\beta) + v(g(x/\alpha), f(x/\alpha, y/\beta))], \quad (8)$$

где f' означает производную по второму аргументу. Общее количество существенных собственных чисел обоих уравнений (2) и (8) определяет коразмерность типа критичности по отношению к классу систем с однонаправленной связью.

Бикритическое поведение (тип B) реализуется в системе с однонаправленной связью, когда первая подсистема демонстрирует фейгенбаумовскую критичность, а критическое поведение во второй подсистеме ассоциируется с нетривиальным решением, найденным в работе [16]: $f(x, y) = 1 - 0,5969x^2 - 0,0321x^4 - 0,8556y^2 - 0,3029x^2y^2 - 0,4317y^4 + \dots$, при этом $\alpha = -2,5029\dots$ и $\beta = -1,505318159$. Численное решение уравнений (2) и (8) дает два существенных собственных числа. Вектор с ненулевой u -составляющей соответствует фейгенбаумовской константе $\delta_1 = 4,6692\dots$. Второй вектор имеет $u \equiv 0$ и ненулевую v -составляющую; собственное число $\delta_2 = 2,39272443$. Таким образом, в классе систем с однонаправленной связью коразмерность бикритической ситуации равна двум. Для модельного отображения

$$\begin{aligned} x_n + 1 &= 1 - \lambda x_n^2, \\ y_n + 1 &= 1 - Ay_n^2 - Bx_n^2, \end{aligned} \tag{9}$$

в пространстве параметров (λ, A, B) существует бикритическая кривая (кривая бикритических точек). При конкретном значении параметра связи $B = 0,375$ бикритическая точка имеет координаты $\lambda_c = 1,401155189\dots$, $A_c = 1,124981403\dots$

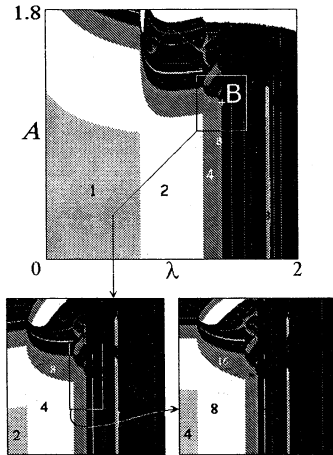


Рис. 4. Топография плоскости параметров (λ, A) для модельного отображения (18), $B = 0,375$, и иллюстрации скейлинга в окрестности бикритической точки. Скейлинговые координаты задаются соотношениями $\lambda - \lambda_c = C_1$ и $A - A_c = C_2$.

В бикритической точке существуют неустойчивые циклы всех периодов 2^k . Универсальные асимптотические значения двух мульти-

пликаторов $\mu_1 = -1,6011913\dots$ и $\mu_2 = -1,17885538\dots$. Бикритический аттрактор является фрактальным множеством, вложенным в плоскость (x, y) , его размерность Хаусдорфа $D_0 = 1,0785514$ [16].

На рис. 4 приведена топография плоскости параметров (λ, A) вблизи бикритической точки B для модельного отображения (9). Свойства скейлинга в данном случае поиллюстрированы с использованием естественных параметров отображения в качестве скейлинговых координат (очевидно, что сдвиг по параметру λ или A приводит к появлению возмущений с собственными числами δ_1 или δ_2 соответственно).

Другие типы критического поведения в системах с однонаправленной связью (двойная фейгенбаумовская точка DF и критическая точка BT на конце бикритической линии) обсуждаются в работах [5, 6, 17].

4. Критическое поведение 2D-отображений

Для того чтобы изучать типы критического поведения, которые возникают вследствие включения дополнительного измерения фазового пространства, необходимо ввести двумерное обобщение РГ-уравнения Фейгенбаума — Цвитановича. Это обобщение получается достаточно просто, если предположить, что в двумерном фазовом пространстве существует такая система координат, что масштабное преобразование, выполняемое по ходу РГ-анализа, является диагональным: $(X \rightarrow X/\alpha, Y \rightarrow Y/\beta)$. В том случае, когда критическое поведение ассоциируется с неподвижной точкой РГ-преобразования, уравнения выглядят следующим образом:

$$\begin{aligned} g(X, Y) &= \alpha g(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta)), \\ f(X, Y) &= \beta f(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta)), \end{aligned} \quad (10)$$

где два фактора скейлинга α и β должны быть найдены вместе с решением уравнений РГ (см. [18, 19]). Подчеркнем, что «скейлинговые переменные» X, Y обычно не совпадают с «естественными» переменными модельного отображения.

Двумерное обобщение задачи на собственные значения для линеаризованного РГ-уравнения в неподвижной точке (g, f) принимает вид

$$\begin{aligned} \nu u(X, Y) &= \alpha [g'_1(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))u(X/\alpha, Y/\beta) \\ &\quad + g'_2(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))v(X/\alpha, Y/\beta) \\ &\quad + u(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))], \\ \nu v(X, Y) &= \beta [f'_1(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))u(X/\alpha, Y/\beta) \\ &\quad + f'_2(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))v(X/\alpha, Y/\beta) \\ &\quad + v(g(X/\alpha, Y/\beta), f(X/\alpha, Y/\beta))], \end{aligned} \quad (11)$$

где индексы 1 и 2 обозначают производные по первому и второму аргументу. Так же, как и раньше, количество существенных собственных чисел с $|\nu| > 1$ определяет тип критичности, а сами собственные значения являются факторами скейлинга вдоль соответствующих направлений в пространстве параметров.

Универсальные мультипликаторы для двумерного случая являются собственными значениями матрицы Якоби, вычисленной в неподвижной точке $X_* = g(X_*, Y_*)$, $Y_* = f(X_*, Y_*)$.

Гамильтоновская критичность (H-тип) является хорошо известным типом поведения, который обнаруживается в сохраняющих площадь 2D-отображениях на пороге перехода к хаосу через удвоения периода [20, 21, 22, 23]. Она ассоциируется с неподвижной точкой РГ-уравнений $g(X, Y) = 1 - 0,1947 - 0,1252Y - 0,9148X^2 - 0,0050XY + 0,0004Y^2 + \dots$, $f(X, Y) = 1 + 4,7901X - 2,0556Y - 14,8638X^2 + 0,1198Y^2 + 0,3204XY + \dots$, при этом где $\alpha = -4,0180767046$ и $\beta = 16,3638968792$ (см. [18, 20, 21, 22]). Отметим, что определитель Якоби для (g, f) тождественно равен единице.

Задача на собственные значения (11) дает два существенных собственных числа: $\delta_1 = 8,7210972$ и $\delta_2 = 2$ [24, 25]. Собственное значение δ_1 связано с возмущениями, оставляющими отображение в классе преобразований, сохраняющих площадь, а δ_2 отвечает за диссипацию. Таким образом, для отображений, сохраняющих площадь, критичность типа H является феноменом коразмерности один. Тем не менее она может возникать и в диссипативных системах, но в этом случае ее коразмерность возрастает до двух, $\text{CoDim}_H = 2$.

В качестве примера рассмотрим отображение Эно (Hénon map):

$$\begin{aligned}x_n + 1 &= 1 - \lambda x_n^2 - by_n, \\y_n + 1 &= x_n.\end{aligned}\tag{12}$$

Известно, что в случае $b < 1$ при увеличении λ наблюдается классический фейгенбаумовский каскад удвоений периода (тип F в наших обозначениях). При $b = 1$ отображение становится сохраняющим площадь. В плоскости параметров (b, λ) критическая точка H расположена при $b_c = 1$, $\lambda_c = 4,13616680390428$. Другими словами, границей хаоса в отображении Эно служит фейгенбаумовская критическая линия, заканчивающаяся при $b = 1$ точкой H . Обсуждение кроссовера между двумя типами критичности проводилось в [24, 25].

Точно в критической точке H существует полный набор неустойчивых циклов с периодами 2^k , значения универсальных мультипликаторов $\mu_1 = -2,057478352$ и $\mu_2 = 1/\mu_1 = -0,486031845$. Динамика в критической точке является консервативной, поэтому кри-

тического аттрактора не существует. Тем не менее пространство состояний устроено самоподобным образом. Для модельного отображения (12) можно наблюдать самоподобие в скейлинговых переменных $X = x - 0,0475282$, $Y = (1 - \lambda x^2)/2 - y$: структура воспроизводит сама себя при увеличении масштаба в α и β вдоль осей X и Y соответственно.

Критическое поведение FQ («Feigenbaum + Quasiperiodicity») ассоциируется с неподвижной точкой РГ-уравнений (10), универсальные функции которой содержат в разложении Тейлора степени X^2 и XY : $g(X, Y) = 1 - 1,0979X^2 + 0,1571X^4 + 0,0018X^2Y^2 - 0,7114XY + 0,0865X^3Y + \dots$, $f(X, Y) = 1 + 0,0680X^2 + 1,5416X^4 + 0,2101X^2Y^2 - 2,7960XY + 1,3619X^3Y + \dots$, при этом $\alpha = -1,90007167$ и $\beta = -4,00815785$ [6, 19]. Задача (11) дает три существенных собственных числа: $\delta_1 = 6,32631925$, $\delta_2 = 3,44470967$, и $\delta_3 = \alpha = -1,90007167$. Следовательно, $\text{CoDim}_{FQ} = 3$.

Соответствующее модельное отображение имеет вид

$$x_n + 1 = 1 - ax_n^2 + dx_n y_n, \quad y_n + 1 = 1 - bx_n y_n. \quad (13)$$

В трехмерном пространстве параметров (a, b, d) существует кривая из точек FQ . При $d = 0,3$ критическая точка FQ расположена при $a_c = 1,767192895$, $b_c = 1,629678013$. Ее координаты получаются посредством оценки предела последовательности терминальных точек для линий бифуркации удвоения периода — точек, в которых оба мультипликатора обращаются в минус единицу.

В критической точке FQ существует полный набор неустойчивых циклов с периодами 2^k , асимптотические значения мультипликаторов $\mu_1 = -1,579739$ и $\mu_2 = -1,057149$. Критический аттрактор является фракталом, который можно представлять как предельный объект — «цикл периода 2^∞ ». Самоподобие аттрактора можно наблюдать в скейлинговых переменных $X = x$, $Y = y - 2,1091x$, при пересчете масштаба вдоль осей X и Y на факторы скейлинга α и β соответственно.

Благодаря специфической форме модельного отображения (13) точка FQ была найдена при двухпараметрическом анализе. В этом отображении в возмущениях, возникающих при сдвиге по параметрам из критической точки, присутствуют лишь собственные моды, соответствующие δ_1 и δ_2 . Топография пространства параметров и ее свойства скейлинга показаны на рис. 5.

Модельное отображение с количеством параметров, равным полной коразмерности, рассмотрено в [6].

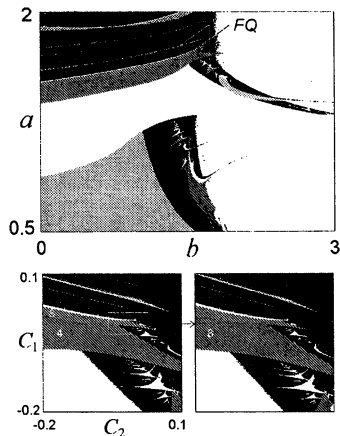


Рис. 5. Топография плоскости параметров (a, b) для отображения (13), $d = 0, 3$. Окрестность критической точки FQ показана отдельно в скейлинговых координатах: $a - a_c = C_1 + 0,47733C_2, b - b_c = C_2$.

Критическое поведение C -типа («cycle») ассоциируется с 2-циклом РГ-преобразования (10) [6, 19]. Решение задается двумя парами функций:

$$\begin{aligned}
 g_2(X, Y) &= \alpha_1 g_1(g_1(X/\alpha_1, Y/\beta_1), f_1(X/\alpha_1, Y/\beta_1)), \\
 f_2(X, Y) &= \beta_1 f_1(g_1(X/\alpha_1, Y/\beta_1), f_1(X/\alpha_1, Y/\beta_1)), \\
 g_1(X, Y) &= \alpha_2 g_2(g_2(X/\alpha_2, Y/\beta_2), f_2(X/\alpha_2, Y/\beta_2)), \\
 f_1(X, Y) &= \beta_2 f_2(g_2(X/\alpha_2, Y/\beta_2), f_2(X/\alpha_2, Y/\beta_2)).
 \end{aligned} \tag{14}$$

Соотношения (1) и (11) обычно называют «преобразованием удвоения» [1, 18]. Соответственно, можно сказать, что пары функций g_1, f_1 и g_2, f_2 являются неподвижными точками «оператора учетверения». Масштабные факторы, найденные численно для этого преобразования учетверения оказываются равны $\alpha = \alpha_1 \alpha_2 = 6,565350$ и $\beta = \beta_1 \beta_2 = 22,120227$, а полиномиальное разложение для универсальных функций см. в [5, 6, 19].

