

НЕЛИНЕЙНЫЕ
ВОЛНЫ' 2006



Российская академия наук
Институт прикладной физики

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2006

Нижний Новгород
2007

Издано по решению редакционно-издательского совета
Института прикладной физики РАН

УДК 534.222.2
ББК 22.312я4

Н49 **Нелинейные волны' 2006** / Отв. ред. А.В. Гапонов-Грехов, В.И. Некоркин. – Нижний Новгород: ИПФ РАН, 2007. – 576 с.

ISBN 978-5-8048-0061-2

В сборник включены обзорные лекции и оригинальные статьи, написанные по материалам лекций, прочитанных на XXIII школе по нелинейным волнам (Нижний Новгород, 1 – 7 марта 2006 г.). Рассматриваются проблемы пространственно-временного хаоса, структурообразования, волновой динамики, нелинейные явления в плазме, оптике и при взаимодействии сверхсильных полей с веществом, математические проблемы нелинейной динамики и другие аспекты нелинейно-волновой тематики.

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

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

Рецензенты
доктор физико-математических наук *В.Н. Белых*,
доктор физико-математических наук *С.Н. Гурбатов*,
доктор физико-математических наук *Л.М. Лерман*,
доктор физико-математических наук *К.А. Постнов*,
доктор физико-математических наук *Г.Ю. Ризниченко*,
доктор физико-математических наук *А.И. Смирнов*

*Проведение школы и издание сборника осуществлены при поддержке
Российской академии наук,
Министерства образования и науки Российской Федерации,
Федерального агентства по науке инновациям,
Российского фонда фундаментальных исследований,
Фонда некоммерческих программ "Династия"*

ISBN 978-5-8048-0061-2

© Институт прикладной физики РАН, 2007 г.

СОДЕРЖАНИЕ

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

Нелинейная динамика сложных систем

<i>С.Г. Флах.</i> Периодические орбиты, локализация в пространстве нормальных мод и проблема Ферми, Паста и Улама.....	9
<i>С.Г. Флах.</i> Симметрии и их нарушения в системах во внешних переменных полях	24
<i>А.С. Пиковский, Ф. Розенау.</i> Фазовые компактоны.....	39
<i>А.Б. Езерский.</i> Пространственно-временная динамика топологических особенностей волновых полей: источники, стоки, дырки, дефекты	55
<i>С.П. Кузнецов.</i> О реализации некоторых классических моделей и феноменов нелинейной динамики на основе связанных неавтономных осцилляторов	68
<i>В.В. Клиньшов, В.И. Некоркин.</i> Фазовые кластеры и краткосрочная память в сети нейронов с последеполаризацией.....	85
<i>А.А. Короновский, А.Е. Храмов.</i> Непрерывный вейвлетный анализ и некоторые его приложения к задачам нелинейной динамики	102
<i>А.С. Дмитриев.</i> Системы связи с хаотическими носителями: проблемы, достижения и перспективы.....	118

Нелинейные колебания в структурированных системах

<i>О.В. Руденко.</i> Математические модели, точные решения, методы анализа.....	139
<i>О.В. Руденко.</i> Актуальные проблемы, связанные с нелинейной акустикой	151
<i>В.Ю. Зайцев.</i> О «неклассических» проявлениях микроструктурно-обусловленной акустической нелинейности на примере контактодержащих сред.....	170
<i>В.В. Курин, А.В. Чигинев.</i> Динамика вихрей в слоистых сверхпроводниках с внутренним эффектом Джозефсона	191
<i>Ф.А. Стариков.</i> Вынужденное рассеяние Мандельштама — Бриллюэна лазерных мод Гаусса — Лагерра: новые явления.....	206
<i>Г.М. Фрайман, А.А. Балакин.</i> Генерация быстрых электронов при электрон-ионных столкновениях в сильных полях	222

Нелинейные эффекты в гравитации и астрофизической плазме

<i>Ю.М. Зиновьев.</i> Гравитация и супергравитация.....	239
<i>Ю.М. Зиновьев.</i> Физика частиц на ускорителях и гравитация.....	248
<i>В.С. Бескин.</i> «Центральная машина» в активных галактических ядрах	254

<i>Е.В. Деришев, В.В. Кочаровский, Вл.В. Кочаровский, В.Ю. Мартьянов.</i> Физика релятивистских джетов	268
<i>И.Н. Топтыгин.</i> Самоорганизация плазмы и магнитного поля в неравновесных астрофизических объектах	289
<i>Ю.А. Фадеев.</i> Пульсации звезд	304
<i>Ю.А. Фадеев.</i> Ударные волны в атмосферах пульсирующих звезд	320

Нелинейные явления в окружающей среде

<i>Е.А. Мареев, В.В. Клименко.</i> Нелинейная динамика высотного грозового разряда	337
<i>Д.И. Иудин, В.Ю. Трахтенгерц.</i> Фрактальные лабиринты: структурная динамика	360
<i>М.А. Носов.</i> Волны цунами: общая характеристика явления, механизмы образования, прогноз	378
<i>Е.Н. Пелиновский.</i> Нелинейно-дисперсионная теория волн цунами: взгляд после катастрофического цунами в Индийском океане	393
<i>Ю.И. Троицкая, С.А. Ермаков.</i> Нелинейные механизмы и модели формирования спутниковых изображений процессов в верхнем слое океана	408

Математические проблемы нелинейной динамики

<i>В.Ф. Ковалев, Д.В. Ширков.</i> Ренормгрупповые симметрии в краевых задачах	433
<i>А.Д. Морозов.</i> Вырожденные резонансы	447
<i>А.И. Нейштадт.</i> Затягивание потери устойчивости при динамических бифуркациях	461

Нелинейные явления в живых системах

<i>А.А. Веденов.</i> Бактериальная клетка с физической точки зрения	479
<i>С.А. Ларионов, А.Ю. Лоскутов, С.Д. Рыбалко, Е.В. Рядченко.</i> Геном как фрактальное множество, порождаемое инверсионно-диффузионной динамикой	491
<i>Е.А. Катруха, Г.Т. Гурия.</i> Роль неустойчивостей, нелинейных колебаний и волн в регуляции тубулинового цитоскелета	509
<i>В.Л. Воейков, С.Ф. Чалкин, Р.Р. Асфарамов.</i> Аэроионы Чижевского и явление свечения влажного воздуха	535
<i>Ю.Н. Кульчин, О.А. Букин, С.С. Вознесенский, А.Н. Галкина, С.В. Гнеденков, А.Л. Дроздов, В.Г. Курявый, Т.Л. Мальцева, С.Л. Синячихин, А.И. Чередничков.</i> Биологический вид волоконных световодов	548

ПРЕДИСЛОВИЕ

Настоящий сборник составлен из обзорных и оригинальных статей, написанных по материалам лекций, прочитанных на XIII научной школе «Нелинейные волны' 2006». Традиционная тематика нижегородских (горьковских) зимних школ по нелинейным волнам является достаточно широкой и охватывает большой круг нелинейных явлений в самых разнообразных областях современной науки, а не только физике и математике. Не явилась исключением и нынешняя школа (Нижний Новгород, март 2006 г.). На наш взгляд, все наиболее актуальные современные направления в нелинейной физике в той или иной степени на школе были представлены. Значительное внимание было уделено проблемам пространственно-временного хаоса, синхронизации, структурообразования, волновой динамики, нелинейных явлений в акустике, оптике, астрофизике и физике плазмы, теории гравитации и супергравитации, нелинейным явлениям в живых системах и окружающей среде и др. Традиционно на школе был прочитан небольшой цикл лекций по математическим проблемам нелинейной динамики.

Направления развития интересов теории нелинейных волн и колебаний, обозначившиеся на двух предыдущих школах, ясно проявились и на нынешней. Относительно новые для нелинейной динамики области – астрофизика, окружающая среда, живые системы и др. привлекают все большее внимание исследователей. Наиболее ярко эта тенденция проявилась в тематике, связанной с явлениями в живых системах. Например, в лекции А. А. Веденова принята попытка рассмотреть с физической точки зрения бактериальную клетку, в лекции А. Ю. Лоскутова на основе представления последовательности ДНК как двумерного блуждания предложен новый подход к изучению различных участков хромосом, в лекции Г. Т. Гурия показана принципиальная роль нелинейных колебаний и волн в регуляции тубулинового цитоскелета клетки и т. д. Такая тенденция развития объясняется, по-видимому, как принципиальной значимостью нелинейных явлений в природе, так и впечатляющими успехами общей теории динамических систем, достигнутыми к настоящему времени. Значительный прогресс

в понимании общих закономерностей формирования диссипативных пространственно-временных структур и механизмов хаотического поведения динамических систем сформировал своего рода «плацдарм», позволивший перейти к изучению сложных динамических процессов в самых разнообразных областях науки. Можно ожидать, что завоевание теорией нелинейных волн и колебаний новых областей приложения продолжится и возникнут новые, совершенно неожиданные «горячие» проблемы. Возможно, в недалёком будущем это произойдет в сфере нанонауки и нанотехнологий.

В сборник не вошли материалы лекций В. Н. Белых «Проблемы синхронизации в сетях квазиидентичных хаотических динамических систем с асимметричными связями и переменной структурой», Л. М. Зеленого «Фрактальная динамика плазменных процессов в хвосте магнитосферы Земли», Н. Г. Макаренко «Фракталы, стохастическая динамика, марковские процессы и прогноз», А. М. Сергеева «Биомедицинские приложения сверхсильных лазерных полей», Д. И. Трубецкова «2005 год в датах нелинейной динамики» и А. М. Фейгина «Нелинейная динамическая история озонной дыры», опубликованные ранее в других изданиях, хотя они также вызвали большой интерес слушателей.

В заключение мы хотим выразить свою признательность авторам за предоставление материалов в этот сборник.

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

Нелинейная динамика сложных систем



ПЕРИОДИЧЕСКИЕ ОРБИТЫ, ЛОКАЛИЗАЦИЯ В ПРОСТРАНСТВЕ НОРМАЛЬНЫХ МОД И ПРОБЛЕМА ФЕРМИ, ПАСТА И УЛАМА

С. Г. Флах

1. Введение

Одна из основных задач физики конденсированных сред заключается в характеристике свойств возбуждений данной системы. Обычное квазичастичное возбуждение, как известно, может быть получено с помощью резонансного взаимодействия с внешними переменными полями. В то время как в квантовых системах сложность проявляется в правильном описании многочастичных волновых функций, в классической нелинейной динамике многочастичных систем трудности могут возникнуть уже на более ранней стадии, так как необходимо определить классы траекторий в фазовом пространстве, эволюция которых будет описывать интересующие возбуждения. Поскольку типичная классическая гамильтонова система неинтегрируема, то почти для всех начальных условий она будет вести себя хаотично на достаточно больших временах. Гамильтонов хаос на больших временных шкалах описывает распад возбуждений, а сами возбуждения характеризуются более регулярными траекториями. Действительно, Колмогоров, Арнольд и Мозер (КАМ) показали, что регулярная динамика на инвариантных N -мерных торах выживает во многих неинтегрируемых системах с конечным числом степеней свободы N . Однако при устремлении этого числа к бесконечности возникают проблемы, ибо частотные спектры могут стать сплошными, и тогда расходящиеся малые знаменатели приводят к исчезновению КАМ-торов. Выход простой: необходимо рассматривать низкоразмерные структуры, аналогичные КАМ-торам. Самая простая из них — это периодическая орбита (ПО). Хорошо известно, что периодические орбиты существуют в сильно хаотических системах и даже плотно заполняют хаотические фазовые объемы, несмотря на то что имеют меру ноль. Именно периодические орбиты и являются самыми естественными регулярными траекториями, описывающими возбуждения даже в макроскопических случаях. Примером таких возбуждений служат дискретные бризеры. Это периодические решения нелинейных решеточных систем, локализованные в прямом пространстве и существующие в бесконечных решетках [1]. Отметим, что дискретные бризеры характеризуются конечной энергией возбуждения.

2. Проблема Ферми, Паста и Улама (ФПУ)

Энрико Ферми, Джон Паста и Станислав Улам в 1953 году рассмотрели задачу о перераспределении энергии между нормальными модами ангармонической атомной цепочки и опубликовали результаты во внутреннем отчете национальной лаборатории в Лос-Аламосе [2]. Уравнения движения цепочки описывают N частиц с нелинейным взаимодействием между ближайшими соседями, содержащим либо квадратичные (α -ФПУ модель)

$$\ddot{x}_n = (x_{n+1} - 2x_n + x_{n-1}) + \alpha[(x_{n+1} - x_n)^2 - (x_n - x_{n-1})^2], \quad (1)$$

либо кубические члены (β -ФПУ модель)

$$\ddot{x}_n = (x_{n+1} - 2x_n + x_{n-1}) + \beta[(x_{n+1} - x_n)^3 - (x_n - x_{n-1})^3]. \quad (2)$$

Переменные $x_n(t)$ описывают смещения n -й частицы из соответствующего положения равновесия. Рассматриваются жесткие граничные условия $x_0 = x_{N+1} = 0$. Каноническое преобразование

$$x_n(t) = \sqrt{\frac{2}{N+1}} \sum_{q=1}^N Q_q(t) \sin\left(\frac{\pi q n}{N+1}\right)$$

позволяет диагонализировать и решить линейную задачу $\alpha = \beta = 0$ при помощи координат нормальных мод $Q_q(t)$. Модовое число $q = 1, \dots, N$ определяет соответствующие частоты нормальных мод $\omega_q = 2 \sin(\pi q/2(N+1))$. Уравнения движения принимают вид

$$\ddot{Q}_q + \omega_q^2 Q_q = -\frac{\alpha}{\sqrt{2(N+1)}} \sum_{l,m=1}^N \omega_q \omega_l \omega_m B_{q,l,m} Q_l Q_m \quad (3)$$

для α -ФПУ цепочки (1) и

$$\ddot{Q}_q + \omega_q^2 Q_q = -\frac{\beta}{2(N+1)} \sum_{l,m,n=1}^N \omega_q \omega_l \omega_m \omega_n C_{q,l,m,n} Q_l Q_m Q_n \quad (4)$$

для β -ФПУ цепочки (2). Коэффициенты взаимодействия

$$B_{q,l,m} = \sum_{\pm} (\delta_{q \pm l \pm m, 0} - \delta_{q \pm l \pm m, 2(N+1)}), \quad (5)$$

$$C_{q,l,m,n} = \sum_{\pm} (\delta_{q \pm l \pm m \pm n, 0} - \delta_{q \pm l \pm m \pm n, 2(N+1)} - \delta_{q \pm l \pm m \pm n, -2(N+1)}) \quad (6)$$

связывают нелинейным образом нормальные моды и не имеют характерного радиуса взаимодействия в пространстве мод. В отсутствие этого взаимодействия имеется N интегралов движения, которые соответствуют энергиям нормальных мод $E_q = \frac{1}{2} (\dot{Q}_q^2 + \omega_q^2 Q_q^2)$.

ФПУ рассмотрели эволюцию нелинейной цепочки с начальными условиями, когда возбуждена только самая низкочастотная мода $q_0 = 1$. Ожидалось, что через сравнительно короткое время (порядка нескольких периодов осцилляции возбужденной моды) начнут возбуждаться другие моды и в итоге энергия распределится равномерно по всем модам. Таким образом Ферми ожидал обнаружить переход к тепловому равновесию, с которым мы сталкиваемся в повседневной жизни. Вместо этого ФПУ обнаружили в точности обратную картину (рис. 1). Энергия остается локализованной на нескольких низкочастотных модах, а также видны возвраты почти всей энергии в исходно возбужденную моду на временах, существенно меньших, чем характерные времена цикла Пуанкаре для всей системы. Последующие исследования показали, что существуют пороги значения исходной энергии, а также размера системы, выше которых может быть достигнуто равномерное распределение (см., например, [5]).

На рис. 2 показаны распределения плотностей энергий мод $\epsilon_q = E_q/N$ для более высокой затравочной энергии в логарифмическом масштабе на временах $t = 10^4, 10^5, 10^6$. Поскольку энергии мод зависят от времени, плотности усреднены на предшествующем интервале времени $\Delta t = 10^4$. Во-первых, очевидно, что локализация энергии в пространстве мод экспоненциальная, и такое распределение устанавливается на довольно малом временном масштабе τ_1 . Во-вторых, видна медленная резонансная накачка мод на краю распределения, которая ведет систему в равновесие на большом временном масштабе $\tau_2 \gg \tau_1$. Вышеупомянутые пороги представляют собой значения энергии, на которых временной масштаб накачки края распределения становится сравнимым со временем установления самого распределения $\tau_2 \approx \tau_1$ [5]. Ниже этих порогов временные масштабы разведены, и эволюция на временах $\tau_1 \ll t \ll \tau_2$ соответствует наблюдениям ФПУ. В то время как ФПУ стояли перед проблемой объяснения отсутствия равномерного распределения, т. е. эргодичности, более поздние исследования позволяют ставить следующие вопросы: «Почему существуют два разных временных масштаба, как оценить длину локализации и как она связана с временными масштабами, сохраняется ли этот эффект в пределе бесконечно большой системы, а также при переходе к двух- и трехмерным системам?»

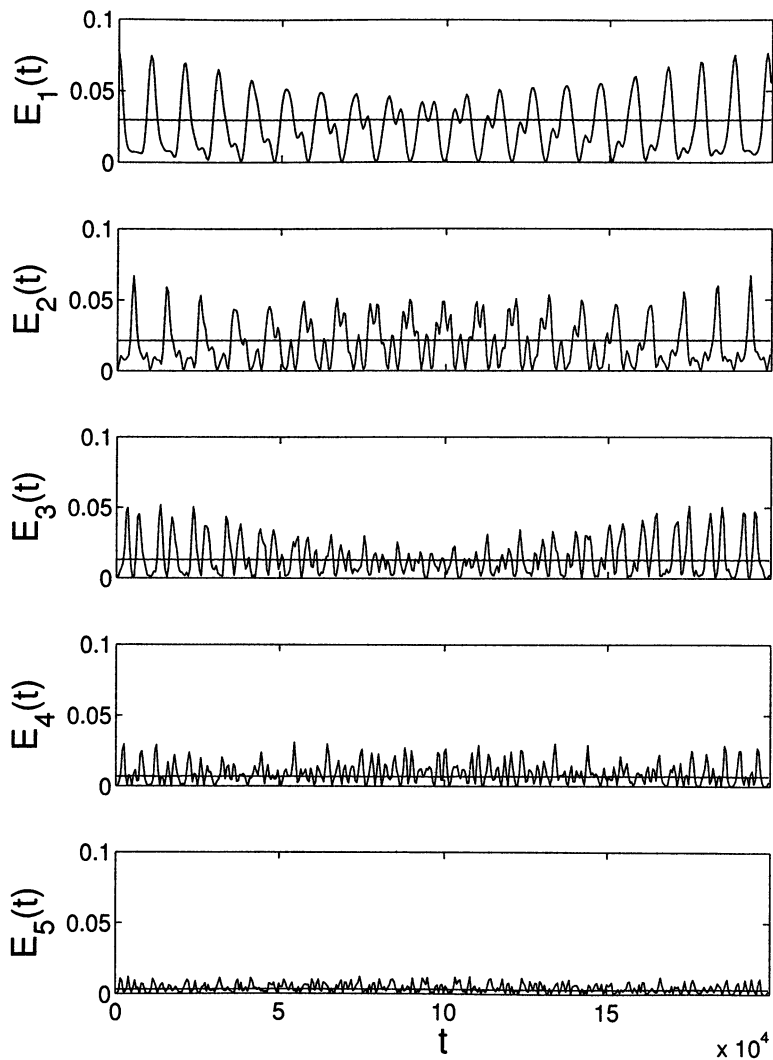


Рис. 1. Эволюция энергий мод $q = 1, 2, 3, 4, 5$ для траектории ФПУ (осциллирующие кривые) и точной ПО типа QV (горизонтальные линии). Параметры $\alpha = 0,25$, $E = 0,077$, $N = 32$ [3].

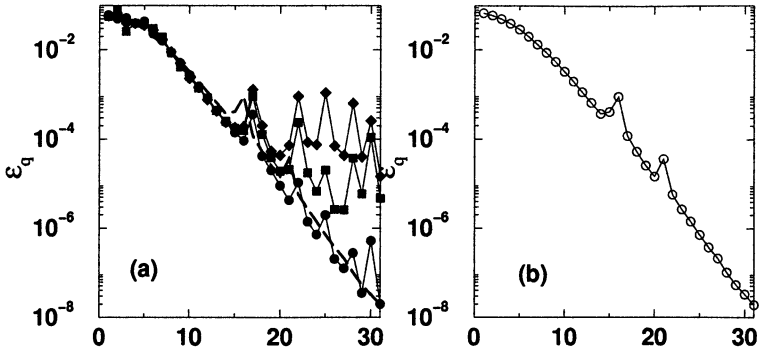


Рис. 2. Распределения плотностей энергий мод при значениях параметров $q_0 = 1$, $N = 31$, $\alpha = 0,33$, $E = 0,32$: для траектории ФПУ (a) (кружки — $t = 10^4$, квадраты — $t = 10^5$, ромбы — $t = 10^6$, штриховая линия — QB из (b) для сравнения); для QB (b) [4].

3. Периодические орбиты — q -breathers

Предположим, что траектория ФПУ на временах $t \ll \tau_2$ находится вблизи ПО. Тогда эволюция будет почти регулярной, что и наблюдается в расчетах. Следует ожидать, что такая ПО будет соответствовать слегка деформированной нормальной моде. Как было показано в работах [3, 6], существует строгая схема построения такой ПО начиная с простого линейного предела $\alpha = \beta = 0$. Рассматривается траектория с заданной энергией E и возбуждением только одной моды с модовым числом q_0 . Нетрудно убедиться, что эта схема однозначно выделяет периодическую орбиту в многомерном фазовом пространстве системы. Далее воспользуемся фактом отсутствия резонансов:

$$\omega_q \neq n\omega_{q_0} \quad (7)$$

для всех целых n и $q \neq q_0$ [7]. В работе [3] используется теорема Ляпунова [8] и доказывается, что исходная ПО продолжает существовать при учете нелинейных членов и фиксированной энергии E . Полученная ПО вовлекает в свою динамику большое количество нормальных мод. Распределение модовых амплитуд и модовых энергий экспоненциально локализовано (см. рис. 1 и рис. 2, b). По аналогии с дискретными бризерами эти решения получили название q -breather (QB), которое подчеркивает, что решение является периодическим во времени и локализованным в пространстве мод.

Численные методы построения QB описаны в работах [3, 6]. Они используют тот факт, что в обобщенных отображениях (сечениях) Пуанкаре ПО соответствует неподвижной точке. Произвольно выбранная начальная точка отобразится на другую точку, и компоненты вектора невязки являются функциями компонент начальной точки. Обобщенный метод Ньютона используется для нахождения нулей всех компонент вектора невязки и, следовательно, для нахождения неподвижной точки — ПО.

3.1. Результаты для α -ФПУ модели, низкие частоты

Применение метода теории возмущения Пуанкаре — Линдштедта для низкочастотных мод приводит к оценке распределения энергий нормальных мод на QB [6]:

$$E_{nq_0} = \gamma^{2n-2} n^2 E_{q_0}, \quad \gamma = \frac{\alpha \sqrt{E_{q_0}} (N+1)^{3/2}}{\pi^2 q_0^2}. \quad (8)$$

Необходимое условие для локализации QB в рамках этой оценки — $\gamma < 1$. Следовательно, уравнения (8) позволяют получить оценки для порога локализации и для соотношений подобия между параметрами QB, оставляющими γ неизменным. Обратная длина локализации ξ^{-1} пропорциональна $q_0^{-1} \ln \gamma$. При условии $\gamma = 1$ QB делокализуется, и следовательно $\tau_2 \approx \tau_1$. На рис. 3 показаны численные результаты.

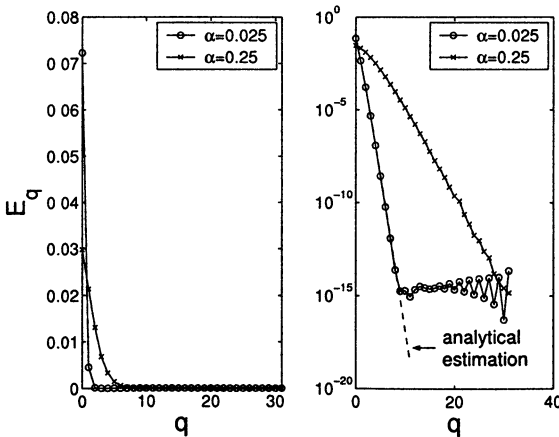


Рис. 3. Устойчивые QB для $\alpha = 0,025$ и $\alpha = 0,25$, $E = 0,077$, $N = 32$, $q_0 = 1$. Прерывистая линия — результат (8) [6].

Полученные QВ-решения линейно устойчивы [3, 6]. Отметим, что вопрос динамической устойчивости ПО, по всей видимости, не имеет прямого отношения к проблеме ФПУ. Действительно, если QВ локализован, то ощутимая неустойчивость на краях распределения энергии исключена — в противном случае соответствующие резонансы привели бы к делокализации самой ПО. Неустойчивость в центре распределения приведет к перестройке динамики в центре, но в силу отсутствия резонансов на краях не приведет к делокализации.

В прямом пространстве распределения энергии делокализованы как для траектории ФПУ, так и для QВ, и не отличаются особенностями [3, 6].

3.2. Результаты для β -ФПУ, низкие частоты

Применение теории возмущения Пуанкаре — Линдштедта для низкочастотных мод приводит к оценке распределения энергий нормальных мод на QВ [6]:

$$E_{(2n+1)q_0} = \lambda^{2n} E_{q_0}, \quad \lambda = \frac{3\beta E_{q_0}(N+1)}{8\pi^2 q_0^2}. \quad (9)$$

Необходимое условие для локализации QВ — $\lambda < 1$. Обратная длина локализации ξ^{-1} пропорциональна $(2q_0)^{-1} \ln \lambda$. При условии $\lambda = 1$ QВ делокализуется, и следовательно $\tau_2 \approx \tau_1$. На рис. 4 показаны численные результаты.

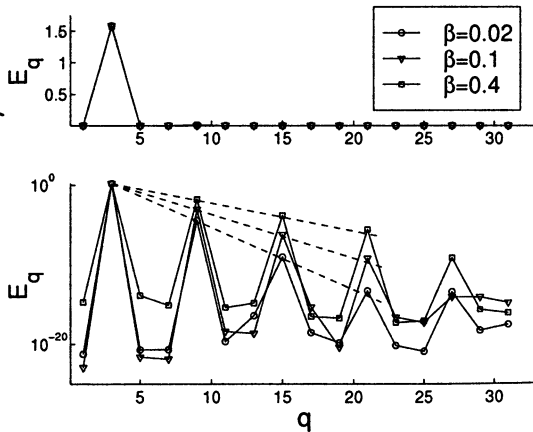


Рис. 4. QВ для разных β и $E = 1, 58$, $N = 32$, $q_0 = 3$. Прерывистые линии — результат (9) [6].

QB становятся неустойчивыми при условии $6\beta E(N+1)/\pi^2 > 1$ [6]. Поскольку сами решения при этом превосходно локализованы, то траектории ФПУ не приводят к делокализации [9] в согласии с вышеизложенным.

3.3. Результаты для высоких частот

Локализация наблюдается также и для случая затравочного волнового числа с большой частотой. На рис. 5 представлены распределения энергии для траектории ФПУ и для QB [4]. Единственное отличие от случая низких частот заключается в том, что для α -ФПУ модели происходит попарное возбуждение высокой и низкой частот, что следует из теории возмущения [4].

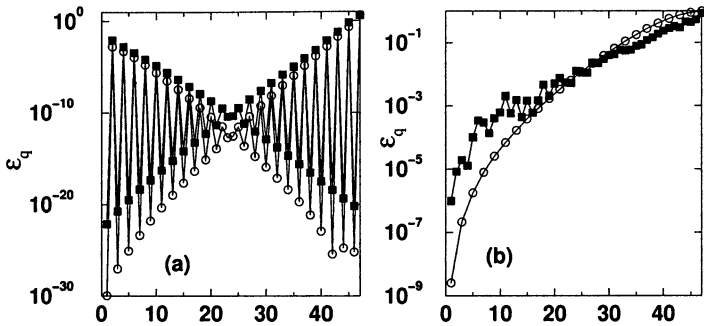


Рис. 5. Распределения плотностей энергии для траектории ФПУ (квадраты, $t = 10^5$) и QB (кружки) при $N = 47$, $E = 4, 7$, $q_0 = 47$: $\alpha = 0, 25$ (a); $\beta = 0, 25$ (b) [4].

4. Обобщения на решетки с размерностью $d = 2; 3$

Основное условие для продолжения QB заключалось в отсутствии резонансов (7). Поскольку бесконечная система имеет сплошной спектр частот нормальных мод, выполнение условия (7) не представляется возможным. Конечная система в прямом пространстве имеет дискретный спектр, и теорема Ляпунова применима. Естественным обобщением вышеизложенных результатов является рассмотрение конечных систем с размерностью $d = 2; 3$. В таком случае модовые числа становятся векторами с d компонентами. В работе [10] были получены QB решения для $d = 2; 3$. Основные свойства решений цепочки оказываются применимыми и здесь. На рис. 6 показан пример QB-решения для двумерной решетки.

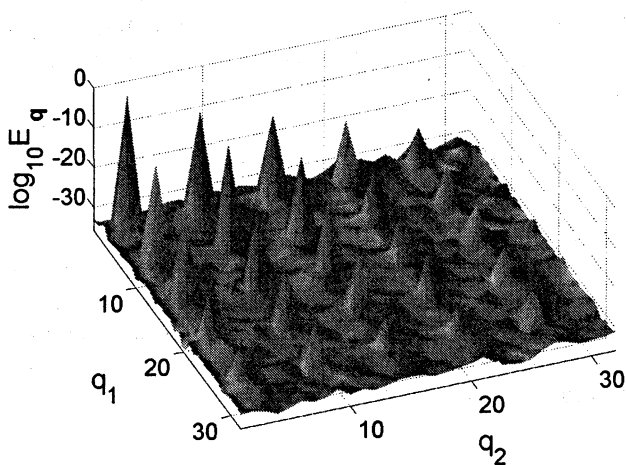


Рис. 6. Распределения энергии для QB в двумерной ФПУ-решетке при $N = 32 \times 32$, $E = 1, 5$, $\mathbf{q}_0 = (3, 3)$, $\beta = 0, 5$ [10].

Для β -ФПУ модели получено обобщение теории возмущения, предсказывающее следующее распределение энергий мод:

$$E_{(2n+1)\mathbf{q}_0} = \lambda_d^{2n} E_{\mathbf{q}_0}, \quad \lambda_d = \frac{3\beta E_{\mathbf{q}_0} N^{2-d}}{2^{2+d} \pi^2 |\mathbf{q}_0|^2}. \quad (10)$$

Численный анализ QB показывает количественное согласие с этими оценками [10].

5. Преобразования подобия и переход к макроскопическим системам

Из результатов теории возмущения (8)–(10) следует, что при переходе от экстенсивных величин (энергии E и модового числа q) к интенсивным величинам (плотности энергии $\epsilon = E/N$ и волновому числу $k = \pi q/(N+1)$) длина локализации перестает зависеть от размера системы N . Следовательно можно ожидать, что QB будут существовать в пределе бесконечно больших, т. е. макроскопических, систем. Ниже мы обсудим метод получения точных ПО-решений путем r -кратного увеличения размера системы. Оказывается, что предельный переход от дискретного к сплошному спектру нормальных мод не влияет на результаты.

В работе [11] было установлено, что достаточно получить QB-решение $Q_q(t)$ для конечной цепочки с размером N . Рассмотрим новую цепочку с размером $\tilde{N} + 1 = r(N + 1)$, где $r = 2, 3, 4, \dots$. Нетрудно убедиться в том, что

$$\tilde{Q}_{\tilde{q}}(t) = \begin{cases} \sqrt{r} Q_q(t), & \tilde{q} = rq \\ 0, & \tilde{q} \neq rq \end{cases} \quad (11)$$

является решением новой цепочки.

Фиксируя r и повторяя эту схему бесконечно много раз, либо просто устремляя r в бесконечность, получим решения в бесконечно больших, а значит, макроскопических системах. Преобразования подобия легко обобщаются на случай $d = 2; 3$, а также на свободные и периодические граничные условия [11].

5.1. Зависимость длины локализации QB от волнового числа в макроскопической системе

Предположим, что QB существуют в бесконечно большой цепочке. Тогда результат (9) принимает следующий вид:

$$\ln \epsilon_k = \left(\frac{k}{k_0} - 1 \right) \ln \sqrt{\lambda} + \ln \epsilon_{k_0}, \quad \sqrt{\lambda} = \frac{3\beta \epsilon_{k_0}}{8 k_0^2}. \quad (12)$$

Для проверки рассмотрим исходную цепочку с $N_0 = 15$ и положим $\bar{k}_0 = \pi/15$. Плавно увеличивая размер системы, будем получать QB-решения численно. В качестве затравочного модового числа q_0 берется целое число, ближайшее к значению $\bar{k}_0(N+1)/\pi$. Из полученного решения оценим коэффициент λ , используя соотношение $\epsilon_{5q_0}/\epsilon_{3q_0}$ (рис. 7). Результаты подтверждают независимость полученного значения λ от размера цепочки для $N + 1 = r(N_0 + 1)$. Каждому модовому числу при данном размере N соответствует волновое число k_0 в некой окрестности \bar{k}_0 . Плавная зависимость $\lambda(k_0)$ подтверждает вышеизложенные предположения для макроскопических систем.

5.2. Закон подобия для длины локализации QB и особенности на краях частотного спектра

При фиксированной плотности энергии ϵ и выполнении закона (9) следует $\epsilon_{k_0} = (1 - \lambda)\epsilon$. Используя (12), получим, что обратная длина локализации $\xi^{-1} = |S|$ задается абсолютным значением величины S ,

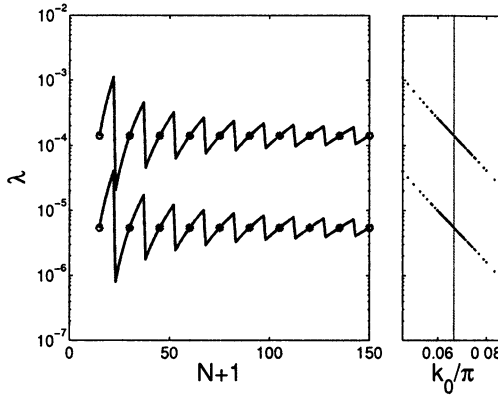


Рис. 7. Зависимость параметра λ от размера цепочки (левая диаграмма) для QB при плотности энергии $\epsilon = 4 \cdot 10^{-4}$ (нижняя кривая) и $2 \cdot 10^{-3}$ (верхняя кривая) и $\beta = 1$. Символы соответствуют $N + 1 = r(N_0 + 1)$, $N_0 = 15$, где r — целое число. На правой диаграмме показана зависимость λ от соответствующего волнового числа k_0 для данных из левой панели [11].

описывающей наклон распределения плотностей энергий в логарифмическом масштабе:

$$S = \frac{1}{k_0} \ln \sqrt{\lambda}, \quad \sqrt{\lambda} = \frac{\sqrt{1 + 4\nu^4/k_0^4} - 1}{2\nu^2/k_0^2}, \quad \nu^2 = \frac{3\beta}{8} \epsilon. \quad (13)$$

S зависит от затравочного волнового числа k_0 и от одного эффективного параметра нелинейности ν , содержащего произведение плотности энергии ϵ и коэффициента β . Длина локализации расходится при $k_0 \rightarrow 0$ и принимает наименьшее значение $\xi_{\min} \approx \nu/0,7432$ при $k_0 \approx 2,577\nu$. С увеличением ν , длина локализации при $k_0 = k_{\min}$ растет. В пределе $k_0 \gg \nu$ имеет место оценка $S \approx 2/k_0 \ln(\nu/k_0)$, а при $k_0 \ll \nu$ получаем $S \approx -k_0/(2\nu^2)$. На рис. 8 проведено сравнение аналитических и численных результатов. Прекрасно подтверждается вывод о независимости длины локализации от размера системы, а также наличие минимума $S(k_0)$, зависимость его глубины и положения от плотности энергии. Систематические различия при малых k_0 связаны с поправками к теории возмущений.

Из результата (13) следует, что величина $S_m(z) = \nu S$ зависит лишь от одной переменной — безразмерного затравочного волнового числа $z = k_0/\nu$. На рис. 9 показана зависимость $S_m(z)$, а также обезразмеренные численные данные из рис. 8.

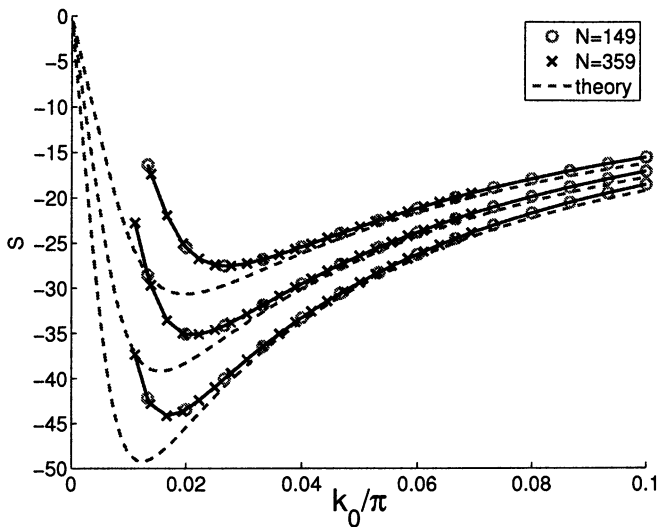


Рис. 8. $S(k_0)$ при $\beta = 1$ и $\epsilon = 6,08 \cdot 10^{-4}$, $9,6 \cdot 10^{-4}$, $1,57 \cdot 10^{-3}$ (штрихованные линии, снизу вверх). Символы и связывающие линии: результаты оценки численных решений QB для $N = 149$ и $N = 359$ [11].

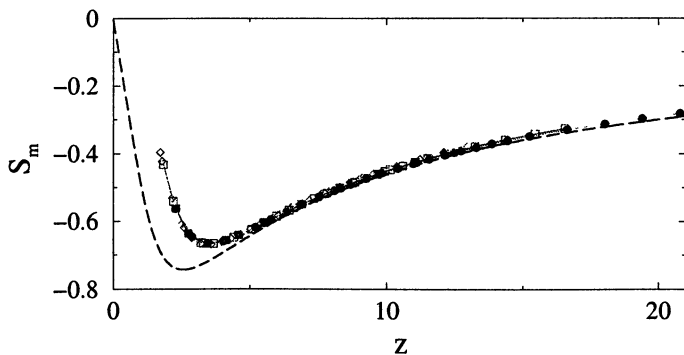


Рис. 9. $S_m(z)$ (штрихованная линия). Символы и связывающая линия: безразмерные результаты оценки численных решений QB из рис. 8 [11].

Несмотря на систематические отклонения на малых z , обнаружено, что все численные данные ложатся на одну кривую. Таким образом, подтверждается наличие закона подобия. Функция $S_m(z)$ описывает степень локализации QB для выбранных плотностей энергии, параметра нелинейности, затравочного модового числа и размера системы. Аналогичные результаты получены для α -ФПУ модели, а также для затравочных волновых чисел с частотами вблизи верхнего края спектра ω_q .

6. Обсуждение результатов

Причиной полной делокализации QB на краях спектра являются частотные резонансы. Следовательно, ожидается, что для макроскопической системы возбужденная мода с волновым числом вблизи края спектра будет быстро распадаться на другие моды, поскольку $\tau_2 \approx \tau_1$. Из полученных результатов следует, что существует целый интервал волновых чисел около краев спектра, где наступает практически полная делокализация. Действительно, неприводимая часть спектра ω_q ограничена по k (в случае цепочки — на величину π). О локализации в пространстве волновых чисел можно говорить, только когда $\xi < \pi$. Это условие определяет критическое расстояние Δk_0 затравочного волнового числа k_0 от границы спектра. При ненулевом значении $\nu \sim \beta\epsilon$ величина Δk_0 будет также ненулевой. Следовательно, имеется целый диапазон значений k_0 , когда возбуждение нормальной моды быстро распадется и нормальная мода практически не определена. При росте ν диапазон Δk_0 также растет, и при некотором критическом значении ν получим $\Delta k_0 \approx \pi$, т. е. нормальные моды более не характеризуют динамику системы независимо от значения волнового числа. Это соответствует режиму сильного взаимодействия между модами, который будем называть турбулентным режимом по аналогии с гидродинамикой. При малых значениях ν турбулентность наступает только для волновых чисел вблизи краев спектра.

Интересно отметить, что в квантовом режиме однофононных низкочастотных возбуждений ангармонизм приводит к распаду фонона на более низкие по частотам фононы. Легко убедиться в том, что классический предел, соответствующий возбуждению многофононного состояния с данной частотой, приведет к слиянию двух или трех фононов в один фонон с более высокой частотой в согласии с изложенными результатами. Также отметим, что в классической макроскопической системе при нулевой плотности энергии все нормальные моды (фононы) обретают бесконечное время жизни, ибо $\xi \rightarrow 0$. Следовательно,

известная теория квазичастиц Ландау дополняется новыми свойствами этих возмущений в классическом пределе.

Вернемся к обсуждению результатов для конечных систем. При данном ν мы имеем конечное значение Δk_0 . Конечная система предполагает дискретный набор волновых чисел для нормальных мод. Теплопроводность ангармонической цепочки обуславливается динамикой низкочастотных волн. Чтобы в численных расчетах обнаружить проводимость, отличающуюся от баллистической проводимости хорошо определенных фононов, необходимо разрешить турбулентный режим, а следовательно, выбрать достаточно большой размер системы, превышающий $N_c \sim 1/(\Delta k_0)$. Изложенная теория позволяет получить соответствующие количественные оценки.

В заключение отметим, что вышеупомянутая теория возмущения Пуанкаре — Линдштедта применима в диапазоне частот, где дисперсионное соотношение ω_q может быть описано разложением, включающим ведущий и следующий за ним члены. Соответственно результаты распространяемы, по крайней мере, на половину всей неприводимой зоны. Следовательно, распад нормальной моды в турбулентном режиме не ограничен диапазоном Δk_0 , а происходит во всем частотном спектре.

Наконец, обратим внимание на резонансные пики на краях распределений на рис. 2, как для траектории ФПУ, так и для QB. В работе [4] получено количественное объяснение возникновения и расположения этих пиков. Они вызваны тем, что для соответствующих модовых чисел почти выполняется равенство в (7). Легко убедиться в том, что пики возникают в верхней половине частотного спектра, а также только в случае низкочастотной затравочной моды. Для объяснения необходимо учесть дисперсионное соотношение более точно. Численные эксперименты показывают, что именно эти резонансные пики приводят к переходу в равновесие. Остается загадкой, каков путь к равновесию для высокочастотных затравочных мод, где аналогичные пики очевидно отсутствуют.

Представленные результаты были получены совместно с М. В. Иванченко, О. И. Канаковым, К. Г. Мишагиным и Т. Пенати. Я благодарен им, а также В. Флерову, Б. Альтшулеру, В. Д. Шалфееву за полезные обсуждения и А. Ярьско за техническую помощь.

Литература

1. *Campbell D. K., Flach S., Kivshar Yu. S. Localizing energy through nonlinearity and discreteness // Phys. Today. 2004. V. 57, № 1. P. 43—49.*

2. *Fermi E., Pasta J., Ulam S.* Studies of nonlinear problems. I // Los Alamos Report LA-1940, 1955.
3. *Flach S., Ivanchenko M. V., Kanakov O. I.* q -breathers in Fermi-Pasta-Ulam chains: existence, localization, and stability // Phys. Rev. E. 2006. V. 73. P. 036618-1-14.
4. *Penati T., Flach S.* Tail resonances of FPU q -breathers and their impact on the pathway to equipartition // nlin.PS/0610006, 2006.
5. *Berchialla L., Giorgilli A., Paleari S.* Exponentially long times to equipartition in the thermodynamic limit // Phys. Lett. A. 2004. V. 321. P. 167–172.
6. *Flach S., Ivanchenko M. V., Kanakov O. I.* q -breathers and the Fermi-Pasta-Ulam problem // Phys. Rev. Lett. 2005. V. 95. P. 064102-1-4.
7. *Conway J. H., Jones A. J.* Trigonometric diophantine equations (On vanishing sums of roots of unity) // Acta Arith. 1976. V. XXX. P. 229–240.
8. *Lyapunov M. A.* The general problem of stability of motion. London: Taylor & Francis, 1992.
9. *De Luca J., Lichtenberg A., Lieberman M. A.* Time scale to ergodicity in the Fermi-Pasta-Ulam system // Chaos. 1995. V. 5. P. 283–297.
10. *Ivanchenko M. V., Kanakov O. I., Mishagin K. G., Flach S.* q -breathers in finite two- and three-dimensional nonlinear acoustic lattices // Phys. Rev. Lett. 2006. V. 97. P. 025505-1-4.
11. *Kanakov O. I., Flach S., Ivanchenko M. V., Mishagin K.G.* Scaling properties of q -breathers in nonlinear acoustic lattices // nlin.PS/0607019, 2006.

СИММЕТРИИ И ИХ НАРУШЕНИЯ В СИСТЕМАХ ВО ВНЕШНИХ ПЕРЕМЕННЫХ ПОЛЯХ

С. Г. Флах

1. Введение

Одним из последствий второго закона термодинамики является невозможность извлечения полезной работы из единственного резервуара с заданной температурой. Разные примеры были предложены для того, чтобы пояснить суть этого высказывания. К ним, бесспорно, принадлежат демон Максвелла, а также система храповика и собачки, которой Фейнман посвятил целую главу в своих лекциях по физике [1].

В случае выведения рассматриваемой системы из равновесия законы равновесной термодинамики не применимы. Ниже под неравновесной системой будем понимать систему под воздействием внешних полей, зависящих от времени. Когда амплитуды полей достаточно малы, теория линейного отклика позволяет описать реакцию системы, основываясь на ее равновесных свойствах. При увеличении амплитуд полей приходится учитывать нелинейные члены в отклике, который в общем случае становится сложным нелинейным функционалом. Получение точного вида такого функционала может оказаться непосильной задачей.

Во многих случаях в первую очередь необходимо определить симметрии функционала отклика, в особенности соответствующие требования к симметриям переменных полей. Это позволяет определить условия нарушения симметрий, что может привести к появлению ненулевых средних от физических наблюдаемых (например, токов, намагниченностей и т. д.). Ниже будет изложен метод построения этих симметричных операций из уравнений движения без необходимости построения функционала отклика.

2. Частица в периодическом потенциале

Рассмотрим движение частицы с массой m и координатой x в периодическом потенциале $U(x) = U(x + \lambda)$:

$$m\ddot{x} + \gamma\dot{x} - f(x) - \chi(t) = 0. \quad (1)$$

Здесь $f(x) = -U'(x)$, $\int_0^\lambda f(x)dx = 0$, γ — коэффициент трения, $\chi(t)$ — случайная сила с нулевым средним. Ставится вопрос, возможен

ли режим с направленным движением, при котором средняя скорость частицы $\langle \dot{x} \rangle \neq 0$.

Очевидно, что если $\chi(t) \equiv \xi(t)$ является реализацией белого шума, то уравнение (1) описывает взаимодействие частицы с термостатом. Из второго закона термодинамики следует, что независимо от конкретного выбора потенциала $U(x)$ направленное движение не будет наблюдаться [2]. Отметим, что $\xi(t)$ не содержит временных корреляций.

Добавка временных корреляций может привести к другим результатам [2]. Такие корреляции проще всего получить с помощью аддитивного периодического поля $E(t)$:

$$\chi(t) = \xi(t) + E(t), \quad E(t) = E(t + T), \quad \int_0^T E(t) dt = 0. \quad (2)$$

Если $\xi(t)$ является реализацией белого шума, то функции $-\xi(t)$, $\xi(-t)$, $\xi(t + \tau)$ также являются реализациями белого шума с одинаковым статистическим весом. Для последующего симметричного анализа шумовой член $\xi(t)$ оказывается несущественным, и достаточно будет изучить симметричные свойства детерминистического уравнения

$$m\ddot{x} + \gamma\dot{x} - f(x) - E(t) = 0. \quad (3)$$

Уравнение (3) содержит две периодические функции $f(x)$ и $E(t)$ с нулевыми средними. Симметричные свойства этих функций решающим образом влияют на свойства решений уравнения (3).

2.1. Симметрии периодической функции с нулевым средним

Рассмотрим периодическую функцию $g(z + 2\pi) = g(z)$ с условием $\int_0^{2\pi} g(z) dz = 0$. Такая функция может иметь три симметрии. Во-первых, допустим что $g(z_0 + z) = g(z_0 - z)$ для некоего значения z_0 . Такие функции будем называть симметричными и обозначать индексом s : g_s . Симметричные функции могут быть разложены в ряд Фурье по косинусам. Во-вторых, допустим что $g(z_1 + z) = -g(z_1 - z)$ для некоего значения z_1 . Такие функции будем называть антисимметричными и обозначать индексом a : g_a . Антисимметричные функции могут быть разложены в ряд Фурье по синусам. Наконец, допустим, что $g(z) = -g(z + \pi)$. Такие функции будем называть сдвигово-симметричными и обозначать индексом sh : g_{sh} . Сдвигово-симметричные функции не содержат четных членов в разложении в ряд Фурье. Легко убедиться, что данная функция $g(z)$ либо обладает всеми тремя симметриями, либо только одной из них, либо не обладает ни одной из указанных симметрий.

Рассмотрим несколько примеров. Функция $\cos(z)$ обладает всеми тремя симметриями. Функция $\cos(z) + \cos(2z + \phi)$ не обладает ни одной из рассмотренных симметрий, если $\phi \neq 0, \pi/2, \pi, 3\pi/2$. Однако она симметрична для $\phi = 0, \pi$ и антисимметрична для $\phi = \pi/2, 3\pi/2$.

2.2. Симметрии уравнения (3)

Фазовое пространство уравнения (3) имеет размерность три. Уравнение можно переписать в виде трех автономных дифференциальных уравнений первого порядка:

$$\dot{x} = \frac{p}{m}, \quad \dot{p} = f(x) + E(\tau) - \frac{\gamma}{m}p, \quad \dot{\tau} = 1. \quad (4)$$

Будем искать симметричные преобразования \hat{S} , не изменяющие сами уравнения. Такие симметричные преобразования по сути есть отображения фазового пространства $\{x, p, \tau\}$ на себя. Допустим, что мы нашли такое преобразование, и рассмотрим данную траекторию в фазовом пространстве. Применив преобразование ко всем точкам этой траектории, получим новое многообразие, которое по определению снова соответствует траектории, т. е. решению уравнений. Исходная и новая траектории могут совпадать, но могут и быть разными. Из всех симметричных преобразований \hat{S} нас интересуют те, которые меняют знак у скорости \dot{x} .

Предположим, что мы нашли такое преобразование. Рассмотрим среднее значение скорости $\bar{v} = \lim_{s \rightarrow \infty} \frac{1}{s}(x(t_0 + s) - x(t_0))$ на исходной траектории. Тогда $\bar{v} = 0$, если исходная и новая траектории совпадают. Если же две траектории разные, то значения \bar{v} на них имеют одинаковые абсолютные значения, но разные знаки. Предполагая, что в присутствии белого шума такие траектории будут иметь одинаковый статистический вес, приходим к выводу, что их вклад в среднюю скорость равен нулю [3]. Следовательно, направленное движение отсутствует. Условие, что симметричное преобразование меняет знак у скорости, ограничивает поиск на случаи, когда меняется знак либо у координаты x , либо у времени t , но не у обоих одновременно.

Следующие симметрии удовлетворяют вышеуказанным условиям [3]:

$$\begin{aligned} \hat{S}_a: & \quad x \rightarrow -x, \quad t \rightarrow t + \frac{T}{2}, \quad \text{if } \{f_a, E_{sh}\}; \\ \hat{S}_b: & \quad x \rightarrow x, \quad t \rightarrow -t, \quad \text{if } \{E_s, \gamma = 0\}; \\ \hat{S}_c: & \quad x \rightarrow x + \frac{\lambda}{2}, \quad t \rightarrow -t, \quad \text{if } \{f_{sh}, E_a, m = 0\}. \end{aligned} \quad (5)$$

Необходимые условия для каждого из трех преобразований — это соответствующие требования к симметричным свойствам функций f и E , которые указаны в скобках в (5). Симметрия \hat{S}_b также предполагает предел нулевой диссипации $\gamma = 0$, а симметрия \hat{S}_c — безынерционный предел сильного трения $m = 0$. Заметим, что во всех трех случаях имеются симметричные требования к функции $E(t)$, в отличие от потенциальной силы $f(x) = -U'(x)$. Во многих экспериментальных реализациях для направленного транспорта конкретный вид временной зависимости поля $E(t)$ может быть подобран в желаемой форме существенно проще, чем вид координатной зависимости $U(x)$. В дальнейшем ограничимся случаем симметричного потенциала $U(x) = 1 - \cos x$ и будем нарушать симметрии (5), используя переменное поле $E(t) = E_1 \cos(t) + E_2 \cos(2t + \phi)$. Если $\phi \neq 0, \pi/2, \pi, 3\pi/2$, то все симметрии (5) нарушены, и мы можем ожидать, что средняя скорость $\bar{v} \neq 0$.

2.3. Результаты

Рассмотрим предельный случай $\gamma = 0$ [3]. Динамика системы изучается с помощью сечения Пуанкаре. Берется начальное условие $\{x, p\}$, уравнения движения интегрируются численно, и полученные через каждый период поля $E(t)$ новые значения $\{x, p\}$ отмечаются точкой в соответствующей фазовой плоскости (рис. 1). Для $E(t) = 0$ сепаратриса будет отделять бегущие решения от осциллирующих. Включение поля разрушает сепаратрису, на месте которой возникает стохастический слой. Динамика в слое эргодична, т. е. средние характеристики всех траекторий из этого слоя совпадают. Следовательно, симметричный анализ применим ко всем траекториям из стохастического слоя. Численный анализ подтвердил эти рассуждения [3, 4]. На рис. 1 показана зависимость $x(t)$, из которой явно следует, что при нарушенных симметриях возникает направленное движение частицы. Динамику траектории в стохастическом слое можно грубо разделить на два режима. Первый режим — это баллистические полеты за счет подлипания траектории к границам слоя. Второй режим — это блуждание внутри слоя. Выпрямление движения частицы происходит за счет нарушения симметричного баланса между вероятностями подлипания к баллистическим полетам в разных направлениях [4]. Численная оценка величины средней скорости может быть получена с помощью обобщенных правил сумм [5]. В частности, следует, что, помимо нарушения соответствующих симметрий, необходимым условием для появления тока является смешанное фазовое пространство, т. е. наличие стохастического слоя и его границ с регулярными областями [5].

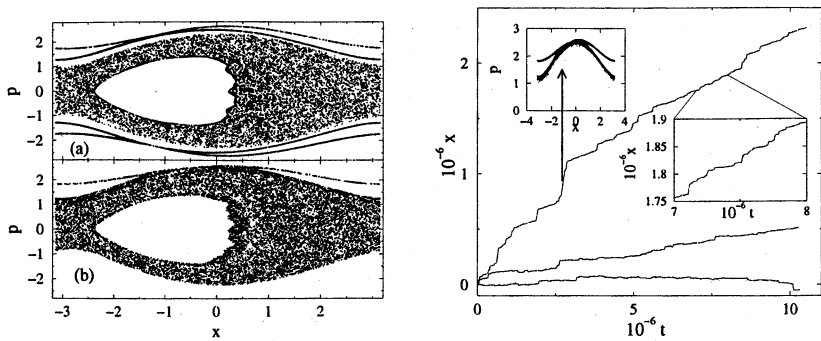


Рис. 1. Левая диаграмма: сечение Пуанкаре для $\gamma = 0$, $E_1 = 0,252$, $E_2 = 0,052$: $a - \phi = 0$, $b - \phi = \pi/2$. Правая диаграмма: $x(t)$ для $\phi = 0, \pi/5, \pi/2$ (снизу вверх). Вставка в верхнем левом углу: отображение Пуанкаре для баллистического полета при $\phi = \pi/2$. Вставка в нижнем правом углу: увеличение для $\phi = \pi/2$ [4].

Трение $\gamma \neq 0$ не влияет на полученные результаты существенным образом. Симметричные рассуждения остаются в силе. Фазовое пространство теперь содержит аттракторы с соответствующими бассейнами притяжения, которые сложным образом переплетаются в районе стохастического слоя и десимметризируются. Добавка малого шума приводит к блужданию траектории по всем бассейнам и, следовательно, к выпрямлению [3, 6]. Сама траектория по-прежнему содержит элементы прилипания, а времена полетов ограничены характерным временем, обратно пропорциональным константе трения γ [6, 7].

Систематический анализ с учетом шума показывает, что при условии полного нарушения всех симметрий (5) максимальное значение \bar{v} достигается в пределе слабой диссипации и задается характеристиками стохастического слоя [6, 7]. Было показано, что в соответствии с симметричным анализом направленный ток исчезает при $\phi = 0, \pi$ в случае слабой диссипации и при $\phi = \pm\pi/2$ для сильной диссипации [6]. Значение фазы ϕ , при которой ток исчезает, является монотонной функцией константы диссипации [6].

Обобщение этих результатов на квантовый режим приводит к возможности резонансного усиления тока [8]. Однако симметричные свойства уравнения Шредингера практически совпадают с классическим пределом [5, 8].

Наконец отметим, что не зависящая от времени компонента поля $\vec{E}(t) = E_{dc} + E(t)$ позволяет получить направленный транспорт против приложенной постоянной силы со средней скоростью, не равной нулю [9].

Эти результаты получили подтверждение в экспериментах группы Ренцони [10, 11]. В этих работах атомы Cs и Rb охлаждаются до температур порядка милликельвина. Последующее подключение лазерных пучков создает периодический потенциал для атомов в виде стоячей волны. Наконец, переменная фаза в одном из пучков моделирует переменное поле $E(t)$. Исследованы и подтверждены результаты симметричного анализа при помощи изменения относительной фазы ϕ и эффективной константы диссипации.

2.4. Квазипериодические переменные поля

Ниже обсудим обобщение симметричного анализа на случай квазипериодических полей. Квазипериодическая функция может быть определена следующим образом:

$$g(z) \equiv \tilde{g}(z_1, z_2, \dots, z_N), \quad \frac{dz_i}{dz} = \Omega_i, \quad (6)$$

где все соотношения Ω_i/Ω_j иррациональны для $i \neq j$ и $\tilde{g}(z_1, z_2, \dots, z_i + 2\pi, \dots, z_N) = \tilde{g}(z_1, z_2, \dots, z_i, \dots, z_N)$ для всех i . Перечислим симметрии квазипериодической функции, которые оказываются существенными для изучения направленного транспорта. Функция \tilde{g} может быть симметричной: $\tilde{g}_s(z_1, z_2, \dots, z_N) = \tilde{g}_s(-z_1, -z_2, \dots, -z_N)$; а также антисимметричной: $\tilde{g}_a(z_1, z_2, \dots, z_N) = -\tilde{g}_a(-z_1, -z_2, \dots, -z_N)$. Для данного подмножества индексов $\tilde{g}_{sh, \{i, j, \dots, m\}}$ функция может обладать сдвиговой симметрией, т. е. \tilde{g} будет менять знак при сдвиге аргументов z_i, z_j, \dots, z_m на π . При этом остальные аргументы не сдвигаются.

Итак, симметричные свойства функции \tilde{g} определены на компактном пространстве переменных $\{z_1, z_2, \dots, z_N\}$. Из несоизмеримости частот следует, что равномерное увеличение переменной z приводит к плотному заполнению компактного пространства с однородной плотностью [12].

Уравнения движения для частицы в периодическом потенциале под воздействием квазипериодического переменного поля $E(t)$ принимают вид [13]

$$\begin{aligned} m\ddot{x} + \gamma\dot{x} - f(x) - E(\phi_1, \phi_2, \dots, \phi_N) &= 0, \\ \dot{\phi}_1 &= \omega_1, \quad \dot{\phi}_2 = \omega_2, \quad \dots, \quad \dot{\phi}_N = \omega_N. \end{aligned} \quad (7)$$

Наконец получим симметричные преобразования, оставляющие уравнения движения (7) инвариантными, но меняющие знак у скорости:

$$\tilde{S}_a : x \rightarrow -x, \quad \phi_{i, j, \dots, m} \rightarrow \phi_{i, j, \dots, m} + \pi, \quad \text{if } \{f_a, E_{sh, \{i, j, \dots, m\}}\}, \quad (8)$$

$$\tilde{S}_b : x \rightarrow x, \quad t \rightarrow -t, \quad \text{if } \{E_s, \gamma = 0\}. \quad (9)$$

Преобразование \tilde{S}_a в действительности представляет собой множество преобразований, каждое из которых задается набором индексов $\{i, j, \dots, m\}$.

Дальнейший анализ показывает, что симметрии (8), (9) могут быть нарушены соответствующим выбором $E(t)$, что приводит к направленному движению [13]. Экспериментальные исследования постепенного перехода от периодических функций к квазипериодическим функциям с использованием холодных атомов в оптических решетках подтвердили результаты симметричного анализа [14].

3. Направленный поток энергии и солитонов

В пределе среднего поля динамика многочастичной системы со взаимодействием описывается нелинейными дифференциальными уравнениями в частных производных. Рассмотрим в качестве примера поле $\varphi(x, t)$, удовлетворяющее уравнению синус-Гордона в присутствии диссипации, шума и внешнего поля:

$$\begin{aligned} \varphi_{,tt} - \varphi_{,xx} + \alpha\varphi_{,t} + \sin \varphi &= f(t, x), \\ f(t, x) &= E(t) + \xi(t, x). \end{aligned} \quad (10)$$

Здесь константа α — коэффициент диссипации. Гауссов белый шум ξ характеризуется корреляционной функцией $\langle \xi(t, x)\xi(t', x') \rangle = 2\frac{\alpha}{\beta}\delta(x - x')\delta(t - t')$, а β соответствует обратной температуре. Зададим граничные условия

$$\varphi(x + L, t) = \varphi(x, t) + Q, \quad \varphi_t(x + L, t) = \varphi_t(x, t). \quad (11)$$

Величина $Q = 2\pi m$ задает целочисленный топологический заряд m , а L — размер системы. Заметим, что поле φ описывает многочастичные возбуждения разных систем, в том числе макроскопические сверхпроводящие зарядовые флуктуации в контактах Джозефсона.

3.1. Симметричный анализ

Поскольку поле является носителем энергии, попытаемся провести симметричный анализ, позволяющий найти условия для направленного переноса энергии в системе (10). Для этого необходимо определить поток энергии и его плотность. Предположим, что роль шума и диссипации сводится к усреднению и не влияет на перенос энергии

вдоль системы. Тогда определим *внутренний поток энергии* [15]

$$J^I(t) = \int_0^L j^I dx, \quad j^I(x, t) = -\varphi_x \varphi_t, \quad (12)$$

являющийся единственной возможностью передачи энергии вдоль системы в отсутствие термостата и переменного поля $E(t)$. Определим симметричные преобразования, оставляющие инвариантными уравнение (10) с учетом граничного условия (11) и меняющие знак у потока энергии (12) (см. [15]):

$$\begin{aligned} \bar{S}_a: & \quad x \rightarrow -x, \quad \text{if } \{Q = 0\}; \\ \bar{S}_b: & \quad x \rightarrow -x, \quad t \rightarrow t + \frac{T}{2}, \quad \varphi \rightarrow -\varphi + Q, \quad \text{if } \{E_{sh}\}; \\ \bar{S}_c: & \quad t \rightarrow -t, \quad \text{if } \{E_s, \alpha = 0, \frac{1}{\beta} = 0\}. \end{aligned} \quad (13)$$

Следовательно при нулевом топологическом заряде $Q = 0$ направленный поток энергии отсутствует независимо от конкретного вида переменного поля $E(t)$ (см. преобразование \bar{S}_a). При ненулевом заряде $Q \neq 0$ поток энергии пропадает, если $E(t)$ имеет сдвиговую симметрию (см. преобразование \bar{S}_b). В пределе слабой диссипации ток будет также пропадать, если $E(t)$ симметрично во времени (см. преобразование \bar{S}_c). Итак, предположим, что $Q \neq 0$, и нарушим все симметрии поля $E(t)$. Тогда можем ожидать, что средний поток энергии в системе будет отличен от нуля.

3.2. Результаты, парадокс и решение

При ненулевом значении топологического заряда в системе существует избыток топологических солитонов (частиц) по сравнению с антисолитонами (античастицами), либо наоборот. В простейшем случае $Q = 2\pi$, $L \rightarrow \infty$, $E(t) = \alpha = 1/\beta = 0$, известно точное солитонное решение

$$\varphi^{kink}(x, t) = 4 \arctan \left\{ \exp \left[\frac{x - Vt}{\sqrt{1 - V^2}} \right] \right\}, \quad (14)$$

двигающееся со скоростью V . Используя определение плотности энергии

$$\rho[\varphi(x, t)] \equiv \rho(x, t) = \frac{1}{2}(\varphi_t^2 + \varphi_x^2) + 1 - \cos \varphi, \quad (15)$$

получим полную энергию солитона:

$$W^{kink} = \int_0^L \rho[\varphi^{kink}] dx, \quad W_{L \rightarrow \infty}^{kink} = \frac{8}{\sqrt{1-V^2}}. \quad (16)$$

Следовательно, можно ожидать, что направленный поток энергии будет осуществляться с помощью направленного движения солитонов. Численный анализ [15, 16] подтвердил эти предсказания. Более того, на кольцевых контактах Джозефсона было получено экспериментальное подтверждение [17].

Однако легко показать [18], что внутренний поток энергии удовлетворяет следующему линейному уравнению:

$$J_t^I(t) = -\alpha J^I(t) - QE(t). \quad (17)$$

Следовательно, для всех периодических полей $E(t)$ с нулевым средним

$$J^I = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t J^I(\tau) d\tau = 0. \quad (18)$$

Внутренний поток энергии строго равен нулю, ибо уравнение (17) — линейное. Это было трудно предположить, глядя на исходные уравнения (10), (11) и (12). Поучительный урок здесь заключается прежде всего в том, что из наличия симметрий следует отсутствие транспорта. Однако отсутствие симметрий есть необходимое, но не достаточное условие для возникновения транспорта. Второй вывод — внутренний поток энергии равен нулю, ибо отклик этой величины на внешнее поле — чисто линейный (17). Для выпрямления транспорта с помощью переменного поля необходим нелинейный отклик.

Внимательный читатель уже заметил, что имеет место парадокс. Мы доказали, что внутренний поток энергии в среднем строго равен нулю. Однако численные расчеты показали, что обнаруживается направленное движение солитонов тогда и только тогда, когда нарушены вышеуказанные симметрии. Поскольку очевидно, что вместе с солитоном переносится энергия (16), то направленный поток энергии все-таки имеет место, в согласии с симметричным анализом! Решение парадокса скрывается в уравнении непрерывности:

$$\rho_t = -j_x^I - \phi(x, t), \quad \phi(x, t) = \alpha \varphi_t^2 - E(t) \varphi_t. \quad (19)$$

Поскольку внутренний поток энергии не дает вклада в полный поток энергии, то величина $\phi(x, t)$ обязана содержать в себе градиент второго — обменного — потока энергии. Подробный анализ подтвердил это предположение [19, 20]. На рис. 2 показана эволюция градиента плотности обменного потока $\phi(x, t)$ и плотности энергии $\rho(x, t)$ для поля $E(t) = 0, 2 \cos(0, 2\pi t) + 0, 2 \cos(0, 4\pi t)$.

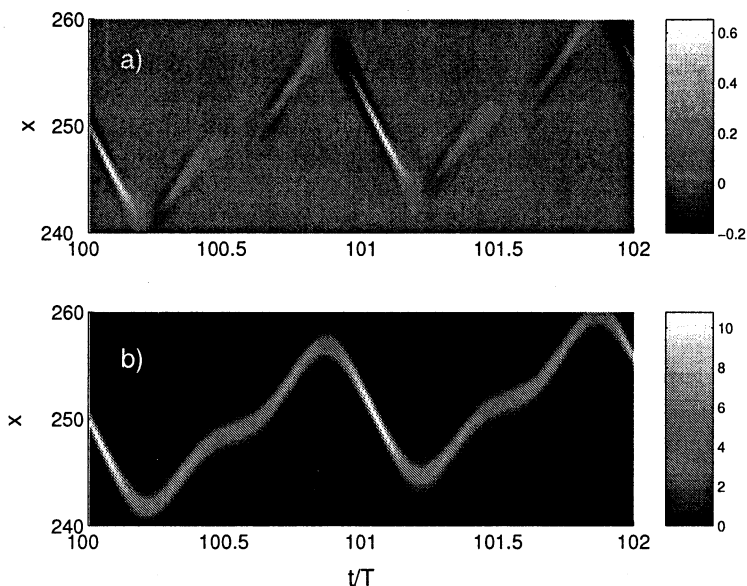


Рис. 2. Пространственно-временная эволюция направленного солитонного движения для $\alpha = 0, 2$, $L = 500$: *a* — $\phi(x, t)$; *b* — $\rho(x, t)$ [19].

Новый путь передачи энергии связан с присутствием термостата. Диссипация приводит к пространственно неоднородному оттоку энергии из системы, а переменное поле обеспечивает как отток, так и приток, который неоднороден в пространстве и времени. Именно эти механизмы и обеспечивают направленное движение солитона [19, 20]. Попытки замены солитона точечной частицей пренебрегают эффектами пространственной неоднородности плотностей. В итоге такие подходы обречены на неудачу, не позволяя получить правильный результат направленного движения.

4. Спин в переменном магнитном поле

Помимо направленного транспорта, переменные поля могут вызывать и другие физические эффекты, например они могут индуцировать поляризацию и намагниченность. Рассмотрим случай квантового спина в переменном магнитном поле в присутствии термостата.

4.1. Полезные определения

Рассмотрим спиновую систему, которая характеризуется гамильтонианом $H(t) = H(t + T)$ и взаимодействует с термостатом при некоторой обратной температуре β . Запишем уравнение Лиувилля для матрицы плотности ρ с учетом термостата:

$$\frac{\partial \rho}{\partial t} = i[H, \rho] - \nu(\rho - \rho_\beta). \quad (20)$$

Здесь коммутатор $[A, B] = AB - BA$, а ρ_β — равновесная матрица плотности, зависящая от обратной температуры β . Феноменологический параметр ν определяет силу взаимодействия между спиновой системой и термостатом и соответствует обратному времени релаксации.

Гамильтониан можно представить в виде суммы усредненной по времени равновесной части $H_0 = 1/T \int_0^T H(t) dt$ и флуктуационного члена $H_1(t) \equiv H(t) - H_0$ с нулевым средним: $\int_0^T H_1(t) dt = 0$. Определим $\rho_\beta = \frac{1}{Z} e^{-\beta H_0}$, где $Z = \text{Tr}(e^{-\beta H_0})$. Неоднозначность определения равновесной матрицы плотности не влияет на последующее рассмотрение симметрий и их нарушений.

Пусть $\bar{A}(t)$ есть ожидаемая оператора A . Тогда $\bar{A}(t) = \text{Tr}(A\rho(t))$. Обозначим среднее по времени от $\bar{A}(t)$ через $\bar{A} = \lim_{t' \rightarrow \infty} \frac{1}{t'} \int_0^{t'} \bar{A}(t) dt$.

Линейное уравнение (20) описывает эволюцию коэффициентов матрицы плотности, сохраняющую след $\text{Tr}(\rho)$. Для $\nu > 0$ и фиксированного следа $\text{Tr}(\rho)$ соответствующее фазовое пространство содержит единственный аттрактор, к которому будет притягиваться любая исходная траектория.

4.2. Уравнения для спина $s = 1/2$

Рассмотрим случай спина $s = 1/2$. Приложим постоянное магнитное поле в направлении z , что приведет к расщеплению двух уровней на величину, равную энергии Зеемана. В присутствии второго, переменного, поля в $x - z$ -плоскости имеем

$$H = h_0 S_z + h(t)(\alpha S_x + \gamma S_z), \quad (21)$$

где $\alpha = \sin \phi$, $\gamma = \cos \phi$ и $\int_0^T h(t) dt = 0$. Для спина $s = \frac{1}{2}$ операторы спиновых компонент определены через матрицы Паули $S_{x,y,z} = \frac{1}{2} \sigma_{x,y,z}$:

$$S_x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{i}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad S_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (22)$$

Матрица плотности является эрмитовой и определена четырьмя независимыми вещественными переменными. Одна из них исключается требованием нормировки $\text{Tr}(\rho) = 1$. Оставшиеся три переменные путем линейного преобразования могут быть выражены через математическое ожидание операторов спиновых компонент. В результате получим [21]

$$\dot{\bar{S}}_x = (h_0 + \gamma h(t))\bar{S}_y - \nu \bar{S}_x, \quad (23)$$

$$\dot{\bar{S}}_y = \alpha h(t)\bar{S}_z - (h_0 + \gamma h(t))\bar{S}_x - \nu \bar{S}_y, \quad (24)$$

$$\dot{\bar{S}}_z = -\alpha h(t)\bar{S}_y - \nu(\bar{S}_z - C), \quad (25)$$

где $C = 0,5 \tanh(h_0\beta/2)$. Эти уравнения являются частным случаем уравнений Блоха.

4.3. Симметрии

Обсудим симметричные операции Q , оставляющие уравнения (23)–(25) неизменными и меняющие знаки у величин $\bar{S}_{x,y}$ [21]. Первая операция

$$Q_1: \bar{S}_{x,y} \rightarrow -\bar{S}_{x,y}, \quad \bar{S}_z \rightarrow \bar{S}_z, \quad t \rightarrow t + T/2, \quad \text{if } \{\gamma = 0, h_{sh}\} \quad (26)$$

позволяет заключить, что при выполнении соответствующих условий $\tilde{S}_x = \tilde{S}_y = 0$, тогда как $\tilde{S}_z \neq 0$. Из применимости второй операции

$$Q_2: \bar{S}_x \rightarrow -\bar{S}_x, \quad \bar{S}_{y,z} \rightarrow \bar{S}_{y,z}, \quad t \rightarrow -t, \quad \text{if } \{\gamma = \nu = 0, h_a\} \quad (27)$$

следует $\tilde{S}_x = 0$. При наличии третьей симметрии

$$Q_3: \tilde{S}_y \rightarrow -\tilde{S}_y, \quad \tilde{S}_{x,z} \rightarrow \tilde{S}_{x,z}, \quad t \rightarrow -t, \quad \text{if } \{\nu = 0, h_s\} \quad (28)$$

закключаем, что $\tilde{S}_y = 0$. Очевидно, что при условии $\nu \neq 0$ и $\gamma \neq 0$ простейший выбор $h(t) = h_1 \cos \omega t$ означает, что $\tilde{S}_y \neq 0$. Следовательно, сумма постоянного и переменного магнитных полей возбуждает компоненту намагниченности, перпендикулярную к приложенным полям, при условии, что угол между двумя полями отличен от прямого. Этот эффект нарушенной симметрии является неадиабатическим, поскольку в пределе бесконечно малой частоты переменного поля намагниченность будет всегда направлена вдоль поля, т. е. не выходит из плоскости (x, z) . Наконец, поперечная намагниченность пропадает в пределе $\nu \rightarrow 0$, следовательно величина этой компоненты служит количественной оценкой силы взаимодействия спина с термостатом [21].

4.4. Результаты

В пределе сильной диссипации $1/\nu \rightarrow 0$ были получены аналитические решения с помощью разложения по малому параметру $1/\nu$ [21]:

$$\begin{aligned} \tilde{S}_x = C\alpha\gamma\langle h^2 \rangle \frac{1}{\nu^2} - C\alpha(-\gamma\langle h\ddot{h} \rangle + 3\gamma h_0^2\langle h^2 \rangle + \\ + (\alpha^2 + 3\gamma^2)h_0\langle h^3 \rangle + \gamma(\gamma^2 + \alpha^2)\langle h^4 \rangle) \frac{1}{\nu^4} + O\left(\frac{1}{\nu^5}\right), \end{aligned} \quad (29)$$

$$\tilde{S}_y = -C\alpha [2\gamma h_0\langle h^2 \rangle + (\gamma^2 + \alpha^2)\langle h^3 \rangle] \frac{1}{\nu^3} + O\left(\frac{1}{\nu^5}\right), \quad (30)$$

где $\langle f(t) \rangle = \frac{1}{T} \int_0^T f(t) dt$. Очевидно, что результаты совпадают с симметричным анализом. Численный расчет (рис. 3) позволяет получить количественные ответы для любых значений ν .

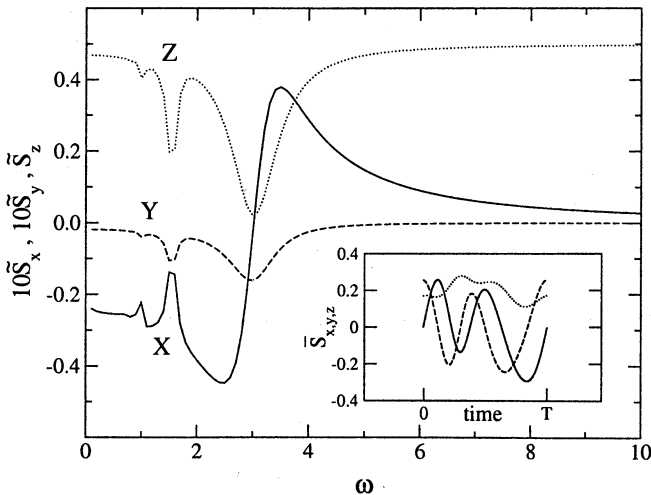


Рис. 3. Зависимость $10\tilde{S}_x$, $10\tilde{S}_y$ и \tilde{S}_z от частоты ω для $\phi = \pi/4$, $h_0 = 3$, $\nu = 0, 1$ и $\beta = 10$. На вставке: зависимость $\tilde{S}_{x,y,z}(t)$ при $\omega = 1, 5$ [21].

В работах [21, 22] симметричный анализ был успешно применен к системе со спином $s = 1$, а также к системе двух взаимодействующих спинов. В 1968 году Аримондо провел ряд экспериментов по нелинейному отклику при парамагнитном резонансе в соединении Diphenylpicrylhydrazyl (DPPH) [23]. Наряду с другими результатами

Аримондо обнаружил поперечную намагниченность при условиях, которые в точности соответствуют нарушениям рассмотренных выше симметрий.

Автор благодарен С. Денисову, О. Евтушенко и Я. Золотарюк за полезные обсуждения, а также А. Ярьско за техническую помощь.

Литература

1. *Фейнман Р., Лейтон Р., Сэндс М.* Фейнмановские лекции по физике. Т. 4. М.: Мир, 1976.
2. *Reimann P.* Brownian motors: noisy transport far from equilibrium // *Phys. Rep.* 2002. V. 361. P. 57–265.
3. *Flach S., Yevtushenko O., Zolotaryuk Y.* Directed current due to broken time-space symmetry // *Phys. Rev. Lett.* 2000. V. 84. P. 2358–2361.
4. *Denisov S., Flach S.* Dynamical mechanisms of dc current generation in driven Hamiltonian systems // *Phys. Rev. E.* 2001. V. 64. P. 056236-1-4.
5. *Schanz H., Otto M. F., Ketzmerick R., Dittrich T.* Classical and quantum Hamiltonian ratchets // *Phys. Rev. Lett.* 2001. V. 87. P. 070601-1-4.
6. *Yevtushenko O., Flach S., Zolotaryuk Y., Ovchinnikov A. A.* Rectification of current in ac-driven nonlinear systems and symmetry properties of the Boltzmann equation // *Europhys. Lett.* 2001. V. 54. P. 141–147.
7. *Denisov S., Flach S., Ovchinnikov A. A., Yevtushenko O., Zolotaryuk Y.* Broken space-time symmetries and mechanisms of rectification of ac fields by nonlinear (non)adiabatic response // *Phys. Rev. E.* 2002. V. 66. P. 041104-1-10.
8. *Denisov S., Morales-Molina L., Flach S.* Quantum resonances and rectification of driven cold atoms in optical lattices // *cond-mat/0607558* 2006.
9. *Denisov S., Flach S., Hänggi P.* Stationary Hamiltonian transport with dc bias // *Europhys. Lett.* 2006. V. 74. P. 588–594.
10. *Schiavoni M., Sanchez-Palencia L., Renzoni F., Grynberg G.* Phase control of directed diffusion in a symmetric optical lattice // *Phys. Rev. Lett.* 2003. V. 90. P. 094101-1-4.
11. *Gommers R., Bergamini S., Renzoni F.* Dissipation-induced symmetry breaking in a driven optical lattice // *Phys. Rev. Lett.* 2005. V. 95. P. 073003-1-4.
12. *Neumann E., Pikovsky A.* Quasiperiodically driven Josephson junctions: strange nonchaotic attractors, symmetries and transport // *Eur. Phys. J. B.* 2002. V. 26. P. 219–228.
13. *Flach S., Denisov S.* Symmetries and transport with quasiperiodic driving // *Acta Phys. Pol. B.* 2004. V. 35. P. 1437–1445.

14. *Gommers R., Denisov S., Renzoni F.* Quasiperiodically driven ratchets for cold atoms // *Phys. Rev. Lett.* 2006. V. 96. P. 240604-1-4.
15. *Flach S., Zolotaryuk Y., Miroshnichenko A. E., Fistul M. V.* Broken symmetries and directed collective energy transport in spatially extended systems // *Phys. Rev. Lett.* 2002. V. 88. P. 184101-1-4.
16. *Salerno M., Zolotaryuk Y.* Soliton ratchetlike dynamics by ac forces with harmonic mixing // *Phys. Rev. E.* 2002. V. 65. P. 056603.
17. *Ustinov A. V., Coqui C., Kemp A., Zolotaryuk Y., Salerno M.* Ratchetlike dynamics of fluxons in annular Josephson junctions driven by biharmonic microwave fields // *Phys. Rev. Lett.* 2004. V. 93. P. 087001-1-4.
18. *Joergensen E., Koshelets V.P., Monaco R., Mygind J., Samuelsen M. R., Salerno M.* Thermal fluctuations in resonant motion of fluxons on a Josephson transmission line — theory and experiment // *Phys. Rev. Lett.* 1982. V. 45. P. 1093.
19. *Denisov S., Flach S., Gorbach A. V.* Energy flow for soliton ratchets // *Europhys. Lett.* 2005. V. 72. P. 183–189.
20. *Gorbach A. V., Denisov S., Flach S.* Energy flow for moving dissipative topological solitons // *Chaos* 2006. V. 16. P. 023125-1-10.
21. *Flach S., Ovchinnikov A. A.* Static magnetization induced by time-periodic fields with zero mean // *Physica A.* 2001. V. 292. P. 268–276.
22. *Flach S., Miroshnichenko A. E., Ovchinnikov A. A.* ac-driven quantum spins: resonant enhancement of transverse dc magnetization // *Phys. Rev. B.* 2002. V. 65. P. 104438-1-6.
23. *Arimondo E.* Angular momentum detection of non-linear phenomena in paramagnetic resonance // *Ann. Phys.* 1968. V. 3. P. 425–447.

ФАЗОВЫЕ КОМПАКТНЫ

А.С. Пиковский, Ф. Розенау

1. Введение

В предмете этой статьи объединены два традиционных объекта нелинейной динамики: связанные автогенераторы и солитоны. Связанные автоколебательные системы обычно исследуются в рамках теории синхронизации (см., например, [1]). При слабой связи взаимодействие можно описать в фазовом приближении [2]. Для двух связанных автогенераторов оно приводит к уравнению Адлера [1]. Соответствующие фазовые модели широко используются при описании цепочек автогенераторов [3—5] и ансамблей с глобальной связью [2, 6, 7].

Для справедливости фазового приближения необходимо, чтобы коэффициент связи был меньше (по модулю), чем старший ненулевой ляпуновский показатель. Тогда можно рассматривать возмущения амплитуд как подчиненные возмущениям фаз. В отсутствие связи получающиеся уравнения для фаз имеют только нулевые ляпуновские показатели, поэтому *диссипативность или дисперсионность фазовой динамики зависят только от типа связи*. Если в центре внимания находятся свойства синхронизации, связь обычно считается диссипативной и стремящейся выровнять фазы.

В отличие от многих предшествующих работ, мы рассматриваем ниже цепочку осцилляторов с дисперсионной связью. Поскольку и в локальной динамике, и в связи диссипация отсутствует, многие свойства такие же, как у гамильтоновских систем общего вида. Аттракторов в таких системах нет, и можно ожидать существование нелинейных волн типа солитонов.

В последние годы наблюдается существенный прогресс в понимании динамики нелинейных гамильтоновских цепочек и полей. Одно из актуальных направлений связано с исследованием существенно нелинейных волновых уравнений. В таких уравнениях возможны уединенные волны, которые в отличие от обычных солитонов не имеют экспоненциальных хвостов, а сосредоточены в компактной области пространства. Эти волны — **компактоны** — были впервые описаны в работах [8, 9]. Обычно уравнения в частных производных, допускающие решения в виде компактонов, не интегрируемы; тем не менее, эти объекты довольно устойчивы, что скорее всего связано не с близостью к солитонной интегрируемости, а основано на проявлении сильной нелинейности, обуславливающей компактность. Многие уравнения не сохраняют энергию и могут приводить к взрывающимся решениям.

Тем не менее, численное моделирование показывает, что компактоны образуются из типичных начальных условий. Примеры можно найти в задачах о раскрытии молекулы ДНК [10], о волнах сжатия в гранулярной среде [11, 12] и об оседании взвеси [13].

В этой статье мы опишем компактоны в цепочке дисперсионно связанных автогенераторов, следуя публикациям [14, 15].

2. Основная модель

2.1. Фазовая цепочка и различные типы связи

Автогенератор с частотой ω описывается с помощью фазы φ , для которой справедливо $\frac{d\varphi}{dt} = \omega$. Уравнение для фаз связанных автогенераторов можно вывести в два этапа [1, 2]. Сначала, используя малость связи по сравнению с модулем наибольшего отрицательного показателя Ляпунова предельного цикла, можно записать уравнение для возмущенной фазы; для цепочки одинаковых элементов оно имеет вид

$$\frac{d\varphi_n}{dt} = \omega + \tilde{q}(\varphi_{n-1}, \varphi_n) + \tilde{q}(\varphi_{n+1}, \varphi_n). \quad (1)$$

Функция связи \tilde{q} 2π -периодична по каждому аргументу. На следующем этапе, используя малость связи по сравнению с частотой колебаний ω , можно усреднить (1) по периоду. Оставшаяся «медленная» компонента связи \tilde{q} зависит только от разностей фаз:

$$\frac{d\varphi_n}{dt} = \omega + q(\varphi_{n-1} - \varphi_n) + q(\varphi_{n+1} - \varphi_n), \quad (2)$$

где $q(\varphi + 2\pi) = q(\varphi)$. Вводя новые переменные

$$v_n = \varphi_{n+1} - \varphi_n, \quad (3)$$

перепишем уравнение для фаз в виде

$$\frac{dv_n}{dt} = q(-v_n) + q(v_{n+1}) - q(-v_{n-1}) - q(v_n). \quad (4)$$

Поскольку частота не фигурирует в (4), перенормируя время, можно сделать связь q величиной порядка единицы.

В общем случае q можно разложить на четную и нечетную составляющие $q(v) = q^o(v) + q^e(v)$, что дает

$$\begin{aligned} \frac{dv_n}{dt} &= q^e(v_{n+1}) - q^e(v_{n-1}) + q^o(v_{n+1}) + q^o(v_{n-1}) - 2q^o(v_n) = \\ &= \nabla_d q^e(v) + \Delta_d q^o(v), \end{aligned} \quad (5)$$

где Δ_d и ∇_d — дискретные операторы Лапласа и дифференцирования:

$$\Delta_d f = f_{n+1} + f_{n-1} - 2f_n, \quad \nabla_d f = f_{n+1} - f_{n-1}. \quad (6)$$

Чаще всего рассматривают случай $q(\varphi) = \sin \varphi$; эта нечетная связь диссипативна и приводит к уравнению $\dot{v}_n = \Delta_d \sin v$, аттрактором в котором является синхронный режим $v_n = 0$. Мы будем рассматривать только четную связь, для которой

$$\frac{dv_n}{dt} = q(v_{n+1}) - q(v_{n-1}) = \nabla_d q(v) \quad (7)$$

(индекс e ниже писать не будем). Эта связь дисперсионна, поскольку (7) сохраняет объем в фазовом пространстве и не имеет аттракторов. В основном мы ограничимся простейшим случаем $q(v) = \cos v$, описывающим цепочку

$$\frac{dv_n}{dt} = \cos v_{n+1} - \cos v_{n-1} = \nabla_d \cos v. \quad (8)$$

Для иллюстрации рассмотрим комплексную цепочку Гинзбурга — Ландау

$$\frac{dA_k}{dt} = A_k(1 - (1 + ic_1)|A_k|^2) + (c_2 + ic_3)(A_{k-1} + A_{k+1} - 2A_k). \quad (9)$$

Это уравнение уже осреднено по быстрому времени, но мы еще должны свести его к уравнению для фаз. Для этого будем считать коэффициенты c_2 и c_3 малыми по сравнению с показателем Ляпунова предельного цикла — в нашем случае -2 . Тогда в нулевом приближении $|A_n| = 1$. Записывая в этом приближении уравнение для фаз, получим

$$\dot{v}_k = (c_3 - c_1 c_2) \nabla_d \cos v_k + (c_2 + c_1 c_3) \Delta_d \sin v_k. \quad (10)$$

Мы видим, что дисперсионная связь пропорциональна $c_3 - c_1 c_2$, а нечетная связь пропорциональна $c_2 + c_1 c_3$.

Обратимся теперь к общим свойствам основной модели (7); при этом будем считать цепочку бесконечной. Кроме сохранения объема в фазовом пространстве, существуют еще несколько законов сохранения:

$$I_1 = \sum_n v_n, \quad (11)$$

$$I_2 = \sum_n (-1)^n v_n, \quad (12)$$

$$I_3 = \sum_n Q(v_n), \quad (13)$$

где $Q(v) = \int_0^v q(u) du$.

Кроме уже предполагавшейся четности $q(v) = q(-v)$, справедлива симметрия

$$v_n \rightarrow -v_{-n}. \quad (14)$$

Кроме того, при $q(v) = \cos v$ справедлива симметрия

$$v_n \rightarrow \pi + v_{-n}. \quad (15)$$

2.2. Малые возмущения однородного состояния

Любая постоянная $v_n = V$ есть решение (7). Для малого возмущения $\tilde{v}_n = v_n - V$ имеем

$$\frac{d\tilde{v}_n}{dt} = q'(V)(\tilde{v}_{n+1} - \tilde{v}_{n-1}). \quad (16)$$

Если $q'(V) = 0$, линейной эволюции нет и нужно рассматривать нелинейные члены. Для четной 2π -периодической функции $q'(0) = q'(\pi) = 0$, так что возмущение однородных состояний $V = 0$ и $V = \pi$ существенно нелинейно. В простейшем случае квадратичных экстремумов вместо (16) получим

$$\frac{d\tilde{v}_n}{dt} = -\frac{1}{2}(\tilde{v}_{n+1}^2 - \tilde{v}_{n-1}^2). \quad (17)$$

Уравнение (17) допускает преобразование подобия

$$\tilde{v} \rightarrow a\tilde{v}, \quad t \rightarrow \frac{t}{a}, \quad (18)$$

что означает медленную эволюцию малых возмущений.

Интересно, что (17) допускает частное аналитическое решение в виде $\tilde{v}_n = A(t) \sin \frac{2\pi n}{3}$. При этом амплитуда эволюционирует согласно выражению

$$\frac{dA}{dt} = \sqrt{3/16}A^2,$$

т. е. уходит в бесконечность — взрывается — за конечное время

$$A(t) = \frac{A_0}{1 - \sqrt{3/16}A_0 \cdot (t - t_0)}.$$

Так что малые возмущения могут нарастать до тех пор, пока не станут существенны старшие нелинейные члены. Существование взрывающихся мод означает относительную негрубость тривиального состояния $v_n = 0$ в уравнении (7).

2.3. Квазинепрерывное приближение

В квазинепрерывном приближении дискретные операторы (6) аппроксимируются производными (пространственный шаг равен h):

$$\nabla_d = 2h \left[\frac{\partial}{\partial x} + \frac{h^2}{6} \frac{\partial^3}{\partial x^3} \right], \quad \Delta_d = h^2 \frac{\partial^2}{\partial x^2}.$$

В результате получается

$$\frac{\partial v}{\partial t} = 2h \left[\frac{\partial}{\partial x} + \frac{h^2}{6} \frac{\partial^3}{\partial x^3} \right] q(v). \quad (19)$$

Подчеркнем, что уравнение (19) не является асимптотическим пределом цепочечной модели, поскольку малый параметр — отношение периода цепочки к длине возмущений — отсутствует. Тем не менее, как будет показано ниже, это приближение дает привильную оценку уединенных волн.

3. Бегущие волны

3.1. Бегущие волны в квазинепрерывном приближении: компактоны и коватоны

Положим $h = 1$ и будем искать решения уравнения в частных производных

$$\frac{\partial v}{\partial t} = \left(2 \frac{\partial}{\partial x} + \frac{1}{3} \frac{\partial^3}{\partial x^3} \right) q(v) \quad (20)$$

в виде бегущих стационарных волн $v = v(s) = v(x - \lambda t)$ на фоне тривиального состояния $V = 0$.

Подставляя выражение для стационарных волн в (20), получим после одного интегрирования

$$\lambda v + 2(q(v) - q(0)) + \frac{1}{3} \frac{d^2}{ds^2} q(v) = 0. \quad (21)$$

Постоянная интегрирования $q(0)$ обеспечивает обнуление «эффективной силы» $\lambda v + 2(q(v) - q(0))$ при $v = 0$. Отметим, что если в дополнение к $v = 0$ производная $q'(v)$ также зануляется при некотором $v = v_*$, то v_* также является критической точкой системы, и «эффективная сила» должна зануляться и в этой точке. Этого можно добиться подбором скорости λ , тогда должно выполняться условие

$$\lambda = \lambda_* \equiv 2 \frac{q(0) - q(v_*)}{v_*}. \quad (22)$$

Умножим теперь (21) на $\frac{dq}{ds}$ и после интегрирования получим

$$\lambda[vq(v) - Q(v)] + [q(v) - q(0)]^2 + \frac{1}{6} \left(\frac{dq}{ds} \right)^2 = 0, \quad (23)$$

где Q определено согласно (13). Перепишем теперь (23) как закон сохранения консервативной системы:

$$(q')^2 \left[\frac{1}{2} \left(\frac{dv}{ds} \right)^2 + U(v) \right] = 0, \quad (24)$$

в которой «потенциальная энергия» $U(v)$ определена как

$$U(v) = 3 \frac{(q(0) - q(v))^2 + \lambda[vq(v) - Q(v)]}{(q'(v))^2}. \quad (25)$$

Для типичной четной функции $q(v)$ с квадратичным экстремумом $\sim av^2$ уравнение (24) при малой амплитуде волн сводится к

$$v^2 \left[\frac{1}{2} \left(\frac{dv}{ds} \right)^2 + \frac{\lambda}{2a}v + \frac{3}{4}v^2 \right] = 0. \quad (26)$$

В скобках стоит «закон сохранения энергии» для квадратичного потенциала, соответствующие решения есть синус и косинус. Симметричное решение с минимумом при $v = 0$ имеет вид

$$v(s) = -\frac{\lambda}{3a} \left(1 + \cos \sqrt{\frac{3}{2}}s \right) = -\frac{2\lambda}{3a} \cos^2 \left(\sqrt{\frac{3}{8}}s \right). \quad (27)$$

Минимум этого решения достигает нуля, где решением является и тривиальное состояние $v = 0$. Обычно два решения обыкновенного дифференциального уравнения не пересекаются. Но в нашем случае точка $v = 0$ сингулярна, поскольку в ней зануляется коэффициент при старшей производной, и в этой точке единственность решения пропадает. В ней можно сшить два решения и получить компактон

$$v(s) = \begin{cases} -\frac{\lambda}{3a} \left(1 + \cos \sqrt{\frac{3}{2}}s \right) & \text{if } |s| \leq \pi \sqrt{\frac{2}{3}}, \\ 0 & \text{if } |s| > \pi \sqrt{\frac{2}{3}}. \end{cases} \quad (28)$$

Отметим, что амплитуда компактона пропорциональна скорости λ , в полном соответствии со свойством подобия (18). Ширина компактона от амплитуды не зависит.

Аналогично строится решение в общем случае (24). При $q(v) = \cos v$ имеем $Q(v) = \sin v$ и (24) принимает вид

$$\sin^2 v \left[\frac{1}{2} \left(\frac{dv}{ds} \right)^2 + U(v) \right] = 0 \quad (29)$$

с потенциалом

$$U(v) = 3 \frac{(\cos v - 1)^2 + \lambda(v \cos v - \sin v)}{\sin^2 v}. \quad (30)$$

Уравнение (29) сингулярно как при $v = 0$, так и при $v = \pi$. Пока $v < \pi$, амплитуда компактона v_m находится из условия $U(v_m) = 0$. Это дает соотношение между амплитудой и скоростью:

$$\lambda = \frac{(\cos v_m - 1)^2}{\sin v_m - v_m \cos v_m}. \quad (31)$$

Чтобы достигнуть второй сингулярной точки $v = \pi$, мы должны потребовать выполнения условия $U(v = \pi) = 0$, таким образом получается критическая скорость λ :

$$\lambda_c = \frac{(\cos \pi - 1)^2}{\sin \pi - \pi \cos \pi} = \frac{4}{\pi}. \quad (32)$$

Поскольку при λ_c потенциальная энергия симметрична $U_c(v) = U_c(\pi - v)$, решение в виде перепада, соединяющего $v = 0$ и $v = \pi$, также симметрично. Это решение можно шить с однородными состояниями $v = 0$ и $v = \pi$ и получить *компактный движущийся перепад*. Отметим, что в силу симметрии $v \rightarrow \pi - v$ перепад, соединяющий состояния $v = \pi$ и $v = 0$, имеет ту же скорость. Поэтому возможно сконструировать из двух перепадов, находящихся на произвольном расстоянии друг от друга, одну стационарную волну. Она выглядит как компактон с плоским верхом и была названа в работе [14] *коватон* (от древнееврейского «кова» — шляпа).

3.2. Бегущие волны в цепочке

В предыдущем разделе мы исследовали стационарные волны в квазинепрерывном приближении. Будем теперь искать стационарные волны вида $v_n(t) = v(t - bn)$, где $b = 1/\lambda$, в полной модели. Подставляя

в (7), получим так называемое уравнение с задержкой и опережением:

$$\dot{v} = q[v(t - b)] - q[v(t + b)]. \quad (33)$$

Интегрируя (33), получим

$$v(t) = \int_{t-b}^{t+b} [q(0) - q(v(s))] ds. \quad (34)$$

Постоянная интегрирования выбрана так, чтобы $v = 0$ было решением. Отметим, что в уравнениях (33), (34) сингулярностей нет.

В дальнейшем ограничимся случаем $q(v) = \cos v$, особенно интересны здесь решения, асимптотические к состояниям $v = 0$ и $v^* = \pi$. Как видно из (34), решение $v^* = \pi$ возможно только при $\lambda^* = 4/\pi$. Для нахождения перепада, соединяющего два однородных состояния, мы решали (34) при $b = \pi/4$ численно, используя простую итеративную процедуру, сходящуюся к искомому решению:

$$v_{k+1}(t) = \int_{t-\pi/4}^{t+\pi/4} [1 - \cos v_k(x)] dx \quad (35)$$

(k — номер итерации). Задав начальную волну $v_0(x)$ и используя схему Лагранжа высокой точности [16], мы получили решение, представленное на рис. 1.

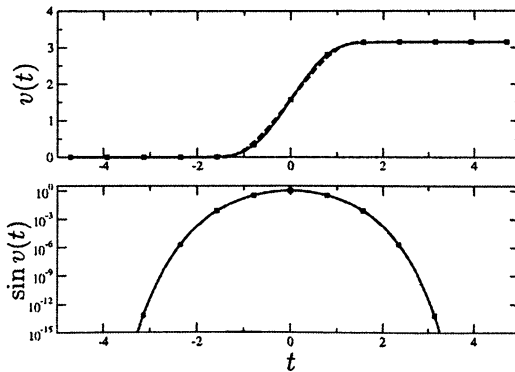


Рис. 1. Перепад в линейном и логарифмическом представлениях. Пунктиром показано решение в квазинепрерывном приближении.

Для того чтобы найти профиль компактона, мы использовали итеративный метод, предложенный В. И. Петвиашвили [17, 18]:

$$\tilde{v}(t) = \int_{t-b}^{t+b} (1 - \cos v_k(x)) dx, \quad v_{k+1} = \left(\frac{\|v_k\|}{\|\tilde{v}\|} \right)^{3/2} \tilde{v} \quad (36)$$

(использовалась норма L_1). Результат при $b = 1$ показан на рис. 2. Используя (35) и (36), можно найти бегущие волны при всех скоростях $0 < \lambda \leq \lambda_c$.

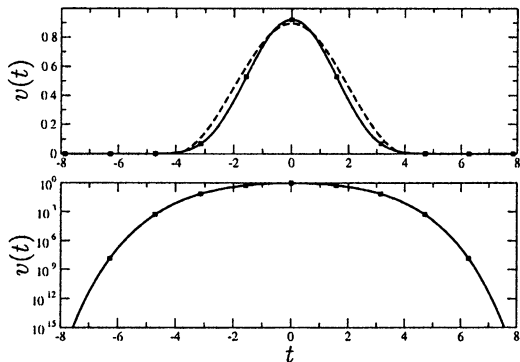


Рис. 2. Компактон при $\lambda = 0,5\lambda_c = 2/\pi$. Пунктиром показано решение в квазинепрерывном приближении.

3.3. Оценка хвостов

Очевидно, что в рамках интегрального уравнения (34) строго компактные решения невозможны. Однако можно оценить скорость убывания хвостов (ср. аналогичный анализ для упругих нелинейных волн в [19]). При малых v можно записать

$$v(t) = \int_{t-b}^{t+b} v^2(\tau) d\tau. \quad (37)$$

Предположим, что $v(t) \sim e^{-f(t)}$, причем $f(t)$ быстро растет при больших t . Тогда, оценивая интеграл по лемме Ватсона, получим

$$v(t) \approx e^{-2f(t-b)} \frac{1}{2f'(t-b)}. \quad (38)$$

Перепишем (38) как

$$e^{-f(t)} = e^{-2f(t-b)} \frac{1}{2f'(t-b)}$$

или как

$$f(t) = 2f(t-b) + \ln 2f'(t-b). \quad (39)$$

Пренебрегая логарифмами, получим уравнение $f_0(t) = 2f_0(t-b)$, решением которого является $f_0(t) = C2^{t/b} = C \exp\left(\frac{\ln 2}{b}t\right)$, где C — произвольная постоянная. Первую поправку $f_1(t)$ найдем из (39):

$$f_1(t) - 2f_1(t-b) = \ln 2f_0'(t-b) = \ln C \frac{\ln 2}{b} + \frac{\ln 2}{b}t.$$

Ища решение в виде $f_1(t) = At + B$, получим

$$At + B - 2A(t-b) - 2B = + \ln C \frac{\ln 2}{b} + \frac{\ln 2}{b}t.$$

Таким образом,

$$A = -\frac{\ln 2}{b}, \quad B = -2 \ln 2 - \ln C \frac{\ln 2}{b} = \ln \frac{b}{C4 \ln 2},$$

и окончательно

$$f(t) = C \exp\left(\frac{\ln 2}{b}t\right) - t \frac{\ln 2}{b} + \ln \frac{b}{C4 \ln 2}.$$

При больших t последним членом можно пренебречь:

$$v \approx e^{-f(t)} = \exp\left[-C \exp\left(\frac{\ln 2}{b}t\right) + t \frac{\ln 2}{b}\right]. \quad (40)$$

Этот результат показывает, что хвосты решения убывают сверхэкспоненциально.

Возможно, наиболее удивительным результатом вышеприведенного анализа является то, что решения, полученные в квазинепрерывном пределе, в котором только один дисперсионный член был принят во внимание, так хорошо аппроксимируют точное решение для цепочки. Можно было бы ожидать, что квазинепрерывное приближение справедливо только для длинных волн, но оно хорошо работает и для коротких.

4. Численное моделирование

В этом разделе мы опишем некоторые результаты численного моделирования уравнения (7).

Эволюция начального импульса. Начальный импульс

$$v_n(0) = A \sin \left(\pi \frac{n - n_0}{M} \right) \quad (41)$$

характеризовался двумя параметрами — амплитудой A и шириной M . На рис. 3, *a* показано, как относительно широкий импульс эволюционирует в цепочку компактонов, упорядоченных по амплитудам. В дополнение к компактонам из начального импульса большой амплитуды образуется коватон (рис. 3, *б*).

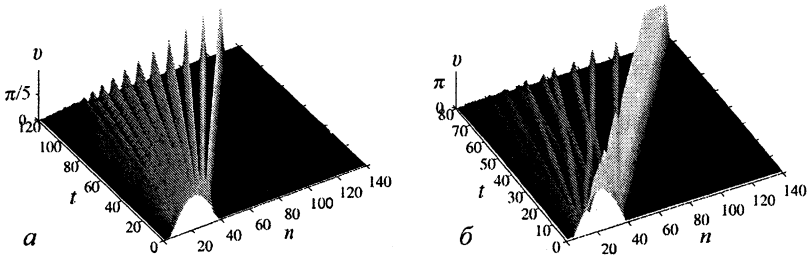


Рис. 3. Эволюция начального импульса (41): *a* — $M = 34$, $A = \pi/4$; *б* — $M = 34$, $A = \pi$.

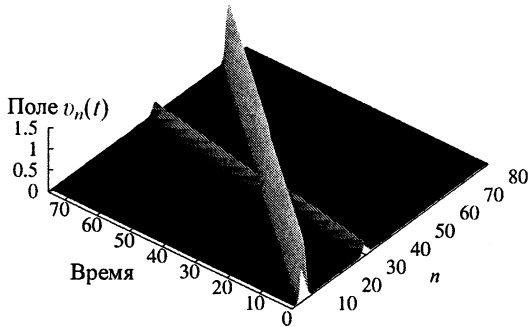


Рис. 4. Столкновение компактонов со скоростями 0, 2 и 0, 9

Столкновения компактонов и коватонов. Столкновение двух компактонов показано на рис. 4. При столкновении компактоны не раз-

рушаются, но их амплитуды немного меняются. Начальные амплитуды были 0,27595 и 1,3955, а после столкновения 0,27407 — 0,27411 (в зависимости от начального расстояния между ними) и 1,3956. Малое возмущение, оставшееся на месте столкновения, медленно эволюционирует в новые компактоны и антикомпактоны. На рис. 5 приведены амплитуды новых компактонов и времена их образования.

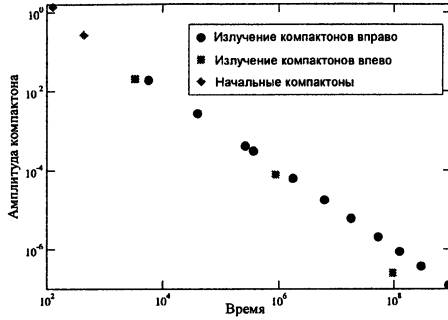


Рис. 5. Амплитуды компактонов, образовавшихся после столкновения, показанного на рис. 4.

На рис. 6 показано столкновение компактона и антикомпактона, имеющих одинаковую скорость 0,5. В то время как оба компактона выходят из столкновения мало изменившимися, дополнительно рождается пара компактон-антикомпактон большой амплитуды. Это связано с неустойчивостью, обсуждавшейся в разд. 2.2.

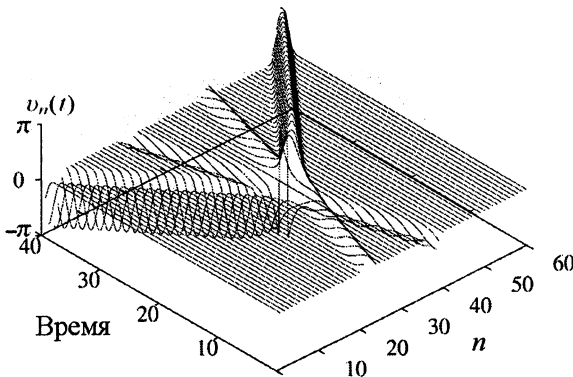


Рис. 6. Столкновение компактона и антикомпактона

Рисунок 7 показывает столкновение компактона и коватоны. Здесь также амплитуда компактона слегка уменьшается (с 0,560164 до 0,556288), и на месте столкновения остаются малые «брызги».

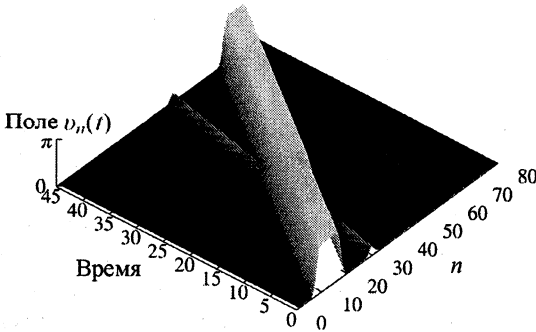


Рис. 7. Столкновение коватоны ($\lambda = 4/\pi$) с компактоном ($\lambda = 0,4$)

Интерпретация в терминах исходных фаз. В заключение этого раздела обсудим полученные решения в терминах исходных фаз автогенераторов. Согласно (3) фазы получаются путем суммирования переменных $v_n(t)$: $\varphi_n = \sum^n v_k$. Поэтому компактное решение для $v_n(t)$ отвечает распространяющемуся перепаду, разделяющему две области постоянной фазы. Состояние цепочки с $v_n = 0$ отвечает синфазному режиму, а при $v_n = \pi$ соседние генераторы оперируют в противофазе. Поэтому коватон описывает участок противофазных колебаний, распространяющихся по синфазной области.

5. Теория возмущений для компактонов

Вернемся к рассмотрению комплексной цепочки Гинзбурга — Ландау (9). Во втором порядке малости по коэффициентам связи фазовое уравнение имеет вид

$$\begin{aligned} \dot{v}_n = & (c_3 - c_1 c_2) \nabla_d \cos v + (c_2 + c_1 c_3) \Delta_d \sin v + \\ & + \frac{c_2^2}{2} \nabla_d (\sin v \nabla_d \cos v) + \frac{c_2 c_3}{2} \Delta_d (\cos v \nabla_d \cos v) - \\ & - \frac{c_2 c_3}{2} \Delta_d (\sin v \Delta_d \sin v) - \frac{c_3^2}{2} \Delta_d (\cos v \Delta_d \sin v). \end{aligned} \quad (42)$$

Члены высших порядков нелегко классифицировать на диссипативные и дисперсионные, в общем случае они приводят к затуханию или

нарастанию компактонов. На рис. 8 мы показываем эволюцию компактона с начальной скоростью $\lambda = 0,5$. Коэффициенты связи фиксированы: $c_2 = 0$, $c_3 = 0,01$, а параметр c_1 меняется. При этих параметрах (42) сводится к уравнению

$$\frac{1}{c_3} \frac{dv_n}{dt} = \nabla_d \cos v + c_1 \Delta_d \sin v - \frac{c_3}{2} \Delta_d (\cos v \Delta_d \sin v). \quad (43)$$

Здесь два неконсервативных слагаемых. При $v \approx 0$ третий член диссипативен. Второй член при $c_1 < 0$ дает нарастание. При больших $|c_1|$ второе слагаемое преобладает, т. е. противофазный режим более устойчив, чем синфазный. Поэтому при больших $|c_1|$ наблюдается ускорение и рост компактона, а при малых $|c_1|$ — замедление и затухание. Для конкретного случая рис. 8 при $c_1 \approx -0,0065$ два противоположных эффекта компенсируют друг друга и компактон практически не изменяется.

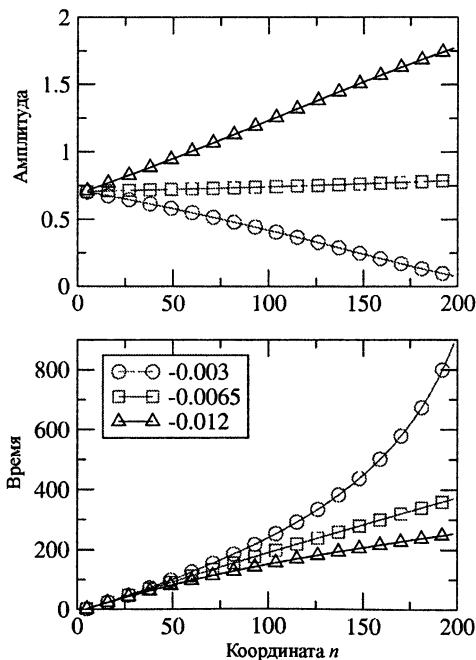


Рис. 8. Амплитуды (верхний график) и траектории компактона с $\lambda = 0,5$ при трех значениях параметра c_1 для цепочки Гинзбурга — Ландау (9) с $c_2 = 0$, $c_3 = 0,01$.

Для компактонов малой амплитуды запишем (43) в виде

$$\dot{v}_n = -\frac{1}{2}\nabla_d v^2 + \alpha\Delta_d v + \beta\Delta_d^2 v, \quad (44)$$

что в квазинепрерывном пределе дает

$$\frac{\partial v}{\partial t} = -\left(\frac{\partial}{\partial x} + \frac{1}{6}\frac{\partial^3}{\partial x^3}\right)v^2 + \alpha\frac{\partial^2}{\partial x^2}v + \beta\frac{\partial^4}{\partial x^4}v. \quad (45)$$

Для построения теории возмущений по малым α , β используем интегралы невозмущенной (т. е. при $\alpha = \beta = 0$) версии (45) (ср. с (11), (13)) $I_1 = \int_{-\infty}^{\infty} dx v$ и $I_3 = \int_{-\infty}^{\infty} dx v^3$. Запишем компактон в виде $v_c(x, t) = \frac{4\lambda}{3} \cos^2 k(x - \lambda t)$, где $k = \sqrt{3/8}$. Эволюция I_3 приводит к уравнению

$$\frac{d}{dt} \int_{-\pi/2k}^{\pi/2k} dx v_c^3 = 3 \int_{-\pi/2k}^{\pi/2k} dx v_c^2 \left(\alpha \frac{\partial^2}{\partial x^2} + \beta \frac{\partial^4}{\partial x^4} \right) v_c. \quad (46)$$

Вычисляя интеграл, получим

$$\frac{d\lambda}{dt} = -\frac{3}{5} \left(\alpha - \frac{3}{2}\beta \right) \lambda. \quad (47)$$

Поскольку интеграл I_1 сохраняется в присутствии возмущения (45), можно вычислить образующийся хвост. Во-первых, запишем сохранение I_1 как

$$\int dx v_c(x, t) + \int dx v_{tail}(x, t) = I_1 = \text{const.} \quad (48)$$

Теперь будем считать, что хвост простирается от начального $x = 0$ до текущего $X = \int_0^t \lambda(\tau) d\tau$ положения компактона. Тогда из (48) и (47) следует

$$\frac{d}{dX} \int_0^X dx v_{tail}(x, t) = -\frac{1}{\lambda} \frac{d}{dt} \int dx v_c(x, t) = \frac{2\pi}{5k} \left(\alpha - \frac{3}{2}\beta \right). \quad (49)$$

Это означает, что хвост постоянен:

$$v_{tail} = \frac{2\pi}{5k} \left(\alpha - \frac{3}{2}\beta \right). \quad (50)$$

Формулы (47) и (50) хорошо согласуются с результатами численных расчетов.

В статье описана интересная динамика цепочки автогенераторов с дисперсионной связью. Основное внимание было уделено бегущим волнам в виде компактонов и коватонов. Их компактность связана с отсутствием линейных волн. Подобные решения возможны и в других системах, описываемых существенно нелинейными уравнениями.

Литература

1. Пиковский А., Розенблюм М., Куртс Ю. Синхронизация. Фундаментальное нелинейное явление. М.: Техносфера, 2003.
2. Kuramoto Y. Chemical Oscillations, Waves and Turbulence. Berlin: Springer, 1984.
3. Ermentrout G. B., Kopell N. Frequency plateaus in a chain of weakly coupled oscillators, I // SIAM J. Math. Anal. 1984. 15(2). P. 215–237.
4. Ren L., Ermentrout B. Phase locking in chains of multiple-coupled oscillators // Physica D. 2000. 143(1–4). P. 56–73.
5. Topaj D., Pikovskiy A. Reversibility versus synchronization in oscillator lattices // Physica D. 2002. 170(2). P. 118–130.
6. Daido H. Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions // Phys. Rev. Lett. 1992. 68(7). P. 1073–1076.
7. Strogatz S. H. From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators // Physica D. 2000. 143(1–4). P. 1–20.
8. Rosenau P., Hyman J. M. Compactons: Solitons with finite wavelength // Phys. Rev. Lett. 1993. 70(5). P. 564–567.
9. Rosenau P. Nonlinear dispersion and compact structures // Phys. Rev. Lett. 1994. 73. P. 1737–1741.
10. Takeno S. Compacton-like modes in model dna systems and their bearing on biological functioning // Phys. Lett. A. 2005. 339. P. 352–360.
11. Нестеренко В. Ф. Распространение нелинейных импульсов сжатия в зернистых средах // ПМТФ. 1983. № 5. С. 136.
12. Лазариди А. Н., Нестеренко В. Ф. Обнаружение уединенных волн нового типа в одномерной зернистой среде // ПМТФ. 1985. № 3. С. 1155.
13. Rubinstein J. Evolution equations for stratified dilute suspensions // Phys. Fluids A. 1990. 2(1). P. 3–6.
14. Rosenau P., Pikovskiy A. Phase compactons in chains of dispersively coupled oscillators // Phys. Rev. Lett. 2005. 94. P. 174102.
15. Pikovskiy A., Rosenau P. Phase Compactons // Physica D. 2006. 218. P. 56–69.
16. Abramowitz M., Stegun I. A. Handbook of Mathematical Functions. Department of Commerce USA, Washington, D.C., 1964.
17. Петвиашвили В. И. Об уравнении обыкновенного солитона // Физика плазмы. 1976. Т. 2, № 3. С. 469.
18. Petviashvili V. I. Multidimensional and dissipative solitons // Physica D. 1981. 3(1–2). P. 329–334.
19. Chatterjee A. Asymptotic solution for solitary waves in a chain of elastic spheres // Phys. Rev. E. 1999. 59(5). P. 5912–5919.

ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ДИНАМИКА ТОПОЛОГИЧЕСКИХ ОСОБЕННОСТЕЙ ВОЛНОВЫХ ПОЛЕЙ: ИСТОЧНИКИ, СТОКИ, ДЫРКИ, ДЕФЕКТЫ

А. Б. Езерский

1. Введение

Во многих случаях, после того как проведены эксперименты и построена соответствующая теоретическая модель, исследователи пытаются сравнить эмпирические данные с расчетом. Если речь идет о пространственно-временной динамике полей, то проведение такого сравнения представляется, вообще говоря, совсем не тривиальной задачей. Строго говоря, нужно определить, насколько отличаются в каждой точке полевые переменные, зависящие от координат и времени $\{U(x, y, z, t), V(x, y, z, t)\dots\}$, в эксперименте и в теории. Гипотетически это сделать можно, а практически нет. Поэтому, когда изучают нелинейную динамику полей в какой-либо системе, то исследуются и сравниваются с моделью не сами поля, а топологические особенности этих полей. Конечно, такое сравнение имеет смысл, если речь идет не о какой-то технической проблеме, для которой именно поля или средние величины этих полей и представляют интерес.

В данной статье речь пойдет о том, какие топологические особенности полей изучаются в настоящее время и как они идентифицируются в экспериментах.

Наиболее полная информация о течениях содержится в трехмерных фильмах. Эксперименты, где снимаются такие фильмы, встречаются достаточно редко, тогда как регистрация двумерных полей — это наиболее распространенный подход к эксперименту. Очень часто из двумерного поля выбирается некоторая линия точек и регистрируется так называемая $(X-T)$ диаграмма [1]. Если движения в системе одномерные, то такая диаграмма адекватно описывает динамику всего поля. Характерные структуры, регистрируемые на $(X-T)$ диаграммах получили определенные названия: источники, стоки, дырки, дефекты и т. д. Эти объекты представляют из себя строительные блоки, на основе которых проводится анализ пространственно-временного хаоса в неравновесных системах. Анализ последних теоретических и экспериментальных данных по динамике этих объектов будет проведен в данной работе.

2. Источники и стоки

В нелинейной среде, как известно, достаточно часто наблюдается конкуренция волновых возмущений, распространяющихся в противополо-

ложных направлениях. В результате такой конкуренции поле разбивается на области, в каждой из которых преобладают волновые возмущения, бегущие в одну сторону. Поэтому возникают участки среды, откуда возмущения расходятся, — источники, и участки, куда возмущения сходятся, — стоки.

Понять структуру источников и стоков можно на очень простой модели. Будем считать, что формообразование происходит в активной среде и у нас есть конкуренция двух встречных волн:

$$u = A(X, T) \exp i(\omega_R t - k_R x) + B(X, T) \exp i(\omega_L t + k_L x) + c.c.$$

При конкуренции таких возмущений уравнения для медленно меняющихся амплитуд представляются в виде [2]

$$\begin{aligned} \frac{\partial A}{\partial T} + V_{g,R} \frac{\partial A}{\partial X} &= \epsilon A - |A|^2 A - \delta |B|^2 A, \\ \frac{\partial B}{\partial T} - V_{g,L} \frac{\partial B}{\partial X} &= \epsilon B - |B|^2 B - \delta |A|^2 B. \end{aligned} \quad (1)$$

В системе (1) все коэффициенты действительны. Если коэффициент взаимодействия δ встречных волн больше единицы, то возникает конкуренция, когда за счет нелинейного затухания одна волна подавляет другую. Для стационарного решения можно построить фазовую плоскость, которая выглядит так, как показано на рис. 1. Для амплитуд A и B существуют периодические по пространству режимы. Для того чтобы понять, как выглядят пространственно-временные диаграммы поля u , надо учесть гармонические волны. При этом вид диаграмм существенно зависит от соотношения знаков групповой и фазовой скоростей волн. Обычно предполагают, что знаки одинаковы, т. е. в системе есть нормальная дисперсия. Тогда для периодических траекторий, близких к сепаратрисам, $(X-T)$ диаграммы имеют вид, показанный на рис. 2, б, где можно выделить области источников и стоков.

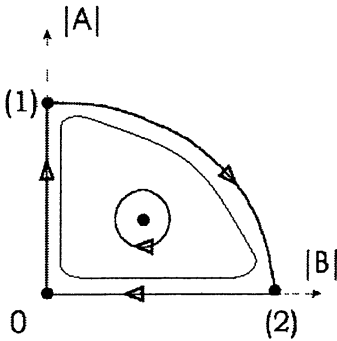


Рис. 1. Фазовая плоскость (1)

Причем стоки, как это видно на рис. 2, б, локализованы в пространстве, а источники размазаны. Если групповая скорость отрицательна, то источники и стоки меняются местами (рис. 2, в) и теперь источники локализованы, стоки нет. Этот вывод получен для чисто действительных коэффициентов в системе (1), что бывает лишь в достаточно ограниченном числе случаев. В общем виде система, описывающая взаимодействие волновых возмущений в неравновесных средах, выглядит как

$$\begin{aligned} & \tau_0 \left(\frac{\partial A}{\partial T} + V_{g,R} \frac{\partial A}{\partial X} \right) = \\ & = \varepsilon(1+ic_0)A + \xi_0^2(1+ic_1) \frac{\partial^2 A}{\partial X^2} - (1+ic_2)|A|^2 A - \delta(1+ic_3)|B|^2 A, \\ & \tau_0 \left(\frac{\partial B}{\partial T} - V_{g,L} \frac{\partial B}{\partial X} \right) = \\ & = \varepsilon(1+ic_0)B + \xi_0^2(1+ic_1) \frac{\partial^2 B}{\partial X^2} - (1+ic_2)|B|^2 B - \delta(1+ic_3)|A|^2 B. \end{aligned}$$

Это два связанных уравнения Гинзбурга — Ландау. Численный расчет режимов в этой системе проведен в большом количестве работ (см., например, [3, 4] и цитируемую там литературу).

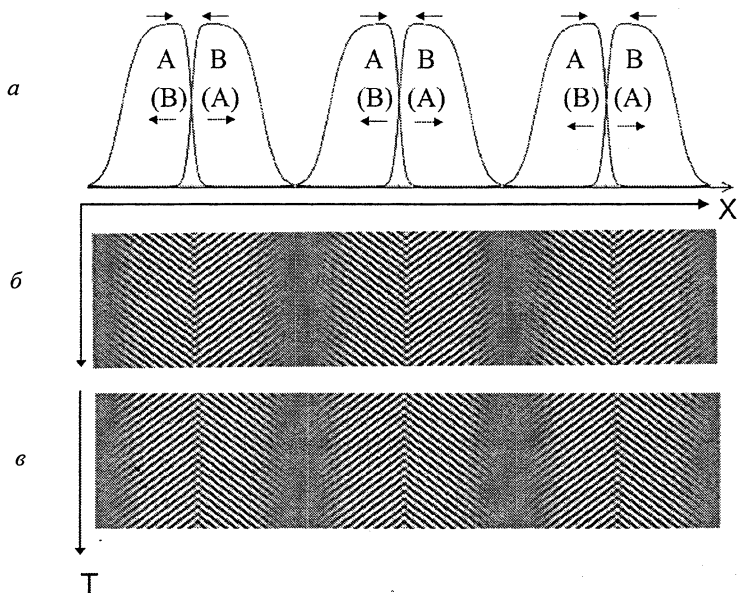


Рис. 2. Источники и стоки на пространственно-временной диаграмме: *a* — амплитуда полей *A* и *B*, без скобок — нормальная дисперсия, в скобках — аномальная; *b* — (*X*-*T*) диаграмма для нормальной дисперсии; *в* — (*X*-*T*) диаграмма для аномальной дисперсии.

Результаты расчетов показали [3, 4], что основной вывод, полученный на простой модели, остается верным и для более сложных уравнений. Стационарными и локализованными в том случае, когда знаки групповой и фазовой скоростей совпадают, могут быть только стоки, а не источники.

Источники и стоки обнаружены во многих экспериментах. Например, в [5] — при исследовании конвекции в слое жидкости, подогреваемом снизу, в [6] — при возникновении возмущений в тонком слое жидкости, стекающей с вращающегося цилиндра, частично погруженного в жидкость, в [7] — при изучении режимов конвекции у тонкого нагретого цилиндра, расположенного параллельно свободной поверхности жидкости, в [8] — для течения Куэтта между цилиндрами, вращающимися в противоположные стороны, и в других системах. Как было обнаружено в экспериментах, при небольшой надкритичности реализуются регулярные структуры, которые становятся неупорядоченными при увеличении надкритичности. Стационарными и локализованными в экспериментах могли быть как источники, так и стоки, однако только в некоторых экспериментах авторы обращали внимание на величину групповой скорости. Так, в [7] дисперсия возмущений была нормальная и стационарными и локализованными были стоки. В [8] дисперсия возмущений была аномальной и были обнаружены стационарные локализованные источники.

3. Дырки

Термин дырки стал широко использоваться в литературе после публикации работ [9, 10]. В этих работах было найдено точное решение уравнения Гинзбурга — Ландау:

$$\tau_0 \left(\frac{\partial A}{\partial T} + V_g \frac{\partial A}{\partial X} \right) = \varepsilon(1 + ic_0)A + \xi_0^2(1 + ic_1) \frac{\partial^2 A}{\partial X^2} - (1 + ic_2)|A|^2 A, \quad (2)$$

связывающее области с разными частотами и волновыми числами. Не вдаваясь в детали нахождения решения, приведем его вид. Если перейти в систему отсчета, движущуюся со скоростью V_g , и сделать перенормировки, то уравнение (2) можно свести к виду

$$\frac{\partial A}{\partial T} = (1 + ic_0)A + (1 + ic_1) \frac{\partial^2 A}{\partial X^2} - (1 + ic_2)|A|^2 A. \quad (3)$$

Это уравнение имеет простое решение в виде гармонической волны с постоянной амплитудой:

$$A = A_0 \exp(i(\Omega t + kx)), \text{ где } \Omega = -c_1 + (c_2 - c_1)k^2.$$

Как показано в [10, 11], есть точное решение этого уравнения в виде

$$A = \frac{b_1 \exp(k\xi) + b_2 \exp(k\xi)}{[\exp(k\xi) + \exp(-k\xi)]^{(1+ia_0)}} \exp[i(Kx - \Omega t)].$$

Здесь $\xi = x - c_{gr}t$ — скорость дырок; b_1, b_2 — комплексные функции параметров c_1, c_2 ; k, K, Ω, a_0 — действительные функции c_1, c_2 .

При $\xi \rightarrow +\infty$ $A = \sqrt{1 - k_1^2} \exp(i(\Omega_1 t - k_1 x))$, а при $\xi \rightarrow -\infty$ $A = \sqrt{1 - k_2^2} \times \exp(i(\Omega_2 t - k_2 x))$, где $\Omega_{1,2} = -c_1 + (c_2 - c_1)k_{1,2}^2$.

Таким образом, дырка связывает два решения в виде плоских волн. Амплитуда и фаза точного решения для дырки показана в виде графиков (рис. 3).

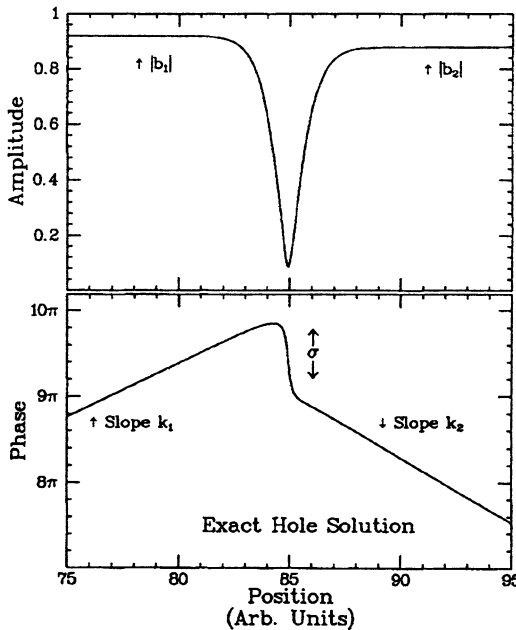


Рис. 3. Амплитуда и фаза точного решения для дырки при $c_1 = 0,5$, $c_2 = 2,3$, $k_1 = 0,392$, $k_2 = -0,477$ [11].

Режимы с дырками были обнаружены во многих гидродинамических экспериментах [11, 12]. Схема одного из таких экспериментов приведена на рис. 4. В протяженном слое жидкости, боковые границы которого поддерживаются при постоянной температуре, при превышении критической разности температур возникает неустойчивость. Неустойчивость приводит к возникновению конвективных роликов, которые расположены под углом к боковым стенкам. Эти ролики не стационарны — они смещаются вдоль кюветы. В эксперименте регистрировались пространственно-временные диаграммы. Линия, на которой в разные моменты времени записывалось изображение, — это центральная ось канала, показанного в нижней части рис. 4. Диаграммы выглядят, как это показано на рис. 5. На рис. 5, а видны периодические волны, распространяющиеся влево и

вправо, а на фоне этих волн есть темные полосы. Амплитуда возмущений (рис. 5, б) в этих темных полосах существенно меньше, чем в других местах.

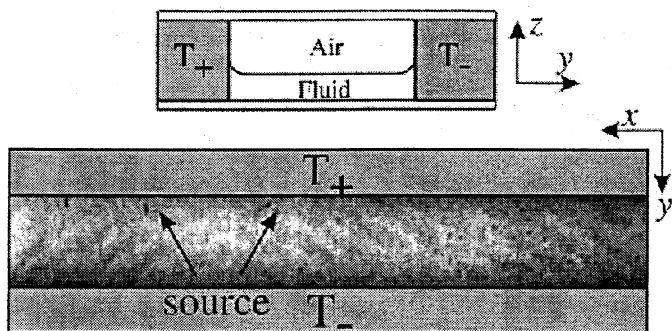


Рис. 4. Схема эксперимента [12] по изучению динамики дырок

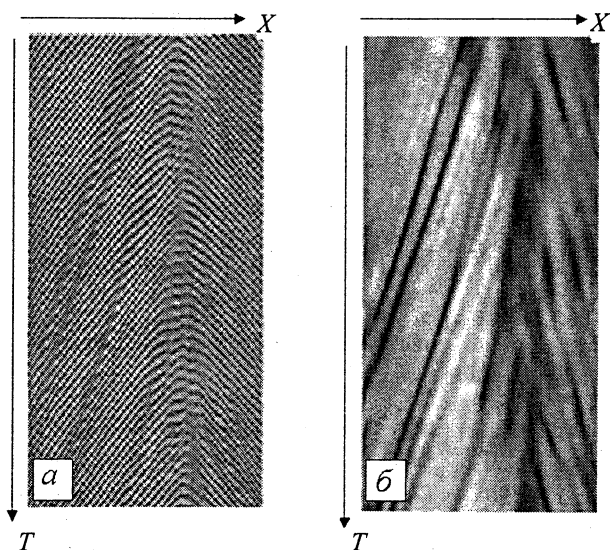


Рис. 5. Пространственно-временные диаграммы в эксперименте с конвекцией при нагреве боковых границ. Записывалась яркость изображения вдоль центральной линии канала: а — изображение и б — амплитуда возмущений.

Темные треки на пространственно-временной диаграмме рис. 5, б были идентифицированы авторами [12] как дырки Нозакки — Беки. Понятно, что в эксперименте фиксировалась яркость изображения, а не поле

скорости или температуры в жидкости. Связь яркости изображения с гидродинамическими полями достаточно сложная и, вообще говоря, нелинейная. Поэтому сравнивать можно только качественные особенности полей. В качестве аргументов, которые выдвинуты в [12] для подтверждения того, что в эксперименте действительно наблюдаются дырки Нозакки — Беки, приведены следующие. Темные треки — это почти стационарные волны затемнения. Понятно, что система ограничена и нет возмущений, определенных от $-\infty$ до $+\infty$. Волны затемнения распространяются с постоянной скоростью, что соответствует решению, полученному Нозакки — Беки. Скорость этих стационарных волн меньше фазовой скорости возмущений. Изменение амплитуды и фазы возмущений в окрестности дырки совпадает с решением Нозакки — Беки. Это показано на рис. 6, взятом из работы [12].

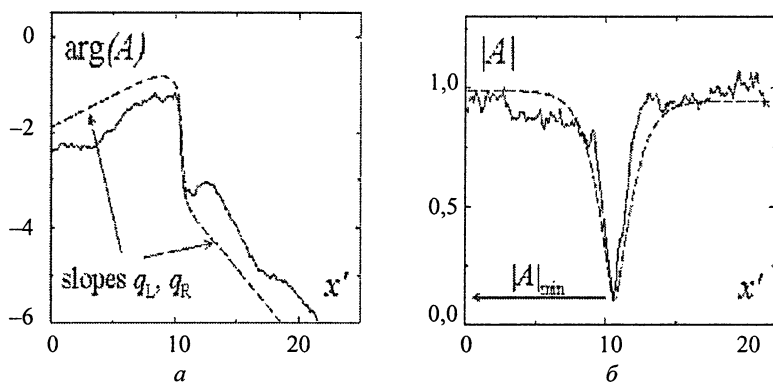


Рис. 6. Сравнение экспериментальных и теоретических распределений фазы (а) и амплитуды (б) в окрестности дырки. Экспериментальные данные представлены непрерывной линией, а теоретические — пунктирной.

Как можно заметить, в экспериментальных данных есть все характерные детали, которые присутствуют в точном решении: скачок фазы, изменение наклона фазы при удалении на большое расстояние, соответствующее изменению волнового числа, уменьшение амплитуды. Именно по этим характерным деталям идентифицировались «дырки». В экспериментах дырки были обнаружены также при исследовании конвекции в кольцевой системе [11].

4. Дефекты

Когда говорят о дефектах, возникающих при развитии неустойчивости в протяженных системах, то подразумевают, по крайней мере, два

различных объекта. Если, например, в двумерном поле есть система роликов, причем в одной полуплоскости на один пространственный период больше, чем в другой, то говорят, что в такой структуре есть топологический дефект. Такой топологический дефект является полным аналогом краевой дислокации в твердых телах. Только в твердых телах роль дополнительного пространственного периода играет слой атомов. Краевые дислокации были постулированы Бюргерсом в начале XX века для объяснения пластических деформаций в твердых телах, а затем были открыты экспериментально после изобретения электронного микроскопа. Макроскопический аналог краевой дислокации активно используется сейчас для анализа нарушений порядка в пространственно-периодических структурах, возникающих при термоконвекции, в параметрически возбуждаемых волнах, в следах за обтекаемыми телами. Динамике топологических дефектов в двумерных структурах была посвящена одна из лекций на школе по нелинейным волнам 2002 года [13]. Другой тип дефекта — это пространственно-временной дефект. На пространственно-временной диаграмме он выглядит, как это показано на рис. 7. Если забыть, что по вертикальной оси на рис. 7 откладывается не координата, а время, то невозможно найти различий между пространственно-временными дефектами и мгновенным снимком двумерной роликовой структуры с дефектом.

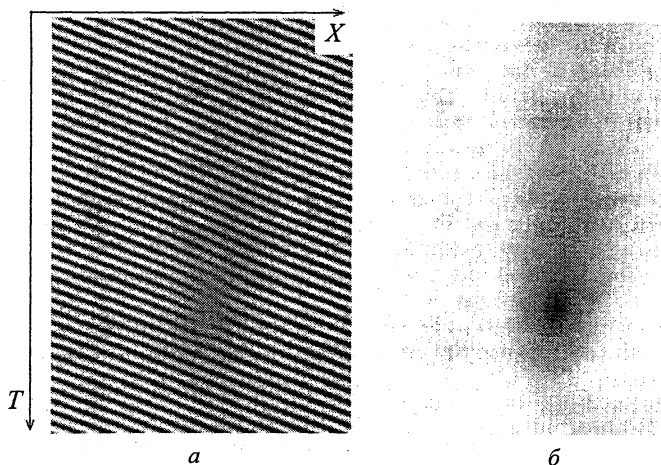


Рис. 7. Пространственно-временная диаграмма с дефектом (а) и амплитуда пространственно-временной диаграммы в яркостном представлении (б). В точке, где располагается дефект, амплитуда равна нулю.

Экспериментально возникновение пространственно-временных дефектов было подробно исследовано в [8]. Дефекты генерировались на фоне волновых возмущений — спиральных волн, возникающих в течении

Куэтта — Тэйлора между цилиндрами, вращающимися в разные стороны. Эксперимент проводился при следующих параметрах: отношение длины цилиндров ($L = 27,4$ см) к размеру зазора между цилиндрами ($d = 0,59$ см) составляло $L/d = 46,5$. Числа Рейнольдса $R_o = 2\pi f_o b d / \nu$ и $R_i = 2\pi f_i a d / \nu$, определенные по частотам вращения внешнего и внутреннего цилиндров f_i , f_o (ν — кинематическая вязкость жидкости, d — диаметр внутреннего цилиндра), были равны соответственно $R_o = -622$ и $R_i = 344$. При таких числах Рейнольдса в системе возникают спирали, которые распространяются в противоположных направлениях вдоль оси цилиндров. В результате нелинейной конкуренции побеждает одна из спиралей. На фоне спиралей периодически во времени могут возникать дефекты. Для визуализации структур в жидкость добавлялись мелкие частицы Kalliroscope AQ1000. Пространственно-временная диаграмма, соответствующая такому режиму, показана на рис. 8. Подобная последовательность дефектов наблюдалась также в экспериментах [14, 15].

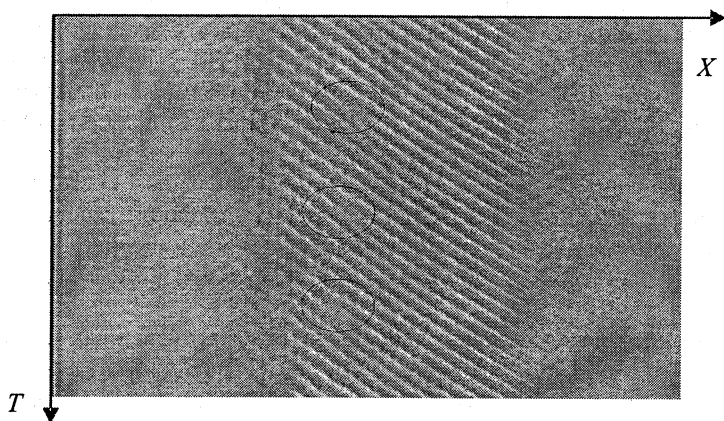


Рис. 8. Пространственно-временная диаграмма для режима спиралей с дефектами в течении Куэтта — Тэйлора. Ось X направлена вдоль оси цилиндров.

Для моделирования таких режимов были проведены численные расчеты уравнения Гизбурга — Ландау. В связи с тем, что в результате конкуренции победила одна из спиралей, использовалось лишь одно из уравнений системы, описывающей взаимодействие встречных волн:

$$\frac{\partial A}{\partial T} = A + (1 + ic_1) \frac{\partial^2 A}{\partial X^2} - (1 + ic_2) |A|^2 A. \quad (4)$$

Принципиальным моментом для получения решений уравнения (4) в виде периодической последовательности дефектов был выбор граничных условий. Следует отметить, что решения уравнения (4) исследовались численно на протяжении длительного времени [16, 17] и к настоящему времени подробно изучены режимы, возникающие при различных параметрах. В большинстве работ это сделано лишь для периодических граничных условий. В данном случае такие граничные условия не отвечают условиям эксперимента. Возмущения, возникающие в эксперименте, имеют амплитуду, равную нулю на границах системы. Поэтому мы искали решение уравнения (4) с граничными условиями $A(-L/2) = A(L/2) = 0$.

Целью численных расчетов было обнаружение режима с периодической генерацией дефектов в уравнении Гинзбурга — Ландау. Необходимо отметить, что область генерации дефектов для периодических граничных условий была хорошо исследована (см., например, [18]). Такой режим наблюдался в области, соответствующей неустойчивости Бенджамина — Фейра ($1 - c_1 c_3 < 0$). Мы обнаружили, что при однородных граничных условиях режим генерации дефектов возникает при коэффициентах, соответствующих устойчивому режиму ($1 - c_1 c_3 < 0$). Полученные в расчетах поле амплитуды $|A|$, фазовое поле $\arg A(x, t)$ и фазовое поле процедуры демодуляции, описанной, например, в [13], показаны на рис. 9. Мы обнаружили, что имеет место мультистабильность: при одних и тех же коэффициентах уравнения (4) режим, возникающий в численном счете, зависит от начальных условий: дефекты могут быть на различном расстоянии от стенок и в системе наблюдается различный временной период. Разбиение плоскости параметров на области, соответствующие разным режимам, показано на рис. 10.

Детальное исследование полей, полученных в численных расчетах, показало, что в окрестности дефектов линии равной амплитуды A представляют собой эллипсы, вытянутые вдоль линий $x - Vt$. В окрестности этой линии сгущаются уровни постоянной фазы.

Структуру такого пространственно-временного дефекта можно понять, используя простое решение уравнения Гинзбурга — Ландау в окрестности точки, где располагается дефект. Мы представили решения (4) в окрестности дефекта в виде

$$A = ix + (x - Vt), \quad (5)$$

где x и t — это координаты и время, отсчитываемые от положения дефекта $x = 0$, $t = 0$. Тогда амплитуда $|A(x, t)| = \sqrt{x^2 + (x - Vt)^2}$, а фаза $\arg(A(x, t)) = \arctan(x/(x - Vt))$. Соответствующие графики показаны на рис. 11.

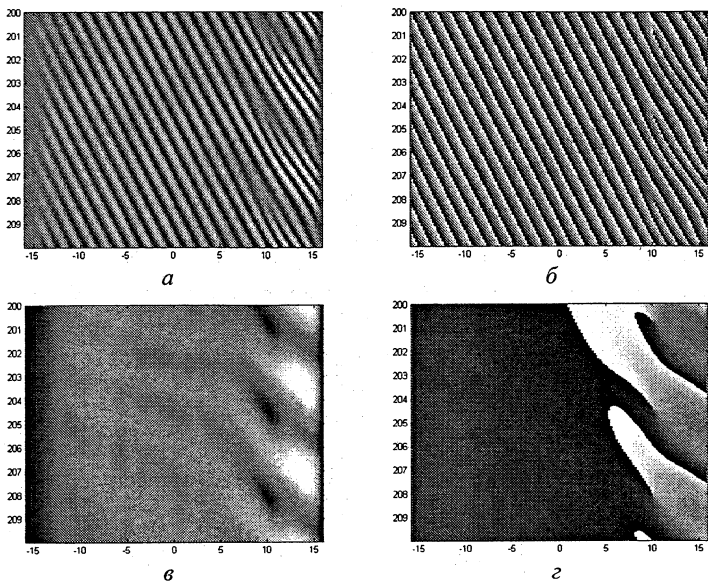


Рис. 9. Пространственно-временная диаграмма поля $\text{Re } u(x,t)$ (а), фаза поля $\text{arg } u(x,t)$ (б), амплитуда поля $|A(x,t)|$ (в) и фаза комплексной амплитуды после демодуляции (г). Периодическая последовательность дефектов наблюдалась при $c_1 = 0,5$; $c_3 = -5,0$, $V = 1,64$, начальных условиях $A(x,0) = A_0(1 + i)\text{sech}(\gamma x)$, $A_0 = 0,01$ и $\gamma = 100$.

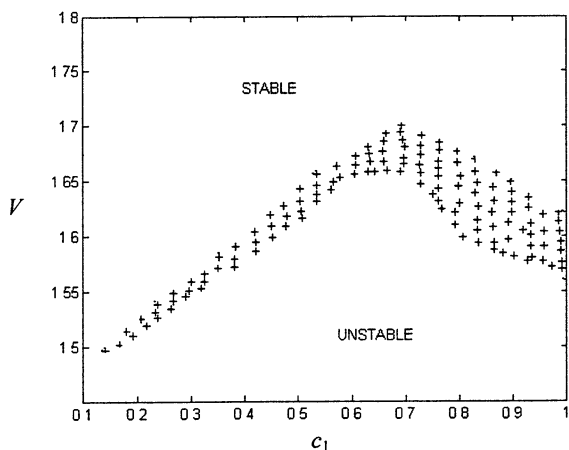


Рис. 10. Плоскость параметров (V, c_1) с различными режимами: $c_3 = -5,0$, + соответствует параметрам, при которых наблюдается периодическая генерация дефектов, выше этой зоны наблюдаются стационарные режимы с пространственно-неоднородным распределением амплитуды, ниже — движение дефектов не периодическое во времени.

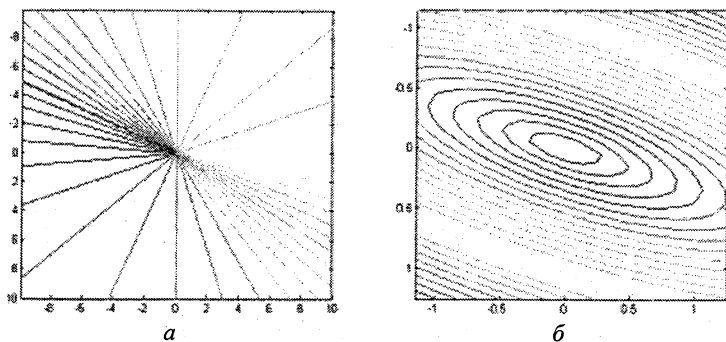


Рис. 11. Линии уровня фазы (а) и амплитуды (б) в окрестности дефекта при $V = 3$.

Можно проверить, что (5) — это главный член разложения решения в окрестности дефекта. Для этого представим комплексную амплитуду $A(x, t)$ в виде степенного ряда:

$$A = ix + (x - Vt) + \beta_2 x^2 + \beta_3 x^3 \dots + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \dots \\ + \gamma_{11} xt + \gamma_{12} xt^2 + \gamma_{21} x^2 t + \dots$$

После подстановки решения в уравнение (4), собирая слагаемые с одинаковыми степенями t и x , можно получить связь коэффициентов разложения с параметрами уравнения. Так, для нулевой степени t и x : $iV_g + \alpha_1 = 2\beta_2(1 + ic_1)$; для первой степени по t : $2\alpha_2 + V_g \gamma_{11} = \alpha_1 - V_g + 2(1 + ic_1)\gamma_{21}$ и т. д. Амплитуда и фаза полей, рассчитанные по такой простой модели, аналогичны характеристикам полей, полученным в физическом эксперименте и численном счете: линии постоянной фазы концентрируются вдоль прямой $x = Vt$, а эллипсы (линии постоянной амплитуды поля на пространственно-временной диаграмме) имеют наклон, равный групповой скорости волн.

На примере нескольких экспериментов, где возникают топологические особенности при обработке результатов по визуализации течений (источники, стоки, дырки, дефекты), мы показали, что возможно построить простые модели пространственных структур. Основой таких моделей являются решения уравнений Гинзбурга — Ландау, обладающие определенными особенностями. Такие решения с особенностями представляют собой строительные блоки, с помощью которых можно конструировать пространственные структуры, обладающие сложной временной динамикой. На языке этих особенностей могут быть сформулированы правила, управляющие

динамикой систем — какие процессы возможны и какие нет. Изучение динамики особенностей может представлять достаточно естественный путь для сравнения экспериментальных данных с результатами расчетов.

Литература

1. *Езерский А. Б.* Пространственно-временная динамика солитонов в резонаторах // *Нелинейные волны*'2004 / Под ред. А.В. Гапонова—Грехова, В.И. Некоркина. Н. Новгород: ИПФ РАН, 2005. С. 23—37.
2. *Coullet P., Frisch T., Plaza F.* Sources and sinks of wave patterns // *Physica D.* 1993. V. 62. P. 75—79.
3. *van Hecke M., Storm C., van Saarloos W.* Sources, sinks and wave number selection in coupled CGL equations and experimental implication for counter-propagating wave systems // *Physica D.* 1999. V. 134. P. 1—47.
4. *Garnier N., Chiffaudel A.* Nonlinear transition to a global mode for traveling — wave instability in a finite box // *Phys. Rev. Lett.* 2001. V. 86, № 1. P. 75—78.
5. *Alvarez R., van Hecke M., van Saarloos W.* Sources and sinks separating domains of left- and right-traveling waves: experiment versus amplitude equation // *Phys. Rev. E.* 1997. V. 56. P. R1306—R1309.
6. *Habdas P., de Bruyn J. R.* Dynamics of defects and traveling waves in an interfacial pattern // *Physica D.* 2005. V. 200. P. 273—286.
7. *Pastu L., Westra M.-T., van de Water W.* Sources and sinks in 1D traveling waves // *Physica D.* 2003. V. 174. P. 71—83.
8. *Ezersky A. B., Crumeyrolle O., Latrache N., Mutabazi I.* Competition of spiral waves with anomalous dispersion in Couette-Taylor flow // *Theor. And Comp. Fluid Dynamics.* 2004. V. 18, № 2—4. P. 85—95.
9. *Nozaki K., Bekki N.* Exact solution of the generalized Ginzburg — Landau equation // *J. Phys. Soc. Japan.* 1984. V. 53, № 5. P. 1581—1582.
10. *Bekki N., Nozaki K.* Formation of spatial patterns and holes in the generalized Ginzburg — Landau equation // *Phys. Lett. A.* 1985. V. 110, № 2. P. 133—135.
11. *Lega J., Janiaund B., Jucquois S., Croquette V.* Localized phase jumps in wave train // *Phys. Rev. A.* 1992. V. 45, № 8. P. 5596—5604.
12. *Burguete J., Chaté H., Daviaud F., Mukolobwiz N.* Bekki — Nozaki amplitude holes in hydrothermal nonlinear waves // *Phys. Rev. Lett.* 1999. V. 82. P. 3252—3255.
13. *Езерский А. Б.* Связанные состояния и пространственно-временной хаос топологических дефектов // *Нелинейные волны*'2002 / Под ред. А.В. Гапонова—Грехова и В.И. Некоркина. Н. Новгород: ИПФ РАН, 2003. С. 189—198.
14. *Bot P., Mutabazi I.* Dynamics of spatio-temporal defects in the Taylor — Dean system // *Eur. Phys. J. B.* 2000. V. 13. P. 141—155.
15. *Nana L., Achba N., Ezersky A. B., Mutabazi I.* Dynamics of spatio-temporal defects in the Couette — Taylor flow (подготовлено для публикации).
16. *Shraiman B. I., Pumir A., van Saarloos W., Hohenberg P. C., Chate H., Holen M.* Spatio-temporal chaos in one-dimensional complex Ginzburg — Landau equation // *Physica D.* 1992. V. 47. P. 241.
17. *Chate H.* Spatiotemporal intermittency regimes of the one-dimensional complex Ginzburg — Landau equation // *Nonlinearity.* 1994. V. 7. P. 185—204.
18. *van Hecke M., Howard M.* Ordered and self-disordered dynamics of holes and defects in the one-dimensional complex Ginzburg — Landau equation // *Phys. Rev. Lett.* 2001. V. 86. P. 2018—2021.

О РЕАЛИЗАЦИИ НЕКОТОРЫХ КЛАССИЧЕСКИХ МОДЕЛЕЙ И ФЕНОМЕНОВ НЕЛИНЕЙНОЙ ДИНАМИКИ НА ОСНОВЕ СВЯЗАННЫХ НЕАВТНОМНЫХ ОСЦИЛЛЯТОРОВ

С. П. Кузнецов

1. Введение

В науке о сложной динамике и хаосе выдвинуто множество интересных и важных идей, теорий, моделей, однако не всегда с ними соотнесены примеры реальных систем или, по крайней мере, привычные для физиков модели в виде дифференциальных уравнений. Сказанное относится, в том числе, к таким классическим объектам, как отображение Бернулли, гиперболические отображения на торе, странные гиперболические аттракторы, феномены комплексной аналитической динамики, ассоциирующиеся с множествами Мандельброта и Жюлиа [1—10]. Настоящая статья посвящена реализации абстрактных моделей и феноменов нелинейной динамики на базе систем, допускающих построение в виде устройств радиотехники и электроники. Это системы, составленные из двух или более связанных автогенераторов, которые благодаря периодической модуляции параметров по очереди выходят за порог генерации и передают друг другу возбуждение как бы по эстафете [11, 12].

2. Отображение Бернулли

Как известно, простейший пример хаоса дается отображением

$$\theta_{n+1} = 2\theta_n \pmod{1}. \quad (1)$$

На рис. 1 приводится график этого отображения и диаграмма, иллюстрирующая динамику на протяжении нескольких итераций.

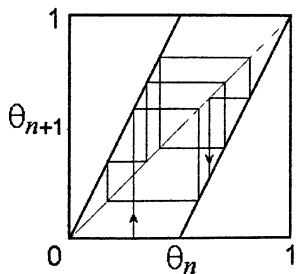


Рис. 1. График отображения Бернулли (1) и динамика на протяжении нескольких итераций.

Представим переменную θ в двоичной системе счисления. При этом цифра 0 на первой позиции после запятой отвечает расположению изображающей точки в левой, а 1 — в правой половине единичного интерва-

ла. Пусть начальное состояние задано случайной последовательностью цифр — нулей и единиц, полученной, скажем, подбрасыванием монеты. Один шаг эволюции в дискретном времени состоит в том, что последовательность сдвигается влево на одну позицию и цифра, оказавшаяся по левую сторону от запятой, отбрасывается:

$$\theta_0 = 0,01011010001010011001010\dots$$

$$\theta_1 = 0,1011010001010011001010\dots$$

$$\theta_2 = 0,011010001010011001010\dots$$

и так далее. Такое преобразование, состоящее в сдвиге последовательности символов на одну позицию, называют *сдвигом Бернулли*. При итерациях отображения (1) изображающая точка будет посещать левую или правую половину единичного интервала, следуя нашей случайной последовательности, — реализуется хаос. Ясно, что малое возмущение состояния увеличивается за один шаг в два раза. Поэтому показатель Ляпунова для данного отображения $\Lambda = \ln 2 = 0,693$.

Имея целью реализовать динамику, описываемую отображением Бернулли, интерпретируем θ как приведенную к единичному интервалу фазу некоторого колебательного процесса, т. е. $\theta = \varphi / 2\pi$, $\varphi \in [0, 2\pi)$. Чтобы описание сводилось к эволюции в дискретном времени, обратимся к неавтономной системе, в которой для фазы имеет место отображение нужного типа за период внешнего воздействия T .

Принципиальный момент состоит в том, чтобы обеспечить принадлежность отображения нужному топологическому классу: при вариации исходной фазы на 2π точка-образ должна совершить двукратный обход окружности. Такая возможность может вызвать сомнение: мы привыкли, что сдвиг фазы на 2π приводит к такому же сдвигу фазы отклика, что несовместимо с указанным топологическим свойством. Выход в том, чтобы реализовать *удвоение фазы, определенной в начале и в конце периода T , но не допускающей определения глобально на всем периоде*.

Рассмотрим систему на базе двух генераторов Ван-дер-Поля с частотами ω_0 и $2\omega_0$:

$$\begin{aligned} \ddot{x} - [A \cos(2\pi t/T) - x^2] \dot{x} + \omega_0^2 x &= \epsilon y \cos \omega_0 t, \\ \ddot{y} - [-A \cos(2\pi t/T) - y^2] \dot{y} + 4\omega_0^2 y &= \epsilon x^2, \end{aligned} \quad (2)$$

в которой обе подсистемы возбуждаются по очереди благодаря принудительному периодическому во времени изменению параметра, ответственного за бифуркацию Андронова — Хопфа. Присутствует также относительно слабое воздействие подсистем друг на друга, характеризующееся параметром ϵ . Предполагается, что $T = 2\pi N / \omega_0$, где N — целое число.

Пусть первый осциллятор имеет на стадии генерации некоторую фазу φ : $x \sim \cos(\omega_0 t + \varphi)$. Его воздействие на второй осциллятор через по-

средство члена εx^2 содержит вторую гармонику $\cos(2\omega_0 t + 2\varphi)$ с фазой 2φ . Это обеспечивает затравку для второго осциллятора, так что при его переходе в активную стадию колебания переменной y получают ту же фазу 2φ . В свою очередь, второй генератор действует на первый через элемент, осуществляющий смешение со вспомогательным опорным сигналом на частоте ω_0 . Порождаемая при этом составляющая на частоте ω_0 с фазой 2φ попадает в резонанс первому осциллятору и обеспечивает затравку для колебаний на очередной его активной стадии. Обе подсистемы по очереди передают возбуждение одна другой. На последовательных стадиях возбуждения первого осциллятора фаза колебаний в рассматриваемом приближении будет даваться отображением

$$\varphi_{n+1} = 2\varphi_n + \text{const} \pmod{2\pi}. \quad (3)$$

Константу, которая возникает из-за добавок к фазе при передаче возбуждения от одного осциллятора к другому, можно устранить сдвигом начала отсчета, и тогда для нормированной переменной $\theta = \varphi/2\pi$ получается в точности отображение (1).

На рис. 2 показаны полученные при численном решении уравнений временные зависимости переменных x и y при $\omega_0 = 2\pi$, $T = N = 10$, $A = 3$, $\varepsilon = 0,5$. Приводится также график эмпирического отображения для фазы, определенной в середине стадий возбуждения первого осциллятора $\theta = \arg(x - ix/\omega_0)/2\pi$. На графике для достаточно большого количества периодов T отложены точки (θ_n, θ_{n+1}) . Хотя вид полученного отображения несколько отличается от отображения (1), оно с очевидностью относится к нужному топологическому классу: график имеет две ветви, расположенные как на рис. 1. Соответствие с классическим отображением Бернулли становится лучше при увеличении отношения периодов N . Хаотический характер динамики выражается в случайном смещении максимумов функций $x(t)$ и $y(t)$ относительно огибающей (см. рис. 2).

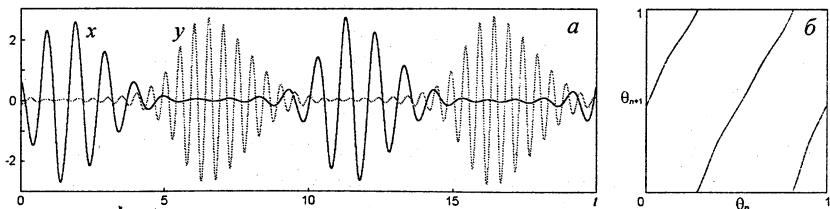


Рис. 2. Образец временной зависимости для переменных x и y , полученной при численном решении уравнений (2) для $\omega_0 = 2\pi$, $T = N = 10$, $A = 3$, $\varepsilon = 0,5$ (а), и диаграмма эмпирического отображения для фаз, определенных в середине стадий возбуждения первого осциллятора (б).

Для расчета показателя Ляпунова с помощью алгоритма Бенеттина [13] выполняется одновременно численное решение уравнений (2) и линеаризованных уравнений

$$\begin{aligned} \ddot{\bar{x}} + 2x\dot{\bar{x}} - [A \cos(2\pi t/T) - x^2] \dot{\bar{x}} + \omega_0^2 \bar{x} &= \varepsilon \bar{y} \cos \omega_0 t, \\ \ddot{\bar{y}} + 2y\dot{\bar{y}} - [-A \cos(2\pi t/T) - y^2] \dot{\bar{y}} + 4\omega_0^2 \bar{y} &= 2\varepsilon x\bar{x} \end{aligned} \quad (4)$$

с пересчетом нормы вектора возмущения к некоторой константе через определенное число шагов (для исключения расходимости). Показатель Ляпунова λ определяется как средняя скорость роста для накапливающейся суммы логарифмов от нормы вектора возмущения.

На рис. 3 представлен график зависимости нормированного показателя Ляпунова $\Lambda = \lambda T$ от амплитуды модуляции A при фиксированных остальных параметрах. В широком интервале значений A величина Λ почти постоянна и близка к $\ln 2 = 0,693$, что отвечает отображению Бернулли. При малых A соответствие нарушается — показатель Ляпунова становится заметно меньше.

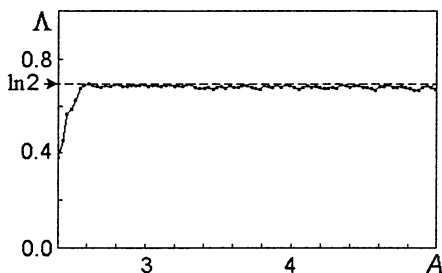


Рис. 3. Зависимость наибольшего показателя Ляпунова системы связанных неавтономных осцилляторов Ван-дер-Поля от параметра A при $\omega_0 = 2\pi$, $T = N = 10$, $\varepsilon = 0,5$.

3. Аттрактор Смейла — Вильямса

Математическая теория хаоса, базирующаяся на строгом аксиоматическом фундаменте, имеет дело с *однородно гиперболическими странными аттракторами* [1—5]. У такого аттрактора все принадлежащие ему фазовые траектории седловые. Устойчивые и неустойчивые многообразия (инвариантные множества, составленные из траекторий, приближающихся к исходной в прямом или обратном времени) могут пересекаться только трансверсально, т. е. без касания. Гиперболические аттракторы являются грубыми (структурно устойчивыми), что представляет очевидный интерес с точки зрения потенциальных приложений хаоса (системы коммуникации, маскировка сигналов и т. д.). К сожалению, аттракторы известных физических систем (простые генераторы хаоса, нелинейные осцилляторы с периодическим воздействием, модели Лоренца и Ресслера) не относятся к классу однородно гиперболических.

В учебниках и монографиях по нелинейной динамике примеры гиперболических аттракторов представлены абстрактными конструкциями.

Например, аттрактор Смейла — Вильямса строится для отображения трехмерного пространства в себя, определенного следующей процедурой. Рассмотрим область в форме тора, растянем ее в длину, сложим вдвое и вложим в исходный тор, как показано на рис. 4. При каждой итерации количество «витков» удваивается. Объект, который получается в пределе большого числа итераций, называют *соленоидом Смейла — Вильямса*. Его поперечная структура имеет вид канторова множества.

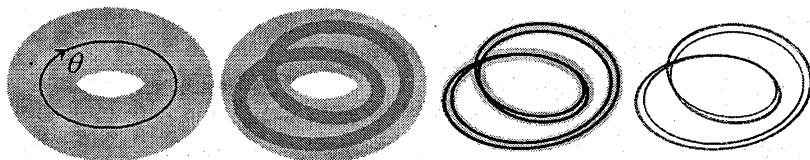


Рис. 4. Построение аттрактора Смейла — Вильямса

Основная проблема конструирования реальных систем с аттрактором Смейла — Вильямса та же самая, что и для отображения Бернулли. В самом деле, если ввести угловую координату θ , то она должна подчиняться отображению того же топологического класса, как отображение Бернулли, в то время как по оставшимся направлениям элемент фазового объема должен сжиматься.

Очевидно, подходящим кандидатом является система связанных неавтономных осцилляторов Ван-дер-Поля, рассмотренная в предыдущем разделе. Для аккуратного описания ее динамики в дискретном времени обратимся к построению отображения Пуанкаре [1—8].

Пусть в момент $t_n = nT$ имеем состояние системы (2), заданное четырьмя переменными — вектором $\mathbf{x}_n = \{x(nT), u(nT), y(nT), v(nT)\}$, где $u = \omega_0^{-1} \dot{x}$, $v = \omega_0^{-1} \dot{y}$. Из решения уравнений с начальными условиями \mathbf{x}_n на интервале времени T однозначным образом получаем новый вектор \mathbf{x}_{n+1} . Тем самым определяем функцию, отображающую четырехмерное пространство $\{x, u, y, v\}$ в себя:

$$\mathbf{x}_{n+1} = \mathbf{T}(\mathbf{x}_n). \quad (5)$$

С геометрической точки зрения, в пятимерном фазовом пространстве исходной неавтономной системы $\{x, u, y, v, t\}$ имеем сечение потока траекторий четырехмерными гиперплоскостями $t = t_n = nT$, где n — целое. В силу периодичности структуры фазового пространства по переменной t эти гиперплоскости естественно отождествить, и можно говорить об отображении гиперплоскости $\{x, u, y, v\}$ в себя.

Отображение Пуанкаре T определяется эволюцией, описываемой дифференциальными уравнениями с гладкими и ограниченными в конечной области переменных $\{x, u, y, v\}$ правыми частями. В соответствии с теоремами существования, единственности, непрерывности и дифференцируемости решения системы дифференциальных уравнений, отображение T представляет собой диффеоморфизм — взаимно однозначное и взаимно дифференцируемое отображение (класса C^∞) [14].

При выполнении отображения $x_{n+1} = T(x_n)$ в четырехмерном пространстве имеет место растяжение в направлении, которое связано с фигурирующей в уравнении (3) фазой, и сжатие по трем остальным направлениям. Интерпретируя действие отображения Пуанкаре геометрически, с учетом цикличности по фазовой переменной, мы можем вообразить вложенный в 4-мерное пространство тороид U (область, топологически представляющая собой прямое произведение одномерной окружности и трехмерного шара). Одна итерация отображения будет отвечать продольному растяжению и поперечному сжатию этого объекта, вкладываемого затем в виде двойной петли внутрь исходной области. Это точно соответствует конструкции Смейла — Вильямса.

Запишем соотношение, задающее область U , в виде

$$\{[(x_0^2 + x_1^2)^{1/2} - r]/d_r\}^2 + (x_2/d)^2 + (x_3/d)^2 \leq 1, \quad (6)$$

где введены новые переменные

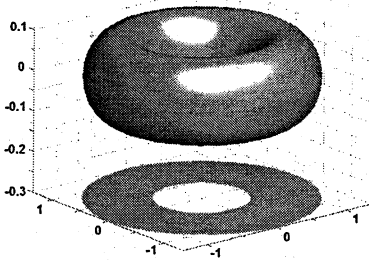
$$\begin{aligned} x_0 &= x/r_0, & x_1 &= (u - c_{ux}x)/r_1, \\ x_2 &= y - c_{yx}x - c_{yu}u, & x_3 &= v - c_{vx}x - c_{vu}u - c_{vy}y, \end{aligned} \quad (7)$$

а значения констант для конкретных параметров исходной системы $\omega_0 = 2\pi$, $T = 6$, $A = 5$, $\varepsilon = 0,5$ подобраны следующим образом:

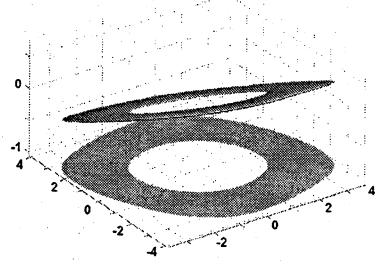
$$\begin{aligned} c_{ux} &= 0,438, & c_{yx} &= -0,042, & c_{yu} &= 0,226, & c_{vx} &= -0,218, \\ c_{vu} &= 0,029, & c_{vy} &= -0,118, & r_0 &= 0,812, & r_1 &= 0,721, \\ r &= 0,94, & d_r &= 0,4, & d &= 0,15. \end{aligned} \quad (8)$$

На рис. 5 приводится серия диаграмм, показывающая, как трансформируется область U в ходе эволюции фазового потока на протяжении времени T . Каждая картинка есть трехмерная проекция четырехмерного объекта в пространство $\{x_0, x_1, x_2\}$ вдоль оси четвертой переменной x_3 . Первая диаграмма (рис. 5, а) отвечает начальному моменту $t = 0$ и изображает исходную область U . Далее происходит сильное сжатие тороида по поперечным направлениям (рис. 5, б), а затем относительное растяжение в продольном направлении с такой деформацией, что образуется двойная петля (рис. 5, в, г). На последней стадии эта петля растягивается и в итоге оказывается вложенной в исходную область (рис. 5, д, е). Это,

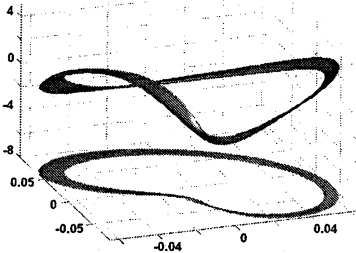
в сущности, воспроизводит ту последовательность непрерывных деформаций пластичного бублика, посредством которой дается популярное описание построения аттрактора Смейла — Вильямса в трехмерном пространстве.



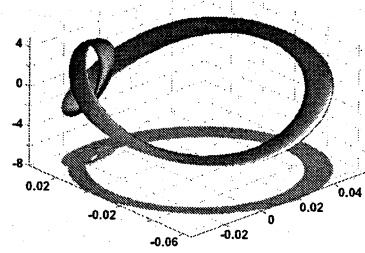
(a) $t = 0$



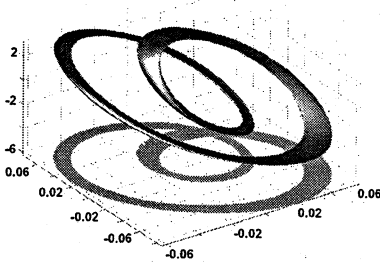
(б) $t/T = 0,1$



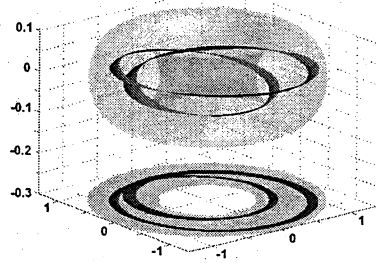
(с) $t/T = 0,6$



(д) $t/T = 0,65$



(е) $t/T = 0,69$



(е) $t/T = 1$

Рис. 5. Серия диаграмм, полученная в результате численного решения уравнений (2) и иллюстрирующая трансформацию образа области U в процессе эволюции за период воздействия T . Каждая картинка представляет собой трехмерную проекцию четырехмерного объекта в пространство переменных $\{x_0, x_1, x_2\}$ вдоль оси четвертой переменной x_3 . Обратите внимание на существенно различающийся масштаб по осям координат на диаграммах.

Таким образом, можно заключить, что внутри тороидальной области U присутствует аттрактор типа Смейла — Вильямса. Желательно, конечно, иметь математическое обоснование гиперболической природы аттрактора. В недавней работе [15] представлены результаты численных расчетов, показывающие, что в поглощающей области U выполнены условия расположения расширяющихся и сжимающихся конусов в пространстве векторов возмущения состояний системы, гарантирующие гиперболичность. Таким образом, вывод о гиперболичности аттрактора, хотя и не доказан пока в классическом математическом смысле, с определенностью вытекает из теоремы [1, 2, 16], условия которой проверены в расчетах на компьютере.

Проиллюстрируем численными расчетами некоторые атрибуты гиперболической динамики в исследуемой системе.

На рис. 6 приводятся диаграммы, изображающие аттрактор системы связанных неавтономных осцилляторов Ван-дер-Поля. На рис. 6, *a* показана проекция аттрактора из пятимерного расширенного фазового пространства на фазовую плоскость первого осциллятора в исходных переменных (x, u) . Аттрактор представлен тонами серого цвета (насыщенность тона определяется относительной частотой пребывания изображающей точки в соответствующем пикселе). Черным отмечены точки, отвечающие сечению Пуанкаре, т. е. моментам времени $t_n = nT$. Рисунок 6, *б* изображает аттрактор в сечении Пуанкаре на плоскости (x_0, x_1) и соответствует фрагменту диаграммы с рис. 6, *a* внутри пунктирного параллелограмма. Заметно очевидное визуальное сходство с аттрактором Смейла — Вильямса. Поперечная канторова структура аттрактора иллюстрируется отдельно на рис. 6, *в* и *г*, где показаны фрагменты исходной картинки с возрастающим увеличением.

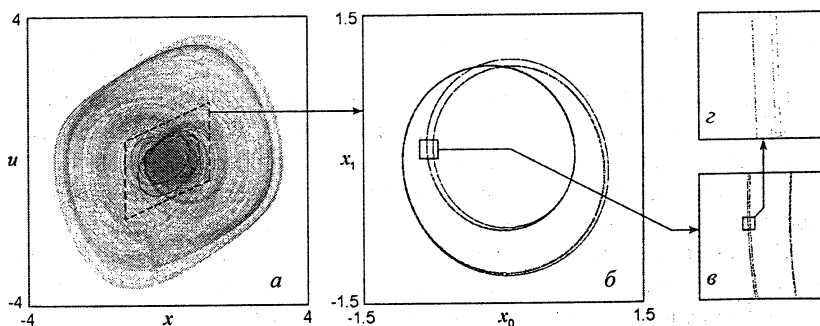


Рис. 6. Аттрактор системы связанных неавтономных осцилляторов Ван-дер-Поля в проекции из пятимерного расширенного фазового пространства на фазовую плоскость первого осциллятора в исходных переменных (*a*), портрет аттрактора в сечении Пуанкаре на плоскости (x_0, x_1) (*б*) и детали поперечной структуры аттрактора (*в*), (*г*).

Согласно результатам компьютерных расчетов¹, показатели Ляпунова для аттрактора отображения Пуанкаре при указанных значениях параметров составили

$$\begin{aligned} \Lambda_1 &= 0,6832 \pm 0,0007, & \Lambda_2 &= -2,602 \pm 0,004, \\ \Lambda_3 &= -4,605 \pm 0,003, & \Lambda_4 &= -6,538 \pm 0,008. \end{aligned} \quad (9)$$

Положительный показатель Λ_1 , наличие которого служит подтверждением присутствия хаоса, близок по величине к $\ln 2 = 0,693$, что естественно в свете возможного приближенного описания эволюции фазовой переменной одномерным отображением Бернулли (3).

Согласно оценке по алгоритму Грассбергера — Прокаччиа с использованием четырехкомпонентного временного ряда $x_n = \{x_0(t_n), x_1(t_n), x_2(t_n), x_3(t_n)\}$, $n = 1 \div M$, $M = 40000$, корреляционная размерность аттрактора в сечении Пуанкаре $D \approx 1,252$. Оценка размерности аттрактора через показатели Ляпунова по формуле Каплана — Йорке приводит к результату $D \approx 1,263$.

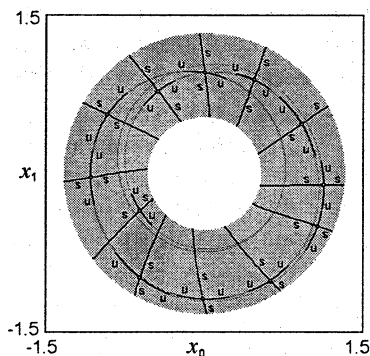


Рис. 7. Диаграмма на плоскости переменных первого осциллятора (x_0, x_1), иллюстрирующая расположение устойчивых (s) и неустойчивых (u) многообразий для некоторого набора точек на аттракторе в сечении Пуанкаре. Светло-серая кольцеобразная зона отвечает проекции поглощающей области U .

С точки зрения теоретического анализа гиперболических аттракторов, принципиальный момент состоит в том, что устойчивые и неустойчивые многообразия принадлежащих аттрактору точек обязаны пересекаться трансверсально. В силу заключения о гиперболической природе аттрактора, это свойство должно иметь место и в нашей системе.

Ситуация расположения устойчивых и неустойчивых многообразий, выявленная в численных расчетах, иллюстрируется на рис. 7 (детали методики вычислений приведены в работе [15]). Для графического пред-

¹ Процедура основана на одновременном численном решении уравнений (2) и четырех комплектов уравнений в вариациях (3). Через определенное число шагов по времени выполняется ортогонализация Грама — Шмидта и нормализация для четырех векторов возмущений. Показатели Ляпунова определяются как средние скорости роста или уменьшения накапливающихся сумм логарифмов норм для векторов возмущения перед нормализацией [13].

ставления используется плоскость переменных, относящихся к первому осциллятору. Неустойчивое многообразие одномерное и изображено в виде проекции на эту плоскость. Для представления устойчивого многообразия (трехмерной гиперповерхности) используем проекцию линии, образующейся от пересечения с двумерной поверхностью $\{x_2 = \text{const}, x_3 = \text{const}\}$, проходящей через точку, для которой строится многообразие. Инвариантное множество, служащее для точек аттрактора неустойчивым многообразием, совпадает с самим аттрактором, т. е. заключено внутри тороида в виде витков соленоида, обходя «дырку от бублика». С другой стороны, устойчивые многообразия располагаются поперек «трубки», образующей поверхность тороида, что в двумерной проекции выглядит как структура типа «спиц колеса». Благодаря такому размещению, пересечения устойчивых и неустойчивых многообразий трансверсальные и касания исключены.

4. Гиперболическое отображение на торе

Одним из традиционных примеров класса систем с хаотической динамикой в учебниках и обзорах математической направленности являются гиперболические автоморфизмы Аносова на торе [1—4]. Наиболее известное отображение этого класса

$$p_{n+1} = p_n + q_n \pmod{1}, \quad q_{n+1} = p_n + 2q_n \pmod{1} \quad (10)$$

часто называют отображением кота Арнольда, поскольку его действие иллюстрировалось в лекциях и книгах Арнольда с использованием изображения головы кота [17]. Графически удобно представлять фазовое пространство на плоскости в виде единичного квадрата с отождествленными противоположными сторонами.

Отображение (10) относится к классу консервативных систем: образ любой области на плоскости (p, q) (например, головы кота) имеет такую же площадь, как и исходная. Известно, что отображение (10) демонстрирует хаотическую динамику в смысле гиперболической теории Смейла и Аносова. Два показателя Ляпунова выражаются через собственные числа соответствующей матрицы 2×2 :

$$\Lambda_1 = \ln(3 + \sqrt{5})/2 = 0,9624, \quad \Lambda_2 = -\ln(3 + \sqrt{5})/2 = -0,9624. \quad (11)$$

Наибольший показатель положительный, что отражает присутствие экспоненциальной неустойчивости по отношению к малому возмущению начальных условий — один из главных атрибутов хаоса.

Полезно заметить, что отображение (10) можно представить как результат двукратного применения более простого отображения:

$$p_{n+1} = q_n \pmod{1}, \quad q_{n+1} = p_n + q_n \pmod{1}. \quad (12)$$

Для реализации динамики, описываемой данным отображением, возьмем четыре связанных неавтономных осциллятора Ван-дер-Поля [18]

$$\begin{aligned} \ddot{x} - [A \cos(2\pi t/T) - x^2] \dot{x} + \omega_0^2 x &= \varepsilon z \cos \omega_0 t, \\ \ddot{y} - [A \cos(2\pi t/T) - y^2] \dot{y} + \omega_0^2 y &= \varepsilon w, \\ \ddot{z} - [-A \cos(2\pi t/T) - z^2] \dot{z} + 4\omega_0^2 z &= \varepsilon xy, \\ \ddot{w} - [-A \cos(2\pi t/T) - w^2] \dot{w} + \omega_0^2 w &= \varepsilon x, \end{aligned} \quad (13)$$

три из которых (x , y , w) имеют равные частоты колебаний ω_0 , а четвертый — частоту $2\omega_0$. Параметр, отвечающий за бифуркацию Андронова — Хопфа в парциальных системах, медленно изменяется с периодом T , причем число $N = \omega_0 T / 2\pi$ предполагается целым. На одном полупериоде активной является первая пара осцилляторов (x , y), а вторая ниже порога генерации. На следующем полупериоде ситуация обратная.

Выведем соотношения, описывающие изменение фаз колебаний в ходе процесса. Предположим, что первый и второй осцилляторы на активной стадии имеют некоторые фазы φ_x и φ_y :

$$x \sim \cos(\omega_0 t + \varphi_x), \quad y \sim \cos(\omega_0 t + \varphi_y). \quad (14)$$

Отвечающий за связь член в третьем уравнении пропорционален

$$\cos(\omega_0 t + \varphi_x) \cos(\omega_0 t + \varphi_y) = \frac{1}{2} \cos(\varphi_x - \varphi_y) + \frac{1}{2} \cos(2\omega_0 t + \varphi_x + \varphi_y). \quad (15)$$

Составляющая на удвоенной частоте находится в резонансе с осциллятором z и при переходе его в активную стадию служит затравкой для колебаний, которые получают фазу $\varphi_z \approx \varphi_x + \varphi_y + \text{const}$. В то же время четвертый осциллятор просто наследует фазу первого: $\varphi_w \approx \varphi_x + \text{const}$. В начале следующей фазы активности для осцилляторов x и y член, задающий воздействие на первый осциллятор,

$$z \cos \omega_0 t \sim \frac{1}{2} \cos(3\omega_0 t + \varphi_x + \varphi_y) + \frac{1}{2} \cos(\omega_0 t + \varphi_x + \varphi_y) \quad (16)$$

обеспечивает затравку для осциллятора x , так что его колебания получают фазу $\varphi'_x \approx \varphi_z + \text{const} \approx \varphi_x + \varphi_y + \text{const}$. Осциллятор y наследует фазу осциллятора w : $\varphi'_y \approx \varphi_w + \text{const} \approx \varphi_x + \text{const}$. Таким образом, приходим к отображению для фаз

$$\varphi'_x = \varphi_x + \varphi_y + \text{const}, \quad \varphi'_y = \varphi_x + \text{const}. \quad (17)$$

При нормализации $q = \varphi_x / 2\pi$ и $p = \varphi_y / 2\pi$ получаем отображение (12) с точностью до постоянных добавок, которые можно устранить сдвигом начала координат. Следовательно, при выборе основного временного интер-

вала $2T$ стробоскопическое описание динамики для фаз осцилляторов (x, y) дается отображением кота Арнольда. Одно из следствий состоит в том, что динамика фаз, наблюдаемая в дискретном времени, будет демонстрировать хаос.

Обратимся к численным расчетам для системы неавтономных дифференциальных уравнений (13). На рис. 8 показаны графики зависимости переменных x, y, z и w от времени при значениях параметров $\omega_0 = 2\pi, N_0 = T = 20, A = 2, \varepsilon = 0,4$. Система функционирует в хаотическом режиме в соответствии с описанным механизмом. Хаос выражается в случайном смещении заполнения относительно огибающей на последовательных периодах модуляции.

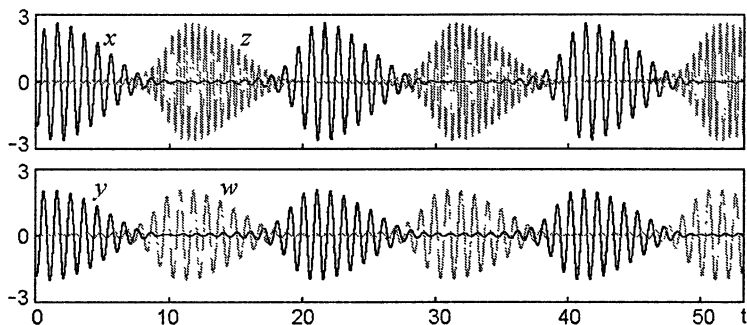


Рис. 8. Зависимость от времени динамических переменных в системе (13) при $\omega_0 = 2\pi, T = 20, A = 2, \varepsilon = 0,4$.

Чтобы продемонстрировать соответствие динамики отображению кота Арнольда, используем следующую процедуру. По ходу численного решения уравнений (13) будем определять фазы для первого и второго осцилляторов в середине активной стадии как $\varphi_x = \arg(x - i\dot{x}/\omega_0)$, $\varphi_y = \arg(y - i\dot{y}/\omega_0)$. Если точка $(q, p) = (\varphi_x/2\pi, \varphi_y/2\pi)$ попадает в область головы кота, нарисованной внутри единичного квадрата, то отмечаем ее на диаграмме. Также отмечаем точки-образы, получаемые через время $2T$ и $4T$ на двух последующих графиках. Накопив достаточно много точек, на первой картинке можно видеть изображение головы кота. На второй и третьей картинках получаются образы, соответствующие одной и двум итерациям отображения (рис. 9, а). Их можно сравнить с представленными на рис. 9, б диаграммами для отображения

$$p' = p + q + 0,07 \pmod{1}, \quad q' = p + 2q - 0,38 \pmod{1}, \quad (18)$$

где введены подобранные эмпирически постоянные добавки, учитывающие сдвиг фаз при передаче возбуждения между осцилляторами.

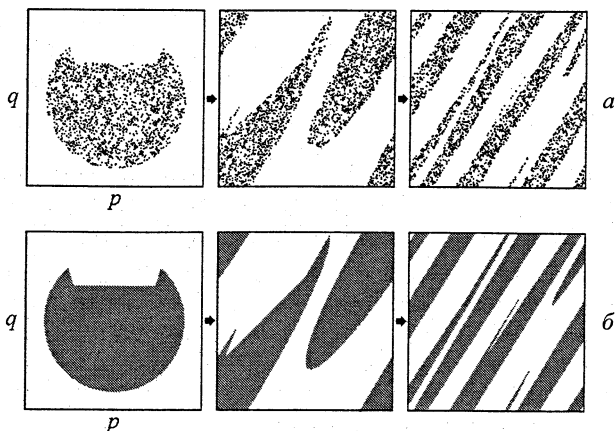


Рис. 9. Трансформация области в форме головы кота на плоскости фазовых переменных в системе связанных неавтономных осцилляторов Ван-дер-Поля (13) при эволюции за время $2T$ и $4T$ (а) и аналогичные картинки, полученные итерациями отображения (18) (б).

На рис. 10 приводится график зависимости двух наибольших показателей Ляпунова от амплитуды медленной модуляции A при фиксированных остальных параметрах, причем за единицу времени принят период $2T$. Как можно видеть, в широком интервале значений параметра оба показателя остаются приблизительно постоянными в хорошем соответствии с величинами, ожидаемыми для отображения кота Арнольда.

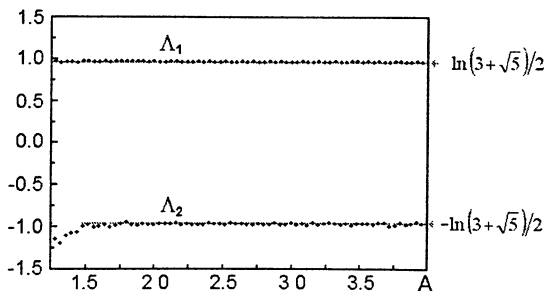


Рис. 10. Два старших показателя Ляпунова системы четырех связанных неавтономных осцилляторов Ван-дер-Поля (13).

При аккуратном описании динамики в дискретном времени задача сводится к восьмимерному отображению Пуанкаре. По двум направлениям, ассоциирующимся с фазовыми переменными φ_x и φ_y , имеет место динамика, описываемая в определенном приближении отображением кота Арнольда, а по остальным — сильное сжатие, характеризуемое большими по абсолютной величине отрицательными показателями Ляпунова. В целом система диссипативная, но аттрактор в восьмимерном фазовом пространстве отображения Пуанкаре располагается вблизи поверхности двумерного тора, динамика на котором хаотическая.

5. Феномены комплексной аналитической динамики: множества Мандельброта и Жюлиа

Один из глубоко проработанных математиками разделов нелинейной динамики содержит исследование итерируемых отображений, заданных аналитическими функциями комплексного переменного [9, 10]. Классическим объектом служит комплексное квадратичное отображение

$$z_{n+1} = c + z_n^2. \quad (19)$$

При $c = 0$ имеет место тривиальное поведение: для $|z_0| > 1$ результат итераций будет стремиться к бесконечности, а для $|z_0| < 1$ оставаться в ограниченной области. Границей раздела между этими двумя типами поведения в данном случае служит единичная окружность $|z_0| = 1$. При других значениях комплексного параметра c множество на комплексной плоскости z , разграничивающее указанные два типа поведения, устроено, вообще говоря, сложным образом и имеет фрактальную природу. Его называют *множеством Жюлиа* (см. примеры на рис. 11).

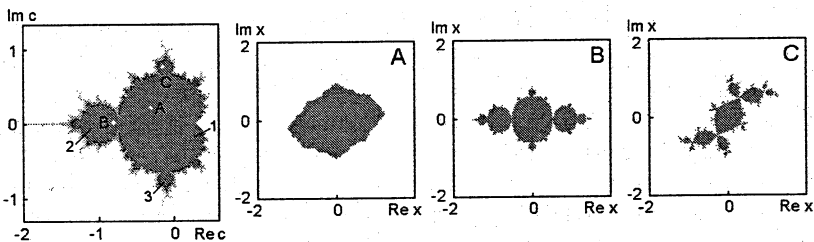


Рис. 11. Множество Мандельброта для квадратичного комплексного отображения (19) (слева) и множества Жюлиа при значениях параметра c , отмеченных точками на первой диаграмме. Белый цвет означает убегание на бесконечность.

С другой стороны, если фиксировать $z_0 = 0$ и рассмотреть поведение последовательности z_n в зависимости от комплексного параметра c , то при одних его значениях итерации будут убегать на бесконечность, а при других оставаться в ограниченной области. Ассоциирующееся с этой последней ситуацией множество точек на комплексной плоскости c называется *множеством Мандельброта* (рис. 11, слева). Область периодических режимов напоминает кактус и представляет собой совокупность касающихся друг друга образований округлой формы, расположенных по периферии большой области в виде кардиоиды. Цифры 1, 2, 3 обозначают периоды динамики, наблюдаемой в нескольких основных лепестках. Хаосу соответствует обрамляющий «кактус» фрактальный узор. Точками отмечены значения параметра, для которых построены изображения множеств Жюлиа.

Идея физической реализации феноменов комплексной динамики опирается на известный в теории колебаний *метод комплексных амплитуд*. Обратимся к следующей системе связанных неавтономных осцилляторов [19]:

$$\begin{aligned} \ddot{x} + F \cdot [g + \sin(2\pi t/T)]\dot{x} + \omega_0^2 x &= \varepsilon y \sin \omega_0 t + \lambda \sin(\omega_0 t + \varphi), \\ \ddot{y} + F \cdot [g - \sin(2\pi t/T)]\dot{y} + 4\omega_0^2 y &= \varepsilon x^2. \end{aligned} \quad (20)$$

Первый осциллятор имеет собственную частоту ω_0 , а второй вдвое большую. Параметр диссипации в обоих осцилляторах медленно изменяется во времени в противофазе с периодом T , составляющим целое число периодов колебаний на частоте ω_0 . Параметр g положительный, меньше единицы. Поэтому диссипация, будучи в среднем положительной, на определенной части периода вариации параметра для одного или другого осциллятора становится отрицательной. При этом осциллятор активный (колебания нарастают).

Предположим, что к моменту начала активной стадии второго осциллятора первый совершает колебания с комплексной амплитудой a : $x(t) \sim \text{Re}[a(t)\exp(i\omega_0 t)]$. Воздействие на второй осциллятор производится через нелинейный квадратичный элемент, так что результатом преобразования будет $x^2(t) = \frac{1}{2}|a(t)|^2 + \frac{1}{2}\text{Re}[a^2 \exp(2i\omega_0 t)]$. Затравкой для колебаний второго осциллятора будет попадающая в резонанс составляющая на второй гармонике, так что комплексная амплитуда его колебаний на активной стадии будет пропорциональна a^2 . Смещение с опорным сигналом (см. правую часть первого уравнения (20)) приводит к появлению составляющей на разностной частоте ω_0 , амплитуда которой пропорциональна a^2 . В сумме с дополнительным колебательным членом на частоте ω_0 , характеризуемым амплитудой λ и фазой φ , это дает затравку для комплексной амплитуды первого осциллятора вида a^2 плюс комплексная константа. Таким образом, стробоскопическое отображение для комплексной амплитуды будет соответствовать (по крайней мере, в определенном приближении) комплексному квадратичному отображению. При этом роль комплексного параметра s играет комплексное число с модулем λ и аргументом φ . В зависимости от этого параметра, решение уравнений связанных неавтономных осцилляторов может оставаться ограниченным при выборе начальных условий в какой-то области фазового пространства или же может оказаться уходящим на бесконечность.

На рис. 12 слева приводится полученная численно диаграмма на комплексной плоскости $\lambda e^{i\varphi}$, где серым цветом показаны зоны, отвечающие динамике в ограниченной области. Объект демонстрирует очевидное сходство с множеством Мандельброта для комплексного квадратичного отображения. Цифры 1, 2, 3 обозначают «лепестки», где реализуется динамика периода, соответственно, T , $2T$, $3T$. Для точек, помеченных

на рисунке «кактуса», построены диаграммы на плоскости переменных первого осциллятора (x, \dot{x}) , изображающие бассейны аттрактора, расположенного в ограниченной области. Их можно сопоставить с множествами Жюлиа для квадратичного отображения на рис. 11.

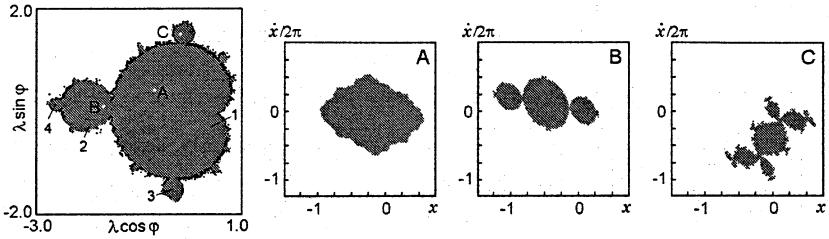


Рис. 12. Зона на плоскости параметра $\lambda e^{i\varphi}$, которой отвечает динамика системы (20) в ограниченной области (слева) и сечения бассейнов аттракторов, соответствующие невозбужденному состоянию второго осциллятора, при значениях комплексного параметра, отмеченных точками на первой диаграмме. Параметры системы: $\omega_0 = 2\pi$, $T = 10$, $F = 7$, $g = 0,5$, $\varepsilon = 1$.

Рисунок 13 иллюстрирует динамику в одном из режимов в лепестке периода 3. Показаны зависимости от времени для переменных x и y , отвечающих связанным неавтономным осцилляторам. Очевидно, тип режима определяется тем, каков будет на участке минимальной амплитуды первого осциллятора сигнал затравки, представляющий собой суперпозицию трансформированного на нелинейном элементе сигнала от осциллятора-партнера и внешнего сигнала. При этом существенны детали соотношения фаз этих двух сигналов, что и определяет тонкую структуру лепестков «кактуса» на плоскости параметров.

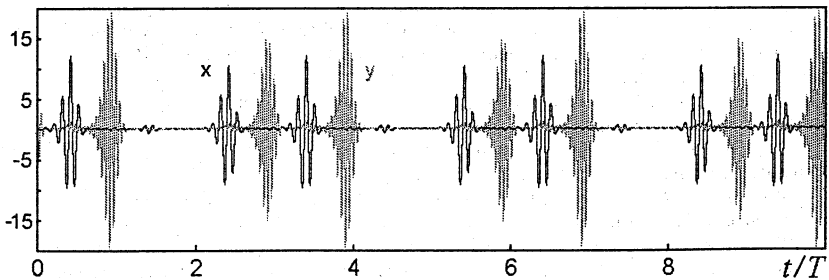


Рис. 13. Временная зависимость обобщенных координат связанных неавтономных осцилляторов (20), отвечающая циклу периода 3 отображения Пуанкаре; $\lambda e^{i\varphi} = -0,2 + 1,5i$, $\omega_0 = 2\pi$, $F = 7$, $g = 0,5$, $N = 10$, $\varepsilon = 1$.

Итак, предложенная система связанных неавтономных осцилляторов демонстрируют характерные для комплексной аналитической динамики

феномены: по крайней мере, в общих чертах воспроизводятся такие объ-екты как множества Мандельброта и Жюлиа¹.

Автор благодарен А. Ю. Жалнину, О. Б. Исаевой и И. Р. Сатаеву за участие в проведении численных расчетов, обсуждение и другую по-мощь. Работа выполнена при поддержке гранта РФФИ 06-02-16619.

Литература

1. *Синай Я. Г.* Стохастичность динамических систем // Нелинейные волны. М.: Нау-ка, 1979. С. 192—212.
2. *Каток А. Б., Хасселблат Б.* Введение в современную теорию динамических сис-тем. М.: Факториал, 1999. 768 с.
3. *Shilnikov L.* Mathematical Problems of Nonlinear Dynamics: A Tutorial // Int. J. of Bi-furcation and Chaos. 1997. V. 7, № 9. P. 1353—2001.
4. *Шильников Л. П., Шильников А. Л., Тураев Д. В., Чуа Л.* Методы качественной теори-и в нелинейной динамике. М.; Ижевск: 2003. 428 с.
5. *Belykh V., Belykh I., Mosekilde E.* The hyperbolic Plykin attractor can exist in neuron models // Int. J. of Bifurcation and Chaos. 2005. V. 15, № 11. P. 3567—3578.
6. *Анищенко В. С., Астахов В. В., Вадиасова Т. Е., Нейман А. Б., Стрелкова Г. И., Шиманский-Гайер Л.* Нелинейные эффекты в хаотических и стохастических системах. М.; Ижевск: 2003. 544 с.
7. *Шустер Г.* Детерминированный хаос. М.: Мир, 1988. 240 с.
8. *Кузнецов С. П.* Динамический хаос. М.: Физматлит, 2006. 356 с.
9. *Пайтген Х.-О., Рихтер П. Х.* Красота фракталов. Образцы комплексных динами-ческих систем. М.: Мир, 1993. 176 с.
10. *Devaney R. L.* An Introduction to Chaotic Dynamical Systems. N.Y.: Addison-Wesley, 1989 336 p.
11. *Kuznetsov S. P.* Example of a Physical System with a Hyperbolic Attractor of the Smale-Williams Type // Phys. Rev. Lett. 2005. V. 95. P. 144101.
12. *Кузнецов С. П., Селезнев Е. П.* Хаотическая динамика в физической системе со странным аттрактором типа Смейла — Вильямса // ЖЭТФ. 2006. Т. 129, № 2. С. 400—412.
13. *Benettin G., Galgani L., Giorgilli A., Strelcyn J.-M.* Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: A method for computing all of them // Meccanica. 1980. V. 15. P. 9—30.
14. *Арнольд В. И.* Обыкновенные дифференциальные уравнения. М.: Наука, 1975. 240 с.
15. *Кузнецов С. П., Сатаев И. Р.* Проверка условий гиперболичности хаотического ат-трактора в системе связанных неавтономных осцилляторов Ван-дер-Поля // Изв. вузов. При-кладная нелинейная динамика. 2006. Т. 15. С. 3—22.
16. *Hunt T. J.* Low Dimensional Dynamics: Bifurcations of Cantori and Realisations of Uni-form Hyperbolicity: PhD Thesis / Univ. of Cambridge, 2000.
17. *Арнольд В. И., Авец А.* Эргодические проблемы классической механики. Ижевск: Регулярная и хаотическая динамика, 1999. 284 с.
18. *Isaeva O. B., Jalnine A. Yu., Kuznetsov S. P.* Arnold's cat map dynamics in a system of coupled nonautonomous van der Pol oscillators // Phys. Rev. E. 2006. V. 74. P. 046207.
19. *Исаева О. Б., Кузнецов С. П., Осбалдестин А.* Феномены комплексной аналитиче-ской динамики в системе связанных неавтономных осцилляторов с эстафетной передачей возбуждения (в печати).

¹ Более тщательный анализ показывает, что соответствие не распространяется на тонкие детали структуры множества Мандельброта. Метод комплексных амплитуд является приближенным, и в следующих порядках теории возмущений отображение Пуанкаре приобретает поправки, не являющиеся комплексно аналитическими функциями. Эти добавки тем меньше, чем больше параметр N , характеризующий отношение периода модуляции пара-метра и периода собственных колебаний [19], так что с ростом N соответствие улучшается.

ФАЗОВЫЕ КЛАСТЕРЫ И КРАТКОСРОЧНАЯ ПАМЯТЬ В СЕТИ НЕЙРОНОВ С ПОСЛЕДЕПОЛЯРИЗАЦИЕЙ

В. В. Клиньшов, В. И. Некошкин

Введение

Краткосрочное хранение информации является типичным свойством нейронных систем, возникающим при решении многих задач на стадии первичной обработки информации. Системы, осуществляющие хранение такой временной информации, называются системами краткосрочной памяти. Существование краткосрочной памяти как важной составляющей системы памяти мозга было постулировано еще Хеббом в 1940-х годах [1]. Аткинсон и Шифрин в своей модели памяти [2] также отводят краткосрочной памяти важное место, определяя ее как память, в которой данные хранятся до тех пор, пока они используются. После прекращения непосредственного использования эта информация может быть либо безвозвратно утрачена (забыта), либо записана в так называемую долгосрочную память, формирование которой происходит с использованием принципиально иных нейронных систем и нейрофизиологических механизмов [3, 18].