Используя это численное решение, мы обнаружили, что отображение $(x, y) \rightarrow (g_1(x, y), f_1(x, y))$ обладает *устойчивой* неподвижной точкой $X_* = 0,25039, Y_* = 1,59489$ с мультипликаторами $\mu_1 = -0,725255$ и $\mu_2 = 0,847450$. Напомним, однако, что отображение (g_1, f_1) , итерированное четырежды и подвергнутое затем перемасштабированию $(X \rightarrow X/\alpha, Y \rightarrow Y/\beta)$, превращается снова в себя. Поэтому наличие устойчивой неподвижной точки означает существование устойчивой

вых циклов всевозможных периодов 4^k , $k = 1, 2, \dots, \infty$, причем все мультипликаторы равны вышеприведенным универсальным значениям. (Заметим, что по крайней мере одна точка из орбиты цикла периода 4^k легко может быть вычислена: $X_*/\alpha^k, Y_*/\beta^k$.)

Итак, отображение $(x, y) \rightarrow (g_1(x, y), f_1(x, y))$ имеет бесконечное (счетное) множество сосуществующих аттракторов — циклов периода 4^k . Мы называем это множество *критическим квазиаттрактором*. Отображение $(x, y) \rightarrow (g_2(x, y), f_2(x, y))$ имеет неустойчивую неподвижную точку с мультипликаторами $\mu_1 = -0,848865$ и $\mu_2 = 1,174459$.

Линеаризуя уравнения (14), можно получить задачу на собственные значения для возмущений найденного решения под действием РГ-преобразования учетверения. Наибольшие собственные числа (исключая те, которые ассоциируются с инфинитесимальными заменами переменных): $\delta_1 = 92,43126348$, $\delta_2 = 4,19244418$, $\delta_3 \approx 0,93$. Только δ_1 и δ_2 больше единицы, поэтому формально коразмерность $\text{CoDim}_C = 2$. Однако δ_3 лишь немного меньше единицы, и в общем случае это собственное число отвечает за очень медленную сходимость: как правило, количественная универсальность C -типа будет наблюдаться только после очень большого числа бифуркаций удвоения периода.

Для модельного отображения

$$x_n + 1 = a - x_n^2 + by_n, \quad y_n + 1 = -x_n^2 + dy_n \quad (15)$$

значение параметра $b = -0,6663$ было подобрано численно таким образом, чтобы оптимизировать скорость сходимости (при этом медленно затухающая мода в решении РГ-уравнения исключается). Чтобы выйти на критическую точку, мы сначала численно нашли последовательность бифуркационных точек коразмерности два — терминальных точек кривых удвоения периода, для которых мультипликаторы равны 1 и минус 1. Критическая точка C определяется как предельная точка этой последовательности: $a_c = 0,24990280$, $d_c = 0,45290288$. В данном случае критический квазиаттрактор состоит из устойчивых циклов с периодами 4^k . Свойство самоподобия критического квазиаттрактора можно наблюдать при использовании скейлинговых переменных $X = x$, $Y = y + 1.31644753$ при пересчете масштаба вдоль осей X и Y на факторы скейлинга α и β соответственно.

На рис. 6 приведена топография плоскости параметров для отображения (15) и показаны отдельно области устойчивости для циклов с периодами 4^k и $2 \cdot 4^k$ в скейлинговых координатах. Хорошо видно самоподобное устройство плоскости параметров при увеличении на масштабные факторы δ_1 и δ_2 вдоль соответствующих осей координат.

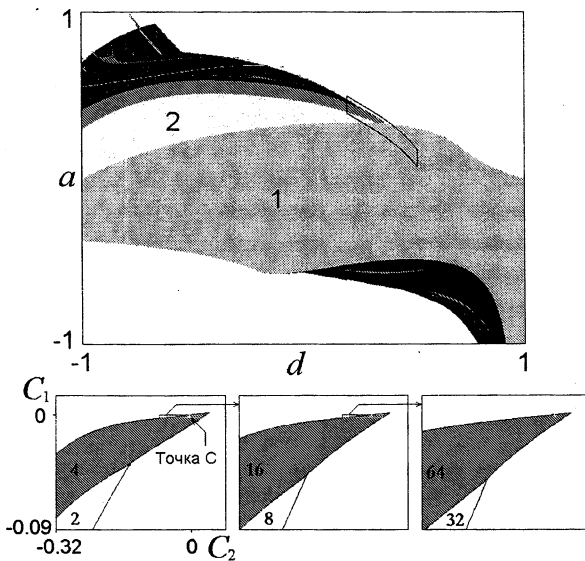


Рис. 6. Топография плоскости параметров (a, d) для отображения (15), $b = -0,6663$ (вверху) и области устойчивости для циклов периодов $2, 4, \dots, 64$ в скейлинговых координатах $a - a_c = C_1 - C_2 - 1,54607C_2^2 - 2,15C_2^3$, $d - d_c = 0,79017C_2$ (внизу). Криволинейный четырехугольник на верхней диаграмме соответствует области, представленной в скейлинговых координатах на левой нижней картинке.

Критическая точка типа C может быть найдена и при положительных b . Наилучшая сходимость наблюдается при $b = 0,6544$, критическая точка расположена при $a_c = 0,566620683$, $d_c = 1,597132592$. Этой точке отвечает цикл РГ-преобразования, осциллирующий в противоположной фазе, так что критический квазиаттрактор состоит из устойчивых циклов с периодами $2 \cdot 4^k$.

5. Заключение

Замечательная идея изучать и классифицировать математические объекты в соответствии с их коразмерностью была выдвинута впервые Анри Пуанкаре. Сначала должен быть рассмотрен случай общего положения, затем ситуации, которые становятся типичными при наличии одного управляющего параметра, потом ситуации, типичные для двухпараметрических семейств, и т. д. В частности, этот подход

эксплуатируется в теории бифуркаций и в теории катастроф. Мы полагаем, что та же самая идея может служить основой для классификации критических ситуаций на границе хаоса. Эту область исследований можно назвать *теорией многопараметрической критичности*. Отображения, перечисленные в разд. 2—4 представляют некоторые классы универсальности и играют по отношению к этим классам такую же роль, что и логистическое отображение для фейгенбаумовского критического поведения. В соответствующих ситуациях эти отображения могут оказаться полезными для феноменологического описания многопараметрических систем вблизи перехода к хаосу через удвоения периода.

В этой работе мы ограничились только такими ситуациями, которые связаны с удвоениями периода, и которые ассоциируются с неподвижными точками PГ-преобразования удвоения и учетверения. Это, конечно, только малая часть некоторой большой картины. Во-первых, кроме каскада удвоений периода, существуют другие пути перехода к хаосу, скажем, через перемежаемость или разрушение квазипериодических режимов (см. [26, 27]). Во-вторых, известно, что критическое поведение на пороге хаоса может ассоциироваться не только с неподвижными точками, но также с более сложными периодическими и непериодическими седловыми орбитами PГ-преобразования [27, 28].

Можно упомянуть также круг вопросов, связанных с анализом свойств универсальности и скейлинга в комплексных аналитических отображениях [29, 30, 37].

Интересной задачей представляется обнаружение и исследование различных типов критического поведения в реальных системах и их моделях, таких как системы дифференциальных уравнений. Несомненно, это более сложная задача, чем наблюдение классического фейгенбаумовского поведения, поскольку коразмерности выше и надо принимать во внимание многие тонкие детали. Однако в тщательно организованных компьютерных и натуральных экспериментах некоторые из обсуждаемых типов критичности могут быть обнаружены.

На двумерных диаграммах плоскостей параметров для систем различной природы, демонстрирующих удвоения периода (например, нелинейный осциллятор с периодическим возбуждением [32, 33], модулированный лазер на ядерном магнитном резонансе [34], электронная автоколебательная система [13], и т. д.), очень часто можно наблюдать структуру, названную «crossroad area» [35]. Это как раз та область пространства параметров, в которой возможно псевдо-трикритическое поведение. Действительно, для ряда системы это поведение было обнаружено численно и детально исследовано [11, 14].

Бикритическое поведение наблюдалось в нескольких демонстрирующих удвоения периода диссипативных системах с однонаправленной связью (см. [11, 15, 36, 37]). Если ввести обратную связь, представляется возможным реализовать типы C , FQ и H [19].

Еще одна возможность получить критическое поведение типа C открывается, если рассматривать диссипативную хаотическую систему под воздействием периодической внешней силы. Если бифуркационная кривая удвоения периода встречается с краем языка Арнольда [38], то в такой системе реализуется критичность C -типа [39].

Работа выполнена при финансовой поддержке РФФИ (грант 00-02-17509). А. П. К. благодарит за поддержку Фонд содействия отечественной науке.

Литература

1. Фейгенбаум М. Универсальность в поведении нелинейных систем // УФН. 1983. Т. 141, N 2. С. 343.
2. Вул Е. Б., Синай Я. Г., Ханин К. М. Универсальность Фейгенбаума и термодинамический формализм // УМН. 1984. Т. 39, N 3. С. 3.
3. Cvitanović P., ed. Universality in Chaos. Adam Hilger, 2nd Edition (1989).
4. Кузнецов С. П. Динамический хаос. М.: Физматлит, 2001. 296 с.
5. Сайт Саратовской группы теоретической нелинейной динамики <http://www.sgtnd.tserv.ru>
6. Kuznetsov A. P., Kuznetsov S. P., Sataev I. R. A variety of period-doubling universality classes in multi-parameter analysis of transition to chaos // Physica D109. 1997. 91–112.
7. Halsey T. S., Jensen M. H., Kadanoff L. P., Procaccia I., Shraiman B. I. Fractal measures and their singularities // Phys. Rev. 1986 V. A33. P. 1141.
8. Chang S. J., Wortis M., Wright J. A. Iterative properties of a one-dimensional quartic map. Critical lines and tricritical behavior // Phys. Rev. 1981. V. A24. P. 2669.
9. Hu B. and Satya I. A spectrum of universality classes in period doubling and period tripling // Phys. Lett. 1983. V. A98. P. 143.
10. Kuznetsov S. P. Tricriticality in two-dimensional maps // Phys. Lett. 1992. V. A169. P. 438.
11. Kuznetsov A. P., Kuznetsov S. P., Sataev I. R., Chua L. O. Multi-parameter criticality in Chua's circuit at period-doubling transition to chaos // Int. J. Bifurcation and Chaos. 1996. V. 6. P. 119.

12. *Fraser S. and Kapral R.* Universal vector scaling in one-dimensional maps // *Phys. Rev.* 1984. V. A25. P. 3223.
13. *Kuznetsov A. P., Kuznetsov S. P., Sataev I. R.* Three-parameter scaling for one-dimensional maps // *Phys. Lett.* 1994. V. A189. P. 367.
14. *Kuznetsov A. P., Kuznetsov S. P., Turukina L. V., Mosekilde E.* Two-parameter analysis of the scaling behavior at the onset of chaos: Tricritical and pseudo-tricritical points // *Physica A*300. 2001. N 3–4. С. 367–385.
15. *Безручко Б. П., Гуляев Ю. В., Кузнецов С. П., Селезнев Е. П.* Новый тип критического поведения связанных систем при переходе к хаосу // *ДАН СССР.* 1986. Т. 287, N 3. С. 619.
16. *Kuznetsov A. P., Kuznetsov S. P., Sataev I. R.* Bicritical dynamics of period-doubling systems with unidirectional coupling // *Int. J. Bifurcation and Chaos.* 1991. V. 1. P. 839.
17. *Kuznetsov A. P., Kuznetsov S. P., Sataev I. R.* Variety of types of critical behavior and multistability in period-doubling systems with unidirectional coupling near the onset of chaos // *Int. J. Bifurcation and Chaos.* 1993 V. 3. P. 139.
18. *Eckmann J.-P., Koch H., Wittwer P.* Existence of a fixed point of the doubling transformation for area-preserving maps of the plane // *Phys. Rev.* 1982. V. A26. P. 720.
19. *Kuznetsov S. P., Sataev I. R.* New types of critical dynamics for two-dimensional maps // *Phys. Lett.* 1992. V. A162. P. 236.
20. *Greene J. M., MacKay R. S., Vivaldi F., Feigenbaum M. J.* Universal behavior in families of area preserving maps // *Physica.* 1981. V. D3. P. 468.
21. *Collet P., Eckmann J.-P., Koch H.* On universality for area-preserving maps of the plane // *Physica.* 1981. V. D3. P. 457.
22. *Widom M., Kadanoff L. P.* Renormalization group analysis of area-preserving maps // *Physica.* 1982. V. D5. P. 287.
23. *Лихтенберг А., Либерман М.* Регулярная и стохастическая динамика. М.: Мир, 1984.
24. *Zisook A. B.* Universal effects of dissipation in two-dimensional mappings // *Phys. Rev.* 1981. V. A24. P. 1640.
25. *Schmidt G.* Universality of dissipative systems // *Direction in Chaos* / Ed Hao Bai-lin. World Scientific, 1988. V. 2. P. 1.
26. *Zisook A. B.* The complete set of Hamiltonian intermittence scaling behaviors // *Commun. Math. Phys.* 1984. V. 96. P. 361.

27. *Ostlund S., Rand D., Sethna J., Siggia E. D.* Universal properties of the transition from quasi-periodicity to chaos in dissipative systems // *Physica*. 1983. V. D8. P. 303.
28. *MacKay R. S., van Zeijts J. B. J.* Period doubling for bimodal maps: a horseshoe for a renormalization operator // *Nonlinearity*. 1988. V. 1. P. 253.
29. *Гольберг А. И., Синай Я. Г., Хачин К. М.* Универсальные свойства последовательности угрюения периода // *УМН*. 1983. Т. 38, N 1. С. 159.
30. *Čvitanović P., Myrheim J.* Complex universality // *Commun. Math. Phys.* 1989. V. 121. P. 225.
31. *Isaeva O. B., Kuznetsov S. P., Ponomarenko V. I.* Mandelbrot set in coupled logistic maps and in electronic experiment // *Phys. Rev E*. 2001. V. 64. P. 055201.
32. *Bezruchko B. P., Prokhorov M. D., Seleznev E. P.* Multiparameter model of a dissipative nonlinear oscillator in the form of a one-dimensional map // *Chaos, Solitons and Fractals*. 1995. V. 11. P. 2095.
33. *Scheffczyk C., Parlitz U., Kurz T., Knop W., Lauterborn W.* Comparison of bifurcation structures of driven dissipative nonlinear oscillators // *Phys. Rev.* 1991. V. A43. P. 6495.
34. *Badu R., Brun E., Finardi M.* Progress in the analysis of experimental chaos through periodic orbits // *Rev. Mod. Phys.* 1994. V. 66. P. 1389.
35. *Carcasses J., Mira C., Bosh M., Simo C., Tatjer J. C.* Crossroad area — spring area transition (1) Parameter plane representation // *Int. J. Bifurcation and Chaos*. 1991. V. 1. P. 183.
36. *Безручко Б. П., Пудовочкин О. Б.* Колебания у порога хаоса в системе однонаправленно связанных нелинейных радиотехнических осцилляторов // *Изв. вузов. Сер. Радиофизика*. 1992. Т. 35, N 1. С. 39.
37. *Kim S.-Y., Lim W.* Bicritical scaling behavior in unidirectionally coupled oscillators // *Phys. Rev. E*. 2001. V. 63. P. 036223.
38. *Anishchenko V. S.* Dynamical chaos — models and experiments. Appearance, routes and structure of chaos in simple dynamical systems. World Scientific, Singapore, 1995.
39. *Kuznetsov S. P., Sataev I. R.* Universality and scaling for the breakup of phase synchronization at the onset of chaos in a periodically driven Rössler oscillator // *Phys. Rev. E*. 2001. V. 64. P. 046214.

ЗАКОНЫ АНОМАЛЬНОЙ ДИФФУЗИИ

А. И. Саичев, С. Г. Уткин

Введение

Разным физическим явлениям, например турбулентной диффузии [1], хаотической динамике гамильтоновых систем [2], движению частиц в плазме [3], толчкам землетрясений [4], присуща аномальная диффузия. Адекватным аппаратом ее описания служат уравнения в дробных производных, обобщающие обычное уравнение диффузии. Однако, хотя статистике аномальной диффузии и дробным диффузионным уравнениям уделено много работ (см., например, обзор [5], а также [3, 6—9]), пока еще нет полного понимания свойств аномальной диффузии, и ее связей с поведением реализаций диффундирующих частиц. Ниже обсуждены простые модели аномальной диффузии, поясняющие ключевые аспекты ее динамики и статистики.

1. Скачущая частица

Многие особенности аномальной диффузии высвечивает процесс $X(t)$, скачком меняющийся на h_k в моменты t_k и постоянный между скачками. Будем трактовать $X(t)$ как координату частицы в одномерном пространстве. Возьмем момент $t = 0$ сразу после очередного скачка и положим $X(t = 0) = 0$. Тогда

$$X(t) = \sum_{k=1}^{M(t)} h_k,$$

где $m = M(t)$ число скачков до текущего момента t .

Пусть интервалы между скачками $\tau_k = t_k - t_{k-1}$ статистически независимы и имеют одинаковое распределение $f(\tau)$, а скачки h_k также независимы с распределением $w(x)$. Тогда по формуле полной вероятности плотность вероятностей $X(t)$ равна

$$W(x; t) = \sum_{m=0}^{\infty} R(m; t) w_m(x), \quad (1)$$

где

$$w_m(x) = w(x) \underbrace{* \cdots *}_m w(x), \quad w_0(x) = \delta(x),$$

а $R(m; t)$ — вероятность, что в интервале $(0, t]$ было m скачков. Выразим ее через статистику обратной к $m = M(t)$ функции

$$t = T(m) = \sum_{k=1}^m \tau_k .$$

Поскольку слагаемые статистически независимы, распределение $T(m)$ и его интегральная функция равны

$$f(t; m) = f(t) \underbrace{* \dots *}_{m-1} f(t), \quad F(t; m) = F(t) \underbrace{* \dots *}_{m-1} f(t) .$$

Здесь и всюду звездочка символизирует свертку. Очевидно, что

$$R(m; t) = P(M(t) < m + 1) - P(M(t) < m) . \quad (2)$$

Сюда входят вероятности

$$P(M(t) < m) = \langle \chi(m - M(t)) \rangle ,$$

что к моменту t было меньше m скачков, $\chi(z)$ — единичная функция, а угловые скобки означают статистическое усреднение. Так как

$$P(M(t) < m) = P(T(m) > t) = 1 - F(t; m) , \quad (3)$$

то из (3), (2), имеем окончательно:

$$R(m; t) = F(t; m) - F(t; m + 1) \quad (m \geq 1); \quad R(0; t) = 1 - F(t; 1) . \quad (4)$$

2. Кинетические уравнения скачущей частицы

Точные формулы (1), (4) для плотности вероятностей $W(x; t)$ координаты скачущей частицы неудобны на практике, так как требуют суммирования ряда из m -кратных интегралов. Кроме того, в (1) заключена чересчур детальная «микроскопическая» информация о статистике $X(t)$. В подобных случаях физики пытаются перейти к грубым макроскопическим соотношениям, справедливым на больших масштабах. Как правило, подобный подход позволяет вычленить универсальные законы, присущие исследуемому объекту.

Используем макроскопический подход для анализа диффузии скачущей частицы. В качестве первого шага выведем уравнение, которому подчиняется плотность вероятностей $W(x; t)$. Просуммируем ряд (1), перейдя от $W(x, t)$ к характеристической функции

$$\Theta(u; t) = \left\langle e^{iuX(t)} \right\rangle = \int_{-\infty}^{\infty} W(x; t) e^{iux} dx ,$$

а затем применим к ней преобразование Лапласа:

$$\hat{\Theta}(u; s) = \int_0^{\infty} \Theta(u; t) e^{-st} dt.$$

Тогда (1) переходит в геометрическую прогрессию. Ее сумма равна

$$\hat{\Theta}(u; s) = \frac{1 - \hat{f}(s)}{s[1 - \hat{f}(s) \tilde{w}(u)]}. \quad (5)$$

Здесь $\hat{f}(s)$ — лаплас-образ распределения $f(\tau)$ интервалов τ_k , а $\tilde{w}(u)$ — фурье-образ распределения $w(x)$ скачков. Перепишем (5) в виде

$$\frac{1}{\hat{f}(s)} \hat{\Theta}(u; s) - \tilde{w}(u) \hat{\Theta}(u; s) = \frac{1 - \hat{f}(s)}{s \hat{f}(s)}. \quad (6)$$

После обратного преобразования Фурье и Лапласа, это равенство превращается в интегральное уравнение для $W(x; t)$. Из него вытекают более простые кинетические уравнения. Выведем их, заметив, что поведение на больших временах и грубых масштабах описывается асимптотиками лаплас- и фурье-образов при $s \rightarrow 0$ и $u \rightarrow 0$.

Обсудим вначале классический случай, когда

$$\langle \tau \rangle = \int_0^{\infty} \tau f(\tau) d\tau < \infty, \quad \langle h^2 \rangle = \int_{-\infty}^{\infty} x^2 w(x) dx < \infty. \quad (7)$$

Тогда справедливы асимптотики

$$\hat{f}(s) \sim 1 - \langle \tau \rangle s \Rightarrow \frac{1}{\hat{f}(s)} \sim 1 + \langle \tau \rangle s \quad (s \rightarrow 0), \quad (8)$$

а также

$$\tilde{w}(u) \sim 1 - \frac{1}{2} \langle h^2 \rangle u^2 \quad (u \rightarrow 0). \quad (9)$$

Здесь принято для простоты, что распределение $w(x)$ четное, и значит $\langle h \rangle = 0$. Подставив (8), (9), в (6), приходим к уравнению

$$s \hat{\Theta} + \frac{D}{2} u^2 \hat{\Theta} = 1, \quad D = \frac{\langle h^2 \rangle}{\langle \tau \rangle}. \quad (10)$$

Так как умножение на s эквивалентно производной по t , а умножение на $-u^2$ эквивалентно 2-й производной по x , то (10) равносильно уравнению

$$\frac{\partial W}{\partial t} = \frac{D}{2} \frac{\partial^2 W}{\partial x^2} + \delta(x) \delta(t). \quad (11)$$

Искомая плотность вероятностей $W(x; t)$ есть вынужденное, удовлетворяющее условию причинности решение неоднородного уравнения (11), совпадающее с решением однородного уравнения:

$$\frac{\partial W}{\partial t} = \frac{D}{2} \frac{\partial^2 W}{\partial x^2}, \quad W(x; t = 0) = \delta(x). \quad (12)$$

Это и есть классическое уравнение диффузии. Оно универсально в том смысле, что не зависит от формы распределений $f(\tau)$ и $w(x)$, а зависит лишь от средних величин, образующих коэффициент диффузии D (10). Уравнение (12) вытекает из закона больших чисел, справедливого при выполнении 1-го неравенства (7), и центральной предельной теоремы, применимой, если имеет место 2-е неравенство (7).

Аномальная диффузия возникает, если нарушен закон больших чисел или (и) неприменима центральная предельная теорема. Выясним вначале, чем чревато нарушение закона больших чисел. Заметим, что $\langle \tau \rangle = \infty$, если распределение $f(t)$ медленно, по закону

$$f(\tau) \sim \kappa^\beta \tau^{-\beta-1} \quad (\tau \rightarrow \infty), \quad (13)$$

стремится к нулю, а асимптотика его лаплас-образа такова:

$$\hat{f}(s) \sim 1 - \kappa^\beta \frac{\Gamma(1-\beta)}{\beta} s^\beta, \quad s \rightarrow 0, \quad 0 < \beta < 1. \quad (14)$$

Для удобства анализа выберем масштаб времени так, чтобы коэффициент при s^β равнялся единице, и перепишем (14) в виде

$$\hat{f}(s) \sim 1 - s^\beta \Rightarrow \frac{1}{\hat{f}(s)} \sim 1 + s^\beta \quad (s \rightarrow 0, \quad 0 < \beta < 1). \quad (15)$$

Аналогично, $\langle h^2 \rangle = \infty$, если

$$\tilde{w}(u) \sim 1 - |u|^\alpha \quad (u \rightarrow 0, \quad 0 < \alpha < 2). \quad (16)$$

Подставив (15), (16) в (6), получим асимптотическое равенство

$$s^\beta \hat{\Theta} + |u|^\alpha \hat{\Theta} = s^{\beta-1}. \quad (17)$$

Известно, что умножение на s^β эквивалентно дробной производной по t порядка β , а умножение на $-|u|^\alpha$ — симметричной дробной производной по x порядка α . Поэтому из (17) вытекает уравнение

$$\frac{\partial^\beta W}{\partial t^\beta} = \frac{\partial^\alpha W}{\partial |x|^\alpha} + \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t) \delta(x). \quad (18)$$

Нелокальность дробной производной по t не дает заменить неоднородность уравнения (18) начальным условием, как это сделано при переходе от (11) к стандартному уравнению диффузии (12). Поэтому $W(x; t)$ ищут как вынужденное, удовлетворяющее условию причинности решение уравнения (18). Для приложений важен частный случай, когда нарушен лишь закон больших чисел, а центральная предельная теорема еще справедлива. Тогда в (16) следует положить $\alpha = 2$, а уравнение (18) переходит в уравнение субдиффузии:

$$\frac{\partial^\beta W}{\partial t^\beta} = \frac{\partial^2 W}{\partial x^2} + \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t) \delta(x). \quad (19)$$

Уравнения (18), (19) будут решены ниже. А пока поясним, почему (19) названо уравнением субдиффузии. Из него следует, что средний квадрат координаты скачущей частицы растет по закону

$$\langle X^2(t) \rangle = \frac{t^\beta}{\Gamma(1+\beta)}. \quad (20)$$

При $0 < \beta < 1$, когда справедливо уравнение (19), $\langle X^2(t) \rangle$ растет медленнее линейного закона $\langle X^2 \rangle \propto t$, присущего классическому уравнению диффузии (12), что и отражает название уравнения (19).