Исследование краткосрочной памяти мозга является важной задачей нейродинамики, интерес к которой наблюдается с 1950-х годов и не ослабевает до сих пор. В начале 70-х годов прошлого века было установлено, что функции краткосрочной памяти связаны с определенными областями коры головного мозга [4–6]. В соответствии с функциональными свойствами различают вербальную, пространственно-визуальную и другие подсистемы краткосрочной памяти [7, 8], которые являются до некоторой степени независимыми [4]. Продолжительность хранения информации в краткосрочной памяти ограничена временем порядка $1 \div 100$ с [9, 18], а емкость системы соответствует запоминанию около 5–9 символов или «кусков» («chunks») информации одновременно [10, 11, 18].

Проблемы краткосрочной памяти привлекают внимание не только нейрофизиологов, но и физиков. В последние годы появился ряд физических моделей, описывающих феномен краткосрочной памяти. Это, например, модель хранения информации в одиночном нейроне с замкнутым на себя синапсом [12], модели на базе ансамблей нейроноподобных элементов [13–15], модель с использованием статистического механизма неокортикальных взаимодействий [16].

В системе краткосрочной памяти хранение информации связано с возникновением продолжительной нейронной активности после прихода короткого внешнего воздействия. Существуют два различных подхода к объяснению нейронного механизма данного феномена. Первый предполагает наличие замкнутых синаптических цепей, сохранение циркуляции возбуждения в которых соответствует хранению информации [12, 17]. Вторая гипотеза ассоциирует краткосрочную память с появлением в коре головного мозга структур периодической нейронной активности [18, 22]. Существует предположение, что достаточно простые сигналы могут храниться в виде циркулярного возбуждения в замкнутых цепях, тогда как более сложная информация (например, видеобразы) сохраняется в виде структур активности [3]. При запоминании информации такие структуры, или кластеры, появляются в коре головного мозга и существуют в течение некоторого времени, сохраняя информацию.

В цикле недавних работ Лисман и др. [18–22] предложили принципиальную схему системы краткосрочной памяти, основанную на особом свойстве некоторых нервных клеток — последеполяризации [18, 23]. Данное свойство состоит во временном повышении возбудимости мембраны нейрона после генерации им потенциала действия. Предложенная система памяти состоит из ансамбля нейронов с последеполяризацией, находящихся под воздействием осцилляторного сигнала и коротких информационных импульсов. Если до прихода информационного импульса на какой-либо элемент (нейрон) он совершал подпороговые колебания, то после воздействия данный элемент начинает периодически генерировать потенциал действия на каждом пике осцилляторного сигнала, и тем самым факт поступления импульса «запоминается» элементом. Однократно подав информационные сигналы на некоторую совокупность нейронов сети, можно запомнить в ней произвольный информационный образ в виде кластера периодической нейронной активности. Добавление в систему подавляющих обратных связей между нейронами приводит к появлению на пике осцилляторного сигнала нескольких «подциклов», на каждом из которых может возбуждаться определенная группа нейронов. Таким образом, в описанной системе могут сосуществовать несколько различных кластеров активности, периодически сменяющих друг друга. Это соответствует запоминанию нескольких образов одновременно.

В данной работе обсуждаются проблемы моделирования сети нейронов с последеполяризацией и межэлементными ингибиторными (подавляющими) обратными связями, осуществляющей функции краткосрочной памяти. Предложена модель нейрона с последеполяризацией,

адекватно описывающая данное свойство. Показано, что система таких нейронов с подавляющими обратными связями способна хранить в виде кластеров периодической активности несколько информационных образов одновременно. Такая возможность связана с наличием в достаточно широкой области параметров режима поочередного возбуждения различных групп элементов. Предложены способы записи, считывания и стирания информации из системы, оценены различные характеристики ее динамики в зависимости от параметров.

Динамика нейрона с последеполяризацией

Рисунок 1 иллюстрирует свойство последеполяризации. В начальный момент нейрон с последеполяризацией находится в покое и имеет нормальный уровень возбудимости. После прихода на нейрон короткого возбуждающего импульса он генерирует потенциал действия. Затем наступает короткий период рефрактерности, после которого мембрана нейрона деполяризуется, т. е. мембранный потенциал приобретает значение больше нормального. Благодаря этому нейрон может быть возбужден повторно с помощью воздействия меньшей амплитуды. Через некоторое время мембранный потенциал и возбудимость нейрона приобретают свои нормальные значения.



Рис. 1. Свойство последеполяризации нейрона: зависимость мембранного потенциала нейрона от времени после однократного возбуждения, полученная с помощью внутриклеточного микроэлектрода (из работы [18]).

Модель

Рассмотрим модель нейрона с последеполяризацией, введенную в [24]:

$$\varepsilon \frac{du}{dt} = v - f(u), \quad (1)$$

$$\frac{dv}{dt} = w - g(u) + x(t), \quad (2)$$

$$\frac{dw}{dt} = -\beta w + \gamma \left[p \left(u - \frac{1}{2} \right) - p(w - w_0) \right], \quad (3)$$

где u — мембранный потенциал нейрона, v описывает ионные токи через мембрану, w характеризует возбудимость мембраны, а $x(t)$ — внешний сигнал (стимул). В (1)–(3) функция $f(u)$ — кусочно-линейная функция следующего вида:

$$f(u) = \begin{cases} u, & u \leq \sigma, \\ \sigma - 2\sigma \frac{u - \sigma}{1 - 2\sigma}, & \sigma \leq u \leq 1 - \sigma, \\ u - 1, & u \geq 1 - \sigma, \end{cases}$$

а функция $g(u)$ определена как

$$g(u) = \begin{cases} u, & u \leq \frac{1}{2}, \\ \frac{1}{2} + \alpha \left(u - \frac{1}{2} \right), & u \geq \frac{1}{2}, \end{cases}$$

где α и σ — положительные параметры, определяющие форму этих функций. Функция $p(u)$ — пороговая функция вида

$$p(u) = \begin{cases} 0, & u \leq 0, \\ u/\kappa, & 0 \leq u \leq \kappa, \\ 1, & u \geq \kappa, \end{cases}$$

где параметр $\kappa \ll 1$ задает ее крутизну. Параметр $\varepsilon \ll 1$ характеризует электрическую емкость мембраны, параметры γ и β определяют скорость роста и спада возбудимости соответственно. Параметр w_0 задает максимальное значение возбудимости нейрона.

При $\beta = \gamma = 0$ имеем $w = \tilde{w} = \text{const}$, и уравнения (1)–(2) определяют известную систему ФитцХью — Нагумо [25, 26], в которой \tilde{w} играет роль параметра. Значение параметра \tilde{w} задает динамический режим, в котором находится система. При $\tilde{w} < \sigma$ нейрон находится в возбудимом состоянии, причем порог возбуждения тем ниже, чем больше значение \tilde{w} (рис. 2, а, б). При $\tilde{w} > \sigma$ нейрон периодически генерирует спайки, и частота их следования растет с увеличением \tilde{w} (рис. 2, в, г). Следовательно, \tilde{w} контролирует степень возбудимости нейрона.

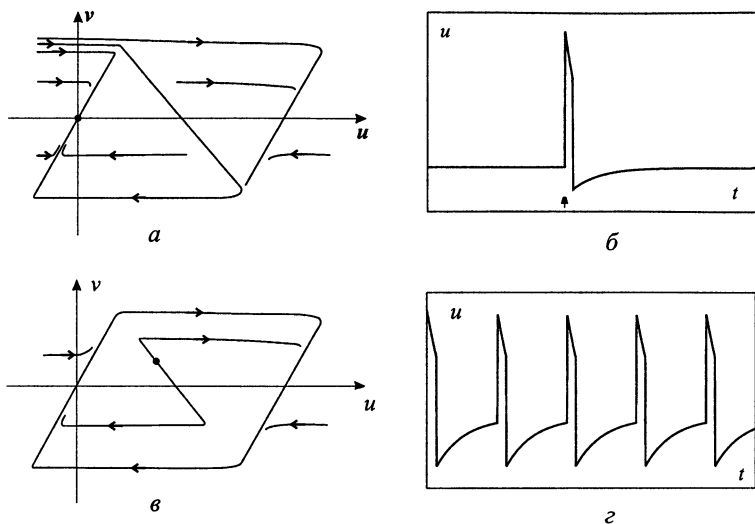


Рис. 2. Фазовый портрет и динамика системы ФитцХью — Нагумо для случаев возбудимого ($w < \sigma$) режима (а, б) и автоколебательного ($w > \sigma$) режима (в, г).

Пусть теперь $\beta > 0, \gamma > 0$. В этом случае возбудимость нейрона меняется со временем. Действительно, если нейрон не возбужден ($u < 0,5$), w со временем уменьшается и стремится к нулю со скоростью, задаваемой параметром β . В противном случае, если нейрон возбужден ($u > 0,5$), w возрастает со скоростью, определяемой величиной γ . Однако, когда w достигает значения w_0 , дальнейший рост возбудимости прекращается.

Динамические свойства элемента

Автономная динамика системы. Рассмотрим автономную динамику системы (1)–(3) при отсутствии внешнего сигнала: $x(t) \equiv 0$. Поскольку $\epsilon \ll 1$, в фазовом пространстве системы существуют так называемые быстрые и медленные движения [27, 28]. Медленные движения происходят в окрестности многообразия медленных движений, которое близко к поверхности $v = f(u)$. Быстрые движения происходят вне тонких (порядка ϵ) слоев, содержащих многообразие медленных движений. Рисунок 3 (вверху) иллюстрирует вид этого многообразия в фазовом пространстве системы (1)–(3). При этом по отношению к быстрым движениям устойчивыми являются «крайние»

части многообразия, т. е. полуплоскости $W_0 = \{u, v, w | u = v, v \leq \sigma\}$ и $W_1 = \{u, v, w | u = 1 + v, v \geq -\sigma\}$. Движение в окрестности многообразия W_0 соответствует динамике нейрона ниже порога возбуждения, а в окрестности W_1 — возбужденному состоянию нейрона. Исследование быстрых и медленных движений системы (1)–(3) показывает, что при $w_0 \leq \sigma$ все траектории системы асимптотически стремятся к единственному состоянию равновесия в начале координат $O(0, 0, 0)$, расположенному на многообразии W_0 .

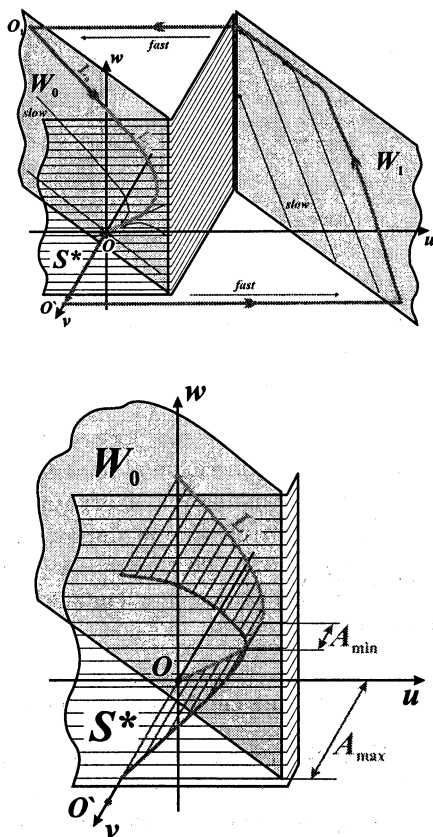


Рис. 3. Многообразия медленных движений системы (1)–(3): W_0 отвечает состоянию нейрона ниже порога возбуждения, W_1 соответствует возбужденному состоянию; поверхность S^* является пороговым множеством порог. Нижний рисунок — увеличенный фрагмент верхнего рисунка и траектория L_1 ; важно отметить, что данная траектория проходит ближе к поверхности S^* , чем точка O .

Отклик системы на короткий импульс. Рассмотрим теперь действие на систему короткого внешнего импульса: $x(t) = A\delta(t)$. Будем считать, что действие импульса сводится к мгновенному изменению координаты v на величину A . Пусть $u(0) = v(0) = w(0) = 0$. Тогда действие импульса перемещает изображающую точку из точки $O(0, 0, 0)$ в точку $O'(0, A, 0)$. Дальнейшая динамика системы зависит от взаимного расположения этой точки и порогового многообразия, представляющего собой плоскость $S^* = \{u, v, w | v = \sigma\}$. Если точка O' расположена по ту же сторону от S^* , что и O (при $A < \sigma$), изображающая точка возвращается в окрестность W_0 и стремится к состоянию равновесия O . Это означает, что возбуждения нейрона в данном случае не происходит. Совершенно по-другому ведет себя система, если O' попадает на противоположную сторону от S^* . В этом случае изображающая точка «перескакивает» в окрестность многообразия W_1 и только после движения около него в течение некоторого времени возвращается к W_0 (см. рис. 3). Данная ситуация соответствует достаточно большой амплитуде внешнего воздействия ($A > \sigma$). Следовательно, достаточно «сильный» импульс может возбудить нейрон, который ответит на возбуждение генерацией потенциала действия (спайка).

Обратим внимание на тот факт, что за время движения в окрестности W_1 переменная w возрастает и, если γ удовлетворяет неравенству $\gamma > \frac{w_0}{2\sigma}(1 - \sigma - w_0)$, это возрастание продолжается вплоть до значения, близкого к w_0 . Поэтому, когда изображающая точка возвращается в окрестность многообразия W_0 , она оказывается вблизи точки $O_1(-\sigma, -\sigma, w_0)$. Траектория L_0 , выходящая из данной точки, стремится к состоянию равновесия O . Обозначим через L_1 часть траектории L_0 , локализованную в области $v > 0$ (см. рис. 3, внизу). Важно то, что L_1 располагается в фазовом пространстве ближе к плоскости S^* , чем точка O . Поэтому, пока изображающая точка движется вдоль L_1 , порог возбудимости системы понижен по сравнению с нормальным состоянием. Это означает, что система обладает свойством последеполаризации (рис. 4, б). Продолжительность действия последеполаризации определяется временем движения изображающей точки вдоль L_1 , которое зависит от параметра β . Нетрудно видеть, что $\tau_{ADP} \sim \beta^{-1}$.

Отклик системы на гармоническое воздействие. Интересные свойства наблюдаются в системе (1)–(3) при ее стимуляции подпороговым периодическим (в частности, гармоническим) сигналом, например $x(t) = V \cos \frac{2\pi t}{T}$. Если изначально нейрон находится в покое, подпороговый сигнал не может его возбудить, и мембранный потенциал будет колебаться ниже порога возбуждения. Однако, если в какой-то момент времени на нейрон поступит дополнительный короткий импульс, то произойдет возбуждение нейрона, и его возбудимость w времен-

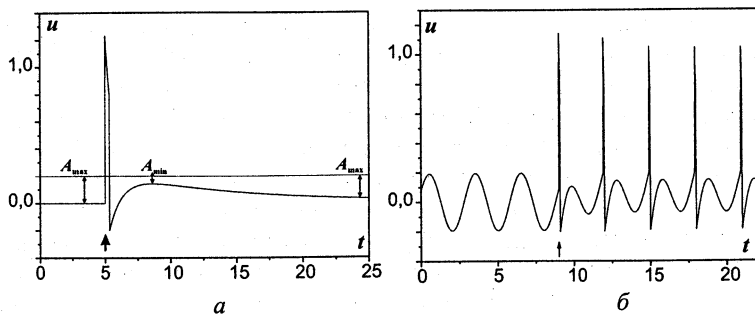


Рис. 4. Динамика системы (1)–(3) после воздействия коротким импульсом (ср. с рис. 1) (а) и динамика системы (1)–(4): краткое импульсное воздействие «переключает» нейрон из режима подпороговых колебаний в режим периодической генерации потенциала действия (б).

но увеличится из-за наличия последеполяризации. Из-за пониженного порога возбуждения сигнал $x(t)$ становится уже достаточным для того, чтобы возбудить нейрон во второй раз вблизи своего пика. Повышенный уровень возбудимости нейрона сохранится, и на следующем периоде внешнего сигнала нейрон снова возбудится и т. д. Следовательно, будет иметь место периодическая генерация спайков в ответ на однократное внешнее возбуждение.

Таким образом, система (1)–(3) при гармонической стимуляции может демонстрировать один из двух динамических режимов: либо режим подпороговых колебаний, либо режим периодической генерации спайков. Другими словами, система (1)–(3) обладает бистабильностью, т. е. в ее фазовом пространстве существует одновременно две устойчивые периодические траектории. Следовательно, под действием дополнительного внешнего импульса возможно «переключение» из одного колебательного режима в другой.

Хранение информации

Отдельный нейрон

Пусть сигнал $x(t)$ имеет вид суммы подпорогового гармонического сигнала и короткого возбуждающего (информационного) импульса

$$x(t) = V \cos \frac{2\pi t}{T} + A\delta(t - \tau), \quad (4)$$

где V — амплитуда гармонического сигнала, T — его период, A — амплитуда информационного импульса, а $\tau > 0$ — момент его поступления. Если в начальный момент нейрон находился в покое ($u = v =$

$= w = 0$ при $t = 0$), то в течение времени $0 < t < \tau$ под действием внешнего стимула он будет совершать подпороговые колебания. Как установлено выше, приходящий в момент $t = t_i$ короткий импульс «переключает» систему на другой режим, и при $t > t_i$ нейрон начинает периодически генерировать спайки (см. рис. 4, б).

Возникновение периодической спайковой активности в ответ на кратковременное воздействие означает динамическое сохранение информации о поступившем информационном импульсе. Другими словами, нейрон «запоминает» факт прихода импульса и тем самым сохраняет 1 бит информации.

Сеть несвязанных нейронов

Для сохранения большего количества данных можно использовать ансамбль нейронов с последеполяризацией. Очевидно, что N несвязанных нейронов способны хранить информацию о поступлении N информационных импульсов, т. е. N бит информации.

Например, представим оцифрованное черно-белое изображение, состоящее из N пикселей. Такое изображение может быть сохранено в сети из N элементов, каждый из которых «запоминает» цвет одного пикселя. Для записи изображения в память необходимо подать возбуждающие импульсы только на элементы, соответствующие белым пикселям, т. е. $A = A_0 > 0$ для «белых» элементов и $A = 0$ для «черных». После прихода импульсов в системе возникает кластер периодической нейронной активности, форма которого соответствует предъявленному изображению (рис. 5, а). Таким образом, изображение запоминается в памяти системы.

Стирание информации

Как известно, система памяти должна обеспечивать: (а) ввод информации, (б) хранение информации и (в) стирание информации. Рассмотрим, как эти функции реализуются в нашей системе. Как показано выше, *ввод* информации осуществляется путем подачи информационных импульсов, а *хранение* данных происходит в виде кластера периодической нейронной активности. *Стирание* информации может быть произведено одним из двух способов. Первый (активный) заключается в подаче импульсов отрицательной полярности на все элементы ансамбля. Второй способ (пассивный) связан с прекращением или уменьшением амплитуды гармонического сигнала. В обоих случаях все нейроны перейдут в режим подпороговых колебаний, и система будет готова к записи следующего образа. Если первый способ можно назвать стиранием в настоящем значении этого слова, то второй скорее можно определить как «забывание» информации. В соответствии

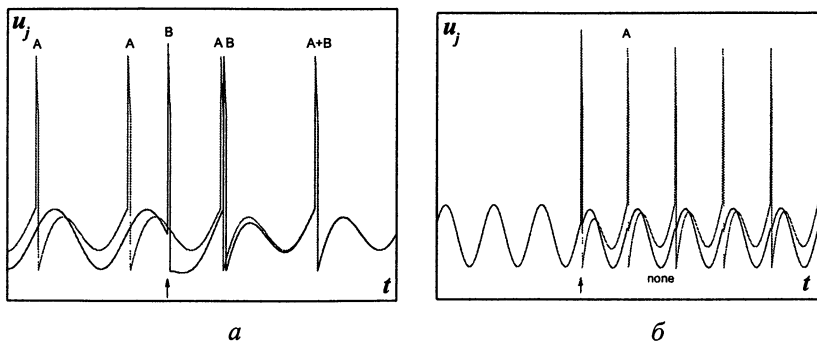


Рис. 5. Хранение информационного образа в ансамбле невзаимодействующих нейронов; в системе возникает кластер периодической нейронной активности, сохраняющий введенный образ (а). Безуспешная попытка записи в систему одновременно двух образов: возникает один кластер, соответствующий «смешению» обоих образов (б).

с этим наличие гармонического сигнала $x(t)$ можно трактовать как работу некоторого механизма «внимания», включающего систему КСП. Ослабление или потеря «внимания» приводит к утрате записанной информации.

Система с подавляющей обратной связью

Мы показали, что ансамбль несвязанных нейронов с последеполяризацией способен хранить информацию. Однако эта система обладает существенным недостатком — она не способна хранить несколько информационных образов (например, изображений) одновременно. С другой стороны, известно, что, например, краткосрочная память человека способна сохранять до 7 образов одновременно [18, 10]. При попытке последовательной записи нескольких образов в рассматриваемую систему в ней образуется *один* кластер, соответствующий *смешению* всех введенных образов (см. рис. 5, б), что приводит к потере информации. Данный недостаток может быть устранен введением глобальной ингибиторной обратной связи между элементами ансамбля. Мы покажем, что такая обратная связь обеспечит временное разделение нескольких образов и позволит осуществить их одновременное хранение (ср. [29]). Структурная схема модифицированной системы памяти представлена на рис. 6, а. Основной частью этой системы является сеть из N нейронов с последеполяризацией. На каждый из них поступает два сигнала: общий для всех гармонический и информаци-

онный. Кроме этого, в системе имеется петля обратной связи, которая активируется при возбуждении хотя бы одного нейрона из N и препятствует возбуждению остальных.

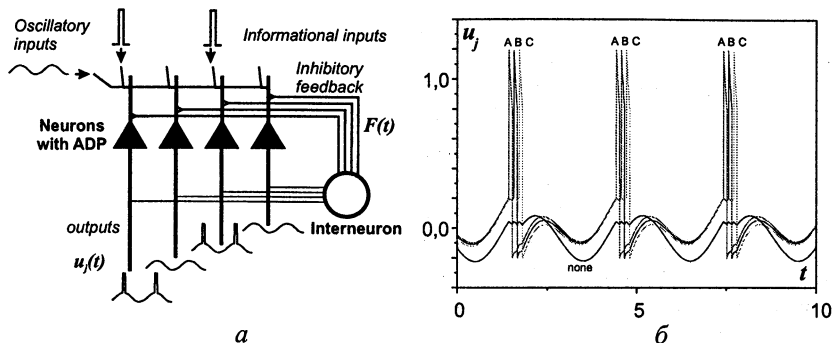


Рис. 6. Структурная схема системы (5)—(9) КСП, основанной на сети нейронов с ПДП (а). Хранение в системе *три* образов одновременно в виде трех последовательно возбуждающихся кластеров (б).

Динамика системы краткосрочной памяти с обратной связью описывается следующей математической моделью:

$$\varepsilon \frac{du_j}{dt} = v_j - f(u_j), \quad (5)$$

$$\frac{dv_j}{dt} = w_j - u_j + x_j(t) - \lambda F, \quad (6)$$

$$\frac{dw_j}{dt} = -\beta w_j + \gamma \left[p \left(u_j - \frac{1}{2} \right) - p(w_j - w_0) \right], \quad (7)$$

$$x_j(t) = V \cos \frac{2\pi t}{T} + A_j \delta(t - \tau_j), \quad (8)$$

$$F = p \left[\sum_{j=1}^N p \left(u_j - \frac{1}{2} \right) \right], \quad (9)$$

где $j = \overline{1, N}$ — порядковый номер элемента в сети; A_j и τ_j — амплитуда и момент поступления информационного импульса для j -го нейрона; F — подавляющий потенциал; λ — коэффициент, определяющий силу подавления, или величину обратной связи.

Введение подавляющей обратной связи существенно влияет на динамику нейронной сети. Теперь в ней может храниться одновременно несколько информационных образов в виде последовательно возбуждающихся кластеров периодической нейронной активности.

На рис. 6, б приведен пример подобной динамики. Все элементы сети ($N = 25$) разделились на четыре кластера, три из которых периодически возбуждаются в различных фазах гармонического сигнала, а четвертый колеблется ниже порога возбуждения. Три периодически возбуждающихся кластера соответствуют трем различным образам, одновременно хранящимся в памяти системы.

Исследуем динамический механизм формирования нескольких кластеров в системе (5)–(9). Для этого рассмотрим процесс последовательной записи нескольких образов в память системы. Изначально все элементы системы находятся в невозбужденном состоянии (рис. 7). При записи первого образа динамика системы будет такой же, как в случае системы несвязанных нейронов. Именно элементы, получившие информационные сигналы, начинают периодически генерировать потенциал действия и тем самым формируют кластер периодической активности, соответствующий записанному образу.

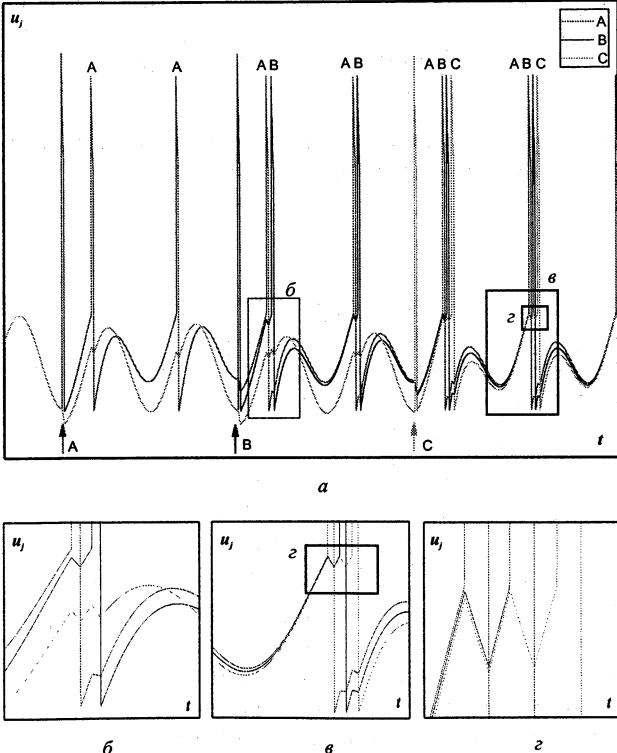


Рис. 7. Процесс последовательной записи образов в систему

Запись второго образа производится аналогично, но при выполнении одного условия — информационные сигналы должны поступить на элементы в момент, когда первый кластер не возбужден. Например, пусть это произойдет немного раньше момента возбуждения первого кластера. Тогда все элементы, соответствующие второму образу, также перейдут в режим периодического возбуждения. Заметим, что во время следующего периода гармонического сигнала второй кластер возбудится немного раньше первого, т. к. подавляющая обратная связь включится и будет препятствовать возбуждению первого кластера.

Дальнейшая динамика сети зависит от «силы» подавляющей обратной связи. Если она достаточно слаба, первый кластер будет возбужден, несмотря на ее действие, и межспайковый интервал между первым и вторым кластерами будет близок к нулю. Это значит, что два кластера объединятся в один, и два образа будут также смешаны в один. С другой стороны, при достаточно сильной обратной связи $\lambda > \lambda_0$ [30] первый кластер не возбудится до тех пор, пока не прекратится генерация потенциалов действия элементами второго кластера. Заметим, что межспайковый интервал в этом случае будет равен некоторому положительному значению τ_{is} , которое больше длительности спайка τ_{sp} .

Таким образом, при $\lambda > \lambda_0$ два кластера, отвечающие двум различным образам, возбуждаются периодически в различных фазах гармонического сигнала, и образы не смешиваются друг с другом. Записывая в память системы тем же способом последующие образы, можно получить несколько последовательно возбуждающихся кластеров, каждый из которых соответствует одному из образов. Это значит, что сеть нейронов с последеполаризацией и глобальной подавляющей обратной связью может хранить несколько образов одновременно.

Важным свойством рассматриваемой системы краткосрочной памяти является устойчивость существующих в ней кластеров. Это значит, что достаточно слабое возмущение начальных условий в системе не приводит к «развалу» одного кластера на несколько различных кластеров, возбуждающихся неодновременно. Данное свойство имеет место из-за некоторой инертности обратной связи, которая включает-ся не мгновенно после возбуждения какого-либо нейрона сети, а только после достижения его мембранным потенциалом значения $u_j = \frac{1}{2}$. Характерное время включения обратной связи $\tau_f \sim \epsilon$. Поэтому, если в результате некоторого возмущения запаздывание между элементами одного кластера все же не превышает τ_f , подавляющая обратная связь не сможет разбить этот кластер. Напротив, временная задержка между различными кластерами перед возбуждением должна быть больше τ_f , иначе они будут смешаны в один.

Следовательно, при обеспечении определенных условий на параметры системы, в ней могут существовать несколько устойчивых кластеров периодической нейронной активности. Все элементы в одном кластере колеблются полностью синхронно, а различные кластеры возбуждаются в разных фазах гармонического сигнала.

Емкость системы

Важнейшим свойством системы краткосрочной памяти является ее емкость C , т. е. максимальное число образов, которые могут одновременно храниться в памяти. Другими словами, это максимальное число кластеров, которые могут сосуществовать в системе, последовательно возбуждаясь на каждом периоде гармонического сигнала. Поэтому очевидно, что емкость напрямую зависит от средней временной задержки между моментами возбуждения соседних кластеров $\langle \tau_{is} \rangle$. Очевидно, что

$$C \sim \frac{T}{\langle \tau_{is} \rangle}.$$

Средняя задержка тем больше, чем сильнее подавляющая обратная связь. Следовательно, увеличение значения λ приводит к уменьшению емкости. С другой стороны, при $\lambda < \lambda_0$ все записанные в систему образы смешиваются в один и $C = 1$. Поэтому существует оптимальный интервал значений $\lambda \in [\lambda_0, \lambda_1]$, для которого емкость C максимальна. Зависимость $C(\lambda)$ приведена на рис. 8, а, ее график имеет ступенчатую форму с ярко выраженным максимумом при $\lambda \in [0,41-0,44]$.

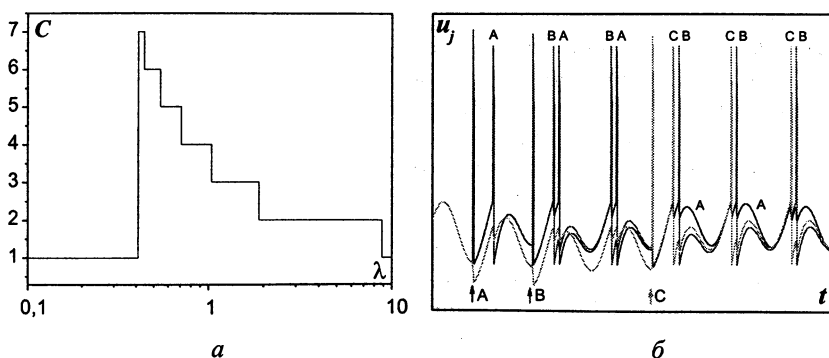


Рис. 8. Емкость C системы в зависимости от силы подавляющей обратной связи λ (а) и процесс «вытеснения» ранее записанного образа для случая $C = 2$ (б).

Рассмотрим процесс записи нового образа в систему с уже заполненной памятью, т. е. в случае, когда уже записано C образов. Все $(C + 1)$ не могут сохраниться в памяти одновременно, поэтому произойдет весьма интересный процесс «вытеснения» лишнего образа (см. рис. 8, б). Новый образ будет записан в память системы на первое место, т. е. соответствующий кластер будет возбуждаться первым, бывший первый кластер станет вторым, бывший второй — третьим и так далее. Последний C -й кластер будет «вытеснен» из памяти системы, и элементы, принадлежащие ему, перейдут в режим подпороговых колебаний.

Заключение

Предложена и исследована модель краткосрочной памяти в виде сети нейронов с последеполяризацией и подавляющей обратной связью. Наиболее интересным свойством модели является возможность одновременного хранения нескольких информационных образов. В этом случае в ней возникает несколько последовательно возбуждающихся кластеров периодической нейронной активности, соответствующих записанным образам. Проведено исследование зависимости емкости системы от ее параметров и показано существование оптимального значения коэффициента обратной связи, соответствующего максимальной емкости.

Полученные результаты подтверждают предположения Лисмана и др. [18—22] о том, что краткосрочная память имеет осцилляторную природу и осуществляется в ансамбле пирамидальных нейронов с последеполяризацией. Гипотеза Лисмана также подтверждается экспериментальным наблюдением нейронных θ -колебаний [21] и согласованного возбуждения интернейронов (обеспечивающих обратную связь) и пирамидальных клеток [31] во время работы КСП.

Работа выполнена при поддержке РФФИ (гранты 06-02-16137, 05-02-17441), гранта Президента РФ (НШ-7309.2006.2) и фонда «Династия».

Литература

1. *Hebb D. O.* The organization of behavior. A neuropsychological theory. N. Y.: Wiley, 1949.
2. *Atkinson R. C., Shiffrin R. M.* The psychology of learning and motivation: advances in research and theory / K. W. Spence (ed.). N. Y.: Academic, 1968.
3. *Хакен Г., Хакен-Крелль М.* Тайны восприятия. М.: Институт компьютерных исследований, 2002.

4. *Ruchkin D. S. et al.* Modality-specific processing streams in verbal working memory: evidence from spatio-temporal patterns of brain activity // *Cognitive Brain Res.* 1997. V. 6. P. 95–113.
5. *Fuster J.* *The Prefrontal Cortex* (2nd edn). Raven Press, 1988.
6. *Goldman-Rakic P. S.* *Handbook of physiology.* APS, 1987.
7. *Baddeley A. D.* The fractionation of working memory // *PNAS.* 1996. V. 93. P. 13468–13472.
8. *Baddeley A. D.* *Working Memory.* Oxford, UK: Clarendon Press, 1986. 289 p.
9. *Peterson L. R., Peterson M. J.* Short-term retention of individual verbal items // *J. Experimental Psychol.* 1959. V. 58. P. 193–198.
10. *Miller G. A.* The magical number seven, plus or minus two: Some limits on our capacity for processing information // *Psychol. Rev.* 1956. V. 63. P. 81–97.
11. *Cowan N.* The magical number 4 in short-term memory: A reconsideration of mental storage capacity // *Behavioral And Brain Sciences.* 2000. V. 24. P. 87–185.
12. *Seung H. S. et al.* The Autapse: A Simple Illustration of Short-Term Analog Memory Storage by Tuned Synaptic Feedback // *J. Computational Neurosci.* 2000. V. 9. P. 171–185.
13. *Stetter M.* Dynamic functional tuning of nonlinear cortical networks // *Phys. Rev. E.* 2006. V. 73. P. 031903.
14. *Batista A. M., Viana R. L., Lopes S. R.* Multiple short-term memories in coupled weakly nonlinear map lattices // *Phys. Rev. E.* 2000. V. 61. P. 5990.
15. *Tanaka S., Okada S.* Functional prefrontal cortical circuitry for visuo-spatial working memory formation: A computational model // *Neurocomputing.* 1999. V. 26–27. P. 891–899.
16. *Ingber L.* Statical mechanics of neocortical interactions: Constraints on 40-Hz models of short-term memory // *Phys. Rev. E.* 1995. V. 52. P. 4561–4563.
17. *Wang X.-J.* Synaptic reverberation underlying mnemonic persistent activity // *TRENDS in Neurosci.* 2001. V. 24. P. 455–463.
18. *Lisman J. E., Idiart M.* Storage of $7/pm2$ short-term memories in oscillatory subcycles // *Science.* 1995. V. 267. P. 1512–1515.
19. *Jensen O., Lisman J. E.* Novel lists of $7/pm2$ known items can be reliably stored in an oscillatory short-term memory network: Interaction with longterm memory // *Learning and Memory.* 1996. V. 3. P. 257–263.
20. *Jensen O., Lisman J. E.* An oscillatory short-term memory buffer model can account for data on the Sternberg task // *J. Neurosci.* 1998. V. 18. P. 10688–10699.
21. *Raghavachari S.* Gating of Human Theta Oscillations by a Working Memory Task // *J. Neurosci.* 2001. V. 21. P. 3175–3183.
22. *Jensen O., Lisman J. E.* Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer // *TRENDS in Neurosci.* 2005. V. 28. P. 67–72.

23. *Izhikevich E. M.* Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting. The M.I.T. Press, 2005.

24. *Клинъшов В. В., Некоржин В. И.* Модель нейрона с последеполяризацией и краткосрочная память // Изв. вузов. Радиофизика. 2005. Т. 48. С. 228–237.

25. *FitzHugh R.* Impulses and physiological states in theoretical models of nerve membrane // Biophys. J. 1961. V. 1. P. 445–466.

26. *Nagumo J. S. et al.* An active pulse transmission line simulating nerve axon // Proc. IRE. 1962. V. 50. P. 2061–2070.

27. *Arnold V. I., Afraimovich V. S., Il'yashenko Yu. S., Shilnikov L. P.* Bifurcation theory // Dynamical systems V. Bifurcation Theory and Catastrophe Theory / V. I. Arnold (ed.). N. Y.: Springer-Verlag, 2004.

28. *Fenichel N.* Geometric singular perturbation theory for ordinary differential equation // J. Diff. Eqns. 1979. V. 31. P. 53–98.

29. *Campbell S. R. et al.* Synchrony and Desynchrony in Integrate-and-Fire Oscillators // Neural Computation, 1999. V. 11. P. 1595.

30. *Клинъшов В. В., Некоржин В. И.* Динамика системы нейронов с последеполяризацией и ингибиторной обратной связью // Изв. вузов. Прикладная нелинейная динамика. 2004. Т. 12, № 6. С. 144–158.

31. *Rao S. G., Williams G. V., Goldman-Rakic P. S.* Isodirectional Tuning of Adjacent Interneurons and Pyramidal Cells During Working Memory: Evidence for Microcolumnar Organization in PFC // J. Neurophysiol. 1999. V. 81. P. 1903–1916.

НЕПРЕРЫВНЫЙ ВЕЙВЛЕТНЫЙ АНАЛИЗ И НЕКОТОРЫЕ ЕГО ПРИЛОЖЕНИЯ К ЗАДАЧАМ НЕЛИНЕЙНОЙ ДИНАМИКИ

А. А. Короновский, А. Е. Храмов

1. Введение

Вейвлетный анализ, возникший сравнительно недавно, как оказалось, является мощным средством анализа динамики систем различной природы и привлекает в последнее время все большее внимание исследователей, работающих в различных областях естествознания. Популярность вейвлетов (кроме чисто спекулятивной, связанной с попытками применить новый «модный» аппарат к своей задаче без каких-либо обоснований и получения действительно новых результатов) определяется большим количеством новых возможностей, которыми обладает вейвлетное преобразование и которые дают в руки исследователю мощный и удобный инструмент по анализу и обработке данных и сигналов, изучению сложных (в том числе, и пространственно-распределенных) систем, созданию новых численных методов и т. д. (см., например, обзорные работы и монографии [1–6], а также библиографию в указанных работах). Вот как пишет о вейвлетах один из ведущих математиков, работающих в области вейвлетного анализа, Ингрид Добеши [7]: «Вейвлеты являются сравнительно новым изобретением в прикладной математике. Это название само по себе возникло около десятилетия назад. . . За последние десять лет интерес к ним вырос взрывообразно. Их нынешний успех объясняется несколькими причинами. С одной стороны, концепция вейвлетов может рассматриваться как синтез идей, возникших за последние двадцать или тридцать лет в технике. . . физике. . . и чистой математике. Вследствие своего междисциплинарного происхождения вейвлеты представляются привлекательными для ученых и инженеров с самыми разными научными интересами. . . »

Заметим, что вопросы приложения вейвлетного анализа к исследованию нелинейных процессов, динамического хаоса и турбулентности в системах различной природы вызывают в настоящее время большой интерес (см., например, [8–19]). Учитывая, что читателями лекции будут в первую очередь специалисты, работающие в области нелинейной теории колебаний и волн, нелинейной динамики, мы постарались в качестве примеров приложений вейвлетного анализа выбрать такие, которые находятся в этой предметной области.

Перед тем как перейти к изложению основного материала, оговоримся, что в данной статье, в полном соответствии с ее названием, рассматривается только непрерывный вейвлетный анализ. Наряду с непрерывным вейвлетным анализом существует дискретное вейвлетное преобразование [7—11], на котором мы здесь не останавливаемся.

Непрерывный вейвлетный анализ: основные понятия и определения. Материнские вейвлеты

Непрерывное вейвлетное преобразование осуществляется путем свертки

$$W(s, t_0) = \int_{-\infty}^{+\infty} f(t) \psi_{s, t_0}^*(t) dt \quad (1)$$

анализируемой функции $f(t)$ с двухпараметрической вейвлетной функцией $\psi_{s, t_0}(t)$, которая получается из материнского вейвлета $\psi_0(t)$:

$$\psi_{s, t_0}(t) = \frac{1}{\sqrt{s}} \psi_0\left(\frac{t - t_0}{s}\right). \quad (2)$$

Параметр s , называемый масштабом вейвлетного преобразования ($s \in \mathbb{R}^+$), отвечает за ширину вейвлета, а $t_0 \in \mathbb{R}$ — параметр сдвига, определяющий положение вейвлета на оси t . Множитель $1/\sqrt{s}$ в соотношении (2) введен для того, чтобы все вейвлетные функции ψ_{s, t_0} имели постоянную (единичную) норму в пространстве $L^2(\mathbb{R})$: $\|\psi_{s, t_0}\|_{L^2} = \|\psi_0\|_{L^2} = 1$. Отметим, что выполнение последнего условия в силу теоремы Парсеваля приводит к выражению

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{\psi}(\omega) \hat{\psi}^*(\omega) d\omega = 1, \quad (3)$$

где $\hat{\psi}(\omega)$ — фурье-образ вейвлетной функции.

Из последних соотношений, в частности, следует, что непрерывное вейвлетное преобразование изометрически отображает пространство функций одной переменной в двумерное вейвлетное пространство:

$$W : L^2(\mathbb{R}) \rightarrow \mathcal{W}(\mathbb{R} \times \mathbb{R}^+),$$

и, значит, информация, содержащаяся в коэффициентах вейвлетного преобразования, является избыточной. Отсюда следует, например, что непрерывное вейвлетное преобразование случайного сигнала будет показывать наличие корреляции, которой нет в сигнале, но которая естественным образом присутствует в самом преобразовании. Это является недостатком вейвлетного преобразования, который необходимо учитывать при интерпретации вейвлетных спектров.

Материнский вейвлет ψ_0 может быть выбран достаточно произвольно, однако при этом он должен удовлетворять ряду условий.

Условие локализации. Базисная вейвлетная функция (материнский вейвлет) ψ_0 должна быть локализована как во временном, так и в частотном представлении. Для этого необходимо, чтобы ψ_0 была задана на конечном интервале и обладала достаточной регулярностью.

Условие допустимости. Материнский вейвлет должен быть выбран таким, чтобы его фурье-образ $\hat{\psi}_0(\omega)$ удовлетворял условию

$$C_\psi = 2\pi \int_{-\infty}^{+\infty} \frac{|\hat{\psi}_0(\omega)|^2}{\omega} d\omega < \infty. \quad (4)$$

Отметим, что для всех практических целей данное условие эквивалентно требованию *нулевого среднего*:

$$\int_{-\infty}^{+\infty} \psi_0(t) dt = 0 \quad \text{или} \quad \hat{\psi}_0(0) = 0. \quad (5)$$

Наконец, следует упомянуть еще *условие ограниченности*:

$$\int_{-\infty}^{+\infty} |\psi_0(t)|^2 dt < \infty. \quad (6)$$

В качестве оценки хорошей локализации и ограниченности могут служить соотношения $|\psi_0(t)| < 1/(1 + |t|^n)$ или $|\hat{\psi}_0(\omega)| < 1/(1 + |\omega - \omega_0|^n)$, где ω_0 — доминантная частота вейвлета, а величина параметра n должна быть как можно больше.

В том случае, если условие допустимости (4) выполняется, существует обратное вейвлетное преобразование:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{ds}{s^2 \sqrt{s}} \int_{-\infty}^{+\infty} \psi_0\left(\frac{t-t_0}{s}\right) W(s, t_0) dt_0. \quad (7)$$

По аналогии со спектром мощности фурье-преобразования $P(\omega) = |\hat{f}(\omega)|^2$ можно ввести в рассмотрение мгновенное

$$E(s, t_0) = |W(s, t_0)|^2 \quad (8)$$

и интегральное

$$\langle E(s) \rangle = \int |W(s, t_0)|^2 dt_0 \quad (9)$$

распределения энергии по масштабам вейвлетного преобразования.

Интегральное распределение энергии по масштабам вейвлетного спектра связано с фурье-спектром мощности $P(\omega)$ соотношением

$$\langle E(s) \rangle \sim s \int P(\omega) |\hat{\psi}_0(s\omega)|^2 d\omega, \quad (10)$$

из которого следует, что распределение энергии по масштабам $\langle E(s) \rangle$ представляет собой сглаженный спектр мощности фурье-преобразования, причем характер сглаживания определяется фурье-образом материнского вейвлета ψ_0 .

Обсудим кратко некоторые важные свойства вейвлетного преобразования, которые часто используются при анализе сигналов с помощью вейвлетов.

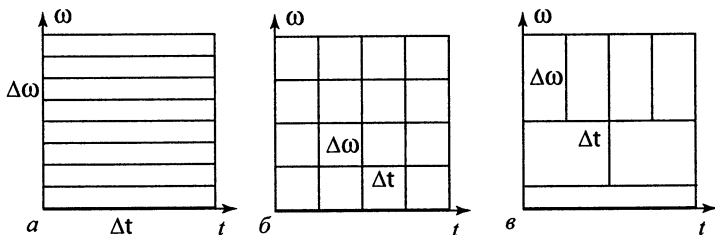


Рис. 1. Частотно-временная локализация в фазовом пространстве (t, ω) для различных преобразований: *a* — преобразование Фурье, *б* — оконное фурье-преобразование, *в* — вейвлетное преобразование.

Напомним, что для получения информации о высокочастотных гармониках исследуемого сигнала с высокой разрешающей способностью во времени нам необходимо извлекать ее из коротких временных интервалов, а не из всего сигнала, в то время как информацию о низкочастотной части спектра необходимо получать, анализируя достаточно длительные интервалы времени. На рис. 1 проведено сравнение возможности частотно-временной локализации вейвлетного и фурье-преобразований, для чего на нем показаны схемы разбиения пространства (t, ω) для выборки дискретных значений сигнала в случае различных преобразований. Рисунок 1, *a* соответствует преобразованию Фурье. Видно, что мы имеем хорошее разрешение по частоте и не имеем локализации во времени. Рисунок 1, *б* соответствует такой разновидности фурье-преобразования, как оконное преобразование Фурье. Из рисунка понятно, что в данном случае временное разрешение на больших и малых масштабах постоянно и определяется размером окна. В случае же вейвлетного преобразования (см. рис. 1, *в*) базисные функции имеют уменьшающееся с увеличением масштаба s временное

разрешение (определяемое шириной вейвлетной функции $\psi(t/s)$) и увеличивающееся с масштабом разрешение по частоте (определяемое шириной фурье-образа вейвлетной функции $\hat{\psi}(s\omega)$). Таким образом, вейвлетное преобразование характеризуется хорошей временной локализацией на малых масштабах и хорошим частотным разрешением на больших масштабах.

Вторым важным свойством вейвлетного преобразования, принципиальным для анализа сложных нестационарных процессов, является способность сохранять локальность представления сигнала и, как следствие, локально реконструировать сигнал. Фактически имеет место связь между локальным поведением сигнала и локальным поведением коэффициентов его вейвлетного преобразования. Под локальностью понимается то, что для реконструкции части сигнала необходимо рассматривать коэффициенты вейвлетного преобразования, относящиеся только к некоторой подобласти вейвлетного пространства (рис. 2, а), так называемому углу влияния. Если базовая вейвлетная функция ψ_0 хорошо локализована в интервале ΔT для $s = 1$, то коэффициенты $|W|$ вейвлетного спектра, соответствующие моменту времени t'_0 , будут содержаться в конусе влияния, ограниченном прямыми $s = t'_0 - (t_0 \Delta T)/2$ и $s = t'_0 + (t_0 \Delta T)/2$. Заметим, что, в свою очередь, коэффициент $W(s, t_0)$ в точке (t'_0, s') зависит от значений ряда из определяемого тем же углом влияния временного диапазона около значения t'_0 (рис. 2, б). Диапазон этот тем больше, чем больше анализируемый масштаб s' , то есть высокочастотная (или, что то же самое, мелкомасштабная) информация определяется на основе малых по длительности отрезков сигнала, а низкочастотная — больших.

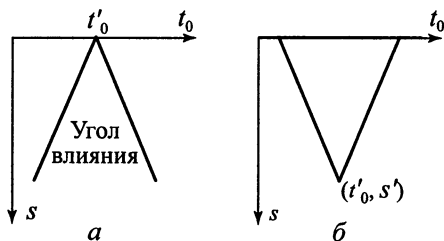


Рис. 2. Угол влияния вейвлетного преобразования.

Например, если функция $f(x)$ локально гладкая, то соответствующие коэффициенты вейвлетного преобразования остаются малыми. Если же $f(x)$ содержит сингулярность, тогда в окрестности этой особенности амплитуда вейвлет-коэффициентов существенно возрастает. Если вейвлет-коэффициенты подвержены случайным ошибкам, то они будут действовать на реконструируемый сигнал только локально вблизи возмущения, в то время как преобразование Фурье будет распространять ошибки по всему реконструируемому сигналу.

Таким образом, вейвлетное преобразование осуществляется с единственной материнской вейвлетной функцией, которая хорошо локализована и быстро стремится к нулю. С помощью нее покрывается вся ось $\mathbb{R}(-\infty, \infty)$ за счет системы сдвигов и масштабирований. Конкретный выбор материнского вейвлета определяется тем, какую информацию необходимо извлечь из сигнала. Каждая базовая вейвлетная функция ψ_0 характеризуется различными свойствами, что позволяет, используя разные вейвлетные функции, выявить все особенности анализируемого сигнала $f(t)$.

Одним из наиболее часто используемых комплексных материнских вейвлетов является морле-вейвлет¹:

$$\psi_0(\eta) = \pi^{-1/4} \exp(j\omega_0\eta) \exp(-\eta^2/2), \quad (11)$$

где ω_0 — параметр вейвлета. Морле-вейвлет обладает базисом, хорошо локализованным в реальном и фурье-пространствах, причем с увеличением ω_0 улучшается разрешение в фурье-пространстве, но ухудшается локализация во времени. Часто рассматривается морле-вейвлет с параметром $\omega_0 = 2\pi$. Выбор данного значения параметра ω_0 обеспечивает разумное соотношение разрешающей способности морле-вейвлета как во временном, так и в частотном пространствах. При этом для $\omega_0 = 2\pi$ имеет место «удобное» соотношение $s \simeq 1/f$ между временными масштабами s и частотами фурье-спектра f , в то время как для других значений ω_0 приходится пользоваться соотношением [5]

$$f = \left(\omega_0 + \sqrt{2 + \omega_0^2} \right) / 4\pi s. \quad (12)$$

Характеристики материнского морле-вейвлета приведены на рис. 3, а.

В качестве действительных базовых вейвлетных функций широко используется семейство DOG-вейвлетов, которые конструируются на базе производных функции Гаусса:

$$\psi_0(\eta) = (-1)^{m+1} \left[\Gamma\left(m + \frac{1}{2}\right) \right]^{-0,5} \frac{d^m}{d\eta^m} \exp\left(\frac{-\eta^2}{2}\right). \quad (13)$$

Материнский вейвлет, соответствующий $m = 1$, называется WAVE-вейвлетом; $m = 2$ — МНАТ-вейвлетом² (см. рис. 3, б).

Информацию о свойствах материнских вейвлетов, а также примеры других базовых вейвлетных функций можно найти в главе 1 работы [5].

¹ Назван по имени французского исследователя J. Morlet, предложившего впервые данный материнский вейвлет.

² От *Mexican Hat* — «мексиканская шляпа».

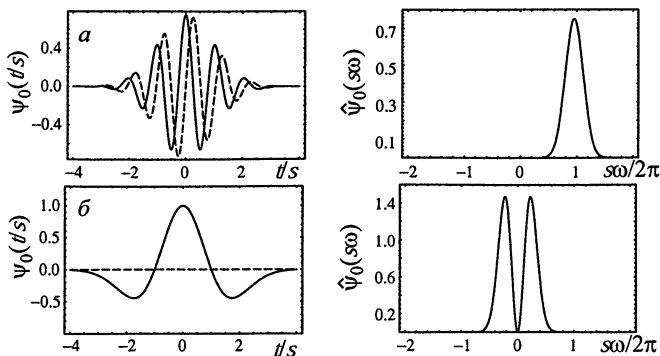


Рис. 3. Характеристики: *a* — морле-вейвлета ($\omega_0 = 2\pi$), *б* — МНАТ-вейвлета. Представлены действительные (сплошные линии) и мнимые (штриховые линии) части материнских вейвлетов (слева) и их фурье-образы (справа).

Приложения непрерывного вейвлетного анализа

Применение непрерывного вейвлетного преобразования для анализа перемежающегося поведения. Многие процессы в природе носят перемежающийся характер. В частности, в гидродинамике в ряде случаев при больших числах Рейнольдса можно наблюдать перемежающуюся структуру течения [20]. Перемежаемость является одним из классических сценариев перехода от периодических колебаний к хаотическим [20]; перемежающееся поведение наблюдается вблизи границы возникновения режимов хаотической синхронизации [21, 22], судорожная активность головного мозга животных и человека также несет на себе черты перемежающегося поведения [23, 24].

При исследовании перемежаемости важной проблемой является задача о выделении в сигналах временных интервалов, соответствующих различным типам динамики систем, демонстрирующих перемежающееся поведение. Как правило, участки временной реализации, на которых наблюдается поведение, близкое к регулярному (ламинарные фазы), выделяются легче, нежели участки с нерегулярным (турбулентным) поведением³. Однако часто характер движения во время различных фаз весьма сложен и разработать эффективные алгоритмы выделения этих фаз оказывается затруднительно. Используя свой-

³Отметим, что для различных типов перемежающегося поведения под «ламинарными» и «турбулентными» фазами понимаются различные типы поведения.

ства вейвлетного преобразования, возможно предложить универсальный метод, позволяющий проводить эффективный анализ временных рядов различных систем, демонстрирующих явление перемежаемости.

Поскольку поведение системы во время турбулентных и ламинарных фаз различается, то и структура вейвлетного спектра $W(t, s)$ в области ламинарных и турбулентных фаз движения благодаря свойству локальности представления сигнала также будет существенно различна. Иными словами, энергия вейвлетного спектра $E(s, t)$ (8) будет распределена по характерным временным масштабам s , которые для разных фаз временной реализации $f(t)$ различны, причем доля энергии, приходящейся на эти характерные временные масштабы, также будет различаться. Таким образом, можно перейти от анализа структуры вейвлетного спектра $W(s, t)$ к анализу распределения энергии вейвлетного спектра по характерным временным масштабам. Для выделения ламинарных и турбулентных фаз в каждый момент времени t определяется интегральное значение энергии вейвлетного спектра $w(t)$, приходящейся на выбранный диапазон характерных временных масштабов $s \in S = (s_1; s_2)$:

$$w(t) = \int_S E(t, s) ds. \quad (14)$$

Диапазон характерных временных масштабов s , по которым будет находиться величина $w(t)$, определяется рассматриваемой системой и в каждом конкретном случае должен выбираться на основании мгновенных распределений энергии вейвлетного спектра для различных фаз перемежающегося поведения. В ряде случаев может сложиться ситуация, что необходимо рассматривать несколько диапазонов временных масштабов S_i , которые однозначно позволяли бы охарактеризовать поведение системы. В этом случае интегрирование (14) должно проводиться по объединению $S = \bigcup_i S_i$ соответствующих диапазонов временных масштабов.

В качестве модельного примера рассмотрим классический случай перемежаемости I типа в системе Лоренца (подробнее см. [20]). На рис. 4 показаны временная реализация $x(t)$ системы Лоренца в режиме перемежаемости, модуль вейвлетного преобразования $|W(t, s)|$ с базовым Морле-вейвлетом сигнала $x(t)$ и функция $w(t)$ (14), приходящаяся на диапазон характерных временных масштабов $s \in (0, 15, 2, 4)$. На рис. 4, б по оси абсцисс отложено время t , по оси ординат — временные масштабы s в логарифмической шкале; интенсивность градаций серого цвета пропорциональна абсолютной величине $|W(s, t)|$ коэффициентов вейвлетного спектра; сплошная линия ограничивает область влияния краевых эффектов, достоверные данные располагаются ниже

этой линии (подробнее см. [5]). На вейвлетном спектре четко выделяются структуры, соответствующие ламинарной и турбулентной фазам временного ряда. Структура вейвлетной поверхности, соответствующая ламинарной фазе движения, имеет профиль $W(t, s)|_{t=\text{const}}$ с двумя глобальными максимумами, соответствующими характерным временным масштабам $s_a \approx 0,4$ и $s_b \approx 1,1$ (им соответствуют две темные области на проекции $|W(t, s)|$, расположенные параллельно оси времени). Профиль поверхности $|W(t, s)|$ не изменяется с течением времени в пределах ламинарной фазы. С входом системы в турбулентную фазу вид спектра $|W(t, s)|$ сильно изменяется. Можно сказать, что в режиме хаотической динамики наблюдается «всплеск» разномасштабных колебательных явлений. Из рис. 4, б видно, что области вейвлетной поверхности, соответствующие турбулентным фазам, четко локализованы во времени. Как следствие этого, в пределах ламинарной фазы функция $w(t)$ практически постоянна, а в течение турбулентной фазы $w(t)$ испытывает резкий провал (см. рис. 4, в), что позволяет точно диагностировать различные фазы движения. Моменты времени, в которые величина $w(t)$ оказывается больше некоторого порогового значения $\Delta = 250$, считаются ламинарными фазами, а интервалы времени, для которых выполняется условие $w(t) < \Delta$, — турбулентными.

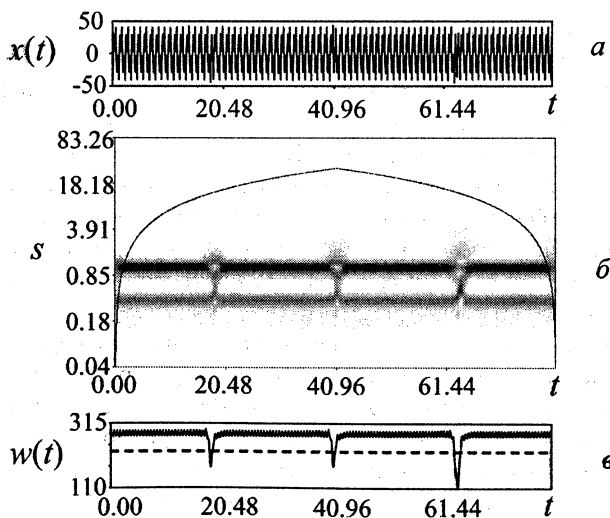


Рис. 4. Временная реализация $x(t)$ системы Лоренца в режиме перемежаемости (а); проекция вейвлетной поверхности $|W(t, s)|$ на плоскость (t, s) для временной реализации $x(t)$ (б); энергия вейвлетного спектра $w(t)$ (14) (в): пороговое значение Δ показано штриховой линией.

На рис. 5 представлена зависимость средней длительности ламинарной L и турбулентной T фаз от параметра надкритичности $(r - r^*)$. Благодаря логарифмическому масштабу видно, что средняя длительность ламинарной фазы зависит от параметра надкритичности как $L \sim (r - r^*)^{-1/2}$. Штриховой линией на рис. 5 показана степенная зависимость с показателем $-1/2$.

Средняя длительность турбулентной фазы практически не меняется с увеличением надкритичности. Это объясняется тем, что турбулентная фаза суть стадия реламинаризации, то есть возвращения к регулярному (ламинарному) движению, длительность которого определяется системой и практически не зависит от параметра надкритичности. Зависимость $L(r - r^*)$, полученная с помощью изложенного выше метода, хорошо согласуется с теорией перемежаемости I рода [20]. Следует отметить, что в литературе зависимости средней длительности турбулентных фаз от надкритичности, как правило, не анализируются, что связано как со сложностью получения аналитических результатов, так и с трудностями выделения турбулентных фаз из анализируемой временной реализации.

В качестве примера анализа перемежающегося поведения с помощью непрерывного вейвлетного преобразования в сложной реальной системе рассмотрим результаты⁴ анализа временной реализации записей электрической активности головного мозга крыс линии WAG/Rij, для которых судорожная активность представлена пик-волновыми разрядами, которые возникают спонтанно и не сопровождаются конвульсиями [25]. В эксперименте запись электрической активности мозга проводили с поверхности коры больших полушарий (электрокортикограмма — ЭКоГ) у интактных свободно-подвижных крыс посредством вживленных электродов (см. подробнее [26]). Длительность ЭКоГ записей составляла от 6 часов до 4 суток, поэтому возникала проблема автоматизации выделения различных паттернов на временном ряде ЭКоГ (пик-волновых разрядов, сонных веретён и т. д.). Типичная

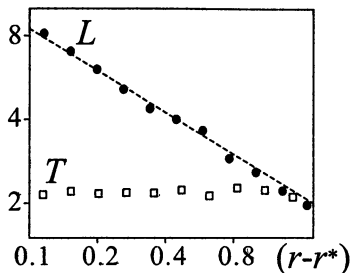


Рис. 5. Зависимости средней длительности ламинарных (●) и турбулентных (□) фаз от параметра надкритичности вблизи порога перемежаемости для системы Лоренца.

⁴ Данные исследования проводились совместно с проф. Г. Д. Кузнецовой, м.н.с. И. С. Мидзяновской и н.с. Е. Ю. Ситниковой (НИИ ВНДиН РАН, г. Москва); Dr. C.M. van Rijn (Radboud University Nijmegen), предоставившими экспериментальные данные.

ЭКоГ (ее фрагмент приведен на рис. 6, а), регистрируемая в ходе наблюдений, представляла собой чередование низкоамплитудных (от 25 до 100 мкВ) полиритмических участков, соответствующих фоновой активности головного мозга (будем называть такие участки ламинарными), и участков высокоамплитудной (100 ÷ 800 мкВ) генерализованной активности с относительно стабильной несущей частотой от 8 до 12 Гц, соответствующих пик-волновым разрядам (эпилептические припадки (турбулентные фазы) показаны серым цветом на рис. 6). Однако в течение ламинарных фаз на ЭКоГ наблюдаются паттерны, характеризующиеся высокоамплитудными колебаниями, но отличающиеся от пик-волновых разрядов (турбулентных фаз), — сонные веретёна (показаны рамками на рис. 6, а). Их амплитуда сравнима с амплитудой колебаний турбулентной фазы, что делает автоматизацию выделения эпилептических припадков из ЭКоГ по амплитудному критерию затруднительным. В то же самое время метод, основанный на непрерывном вейвлетном преобразовании, позволяет решить данную проблему достаточно просто, так как эти структуры ЭКоГ характеризуются различными распределениями энергии вейвлетного спектра по временным масштабам. Следует отметить, что анализ и выделение сонных веретён также могут являться самостоятельной задачей, однако здесь мы остановимся только на выделении эпилептических припадков с помощью непрерывного вейвлетного преобразования.

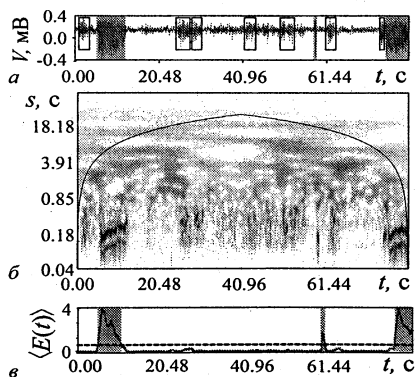


Рис. 6. Типичный фрагмент ЭКоГ, представляющий собой ламинарные фазы (нормальное поведение), перемежающиеся турбулентными всплесками (эпилептическая активность показана серыми прямоугольниками) (а); рамками выделены сонные веретёна, наблюдающиеся в течение ламинарных фаз. Проекция вейвлетного спектра $|W(s, t)|$, соответствующего ЭКоГ (б). Зависимость энергии $w(t)$, приходящейся на диапазон характерных временных масштабов (s_1, s_2) (в); величина порога $\Delta = 0,5$ показана штриховой линией.

Пик-волновые разряды были выделены экспертом на ЭКоГ на основании критериев, описанных в [25]. Из рис. 6, б видно, что фрагменты вейвлетного спектра ЭКоГ, соответствующие ламинарным и турбулентным участкам, оказываются принципиально различными. Зависимость от времени энергии $w(t)$ вейвлетного спектра, приходящейся на диапазон временных масштабов $s \in (s_1, s_2)$ ($s_1 = 70$ мс, $s_2 = 110$ мс), показана на рис. 6, в. При наступлении турбулентной фазы (пик-волнового разряда) величина $w(t)$ превышает порог $\Delta = 0,5$, что позволяет легко разделять различные типы поведения во временной реализации. Выбор диапазона временных масштабов (s_1, s_2) обусловлен структурой вейвлетной поверхности $W(s, t)$ во время ламинарных и турбулентных фаз. Отметим, что предложенный метод позволяет весьма эффективно разделить различные паттерны на ЭКоГ, в частности сонные веретена и пик-волновые разряды (ср. рис. 6, а и рис. 6, в).

Анализ распределений длительностей ламинарных фаз, выделенных из ЭКоГ с помощью вейвлетного преобразования, показал, что подобные распределения характеризуются степенными законами с показателем степени $-3/2$, что свидетельствует о присутствии поведения, характерного для on-off-перемежаемости [27]. Аналогичные результаты были получены при анализе судорожной активности данного типа для ЭКоГ всех экспериментальных животных (11 крыс), что позволило сделать вывод о том, что чередование судорожной и фоновой активности головного мозга крыс WAG/Rij является перемежаемостью on-off-типа и подчиняется универсальной закономерности, свойственной для данного типа перемежаемости (см. подробнее [24, 26]).

Анализ хаотической синхронизации с помощью непрерывного вейвлетного преобразования с комплексным базисом (синхронизация временных масштабов)⁵. Остановимся кратко на еще одном примере применения непрерывного вейвлетного преобразования в задачах нелинейной динамики, а именно, к проблеме хаотической синхронизации, где использование вейвлетов позволило предложить новый способ описания хаотической синхронизации, обобщающий все известные ранее типы синхронного поведения хаотических осцилляторов.

Рассмотрим вейвлетное преобразование (1) с комплексным материнским вейвлетом (например, вейвлетом Морле (11)) сигнала $x(t)$. Вейвлетный спектр $W(s, t) = |W(s, t)| \exp[j\phi(s, t)]$ (1) в этом случае характеризуется как амплитудой $|W|$, определяющей интенсивность

⁵ См. подробнее работы [16–19].

соответствующего масштаба s в момент времени t , так и фазой

$$\phi_s(t) = \phi(s, t) = \arg W(s, t), \quad (15)$$

которая естественным образом оказывается определенной для каждого временного масштаба s .

Рассмотрим теперь два связанных хаотических осциллятора $x_1(t)$ и $x_2(t)$ и осуществим вейвлетные преобразования сигналов каждого из осцилляторов. Если существует некоторый диапазон временных масштабов $[s_1, s_2]$, такой, что для любого временного масштаба $s \in [s_1, s_2]$ выполняется условие захвата фаз

$$|\phi_{s1}(t) - \phi_{s2}(t)| < \text{const} \quad (16)$$

и доля энергии вейвлетного спектра, приходящаяся на этот диапазон, отлична от нуля

$$\mathcal{E} = \int_{s_1}^{s_2} \langle E(s) \rangle ds > 0, \quad (17)$$

то временные масштабы $s \in [s_1, s_2]$ являются синхронизованными, а хаотические осцилляторы находятся в режиме синхронизации временных масштабов [17]. В соотношении (16) $\phi_{s1,2}(t)$ — непрерывные фазы (15) первого и второго осцилляторов, соответствующие рассматриваемому временному масштабу s .

Как было показано в работах [16—18], подход с позиций синхронизации временных масштабов, во-первых, позволяет уйти от сложностей введения мгновенных фаз хаотических осцилляторов в случае фазонекогерентного аттрактора [28]. Во-вторых, он может быть использован для анализа экспериментальных данных, поскольку он не требует каких-либо априорных знаний об изучаемой системе. Более того, в ряде случаев применение вейвлетного преобразования может уменьшить влияние шумов (см. [5]). Анализ с точки зрения синхронизации временных масштабов может быть эффективным при анализе живых систем (см., например, [19]). В-третьих, анализ динамики системы на всех временных масштабах, основанный на введении непрерывного множества фаз сигнала с помощью вейвлетного преобразования, позволяет рассматривать различные типы поведения связанных осцилляторов (синхронизация полная, фазовая, обобщенная, лаг-синхронизация, асинхронные колебания [28]) с единых позиций. Количество синхронизованных масштабов определяет тип поведения связанных осцилляторов, что позволяет ввести количественную характеристику меры хаотической синхронизации как относительную долю

энергии \mathcal{E} вейвлетного спектра, приходящуюся на синхронные временные масштабы $[s_1, s_2]$:

$$\gamma = \mathcal{E} / \int_0^\infty \langle E(s) \rangle ds. \quad (18)$$

Значение величины $\gamma = 0$ свидетельствует об отсутствии режима хаотической синхронизации связанных подсистем. Величина $\gamma \neq 0$ означает, что в системе реализуется режим хаотической синхронизации, означающий наличие синхронизированных масштабов. Значение $\gamma = 1$ показывает, что все временные масштабы синхронизированы. Рост γ от 0 до 1 свидетельствует об увеличении доли энергии, приходящейся на синхронные временные масштабы s .

На рис. 7 показана зависимость меры синхронизации (18) от интенсивности ϵ диффузионной связи, взятая из [16], для двух осцилляторов Рёсслера. В данной системе с ростом ϵ наблюдается последовательность переходов от асинхронной динамики к фазовой и далее полной хаотической синхронизации [28]. Видно, что с ростом ϵ доля энергии колебательного процесса, приходящаяся на синхронизованные масштабы, увеличивается. Для значения $\epsilon = 0,039$ (показано стрелкой 1), соответствующего возникновению фазовой синхронизации, мера синхронизации равна $\gamma = 0,21$. При больших интенсивностях связи ($\epsilon > 0,2$) мера синхронизации стремится к единице (отмечено на рис. 7 стрелкой 2). Последнее означает, что вся энергия хаотических автоколебаний приходится на синхронизованные временные масштабы и в системе возникает режим полной синхронизации.

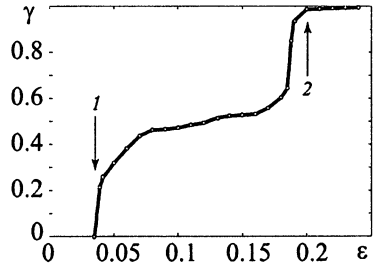


Рис. 7. Мера хаотической синхронизации, вводимая с помощью непрерывного вейвлетного преобразования, двух связанных систем Рёсслера [16].

Вместо заключения

К сожалению, объем лекции, прочитанной на Школе, не позволяет остановиться на многих вопросах приложения непрерывного вейвлетного преобразования к обработке различных сигналов, в том числе и в задачах, связанных с нелинейной динамикой. Поэтому вместо заключения позволим себе привести список тех проблем, которые не были освещены, но которые имеют несомненный интерес для специалистов,

занимающихся обработкой сигналов и работающих в области нелинейной физики (обсуждение некоторых из них можно найти уже в упоминавшихся ранее обзорных работах и монографиях [1–7]):

- эффективная численная реализация непрерывного вейвлетного преобразования;
- применение адаптивных вейвлетных базисов;
- бикогерентное вейвлетное преобразование и его приложения к анализу пространственно-временных структур;
- фильтрация сигналов с помощью вейвлетов;
- вейвлетный анализ случайных процессов.

В заключение авторы выражают искреннюю признательность чл.-корр. РАН, проф. Д. И. Трубецкову за обсуждение материала и рукописи статьи, а также благодарят проф. Г. Д. Кузнецову, м.н.с. И. С. Мидзяновскую, н.с., к.б.н. Е. Ю. Ситникову и Dr. С. М. van Rijп за предоставление экспериментальных данных. Работа выполнена при поддержке Программы поддержки ведущих научных школ РФ (НШ-4167.2006.2), РФФИ (проекты 05-02-16273 и 06-02-16451), CRDF (REC-006) и ФНП «Династия».

Литература

1. *Meyer Y. Wavelets: Algorithms and Applications.* Philadelphia: SIAM, 1993.
2. *Wavelets in Medicine and Biology / Ed. by A. Aldroubi, M. Unser. Boca Raton: CRC Press. FL., 1994.*
3. *Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения // УФН. 1996. Т. 166. С. 1145.*
4. *Ogden R. T. Essential wavelets for statistical applications and data analysis.* Boston; Berlin: Birkhäuser, 1997.
5. *Короновский А. А., Храмов А. Е. Непрерывный вейвлетный анализ и его приложения.* М.: Физматлит, 2003.
6. *Wavelets in Physics / Ed. by Van den Berg J.C. Cambridge, 2004.*
7. *Добеш И. Десять лекций по вейвлетам.* Ижевск: РХД, 2001.
8. *Argoul F. et al. Wavelet analysis of turbulence reveals the multifractal nature of the Richardson cascade // Nature. 1989. V. 338. P. 51.*
9. *Farge M. Wavelet transform and their applications to turbulence // Annu. Rev. Fluid. Mech. 1992. V. 24. P. 395.*
10. *van Milligen B. Ph. et al. Wavelet bicoherence: A new turbulence analysis tool // Phys. Plasmas. 1995. V. 2. P. 3017.*
11. *Дремин И. М., Иванов О. В., Нечитайло В. А. Вейвлеты и их применение // УФН. 2001. Т. 171. С. 465–501.*
12. *Ватрунин В. Е., Дубинов А. Е., Селемир В. Д. Самоподобие и ветвление вейвлет-образов СВЧ-сигналов в виркаторе // Письма в ЖТФ. 1996. Т. 22. С. 92.*

13. Quiroga R. Q., Kraskov A., Kreuz T., Grassberger P. Performance of different synchronization measures in real data: a case study on electroencephalographic signals // *Phys. Rev. E*. 2002. V. 65. P. 041903.

14. Короновский А. А., Храмов А. Е. Исследование когерентных структур в электронном пучке со сверхкритическим током с помощью вейвлетной бикогерентности // *Физика плазмы*. 2002. Т. 28. С. 722—738.

15. Lachaux J. P. et al. Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence // *Neurophysiol. Clin.* 2002. V. 32. P. 157—174.

16. Hramov A. E., Koronovskii A. A. An approach to chaotic synchronization // *Chaos*. 2004. V. 14. P. 603—610.

17. Hramov A. E., Koronovskii A. A. Time scale synchronization of chaotic oscillators // *Physica D*. 2005. V. 206, № 3—4. P. 252—264.

18. Храмов А. Е., Короновский А. А., Левин Ю. И. Синхронизация временных масштабов хаотических осцилляторов // *ЖЭТФ*. 2005. Т. 127. С. 886.

19. Hramov A. E., Koronovskii A. A., Ponomarenko V. I., Prokhorov M. D. Detecting synchronization of self-sustained oscillators by external driving with varying frequency // *Phys. Rev. E*. 2006. V. 73. P. 026208.

20. Берже П., Помо И., Вудаль К. Порядок в хаосе. М.: Мир, 1991.

21. Boccaletti S., Valladares D. L. Characterization of intermittent lag synchronization // *Phys. Rev. E*. 2000. V. 62. P. 7497—7500.

22. Hramov A. E., Koronovskii A. A. Intermittent generalized synchronization // *Eur. Lett.* 2005. V. 70. P. 169—175.

23. Perez Velazquez J. L. et al. Type III intermittency in human partial epilepsy // *Eur. J. Neurosci.* 1999. V. 11. P. 2571—2576.

24. Короновский А. А. и др. Закономерности перемежающегося поведения в спонтанной неконвульсивной судорожной активности у крыс // *ДАН*. 2006. Т. 409.

25. van Ligtelaar E. L., Coenen A. M. Two types of electrocortical paroxysms in an inbred strain of rats // *Neurosci. Lett.* 1986. V. 70. P. 393—397.

26. Hramov A. E., Koronovskii A. A., Midzyanovskaya I. S., Sitnikova E., van Rijn C. M. On-Off intermittency in time series of spontaneous paroxysmal activity in rats with genetic absence epilepsy // *Chaos* (to be published).

27. Platt N., Spiegel E. A., Tresser C. On—Off intermittency: a mechanism for bursting // *Phys. Rev. Lett.* 1993. V. 70. P. 279—282.

28. Пиковский А., Розенблюм М., Куртс Ю. Синхронизация. Фундаментальное нелинейное явление. М.: Техносфера, 2003.

СИСТЕМЫ СВЯЗИ С ХАОТИЧЕСКИМИ НОСИТЕЛЯМИ: ПРОБЛЕМЫ, ДОСТИЖЕНИЯ И ПЕРСПЕКТИВЫ

А. С. Дмитриев

Введение

Работы по использованию хаоса в системах связи проводились еще в 80-е годы прошлого века [1], интенсивные исследования в этом направлении стартовали в начале 90-х годов. Возникший интерес был во многом связан с открытием явлений хаотической синхронизации [2—4] и хаотического синхронного отклика [5].

Первые серьезные успехи были связаны с тем, что для ряда модельных схем была продемонстрирована возможность передачи цифровых и аналоговых сообщений с использованием хаотических сигналов [6—11]. В схеме с нелинейным подмешиванием информационного сигнала в хаотический, передача речевых и музыкальных сигналов, в низкочастотном и в радиодиапазонах была продемонстрирована экспериментально [12].

Разработанные подходы и модели передачи информации с использованием хаотической синхронизации и их экспериментальная проверка заложили основу для развития хаотических коммуникаций. Вместе с тем дальнейшие исследования показали, что системы связи, использующие хаотическую синхронизацию (или хаотический синхронный отклик), предъявляют высокие требования к качеству канала связи и не являются на ближайшую перспективу практически применимыми.

Следующим важным шагом было осознание того, что, возможно, следует отказаться от использования хаотической синхронизации для того, чтобы улучшить характеристики связных систем, использующих хаос [13—14]. Отмеченные работы произвели отрезвляющий эффект на первоначальные радужные надежды исследователей, (в первую очередь в отношении применения хаотической синхронизации), но в то же время показала, что при использовании динамического хаоса можно получить неплохие характеристики связных систем в каналах с белым шумом.

Анализ накопленной к этому времени информации позволил сделать вывод о том, что предложения, которые рассматривались в подавляющем большинстве публикаций, посвященных связи на основе динамического хаоса, страдали отсутствием целевой установки: как, зачем и при каких условиях такие системы могут быть использованы. При этом даже в наиболее продвинутых работах авторы ограничивались сугубо модельными ситуациями, такими как рассмотрение характеристик системы в канале с белым шумом или в канале с простейшей многолучевой структурой. Все это находится весьма далеко от конкретных коммуникационных задач и практических инженерных решений.

При построении реальных коммуникационных каналов на основе хаоса, в том числе каналов беспроводной связи, нужно отчетливо понимать, что:

а) хаотические системы связи будут применяться только там и только в том случае, когда они будут иметь совокупность свойств, делающих их конкурентоспособными по отношению к другим типам беспроводных систем. В список этих свойств могут входить скорость передачи информации, простота и стоимость системы, устойчивость работы в конкретных условиях, многопользовательский доступ, возможность удовлетворения определенным правилам частотного регулирования и т. д.;

б) техника передачи информации с помощью хаотических сигналов находится в зарождающейся фазе, и эффективные инженерные решения достаточно ограничены по элементной базе.

С практической точки зрения за кадром долгое время оставался также вопрос о реализации самих источников динамического хаоса. Неявно предполагалось, что скорость передачи не будет превышать нескольких сотен килобит в секунду, а хаос будет использоваться в качестве промежуточного носителя. При этом хаотический сигнал может быть сгенерирован в цифровом виде, а его полоса будет составлять от нескольких мегагерц до нескольких десятков мегагерц. Подобный подход не дает возможности использовать такие наиболее притягательные свойства хаоса, как широкая полоса и простота устройства связи на основе аналогового хаотического генератора. Именно такого типа преимущества могли бы служить причиной использования систем связи на основе динамического хаоса вместо традиционных систем с регулярными носителями информации.

В 2000 году в ИРЭ РАН была предложена схема *прямохаотической радиосвязи* [15—27, 40]. Ключевым понятием предлагаемой технологии является хаотический радиоимпульс. Он представляет собой фрагмент сигнала с длиной, превышающей длину квазипериода хаотических колебаний. Полоса частот хаотического радиоимпульса определяется полосой частоты исходного хаотического сигнала, генерируемого источником хаоса, в широких пределах изменения длины импульса не зависит от его длительности. Это существенно отличает хаотический радиоимпульс от классического, заполненного фрагментом периодической несущей, полоса частот которого определяется его длиной. В основу прямохаотических схем связи заложены три базовые идеи:

- источник хаоса генерирует хаотические колебания непосредственно в заданной полосе СВЧ-диапазона;
- ввод информационного сигнала в хаотический осуществляется путем формирования соответствующего потока хаотических радиоимпульсов;

- извлечение информации производится из хаотического СВЧ-сигнала без промежуточного преобразования частоты.

Ниже на примере прямохаотической схемы рассматриваются достижения и перспективы применения динамического хаоса в беспроводных системах связи.

1. Генераторы динамического хаоса как эффективные источники шумоподобных сигналов

Основным элементом хаотической системы связи является генератор хаоса. Не будем рассматривать схемы связи, использующие хаос как промежуточный носитель, речь пойдет о генераторах хаоса радио- и СВЧ-диапазонов.

К настоящему времени известен не один десяток различных динамических систем, демонстрирующих хаотическое поведение. Что стоит взять любую из них и с учетом возможностей современной технологии создать на ее основе генератор хаоса?

Однако решение не является столь простым и очевидным.

Во-первых, нам понадобится не просто динамическая система, порождающая хаос, а система, порождающая хаос с определенными, как минимум, спектральными свойствами. Например, система, порождающая хаос с относительно равномерной спектральной плотностью в заданном диапазоне частот.

Во-вторых, это должна быть система, реализуемая на более или менее стандартных элементах, используемых в радиотехнике.

В-третьих, это должна быть система, генерирующая хаос на высоких или очень высоких частотах, возможно близких к предельным характеристикам используемой технологии.

Наконец, в-четвертых, могут быть дополнительные ограничения со стороны технологии, например, требование, чтобы система была реализована на КМОП-технологии.

Рассмотрим ситуацию с генераторами хаоса, сложившуюся на сегодняшний день.

История генераторов хаоса радио- и СВЧ-диапазонов восходит, как минимум, к началу шестидесятых годов двадцатого века. О генераторе шумоподобных колебаний, использующем в качестве активного элемента лампу бегущей волны, шла речь в патенте США [28]. Однако там еще отсутствовало ясное объяснение природы этих колебаний. Серия теоретических и экспериментальных работ, посвященных генераторам стохастических (хаотических) колебаний на основе лампы бегущей волны, была проведена в ИРЭ АН СССР [29—32]. В этих работах были предложены устройства, названные «шумотронами», дано объяснение причинам возникновения шумоподобных колебаний и построены простые модели на

основе одномерных отображений и дифференциально-разностных уравнений, достаточно адекватно отражающие основные процессы возникновения и развития хаотической динамики. Аналогичные исследования проводились и с генераторами на основе других электронных приборов, в частности, на основе лампы обратной волны [33—34].

В начале восьмидесятых годов возник естественный интерес к маломодовым генераторам хаоса радио- и СВЧ-диапазонов на основе транзисторов и других полупроводниковых активных элементов. Практические достижения в этой области отображены в обзоре [35]. Следует отметить, что математические модели транзисторных генераторов хаоса стали появляться уже после разработки реальных устройств и начала их практического использования [36]. Среди этих моделей удачной с теоретической точки зрения оказалась модель генератора с полутора степенями свободы, построенная на основе трехточечной схемы Колпитца. Эта модель [37] детально исследована в работе [38]. Отметим, однако, что она изучалась в соответствии с существующей традицией как система обыкновенных дифференциальных уравнений третьего порядка и поэтому принципиально была низкочастотной моделью транзисторного генератора. В работе [39] предложен метод построения на ее основе моделей источников хаоса радио- и СВЧ-диапазона.

Он заключается в переходе от упрощенной математической модели к использованию специальных симулирующих пакетов типа ADS (Advanced Design System), в которых уже сами активные элементы представляются в виде достаточно сложных моделей (для потребителей это модели типа черного ящика). Модели активных элементов либо имеются в самих пакетах, либо поставляются производителями соответствующих компонентов. Кроме того, в симулирующих пакетах используются модели как активных, так и пассивных компонентов, учитывающие появления артефактов на высоких и сверхвысоких частотах.

Могут ли эти источники хаоса рассматриваться как генераторы хаоса, и насколько они эффективны как генераторы хаоса?

Имеющиеся результаты свидетельствуют, что с помощью таких источников можно получить более или менее равномерную спектральную плотность мощность в заданном диапазоне частот. При этом характеристики такого рода устройства оказываются достаточно чувствительными к вариации параметров. Другим вопросом, чрезвычайно важным с точки зрения приложений, является энергетическая эффективность источника. Оценки, проведенные для рассматриваемой модели генератора, показывают, что типичная доля мощности, которая может быть передана от источника в передатчик, составляет около 1%.

Много это или мало? С точки зрения стандартного электронного устройства СВЧ-диапазона, это, конечно, немного. Но нужно учитывать, что оптимизация по энергетическим характеристикам для рассматриваемой

модели не проводилась. В то же время известны генераторы хаоса с КПД до 15—20%.

Посмотрим на проблему КПД с несколько другой стороны. Будем рассматривать генератор хаоса как источник псевдошумового аналогового сигнала, т. е. как источник специфического шума. Здесь обнаруживается очень интересная особенность источников хаоса: как источники псевдошумового сигнала они даже в случае низкой энергетической эффективности, при рассмотрении их как электронных устройств, оказываются на 4—7 порядков более эффективными при генерации аналоговых псевдошумовых сигналов, чем шумовые диоды, и имеют существенно большую спектральную плотность мощности.

Здесь напрашивается следующая аналогия. Подобно тому, как оптический квантовый генератор — лазер — представляет собой эффективный источник специального (узкополосного) излучения оптического (инфракрасного, ультрафиолетового) диапазона, генератор хаоса представляет собой, как минимум, эффективный источник специальных широкополосных и сверхширокополосных шумоподобных колебаний.

Рассмотрим принципы применения этих уникальных устройств в системах связи.

2. Сверхширокополосные прямохаотические системы связи

Структура прямохаотической системы связи показана на рис. 1.

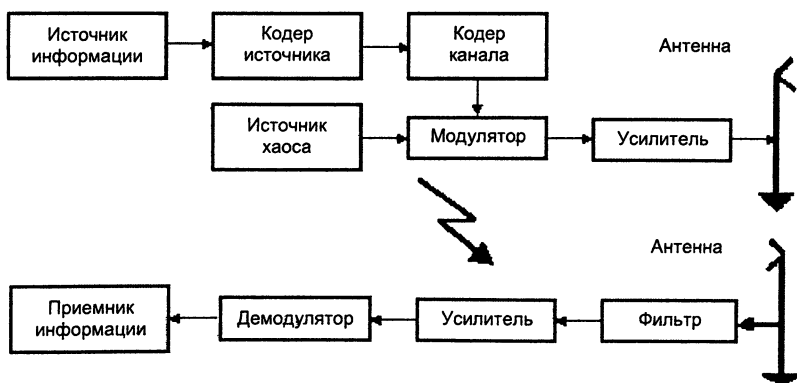


Рис. 1. Структура прямохаотической системы связи (ПХСС)

Сверхширокополосными прямохаотическими системами связи (СШП ПХСС) будем называть системы с относительной полосой частот DF :

$$DF = \frac{F_U - F_L}{2(F_U + F_L)} > 0,2, \quad (1)$$

где F_U и F_L — соответственно верхняя и нижняя частоты в спектре сигнала или с полосой частот

$$\Delta F > 500 \text{ МГц.} \quad (2)$$

В ПХСС могут использоваться различные виды модуляции: наличие или отсутствие хаотического импульса на информационной позиции (*chaotic on-off keying* — *chaotic OOK*), относительная хаотическая манипуляция (*differential chaotic shift keying* — *DCSK*), модуляция позиций импульсов (*pulse position modulation* — *PPM*) и т. д. [41]. Существенно, что для передачи информации здесь используется не непрерывный сигнал, а поток импульсов. Поэтому наряду с методом модуляции важными характеристиками являются длина импульса и скважность. Вариация этих характеристик определяет скоростные свойства системы связи и ее устойчивость для различных типов каналов связи.

В качестве приемника в ПХСС используется детектор огибающей.

Длина импульса T вместе с полосой частот ΔF определяет коэффициент процессинга (базу) сигнала:

$$B = 2 \cdot \Delta F \cdot T. \quad (3)$$

Характеристики прямохаотической системы связи в каналах с белым шумом

Рассмотрим типичные характеристики СШП ПХСС в канале с белым шумом для двух типов модуляции — *OOK* и *PPM*. Оба типа модуляции оперируют ортогональными сигналами. Поэтому можно ожидать, что при некогерентном приеме устойчивость передачи в канале с белым шумом будет примерно одинакова. Однако она будет зависеть от базы сигнала (3).

Результаты моделирования для длины импульсов равной 100 нс приведены на рис. 2. Расчеты проводились для сигналов с полосами 500 МГц и 2 ГГц. В первом случае база сигнала равна $B = 100$, во втором $B = 400$. На рис. 2 видно, что разница в устойчивости обоих методов модуляции к белому шуму не превышает 1 дБ. Сигналы с меньшим процессингом выигрывают примерно 2 дБ в устойчивости по отношению к сигналам с базой $B = 400$.

Для практических приложений, наряду с характеристиками в канале с белым шумом, важны характеристики в реальных каналах с многолучевым распространением, характерных для внутриофисных внутрикомнатных приложений. Рассмотрим этот вопрос на примере низкоскоростной СШП беспроводной связи для сенсорных сетей [42, 43].

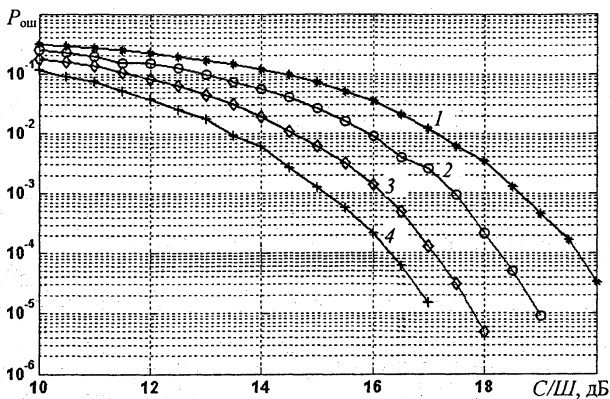


Рис. 2. Зависимость вероятности ошибки на бит ($P_{ош}$) от отношения сигнал/шум:

- 1 — ООК-модуляция, ширина полосы — 2 ГГц;
- 2 — PPM-модуляция, ширина полосы — 2 ГГц;
- 3 — ООК-модуляция, ширина полосы — 500 МГц;
- 4 — PPM-модуляция, ширина полосы — 500 МГц.

Влияние многолучевых каналов

При расчетах характеристик систем беспроводной связи внутри помещений на сверхширокополосных сигналах необходимо использовать реалистичные модели многолучевых каналов, отражающие сложные условия распространения, связанные с большим числом переотражений сигнала.

Эти каналы могут быть с наличием прямого луча (*LOS, Line of Sight*) или без прямого луча (*NLOS, no Line of Sight*).

В течение последних нескольких лет был проведен цикл экспериментальных работ, на основании результатов которых были предложены статистические модели внутриофисного многолучевого распространения сигналов [44—45]. Они позволяют оценивать основные характеристики канала связи для различных ситуаций (наличии/отсутствии прямого луча, типичных расстояний, характерных групповых запаздываний и т. д.). Именно эти модели рекомендованы комиссией (*IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs)*) [42, 43] для дальнейшего использования при исследовании систем сверхширокополосной связи.

Отличительной особенностью предложенных моделей распространения является то, что они по сути своей являются двухмасштабными: все приходящие в приемник лучи собираются в относительно плотные пакеты («кластеры»), характеризующиеся запаздыванием головного луча. Внут-

ри кластеров располагаются группы лучей с близкими запаздываниями. Наличие двух масштабов позволяет учитывать особенности распространения электромагнитных волн внутри помещений, в которых могут присутствовать относительно крупные однородные отражательные поверхности (стены, крупные предметы и др.), формирующие близкие по своим свойствам группы лучей. Различие в лучах внутри каждого кластера обусловливается особенностями данной конкретной поверхности (неоднородность, характерная геометрия и др.). Как известно, в канале с многолучевым распространением радиоволны движутся по многим лучам и приходят в приемник с разных направлений, с разными задержками, испытывая при этом еще и разное затухание. На входе приемника они складываются. Это может приводить к двум основным негативным факторам многолучевого распространения: затуханию сигнала и межсимвольной интерференции. Рассмотрим эти влияния этих факторов на систему сверхширокополосной связи с хаотическим носителем.

Затухание сигнала. В общем случае сложение волн приводит к их интерференции. В случае гармонического носителя это может вызывать как усиление, так и ослабление сигнала в зависимости от частоты. Однако в случае сверхширокополосного хаотического сигнала лучи, пришедшие с разными запаздываниями на вход приемника, оказываются некоррелированными, поскольку автокорреляционная функция хаотического сигнала очень быстро спадает. Это означает, что при сложении лучей мощность принимаемого сигнала возрастает.

Межсимвольная интерференция. Вторым негативным фактором многолучевого распространения является межсимвольная интерференция. Это явление наблюдается в тех схемах связи, где передаваемая информация кодируется положением на временной оси, например, позицией последовательных радиоимпульсов. Если импульсная переходная характеристика многолучевого канала очень длинная (задержка откликов велика), то энергия, переносимая лучами одного символа, может попадать на позицию другого символа, что может вызывать ошибки. Часто с этим борются, вводя защитный интервал. Тогда задержанные лучи приходят не на позицию следующего информационного символа, а на защитный временной интервал между ними. Искажений сигнала на информационных интервалах становится меньше, однако платой за это служит потеря части энергии.

В случае, когда канал является существенно «многолучевым», т. е. переходная характеристика канала содержит много сильных, задержанных на большое время откликов, длительность переходной характеристики сравнима или превышает длительность импульсов, форма принимаемого сигнала сильно искажается.

При этом, если длина радиоимпульса достаточно велика (сравнима с длительностью многолучевого отклика), то большая часть энергии будет

попадать на позицию символа, что не только уменьшает ошибку, но и увеличивает энергетическую эффективность приема.

Для количественной оценки эффекта усиления сигнала при многолучевом распространении было введено понятие коэффициента многолучевого усиления $K_{\text{мл ус}}$. Будем понимать под этим отношение энергии радиоимпульса, принесенной всеми лучами на определенном временном интервале, к энергии радиоимпульса, принесенной одним самым сильным лучом.

Более подробно эффект многолучевого усиления описан в статье [48].

3. Экспериментальные макеты

Развиваемая технология апробирована в ряде лабораторных макетов беспроводных коммуникационных систем. Ниже описываются некоторые из них.

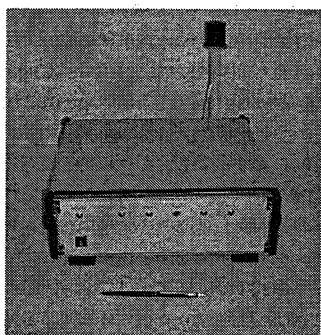
Эксперименты в диапазоне 0,5—3,5 ГГц

Первые эксперименты по сверхширокополосной беспроводной связи проводились в диапазоне 0,5—3,5 ГГц. По радиоканалу передавались пакеты данных в стандарте Ethernet. Для проведения экспериментов был разработан адаптер, который формировал потоки данных со скоростью 10 Мбит/с (стандарт 10 BaseT) и 100 Мбит/с (стандарт 100 BaseTx) [23]. Была реализована устойчивая передача информации при длинах тестовых пакетов от 500 до 60000 байт. Помимо тестовых пакетов по эфиру производилась также передача файла размером 10 Мбайт (видеосюжет). Время этой передачи практически совпало со временем передачи того же файла по проводной линии, что свидетельствует о низком уровне ошибок при передаче файла с помощью хаотической беспроводной системы связи.

Лабораторный макет системы связи в диапазоне 3—5 ГГц

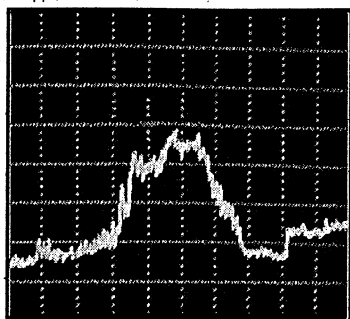
Следующий этап экспериментальных исследований заключался в создании автономного лабораторного макета радиопередатчика, способного передавать цифровую информацию со скоростями 100 Мб/с в диапазоне 3—5 ГГц. Выбор частотного диапазона был продиктован требованиями комиссии IEEE, предъявляемыми к стандарту беспроводной сверхширокополосной связи для мультимедийных приложений IEEE 802.15.3a.

В состав радиопередатчика (внешний вид которого изображен на рис. 3, а) входят: передающий модуль, содержащий генератор сверхширокополосных хаотических сигналов в указанном диапазоне; модулятор; приемная часть и приемопередающая антенна. Лабораторный макет включал также цифровой модуль, предназначенный для сопряжения приемопередатчика с выходной шиной компьютера и осуществляющий формирование управляющего потока модулирующих импульсов, а также обработку сигнала на выходе аналоговой приемной части.



а

P/10 дБ



б

Рис. 3. Внешний вид лабораторного макета (а) и форма спектральной характеристики сверхширокополосного хаотического сигнала (б).

Эксперименты проводились в соответствии со схемой, показанной на рис. 4. Расстояние между макетами составляло 3—5 м. Для передачи использовался метод кодирования без возвращения к нулю (*non return to zero — NRZ*). Спектр мощности формируемого передатчиком потока хаотических радиоимпульсов изображен на рис. 3, б. По уровню 10 дБ сигнал лежит в полосе ~ 2 ГГц (от 3 до 5 ГГц).

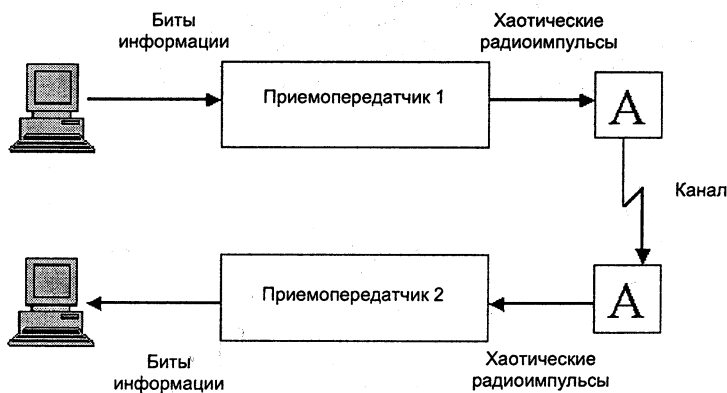


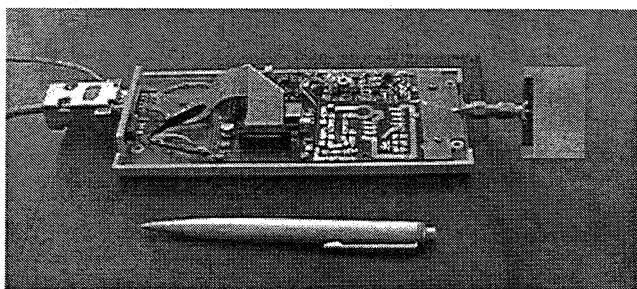
Рис. 4. Схема экспериментов с макетами приемопередатчиков

Передача производилась в стандарте *Ethernet* со скоростью до 100 Мбит/с. Вероятность ошибки составляла 10^{-8} — 10^{-9} при скорости передачи 10—20 Мбит/с и 10^{-3} — 10^{-4} при скорости передачи 100 Мбит/с.

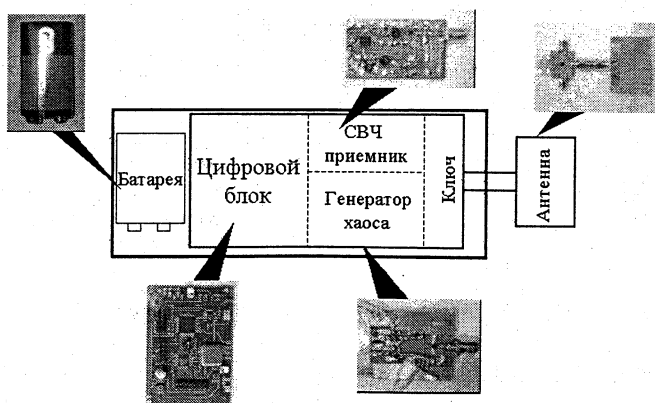
Макет малогабаритного радиопередатчика

Еще одной возможной областью применения прямохаотических систем являются беспроводные персональные коммуникационные системы, ориентированные на обмен цифровой информацией между различными периферийными устройствами, датчиками, сенсорами, бытовыми приборами и т. д., расположенными на относительно небольших расстояниях (от единиц до нескольких десятков метров). Среди основных требований, выдвигаемых к такого класса устройствам, — простота, технологичность и малое энергопотребление. Предварительные исследования показали, что на основе прямохаотической технологии связи могут быть созданы приемопередатчики с вышеперечисленными свойствами. Для проверки этих предварительных выводов был создан лабораторный макет малогабаритного приемопередатчика и проведено его экспериментальное исследование.