3. Дробно-диффузионный процесс

Выше применен традиционный физический подход, состоящий в переходе от микроскопических соотношений к универсальным макроскопическим законам. Другой плодотворный метод основан на детальном анализе модельных процессов, плотности вероятностей которых есть точные решения тех или иных эталонных уравнений. Ярким примером здесь служит винеровский процесс, чья плотность вероятностей подчиняется уравнению диффузии (12). В теории аномальной диффузии также есть подобные модельные процессы. Например, описанный ниже дробно-диффузионный процесс. Главным элементом его конструкции служит β -устойчивое время $t(\varrho)$, зависящее от ϱ – непрерывного аналога дискретного числа событий m :

$$t(\varrho) = \lim_{n \rightarrow \infty} \sum_{k=1}^{[\varrho n]} T_k^{(1/n)}.$$

Здесь $T_k^{(1/n)}$ – неотрицательные независимые величины, лаплас-образы распределений которых равны $\hat{f}_\beta^{1/n}(s)$, где

$$\hat{f}_\beta(s) = e^{-s^\beta} \quad (0 < \beta < 1) \quad (21)$$

Лаплас-образ безгранично-делимого устойчивого распределения $f_\beta(\tau)$. В силу его устойчивости распределение $t(\varrho)$ равно

$$f(t; \varrho) = \frac{1}{\varrho^{1/\beta}} f_\beta \left(\frac{t}{\varrho^{1/\beta}} \right). \quad (22)$$

Введем обратную к $t(\varrho)$ случайную функцию $\varrho(t)$. Найдем связь $f(t; \varrho)$ с распределением $Q(\varrho; t)$ функции $\varrho(t)$, опираясь на эквивалентность неравенств:

$$t(\varrho) < t \quad \iff \quad \varrho(t) \geq \varrho.$$

На вероятностном языке это означает, что

$$F(t; \varrho) = P(t(\varrho) < t) = P(\varrho(t) \geq \varrho) = \int_{\varrho}^{\infty} Q(\varrho'; t) d\varrho'.$$

Отсюда имеем соотношение между распределениями $t(\varrho)$ и $\varrho(t)$:

$$Q(\varrho; t) = -\frac{\partial F(t; \varrho)}{\partial \varrho} = -\frac{\partial}{\partial \varrho} \int_{-\infty}^t f(t'; \varrho) dt'.$$

Подставив сюда (22), найдем распределение функции $\varrho(t)$:

$$Q(\varrho; t) = \frac{1}{t^\beta} Q_\beta \left(\frac{\varrho}{t^\beta} \right), \quad Q_\beta(\varrho) = \frac{1}{\beta \varrho^{1+1/\beta}} f_\beta \left(\frac{1}{\varrho^{1/\beta}} \right). \quad (23)$$

Из определения $Q_\beta(\varrho)$ и из вида лаплас-образа (21) распределения $f_\beta(t)$ следует, что

$$\exp(-\varrho s^\beta) = \beta \int_0^\infty e^{-st} \frac{\varrho}{t^{\beta+1}} Q_\beta \left(\frac{\varrho}{t^\beta} \right) dt. \quad (24)$$

Пусть величина \mathcal{R} имеет распределение $Q_\beta(\varrho)$. Найдем моменты \mathcal{R} , умножив (24) на $\varrho^{\nu-1}$ и интегрируя по ϱ от 0 до ∞ . Это дает

$$\frac{\Gamma(\nu)}{s^{\nu\beta}} = \beta \langle \mathcal{R}^\nu \rangle \int_0^\infty e^{-st} t^{\nu\beta-1} dt = \beta \langle \mathcal{R}^\nu \rangle \frac{\Gamma(\nu\beta)}{s^{\nu\beta}}.$$

Отсюда имеем окончательно:

$$\langle \mathcal{R}^\nu \rangle = \int_0^\infty \varrho^\nu Q_\beta(\varrho) d\varrho = \frac{\Gamma(\nu)}{\beta \Gamma(\nu\beta)} = \frac{\Gamma(\nu+1)}{\Gamma(\nu\beta+1)}. \quad (25)$$

Следовательно, ряд Тейлора лаплас-образа $Q_\beta(\tau)$ имеет вид

$$\hat{Q}_\beta(s) = \langle e^{-s\mathcal{R}} \rangle = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n\beta+1)} s^n = E_\beta(-s). \quad (26)$$

Здесь $E_\beta(z)$ — функция Миттаг — Леффлера, естественным образом обобщающая экспоненту $E_1(z) = e^z$ на дробные β . Из (23), (26) видно, что лаплас-образ по ϱ распределения $Q(\varrho; t)$ (23) равен

$$\hat{Q}_\beta(\gamma; t) = \langle e^{-\gamma e(t)} \rangle = E_\beta(-\gamma t^\beta). \quad (27)$$

Для дальнейшего важно, что $E_\beta(-\gamma t^\beta)$ есть вынужденное решение уравнения в дробных производных [10]:

$$\frac{d^\beta \hat{Q}_\beta(\gamma; t)}{dt^\beta} + \gamma \hat{Q}_\beta(\gamma; t) = \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t). \quad (28)$$

Назовем дробно-диффузионным процесс

$$\mathcal{X}(t) = X(\varrho(t)). \quad (29)$$

Здесь $X(\varrho)$ так называемый Levy flights с плотностью вероятностей

$$W(x; \varrho) = \frac{1}{\varrho^{1/\alpha}} w_\alpha \left(\frac{x}{\varrho^{1/\alpha}} \right),$$

$w_\alpha(x)$ — симметричное устойчивое распределение с характеристической функцией

$$\tilde{w}_\alpha(u) = e^{-|u|^\alpha}, \quad (30)$$

а $\varrho(t)$ — независимая от $X(\varrho)$ функция с распределением $Q(\varrho; t)$ (23). Очевидно, плотность вероятностей $W(x; t)$ дробно-диффузионного процесса $\mathcal{X}(t)$ (29) автомодельна и равна

$$W(x; t) = \frac{1}{\sigma t^\mu} W_{\alpha, \beta} \left(\frac{x}{\sigma t^\mu} \right). \quad (31)$$

Здесь

$$W_{\alpha, \beta}(y) = \int_0^\infty Q_\beta(z) w_\alpha \left(\frac{y}{z^{1/\alpha}} \right) \frac{dz}{z^{1/\alpha}}, \quad \mu = \frac{\beta}{\alpha}. \quad (32)$$

Характеристическую функцию дробно-диффузионного процесса найдем, применив к (31) преобразование Фурье по x и привлекая равенства (27), (30), (32). В итоге получим

$$\Theta(u; t) = E_\beta(-t^\beta |u|^\alpha).$$

Отсюда и из (27), (28) следует, что характеристическая функция дробно-диффузионного процесса удовлетворяет уравнению

$$\frac{d^\beta \Theta(u; t)}{dt^\beta} + |u|^\alpha \Theta(u; t) = \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t),$$

а его плотность вероятностей (31) есть точное решение дробно-диффузионного уравнения (18).

Опуская громоздкие выкладки, опирающиеся на свойства устойчивых распределений и функции Миттаг — Леффлера, дадим краткую сводку свойств решения (31) уравнения (18). Так, моменты дробно-диффузионного процесса растут по закону

$$\langle |\mathcal{X}(t)|^\gamma \rangle = \frac{2}{\beta} \frac{\sin\left(\frac{\pi\gamma}{2}\right)}{\sin\left(\frac{\pi\gamma}{\alpha}\right)} \frac{\Gamma(\gamma)}{\Gamma\left(\frac{\gamma\beta}{\alpha}\right)} t^{\gamma\beta/\alpha} \quad (\gamma < \alpha < 2),$$

значение $W(x; t)$ при $x = 0$ равно

$$W(0; t) = \frac{t^{-\mu}}{\alpha \Gamma\left(1 - \frac{\beta}{\alpha}\right) \sin\left(\frac{\pi}{\alpha}\right)} \quad (\alpha > 1),$$

а его асимптотика при $x/\sigma t^\mu \rightarrow \infty$ такова:

$$W(x; t) \sim \frac{1}{\pi} \frac{t^\beta}{|x|^{\alpha+1}} \frac{\Gamma(1+\alpha)}{\Gamma(1+\beta)} \sin\left(\frac{\pi\alpha}{2}\right) \quad (|x| \gg t^\mu).$$

При $\alpha \leq 1$ $W(0; t) = \infty$. Сингулярность $W(x; t)$ при $\alpha \leq 1$ связана с аномально большим вкладом в статистику $\mathcal{X}(t)$ значений Levy flights $X(\varrho)$ при $\varrho \simeq 0$. При некоторых α и β имеются явные решения уравнения (18). Так, при $\alpha = \beta$ имеем

$$W(x; t) = \frac{1}{\pi |y| t} \frac{\sin\left(\frac{\pi\beta}{2}\right)}{|y|^\beta + |y|^{-\beta} + 2 \cos\left(\frac{\pi\beta}{2}\right)} \quad \left(y = \frac{x}{t}\right),$$

а при $\alpha = 1$, $\beta = 1/2$ справедливо равенство

$$W_{\alpha,\beta}(y) = -\frac{1}{2\pi\sqrt{\pi}} \exp\left(\frac{y^2}{4}\right) \text{Ei}\left(-\frac{y^2}{4}\right).$$

4. Процессы дробного сноса и субдиффузии

Назовем дробным сносом процесс $\mathcal{X}(t) = \varrho(t)$ с плотностью вероятностей

$$W(x; t) = \frac{1}{t^\beta} Q_\beta\left(\frac{\varrho}{t^\beta}\right). \quad (33)$$

Его характеристическую функцию найдем, подставив в (27) $\gamma = -iu$:

$$\Theta(u; t) = \langle e^{iu\varrho(t)} \rangle = E_\beta(iu t^\beta). \quad (34)$$

Согласно (28), она подчиняется уравнению

$$\frac{\partial^\beta \Theta}{\partial t^\beta} = iu\Theta + \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t). \quad (35)$$

Распределение $Q_\beta(z)$ найдем обратным преобразованием Фурье от $E_\beta(iu)$. Из свойств функции Миттаг – Леффлера следует, что

$$Q_\beta(z) = \frac{1}{\pi\beta} \operatorname{Im} e^{i\pi\beta} \int_0^\infty \exp(-x^{1/\beta} - xz e^{i\pi\beta}) dx$$

[0.3cm]($0 < \beta \leq 1/2$)

и

$$Q_\beta(z) = \frac{1}{\pi\beta} \operatorname{Re} \int_0^\infty \exp(ixz + x^{1/\beta} e^{-i\pi/2\beta}) dx$$

[0.3cm]($1/2 \leq \beta \leq 1$).

Построенные с их помощью графики $Q_\beta(z)$ показаны на рисунке. Заметим еще, что имеются явные аналитические выражения:

$$Q_{1/2}(z) = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{z^2}{4}\right), \quad Q_{1/3}(z) = \sqrt[3]{9} \operatorname{Ai}\left(\frac{z}{\sqrt[3]{3}}\right). \quad (36)$$

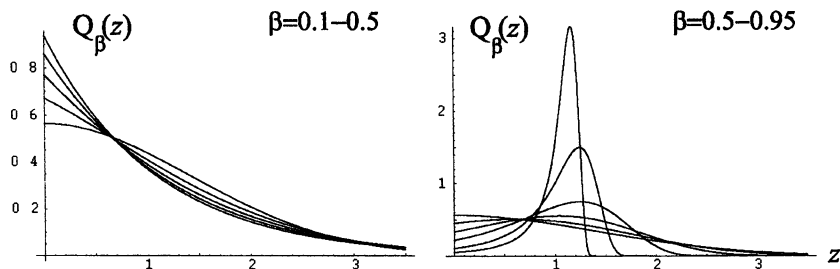


График распределения $Q_\beta(z)$ при разных β

С помощью процесса дробного сноса удается найти плотность вероятностей субдиффузионного процесса $\mathcal{X}(t) = V(\varrho(t))$. Здесь $V(\varrho)$ – винеровский процесс ($\langle V^2(\varrho) \rangle = 2\varrho$). Искомая плотность вероятностей подчиняется субдиффузионному уравнению (19), а характеристическая функция – уравнению

$$\frac{d^\beta \Theta}{dt^\beta} + u^2 \Theta = \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t).$$

Решим его, опираясь на характеристическую функцию дробного сноса. Разобьем ее на четную и нечетную составляющие

$$\Theta = \Theta_{\text{even}} + \Theta_{\text{odd}}$$

и подставим последнее равенство в (35). Это дает два уравнения:

$$\frac{d^\beta \Theta_{\text{even}}}{dt^\beta} = iu\Theta_{\text{odd}} + \frac{t^{-\beta}}{\Gamma(1-\beta)} \chi(t), \quad \frac{d^\beta \Theta_{\text{odd}}}{dt^\beta} = iu\Theta_{\text{even}}.$$

Поддействовав на первое из них производной по t порядка β и приняв во внимание, что (см. приложение)

$$\frac{d^\beta t^\delta \chi(t)}{dt^\beta} = \frac{\Gamma(\delta+1)}{\Gamma(1+\delta-\beta)} t^{\delta-\beta} \chi(t),$$

после исключения нечетной компоненты Θ_{odd} придем к замкнутому уравнению относительно $\Theta_{\text{even}}(u; t)$:

$$\frac{d^{2\beta} \Theta_{\text{even}}}{dt^{2\beta}} + u^2 \Theta_{\text{even}} = \frac{t^{-2\beta}}{\Gamma(1-2\beta)} \chi(t).$$

Соответственно, четная компонента

$$W_{\text{even}}(x; t) = \frac{1}{2} [W(x; t) + W(-x; t)] = \frac{1}{2} W(|x|; t) \quad (37)$$

плотности вероятностей дробного сноса подчиняется уравнению

$$\frac{\partial^{2\beta} W_{\text{even}}}{\partial t^{2\beta}} = \frac{\partial^2 f}{\partial x^2} + \frac{t^{-2\beta}}{\Gamma(1-2\beta)} \chi(t) \delta(x).$$

Замена $\beta \mapsto \beta/2$ сводит его к (19). Таким образом, согласно (37), (33), решение уравнения субдиффузии (19) выглядит как

$$W(x; t) = \frac{1}{2 t^{\beta/2}} Q_{\beta/2} \left(\frac{|x|}{t^{\beta/2}} \right). \quad (38)$$

Положив $\beta = 2/3$ и вспомнив 2-ю формулу (36), в частности, имеем:

$$W(x; t) = \frac{1}{2} \sqrt[3]{\frac{9}{t}} \text{Ai} \left(\frac{|x|}{\sqrt[3]{3t}} \right), \quad \beta = \frac{2}{3}.$$

5. Супердиффузия летящих частиц

К супердиффузии ведет процесс $X(t)$, аналогичный координате частиц газа. Пусть $\{\dots t_0, t_1, \dots t_n, \dots\}$ — моменты «столкновений». Между ними частица летит равномерно, а при $t = t_k$ ее скорость становится v_{k+1} и не зависит от прежней скорости. Начальный момент выберем в момент столкновения ($t_0 = 0$) и будем считать, что частица в этот момент была в начале координат ($X(t = 0) = 0$). Кроме того, интересуясь «макроскопическими» свойствами супердиффузии, и для простоты выкладок, применим «макроскопическую процедуру измерения»: за координату частицы в текущий момент t примем координату ее последнего столкновения:

$$X_-(t) = X(T(m = M(t))) = \sum_{k=1}^{M(t)} v_k \tau_k.$$

Пусть v_k и τ_k независимы и имеют распределения $w(v)$ и $f(\tau)$. Тогда выкладки, аналогичные тем, что привели к формуле (5), дают лаплас-образ характеристической функции летящей частицы:

$$\hat{\Theta}_-(u; s) = \frac{1}{s} \frac{1 - \hat{f}(s)}{1 - g(u, s)}, \quad g(u, s) = \langle \tilde{w}(u\tau) e^{-s\tau} \rangle. \quad (39)$$

Изучим законы диффузии летящих частиц, полагая для определенности, что $\langle v \rangle = 0$, а $\langle v^2 \rangle = \sigma^2$. В этом случае имеем

$$g(u, s) = \hat{f}(s) - \frac{\sigma^2}{2} u^2 \hat{f}''(s) + \dots,$$

где штрихи означают производные по s . Подставив последнее равенство в (39), придем к формуле, справедливой при $u \rightarrow 0$:

$$\hat{\Theta}_-(u; s) = \frac{\varphi(s)}{s} \frac{1}{\varphi(s) + \frac{1}{2} \sigma^2 u^2}, \quad \varphi(s) = \frac{1 - \hat{f}(s)}{\hat{f}''(s)}. \quad (40)$$

Физиков больше всего интересует средний квадрат $\langle X^2(t) \rangle$ координаты диффундирующей частицы. Обсудим его в первую очередь. Из (40) следует, что лаплас-образ среднего квадрата равен

$$\hat{X}_-^2(s) = \frac{\sigma^2}{s \varphi(s)}. \quad (41)$$

Найдем его асимптотику при $s \rightarrow 0$ для разных типов распределений $f(\tau)$. Пусть, во-первых, $f(\tau)$ при $t \rightarrow \infty$ быстро спадает к нулю, а

первые 2 момента интервалов между столкновениями ограничены:

$$\langle \tau \rangle < \infty, \quad \langle \tau^2 \rangle < \infty. \quad (42)$$

Тогда справедливо разложение:

$$\hat{f}(s) = 1 - \langle \tau \rangle s + \frac{1}{2} \langle \tau^2 \rangle s^2 + \dots \Rightarrow \hat{f}(s) = \langle \tau^2 \rangle + \dots \quad (43)$$

Подставив (43) в (41), будем иметь

$$\hat{X}_-^2(s) \sim \sigma^2 \frac{\langle \tau^2 \rangle}{\langle \tau \rangle} \frac{1}{s^2}, \quad s \rightarrow 0. \quad (44)$$

Пользуясь табличной формулой преобразования Лапласа

$$\frac{1}{s^\gamma} \iff \frac{t^{\gamma-1}}{\Gamma(\gamma)}, \quad (45)$$

из (44) имеем стандартный линейный закон диффузии:

$$\langle X_-^2(t) \rangle \sim Dt, \quad D = \sigma^2 \frac{\langle \tau^2 \rangle}{\langle \tau \rangle}.$$

Супердиффузия возникает, если нарушены неравенства (42), а $f(\tau)$ имеет асимптотику (13). Отдельно обсудим случаи $0 < \beta < 1$ и $1 < \beta < 2$. В первом из них справедлива асимптотика (14), а

$$\hat{f}''(s) \sim \kappa^\beta \Gamma(2 - \beta) s^{\beta-2}, \quad s \rightarrow 0, \quad 0 < \beta < 1. \quad (46)$$

Подставив (14), (46) в (41), имеем

$$\hat{X}_-^2(s) \sim \sigma^2 \frac{\beta \Gamma(2 - \beta)}{\Gamma(1 - \beta)} \frac{1}{s^3}, \quad s \rightarrow 0, \quad 0 < \beta < 1.$$

Отсюда и из (45) получим баллистический закон диффузии:

$$\langle X_-^2(t) \rangle \sim \sigma^2 \frac{\beta}{2} \frac{\Gamma(2 - \beta)}{\Gamma(1 - \beta)} t^2, \quad t \rightarrow \infty, \quad 0 < \beta < 1. \quad (47)$$

Его физический смысл прост. При $0 < \beta < 1$ столкновения редки, а частица летит почти равномерно: $X(t) \propto vt \rightarrow \langle X^2(t) \rangle \propto \sigma^2 t^2$.

Перейдем к случаю $1 < \beta < 2$. Здесь справедлива асимптотика

$$\hat{f}(s) \sim 1 - \langle \tau \rangle s + \kappa^\beta \frac{\Gamma(2 - \beta)}{\beta(\beta - 1)} s^\beta, \quad s \rightarrow 0, \quad 1 < \beta < 2, \quad (48)$$

а асимптотика \hat{f}'' задана формулой (46). Соответственно,

$$\hat{X}_-^2(s) \sim \sigma^2 \frac{\kappa^\beta}{\langle \tau \rangle} \Gamma(2 - \beta) s^{\beta-4}, \quad s \rightarrow 0, \quad 1 < \beta < 2. \quad (49)$$

Пользуясь табличной формулой (45), получим окончательно:

$$\langle X_-^2(t) \rangle \sim \sigma^2 \frac{\kappa^\beta}{\langle \tau \rangle} \frac{\Gamma(2 - \beta)}{\Gamma(4 - \beta)} t^{3-\beta} \quad (1 < \beta < 2). \quad (50)$$

Таким образом, при $1 < \beta < 2$ средний квадрат координаты частицы подчиняется супердиффузионному закону:

$$\langle X_-^2(t) \rangle \propto t^\gamma, \quad \gamma = 3 - \beta \quad (1 < \beta < 2). \quad (51)$$

Изучим вероятностные свойства супердиффузии. Выпишем асимптотику $\varphi(s)$ (40) при $s \rightarrow 0$:

$$\varphi(s) \sim \delta s^\gamma, \quad \delta = \frac{\langle \tau \rangle}{\kappa^\beta \Gamma(2 - \beta)}, \quad \gamma = 3 - \beta \quad (1 < \beta < 2). \quad (52)$$

Подберем масштабы времени и координаты так, что $\sigma^2 = 2$, а $\delta = 1$. Подставив затем (52) в первое равенство (40), будем иметь

$$s^\gamma \hat{\Theta} + u^2 \hat{\Theta} = s^{\gamma-1}.$$

Отсюда видно, что плотность вероятностей супердиффузии подчиняется тому же уравнению (19), что и субдиффузии, но с заменой β на $\gamma = 3 - \beta$, а плотность вероятностей супердиффузии задана выражением (38) с той же заменой β на γ . Иначе говоря, форма плотностей вероятностей субдиффузии изображена на рисунке 1 слева, а супердиффузии — справа.

Приложение. Дробные производные

Статистическое описание аномальной диффузии опирается на уравнения в дробных производных. Поэтому дадим краткую справку о их свойствах. Операцию дробного дифференцирования вводят как обратную к дробному интегрированию. Формула же дробного интегрирования постулируется как естественное обобщение *формулы Коши* для n -кратного интеграла. Выведем ее, для чего введем оператор n -кратного интегрирования:

$$I^n g(t) = \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \dots \int_{-\infty}^{t_{n-1}} dt_n g(t_n). \quad (53)$$

Результат действия I^n на $g(t)$ есть линейный функционал. Любой линейный функционал представим в виде однократного интеграла, а с учетом инвариантности оператора I^n к сдвигам — в виде свертки $g(t)$ и ядра $k_n(t)$:

$$I^n g(t) = \int_{-\infty}^t k_n(t-s)g(s) ds = k_n(t) * g(t). \quad (54)$$

Формулу Коши получим, найдя ядро $k_n(t)$ линейного функционала (54). Подставив в (53) и (54) $g(t) = \delta(t)$, получим

$$k_n(t) = \frac{t^{n-1}}{(n-1)!} \chi(t).$$

Естественное обобщение этого ядра на дробные n имеет вид

$$k_\beta(t) = \frac{t^{\beta-1}}{\Gamma(\beta)} \chi(t).$$

Оператор дробного дифференцирования (обозначим его здесь D^β) определяют как оператор, обратный дробному интегрированию, т. е. такой, что

$$D^\beta \cdot I^\beta = \text{Identity}.$$

Теория обобщенных функций (см., например, [10]) дает следующую конструкцию оператора D^β . Пусть $n = [\beta]$ — наименьшее целое, большее β . Производная порядка β равна свертке n -й производной функции $g(t)$ с ядром оператора дробного интегрирования кратности $n - \beta$:

$$D^\beta g(t) = k_{n-\beta}(t) \cdot g^{(n)}(t).$$

Образно говоря, чтобы найти дробную производную функции $g(t)$ порядка β , надо «передифференцировать» ее, взяв целую производную порядка $n > \beta$, а затем подынтегрировать дробное число $n - \beta$ раз. В случае $0 < \beta < 1$ дробная производная равна

$$D^\beta g(t) = \frac{1}{\Gamma(1-\beta)} \int_{-\infty}^t \frac{g'(s)}{(t-s)^\beta} ds. \quad (55)$$

Продифференцируем, к примеру, $\beta < 1$ раз произведение гладкой функции $\varphi(t)$ и единичной функции $\chi(t)$: $g(t) = \varphi(t) \chi(t)$. Так как

$$g'(t) = \varphi(0) \delta(t) + \chi(t) \varphi'(t),$$

то из (55) имеем

$$D^\beta g(t) = \varphi(0) \frac{t^{-\beta}}{\Gamma(1-\beta)} + \frac{1}{\Gamma(1-\beta)} \int_0^t \frac{\varphi'(s)}{(t-s)^\beta} ds \quad (t > 0).$$

В частности, положив здесь $\varphi(t) \equiv 1$, найдем

$$D^\beta \chi(t) = \frac{t^{-\beta}}{\Gamma(1-\beta)} \quad (t > 0).$$

Из приведенных формул видно главное отличие дробной производной от обычных производных – это нелокальная операция. Так, скачок $\chi(t)$ при $t = 0$ рождает спадающий как $t^{-\beta}$ хвост дробной производной. Роднит же операторы дробных и целых производных их масштабная инвариантность. Заметим еще, что из-за нелокальности дробной производной по t дробное уравнение (18) содержит источник, выполняющий роль начальных условий уравнений с целыми производными по времени. Осталось добавить, что лаплас-образ производной порядка β от функции $g(t)$ равен $s^\beta \hat{g}(s)$.