Внешний вид и блок-схема макета приёмопередатчика приведены на рис. 5. Он состоит из цифрового блока, генератора хаоса, приемника СВЧ, ключа и антенны.



а



б

Рис. 5. Внешний вид (а) и структура (б) макета приёмопередатчика

Принцип работы радиопередатчика следующий. Генератор хаоса вырабатывает непрерывный хаотический сигнал в полосе частот 3—5 ГГц. Сигнал с выхода генератора поступает далее на один из СВЧ-входов ключа. Второй СВЧ-вход ключа соединён с приёмником, а СВЧ-выход ключа подключён к антенне. Работой ключа управляет сигнал с выхода цифрового блока. Если он соответствует единичному биту («1»), то сигнал с выхода генератора хаоса поступает в антенну и излучается в пространство. Напротив, нулевой информационный бит («0») переключает ключ в положение соединения антенны с приёмником, прерывая соединение антенны с генератором хаоса.

Цифровой блок, с одной стороны, преобразует сигнал с выхода источника информации в форму, необходимую для управления ключом, а с другой стороны, содержит вторичный источник питания, вырабатывающий напряжение для питания генератора хаоса и приёмника. Моментами подачи питания на тот или иной из указанных блоков также управляет цифровой блок. Так, в режиме «Передача» необходимое напряжение питания подается только на генератор хаоса и, наоборот, в режиме «Приём» генератор отключается, а питание поступает только на приёмник макета. В последнем случае, на выходе приёмника формируется сигнал, представляющий собой последовательность видеоимпульсов, соответствующих передаваемой информации. Видеоимпульсы поступают в цифровой блок, где претерпевают обратное преобразование информационного сигнала в исходный вид, воспринимаемый источником информации.

Структура сигналов в различных точках приемопередатчика изображена на рис. 6.

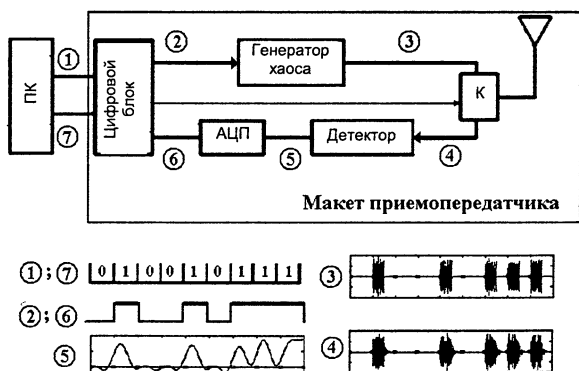


Рис. 6. Структура сигнала в различных точках макета: 1, 7 — сигнал на входе/выходе устройства; 2 — модулирующий сигнал на входе генератора хаоса; 3 — поток хаотических радиоимпульсов, излучаемых в беспроводной канал; 4 — принимаемый сигнал; 5 — сигнал на выходе детектора; 6 — сигнал на входе цифровой платы трансивера.

С использованием созданных приемопередатчиков были проведены эксперименты по передаче звуковых фрагментов MP3-формата. MP3-файлы состояли из серии независимых друг от друга фреймов длиной до килобайта каждый. Для передачи из фреймов или их частей приготавливались пакеты, содержащие, кроме собственно музыкальной информации, служебную информацию. Компьютер с проигрывателем запрашивает у компьютера-источника определенные пакеты и, получая пакеты, восстанавливает из них фреймы и воспроизводит музыкальный трек.

Эксперименты продемонстрировали устойчивую связь, позволяющую без искажения прослушивать музыкальные MP3-файлы на приемной стороне. В ходе экспериментов показана работоспособность устройств на расстоянии до 30 м при скорости передачи информации 100—200 Кбит/с. Кратковременное нарушение условий распространения (экранирование принимаемого сигнала в непосредственной близости от приемной антенны) прерывало процесс воспроизведения музыкальных файлов. Тем не менее, восстановление беспроводного канала практически мгновенно возобновляло процесс воспроизведения без каких-либо дополнительных процедур инициализации или перезапуска.

4. Применения

С тех пор как в начале 2002 года Федеральная комиссия по коммуникациям США ввела частотный диапазон и спектральную маску мощности для нелицензируемых СШП-систем связи [47], был предложен и проанализирован ряд коммерческих приложений для систем сверхширокополосной связи в полосе 3,1—10,6 ГГц.

Анализ показал, что основной областью применения СШП-сигналов являются беспроводные персональные локальные сети (*wireless personal area networks* — *WPAN*), для которых особое значение имеют низкая стоимость аппаратуры и малое энергопотребление. Такие сети разрабатываются, в частности, в рамках упоминавшегося выше стандарта IEEE 802.15.4a, предназначенного для передачи данных со скоростями до 1—3 Мбит/с (агрегированная по сети скорость до 10 Мбит/с) на расстоянии до 100 м. Кроме того, устройства стандарта должны иметь возможность определять свое местоположение в сети. Для реализации задач стандарта предложено использовать несколько типов сигналов: ультракороткие импульсы [46] и последовательности таких импульсов, сигналы с расширением спектра, хаотические радиоимпульсы, сигналы с ортогональным частотным разделением (*OFDM*) и сигналы с линейной частотной модуляцией (ЛЧМ).

Беспроводные сверхширокополосные средства связи для «умных домов и офисов» рассматриваются как одни из наиболее привлекательных продуктов, которые могут существенно расширить рынок *WPAN*. Кроме

того, множество вспомогательных коммуникационных устройств также должны быть беспроводными и едиными, например, с сотовыми телефонами для более удобного использования.

При разработке сверхширокополосных *WPAN* необходимо учитывать ряд серьезных ограничений на излучаемый сигнал.

Во-первых, максимальная спектральная плотность сверхширокополосных сигналов ограничена величиной $-41,3$ дБ·м/МГц, поэтому даже при использовании всего разрешенного диапазона интегральная излучаемая мощность не должна превышать $-2,4$ дБ·м.

Вторым ограничением, накладываемым на параметры излучения, является ограничение на пиковую излучаемую мощность, которая не должна превышать 0 дБ·м в полосе 50 МГц. Таким образом, для полосы 500 МГц мы имеем ограничение в 20 дБ·м и для полосы 2 ГГц — 32 дБ·м.

Наконец, третье ограничение на параметры накладывается технологией. Как правило, напряжение питания в СШП нелицензируемых системах связи не превышает 2 В. Поэтому и напряжение на антенне не будет превышать 1 В. Если антенна согласована на 50 Ом, то пиковая мощность будет ограничена 10 мВт.

Исходной идеей для реализации сверхширокополосной связи была идея использования в качестве носителя информации сверхкоротких импульсов (при полосе частот $7,5$ ГГц длина импульса около 150 пс). Согласно этой идее каждому импульсу соответствует один бит информации, что делает обработку сигнала в приемнике очень простой.

Ряд предложений по использованию этой технологии был внесен при разработке стандарта IEEE 802.15.3a. Однако расчеты показывают, что дальность передачи при технологическом ограничении на амплитуду будет не более 10 м.

Это обстоятельство не было принципиальным при разработке стандарта для высокоскоростной передачи данных, но является очень существенным при разработке стандарта для низкоскоростной передачи данных IEEE 802.15.4a. В этом стандарте дальность передачи должна быть, как минимум, 30 м. Для того чтобы реализовать такие дальности, нужно иметь на входе антенны напряжение ~ 3 В, что входит в противоречие с технологическими тенденциями к снижению напряжения питания и потребляемой мощности. Еще более драматичной ситуация будет, если потребуется обеспечить дальность передачи на 50 — 60 и более метров.

Заметим, что из трех ограничений на излучаемую мощность именно третье ограничение является наиболее тяжелым для импульсных систем. Действительно, первое ограничение можно обойти, снижая скорость передачи, т. е. уменьшая число излучаемых в единицу времени импульсов. Второе ограничение оставляет возможность обеспечить дальность передачи до 1000 м (в полосе 2 ГГц и при скорости 1000 бит/с). Но третье ограничение снижает максимальную дальность передачи до 10 м.

Выходом из этого положения является переход от передачи одного бита одним импульсом к передаче одного бита серией импульсов. В этом случае при сохранении энергии, излучаемой на один бит, можно снизить энергию, приходящуюся на один импульс, обратно пропорционально числу импульсов в серии. Это означает переход от сигналов с единичной базой (процессингом) к сигналам с большой базой.

Переход к сигналам с большой базой содержится в той или иной форме в большинстве предложений по стандарту IEEE 802.15.4a. Однако при этом исходная идея простоты короткоимпульсных систем начинает размываться.

Прямохаотические системы, у которых спектр мощности не зависит от длины импульса, а следовательно и от величины базы сигнала, не имеют энергетических ограничений, свойственных системам с короткими импульсами, и при любой скорости передачи сохраняют в системах связи идеологию «один бит — один импульс». Это является одним из главных их естественных преимуществ по сравнению со сверхширокополосными системами, использующими сверхкороткие импульсы, и существенно упрощает аппаратные решения.

В современных беспроводных сетях связи выбор той или иной технологии для использования определяется совокупностью характеристик, специфических для области применения. Для сверхширокополосных систем основные из этих характеристик сведены в таблицу, в которой приводится сравнительный анализ трех сверхширокополосных технологий: прямохаотической (СШП ПХСС); *OFDM* и технологии, основанной на использовании ультракоротких видеоимпульсов. Сравнительная оценка технологий для каждого критерия приводится по трехбалльной шкале качества (хорошо — «+»; удовлетворительно — «0»; плохо — «-»).

№ п/п	Характеристика	ПХСС	<i>OFDM</i>	Импульсы
1	Потребление энергии	+	0	+
2	Простота реализации	+	-	0
3	Масштабирование по скоростям передачи	+	-	+
4	Определение местоположения	+	-	+
5	Устойчивость к многолучевому распространению	+	+	0
6	Критичность к антенне	+	+	-
7	Масштабирование по дальности	+	-	0
8	Синхронизация	+	0	0

Потребление энергии. Экономия в потреблении энергии ПХСС осуществляется за счет полномасштабного включения передатчика и приемника только на время их непосредственной работы. При пакетной передаче информации приемник и передатчик включаются только на время посылки или приема пакета. Кроме того, если производится передача информации со скважностью между битами больше, чем единица, то на время между информационными позициями передатчик может быть выключен. Такой режим возможен благодаря большой длине импульсов (~ 100 нс). Этой длительности хватает, чтобы в основном справиться с переходными процессами в системах питания и обеспечить выход работы элементов приемопередатчика на рабочий режим. Отметим, что добиться подобного результата с короткоимпульсными системами значительно сложнее.

Простота реализации. Технология ПХСС реализуема на существующей технологической базе начиная с 0,35 мкм. В ней используется приемник прямого усиления.

Масштабирование по скоростям передачи. Технология ПХСС позволяет перекрыть диапазон скоростей передачи информации от единиц бит в секунду до 2—3 Гбит/с. В короткоимпульсных системах увеличение скорости передачи выше 100 Мбит/с приводит к значительному усложнению структуры системы.

Определение местоположения. Технология имеет хорошие характеристики по определению местоположения передающего устройства. Эти характеристики определяются полосой частот хаотического сигнала (мы предлагаем использовать полосы от 2 ГГц и выше). Таким образом, точность определения местоположения примерно такая же, как у короткоимпульсных систем. Системы на основе OFDM не могут быть использованы для определения местоположения с высокой точностью в силу большой длительности базовых сигналов.

Устойчивость к многолучевому распространению. Защита от замираний в ПХСС обеспечивается автоматически в силу большой базы сигнала. При малых и средних скоростях передачи защита от межсимвольной интерференции, как показано выше, эффективно обеспечивается путем введения защитных интервалов. Кроме того, в случае многолучевого распространения возникает значительный энергетический выигрыш по отношению к однолучевому из-за эффекта многолучевого усиления.

Критичность к антенне. Диаграмма направленности антенны для хаотического сигнала определяется диаграммами направленности на отдельных частотах внутри частотного диапазона. Здесь не может происходить развалов диаграммы, характерных для ультракороткоимпульсных и короткоимпульсных систем. Таким образом, технология не критична к антеннам. Могут быть использованы антенны обычных типов, обеспечивающие необходимую широкополосность. При этом достаточно просто реализовать как ненаправленную, так и направленную диаграммы.

Масштабирование по дальности. Технология позволяет реализовать системы передачи, эффективно работающие на очень малых, малых, средних, больших и очень больших дальностях. Эти возможности определяются большим диапазоном изменения скоростей передачи информации, возможностью использования как всенаправленных, так и направленных антенн, а также возможностью применения к хаотическим сигналам стандартной усилительной техники. Ограничения на масштабирование короткоимпульсных систем связаны с упоминавшимися технологическими ограничениями.

Синхронизация. Требования к синхронизации в СШ ПХСС определяются длительностью хаотических радиоимпульсов. Требуемая точность синхронизации составляет примерно 10% от длительности импульса. Так, например, при длительности импульса 100 нс требуемая точность синхронизации составит 10 нс. В системах с ультракороткими импульсами требуемая точность синхронизации составляет 10—20 пс, а в короткоимпульсных системах 0,1—0,2 нс. Причем эти требования должны выполняться независимо от того, решает ли система задачу определения местоположения или нет. Потому выигрыш СШ ПХСС по требованиям к синхронизации составляет от 50 до 500 раз.

Таким образом, по совокупности характеристик беспроводные сверхширокополосные прямохаотические системы имеют хорошие перспективы.

В заключение отметим, что технология прямохаотической связи предложена Институтом радиотехники и электроники РАН совместно с Институтом передовых технологий (SAIT) корпорации *Samsung Electronics* в январе 2005 года в качестве технологии для нового стандарта беспроводных персональных коммуникационных сетей (WPAN) — IEEE802.15.4a. В марте 2006 года международной рабочей группой принято решение о ее включении в разрабатываемый стандарт в качестве опционального варианта.

Литература

1. *Дмитриев А. С., Кислов В. Я., Панас А. И. и др.* Система связи с шумовой несущей: А. с. 279024. СССР, 1985.
2. *Fujisaka H., Yamada T.* // Prog. Theor. Phys. 1983. V. 69. P. 32.
3. *Pikovsky A.* // Z. Physik B. 1984. V. 55. P. 149.
4. *Афраймович В. С., Веричев Н. И., Рабинович М. И.* // Изв. вузов. Сер. Радиофизика. 1986. Т. 29, № 9. С. 1050.
5. *Pecora L. M., Carroll T. L.* // Phys. Rev. Lett. 1990. V. 64, № 8. P. 821.
6. *Chua L. O., Kocarev L., Eckert K. et al.* // Int. J. Bifurcation and Chaos. 1992. V. 2. P. 705.
7. *Andreyev Yu. V., Belsky Yu. L., Dmitriev A. S.* // Proc. Int. Seminar Nonlinear Circuits and Systems. Moscow, Russia, 1992. V. 1. P. 51.
8. *Cuomo K., Oppenheim A.* // Phys. Rev. Lett. 1993. V. 71, № 1. P. 65.

9. Halle K. S., Chai W. W., Itoh M. et al. // Int. J. Bifurcation and Chaos. 1993. V. 3, № 2. P. 469.
10. Hasler M., Dedieu H., Kennedy M., Schweizer J. // Proc. of Int. Symp. on Nonlinear Theory and Application. Hawaii, USA, 1993. P. 87.
11. Волковский А. Р., Рульников Н. В. // Письма в ЖТФ. 1993. Т. 19, № 3. С. 71.
12. Дмитриев А. С., Панас А. И., Старков С. О. Эксперименты по передаче музыкальных и речевых сигналов с использованием динамического хаоса: Препринт № 12(600). М.: ИРЭ РАН, 1994.
13. Kolumban G., Kennedy M. P., Chua L. O. // IEEE Trans. 1997. V. CS-44. P. 927.
14. Kolumban G., Kennedy M. P., Chua L. O. // IEEE Trans. 1998. V. CS-45. P. 1129.
15. Дмитриев А. С., Кяргинский Б. Е., Максимов Н. А. и др. // Радиотехника. 2000. № 3. С. 9.
16. Дмитриев А. С., Кяргинский Б. Е., Максимов Н. А. и др. Прямохаотическая передача информации в СВЧ-диапазоне: Препринт № 1 (625). М.: ИРЭ РАН, 2000.
17. Дмитриев А. С., Старков С. О. // Зарубеж. радиоэлектроника. Усп. совр. радиоэлектроники. 1998. № 11. С. 4.
18. Дмитриев А. С., Панас А. И., Старков С. О. и др. Способ передачи информации с помощью хаотических сигналов: Пат. РФ № 2185032 от 27.07.2000.
19. Dmitriev A. S., Panas A. I., Starkov S. O. Electronic NonLinear Science Preprint. <http://arxiv.org/abs/nlin.CD/0110047>.
20. Дмитриев А. С., Кузьмин Л. В., Панас А. И. и др. // Зарубеж. радиоэлектроника. Усп. совр. радиоэлектроники. 2003. № 9. С. 26.
21. Dmitriev A. S., Kyarginsky B. Ye., Panas A. I. et al. // Int. J. Bifurcation and Chaos. 2003. V. 13, № 6. P. 1495.
22. Дмитриев А. С., Панас А. И., Старков С. О. и др. // РЭ. 2001. Т. 46, № 2. С. 224.
23. Дмитриев А. С., Кяргинский Б. Е., Панас А. И. и др. // РЭ. 2002. Т. 47, № 10. С. 1219.
24. Dmitriev A. S., Hasler M., Panas A. I. et al. // Nonlinear Phenomena in Complex Systems. 2003. V. 6, № 1. P. 488.
25. Дмитриев А. С., Кяргинский Б. Е., Панас А. И. и др. // Письма в ЖТФ. 2003. Т. 29, № 2. С. 70.
26. Andreyev Yu. V., Dmitriev A. S., Efremova E. V. et al // Int. J. Bifurcation and Chaos. 2005. V. 15, № 11.
27. Дмитриев А. С., Кинев А. В., Клецов А. В. и др. Дистанционное управление мобильными объектами с помощью сверхширокополосных хаотических СВЧ сигналов: Препринт № 1 (639). М.: ИРЭ РАН, 2005.
28. Reis C. A., Zellenz J. E. Pat. USA №3178655. 1961. С1. 331—78.
29. Мясин Е. А., Кислов В. Я., Богданов Е. В. Способ генерирования электромагнитных шумовых колебаний: А.с. 1125735 СССР. Приоритет 31.07.67, 1967, 23.11.84 // Бюл. изобретений. 1984. № 43.
30. Кислов В. Я., Залогин Н. Н., Мясин Е. А. // РЭ. 1979. Т. 24, № 6. С. 118.
31. Кислов В. Я. // РЭ. 1980. Т. 25, № 8. С. 1683.
32. Анисимова Ю. В., Дмитриев А. С., Залогин Н. Н. и др. // Письма в ЖЭТФ. 1983. Т. 37, № 8. С. 387.
33. Безручко Б. П., Кузнецов С. П., Трубецков Д. И. // Письма в ЖЭТФ. 1979. Т. 29, № 3. С. 180.
34. Безручко Б. П., Булгакова Л. В., Кузнецов С. П. и др. // РЭ. 1983. Т. 28, № 6. С. 1136.
35. Кислов В. Я. // РЭ. 1993. Т. 38, № 10. С. 1783.
36. Дмитриев А. С., Иванов В. П., Лебедев М. Н. // РЭ. 1988. Т. 33, № 5. С. 1085.
37. Kennedy M. // IEEE Trans. 1994. V. CS-41, № 11. P. 771.
38. Feo O., Maggio G., Kennedy M. // Int. J. of Bifurcation and Chaos. 2000. V. 10, № 5. P. 935.

39. *Дмитриев А. С., Ефремова Е. В., Хилинский А. Д.* Принципы компьютерного моделирования транзисторных генераторов хаоса в пакете ADS (Advanced Design System). Препринт № 5(633). М.: ИПЭ РАН, 2003. 58 с.

40. *Дмитриев А. С., Панас А. И.* Динамический хаос. Новые носители информации для систем связи. М: Физматлит, 2002.

41. *Proakis John G.* Digital communications. New York: MCGraw-Hill, 3rd ed , 1995.

42 Channel Modeling Sub-committee Report Final.

http://grouper.ieee.org/groups/802/15/pub/2002/Nov02/02490r0P802-15_SG3a-Channel-Modeling-Subcommittee-Report-Final.zip, 2002.

43. Channel Modeling Sub-committee Report Final.

<http://grouper.ieee.org/groups/802/15/pub/04/15-04-0662-02-004a-channel-model-final-report-r1.pdf>, 2004.

44. *Cramer J. M., Scholtz R. A., Win M. Z.* // Proc. of MILCOM'99. Atlantic city, USA, 31 oct — 3 nov, 1999. V. 2. P. 1191.

45. *Ghassemzadeh S., Jana R., Rice C. et al.* // IEEE UWBST'02. Baltimore, USA, 2002.

46. *Win M. Z., Scholtz R. A.* // IEEE Comm. Lett. 1998. V. 2, № 2. P. 10.

47. FCC News Release. Feb. 14. 2002. <http://www.fcc.gov/headlines2002.html>.

48. *Андреев Ю. В., Дмитриев А. С., Клецов А. В.* // РЭ. 2007, № 52 (в печати).

Нелинейные колебания в структурированных системах

МАТЕМАТИЧЕСКИЕ МОДЕЛИ, ТОЧНЫЕ РЕШЕНИЯ, МЕТОДЫ АНАЛИЗА

О. В. Руденко

Обсуждается ряд моделей нелинейной акустики, представляющих интерес и для других разделов физики нелинейных волн. Приводятся точные решения, имеющие физический смысл. Описываются некоторые способы их нахождения или угадывания. На примерах показана связь «нестрогих» подходов с корректным математическим анализом нелинейных моделей на основе их симметрий, которому посвящена статья в данном сборнике В. Ф. Ковалева и Д. В. Ширкова (с. 433 — 446). Для привязки к содержанию этих лекций обсуждался смысл принципа априорного использования симметрий.

Среди примеров: модели и решения для бегущих и стоячих волн, волновых пучков, возбуждения нелинейных волн, нелинейных статистических задач, течений в акустическом поле, кинетики ударных волн. Рассматривались, в частности, дифференциальные уравнения высоких порядков и функциональные уравнения.

Начнем с «базовой» модели нелинейной акустики — уравнения Бюргера, описывающего недиспергирующие волны в диссипативной среде:

$$\frac{\partial V}{\partial z} - V \frac{\partial V}{\partial \theta} = \Gamma \frac{\partial^2 V}{\partial \theta^2}. \quad (1)$$

Здесь $\theta = \omega(t - x/c)$ — время в системе координат, «бегущей» вместе с волной со скоростью звука, и $V = u/u_0$ — нормированное возмущение колебательной скорости или давления. Расстояние z измерено в единицах длин образования разрыва; Γ — обратное акустическое число Рейнольдса (число Гольдберга) — отношение характерных «нелинейной» и «диссипативной» длин [1]. Общее решение уравнения Бюргера (1) имеет сложную структуру и не всегда удобно для анализа конкретных задач. Помимо тривиальных частных решений, получаемых из стационарного решения

$$\partial V / \partial z = 0, \quad V = \tanh[(\theta - \theta_0) / 2\Gamma] \quad (2)$$

(это ударная волна с конечной шириной фронта) путем сдвигов вдоль z, θ, V , полезными оказались другие, «нетривиальные» точные решения.

«Решение Хохлова» [2]

$$V = \frac{1}{1+z} \left[-\theta + \pi \tanh \frac{\pi \theta}{2\Gamma(1+z)} \right], \quad -\pi < \theta < \pi \quad (3)$$

описывает один период пилообразной волны в области существования развитых ударных фронтов. Формула была фактически «угадана» в 1960 году. Она получается также методом сращиваемых асимптотических разложений; при этом 1-е приближение дает точное решение.

Автомодельное решение уравнения Бюргера [3]

$$V = \sqrt{\frac{4\Gamma}{\pi z}} \frac{\exp(-\theta^2/4\Gamma z)}{C + \operatorname{erf}(\theta/\sqrt{4\Gamma z})} \quad (4)$$

описывает однополярный импульс с крутым передним фронтом; оно получается из инвариантности (1) относительно преобразования

$$V \rightarrow V/C, \quad z \rightarrow C^2 z, \quad \theta \rightarrow C\theta. \quad (5)$$

Соответственно, конструируя подстановку

$$V = z^{-1/2} Q(z^{-1/2}\theta), \quad (6)$$

инвариантную относительно (5), удается свести (1) к обыкновенному дифференциальному уравнению и найти решение (4). Нетривиально и угаданное С. Н. Гурбатовым [4] преобразование: если функция $V = \Phi(z, \theta)$ есть точное решение уравнения Бюргера, то точным решением будет также

$$V = \frac{1}{1-z/z_0} \Phi\left(\frac{z}{1-z/z_0}, \frac{\theta}{1-z/z_0}\right) + \frac{\theta/z_0}{1-z/z_0}. \quad (7)$$

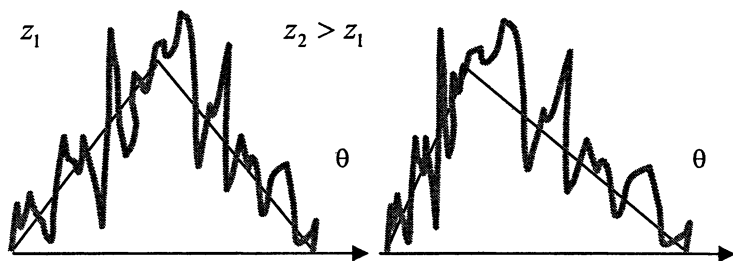


Рис. 1

Эта формула описывает, например, взаимодействие ВЧ-шума и НЧ-сигнала [5]. С увеличением расстояния z частота шума на переднем фронте регулярной волны растет, на заднем — уменьшается (рис. 1).

Однако рассказанная история «угадывания» точных результатов свидетельствует о недостаточной известности методов теории групп Ли применительно к решению дифференциальных уравнений [6, 7].

Покажем, как эти результаты получаются регулярными методами. Изменяя нормировку переменных, запишем уравнение Бюргера в самой простой форме:

$$u_x = uu_t + u_{tt}. \quad (8)$$

Инфинитезимальные симметрии [6, 7] уравнения Бюргерса образуют 5-мерную алгебру Ли, «натянутую» на линейно независимые операторы:

$$\begin{aligned} X_1 &= \frac{\partial}{\partial x}, & X_2 &= \frac{\partial}{\partial t}, & X_3 &= x \frac{\partial}{\partial t} - \frac{\partial}{\partial u}, \\ X_4 &= 2x \frac{\partial}{\partial x} + t \frac{\partial}{\partial t} - u \frac{\partial}{\partial u}, & X_5 &= x^2 \frac{\partial}{\partial x} + tx \frac{\partial}{\partial t} - (t + xu) \frac{\partial}{\partial u}. \end{aligned} \quad (9)$$

Техника нахождения операторов (9) стандартна; существуют даже компьютерные программы для их вычисления.

При известной симметрии инвариантное решение уравнения Бюргерса получается так. Вычисляются два независимых инварианта как решения уравнения $X(J) = 0$: $J_1 = \lambda(t, x)$, $J_2 = \mu(t, x, u)$. Затем один из инвариантов выражается как функция другого, $\mu = \phi(\lambda)$. Далее разрешенное относительно u выражение подставляется в уравнение Бюргерса. В результате получается уравнение для функции одной переменной $\phi(\lambda)$. Эта процедура уменьшает число независимых переменных на единицу.

Решения, инвариантные относительно группы растяжений, генерируемой оператором X_4 , часто называют автомодельными. В этом случае уравнение $X(J) = 0$ имеет вид

$$X_4(J) = 2x \frac{\partial J}{\partial x} + t \frac{\partial J}{\partial t} - u \frac{\partial J}{\partial u} = 0. \quad (10)$$

Его характеристическая система и инварианты таковы:

$$\frac{dx}{2x} = \frac{dt}{t} = -\frac{du}{u}; \quad \lambda = \frac{t}{\sqrt{x}}, \quad \mu = u\sqrt{x}. \quad (11)$$

Следовательно инвариантное решение нужно искать в виде

$$u = x^{-1/2} \Phi(\lambda), \quad \lambda = x^{-1/2} t. \quad (12)$$

Подставляя (12) в (8), получаем обыкновенное уравнение

$$\Phi'' + \Phi\Phi' + 0,5(\lambda\Phi' + \Phi) = 0 \quad (13)$$

и находим его точное решение:

$$u = \frac{2}{\sqrt{\pi x}} \frac{\exp(-t^2/4x)}{C + \operatorname{erf}(t/2\sqrt{x})}, \quad \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-s^2) ds. \quad (14)$$

Формула (14) совпадает с (4). Для обобщенного уравнения Бюргерса (его коэффициенты зависят от координат) автомодельные решения найдены в работе [8].

Для проективной группы, генерируемой оператором X_5 , получаем

$$X_5(J) = x^2 \frac{\partial J}{\partial x} + tx \frac{\partial J}{\partial t} - (t + xu) \frac{\partial J}{\partial u} = 0. \quad (15)$$

Характеристическая система и два инварианта имеют вид

$$\frac{dx}{x^2} = \frac{dt}{tx} = -\frac{du}{t+xu}; \quad \lambda = \frac{t}{x}, \quad \mu = t+xu. \quad (16)$$

Следовательно, инвариантное решение нужно искать в форме

$$u = -\frac{t}{x} + \frac{1}{x} \Phi(\lambda), \quad \lambda = \frac{t}{x}. \quad (17)$$

Подставляя это выражение в уравнение Бюргерса, получим для $\Phi(\lambda)$ уравнение $\Phi'' + \Phi\Phi' = 0$, из которого сразу находим решение Хохлова (3):

$$u = \frac{1}{x} \left[-t + \pi \tanh \left(C + \frac{\pi t}{2x} \right) \right]. \quad (18)$$

Используя проективное преобразование, из любого известного решения $u = \Phi(t, x)$ уравнения Бюргерса удастся получить однопараметрическое семейство новых решений:

$$u = \frac{ax}{1-at} + \frac{1}{1-at} \Phi \left(\frac{t}{1-at}, \frac{x}{1-at} \right). \quad (19)$$

Формула (19) отличается от (7) лишь обозначениями. Если применить преобразование (19) к стационарному решению уравнения Бюргерса, получается ряд новых нестационарных решений, в том числе (18) [9].

Решения можно найти не только из основных симметрий (9), но и из их линейных комбинаций. Рассмотрим, к примеру, оператор

$$X_1 + X_5 = (1+x^2) \frac{\partial}{\partial x} + tx \frac{\partial}{\partial t} - (t+xu) \frac{\partial}{\partial u}. \quad (20)$$

Решая характеристическое уравнение, находим инвариантную форму

$$u = -\frac{tx}{1+t^2} + \frac{1}{\sqrt{1+t^2}} \Phi(\lambda), \quad \lambda = \frac{x}{\sqrt{1+t^2}} \quad (21)$$

и соответствующее новое точное решение уравнения Бюргерса [9]:

$$\Phi = 2 \frac{d}{d\lambda} \ln \left[\sqrt{\lambda} J_{1/4} \left(\frac{\lambda^2}{4} \right) + C \sqrt{\lambda} Y_{1/4} \left(\frac{\lambda^2}{4} \right) \right]. \quad (22)$$

Несмотря на мощь групповых методов, они в ряде случаев трудны для нахождения точных решений, которые проще получить другими способами. Такая ситуация типична для уравнений с низкой симметрией.

Иногда точные решения для таких уравнений удается найти, понижая их порядок. Приведем примеры.

Пример 1. Нелинейное уравнение 2-го порядка для бегущих волн

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \left(p - \frac{\varepsilon}{c^2 \rho} p^2 \right) \quad (23)$$

точно сводится к паре уравнений 1-го порядка:

$$\frac{\partial p}{\partial x} = \pm \frac{1}{c} \sqrt{1 - \frac{2\varepsilon}{c^2 \rho}} p \frac{\partial p}{\partial t}. \quad (24)$$

Это легко проверить. Дифференцируя уравнение (24) по переменной x и заменяя производные по x в правой части производными по времени, получим исходное уравнение (23). Решить (24) гораздо проще, чем (23). Его особое решение, вычисляемое методом характеристик, имеет вид

$$p = \Phi \left(t \pm \frac{x}{c} \sqrt{1 - \frac{2\varepsilon}{c^2 \rho}} p \right). \quad (25)$$

Оно описывает нелинейные волны, бегущие вправо или влево (в зависимости от знака) вдоль оси x .

Пример 2. Рассмотрим теперь уравнение 4-го порядка [10]:

$$\frac{\partial V}{\partial z} - V \frac{\partial V}{\partial \theta} = -R \frac{\partial^4 V}{\partial \theta^4}, \quad (26)$$

которое продолжает серию замечательных уравнений: уравнения Бюргерса (1) — 2-й порядок, Кортевега — де Вриза — 3-й порядок. Для волн в рассеивающей среде

$$R = \frac{x_{\text{nonl}}}{x_{\text{scatt}}} = \frac{8 \langle \mu^2 \rangle a^3}{c^4} \frac{\omega_0^3 c^3 \rho}{\varepsilon p_0} \quad (27)$$

есть отношение характерных масштабов нелинейности и рассеяния, $\langle \mu^2 \rangle$ — средний квадрат флуктуаций показателя преломления, a — радиус корреляции, ε — нелинейный параметр, p_0 — амплитуда волны давления. Уравнение для стационарных волн и одно из его решений имеют вид:

$$R \frac{d^4 V}{d\theta^4} = V \frac{dV}{d\theta}, \quad \theta = \left(\frac{40}{9} R \right)^{1/3} \int_0^V \frac{dy}{(1-y^2)^{2/3}}. \quad (28)$$

Формула (28) получена так. Уравнению 4-го порядка сопоставлялось уравнение 1-го порядка $dV/d\theta = f(V)$, правая часть которого неизвестна. Путем цепочки преобразований находилось уравнение 4-го порядка, из которого определялась функция $f(V)$. Нетрудно повторить эти преобразования, если $f(V)$ известна:

$$\frac{dV}{d\theta} = \left(\frac{9}{40R} \right)^{1/3} (1-V^2)^{2/3}, \quad \frac{d^2 V}{d\theta^2} = -\frac{4}{3} \left(\frac{9}{40R} \right)^{2/3} V (1-V^2)^{1/3}, \quad \frac{d^3 V}{d\theta^3} = \dots \quad (29)$$

Другое решение — ограниченная стационарная волна — описывает ударный фронт с быстро затухающими осцилляциями [10]. В отличие от осцилляций на фронте решения уравнения Кортевега — де Вриза — Бюр-

герса, связанных с дисперсией, здесь колебания вызваны «явлением Фурье» — «обрезанием» высших гармоник, образующих фронт.

Еще один способ, известный на примере уравнения Бюргерса — нахождение линеаризующей замены переменных. Эти замены часто «угадываются», но и находятся методами теории групп Ли. Примеры таких, в общем-то исключительных моделей приводятся ниже.

Пример 3. Неоднородное уравнение Бюргерса

$$\frac{\partial U}{\partial T} + \Delta \frac{\partial U}{\partial \xi} - \pi \epsilon U \frac{\partial U}{\partial \xi} - D \frac{\partial^2 U}{\partial \xi^2} = F(\xi) \quad (30)$$

используется в разнообразных задачах, например для описания лазерного возбуждения интенсивного гиперзвука [11], обтекания тел и нагретых зон трансзвуковым потоком газа [12], нелинейных стоячих волн [13]. В последней задаче T — «медленное» время, описывающее эволюцию поля; ξ — «быстрое» время, описывающее осцилляции; D — параметр диссипации; Δ — расстройка от резонанса; $F(\xi)$ задает колебания одной из стенок резонатора. Уравнение (30) линеаризуется той же подстановкой Хорфа — Коула (Флорина), что и обычное уравнение Бюргерса (1):

$$U = \frac{2D}{\pi \epsilon} \frac{\partial}{\partial \xi} \ln W, \quad \frac{\partial W}{\partial T} + \Delta \frac{\partial W}{\partial \xi} - D \frac{\partial^2 W}{\partial \xi^2} = \frac{1}{2} q D \cos \xi \cdot W. \quad (31)$$

Здесь для определенности $F = -(M/2) \cdot \sin \xi$. Интересно, что средняя за период интенсивность волны выражается через собственное число λ_0 функции Матье ce_0 [11]:

$$\overline{U^2} = -\left(\frac{D}{\pi \epsilon}\right)^2 \cdot \lambda_0 \left(q = \frac{\pi \epsilon M}{2D^2}\right), \quad U = \frac{2D}{\pi \epsilon} \frac{d}{d\xi} \ln ce_0 \left(\frac{\xi}{2}, q\right). \quad (32)$$

Резонансная кривая для положительного пикового давления в профиле волны имеет конечный (из-за нелинейного затухания) максимум, сдвинутый в область $\Delta > 0$ на величину порядка \sqrt{M} . Точно так же, при нелинейном обтекании тел максимум коэффициента сопротивления оказывается конечным и сдвинутым в сверхзвуковую область [12].

Пример 4. Уравнение для волновых пучков в кубично нелинейных недиспергирующих средах [14]

$$\frac{\partial}{\partial \tau} \left[\frac{\partial p}{\partial x} + \gamma p^2 \frac{\partial p}{\partial \tau} - \beta \frac{\partial^2 p}{\partial \tau^2} \right] = \frac{c}{2} \Delta_{\perp} p \quad (33)$$

используется, в частности, для описания дифракции сильно искаженных волн и их самофокусировки. Если дифракция слабая, гармоническая волна превращается в пилообразную, каждый период которой имеет трапециевидную форму и содержит два ударных фронта: сжатия и разрежения.

В приближении нелинейной геометрической акустики уравнение (33) сводится к системе

$$\frac{\partial I}{\partial x} + \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) (I \cdot V) = -\alpha \omega \gamma I^2, \quad \frac{\partial V}{\partial x} + V \frac{\partial V}{\partial r} = \gamma c \frac{\partial I}{\partial r} \quad (34)$$

для средней интенсивности $I = \langle p^2 \rangle$ и наклона лучей $V = \partial \psi / \partial r$ к оси пучка. Первое линейризуется заменой $I = B^{-1}$; для параболического волнового фронта

$$\psi(x, r, t) = \varphi(x, t) + \frac{r^2}{2} \frac{\partial}{\partial x} \ln f(x, t) \quad (35)$$

решение имеет вид

$$I = \frac{1}{f^2(x)} I_0 \left(\frac{r}{af} \right) \left[1 + \alpha \omega \gamma I_0 \left(\frac{r}{af} \right) \int_0^x \frac{dx'}{f^2(x')} \right]^{-1}. \quad (36)$$

Функция f , которая описывает изменение амплитуды волны и ширины пучка, удовлетворяет сложному нелинейному уравнению

$$f^3 \frac{d^2 f}{dx^2} + \frac{1}{2x_d x_s} \left[1 + \frac{\alpha}{x_s} \int_0^x \frac{dx'}{f^2(x')} \right]^{-2} = \frac{1}{x_d^2}. \quad (37)$$

Граничные условия при $x = 0$: $f = 1$, $df/dx = R^{-1}$, где R — радиус кривизны фронта. Если правая часть (37) равна нулю, эта задача Коши имеет точное решение:

$$f(x) = \left(1 + \frac{x}{R} + \delta_1 \frac{x}{x_s} \right)^{\delta_2} \cdot \left(1 + \frac{x}{R} - \delta_2 \frac{x}{x_s} \right)^{\delta_1}, \quad \delta_{1,2} = \frac{\sqrt{\alpha^2 + 2x_s/x_d} \pm \alpha}{2}. \quad (38)$$

Оно получено так. Вначале угадано точное частное решение $f = \exp(-x/\alpha x_d)$, которое затем обобщено (38) с помощью «линзового преобразования», использующего свойство симметрии уравнения (33). Недостаток решения (38) в том, что описываемый им пучок самофокусируется в точку на расстоянии $x_{sf} = (\delta_2/x_s - 1/R)^{-1}$. Для устранения сингулярности нужно учесть дифракцию, т. е. правую часть (37). Удивительно, что и в этом случае есть точное решение. Его удастся найти, линейризуя (37) заменой переменных:

$$q = f^{-1}, \quad \xi x_d = x_s/\alpha + \int_0^x f^{-2}(x') dx', \quad \frac{d^2 q}{d\xi^2} + \left(1 - \alpha^{-2} (x_s/x_d) \frac{1}{\xi^2} \right) \cdot q = 0. \quad (39)$$

Решение линейного уравнения дается цилиндрическими функциями

$$f(\xi) = \frac{\xi^{-1/2}}{C_1 J_\nu(\xi) + C_2 Y_\nu(\xi)}, \quad \nu = \frac{1}{2} \sqrt{1 + \frac{2x_s}{\alpha^2 x_d}}. \quad (40)$$

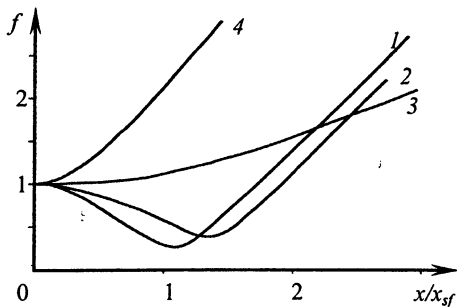


Рис. 2

Кривые 1—4 на рис. 2 соответствуют параметру $x_g/x_d = 0,01, 0,1, 0,5, 2$. При слабой дифракции пучок заметно сужается (кривые 1 и 2), а при сильной дифракции самофокусировка не проявляется (кривые 3 и 4). В отсутствие дисперсии самофокусировка не приводит к сильному росту амплитуды. Пучок заметно сужается и имеет нелинейную перетяжку, но фактор усиления невелик из-за принципиально неустранимого поглощения на ударных фронтах «пилы». Наибольшего усиления в фокусе 1,65 удается достичь при отношении длин нелинейности и дифракции, равном 0,06.

Обсудим теперь модели, которые удалось не линеаризовать, но существенно упростить и получить точные решения.

Пример 5. Рассмотрим ансамбль слабых ударных волн (рис. 3), имеющих случайное расположение и «амплитуду». Каждой «ступеньке» сопоставим частицу, масса и скорость которой даются формулами

$$m_i = u_{i+1} - u_i, \quad v_i = -(\varepsilon/c^2) \cdot (u_{i+1} + u_i). \quad (41)$$

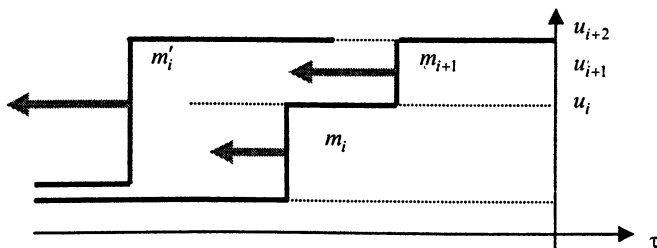


Рис. 3

При слиянии двух ступенек образуется третья; при этом сохраняются «масса» и «импульс»:

$$m'_i = m_i + m_{i+1}, \quad m'_i v'_i = m_i v_i + m_{i+1} v_{i+1}. \quad (42)$$

Аналогия с неупругими частицами позволяет получить кинетическое уравнение [15]

$$\frac{\partial G}{\partial x} - \frac{\varepsilon}{2c^2} (m + \langle m \rangle) \frac{\partial G}{\partial x} = \frac{\varepsilon}{2c^2} [*], \quad (43)$$

$$[*] \equiv \int_0^m G(x, t, \xi) \cdot G(x, 0, m - \xi) \cdot d\xi - (m - \langle m \rangle) \cdot G \cdot \int_0^\infty G(x, 0, \xi) \cdot d\xi.$$

Здесь G — вероятность того, что промежуток времени между двумя соседними разрывами равен t и величина 2-го из них равна m . Нелинейные члены (43) образуют интеграл столкновений. Совершая замену переменной и преобразуя новую функцию по Лапласу

$$G = \frac{1}{t_0} \exp\left(-\frac{t}{t_0} - \frac{\varepsilon x}{c^2 t_0} m\right) \cdot F(x, t, m), \quad f = \int_0^\infty F \exp(-sm) dm, \quad (44)$$

получим уравнение типа простых волн, которое решается точно:

$$\frac{\partial f}{\partial x} + \frac{\varepsilon}{c^2 t_0} f \frac{\partial f}{\partial s} = 0. \quad (45)$$

Аналогичный подход применим к пилообразной волне — ансамблю треугольных ударных импульсов со случайными пиковым значением и длительностью. Нелинейное затухание эквивалентно «испарению» (потерям «массы») частиц между их соударениями.

Пример 6. Систему уравнений нелинейной геометрической акустики

$$\frac{\partial u}{\partial x} - \frac{\varepsilon}{c_0^2} u \frac{\partial u}{\partial T} + \alpha \frac{\partial u}{\partial y} + \beta \frac{\partial u}{\partial z} + \left(\frac{\partial \alpha}{\partial y} + \frac{\partial \beta}{\partial z} \right) \frac{u}{2} = 0, \quad (46)$$

$$\frac{\partial \alpha}{\partial x} + \alpha \frac{\partial \alpha}{\partial y} + \beta \frac{\partial \alpha}{\partial z} = 0, \quad \frac{\partial \beta}{\partial x} + \alpha \frac{\partial \beta}{\partial y} + \beta \frac{\partial \beta}{\partial z} = 0, \quad (47)$$

где $\alpha = \partial \psi / \partial y$, $\beta = \partial \psi / \partial z$ — наклоны лучей, удается свести [16] к точно решаемому уравнению простых волн

$$\frac{\partial u}{\partial x} + \frac{u}{2} \frac{d}{dx} \ln S - \frac{\varepsilon}{c_0^2} u \frac{\partial u}{\partial T} = 0 \quad (48)$$

путем перехода к лагранжевым координатам $\xi = y - \alpha x$, $\eta = z - \beta x$, где S — якобиан. При этом переменные ξ, η становятся параметрами: по ним производных нет. Заметим также, что из уравнения переноса (46) удастся исключить время, считая профиль пилообразным. После этого его удастся линеаризовать.

Пример 7. В нелинейном резонаторе нужно рассматривать две встречные, искажающиеся и взаимодействующие волны. Но для перио-

дических во времени возмущений можно ограничиться суперпозицией: учесть «самовоздействие», а «перекрестными» (нерезонансными) взаимодействиями пренебречь [13]. При этом поле в резонаторе описывается нелинейным функциональным уравнением

$$F\left[\omega t + kL - \frac{\varepsilon}{c}kLF\right] - F\left[\omega t - kL + \frac{\varepsilon}{c}kLF\right] = A \sin(\omega t), \quad (49)$$

в котором F — неизвестная. Однако в практически интересном случае, когда длина резонатора мала по сравнению с нелинейной длиной, частота близка к частоте одной из мод, а затухание на длине L резонатора мало, (49) сводится к точно решаемому уравнению (30).

Обсудим в заключение подход к решению нелинейных уравнений путем «погружения» в класс более общих (и сложных) моделей [9]. Опыт, казалось бы, противоречит идее усложнения: «Чтобы упростить задачу, мы должны до известных пределов идеализировать свойства системы, сделать ряд упрощающих предположений...» (А. А. Андронов, А. А. Витт, С. Э. Хайкин «Теория колебаний»). Но за внешними различиями скрыто единство двух подходов. Ясно, что более симметричная модель должна иметь больше точных решений. Как добиться большей симметрии? Можно идти по традиционному пути упрощения, «отсекая» элементы модели, нарушающие ее симметрию (пренебрегая некоторыми членами уравнения или как-то видоизменяя их). С другой стороны, можно «достроить» модель до более симметричной путем ее усложнения. Если сложная модель допускает подходящее точное решение, упрощения можно произвести в конечных формулах.

Пример 8. Рассмотрим уравнение

$$\frac{\partial}{\partial t} \left[\frac{\partial u}{\partial x} - u \frac{\partial u}{\partial t} \right] = -\beta u, \quad (50)$$

моделирующее нелинейные процессы с низкочастотной дисперсией (в том числе, дифрагирующие пучки). Это уравнение допускает 3-мерную алгебру Ли с базисом

$$X_1 = \frac{\partial}{\partial t}, \quad X_2 = \frac{\partial}{\partial x}, \quad X_3 = t \frac{\partial}{\partial t} - x \frac{\partial}{\partial x} + 2u \frac{\partial}{\partial u}. \quad (51)$$

Класс инвариантно-групповых решений ограничен бегущими волнами (они строятся с помощью генераторов переноса X_1, X_2) и автомодельными решениями (генератор растяжения X_3). Поэтому, используя идею погружения, рассмотрим обобщающую модель

$$\frac{\partial}{\partial t} \left[\frac{\partial u}{\partial x} - Q(x, u) \frac{\partial u}{\partial t} \right] = F(x, u). \quad (52)$$

Чтобы установить вид функций Q, F , при которых (52) обогащается симметриями и, следовательно, нетривиальными точными решениями, нужно

воспользоваться понятием алгебры эквивалентности, провести вычисления с помощью инфинитезимальной техники Ли (при этом Q, F рассматриваются как новые независимые переменные наряду с физическими переменными t, x, u). С помощью теоремы Н. Х. Ибрагимова о проекциях найдена конкретная форма (52):

$$\frac{\partial}{\partial t} \left[\frac{\partial u}{\partial x} - \Phi_1 \left(\frac{u}{x-k} \right) \frac{\partial u}{\partial t} \right] = \frac{1}{x-k} \Phi_2 \left(\frac{u}{x-k} \right) \quad (53)$$

и показано, что (53) допускает, наряду с очевидным генератором переноса по времени, дополнительный оператор

$$X' = t \frac{\partial}{\partial t} + (x-k) \frac{\partial}{\partial x} + u \frac{\partial}{\partial u}. \quad (54)$$

Частный случай линейных функций Φ_1, Φ_2 дает точное решение для сфокусированных волновых пучков в квадратично-нелинейной среде.

Пример 9. Уравнение Ирншоу

$$\frac{\partial^2 v}{\partial t^2} - c^2 \left(1 + \frac{\partial v}{\partial \xi} \right)^{-2\epsilon} \frac{\partial^2 v}{\partial \xi^2} = 0 \quad (55)$$

в лагранжевом представлении описывает одномерное движение сжимаемого газа. Оно линеаризуется преобразованием годографа $x = v_\xi, y = v_t, \xi = X(x, y), t = u(x, y)$; теперь координата и время — функции новых независимых переменных — первых производных от искомой функции. Линейное уравнение и его упрощенная (для малых чисел Маха) версия имеют вид

$$\frac{\partial^2 u}{\partial x^2} = c^2 (1+x)^{-2\epsilon} \frac{\partial^2 u}{\partial y^2} \approx c^2 [1 - 2\epsilon x] \frac{\partial^2 u}{\partial y^2}. \quad (56)$$

Однако уравнение (56) не богато симметриями и точно не решается. Зато более общее уравнение

$$\frac{\partial^2 u}{\partial x^2} - c^2 \psi^2(x) \frac{\partial^2 u}{\partial y^2} = 0 \quad (57)$$

имеет максимально широкую группу симметрии при $\psi = [l + sx]^{-2}$. Решая (57), полагая $l=1, s=\epsilon/2$ и разлагая в ряд по малым x , найдем точное решение (56) (см. [9]).

Автор благодарен В. Ф. Ковалеву за полезные обсуждения и замечания.

Литература

1. Руденко О. В., Солуян С. И. Теоретические основы нелинейной акустики. М.: Наука, 1975.

2. *Виноградова М. Б., Руденко О. В., Сухоруков А. П.* Теория волн. 2-е изд. М.: Наука, 1990.
3. *Руденко О. В., Солуян С. И.* Некоторые нестационарные задачи теории волн конечной амплитуды в диссипативных средах // ДАН. 1970. Т. 190, № 4. С. 815—818.
4. *Гурбатов С. Н., Саичев А. И., Якушкин И. Г.* Нелинейные волны и одномерная турбулентность в средах без дисперсии // УФН. 1983. Т. 141. С. 221—255.
5. *Руденко О. В.* Взаимодействия интенсивных шумовых волн // УФН. 1986. Т. 149, № 3. С. 413—447.
6. *Ибрагимов Н. Х.* Азбука группового анализа // Математика, кибернетика. 1989. № 8. С. 1—47; Опыт группового анализа обыкновенных дифференциальных уравнений // Математика, кибернетика. 1991. № 7. С. 1—47.
7. CRC Handbook of Lie group analysis of differential equations. In 3 vol. / *N. H. Ibragimov* (ed.). Boca Raton, Florida: CRC Press, 1994—1996.
V. 1: Symmetries, exact solutions and conservation laws, 1994.
V. 2: Applications in engineering and physical sciences, 1995.
V. 3: New trends in theoretical developments and computational methods, 1996.
8. *Enflo B. O., Rudenko O. V.* To the Theory of Generalized Burgers' Equations // Acta Acustica. 2002. V. 88. P. 155—162.
9. *Ибрагимов Н. Х., Руденко О. В.* Принцип априорного использования симметрий в теории нелинейных волн // Акуст. журн. 2004. Т. 50, № 4. С. 406—419.
10. *Руденко О. В., Робсман В. А.* Уравнение нелинейных волн в рассеивающей среде // ДАН. 2002. Т. 384, № 6. С. 755—759.
11. *Руденко О. В.* О возможности генерации мощного гиперзвука с помощью лазерного излучения // Письма ЖЭТФ. 1974. Т. 20, № 7. С. 445—448.
12. *Карабутов А. А., Руденко О. В.* Модифицированный метод Хохлова для исследования нестационарных трансзвуковых течений сжимаемого газа // ДАН СССР. 1979. Т. 248, № 5. С. 1082—1085.
13. *Enflo B. O., Hedberg C. M., Rudenko O. V.* Resonant properties of a nonlinear dissipative layer excited by a vibrating boundary: Q-factor and frequency response // J. Acoust. Soc. Am. 2005. V. 117, № 2. P. 601—612.
14. *Руденко О. В., Сапожников О. А.* Явления самовоздействия пучков волн, содержащих ударные фронты // УФН. 2004. Т. 174, № 9. С. 973—989.
15. *Руденко О. В., Хохлова В. А.* Кинетика одномерных пилообразных волн // Акуст. журн. 1991. Т. 37, № 1.
16. *Руденко О. В., Гурбатов С. Н., Хедберг К.* Нелинейная акустика в задачах и примерах М.: Физматлит, 2006.

АКТУАЛЬНЫЕ ПРОБЛЕМЫ, СВЯЗАННЫЕ С НЕЛИНЕЙНОЙ АКУСТИКОЙ

О. В. Руденко

На предшествующей школе «Нелинейные волны-2004» был прочитан цикл из трех лекций по нелинейной акустике [1]. В качестве лекторов последовательно выступили О. В. Руденко, С. Н. Гурбатов и Л. А. Островский. Поскольку вторая и третья лекции были посвящены конкретным проблемам, автор первой посчитал уместным сделать популярное введение, с тем чтобы подготовить аудиторию к более сложной информации своих коллег. При этом пришлось немного сократить «научную» часть в пользу учебно-методического материала.

Однако А. В. Гапонов-Грехов заметил, что именно во вводной лекции целесообразно дать обзор современных проблем и тенденций. Спустя два года настоящая статья готовилась на основе именно этих пожеланий: в ней рассказано о направлениях, привлечших наибольшее внимание в последние годы.

При отборе материала оказалось трудным уйти от стандартной формулы: «Актуально то, чем сейчас занимаюсь я, мои сотрудники и друзья из Нижнего». Для оценки актуальности пришлось выбрать другие критерии: 1) число публикаций по данной проблеме; 2) активность на международных конференциях (специальные секции, симпозиумы, круглые столы и т. п.); 3) «шум» в прессе (научно-популярные статьи и реакция на них); 4) число объявленных программ и грантов; усилия, затраченные на получение финансовой поддержки. По этим критериям, согласно нашей субъективной оценке, лидируют следующие проблемы.

1. Звуковой удар от тел, движущихся со сверхзвуковой скоростью. Явления при формировании N -волны в атмосфере. Выбросы, их экологические последствия. Удар и перспективы сверхзвуковой пассажирской авиации.

2. «Сонофьюжн» («акустический термояд») — проблема реализации термоядерного синтеза в коллапсирующем пузырьке при кавитации. «Нашумевшая» история (после 2002 года) и давние работы, с более скромными задачами — наблюдать не поток нейтронов, а излучение более коротковолновое, чем при сонолюминесценции одиночного пузырька. Явление неограниченной кумуляции.

3. Известное в общих чертах, но до сих пор не понятое свойство скелетных мышц «затягивать» и демпфировать удары. Нелинейная динамика мышечного сокращения. Данные по измерениям сдвиговой упругости мышц, новые модели.

4. Необычные нелинейные эффекты при распространении волн в геофизических структурах.

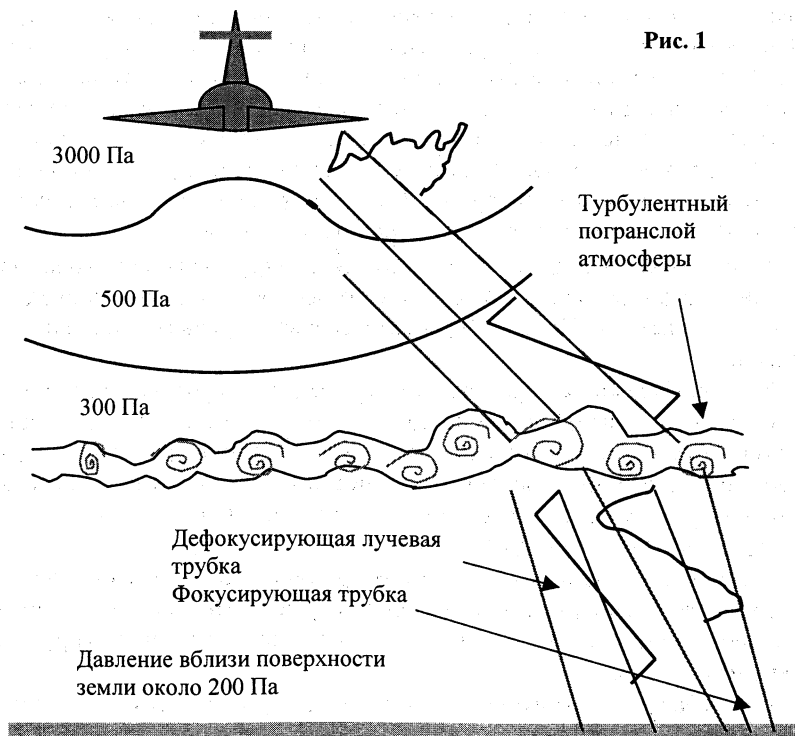
5. Новые результаты и приложения в области медицинского ультразвука, где нелинейность — вполне обычное явление.

Время, отведенное на лекцию, позволило коротко рассказать только о проблемах 1, 2 и 3. Проблема 4 опущена; с ней можно ознакомиться по журналу УФН [2], где опубликованы материалы соответствующей сессии Отделения физических наук РАН. Текущее состояние дел по «вечно актуальной» медицинской проблеме 5 неоднократно докладывалось с учетом новых результатов; в последний раз — на устном выпуске УФН [3] в честь 90-летия В. Л. Гинзбурга.

1. Звуковой удар

При движении со сверхзвуковой скоростью летательный аппарат возбуждает в среде сильную звуковую волну. На расстояниях порядка десятков метров от самолета акустические давления составляют обычно 3—5 кПа (рис. 1).

Волна во времени имеет сложный вид, определяемый формой обтекаемого профиля. На больших расстояниях, если полет происходит в од-



народной среде, вдоль прямой и с постоянной скоростью, волновые фронты приобретают вид конусов Маха, а импульс принимает универсальную *N*-образную форму. Его передний и задний фронты образованы слабыми ударными волнами (типичные пиковые давления здесь порядка 100—200 Па).

Качественная картина формирования *N*-волны изображена на рис. 2.

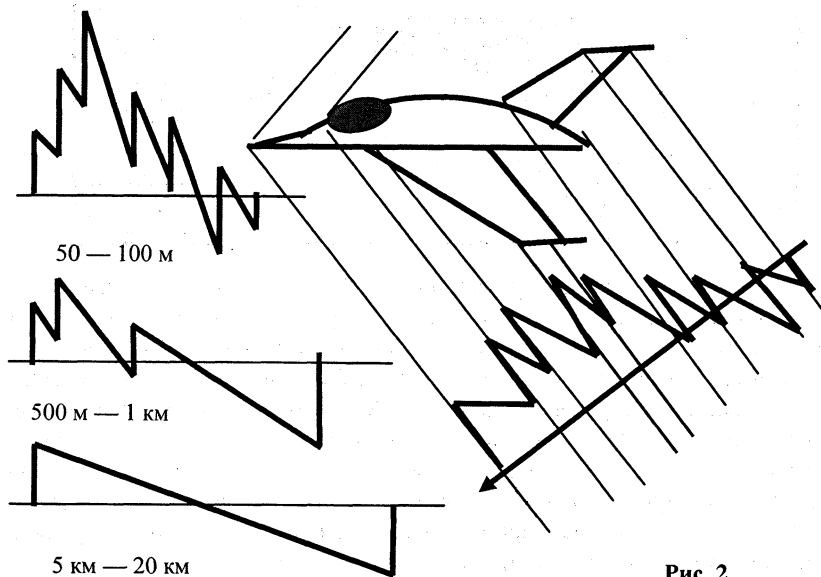


Рис. 2

В ближней зоне наблюдается ансамбль ударных фронтов, каждый из которых возникает благодаря «излому» функции, описывающей аэродинамический профиль. В системе координат, движущейся со скоростью звука, фронты перемещаются навстречу друг другу и сталкиваются по закону парного соударения абсолютно неупругих частиц. В промежутках между соударениями происходит нелинейное затухание ударных волн, которое на языке аналогии с частицами можно интерпретировать как их «испарение». В результате многих столкновений образуются две «крупные частицы»: передний и задний ударные фронты. Этот процесс («акустическая турбулентность») описывается нелинейными кинетическими уравнениями [4, 5]. Его асимптотические свойства также хорошо изучены [6].

Обратим внимание на то, что акустическая (потенциальная) турбулентность вблизи самолета бывает настолько сильной, что приводит к оптическим эффектам, подобно сильной гидродинамической (вихревой) турбулентности (рис. 3).

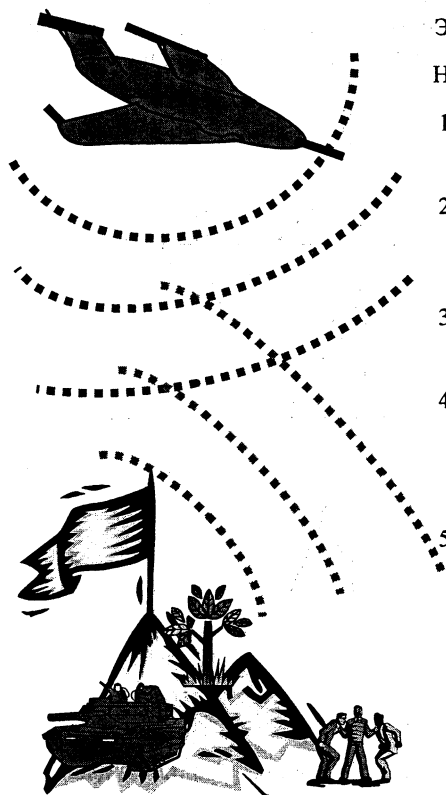


Вихревая турбулентность, наиболее существенная в приземном пограничном слое атмосферы (см. рис. 1), формирует области локальных фокусировок. В последних возникают флуктуационные выбросы пикового давления. Такие «сверхудары» содержат большие градиенты давления, опасные для людей и животных. По этой причине сверхзвуковые пассажирские самолеты (СПС) «Конкорд» много лет эксплуатировались только на трассах, проложенных над океаном. Небольшие участки континентальных трасс пролегли над малонаселенной местностью; при этом использовались режимы маневрирования, не приводящие к пространственно-временной концентрации акустической энергии.

В конце 80-х гг. в США был сделан анализ тенденций развития сверхзвуковой авиации, который показал, что примерно к 2020 г. основной недостаток первых СПС типа «Конкорд» и ТУ-144 — их низкая экономичность — будет преодолен благодаря появлению новых высокоэффективных двигателей. Таким образом, основной проблемой, препятствующей широкому использованию СПС, станет проблема достижения приемлемых уровней звукового удара на местности.

Эти соображения вызвали в начале 90-х гг. новую волну активности в изучении звукового удара (первая волна [7] «затухла» в 70-х гг.): выросло число публикаций, грантов и тематических конференций. Так, соответствующие секции выделены в программе многих съездов Акустического общества США, проводимых 2 раза в год. Волна не угасает до сих пор. На последнем, 17-м симпозиуме по нелинейной акустике (Пенн Стейт, 2005) прошел симпозиум-сателлит по звуковому удару. Наряду с важными новыми результатами (прежде всего, данными натурных и лабораторных экспериментов) в зарубежных изданиях появилось немало работ, воспроизводящих старые результаты или содержащих фантастические предложения. Примером может служить идея сгладить ударный

фронт путем ухудшения аэродинамических характеристик: затупления передней части фюзеляжа или использования конструкций типа биплана Буземана. Если бы «черенковское» («маховское») излучение каким-то образом удалось подавить, то исчезли бы как волновое сопротивление сверхзвуковому движению, так и связанная с реакцией на излучение подъемная сила.



Экспертная система

Необходимые для учета факторы

1. СПС: аэродинамические характеристики, скорость полета, маневрирование, трасса
2. Атмосфера: стратификация, регулярные неоднородности, турбулентность, влажность, молекулярный состав, ветер
3. Волна: нелинейность, дифракция, рефракция, рассеяние, затухание, релаксация
4. Местность: рельеф, акустические свойства границ, отклик на импульсное воздействие, проникновение в океан через неровную поверхность
5. Действие на: живые организмы (человек, сухопутные и морские животные), строения, технику

Рис. 4

Российские работы середины 90-х годов были нацелены на развитие математического аппарата для адекватного описания нелинейных волн в реальной атмосфере, численных методов анализа нелинейных моделей, исследование биологических эффектов воздействия удара на живые организмы, изучение особенностей аэродинамики и режимов маневрирования на форму волны, расчет вероятности появления больших выбросов, анализ условий прохождения удара через водную поверхность и в грунт. Конечной целью виделось создание экспертной системы для выбора оптимальных режимов и трасс полета СПС. Многообразие проблем, которые было необходимо решить, иллюстрирует рис. 4.

На рис. 5 изображены формы импульсов звукового удара, часто наблюдаемые на местности. Кривая 1 представляет классическую N -образную форму; отличие гладкого участка профиля от прямой и асимметрия областей сжатия и разрежения обусловлены дифракцией. Кривыми 2 и 3 показаны «сглаженный» и «обостренный» импульсы. Кривая 4 изображает так называемую U -волну, которая формируется в результате дифференцирования N -волны в окрестности фокусов или каустик в неоднородной атмосфере, а также ее нелинейного искажения. Импульс 5 содержит два пика на переднем фронте, импульс 6 — несколько таких пиков.

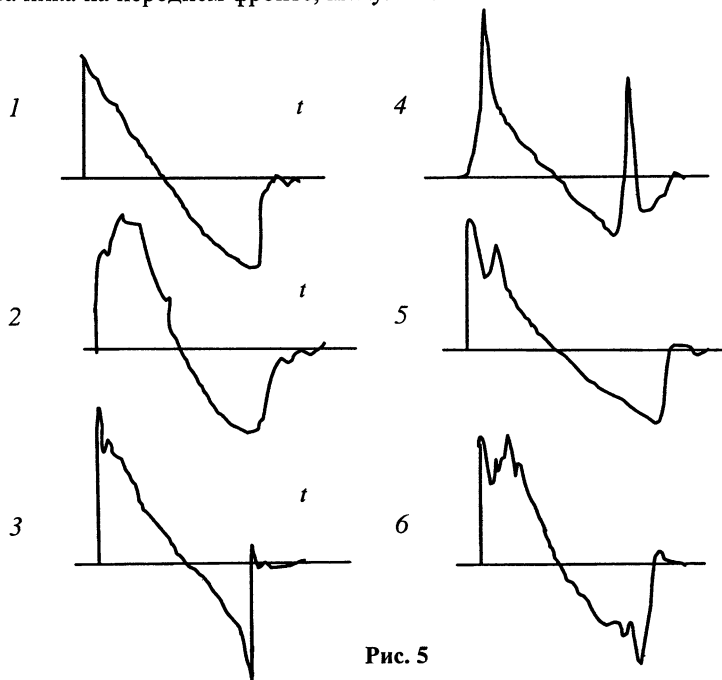


Рис. 5

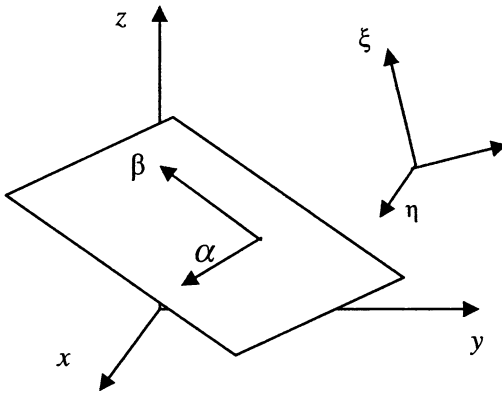
Длительность ударного фронта, определяющая градиент давления, зависит как от пикового давления, так и от многих других причин, в частности, интенсивности турбулентных пульсаций. Интересно, что среди диссипативных механизмов, влияющих на ширину фронта, основным является не вязкость, а потери из-за релаксационных процессов, идущих при акустическом возбуждении молекул газов.

Покажем несколько нелинейных моделей, использованных при расчетах волн звукового удара [5, 8]. Уравнение

$$\frac{\partial p}{\partial s} + \frac{1}{2} p \frac{\partial \ln D}{\partial s} - \frac{\varepsilon}{c^3 \rho} p \frac{\partial p}{\partial \tau} - \gamma \frac{\partial^2}{\partial \tau^2} \int_0^{\infty} K(\zeta) p(x, \tau - \zeta) d\zeta = 0 \quad (1)$$

записано в приближении нелинейной геометрической акустики.

Здесь p , ρ , c — акустическое давление, плотность и скорость звука в неоднородной среде, ε — нелинейный параметр, $\tau = t - s/c$, γ — «сила» релаксации, $K(t)$ — ядро, описывающее динамику релаксационных процессов, $D = \partial(x, y, z)/\partial(\alpha, \beta, s)$ — якобиан перехода от декартовых (x, y, z) к лучевым (α, β, s) координатам (рис. 6). Заметим, что по лагранжевым координатам (α, β) , фиксирующим точку выхода луча с исходной поверхности, дифференцирование в уравнении (1) не производится. Пучок лучей может быть широким; удастся рассмотреть произвольные фронты, задавая различные сдвиги фаз на разных лучах. Однако в области пересечения лучей уравнение (1) несправедливо.



Можно использовать уравнение типа Хохлова — Заболотской в лагранжевых координатах:

$$\frac{\partial}{\partial \tau} \left[\frac{\partial p}{\partial s} + \frac{1}{2} p \frac{\partial}{\partial s} \ln \frac{D}{c\rho} - \frac{\varepsilon}{c^3 \rho} p \frac{\partial p}{\partial \tau} - \gamma \frac{\partial^2}{\partial \tau^2} \int_0^{\infty} K(\zeta) p(x, \tau - \zeta) d\zeta \right] = \frac{c\rho}{2} \nabla_{\alpha, \beta} \left(\frac{\nabla_{\alpha, \beta} p}{\rho} \right). \quad (2)$$

Это точное уравнение. Однако его использование не всегда возможно из-за неудобства лагранжевых координат и появления особенностей в якобиане. Этих недостатков лишено уравнение, записанное в квазиоптическом приближении в окрестности выбранного осевого луча:

$$\frac{\partial}{\partial \tau} \left[\frac{\partial p}{\partial s} - \frac{1}{2} p \frac{\partial}{\partial s} \ln c\rho - \frac{\varepsilon}{c^3 \rho} p \frac{\partial p}{\partial \tau} - \gamma \frac{\partial^2}{\partial \tau^2} \int_0^{\infty} K(\zeta) p(x, \tau - \zeta) d\zeta \right] - \frac{1}{2c^2} \left[\left(\xi \nabla_{\perp} \right)^2 c \right]_{\xi=0} \frac{\partial^2 p}{\partial \tau^2} = \frac{\rho}{2} \nabla_{\perp} \left(\frac{c}{\rho} \nabla_{\perp} p \right). \quad (3)$$

Оно позволяет описать поля в окрестности фокусов и каустик, но лишь для узкого пучка лучей. При записи (3) использованы локальные криволинейные координаты, причем поперечные $\bar{\xi} = (\xi, \eta)$ (рис. 6) введены специальным образом: этот базис повернут относительно трехгранника Френе на угол, определяемый криволинейным интегралом от кручения луча.

Простейшей моделью для оценки вероятности выбросов удара на местности является модель случайного фазового экрана для турбулентного погранслоя атмосферы (см. рис. 1). Двумерная нелинейная система, состоящая из уравнений для наклонов лучей

$$\begin{aligned} \frac{\partial \alpha}{\partial z} + \alpha \frac{\partial \alpha}{\partial x} + \beta \frac{\partial \alpha}{\partial y} &= 0, & \alpha &= \frac{\partial \psi}{\partial x}, \\ \frac{\partial \beta}{\partial z} + \alpha \frac{\partial \beta}{\partial x} + \beta \frac{\partial \beta}{\partial y} &= 0, & \beta &= \frac{\partial \psi}{\partial y}, \end{aligned} \quad (4)$$

и уравнения переноса

$$\frac{\partial p}{\partial z} - \frac{\varepsilon}{\rho_0 c_0^3} p \frac{\partial p}{\partial T} + \alpha \frac{\partial p}{\partial x} + \beta \frac{\partial p}{\partial y} + \frac{1}{2} p \left(\frac{\partial \alpha}{\partial x} + \frac{\partial \beta}{\partial y} \right) = 0, \quad (5)$$

имеет точное аналитическое решение, которое использовано для анализа как статистических, так и динамических задач.

Волна, падающая на экран $z = 0$, задается в виде $p = P_0(x, y, t)$, где P_0 — произвольная функция времени, имеющая, в частности, N -образную форму. В уравнении (5) $T = t - x/c_0 - \psi(x, y, z)/c_0$, где ψ — эйконал.

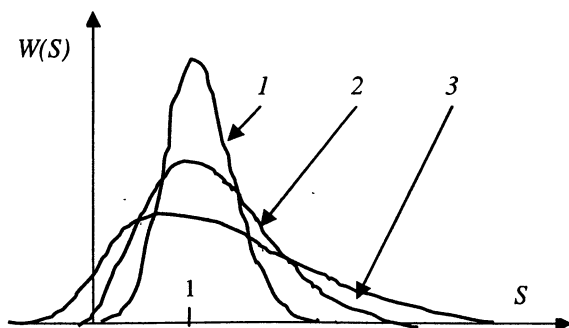


Рис. 7

На рис. 7 изображены вероятностные распределения для площади сечения лучевой трубки, рассчитанные для различных значений $l = 8\sigma z/r_0^2 = 0,4; 0,7; 1$, пропорциональных расстоянию z от экрана, а также дисперсии σ турбулентных пульсаций. Среднее сечение постоян-

но, но флуктуации с ростом l увеличиваются; это означает, что растет вероятность появления как фокусирующих, так и дефокусирующих лучевых трубок.

На рис. 8, *a* изображены рассчитанные плотности вероятности пиковых давлений. Появление «крыльев» $p/p_0 > 1$ за экраном свидетельствует о росте вероятности появления больших выбросов [9]. Теоретические результаты согласуются с данными экспериментов (см., например, [10], рис. 8, *б*), в которых наблюдалось уменьшение среднего пикового давления (кривая 1), рост средней длительности ударного фронта (кривая 2) и одновременно увеличение дисперсии флуктуаций этих параметров (кривые 3 и 4).

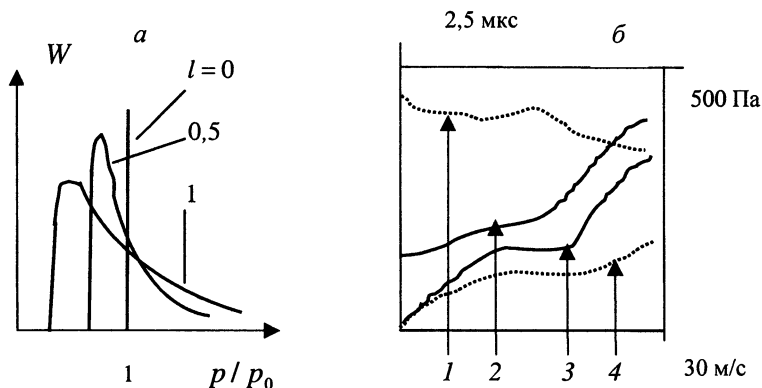


Рис. 8

Разнообразие задач, трудность их решения и прикладная значимость проблемы звукового удара позволяют сделать правдоподобный прогноз о продолжении активных исследований в этой области, по крайней мере до тех пор, пока создание СПС стоит в планах ведущих авиационных корпораций и пользуется (в ряде мест) бюджетной поддержкой.

2. «Акустический термояд»

Как известно, при слиянии некоторых легких ядер, например изотопов водорода дейтерия и трития, выделяется большая энергия. Реакция синтеза распространена в природе, будучи источником энергии звезд; в земных условиях она происходит при термоядерном взрыве. Чтобы реакция пошла, нужны температуры порядка 10^7 К. При этом частицы могут преодолеть электростатическое отталкивание и сблизиться на расстояние, при котором возможно слияние ядер. Кроме высокой температуры, для положительного выхода энергии нужно, чтобы произведение времени

жизни плазмы τ и плотности ионов n было достаточно велико, более $5 \cdot 10^{14}$ см/см³ (критерий Лоусона). Существуют большие запасы топлива для термоядерной энергетики. Дейтерий — распространенный изотоп, который может добываться из воды. Тритий может производиться в реакторе из лития. Запасы дейтерия и лития достаточны для производства энергии в течение тысяч лет. Теоретически один литр воды может дать столько же энергии, сколько 300 литров бензина.

Начиная с 1950 г. ведутся работы по реализации управляемого синтеза. Основным направлением было создание установок типа ТОКАМАК (ТОроидальная КАмера с МАгнитными Катушками) — см.: Л.А. Арцимович, С.В. Миронов, В.С. Стрелков // Атомная энергия. 1964. Т. 17. С. 170. В СССР был построен ряд установок Токамак (Т-7, Т-10, Т-15). В США создан испытательный реактор — токамак «TFTR»; на нем достигнута температура плазмы порядка 200 миллионов градусов. Оценивая перспективы этого направления, Е. П. Велихов в 2000 г. сказал примерно так: в 2011 г. будет зажжена первая плазма — сначала мы будем работать с водородом, потом с дейтерием, потом с тритием, потом, в 2020—2022 гг., начнется эксплуатация реактора с выяснением всех деталей, изучением устойчивости работы и т. д. Если все пойдет, как намечено, то к 2030 г. можно будет построить первую демонстрационную электростанцию. 21 ноября 2006 г. в Париже подписано соглашение о создании Международной организации по строительству термоядерного экспериментального реактора ИТЭР.

Второе развиваемое направление — лазерный термоядерный синтез. (Н.Г. Басов, О.Н. Крохин // ЖЭТФ. 1964. Т. 46. С. 171). Как писал В. Л. Гинзбург, суть его в следующем. Небольшую ампулу со смесью дейтерия и трития со всех сторон облучают мощными лазерными импульсами. Ампула испаряется, а световое давление сжимает ее содержимое так, что в смеси «зажигается» реакция. Обычно она проходит со взрывом, эквивалентным 100 кг тротила. Строятся гигантские установки, но о них мало известно в силу засекреченности: на них, видимо, надеются имитировать термоядерные взрывы.

В США в 2001 г. проведен эксперимент по имплозии (направленный внутрь взрыв) шарика из сжиженной смеси дейтерия и трития с помощью мощного лазера «Omega», находящегося в Ливерморской национальной лаборатории. Шарик диаметром 1 мм, в пластиковой оболочке и охлажденный до температуры -253 °С, облучался с 60 направлений импульсами продолжительностью меньше 1 нс. За это время его температура поднялась примерно до $28 \cdot 10^6$ К. При этой температуре происходило слияние ядер дейтерия и трития с выделением энергии. Однако длительности импульса было недостаточно, чтобы баланс энергии был положительным.

Третье направление — акустический термояд — связано с явлением сонолюминесценции, обнаруженным в 1934 г. Оно состоит в излучении

света при схлопывании паровых или газовых пузырьков в жидкости под действием ультразвука. Вначале акустическая волна большой интенсивности в фазе отрицательного давления разрывает жидкость. Образуется облако кавитационных пузырьков. Затем, в фазе положительного давления, пузырьки схлопываются, концентрируя запасенную энергию в очень малом объеме.

В 1998 г. Гэтан и Крам наблюдали схлопывание одиночных пузырьков, созданных в фокусе сферического резонатора. Сонолюминесценция одиночного пузырька — стабильный эффект; свет излучается серией наносекундных импульсов, повторяющихся в течение каждого периода волны. Согласно оценкам, температура излучающей свет плазмы в пузырьках не менее 25000 К.

На протяжении многих лет дискутировался вопрос: можно ли достичь при схлопывании пузырька температуры порядка 10^7 К и осуществить термоядерную реакцию? Идея даже была запатентована в США в 1982 г. [11]. Однако взрыв интереса к проблеме произошел в 2002 г., после появления публикации [12]. В эксперименте [12] использовалась рабочая жидкость — дейтерированный ацетон C_3D_6O и обычный ацетон C_3H_6O (для контроля). Частота ультразвука — 19,3 кГц, амплитуда — 15 атм. Зародыши создавались пучком нейтронов с энергией 14 МэВ; при этом порог кавитации был 7 атм. Схлопывание пузырька сопровождалось излучением нейтронов с энергией около 2,5 МэВ, что могло быть следствием реакции $D + D \rightarrow {}^3He + n$. Удалось зарегистрировать повышение уровня трития — след реакции $D + D \rightarrow t + p$.

Высказанные сомнения основывались на следующем: а) для создания зародышей кавитации использовался пучок нейтронов, которые после замедления в жидкости могли быть приняты за нейтроны реакции синтеза; б) сразу же предпринятые попытки повторить эксперимент были безуспешными. После бурных дискуссий (см.: Science, № 295, 8 марта 2002, с. 1793; D. Kennedy «To publish or not to publish»; C. Seife «Bubble fusion paper generates a tempest in a beaker», с. 1808) статью [12] опубликовали, хотя многие относились к предмету скептически. Появился ряд статей, например, B. G. Levi «Skepticism greets claim of bubble fusion», Physics Today, April 2002, p. 16—184; T. D. Rossing «Sonofusion?», Echoes, v. 12, №2, Spring, 2002. Однако уничтожающей публичной критики не было. Как говорили некоторые известные специалисты, «практически значимого эффекта, видимо, нет, эксперимент неточный, но шум вокруг этой проблемы поможет получить финансирование».

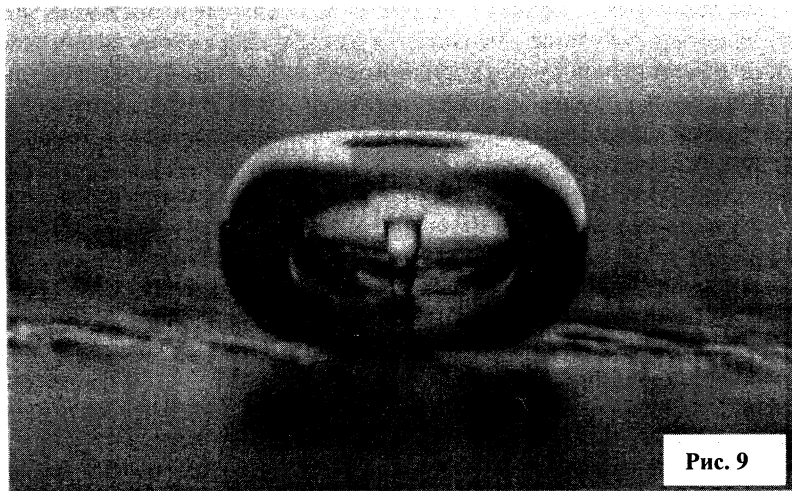
Для оценки достоверности этих результатов на Международном симпозиуме по нелинейной акустике (Москва, МГУ, август 2002), «по горячим следам» была организована специальная секция по проблеме «сонофьюжн» (ее вел L. Crum). Присутствовали специалисты как по ди-

намике пузырьков (Р. И. Нигматулин, В. В. Кедринский, С. А. Рыбак и др.), так и по ядерной физике (А. М. Дыхне, группы из НИИЯФ МГУ и ТРИНИТИ, ОИФЗ). Дискуссии на сессии и вокруг нее были интересными, но ни «закрыть» проблему, ни оценить ее как реальную не удалось. Более того, оказалось, что и у нас проблема имеет давнюю и интересную историю.

Еще в 1973 г. К. П. Станюкович и О. Б. Хаврошкин подали две заявки на открытие в схлопнувшейся кавитационной полости рентгеновского излучения и импульсной термоядерной реакции [13]. Позднее М.А. Маргулису был выдан патент [14]. Имеется ряд публикаций в солидных журналах (см., например, [15]).

Но это не вся история. Еще в конце 40-х гг. Г. И. Покровский думал об использовании кавитации для инициирования термояда. Он высказал эту идею студенту МФТИ В. А. Белоконю в начале 50-х гг., а в 1955 г. предложил эту тему в качестве дипломной работы (частное сообщение). Он выражал оптимизм относительно инициирования дейтерия при отсутствии газа или пара в полости, т. е. в модели «холодного» разгона стенок внутрь, хотя отчетливо понимал проблемы, возникающие в связи с реальным присутствием газа. Возбудить кавитацию он предлагал тривиальным способом — за счет кипения тяжелой воды в контейнере со стенками, содержащими соединение серебра, эффективно реагирующее на нейтроны; многочасовое кипение дало бы возможность подтвердить факт термояда. Именно Г. И. Покровский был одним из учителей Е. И. Забабахина, классика советских прикладных исследований в области термояда и его инициирования за счет явлений кумуляции. У специалистов есть основания считать, что в области «неограниченной кумуляции» (НК) (НК — термин Е. И. Забабахина) сделано и понято очень многое, однако эти результаты не опубликованы. С глубокими (но, возможно, известными не в полной мере) соображениями Е. И. Забабахина можно познакомиться в его открытых публикациях [16]. Вот выдержка из них: «Явления НК — это процессы, в которых в каком-то виде происходит концентрация энергии, и притом неограниченно сильно». Примером НК является схлопывание сферического пузырька; это нелинейная задача с подвижной границей. Для невязкой несжимаемой жидкости (Рэлей) скорость движения поверхности неограниченно растет при приближении к фокусу. Когда радиус пузырька уменьшается с 1 см до 1 мм, давление вблизи границы возрастает с 1 атм до 150000 атм. Что может устранить особенность в фокусе? НК «убивается» вязкостью только для очень маленьких пузырьков (Е. И. Забабахин), исходный радиус которых при 1 атм менее 0,8 мкм (вода). Учет сжимаемости проведен в работах Я. Б. Зельдовича, И. М. Гельфанда, К. Хантера и других выдающихся исследователей; показано, что явление НК сжимаемостью не устраняется. Следовательно, НК — это свойство явления схлопывания, а не следствие идеализации

задачи. Е. И. Забабахин в 1965 г. сформулировал «гипотезу о неустойчивости кумуляции»: «Всякая неограниченная кумуляция неустойчива; неустойчивость не только видоизменяет ее, но и устраняет вообще (из неограниченной делает ограниченной)... Это соображение интуитивно». Что могло бы устранить НК? Несферичность (асимметрия) сжатия, проявляющаяся у больших пузырьков, в которой концентрация энергии может быть максимальной. (Пример сжатия пузырька вблизи границы, с образованием высокоскоростной кумулятивной струи, дан на рис. 9 — L. Crum). Кроме того, несправедливость приближения сплошной среды при малых размерах перед схлопыванием, особенности в уравнении экстремального состояния пара в полости, кинетические явления.



Что могло бы помочь развитию НК? Ударная волна (УВ), бегущая по газу перед стенкой полости и сжимающая газ. Многократные переотражения УВ от центра и движущейся стенки, которая усиливает УВ. Вопрос о НК в кавитационном пузырьке открыт. Вдохновляющий пример — «черные дыры» в астрофизике.

И вот совсем «свежие» результаты. В работе [17] проведен детальный численный расчет с ответами на критику идеи и первого эксперимента [12]. Расчет, в частности, для финальной стадии коллапса (диаметр пузырька 100 нм) дает оценку температуры 100 млн °С и плотность 10 г/см³. Проведен новый эксперимент [18]. В смеси дейтерированный бензол + дейтерированный ацетон с растворенной солью урана-238 обнаружено заметное возрастание эмиссии гамма-лучей и нейтронов (до 10⁴ в секунду) с энергией 2,45 МэВ. Эффекты исчезали в контрольных недейтерированных жидкостях, при отсутствии кавитации и в тяжелой воде.

...События развиваются. Интрига не ослабевает.

3. Нелинейность и динамические свойства скелетных мышц

Известно, что спортсмены умеют «гасить» удар, напрягая группу мышц в районе контакта. Этот прием используется борцами при падении на ковер, футболистами при приеме мяча головой, каратистами, выдерживающими очень сильный удар. Способность скелетных мышц защищать суставы от травм за счет рассеяния и поглощения волновой энергии сформировалась в результате эволюции. По-видимому, в процессе адаптации части океанских обитателей к жизни на суше появилась необходимость обезопасить организмы от ударных нагрузок, вызванных бегом, прыжками, а также активностью, связанной с охотой, схватками самцов и другими действиями в рамках межвидовой и внутривидовой конкуренции. Вероятность травмы уменьшается за счет напряжения мышц в момент, предшествующий удару; при этом субъект как бы «управляет» диссипативными и упругими характеристиками своих мышц, повышая их эффективную динамическую вязкость и «растягивая» время ударного воздействия.

Вязкоупругие свойства мышечной ткани, составляющей 40% массы тела, изучены недостаточно. Разброс измерений модуля Юнга: от 1 кПа до 1000 кПа. Мало данных по сдвиговому модулю упругости — наиболее информативному параметру при диагностике патологических изменений в мягких тканях. Еще меньше данных по динамической вязкости мышц. Однако недавно наблюдалось [19] интересное явление: значительный (на 2 порядка) рост поглощения волны, усиливающейся с ростом напряжения мышцы. Это явление пока не имеет объяснения; нет и физической модели, адекватно описывающей процесс.

Акустические параметры биотканей определяются силами взаимодействия между молекулами, а не клеточной и надклеточной структурой. Однако ткани различаются по составу гораздо меньше, чем по структуре. Большинство мягких тканей содержат до 80% воды. Белки, а также иные органические и низкомолекулярные компоненты — это в основном одни и те же вещества, их концентрация слабо дифференцирована. Поэтому модуль всестороннего сжатия и скорость звука примерно одинаковы (с точностью 10%) для всех мягких тканей. Столь же малы и вариации акустической нелинейности.

Напротив, пределы изменения структурных особенностей (геометрических параметров клеток, степени неоднородности и анизотропии) несравнимо шире. Поэтому модули сдвига и Юнга, сдвиговая вязкость отличаются в разных тканях не на единицы процентов, а на порядки. Модули изменяются в пределах трех-четырех порядков; даже для одной ткани изменения достигают тысяч процентов, например, при росте опухоли или при обычном сокращении мышц. Поэтому в последние годы возникло направление, основанное на использовании сдвиговых акустических (СА) волн для диагностики мышечных патологий [20, 21]. Примером

физиологического процесса, который важно характеризовать СА-методами, является саркопения — уничтожение мышечных клеток при старении. В последние годы стала ясна важнейшая роль саркопении в медицинских проблемах стареющего организма. Саркопения, в дополнение к остеопорозу, является одним из факторов роста числа переломов у пожилых людей. По данным вибромиографии, остеопения — это не только потеря мышечной массы, но и деградация функциональных характеристик мышечных волокон, которую также можно оценить СА-методами. Потенциальные области применения СА-методов включают также: патологии, связанные с мышечной атрофией и дистрофией при невралгических заболеваниях, миопатиях, иммобилизации, длительном постельном режиме и т. п.; это также геронтология; спортивная и космическая медицина.

Мышечное волокно состоит примерно из 10000 последовательных саркомеров (рис. 10), каждый из которых, содержит 10^6 нитей: тонких (белок актин) и толстых (миозин). Пары тонких и толстых нитей параллельны и контактируют через «миозиновые мостики». Эти мостики расположены упорядоченно на толстых нитях. На рис. 10, а показана схема саркомера и его сокращения: решетки нитей вдвигаются друг в друга. На рис. 10, б выделено сечение саркомера по АА, показывающее его гексагональную симметрию. Каждый мостик может находиться в двух состояниях: разомкнутом и замкнутом. Мостик работает циклически: замыкается, развивает силу на некотором интервале относительного перемещения нитей и затем размыкается. В. И. Дещеревский предложил модель [22], основанную на следующих предположениях.

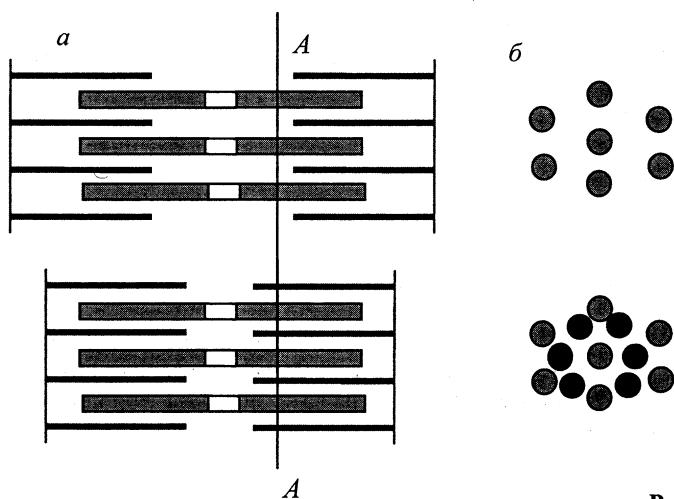


Рис. 10

Замыкание мостика (образование связи в паре актин-миозин в точке $\xi = 0$) приводит к скольжению места его присоединения к актину вдоль толстой нити к новому равновесному положению $\xi = \delta$. При этом замкнутый мостик развивает на расстоянии δ положительную (для актина) силу, направленную к центру саркомера. При $\xi = \delta$ сила проходит через ноль; тянущий мостик становится тормозящим. Тормозящий мостик замыкается и мгновенно возвращается в исходное состояние $\xi = 0$. Уравнения [22], модифицированные для акустических задач, имеют вид

$$\begin{aligned} \frac{dn}{dt} &= k_1 \alpha_0 \beta l - k_1(n+m) - \frac{n}{\delta} \frac{dl}{dt}, \\ \frac{dm}{dt} &= -k_2 m + \frac{n}{\delta} \frac{dl}{dt}, \quad P(t) = (n-m)f. \end{aligned} \quad (6)$$

Здесь $\gamma = \alpha(l) - n - m$ — число свободных мостиков, n , m — числа тянущих и тормозящих мостиков, $\alpha(l) = \alpha_0 \beta l$ — число всех активных мостиков, зависящее от укорочения l половинки саркомера, где β — константа укорочения, P — приложенное к мышце напряжение. Константа k_1 характеризует скорость замыкания свободных мостиков, k_2 — скорость размыкания замкнутых мостиков.

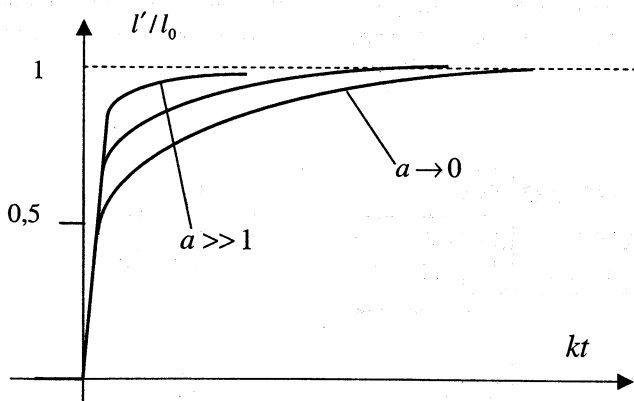


Рис. 11

Точное (при $k_1 = k_2 = k$) решение (6), моделирующее «быстрые» импульсные нагрузки длительностью меньше 0,01 с или периодические воздействия с частотами выше 100 Гц, из-за замыкания дополнительных мостиков имеет вид

$$\frac{l_s - l}{\delta} = \ln \frac{n}{n_s} = \ln \left[e^{-\frac{k}{2}t} + \frac{1}{2P_s} \int_0^t e^{-\frac{k}{2}(t-\tau)} \left(\frac{d}{d\tau} + k \right) P(\tau) d\tau \right]. \quad (7)$$

Здесь $a = P_0 / P_s$ — отношение напряжения на падающем фронте к статическому напряжению мышцы. Видно, что происходит увеличение продолжительности воздействия, приводящее к уменьшению максимальных значений силы, действующей на суставы. Если внешняя нагрузка изменяется по гармоническому закону, укорочение саркомера содержит высшие гармоники. Но при этом нелинейность отклика слабее, если мышца напряжена; это связано с большей асимметрией реакции расслабленной мышцы на сжимающую и растягивающую нагрузку. При сильном изометрическом напряжении мышцы режим ее колебаний почти линеен. В этом случае нетрудно рассчитать динамическую вязкость; на частотах порядка сотен герц она равна $\eta = \eta_0 (1 + P_s / P_*)$, где для портняжной мышцы лягушки $P_* \approx 2,5 \cdot 10^4$ дин/см², $\eta_0 = 0,045$ П. Видно, что уже при нагрузке 1 кГ/см² вязкость возрастает в 40 раз.

Саркомер имеет гексагональную симметрию. Его форма не изменяется при повороте вокруг оси на угол, кратный $\pi/3$ (см. рис. 10). Но соседние саркомеры повернуты на малый случайный угол, саркомеров много, дисперсия поворотов велика и эта симметрия исчезает. Тем не менее в мышце сохраняется анизотропия упругих свойств в направлении оси волокна z и в ортогональном направлении; в поперечном сечении мышцы, очевидно, все направления равноправны. В такой среде тензор модулей упругости, входящий в волновое уравнение

$$\rho \frac{\partial^2 U_i}{\partial t^2} = \lambda_{iklm} \frac{\partial^2 U_m}{\partial x_k \partial x_l}, \quad (8)$$

имеет только 5 независимых компонент:

$$\begin{aligned} \lambda_{xxxx} = \lambda_{yyyy} = a, \quad \lambda_{xyxy} = b, \quad \lambda_{xxyy} = a - 2b, \\ \lambda_{xzzz} = \lambda_{yzzz} = c, \quad \lambda_{xzxz} = \lambda_{yzyz} = d, \quad \lambda_{zzzz} = f. \end{aligned} \quad (9)$$

Удалось показать, что в мышце между модулями имеются жесткие связи $a - c = d + b$, $f - c = 2d$. Таким образом, для сдвиговых волн в мышечной ткани формируется более высокая симметрия. Она характеризуется лишь двумя модулями упругости b , d . Вместо уравнения (8), содержащего тензор λ_{iklm} 4-го ранга, теперь достаточно использовать более простое волновое уравнение

$$\rho \frac{\partial^2 \bar{U}}{\partial t^2} + \text{rot rot} (\hat{R} \bar{U}) = 0, \quad (10)$$

содержащее симметричный тензор 2-го ранга R_{ij} . Возникает такая же ситуация, как для световых волн в одноосных оптических кристаллах.

Результаты этого раздела недавно опубликованы (*Руденко О. В., Сарвазян А. П.* Волновая биомеханика скелетной мышцы // Акуст. журн. 2006. Т. 52. №6. С.833—846). В работе показано, что диссипативными свойствами скелетных мышц можно управлять, напрягая их изометрически. Импульсные нагрузки (удары) и вибрации демпфируются по-разному, в зависимости от их частотного спектра и направления действия. Характерные частоты порядка десятков и сотен герц поглощаются тормозящими миозиновыми мостиками, за счет активной реакции напряженной мышцы. На килогерцовых частотах вязкость ткани увеличивается при напряжении мышцы в десятки раз за счет увеличения трения при относительном движении нитевидных структур в окружающей их жидкости; натяжение нитей изменяет условия их гидродинамического обтекания. Наконец, на более высоких частотах поглощение связано с реологическими свойствами биомолекул, в частности с их конформационной динамикой в волновом поле.

В настоящее время активно идет подготовка экспериментов для проверки сформулированных выше представлений.