Помимо «ориентированной вперед по времени» дробной производной по t , в уравнение (18) входит симметричная производная по x . Ее смысл наглядней всего пояснить не в координатном, а в фурье-представлении. Для этого введем оператор преобразования Фурье:

$$\mathcal{F}[w(x)] = \tilde{w}(q) = \int_{-\infty}^{\infty} w(x) e^{iqx} dx.$$

По определению фурье-образ симметричной пространственной производной функции $w(x)$ равен фурье-образу исходной функции $\tilde{w}(q)$, умноженному на минус $|q|^\alpha$:

$$\mathcal{F} \left[\frac{\partial^\alpha}{\partial |x|^\alpha} w(x) \right] = -|q|^\alpha \tilde{w}(q). \quad (56)$$

Минус взят для того, чтобы при $\alpha \rightarrow 2$ симметричная дробная производная переходила в обычную 2-ю производную.

Интегральный оператор симметричного дифференцирования можно найти, например, обращением равенства (56) по Фурье. К примеру, симметричная производная 1-го порядка равна

$$\frac{dw(x)}{d|x|} = \frac{1}{\pi} \int_0^\infty \frac{w(x-y) + w(x+y) - 2w(x)}{y^2} dy.$$

Литература

1. *Klafter J., Shlesinger M. F., Zumofen G.* Beyond Brownian Motion // *Physics Today*. 1996. V. 49. P. 33–39.
2. *Zaslavsky G. M., Edelman M., Niyazov D. A.* Self-similarity, renormalization, and space nonuniformity of Hamilton chaotic dynamics // *Chaos*. 1997. V. 7. P. 159–181.
3. *Забурдаев В. Ю., Чукбар К. В.* Ускоренная супердиффузия и конечная скорость полетов Леви // *ЖЭТФ*. 2002. Т. 121, вып. 2. С. 299–307.
4. *Sornette A., Sornette D.* Sub-critical and Super-critical Regimes in Epidemic Models of Earthquake Aftershocks // *Geophys. Res.* (<http://ArXiv.org/abs/cond-mat/01093180>).
5. *Metzler R., Klafter J.* The random walk's guide to anomalous diffusion: a fractional dynamics approach // *Physics Reports*. 2000. V. 339. P. 1–77.
6. *Saichev A. I., Zaslavsky G. M.* Fractional kinetic equations: solutions and applications // *Chaos*. 1997. V. 7. P. 753–764.
7. *Золотарев В. М., Учайкин В. В., Саенко В. В.* Супердиффузия и устойчивые законы // *ЖЭТФ*. 1999. Т. 115, вып. 4. С. 1411–1425.
8. *Metzler R., Barkai E., Klafter J.* Anomalous diffusion and relaxation close to thermal equilibrium: A fractional Fokker-Planck equation approach // *Phys. Rev. Lett.* 1999. V. 82. P. 3563–3567.
9. *Barkai E.* Fractional Fokker-Planck equation, solution, and application // *Phys. Rev. E*. 2001. V. 63. P. 046118–1/17.
10. *Saichev A. I., Woyczyński W. A.* Distributions in the physical and engineering sciences. Boston: Birkhäuser, 1997. 336 p.

ПРОСТЕЙШИЕ МОДЕЛИ ПОТОКОВ ЧАСТИЦ

С. Н. Гурбатов, А. И. Саичев

Введение

Обсудим простейшие, допускающие полное аналитическое исследование модели потоков частиц. Сюда относится, прежде всего, одномерное движение слипающихся частиц. На их важность применительно к астрофизике указал Зельдович, заметивший, что формирование крупномасштабных структур вещества во Вселенной носит квазиодномерный характер [1, 2]. Рассмотрим также модельные уравнения одномерного газа, учитывающие конкурентное влияние инерционной нелинейности, давления и диссипации. Анализ подобных точно решаемых моделей полезен для тестирования явлений и эффектов, присущих реальным потокам частиц.

1. Динамика слипающихся частиц

Удобный способ анализа одномерного газа без давления, родственному принципу Олейник — Лэкса¹, основан на вариационном принципе, предложенном в [4]. Обсудим некоторые его следствия.

1.1. Одномерный поток невзаимодействующих частиц

Пусть по оси x движется поток слипающихся частиц (материальных точек), начальные плотность и скорость которого равны $\rho_0(x)$ и $v_0(x)$. Найдем поля скорости $v(x, t)$ и плотности $\rho(x, t)$ потока в текущий момент $t > 0$, полагая на время, что $v_0(x)$ и $\rho_0(x)$ — гладкие функции. Тогда имеется такое $t_n > 0$, что в интервале $t \in [0, t_n)$ частицы не успевают столкнуться и движутся равномерно:

$$x = X(y, t) = y + v_0(y) t. \quad (1)$$

Здесь y — начальные (лагранжевы) координаты частиц. Равенство (1) отображает лагранжевы координаты в эйлеровы. До момента t_n это гладкая строго монотонная функция y . Поэтому при $t \in [0, t_n)$ имеется обратная функция $y = y(x, t)$, отображающая эйлеровы координаты в лагранжевы. Скорость $v(x, t)$ равномерно движущихся частиц

¹Наглядное изложение этого принципа и его астрофизические приложения даны в [3].

найдем, поделив пройденный ими путь на время движения

$$v(x, t) = \frac{x - y(x, t)}{t}. \quad (2)$$

Введем еще массу частиц слева от точки x :

$$m(x, t) = \int_{-\infty}^x \rho(z, t) dz, \quad M(x) = \int_{-\infty}^x \rho_0(z) dz. \quad (3)$$

Так как обгон частиц отсутствует, их масса слева и плотность равны

$$m(x, t) = M(y(x, t)) \quad \Rightarrow \quad \rho(x, t) = \frac{\partial}{\partial x} M(y(x, t)). \quad (4)$$

При $t \in [0, t_n)$, пока частицы не слипаются, равенства (2), (4) дают классические решения уравнений одномерного газа без давления

$$\begin{cases} \frac{\partial \rho v}{\partial t} + \frac{\partial \rho v^2}{\partial x} = 0, \\ [2mm] \frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x} = 0. \end{cases} \quad (5)$$

Ниже описан вариационный принцип [4], дающий обобщенные решения этих уравнений с учетом слипания частиц.

1.2. Учет слипания частиц

Соотношения (2), (4) задают движение одномерного газа и при слипании частиц. Надо лишь дать более общее определение отображения $y = y(x, t)$. Это достигается введением функции

$$\mathcal{S}(y; x, t) = \int^y \mathcal{R}(z; x, t) \rho_0(z) dz, \quad (6)$$

где

$$\mathcal{R}(y; x, t) = X(y, t) - x = v_0(y) t + y - x. \quad (7)$$

Назовем y аргументом функций (6), (7), а x и t — их параметрами. Нижний предел в (6) опущен, поскольку, подобно любому потенциалу, \mathcal{S} определена с точностью до произвольной постоянной. Укажем ключевые свойства функций \mathcal{R} и \mathcal{S} , полагая пока, что $t \in [0, t_n)$, а частицы еще не успели столкнуться. Тогда \mathcal{R} — монотонная функция, равная нулю в точке $y = y(x, t)$, а \mathcal{S} — выпуклая функция, координата минимума которой есть отображение эйлеровых в лагранжевы

координаты $y = y(x, t)$. При $t > t_n$ \mathcal{R} (7) теряет монотонность, и может обращаться в нуль в нескольких точках

$$y_1(x, t) < y_2(x, t) < \dots < y_k(x, t) \quad (8)$$

оси y . Здесь $k = k(x, t) \geq 1$. Соответственно, функция \mathcal{S} (6) имеет $k(x, t)$ локальных экстремумов в точках с координатами (8). Их физический смысл очевиден: это лагранжевы координаты не взаимодействующих при обгоне частиц, попавших в момент t в точку с эйлеровой координатой x . Иначе говоря, функции (8) есть разные ветви многозначного отображения $y = y(x, t)$, описывающего многопоточковое движение частиц. Подставив $y = y(x, t)$ в (2), (4), получим многопоточковые поля скорости и плотности.

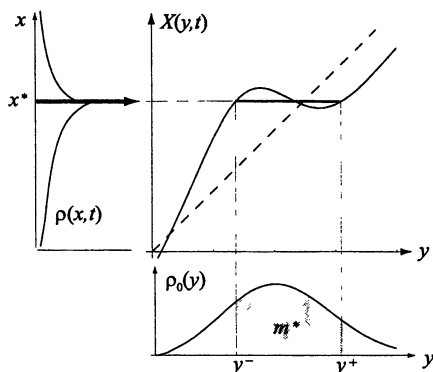
Сконструируем из многозначной функции $y = y(x, t)$ однозначную функцию $y = y_w(x, t)$, оставляя лишь ту ее ветвь, на которой потенциал $\mathcal{S}(y; x, t)$ принимает наименьшее значение. В [4] доказано, что монотонно растущая разрывная функция $y_w(x, t)$, будучи подставленной в (2), (4), дает разрывное поле скорости $v_w(x, t)$ и сингулярную плотность $\rho_w(x, t)$ слипающихся частиц. Обсудим структуру этих обобщенных (weak) решений уравнений (5). Пусть $y = y_w(x, t)$ терпит разрывы в точках $\{x_i^*(t)\}$, и принимает слева и справа от них значения $\{y_i^\mp(t)\}$. Тогда масса слева $m_w(x, t)$ (4) имеет в точках $\{x_i^*(t)\}$ разрывы величиной

$$m_i^*(t) = \int_{y_i^-(t)}^{y_i^+(t)} \rho_0(y) dy. \quad (9)$$

Очевидно, это массы «макрочастиц», возникших из-за слипания частиц, заключенных при $t = 0$ в интервалах $[y_i^-(t), y_i^+(t)]$. Соответственно, обобщенная плотность равна

$$\rho_w(x, t) = \rho_c(x, t) + \sum_i m_i^*(t) \delta(x - x_i^*(t)), \quad (10)$$

где ρ_c — плотность еще не слипшихся частиц, а сумма дает сингулярную плотность макрочастиц. Аналогично, $v_w(x, t)$ (2) имеет разрывы в точках $\{x_i^*(t)\}$, а между ними описывает равномерное движение не слипшихся частиц. Наглядное представление о поведении слипающихся частиц дает обратное к $y = y_w(x, t)$ отображение $x = X_w(y, t)$. Это непрерывная функция, постоянная в интервалах $[y_i^-(t), y_i^+(t)]$, где $X_w(y, t)$ равна координатам макрочастиц $x_i^*(t)$. Типичный график отображения $x = X_w(y, t)$ и порожденного им сингулярного поля плотности дан на рисунке.



Отображение $x = X_w(y, t)$ и отвечающее ему преобразование начальной плотности $\rho_0(y)$ в текущую сингулярную плотность $\rho_w(x, t)$ слипающихся частиц.