Литература

1. Нелинейные волны — 2004 /Под ред. А.В. Гапонова-Грехова и В.И. Некоркина. Нижний Новгород: ИПФ РАН, 2005.
2. УФН. 2006. Т. 176, №1.
3. *Руденко О. В.* Нелинейные волны: некоторые биомедицинские приложения. Специальное заседание редколлегии УФН к 90-летию В. Л. Гинзбурга (Устный выпуск УФН 03.10.2006).
4. *Руденко О. В., Хохлова В. А.* Кинетика одномерных пилообразных волн // Акуст. журн. 1991. Т. 37, №1.
5. *Руденко О. В.* Нелинейные пилообразные волны // УФН. 1995. Т.165, вып. 9.
6. *Гурбатов С. Н., Малахов А. Н., Саичев А. И.* Нелинейные случайные волны в средах без дисперсии. М.: Наука, 1990.
7. *Pierce A. D. J.* Acoust. Soc. Am., 1968. Т. 44. С. 1052.
8. *Руденко О. В., Сухорукова А. К., Сухоруков А. П.* Уравнения высокочастотной нелинейной акустики неоднородных сред // Акуст. журн. 1994. Т. 40, № 2; Двумерные нелинейные волны с разрывами в стратифицированных средах // Акуст. журн. 1995. Т. 41, №2; Полные решения уравнения геометрической акустики в движущихся стратифицированных средах // Акуст. журн. 1997. Т. 43, №3.
9. *Гусев В. А., Руденко О. В.* Статистические характеристики интенсивной волны за двумерным фазовым экраном // Акуст. журн. 2006. Т. 52, №2; *Rudenko O. V., Enflo B. O.* Nonlinear N-wave Propagation Through a One-dimensional Phase Screen // *Acustica — Acta Acustica*. 2000. Т. 86, №2.
10. *Lipkens B., Blackstock D. T.* // *J. Acoust. Soc. Am.* 1998. Т. 104. С. 148; // *J. Acoust. Soc. Am.* 1998. Т. 104. С. 1301.

11. *Flynn H. G.* Method of Generation of Energy by Acoustically Induced Cavitation Fusion and Reactor Therefore. United State Patent 4,333,796 (Jun. 8, 1982).
12. *Taleyarkhan R. P., West C. D., Cho J. S., Lahey Jr. R. T., Nigmatulin R. I., Block R. C.* Evidence for Nuclear Emissions During Acoustic Cavitation // *Science*. 2002. Т. 295. С. 1868—1873.
13. *Станюкович К. П., Хаврошкин О. Б.* Явление рентгеновского излучения кавитационной полостью. Госкомизобретений. Заявка на открытие, ОТ №8576 от 28.09.1973; Явление возникновения импульсной термоядерной реакции в ядре схлопнувшейся кавитационной полости в обычной воде. Госкомизобретений. Заявка на открытие, ОТ №8589 от 22.10.1973.
14. *Марзулис М. А.* Способ получения ВТ плазмы и осуществления термоядерных реакций: Патент РФ №2045715, 1997.
15. *Липсон А. Г., Дерягин Б. В., Клюев В. А. и др.* Инициирование ядерных реакций синтеза при кавитационном воздействии на дейтерийсодержащие среды // *ЖТФ*. 1992. Т. 62, вып. 12. С. 121—130.
16. *Забабихин Е. И.* Явления неограниченной кумуляции // *Механика в СССР за 50 лет*. М.: Наука, 1970; *Кумуляция энергии и ее пределы* // *УФН*. 1965. Т. 85, №4. С. 721—726. *Unlimited Cumulation Phenomena* (Ed. Ya. V. Zeldovich), Moscow: Nauka, 1990.
17. *Nigmatulin R. I., et al.* Theory of Supercompression of Vapor Bubbles and Nanoscale Thermonuclear Fusion // *Physics of Fluids*. 2005. Т. 17, 107106.
18. *Taleyarkhan R. P., Nigmatulin R. I., et al.* Nuclear Emission during Self-Nucleated Acoustic Cavitation // *Phys. Rev. Letters*. 2006. Т. 96, 034301.
19. *Levinson S., Catheline S., Fink M.* Anisotropic elasticity & viscosity deduced from supersonic shear imaging in muscle // *2004 Annual Report*. 2004. Rochester Center for Biomedical Ultrasound. P. 17.
20. *Руденко О. В.* Мощный фокусированный УЗ: нелинейные эффекты, возбуждение сдвиговых волн и медицинская диагностика // *Вестник Моск. ун-та. Сер. 3 (физ., астрон.)*. 1996. Т. 32. №6. С. 3—17.
21. *Sarvazyan A. P., Rudenko O. V., Swanson S. D., Fowlkes J. B., Emelianov S. Y.* Shear Wave Elasticity Imaging — a new ultrasonic technology of medical diagnostics // *Ultrasound in Medicine and Biology*. 1998. V. 24. P. 1419—1435.
22. *Децеревский В. И.* Математические модели мышечного сокращения. М.: Наука, 1977.

О «НЕКЛАССИЧЕСКИХ» ПРОЯВЛЕНИЯХ МИКРОСТРУКТУРНО-ОБУСЛОВЛЕННОЙ АКУСТИЧЕСКОЙ НЕЛИНЕЙНОСТИ НА ПРИМЕРЕ КОНТАКТСОДЕРЖАЩИХ СРЕД

В. Ю. Зайцев

1. «Классическая» решеточная и «неклассическая» нелинейности

Исследования «неклассической» акустической нелинейности твердых тел (которую в литературе называют также мезоскопической, аномальной, гигантской, микроструктурно-обусловленной) стали особенно популярны в последнее десятилетие, несмотря на то что многие важные экспериментальные данные, необходимые для развития соответствующих моделей, были доступны еще с 1960—1970 годов. Фактически все отмеченные выше «аномальные» названия нелинейности связаны со стремлением подчеркнуть ее качественные и количественные отличия от традиционно обсуждаемой в физике твердого тела упругой решеточной нелинейности, связанной с проявлением слабого ангармонизма межатомного потенциала [1]. Как это традиционно делается в теории нелинейной упругости, для описания такой «обычной» нелинейности достаточно к линейному слагаемому в законе Гука добавить квадратичные и кубические по компонентам тензора деформации слагаемые: $\sigma = E[\varepsilon + \gamma^{(2)}; \varepsilon^2 + \gamma^{(3)}; \varepsilon^3 + \dots]$, где σ — упругое напряжение, ε — деформация среды, E — ее модуль упругости. Чтобы не загромождать рассмотрение второстепенными деталями, здесь мы ограничимся случаем продольных деформаций. Характерные величины безразмерных квадратичного и кубического коэффициентов нелинейности $\gamma^{(2)}$ и $\gamma^{(3)}$ для однородных аморфных материалов и монокристаллов обычно лежат в диапазоне нескольких единиц, а характерные величины деформаций до порога разрушения твердых тел обычно весьма малы. Например, в акустике $\varepsilon < 10^{-5}$ даже при весьма интенсивных акустических нагрузках и $\varepsilon < 10^{-3}$ при «обычных» механических деформациях. В связи с этим по сравнению с линейным слагаемым нелинейные поправки в уравнении состояния среды, как правило, очень малы, хотя именно их наличием обусловлены такие известные явления, как тепловое расширение и зависимость скорости упругих волн от механического напряжения и температуры. Оценки величин коэффициентов нелинейности по таким эффектам прекрасно согласуются с представлениями о форме межатомного потенциала и его слабой ангармоничности [1, с. 168].

В настоящее время является общепризнанным, что наличие нарушений структуры твердых тел может очень существенно менять их нелинейные акустические свойства по сравнению с идеальными кристаллами и однородными аморфными материалами, хотя сопутствующие измене-

ния линейных упругих модулей при этом могут оставаться очень незначительными. Такая микро-, или мезоструктурная, неоднородность является скорее правилом, чем исключением, для большинства реальных сред. Получаемые для них экспериментальные данные часто демонстрируют не просто многократный количественный рост их нелинейных параметров («гигантская» нелинейность), но и качественные изменения самого характера нелинейно-акустических эффектов (например, наблюдаются немонотонные или дробно-степенные амплитудные зависимости гармоник, появляются выраженные нелинейно-диссипативные свойства и т. п.), что оправдывает использование терминов «неклассическая» или «аномальная» нелинейность. Ниже будут кратко охарактеризованы основные модельные представления, развитые в процессе исследования необычных нелинейных свойств таких сред, и приведены некоторые показательные экспериментальные факты. Для получения дополнительных сведений могут быть рекомендованы, например, работы [2—4], так как в силу ограниченности доступного объема данный обзор не претендует на полноту и представляет, в основном, результаты, не перекрывающиеся с [2—4].

2. Почему изменчивость именно нелинейных акустических свойств столь высока?

Отмеченные выше необычные нелинейно-акустические свойства наблюдаются у очень широкого класса сред, к которым относятся практически все горные породы, многие металлы с выраженной поликристаллической структурой и/или усталостными и термическими повреждениями, многие конструкционные материалы (типа бетонов), гранулированные среды и т. п. При всех различиях характерной общей особенностью таких сред является наличие в их микроструктуре некоторых структурных неоднородностей (например, дефектов типа выраженных границ между кристаллитами в поликристаллах, где агрегаты дислокаций фактически становятся зародышами трещин, макроскопических трещин и расслоений, которые часто содержат внутренние контакты, межзеренные контакты в гранулированных материалах и т. п.). Характерные размеры этих неоднородностей соответствуют мезомасштабу, т. е. значительно превышают атомарный размер, но являются малыми в масштабе длины упругой волны (что отражается в названиях «мезоскопические», или «микронеоднородные», среды).

Общим свойством отмеченных выше мезо- (или микро-) структурных особенностей является их многократно повышенная «мягкость» по сравнению с окружающим однородным упругим материалом. Например, различные модели трещин согласованно предсказывают, что трещина с характерными раскрытием d и диаметром L может быть полностью закрыта при создании в окружающем материале средней деформации сжатия $\varepsilon \sim d/L \ll 1$. Это фактически означает, что податливость (мягкость)

трещины примерно в L/d раз выше, чем окружающего однородного материала, причем вполне типичными для реальных трещин являются значения $L/d \sim 10^3 \div 10^5$. Аналогичным образом, в силу малости площади соприкосновения, у контактов между гранулами локальная жесткость оказывается многократно ниже, чем у объема зерен. Повышенная мягкость таких структурных особенностей приводит к многократному возрастанию на них локальной деформации (и скорости деформации) в поле упругих волн, в связи с чем на таких мягких дефектах оказываются локализованы как заметные отклонения от линейного соотношения между упругим напряжением и деформацией, так и потери энергии упругих возмущений. В результате при не слишком малой (и, как будет видно ниже, не слишком большой) концентрации дефектов микронеоднородный материал демонстрирует ярко выраженные нелинейные акустические свойства, причем не только нелинейно-упругие, но и нелинейно-диссипативные и нелинейно-дисперсионные.

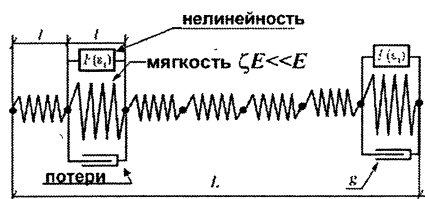


Рис. 1. Реологическая модель среды, содержащей контрастно-мягкие дефекты-включения, на которых локализованы нелинейность и диссипация.

Отразить отмеченные выше структурные особенности удастся, не конкретизируя детально тип дефектов, в рамках наглядной реологической модели микронеоднородной среды с контрастно-мягкими включениями [5—8]. Эта модель допускает обобщение на трехмерный случай (с учетом как тангенциальной, так и нормальной податливости дефектов) [9, 10], однако и

в одномерном варианте она позволяет получить ряд нетривиальных выводов. Схематически такая модель микронеоднородной среды показана на рис. 1. Однородную среду-матрицу, окружающую дефекты, при этом можно рассматривать как линейно-упругий материал, описываемый законом Гука $\sigma = E\varepsilon$. Дефектами в такой модели служат мягкие включения. Важнейшими параметрами дефектов являются их относительная (по сравнению с модулем упругости E среды-матрицы) податливость, характеризующая малым параметром $\zeta \ll 1$, а также концентрация ν (т. е. линейная концентрация мягких включений в 1D-случае либо их объемная доля при обобщении на 3D-случай). Размер дефекта считаем много меньшим длины упругой волны, так что высокая сжимаемость дефектов приводит к отмеченным выше локально повышенным деформации и скорости деформации. Именно поэтому в модели однородную среду-матрицу можно рассматривать как идеально линейную и непоглощающую и считать, что отклонения от линейного закона Гука и диссипация локализованы на дефектах-включениях. При этом, если деформацию ε_1 дефектов

измерять в их собственном масштабе (при средней деформации среды $\varepsilon \ll \varepsilon_1$), то как собственная нелинейность дефекта, так и поглощение на нем характеризуются вполне «обычным», а не повышенным уровнем. Уравнение состояния, учитывающее отмеченные свойства отдельных дефектов с параметром мягкости ζ , имеет вид

$$\sigma = \zeta E [\varepsilon_1 + F(\varepsilon_1)] - g d \varepsilon_1 / dt. \quad (1)$$

Коэффициент g характеризует эффективную «вязкость» дефектов, а функция $F(\cdot)$ описывает их нелинейность. Например, для квадратично-нелинейных дефектов $F(\varepsilon_1) = \gamma \varepsilon_1^2$, причем следует еще раз подчеркнуть, что параметр γ , характеризующий нелинейность дефекта в масштабе его собственной (локальной) деформации ε_1 , имеет «обычную» величину порядка нескольких единиц. Только присутствие окружающей относительно более жесткой среды-матрицы обеспечивает многократное возрастание деформации на дефектах, что может приводить и к значительному повышению средней (макроскопической) нелинейности материала. Особенно наглядно такой «контрастно-упругий» механизм возрастания упругой нелинейности (в англоязычной литературе в последнее время называемый «soft-rigid paradigm») проявляется в квазистатическом пределе, когда вязким слагаемым в (1) можно пренебречь. В таком случае, например для степенной нелинейности дефектов $F(\varepsilon_1) = \gamma^{(n)} \varepsilon_1^n$, макроскопическая связь напряжение — деформация принимает следующую простую форму (справедливую во всем возможном диапазоне концентрации дефектов $0 \leq v \leq 1$) [5]:

$$\sigma = \varepsilon \cdot E_{\text{eff}} \{1 + \varepsilon^{n-1} \cdot \gamma_{\text{eff}}^{(n)}\},$$

где

$$E_{\text{eff}} / E = 1 / (1 - v + v / \zeta), \quad \gamma_{\text{eff}}^{(n)} / \gamma^{(n)} = (1 - v + v / \zeta^n) / (1 - v + v / \zeta)^n. \quad (2)$$

Эти выражения иллюстрируют (рис. 2), что локальное повышение деформации на мягких дефектах (связанное с их показателем относительной мягкости $\zeta \ll 1$) существенно сильнее проявляется в возрастании нелинейных членов в уравнении состояния, нежели в уменьшении линейного модуля упругости. Для данного показателя мягкости дефектов ζ существует диапазон очень малых их концентраций $v < \zeta \ll 1$, где линейный модуль все еще остается практически неизменным, а нелинейный параметр материала уже многократно возрастает (см. рис. 2). Именно в этой области малых концентраций, где чувствительность линейных методов еще недостаточна для получения определенного результата, представляет основной интерес использование нелинейных эффектов для раннего обнаружения «слабых мест» в материале (в первую очередь, зарождения трещин). Интересно подчеркнуть, что увеличение коэффициентов нели-

нейности $\gamma_{\text{eff}}^{(n)} / \gamma^{(n)}$ разных порядков в зависимости от концентрации дефектов ν демонстрирует выраженный максимум $\gamma_{\text{eff}}^{(n)} / \gamma^{(n)} = [(n-1)/\zeta]^{n-1} / n^n \gg 1$, причем, чем выше порядок нелинейности, тем при меньшей концентрации дефектов $\nu_{\text{опт}} \approx \zeta / (n-1) \ll 1$ достигается этот максимум и тем выше его величина. Такое немонокотное возрастание физически обусловлено игрой между фактором локального повышения деформации на мягких включениях и величиной их объемного содержания (концентрацией). Аналогичное явление известно и в акустике газожидкостных смесей, демонстрирующих яркий пример «контрастного» механизма возрастания нелинейности, в то время как чистые жидкость и газ являются классическими слабнонелинейными средами [11].



Рис. 2. Относительные изменения модуля упругости E_{eff}/E и параметра квадратичной нелинейности $\gamma_{\text{eff}}^{(2)}/\gamma^{(2)}$ микронеоднородной среды, содержащей дефекты с типичным для трещин показателем относительной мягкости $\zeta = 10^{-4}$.

Анализ показывает также, что возможная иерархическая структура среды (когда отдельные элементы среды сами имеют вложенные уровни микронеоднородности) и распределение дефектов по параметру мягкости не приводят к дополнительному возрастанию макроскопической нелинейности по сравнению с оценкой (2) [6, 7]. С другой стороны, учет широкого распределения реальных дефектов по их упруго-вязким параметрам позволяет естественным образом описать типичные для многих горных пород почти частотно-независимую добротность и сопутствующую ей слабую (логарифмическую по частоте) дисперсию скорости упругих волн без привлечения гипотез о каком-либо специальном (неэкспоненциальном) характере релаксации в среде [12, 13]. Не останавливаясь подробнее на следствиях такой модели в линейном приближении, отметим лишь один несколько неожиданный ее результат: в широкой области частот, где добротность среды с распределенными по параметру мягкости дефектами остается приблизительно постоянной, значение этой добротности вообще не зависит от параметра вязкости дефектов, а определяется только их концентрацией [12].

Далее, совместный учет и нелинейных, и релаксационных свойств дефектов [8, 13] приводит (в приближении их малой концентрации и при одинаковом показателе мягкости ζ) к следующему соотношению между упругим напряжением и средней (макроскопической) деформацией материала:

$$\begin{aligned} \sigma(\varepsilon) = E\varepsilon - E\nu\Omega \int_{-\infty}^t \varepsilon(\tau) e^{-\zeta\Omega(t-\tau)} d\tau + \\ + \nu E\Omega\zeta \int_{-\infty}^t e^{-\zeta\Omega(t-\tau)} F\left\{\Omega \int_{-\infty}^t \varepsilon(\tau') e^{-\zeta\Omega(\tau-\tau')} d\tau'\right\} d\tau. \end{aligned} \quad (3)$$

Здесь использовано обозначение $\Omega = E/g$, так что $\zeta\Omega$ имеет смысл релаксационной частоты дефекта, причем в случае различных дефектов это уравнение следует дополнить усреднением по их свойствам, т. е. по распределению $\nu(\zeta, g)$. Первое слагаемое в правой части уравнения (3) обязано своим происхождением линейной среде-матрице, второе учитывает привносимые дефектами линейное уменьшение упругого модуля и диссипацию. Третье, нелинейное, слагаемое описывает совместное влияние нелинейных и релаксационных свойств дефектов, причем в квазистатическом пределе уравнение (3) переходит в уже рассмотренное выше уравнение (2).

Комбинация в последнем слагаемом уравнения (3) линейной экспоненциальной релаксации и упругой нелинейности, традиционно рассматриваемых в акустике, приводит к появлению таких микроструктурно-обусловленных нелинейных свойств, которые представляются, на первый взгляд, совершенно неожиданными с точки зрения интуиции, выработанной на основе «классической» теории нелинейной упругости. Например, структура нелинейного слагаемого в (3) показывает, что для вполне обычного (в частности, упруго-квадратичного, описываемого функцией $F(\varepsilon_1) = \gamma\varepsilon_1^2$) характера нелинейности дефекта влияние релаксационных свойств приводит к выраженной частотной зависимости эффективных нелинейных параметров среды [8, 14], что совершенно не характерно для решеточной нелинейности идеальных кристаллов и однородных аморфных сред. Для микронеоднородных сред, однако, ситуация оказывается существенно иной, поскольку в них как нелинейность, так и релаксационные свойства физически локализованы, в основном, в одних и тех же местах — на мягких дефектах. Поэтому релаксационное «замораживание» линейной реакции дефектов на акустические воздействия с частотами много выше их релаксационной частоты $\zeta\Omega$ одновременно «выключает» и их нелинейный отклик. В результате при наблюдении, например, обычного эффекта смешения волн с частотами ω_1 и ω_2 эффективные зна-

чения квадратичного нелинейного параметра, оцениваемого по уровню возникающих суммарных гармоник $2\omega_{1,2}$ и $\omega_1 + \omega_2$, могут оказаться существенно ниже, чем для гармоник на разностной частоте $\omega_1 - \omega_2$. Кроме того, становятся возможными необычные эффекты, связанные с интерференцией нелинейно-порожденных компонент сигнала от дефектов с различными релаксационными частотами. Действительно, вспомним, что и в линейном приближении отклик релаксатора запаздывает по фазе по отношению к возбуждающему гармоническому воздействию. Однако для линейного отклика релаксаторов такая фазовая задержка (см. второе слагаемое в уравнении (3)) не превышает $\pi/2$ радиан, да и сам этот отклик является лишь небольшой поправкой к основному (первому) линейному слагаемому в (3). Ситуация оказывается совершенно иной для высших гармоник, возникновение которых описывается последним нелинейно-релаксационным слагаемым в (3). Как видно из его структуры, релаксационный поворот фазы в нем может «срабатывать» несколько раз и потому превышать π радиан. В результате вклады от дефектов с различными релаксационными частотами могут интерферировать как конструктивно, так и деструктивно, приводя к немонотонной частотной зависимости уровня высших гармоник и даже к возможности полного гашения выбранной гармоники при некотором значении частоты [6, 14]. В целом традиционная для нелинейной акустики интуиция, основанная на возможности аддитивного учета нелинейности и релаксации (что правомерно для однородных сред), не работает в случае микронеоднородной среды. Аддитивный подход становится неправомерен в силу уже отмеченного релаксационного «замораживания» реакции дефектов на высоких частотах, что одновременно «выключает» и их нелинейность, так что независимый учет нелинейности и поглощения (как это имеет место в структуре, например, уравнения Бюргерса) оказывается невозможным.

Еще одним нетривиальным проявлением такой «контрастной» микронеоднородности среды является возможность сильного (в разы) изменения уровня диссипации (декремента) акустического возмущения в микронеоднородном материале при вполне умеренном статическом или динамическом (скажем, акустическом с деформацией $\varepsilon \sim 10^{-5}$) воздействии на среду [10, 15]. Такой необычный эффект возникает опять-таки благодаря комбинированному действию обычного линейного релаксационного поглощения на дефектах и их упругой нелинейности. Например, для простейшей квадратичной упругой нелинейности дефектов $F(\varepsilon_1) = -\gamma\varepsilon_1^2$, считая выполненным условие $\varepsilon_0 \gg \varepsilon_a$ для статической деформации ε_0 и осциллирующей (акустической) деформации $\tilde{\varepsilon}_a = (\varepsilon_a e^{i\omega t} + \varepsilon_a^* e^{-i\omega t})/2$, из уравнения (3) получаем для акустической составляющей напряжения $\sigma_a(\varepsilon_a)$:

$$\sigma_a(\varepsilon_a) = Ea - \frac{\nu E}{\zeta} \varepsilon_a \frac{1}{1 + \bar{\omega}^2} - \frac{\nu E}{\zeta} \varepsilon_a F' \left(\frac{\varepsilon_0}{\zeta} \right) \frac{1 - \bar{\omega}^2}{(1 + \bar{\omega}^2)^2} +$$

$$+ i \frac{\nu E}{\zeta} \varepsilon_a \frac{\bar{\omega}}{1 + \bar{\omega}^2} + i \frac{\nu E}{\zeta} \varepsilon_a F' \left(\frac{\varepsilon_0}{\zeta} \right) \frac{2\bar{\omega}}{(1 + \bar{\omega}^2)^2},$$
(4)

где $\bar{\omega} = \omega / (\zeta \Omega)$ — нормированная частота. Отсюда для эффективного модуля упругости среды и декремента затухания следует [15]

$$E_{\text{eff}} / E \approx 1 - \frac{\nu}{\zeta} \frac{1}{1 + \bar{\omega}^2} - 2 \frac{\gamma \varepsilon_0 \nu}{\zeta^2} \frac{1 - \bar{\omega}^2}{(1 + \bar{\omega}^2)^2},$$
(5)

$$\theta = \theta_{\text{lin}} + \theta_{\text{nl}} \approx \pi \frac{\nu}{\zeta} \frac{\bar{\omega}}{1 + \bar{\omega}^2} + 2\pi \frac{\nu \gamma \varepsilon_0}{\zeta^2} \frac{2\bar{\omega}}{(1 + \bar{\omega}^2)^2}.$$
(6)

Выражение (6) показывает, что отношение $\theta_{\text{nl}} / \theta_{\text{lin}}$ нелинейного (т. е. зависящего от ε_0) вклада в декремент к линейному вкладу по порядку величины равно $\gamma \varepsilon_0 / \zeta$. Отсюда следует, что для достаточно мягких ($\zeta \ll 1$) дефектов нелинейный вклад θ_{nl} может оказаться одного порядка с линейным даже при вполне умеренных деформациях ε_0 среды. Примем для иллюстративной оценки значение параметра мягкости дефектов $\zeta = 10^{-4}$ (достаточно типичное для тонких трещин), выберем заведомо «скромное» значение параметра собственной нелинейности дефектов $\gamma = 2$, а для их концентрации возь-

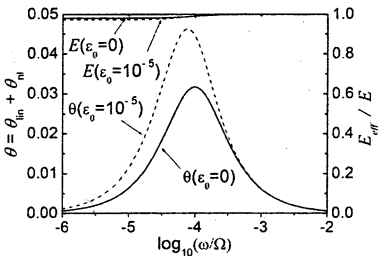


Рис. 3. Иллюстрация резкого различия одновременных изменений декремента θ и модуля упругости E_{eff} , вызванных приложением статической деформации $\varepsilon_0 = 10^{-5}$. Видно, что в окрестности максимума декремента на частоте $\omega = \zeta \Omega$ релаксации дефектов изменение декремента превышает 1,5 раза при сопутствующем относительно изменении $\Delta E_{\text{eff}} / E < 10^{-2}$.

мем значение $\nu = 2 \cdot 10^{-6}$. Такая концентрация для линейных потерь (в окрестности релаксационного максимума линейного слагаемого θ_{lin}) соответствует типичной для сейсмических волн величине добротности $Q \equiv \pi / \theta \approx 100$. На рис. 3 приведены построенные при этих параметрах на основе выражений (5) и (6) частотные зависимости для относительного изменения модуля упругости $\Delta E_{\text{eff}} / E$ и декремента θ при значениях статической деформации $\varepsilon_0 = 0$ и умеренной деформации $\varepsilon_0 = 10^{-5}$ (здесь положительный знак ε_0 соответствует растяжению). Как видно из этого примера, при

вариациях упругого модуля менее процента амплитудно-зависимое изменение декремента превышает 50%. Отметим также, что для принятых параметров «макроскопический» коэффициент упругой квадратичной нелинейности в уравнении (5) $2\gamma v/\zeta^2 \approx 800 \gg 1$, что вполне типично для горных пород.

Интересно отметить, что в окрестности релаксационной частоты дефектов ($\omega = \zeta\Omega$) амплитудно-зависимая поправка к упругому модулю обращается в ноль, меняя знак из-за изменения фазы отклика дефектов-релаксаторов. Таким образом, в окрестности релаксационной частоты может иметь место несколько парадоксальный эффект появления очень заметного амплитудно-зависимого поглощения при отсутствии одновременных изменений упругого модуля («пропадание» упругой компоненты нелинейности). Этот эффект не исчезает и при учете распределения реальных дефектов по параметру мягкости [15] и выглядит как проявление чисто диссипативной нелинейности, хотя физически он возникает за счет комбинации чисто реактивной (упругой) нелинейности дефектов и обычных линейных потерь релаксационного типа. Подобные ярко выраженные амплитудно-зависимые изменения диссипации одной упругой волны под влиянием другого статического или динамического воздействия действительно наблюдаются экспериментально в различных материалах (см., например, [16, 10]).

Важно отметить также, что рассмотренный выше механизм проявляется и для самовоздействия одной волны, приводя к одновременным вариациям одного порядка величины для относительного изменения упругого модуля (дефект модуля) и абсолютной величины декремента. В нелинейной акустике такие одновременные изменения традиционно приписывают влиянию гистерезисной нелинейности (возникающей за счет фрикционных или адгезионных явлений на тех же самых микроконтактах и трещинах, соответствующих мягким дефектам в обсуждаемой модели) [2, 3, 16—18], причем количественное соотношение между этими вариациями предлагается использовать для уточнения формы гистерезисной петли [19]. Оценки, выполненные на основе рассмотренного выше негистерезисного (т. е. комбинированного релаксационно-нелинейно-упругого) механизма амплитудно-зависимых потерь, показывают, что его вклад может быть количественно сравним со вкладом параллельно действующей гистерезисной нелинейности. Этот факт необходимо учитывать при интерпретации нелинейно-акустических эффектов в микронеоднородных средах, в которых гистерезисны и рассмотренный негистерезисный механизмы должны действовать параллельно, давая сравнимые по величине вклады [15].

3. Трещины с внутренними контактами как мягкие дефекты в упругой матрице

После проведенного на реологическом уровне обсуждения причин появления аномально высокой (причем не только реактивной, упругой, но и диссипативной) нелинейности микронеоднородных сред остановимся подробнее на проявлениях контактной нелинейности в трещиносодержащих средах. Отметим, что важность роли трещин, являющихся ярким примером «контрастно-мягких» включений, качественно была понята достаточно давно. Еще в 1960-е годы было понятно, что трещины дают очень существенный вклад в поглощение сейсмических волн. Например, в работе [17] потери приписывались трению между поверхностями трещин, хотя сами авторы высказывали сомнения в собственном подходе, отмечая, что обсуждаемый диапазон *«амплитуд смещений [на поверхностях трещин] начинается с 10^{-12} см. Эти смещения столь малы [на 3—4 порядка меньше атомного размера], что характеристики трения контактирующих поверхностей, по-видимому, должны сильно отличаться от трения в макроскопических экспериментах»*. Действительно, фрикционные эффекты являются пороговыми, так как подразумевают, что смещения на контактирующих поверхностях превышают хотя бы один атомный размер (т. е. 3×10^{-8} см). Другие авторы (например, [20]) рассматривали непороговый механизм термоупругих потерь на трещинах. Однако, так как релаксационные частоты для трещин диаметром $L \sim 1$ мм имеют порядок единиц и долей герц и пропорциональны L^{-2} , а амплитуды пиков спадают как L^{-3} , такая модель требовала предположения о нереально большой концентрации микротрещин (микронных масштабов) для объяснения величины потерь сейсмических волн на килогерцовых и более высоких частотах. Для объяснения упругой нелинейности горных пород также традиционно привлекались соображения о высокой сжимаемости трещин, например, в работе [21], где обсуждался эффект модуляции скоростей сейсмических волн полем приливных деформаций (с типичными амплитудами деформации 10^{-8}). Однако для получения количественного согласования с данными [21] и другими аналогичными высокими оценками уровня нелинейности требуется предположить очень малые значения аспект-отношения трещин, $d/L \sim 10^{-5}—10^{-6}$. Более того, такое значение необходимо принять практически для всех трещин в среде, что представляется заведомо нереалистичным, так как для большинства трещин характерны более высокие значения $d/L \sim 10^{-4}—10^{-3}$, что, соответственно, на порядки понижает ожидаемую чувствительность таких дефектов к упругим напряжениям в среде.

Разрешить отмеченные выше сложности и противоречия, возникающие при объяснении нелинейно-упругих и диссипативных свойств горных пород присутствием трещин, удастся естественным образом, если

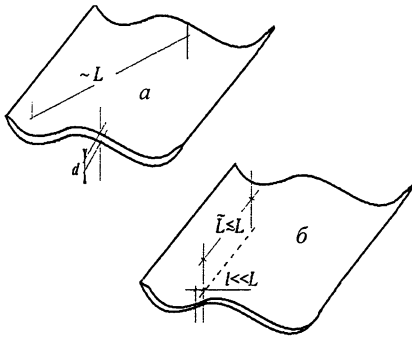


Рис. 4. Трещина с волнистыми поверхностями без (а) и при наличии (б) внутреннего контакта размером $\tilde{L} \times l$. При $\tilde{L} \rightarrow l$ полосковый контакт переходит в точечный.

учесть неровность поверхностей трещин и наличие в них внутренних контактов. Существенно, что в силу двумерного характера инициации трещин эти неровности (и их контакты) являются скорее не точечными, а удлиненными, с локально цилиндрической геометрией, так что длина \tilde{L} такого контакта может быть сопоставима с диаметром L самой трещины (рис. 4). Как по периметру трещины, так и на контакте имеет место концентрация напряжений (и деформаций), кроме того, по сравнению с однородной средой оказываются резко повышены температурные градиенты в поле упругих волн, определяемые в данном случае не их длиной, а значительно меньшими размерами трещины и контакта, а также длиной тепловой волны. На частоте, для которой длина тепловой волны совпадает с размером либо всей трещины [20], либо контакта [22, 23], имеет место релаксационный максимум термоупругого поглощения. Для потерь упругой энергии на таком контакте длиной \tilde{L} и шириной d можно получить [22]:

$$W_{LF} \approx 2\pi\omega T(\alpha^2 K^2 / \kappa) l^2 \tilde{L}^2 \varepsilon^2, \quad \omega \ll \omega_l \approx \kappa / (\rho C l^2), \quad (7)$$

$$W_{HF} \approx (2\pi / \omega) \kappa T (\alpha K / C \rho)^2 \tilde{L} (L / l)^2 \varepsilon^2, \quad \omega \gg \omega_l, \quad (8)$$

$$W_{\text{cont}}^{\text{max}} \approx 2\pi T (\alpha^2 K^2 / \rho C) \tilde{L}^2 \varepsilon^2, \quad \omega \approx \omega_l, \quad (9)$$

где ω — круговая частота; T — температура; α — коэффициент объемного теплового расширения; K — модуль объемного сжатия; ρ — плотность; C — удельная теплоемкость; ε — средняя деформация в материале; κ — коэффициент теплопроводности; ω_l — характерная частота терморелаксации, соответствующая размеру l контакта в трещине. Отметим, что аналогичные выражения термоупругих потерь на трещине в целом (рассмотренных в [20]) можно привести [22] к аналогичному виду:

$$W_{LF} \approx 2\pi\omega T (\alpha^2 K^2 / \kappa) L^2 \varepsilon^2, \quad \omega \ll \omega_L \approx \kappa / (\rho C L^2), \quad (10)$$

$$W_{HF} \approx 2\pi T (\alpha^2 K^2 / \rho C) [\kappa / (\rho C \omega)]^{1/2} L^2 \varepsilon^2, \quad \omega \gg \omega_L, \quad (11)$$

$$W_{\text{crack}}^{\text{max}} \approx 2\pi T (\alpha^2 K^2 / \rho C) L^2 \varepsilon^2, \quad \omega \approx \omega_L. \quad (12)$$

Сравнение выражений (12) и (9) для потерь в области релаксационного максимума показывает неожиданный, на первый взгляд, результат, что для полосковых контактов с длиной $\tilde{L} \sim L$ максимальные потери на трещине в целом и на узком контакте, площадь которого во много раз меньше, оказываются одного порядка, хотя при этом релаксационный максимум для узкого контакта $l \ll L$ может лежать на 4—6 порядков выше по оси частот.

Приведенные выше выражения показывают, что традиционное представление о несущественности термоупругих потерь для, например, сейсмических волн требует существенного пересмотра уже в линейном приближении, причем для волн малой амплитуды ($\epsilon \leq 10^{-7}—10^{-9}$) в рассмотренном варианте термоупругий механизм может доминировать. Из выражений (7)—(12) следует, что трещины с несколькими мягкими контактами могут обеспечить заметное термоупругое поглощение в диапазоне частот шириной в несколько порядков (от релаксационной частоты трещины в целом до релаксационных частот контактов). При этом находящийся в трещине размера L единичный полосковый контакт ширины $l \ll L$ при длине $\tilde{L} \sim L$ обеспечивает в окрестности своей релаксационной частоты ω_l примерно такую же диссипацию, какую производят в соответствии с «глобальным» механизмом [20] $N = (L/l)^3 \gg 1$ микро-трещины размера l , имеющие ту же частоту релаксации. Причем «эквивалентное количество» N микротрещин может быть очень большим, например $N \sim 10^6$ при типичном для реальных трещин отношении $L/l \sim 10^2$.

Далее, для нелинейных эффектов существенно, что в окрестности контактов локальное расстояние между поверхностями трещины (или их взаимопроникновение) \tilde{d} значительно меньше среднего значения d раскрытия трещины. В связи с такой геометрией окрестность контакта многократно (примерно в $d/\tilde{d} \gg 1$ раз) более чувствительна к внешнему напряжению, чем трещина в целом, так что состояние контактов в трещине может быть существенно изменено уже относительно слабой средней деформацией в среде $\epsilon \sim 10^{-6}—10^{-5}$ (которая еще много меньше деформации $\epsilon \sim d/L \sim 10^{-4} \div 10^{-3}$, необходимой для полного закрытия трещины). Так как жесткость возникающего полоскового контакта оказывается сравнимой с арочной жесткостью трещины, это означает, что введенный выше эффективный параметр мягкости ζ (характеризующий чувствительность дефектов к средней деформации в материале) может для тонких трещин с $d/L \sim 10^{-4} \div 10^{-5}$ достигать величины $\zeta \sim 10^{-6} \div 10^{-7}$ без привлечения гипотез о нереалистичных значениях $d/L \sim 10^{-6} \div 10^{-7}$. Оказывается, что таких значений эффективной мягкости и уровня диссипации, оцененных для трещин с реальными параметрами, с запасом хватает для того, чтобы объяснить наблюдаемую глубину (~5—10%) модуляции высоко-

частотных сейсмических шумов [24] приливными деформациями (с амплитудами $\sim 10^{-8}$). При этом достаточно учесть только модуляцию приливами поглощения в среде без привлечения гипотез о микроразрушениях и сопутствующей им эмиссии, инициируемых слабыми приливыми деформациями. До настоящего времени реалистичных физических моделей этого эффекта предложено не было.

Еще одной интересной особенностью термоупругого поглощения на полосковых контактах в трещинах (см. уравнение (9)) является независимость высоты релаксационного пика от ширины контакта l , которая может меняться под действием прилагаемых статических или динамических нагрузок. Это означает, что такой сдвиг релаксационного пика может приводить одновременно к индуцированному дополнительному поглощению или «просветлению» для пробных волн, частоты которых лежат по разные стороны пика. На рис. 5 приведено схематическое пояснение эффекта для динамического воздействия на контакт акустической «накачки», и на рис. 6 дан соответствующий экспериментальный пример индуцированных просветления/поглощения для разных мод в стержневом стеклянном резонаторе с одиночной трещиной [26].

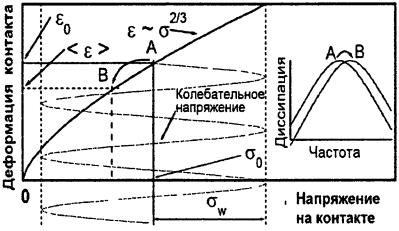
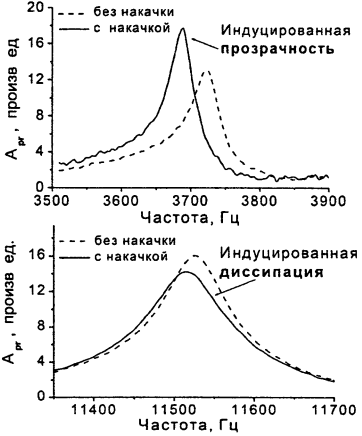


Рис. 5. Схематически показано умягчение контакта Герца осциллирующим воздействием из-за асимметрии зависимости поджатия контакта от прилагаемого упругого напряжения. Невозмущенное статическое положение равновесия $A = (\sigma_0, \epsilon_0)$, усредненное по периоду измененное положение $B = (\langle \sigma \rangle, \langle \epsilon \rangle)$ и соответствующий сдвиг релаксационного пика.

Рис. 6. Примеры одновременных изменений резонансных кривых пробной волны на разных модах под действием накачки. Мода на частоте $F = 3,7$ кГц испытывает уменьшение диссипации, а для моды с частотой $F = 11,5$ кГц, наоборот, потери увеличиваются. В отличие от разного знака изменения потеря частота сдвигается вниз для обеих мод. Уровень накачки (на частотах $F = 40 \div 50$ кГц) $\epsilon \sim 10^{-6}$.



Следующей также «неклассической» особенностью контактной нелинейности акустически активируемого контакта является немонотонность амплитудной зависимости некоторых его свойств, например средней ширины и жесткости. С ростом амплитуды колебаний при переходе к «хлопающему» режиму эти величины сначала уменьшаются, а затем вновь увеличиваются. Такая немонотонность объясняет наблюдаемый для образцов с трещинами немонотонный сдвиг резонансных пиков сначала влево (понижение среднего модуля упругости), а затем вправо (вновь увеличение жесткости) при одновременном немонотонном поведении индуцированной диссипации, т. е. сначала увеличение, а затем некоторое просветление [25].

До недавних пор в акустике не обсуждалось еще одно интересное проявление выраженных нелинейно-диссипативных свойств, вносимых присутствием дефектов, а именно, эффект кросс-модуляции в образцах с трещинами. Этот акустический аналог известного для радиоволн в ионосфере Люксембург-горьковского (ЛГ) эффекта представляет интерес в качестве высокочувствительного метода мониторинга состояния среды и диагностики появления трещин [22, 26] наряду с использованием в диагностических схемах обычной модуляции пробной акустической волны другой более низкочастотной волной или колебанием (рис. 7).

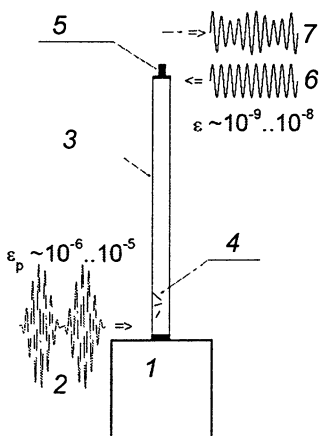


Рис. 7. Схема эксперимента по наблюдению ЛГ-эффекта: 1 — нагрузочная масса с пьезоизлучателем; 2 — модулированная накачка; 3 — стеклянный стержень; 4 — трещины, инициированные термическим шоком; 5 — акселерометр и источник пробной волны; 6 — форма немодулированного возбуждаемого и 7 — модулированного принимаемого пробного сигнала.

Кроме обсуждавшихся выше «мгновенных» и «быстрых» релаксационных проявлений нелинейности дефектов в последние годы привлекают большое внимание эффекты медленной динамики акустически активируемых образцов микронеоднородных сред. После выключения интенсивного акустического воздействия на образец при дальнейшем наблюдении значений его упругого модуля (и иногда добротности) с помощью слабой пробной волны часто обнаруживается, что начальные значения параметров материала (до активации) восстанавливаются не

мгновенно, а демонстрируют медленное (на масштабах 10^2 — 10^4 с) логарифмическое по времени восстановление [27]. При этом обычно исследуются образцы горных пород, содержащие очень большие ансамбли трещин, а наблюдаемая медленная динамика приписывается восстановлению под действием термических флуктуаций неких внутренних связей, разрушенных при акустическом активировании образца. Характер такой медленной релаксации определяется статистикой соответствующих этим связям энергетических барьеров, подбирая функции распределения которых, удается получить и логарифмическое по времени поведение [27]. Вопрос о физическом соответствии таких барьеров и их статистики структурным особенностям среды остается открытым.

В ходе детальных исследований нелинейно-акустического отклика образцов с одиночными трещинами (типа экспериментов, описанных в [22]) было обнаружено, что образцы даже с одиночными трещинами также демонстрируют выраженную, почти логарифмическую по времени, медленную динамику на масштабах в сотни секунд [25, 26]. Более того, применение волн различных частот в качестве пробного воздействия и интенсивной «накачки» позволило исследовать поведение образца не только после, но и во время его активации. При этом выяснилось, что во время активации параметры образца также логарифмически по времени медленно изменяются, причем практически с той же скоростью, с какой затем происходит их восстановление (рис. 8). Известные объяснения подобной динамики за счет разрыва/восстановления большого ансамбля связей [27] явно не подходили для одиночного дефекта и не согласовывались с симметрией между стадиями активации и релаксации. В то же время локально-цилиндрическая геометрия контактов в трещинах (на которую первоначально внимание было обращено по совершенно другим соображениям [22]) позволила предложить естественное объяснение этим наблюдениям за счет локального нагрева (и затем охлаждения) областей концентрации напряжений на контакте и периметре трещины интенсивной акустической накачкой [25]. Промежуточная логарифмическая асимптотика и симметрия в решениях для стадий нагрева и охлаждения непосредственно вытекает из локально-цилиндрической геометрии задачи. Эта модель также предсказывает, что, когда размер области нагрева срав-

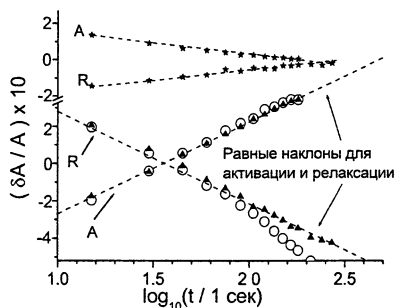


Рис. 8. Логарифмическая и симметрично-обратимая во времени динамика амплитуды пробной волны (т. е. добротности) для двух мод в стеклянном резонаторе с трещиной (значок А — активация, R — релаксация) [26].

нивается с диаметром трещины, логарифмический дрейф параметров насыщается, и дальнейшее распространение тепла происходит как от точечного источника в трехмерной геометрии. И качественно (насыщение логарифмического поведения), и количественно (оценки характерных времен) эти предсказания модели прекрасно согласуются с наблюдениями [25]. Аналогичное нетипичное для традиционных задач нелинейной акустики логарифмическое поведение было затем обнаружено и для металлических образцов с трещинами, причем с меньшими характерными временами, как того и следовало ожидать в силу большей теплопроводности металла [26].

Подробнее обсуждение упомянутых выше и некоторых других (например, «памяти» нелинейности) «неклассических» нелинейных эффектов в образцах с трещинами можно найти в работах [22, 23, 25, 26, 28] и цитированной в них литературе.

4. Эффекты контактной нелинейности в гранулированных средах

На диаграмме микроструктурно-обусловленных вариаций параметров (см. рис. 2) материалы с небольшим содержанием трещин (вдали от порога разрушения) лежат слева от максимума нелинейности среды, где ее модуль упругости еще незначительно снижен по сравнению с однородной средой-матрицей. Гранулированные среды, представляющие собой сеть контактов (предварительно нагруженных в той или иной степени), лежат на правой стороне этой диаграммы, где линейный модуль среды уже значительно ниже модуля однородного материала гранул [6]. Базовой моделью при описании нелинейности гранулированных сред служит герцевская модель контакта двух упругих сфер, предсказывающая, что сила на контакте пропорциональна величине сближения центров в степени $3/2$ (причем эта функциональная зависимость сохраняется и для других тел конечных размеров, форма которых вблизи контакта может быть аппроксимирована поверхностями второго порядка) [29]. Происхождение такой нелинейности чисто геометрическое, так что она имеет место и для идеально линейного материала гранул. Для среды с одинаково нагруженными контактами их герцевская нелинейность обуславливает функционально такой же вид связи «напряжение и деформация» $\sigma = \sigma(\epsilon)$ в среде:

$$\sigma(\epsilon) = bn\epsilon^{3/2}H(\epsilon), \quad (13)$$

где множитель b зависит от упругих модулей материала зерен; n — среднее количество контактов, приходящихся на одно зерно; $H(\epsilon)$ — функция Хевисайда, показывающая, что только при сжатии ($\sigma, \epsilon > 0$) контактов возникает напряжение в среде (здесь мы по-прежнему ограничились продольной деформацией). В реальной зернистой среде, однако, нагружен-

ность контактов, как известно, может отличаться очень сильно, что приводит к изменению макроскопического уравнения (13) [30, 31]. Для иллюстрации роли различных контактов в линейных и нелинейных акустических проявлениях предположим, что гранулированная среда содержит только две фракции контактов, имеющие различные статические деформации: среднюю ε_0 (близкую к макроскопической деформации) и $\mu\varepsilon_0$, где безразмерный параметр нагружения $\mu \ll 1$. Разделяя статическую (σ_0, ε_0) и динамическую ($\tilde{\sigma}, \tilde{\varepsilon}$) части в общих напряжениях и деформациях для обеих фракций, из уравнения (13) получим

$$\sigma_0 + \tilde{\sigma} = bn_1(\varepsilon_0 + \tilde{\varepsilon})^{3/2} H(\varepsilon_0 + \tilde{\varepsilon}) + bn_2(\mu\varepsilon_0 + \tilde{\varepsilon})^{3/2} H(\mu\varepsilon_0 + \tilde{\varepsilon}), \quad (14)$$

где n_1 и n_2 характеризуют число контактов в разных фракциях. Заметим, что, в отличие от сильно различной статической, *динамическая (акустическая) деформация $\tilde{\varepsilon}$ для обеих фракций одинакова* [30, 31]. Для предварительно сжатых контактов и не слишком сильных динамических деформаций $|\tilde{\varepsilon}| \ll \mu\varepsilon_0$, уравнение (14) может быть разложено в ряд Тейлора с коэффициентами $d^m \tilde{\sigma}(\varepsilon_0) / d\tilde{\varepsilon}^m$. Эти коэффициенты характеризуют линейный ($m = 1$) и нелинейные ($m = 2, 3, \dots$) упругие модули M_m среды, определяющие, соответственно, линейную скорость и нелинейные поправки к скорости распространения акустической волны и интенсивность других нелинейных эффектов m -го порядка:

$$M_m \sim \frac{d^m \tilde{\sigma}(\varepsilon_0)}{d\tilde{\varepsilon}^m} \sim bn_1 \left(1 + \frac{n_2}{n_1} \mu^{(3/2)-m} \right) \varepsilon_0^{(3/2)-m}. \quad (15)$$

Выражение (15) показывает, что вклад в линейный модуль M_1 от слабых контактов пропорционален $\mu^{1/2} \ll 1$ и пренебрежимо мал при $n_1 \sim n_2$. В противоположность этому, вклад слабой фракции в нелинейные модули M_m (с $m = 2, 3, \dots$) пропорционален $\mu^{(3/2)-m} \gg 1$. Следовательно, чем слабее контакт, тем сильнее может быть его вклад в нелинейные компоненты акустического сигнала. Таким образом, наблюдение нелинейных эффектов открывает уникальную возможность диагностики наиболее слабых межзеренных контактов в упаковках (с параметром нагружения $\mu \leq 10^{-1} \div 10^{-2}$ и еще меньшими по сравнению со средним значением силами). Такие малые силы находятся вне диапазона чувствительности известных методов (например, оптических на основе фотоупругого эффекта, механической оценки контактных сил на основе размера отпечатков зерен на дне контейнера и т. п.), так как все эти методы основаны на измерении величин, для которых вклад сильно сжатых контактов заведомо доминирует.

В нелинейных методиках очень полезными оказываются уже самые простые нелинейные эффекты, такие как генерация второй гармоники или самодетектирование (демодуляция) модулированной первичной волны,



Рис. 9. Амплитуда демодулированного сигнала как функция амплитуды сдвиговой волны накачки при ее горизонтальном распространении для вертикальной (*V*) и горизонтальной (*H*) поляризации. Четко видно амплитуды перехода показателя степени $2 \rightarrow 3/2$ (отмечены стрелками) при переходе в режим «хлопания» слабых контактов. Для *V*-поляризации в силу анизотропии нагруженности контактов переход происходит при большей амплитуде, а уровень детектированного сигнала, соответственно, ниже, чем для *H*-поляризации накачки [31].

волны накачки (для которой нелинейность не имеет четной составляющей, что понятно из соображений симметрии) в гранулированной среде позволяет легко наблюдать детектирование с одновременной конверсией в продольную моду (что симметрия не запрещает). В нескольких иных терминах можно сказать, что эта продольная составляющая возникает из-за дилатансии гранулированной упаковки при сдвиговом (акустическом в нашем случае) воздействии на нее. То есть измерение амплитуды детектированного сигнала от сдвиговой волны накачки позволяет исследовать амплитудный закон дилатансии в недоступном другим методам диапазоне малых (акустических) деформаций [31]. Наконец, использование сдвиговой волны с различной поляризацией позволяет непосредственно наблюдать анизотропию нагруженности контактов (проявляющуюся в анизотропии нелинейности среды). Эти эффекты проиллюстрированы на рис. 9.

Тот же эффект «хлопания» контактов при изменении амплитуды первичного пучка приводит к эффективному изменению длины области взаимодействия (так как в ее пределах часть контактов вблизи излучателя уже «хлопает», а более даленные контакты — еще нет). При этом из-за различия скоростей сдвиговой накачки и возникающей 2-й гармоники на продольной моде изменение амплитуды меняет условия синхронизма (эф-

причем даже эти эффекты демонстрируют «неклассические» черты в силу своего микроструктурного происхождения. Например, степенное разложение (15) теряет силу для очень умеренных акустических амплитуд ($\tilde{\epsilon}/\epsilon_0 \ll 1$, где ϵ_0 — макроскопическая статическая деформация среды), поскольку уже при $\tilde{\epsilon}/\epsilon_0 \sim \mu \ll 1$ слабые контакты переходят в хлопающий режим. В результате вместо квадратичного детектирования $\sim \tilde{\epsilon}^2$ в этом диапазоне амплитуд наблюдается детектирование на «хлопающей» герцевской нелинейности, дающей сигнал $\sim \tilde{\epsilon}^{3/2}$. По амплитуде перехода $2 \rightarrow 3/2$ можно оценить степень разгруженности слабых контактов и их долю в паковке. Кроме того, использование поперечной

фактивную длину взаимодействия), что приводит к биениям уровня гармоник с ростом амплитуды накачки (см. детали в [32, 33]). В оптике подобные биения связаны с изменением геометрической длины области взаимодействия.

В силу ограниченности объема отметим в качестве еще одного и последнего примера показательную лабораторную демонстрацию [34] возможности использования акустического аналога ЛГ-эффекта для высокочувствительного мониторинга тонких изменений микроструктуры гранулированной среды, возмущаемой короткими импульсными воздействиями (см. рис. 10).

На вставке показаны примеры спектров 100% амплитудно-модулированной волны накачки (7 кГц несущая, 30 Гц модуляция) и пробной волны (10 кГц) с наведенной модуляцией. Графики на рис. 10 показывают уровни основной гармоники пробной волны и модуляционных лепестков 1-го и 2-го порядков в процессе наблюдения. Как видно, основная гармоника почти не чувствует изменения микроструктуры среды, вызванные «сейсмическими событиями», в то время как модуляционные лепестки очень чувствительны к структурным изменениям, связанным прежде всего с изменением состояния наиболее «слабых мест» в среде. Медленные изменения уровня лепестков после прохождения возмущающих среду импульсов связаны с процессами медленной релаксации нарушенной сети слабых контактов в среде. Возможность уверенного наблюдения нелинейных эффектов для сейсмических волн в полевых условиях (например, эффектов самовоздействия [35]) на дистанциях более сотни метров свидетельствует о перспективности использования нелинейных методов в диагностических целях, по крайней мере на инженерно-сейсмических масштабах, для мониторинга состояния объектов типа плотин, фундаментов крупных сооружений и т. п. Более детальное обсуждение этих и ряда других аспектов акустической нелинейности гранулированных сред можно найти в работах [30—34] и цитированной в них литературе.

Заключение

О продолжающемся росте интереса к исследованию проявлений микроструктурно-обусловленной нелинейности свидетельствует значи-

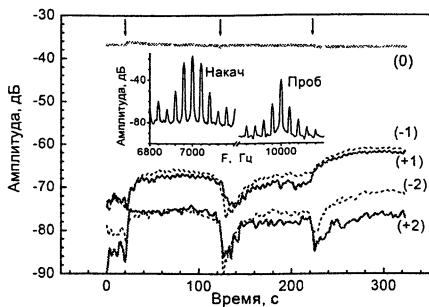


Рис. 10. Временные зависимости амплитуд основной гармоники пробной волны (отмечена цифрой 0), а также ее правых и левых модуляционных лепестков 1-го и 2-го порядков (отмечены цифрами ± 1 и ± 2). Стрелками отмечены моменты «сейсмических событий». На вставке — примеры спектров, прошедших через среду модулированной волны накачки и пробной волны с наведенной модуляцией.

тельный поток публикаций по этой тематике в последние годы, в том числе появление ряда обзорных работ (например, [2—4]). С учетом содержания [2—4] в предыдущих разделах был отобран материал, не перекрывающийся с этими обзорами, значительное внимание в которых было уделено гистерезисной нелинейности микронеоднородных сред. При этом обсуждение ряда интересных особенностей [37, 38], относящихся именно к гистерезисной нелинейности и не затронутых в [2—4], за неимением места в данном кратком обзоре пришлось опустить.

Представленные выше экспериментальные примеры и наглядные модели особенностей микронеоднородных сред показывают, как сочетание, казалось бы, хорошо известных факторов может приводить к очень необычным нелинейным акустическим эффектам и свойствам, вполне заслуживающим названий «аномальные» или «неклассические» (как, например, ярко выраженные амплитудно-зависимая диссипация или дисперсия нелинейности, возникающие за счет комбинации линейной релаксации на дефектах и их упругой нелинейности).

Работа подготовлена при поддержке грантов РФФИ № 05-02-17355, 05-05-16491 и 06-02-72550-НЦНИЛ.

Литература

1. Киттель Ч. Введение в физику твердого тела. М.: Наука, 1978. 791 с.
2. Guyer R., Johnson P. Nonlinear mesoscopic elasticity: Evidence for a new class of materials // *Physics Today*. 1999. № 4. P. 30—36.
3. Ostrovsky L. A., Johnson P. A. Nonlinear elasticity in geomaterials // *Rivista del Nuovo Chimento*. 2001. V. 24, № 7. P. 1—46.
4. Руденко О. В. Гигантские нелинейности структурно-неоднородных сред и основы методов нелинейной акустической диагностики // УФН. 2006. Т. 176, № 1. С. 97—102.
5. Zaitsev V. Yu. A model of anomalous acoustic nonlinearity of micro-inhomogeneous media // *Acoust. Lett.* 1996. V. 19, № 9. P. 171—176.
6. Беляева И. Ю., Зайцев В. Ю. Упругие нелинейные свойства микронеоднородных сред с иерархической структурой // *Акуст. журн* 1997. Т. 43, № 5. С. 594—599.
7. Беляева И. Ю., Зайцев В. Ю. О предельном значении параметра упругой нелинейности структурно-неоднородных сред // *Акуст. журн* 1998. Т. 44, № 6. С. 731—737.
8. Зайцев В. Ю., Назаров В. Е., Беляева И. Ю. Уравнение состояния микронеоднородных сред и частотная зависимость их упругой нелинейности // *Акуст. журн*. 2001. Т. 47, № 2. С. 220—226.
9. Zaitsev V., Sas P. Elastic moduli and dissipative properties of microinhomogeneous solids with isotropically oriented defects // *Acustica-Acta Acustica*. 2000. V. 86. P. 216—228.
10. Zaitsev V. Yu., Sas P. Dissipation in microinhomogeneous solids: inherent amplitude-dependent attenuation of a non-hysteretical and non-frictional type // *Acustica-Acta Acustica*. 2000. V. 86. P. 429—445.
11. Руденко О. В., Солуян С. И. Теоретические основы нелинейной акустики. М.: Наука, 1975. 287 с.
12. Зайцев В. Ю., Назаров В. Е. О линейной частотной зависимости коэффициента поглощения упругих волн в микронеоднородных средах // *Акуст. журн*. 1999. Т. 45, № 5. С. 552—556.
13. Зайцев В. Ю., Назаров В. Е., Шульга А. Е. Дисперсионные и диссипативные свойства микронеоднородных сред // *Акуст. журн*. 2000. Т. 46, № 3. С. 348—355.
14. Nazarov V. E., Zaitsev V. Yu., Belyaeva I. Yu. Nonlinear transformation of acoustic waves in microinhomogeneous media with relaxation // *Acustica-Acta Acustica*. 2002. V. 88. P. 40—49.

15. *Зайцев В. Ю., Матвеев Л. А.* Амплитудно-зависимая диссипация в микронеоднородных средах с линейным поглощением и упругой нелинейностью // *Геология и геофизика*. 2006. Т. 47, № 5. С. 695—710.
16. *Назаров В. Е., Радостин А. В., Соустова И. А.* Влияние интенсивной звуковой волны на акустические свойства резонатора из песчаника // *Акуст. журн.* 2002. Т. 48, № 1. С. 85—90.
17. *Gordon R. B., Davis L. A.* Velocity and attenuation of seismic waves in imperfectly elastic rock // *J. Geophys. Res.* 1968. V. 73(12). P. 3917—3935.
18. *Sharma M. M., Tutuncu A. N.* Grain Contact Adhesion Hysteresis: A Mechanism for Attenuation of Seismic Waves // *Geophys. Res. Lett.* 1994. V. 21, № 21. P. 2323—2326.
19. *Лебедев А. Б.* Амплитудно-зависимый дефект модуля упругости в основных моделях дислокационного гистерезиса // *ФТТ*. 1999. Т. 41, № 7. С. 1214—1221.
20. *Savage J. S.* Thermoelastic attenuation of elastic waves by cracks // *J. Geophys. Res.* 1966. V. 71(16). P. 3929—3938.
21. *Reasenberg P., Aki K.* A precise, continuous measurement of seismic velocity for monitoring in situ stress // *J. Geophys. Res.* 1974. V. 79(2). P. 399—406.
22. *Zaitsev V., Gusev V., Castagnède B.* Luxemburg-Gorky Effect Retooled for Elastic Waves: A Mechanism and Experimental Evidence // *Phys. Rev. Lett.* 2002. V. 89(10). P. 105502(1—4).
23. *Fillinger L., Zaitsev V., Gusev V., Castagnède B.* Nonlinear relaxational absorption/transparency for acoustic waves due to thermoelastic effect // *Acustica-Acta Acustica*. 2006. V. 92, № 1. P. 24—34.
24. *Saltykov V., Chebrov V., Kugaenko Yu., Sinityn V.* High-frequency seismic noise: Results of investigation in Kamchatka // *Phys. Chem. Earth*. 2006, V. 31. P. 132—137.
25. *Zaitsev V., Gusev V., Castagnède B.* Thermoelastic mechanism for logarithmic slow dynamics and memory in elastic wave interaction with individual cracks // *Phys. Rev. Lett.* 2003. V. 90(7). P. 075501(1—4).
26. *Zaitsev V., Nazarov V., Gusev V., Castagnède B.* Novel nonlinear-modulation acoustic technique for crack detection // *NDT&E Int.* 2006. V. 39, № 3. P. 184—194.
27. *TenCate J. A., Smith E., Guyer R. A.* Universal Slow Dynamics in Granular Solids // *Phys. Rev. Lett.* 2000. V. 85(5). P. 1020—1023.
28. *Зайцев В. Ю., Гусев В. Э., Назаров В. Е., Кастаньеде Б.* Взаимодействие акустических волн с трещинами: упругие и неупругие механизмы нелинейности с различными временными масштабами // *Акуст. журн.* 2005. Т. 51, доп. вып. «Геоакустика». С. 80—91.
29. *Ландау Л. Д., Лифшиц Е. М.* Теория упругости. М.: Наука, 1965. 204 с.
30. *Зайцев В. Ю.* Численное моделирование нелинейных упругих свойств зернистых сред с неидеальной упаковкой // *Акуст. журн.* 1995. Т. 41, № 3. С. 439—445.
31. *Tournat V., Zaitsev V., Gusev V., Nazarov V., Bequin P., Castagnède B.* Probing granular media by acoustic parametric emitting antenna: clapping contacts, nonlinear dilatancy and polarization anisotropy // *Phys. Rev. Lett.* 2004. V. 92(8). P. 085502(1—4).
32. *Tournat V., Gusev V., Zaitsev V., Castagnède B.* Acoustic second harmonic generation with mode conversion in granular media: nonlinear transformation of the emitting antenna // *Europhys. Lett.* 2004. V. 66(6). P. 798—804.
33. *Турна В., Зайцев В. Ю., Назаров В. Е., Гусев В., Кастаньеде Б.* Экспериментальное исследование нелинейных акустических эффектов в зернистых средах // *Акуст. журн.* 2005. Т. 51, № 5. С. 633—644.
34. *Zaitsev V., Nazarov V., Tournat V., Gusev V., Castagnède B.* Luxemburg-Gorky effect in a granular medium: probing perturbations of the material state via cross-modulation of elastic waves // *Europhys. Lett.* 2005. V. 70(5). P. 607—613.
35. *Зайцев В. Ю., Назаров В. Е., Таланов В. И.* Экспериментальное исследование самовоздействия сейсмоакустических волн // *Акуст. журн.* 1999. Т. 45, № 6. С. 799—806.
36. *Zaitsev V., Gusev V., Zaitsev Yu.* Mutually induced variations in dissipation and elasticity for oscillations in hysteretic materials: non-simplex interaction regimes // *Ultrasonics*. 2005. V. 43. P. 699—709.
37. *Zaitsev V., Gusev V.* On the differences between «adhesion-type» and «friction-type» hysteresis: theoretical description and experimental indications / CD-Proc. 8-eme Congrès Français d'Acoustique. Tours, 2006. P. 513—516.

ДИНАМИКА ВИХРЕЙ В СЛОИСТЫХ СВЕРХПРОВОДНИКАХ С ВНУТРЕННИМ ЭФФЕКТОМ ДЖОЗЕФСОНА

В. В. Курин, А. В. Чигинев

1. Введение

На протяжении многих лет высокотемпературные сверхпроводники (ВТСП) с сильной анизотропией являются предметом интенсивных исследований. Слоистое строение, внутренний эффект Джозефсона и сложный химический состав соединений данного типа обуславливают богатство и разнообразие физических свойств этих материалов. Большое внимание уделялось и продолжает уделяться изучению джозефсоновской динамики слоистых ВТСП.

В настоящее время считается общепринятым, что сильно анизотропные ВТСП типа BiSrCaCuO представляют собой естественные сверхрешетки распределенных джозефсоновских контактов, связанных между собой связями различных типов. Наиболее известным является магнитный тип связи за счет проникновения магнитного поля через сверхпроводящий слой [1]. В литературе описаны также зарядовая связь [2], комбинация магнитной и зарядовой связей [3], связь посредством квазичастиц [4], пространственная дисперсия в направлении поперек слоев [5], «глобальная» связь через внешнюю волноведущую систему [6] и т. д. Благодаря наличию связей многослойные джозефсоновские структуры обладают чрезвычайно сложной и богатой динамикой и огромным разнообразием нелинейных решений. В дальнейшем под многослойными или слоистыми структурами мы будем понимать как слоистые ВТСП, так и искусственные джозефсоновские сверхрешетки.

Особое место в исследованиях джозефсоновской динамики слоистых структур занимают исследования устойчивости различных конфигураций джозефсоновских вихрей. Большой интерес вызывает прямоугольная вихревая решетка, или синфазный режим движения джозефсоновских вихрей, сопровождаемый сильным излучением. Проведенные исследования устойчивости прямоугольной вихревой решетки показывают, что при учете только магнитной связи [7] данное решение является неустойчивым. Для стабилизации синфазного режима движения вихрей было предложено соединить слоистую структуру с внешней электродинамической системой (волноводом). Показано, что в этом случае устойчивость синфазного режима может быть достигнута [6]. Данный результат иллюстрирует тот факт, что линейные свойства системы уравнений, описывающей слоистую структуру, оказывают существенное влияние на свойства нелинейных решений этой системы.

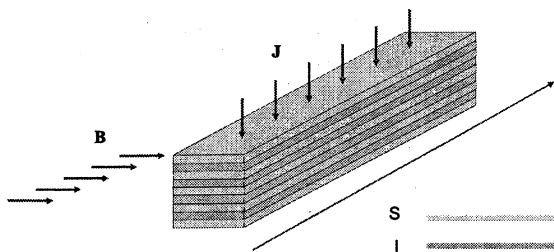


Рис. 1. Схематическое изображение слоистой джозефсоновской структуры

Настоящая статья посвящена демонстрации влияния свойств линейных волн в слоистой джозефсоновской структуре (рис. 1) на свойства движущихся решеток джозефсоновских вихрей, в частности на устойчивость синфазного режима движения вихрей. Подход, использованный здесь, основан на анализе системы уравнений общего вида, описывающей джозефсоновскую динамику слоистой структуры, в которой могут быть учтены, в принципе, любые типы связей и их комбинации.

2. Система нелинейных уравнений, описывающих слоистую джозефсоновскую структуру.

Типы связей

Простейшим примером структуры, в которой возможно формирование джозефсоновского вихря, является распределенный джозефсоновский контакт. В одномерном случае его динамика описывается уравнением типа синус-Гордона [8]:

$$\varphi_{tt} + \gamma\varphi_t - \varphi_{xx} + \sin \varphi = j_{ext},$$

где φ — джозефсоновская разность фаз. Данное уравнение выводится из уравнений Максвелла с учетом джозефсоновских соотношений. Происхождение нелинейного члена в уравнении синус-Гордона связано с синусоидальной зависимостью джозефсоновского сверхтока от разности фаз $j = j_c \sin \varphi$. Аналогичным образом записывается система уравнений, описывающая многослойную джозефсоновскую структуру. Приведем общий вид такой системы:

$$\hat{L}_{nm}\varphi_m + \sin \varphi_n = j_{ext}, \quad (1)$$

где \hat{L}_{nm} — линейный оператор, φ_n — джозефсоновская разность фаз на n -м контакте структуры; запись $\hat{L}_{nm}\varphi_m$ означает

$$\hat{L}_{nm}\varphi_m \equiv \sum_{m=0}^l \int_0^t dt' \int_{-\infty}^{\infty} dx' L_{n-m}(x-x', t-t') \varphi_m(x', t'). \quad (2)$$

Для исследования слоистых джозефсоновских структур обычно используется подход, основанный на теории длинных линий. Предполагая, что характерные масштабы изменения $\varphi_m(x, t)$ много больше соответствующих масштабов ядра L интегрального оператора в (2), разложим ядро $L_{n-m}(x, t)$ по степеням x и t и получим линейный оператор \hat{L}_{nm} в виде

$$\hat{L}_{nm} = (\hat{C}\hat{t}^{-1})_{nm}\partial_t^2 + (\hat{r}^{-1}\hat{t}^{-1})_{nm}\partial_t + \hat{G}\partial_t - (\hat{K}^{-1})_{nm}\partial_x^2, \quad (3)$$

где коэффициенты \hat{C} , \hat{t} , \hat{r} , \hat{K} , \hat{G} являются соответственно операторами погонной емкости, нелокальной зарядовой связи, сопротивления, магнитной связи, «глобальной» связи. Мы будем пользоваться выражением (3), имея в виду, однако, что операторы \hat{C} , \hat{t} , \hat{r} , \hat{K} , \hat{G} сами, вообще говоря, являются функциями ω , k , q . При этом разложение линейного оператора является в известной степени условным, позволяющим, тем не менее, придать операторам, входящим в выражение для \hat{L}_{nm} , указанный выше физический смысл.

Оператор \hat{K} , входящий в выражение (3), связывает джозефсоновскую разность фаз со сверхтоком, текущим вдоль слоев:

$$\sum_m K_{nm} I_m = \Phi_0 (2\pi)^{-1} \partial_x (\varphi_n - \varphi_{n-1}).$$

С его помощью может быть описана связь между соседними слоями за счет проникновения магнитного поля сквозь сверхпроводящий слой, или магнитная связь [1]. Оператор \hat{t} связывает напряжение в n -м контакте структуры с джозефсоновской разностью фаз посредством формулы

$$\sum_m t_{nm} V_m = \Phi_0 (2\pi c)^{-1} \partial_t \varphi_n,$$

являющейся обобщением нестационарного джозефсоновского соотношения. С помощью данной формулы может быть описана связь между соседними контактами за счет неполной экранировки продольного электрического поля сверхпроводящим слоем, или зарядовая связь [2]. Оператор сопротивления \hat{r} входит в соотношение, связывающее напряжение в n -м контакте с током квазичастиц через n -й контакт

$$\sum_m r_{nm}^{-1} V_m = j_n^{qp};$$

этим соотношением может быть описана квазичастичная связь между контактами [4]. Оператор емкости \hat{C} связывает плотность заряда на слое n с напряжением на n -м контакте

$$\sum_m C_{nm} V_m = Q_n$$

и может описывать связь между контактами за счет продольного фонона, поляризованного поперек слоев [5]. Оператор \hat{G} описывает «глобальную» связь через внешнюю электродинамическую систему [6].

В предположении об отсутствии магнитной, зарядовой, квазичастичной или фононной связей операторы \hat{K} , \hat{t} , \hat{r} , \hat{C} соответственно являются диагональными по номерам слоев. В отсутствие «глобальной» связи за счет внешней электродинамической системы $\hat{G} = 0$.

Наконец, выпишем выражение для оператора \hat{L}_{nm} в (ω, k, q) -представлении, которое понадобится нам в дальнейшем:

$$L(\omega, k, q) = \sum_n \int_0^\infty dt \int_{-\infty}^\infty dx L_n(x, t) e^{i\omega t - ikx - iqn}. \quad (4)$$

3. Нелинейные решения в виде решеток джозефсоновских вихрей

Система уравнений (1) имеет нелинейные решения, описывающие движущиеся решетки джозефсоновских вихрей в слоистой структуре. В сильном внешнем магнитном поле, т. е. когда расстояние между соседними вихрями не превышает размеров самих вихрей, эти решения могут быть записаны в виде

$$\varphi_n = hx - \Omega t + \psi_n + \theta_n, \quad (5)$$

где h — безразмерное внешнее магнитное поле; $\Omega = hu$ — частота, пропорциональная скорости вихревой решетки u ; $\psi_n, |\psi_n| \ll 1$ — малые добавки; θ_n — плавные функции, медленно меняющиеся на масштабах h^{-1} и Ω^{-1} .

Подставляя решение (5) в систему (1) и используя стандартную процедуру усреднения, можно получить систему укороченных уравнений для θ_n . Данная система имеет решения в виде наборов не зависящих от координаты и времени констант, соответствующие установившемуся режиму движения вихрей в структуре. Разные наборы таких констант соответствуют разным типам вихревых решеток. Типичные вихревые решения системы показаны на рис. 2 и 3. Для системы, состоящей из трех контактов ($N = 3$), возможны решения в виде прямоугольной вихревой решетки (синфазный режим), треугольной решетки, а также в виде наклонных рядов джозефсоновских вихрей (см. рис. 2). Для $N = 4$ существует много различных решений, однако, как и для случая $N = 3$, есть решения, описывающие синфазный режим и треугольную вихревую решетку (см. рис. 3).

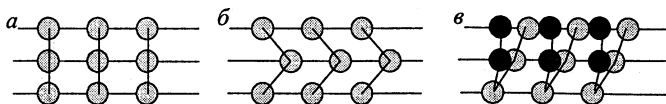


Рис. 2. Типичное взаимное расположение джозефсоновских вихрей в системе из трех слоев: *a* — прямоугольная вихревая решетка (синфазный режим), *б* — треугольная решетка, *в* — решение в виде наклонных рядов.

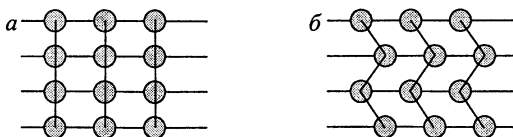


Рис. 3. Взаимное расположение джозефсоновских вихрей в системе из четырех слоев: *a* — синфазный режим, *б* — треугольная решетка.

Если известен набор θ_n , описывающих какую-либо вихревую решетку, то, используя выражение для высокочастотных возмущений

$$\psi_n = \text{Re} \left(iN^{-1} \sum_{q,m} L^{-1}(\Omega, h, q) e^{i\theta_m} e^{iq(n-m)} e^{i(hx - \Omega t)} \right), \quad (6)$$

можно найти вольт-амперную характеристику (ВАХ) слоистой джозефсоновской структуры с движущейся вихревой решеткой

$$\sum_{m,n} \hat{L}_{nm}(hx - \Omega t) - \text{Im} \left((2N)^{-1} \sum_{q,m,n} L^{-1}(\Omega, h, q) e^{i(\theta_m - \theta_n)} e^{iq(n-m)} \right) = Nj_{ext}, \quad (7)$$

а также переменное напряжение на структуре

$$V = \Omega \text{Re} \left(L^{-1}(\Omega, h, 0) \sum_m (\hat{t}^{-1})_{nm} \sum_k e^{i\theta_k} e^{i(hx - \Omega t)} \right). \quad (8)$$

Как видно из формулы (8), наибольшая амплитуда переменного напряжения на слоистой структуре достигается в том случае, когда все θ_n равны друг другу. Это решение соответствует тому, что вихри в соседних слоях структуры расположены точно один над другим (см. рис. 2, *a*, рис. 3, *a*). Такая вихревая конфигурация называется прямоугольной вихревой решеткой, или синфазным режимом движения джозефсоновских вихрей.

Примеры ВАХ слоистой структуры с движущейся решеткой джозефсоновских вихрей показаны на рис. 4. Линиями 1 и 2 изображены ВАХ для случаев, когда в структуре сформированы соответственно треугольная и прямоугольная вихревые решетки. Пики на ВАХ соответствуют характеристическим скоростям линейных мод в системе (см. далее разд. 5), причем для случая, когда доминирует магнитная связь, пик для треугольной решетки имеет место при более низком напряжении

(см. рис. 4, *a*). Для случая же, когда преобладает зарядовая связь, при наиболее низком напряжении расположен пик, соответствующий синфазному режиму движения джозефсоновских вихрей (см. рис. 4, *б*). Забегая вперед, скажем, что именно этим обстоятельством объясняется выход на устойчивый синфазный режим в случае, когда симметричная мода линейных волн в структуре является самой медленной.

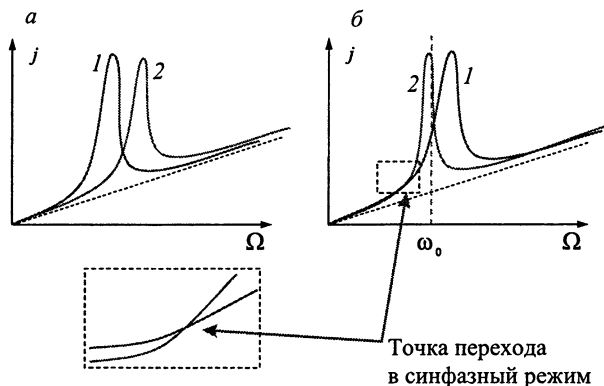


Рис. 4. ВАХ многослойной джозефсоновской структуры с движущейся решеткой джозефсоновских вихрей в случае доминирования магнитной (*a*) или зарядовой (*б*) связи. Кривые 1 и 2 построены при условии, что в структуре сформированы треугольная и прямоугольная вихревые решетки. Отдельно вынесен и увеличен участок, соответствующий спонтанному переходу от треугольной решетки к синфазному режиму.

Итак, уравнение (1) имеет множество нелинейных решений, описывающих решетки джозефсоновских вихрей. Однако для того, чтобы какое-либо решение было реализовано на практике, оно должно быть устойчивым. Ниже мы покажем, как анализируется устойчивость одного из наиболее интересных решений — синфазного режима движения джозефсоновских вихрей.

4. Устойчивость синфазного режима движения джозефсоновских вихрей

Синфазный режим движения джозефсоновских вихрей в слоистой структуре (см. рис. 2, *a*, рис. 3, *a*) сопровождается сильным излучением и может быть использован при построении генераторов электромагнитного излучения на ВТСП в диапазоне вплоть до частоты энергетической щели, составляющей величину порядка 10 ТГц. Поэтому изучение синфазного режима движения вихрей в ВТСП представляется важным с точки зрения

приложений. В настоящем разделе мы описываем метод анализа нелинейных решений на устойчивость на примере синфазного режима.

В сильном внешнем магнитном поле распределение джозефсоновской разности фаз в синфазном режиме выглядит следующим образом:

$$\varphi_n^0 = \varphi^0 = hx - \Omega t + \text{Im} \frac{\exp i(hx - \Omega t)}{L(\Omega, h, 0)}. \quad (9)$$

Проанализируем это решение на устойчивость. Для этого будем искать возмущенное решение системы уравнений (1) в виде суммы решения, описывающего синфазный режим, и малой поправки: $\varphi_n^0 = \varphi^0 + \delta\varphi_n$. Подставляя это решение в исходное уравнение (1) и диагонализуя линейный оператор в этом уравнении с помощью линейного преобразования $\chi_q = \sum_n \delta\varphi_n \exp(iqn)$, получаем уравнения для малых поправок

$$\hat{L}_q \chi_q + \chi_q \cos \varphi^0 (hx - \Omega t) = 0, \quad (10)$$

где линейный оператор $\hat{L}_q = \sum_n \hat{L}_n e^{-iqn}$.

Опишем кратко решение уравнения (10) для случая, когда \hat{L}_q является оператором Даламбера $\hat{L}_q = \partial_t^2 - v^2(q) \partial_x^2$, где $v(q)$ — характеристическая скорость q -й моды (см. также разд. 5). Данный частный случай был подробно рассмотрен в [6]. Поскольку уравнение (10) содержит периодический множитель, его решение может быть представлено в виде разложения по собственным модам с определенными квазиимпульсами и квазичастотами, связанными дисперсионной зависимостью. Ввиду малости периодического множителя возмущения, испытываемые спектром из-за брэгговских резонансов, будут особенно заметны вблизи центра и краев зоны Бриллюэна и могут быть вычислены во втором порядке по малому параметру $\mu \sim h^{-2}$. На рис. 5 показаны возможные варианты расщепления спектра решений уравнения (9) в центре зоны Бриллюэна и на ее краях. Ход дисперсионной характеристики, обозначенный на рис. 5 цифрой 1, соответствует устойчивости синфазного режима по отношению к длинноволновым возмущениям. Поведение кривой, обозначенное цифрой 2, соответствует неустойчивости по отношению к длинноволновым возмущениям. Цифрой 3 обозначено расщепление дисперсионной кривой, соответствующее устойчивости по отношению к возмущениям с малой длиной волны, и, наконец, цифрой 4 — ход кривой, соответствующий коротковолновой неустойчивости.

С помощью метода, описанного выше, в работе [7] показано, что в структуре, состоящей из двух распределенных джозефсоновских контактов, связанных только магнитной связью, синфазный режим движения джозефсоновских вихрей является неустойчивым. Чтобы обеспечить устойчивость синфазного режима движения джозефсоновских вихрей,

в работе [6] предложено соединить многослойную структуру с внешней электродинамической системой (ввести «глобальную» связь), т. е. соответствующим образом подкорректировать оператор $L(\omega, k, q)$. Исследование устойчивости синфазного режима методом, описанным выше, показало, что «глобальная» связь способна при определенных условиях обеспечить устойчивость данного режима [6].

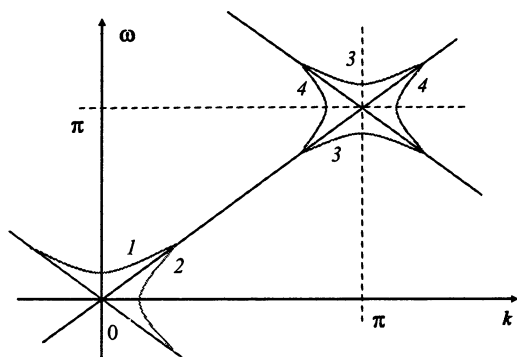


Рис. 5. Возможные варианты поведения дисперсионной характеристики линейных волн возмущений прямоугольной вихревой решетки, соответствующие: 1 — длинноволновой устойчивости, 2 — длинноволновой неустойчивости, 3 — коротковолновой устойчивости, 4 — коротковолновой неустойчивости.

Численный эксперимент продемонстрировал спонтанное установление синфазного режима движения джозефсоновских вихрей в многослойной структуре, охваченной «глобальной» связью [6]. При достаточно малом внешнем токе дисперсионная кривая малых возмущений соответствует длинноволновой неустойчивости (см. рис. 5, кривая 2). Это означает, что вихри в соседних слоях отталкиваются друг от друга и образуют треугольную решетку (рис. 6, а). При повышении внешнего тока и достижении им некоторого значения длинноволновая неустойчивость сменяется устойчивостью (см. рис. 5, кривая 1) и отталкивание между соседними вихрями сменяется притяжением (рис. 6, б). Две вихревые подрешетки в структуре сдвигаются друг относительно друга (рис. 6, в), образуя прямоугольную вихревую решетку, или синфазный режим движения джозефсоновских вихрей (рис. 6, г).

Таким образом, введение «глобальной» связи способно привести к устойчивости синфазного режима. Этот результат указывает на то, что свойства нелинейных решений, в частности устойчивость того или иного режима движения вихрей, определяются линейным оператором в системе (1). Свойства линейного оператора могут быть выяснены путем решения линеаризованной системы уравнений (1). В двух последующих разделах

мы исследуем линейные решения (1) и приведем конкретный пример построения линейного оператора \hat{L}_{nm} в континуальном пределе.

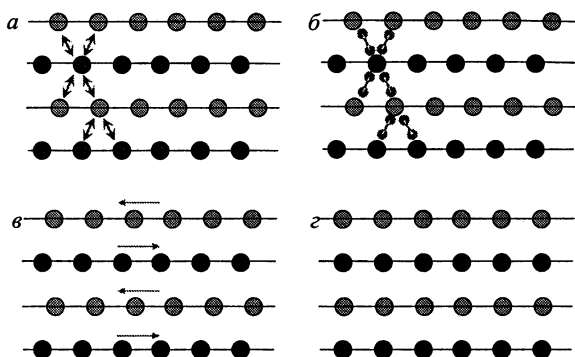


Рис. 6. Спонтанное установление синфазного режима движения джозефсоновских вихрей: *a* — треугольная вихревая решетка — вихри в соседних слоях отталкиваются; *b* — смена отталкивания притяжением; *в* — сдвиг двух вихревых подрешеток относительно друг друга; *г* — синфазный режим движения джозефсоновских вихрей.

5. Линейные волны в многослойных джозефсоновских структурах

Исследуем дисперсионные характеристики линейных волн в слоистой структуре в предположении о преобладании того или иного типа связи в системе.

Система уравнений (1) имеет решение в виде линейных волн малой амплитуды. Если в системе (1) положить $|\varphi_n| \ll 1$ и $\varphi_n(x, t) \sim \exp(ikx + iqn - i\omega t)$, а также $j_{ext} = 0$, то дисперсионная характеристика линейных волн получится как условие существования нетривиального решения системы (1):

$$L(\omega, k, q) + 1 = 0. \quad (11)$$

Дисперсионная характеристика может быть описана семейством поверхностей волновых векторов, или линий постоянной частоты $\omega(k, q) = \text{const}$. Форма линии постоянной частоты зависит от того, какой тип связи имеет место в слоистой структуре. На рис. 7 изображены варианты линии постоянной частоты в пределе малых q . Если в системе преобладает магнитная связь, то линии постоянной частоты имеют форму гипербол (см. рис. 7, *a*). Если же в структуре доминирует зарядовая связь, то линии постоянной частоты имеют форму эллипсов (см. рис. 7, *б*).

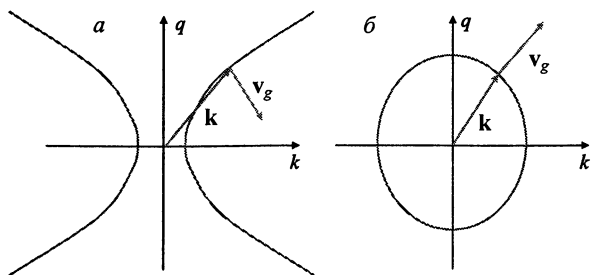


Рис. 7. Линии постоянной частоты дисперсионной характеристики при преобладании магнитной (а) или зарядовой (б) связи. Стрелками обозначены волновой вектор и вектор групповой скорости.

Если операторы K, C и t (см. формулу (3)) не зависят от ω , а связь через квазичастицы и «глобальная» связь отсутствуют ($r = G = 0$), то дисперсионная кривая $\omega(k)$ при фиксированном q имеет форму гиперболы:

$$\omega^2 - v^2(q)k^2 = \omega_j^2(q), \quad (12)$$

где $\omega_j(q)$ — плазменная частота, $v(q)$ — характеристическая скорость q -й моды. Зависимость характеристической скорости $v(q)$ от поперечного волнового числа q определяется, в частности, тем, какой тип связи — магнитный или зарядовый — является доминирующим. При преобладании магнитной связи характеристическая скорость линейной моды убывает с ростом ее номера (рис. 8, а). Такая ситуация является типичной для слоистых сверхпроводников. В противоположном случае, при преобладании зарядового типа связи, характеристическая скорость растет с ростом номера моды (рис. 8, б). Как будет продемонстрировано далее, такая ситуация может иметь место в слоистых сверхпроводниках вблизи частот оптических фононов.

Рассмотрим теперь влияние «глобальной» связи на спектр характеристических скоростей линейных мод в системе. Эту связь можно обеспечить, соединив многослойную джозефсоновскую структуру с внешней электродинамической волноведущей системой.

Из выражения (3) видно, что оператор \hat{G} , описывающий влияние электродинамической системы на слоистую структуру, воздействует только на симметричную моду линейных волн в системе ($q = 0$). Меняя параметры внешней системы, можно менять характеристическую скорость симметричной моды и, в частности, сделать ее самой медленной (рис. 9).

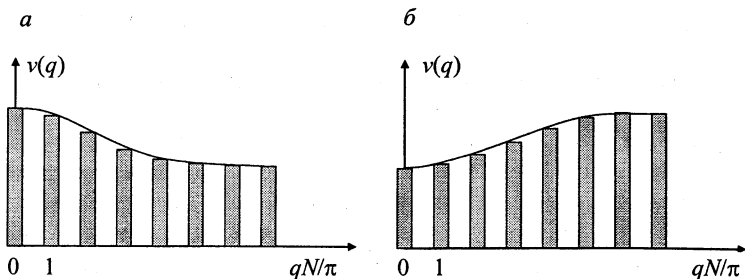


Рис. 8. Спектр характеристических скоростей в зависимости от номера линейной моды в случае доминирования магнитной (а) и зарядовой (б) связи между слоями.

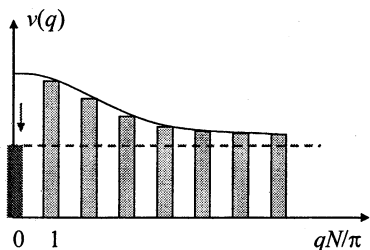


Рис. 9. Влияние внешней замедляющей системы на скорость симметричной линейной моды в слоистой структуре.

В работе [6] показано, что если сделать симметричную линейную моду самой медленной (см. рис. 9), то синфазный режим движения вихрей станет устойчивым. Данный результат не зависит от природы замедления симметричной моды и является справедливым также и в случае преобладания зарядовой связи (см. рис. 8, б). Ниже мы покажем, что замедление симметричной моды Свихарта достигается и в ВТСП вблизи частоты оптического фонона. Это наводит на мысль о том, что взаимодействие джозефсоновских вихрей с оптическими фононами способно приводить к устойчивости синфазного режима движения вихрей в ВТСП.

6. Влияние оптических фононов на движение джозефсоновских вихрей в ВТСП

Приведем пример построения линейного оператора в уравнении (1) для слоистого сверхпроводника с учетом оптических фононов и проанализируем дисперсионную характеристику связанных линейных мод в такой системе. На основе анализа дисперсионной характеристики мы покажем, что вблизи частоты оптического фонона может быть достигнуто замедление симметричной моды Свихарта в ВТСП, что должно приво-

доть к устойчивости синфазного режима движения джозефсоновских вихрей в ВТСП.

Как известно [9], в слое ВТСП существуют «мягкие» фоновые моды с частотами, меньшими частоты энергетической щели в слоистых сверхпроводниках. Такие фононы при наличии у них ненулевого дипольного момента могут взаимодействовать с джозефсоновскими колебаниями в системе, а значит, и с движущейся решеткой джозефсоновских вихрей. Для учета влияния оптических фононов мы построим оператор $L(\omega, k, q)$ в континуальном пределе, когда длина волны много больше расстояния между слоями. В этом случае слоистый сверхпроводник представляет собой анизотропную среду, дисперсионная характеристика волн в которой описывается уравнением Френеля

$$\det \left\| \frac{\omega^2}{c^2} \varepsilon_{ij}(\omega, \mathbf{k}) + k_i k_j - k^2 \delta_{ij} \right\| = 0. \quad (13)$$

Опишем построение диэлектрической проницаемости $\hat{\varepsilon}$ слоистого сверхпроводника, учитывающей вклады от электронной $\hat{\chi}_e$ и ионной $\hat{\chi}_{ph}$ подсистем сверхпроводника $\hat{\varepsilon} = \hat{I} + 4\pi\hat{\chi}_e + 4\pi\hat{\chi}_{ph}$. Найдем вклад электронной подсистемы слоистого сверхпроводника в рамках гидродинамической модели анизотропной плазмы. Запишем гидродинамические уравнения

$$\dot{n} + n_0 \operatorname{div} \mathbf{v} = 0, \quad \hat{m} \dot{\mathbf{v}} + \varepsilon_F n_0^{-1} \nabla n = -e \mathbf{E}, \quad (14)$$

с анизотропной массой $m_{xx} = m\Gamma^{-1}$, $m_{zz} = m\Gamma$, $m_{xz} = m_x = 0$, где Γ — фактор анизотропии. Из уравнений (14) находится отклик электронной системы сверхпроводника:

$$\hat{\chi}_e = \frac{-\omega_p^2}{4\pi\omega^2(\omega^2 - \Gamma v_F^2 k_x^2 - \Gamma^{-1} v_F^2 k_z^2)} \begin{pmatrix} \Gamma\omega^2 - v_F^2 k_z^2 & v_F^2 k_x k_z \\ v_F^2 k_x k_z & \Gamma^{-1}\omega^2 - v_F^2 k_x^2 \end{pmatrix}, \quad (15)$$

где $\omega_p^2 = 4\pi e^2 n_0 / m$ — квадрат плазменной частоты электронов, $v_F^2 = \varepsilon_F / m$ — квадрат скорости Ферми.

Формула (15) выражает вклад только сверхпроводящих электронов в диэлектрическую проницаемость среды. Для нахождения вклада нормальных электронов можно воспользоваться теми же уравнениями сплошной среды (14), только с учетом столкновений и с другим фактором анизотропии. Будем пренебрегать нормальными электронами, считая температуру сверхпроводника достаточно низкой, так что концентрация квазичастиц исчезающе мала.

Для того чтобы найти вклад в диэлектрическую проницаемость от фоновой подсистемы слоистого сверхпроводника, воспользуемся уравнением движения иона в адиабатическом приближении Борна — Оппенгеймера:

$$M_\nu \ddot{z}_N^\nu = - \sum_{\mu, M} \hat{G}_{N-M}^{\nu\mu} z_M^\mu + q_\nu E_N^\nu. \quad (16)$$

Учет отличия действующего поля E_N^ν от среднего макроскопического поля E_N производится с помощью тензора Лоренца \hat{L}_N^ν :

$$E_N^\nu = E_N + 4\pi \sum_{\mu, M} q_\mu \hat{L}_{N-M}^{\nu\mu} z_M^\mu. \quad (17)$$

С учетом данного отличия уравнение движения иона запишется в виде

$$M_\nu \ddot{z}_N^\nu = - \sum_{\mu, M} \hat{F}_{N-M}^{\nu\mu} z_M^\mu + q_\nu E_N, \quad (18)$$

где $\hat{F}_{N-M}^{\nu\mu} = \hat{G}_{N-M}^{\nu\mu} - 4\pi q_\mu \hat{L}_{N-M}^{\nu\mu} q_\nu$.

Из уравнения (18) можно получить выражение для фоновой восприимчивости слоистого сверхпроводника:

$$\hat{\chi}_{ph} = \frac{1}{V} \sum_{a=1}^{3L} \frac{1}{-\omega^2 + \Omega^2(\mathbf{k}, a)} \frac{\sum_{\nu, \mu} q_\nu q_\mu \mathbf{e}_\nu \otimes \mathbf{e}_\mu^*}{\sum_\nu M_\nu |\mathbf{e}_\nu|^2}. \quad (19)$$

Фоновые частоты $\Omega(\mathbf{k}, a)$ и вектора поляризации фононов $\mathbf{e}_\nu(\mathbf{k}, a)$ находятся из уравнения на собственные числа и собственные векторы тензора $\hat{F}^{\nu\mu}(\mathbf{k})$

$$[-M_\nu \Omega^2(\mathbf{k}, a) \delta_{\nu\mu} \hat{I} + \hat{F}^{\nu\mu}(\mathbf{k})] \mathbf{e}_\nu(\mathbf{k}, a) = 0. \quad (20)$$

Для иллюстрации влияния фононов на дисперсию линейных мод в слоистом сверхпроводнике, а следовательно, и на динамику джозефсоновских вихрей в системе, рассмотрим простейший случай. Будем считать, что, во-первых, пространственная дисперсия электронов отсутствует ($v_F \rightarrow 0$) и, во-вторых, в системе есть один диагональный фонон с фиксированной частотой ω_0 :

$$\hat{\chi}_{ph} = - \frac{\omega_{pi}^2}{4\pi(\omega^2 - \omega_0^2)} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}. \quad (21)$$

Подставляя $\hat{\epsilon}$ с учетом (21) в уравнение (13), получаем уравнение Френеля 2-го порядка по \mathbf{k} , поэтому поверхности волновых векторов, описываемые данным уравнением, имеют форму гиперболы или эллипса.

Опишем последовательность бифуркаций поверхности волновых векторов уравнения (13) с увеличением частоты (рис. 10). Вначале идет область 1, в которой нет решений дисперсионного уравнения для действительных k_x, k_z . Затем начиная с частоты порядка плазменной частоты электронов поперек слоев $\Gamma^{-1/2} \omega_p$ идет область 2, в которой линии постоянной частоты являются гиперболами. Вблизи фоновой частоты ω_0

есть узкая область 3, в которой изочастоты имеют форму эллипсов. При дальнейшем повышении частоты последовательно идут область 4, где нет решения для действительных k_x, k_z , область 5 с гиперболическими поверхностями волновых векторов, и, наконец, при частотах, больших плазменной частоты электронов вдоль слоев $\Gamma^{1/2}\omega_p$, вновь появляется область 6 с эллиптическими линиями постоянной частоты.

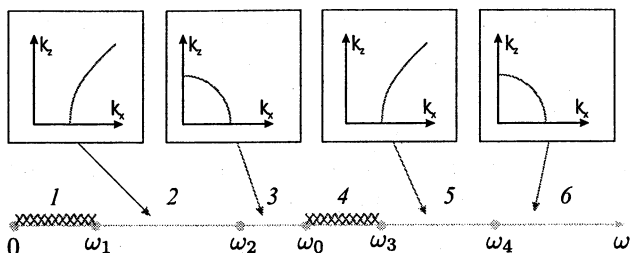


Рис. 10. Изменение формы линии постоянной частоты в зависимости от частоты. Штриховкой показаны области, где нет решений для действительных k_x, k_z .

Как уже отмечалось выше, изочастоты в виде гиперболы (см. рис. 7, а) соответствуют тому, что в системе преобладает магнитная связь и симметричная мода линейных волн обладает наибольшей характеристической скоростью (см. рис. 8, а). И напротив, линии постоянной частоты в виде эллипса (см. рис. 7, б) соответствуют преобладанию зарядовой связи и тому, что симметричная мода обладает наименьшей характеристической скоростью (см. рис. 8, б), т. е. устойчивости синфазного режима движения джозефсоновских вихрей в структуре. Мы видим, что вблизи частоты оптического фона характер ранее преобладавшей магнитной связи изменился таким образом, что спонтанное установление синфазного режима движения вихрей стало возможным.

7. Заключение

Свойства линейного оператора в уравнении, описывающем джозефсоновскую динамику многослойных структур и слоистых сверхпроводников, определяющим образом влияют на свойства джозефсоновских вихревых решеток в данных слоистых структурах. Например, спонтанное установление синфазного режима движения джозефсоновских вихрей с сильным излучением в многослойной структуре возможно в том случае, если симметричная мода линейных волн в системе будет обладать наименьшей характеристической скоростью. Это возможно либо в случае сильного замедления этой моды внешней замедляющей системой, либо вблизи частот оптических фононов в слоистом сверхпроводнике при ус-

ловии их достаточно высокой добротности. В последнем случае происходит эффективное изменение характера магнитной связи между соседними джозефсоновскими контактами из-за взаимодействия с оптическим фононом.

Работа выполнена при поддержке РФФИ (грант № 06-02-16592а), а также программ РАН «Нелинейная динамика», «Квантовая макрофизика», «Проблемы радиофизики».

Литература

1. Sakai S., Bodin P., Pedersen N. F. Fluxons in thin-film superconductor-insulator superlattices // *J. Appl. Phys.* 1993. V. 73, № 5. P. 2411—2418; Bulaevskii L. N. et al. Time-dependent equations for phase differences and a collective mode in Josephson-coupled layered superconductors // *Phys. Rev. B* 1994. V. 50, № 17. P. 12831—12834.
2. Koyama T., Tachiki M. I-V characteristics of Josephson-coupled layered superconductors with longitudinal plasma excitations // *Phys. Rev. B* 1996. V. 54, № 22. P. 16183—16191.
3. Kim Ju H., Pokharel J. Collective Josephson vortex dynamics in long Josephson junction stacks // *Physica C* 2003. V. 384. P. 425—436.
4. Ryndyk D. A. Quasiparticle dynamics and phase locking in a S-I-S multilayer Josephson junction // *Pis'ma v ZhETF*. V. 65, № 10. P. 755—759; Ryndyk D. A. Collective dynamics of intrinsic Josephson junctions in high- T_c superconductors // *Phys. Rev. Lett.* 1998. V. 80, № 15. P. 3376—3379; Рындык Д. А. Неравновесный эффект Джозефсона в системах туннельных сверхпроводниковых контактов и в слоистых сверхпроводниках // *ЖЭТФ*. 1999. Т. 116, вып. 5(11). С. 1—20.
5. Helm Ch., Preis Ch., Walter Ch., Keller J. Theory for the coupling between longitudinal phonons and intrinsic Josephson oscillations in layered superconductors // *Phys. Rev. B* 2000. V. 62, № 9. P. 6002—6014.
6. Chiginev A. V., Kurin V. V. In-phase vortex flow and superradiance in Josephson superlattice embedded in a waveguide // *Phys. Rev. B* 2004. V. 70, № 21. P. 214523-1—214523-10.
7. Chiginev A. V., Kurin V. V. Instability of a rectangular vortex lattice in a stack of two long Josephson junctions // *Phys. Rev. B* 2002. V. 66, № 5. P. 052510-1—052510-4.
8. Barone A., Paternó G. Эффект Джозефсона. Физика и применения. М.: Мир, 1984. 640 с.
9. Prade J. et al. Calculation of Raman- and infrared-active modes of $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$ // *Phys. Rev. B* 1989. V. 39, № 4. P. 2771—2774.

ВЫНУЖДЕННОЕ РАССЕЯНИЕ МАНДЕЛЬШТАМА — БРИЛЛЮЭНА ЛАЗЕРНЫХ МОД ГАУССА — ЛАГЕРРА: НОВЫЕ ЯВЛЕНИЯ

Ф. А. Стариков

Введение

Оптические преобразования лазерных пучков с винтовой дислокацией волнового фронта (вихревых пучков) [1], в частности мод Гаусса — Лагерра, привлекают большое внимание в связи с их потенциальным применением при оптической обработке информации, тонких технологиях и в связи с проблемой турбулентной атмосферы (см., например, [2—5]). Интерес к обращению волнового фронта (ОВФ) таких пучков усиливается тем, что их трудно корректировать с помощью гибких адаптивных

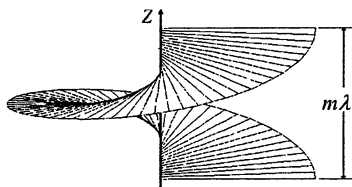


Рис. 1. Фазовая поверхность оптического вихря.

зеркал ввиду разрывной фазовой поверхности [2, 6] (рис. 1). ОВФ вихревых пучков при четырехволновом смешении в фоторефрактивном кристалле демонстрировалось в [7]. Представляется интересным получить ОВФ при вынужденном рассеянии Манделъштама — Бриллюэна (ВРМБ) вихревых пучков. ОВФ при ВРМБ обеспечивает автоматическую

юстировку лазерной системы, не требует сложной оптической настройки и дорогостоящих оптических компонентов. Предполагалось [2], что объемное ВРМБ-зеркало в отличие от непрерывных изгибаемых зеркал может решить проблему ОВФ пучка с дислокациями волнового фронта. Это отвечало представлению об универсальности эффекта ОВФ при ВРМБ (по крайней мере, для плоскополяризованных волн), основанного на преимущественном усилении сопряженной стоксовой волны в поле лазерной накачки [6].

Однако анализ по теории возмущений и расчеты при линейном стационарном [8, 9] и нестационарном [10] ВРМБ вихревых лазерных пучков (мод Гаусса — Лагерра), непосредственно сфокусированных в ВРМБ-среду, привели к обнаружению новых интересных явлений. Во-первых, было показано, что ОВФ вихревых пучков не наблюдается как явление из-за отсутствия селекции обращенной стоксовой моды. Это происходит вследствие того, что коэффициенты усиления обращенной стоксовой моды и аналогичной моды с обратной спиральностью одинаковы. Поэтому (например, в случае простейшей торообразной моды LC_0^1) стоксов пучок представляет собой случайную комбинацию нескольких мод, включая

сопряженную. Во-вторых, для достаточно широкого класса вихревых пучков (например, в случае лазерной моды с двумя кольцами LG_1^1) наблюдается явление, которое может быть названо преобразованием волнового фронта (ПВФ) при ВРМБ. Его суть заключается в том, что в стоксовом пучке селекционируется единственная мода, но она является основной гауссовой модой, которая ортогональна лазерной моде. Расчеты показали, что при ВРМБ вихревых лазерных пучков происходит генерация вихревых гиперзвуковых волн в ВРМБ-среде, структура которых наиболее близка к структуре лазерной волны в случае ПВФ. Экспериментальное формирование оптического вихря (моды LG_0^1) с помощью специальных киноформных спиральных фазовых пластин было проведено в [11—14]. В результате предсказанное явление отсутствия ОВФ при ВРМБ вихревого пучка было полностью подтверждено в экспериментах [15, 16]. Расчеты [16] также показали, что высококачественное ОВФ вихревого пучка можно получить в более сложном ВРМБ-зеркале с растром микролинз [17—19]. В настоящей статье проводится обзор вышеуказанных работ, выполненных в РФЯЦ-ВНИИЭФ.

1. Уравнения ВРМБ лазерного излучения

Рассеянное в результате ВРМБ стоксово излучение имеет частотный сдвиг по отношению к падающему лазерному излучению, равный частоте гиперзвука $\Omega = \omega_L - \omega_S = (2n_0\omega_L v_s/c)\sin(\theta/2)$, где n_0 — средний показатель преломления, ω_L и ω_S — частоты лазерного и стоксова излучения, v_s — адиабатическая скорость гиперзвука, θ — угол рассеяния [20]. Выбор активной среды, вытянутой в направлении распространения лазерной накачки, обеспечивает отражение стоксовой компоненты почти строго назад. Уравнения ВРМБ для медленно меняющихся амплитуд лазерного A_L , стоксова A_S и гиперзвукового p полей выводятся из волновых уравнений для света и гиперзвука [20] и имеют вид [21, 22]

$$\frac{n_0}{c} \frac{\partial A_L}{\partial t} + \frac{\partial A_L}{\partial z} - \frac{1}{2ik_L} \Delta_{\perp} A_L = -\frac{i}{2} p A_S, \quad (1)$$

$$\frac{n_0}{c} \frac{\partial A_S}{\partial t} - \frac{\partial A_S}{\partial z} - \frac{1}{2ik_S} \Delta_{\perp} A_S = -\frac{i}{2} p^* A_L, \quad (2)$$

$$\frac{\partial p}{\partial t} + \frac{p}{\tau} = -i \frac{g}{\tau} A_L A_S^* + S, \quad (3)$$

где Δ_{\perp} — поперечный лапласиан, g — коэффициент усиления ВРМБ, τ — время релаксации гиперзвука. Предполагается, что время изменения A_L заметно превышает период гиперзвуковых колебаний $T = 2\pi/\Omega$ (противоположный случай рассмотрен в [23, 24]).

Флуктуационная сила Ланжевена S в уравнении (3) дельта-коррелирована во времени:

$$\langle S(\mathbf{r}_1, z_1, t_1)S^*(\mathbf{r}_2, z_2, t_2) \rangle = B_p(\mathbf{r}_1 - \mathbf{r}_2, z_1 - z_2)\delta(t_1 - t_2), \quad (4)$$

где B_p — пространственная корреляционная функция флуктуаций источника S . Ее вид предполагает статистическую однородность ВРМБ активной среды.

Более подробно модель развития стокового излучения при ВРМБ с учетом переходных процессов, нестационарности, гиперзвукового шума и насыщения ВРМБ описана в [21, 22]. Получено хорошее согласие результатов расчетов с экспериментальными данными в случае ОВФ регулярных лазерных пучков в ВРМБ-зеркалах разного типа [19, 21, 22].

Если время изменения A_L заметно превышает и $T = 2\pi/\Omega$, и τ , то из уравнения (3), пренебрегая шумом, получаем асимптотическое решение $p \cong -igA_L A_S^*$. Если интенсивность усиленного стокового излучения относительно невысока по сравнению с лазерной интенсивностью, коэффициент отражения стокового излучения не превышает нескольких процентов [6], то применимо приближение линейного ВРМБ (заданной накачки). В его рамках правая часть уравнения (1) зануляется. Тогда из (1) и (2) получим два уравнения для исследуемых полей в стационарном приближении:

$$\frac{\partial A_L}{\partial z} - \frac{1}{2ik} \Delta_{\perp} A_L = 0, \quad (5)$$

$$-\frac{\partial A_S}{\partial z} - \frac{1}{2ik} \Delta_{\perp} A_S = \frac{g}{2} J_L A_S. \quad (6)$$

В схемах ВРМБ-зеркал используется острая фокусировка лазерного излучения в кювету ВРМБ с помощью линзы с фокусным расстоянием F . В кювете ВРМБ область поперечной локализации лазерного пучка изменяется существенно, что представляет трудность для моделирования. Поэтому осуществляется замена переменных типа [25]:

$$\rho = \frac{\mathbf{r}}{w} \sqrt{2N_L}, \quad \tan \eta = z/z_0,$$

$$A_L(r, \varphi, z) = A_0 \left(\frac{w_0}{w} \cos \eta \right) \exp \left(-i \frac{\rho^2}{2} \operatorname{tg} \eta \right) \cdot \tilde{A}_L(\rho, \varphi, \eta), \quad (7)$$

$$A_S(r, \varphi, z) = A_0 \left(\frac{w_0}{w} \cos \eta \right) \exp \left(i \frac{\rho^2}{2} \operatorname{tg} \eta \right) \cdot \tilde{A}_S(\rho, \varphi, \eta).$$

Для перехода к новым переменным требуется задание радиуса лазерного пучка на входе в среду a и в перетяжке w_0 . При заданном w_0 определяем дифракционную длину $z_d = F(a^2/w_0^2 - 1)^{1/2}$, длину перетяжки $z_0 = z_d/(1 + z_d^2/F^2)$. В переменных (7) уравнения (5), (6) имеют вид [18]

$$\left[\frac{\partial}{\partial \eta} - \frac{i}{2} \Delta_{\perp} + \frac{i}{2} \rho^2 \right] \varepsilon_L = 0, \quad (8)$$

$$\left[\frac{\partial}{\partial \eta} + \frac{i}{2} \Delta_{\perp} - \frac{i}{2} \rho^2 \right] \varepsilon_S = \frac{1}{2} G_0 |\varepsilon_L|^2 \varepsilon_S, \quad (9)$$

где $G_0 = g|A_0|^2 z_0$.

2. Моды Гаусса — Лагерра

Лазерные пучки Гаусса — Лагерра LG_n^m являются собственными модами параболического квазиоптического уравнения (5), поэтому при распространении в свободном пространстве и при линзовых преобразованиях не меняют свою форму. Соответствующее решение в цилиндрических координатах имеет вид

$$A_L(r, \varphi; z) = A_0 \frac{w_0}{w} \left(\frac{r}{w} \right)^m \Phi_m(\varphi) L_n^m \left(2 \frac{r^2}{w^2} \right) \times \exp \left(-\frac{r^2}{w^2} + \frac{izr^2}{z_0 w^2} - i(2n+m+1) \arctg \frac{z}{z_0} \right), \quad (10)$$

где в радиальную часть распределения входит обобщенный полином Лагерра L_n^m . Угловая часть распределения (10) представляется в виде линейной комбинации гармонических функций

$$\Phi_m(\varphi) = c_1 \Omega_m(\varphi) + c_2 \Omega_{-m}(\varphi), \quad (11)$$

где $\Omega_m(\varphi) = \exp(im\varphi)$.

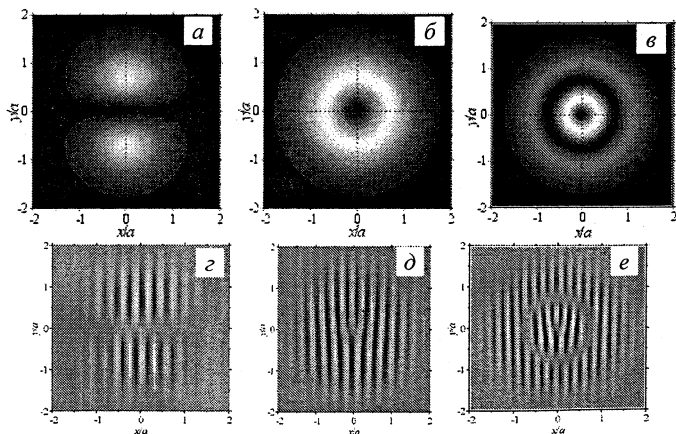


Рис. 2. Распределения интенсивности в лазерном пучке (а—в) и картина его интерференции с наклонно падающей плоской волной (г—е) при $n = 0$, $m = 1$, $\Phi_m = \sin \varphi$ (а, г), $n = 0$, $m = 1$, $\Phi_m = \exp(i\varphi)$ (б, д) и $n = m = 1$, $\Phi_m = \exp(i\varphi)$ (в, е).

Рассмотрим два случая распределения угловой функции $\Phi_m(\varphi)$ из (11). В первом случае $c_1 = -c_2 = 1/2i$, т. е. $\Phi_m(\varphi) = \sin(m\varphi)$. При этом интенсивность излучения не обладает аксиальной симметрией. На рис. 2, а показано поперечное распределение лазерной интенсивности при $\Phi_m = \sin\varphi$ для $n = 0$, $m = 1$. Во втором случае, когда $c_1 = 1$ и $c_2 = 0$, имеем $\Phi_m(\varphi) = \exp(im\varphi)$, и распределение интенсивности в пучке является аксиально-симметричным. На рис. 2, б, в показаны распределения интенсивности при $\Phi_m = \exp(i\varphi)$ для $n = 0$, $m = 1$ (торообразная мода LG_0^1) и $n = m = 1$ (мода с двумя кольцами LG_1^1). На рис. 2, г—е даны соответствующие картины интерференции лазерных пучков с наклонно падающей плоской волной. При $\Phi_m = \exp(i\varphi)$ в картине наблюдается ветвление полос в центре пучка с образованием «вилки», характерной для винтовой дислокации: на оси интенсивность равна нулю, а фаза не определена. Это признак вихревого пучка [1—5]. Фазовая поверхность пучка при $\Phi_m = \exp(im\varphi)$ показана на рис. 1. При круговом обходе особой точки по фазовой поверхности невозможно попасть в исходную точку. Число m называется топологическим зарядом оптического вихря. При $\Phi_m = \sin\varphi$ особой точки нет, имеет место краевая дислокация со сдвигом полос, что демонстрирует антисимметрию фазы поля относительно оси x .

3. Анализ ВРМБ оптических вихрей по теории возмущений

В переменных (7) нормированный пучок Гаусса — Лагерра (10) имеет вид $\tilde{A}_L(\rho, \varphi; z) = \psi_L^{n,m} \exp\{-i(2n + m + 1)\eta\}$, где

$$\psi_L^{n,m}(\rho, \varphi) = \left\{ \frac{2n!}{B_m(n+m)!} \right\}^{1/2} \rho^m L_n^m(\rho^2) \Phi_m(\varphi) \exp\left(-\frac{\rho^2}{2}\right), \quad (12)$$

где $B_m = \int_{-\pi}^{\pi} |\Phi_m|^2 d\varphi = 2\pi[|c_1|^2 + |c_2|^2]$ при $m \neq 0$.

Решение (9) ищем в виде $\tilde{A}_S(\rho, \varphi; \eta) = \tilde{\psi}_S(\rho, \varphi) \exp(i\tilde{\mu}\eta)$, тогда получаем уравнение

$$\left[-\frac{1}{2}\Delta_{\perp} + \frac{1}{2}\rho^2 + V \right] \tilde{\psi}_S = \tilde{\mu}\tilde{\psi}_S, \quad (13)$$

которое имеет вид стационарного уравнения Шредингера для частицы в осцилляторной потенциальной яме на плоскости при наличии возмущающего потенциала

$$V \equiv V_n^m = -\frac{i}{2}G_0 |\varepsilon_L|^2 = -\frac{i}{2}G_0 |\psi_L^{n,m}|^2.$$

Собственные функции и собственные значения уравнения (13) по теории возмущений имеют вид [26]

$$\tilde{\Psi}_S^{n',m'}(\rho, \varphi) = \Psi_S^{n',m'}(\rho, \varphi) + \sum_{\substack{\mu_n'' \neq \mu_n'' \\ \mu_n'' = \mu_n''}} \frac{V_{n',n,n''}^{m',m,m''}}{\mu_n'' - \mu_n''} \Psi_S^{n'',m''}(\rho, \varphi),$$

$$\tilde{\mu}_n^{m'} = \mu_n^{m'} + V_{n',n,n''}^{m',m,m''},$$

где $V_{n',n,n''}^{m',m,m''} = \int_{-\pi}^{\pi} d\varphi \int_0^{\infty} \rho d\rho \Psi_S^{n',m'} * V_n^m \Psi_S^{n'',m''}$, а

$$\Psi_S^{n',m'}(\rho, \varphi) = \left\{ \frac{n'!}{\pi(n' + |m'|)!} \right\}^{1/2} \rho^{|m'|} L_{n'}^{|m'|}(\rho^2) \Omega_{m'}(\varphi) \exp\left(-\frac{\rho^2}{2}\right), \quad (14)$$

$$\mu_n^{m'} = 2n' + |m'| + 1 \quad (15)$$

суть собственные функции и собственные значения в отсутствие возмущения. Сумма $2n' + |m'|$ — аналог главного квантового числа, а $|m'|$ — аналог орбитального квантового числа. Заметим, что согласно (15) уровень с «энергией» $\mu_n^{m'}$ имеет $\mu_n^{m'}$ -кратное вырождение. В частности, оно выражается в том, что угловая часть волновой функции (14) $\Omega_{\pm|m'|} = \exp(\pm i|m'|\varphi)$ дает два типа закручивания фазовой спирали.

Физический смысл вырождения в данной задаче заключается в том, что вырожденные стоксовы (n', m') -моды из базиса Гаусса — Лагерра (14) при ВРМБ лазерной (n, m) -моды не будут в общем случае усиливаться экспоненциально с коэффициентом $M_{n',n,n''}^{m',m,m''} = -2iV_{n',n,n''}^{m',m,m''}$. Экспоненциальное усиление будет наблюдаться для линейных комбинаций этих вырожденных мод. Коэффициенты разложения в этих комбинациях и коэффициенты усиления \tilde{M} новых состояний находятся из решения секулярного уравнения $\det(M_{n',n,n''}^{m',m,m''} - \tilde{M}\delta_{n,n'}\delta_{m,m'}) = 0$, где

$$M_{n',n,n''}^{m',m,m''} = -2iV_{n',n,n''}^{m',m,m''} = \frac{G_0 n! \sqrt{n'! n''!} \mathfrak{G}_{n',m,m''} / \pi}{(n+m)! \sqrt{(n'+|m'|)! (n''+|m''|)!}} \times$$

$$\times \int_0^{\infty} x^{m+(|m'|+|m''|)/2} \left\{ L_n^m(x) \right\}^2 L_{n'}^{|m'|}(x) L_{n''}^{|m''|}(x) e^{-2x} dx, \quad (16)$$

$$\mathfrak{G}_{m',m,m''} = \int_{-\pi}^{\pi} \Omega_{m'}^* |\Phi_m|^2 \Omega_{m''} d\varphi / B_m.$$

При $\Phi_m = \exp(im\varphi)$ получаем $\mathfrak{G}_{m',m,m''} = \delta_{m',m''}$, т. е. новые состояния совпадают со старыми, так как для вырожденных состояний либо $m' = -m''$, либо $|m' - m''|$ кратно двум. Таким образом, в первом порядке теории возмущений в случае вихревых лазерных пучков вырождение не снимается. Коэффициенты усиления стоксовых мод описываются вели-

чиной $M_{n',n,n'}^{m',m,m'}$. Для невихревого пучка при $\Phi_m = \sin(m\varphi)$ имеем $\mathcal{G}_{m',m,m''} = (\delta_{m',m''} - \delta_{|m'-m'|, 2m})/2$, т. е. здесь вырождение снимается.

В случае $n = 0, m = 1$ зависимость от $\mu_n^{m'}$ коэффициента усиления $M = \tilde{M}$ стоксовых мод при $\Phi_m = \sin \varphi$ показана на рис. 3, а, а зависимость коэффициента усиления $M = M_{n',0,n'}^{m',1,m'}$ при $\Phi_m = \exp(i\varphi)$ дана на рис. 3, б. Для невихревой лазерной моды коэффициент усиления сопряженной ей стоксовой моды в $\sim 1,5$ раза больше, чем для ближайшей основной моды $n' = m' = 0$. Ситуация на рис. 3, а, когда сопряженная мода испытывает наибольшее усиление, является основой для явления ОВФ. И именно для ОВФ сфокусированных пучков характерно такое превышение коэффициента усиления сопряженной моды [6].

Как видно из рис. 3, б, для вихревой лазерной моды коэффициенты усиления сопряженной вихревой моды ($n' = 0, m' = 1$) и вихревой моды с противоположным закручиванием «фазовой спирали» ($n' = 0, m' = -1$) одинаковы, что прямо следует из формулы (16). Это происходит оттого, что усиление ВРМБ определяется лазерной интенсивностью и стоксовы моды не «чувствуют», куда закручена фазовая спираль в пучке накачки. Более того, такой же коэффициент усиления имеет и основная стоксова мода ($n' = m' = 0$). Остальные стоксовы моды подавлены. Ситуация, когда сопряженная мода имеет равноправные конкурирующие моды, не является типичной для ОВФ. ОВФ как явление в данном случае не имеет места из-за отсутствия селекции сопряженной компоненты [8—10].

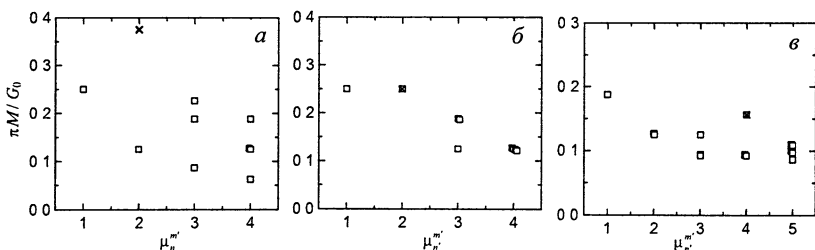


Рис. 3. Зависимость коэффициента усиления стоксовых мод от $\mu_n^{m'}$ при лазерных модах ($n = 0, m = 1$) с $\Phi_m = \sin \varphi$ (а), $\Phi_m = \exp(i\varphi)$ (б) и ($n = m = 1$) с $\Phi_m = \exp(i\varphi)$ (в). Символ \times отвечает обращенной моде.

Для случая вихревой лазерной моды с $n = m = 1$ и $\Phi_m = \exp(i\varphi)$ зависимость от $\mu_n^{m'}$ коэффициента усиления стоксовых мод показана на рис. 3, в. Коэффициенты усиления двух стоксовых мод — сопряженной вихревой моды ($n' = 1, m' = 1$) и вихревой моды обратной спиральности ($n' = 1, m' = -1$) — одинаковы, т. е. опять-таки явление ОВФ не имеет места. Но дополнительным неожиданным обстоятельством является то,

что имеется мода, коэффициент усиления которой оказывается даже большим (на 20%), чем у сопряженной моды! Такой модой является основная мода ($n' = m' = 0$). Итак, в ситуации, показанной на рис. 3, в, реализуется селекция основной гауссовой моды. Такой уникальный случай, когда селектируемой является ортогональная необращенная мода, позволяет говорить о новом явлении, которое может быть названо преобразованием волнового фронта (ПВФ) при ВРМБ [8—10]. Его причина, наряду с вырождением мод, кроется в их факторизованной форме по ρ и ϕ , что распространяется и на интеграл перекрытия (16). Поэтому в $M_{n',n,n'}^{m',m,m'}$ мы имеем фактически интеграл перекрытия радиальных частей функций $|\psi_L^{n,m}|^2$ и $\rho |\psi_S^{n',m'}|^2$, максимальный не обязательно при $n = n', m = m'$.

4. Результаты экспериментов и расчетов

Оптический вихрь получается с использованием спиральных фазовых пластин, изготовленных с помощью травления поверхности кварцевой подложки по киноформной технологии [14].

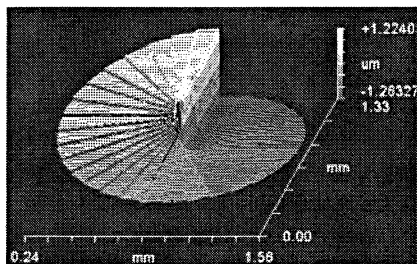


Рис. 4. Изображение приосевой области фазовой пластины.

Количество уровней пластин 16 и 32. Как показали расчеты, такое число уровней достаточно для аппроксимации гладкой поверхности. На рис. 4 показано изображение 32-уровневой пластины для $\lambda = 1,064$ мкм, полученное с помощью прецизионного профилометра.

Прошедшая через пластинку основная гауссова мода дает в дальней зоне оптический вихрь — моду Гаусса — Лагерра LG_0^1 — с коэффициентом преобразования более 90% [11—13]. Схема экспериментальной установки приведена на рис. 5 [15, 16]. На выходе неодимового лазера ($\lambda = 1,064$ мкм) формировался лазерный гауссов пучок с практически дифракционной расходимостью, диаметром 0,4 см, длительностью импульса 60 нс и энергией 170 мДж. Далее излучение с помощью полупрозрачного зеркала 3 делилось на два пучка. В первый пучок ставилась спиральная фазовая пластина 5. Фокусирующая линза 7 с фокусным расстоянием $F = 50$ см, располагавшаяся на расстоянии около 10 м от пластины 5, т. е. фактически в дальней зоне, фокусировала моду LG_0^1 в кювету ВРМБ 11. Экспериментальное распределение плотности энергии лазерного пучка на линзе 7 и картина его интерференции с наклонно падающей плоской волной показаны на рис. 6. Они полностью отвечают моде Гаусса — Лагерра LG_0^1 (ср. с рис. 1, б, д).

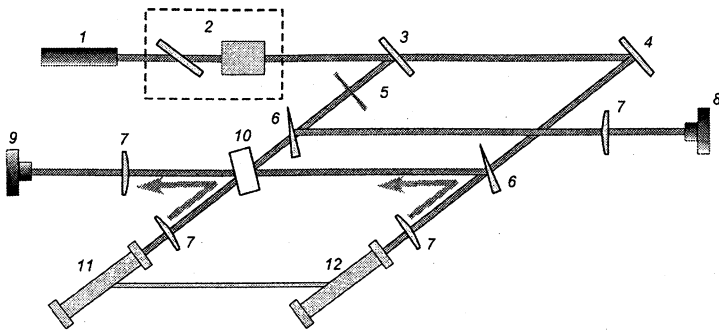


Рис. 5. Схема эксперимента: 1 — Nd:YAG лазер; 2 — оптический вентиль; 3 — зеркало $R = 50\%$; 4 — поворотное зеркало; 5 — фазовая пластинка; 6 — клин; 7 — линза; 8, 9 — CCD-камеры; 10 — делительная пластина; 11, 12 — ВРМБ-кюветы.

Второй пучок, не подвергавшийся фазовым искажениям, направлялся в кювету ВРМБ 12. Обе ВРМБ-кюветы длиной $L = 100$ см наполнены смесью газов Хе и SF_6 ($\tau = 22$ нс). CCD-камера 8 измеряла распределение энергии лазерного излучения, прошедшего через фазовую пластину и падающего на линзу 7. CCD-камера 9 измеряла распределение энергии в стоксовом пучке при ВРМБ вихревого пучка, если закрыта ВРМБ-кювета 12. В случае, когда обе кюветы были открыты, на CCD-камеру 9 проецировалась интерференционная картина двух стоксовых пучков. Для получения интерференционной картины необходимо уравновесить давление в обеих кюветках, поэтому они были соединены между собой гибким трубопроводом (общее давление смеси 30 атм).

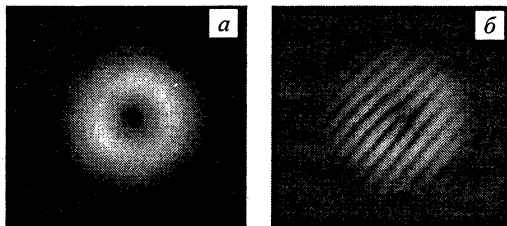


Рис. 6. Экспериментальное распределение лазерной энергии на фокусирующей линзе (а) и картина его интерференции с наклонной плоской волной (б).

Как расчеты уравнений (1)—(3), так и эксперименты проводились в околороговых условиях и при глубоком насыщении ВРМБ. Качество ОВФ в расчетах характеризуется коэффициентом обращения на входе кюветы [6] $h = \left| \int A_L A_S dr \right|^2 / (P_S P_L)$, где $P_L = \int J_L dr$ и $P_S = \int J_S dr$ — мощность лазерного и стоксова излучения. При идеальном ОВФ имеем $h = 1$.

Случай ВРМБ невихревого лазерного пучка при $n = 0$, $m = 1$, $\Phi_m = \sin \varphi$, структура которого показана на рис. 2, а, з, рассмотрен в расчетах [8—10]. В этом случае реализуется классическое ОВФ. В околопороговых условиях коэффициент обращения $h \sim 75\%$. Это характерно для условий линейного ВРМБ [21, 22]. При этом наблюдается качественное сходство распределений энергии и интерференционных картин лазерного и стокова излучения, что согласуется с вышеприведенным теоретическим анализом для условий линейного ВРМБ (см. рис. 3, а). Насыщение ВРМБ и рост коэффициента отражения ВРМБ-зеркала способствует улучшению качества ОВФ (сравните рис. 7, а, б с рис. 2, а, з). При коэффициенте отражения 50% интегральное за время импульса значение h составляет уже 90% (см. рис. 7, в). При этом картина выходного стокова излучения практически не зависит от реализации гиперзвукового шума.

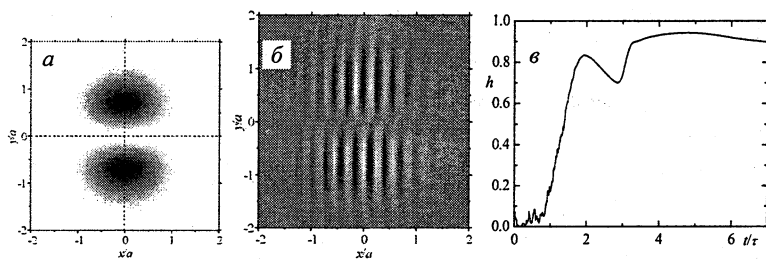


Рис. 7. Распределение энергии стокова пучка (а), картина его интерференции с наклонно падающей плоской волной (б) и динамика коэффициента обращения h (в) на входе кюветы при $n = 0$, $m = 1$, $\Phi_m = \sin \varphi$ и насыщении ВРМБ.

Теперь возьмем вихревой лазерный пучок при $n = 0$, $m = 1$, $\Phi_m = \exp(i\varphi)$, т. е. моду Гаусса — Лагерра LG_0^1 (см. рис. 2, б, д). На рис. 8 показана расчетная динамика мощности при коэффициенте отражения $R \sim 7—10\%$, т. е. в режиме ВРМБ, близком к линейному, при различных реализациях гиперзвукового шума. Из рис. 9 видно, что коэффициент обращения h составляет малую величину, равную в среднем 5—10% (ср. с рис. 7, в). Зависящие от реализации шума пространственные распределения энергии и интерференционные картины стокова излучения на входе кюветы показаны на рис. 10. Как предсказывает теория (см. рис. 3, б) и как видно из рис. 10, а, б, стоков пучок представляет собой случайную суперпозицию вихревых мод $n = 0$, $m = \pm 1$ с менее заметной добавкой моды $n = m = 0$. Это подтверждают интерференционные картины на рис. 10, в, г, где отсутствует характерная «вилка», зато имеется фрагмент краевой дислокации, характерный для комбинации мод $n = 0$, $m = \pm 1$ (ср. с рис. 2, з).

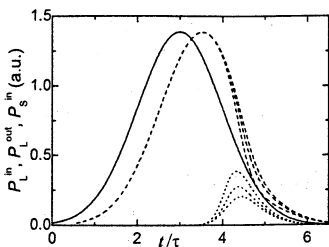


Рис. 8. Динамика мощности лазерного излучения на входе (—) и выходе (---) кюветы и стоксова излучения на входе (····) при $n = 0, m = 1, \Phi_m = \exp(i\varphi)$ вблизи порога ВРМБ.

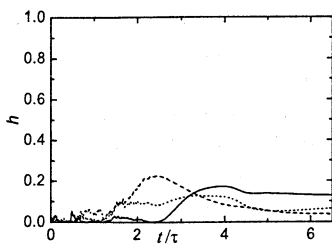


Рис. 9. Динамика коэффициента обращения h на входе кюветы при $n = 0, m = 1, \Phi_m = \exp(i\varphi)$ вблизи порога ВРМБ.

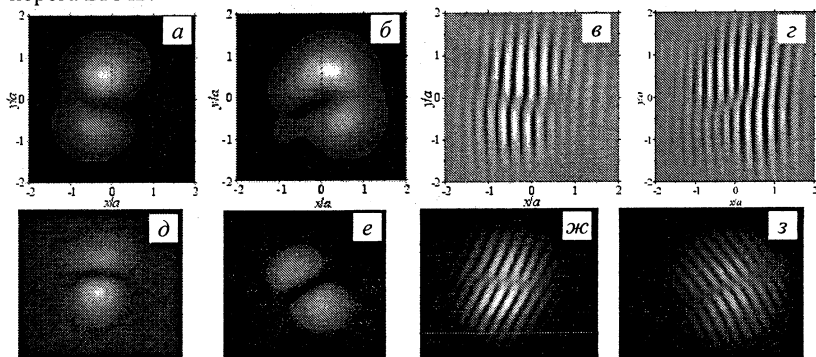


Рис. 10. Расчетные (верхний ряд) и экспериментальные (нижний ряд) распределения плотности энергии в стоксовом пучке и картины его интерференции с наклонно падающей плоской волной при ВРМБ моды LG_0^1 вблизи порога.

На рис. 10, δ — $з$ показаны соответствующие экспериментальные данные, которые абсолютно согласуются с результатами расчетов и являются типичными среди множества проведенных экспериментов.

Глубина насыщения ВРМБ не оказывает принципиального влияния на результаты. При увеличении лазерной энергии коэффициент отражения R вырастает до 60% (рис. 11), тогда как ОВФ качественно не меняется (рис. 12). В распределениях стоксовой интенсивности (рис. 13, $a, б$), отвечающих разным реализациям гиперзвукового шума, в большей степени подчеркивается кольцевая форма пучка, т. е. стоксово излучение заметнее «втягивается» в «бублик» лазерной накачки. Это может создать иллюзию наличия ОВФ. Но на фазовую структуру (рис. 13, $в, г$) насыщение ВРМБ влияет слабо. Интерференционная картина опять-таки может содержать фрагмент краевой дислокации. Особенностью является то, что в некоторых случаях на периферии пучка образуется вихрь с тем же зарядом, что и у лазерного пучка (см. рис. 13, $з$).

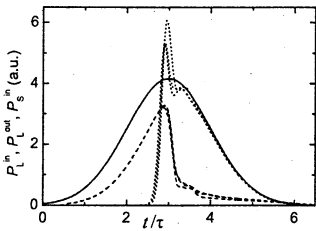


Рис. 11. Динамика мощности лазерного излучения на входе (—) и выходе (---) кюветы и стоксова излучения на входе (....) при ВРМБ моды Гаусса — Лагерра LG_0^1 в условиях насыщения ВРМБ.

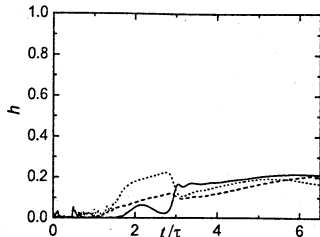


Рис. 12. Динамика коэффициента обращения h на входе кюветы при ВРМБ моды Гаусса — Лагерра LG_0^1 в условиях насыщения ВРМБ.

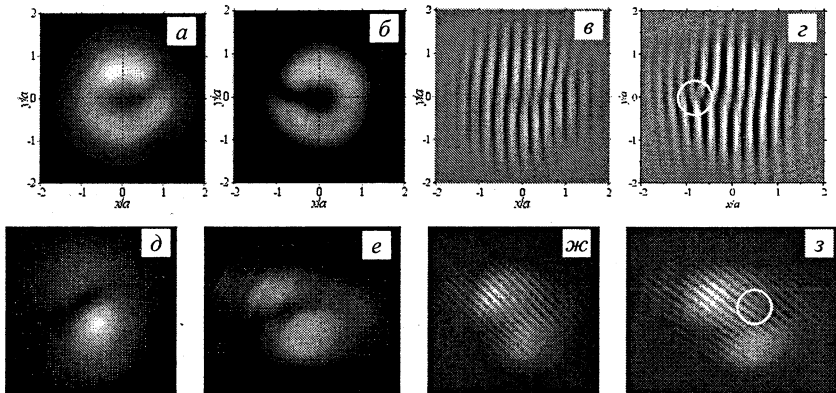


Рис. 13. Расчетные (верхний ряд) и экспериментальные (нижний ряд) распределения плотности энергии в стоксовом пучке и картины его интерференции с наклонно падающей плоской волной при ВРМБ моды LG_0^1 в условиях насыщения ВРМБ.

На рис. 13, δ — $з$ даны экспериментальные данные, полученные в условиях влияния насыщения ВРМБ, которые абсолютно согласуются с результатами расчетов. Демонстрируются распределения, которые являются типичными среди множества проведенных экспериментов, которые не показали наличия осевых вихрей в стоксовом пучке. В ряде экспериментов в стоксовом пучке регистрировался внеосевой вихрь (см. рис. 13, $з$) в полном соответствии с расчетами.

ВРМБ вихревого лазерного пучка с двумя кольцами интенсивности при $n = m = 1$, $\Phi_m = \exp(i\varphi)$, структура которого показана на рис. 2, $в$, $е$, рассматривалась в расчетах [8—10]. При коэффициенте отражения $R \sim 2$ —4% коэффициент обращения мал, $h \sim 4$ —10%. В соответствии с теорией линейного ВРМБ данного вихревого пучка (см. рис. 3, $в$) в стоксовом поле наибольшее усиление испытывает основная гауссова мода $n = m = 0$, т. е. наблюдается эффект ПВФ. В самом деле, при учете пере-

ходных процессов и нестационарности при слабом насыщении ВРМБ расчетный стоксов пучок на входе кюветы ВРМБ представляет собой искаженную основную гауссову моду (рис. 14, а). Характер искажений зависит от реализации гиперзвукового шума. В нашем случае превышение коэффициента усиления основной гауссовой моды над ближайшими модами ($n' = 1, m' = \pm 1$) относительно мало и составляет 20%, что и является причиной ее искаженного вида. Но важно, что интерференционная картина стоксова излучения не содержит «вилки» (рис. 14, б), что указывает на регулярность ее фазового фронта. При влиянии насыщения ВРМБ коэффициент отражения R вырос до 70%, но коэффициент обращения остался малым, $h \sim 1-10\%$. Интенсивность стоксова излучения при глубоком насыщении ВРМБ «втягивается» в кольца накачки с образованием осевого провала (рис. 14, в). Стоксов пучок, таким образом, более далек от основной гауссовой моды, но, тем не менее, все-таки не содержит вихря на оптической оси (см. рис. 11, з).

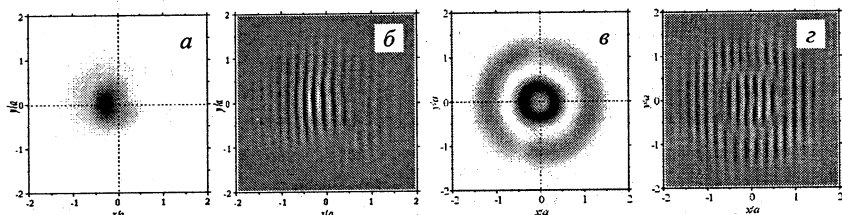


Рис. 14. Распределение плотности энергии в стоксовом пучке (а, в) и картина его интерференции с наклонно падающей плоской волной (б, з) на входе кюветы при $n = 1, m = 1, \Phi_m = \exp(i\varphi)$ без (а, б) и в условиях (в, з) насыщения ВРМБ.

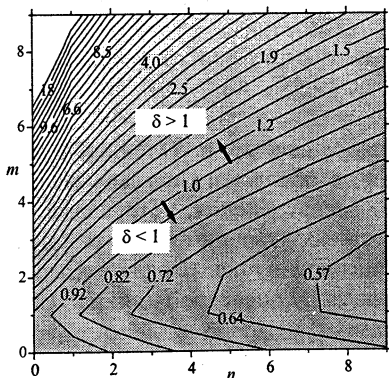


Рис. 15. Отношение коэффициентов усиления обращенной моды LG_n^m и основной моды LG_0^0 .

Как показывают расчеты (16), с ростом n степень селекции основной гауссовой моды при ПВФ возрастает [8, 9]. На рис. 15 показана зависимость отношения коэффициентов усиления обращенной моды LG_n^m и основной моды LG_0^0 , т. е. $\delta = M_{n,n,n}^{m,m,m} / M_{0,0,0}^{0,0,0}$, от n и m . Как видно из

рис. 15, ПВФ (область $\delta < 1$) может наблюдаться и для невихревых пучков с $m = 0$ без осевого провала в распределении интенсивности. Анализ показывает, что существует возможность ПВФ вихревой лазерной моды в более сложную, не основную, гауссову моду при рассмотрении режима ВРМБ-усилителя, когда стоксов пучок формируется не из шума, а от входного стокова сигнала в виде одной конкретной моды.

Полученные результаты [8—10] показывают генерацию вихревых гиперзвуковых волн в ВРМБ-среде и демонстрируют закон сохранения топологического заряда: суммарный заряд в системе «стоксова волна + волна гиперзвука» равен заряду входной лазерной моды.

5. ОВФ вихревого пучка

Более сложное ВРМБ-зеркало с растром малых дифракционных линз Френеля и угловым селектором [17] (рис. 16, а) позволяет провести эффективную угловую фильтрацию необращенной стоковой компоненты, уменьшить локальные световые нагрузки и избежать появления нежелательных попутных нелинейных эффектов. В [18, 19] расчетно и экспериментально определена оптимальная геометрия данного ВРМБ-зеркала, которая устойчиво дает $h \geq 90$ —95% при коэффициенте пропускания селектора 50—70% и любом уровне насыщения ВРМБ, т. е. любом коэффициенте отражения.

Разработанная концепция ВРМБ-зеркала применялась для получения ОВФ вихревого пучка [16]. Фокусное расстояние главной линзы $F = 50$ см, размер малой линзы растра $d = 0,1$ см, фокусное расстояние малой линзы $f = 16$ см. На рис. 16, б показано распределение интенсивности стокова излучения в ближней зоне после прохождения селектора. Оно получено в условиях линейного ВРМБ при оптимальной геометрии ВРМБ-зеркала. Как показали расчеты, использование растра в ВРМБ-зеркале предпочтительнее применения случайной фазовой пластины.

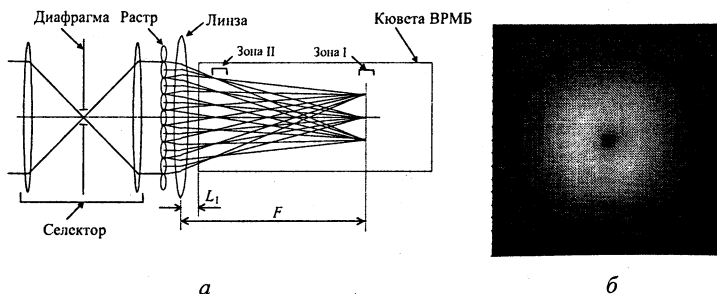


Рис. 16. Схема ОВФ-зеркала с растром дифракционных линз (а) и расчетное распределение интенсивности стокова излучения в ближней зоне (б).

Теория, расчеты и эксперимент показали, что ОВФ при ВРМБ сфокусированных вихревых лазерных пучков (мод Гаусса — Лагерра) отсутствует как явление, так как коэффициенты усиления обращенной стоксовой моды и аналогичной моды с обратной спиральностью одинаковы. Для широкого класса вихревых пучков (например, в случае моды LG_1^1) предсказывается явление, которое может быть названо преобразованием волнового фронта при ВРМБ. Его суть заключается в том, что в стоксовом пучке селективируется единственная мода, но она является основной гауссовой модой, ортогональной лазерной моде. Показано, что получить ОВФ вихревого пучка можно при разрушении его структуры в ВРМБ-среде, например в ВРМБ-зеркале с растром микролинз.

Литература

1. Nye J. F., Berry M. V. Dislocations in wave trains // Proc. Roy. Soc. A. 1974. V. 336. P. 165—190.
2. Баранова Н. Б. Зельдович Б. Я., Мамаев А. В., Пилипецкий Н. В., Шкунов В. В. Дислокации волнового фронта спекл-неоднородного поля (теория и эксперимент) // Письма в ЖЭТФ. 1981. Т. 33, вып. 4. С. 206—209.
3. Баженов В. Ю., Васнецов М. В., Соскин М. С. Лазерные пучки с дислокациями волнового фронта // Письма в ЖЭТФ. 1990. Т. 52, вып. 8. С. 1037—1039.
4. Fried D. L. Branch point problem in adaptive optics // JOSA. 1998. V. 15. № 10. P. 2759—2768.
5. Optical Vortices // Horizons in World Physics, V. 228 / Eds. M. Vasnnetsov, K. Staliunas. N. Y.: Nova Science, 1999. 218 p.
6. Зельдович Б. Я., Пилипецкий Н. Ф., Шкунов В. В. Обращение волнового фронта. М.: Наука, ГРФМЛ, 1985.
7. Marienko I. G., Soskin M. S., Vasnetsov M. V. Phase reversal of light beams carrying optical vortices // Asian J. Phys. 1998. V. 7. P. 495.
8. Starikov F. A., Kochemasov G. G. Novel phenomena at stimulated Brillouin scattering of vortex laser beams // Opt. Commun. 2001. V. 193. P. 207—215.
9. Стариков Ф. А., Кочемасов Г. Г. Новые явления при вынужденном рассеянии Мандельштама — Бриллюэна вихревых пучков // Тр. РФЯЦ-ВНИИЭФ. Саров, 2001. Вып. 1. С. 528—535.
10. Starikov F. A., Kochemasov G. G. Investigation of stimulated Brillouin scattering of vortex laser beams // Proc. SPIE. 2001. V. 4403. P. 217—228.
11. Starikov F. A., Atuchin V. V., Cherezova T. Yu. et al. Generation of optical vortex for an adaptive optical system for phase correction of laser beams with wave front dislocations // Proc. SPIE. 2004. V. 5572. P. 400—407.
12. Starikov F. A., Atuchin V. V., Dolgoplov Yu. V. et al. Development of an adaptive optical system for phase correction of laser beams with wave front dislocations: generation of an optical vortex // Proc. SPIE. 2004. V. 5777. P. 784—787.
13. Starikov F. A., Atuchin V. V., Kochemasov G. G. et al. Wave front registration of an optical vortex generated with the help of spiral phase plates // Proc. SPIE. 2005. V. 5894. P. O7 1—11.
14. Atuchin V. V., Permyakov S. L., Soldatenkov I. S., Starikov F. A. Kinoform generator of vortex laser beams // Proc. SPIE. 2006. V. 6054. P. 188—191.
15. Starikov F. A., Dolgoplov Yu. V., Kopalkin A. V., Kochemasov G. G., Kulikov S. M., Sukharev S. A. Stimulated Brillouin scattering of Laguerre-Gaussian laser modes: new phenomena

// Proc. 2nd Int. Conf. on Advanced Optoelectronics and Lasers (CAOL 2005, Yalta, 12—17 Sept. 2005). V. 2. P. 113—116

16. *Starikov F. A., Dolgoplov Yu. V., Kopalkin A. V., Kochemasov G. G., Kulikov S. M., Sukharev S. A.* About the correction of laser beams with phase front vortex // J. Phys. IV. 2006. V. 133. P. 683—685.

17. *Бобров С. Т., Грацианов К. В., Корнев А. Ф., Любимов В. В., Панков В. Г., Степанов А. И., Туркевич Ю. Г.* // Оптика и спектроскопия. 1987. Т. 62. С. 402.

18. *Kochemasov G. G., Starikov F. A.* Novel features of phase conjugation at SBS of beams passed through an ordered phase plate // Opt. Commun. 1999. V. 170. P. 161—174.

19. *Starikov F. A., Dolgoplov Yu. V., Kovaldov S. A. et al.* Superhigh fidelity of phase conjugation at SBS with using the kinoform optics of new generation // Proc. SPIE. 2001. V. 4353. P. 202—213.

20. *Старунов В. С., Фабелинский И. Л.* Вынужденное рассеяние Манделъштама — Бриллюэна и вынужденное энтропийное (температурное) рассеяние света // УФН. 1969. Т. 98. С. 441.

21. *Starikov F. A., Dolgoplov Yu. V., Kochemasov G. G. et al.* Investigation of phase conjugation at SBS of short laser pulses // Proc. Int. Conf. LASERS'99 (Quebec, Canada, December 11—16, 1999) / Eds. V. J. Corcoran & T. A. Corcoran, McLean, VA. STS Press, 2000. P. 482—489.

22. *Starikov F. A., Dolgoplov Yu. V., Kochemasov G. G. et al.* Phase conjugation at SBS of pulse laser radiation with using the kinoform optics // Proc. SPIE. 2000. V. 3930. P. 12—23.

23. *Kochemasov G. G., Starikov F. A., Ladagin V. K., Eroshenko V. A.* Phase conjugation at SBS of profiled laser pulse // Proc. Int. Conf. LASERS'99 (Quebec, Canada, December 11—16, 1999) / Eds. V. J. Corcoran & T. A. Corcoran, McLean, VA: STS Press, 2000. P. 474—481.

24. *Kochemasov G. G., Starikov F. A., Ladagin V. K., Eroshenko V. A.* Phase conjugation at SBS of CPA modulated profiled laser pulses // Proc. SPIE. 2000. V. 3929. P. 312—322.

25. *Таланов В. И.* О фокусировке света в кубичных средах // Письма в ЖЭТФ. 1970. Т. 11. С. 303.

26. *Ландау Л. Д., Лифшиц И. М.* Квантовая механика (нерелятивистская теория). М.: ГИФМЛ, 1963.

ГЕНЕРАЦИЯ БЫСТРЫХ ЭЛЕКТРОНОВ ПРИ ЭЛЕКТРОН-ИОННЫХ СТОЛКНОВЕНИЯХ В СИЛЬНЫХ ПОЛЯХ

Г. М. Фрайман, А. А. Балакин

1. Введение

В последние годы наблюдается значительный прогресс в генерации ультракоротких сверхсильных лазерных импульсов с интенсивностью в фокусе 10^{18} — 10^{21} Вт/см² и выше. При таких интенсивностях в лазерном диапазоне длин волн электроны плазмы осциллируют со скоростями (вплоть до релятивистских), существенно превышающими их тепловые скорости, что открывает совершенно новые, не исследованные ранее возможности взаимодействия излучения с плазмой. Одним из ключевых вопросов при этом является вопрос о том, насколько существенны столкновительные процессы при таких взаимодействиях.

Исследование столкновительного поглощения сверхсильных лазерных импульсов в плазме необходимо и для ряда приложений. К таковым можно отнести термоядерный синтез с использованием ультракоротких интенсивных лазерных импульсов для быстрого поджига термоядерной реакции, рентгеновские лазеры, преобразование энергии лазерного излучения в когерентное излучение на гармониках основной частоты при взаимодействии сверхсильных лазерных импульсов с плазмой.

Традиционно теоретическое исследование электрон-ионных соударений в сильных ЭМ-полях проводилось на основе трех моделей: «прямолинейных траекторий» (малоуглового) [1–3], борновского [4] и низкочастотного [5] приближений. Приближение [1–3] исторически называют малоугловым, но правильное его было бы называть приближением прямолинейных траекторий, поскольку в нем предполагается, что дрейфовые траектории¹ рассеивающихся частиц являются почти прямыми линиями. Во всех этих приближениях получались результаты, различающиеся только логарифмическими множителями. Основной причиной подобного совпадения, по-видимому, были общие

¹Дрейфовой траекторией частицы назовем траекторию, получающуюся вычитанием осцилляций частицы из ее траектории в лабораторной системе координат. Отметим, что в дрейфовых координатах невозмущенное кулоновским потенциалом движение происходит по прямой линии, а ион осциллирует. Наоборот, в лабораторной системе координат ион покоится, а движение электрона представляет собой медленный дрейф осцилляций.

для всех предположения о некоррелированности моментов столкновений и невозможности электрону искривить свою дрейфовую траекторию (притянуться к иону) в процессе многократных осциллиаций около иона. Особенно наглядно это продемонстрировано в недавних работах [2], где автор непосредственно из кинетического уравнения с интегралом столкновений в форме Ландау² получает, опять-таки с логарифмической точностью, те же результаты. По-видимому, совпадение результатов трех различных, на первый взгляд, приближений стало причиной угасания интереса к этой тематике более чем на тридцать лет.

В последнее время предпринимаются попытки создания численных кодов для моделирования процессов энергообмена в плазме с учетом электрон-ионных столкновений в сильных лазерных полях [6, 7]. В частности, уже результаты [6] не совпадают с традиционными. В то же время появились экспериментальные данные, не получившие удовлетворительного объяснения в рамках традиционных представлений.

Какие же эффекты, на наш взгляд, опускаются в традиционных моделях? Рассмотрим, как обычно описывалось столкновение. Напомним, что в отсутствие кулоновского поля иона движение частицы представляет собой осцилляции в поле ЭМ-волны с осцилляторной скоростью $v_{osc} = eE/m\omega_0$ (ω_0 и E — частота и амплитуда поля ЭМ-волны), медленно дрейфующие с постоянной дрейфовой скоростью v . Обычно полагается [1], что дрейф происходит по прямой линии за все время столкновения и осуществляет только «доставку» частиц в область взаимодействия. Затем происходит столкновение частицы с ионом, по-разному описываемое в разных моделях. При этом частицы перед столкновением считаются равномерно распределенными в пространстве. После столкновения частица считается покинувшей область рассеяния.

Предположение о равномерности частиц оправдано в случае быстрых частиц $v \gg v_{osc}$, когда частица пролетает область существенного рассеяния³ $b_v = e^2 Z/mv^2$ за время много меньше периода поля и больше не возвращается к этому иону. Формальное перенесение этого предположения на случай медленных частиц $v < v_{osc}$ с соответствующей оценкой области рассеяния как $b_{osc} = e^2 Z/mv^2$ неправомерно, поскольку частица пролетает над ионом много раз (много

²Напомним [3], что вывод интеграла столкновений Ландау базируется на предположениях о равномерности моментов столкновений и определяющем вкладе дальних (почти прямолинейных, без изменения прицельного параметра) рассеяний.

³Из задачи Резерфорда известно, что из всех частиц с начальной скоростью v на большой угол рассеиваются только частицы, пролетающие мимо иона на расстоянии меньше резерфордовского радиуса $b_v = e^2 Z/mv^2$.

периодов поля) и имеет возможность «подтянуться» к иону. Как следствие, и предположение о равномерности частиц может быть неприменимо.

Действительно, численное моделирование показывает другую динамику рассеяния частицы. В сильных полях из-за большого размаха осцилляций $r_{osc} = eE/m\omega_0^2 \gg b_{osc}$ частица многократно возвращается к одному и тому же иону и испытывает много далеких, малоугловых по полной скорости столкновений за время рассеяния. Энергия частицы при этом практически не меняется, но происходит искривление дрейфовой траектории частицы.

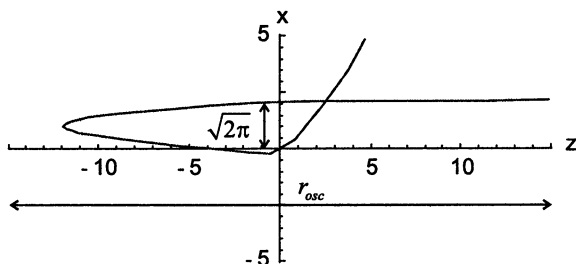


Рис. 1. Траектория частицы, притянувшейся к иону за одну осцилляцию

В качестве примера на рис. 1 изображена траектория частицы, «подтянувшейся» к иону за половину осцилляции ЭМ-поля (за один пролет над ионом). В результате, перед последним ударом, фактически и изменяющим энергию частиц, частицы оказываются значительно ближе к иону, чем при прямолинейном дрейфовом движении. Поэтому изменение энергии частиц при рассеянии существенно возрастает и появляется множество новых эффектов, анализу которых и посвящена эта работа.

2. Низкочастотное приближение

Будем далее рассматривать случай сильного поля накачки и малой тепловой скорости:

$$b_{osc} \ll r_{osc}, \quad v \ll v_{osc}. \quad (1)$$

В таких полях размах осцилляций r_{osc} очень большой. Другими словами, электрон (в лабораторной системе координат) большую часть времени движется вдали от кулоновского центра по осциллирующей

под действием поля накачки траектории. И только один или два раза за период частица может подлететь к иону и за очень малое время (практически мгновенно)

$$t_{coll} \leq b_v/v_{osc} \ll 2\pi/\omega_0$$

рассеяться на ионе. Соответственно дрейфовая траектория электрона представляет собой ломаную линию.

Используя условие малости времени столкновения по сравнению с периодом поля накачки, можно разделить два временных масштаба: а) мгновенное рассеяние на кулоновском центре, в процессе которого поле накачки не существенно; б) движение по осциллирующей траектории без учета кулоновского притяжения. В этом заключается идея низкочастотного приближения, предложенного Кролом и Ватсоном [5]. Однако в отличие от модели [5], учитывающей только однократное столкновение (удар) электрона с ионом, при рассмотрении движения частицы ограничение на количество «ударов» делаться не будет. Более того, все качественно новые, необычные эффекты получаются только при учете многих ударов. На траекторию частицы не накладывається каких-либо ограничений кроме условия того, что каждая частица пролетает мимо иона за время много меньшее периода поля, т. е. выполняется условие на прицельный параметр:

$$\rho/v \ll 2\pi/\omega_0 \iff \rho \ll r_{osc}. \quad (2)$$

Для описания системы в рамках низкочастотного приближения удобно использовать точечное отображение. Его идея заключается в том, что на непрерывной траектории $x(t)$ (здесь $x = \{\mathbf{r}, \mathbf{p}\}$ — совокупность координат и импульсов частицы) выделяется дискретный набор точек $x_n = x(nT)$, где n — целое число, $T = \pi/\omega_0$ — половина периода поля накачки⁴. Связь точки с номером $n + 1$ с точкой n описывается точечным отображением:

$$x_{n+1} = \hat{M}[x_n] \text{ или } \bar{x} = \hat{M}[x], \quad (3)$$

где символом \bar{x} обозначено значение x через итерацию.

В диапазоне параметров (1), (2) точечное отображение удастся найти в явном виде, но оно имеет достаточно сложный вид и здесь не приводится. Однако ряд свойств этого отображения легко обнаружить на простейшем варианте точечного отображения.

⁴Для линейно поляризованного поля накачки удобно использовать в качестве интервала времени T не период, а половину периода поля накачки, поскольку за период поля частица испытывает два «удара» об ион. Различие между этими ударами проявляется в отображении для продольных (вдоль z) координаты и скорости, которые в дальнейшем рассматриваться не будут.

3. Простейшее отображение

Пренебрегая смещением частицы поперек электрического поля из-за дрейфовой скорости $v_{\perp} \ll 2\pi/\rho v_{osc}$ (наиболее легко это реализуется при продольном рассеянии или при поперечном падении после первого столкновения — см. п. 4), отображение можно разбить на два: сдвиговое отображение для продольной координаты, ответственное за пролет частиц мимо иона и эффект группировки [10], и отображение для изменения поперечной координаты — прицельного параметра $\rho = \sqrt{x^2 + y^2}$:

$$\bar{\rho} = |\rho - 2\pi/\rho|. \quad (4)$$

При этом использована связь изменения скорости частицы с ее прицельным параметром:

$$\Delta v_{\perp} \approx v_{osc} b_{osc}/\rho. \quad (5)$$

Здесь и далее пространственные масштабы нормированы на характерные радиус $r_E = \sqrt{b_{osc} r_{osc}}$ — радиус, на котором осциллирующая сила поля накачки равна силе со стороны кулоновского потенциала иона, — и частоту $\omega_E = \sqrt[4]{\frac{eE^3}{m^2 Z}}$. При этом задача столкновений электронов с ионом описывается *единственным* безразмерным параметром

$$\Omega = \omega_0/\omega_E = (\omega_0^4 m^2 Z / eE^3)^{1/4}. \quad (6)$$

Графически отображение (4) представлено на рис. 2.

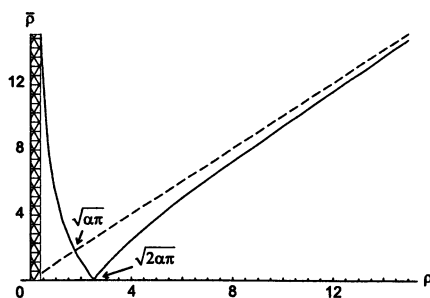


Рис. 2. Зависимость прицельного параметра $\bar{\rho}$ через итерацию отображения (4) от прицельного параметра ρ . Заштрихована область потерь частиц при рассеянии на большие углы.

Отображение (4) не описывает рассеяние частиц на большие углы и их уход из области рассеяния. Для учета таких частиц можно ввести

область потерь с малой шириной порядка $b_{osc} \ll r_E$ (заштрихованная область на рис. 2), при попадании в нее частицы считаются ушедшими из области рассеяния. Рассмотрим свойства отображения (4).

4. Многопотоковость

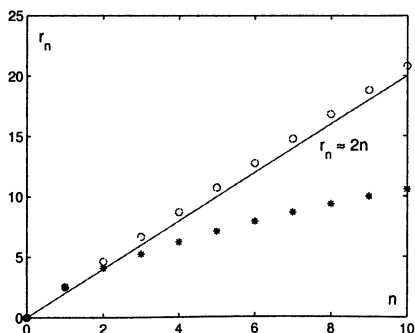
Во-первых, необходимо отметить появление многопотоковости, состоящей в том, что в одну и ту же точку пространства одновременно приходят частицы, стартовавшие с разных прицельных параметров. Найдем множество начальных прицельных параметров ρ_{ini} таких, что частицы из них попадают в область $\rho \leq b_{osc}$ и, соответственно, испытывают рассеяние на большие углы, сильное изменение энергии и т. д. Ранее [8] такие электроны были названы представительными.

Простейший случай — частицы, имеющие малые начальные параметры $\rho_{ini} \leq b_{osc}$, т. е. частицы, «столкнувшиеся» бы с ионом и без учета осцилляций под действием поля накачки. Именно эти частицы описываются прямолинейным (малоугловым) приближением. Кроме них есть частицы, пролетевшие мимо иона один раз, рассеявшиеся на малый угол и попавшие в ион при возврате (рис. 1). То есть в одну и ту же точку пространства приходят частицы, стартовавшие с разных прицельных параметров. Это и есть многопотоковость. Прицельный параметр ρ_1 , соответствующий этим частицам, можно найти из уравнения

$$\rho_1 - 2\pi/\rho_1 = 0 \implies \rho_1 = \sqrt{2\pi}. \quad (7)$$

Именно таково значение прицельных параметров на первом максимуме (рис. 3).

Рис. 3. Координаты максимумов, полученные в рамках простейшего отображения (4) (звездочки) и более сложного варианта отображения, учитывающего наличие конечной поперечной скорости у частицы (кружки), которое хорошо аппроксимируется прямой линией $r = 2n$.



Аналогичным образом можно найти значения начальных прицельных параметров ρ_2, ρ_3, \dots для 2-го, 3-го и т. д. максимумов, появляющихся после 2-го, 3-го и т. д. возврата частицы к иону. Полное число максимумов N (или, что то же самое, пролетов над ионом) можно оценить из формулы

$$N \approx 2r_{osc}/vT = 2v_{osc}/\pi v. \quad (8)$$

При этом из анализа полного отображения, учитывающего наличие конечной поперечной скорости у частицы, можно получить координату последнего максимума, являющегося границей области притяжения

$$\rho_N \equiv \rho_{attr} \approx 2/\Omega v, \quad (9)$$

хорошо совпадающую с результатами численного моделирования [8].

Координаты максимумов ρ_n (см. рис. 3) обладают еще одним интересным свойством: все частицы с прицельными параметрами $\rho \in [\rho_n, \rho_{n+1}]$ через один пролет над ионом оказываются в диапазоне $\rho \in [\rho_{n-1}, \rho_n]$. Это означает, что все частицы с прицельными параметрами $\rho \leq \rho_{attr}$ оказываются в области $\rho \leq \rho_1 = \sqrt{2\pi}$. Таким образом, вместо исходной задачи рассеяния бесконечно широкого пучка частиц с концентрацией n_0 достаточно рассмотреть рассеяние пучка частиц диаметром $2\rho_1$ и концентрацией

$$n_{eff} = n_0 \cdot \rho_{attr}^2 / \rho_1^2, \quad r_D \geq \rho_1. \quad (10)$$

5. Сингулярность в функции корреляции

Вторым важным свойством отображения (4) является возникновение сингулярности в концентрации перед «последним» ударом. Как уже отмечалось выше, все частицы с прицельными параметрами $\rho \leq \rho_{attr}$ стягиваются в процессе рассеяния в область $\rho \leq \rho_1$. Представляется интересным, каково 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 (4) из закона сохранения числа частиц:

$$\bar{n} \bar{\rho} d\bar{\rho} d\bar{z} = n \rho d\rho dz.$$

Выражая из отображения (4) старую координату ρ через новую $\bar{\rho}$ и используя неравенство $\rho > \rho_1$ (рассматриваются частицы, еще не попавшие в область $[0, \rho_1]$), находим⁵:

$$\bar{n} = n(\rho) \frac{\rho^3}{\bar{\rho}(\rho^2 + 2\pi)}, \quad \rho(\bar{\rho}) = \frac{1}{2} \left(\bar{\rho} + \sqrt{\bar{\rho}^2 + 8\pi} \right). \quad (11)$$

В частности, при больших прицельных параметрах $\bar{\rho} \gg \sqrt{8\pi} = 2\rho_1$, как и следует ожидать, концентрация практически не меняется. Однако при малых прицельных параметрах в концентрации появляется особенность:

$$\bar{n} \approx n \cdot \rho_1 / 2\bar{\rho}. \quad (12)$$

Учет формулы (10) позволяет найти распределение частиц по прицельным параметрам после N пролетов:

$$n_{eff} \approx n_0 \cdot \left(\frac{v_{osc}}{v} + \frac{a}{\rho} \right) \approx n_0 \frac{a}{\rho}, \quad a = \frac{\rho_{attr}^2}{2\rho_1}. \quad (13)$$

Появление первого слагаемого связано с тем, что далекие частицы, не успевшие подтянуться к кулоновскому центру, испытают, тем не менее, за время пролета мимо иона $N = v_{osc}/v$ малоугловых по полной скорости столкновений. Поэтому эффективная концентрация для одного столкновения с ионом для таких частиц будет иметь вид однородной концентрации n_0 , умноженной на число пролетов частицы мимо иона

$$n_{eff} \approx n_0 \cdot v_{osc}/v.$$

Знание распределения частиц перед последним «ударом» и использование аналитических формул рассеяния на кулоновском потенциале (см. [9], нерелятивистское движение) для последнего удара позволяют найти все характеристики рассеяния. Рассмотрим вначале распределение частиц по изменению энергии. Для этого отметим, что изменение энергии частицы с прицельным параметром ρ можно найти по формуле (см. далее сноску 7)

$$\Delta w = 4v_{osc}^2 / (1 + \rho^2/b_{osc}^2), \quad (14)$$

которая дает однозначную связь между изменением энергии частицы и прицельным параметром⁶ ρ . Это соотношение позволяет связать

⁵ В действительности норму отображения (4) 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.

⁶ В данном случае обсуждается только последний удар, поскольку полагается, что «мягкие» удары приводят только к изменению пространственного распределения частиц.

распределение частиц по прицельным параметрам и распределение частиц по изменению энергии в единицу времени:

$$p(\Delta w) = vn(\rho)\rho d\rho/d\Delta w = n(\rho) \cdot c2b_{osc}^2 w_{osc}/\Delta w^2, \quad w_{osc} = v_{osc}^2. \quad (15)$$

Используя распределение вида (13), получим при $\Delta w \ll w_{osc}$

$$p(\Delta w) \approx n_0 \left(v_{osc} \frac{4b_{osc}^2 w_{osc}}{\Delta w^2} + v \frac{2ab_{osc}\sqrt{w_{osc}}}{\Delta w^{3/2}} \right). \quad (16)$$

Видно, что зависимость распределения частиц по изменению энергии от первого слагаемого в (16) спадает как $\sim 1/\Delta w^2$. Напомним, что это слагаемое получено в рамках прямолинейного (малоуглового) приближения. Второе слагаемое спадает медленнее: $\sim 1/\Delta w^{3/2}$. Заметим здесь, что чем выше степень сингулярности в функции корреляции, тем более медленный закон спадания в распределении по изменению энергии.

Таким образом, анализ распределения частиц по изменению энергии позволяет определить, применимо прямолинейное (малоугловое) приближение или нет. Если прямолинейное (малоугловое) приближение применимо, то распределение по изменению энергии спадает как $\sim 1/\Delta w^2$ или быстрее. Если распределение спадает медленнее, то прямолинейное (малоугловое) приближение не применимо, и в функции корреляции существенна сингулярность. Все вышесказанное относится только к малым дрейфовым скоростям $v \ll v_{osc}$.

Аналогичным образом из формул (13) и (14) можно найти эффективное сечение

$$\sigma_{eff} = \frac{1}{v_{osc}^2} \int_0^{r_D} \frac{4n(\rho)v_{osc}^2}{1 + \rho^2/b_{osc}^2} \rho d\rho \approx 2b_{osc}^2 \frac{v_{osc}}{v} \ln + 2\pi ab_{osc} \approx 2\pi ab_{osc} \quad (17)$$

и эффективную частоту столкновений:

$$\nu_{eff} = vn_i \sigma_{eff} \approx 2n_i v_{osc} b_{osc}^2 \ln + 2\pi n_i v a b_{osc} \approx 2\pi n_i v a b_{osc}. \quad (18)$$

Учет выражения для a в случае продольного падения (13) приводит к тому же результату

$$\sigma_{eff} \approx 2\pi^2/v^2, \quad (19)$$

которые дают и качественные оценки, и численное интегрирование [8].

Границу между двумя моделями (прямолинейное приближение и учет притяжения) можно провести на базе сравнения различных членов в формуле для эффективного сечения (17). Если $\pi a < b_{osc} \ln$, то

можно пользоваться прямолинейным (малоугловым) приближением. В противном случае $\pi a > b_{osc} \ln$ прямолинейное (малоугловое) приближение не применимо. В частном случае продольного падения имеем

$$a = 1/\Omega^2 v^2 \implies v \leq v_{osc}, \quad (20)$$

т. е. при $\Omega \leq 0,3$ условие применимости прямолинейного (малоуглового) приближения нарушается уже при $v \leq v_{osc}$. В случае поперечного падения (см. п. 6) $a = 1/v^2$ и условие применимости прямолинейного (малоуглового) приближения имеет вид

$$v \leq v_{osc}/\sqrt{\ln}. \quad (21)$$

Подставляя $\ln \sim 10$, получаем, что вплоть до скоростей $v > v_{osc}/3$ прямолинейное (малоугловое) приближение применимо.

6. Особенности поперечного рассеяния

В общем случае поперечного падения при линейной поляризации поля накачки ($\mathbf{v} \perp \mathbf{v}_{osc}$) для определения эффективного сечения необходимо использовать полное отображение (3). При малых скоростях $v \leq \Omega$ (или $b_v \geq r_{osc} \gg b_{osc}$), как уже отмечалось выше, можно использовать упрощенное одномерное отображение (4). Соответственно, и результаты будут близки к результатам п. 3–4. В обратном предельном случае $b_{osc} < b_v < r_{osc}$ динамика частицы другая, однако, как и ранее, для определения эффективного сечения необходимо учитывать рассеяние на большие углы.

Чтобы оценить величину эффективного сечения в этом случае, рассмотрим процесс рассеяния частицы. На рис. 4 изображены частицы, налетающие на ион перпендикулярно полю, осциллирующему вдоль оси z . Начальные прицельные параметры частиц ρ примерно равны резерфордскому радиусу $b_v \gg b_{osc}$, так что они рассеиваются на покоящемся ионе⁷ с изменением угла по дрейфовой скорости на величину больше чем $\pi/2$.

После первого столкновения, мало изменяющего энергию частиц, частицы движутся по прямой, проходя за период поля расстояние порядка $2\pi r_a = 2\pi v/\Omega$. Через период частица в точке $z = r_{osc} - 2\pi r_a$ сталкивается с «возвращающимся» ионом, имеющим скорость $v_* =$

⁷ Речь идет о движении вблизи точек останова, где ион практически неподвижен большую часть периода.

$= \sqrt{4\pi v v_{osc}} \gg v$. Происходит второе столкновение, изменяющее энергию частицы до величины порядка $v_*^2 \approx 4\pi v v_{osc}$. Заметим, что фактически ко второму столкновению частицы попадают в режим продольного рассеяния, описанный в п. 2, поскольку поперечная скорость у таких частиц мала. Это также означает, что реально частица может испытывать не одно, а много столкновений, прежде чем существенно изменит свою энергию и покинет область рассеяния (см. рис. 4).

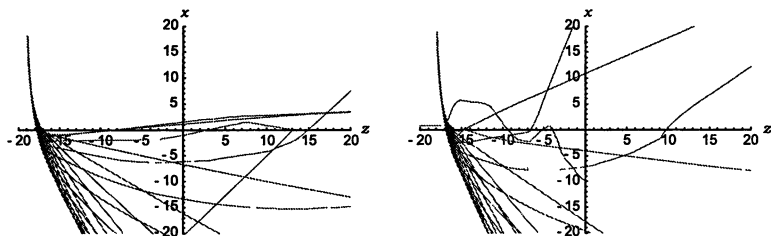


Рис. 4. Набор дрейфовых траекторий с различными начальными фазами для частиц, летящих перпендикулярно полю накачки, на границе (слева) и внутри (справа) области притяжения.

Однако нужно учесть, что такое изменение энергии получают не все частицы, а лишь те, которые попали в окрестность $b_* = 1/v_* v_{osc}$ иона при втором столкновении. В силу условия $b_* \ll b_v$ это возможно только для частиц из полосы шириной b_* . Последнее позволяет оценить эффективное сечение как произведение площади, с которой собираются частицы $b_v b_*$, на их относительное изменение энергии $4\pi v v_{osc}/v_{osc}^2$:

$$\sigma_{eff} = 4\pi b_v b_* \cdot v/v_{osc} = 4\pi b_v b_{osc}. \quad (22)$$

Точно такой же результат дает и численное интегрирование при поперечном падении.

7. Релятивистские столкновения

Столкновения в релятивистски сильных полях $v_{osc} \geq c$ сложнее и богаче, чем столкновения в полях дорелятивистской интенсивности. Причина состоит в том, что появляется еще один безразмерный параметр $\alpha = p_{osc}/p_0 = eA/mc^2$, ответственный за утяжеление массы частицы (A — вектор-потенциал волны накачки, $p_0 = mc$ — импульс покоя). В частности, это позволяет частице достичь импульсов после рассеяния много больших (точнее в α^2 раз больших) осцилляторного импульса. Подробнее об этом расскажем чуть ниже.

Здесь же стоит заметить, что все эффекты, рассмотренные ранее, носят чисто кинематический характер и связаны с притяжением осциллирующего электрона к иону. При этом существенно наличие двух условий, сохраняющихся и в релятивистском случае: притягивающего центра иона и периодического внешнего поля. Осцилляции частицы при этом будут очень сложные, но качественный характер столкновений не изменится.

Действительно, из точных формул рассеяния релятивистской частицы в поле иона [11] следует простая связь изменения импульса частицы Δp_{\perp} с ее прицельным параметром ρ (сравните с (5)):

$$\Delta p_{\perp} = p_{osc} b_{osc} / \rho \ll p_{osc}. \quad (23)$$

Однако для $b_{osc} = e^2 Z / p_{osc} v_{osc}$ надо подставлять релятивистское выражение, учитывающее релятивистскую связь скорости и импульса $v_{osc} = p_{osc} c / \sqrt{p_{osc}^2 + p_0^2}$. Находя отсюда величину изменения поперечной скорости $\Delta v_{\perp} = \Delta p_{\perp} c / \sqrt{p_{osc}^2 + p_0^2}$, легко прийти к отображению вида (4). При этом сохраняется и определение для $r_E = \sqrt{r_{osc} b_{osc}}$, в которое надо подставить релятивистское выражение для радиуса осцилляций и резерфордского радиуса:

$$r_{osc} = \frac{p_{osc} c}{\omega \sqrt{p_{osc}^2 + p_0^2}}, \quad b_{osc} = \frac{e^2 Z \sqrt{p_{osc}^2 + p_0^2}}{p_{osc}^2 c}. \quad (24)$$

Таким образом, все представленные выше выводы остаются в силе с учетом релятивистского определения параметров задачи рассеяния (24). В частности, отсюда следует вывод о нарастании Джоулева нагрева в релятивистски сильных полях. Подставляя в формулу (22) релятивистское значение резерфордского радиуса (24), легко видеть зависимость $\sigma_{eff} \sim 1/p_{osc}$ при $p_{osc} \gg p_0$, которая приводит к росту частоты столкновений (18) пропорционально амплитуде поля накачки в ультрарелятивистских полях.

В релятивистских полях присутствует еще один «забавный» эффект — ускорение рассеявшихся частиц в поле волны накачки в α^2 раз. Подробно этот вопрос рассмотрен в [12]. Здесь же мы выполним простейшие оценки. Зададимся вопросом о величине максимальной энергии, которую может получить частица в релятивистски сильных полях. Например, известно [13], что при ионизации газа в релятивистски сильном лазерном поле частицы могут иметь энергию порядка $p_{osc}^2 / m = mc^2 \alpha^2$. Возникает вопрос, можно ли при столкновениях получить большую энергию?

Ответ тривиально получается из формул для движения частицы в поле плоской монохроматической волны [11]:

$$\mathbf{p}_\perp = \mathbf{q} + \mathbf{p}_{osc}, \quad p_{||} = \frac{p_0^2 + q^2 + 2\mathbf{q}\mathbf{p}_{osc} + p_{osc}^2}{2\zeta} - \frac{\zeta}{2}, \quad (25)$$

где \mathbf{q} , ζ — интегралы движения, $\mathbf{p}_\perp, p_{||}$ — кинематический импульс частицы поперек и вдоль волнового вектора волны в лабораторной системе координат.

Считая удар мгновенным на периоде поля, т. е. фактически рассматривая скачкообразное изменение импульса частицы («инжекцию» частицы с новыми импульсами) и *учитывая* сохранение в адиабатическом приближении усредненного по фазе импульса после прохождения волнового пакета, легко получить ответ. Оптимальный режим столкновения достигается при «рождении» частицы с начальными импульсами $p_\perp = 0$, $p_{||} = p_{osc}$, что соответствует $\mathbf{q} = -\mathbf{p}_{osc}$ и $\zeta \approx p_0^2/2p_{osc}$. Значение усредненного по фазе волны (т. е. дрейфового) продольного импульса будет равно

$$\langle p_{||} \rangle \approx \frac{p_{osc}^2 + \langle p_{osc}^2 \rangle}{2\zeta} = \frac{p_{osc}(p_{osc}^2 + \langle p_{osc}^2 \rangle)}{p_0^2}. \quad (26)$$

Переписывая это выражение через величину α

$$w_{max} = c\langle p \rangle_{max} \approx mc^2 \cdot \alpha(\alpha^2 + \langle \alpha^2 \rangle) \sim 2mc^2\alpha^3,$$

получим рост максимальной энергии, пропорциональный α^3 ! Интересно, что в релятивизме с ростом амплитуды электромагнитной волны p_{osc} порог по энергии растет даже быстрее, чем в нерелятивистском случае ($w_{nonrel}^{max} \leq 2cp_{osc}^2/p_0 = 2mc^2\alpha^2$).

Вообще говоря, отсюда можно получить и достаточно неожиданный результат о возможности еще более сильного ускорения частицы в поле электромагнитной волны. Пусть частица рассеялась (например, на ионе) и получила импульс \mathbf{p} . Тогда вновь используя формулы (25), (26), найдем, что средний по фазе импульс будет равен:

$$\langle \mathbf{p}_\perp \rangle = \mathbf{q}, \quad \langle p_{||} \rangle = p_{||} - \frac{p_{osc}^2 + 2\mathbf{q}\mathbf{p}_{osc} - \langle p_{osc}^2 \rangle}{2(p_{||} - w/c)}. \quad (27)$$

Интегралы движения равны $\mathbf{q} = \mathbf{p}_\perp - \mathbf{p}_{osc}$, $\zeta = w/c - p_{||}$, энергия $w = c\sqrt{p^2 + p_0^2}$. В качестве значения p_{osc} берутся значения осциллирующего импульса в момент рассеяния.

Полагая $p_\perp = 0$ (частица рассеялась вдоль \mathbf{k}), а величину импульса достаточно большой $p \gg p_0$, сразу получим, что средний импульс

(и, соответственно, конечная энергия частицы) станет в α^2 раз больше, чем после рассеяния! Такой процесс можно назвать «кулоновским ускорителем».

Наконец найдем распределение таких частиц по энергии. Из задачи рассеяния релятивистской частицы [11] известно, что при малых прицельных параметрах $\rho \sim \beta = e^2 Z/mc^2$ рассеяние становится почти изотропным. Тогда с учетом формул (27) и (13) распределение по энергиям принимает простой вид для высокоэнергичных частиц $p > p_{osc}$:

$$g(w) = 2\pi\nu n_i n_e a\beta \int \delta\left(w - \frac{cp_{osc}}{1 + \mu - \cos\vartheta}\right) \sin\vartheta d\vartheta.$$

Здесь ϑ — угол между волновым вектором и кинематическим импульсом частицы после столкновения $p_x = p_{osc} \cos\vartheta$, $\mu = p_0^2/2p_{osc}^2 = 1/2\alpha^2 \ll 1$ — малый параметр задачи. Интеграл легко вычисляется с помощью определения δ -функции при условиях на энергию частицы $w > p_{osc}c$, $w < 2p_{osc}c\alpha^2$:

$$g(w) = 4\pi\nu n_i n_e a\beta \frac{p_{osc}c}{w^2} \sim \frac{n_i^2 Z^3}{w^2}. \quad (28)$$

Интересно отметить, что в результате грубого предположения об изотропии рассеяния был «утрачен» логарифмический фактор (в сравнении с [12]), но зависимость от основных параметров осталась правильной.

8. Заключение

Обсудим в заключение область применимости представленных выше результатов. Они получены в рамках приближения парных столкновений, когда считается, что вероятность столкновения одновременно трех и более частиц в одной точке мала. Условием выполнения этого приближения является условие малости объема взаимодействия:

$$nV_{int} \ll 1. \quad (29)$$

Традиционно (без поля накачки) объем взаимодействия оценивался как $V_{int} = b_v^3$, что приводило к оценке

$$nb_v^3 \ll 1 \iff nr_D^3 \gg 1, \quad (30)$$

где $r_D = \sqrt{2\pi e^2 n/T_e}$ — радиус Дебая. В сильных полях объем взаимодействия есть $V_{int} = \sigma_{eff} r_{osc}$, что приводит к условию

$$r_E \ll r_D. \quad (31)$$

Но это условие, очевидно, может быть получено и из других соображений. Действительно, новый масштаб, появляющийся при учете притяжения частиц r_E , — масштаб, с которого частица через осцилляцию попадает точно в ион (точнее, $\sqrt{2\pi}r_E$). Эффект притяжения не будет «замыт» соседними частицами, если этот масштаб будет меньше радиуса дебаевской экранировки r_D . Отсюда снова получаем условие (31).

Итак, предположение о двух видах рассеивающихся частиц (13) позволяет получить выражение для эффективной частоты столкновений и распределение горячих частиц по энергиям, хорошо согласующееся с экспериментом. Более того, учет «представительных» электронов (сингулярной части (13)) необходим для адекватного объяснения получаемых экспериментальных результатов. Подчеркнем, что использование описания, учитывающего только малоугловые столкновения (прямолинейные траектории), привело бы к одномасштабной функции распределения горячих частиц, что противоречит эксперименту.

Работа выполнена при поддержке грантов РФФИ № 05-02-17367 и 04-02-16684.

Литература

1. Dawson J., Oberman C. // Phys. Fluids. 1962. V. 5. P. 517; Силин В. П. // ЖЭТФ. 1964. Т. 47. С. 2254.
2. Силин В. П. // ЖЭТФ. 1998. Т. 114. С. 864.
3. Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. М.: Наука, 1979.
4. Бункин Ф. В., Казаков А. Е., Федоров М. В. // УФН. 1973. Т. 107. С. 559; Pert G. J. // Phys. Rev. E. 1995. V. 51. P. 4778.
5. Kroll N. M., Watson K. M. // Phys. Rev. A. 1973. V. 8. P. 804.
6. Decker C. D., Mori W. B., Dawson J. M., Katsouleas T. // Phys. Plasmas. 1994. V. 1. P. 4043; Wiesenfeld L. // Phys. Rev. A. 1990. V. 144. P. 467.
7. Pfalzner S., Gibbon P. // Phys. Rev. E. 1998. V. 57. P. 4698; Bornath T., Schlanges M., Hilse P., Kremp D. // Phys. Rev. E. 2001. V. 64. P. 026414.
8. Fraiman G. M., Mironov V. A., Balakin A. A. // Phys. Rev. Lett. 1999. V. 82. P. 319; Фрайман Г. М., Миронов В. А., Балакин А. А. // ЖЭТФ. 1999. Т. 115. С. 463; Fraiman G. M., Mironov V. A., Balakin A. A. // Phys. Plasma. 2001. V. 8. P. 2502.
9. Ландау Л. Д., Лифшиц Е. М. Механика. М.: Наука, 1988.
10. Фрайман Г. М., Миронов В. А., Балакин А. А. // ЖЭТФ. 2001. Т. 120. С. 797.
11. Ландау Л. Д., Лифшиц Е. М. Теория поля. М.: Наука, 1988.
12. Балакин А. А., Фрайман Г. М. // ЖЭТФ. 2006. Т. 130. С. 426.
13. Dodin I. Y., Fisch N. J., Fraiman G. M. // JETP Lett. 2003. V. 78. P. 238.

**Нелинейные эффекты
в гравитации
и астрофизической плазме**

ГРАВИТАЦИЯ И СУПЕРГРАВИТАЦИЯ

Ю. М. Зиновьев

Тема настоящей статьи — гравитация и супергравитация с точки зрения физики элементарных частиц. При этом очень полезной будет аналогия с электродинамикой.

1. Электродинамика

Уравнения Максвелла. Основу электродинамики составляют уравнения Максвелла, которые в трехмерных обозначениях (т. е. в терминах напряженности электрического и индукции магнитного полей) выглядят как:

$$\vec{\nabla} \cdot \vec{E} = \rho, \quad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t} + \vec{j}.$$

Второе и третье уравнения могут быть решены в терминах скалярного и векторного потенциалов:

$$\vec{E} = -\vec{\nabla} A_0 - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A}.$$

Потенциалы определены с точностью до калибровочного преобразования:

$$\vec{A}' = \vec{A} - \vec{\nabla} \Lambda, \quad A_0' = A_0 + \frac{\partial \Lambda}{\partial t}.$$

Как мы увидим далее, калибровочные преобразования играют исключительно важную роль не только в электродинамике, но и в (супер)гравитации.

Все дальнейшее изложение будет использовать четырехмерные (лоренц-ковариантные) обозначения. Для того чтобы переписать уравнения Максвелла в четырехмерных обозначениях, введем тензор напряженности электромагнитного поля:

$$F_{\mu\nu} = -F_{\nu\mu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}.$$

Тогда уравнения Максвелла можно записать в виде

$$\partial_\nu F^{\mu\nu} = j^\mu, \quad \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} + \partial_\alpha F_{\mu\nu} = 0.$$

Если ввести четырехмерный вектор-потенциал $A_\mu = (A_0, \vec{A})$, так что

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

то второе уравнение становится тождеством, а первое — лагранжевым. Действительно, если ввести лагранжиан электромагнитного поля

$$S = \int d^4x \mathcal{L}, \quad \mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + A_\mu \cdot j^\mu,$$

то вариация лагранжиана дает уравнение Максвелла:

$$\frac{\delta \mathcal{L}}{\delta A_\mu} = -\partial_\nu F^{\mu\nu} + j^\mu.$$

Лагранжиан инвариантен относительно калибровочных преобразований $A_\mu' = A_\mu + \partial_\mu \Lambda$ при условии сохранения тока $\partial_\mu j^\mu = 0$.

Фотон — безмассовая частица со спином 1. Пространство Минковского однородно и изотропно, что приводит к требованию инвариантности относительно трансляций и лоренцевских преобразований: $\delta x_\mu = a_\mu + \Lambda_\mu{}^\nu x_\nu$. Эти преобразования образуют группу, называемую группой Пуанкаре. Генераторы трансляций P_μ и лоренцевских вращений $M_{\mu\nu} = -M_{\nu\mu}$ образуют алгебру:

$$\begin{aligned} [P_\mu, P_\nu] &= 0, \\ [M_{\mu\nu}, P_\alpha] &= \eta_{\nu\alpha} P_\mu - \eta_{\mu\alpha} P_\nu, \\ [M_{\mu\nu}, M_{\alpha\beta}] &= \eta_{\nu\alpha} M_{\mu\beta} - \eta_{\nu\beta} M_{\mu\alpha} - \eta_{\mu\alpha} M_{\nu\beta} + \eta_{\mu\beta} M_{\nu\alpha}. \end{aligned}$$

Здесь $\eta_{\mu\nu}$ — метрика пространства Минковского.

Важнейшая роль, которую играет группа Пуанкаре в физике частиц, в первую очередь связана с тем, что именно группа Пуанкаре дает строгое определение понятия частицы как неприводимого представления группы Пуанкаре.

Группа Пуанкаре имеет два оператора Казимира (операторы, коммутирующие со всеми генераторами алгебры) — P^2 и W^2 , где $W^\mu = \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} P_\nu M_{\alpha\beta}$. Два типа неприводимых представлений, которые интерпретируются как частицы:

- массивные частицы

$$P^2 = m^2, \quad W^2 = m^2 s(s+1), \quad s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$$

(число независимых степеней свободы $2s+1$);

- безмассовые частицы

$$P^2 = 0, \quad W^2 = 0, \quad W_\mu = \lambda P_\mu, \quad \lambda = \pm s$$

(число независимых степеней свободы 2).

С точки зрения группы Пуанкаре вектор-потенциал A_μ (векторное поле) содержит состояния, соответствующие спинам 1 и 0 (или четырем спиральностям $\pm 1, 0, 0'$). Однако поперечность электромагнитных волн означает, что фотон имеет только две физические степени свободы, соответствующие спиральностям ± 1 , и является безмассовой частицей со спином 1. Для массивного векторного поля спин 1 можно выделить лоренц-ковариантным образом. Действительно, в этом случае

$$(W^2)_\mu{}^\nu = 2(m^2 \delta_\mu{}^\nu - \partial_\mu \partial^\nu)$$

и необходимое условие — это просто условие Лоренца: $\partial^\mu A_\mu = 0$. Для безмассовой частицы лоренц-ковариантным образом выделить спиральности ± 1 невозможно, поэтому единственный способ лоренц-ковариантного описания фотона — использование калибровочной инвариантности: $\delta A_\mu = \partial_\mu \Lambda$. Это относится и ко всем безмассовым частицам со спином $s > 1$. Более того, калибровочная инвариантность во многом фиксирует возможные взаимодействия таких частиц.

Нелинейные эффекты. Несколько слов о терминологии. То, что в волновой теории называется нелинейными эффектами, в физике частиц соответствует взаимодействию. Логика здесь следующая:

- Линейная теория \Rightarrow квадратичный по полям лагранжиан \Rightarrow линейные уравнения \Rightarrow плоские волны \Rightarrow свободные частицы.
- Нелинейная теория \Rightarrow кубические и т. д. члены в лагранжиане \Rightarrow квадратичные и т. д. члены в уравнениях \Rightarrow взаимодействующие частицы. -

Пример: взаимодействие фотонов с электронами: $\mathcal{L}_{int} = ie(\bar{\Psi}\gamma^\mu\Psi)A_\mu$.

В качестве иллюстрации того, как калибровочная инвариантность фиксирует возможные взаимодействия, рассмотрим возможные нелинейные эффекты в электродинамике.

Трехфотонное взаимодействие. Лоренц-инвариантность требует, чтобы число производных было нечетное. Для случая одной производной легко показать, что невозможно построить лагранжиан вида $\mathcal{L} \sim A\partial A$, инвариантный относительно (модифицированных) калибровочных преобразований $\delta A_\mu = \partial_\mu \Lambda + \kappa A_\mu \Lambda$. В случае взаимодействия с тремя производными единственное калибровочно инвариантное выражение — $\mathcal{L}_{int} \sim F^{\mu\nu} F_{\nu\alpha} F^\alpha{}_\mu \equiv 0$ в силу антисимметричности тензора $F_{\mu\nu}$. Таким образом, вместе лоренц-инвариантность и калибровочная инвариантность запрещают нелинейности такого порядка.

Четырехфотонное взаимодействие. Лоренц-инвариантность требует, чтобы число производных было четное. Для случая взаимодействия с двумя производными наиболее общий калибровочно инвариантный лагранжиан вида $\mathcal{L}_{int} \sim A\partial A\partial A$ заменой переменных вида $A_\mu \rightarrow A_\mu + \kappa A^2 A_\mu$ сводится к свободному! Для взаимодействия с четырьмя производными наиболее общий калибровочно инвариантный лагранжиан: $\mathcal{L}_{int} = \kappa_1 (F^{\mu\nu} F_{\mu\nu})^2 + \kappa_2 (\tilde{F}^{\mu\nu} F_{\mu\nu})^2$.

Здесь $\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$.

Заметим, что калибровочная инвариантность никак не фиксирует значения параметров κ_1 и κ_2 . Два примера того, как эти параметры могут фиксироваться.

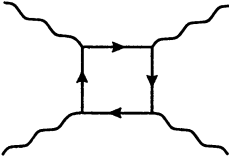
Электродинамика Борна – Инфельда. Это существенно нелинейный вариант электродинамики с лагранжианом

$$\mathcal{L}_{BI} = \sqrt{-\det(\eta_{\mu\nu} + \kappa F_{\mu\nu})}.$$

Разложение в случае слабого поля дает

$$\mathcal{L}_{BI} \sim 1 + \frac{\kappa^2}{4} F^2 - \frac{\kappa^4}{8} (F^4 - (F^2)^2) + \dots$$

Квантовые поправки от фермионных петель. Фотон-фотонное рассеяние на виртуальных парах (см. рисунок).



Этой теме была посвящена лекция В. В. Кочаровского на школе «Нелинейные волны – 2004» (Н. Новгород, 2005, с. 202–216).

2. Гравитация

Гравитон — безмассовая частица со спином 2. К сожалению, до сих пор ни в одном эксперименте не удалось непосредственно зарегистрировать гравитационные волны. Косвенное подтверждение их существования дают наблюдения за изменением орбитального периода двойных пульсаров. Например, для пульсара PSR-1913+16 (две нейтронные звезды, одна из которых пульсар)

$$\frac{P'_{exp}}{P'_{theor}} = 1,0013, \quad P' \sim 10^{-12}.$$

Более того, нет никакого аналога фотоэффекта, который показал бы, что гравитационные волны представляют собой поток частиц — гравитонов. В этих условиях попробуем решить обратную задачу. А именно, предположим, что гравитоны (как безмассовые частицы со спином 2) действительно существуют, и попробуем определить, как должна выглядеть теория, описывающая такие частицы.

Простейшее поле, которое может описывать частицы со спином 2, — симметричный тензор второго ранга $h_{\mu\nu} = h_{\nu\mu}$. Эти 10 степеней свободы соответствуют комбинации спинов 2, 1, 0 и 0'. 5 степеней свободы, соответствующих массивному спину 2, можно было бы выделить, наложив ковариантные условия (аналог условий Лоренца):

$$\partial_\mu h^{\mu\nu} = 0, \quad h = \eta_{\mu\nu} h^{\mu\nu} = 0.$$

Однако 2 степени свободы (спиральности ± 2), соответствующие безмассовой частице со спином 2, ковариантно выделить невозможно \Rightarrow требуется калибровочная инвариантность $\delta h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$.

Лоренц-инвариантность и калибровочная инвариантность однозначно определяют свободный (т. е. квадратичный) лагранжиан:

$$\mathcal{L}_0 = \frac{1}{2} \partial^\alpha h^{\mu\nu} \partial_\alpha h_{\mu\nu} - (\partial h)^\mu (\partial h)_\mu + (\partial h)^\mu \partial_\mu h - \frac{1}{2} \partial^\mu h \partial_\mu h.$$

Отметим здесь, что невозможно построить калибровочно инвариантный объект из первых производных поля $h_{\mu\nu}$, поэтому лагранжиан не является квадратом какого-либо калибровочно инвариантного тензора (аналогичного напряженности электромагнитного поля). Однако из вторых производных поля $h_{\mu\nu}$ можно построить калибровочно инвариантный тензор:

$$R_{\mu\nu,\alpha\beta} = \partial_\mu \partial_\alpha h_{\nu\beta} - \partial_\mu \partial_\beta h_{\nu\alpha} - \partial_\nu \partial_\alpha h_{\mu\beta} + \partial_\nu \partial_\beta h_{\mu\alpha},$$

причем $R_{\mu\nu,\alpha\beta} = -R_{\nu\mu,\alpha\beta} = -R_{\mu\nu,\beta\alpha} = R_{\alpha\beta,\mu\nu}$. Свойства симметрии этого тензора совпадают со свойствами симметрии тензора Римана, что может служить первым указанием на то, что наша (будущая) теория каким-то образом связана с геометрией.

Самодействие (нелинейность). Рассмотрим минимальное (кубическое) взаимодействие. Лоренц-инвариантность требует, чтобы число производных было четное. Легко показать, что калибровочно инвариантное взаимодействие без производных $\mathcal{L} \sim hhh$ отсутствует. Важно понимать, что взаимодействие с двумя производными $\mathcal{L} \sim h\partial h\partial h$ (как любое взаимодействие, в котором число производных

такое же, как в свободном лагранжиане) всегда определено с точностью до замены переменных вида

$$h_{\mu\nu} \rightarrow h_{\mu\nu} + \kappa_1 h_{\mu}{}^{\alpha} h_{\alpha\nu} + \kappa_2 h h_{\mu\nu} + \kappa_3 \eta_{\mu\nu} h^{\alpha\beta} h_{\alpha\beta} + \kappa_4 \eta_{\mu\nu} h^2.$$

Калибровочная инвариантность достигается за счет модификации калибровочных преобразований:

$$\delta' h_{\mu\nu} = \kappa(\xi^{\alpha} \partial_{\alpha} h_{\mu\nu} + \partial_{\mu} \xi^{\alpha} h_{\alpha\nu} + \partial_{\nu} \xi^{\alpha} h_{\mu\alpha}).$$

В системе единиц $h = c = 1$ константа взаимодействия имеет размерность $\kappa \sim \frac{1}{m_{pl}}$, где $m_{pl} \sim 10^{18}$ ГэВ.

Полная теория с необходимостью оказывается существенно нелинейной, $\mathcal{L} \sim (1 + \kappa h + \kappa^2 h^2 + \dots) \partial h \partial h$, при этом с точностью до возможных замен переменных она определена совершенно однозначно. Если ввести эффективную метрику $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$, то калибровочные преобразования принимают вид общекоординатных:

$$\delta g_{\mu\nu} = \xi^{\alpha} \partial_{\alpha} g_{\mu\nu} + \partial_{\mu} \xi^{\alpha} g_{\alpha\nu} + \partial_{\nu} \xi^{\alpha} g_{\mu\alpha}.$$

Пользуясь знанием дифференциальной геометрии можно «угадать» полный лагранжиан:

$$\mathcal{L} = -\frac{1}{2\kappa^2} \sqrt{-g} g^{\mu\nu} R_{\mu\nu}(g).$$

Здесь $g^{\mu\nu}$ — обратная метрика $g_{\mu\nu} g^{\nu\alpha} = \delta_{\mu}{}^{\alpha}$, $g = \det(g_{\mu\nu})$, а $R_{\mu\nu}$ — свертка тензора Римана для эффективной метрики $g_{\mu\nu}$. В полную теорию $\eta_{\mu\nu}$ и $h_{\mu\nu}$ входят только через $g_{\mu\nu}$, в результате мы «теряем» связь с исходным пространством Минковского!

Взаимодействие с фотонами. В качестве примера взаимодействия с материей рассмотрим взаимодействие с фотонами. Минимальное взаимодействие — кубическое с двумя производными $\mathcal{L}_{int} \sim h \partial A \partial A$. Так же как и самодействие, оно определено с точностью до замены переменных $A_{\mu} \rightarrow A_{\mu} + \kappa_1 h_{\mu\nu} A^{\nu} + \kappa_2 h A_{\mu}$. Удобный выбор (соответствует тому, что A_{μ} — ковариантный вектор):

$$\mathcal{L}_{int} = \frac{\kappa}{2} h^{\mu\nu} F_{\mu\alpha} F_{\nu\alpha} - \frac{\kappa}{8} h F^{\alpha\beta} F_{\alpha\beta}, \quad \delta A_{\mu} = \kappa \xi^{\alpha} F_{\alpha\mu}.$$

Отметим, что этого приближения уже достаточно, например, чтобы описать отклонение света в поле солнца. Полная нелинейная теория также с необходимостью существенно нелинейна и требует введения эффективной метрики $g_{\mu\nu}$: $\mathcal{L} = -\frac{1}{4} \sqrt{-g} g^{\mu\alpha} g^{\nu\beta} F_{\mu\nu} F_{\alpha\beta}$.