1.3. Кинематика слипающихся частиц

Убедимся, что поля v_w и ρ_w правильно описывают слипание частиц, обсудив движение макрочастиц. По закону сохранения импульса, их скорость равна

$$v_i^*(t) = \frac{p_i^*(t)}{m_i^*(t)}, \quad (11)$$

где $m_i^*(t)$ — масса (9) макрочастицы, а

$$p_i^*(t) = \int_{y_i^-(t)}^{y_i^+(t)} \rho_0(y) v_0(y) dy \quad (12)$$

— ее импульс. Перепишем потенциал (6) в более наглядной форме, положив нижний предел интеграла равным $-\infty$. Тогда

$$S(y; x, t) = P(y)t - xM(y) + X^c(y)M(y). \quad (13)$$

Сюда входит начальная масса $M(y)$ (3) и импульс частиц слева

$$P(y) = \int_{-\infty}^y \rho_0(z) v_0(z) dz. \quad (14)$$

Кроме того, в (13) фигурирует начальный центр масс частиц, расположенных левее точки y :

$$X^c(y) = \frac{N(y)}{M(y)}, \quad N(y) = \int_{-\infty}^y z \rho_0(z) dz. \quad (15)$$

Заметим далее, что при изменении x от $x_i^*(t) - 0$ до $x_i^*(t) + 0$ координата минимума функции $S(y; x, t)$ скачком меняется от $y_i^-(t)$ до $y_i^+(t)$, а $S(y; x^*, t)$ имеет, в силу непрерывности, одинаковые минимумы в точках $y_i^-(t)$ и $y_i^+(t)$. Иначе говоря, справедливо равенство

$$S(y_i^+; x_i^*, t) - S(y_i^-; x_i^*, t) = \int_{y_i^-}^{y_i^+} [x_i^* - y - v_0(y)t] \rho_0(y) dy = 0. \quad (16)$$

Почленно интегрируя, получаем отсюда

$$x_i^*(t) = x_i^c(t) + v_i^*(t)t. \quad (17)$$

Здесь $v_i^*(t)$ задано выражением (11), где теперь

$$p_i^*(t) = P(y_i^+(t)) - P(y_i^-(t)), \quad m_i^*(t) = M(y_i^+(t)) - M(y_i^-(t))$$

— импульс и масса частиц, находившихся в начальный момент времени $t = 0$ в интервале $[y_i^-(t), y_i^+(t)]$, а

$$x_i^c(t) = \frac{1}{m_i^*(t)} \int_{y_i^-(t)}^{y_i^+(t)} z \rho_0(z) dz$$

— их центр масс. Равенство (17) дает центр масс частиц, находившихся при $t = 0$ в интервале $[y_i^-(t), y_i^+(t)]$. Если они слиплись в одну макрочастицу, то (17) — ее координата. Окончательно убедимся, что (17) задает движение макрочастиц, показав, что входящие в (17) координата $x_i^*(t)$ и скорость $v_i^*(t)$ связаны соотношением

$$\frac{dx_i^*(t)}{dt} = v_i^*(t). \quad (18)$$

Покажем это, дифференцируя по t условие разрыва (16):

$$\frac{d}{dt} \int_{y_i^-(t)}^{y_i^+(t)} [x_i^*(t) - y - v_0(y)t] \rho_0(y) dy = 0. \quad (19)$$

Производные по пределам пропадают в силу очевидных равенств:

$$\begin{cases} x_i^*(t) = X(y_i^+(t), t) = y_i^+(t) + v_0(y_i^+(t))t, \\ x_i^*(t) = X(y_i^-(t), t) = y_i^-(t) + v_0(y_i^-(t))t, \end{cases} \quad (20)$$

означающих, что на границах интервала интегрирования подинтегральное выражение в (19) равно нулю. Дифференцирование же подинтегрального выражения дает требуемое соотношение (18).

1.4. Условие абсорбции

Отметим еще одно свойство неупругих слипаний: частицы могут лишь прилипнуть к макрочастицам, а не отрываться от них. Это действительно так, если выполнены условия абсорбции

$$v_i^-(t) \geq v_i^*(t) \geq v_i^+(t), \quad (21)$$

где

$$\begin{cases} v_i^-(t) = v_0(y_i^-(t)) = v_w(x_i^* - 0, t), \\ v_i^+(t) = v_0(y_i^+(t)) = v_w(x_i^* + 0, t) \end{cases}$$

— значения поля скорости слева и справа от разрывов. Покажем, что построенные выше обобщенные решения уравнений (5) выполняют условие (21). Для этого вначале вычтем первое равенство (20) из второго. После элементарных выкладок найдем полуразность значений поля скорости $v_w(x, t)$ слева и справа от разрыва:

$$V_i = \frac{v_i^- - v_i^+}{2} = \frac{y_i^+ - y_i^-}{2t} > 0. \quad (22)$$

Отсюда видно, что скорость v_i^- частиц слева от разрыва всегда больше скорости v_i^+ справа от него. Осталось доказать, что скорость самого разрыва (макрочастицы) лежит в интервале $v_i^* \in [v_i^+, v_i^-]$. Сложив равенства (19), будем иметь

$$2x_i^* = y_i^+ + y_i^- + (v_i^+ + v_i^-)t.$$

Подставив сюда (17) и выразив из полученного соотношения скорость макрочастицы, найдем

$$v_i^* = U_i + W_i, \quad (23)$$

где

$$U_i = \frac{v_i^+ + v_i^-}{2}, \quad W_i = \frac{1}{t} \left[\frac{y_i^+ + y_i^-}{2} - x_i^c \right]. \quad (24)$$

Поскольку центр масс частиц, заключенных в интервале $[y_i^-, y_i^+]$, подчиняется очевидным неравенствам

$$y_i^- \leq x_i^c \leq y_i^+,$$

то из (24) и (22) имеем

$$-\frac{V_i}{2} \leq W_i \leq \frac{V_i}{2} \quad \Rightarrow \quad v_i^+ \leq U_i + W_i \leq v_i^-.$$

Последние же неравенства эквивалентны условию абсорбции (21).

2. Газ с давлением

Найденные выше поля скорости и плотности не учитывали давления. Обсудим модельные уравнения газа с давлением.

2.1. Уравнения политропного газа

Уравнения движения одномерного политропного газа, учитывающие давление, вязкость и молекулярную диффузию, таковы:

$$\begin{cases} \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + \frac{1}{\rho} \frac{\partial \mathcal{P}}{\partial x} = \frac{\nu}{\rho} \frac{\partial^2 v}{\partial x^2}, \\ \frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x} = \mu \frac{\partial^2 \rho}{\partial x^2}. \end{cases} \quad (25)$$

Здесь ν и μ — соответственно коэффициенты вязкости и молекулярной диффузии, а \mathcal{P} — давление газа, равное

$$\mathcal{P}(x, t) = \frac{\kappa^2}{\gamma} \rho^\gamma(x, t), \quad \kappa > 0, \quad \gamma > 0. \quad (26)$$

Решения уравнений (25) подчиняются закону сохранения импульса

$$\int_{-\infty}^{\infty} \rho(x, t) v(x, t) dx = \text{const}, \quad (27)$$

а при $\nu = 0$, $\mu = 0$ и $\kappa = 0$ система (25) переходит в (5). Поэтому естественно трактовать указанные выше обобщенные решения уравнений (5) как слабый предел при $\nu \rightarrow 0$, $\mu \rightarrow 0$ и $\kappa \rightarrow 0$ решений уравнений (25). Ниже будем, в первую очередь, интересоваться влиянием давления ($\kappa \neq 0$) в пределе $\nu \rightarrow 0$ и $\mu \rightarrow 0$. Решение уравнений (25) неизвестно. Поэтому заменим их модельными уравнениями

$$\begin{cases} \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + \frac{1}{\rho} \frac{\partial \mathcal{P}}{\partial x} = \mu \frac{\partial^2 v}{\partial x^2}, \\ \frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x} = \mu \frac{\partial^2 \rho}{\partial x^2}, \end{cases} \quad (28)$$

и обсудим свойства их решений при $\mu \rightarrow 0$. Критерием физической адекватности уравнений (28) поведению политропного газа будет служить выполнение инварианта (27).

2.2. Газ с показателем политропы $\gamma = 3$

Найдем общее решение уравнений (28) при $\gamma = 3$. Для этого перепишем (28) в виде

$$\begin{cases} \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + c \frac{\partial c}{\partial x} = \mu \frac{\partial^2 v}{\partial x^2}, \\ [2mm] \frac{\partial c}{\partial t} + v \frac{\partial c}{\partial x} + c \frac{\partial v}{\partial x} = \mu \frac{\partial^2 c}{\partial x^2}. \end{cases} \quad (29)$$

Здесь введено поле локальной скорости звука

$$c(x, t) = \kappa \rho(x, t). \quad (30)$$

Введем еще два вспомогательных поля

$$u_{\pm}(x, t) = v(x, t) \pm c(x, t), \quad (31)$$

такие, что скорость и плотность выражаются через них равенствами

$$v(x, t) = \frac{u_+(x, t) + u_-(x, t)}{2}, \quad \rho(x, t) = \frac{u_+(x, t) - u_-(x, t)}{2\kappa}. \quad (32)$$

Складывая и вычитая уравнения (29), обнаруживаем, что вспомогательные поля (31) удовлетворяют уравнению Бюргерса

$$\frac{\partial u_{\pm}}{\partial t} + u_{\pm} \frac{\partial u_{\pm}}{\partial x} = \mu \frac{\partial^2 u_{\pm}}{\partial x^2} \quad (33)$$

с начальными условиями

$$u_{\pm}(x, t = 0) = v_0(x) \pm \kappa \rho_0(x). \quad (34)$$

Физикам наиболее интересны решения (32) в частном случае

$$\rho(x, t = 0) = \rho_0 = \text{const}, \quad (35)$$

так как тогда решения (32) удовлетворяют закону сохранения импульса (27). В данном случае

$$u_{\pm}(x, t) = u(x \mp c_0 t, t) + c_0 \quad (c_0 = \kappa \rho_0), \quad (36)$$

где $u(x, t)$ удовлетворяет уравнению Бюргерса:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \mu \frac{\partial^2 u}{\partial x^2}, \quad u(x, t = 0) = v_0(x). \quad (37)$$

Подставив (36) в (32), выразим решения уравнений (29) через решение уравнения Бюргера (37):

$$\begin{cases} v(x, t) = \frac{1}{2} [u(x - c_0 t, t) + u(x + c_0 t, t)] , \\ [2mm]\rho(x, t) = \rho_0 \left[1 - \frac{u(x + c_0 t, t) - u(x - c_0 t, t)}{2c_0} \right] . \end{cases} \quad (38)$$

Убедимся, что данные решения подчиняются закону сохранения импульса (27). Перемножив правые части (38), будем иметь

$$\begin{aligned} \rho(x, t) v(x, t) &= \frac{\rho_0}{2} [u(x - c_0 t, t) + u(x + c_0 t, t)] + \\ &\quad \frac{1}{2\kappa} [u^2(x - c_0 t, t) - u^2(x + c_0 t, t)] . \end{aligned}$$

После интегрирования этого равенства по всем x , последнее слагаемое, в силу симметрии, пропадет. Интеграл же первого слагаемого справа постоянен, так как уравнение Бюргера имеет инвариант:

$$\int_{-\infty}^{\infty} u(x, t) dx = \int_{-\infty}^{\infty} v_0(x) dx = \text{const} .$$

Таким образом, полный импульс модельного газа равен

$$\int_{-\infty}^{\infty} \rho(x, t) v(x, t) dx = \rho_0 \int_{-\infty}^{\infty} v_0(x) dx = \text{const} .$$

2.3. Обобщенные решения уравнений политропного газа

Напомним, решение уравнения (37) имеет вид, аналогичный (2):

$$v(x, t) = \frac{x - \{y\}(x, t)}{t} , \quad (39)$$

где фигурные скобки означают пространственное усреднение с помощью распределения

$$f(y; x, t) = \frac{\exp \left[-\frac{1}{2\mu} \mathcal{G}(y; x, t) \right]}{\int_{-\infty}^{\infty} \exp \left[-\frac{1}{2\mu} \mathcal{G}(y; x, t) \right] dy} , \quad (40)$$

а \mathcal{G} — потенциал (6) в случае $\rho_0 = \text{const}$, равный

$$\mathcal{G}(y; x, t) = s_0(y)t + \frac{1}{2}(y - x)^2 , \quad s_0(y) = \int^y v_0(z) dz . \quad (41)$$

При $\mu \rightarrow 0$ среднее $\{y\}$ сходится к $y(x, t)$ — координате абсолютного минимума потенциала \mathcal{G} (41), а выражения (38) принимают вид

$$\begin{cases} v_w(x, t) = \frac{2x - y(x - c_0 t) - y(x + c_0 t)}{2t}, \\ [2mm]\rho_0(x, t) = \rho_0 \frac{y(x + c_0 t) - y(x - c_0 t)}{2ct}, \end{cases} \quad (42)$$

и дают обобщенные, подчиняющиеся закону сохранения импульса (27) решения уравнений политропного газа

$$\begin{cases} \frac{\partial \rho v}{\partial t} + \frac{\partial \rho v^2}{\partial x} + \frac{\partial \mathcal{P}}{\partial x} = 0, \\ [2mm]\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x} = 0, \end{cases} \quad (43)$$

при $\gamma = 3$ и $\rho_0 = \text{const}$.

3. Перенос активной примеси

В завершение краткого обзора простейших моделей газовой динамики обсудим еще одну, точно решаемую систему нелинейных уравнений, моделирующую перенос активной примеси.