3. Супергравитация

Бозоны и фермионы. В физике частиц (не только элементарных) важную роль всегда играло деление частиц на два больших класса: бозоны — частицы с целыми спинами — и фермионы — с полуцелыми. Во многих отношениях эти частицы ведут себя совершенно по-разному вследствие того, что они подчиняются разной статистике: Бозе — Эйнштейна и Ферми — Дирака соответственно. До сих пор и роли, которые они играют в физике элементарных частиц, существенно различаются: фермионы (кварки, лептоны) играют роль «материи», а бозоны (фотон, глюоны, W- и Z-бозоны, гравитон) — роль «переносчиков взаимодействия».

Как мы видели, гравитация с необходимостью имеет размерную константу взаимодействия и, как следствие, неперенормируема. При этом мы знаем, что вследствие разной статистики в квантовые поправки бозонные и фермионные петли дают вклады разного знака. Вопрос: нельзя ли добиться сокращения расходимостей, используя то, что фермионные петли дают вклады противоположного знака?

Если «рассортировать»

известные нам элементарные частицы, которые на сегодняшний день считаются элементарными) по спинам, то мы увидим, что в этой таблице есть «недостающее звено» — см. след. подраздел.

Спин	Частицы
2	гравитон
3/2	
1	фотон, глюоны, W- и Z-бозоны
1/2	кварки, лептоны
0	частица Хигса (?)

Гравитино — частица со спином 3/2. Так же как и в случае с гравитоном, до сих пор нет никаких экспериментальных данных, которые указывали бы на существование элементарной частицы со спином 3/2. Поэтому предположим, что такая частица существует и попробуем понять, какая теория должна ее описывать.

Простейшее поле, которое содержит спин 3/2 — вектор-спинор (или спин-вектор) Ψ_μ . Его 16 степеней свободы соответствуют частицам со спинами 3/2, 1/2 и 1/2' (и соответствующим античастицам). Массивную частицу со спином 3/2 можно выделить, наложив ковариантные условия $\partial^\mu \Psi_\mu = 0$, $\gamma^\mu \Psi_\mu = 0$, где γ_μ — матрицы Дирака: $\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}$. Безмассовую частицу ковариантно выделить невозможно, поэтому нам нужна калибровочная инвариантность (ζ — спинор!): $\delta \Psi_\mu = \partial_\mu \zeta$. Калибровочно инвариантный квадратичный лагранжиан определен однозначно:

$$\mathcal{L} = \frac{i}{2} \varepsilon^{\mu\nu\alpha\beta} \bar{\Psi}_\mu \gamma_5 \gamma_\nu \partial_\alpha \Psi_\beta.$$

Попробуем включить минимальное электромагнитное взаимодействие по аналогии с электродинамикой (заменяв производную на калибровочно ковариантную $\partial_\mu \rightarrow D_\mu = \partial_\mu + ieA_\mu$):

$$\mathcal{L}_{int} = -\frac{e}{2} \varepsilon^{\mu\nu\alpha\beta} \bar{\Psi}_\mu \gamma_5 \gamma_\nu A_\alpha \Psi_\beta.$$

Такая теория инвариантна относительно калибровочных преобразований со скалярным параметром $\Psi_\mu \rightarrow \exp(ie\Lambda) \Psi_\mu$, $A_\mu \rightarrow A_\mu + \partial_\mu \Lambda$, но неинвариантна относительно преобразований со спинорным параметром $\delta\mathcal{L} = e \bar{\Psi}_\mu \gamma_5 \gamma_\nu \tilde{F}^{\mu\nu} \zeta$, причем инвариантности невозможно добиться ни включением неминимальных взаимодействий, ни поправками к законам калибровочных преобразований! Это первый (но, к сожалению, не единственный) случай, когда нам не удастся включить электромагнитные взаимодействия для заряженной частицы. В данном случае причина оказалась в том, что непротиворечивое описание частицы со спином 3/2 требует присутствия частицы со спином 2 — гравитона! Теория, описывающая одновременно обе эти частицы, получила название супергравитации.

Супергравитация ($h_{\mu\nu}$, Ψ_μ). Минимальное взаимодействие с гравитоном имеет структуру $\mathcal{L}_{int} \sim \kappa h \bar{\Psi} \partial \Psi + \kappa \partial h \bar{\Psi} \Psi$. Калибровочная инвариантность возможна только при нетривиальных законах преобразований для полей: $\delta h_{\mu\nu} = i\kappa(\bar{\Psi}_{(\mu} \gamma_{\nu)} \zeta)$, $\delta \Psi_\mu = -\kappa \sigma^{\alpha\beta} \partial_\alpha h_{\beta\mu} \zeta$. ζ -преобразования дают первый пример суперсимметрий, т. е. симметрий, которые объединяют в один (супер)мультиплет частицы с разными спинами (целыми и полуцелыми!). Если ввести генератор таких преобразований Q (спинор), то получим алгебру:

$$\{Q, Q\} = 2i\kappa(\gamma_0 \gamma_\mu) P^\mu, \quad [P_\mu, Q] = 0, \quad [M_{\mu\nu}, Q] = \frac{1}{2} \sigma_{\mu\nu} Q,$$

которая представляет собой нетривиальное расширение алгебры Пуанкаре спинорными (антикоммутирующими в соответствии со своей статистикой) генераторами.

Полная нелинейная теория помимо гравитационных взаимодействий содержит четырехфермионные взаимодействия для гравитино $\mathcal{L} \sim \kappa^2 (\bar{\Psi} \Psi)^2$. Все взаимодействия в такой теории определяются одной константой — константой гравитационного взаимодействия $\kappa \sim 1/m_{pl}$. Как следствие единственности константы взаимодействия и сочетания бозонных и фермионных степеней свободы, однопетлевые и двухпетлевые квантовые поправки оказываются не просто перенормируемыми, а конечными!

В старших петлях расходимости все-таки появляются, поэтому для построения квантовой теории необходимо дальнейшее расширение теории. Наиболее перспективными считаются расширения, использующие конформные симметрии и переход от точечных частиц к (супер)струнам.

Суперпартнеры. На первый взгляд, наши рассуждения о расширении гравитации до супергравитации (и далее) не имеют отношения к остальным частицам и остальным взаимодействиям.

Частицы	Суперпартнеры
гравитон ($s = 2$)	гравитино ($s = 3/2$)
фотон ($s = 1$)	фотино ($s = 1/2$)
глюон ($s = 1$)	глюино ($s = 1/2$)
кварк ($s = 1/2$)	скварк ($s = 0$)
лептон ($s = 1/2$)	слептон ($s = 0$)

Однако в таком варианте гравитации материя тоже должна быть суперсимметрична, т. е. все частицы должны группироваться в супермультиплеты. Но так как суперпреобразования (как и преобразования группы Пуанкаре)

коммутируют со всеми внутренними симметриями, в мультиплеты должны объединяться частицы с одинаковыми квантовыми числами (электрический заряд, изоспин и т. д.). Это приводит к необходимости удвоения спектра частиц! Таким образом, суперсимметрия предсказывает существование целой плеяды новых элементарных частиц (см. таблицу). Более того, все взаимодействия новых частиц определяются суперсимметрией, произвол остается только в массах.

Заключение

Гравитация и супергравитация (как и электродинамика) могут быть поняты просто как теории, описывающие взаимодействия безмассовых частиц со спинами 2 и 3/2 (и 1) на основе лоренц-инвариантности и калибровочной инвариантности. Экспериментально супергравитация открывает возможность появления целого поколения новых частиц, и поиск таких частиц составляет существенную часть программы будущих экспериментов на новом ускорителе в ЦЕРНе (Женева), который должен начать работать в 2007 году.

Но даже если суперсимметрия будет подтверждена экспериментально, это будет лишь косвенным указанием на существование супергравитации. Остается вопрос: возможно ли непосредственное исследование гравитационных взаимодействий в ускорительных экспериментах?

ФИЗИКА ЧАСТИЦ НА УСКОРИТЕЛЯХ И ГРАВИТАЦИЯ

Ю. М. Зиновьев

В настоящей статье обсуждается возможность исследования гравитационных взаимодействий в экспериментах на ускорителях. Общепринятым является мнение, что гравитационные взаимодействия слишком слабы для того, чтобы это стало возможным. Чтобы понять, насколько это утверждение окончательно, полезно, прежде всего, разобраться, что означают слова: «Гравитация — сильная или слабая?»

1. Гравитация — сильная или слабая?

В системе единиц, в которой $\hbar = 1$ и $c = 1$, размерности всех физических величин являются степенями массы (или длины), например, $[x] \sim m^{-1}$, $[\partial] = m$ и т. д. Размерности полей определяются из требования безразмерности действия $S = \int d^4x \mathcal{L} \Rightarrow [\mathcal{L}] \sim m^4$. Для всех бозонных полей (фотон, гравитон) свободный лагранжиан имеет структуру $\mathcal{L} \sim \partial\Phi\partial\Phi$, следовательно, размерность любого бозонного поля $[\Phi] \sim m$. В то же время для фермионов (электрон, гравитино) $\mathcal{L} \sim \bar{\Psi}\hat{\partial}\Psi$, а значит, все фермионы имеют размерность $[\Psi] \sim m^{3/2}$.

Зная размерности полей, мы можем определять размерности констант взаимодействия. Например, константа взаимодействия фотонов и электронов (электрический заряд) $\mathcal{L}_{int} = ie(\bar{\Psi}\gamma^\mu\Psi)A_\mu$ безразмерна, и ее можно определить из постоянной тонкой структуры $\alpha = \frac{e^2}{4\pi} \sim \frac{1}{137}$. Именно малость этой безразмерной величины и имеют в виду, когда говорят о слабости электромагнитных взаимодействий.

Константа гравитационных взаимодействий, например, гравитона с фотонами $\mathcal{L}_{int} \sim \kappa h\partial A\partial A$ имеет размерность $\kappa \sim \frac{1}{m_{pl}}$, где $m_{pl} \sim 10^{18}$ ГэВ — масса Планка. Поэтому сила или слабость гравитационных взаимодействий определяется отношением $\frac{E}{m_{pl}}$, где E — характерные энергии частиц, участвующих во взаимодействии. Приведем в таблице некоторые характерные значения энергий.

Здесь ЛНС — новый строящийся в Женеве ускоритель, а SSA — гипотетический ускоритель размером с нашу солнечную систему (решение о строительстве пока не принято).

Сравнивая энергию ЛНС с массой Планка, которая и определяет характерный масштаб гравитационных взаимодействий, видим, что действительно гравитационные эффекты в ускорительных экспериментах должны быть пренебрежимо малы. Тем не менее, в наших рассуждениях есть естественные (даже очевидные) предположения, которые, как это не раз случалось в физике, могут оказаться неверными.

10^{18} ГэВ	масса Планка
10^{15} ГэВ	«великое объединение»
10^{13} ГэВ	SSA — Solar System Accelerator
10^4 ГэВ	LHC — Large Hadronic Collider
10^3 ГэВ	Tevatron
10^2 ГэВ	W - и Z -бозоны, t -кварк
1 ГэВ	протон, нейтрон

2. Идея Калуцы и Клейна

Все наши рассуждения о размерностях физических величин исходили из предположения о том, что наше пространство — время четырехмерно: $S = \int d^4x \mathcal{L}$. Однако с уверенностью мы можем говорить только о числе макроскопических измерений (например, в пятимерном пространстве в законе всемирного тяготения и в законе Кулона зависимость от расстояния была бы R^{-3} , а не R^{-2}). Идея Калуцы и Клейна состоит в том, что могут быть компактные микроскопические измерения.

На примере электродинамики и гравитации рассмотрим, что это означало бы с точки зрения физики частиц и как такие микроскопические измерения проявлялись бы.

2.1. Гравитация и электродинамика. Действие для электродинамики в пятимерном пространстве

$$S = \int d^4x \int dx_5 \left\{ -\frac{1}{4} F^{MN} F_{MN} \right\}, \quad F_{MN} = \partial_M A_N - \partial_N A_M,$$

где $M, N = 0, 1, 2, 3, 5$. При этом с четырехмерной точки зрения пятимерный вектор $A_M = (A_\mu, A_5)$ эквивалентен четырехвектору и скаляру. Пусть координата x_5 соответствует компактному измерению — окружности радиуса R , и следовательно все функции должны быть периодическими по x_5 с периодом $2\pi R$. Тогда все поля можно разложить в ряд Фурье:

$$A_\mu(x_\mu, x_5) = A_\mu^0(x) + \sum_{n=1}^{\infty} A_\mu^n(x) \exp\left(\frac{inx_5}{R}\right),$$

$$A_5(x_\mu, x_5) = \varphi(x) + \sum_{n=1}^{\infty} A_5^n(x) \exp\left(\frac{inx_5}{R}\right).$$

Подставив эти разложения в исходное действие и выполнив интегрирование по координате x_5 , получим эффективное *четырёхмерное* действие

$$S = \int d^4x \left[-\frac{1}{4} F_{\mu\nu}^0 F_{\mu\nu}^0 + \frac{1}{2} \partial^\mu \varphi \partial_\mu \varphi + \sum_{n=1}^{\infty} \left(-\frac{1}{4} F_{\mu\nu}^n F_{\mu\nu}^n + \frac{n^2 M^2}{2} A_\mu^n A_\mu^n \right) \right].$$

Таким образом, теория содержит безмассовое векторное и скалярное поля, а также *бесконечный* набор массивных векторных полей, массы которых определяются радиусом компактного измерения $M_n = \frac{n}{R}$. При этом, пока энергии наших ускорителей малы по сравнению с характерным масштабом $E \ll M$, мы будем видеть только безмассовые поля.

Пятимерная гравитация описывается симметричным тензором второго ранга h_{MN} . С четырёхмерной точки зрения это эквивалентно тензору $h_{\mu\nu}$, вектору $A_\mu = h_{\mu 5}$ и скаляру $\varphi = h_{55}$. В случае, когда пятое измерение — компактное, такая теория эквивалентна четырёхмерной гравитации (с безмассовым гравитоном), безмассовым векторному и скалярному полям, а также бесконечному набору массивных частиц со спином 2 с массами $M_n = \frac{n}{R}$. Однако отождествить такое безмассовое векторное поле с фотоном (это означало бы объединение гравитации и электродинамики, что и было исходным замыслом Калуцы и Клейна) не удастся, поскольку константа взаимодействия этого поля с другими полями (аналог электрического заряда) определяется отношением $\frac{m}{m_{pl}}$, где m — масса частицы, а m_{pl} — масса Планка. Так же как и в случае электродинамики, на низких энергиях мы наблюдали бы только безмассовые частицы.

2.2. Супергравитация. После неудачной попытки объединения гравитации и электромагнетизма идеи Калуцы и Клейна долгое время практически не обсуждались. С развитием супергравитации они снова стали востребованными, но уже на новом уровне в виде идеи о спонтанной компактификации.

В гравитации (и супергравитации) гравитон входит только в виде метрики эффективного Риманова пространства $g_{\mu\nu}$ без явного деле-

ния на метрику пространства Минковского $\eta_{\mu\nu}$ и гравитационное поле $h_{\mu\nu}$. В качестве «фоновой» метрики, т. е. метрики, определяющей свойства пространства, может выступать любое решение лагранжевых уравнений теории $\bar{g}_{\mu\nu}: R_{\mu\nu}(g) - \frac{1}{2}g_{\mu\nu}R = \dots$. Если ввести разложение $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$, то получим теорию, описывающую безмассовый гравитон (и другие частицы), «живущие» в пространстве с метрикой $\bar{g}_{\mu\nu}$.

Во многих моделях супергравитации действительно существуют решения гравитационных уравнений, соответствующие пространствам вида $M_4 \otimes K_d$, где M_4 — четырехмерное пространство (Минковского или (анти) де Ситтера), а K_d — d -мерное компактное пространство. Это значит, что специфическая структура пространства с компактными измерениями не постулируется, а возникает как решение уравнений супергравитации — спонтанная компактификация!

Вместе с тем, супергравитация и спонтанная компактификация возродили надежды на суперобъединение — объединение всех частиц и всех взаимодействий в рамках единой теории. При компактификации в качестве «побочного продукта» возникают безмассовые поля с низкими спинами. Например, при компактификации вида $M_{4+d} \rightarrow M_4 \otimes K_d$ от гравитона h_{MN} мы получим (помимо 4-мерного гравитона) d векторных полей и $\frac{d(d+1)}{2}$ скалярных. При компактификации гравитино (с учетом разной размерности спиноров) получим 2^k гравитино и $d \times 2^k$ спиноров, где $k = \lfloor \frac{d}{2} \rfloor$. Появляется возможность сформулировать теорию с небольшим числом полей и произвольных констант, в которой все известные нам частицы и их взаимодействия возникают в результате спонтанной компактификации \Rightarrow суперединая теория!

Более того, суперсимметрия с ее требованием равного числа бозонных и фермионных степеней свободы приводит к ограничению на размерность пространства $d \leq 11$!

2.3. Суперструны. Дальнейшую поддержку идея о возможности существования дополнительных измерений получила в теории суперструн, интерес к которой связан с возможностью решить проблемы квантовой гравитации. В теории струн на смену точечным частицам приходит протяженный одномерный объект — струна. В силу одномерности и конечной длины такая теория эквивалентна локальной теории, описывающей бесконечный набор точечных частиц со всевозможными спинами и массами. В суперструне они группируются в супермультиплеты. Спектр безмассовых состояний любой (супер)струны всегда конечен и не содержит частиц со спинами больше 2, следовательно, эффективной низкоэнергетической теорией для суперструны является

ся супергравитация. Замечательно, что суперструна является конечной теорией (т. е. все квантовые поправки вычисляемы)! При этом суперструна может быть непротиворечиво сформулирована только в пространстве размерности $d = 10$!

Немаловажно и то, что суперструна представляет собой пример суперединой теории, т. е. теории с небольшим числом исходных объектов и произвольных параметров. При этом суперструна действительно допускает варианты компактификации вида $M_{10} \rightarrow M_4 \otimes K_6$, где M_4 — обычное пространство Минковского. Однако число вариантов таких компактификаций огромно, при этом ни один из известных не удовлетворяет всем требованиям к реалистической теории. Остается открытым и вопрос о критериях выбора одного из огромного числа возможных вариантов такой компактификации.

3. Жизнь на «бране»

Вернемся теперь к вопросу о возможности исследования гравитации в ускорительных экспериментах и о том, могут ли дополнительные измерения повлиять на наши рассуждения.

В пятимерном пространстве размерность бозонных полей $\Phi \sim m^{\frac{3}{2}}$. Соответственно размерность константы гравитационного взаимодействия $\kappa_5 \sim M^{-\frac{3}{2}}$, где M — характерный масштаб энергий. При компактификации одного измерения эффективная константа определяется соотношением $m_{pl}^2 \sim M^3 R$, где R — радиус компактификации. Таким образом, в этом случае масса Планка — это просто некий эффективный параметр, а истинный масштаб гравитационных явлений определяется новым параметром M , который, в принципе, может быть существенно меньше.

Электромагнитные (а также сильные и слабые) взаимодействия проверены уже до расстояний порядка 10^{-18} м, при этом никаких отклонений от предсказаний «обычной» четырехмерной теории не наблюдается. Казалось бы, это значит, что $M > 10^{13}$ ГэВ, что, конечно, существенно ниже массы Планка (и даже сопоставимо с энергией нашего ускорителя размером с солнечную систему), но все-таки еще далеко от наших экспериментальных возможностей.

Прежде чем сформулировать следующую идею, приведем простую аналогию. Предположим, что мы изучаем волны в жидкости, объем которой ограничен некоторой поверхностью. При этом мы немедленно увидим, что в нашей задаче есть два очень различных типа волн. Одни — это продольные волны, «живущие» во всем объеме жидкости, а другие — поперечные, «живущие» только на ее поверхности. При

этом ничто не мешает этим двум типам волн взаимодействовать между собой.

В отличие от других взаимодействий, где есть ускорительные эксперименты, гравитация (R^{-2} в законе всемирного тяготения) проверена только до расстояний порядка 0,1 мм. Предположим теперь, что наше пространство — пятимерно и ограничено с одной или двух сторон (гипер)поверхностями. Предположим также, что гравитон (а также гравитино и, возможно, некоторые другие частицы) «живет» во всем пятимерном пространстве, а остальные частицы (кварки, лептоны, фотоны, а значит, и мы сами) «живут» на поверхности. Тогда размер пятого измерения мог бы быть существенно больше, чем все, что мы обсуждали до сих пор, вплоть до 10^{-6} м!

Отметим, что в современной терминологии (жаргоне) d -мерная гиперповерхность в пространстве большей размерности называется d -браной (т. е. струна — это 1-брана, мембрана — 2-брана и т. д.). Таким образом, если такой сценарий действительно реализуется, то все мы живем на некоторой 3-бране в пространстве большей, чем привычные 4 размерности! Конечно, такая гипотеза кажется более чем экзотической, но на сегодняшний день это единственная возможность того, что гравитационные явления со временем могут стать доступны для экспериментов на ускорителях.

Какие наблюдаемые проявления при этом возможны? Самое непосредственное — это отклонения в законе всемирного тяготения на микроскопических (микронных?) расстояниях: $R^{-2} \rightarrow R^{-3}$. В ускорительных экспериментах это, в первую очередь, «потерянные» энергия и импульс в экспериментах по рассеянию обычных частиц, однофотонная аннигиляция (аннигиляция электрона и позитрона в один фотон) с испусканием нерегистрируемого («улетающего в пятое измерение») гравитона, вклады от обменов виртуальными гравитонами и их интерференция с обычными процессами. Наконец, на самых высоких энергиях начнут проявляться эффекты квантовой гравитации типа рождения микроскопических черных дыр и т. д.

Заключение

Есть целый ряд серьезных (но пока чисто теоретических) аргументов в пользу того, что наше пространство имеет размерность больше 4. Наиболее важный — теория суперструн, единственная на сегодняшний день конечная квантовая теория, включающая гравитацию. Это открывает возможность того, что эффекты гравитации начнут проявляться на гораздо менее высоких энергиях, чем масса Планка, и могут быть доступны в экспериментах на ускорителях.

«ЦЕНТРАЛЬНАЯ МАШИНА» В АКТИВНЫХ ГАЛАКТИЧЕСКИХ ЯДРАХ

В. С. Бескин

1. Струйные выбросы

Наблюдения показывают, что существование струйных выбросов (джетов), т. е. узконаправленных потоков вещества, является общим свойством, характерным для очень многих астрономических источников (рис. 1). Они наблюдаются как в активных ядрах удаленных галактик, так и в более близких объектах — микрокваразах и радиопульсарах — расположенных в нашей Галактике. Еще одним классом объектов, обладающих струйными выбросами, являются молодые звезды. Есть все основания полагать, что природа активности гамма-всплесков — космических источников, находящихся на космологических расстояниях, — также связана с направленными выбросами вещества.

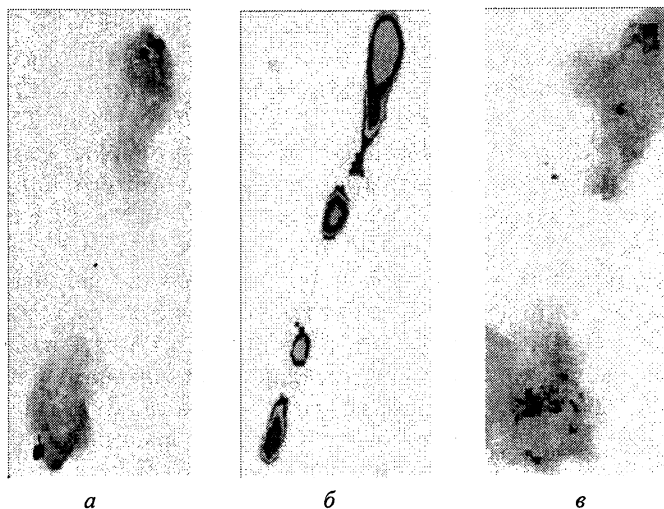


Рис. 1. Струйные выбросы, наблюдаемые в активных галактических ядрах (*а*), микрокваразах (*б*) и молодых звездах (*в*).

Уже через несколько лет после открытия на рубеже шестидесятых годов радиогалактик — диффузионных областей, излучающих в радиодиапазоне (рис. 1, *а*), — стало ясно, что их активность связана со струйными выбросами, истекающими из квазаров (по-английски QSO — квазизвездный объект). Квазары представляют собой ядра активных галактик, столь далеких, что их оптическое изображение неотличимо от обыч-

ных звезд. Именно джеты поставляют в эти области вещество и энергию, которые сам джет черпает из активного ядра галактики. Вещество джетов движется со скоростями v , близкими к скорости света c , так что лоренц-фактор джета как целого $\gamma = (1 - v^2/c^2)^{-1/2}$ достигает как минимум нескольких единиц. Характерные продольные размеры струйных выбросов составляют сотни тысяч световых лет, что сравнимо с расстояниями между галактиками. Другим удивительным свойством джетов является высокая степень их коллимации — угол расхождения составляет всего несколько градусов. При этом потери энергии, связанные со струйными течениями, могут достигать 10^{45} эрг/с, что по порядку величины совпадает с полной светимостью самой галактики. В настоящее время уже мало кто сомневается в том, что источник активности квазаров связан с аккрецией (падением) вещества на сверхмассивные черные дыры (их массы оценивают в миллиарды солнечных масс), поскольку лишь в этом случае удастся объяснить высокую эффективность энерговыделения этих удивительных источников [1].

Напротив, недавно открытые микроквазары (рис. 1, б) — это объекты, расположенные в нашей Галактике. Формирование их струйных выбросов связывают со сверхкритической аккрецией на черную дыру или нейтронную звезду солнечных масс, когда компактный объект не в состоянии поглотить все падающее на него вещество [2]. В этом случае часть аккрецирующего вещества неизбежно должна быть отброшена в окружающее пространство. Характерный продольный размер таких струйных выбросов составляет обычно доли светового года, и при этом угол раствора также не превышает нескольких градусов. Общая же мощность излучения оценивается в 10^{37} эрг/с, что лишь в тысячу раз больше светимости Солнца. Несмотря на большое различие в размерах и энерговыделении, внешне струйные выбросы у квазаров и микроквазаров оказываются чрезвычайно похожими, что и объясняет возникновение самого термина микроквазар.

Наконец, молодые звезды были впервые обнаружены в начале пятидесятых годов, когда Дж. Хербиг и Дж. Аро открыли новый класс протяженных диффузионных образований (рис. 1, в), существующих обычно парами и, как стало ясно значительно позже, также связанных струйными течениями с молодыми быстровращающимися звездами. У 60% источников видны оба струйных выброса, а в остальных случаях удаленный от нас выброс закрыт аккреционным диском. Протяженность джетов может достигать нескольких световых лет при характерной скорости движения вещества 600 км/с. Как мы видим, ситуация здесь вполне аналогична активным галактическим ядрам, когда вначале был обнаружен целый ряд разнородных источников (квазары, радиогалактики) и лишь затем стало ясно, что активность всех этих объектов имеет единую природу. Более того, подобие наблюдаемых свойств наводит на мысль, что и физический

механизм формирования направленных течений у молодых звезд может быть подобен механизму формирования струйных выбросов в активных ядрах. И это несмотря на то, что физические условия вблизи молодых звезд (их массы порядка 3—10 масс Солнца, а полная излучаемая мощность составляет 10^{31} — 10^{36} эрг/с) существенно отличаются от условий в центре активных галактических ядер. И одним из главных отличий является нерелятивистский характер истечения вещества из молодых звезд.

Если же говорить об источниках космологических гамма-всплесков, то здесь имеются лишь косвенные, хотя и достаточно надежные аргументы в пользу существования у них струйных выбросов, причем именно связанных с релятивистскими течениями, о которых пойдет речь в этой статье. Как известно, обнаружение оптического послесвечения (afterglow), позволившее по красному смещению определить расстояния до этих объектов, накладывает очень серьезные ограничения на их энергетику [3]. Если считать, что излучение в гамма-диапазоне происходит изотропно, то для расстояний порядка 10^9 световых лет, характерных для таких источников, приходится предположить, что их полная светимость достигает 10^{54} эрг/с. Однако в настоящее время процессы с таким огромным энерговыделением нам не известны. Если же считать, что излучение происходит в узкий конус углов θ порядка 1 градуса, то излучаемая мощность может быть понижена до 10^{51} эрг/с, что по порядку величины уже близко к энергии, выделяющейся во время взрывов сверхновых. Но это означает, что характерный лоренц-фактор частиц, с которыми связано наблюдаемое излучение, должен составлять величину в несколько сотен ($\gamma \sim 100$ — 300). Что же касается самой центральной машины, приводящей к образованию струйных выбросов, то здесь обычно обсуждается либо столкновение двух нейтронных звезд или нейтронной звезды и черной дыры, либо коллапс массивного ядра необычной сверхновой (гиперновой, супрановой и т. д.). Однако в большинстве моделей в итоге все равно образуется быстровращающаяся черная дыра с массой порядка массы Солнца, окруженная аккреционным диском. Иными словами, модель вновь строится по той же схеме, что и для активных галактических ядер [3].

2. Источник активности — вращающаяся черная дыра

Таким образом, струйные выбросы существуют не только в релятивистских объектах, в которых истекающее вещество имеет скорости, близкие к скорости света, но и в молодых звездах, где скорости истекающего вещества не являются релятивистскими. Поэтому причина их образования должна быть достаточно универсальной. Как мы видели, во всех рассмотренных выше примерах таким общим свойством была аккреция вещества на компактный объект. При этом наличие аккреционных дисков

однозначно указывало на то, что падающее вещество обладает большим угловым моментом (в случае молодых звезд, помимо диска, вращается и сама молодая звезда). Но хорошо известно, что существование углового момента препятствует падению на гравитирующий центр. Если потери энергии и углового момента отсутствуют, то вращение может продолжаться бесконечно долго. Именно такая ситуация имеет место в солнечной системе, когда планеты в среднем находятся на постоянном расстоянии от Солнца.

С другой стороны, было ясно, что единственным механизмом, способным легко объяснить высокую энергетику компактных объектов, могло быть выделение гравитационной (т. е. попросту потенциальной) энергии при падении вещества на компактный объект. Поэтому для объяснения высокой активности компактных объектов (которая как раз и показывает, что темп аккреции вещества во всех случаях достаточно велик) необходимо было найти эффективный механизм отвода углового момента за пределы аккреционного диска. Как оказалось, именно эту роль и играют струйные выбросы. Для молодых звезд при этом более важно, что, помимо отвода углового момента из аккреционного диска, истечение вещества уменьшает также и угловой момент вращающейся молодой звезды, помогая ей сжаться до необходимых размеров.

Построение модели центральной машины распадается на несколько частей. Необходимо объяснить полные потери энергии в компактном объекте, механизм передачи углового момента от аккрецирующего вещества к струйным выбросам и саму коллимацию струйных выбросов. Ответить здесь все эти вопросы не представляется возможным. Поэтому мы обсудим лишь модель, в которой основное место отводится вращающейся черной дыре. Как мы видели, этот случай реализуется как в случае активных ядер, так и в микрокварах и источниках гамма-всплесков. При этом мы сосредоточимся лишь на механизме потерь энергии этих объектов, который связан с сильными электрическими токами, текущими в магнитосфере черной дыры.

Прежде всего, постараемся понять, почему для объяснения релятивистских струйных выбросов необходим столь необычный объект, каким является черная дыра. Как уже отмечалось, в перечисленных выше случаях скорости частиц в струйных выбросах оказываются близки к скорости света. В принципе, частицы могут быть ускорены до больших скоростей за счет газового давления или давления света. Но для этого светимость компактного объекта должна быть достаточно велика, что не наблюдается в действительности. При аккреции же релятивистские скорости могут появиться лишь в том случае, если гравитационный потенциал $\phi_g = -GM/R$ сравнивается с квадратом скорости света c^2 . А это и означает, что радиус компактного объекта R должен быть близок к радиусу горизонта черной дыры

$$r_g = 2GM / c^2 \quad (1)$$

(это выражение может быть получено и из условия, что вторая космическая скорость равна скорости света). Конечно же, для эффективного отвода энергии и углового момента необходимо, чтобы скорости частиц были направлены от черной дыры. Для этого и приходится привлекать сильные электромагнитные поля.

Имеется также ряд косвенных аргументов, говорящих в пользу существования черной дыры у основания струйных выбросов. Как известно, черная дыра по определению не может быть источником плазмы. Поэтому истечение вещества может стать возможным лишь в том случае, если плазма сама будет рождаться в магнитосфере черной дыры. Оказалось, что эффективный механизм рождения плазмы действительно существует. Дело в том, что благодаря сильному разогреву внутренних областей аккреционные диски должны генерировать большое количество жестких гамма-квантов, энергии которых превышают массу покоя электрона. Столкновение таких энергичных фотонов друг с другом и приводит к эффективному рождению электронно-позитронных пар, заполняющих магнитосферу черной дыры. Таким образом, мы приходим к заключению, что струйные выбросы, связанные не с аккреционным диском, а с черной дырой, должны состоять не из обычной, а из электронно-позитронной плазмы. В пользу этой модели говорит, например, тот факт, что в релятивистских выбросах микроквazarов не наблюдаются эмиссионные линии, которые излучает обычное вещество. Ультрарелятивистский характер течения в источниках гамма-всплесков также накладывает ограничение на состав частиц в струйном выбросе, поскольку существование протонов с такой энергией противоречило бы полному энергосыделению гамма-всплеска. Как показывают оценки, вклад протонов должен составлять лишь 10^{-5} от полного числа частиц, так что и здесь речь может идти именно об электронно-позитронных джетах.

Наконец, отметим, что аккрецирующее вещество, обладающее собственным угловым моментом, будет увеличивать угловую скорость вращения черной дыры Ω_H . Поэтому не исключено, что черная дыра сможет раскрутиться до предельных значений $\Omega_H r_g \sim c$. Во всяком случае, как полагают, именно за счет аккреции были раскручены миллисекундные радиопульсары, так что этот процесс действительно может быть очень эффективным. В результате, в обсуждаемой модели источником активности будет не только гравитационная энергия падающего вещества, но и энергия вращения черной дыры. Тогда для случая предельного вращения запасы энергии будут сравнимы с ее энергией покоя: $E_{\text{вр}} \sim Mc^2$. Этой энергии вполне достаточно, чтобы объяснить наблюдаемое энергосыделение.

Более того, была доказана строгая теорема, согласно которой выделение энергии вращения не противоречит тому факту, что ни вещество,

ни излучение не могут покинуть черную дыру. Оказалось, что если достаточно близко от горизонта событий произойдет распад какого-нибудь тела, так что одна его часть упадет на вращающуюся черную дыру, то в этом случае скорость вращения черной дыры уменьшится, а ее энергия будет передана другому телу, которое сможет уйти на бесконечность. Иными словами, при определенных условиях энергия тела, падающего на черную дыру, может быть отрицательной, так что падение тела не увеличит, а уменьшит энергию черной дыры. Этот эффект, открытый Р. Пенроузом, связан с дальнодействием сил гравитации.

3. Униполярный индуктор

Чтобы понять, почему существование магнитного поля может приводить к потерям энергии вращающейся черной дыры, обратимся к известному примеру униполярного индуктора, представляющего собой вращающийся намагниченный шар, к которому двумя концами присоединен проводник (рис. 2). При этом обязательно должны быть выполнены два условия.

1) Проводник не должен вращаться вместе с шаром (точнее, он не должен вращаться с той же угловой скоростью).

2) Два конца проводника должны подсоединяться к шару на разных широтах (наиболее выгодный способ — полюс и экватор).

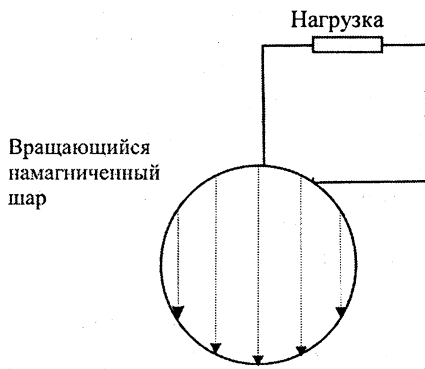


Рис. 2. Униполярный индуктор как модель магнитосферы компактного объекта. Вращающийся намагниченный шар играет роль нейтронной звезды или черной дыры, тогда как невращающийся проводник моделирует магнитные силовые линии.

В результате между концами проводника появляется разность потенциалов U , которая и является источником электрического тока. Ее величину легко оценить, зная, что хорошо проводящая среда, движущаяся со скоростью v в магнитном поле B , создает электрическое поле $E \sim (v/c)B$.

Поэтому по порядку величины разность потенциалов $U \sim ER$ на концах проводника будет равна

$$U \sim (\Omega R/c) RB, \quad (2)$$

где Ω — угловая скорость вращения шара, R — его радиус, а B — магнитное поле на его поверхности. При этом электрический ток I , текущий вдоль провода, будет отводить энергию от батареи (т. е. от вращающегося шара) к внешней нагрузке, где и происходит выделение тепла.

Подчеркнем, что появление электродвижущей силы, благодаря которой вращающийся шар становится батареей постоянного тока, не связано с эффектом Фарадея. Магнитный поток через внешний контур, как видно из рис. 2, не изменяется во времени. Электрический ток становится возможным благодаря появлению сторонних сил, действующих на проводящие электроны при их взаимодействии с твердотельно вращающимся шаром. Стараясь выполнить условие вмороженности $\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c$ в объеме шара, эти силы перемещают электроны против действия электрического поля.

В настоящее время точно установлено, что именно таким образом происходит выделение энергии у радиопульсаров — вращающихся намагниченных нейтронных звезд (рис. 3). При этом механизм замедления вращения связан с поверхностными токами, которые неизбежно должны появиться для замыкания токов, текущих вне звезды. Сила Ампера $\mathbf{F}_A = \mathbf{J} \times \mathbf{B}/c$, действующая со стороны тока на поверхность звезды, и приводит к ее торможению. С другой стороны, от поверхности звезды энергия будет уноситься за счет потока вектора Пойнтинга $\mathbf{S} = (c/4\pi) \mathbf{E} \times \mathbf{B}$. Действительно, вращение звезды приводит к возникновению полоидального электрического поля \mathbf{E} , тогда как продольный ток генерирует тороидальное магнитное поле B_ϕ . Как показано на рис. 3, эти поля и формируют поток электромагнитного поля. Подчеркнем, что эта энергия переносится на нулевой частоте и поэтому уносящее энергию электромагнитное поле не является электромагнитной волной в обычном понимании этого слова.

Что же касается величины электрического тока I , то она определяется из следующих соображений. Плотность электрического заряда ρ_c , необходимая для генерации индукционного электрического поля $E = (v/c)B$, может быть оценена как [4]

$$\rho_c = -\Omega B/2\pi c. \quad (3)$$

Умножая плотность заряда на скорость света c , а также на площадь πR_0^2 , где R_0 — размер области, в котором течет ток I , получаем

$$I \sim \Omega R_0^2 \rho_c c. \quad (4)$$

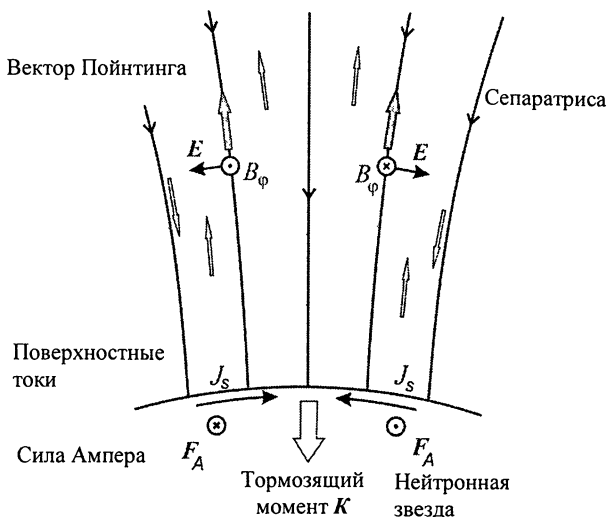


Рис. 3. Токи, текущие в магнитосфере пульсара, замыкаются на его поверхности. Сила Ампера, связанная с поверхностными токами, приводит к замедлению вращения нейтронной звезды. При этом токи текут лишь в ограниченной области магнитосферы, которая не может вращаться вместе со звездой.

Полные же потери энергии W могут быть оценены как $W = IU$. Для радиопулсаров радиус R_0 определяется размером полярной шапки, из которой магнитные силовые линии дипольного магнитного поля уходят за пределы так называемого светового цилиндра $R_L = c/\Omega$. Дело в том, что плазма, заполняющая магнитосферу нейтронной звезды, стремится закоротить продольное электрическое поле. При этом она начинает вращаться твердотельно вместе со звездой. Такая коротация хорошо известна в магнитосфере Земли и больших планет. Ясно, однако, что коротация становится невозможной за пределами светового цилиндра (скорость вращения превышала бы скорость света). Поэтому, как показано на рис. 4, в магнитосфере выделяются две группы магнитных силовых линий. Одна группа, выходящая из пределов полярной шапки, пересекает световой цилиндр и уходит на бесконечность, а другая, находящаяся внутри от магнитной оси, замыкается в пределах светового цилиндра. При этом плазма, находящаяся на замкнутых магнитных силовых линиях, будет вращаться вместе с нейтронной звездой и, следовательно, не сможет участвовать в работе униполярного индуктора. Продольные токи смогут течь лишь в области открытых силовых линий.

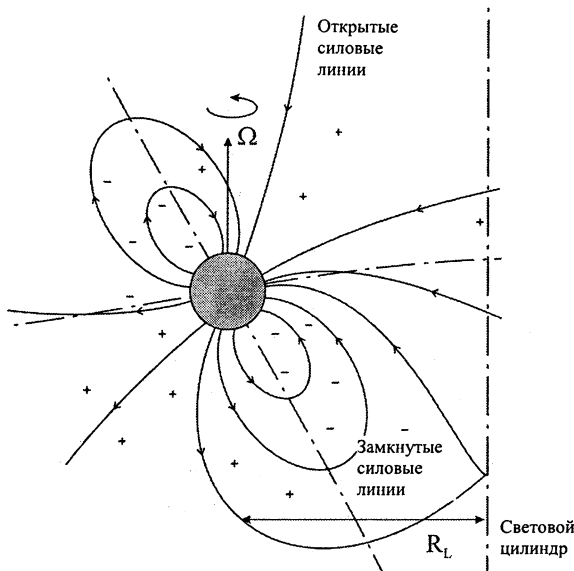


Рис. 4. Открытые и замкнутые силовые линии в магнитосфере радиопульсара

Ясно, что за пределами светового цилиндра плазма не может твердо-вращаться вместе с нейтронной звездой, что как раз и необходимо для работы униполярного индуктора. Воспользовавшись теперь оценкой $R_0 \sim (\Omega R/c)^{1/2} R$, справедливой для дипольного магнитного поля, получаем окончательно хорошо известное выражение для потерь энергии радиопульсаров [4]:

$$W \sim (\Omega R/c)^4 B_0^2 R^2 c. \quad (5)$$

Для стандартных параметров нейтронных звезд (радиус $R \sim 10$ км, $\Omega R/c \sim 10^{-2}—10^{-4}$, магнитное поле $B_0 \sim 10^{12}$ Гс) эта формула полностью соответствует наблюдаемым потерям энергии радиопульсаров $W \sim 10^{31}—10^{38}$ эрг/с.

4. Черная дыра в магнитном поле

Простота и естественность, с которой удалось объяснить активность радиопульсаров, привела к тому, что этот процесс стал рассматриваться как универсальный механизм, применимый и для случая вращающейся черной дыры. Для этого, как мы видели, необходимо существование

- 1) регулярного магнитного поля, играющего роль проводника,
- 2) вращения центрального тела, создающего разность потенциалов U ,

3) электрического тока I , уносящего энергию от вращающегося тела и уменьшающего скорость его вращения.

А основной задачей, которую нужно было решить при определении природы активности компактного объекта, стала задача нахождения структуры электрических токов, текущих в его магнитосфере.

Здесь следует сделать два важных замечания. Во-первых, необходимо объяснить, почему магнитное поле может играть роль проводника (этот вопрос относится и к магнитосфере радиопульсаров). Действительно, в земных условиях внешнее магнитное поле (например, собственное магнитное поле Земли) никак не влияет на направление распространения токов — они будут течь лишь вдоль проводящих тел. Однако если магнитное поле достаточно велико, а плотность среды достаточно мала (так что длина свободного пробега частиц становится меньше радиуса окружности r_L , по которой заряженная частица движется в магнитном поле), и при этом радиус r_L будет много меньше размера системы, то в этом случае проводимость среды поперек магнитного поля должна быть существенно подавлена. В результате с хорошей точностью можно считать, что заряженная частица (а значит, и электрический ток) будет двигаться лишь вдоль внешнего магнитного поля.

Во-вторых, согласно так называемой «теореме об отсутствии волос» черная дыра не может иметь своего собственного магнитного поля. Магнитное поле любого источника (например, контура с током), приближающегося к поверхности черной дыры, будет искажаться ее гравитационным полем, в результате чего удаленный наблюдатель увидит, что магнитное поле в его окрестности будет быстро уменьшаться с течением времени. В итоге, в момент прохождения источником горизонта событий магнитное поле вне черной дыры окажется равным нулю. Единственным собственным полем, которым может обладать черная дыра, будет электрическое поле точечного заряда, расположенного в ее центре.

Чтобы понять причину этого явления, удобнее рассмотреть не магнитный, а электрический диполь, т. е. два близко расположенных разноименных электрических заряда. На рис. 5 показано, как искажается электрическое поле одного заряда q , находящегося вблизи поверхности черной дыры. Несмотря на то что заряд q находится вблизи горизонта, удаленному наблюдателю будет казаться, что его положение совпадает с центром черной дыры. А это значит, что при приближении электрического диполя к горизонту событий изображения зарядов, по мнению удаленного наблюдателя, будут приближаться к центру черной дыры, так что величина дипольного момента $\mathbf{d} = q\mathbf{l}$ за счет уменьшения расстояния l между зарядами будет стремиться к нулю. Следовательно, будет стремиться к нулю и электрическое поле диполя, регистрируемое вдали от черной дыры. Точно также будет уменьшаться и магнитное поле магнитного диполя по мере его приближения к горизонту событий.

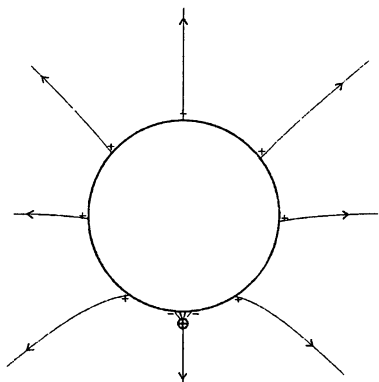


Рис. 5. Силовые линии электрического поля заряда q , находящегося вблизи горизонта черной дыры.

Таким образом, собственное магнитное поле черной дыры должно быть равно нулю. Однако «теорема об отсутствии волос» не запрещает черной дыре быть погруженной во внешнее магнитное поле. Такое магнитное поле вполне может быть создано в аккреционном диске, который способен играть роль динамо-машины. Поэтому возможность существования сильного магнитного поля в рассматриваемых системах не вызывает сомнения.

5. Процесс Блэндфорда — Знайека

Мы видим, что идея униполярного индуктора может быть напрямую применена и к магнитосфере черной дыры. В частности, потери энергии можно оценивать точно так же, как это было сделано для магнитосферы пульсаров. Единственное отличие состоит в том, что радиус полярной шапки R_0 следует заменить на гравитационный радиус r_g , поскольку здесь в отличие от радиопульсаров вращение плазмы не обязательно происходит с той же угловой скоростью, что и угловая скорость черной дыры. Следовательно, роль униполярного индуктора может играть вся поверхность горизонта событий. В результате оценка для потерь энергии по формуле $W = IU$ немедленно приводит к выражению

$$W_{BZ} \sim (\Omega r_g / c)^2 B^2 r_g^2 c, \quad (6)$$

полученному Р. Блэндфордом и Р. Знайеком в 1977 г. при анализе строения магнитосферы вращающейся черной дыры. Для параметров, характерных для активных ядер (магнитное поле $B \sim 10^4$ Гс, радиус черной дыры $r_g \sim 10^{14}$ см, $\Omega r_g / c \sim 1$), эта формула хорошо согласуется с наблюдаемыми потерями энергии $W \sim 10^{45}$ эрг/с. Именно Блэндфорд и Знайек первыми показали, что вращающаяся черная дыра, погруженная во внешнее магнитное поле, может работать как униполярный индуктор, и поэтому этот процесс носит их имя.

Вместе с тем, несмотря на повсеместное использование механизма Блэндфорда — Знайека для объяснения природы активности разных астрофизических объектов, сама возможность электромагнитного выделения энергии вращающейся черной дырой много лет подвергалась сомнению. При этом вопрос оказался столь серьезным, что в некоторых работах был даже сделан вывод о том, что черная дыра вообще не может работать как униполярный индуктор. Дело в том, что в 1982 г. К. Горном и Д. Макдональдом был сформулирован подход, который полностью закреплял аналогию между магнитосферой нейтронной звезды и черной дыры. Для этого было введено понятие проводящей мембраны, расположенной у горизонта черной дыры, свойства которой полностью эквивалентны свойствам поверхности нейтронной звезды [5]. В частности, по поверхности мембраны могли течь электрические токи, замыкающие токи, текущие в магнитосфере. При этом было показано, что тормозящее действие силы Ампера, связанной с поверхностными токами, действительно приводит к выражению (6) для потерь энергии вращающейся черной дыры.

Казалось бы, такое совпадение можно было рассматривать как прямое подтверждение существования процесса Блэндфорда — Знайека. Однако вскоре стало ясно, что многие элементы теории, связанной с поверхностными токами, не выдерживают критики. Например, поверхность черной дыры по определению причинно не связана с внешним пространством, так что информация о поверхностном токе (а значит, и о величине потерь энергии) не может быть передана удаленному наблюдателю. И введение мембраны над поверхностью горизонта ситуацию не проясняло, поскольку в оригинальной работе Блэндфорда и Знайека граничное условие на горизонте существенно использовалось. Кроме того, хорошо известно, что черная дыра не имеет твердой поверхности. Поэтому совершенно непонятно, какой смысл имеют поверхностные токи и на что действует сила Ампера, которая должна тормозить черную дыру.

Лишь в самое последнее время стало ясно, что критика касалась лишь мембранной трактовки, и не затрагивала самих основ процесса Блэндфорда — Знайека. Прежде всего отметим, что в случае черной дыры для замыкания тока совсем не нужно, чтобы заряженные частицы покидали черную дыру (рис. 6). Для этого достаточно, чтобы суммарный заряд, пересекающий горизонт событий, был равен нулю. Более того, для замыкания тока заряды фактически не должны пересекать и магнитные силовые линии, как это имело место на поверхности нейтронной звезды. Как легко понять из рис. 5, удаленному наблюдателю будет казаться, что два разноименных заряда, приблизившихся к горизонту, будут находиться в центре черной дыры, где, с его точки зрения, и произойдет замыкание тока. Этот пример как раз показывает, что поверхностные токи вообще не должны играть никакой роли при торможении черной дыры.

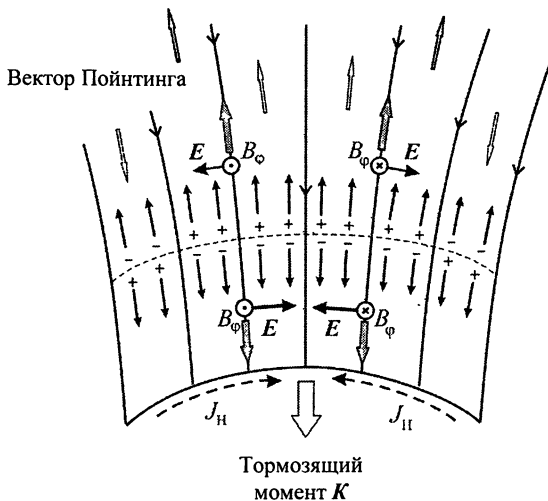


Рис. 6. Структура электрических токов вблизи поверхности черной дыры. Источником тока является область рождения электронно-позитронных пар. Замыкание тока связано с тем, что в разных частях горизонта событий его пересекают частицы разного знака заряда.

Однако тогда в электрическом контуре, содержащем черную дыру, необходимо было найти другой источник электродвижущей силы, поскольку теперь черная дыра уже не может быть источником сторонних сил и, следовательно, играть роль батареи постоянного тока. Оказалось, что в магнитосфере черной дыры появление ЭДС должно быть связано с эффектом Лензе — Тирринга (эффектом увлечения систем отсчета), который возникает за счет вращения черной дыры. Попросту говоря, можно считать, что тело, находящееся в покое относительно вращающейся черной дыры, на самом деле движется относительно выделенной системы отсчета, вращающейся с угловой скоростью Лензе — Тирринга $\omega \sim \Omega_H (r_g/r)^3$. Поэтому, как и в случае любых тел, движущихся в магнитном поле, относительное движение приводит к появлению электрического поля $\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c$, где $\mathbf{v} = \omega \times \mathbf{r}$ — скорость тела относительно выделенной системы отсчета. При этом важно, что угловая скорость вращения будет различной на разных расстояниях от черной дыры. Поэтому, как показано на рис. 7, циркуляция электрического поля в контуре будет отлична от нуля. В результате даже в том случае, если электрический контур покоится относительно черной дыры, так что магнитный поток через него остается постоянным, «движение пространства» приведет к появлению электродвижущей силы в этом контуре.

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B}/c$$

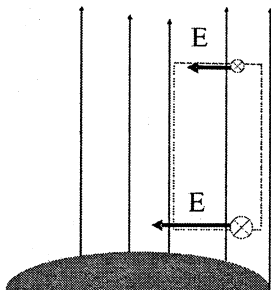


Рис. 7. Появление электродвижущей силы в контуре, покоящемся относительно вращающейся черной дыры.

С другой стороны, мы видели, что для того, чтобы в магнитосфере черной дыры могли течь электрические токи (и, значит, она могла работать как униполярный индуктор), необходимо, чтобы существовал механизм рождения плазмы. Так что и в этом случае одна частица должна упасть на черную дыру, а вторая уйти на бесконечность. Таким образом, процесс Блэндфорда — Знайека является фактически просто электромагнитной реализацией процесса Пенроуза. Только он касается не самих заряженных частиц, а создаваемого ими электромагнитного поля. Иными словами, замедление вращения черной дыры связано не с электрическими токами, текущими по поверхности горизонта, а с отрицательным потоком энергии электромагнитного поля, падающего на поверхность черной дыры. Эта интерпретация представляется сейчас наиболее правдоподобной, так что к ней склоняется большинство исследователей, занимающихся этим кругом вопросов.

Таким образом, можно надеяться, что мы сейчас правильно понимаем механизм выделения энергии релятивистских компактных объектов. С другой стороны, мы видели, как следует быть осторожным при попытке перенести наше понимание электродинамических процессов в плоском пространстве на случай магнитосферы черной дыры. Несмотря на то что вращающаяся черная дыра, погруженная во внешнее магнитное поле, действительно может работать как униполярный индуктор, сам механизм замедления вращения оказывается не связанным с электрическими токами, текущими по поверхности черной дыры.

Литература

1. Черепашук А. М. Сверхмассивные черные дыры в ядрах галактик // СОЖ. 2000. Вып. 11. С. 69—76.
2. Черепашук А. М. Поиски черных дыр // УФН. 2003. Т. 173, вып. 4. С. 345—385.
3. Постнов К. А. Космические гамма-всплески // УФН. 1999. Т. 169, вып. 5. С. 545—558.
4. Бескин В. С. Радиопульсары // УФН. 1999. Т. 169, вып. 11. С. 1169—1198.
5. Торн К. С., Прайс Р., Макдональд Д. Черные дыры. Мембранный подход. М.: Мир, 1988. 630 с.

ФИЗИКА РЕЛЯТИВИСТСКИХ ДЖЕТОВ

*Е. В. Деришев, В. В. Кочаровский,
Вл. В. Кочаровский, В. Ю. Мартъянов*

Основные наблюдательные данные и физические проблемы

В данной статье речь пойдёт о некоторых проблемах физики неравновесной, бесстолкновительной, эффективно излучающей плазмы, собственное движение которой, как и относительное движение составляющих её частиц, является релятивистским, т. е. характеризуется скоростями, близкими к скорости света в вакууме c . Подобная плазма часто наблюдается в виде струйных выбросов, называемых джетами (от англ. jet), которые образуются вблизи компактных объектов, например чёрных дыр, нейтронных звезд или коллапсирующих ядер массивных звезд. О некоторых свойствах джетов и их источников, а также о работе униполярного индуктора вблизи чёрной дыры рассказывается в статье В. С. Бескина в настоящем сборнике. В статье И. Н. Топтыгина рассматривается близкая по характеру задача об ускорении частиц и вызванной ими генерации самосогласованного магнитного поля и турбулентности на нелинейной стадии эволюции нерелятивистской ударной волны.

Наиболее многочисленными из известных источников релятивистских джетов являются активные ядра галактик (см., например, рис. 1), подразделяющиеся на несколько типов [1—4]. Тип определяется степенью активности ядра и ориентацией его структурных элементов (рис. 2) относительно луча зрения. Излучение активного ядра составляет заметную долю излучения остальной части галактики, а в квазарах даже превышает этот уровень. Мощное излучение активных ядер галактик связано с аккрецией, т. е. падением вещества (газа) на их центральную часть. По современным представлениям, там должна находиться сверхмассивная черная дыра с массой, достигающей миллиардов масс Солнца. Падающее на черную дыру вещество разогревается, ионизуется и образует аккреционный диск. В диске вещество (плазма) движется практически по круговым орбитам, медленно приближаясь к черной дыре. Из центральной части ядра галактики бьют две симметричные плазменные струи — джеты, выметающие межгалактический газ на своём пути (рис. 3). Они излучают в широком диапазоне длин волн: от радио- до рентгеновского и даже гамма-диапазона.

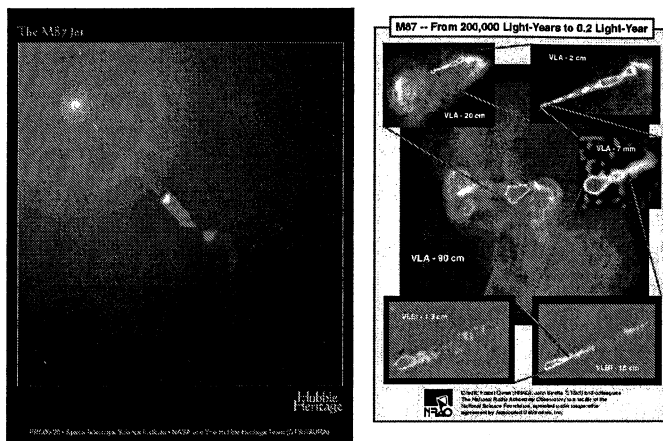


Рис. 1. Джет в галактике M87. Оптическое изображение, полученное космическим телескопом «Хаббл» (слева), и радиоизображения, полученные радиотелескопами VLA и VLBI (справа).

Рис. 2. Основные структурные элементы активного ядра галактики и микроквара: джеты, аккреционный диск и центральная черная дыра.

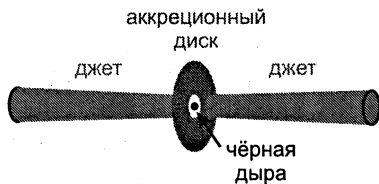


Рис. 3. Столкновение джетов радиогалактики Лебедь А с межгалактической средой. Изображение получено на телескопах Национальной радиоастрономической обсерватории в США (NRAO/AUI/NSF).

Большая часть джетов, видимых в активных ядрах галактик, направлена на наблюдателя (такие объекты называют блазарами). На самом деле это результат наблюдательной селекции: видимая яркость релятивистского джета сильно возрастает для того наблюдателя, чей луч зрения близок к его оси, позволяя регистрировать более далекие — и более многочисленные — объекты. Данный эффект обусловлен релятивистской абберацией света и имеет место для излучения, порождаемого высокоэнергичными частицами, которые движутся вместе с плазмой джета и в её системе отсчета распределены по скоростям изотропно. Соответственно излучение джета тоже изотропно в движущейся системе отсчета. При переходе в неподвижную систему отсчета уменьшается угол между направлением движения фотонов и скоростью источника. Диаграмма направленности излучения приобретает ширину порядка обратного лоренц-фактора джета, т. е. порядка нескольких градусов для типичных активных ядер, а наблюдаемая яркость вне основного лепестка диаграммы направленности спадает обратно пропорционально шестой степени угла. Однако в рентгеновском и гамма-диапазоне излучение джетов можно видеть и под большим углом к их оси (см. ниже).

В 1992 году в нашей галактике был обнаружен объект GRS1915+105, удивительно напоминающий квазар и имеющий несимметричные джеты, яркие пятна в которых двигаются со «сверхсветовой» скоростью [5]. (Напомним, что для релятивистского джета с большим лоренц-фактором $\Gamma = (1 - v^2/c^2)^{-1/2}$, движущегося со скоростью v под углом θ к наблюдателю, кажущаяся скорость смещения излучающего сгустка в картинной плоскости равна $v_{\text{app}} = v \sin \theta / (c - v \cos \theta)$ и может превышать скорость света примерно в Γ раз при $\theta \sim 1/\Gamma$.) Объект оказался аккрецирующей черной дырой в составе двойной звездной системы. Сейчас известно более десяти подобных систем (их относят к классу так называемых рентгеновских двойных, причем роль аккрецирующей компоненты может играть и нейтронная звезда) [6]. По своему строению они похожи на активные ядра галактик, однако масса черной дыры в них составляет лишь несколько масс Солнца, и источником аккрецируемого вещества служит не межзвездный газ, а звезда-компаньон. Ввиду своего сходства с квазарами (при огромной разнице в размерах) такие объекты называли микроквазарами. Для их джетов также характерно мощное электромагнитное излучение в диапазоне от радиоволн до рентгеновских лучей.

Совершенно иной пример релятивистских джетов представляют источники гамма-всплесков [7]. В этом случае джеты никогда не наблюдались, а их существование следует из косвенных, хотя и весьма

убедительных аргументов. Яркие вспышки гамма-излучения космической природы, открытые еще в 1968 году, долгое время не удавалось связать с излучением в каком-либо ином диапазоне электромагнитного спектра, чем и обусловлено их название. Прорыв произошел в 1997 году, когда впервые удалось наблюдать оптическое послесвечение гамма-всплесков, что позволило определить красное смещение линий в их спектрах и, соответственно, расстояние до источников. Оказалось, что гамма-всплески происходят на расстояниях, сравнимых с размерами видимой части Вселенной. Вытекающая отсюда оценка их светимости порядка 10^{50} – 10^{52} эрг/с, на первый взгляд, не согласуется с малыми размерами источников, диктуемыми миллисекундным временем переменности: при этом создается настолько высокая плотность фотонов, что они взаимодействуют друг с другом, создавая непрозрачное облако электрон-позитронных пар. Противоречия можно избежать, если предположить, что порождающий излучение поток плазмы движется на наблюдателя с релятивистской скоростью и является узконаправленным. В противном случае за время типичного всплеска продолжительностью несколько десятков секунд должна была бы выделяться энергия, сравнимая с энергией покоя Солнца (и это только в виде гамма-излучения). Природа источников гамма-всплесков достоверно не известна. Наиболее популярная модель связывает образование релятивистских джетов с быстрой аккрецией звездного вещества на черную дыру, возникающую при коллапсе ядра сверхновой определённого класса.

Круг физических проблем, возникающих при описании релятивистских джетов, чрезвычайно обширен. Сюда относятся задачи динамики разлёта многокомпонентной плазмы, эволюции релятивистских ударных волн с потерями, эффективного ускорения частиц, объяснения многообразия наблюдаемых спектров их излучения, формирования магнитного поля, развития волновой турбулентности в плазме и т. д. Ниже мы кратко коснёмся лишь небольшой части этих проблем.

Влияние нейтронов на динамику разлёта плазмы

Чтобы дать представление о богатстве динамики джетов, расскажем о релятивистских джетах с нейтронной компонентой, присутствие которой, по-видимому, типично для плазмы в источниках космических гамма-всплесков [8, 9]. Хотя время жизни свободного нейтрона составляет всего около 900 секунд, оно превосходит динамическое время гамма-всплесков, так что нейтроны в этих условиях могут считаться стабильными частицами. Светимость гамма-всплесков так велика, что

температура в основании их джетов достаточна для диссоциации ядер гелия или более тяжелых элементов на протоны и нейтроны. В основании джета нейтроны сталкиваются с протонами достаточно часто, чтобы оказаться увлеченными общим потоком, но при ускорении и уменьшении плотности джета (с удалением от центрального источника) частота нейтрон-протонных столкновений падает гораздо быстрее, чем увеличивается временной масштаб ускорения. Начиная с некоторого расстояния нейтроны разлетаются практически свободно. Если разгон джета к этому моменту еще не завершился, то плазменная (электрон-протонная) компонента приобретает большой лоренц-фактор (Γ_p), чем нейтронная (Γ_n), т. е. происходит их расцепление.

С увеличением относительного лоренц-фактора протонов и нейтронов кинетическая энергия их взаимного движения превосходит порог рождения пионов, после чего столкновения протонов с нейтронами становятся неупругими. В таких столкновениях с близкой вероятностью рождаются как нейтральные, так и заряженные пионы¹. Распад заряженных пионов приводит к появлению трех нейтрино и позитрона (электрона), причем каждая частица уносит примерно одинаковую энергию — приблизительно 35 МэВ. Нейтральные пионы распадаются на два гамма-кванта с энергией около 70 МэВ. Таким образом, в результате расцепления возникает нейтринное излучение, идущее в основном из области вблизи радиуса расцепления, и жесткое гамма-излучение, появляющееся в результате электромагнитного каскада и выходящее с более далекой фотосферы. В сравнении с нейтринным излучением электромагнитное слабее, но из-за относительной простоты регистрации наблюдать электромагнитный сигнал намного легче. Однако динамическое воздействие нейтринного излучения на джет за счет импульса отдачи столь велико, что переданный излучению импульс может даже превосходить конечный импульс самого джета, что приводит к его значительному торможению.

Оценим нейтринные потери в случае больших факторов расцепления $D \equiv \Gamma_p/\Gamma_n \gg 1$. Когда $\Gamma \gg \Gamma_n$ (Γ — текущий лоренц-фактор плазменного потока), все три разновидности пионов рождаются с одинаковой вероятностью и их суммарная энергия в системе центра масс сталкивающихся нуклонов приблизительно равна $mc^2 \sqrt{\Gamma/\Gamma_n}$. Нейтринные потери возрастают, если пионы до своего распада успевают частично или полностью изменить свой начальный импульс, например из-за вза-

¹В числе прочих неупругих реакций идут процессы с образованием дейтерия. Получившиеся дейтроны имеют неплохие шансы избежать новых столкновений, а следовательно, и диссоциации. В итоге релятивистский джет становится эффективным источником дейтерия — в него превращается до 20% всего выбрасываемого нуклонного вещества.

имодействия с магнитным полем. Время распада пиона ($\approx 2,6 \cdot 10^{-8}$ с), становится больше обратной гирочастоты, когда магнитное поле сильнее, чем 600 Гс, причем плотность энергии такого поля на много порядков меньше плотности тепловой энергии плазмы в области расщепления. Поскольку требования к однородности магнитного поля при этом невысоки, будем считать, что эффект присутствует всегда.

С учетом влияния магнитного поля нейтрино уносят энергию $\epsilon_\nu \approx mc^2\Gamma^2/(4\Gamma_n)$ (в лабораторной системе отсчета). Вообще говоря, это больше, чем суммарная энергия сталкивающихся протона и нейтрона. Разумеется, никакого нарушения законов сохранения здесь нет — лишняя энергия черпается из кинетической энергии плазменного потока посредством вмороженного в него магнитного поля.

В расчете на один протон суммарная энергия излученных нейтрино составляет

$$\delta E_\nu \approx 1,5 \eta \Gamma_p mc^2. \quad (1)$$

В силу использованных приближений выражение (1) применимо только при условии $\Gamma_p/\Gamma_n \gg 1$. Результат, отнесенный к конечной энергии протона $\Gamma_p mc^2$, показывает, насколько велики нейтринные потери по сравнению с энергией, затраченной на разгон релятивистского ветра. Отметим, что когда относительная концентрация нейтронов $\eta \approx 1$, то более половины мощности источника теряется на излучение нейтрино. Типичная энергия нейтрино составляет порядка 30 ГэВ.

Конверсионный механизм ускорения и внеосевое излучение

Излучение джетов обязано своим происхождением высокоэнергичным релятивистским электронам (и позитронам), чья энергия на многие порядки превышает тепловую. Такие электроны часто называют ускоренными, имея в виду, что их существование можно объяснить только действием существенно неравновесных механизмов передачи энергии малой доле частиц плазмы. Наиболее универсальным поставщиком высокоэнергичных частиц является механизм Э. Ферми, а источником ускорения — крупномасштабные перепады гидродинамической скорости плазмы. В частности, давно известно ускорение в сдвиговых течениях и бесстолкновительных ударных волнах [1, 3, 10], которые с необходимостью присутствуют в релятивистских джетах. Сдвиговые течения образуют кокон вокруг джета там, где он граничит с прилегающим межзвездным веществом. Ударные волны возникают в области торможения джета или даже внутри него из-за быстрых изменений скорости потока в основании джета.

Механизм ускорения удобнее пояснить на примере ударной волны. В бесстолкновительной плазме частицы взаимодействуют друг с другом посредством магнитного поля, создаваемого текущими в этой плазме токами. Магнитное поле в заморожено в плазму по обе стороны от фронта ударной волны, где гидродинамическая скорость плазмы испытывает скачок; таким образом, для частицы с достаточно большой длиной свободного пробега (что означает — с достаточно большим импульсом) прохождение через фронт подобно столкновению с движущейся стенкой. В зависимости от направления своего движения частицы могут как приобретать, так и терять энергию, однако в среднем энергия возрастает. Такое ускорение аналогично нагреву газа в сосуде со сближающимися стенками с той лишь разницей, что отсутствует вторая стенка, которая отражала бы частицы обратно к фронту ударной волны. Продолжительное ускорение и значительный набор энергии возможны только для тех частиц, которые могут многократно пересекать фронт ударной волны в обратном направлении, обгоняя его за счет диффузионного движения в неоднородном магнитном поле.

Ускорение продолжается до тех пор, пока потери энергии на излучение не сравняются с темпом ускорения или темп диффузии высокоэнергичных частиц не станет слишком большим для удержания их в области фронта ударной волны. Первая причина типична для электронов, вторая — для протонов, которые тяжелее и излучают слабее. В начале процесса ускорение тепловых электронов затруднено, поскольку их длина свободного пробега значительно меньше, чем у более тяжелых ионов, а именно последние определяют ширину фронта ударной волны. По-видимому, это обстоятельство и объясняет малую долю электронов в космических лучах даже самых низких энергий. Среднестатистический темп ускорения частиц в нерелятивистской ударной волне увеличивается пропорционально квадрату ее скорости. Чем выше скорость — тем больше максимальная энергия ускоренных частиц, тем выше частота и больше мощность их излучения.

Однако в релятивистской ударной волне, скорость которой близка к скорости света, рассмотренный механизм ускорения оказывается неэффективным. Причин этому две. Во-первых, гидродинамическая скорость течения за фронтом релятивистской ударной волны равна одной трети скорости света, так что регулярный снос частиц преобладает над их диффузионным смещением и поэтому они практически не имеют шансов вернуться к фронту и продолжить ускорение. Во-вторых, при сжатии плазмы в ударной волне увеличивается параллельная фронту компонента замороженного магнитного поля, тогда как перпендикулярная компонента остается неизменной. В результате

дрейфовая скорость частиц, связанная с неоднородностью магнитного поля, направлена приблизительно параллельно фронту, а частицы движутся по циклоидам, не пересекающим его.

Тем не менее эффективное ускорение частиц в релятивистских ударных волнах и сдвиговых течениях все же возможно [11]. Парадоксально, но главную роль здесь играет взаимодействие ускоряемых частиц с фотонными полями в джетях, которое обычно рассматривается лишь как помеха, дополнительный канал потери энергии. Дело в том, что при определенных условиях столкновения с фотонами можно трактовать как механизм случайного «выключения» и «включения» электрического заряда частиц (рис. 4). Проходя через фронт ударной волны и отражаясь от неоднородностей магнитного поля, заряженная частица увеличивает свою энергию, а затем, после встречи с фотоном, становится нейтральной, что позволяет ей вернуться к фронту без препятствий со стороны магнитного поля. Оказавшись перед ударной волной, частица вновь становится заряженной, так что весь цикл ускорения повторяется снова и снова. Такой механизм ускорения частиц называется конверсионным; вычисления показывают, что он играет важнейшую роль в излучении и динамике релятивистских джетов [11, 12].

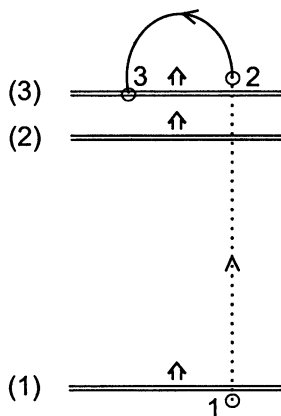


Рис. 4. В конверсионном механизме ускорения частица покидает ударную волну, будучи нейтральной (момент времени 1), и движется прямолинейно до своего превращения в заряженную (момент 2). За это время частица обгоняет фронт ударной волны на некоторое расстояние, так что у нее появляется достаточный запас времени, чтобы развернуться в магнитном поле прежде, чем она будет вновь подхвачена ударной волной (момент 3). Положения частицы и фронта ударной волны в соответствующие моменты времени отмечены цифрами 1, 2 и 3.

Разумеется, выключение и включение электрического заряда — это лишь условный способ описания происходящих процессов. Проследим, например, за цепочкой взаимодействий, начинающейся рассеянием фотона на релятивистском электроне. Если энергия фотона в системе отсчета электрона много меньше его энергии покоя, то рассеяние происходит в классическом, или томсоновском, режиме. В противоположном случае говорят о рассеянии в квантовом, или клейн-нишиновском, режиме. В результате обратного комптоновского эффекта (комptonизации) энергия фотона после рассеяния многократно возрастает, но при томсоновском рассеянии составляет лишь малую долю энергии электрона. Такое взаимодействие понемногу отбирает энергию у электронов и их ускорению не способствует. Однако в широкополосном спектре излучения джетов есть много высокочастотных фотонов. Их рассеяние происходит в клейн-нишиновском режиме, а спектр комптонизированного излучения имеет два пика: с равной вероятностью электрон либо сохраняет основную часть своей энергии, либо передаёт ее фотону. В последнем случае можно говорить о смене заряда ускоряемой частицы, поскольку получившийся фотон движется в том же направлении, что и исходный электрон, и несет почти всю его энергию. Высокоэнергичные комптонизированные фотоны взаимодействуют с относительно низкочастотным излучением джета, образуя электрон-позитронные пары, причём энергия распределяется между электроном и позитроном примерно поровну. Результатом описанной цепочки взаимодействий является замкнутый цикл выключения и включения заряда ускоряемой частицы. То обстоятельство, что количество ускоряемых частиц при этом удваивается, не играет принципиальной роли.

Аналогичная цепочка превращений существует и для протонов. Фотоны, энергия которых превышает 350 МэВ в системе покоя протона, могут возбуждать его внутренние (кварковые) степени свободы. Образовавшийся возбужденный адрон моментально распадается на нуклон и пион. Приблизительно в трети случаев образуется заряженный пион, а протон превращается в нейтрон. Обладая аналогичной кварковой структурой, нейтроны взаимодействуют с фотонами по той же схеме, т. е. переходят в протоны. Побочным продуктом протон-нейтронного цикла становятся высокоэнергичные заряженные пионы, распад которых приводит к появлению вторичных электронов и позитронов, а также нейтринного излучения.

Вероятность того, что конверсия частиц из заряженных в нейтральные и обратно происходит при взаимодействии с фотонами подходящей энергии и в правильной последовательности, много меньше

единицы. С другой стороны, частица, прошедшая полный цикл ускорения (см. рис. 4), увеличивает свою энергию приблизительно в $g \sim \Gamma^2$ раз, что составляет величину от нескольких сотен для джетов в активных ядрах галактик до сотен тысяч для гамма-всплесков. Когда произведение вероятности прохождения цикла k на фактор увеличения энергии g превосходит единицу, конверсионный механизм достигает наибольшей эффективности: суммарное энергосодержание в ускоренных частицах возрастает с каждым циклом, причем основной вклад приходится на те частицы, чья энергия близка к максимально достижимой. Общее представление о распределении ускоренных частиц по энергиям дает его огибающая, которую можно приближенно представить степенной функцией $\frac{dN}{d\varepsilon} \propto \varepsilon^{-\alpha}$ с показателем степени

$$\alpha = 1 - \frac{\ln k}{\ln g}, \quad (2)$$

который слабо зависит от энергии частиц ε .

По оценкам, релятивистские джеты в гамма-всплесках и активных ядрах галактик в таком режиме ускорения способны обеспечить существование наблюдаемых космических лучей (протонов) вплоть до сверхвысоких энергий 10^{20} эВ. Интересно отметить, что при конверсионном ускорении частицы достигают такой энергии всего за 2—4 прохода через фронт ударной волны. При обычном диффузионном ускорении для этого потребовалось бы много десятков проходов; соответственно умножаются и трудности реализации такого механизма. Пояснить различие можно следующим образом.

При ускорении без конверсии пересечь фронт ударной волны в обратном направлении могут только те частицы, которые не успевают потерять свою энергию (например, на синхротронное излучение) за время, требуемое для совершения половины оборота в магнитном поле за фронтом. Назовем критическими те частицы, которые теряют на этом развороте половину своей энергии. Как следует из преобразований Лоренца, скорость частиц, покидающих релятивистскую ударную волну, в неподвижной системе отсчета направлена почти параллельно скорости ударной волны. Однако траектория заряженных частиц немедленно начинает искривляться под действием магнитного поля, и, как только угол между скоростью частицы и направлением движения ударной волны становится порядка ее обратного лоренц-фактора, фронт волны догоняет убежавшую частицу и завершает цикл ускорения. Результатом цикла становится увеличение энергии частицы примерно вдвое, так что критическая энергия одновременно является и

максимально достижимой, при которой темп потерь энергии сравнивается с темпом ускорения.

При конверсионном ускорении частица покидает ударную волну, будучи нейтральной, и движется прямолинейно до своего превращения в заряженную. За это время частица обгоняет фронт ударной волны на некоторое расстояние, так что у нее появляется достаточный запас времени, чтобы развернуться в магнитном поле прежде, чем она будет вновь подхвачена ударной волной. В итоге конверсионный цикл ускорения увеличивает энергию частицы не вдвое, а в квадрат лоренц-фактора ударной волны раз.

Если энергия покидающих ударную волну частиц ограничена критическим значением, то энергия частиц, попадающих в нее из внешней среды, может быть многократно больше. Сверхкритические частицы теряют свою энергию, проходя путь много меньше их гирорадиуса, т. е. мало отклоняясь от первоначального направления своего движения (почти строго назад в системе отсчета, связанной с веществом за фронтом ударной волны). Соответственно, диаграмма направленности излучения сверхкритических частиц сильно вытянута в направлении против скорости ударной волны, т. е. сильно анизотропна, как и распределение частиц по скоростям в системе отсчета джета [13]. Из-за релятивистской абберации света неподвижный наблюдатель все-таки увидит диаграмму направленности вытянутой в направлении движения джета, но она остается значительно шире, чем диаграмма направленности менее энергичного излучения докритических частиц. Излучение сверхкритических частиц, сосредоточенное в основном в рентгеновском и гамма-диапазонах, можно наблюдать под большим углом к оси джета, т. е. оно является внеосевым (рис. 5).

Внеосевое излучение объясняет, например, феномен запаздывающего жесткого излучения гамма-всплесков. Задержка в данном случае имеет геометрическую природу: обладая уширенной диаграммой направленности, внеосевое излучение наблюдается от более широкого сегмента сферической ударной волны, края которого находятся дальше от наблюдателя, чем центральная часть. Уширение диаграммы направленности может также объяснить регистрацию так называемых неидентифицированных источников жесткого гамма-излучения (большое количество таких источников обнаружено космическим телескопом EGRET в 1992 — 1994 гг.). Они, вероятно, связаны с теми квазарами, джеты которых ориентированы под большими углами к направлению на Землю и поэтому не создают заметного излучения в более низкочастотных диапазонах, где излучение джета является узконаправленным. Широконаправленным может быть и высокоэнергичное ней-

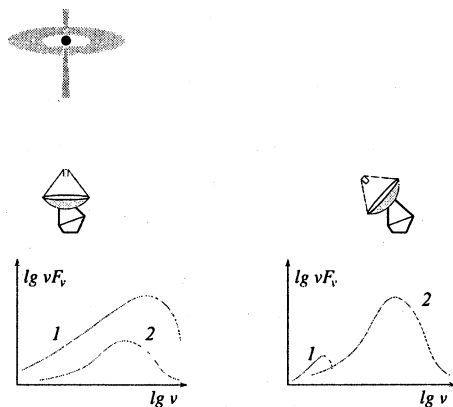


Рис. 5. Относительные изменения спектра излучения умеренно ускоренных (кривые 1) и предельно ускоренных (кривые 2) частиц при переходе от осевого (слева) к внеосевому (справа) наблюдению релятивистского джета.

тринное излучение релятивистских джетов, сопровождающее распад пионов, которые образуются при неупругих столкновениях протонов и нейтронов.

Синхротронно-комптоновское излучение

Высокоэнергичные электроны (и позитроны) могут излучать посредством множества самых разнообразных механизмов, однако практический интерес представляют только те из них, в которых условия реализации являются типичными для релятивистских джетов, а ускоренные частицы излучают значительную долю своей энергии за время, не превышающее время переменности источника. Последнее условие связано с высокой светимостью активных ядер галактик, и особенно источников гамма-всплесков. Неэффективный механизм излучения приводит к завышенным и нереализуемым требованиям к мощности источника ускоренных частиц. Указанным условиям удовлетворяют синхротронное излучение, которое связано с повсеместно присутствующими магнитными полями и обладает быстро растущей с увеличением энергии частиц эффективностью, и так называемое комптоновское излучение. Последнее не является излучением в строгом смысле этого слова: передача энергии от ускоренных электронов к электромагнитному полю происходит за счет обратного комптон-эффекта при рассеянии низкоэнергичных (затравочных) фотонов на высокоэнергичных

электронах. Число фотонов при этом не изменяется, а источником затравочных фотонов может служить как излучение аккреционного диска, так и собственное синхротронное излучение ускоренных частиц джета.

Как правило, мощность синхротронного излучения релятивистских джетов настолько велика, что плотность его энергии сравнима с плотностью энергии магнитного поля. При этом комптоновский механизм вносит большой, а во многих случаях определяющий вклад в полную мощность излучения, и само излучение правильнее называть синхротронно-комптоновским. Совместное действие обоих механизмов излучения приводит к появлению в спектрах активных ядер галактик двух далеко отстоящих по частоте компонент: низкочастотной, синхротронной (доли электрон-вольт — единицы мегаэлектрон-вольт) и высокочастотной, комптоновской (единицы гигаэлектрон-вольт — десятки тераэлектрон-вольт). Можно ожидать, что комптоновская компонента присутствует и в спектрах гамма-всплесков (в диапазоне от единиц до сотен тераэлектрон-вольт), однако наблюдательных доказательств этому нет. Увидеть ее мешает сильное поглощение высокоэнергичных фотонов при взаимодействии с межгалактическим фоном оптического излучения: горизонт видимости для фотонов с энергией выше 100 ГэВ расположен ближе подавляющего большинства источников гамма-всплесков.

Вообще говоря, синхротронную и комптоновскую компоненты нельзя рассматривать независимо [14]. Между ними существует тесная нелинейная связь: обе они суть излучение одних и тех же высокоэнергичных электронов и позитронов, функция распределения которых испытывает обратное влияние со стороны излучения обоих видов. Наибольший интерес представляет случай эффективного излучения, когда за характерное время эволюции джета частицы успевают излучить значительную часть своей энергии. Тогда формируется квазистационарное распределение частиц, вид которого определяется как зависимостью темпа радиационных потерь от их энергии, так и свойствами источника частиц (его роль играет, например, ударная волна, поставляющая ускоренные электроны внутрь джета).

Темп синхротронных потерь пропорционален квадрату энергии излучающих частиц, тогда как комптоновские потери зависят еще и от спектра затравочного излучения, что объясняется эффектом Клейна — Нишины. Он состоит в том, что для фотонов, энергия которых в системе отсчета электрона превышает его энергию покоя, сечение рассеяния быстро уменьшается с ростом частоты, так что их вклад в темп комптоновских потерь мал. Поскольку пороговая частота обрат-

но пропорциональна энергии электронов, эффект Клейна — Нишины максимален для высокоэнергичных электронов и практически пропадает для низкоэнергичных. Соответственно, для электронов с энергией ниже порога Клейна — Нишины преобладают комптоновские потери; для остальных главным каналом потерь является синхротронное излучение. Величина указанного порога зависит от нескольких факторов, в том числе и от мощности источника частиц. Таким образом, даже простое изменение количества ускоряемых электронов и позитронов при прочих равных условиях может приводить к изменению формы спектра и положения максимумов как для синхротронного, так и для комптоновского излучения.

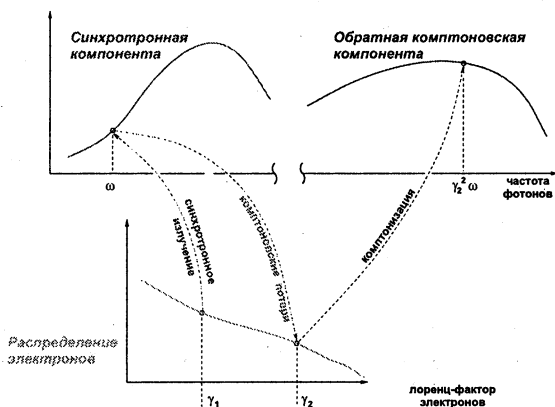


Рис. 6. Схема синхротронно-комптоновского механизма излучения. По известному энергетическому распределению электронов (нижний график) можно однозначно восстановить спектр их синхротронного излучения (левый график), причем на каждой частоте основной вклад приходится на группу электронов (γ_1), у которых максимум синхротронного спектра близок к этой частоте. В то же время фотоны частоты ω соответствуют порогу Клейна — Нишины для другой группы электронов (γ_2) и, если он расположен на растущей части спектра, определяют их комптоновские потери. Последние, в свою очередь, диктуют спектр комптоновского излучения в области частот порядка $\gamma_2^2\omega$ и влияют на вид функции распределения электронов.

Как показано на рис. 6, взаимосвязь синхротронного и комптоновского механизмов излучения ведет к изменению спектров обеих компонент, причем она имеет нелокальный по частоте, интегральный характер. В частности, комптоновские потери у данной группы электронов, а значит, и вид функции распределения в этой области, определяются

фотонами с частотой вблизи порога Клейна — Нишины (если он расположен на растущей части спектра), то есть синхротронным излучением другой группы электронов. В итоге даже простые (например, с единственной особенностью) функции распределения электронов у источника ускоренных частиц порождают сложные спектры синхротронного и комптоновского излучения с несколькими изломами. Это позволяет объяснить основные спектральные особенности излучения релятивистских джетов.

Несмотря на кажущуюся сложность, задача об определении спектра синхротронно-комптоновского излучения при заданной функции инжекции электронов сводится к решению единственного уравнения для синхротронной эффективности $\eta(\gamma)$, которая есть отношение мощности синхротронных потерь для электрона с лоренц-фактором γ к суммарной мощности его излучения. Если считать синхротронный спектр отдельного электрона монохроматическим, а сечение рассеяния фотонов в режиме Клейна — Нишины — нулевым (что дает достаточную для практических целей точность), то указанное уравнение принимает совсем простой вид:

$$\frac{1}{\eta(x)} = 1 + \mathcal{K} \int_0^{1/\sqrt{x}} p(x') \eta(x') dx'. \quad (3)$$

Здесь x — лоренц-фактор электронов, нормированный на величину γ_0 , для которой положение порога Клейна — Нишины совпадает с частотой их собственного синхротронного излучения; $p(x)$ — вероятность того, что электрон был инжектирован с лоренц-фактором больше x , а \mathcal{K} — комптоновский потенциал, который содержит в себе информацию о мощности источника ускоренных частиц, средней по спектру синхротронной эффективности и геометрии излучающей области.

Происхождение и структура магнитного поля

Проблема генерации и пространственно-временной эволюции магнитного поля важна не только для выяснения структуры и устойчивости релятивистских джетов. В бесстолкновительной плазме именно магнитное поле определяет характер движения и ускорения заряженных частиц. Последнее, как уже отмечалось, происходит в областях сильной неоднородности гидродинамической скорости плазмы, прежде всего в окрестности ударных волн и сдвиговых течений. В этих областях функция распределения частиц по импульсам обладает значительной анизотропией, в том числе благодаря эффективному ускорению частиц.

В таких условиях плазма подвержена так называемой вейбелевской неустойчивости [15, 16] (рис. 7), обуславливающей появление сильного магнитного поля. Согласно теореме о его вмороженности, при сжатии плазмы вблизи фронта ударной волны вектор магнитного поля прижимается к плоскости фронта волны и, по-видимому, ориентирован там случайным образом. Характерный размер неоднородности поля в плоскости фронта при этом оказывается порядка величины c/ω_p и много меньше пространственного масштаба ударной волны, а характерный инкремент может быть порядка ω_p . Здесь ω_p — плазменная частота.

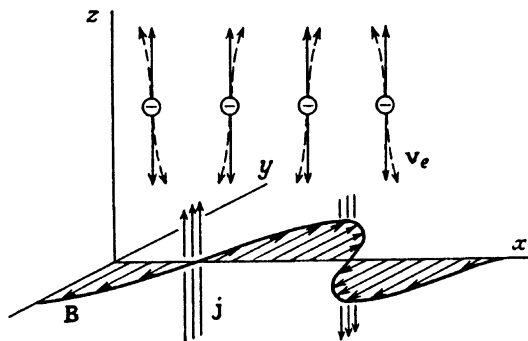


Рис. 7. Схема физического механизма вейбелевской неустойчивости: отклоняясь в затравочном магнитном поле, потоки электронов группируются в токовые слои, тем самым усиливая магнитное возмущение.

Образование магнитного поля в результате вейбелевской неустойчивости было численно проанализировано в нескольких работах как для начальной, так и для граничной постановки задачи; см., например, [17 — 19]. Трехмерное релятивистское численное моделирование столкновения двух плазменных оболочек методом частиц в ячейках [19] показало, что возможно образование близких к равномерным квазистатическим долгоживущим магнитным полям. После быстрой линейной стадии развития неустойчивости наблюдается медленное затухание магнитного поля на масштабах гирорадиуса электронов, много большего величины c/ω_p , и эволюция его структуры от малых масштабов к большим. Поэтому магнитное поле получается сильно неоднородным и представляет собой набор филаментов, ориентированных в плоскости фронта (плоскость $x - y$ на рис. 7). Имеется и аналитическое описание подобных токовых филаментов (рис. 8 и ниже).

Для ряда модельных функций распределения частиц возможно аналитическое исследование релятивистской вейбелевской неустойчивости [16, 20, 21] — нахождение дисперсионных соотношений и инкремента. Удастся также установить общие условия существования неустойчивости для плазмы с произвольной степенью анизотропии, найти уровень магнитного поля, насыщающего неустойчивость, провести оценки скорости слияния токовых филаментов и затухания магнитного поля. Оказывается, что экспоненциальный рост магнитного поля прекращается либо когда траектории движения частиц плазмы отклонятся в возникшем поле от первоначального направления на угол порядка единицы, либо когда смещение частиц вдоль направления возмущения станет порядка длины волны, либо когда возникающее электрическое поле отнимет у частиц значительную часть их энергии. Магнитное поле при этом может дорасти до равномерного, т. е. такого, для которого средняя плотность энергии порядка плотности энергии частиц плазмы.

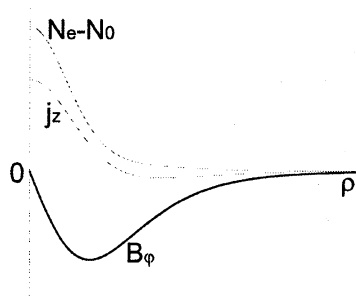


Рис. 8. Типичные профили магнитного поля, тока и возмущения концентрации плазмы в локализованном цилиндрическом филаменте.

Для описания квазистационарных самосогласованных конфигураций магнитного поля и тока в бесстолкновительной плазме часто используется так называемый нейтральный токовый слой Харриса с максвелловской функцией распределения электронов, сдвинутой относительно функции распределения ионов [22, 23]. Динамика таких токовых слоев, сопровождающаяся, в частности, перезамыканием силовых линий магнитного поля и генерацией ускоренных частиц и волновой турбулентности, обусловлена не только быстрыми диссипативными процессами вблизи особых точек, но и медленными процессами филаментации и структурной перестройки — и спонтанной, и вызванной внешними факторами. Однако до сих пор класс аналитически рассматриваемых (квази)стационарных токовых конфигураций был невелик и включал лишь простейшие обобщения слоя Харриса, скажем, с цилиндрической симметрией или с релятивистской модификацией максвел-

ловского распределения. Магнитогидродинамическое описание цилиндрически симметричных стационарных течений, в которых магнитное поле играет существенную роль, с использованием уравнения Грэда — Шафранова приведено в книге [24].

Вместе с тем в одномерном случае давно известен целый класс нелинейных стационарных продольных БГК-волн (БГК — Бернштейн, Грин, Крускал [25]), в которых по существу произвольный профиль электростатического потенциала согласован с сильно неравновесной функцией распределения частиц [25]. Используемый для их нахождения подход, основанный на инвариантах движения частиц, можно модифицировать и применить для двумерного движения частиц при наличии самосогласованного магнитного поля. В итоге аналитически находится новый обширный класс стационарных плоскостных и цилиндрически симметричных токовых структур, которые характеризуются существенно неоднородной и немаксвелловской функцией распределения частиц и могут одинаково хорошо реализоваться и в нерелятивистской, и в релятивистской бесстолкновительной многокомпонентной плазме.

В частности, для полиномиальной зависимости функции распределения от одной из проекций обобщенного импульса частиц (в том числе, для случая отрицательных степеней полинома) удаётся в конечной форме записать источники в уравнении для векторного потенциала и явно выразить их через этот потенциал. Выведенное нелинейное обыкновенное дифференциальное уравнение второго порядка аналогично уравнению осциллятора с произвольным профилем потенциальной энергии, задаваемым определёнными моментами функции распределения, и в общем виде решается в квадратурах, а во многих случаях — до конца [26].

Континуум найденных таким образом конфигураций включает, например, гармоническое магнитное поле с энергией порядка энергии всех частиц, сосредоточенный токовый слой (аналог слоя Харриса), пару противоположно направленных токовых слоев с локализованным между ними магнитным полем, токовый слой с противотоками, экспоненциально экранирующими магнитное поле на бесконечности, и др. Электрическое поле в них может быть скомпенсировано подходящим распределением разноимённо заряженных компонент плазмы, например, электронами и позитронами, движущимися по одинаковым встречным траекториям в электрон-позитронной плазме. Размеры токовых филаментов могут быть как меньше, так и значительно больше минимального гирорадиуса частиц плазмы.

Для цилиндрически симметричного случая соответствующее уравнение на компоненту векторного потенциал A как функцию цилиндрического радиуса ρ имеет вид уравнения нелинейного маятника с затуханием

$$\frac{\partial^2 A}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial A}{\partial \rho} = F(A), \quad (4)$$

где функция F определяется конкретным видом функции распределения частиц по импульсам. Поскольку в выборе функции распределения имеется функциональная степень свободы, форма «потенциала», соответствующего «силе» $F(A)$, а с ней и вид решения допускают широкий произвол. Профили магнитного поля, тока и возмущения концентрации плазмы для одного из простейших решений указанного уравнения показаны на рис. 8 (см. подробнее в [26]). Самосогласованная функция распределения частиц может быть сильно анизотропной и неравновесной, немаксвелловской (в отличие от слоя Харриса), что при определённых условиях ведёт к эволюции и различного рода неустойчивостям в тех или иных подклассах рассмотренных токовых конфигураций. С учётом этого обстоятельства полученные результаты открывают широкие возможности анализа филаментации и динамики бесстолкновительных токовых конфигураций в плазме. Так или иначе, найденный широкий класс токовых слоёв и филаментов важен для интерпретации наблюдений соответствующих магнитных структур в релятивистских ударных волнах и джетах, свойственных ядрам активных галактик, микроквасарам и источникам гамма-всплесков. На основе полученных решений возможно единое описание различных токовых конфигураций в указанных астрофизических объектах.

Вместо заключения

Вся доступная нам информация о релятивистских джетах извлекается из их излучения, его спектра, поляризации, переменности. На этой основе можно судить, хотя и не вполне однозначно, о свойствах ускоренных частиц, породивших излучение. Гораздо сложнее построить полную физическую картину данного явления, охарактеризовать динамику джетов, процессы их формирования и ускорения до релятивистских скоростей. Об этом приходится судить исходя из общезначимых физических соображений. Затронутые в настоящей статье процессы и явления, конечно, представляют собой лишь отдельные фрагменты требуемой картины. Однако даже эти, уже выясненные фрагменты демонстрируют разнообразную и богатую физику релятивистских джетов, для понимания которой необходима не только физика плаз-

мы, гидродинамика, ядерная физика, электродинамика, но и глубокий нелинейно-волновой подход.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект 05-02-17525), Совета по грантам Президента Российской Федерации для поддержки молодых российских ученых и ведущих научных школ Российской Федерации (грант НШ-4588.2006.2), программы «Происхождение и эволюция звезд и галактик» Президиума РАН, гранта Президента Российской Федерации МК-2752.2005.2.

Литература

1. Relativistic Jets in AGNs // Proc. Int. Conf. (Cracow, May 27–30, 1997) / Eds. M. Ostrowski, M. Sikora, G. Madejski, M. Begelman.
2. Whitney A. R., Shapiro I. I., Rogers A. E. E. et al. // Science. 1971. V. 173. P. 225–230.
3. Urry C. M., Padovani P. // Publications of the Astronomical Society of the Pacific. 1995. V. 107. P. 803–845.
4. Blandford R. D. // Progress of Theoretical Physics Supplement. 2001. № 143. P. 182–201.
5. Mirabel I. F., Rodriguez L. F. // Nature. 1994. V. 371. P. 46–48.
6. Fender R. P. Jets from X-ray binaries // Compact Stellar X-ray Sources / W. H. G. Lewin, M. van der Klis (Eds.). Cambridge, 2006. P. 381; astro-ph/0303339.
7. Постнов К. А. // УФН. 1999. Т. 169, № 5. С. 545–558.
8. Derishev E. V., Kocharovskiy V. V., Kocharovskiy Vl. V. // Astrophys. J. 1999. V. 521. P. 640–649.
9. Белянин А. А., Дерисhev Е. В., Кочаровский В. В., Кочаровский Вл. В. О влиянии свободных нейтронов на динамику и излучение астрофизической плазмы // Изв. ВУЗов. Радиофизика. 2001. Т. 44, №1–2. С. 3–16.
10. Бережко Е. Г., Крымский Г. Ф. // УФН. 1988. Т. 154, № 1. С. 49–91.
11. Derishev E. V., Aharonian F. A., Kocharovskiy V. V., Kocharovskiy Vl. V. // Phys. Rev. D. 2003. V. 68. P. 043003(1–10).
12. Stern B. E. // Monthly Notices of the Royal Astronomical Society. 2003. V. 345, № 2. P. 590–600.
13. Derishev E. V., Aharonian F. A., Kocharovskiy Vl. V. // Astrophys. J. 2007. V. 655, № 1 (in press); astro-ph/0605456.
14. Derishev E. V., Kocharovskiy V. V., Kocharovskiy Vl. V. Cosmological Gamma-Ray Bursts: the Modern Scenario // Proc. 3rd Int. Sakharov Conf. on Physics / A. Semikhatov, M. Vasiliev, V. Zaikin (Eds.). 2003. V. 1. P. 373–385.
15. Чен Ф. Введение в физику плазмы. М.: Мир, 1987.
16. Yoon P., Davidson R. C. // Phys. Rev. A. 1987. V. 35. P. 2718.
17. Frederiksen J. T. et al. // Ap. J. 2004. V. 608. L13.
18. Medvedev M., Loeb A. // Ap. J. 1999. V. 526. P. 697.

19. *Spitkovsky A.* // AIP Conf. Proc. 2006. V. 801. P. 345; astro-ph/0603211.
20. *Wiersma J., Achterberg A.* // Astron. Astrophys. 2004. V. 428. P. 365.
21. *Lyubarsky Y., Eichler D.* // astro-ph/0512579.
22. *Harris E. G.* // Nuovo Cimento. 1962. V. 23. P. 117.
23. *Yoon P., Lui A.* // JGR. 2005. V. 110. A01202.
24. *Бескин В. С.* Осесимметричные стационарные течения в астрофизике. М.: Физматлит, 2005.
25. *Bernstein I. B., Greene J. M., Kruskal M. D.* // Phys. Rev. 1957. V. 108. P. 546.
26. *Derishev E. V., Kocharovsky V. V., Kocharovsky Vl. V., Martyanov V. Ju.* // Frontiers of Nonlinear Physics: Proc. of the 2nd Int. Conf. / Ed. by A. Litvak. N. Novgorod: Institute of Applied Physics RAS, 2005.