3.1. Подстановка Хопфа — Коула

Напомним, подстановка Хопфа — Коула сводит нелинейное уравнение Бюргерса к линейному уравнению диффузии (см., например, [2]). Именно поэтому хорошо изучены свойства решений уравнения Бюргерса. Оказывается, родственная подстановка, примененная к паре линейных уравнений диффузии, дает решения нелинейных уравнений, моделирующих перенос активной примеси [5, 6]. Покажем это, взяв систему диффузионных уравнений:

$$\begin{cases} \frac{\partial \varphi}{\partial t} = \mu \Delta \varphi + a \varphi + b \psi, \\ [2mm]\frac{\partial \psi}{\partial t} = \mu \Delta \psi + d \varphi + e \psi. \end{cases} \quad (44)$$

Применим к первому уравнению подстановку Хопфа — Коула, умножив его на $2\mu/\varphi$. В итоге получим:

$$\frac{\partial h}{\partial t} = \frac{1}{2} (\nabla h)^2 + \mu \Delta h + 2\mu a + 2\mu b N. \quad (45)$$

Здесь использованы обозначения:

$$h(x, t) = 2\mu \ln \varphi(x, t), \quad N(x, t) = \frac{\psi(x, t)}{\varphi(x, t)}. \quad (46)$$

Найдем уравнение, которому подчиняется $N(\mathbf{x}, t)$. Очевидно,

$$\frac{\partial N}{\partial t} = \frac{1}{\varphi} \frac{\partial \psi}{\partial t} - \frac{\psi}{\varphi^2} \frac{\partial \varphi}{\partial t}. \quad (47)$$

Подставив сюда правые части уравнений (44), будем иметь

$$\frac{\partial N}{\partial t} = \mu \left(\frac{\Delta \psi}{\varphi} - \frac{\psi \Delta \varphi}{\varphi^2} \right) + d + (e - a) N - b N^2. \quad (48)$$

Выразим через N и h первое слагаемое справа. Для этого заметим, что имеет место аналогичное (47) равенство:

$$\nabla N = \frac{\nabla \psi}{\varphi} - \frac{\psi \nabla \varphi}{\varphi^2}. \quad (49)$$

Скалярно умножив его на вектор набла ∇ , получим

$$\Delta N = \left(\frac{\Delta \psi}{\varphi} - \frac{\psi \Delta \varphi}{\varphi^2} \right) - \frac{2}{\varphi} \left[\frac{(\nabla \varphi \cdot \nabla \psi)}{\varphi} - \frac{\psi (\nabla \varphi \cdot \nabla \varphi)}{\varphi^2} \right]. \quad (50)$$

Принимая во внимание (49), а также выражение

$$\nabla h = 2\mu \frac{\nabla \varphi}{\varphi},$$

перепишем (50) в форме

$$\frac{\Delta \psi}{\varphi} - \frac{\psi \Delta \varphi}{\varphi^2} = \Delta N + \frac{1}{\mu} (\nabla h \cdot \nabla) N.$$

С учетом этого соотношения уравнение (48) преобразуется к виду

$$\frac{\partial N}{\partial t} = \mu \Delta N + (\nabla h \cdot \nabla) N + d + (e - a) N - b N^2, \quad (51)$$

содержащему только поля h и N .

Введем потенциальное векторное поле

$$\mathbf{v}(\mathbf{x}, t) = -\nabla h(\mathbf{x}, t), \quad (52)$$

а также физически более прозрачные обозначения

$$a = \frac{1}{2\mu} F(\mathbf{x}, t), \quad d = g(\mathbf{x}, t), \quad e = \kappa + \frac{1}{2\mu} F(\mathbf{x}, t). \quad (53)$$

В итоге приходим к системе уравнений

$$\begin{cases} \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \mu \Delta \mathbf{v} + \mathbf{f}(\mathbf{x}, t) - 2\mu b \nabla N \\ [2mm] \frac{\partial N}{\partial t} + (\mathbf{v} \cdot \nabla) N = \mu \Delta N + g(\mathbf{x}, t) + \kappa N - b N^2 \end{cases} \quad (54)$$

для полей $\mathbf{v}(\mathbf{x}, t)$ и $N(\mathbf{x}, t)$. Первое из них естественно трактовать как поле скорости некоторой сплошной среды, а второе – как поле концентрации примеси, переносимой средой. Здесь $g(\mathbf{x}, t)$ источник примеси, а $\mathbf{f}(\mathbf{x}, t) = -\nabla F(\mathbf{x}, t)$ – сила, порождающая потенциальное поле скорости $\mathbf{v}(\mathbf{x}, t)$. Последние два слагаемых во 2-м уравнении могут описывать, например, процессы химических реакций. Последнее слагаемое в 1-м уравнении (54) учитывает «силы давления» в областях повышенной концентрации.

Из приведенных выкладок следует, что решение пары уравнений (54) с начальными условиями

$$\mathbf{v}(\mathbf{x}, t = 0) = \mathbf{v}_0(\mathbf{x}), \quad C(\mathbf{x}, t = 0) = C_0(\mathbf{x}), \quad (55)$$

подстановками

$$\mathbf{v} = -2\mu \frac{\nabla \varphi}{\varphi}, \quad C = \frac{\psi}{\varphi} \quad (56)$$

сводится к решению линейных уравнений

$$\begin{cases} \frac{\partial \varphi}{\partial t} = \mu \Delta \varphi + \frac{1}{2\mu} F(\mathbf{x}, t) \varphi + b \psi \\ [2mm] \frac{\partial \psi}{\partial t} = \mu \Delta \psi + \frac{1}{2\mu} F(\mathbf{x}, t) \psi + g(\mathbf{x}, t) \varphi + \kappa \psi \end{cases} \quad (57)$$

с начальными условиями

$$\varphi_0(\mathbf{x}, t = 0) = \exp\left(-\frac{s_0(\mathbf{x})}{2\mu}\right), \quad \psi_0(\mathbf{x}) = C_0(\mathbf{x}) \exp\left(-\frac{s_0(\mathbf{x})}{2\mu}\right). \quad (58)$$

Здесь $s_0(\mathbf{x})$ – потенциал начального поля скорости, связанный с ним равенством $\mathbf{v}(\mathbf{x}) = \nabla s_0(\mathbf{x})$.

В частности, в случае пассивной примеси ($b = \kappa = 0$), при отсутствии внешней силы $\mathbf{f} \equiv 0$ и источников примеси $g \equiv 0$ решения уравнений (54) принимают вид

$$\mathbf{v}(\mathbf{x}, t) = \frac{\mathbf{x} - \{\mathbf{y}\}(\mathbf{x}, t)}{t}, \quad C(\mathbf{x}, t) = \{C_0(\mathbf{y})\}(\mathbf{x}, t). \quad (59)$$

Здесь использована операция пространственного усреднения с помощью многомерного аналога распределения (40) [5, 6].

Работа выполнена при финансовой поддержке Гранта «Ведущие Научные Школы» N 00-15-96619, Гранта «Университеты России», а также Грантов РФФИ N 00-02-16167 и N 00-02-17347.

Литература

1. *Shandarin S. F., Zeldovich Ya. B.* The large-scale structures of the universe: Turbulence, intermittency, structures in a self-gravitating medium // *Rev. Mod. Phys.* 1989. V. 61. P. 185–220.
2. *Gurbatov S., Malakhov A., Saichev A.* Nonlinear random waves and turbulence in nondispersive media: waves, rays and particles. Manchester — New York: Manchester University Press, 1991. 308 p.
3. *Vergassola M., Dubrulle B., Frisch U., Noullez A.* Burgers' equation, Devil's staircases and the mass distribution for the large-scale structures // *Astron. Astrophys.* 1994. V. 289. P. 325–356.
4. *E W., Rykov Yu. G., Sinai Ya. G.* Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics // *Comm. Math. Phys.* 1996. V. 177. P. 349–380.
5. *Saichev A. I., Woyczynski W. A.* Density fields in burgers and KdV-burgers turbulence // *SIAM J. Appl. Math.* 1996, V. 56, N 4. P. 1008–1038.
6. *Saichev A. I., Woyczynski W. A.* Advection of passive and reactive tracers in multi-dimensional Burgers' velocity field // *Physica D.* 1997. V. 100. P. 119–141.

CONTENTS

<i>Preface</i>	5
----------------------	---

Information aspects of nonlinear dynamics

<i>V. B. Kazantsev, V. I. Nekorkin.</i> Dynamics of oscillatory neurons. Information aspects	9
<i>Ya. I. Molkov, A. M. Feigin.</i> Prognosis of qualitative behavior of a dynamical system by chaotic time series	34
<i>A. S. Dmitriev.</i> Dynamical chaos and information	53
<i>V. D. Shalfeev, V. V. Matrosov.</i> Chaotically modulated oscillations in coupled phase systems	77
<i>V. G. Yakhno.</i> Models of neuron-like systems. Dynamical regimes of transformation of information	90

General problems of nonlinear dynamics

<i>G. S. Golitsyn.</i> White noise as a basis for explaining many statistical regularities in the nature	117
<i>A. M. Fridman, A. V. Klimenko.</i> Two components of seismic activity of the Earth and their relation to features of diurnal rotation	133
<i>E. A. Kuznetsov.</i> Collapse in hydrodynamics	156
<i>Yu. I. Troitskaya.</i> Nonlinear waves and structures in stratified shear flows.....	179
<i>A. B. Ezersky.</i> Bound states and spatio-temporal chaos of topological defects	189
<i>E. N. Pelinovsky.</i> Nonlinear models of tsunami generation by moving sources	199
<i>E. I. Yakubovich.</i> A new approach to the hydrodynamics of an incompressible liquid.....	211
<i>S. Flach, A. E. Miroshnichenko.</i> Discrete breathers	223

Nonlinear phenomena in plasma and optics

<i>M. Yu. Ryabikin, A. M. Sergeev.</i> Dynamics of an atom in a superstrong field and generation of attosecond pulses.....	235
<i>M. D. Tokman, A. G. Litvak, M. A. Erukhimova, A. Yu. Kryachko.</i> Stimulated emission without inversion and electromagnetically induced transparency in a plasma and in electron fluxes.....	250
<i>G. M. Fraiman, A. A. Balakin.</i> Electron–ion collisions in a strong laser field.....	268
<i>V. Yu. Trakhtengerts, D. I. Iudin, A. N. Grigoriev.</i> On fractal dynamics of active media.....	287
<i>L. S. Mkhitarian.</i> Nonlinear stage of development of disturbances at shell acceleration.....	303
<i>G. G. Kochemasov.</i> Some nonlinear problems of laser fusion.....	314
<i>A. S. Kingsep.</i> Nonlinear waves in electron magnetic hydrodynamics.....	329

Nonlinear phenomena in medicine and biology

<i>V. L. Voeikov.</i> Nonlinear oscillations at molecular and cellular levels in biological systems: possible sources of origin and functional role in processes of vital functions.....	345
<i>Yu. M. Romanovsky, A. V. Netrebko, O. A. Chichigina.</i> The problem of damping of subglobular oscillations of protein molecules in water.....	359
<i>Yu. I. Neimark.</i> Mathematical model of immune response of the organism to the invasion of infection.....	372

Mathematical problems of nonlinear dynamics

<i>N. G. Makarenko.</i> Fractals, multifractal measures and attractors.....	381
<i>A. P. Kuznetsov, S. P. Kuznetsov, I. R. Sataev.</i> Critical behavior in the transition to chaos through period doubling. Model mapping and renormgroup analysis.....	395
<i>A. I. Saichev, S. G. Utkin.</i> The laws of anomalous diffusion.....	416
<i>S. N. Gurbatov, A. I. Saichev.</i> Simplest models of the flows of particles.....	432

NONLINEAR WAVES' 2002

Edited by A. V. Gaponov-Grekhov and V. I. Nekorkin

This volume includes papers that are based on the lectures given at the XI-th School in Nonlinear Waves (Nizhny Novgorod, 2002). Collective dynamics of complex systems, processing, transformation and transmission of information, wave dynamics, nonlinear optics, nonlinear phenomena in plasma, and other problems of nonlinear science are considered.

The book is intended for specialists involved with investigation of nonlinear phenomena, as well as for graduate and postgraduate students of the corresponding specialities.

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2002

Утверждено к печати Институтом прикладной физики РАН,
603950 Нижний Новгород, ул. Ульянова, 46

Редактор *Н. Н. Кралина*
Верстка *М. В. Башевой, С. Н. Новиковой*
Корректоры *Л. О. Кузнецова, Д. П. Семенова*

Подписано к печати 18.12.2002 г. Формат 60 x 90 1/16. Бумага офсетная № 1.
Усл. печ. л. 28. Уч.-изд. л. 25,4. Темплан 2003 г. Поз. 1. Тираж 400 экз.

Отпечатано в полном соответствии с качеством
предоставленных диапозитивов в ППП «Типография «Наука»
121099, Москва, Шубинский пер., 6

Заказ № 7317