САМООРГАНИЗАЦИЯ ПЛАЗМЫ И МАГНИТНОГО ПОЛЯ В НЕРАВНОВЕСНЫХ АСТРОФИЗИЧЕСКИХ ОБЪЕКТАХ

И. Н. Топтыгин

Для многих астрофизических объектов характерны быстрые процессы эволюции, сопровождающиеся выделением больших запасов энергии в форме механического движения вещества (солнечные и звездные вспышки, аккреция, взрывы сверхновых, звездные ветры, столкновения облаков и др.). При этом происходит генерация сильных ударных волн. В статье обсуждаются основные нелинейные явления, присущие сильным бесстолкновительным магнитогидродинамическим (МГД) ударным волнам в разреженных астрофизических средах с магнитным полем. Одним из важных факторов в этих условиях выступают надтепловые (часто релятивистские и ультрарелятивистские) частицы, которые ускоряются вблизи ударных фронтов, отбирая энергию от механического движения окружающей фоновой среды.

Процесс ускорения обладает высокой эффективностью [1–3], благодаря чему ускоренным частицам может быть передана значительная доля (десятки процентов) начальной механической энергии. Это приводит к существенной модификации профиля газодинамического течения по сравнению с ударной волной без ускоренных частиц. Надтепловая компонента, взаимодействуя с фоновой многокомпонентной средой (ионы, электроны, нейтральные атомы), создает нестационарные электрические токи, которые, в свою очередь, могут генерировать турбулентные пульсации МГД-типа, а также крупномасштабные квазистатические магнитные поля. Эти вторичные поля и турбулентность оказывают существенное влияние на процесс ускорения частиц, повышая, как правило, его эффективность.

Таким образом, в сильно неравновесных астрофизических объектах с ударными волнами исследователь имеет дело с богатым набором взаимосвязанных нелинейных явлений, анализ которых труден, и до полного понимания этих процессов еще далеко. В то же время их анализ очень важен для оценки физических условий в соответствующих объектах и для интерпретации нетепловых излучений в широком диапазоне частот, которые интенсивно исследуются в последние годы многими международными обсерваториями. В данной статье произведем обсуждение (в основном на качественном уровне) перечисленных выше физических явлений и моделей, которые можно привлечь для их

интерпретации. Количественные результаты и их обоснование можно найти в статьях, указанных в списке литературы.

1. Ударные волны (характерные параметры)

Межпланетное пространство: ($n \approx 5$ протонов/см³, $B \approx 5 \cdot 10^{-5}$ Гс, $u_w \approx 400$ км/с — скорость солнечного ветра вблизи орбиты Земли). Главный источник ударных волн: солнечные вспышки ($E \lesssim 10^{32}$ эрг). Наиболее сильные ударные волны: $u \approx (2 \div 3) \cdot 10^3$ км/с, число Маха $> M \approx 50$; типичные: $M \approx 2 \div 3$; толщина фронта $500 \lesssim \delta \lesssim 10^4$ км. Пробег фоновых протонов относительно кулоновских столкновений $\Lambda_p \approx 10^8$ км. Из сравнения последних величин однозначно следует: ударные волны в солнечном ветре бесстолкновительны.

Галактический диск («теплая» фаза: $n \approx 0,2$ атомов/см³, $B \approx 2 \cdot 10^{-6}$ Гс, $T \approx 10^4$ К, степень ионизации $\approx 0,1$). Вспышки сверхновых: $E \lesssim 10^{51}$ эрг. Начальная фаза расширения остатка (свободный разлет оболочки): $u \approx (10 \div 30) \cdot 10^3$ км/с, $M \approx 10^3$. Стадия Седова: $u \approx (2 \div 3) \cdot 10^3$ км/с, $M \approx 100$. Звездные ветры (O , B , WR): $u \approx (2 \div 3) \cdot 10^3$ км/с, $M \approx 100$, полная выделяемая энергия $E \lesssim 10^{50}$ эрг.

Другие источники ударных волн: а) столкновения межзвездных облаков; б) спиральная ударная волна в галактическом диске; в) ударная волна в галактическом ветре.

Во всех перечисленных объектах плотность среды очень низкая, ударные волны бесстолкновительны.

2. Ускорение частиц (линейная теория)

Простейшая модель — плоский стационарный ударный фронт в турбулентной среде. Фоновая плазма — нерелятивистская. Профиль течения — прямоугольная ступенька. Относительное сжатие

$$\sigma = \frac{u_1}{u_2} = \frac{\rho_2}{\rho_1}, \quad 1 \leq \sigma \leq \frac{\gamma_g + 1}{\gamma_g - 1} = 4 \quad (1)$$

для одноатомного нерелятивистского газа. Пробные надтепловые ($r_g \gg \delta$) частицы (их гирорадиус больше толщины фронта) совершают случайные блуждания, свободно пересекают фронт и ведут себя как разреженный газ в сосуде между стенками, сближающимися с относительной скоростью, равной скачку скорости на ударном фронте $\Delta u = u_1 - u_2 > 0$. Происходит нагрев частиц, т. е. ускорение механизмом Ферми.

В линейной постановке, т. е. в пренебрежении влиянием ускоряемых частиц на газодинамическое течение вещества, можно получить точное решение для функции распределения ускоренных частиц [4]. За фронтом

$$N(p) = \frac{3Q_0}{\Delta u p_0^3} \left(\frac{p_0}{p} \right)^\alpha, \quad \alpha = \frac{3\sigma}{\sigma - 1}, \quad z > 0; \quad (2)$$

перед фронтом

$$N(z, p) = N(p) \exp \int_0^z \frac{u_1 dz'}{\kappa_1(z', p)}, \quad z < 0. \quad (3)$$

Здесь N — функция распределения (плотность в фазовом пространстве), $\kappa_1(z, p)$ — коэффициент пространственной диффузии ускоряемых частиц, Q_0 — мощность инжекции частиц с импульсом p_0 в режим ускорения (на единицу площади фронта). Эта величина рассматривается здесь как внешний параметр.

Приведенное решение изображает «универсальный» и однородный степенной спектр ускоренных частиц за фронтом и их неоднородное распределение в области толщиной $l(p) \approx \kappa_1(p)/u_1$ перед фронтом.

Важность нелинейных эффектов в ускорении обнаруживается при рассмотрении предельно сильного ударного фронта (число Маха $M \gg 1$, $\sigma \rightarrow 4$, $\alpha \rightarrow 4$). При этом давление ускоренных частиц растет неограниченно:

$$P_c = \frac{1}{3} \int_{p_0}^{p_m} p v N(p) p^2 dp \approx P_0 \ln \frac{p_m}{p_0} \rightarrow \infty \quad \text{при} \quad p_m \rightarrow \infty. \quad (4)$$

Это означает, что значительная часть энергии потока перейдет к ускоренным частицам, произойдет обрезание спектра на некоторой энергии и нелинейное искажение фронта.

Еще два фактора усиливают нелинейный эффект:

1) уменьшение показателя адиабаты Пуассона. Для релятивистских частиц $\gamma_c < \gamma_g = 5/3$, для ультрарелятивистских частиц $\gamma_c \rightarrow 4/3$. Это приводит к относительному сжатию $\sigma = (\gamma_c + 1)/(\gamma_c - 1) \rightarrow 7$, $\alpha \rightarrow 3, 5$, $P_c \propto p_m^{4-\alpha}$, т. е. переход энергии к ускоренным частицам возрастает;

2) убегающие частицы: с ростом импульса растет пробег частиц, и они перестают удерживаться возле фронта, убегают и уносят с собой энергию. Это ведет к росту сжатия: $\sigma > 7$, $\alpha \rightarrow 3$ при $\sigma \gg 1$.

Вывод: процесс ускорения на сильном фронте по своей природе нелинеен, он может заметно изменить структуру ударного перехода и эффективность ускорения.

3. Нелинейные эффекты в ускорении

Рассмотрим только взаимовлияние ускоренных частиц и газодинамического потока. Считаем, что магнитное поле перпендикулярно фронту и задано, турбулентность тоже задана. При одновременном определении макроскопических характеристик ударного перехода (теплового скачка скорости, формы профиля течения в предфронте, сжатия среды) и функции распределения ускоренных частиц используем сочетание газодинамического и кинетического подходов [5].

Используется система нелинейных уравнений в частных производных, включающая в себя макроскопические законы переноса массы, импульса и энергии на ударном переходе и кинетическое уравнение для ускоряемых частиц. Изначально задаются все характеристики потока набегающих нерелятивистских частиц (плотность вещества, скорость, температура), число Маха, скорость инжекции частиц в режим ускорения, импульсы инжекции (p_0) и убегания (p_m). Все остальные величины определяются самосогласованным способом (профиль скорости $u(z)$, функция распределения ускоренных частиц на ударном переходе со сложным профилем, их показатель адиабаты γ_c , сжатие σ , поток энергии убегających частиц q_m , доля динамического давления $\eta = P_{c2}/\rho_1 u_1^2$, преобразующаяся в давление ускоренных частиц).

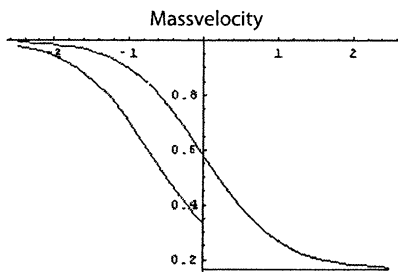


Рис. 1. Схематическое изображение профилей размытого фронта (верхняя кривая) и фронта с тепловым скачком (нижняя кривая).

Используемые уравнения позволяют найти два качественно различных профиля скорости: а) модифицированный фронт с тепловым скачком; б) полностью размытый фронт (рис. 1). Профиль ударного перехода для фронта с тепловым скачком:

$$\frac{u(\zeta)}{u_1} = \frac{\sigma_m - 1 + [2\sigma_* - \sigma_m - 1] \tanh \zeta}{(\sigma_m - 1)\sigma_* + [(\sigma_m + 1)\sigma_* - 2\sigma_m] \tanh \zeta}, \quad \zeta \leq 0; \quad (5)$$

$$u(\zeta) = u_2 = \text{const}, \quad \zeta > 0, \quad \sigma_m = \frac{\gamma_c + 1}{(\gamma_c - 1)(1 - 2q_m/u_1 P_{c2})}, \quad (6)$$

где $\sigma_* = \frac{u_*}{u_1}$ — сжатие среды в предфронте, ζ — безразмерная координата z .

Профиль размытого фронта имеет обычную форму:

$$\frac{u(\zeta)}{u_1} = \frac{1}{2} \left(1 + \frac{1}{\sigma_m} \right) - \frac{1}{2} \left(1 - \frac{1}{\sigma_m} \right) \tanh \zeta. \quad (7)$$

Его толщина, как и толщина предфронта при наличии теплового скачка, определяется транспортным пробегом релятивистских частиц, тогда как толщина теплового скачка определяется свойствами фоновой среды.

Уравнение переноса для релятивистских частиц решается с помощью преобразования Меллина по импульсной переменной p и перехода к новой координатной переменной $\xi(\zeta) = u(\zeta)/u_1$. Решение выражается через гипергеометрическую функцию $F(a, b; c; \xi)$. В итоге за фронтом формируется степенной спектр

$$N(p) = (\alpha - 3)N_0 \frac{p_0^{\alpha-3}}{p^\alpha} \propto p^{-\alpha}. \quad (8)$$

Показатель α вычисляется из трансцендентного уравнения, содержащего гиперфункцию. Его наименьшее значение в рассматриваемой модели равно трем. Результаты приведены на рис. 2, 3.

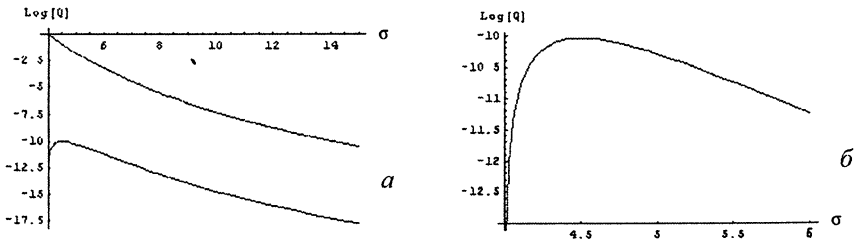


Рис. 2. Схематическая зависимость безразмерной скорости инъекции Q (произвольные единицы) от полного сжатия среды σ для размытого фронта (верхняя кривая) и фронта с тепловым скачком (нижняя кривая) (а); начальный участок кривой, которая относится к фронту с тепловым скачком) (б).

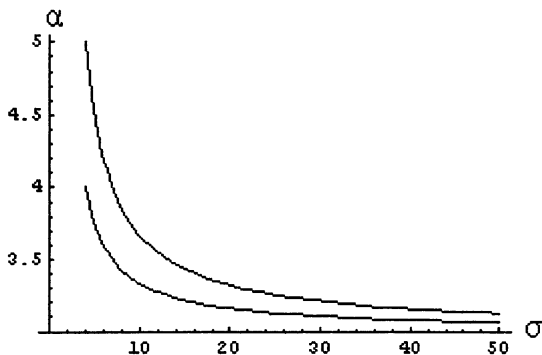


Рис. 3. Показатель α спектра частиц, ускоренных на размытом фронте (верхняя кривая) и на фронте с тепловым скачком (нижняя кривая).

Отметим характерные нелинейные особенности рассмотренной физической системы.

1. В некотором диапазоне скоростей инжекции (см. рис. 2) имеются две ветви решений и три разных состояния ударного перехода. В линейном режиме (восходящая ветвь нижней кривой) сжатие растет с ростом инжекции. В нелинейных режимах сжатие растет с уменьшением инжекции. Наличие нескольких ветвей решения — характерная черта самых разных нелинейных систем.

2. Ветвь с тепловым скачком существует в ограниченной области инжекций. Вблизи того значения скорости инжекции, которое соответствует максимуму нижней кривой, по-видимому, возможны переходы колебательного типа между состояниями, лежащими на верхней и нижней кривых. При этом сжатие среды существенно изменяется.

3. Состояния ударного перехода были исследованы на устойчивость [6]. Неустойчивость обычной нерелятивистской ударной волны относительно искажений ударного фронта (гофрировочную неустойчивость) исследовали ранее Дьяков [7] и Конторович [8]. Такая неустойчивость проявляется лишь при весьма специфической форме адиабаты Гюгонио и отсутствует в случае политропного газа. В бесстолкновительных ударных волнах с релятивистскими частицами устойчивы состояния с умеренным сжатием ($\sigma < 8 \div 9$). При сжатиях $4 < \sigma < 7$ возможна спонтанная генерация звука затухающими колебаниями ударного фронта, не приводящая к его дезинтеграции. При больших сжатиях фронт разрушается (распадается на отдельные фрагменты, размер которых превышает толщину предфронта). Отметим, что устой-

чивым состояниям соответствуют бóльшие значения полной энтропии системы.

4. Полная энергия ускоренных частиц и форма спектра заметно различаются на двух ветвях полученных решений. Температура за фронтом в отсутствие теплового скачка намного ниже, так как она определяется адиабатой Пуассона, а не Гюгонио, и весь избыток энергии передается ускоренным частицам. Поэтому в отсутствие и при наличии теплового скачка должны наблюдаться сильные различия в интенсивностях излучения релятивистских частиц и нагретой плазмы из-за фронта.

4. Генерация магнитного поля

Рассмотренные эффекты не исчерпывают всего богатства нелинейных явлений, присущих сильному ударному фронту с ускоренными частицами. Важнейшим компонентом общей картины выступают магнитные поля, крупномасштабные и турбулентные. Они необходимы для длительного удержания частиц в области ускорения (т. е. в окрестности фронта). Требуемые поля могут генерировать сами ускоряемые частицы, поскольку они вместе с возбуждаемыми ими электронами и ионами фоновой плазмы создают электрический ток, необходимый для генерации магнитных полей. Энергетический ресурс возбуждаемых таким механизмом полей можно оценить из закона сохранения энергии. Принимая умеренную эффективность ускорения частиц на фронте ($\eta = 0,1$ — доля энергии потока, переходящая ускоренным частицам), оцениваем верхний предел возможного поля:

$$\frac{1}{2} n_0 m_p u_1^2 \eta \gtrsim \frac{B^2}{8\pi}, \quad \eta \approx 0,1. \quad (9)$$

Задавая характерное значение скорости вещества в предфронте ($u_1 \approx 3 \cdot 10^8$ см/с, стадия Седова для сверхновой), получим оценки поля для основных фаз межзвездной среды:

$n_0, \text{ см}^{-3}$	$B_{\text{max}}, \text{ Гс}$	$B_{\text{obs}}, \text{ Гс}$	Примечание
20	$2 \cdot 10^{-3}$	$1,5 \cdot 10^{-5}$	Нейтральные облака
0,2	$2 \cdot 10^{-4}$	$2 \cdot 10^{-6}$	«Теплая» среда
$2 \cdot 10^{-3}$	$2 \cdot 10^{-5}$?	Горячие каверны

Как следует из этих оценок, генерация поля ускоренными частицами может поднять его значение по сравнению с наблюдаемыми средними по галактическому диску значениями приблизительно на 2 порядка величины. Но нужно подтвердить эту возможность путем решения уравнений Максвелла с соответствующими токами.

Система уравнений Максвелла в МГД-приближении имеет вид

$$\frac{\partial \mathbf{B}}{\partial t} = -c \nabla \times \mathbf{E}, \quad \Delta \mathbf{B} = -\frac{4\pi}{c} \nabla \times (\mathbf{j} + \mathbf{j}^{ext}), \quad \nabla \cdot \mathbf{B} = 0. \quad (10)$$

В уравнения входит ток плазмы \mathbf{j} и сторонний ток (ток ускоренных частиц) \mathbf{j}^{ext} . В простейшем варианте ток плазмы подчиняется закону Ома

$$\mathbf{E} + \frac{1}{c} \mathbf{u} \times \mathbf{B} = \frac{\mathbf{j}}{\sigma}. \quad (11)$$

Для наших задач требуется более точное выражение для четырехкомпонентной среды: фоновые электроны (e), фоновые ионы (i), нейтральные атомы (a), релятивистские протоны. Нейтральные атомы во многих случаях могут сильно модифицировать закон Ома (см. монографии [9, 10]). Вместо (11) будем иметь

$$\mathbf{E} + \frac{1}{c} \mathbf{u} \times \mathbf{B} = \frac{\mathbf{j}_{\parallel}}{\sigma} + \frac{\mathbf{j}_{\perp}}{\sigma_{\perp}^{ef}} + \frac{1}{n_i e c} \mathbf{j} \times \mathbf{B}, \quad (12)$$

$$\frac{1}{\sigma_{\perp}^{ef}} = \frac{1}{\sigma} + \frac{F^2 \mathcal{B}^2 \tau_{ia}}{n_i m_i c^2}, \quad F = \frac{n_a m_a}{n_a m_a + n_i m_i}. \quad (13)$$

Появились две магнитные вязкости, $\nu_m = c^2/4\pi\sigma$, $\nu_{ef} = c^2/4\pi\sigma_{\perp}^{ef}$, их отношение

$$\frac{\nu_{ef}}{\nu_m} = 1 + F^2 (\omega_i \tau_{ia}) (\omega_e \tau_{ei}) = 1 + K. \quad (14)$$

Здесь ω_i , ω_e — циклотронные частоты, τ — времена между столкновениями соответствующих частиц. Роль нейтральных атомов характеризует величина K . Для разреженных сред $K \approx 10^7 \div 10^{13}$, т. е. не просто большое, а *очень большое* число. В уравнение для магнитного поля (10) ток входит под знаком ротора, его продольная и поперечная компоненты перемешаны. Поэтому при вычислении магнитного поля вклад от поперечного тока превалирует во всех тех случаях, когда он не равен точно нулю, и можно положить $\mathbf{j} = \mathbf{j}_{\perp}$. Тогда из системы (10) после исключения электрического поля \mathbf{E} получим уравнение для магнитного поля, содержащее сторонний ток:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times (\nu_{ef} \nabla \times \mathbf{B}) + \frac{4\pi}{c} \nabla \times \nu_{ef} \mathbf{j}^{ext}. \quad (15)$$

Если продольным током нельзя пренебречь, то в (15) надо добавить члены с ν_m (а также холловское слагаемое). При $\nu_m \rightarrow 0$, $\nu_{ef} \rightarrow 0$ исчезает джоулева диссипация, а вместе с ней и эффект от стороннего тока. Это отражает экранировку внешнего тока плазмой в среде с идеальной электропроводностью.

Вычислим теперь статическое (крупномасштабное) магнитное поле перед плоским стационарным ударным фронтом с учетом вторичного поля, создаваемого ускоренными частицами [11]. Первичное поле \mathbf{V}_1 считаем однородным и ориентированным под произвольным углом к фронту. Полный сторонний ток в плазме создается релятивистскими протонами, а также возбуждаемыми ими фоновыми протонами и электронами: $\mathbf{j}^{ext} = \mathbf{j}^p + \mathbf{j}^e$. Токи, создаваемые протонами и электронами, можно записать в виде ($z \neq 0$)

$$j_\alpha^p = -e\kappa_{\alpha\beta} \frac{\partial N}{\partial x_\beta} - eD_{\alpha\beta}^p \frac{\partial n_p}{\partial x_\beta} + eu_\alpha(N + n_p) + \sigma_{\alpha\beta}^p E_\beta, \quad \nabla \cdot \mathbf{j}^p = 0, \quad (16)$$

$$j_\alpha^e = eD_{\alpha\beta}^e \frac{\partial n_e}{\partial x_\beta} - eu_\alpha n_e + \sigma_{\alpha\beta}^e E_\beta, \quad \nabla \cdot \mathbf{j}^e = 0, \quad (17)$$

где

$$N = N_0 \exp \left\{ \int_0^z \frac{u_z(z') dz'}{\kappa(z')} \right\} \quad (z \leq 0) \quad (18)$$

— концентрация ускоренных протонов со всеми энергиями; n_p , n_e — неравновесные добавки к концентрациям фоновых протонов и электронов; \mathbf{E} — электрическое поле, возникающее из-за разделения зарядов и удовлетворяющее уравнению электростатики

$$\nabla \cdot \mathbf{E} = 4\pi e(N + n_p - n_e). \quad (19)$$

Рассматриваем статический случай, однородную вдоль фронта среду, зависимость от одной координаты z , отсчитываемой по нормали к фронту. Из приведенных уравнений находим ток:

$$\begin{aligned} j_x^{ext} &= eu_z(z) \left\{ \frac{D_{xz}^e + D_{xz}^p}{D_1^e + D_1^p} - \frac{\kappa_{xz}}{\kappa_1} \right\} N(z), \\ j_y^{ext} &= eu_z(z) \left\{ \frac{D_{yz}^e + D_{yz}^p}{D_1^e + D_1^p} - \frac{\kappa_{yz}}{\kappa_1} \right\} N(z), \\ j_z^{ext} &= 0. \end{aligned} \quad (20)$$

Рассматриваемые токи текут в области, где имеются ускоренные частицы, т. е. в слое толщины порядка $l \approx \kappa_1/u_1$ перед фронтом.

Характерные особенности токов:

1. Для генерации поля нужна анизотропия среды. В изотропной среде (при $\kappa_{xz} = \kappa_{yz} = 0$) ток исчезает. Ток вызван неоднородным распределением ларморовских кружков в предфронте.

2. Токи релятивистских и фоновых частиц входят с противоположными знаками и в принципе могут взаимно сокращаться. Но вероятность этого крайне мала (не известны механизмы полной подстройки движения фоновых частиц под релятивистскую компоненту).

Магнитное поле на фронте вычисляется самосогласованным образом из нелинейного уравнения (15) с учетом влияния генерируемого поля на магнитную вязкость и коэффициенты диффузии частиц. Оценку поля на фронте ($z \rightarrow 0$) можно получить аналитическим путем:

$$B_z = B_{z1}, \quad B_{\perp} \approx \frac{B_{\perp 1}}{\xi^2}. \quad (21)$$

Величина

$$\xi^2 \approx \frac{2\pi\gamma_c n_0 m_p u_1^2 \eta}{B_{z1}^2} \approx 10^4$$

весьма велика при типичных условиях. Вывод: если первоначальное поле имеет нормальную к фронту составляющую, то его параллельные фронту компоненты подавляются ускоренными частицами на несколько порядков величины. Остается лишь компонента, перпендикулярная фронту.

Такое состояние сильно неравновесной плазменной системы представляет собой пример самоорганизации плазмы с магнитным полем. Переход системы в это состояние имеет простой физический смысл. При нормальной ориентации магнитного поля к фронту дальнейшая генерация поля в силу симметрии системы прекращается, и в этом смысле такая конфигурация поля устойчива. Именно в это устойчивое состояние стремится перейти система. Если вдали от фронта имеется тангенциальная компонента поля, то она убывает по величине по мере приближения к фронту. Степень подавления определяется запасом энергии ускоренных частиц, создающих вторичное поле.

Но полученный результат (подавление поля) не следует абсолютизировать. Он специфичен для рассмотренной выше модели, которой присуща высокая степень симметрии (однородная вдоль фронта инжекция частиц в режим ускорения, неограниченный размер фронта). При неоднородной инжекции (например, за счет неоднородности плотности перед фронтом

$$n(\mathbf{r}, t) = n_0 \left(1 + \sum_{\mathbf{k}} \frac{\Delta n_{\mathbf{k}}}{n_0} e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{u}t)} \right), \quad k < \frac{2\pi}{l},$$

где $l = \kappa/u$ — толщина предфронта, κ — коэффициент диффузии релятивистских протонов) возникает ток, параллельный фронту, который

генерирует магнитные неоднородности, в том числе тангенциальные, в предфронте. Этим механизмом генерируются неоднородности с пространственными масштабами порядка и больше толщины предфронта. Подробнее этот вопрос рассмотрен в статьях автора [12, 13], а наблюдательные подтверждения имеются в [14].

Существенно также предположение о неограниченных размерах плоского фронта и вытекающая из него одномерность задачи. В действительности из-за ограниченных размеров ударных фронтов в реальных объектах каждый такой фронт выступает источником крупномасштабного электрического тока, растекающегося по всей Галактике и способного генерировать магнитное поле вдали от фронта, а не только в предфронте [15].

5. Генерация МГД-турбулентности

Стохастичность движения ускоряемых частиц и их блуждания в окрестности фронта возникают из-за их рассеяния турбулентными магнитными полями. Рассеяние носит резонансный характер: главный эффект рассеяния вызван магнитными неоднородностями с масштабами порядка гирорадиуса $r_g = cp/eB$ частицы. В среднегалактическом поле протон с энергией 1 ГэВ (на максимуме энергетического спектра) имеет гирорадиус 10^{12} см, с энергией $3 \cdot 10^6$ ГэВ (характерная особенность — загиб — в энергетическом спектре) имеет гирорадиус $3 \cdot 10^{18}$ см = 1 пк. Таков масштаб стохастических неоднородностей, способных обеспечить эффективное рассеяние. Это могут быть только крупномасштабные моды МГД-типа. Поэтому анализ возможностей генерации крупномасштабных возмущений очень важен для объяснения происхождения частиц сверхвысоких энергий.

Как и в предыдущем случае, начнем с вычисления электрического тока. При взаимодействии частиц с турбулентностью существенны резонансные эффекты, поэтому недостаточно ограничиться рассмотрением только полного тока ускоренных частиц. Следует учесть влияние конкретной МГД-волны на частицы с заданной энергией, т. е. фактически вычислить диэлектрическую проницаемость четырехкомпонентной среды [16]. Используем кинетическое уравнение для ускоренных частиц в предфронте:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + e\mathbf{E} \cdot \frac{\partial f}{\partial \mathbf{p}} - \frac{ec}{\mathcal{E}} (\mathbf{B}_0 + \mathbf{b}) \cdot \mathcal{O}f = I[f], \quad (22)$$

где

$$\mathcal{O} = \mathbf{p} \times \frac{\partial}{\partial \mathbf{p}} \quad (23)$$

— оператор поворота импульса, \mathbf{E} , \mathbf{b} — внешнее электромагнитное поле МГД-волны.

Линеаризуем кинетическое уравнение (22), считая внешнее поле слабым и выделяя из функции распределения $f = f_0 + \delta f$ малую часть δf , обусловленную этим полем:

$$\frac{\partial \delta f}{\partial t} + \mathbf{v} \cdot \frac{\partial \delta f}{\partial \mathbf{r}} - \frac{ec}{\mathcal{E}} \mathbf{B}_0 \cdot \mathcal{O} \delta f = -e\mathbf{E} \cdot \frac{\partial f_0}{\partial \mathbf{p}} + \frac{ec}{\mathcal{E}} \mathbf{b} \cdot \mathcal{O} f_0 + I[\delta f]. \quad (24)$$

Функция распределения f_0 , входящая в правую часть равенства, представляет собой степенную функцию распределения ускоренных частиц, вычисленную ранее.

Дальнейший аналитический расчет удается произвести только для волн, распространяющихся вдоль поля \mathbf{B}_0 или под малым углом к нему. С помощью найденной функции распределения δf получим выражение для электрического тока

$$\mathbf{j}^{cr} = -(\sigma'_{cr} + i\sigma''_{cr})\mathbf{e}_{\parallel} \times \mathbf{b} + (g' + ig'')\mathbf{b} \quad (25)$$

с кинетическими коэффициентами

$$\sigma'_{cr} = \frac{(\alpha - 3)^2 \omega_0^2 \omega}{8(\alpha - 1)\omega_i^2} \left(\frac{ck}{\omega_i}\right)^{\alpha-4} \frac{u_0}{c}, \quad \sigma''_{cr} = \frac{(\alpha + 1)(\alpha - 3)^2 \omega_0^2 \omega u_0}{60\pi(\alpha - 4)\omega_i^2 c},$$

$$g' = \frac{\omega_0^2}{12\pi\omega_i} \left(\frac{u_0}{c} - \frac{\alpha\omega k_{\parallel}}{ck^2}\right), \quad g'' = -\frac{(\alpha - 3)\omega_0^2}{8\alpha(\alpha - 2)\omega_i} \left(\frac{ck}{\omega_i}\right)^{\alpha-3} \left(\frac{u_0}{c} - \frac{\alpha\omega k_{\parallel}}{ck^2}\right). \quad (26)$$

Все они пропорциональны полной концентрации N_0 релятивистских частиц,

$$\omega_0^2 = \frac{4\pi e^2 N_0}{m_p},$$

и зависят от частоты ω и волнового вектора \mathbf{k} МГД-моды.

После подстановки найденного тока в МГД-уравнения получим дисперсионное уравнение для собственных мод, учитывающее диссипацию колебаний за счет эффективной магнитной вязкости и влияние на них популяции релятивистских частиц:

$$\omega^2 - v_A^2 k^2 + \frac{4\pi\nu_{ef} k_{\parallel} \omega}{c} (\sigma'_{cr} \mp g'') - i\nu_{ef} k^2 \omega \left(1 - \frac{4\pi k_{\parallel}}{ck^2} (\sigma''_{cr} \mp g')\right) = 0. \quad (27)$$

Мнимая часть частоты имеет вид

$$\gamma_{nr} = -\frac{1}{2}\nu_{ef}k^2 \left[1 - \frac{4\pi k_{\parallel}}{ck^2}(\sigma_{cr}'' \pm g') \right], \quad (28)$$

где второе слагаемое в квадратных скобках — вклад от релятивистских частиц. Эта величина может иметь любой знак и при

$$k < k_c = \frac{4\pi|g'|}{c} \quad (29)$$

превосходит единицу по абсолютной величине. Поэтому в ограниченной неравенством (29) области малых k и больших длин волн имеет место раскачка турбулентности. Мнимая часть частоты $\gamma_{nr} = \Im \omega$ в этом случае положительна и при $\gamma \ll |\omega|$ имеет значение

$$\gamma_{nr} = \frac{1}{2} \left(\frac{4\pi|g'|}{ck} - 1 \right) \nu_{ef}k^2 \approx \frac{2\pi|g'| \nu_{ef}k}{c}. \quad (30)$$

Инкремент можно вычислить и из баланса магнитной энергии, минуя уравнение дисперсии для волн. Усиление колебаний в данном случае носит нерезонансный характер и не ограничено никакими резонансными условиями.

Ранее многие авторы [17–19] рассматривали резонансную раскачку низкочастотных МГД-волн релятивистскими частицами в полностью ионизованной бесстолкновительной плазме. Резонансный характер процесса проявляется в том, что волна с длиной λ раскачивает частицами, гирорадиусы r_g которых удовлетворяют резонансному условию $r_g \geq \lambda/2\pi$. Число таких частиц убывает с ростом энергии. В плазме с примесью нейтральной компоненты такой механизм работает для волн с длинами $\lambda < \Lambda_{ia}$. Инкремент для резонансного случая вычисляется аналогично предыдущему и имеет вид

$$\gamma_{res} \approx \frac{(\alpha - 3)\pi \omega_0^2 v_A N(p \geq p_r)}{\alpha(\alpha - 2) \omega_i c N_0} \left(\frac{u}{c} + \alpha \frac{v_A}{c} \right), \quad (31)$$

где $p_r = eB_0/ck$ — резонансный импульс частиц, который определяется условием $r_g(p_r) = k^{-1}$, а $N(p \geq p_r)$ — плотность числа частиц с импульсами, превышающими его резонансное значение.

Сравним эффективности резонансной и нерезонансной раскачек турбулентности вблизи ударного фронта, пользуясь развитой теорией. Пусть ускорение происходит до энергии $\mathcal{E}_m \approx 3 \cdot 10^6$ ГэВ (энергия «колена» в спектре протонов) в теплой межзвездной среде. Примем умеренный темп ускорения ($\eta = 0, 1$ — доля энергии, переходящая к

ускоренным частицам) и слабую модификацию фронта ускоренными частицами ($\alpha = 4$). Из закона сохранения энергии следует, что при этом в релятивистскую компоненту превращается доля

$$\frac{N_0}{n_0} \approx \frac{0,2}{\ln(p_m/p_0)} \left(\frac{u}{c}\right)^2 \approx 1,4 \cdot 10^{-2} \left(\frac{u}{c}\right)^2$$

от полной концентрации n_0 протонов перед фронтом. Пользуясь этой оценкой и вычисляя нужные частоты для типичных значений параметров галактической среды, получим

$$\left(\frac{\gamma}{\omega}\right)_{res} \approx 7,5 \cdot 10^5 \left(\frac{u}{c}\right)^3 \quad (32)$$

для резонансного случая и

$$\left(\frac{\gamma}{\omega}\right)_{nr} \approx 7,5 \cdot 10^8 \left(\frac{u}{c}\right)^3 \quad (33)$$

для нерезонансной раскачки. В рассмотренном случае нерезонансный эффект оказывается значительно более сильным.

В применении к сверхновым на стадии свободного разлета обе формулы, резонансная и нерезонансная, приводят к соотношению $\gamma/\omega \gg 1$, т. е. происходит очень быстрая генерация турбулентности и поля. При этом обе формулы для инкремента, полученные с использованием условия $\gamma \ll \omega$, неприменимы. Генерируемые в этом режиме возмущения нельзя считать обычными МГД-волнами. На стадии Седова, когда нагребенная масса среды превысит массу оболочки, скорость фронта убывает и ее характерные значения становятся порядка $u \approx (2 \div 3) \cdot 10^8$ см/с. При этом $u/c \approx 10^{-2}$, резонансный инкремент становится меньше частоты, а нерезонансный инкремент по-прежнему велик по сравнению с частотой, если имеется достаточно нейтральных атомов и множитель $F/(1-F)$ в формуле для ν_{ef} не слишком мал.

В итоге можно сделать полукачественный вывод, что нет препятствий для ускорения частиц в остатках сверхновых до энергий «колена» (10^6 ГэВ) и более высоких. Но требует дальнейшей разработки трудный вопрос о спектре генерируемой частицами турбулентности. Необходимо также корректный учет неоднородности предфронта. Более подробное рассмотрение вопросов, затронутых во второй части настоящей статьи, проведено в обзоре [20].

Работа поддержана Российским фондом фундаментальных исследований, гранты 04-02-16595, 06-02-16844а, 07-02-00245а.

Литература

1. Аксфорд В., Лир Е., Скадрон Дж. Ускорение космических лучей ударными волнами // X Ленинградский семинар по космофизике. Л.: Изд-во ФТИ им. А.Ф. Иоффе, 1978. С. 125–134.
2. Bell A. R. MNRAS 1987. V. 182. P. 147–157.
3. Бережко Е. Г., Крымский Г. Ф. Ускорение космических лучей ударными волнами // УФН. 1988. Т. 154. С. 49–91.
4. Toptygin I. N. Cosmic Rays in Interplanetary Magnetic Fields. Dordrecht: D. Reidel, 1985. 375 p.
5. Топтыгин И. Н. Структура фронта бесстолкновительной ударной волны с релятивистскими ускоренными частицами // ЖЭТФ. 1997. Т. 112. С. 1584–1602.
6. Топтыгин И. Н. Об устойчивости сильных бесстолкновительных ударных волн с релятивистскими ускоренными частицами // Письма в Астрон. журн. 1999. Т. 25. С. 40–48.
7. Дьяков С. П. // ЖЭТФ. 1954. Т. 27. С. 288–295.
8. Конторович В. М. // ЖЭТФ. 1957. Т. 33. С. 1527–1529.
9. Каулинг Т. Магнитная гидродинамика. М.: Атомиздат. 1978. 142 с.
10. Пикельнер С. Б. Основы космической электродинамики. М.: Наука ФМ, 1966. 408 с.
11. Долгинов А. З., Подольская Н. И., Топтыгин И. Н. Перестройка магнитного поля ускоренными частицами вблизи фронта ударной волны // Письма в Астрон. журн. 2004. Т. 30. С. 196–206.
12. Топтыгин И. Н. Влияние ускоренных частиц на крупномасштабное магнитное поле в молодых остатках сверхновых // Письма в Астрон. журн. 2004. Т. 30. С. 196–206.
13. Топтыгин И. Н. Нерезонансная генерация МГД-возмущений при взрывах сверхновых в неоднородной межзвездной среде // Письма в Астрон. журн. 2007. Т. 33. № 3 (в печати).
14. Reynolds, S. P., Gilmore D. M. // Astron. J. 1993. V. 106. P. 272–285.
15. Dolginov A. Z., Toptygin I. N. Galactic magnetic field generation by the electric current of cosmic rays. // Astron. Astrophys. 2004. V. 416. P. 411–422.
16. Быков А. М., Топтыгин И. Н. Генерация магнитных флуктуаций вблизи ударного фронта в частично ионизованной среде // Письма в Астрон. журн. 2005. Т. 31. С. 839–846.
17. Wentzel D. G. // Astrophys. J. 1968. V. 152. P. 987–995.
18. Kulsrud R. M., Pears W. F. // Astrophys. J. 1969. V. 156. P. 445–459.
19. Skilling J. MNRAS 1975. V. 173. P. 255–265.
20. Быков А. М., Топтыгин И. Н. Неустойчивости многокомпонентной плазмы с ускоренными частицами и генерация магнитных полей в астрофизических объектах // УФН. 2007. Т. 177, № 2 (в печати).

ПУЛЬСАЦИИ ЗВЕЗД

Ю. А. Фадеев

1. Введение

Наиболее продолжительную часть своей жизни звезды находятся в состоянии гидростатического равновесия, при котором в каждой точке звезды отрицательный градиент давления уравнивает силу тяжести вышележащих слоев. Однако при гидростатическом равновесии, существующем главным образом благодаря выделению энергии в реакциях термоядерного синтеза, есть возможность возникновения звездных пульсаций — периодических движений газа около состояния равновесия. В наиболее простом случае это радиальные пульсации, когда амплитуда и фаза смещения зависят только от радиуса, т. е. звезда всегда сохраняет сферическую форму. По порядку величины период радиальных колебаний сопоставим с удвоенным временем пробега звуковой волны от центра звезды до ее поверхности и, в зависимости от распределения плотности газа вдоль радиуса, составляет в различных звездах от нескольких минут до сотен суток. Амплитуда смещения внешних слоев может быть как пренебрежимо малой, так и сравнимой с радиусом звезды. При нелинейных радиальных пульсациях болометрическая светимость звезды может изменяться в два-три раза, и несколько таких звезд были известны уже в древности, хотя природа переменности их блеска стала понятна лишь в XX в. В звездах могут возникать и нерадиальные пульсации, однако вследствие малой амплитуды эти колебания с хорошей точностью описываются в рамках линейной теории. По этой причине все дальнейшее изложение касается радиально пульсирующих звезд, хотя следует заметить, что в подавляющем большинстве случаев неустойчивость радиальных и нерадиальных мод связана с одним и тем же обсуждаемым ниже физическим механизмом.

2. Уравнения газовой динамики

В радиально пульсирующих звездах эффекты осевого вращения и собственного магнитного поля обычно достаточно малы, что позволяет рассматривать звезду в приближении сферической симметрии. В этом случае лагранжева координата определяется массой вещества внутри сферы радиуса r :

$$M_r = 4\pi \int_0^r (r')^2 \rho dr', \quad (1)$$

где ρ — плотность газа. Будем полагать, что вязкость газа пренебрежимо мала и движение обусловлено лишь силой тяжести и градиентом давления. Тогда уравнения неразрывности, сохранения импульса и энергии записываются в виде

$$\frac{\partial r}{\partial M_r} = \frac{1}{4\pi r^2 \rho}, \quad (2)$$

$$\frac{\partial^2 r}{\partial t^2} = -4\pi r^2 \frac{\partial P}{\partial M_r} - \frac{GM_r}{r^2}, \quad (3)$$

$$\frac{\partial E}{\partial t} + P \frac{\partial}{\partial t} \left(\frac{1}{\rho} \right) = \varepsilon - \frac{1}{\rho} \nabla \cdot \mathbf{F}, \quad (4)$$

где t — время, P — полное давление, E — удельная внутренняя энергия, ε — скорость выделения энергии в термоядерных реакциях, \mathbf{F} — поток энергии, G — гравитационная постоянная.

Уравнения (2) — (4) должны быть дополнены соотношениями, определяющими давление, внутреннюю энергию и скорость генерации энергии в функции независимых термодинамических переменных T и ρ :

$$E = E_g(T, \rho) + E_R/\rho, \quad P = P_g(T, \rho) + P_R, \quad \varepsilon = \varepsilon(T, \rho), \quad (5)$$

где $E_g(T, \rho)$ и $P_g(T, \rho)$ — внутренняя энергия и давление, определяемые уравнением состояния, E_R — плотность лучистой энергии, P_R — давление излучения.

2.1. Уравнение лучистой теплопроводности. Перенос энергии в звездах может происходить посредством излучения, конвективного теплообмена и теплопроводности вырожденного электронного газа. В наблюдаемых радиально пульсирующих звездах вещество с высокой точностью подчиняется уравнению идеального газа, и при вычислении полного потока энергии \mathbf{F} достаточно учитывать лишь излучение и конвекцию. Более того, у значительной группы пульсирующих звезд внешние слои газа, участвующие в пульсациях, устойчивы относительно конвекции, и перенос энергии в этих слоях происходит за счет излучения. Дальнейшее обсуждение ограничивается рассмотрением именно этого случая.

Плотность лучистой энергии E_R , поток излучения F_R и давление излучения P_R являются угловыми моментами интенсивности излучения $I_\nu(\mathbf{n})$ и определяются соотношениями

$$E_R = \frac{1}{c} \int_0^\infty d\nu \oint I_\nu(\mathbf{n}) d\omega, \quad (6)$$

$$F_R = \int_0^\infty d\nu \oint I_\nu(\mathbf{n}) \mathbf{n} d\omega, \quad (7)$$

$$P_R = \frac{1}{c} \int_0^\infty d\nu \oint I_\nu(\mathbf{n}) \mathbf{n} \mathbf{n} d\omega, \quad (8)$$

где c — скорость света, ν — частота излучения, ω — телесный угол, \mathbf{n} — направляющий вектор. Следует заметить, что в общем случае P_R является тензором, однако в случае плоской или сферической симметрии определенные выше скалярные величины E_R и P_R полностью описывают анизотропию поля излучения (подробнее см. [1]).

Интенсивность $I_\nu(\mathbf{n})$ является решением уравнения переноса

$$\frac{1}{c} \frac{\partial}{\partial t} I_\nu(\mathbf{n}) + \mathbf{n} \nabla I_\nu(\mathbf{n}) = \eta_\nu - \kappa_\nu I_\nu(\mathbf{n}), \quad (9)$$

где η_ν и κ_ν — коэффициенты эмиссии и поглощения, а указания зависимости переменных от пространственной координаты для сокращения письма опущены. В пульсирующих звездах скорость течения газа по сравнению со скоростью света пренебрежимо мала, и первое слагаемое в левой части (9) может быть отброшено.

Коэффициент эмиссии η_ν зависит от интегралов интенсивности излучения по всей рассматриваемой области пространства, и уравнение переноса (9) принадлежит к интегро-дифференциальным уравнениям. В рассматриваемом случае трудности решения уравнения переноса удается преодолеть благодаря следующим двум обстоятельствам. Во-первых, длина пробега фотонов в звездных недрах значительно меньше масштабов, в пределах которых заметно изменяются характеристики среды. Это позволяет перейти от интегро-дифференциального уравнения переноса (9) к уравнению диффузии излучения, связывающему монохроматический поток излучения с градиентом плотности энергии излучения на частоте ν :

$$\mathbf{F}_\nu = -\frac{c}{3\kappa_\nu} \nabla E_\nu, \quad (10)$$

где величина $1/\kappa_\nu$ представляет собой среднюю длину пробега фотонов. Во-вторых, в звездных недрах излучение находится в термодинамическом равновесии с веществом и зависимость потока излучения

от частоты ν описывается функцией Планка $B_\nu(T)$. Это позволяет перейти от монохроматического уравнения диффузии излучения (10) к уравнению лучистой теплопроводности для болометрического потока:

$$F_R = -4\pi r^2 \frac{4\sigma}{3k_R} \frac{dT^4}{dM_r}, \quad (11)$$

где σ — постоянная Стефана — Больцмана, k_R — Росселандово среднее коэффициента поглощения, определяемое соотношением

$$\frac{1}{k_R} = \frac{\int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B_\nu(T)}{\partial T} d\nu}{\int_0^\infty \frac{\partial B_\nu(T)}{\partial T} d\nu} = \frac{\pi}{4\sigma T^3} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B_\nu(T)}{\partial T} d\nu. \quad (12)$$

В условиях термодинамического равновесия поле излучения изотропно и

$$E_R = aT^4, \quad P_R = \frac{a}{3}T^4, \quad (13)$$

где $a = 4\sigma/c$ — постоянная плотности излучения.

2.2. Граничные условия. При радиальных пульсациях амплитуда смещения элемента газа быстро убывает по направлению к центру звезды и в качестве внутреннего граничного условия используется приближение жесткой постоянно излучающей сферы:

$$\frac{\partial r_0}{\partial t} = 0, \quad \frac{\partial L_0}{\partial t} = 0, \quad r_0 \lesssim 10^{-2}R, \quad (14)$$

где r_0 и L_0 — радиус и светимость на внутренней границе, R — радиус звезды. При задании радиуса внутренней границы r_0 желательно выполнить условие, чтобы температура в этом слое составляла $T \lesssim 10^6$ К. Тогда выделение энергии термоядерного синтеза во всех вышележащих слоях пульсирующей звезды пренебрежимо мало по сравнению с общей генерацией энергии, и это позволяет исключить ε из уравнения сохранения энергии (4).

Внешняя граница модели звезды определяется для слоя с радиусом R , за пределами которого остается пренебрежимая доля вещества. Таким образом, в уравнении сохранения импульса (3) можно положить, что полное давление на внешней границе $P = P_R$. Внешнее граничное условие для уравнения сохранения энергии может быть

сформулировано в приближении чернотельного излучения:

$$F_R = \frac{\sigma}{W} T^4, \quad (15)$$

где

$$W = \frac{1}{2} \left[1 - \sqrt{1 - (R_{ph}/R)^2} \right] \quad (16)$$

— коэффициент дилуции излучения, R_{ph} — радиус фотосферы. Удобство чернотельного приближения заключается в том, что оно связывает простой зависимостью болометрическую светимость звезды L с ее эффективной температурой T_{eff} , соответствующей непрозрачному слою (фотосфере), испускающему непрерывное излучение:

$$L = 4\pi R_{ph}^2 \sigma T_{\text{eff}}^4. \quad (17)$$

Для звезд с непротяженными атмосферами коэффициент дилуции $W \approx 1/2$.

Следует заметить, что предположение о чернотельном характере излучения, испускаемого внешними слоями звезды, не всегда оказывается достаточным, и тогда приходится обращаться к решению уравнения переноса (9) для внешних оптически тонких слоев, которое свивается с решением уравнения лучистой теплопроводности (11) в более глубоких слоях звезды.

3. Линейное адиабатическое волновое уравнение

Поскольку в течение одного пульсационного цикла звезда не может потерять заметное количество энергии, будем полагать, что каждый элемент газа внутри звезды как не теряет энергию, так и не поглощает ее извне. В приближении адиабатичности линеаризованное уравнение сохранения энергии сводится к двум простым соотношениям, связывающим вариации величин с их невозмущенными значениями:

$$\frac{\delta P}{P_0} = \Gamma_1 \frac{\delta \varrho}{\varrho_0}, \quad \frac{\delta T}{T_0} = (\Gamma_3 - 1) \frac{\delta \varrho}{\varrho_0}, \quad (18)$$

где

$$\Gamma_1 = \left(\frac{\partial \ln P}{\partial \ln \rho} \right)_S, \quad \Gamma_3 - 1 = \left(\frac{\partial \ln T}{\partial \ln \rho} \right)_S \quad (19)$$

— обобщенные показатели адиабаты, S — удельная энтропия вещества.

Для малого радиального смещения элемента газа $\delta r/r = \xi(r)e^{i\sigma t}$ с периодом Π и круговой частотой $\sigma = 2\pi/\Pi$ линеаризация уравнений

неразрывности (2) и движения (3) вместе с соотношениями (18) дает линейное адиабатическое волновое уравнение

$$\sigma_k^2 \xi_k = -\frac{1}{\rho r^4} \frac{d}{dr} \left(\Gamma_1 P r^4 \frac{d\xi_k}{dr} \right) - \frac{1}{\rho r} \left\{ \frac{d}{dr} \left[(3\Gamma_1 - 4) P \right] \right\} \xi_k, \quad (20)$$

где ξ_k — собственные функции, описывающие радиальную зависимость смещения с точностью до нормировочного множителя. Порядок моды $k \geq 0$ равен числу узлов собственной функции, и колебания с самым длинным периодом соответствуют фундаментальной моде $k = 0$. Собственные функции ξ_k при $k = 0, 1, 2$ показаны на рис. 1.

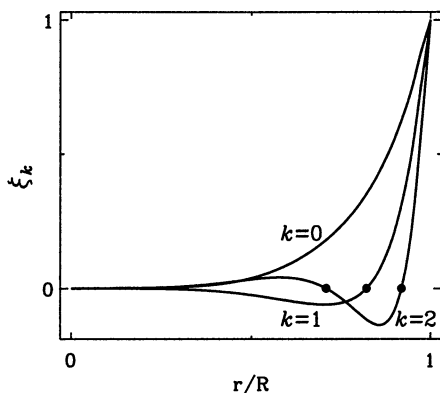


Рис. 1. Собственные функции ξ_k линейного адиабатического волнового уравнения (20) в функции расстояния от центра звезды r . Заполненные кружки указывают положение узлов собственных функций.

Собственные функции ξ_k линейного адиабатического волнового уравнения (20) описывают движения типа стоячей волны со свободной внешней границей, и для моды k -го порядка вдоль радиуса звезды R укладывается $k + 1/2$ длин волн. Существование стоячей волны предполагает идеальное отражение бегущей волны на обеих границах, и в звездах такое отражение происходит в слоях с резким изменением плотности газа на пространственном масштабе меньше длины волны λ :

$$4\pi H_P < \lambda, \quad (21)$$

где $H_P = -(d \ln P / dr)^{-1} = P / (\rho g)$ — локальная шкала высот по давлению, $g = GM_r / r^2$ — ускорение силы тяжести.

Благодаря большому температурному градиенту в окрестности центра звезды условие отражения (21) всегда выполняется при $r = 0$. Отражение пульсационной волны на внешней границе происходит в звездной атмосфере — оптически тонких слоях с приблизительно экспоненциальным убыванием плотности газа в зависимости от радиуса. Однако в звездах-гигантах с протяженными атмосферами шкала высот однородной атмосферы может быть сравнима с длиной волны, отражение пульсационной волны оказывается неполным и часть кинетической энергии пульсаций трансформируется в периодические ударные волны, распространяющиеся во внешних разреженных слоях звездной атмосферы.

4. Соотношение период — средняя плотность

Первое слагаемое в правой части уравнения (20) по абсолютной величине заведомо меньше второго, поэтому приближенно можно записать

$$\sigma^2 \xi = -\frac{3\bar{\Gamma}_1 - 4}{r\rho} \frac{dP}{dr} \xi, \quad (22)$$

где $\bar{\Gamma}_1$ — среднее значение обобщенного показателя адиабаты (для идеального одноатомного газа $4/3 < \Gamma_1 \leq 5/3$). С помощью уравнения гидростатического равновесия

$$\frac{dP}{dr} = -\rho \frac{GM}{R^2} \quad (23)$$

выражение (22) приобретает вид

$$\sigma^2 = (3\bar{\Gamma}_1 - 4) \frac{GM}{R^3}, \quad (24)$$

а окончательно преобразуется к соотношению, связывающему период радиальных колебаний Π со средней плотностью вещества $\bar{\rho}$:

$$\Pi = Q \sqrt{\bar{\rho}_\odot / \bar{\rho}} = Q (R/R_\odot)^{3/2} (M/M_\odot)^{-1/2}, \quad (25)$$

где $\bar{\rho}_\odot = 1,41 \text{ г/см}^3$; M_\odot , R_\odot — средняя плотность, масса и радиус Солнца. В силу исторических причин величина Q получила название пульсационной константы, хотя на самом деле ее значение является медленно меняющейся функцией массы M и радиуса R звезды и определяется радиальной зависимостью обобщенного показателя адиабаты Γ_1 . Для фундаментальной моды $0,03 \text{ сут} \leq Q \leq 0,1 \text{ сут}$.

Существующие в настоящее время методы астрономических наблюдений позволяют измерять радиусы пульсирующих звезд с погрешностью $\sim 10\%$, и при известном среднем значении радиуса звезды R подстановка полученной из теоретических расчетов зависимости $Q = Q(M, R)$ в соотношение период — средняя плотность (25) превращает последнее в алгебраическое уравнение относительно массы звезды M . В настоящее время этот метод является наиболее надежным способом определения масс одиночных звезд.

5. Физический механизм пульсационной неустойчивости

Оценки собственных частот радиальных мод, полученные из решения линейного адиабатического волнового уравнения (20), в пределах от одного до нескольких процентов совпадают с результатами более точной теории. Это указывает на достаточно малую роль неадиабатических эффектов в пульсирующих звездах, однако именно они приводят в конечном счете к затуханию или росту амплитуды пульсаций. Поскольку обмен энергией между отдельными элементами газа происходит посредством излучения, естественно предположить, что неустойчивость возникает при взаимодействии звездного вещества с излучением, распространяющимся от центра звезды к ее поверхности. Впервые эта гипотеза была выдвинута Эддингтоном [2], однако исчерпывающее объяснение явления звездных пульсаций как автоколебаний было дано позднее в работах Жевакина [3, 4]. Достаточно подробное изложение неадиабатической теории звездных пульсаций приводится в монографии Кокса [5], поэтому далее мы рассмотрим лишь физическую сущность явления.

Взаимодействие газа с проходящим через него излучением описывается росселандовым средним коэффициентом поглощения, приближенная зависимость которого от плотности и температуры выражается формулой

$$k_R \approx \kappa_0 \rho^n T^{-s}, \quad (26)$$

где $n \approx 0,8$; $s \approx 3,5$; κ_0 — постоянная величина, зависящая от химического состава звездного вещества. Предположим, что вследствие малого возмущения элемент газа был выведен из равновесия и в ходе замкнутого цикла совершил механическую работу $A = \oint PdV$. В нейтральном или полностью ионизованном газе работа адиабатического сжатия превращается главным образом в кинетическую энергию хаотического движения частиц, т. е. сопровождается увеличением температуры, и изменения коэффициента поглощения k_R связаны пре-

имущественно с вариациями температуры. Таким образом, при сжатии нейтральный или полностью ионизованный газ становится более прозрачным, что ведет к увеличению потерь тепла и затуханию колебаний, т. е. $A < 0$. Однако в частично ионизованном газе, обладающем более высокой теплоемкостью, существенная часть работы адиабатического сжатия расходуется на ионизацию оставшихся нейтральных атомов и основным фактором, определяющим изменение непрозрачности, является плотность газа ρ . При сжатии слой частично ионизованного газа становится более непрозрачным, что ведет к накоплению тепла и положительной механической работе $A > 0$, т. е. к усилению колебаний.

Каждый сферический слой звезды может рассматриваться как элементарная тепловая машина, способная совершать положительную (нарастание колебаний) или отрицательную (затухание колебаний) механическую работу за счет проходящего через этот слой излучения. Представляя звезду как совокупность большого числа тепловых машин, мы приходим к выводу, что звезда будет неустойчивой лишь в том случае, если суммарная работа всех слоев положительна.

Звездное вещество состоит из различных химических элементов, и по числу атомов относительное содержание водорода обычно составляет порядка 90%, а относительное содержание атомов гелия около 7%. Несмотря на преобладающее содержание атомов водорода, положительная механическая работа в зоне частичной ионизации водорода оказывается пренебрежимой по сравнению с отрицательным вкладом более глубоких слоев подавляющих колебаний. Это связано как с малой геометрической толщиной зоны частичной ионизации водорода, так и с низкой плотностью газа вблизи фотосферы, где эти слои находятся. Необходимыми для неустойчивости свойствами обладают зоны частичной ионизации нейтрального и однократно ионизованного гелия: $\text{He} \rightarrow \text{He}^+$ и $\text{He}^+ \rightarrow \text{He}^{++}$. На графике, представляющем механическую работу в зависимости от номера лагранжева слоя газодинамической модели, обе зоны частичной ионизации гелия обнаруживаются отдельными максимумами (рис. 2).

Термодинамические свойства частично ионизованного газа являются также причиной прекращения роста амплитуды и перехода к автоколебаниям. Действительно, возрастание непрозрачности вещества с усилением сжатия продолжается лишь до тех пор, пока в газе остаются нейтральные атомы. Как только все атомы переходят в ионизованное состояние, дальнейшее усиление сжатия ведет к уменьшению непрозрачности и возрастанию потерь тепла, т. е. к уменьшению механической работы. Таким образом, механизм возбуждения неустойчи-

ности достигает насыщения и происходит переход к автоколебаниям, когда суммарный вклад всех слоев, возбуждающих неустойчивость, уравновешивается отрицательным вкладом слоев, подавляющих колебания.

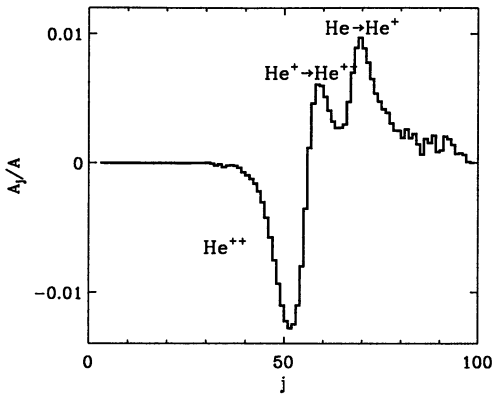


Рис. 2. Нормированная работа $A = \oint PdV$, совершаемая отдельным сферическим слоем в течение замкнутого цикла. По горизонтальной оси — номер лагранжева слоя j , где $j = 1$ на внутренней границе модели.

Тот факт, что одни звезды пульсируют в фундаментальной моде, тогда как другие — в первом или втором оберitone, находит простое объяснение во взаимном расположении зон ионизации относительно узлов собственных функций. Например, для неустойчивости первого обертона необходимо, чтобы зона частичной ионизации гелия находилась между поверхностью звезды и узлом, т. е. в слоях с радиусом $r > 0,8R$. В процессе пульсаций зона частичной ионизации смещается вдоль радиуса, и если, например, в ходе роста амплитуды она захватывает узел первого обертона, то происходит переход к колебаниям в фундаментальной моде. Неустойчивость звезд в радиальных модах порядка $k > 2$ встречается редко, поскольку слои частично ионизованного гелия должны находиться очень близко к поверхности звезды и в то же самое время их масса должна быть достаточно велика для того, чтобы затухание неустойчивости в более глубоких слоях не подавило колебания звезды в целом.

В последние годы благодаря значительному прогрессу в методах расчета уравнения состояния звездного вещества и созданию баз данных линий поглощения ионов в различных стадиях ионизации было

установлено, что в определенных случаях источником пульсационной неустойчивости звезд являются также слои частичной ионизации элементов группы железа в области температур $T \sim 2 \cdot 10^5$ К.

6. Нелинейные звездные пульсации

Исследование нелинейных звездных пульсаций заключается в решении задачи Коши для уравнений радиационной газовой динамики (2)—(4). В качестве исходной используется гидростатически равновесная модель звезды, структура которой однозначно определяется ее массой M , светимостью L , радиусом R и распределением содержаний химических элементов от центра до поверхности. Численное интегрирование уравнений газовой динамики проводится с помощью конечно-разностных методов второго порядка точности по лагранжевой координате M_r и времени t . Счетная лагранжева область $[M_1, M_N]$, простирающаяся от внутренней границы до поверхности, разбивается на $N - 1$ интервалов при $10^2 \lesssim N \lesssim 10^3$. Поскольку температура, давление и плотность наиболее быстро изменяются вблизи внешней границы, величина лагранжева интервала $\Delta M_{j-1/2} = M_j - M_{j-1}$ возрастает по направлению к центру по закону геометрической прогрессии.

Уравнения неразрывности (2) и движения (3) аппроксимируются явными разностными схемами, и в этом случае предельное значение шага интегрирования по времени Δt определяется условием устойчивости Куранта, т. е. наименьшим временем пробега звуковой волны между двумя соседними узлами разностной сетки:

$$\Delta t < \min_{2 \leq j \leq N} \frac{r_j - r_{j-1}}{a_{j-1/2}}, \quad (27)$$

где $a_{j-1/2}$ — скорость звука в интервале $[j - 1, j]$. Основной интерес в данной задаче представляют процессы со шкалой времени порядка пробега звуковой волны, поэтому ограничение на шаг интегрирования по времени, налагаемое условием устойчивости (27), оказывается приемлемым.

Уравнение лучистой теплопроводности (11) принадлежит к уравнениям параболического типа, и использование для его решения явной разностной схемы приводит к слишком жестким ограничениям, налагаемым условием устойчивости на величину шага по времени Δt . Поэтому решение уравнений (4) и (11) проводится с помощью неявной разностной схемы, причем предпочтительной является без-

условно устойчивая схема Кранка — Никольсона. Подробное описание метода расчета нелинейных звездных пульсаций приводится в работах [6, 7].

При задании начальных условий в качестве малого возмущения можно использовать поле скоростей с радиальной зависимостью, пропорциональной собственной функции какой-либо моды, полученной, например, из решения линейного адиабатического волнового уравнения (20). Однако малым начальным возмущением является также вычислительный шум, связанный с погрешностями аппроксимации разностных уравнений. В любом случае, независимо от вида начального возмущения, газодинамическая модель пульсирующей звезды демонстрирует экспоненциальный рост амплитуды с последующим переходом к автоколебаниям. Типичная зависимость изменения кинетической энергии

$$E_K = \frac{1}{2} \sum_1^N U_j^2 \Delta M_j, \quad (28)$$

j -м лагранжевом слое, приведена на рис. 3.

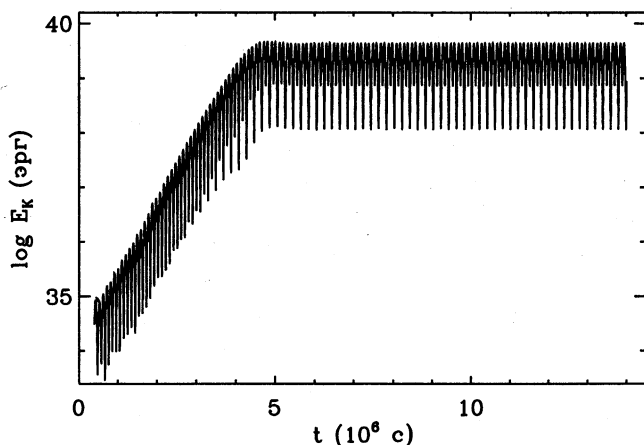


Рис. 3. Кинетическая энергия пульсирующей звезды E_K в зависимости от времени t .

При колебаниях в виде стоячей волны кинетическая энергия E_K обращается в нуль дважды за период: при максимальном сжатии и максимальном расширении звезды. Отсутствие нулевых значений кинетической энергии на рис. 3 связано с присутствием малой бегущей

составляющей, благодаря которой происходит перенос энергии пульсаций из области возбуждения неустойчивости в другие слои звезды.

От распределения плотности газа вдоль радиуса зависит не только период, но также и амплитуда радиальных пульсаций. Например, с увеличением радиуса R при фиксированной массе M плотность газа внутри звезды убывает и вследствие уменьшения скорости распространения звука возрастает длина периода радиальных пульсаций. Одновременно с этим происходит возрастание фазовых сдвигов между изменениями давления и плотности, что приводит к возрастанию по абсолютной величине механической работы $A = \oint PdV$, совершаемой в каждом лагранжевом слое. Эффекты неадиабатичности особенно велики вблизи поверхности, где находятся зоны частичной ионизации гелия. Поэтому в звездах с меньшей средней плотностью вещества рост неустойчивости происходит обычно быстрее, а режим автоколебаний устанавливается при большей амплитуде.

Нелинейность звездных пульсаций наиболее велика во внешних оптически тонких слоях, где она приводит к заметным дополнительным потерям энергии колебаний. При расширении внешних слоев происходит заметное уменьшение действующей на них силы тяжести, и по истечении периода колебаний эти слои не успевают вернуться в исходную точку. В области столкновения внутренних расширяющихся слоев газа и продолжающих падать на них внешних слоев возникает ударная волна, которая в лагранжевой системе координат распространяется в сторону внешней границы звезды. Спектроскопические наблюдения свидетельствуют, что ударные волны в протяженных атмосферах пульсирующих звезд-гигантов действительно возникают на начальной стадии расширения фотосферных слоев, а продолжительность существования ударной волны составляет от нескольких десятых долей до приблизительно целого периода пульсаций.

Скорость течения газа за ударным разрывом обычно не превосходит локальной скорости убегания, поэтому одиночная ударная волна не способна привести к выбросу вещества из сферы гравитационного влияния звезды. Однако вследствие периодически распространяющихся в звездной атмосфере ударных волн происходит увеличение среднего радиуса лагранжевых слоев, и осредненное за цикл колебаний распределение плотности газа вдоль радиуса описывается уже не экспоненциальным законом, как в случае гидростатически равновесной атмосферы, а зависимостью вида $\rho \propto r^{-2}$.

Понимание перестройки структуры атмосферы пульсирующей звезды вследствие периодических ударных волн позволило найти естественное объяснение существованию околозвездной пыли, наблюдаемой вокруг красных гигантов — звезд с массой, приблизительно равной солнечной, эффективной температурой $T_{\text{эф}} \approx 3000 \text{ K}$ и радиусами, в сотни раз превосходящими солнечный. Действительно, в гидростатически равновесной звездной атмосфере плотность газа убывает с удалением от звезды столь быстро, что ни одно молекулярное соединение не может перейти в твердую фазу, так как его парциальное давление оказывается много меньше давления насыщенного пара. Однако вследствие периодически распространяющихся в звездной атмосфере ударных волн плотность газа возрастает на несколько порядков величины и на расстоянии, в три — пять раз превышающем радиус фотосферы, происходит фазовый переход некоторых молекулярных соединений.

По сравнению с газом пылевые частицы обладают значительно более высокой поглощательной способностью, и сила лучистого давления, действующая на пыль, оказывается больше силы тяжести. Вследствие дрейфа пылевых частиц сквозь газ молекулам и атомам газовой составляющей передается часть импульса от пылевых частиц, которые увлекают за собой газ. Таким образом, звездный ветер красных гигантов связан с взаимодействием двух физических механизмов: периодических ударных волн и давлением излучения на пылевые частицы. Наблюдательные оценки скорости потери массы красных гигантов составляют $\dot{M} \lesssim 10^{-6} M_{\odot}/\text{год}$, поэтому по истечении нескольких сотен тысяч лет звезда теряет около половины своей массы и превращается в остывающий белый карлик.

Результаты газодинамических расчетов, приведенные на рис. 4, наглядно иллюстрируют эффект потери массы в результате взаимодействия периодических ударных волн и давления излучения на пыль с последующей передачей импульса газовой составляющей звездной атмосферы. Следует, однако, заметить, что в подобных расчетах не учитывается детальная структура ударной волны и не рассматривается кинетика фазового перехода, поэтому представленные на рис. 4 результаты следует рассматривать лишь как качественную модель звездного ветра нелинейно пульсирующих красных гигантов.

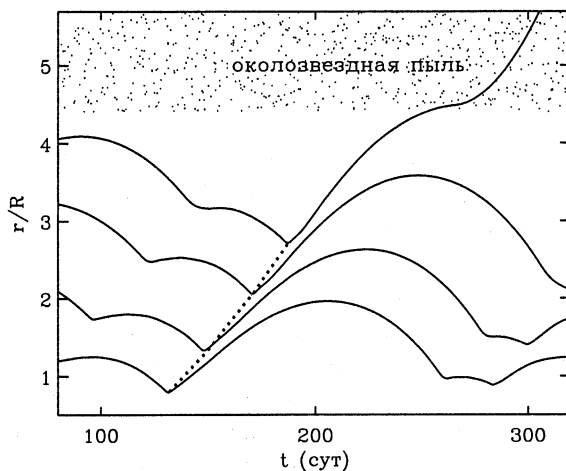


Рис. 4. Радиусы отдельных лагранжевых слоев газодинамической модели нелинейно пульсирующего красного гиганта (сплошные линии) в единицах равновесного радиуса фотосферы звезды. Траектория ударной волны показана пунктирной линией.

7. Заключение

Используемые при расчетах звездных пульсаций газовые характеристики основываются на данных уравнения состояния и росселандова среднего коэффициента непрозрачности для смеси ионов различных химических элементов. В последние полтора десятилетия благодаря замечательным успехам в области атомной физики были рассчитаны обширные таблицы термодинамических функций и коэффициента непрозрачности, что привело к многим успешным приложениям теории звездных пульсаций. Вместе с тем продолжает существовать нерешенная проблема пульсационной неустойчивости красных гигантов. Внешние слои этих звезд конвективно неустойчивы, и до сих пор остается неясным, каким именно образом следует описывать взаимодействие звездных пульсаций и турбулентной конвекции. В некоторых работах конвекция рассматривалась в рамках теории длины перемешивания, однако результаты этих расчетов так и не позволили объяснить наблюдаемые особенности пульсирующих красных гигантов. Предварительные результаты трехмерного численного моделирования крупномасштабной конвекции в красных гигантах заставляют предпо-

лагать как неприменимость приближения сферической симметрии при исследованиях нелинейных пульсаций этих звезд, так и необходимость учета взаимодействия магнитного поля с конвективными элементами.

Литература

1. *Mihalas D., Weibel Mihalas B.* Foundations of radiation hydrodynamics. N. Y.: Oxford University Press, 1984. 717 p.
2. *Eddington A. S.* The internal constitution of the stars. Cambridge: Cambridge University Press, 1926. 422 p.
3. *Жевакин С. А.* К теории цефеид. I // Астрон. журн. 1953. Т. 30, № 2. С. 161—179.
4. *Жевакин С. А.* К теории звездной переменности. II // Астрон. журн. 1954. Т. 31, № 2. С. 141—153.
5. *Кокс Дж. П.* Теория звездных пульсаций. М.: Мир, 1983. 328 с.
6. *Christy R. F.* Computational methods in stellar pulsation // Methods in computational physics. N. Y.: Academic Press, 1967. V. 7. P. 191—218.
7. *Фадеев Ю. А.* Методы расчета нелинейных колебаний звезд // Научн. информ. Астрон. совета АН СССР. 1983. Вып. 52. С. 3—28.

УДАРНЫЕ ВОЛНЫ В АТМОСФЕРАХ ПУЛЬСИРУЮЩИХ ЗВЕЗД

Ю. А. Фадеев

1. Введение

При нелинейных пульсациях в течение каждого цикла колебаний в звездной атмосфере возникает ударная волна. Основные наблюдаемые признаки ударных волн в пульсирующих звездах — это интенсивные эмиссионные линии бальмеровской серии водорода и доплеровское расщепление линий поглощения некоторых атомов. Бальмеровскую эмиссию связывают с радиативным охлаждением газа в области релаксации за ударным скачком, а величина доплеровского расщепления свидетельствует об амплитуде разрыва скоростей, измеряемой несколькими десятками километров в секунду. Ударные волны, возникающие при нелинейных звездных пульсациях, являются источником значительных потерь энергии колебаний, а вследствие их периодического возникновения в звездной атмосфере скорость движения самых внешних слоев газа может достигать скорости убегания. Несмотря на столь важную роль ударных волн, все теоретические исследования нелинейных звездных пульсаций ограничиваются использованием метода искусственной вязкости, который предназначен лишь для подавления резких колебаний сеточных функций за ударным фронтом. Это связано с тем, что рассмотрение релаксационных процессов с масштабом меньше размера лагранжевых интервалов, используемых в газодинамических расчетах, практически невозможно. Попытки использования адаптивных сеток с целью выделения ударных волн в расчетах звездных пульсаций также не привели к положительным результатам.

Однако возможность детального анализа структуры ударных волн в звездных атмосферах все же существует. Протяженность ударной волны ($\mathcal{L} \lesssim 10^5$ см) на несколько порядков величины меньше шкалы высот однородной атмосферы, поэтому за время пробега элемента газа сквозь ударную волну характеристики окружающей среды остаются практически неизменными. Таким образом, структура ударной волны может рассматриваться в приближении стационарности и плоской геометрии (подробное обсуждение применимости приближения стационарности для ударных волн приводится в монографии Зельдовича и Райзера [1]), а уравнения газовой динамики и атомной кинетики сводятся к системе обыкновенных дифференциальных уравнений. Очевидные преимущества такого подхода связаны также с тем,

что исходные уравнения являются жесткими, а наиболее эффективные методы интегрирования жестких систем разработаны для обыкновенных дифференциальных уравнений.

Ниже обсуждаются свойства стационарных ударных волн, распространяющихся в частично ионизованном водородном газе, характеристики которого соответствуют физическим условиям в атмосферах нелинейно пульсирующих звезд-гигантов. За недостатком места приводятся лишь наиболее важные уравнения, и для более детального ознакомления читателю следует обратиться к работам [2—8].

2. Модель стационарной ударной волны

Рассмотрим однородный атомарный водородный газ с плотностью ρ_1 и температурой T_1 , в котором с постоянной скоростью U_1 движется ударная волна. Задача заключается в том, чтобы для физических условий, определяемых значениями ρ_1 , T_1 , U_1 , установить структуру ударной волны (т. е. распределение физических величин вдоль пространственной координаты) и определить спектральные особенности испускаемого волной излучения.

Ударное сжатие и резкое торможение газодинамического потока происходят в пределах вязкого скачка уплотнения, ширина которого измеряется длиной пробега атомов газа. Этот масштаб значительно меньше протяженности ударной волны, и в данной задаче вязкий скачок может рассматриваться как бесконечно тонкий слой, в пределах которого происходит скачкообразное изменение газодинамических переменных. Совместим начало сопутствующей системы координат $x = 0$ с разрывом, а через x_1 и x_n обозначим координаты передней и задней границ модели ударной волны. В общем случае координаты границ модели заранее не известны и должны определяться из пробных расчетов таким образом, чтобы искомое решение не зависело от этих значений.

Уравнения газовой динамики. Для установившегося плоскопараллельного течения газа уравнение неразрывности сводится к соотношению

$$U\rho = \dot{m} \equiv \text{const}, \quad (1)$$

где U — скорость течения газа в системе координат, связанной с волной, ρ — плотность газа. Предполагается, что силы, действующие на элемент газа, обусловлены градиентом газового давления P_g и лучистым давлением, поэтому уравнение движения имеет вид

$$\frac{dU}{dt} = -\frac{1}{\dot{m}} \frac{dP_g}{dt} + \frac{1}{c\rho} \int_0^\infty \chi_\nu F_\nu d\nu, \quad (2)$$

где c — скорость света, ν — частота излучения, F_ν — монохроматический поток излучения, χ_ν — монохроматический коэффициент ослабления излучения (т. е. сумма коэффициентов поглощения и рассеяния).

Уравнение баланса энергии записывается отдельно для электронного газа с температурой T_e и для тяжелых частиц (нейтральные атомы и ионы водорода) с температурой T_a :

$$\frac{3}{2} \frac{d}{dt} \left(\frac{n_e}{\rho} k T_e \right) = -n_e k T_e \frac{d}{dt} \left(\frac{1}{\rho} \right) + Q_{elc} + Q_{inc} - \frac{d}{dt} \left(\frac{E_R}{\rho} \right), \quad (3)$$

$$\frac{3}{2} \frac{d}{dt} \left(\frac{n_a}{\rho} k T_a \right) = -n_a k T_a \frac{d}{dt} \left(\frac{1}{\rho} \right) - Q_{elc}, \quad (4)$$

где n_e и n_a — концентрации электронов и тяжелых частиц, E_R — плотность лучистой энергии, Q_{elc} — скорость передачи энергии электронам при упругих столкновениях с ионами и нейтральными атомами. Скорость приобретения энергии электронами при неупругих столкновениях

$$Q_{inc} = -\frac{dE_{ex}}{dt} - \frac{dE_I}{dt} - \frac{1}{\rho} \nabla \cdot \mathbf{F}_R - \frac{1}{\rho} \nabla \cdot \mathbf{F}_e, \quad (5)$$

где E_{ex} — энергия возбуждения атомных уровней, E_I — энергия ионизации, \mathbf{F}_R — поток излучения, \mathbf{F}_e — поток электронной теплопроводности. В силу приближения плоской геометрии векторное представление потоков может быть заменено скалярным, и в соответствии с принятой системой координат положительное значение потока означает направление по ходу течения газа и, следовательно, противоположное направлению движения ударной волны.

Уравнения атомной кинетики. Для атома водорода с L уровнями кинетические уравнения, описывающие изменение населенностей атомных уровней n_i и концентрации свободных электронов n_e , имеют вид

$$\begin{aligned} \frac{dn_i}{dt} = & \sum_{j=1}^{i-1} n_j (n_e C_{ji} + R_{ji}) + \sum_{j=i+1}^L n_j \frac{n_i^*}{n_j^*} (n_e C_{ij} + R_{ij}^\dagger) + \\ & + n_e \frac{n_i^*}{n_e^*} (n_e C_{ik} + R_{ik}^\dagger) - \\ & - n_i \sum_{j=1}^{i-1} \frac{n_j^*}{n_i^*} (n_e C_{ji} + R_{ji}^\dagger) - n_i \sum_{j=i+1}^L (n_e C_{ij} + R_{ij}) - \\ & - n_i (n_e C_{ik} + R_{ik}), \quad (i = 1, \dots, L), \end{aligned} \quad (6)$$

$$\frac{dn_e}{dt} = \sum_{i=1}^L n_i (n_e C_{ik} + R_{ik}) - n_e \sum_{i=1}^L \frac{n_i^*}{n_e^*} (n_e C_{ik} + R_{ik}^\dagger), \quad (7)$$

где C_{ij} и C_{ik} — скорости возбуждения при переходе на верхний уровень $i \rightarrow j$ и ионизации с i -го уровня электронным ударом; R_{ij} и R_{ik} — скорости фотовозбуждения и фотоионизации; n_i^* и n_e^* — концентрации возбужденных атомов и свободных электронов, рассчитываемые по формуле Больцмана — Саха; $n_j(n_i^*/n_j^*)R_{ij}^\dagger$ — скорость радиативных девозбуждений; $n_e(n_i^*/n_e^*)R_{ik}^\dagger$ — скорость фоторекомбинации.

Уравнение переноса излучения. В плоской геометрии и при $U \ll c$ уравнение переноса излучения имеет вид

$$\mu \frac{d}{d\tau_\nu} I_\nu(x, \mu) = S_\nu(x) - I_\nu(x, \mu), \quad (8)$$

где $I_\nu(x, \mu)$ — интенсивность излучения; $\mu = \cos \theta$ — направляющий косинус; θ — угол между нормалью к плоскости ударной волны и направлением распространения излучения; τ_ν — оптическая глубина на частоте ν , определяемая как $d\tau_\nu = \chi_\nu dx$; $S_\nu(x) = \eta_\nu / \chi_\nu$ — функция источников; η_ν — коэффициент излучения. Наиболее эффективный метод решения уравнения (8) заключается в рассмотрении переноса излучения как краевой задачи (подробнее см. [9]), и для модели ударной волны граничные условия формулируются в предположении чернотельного изотропного излучения со стороны окружающей среды на обеих границах модели:

$$I_\nu(\mu > 0, x_1) = I_\nu(\mu < 0, x_n) = B_\nu(T_1), \quad (9)$$

где $B_\nu(T)$ — функция Планка.

Самосогласованное решение уравнений. Решение уравнений газовой динамики (2), (3) и атомной кинетики (6), (7) представляет собой задачу Коши для системы жестких обыкновенных дифференциальных уравнений, интегрирование которых должно проводиться с применением специально разработанных методов (см., например, [10]). Вычисление характеристик поля излучения предполагает решение краевой задачи (8) для всей ударной волны, поэтому самосогласованное решение всех уравнений может быть получено с помощью итерационной процедуры.

При заданных характеристиках поля излучения (т. е. интенсивности излучения, известной в функции пространственной координаты x , направляющего косинуса μ и частоты излучения ν) интегрирование

системы обыкновенных дифференциальных уравнений (2)—(4), (6), (7) проводится на отрезке от передней границы до точки разрыва $[x_1, 0]$. В самой точке разрыва скачкообразное изменение газодинамических величин рассчитывается с помощью соотношений Гюгонио, учитывающих вклад поля излучения [11]:

$$\begin{cases} \rho U = \dot{m} = C_0, \\ \dot{m} + P_g + P_R = C_1, \\ \frac{1}{2} \dot{m} U^2 + \dot{m} h + F_R + U(E_R + P_R) = C_2, \end{cases} \quad (10)$$

где h — удельная энтальпия. Величины C_0 , C_1 и C_2 в правой части (10) представляют собой потоки массы, импульса и энергии, которые сохраняют постоянное значение на всем интервале $[x_1, x_n]$. Затем интегрирование дифференциальных уравнений продолжается на отрезке от точки разрыва до задней границы модели $[0, x_n]$. Полученный вектор решения

$$y(x) = \{U, T_e, T_a, n_i, \dots, n_L, n_e\}^\dagger \quad (11)$$

(здесь \dagger — знак транспозиции) используется для вычисления оптической глубины $\tau_\nu(x)$ и функции источников $S_\nu(x)$, которые необходимы для решения уравнения переноса излучения. В начальной итерации поле излучения может быть задано в приближении термодинамического равновесия с планковской функцией распределения, соответствующей достаточно произвольному пространственному распределению электронной температуры.

3. Структура ударной волны

В ударной волне наиболее быстрые процессы протекают в окрестности вязкого скачка уплотнения, поэтому при графическом представлении независимую переменную удобно выражать в логарифмическом масштабе, а пространственную зависимость каждой величины изображать в виде двух отдельных графиков для $x < 0$ и $x > 0$ соответственно. Приведенные на рис. 1 температура, степень ионизации и поток излучения в функции пространственной координаты иллюстрируют главную особенность радиативных ударных волн, заключающуюся в существовании тесной взаимосвязи между газодинамическим потоком и полем излучения, порождаемым этим потоком.

В лаймановском континууме (т. е. на длинах волн $\lambda < 912\text{\AA}$, соответствующих фотоионизации с основного уровня) частично ионизован-

ный водородный газ значительно менее прозрачен по сравнению с другими участками спектра и коротковолновое излучение ударной волны оказывается запертым в пределах области, ограниченной несколькими пробегами лаймановских фотонов. Излучение лаймановского континуума формируется позади разрыва, однако зона более высокой плотности лучистой энергии расположена приблизительно симметрично относительно разрыва, и та часть ее, которая простирается перед разрывом, называется радиативным прекурсором (или зоной лучистого прогрева). Таким образом, оказываясь в прекурсоре, приближающийся к разрыву элемент газа попадает в сферу влияния процессов, протекающих за разрывом.

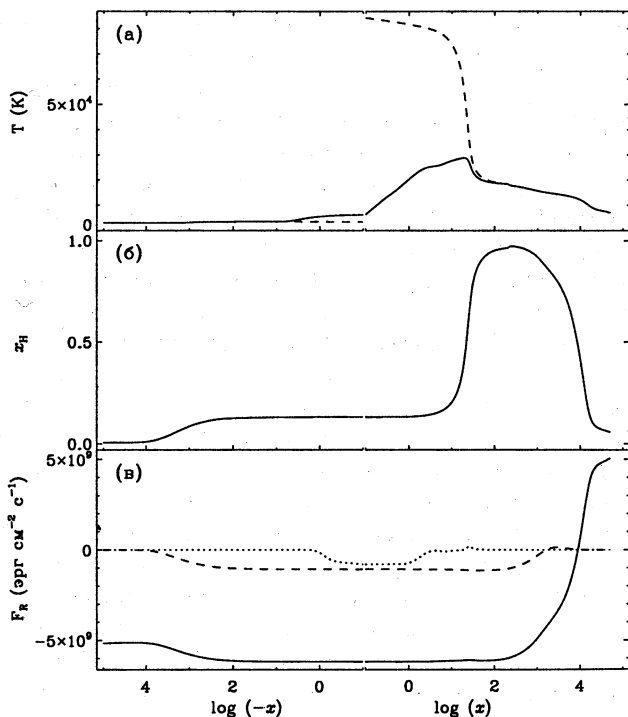


Рис. 1. Структура ударной волны при $\rho_1 = 10^{-10} \text{ г/см}^3$, $T_1 = 3000 \text{ К}$, $U_1 = 60 \text{ км/с}$: а — электронная температура T_e (сплошная линия) и температура тяжелых частиц T_a (штриховая линия); б — степень ионизации водорода; в — интегральный поток излучения F_R (сплошная линия), поток излучения в лаймановском континууме F_{Ly} (штриховая линия) и поток электронной теплопроводности F_e (пунктирная линия).

Наиболее существенными процессами в зоне радиативного прекурсора являются фотоионизация атомов водорода с основного уровня и нагревание газа вследствие поглощения лаймановских фотонов. С увеличением скорости ударной волны возрастает плотность лучистой энергии в лаймановском континууме и, соответственно, возрастают степень ионизации водорода и температура газа перед вязким скачком уплотнения. При температуре невозмущенного газа $T_1 = 3000$ К равновесная степень ионизации водорода пренебрежимо мала ($x_H \sim 10^{-8}$), и концентрация электронов перед вязким скачком является функцией скорости ударной волны. Состояние полной ионизации водородного газа достигается при $U_1 > 90$ км/с (т. е. при числах Маха $M_1 > 14$).

Основная доля кинетической энергии газа, набегающего на разрыв, вследствие эффектов вязкости переходит в энергию хаотического поступательного движения частиц. По порядку величины скачкообразное возрастание температуры составляет $\Delta T \sim U_1^2 m/k$, где m — масса частиц. Таким образом, на ударном разрыве значительно возрастает температура нейтральных атомов и ионов водорода, тогда как приращение электронной температуры за счет эффектов вязкости в $m_H/m_e \approx 1,8 \cdot 10^3$ раз меньше, т. е. пренебрежимо мало. Однако электронный газ жестко связан с ионами электрическими силами, поэтому его сжатие происходит адиабатически и значения электронной температуры по обе стороны от разрыва связаны соотношением $T_e^+/T_e^- = (\rho^+/\rho^-)^{\gamma-1}$, где $\gamma = 5/3$, а верхние индексы минус и плюс соответствуют положениям по обе стороны от поверхности разрыва.

В водородном газе при равных температурах электронного газа и тяжелых частиц длины свободного пробега электронов и ионов оказываются одного порядка, однако скорость электронов и, следовательно, коэффициент теплопроводности, в $\sqrt{m_H/m_e}$ раз больше. Поэтому в окрестности разрыва возникает поток электронной теплопроводности, обеспечивающий перенос тепла в направлении, противоположном течению газа. На рис. 1 область электронной теплопроводности обнаруживается ненулевыми значениями потока $F_e = -\kappa_e \nabla T_e$ (здесь κ_e — коэффициент электронной теплопроводности). Перед ударным разрывом температура электронного газа превосходит температуру атомов и в зоне прогрева электронной теплопроводности вследствие упругих столкновений происходит обмен энергией между тяжелыми частицами и электронами.

Сразу за ударным разрывом газ оказывается в неравновесном состоянии, и установление в нем термодинамического равновесия происходит в виде последовательности релаксационных процессов, харак-

теризующихся наиболее существенным вкладом в теплоемкость газа. Для водородного газа такими процессами являются выравнивание поступательных степеней свободы (т. е. выравнивание температур тяжелых частиц и электронов), а также ионизация и рекомбинация атомов водорода. Вклад эффектов возбуждения атомных уровней в теплоемкость газа незначителен, однако от заселенностей атомных уровней зависят как скорости ионизации и рекомбинации, так и поток излучения, формирующийся в результате связанно-связанных и связанно-свободных переходов.

Быстрее всего в газе происходит перераспределение энергии между поступательными степенями свободы, поэтому сразу за разрывом вследствие значительного различия температур тяжелых частиц и электронного газа начинается процесс передачи тепловой энергии к электронам от нейтральных атомов и ионов водорода. При относительно малой скорости ударной волны ($U_1 < 40$ км/с, $M_1 < 6$) степень ионизации водорода в зоне радиативного прекурсора $x_H < 10^{-2}$ и обмен энергией происходит в результате упругих столкновений электронов с нейтральными атомами водорода. При более высокой скорости ударной волны увеличивается степень ионизации и начинают играть роль упругие столкновения электронов с ионами водорода. Сечение этих столкновений примерно в 10^3 раз больше, чем сечение столкновений электронов с нейтральными атомами, поэтому процесс выравнивания температур значительно ускоряется с увеличением скорости ударной волны. Таким образом, физические условия за ударным разрывом, в которых происходит выравнивание поступательных степеней свободы, оказываются тесно связанными с состоянием газа в зоне лучистого прогрева перед ударным разрывом.

Ионизация атомов водорода за вязким скачком уплотнения происходит за счет столкновений с электронами. Для ионизации электронным ударом требуется большое число соударений, поэтому рост степени ионизации за ударным разрывом происходит с заметным запаздыванием относительно изменений температуры газа. Последующая фоторекомбинация водорода протекает также с запаздыванием относительно убывающей температуры газа. Как видно из рис. 1, именно в области рекомбинации происходит формирование излучения ударной волны, испускаемого ею в равной мере как по ходу распространения, так и в противоположном направлении.

Слои газа, в окрестности которых формируется интегральный по всему спектру поток излучения, определяются условием $F_R = 0$, и для модели, представленной на рис. 1, координата этих слоев равна $x \approx 9 \cdot 10^3$ см. Однако в зависимости от частоты излучения опти-

ческая толщина ударной волны τ_ν претерпевает значительные изменения и координата нулевого монохроматического потока излучения изменяется в значительных пределах. Поскольку оптическая глубина особенно велика в лаймановском континууме, область формирования лаймановского излучения оказывается заметно ближе к ударному разрыву (для модели на рис. 1 $x \approx 1,7 \cdot 10^3$ см). Таким образом, в области $1,7 \cdot 10^3 \text{ см} < x < 9 \cdot 10^3 \text{ см}$ лаймановский поток излучения направлен против движения ударной волны (т. е. $F_{Ly} > 0$), тогда как фотоны меньшей частоты распространяются в противоположном направлении. Однако, как видно из рис. 1, вклад противоположно направленных лаймановских фотонов достаточно мал.

При адиабатическом сжатии плотность газа возрастает не беспрельдно, а стремится к значению ρ^* , определяемому соотношением

$$\frac{\rho^*}{\rho_1} = \frac{\gamma + 1}{\gamma - 1}, \quad (12)$$

и для одноатомного газа предельное сжатие $\rho^*/\rho_1 = 4$. В радиативных ударных волнах с числом Маха $M_1 < 15$ коэффициент сжатия на ударном разрыве $\rho^+/\rho^- < 4$, тогда как конечное сжатие на хвосте волны $10 \lesssim \rho^*/\rho_1 \lesssim 10^2$. Возрастание плотности газа за ударным разрывом (и, соответственно, замедление скорости течения газа U) происходит вследствие значительного запаздывания ионизации и рекомбинации относительно изменений температуры газа.

Наибольший интерес в структуре радиативной ударной волны представляет превращение кинетической энергии направленного газодинамического потока в излучение. Последовательность процессов, в результате которых это происходит, можно наглядно проследить, рассматривая отдельные слагаемые третьего соотношения Гюгонио (10), выражающего закон сохранения энергии для плоскопараллельного стационарного потока газа. На рис. 2 эти слагаемые приведены в функции пространственной координаты x и представляют собой поток кинетической энергии F_K , поток энтальпии F_h и поток лучистой энергии $F_R + U(E_R + P_R)$. Сумма этих величин является постоянной величиной в пределах всей ударной волны.

Возрастание потока энтальпии F_h в зоне радиативного прекурсора связано с нагреванием газа вследствие поглощения лаймановских фотонов и фотоионизации атомов водорода с основного уровня. Величина этих изменений F_h приблизительно равна соответствующим изменениям лаймановского потока излучения F_{Ly} . Отсутствие точного равенства в изменениях потоков F_h и F_{Ly} обусловлено малыми изменениями плотности газа, сопровождающими рост температуры,

и является причиной незначительного замедления скорости газа, приближающегося к разрыву. Например, при скорости ударной волны $U_1 = 70$ км/с ($M_1 = 10,9$) замедление потока газа составляет ≈ 1 км/с.

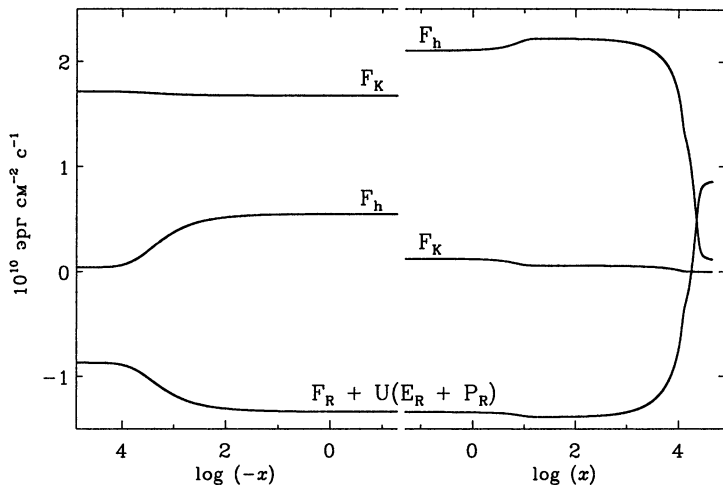


Рис. 2. Поток кинетической энергии F_K , поток энтальпии F_h и поток лучистой энергии $F_R + U(E_R + P_R)$ в ударной волне со скоростью $U_1 = 70$ км/с при $\rho_1 = 10^{-10}$ г/см³, $T_1 = 3000$ К.

На ударном разрыве поток лучистой энергии остается постоянным и скачкообразное возрастание температуры атомов и ионов является причиной равных по величине (но обратных по знаку) изменений потока кинетической энергии F_K и потока энтальпии F_h . Увеличение потока энтальпии за разрывом связано с эффектами возбуждения и ионизации атомов, а на хвосте ударной волны поток энтальпии убывает вследствие радиативного охлаждения газа и фоторекомбинации атомов водорода. Именно в этих слоях происходит формирование основной доли излучения, испускаемого ударной волной.

4. Бальмеровский спектр излучения ударной волны

В системе отсчета наблюдателя уравнение переноса излучения для движущейся среды имеет тот же вид, что и для неподвижной среды (8), однако вследствие доплеровских сдвигов по частоте появляется дополнительная зависимость коэффициентов поглощения и из-

лучения от угловой переменной μ [11]. При скорости ударной волны $U_1 \leq 10^2$ км/с эффекты движущейся среды заметны лишь в частотах спектральных линий, где изменения интенсивности излучения с частотой наиболее велики, тогда как для участков непрерывного спектра доплеровским сдвигом по частоте можно пренебречь. Таким образом, при решении уравнения переноса, учитывающего эффекты движущейся среды, достаточно ограничиться рассмотрением относительно узких участков спектра, в пределах которых находятся представляющие интерес спектральные линии.

Коэффициент поглощения в спектральной линии, связанный с переходом электрона с нижнего уровня i на верхний уровень j , определяется выражением

$$\kappa_{ij}(\nu, \mu) = n_i B_{ij} \frac{h\nu_{ij}}{4\pi} \left(1 - \frac{n_j g_i}{n_i g_j} \right) \phi_\nu(\mu), \quad (13)$$

где B_{ij} — эйнштейновский коэффициент вероятности поглощения при переходе $i \rightarrow j$; ν_{ij} — частота излучения, связанная с этим переходом; g_i и g_j — статистические веса уровней; $\phi_\nu(\mu)$ — профиль поглощения в линии. Следует заметить, что выражение (13) справедливо лишь при полном перераспределении по частотам, т. е. при совпадении профилей поглощения и излучения (подробнее см. [11]). При доплеровском (т. е. связанном с хаотическим тепловым движением атомов) уширении профиль спектральной линии определяется выражением

$$\phi_\nu(\mu) = \frac{1}{\Delta\nu_D \sqrt{\pi}} \exp \left[-\frac{(\mu V \nu_0 / c + \nu - \nu_0)^2}{\Delta\nu_D^2} \right], \quad (14)$$

где V — скорость течения газа в системе отсчета наблюдателя,

$$\Delta\nu_D = \frac{\nu_0}{c} \left(\frac{2kT_a}{m_H} \right)^{1/2} \quad (15)$$

— доплеровская ширина линии, ν_0 — частота центра линии, m_H — масса атома водорода.

При наблюдениях звезды ударная волна движется по направлению к наблюдателю и скорость течения газа в системе отсчета, связанной с наблюдателем, определяется соотношениями

$$\begin{cases} V = 2U^- - U_1 - U, & (x < 0), \\ V = 2U^+ - U_1 - U, & (x > 0), \end{cases} \quad (16)$$

где U^- и U^+ — скорости течения газа по обе стороны поверхности разрыва в системе отсчета ударной волны. Таким образом, при астрономических измерениях для неподвижного газа $V = 0$, тогда как

скорость течения газа, увлекаемого ударной волной, $V < 0$, что соответствует доплеровскому сдвигу в коротковолновую сторону спектра. Далее с помощью соотношения $\delta\nu = \nu_0 V/c$ величину доплеровского сдвига будем выражать в километрах в секунду.

Формирование основной доли излучения в линиях бальмеровской серии происходит в зоне рекомбинации водорода, причем переходы с уровней $j \geq 3$ на второй уровень обусловлены как столкновительными, так и радиативными переходами. Скорость радиативного девозбуждения наиболее велика для перехода $3 \rightarrow 2$, и действительно, в пульсирующих красных гигантах наиболее интенсивной эмиссионной линией, возникающей при распространении в звездной атмосфере ударной волны, является линия $H\alpha$. На рис. 3 приведены профили этой линии, полученные из расчетов структуры ударной волны для двух случаев: с решением уравнения переноса для неподвижной среды и с учетом эффектов движущейся среды. Для удобства сравнения с результатами спектроскопических измерений монохроматический поток излучения F_ν приведен с обратным знаком.

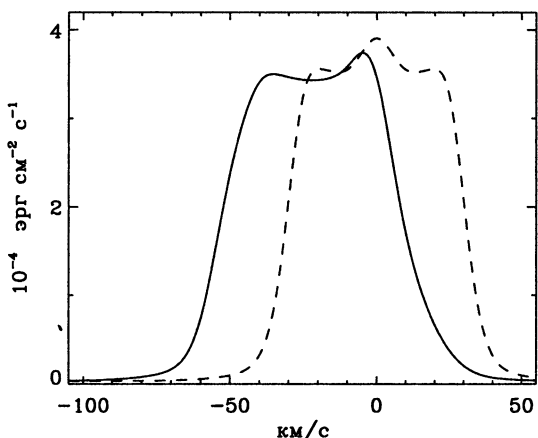


Рис. 3. Профили эмиссионной линии $H\alpha$, рассчитанные для движущейся (сплошная линия) и неподвижной (штриховая линия) среды при $\rho_1 = 10^{-10}$ г/см³, $T_1 = 3000$ К, $U_1 = 60$ км/с.

Как видно из рис. 3, монохроматический поток излучения в частотах линии $H\alpha$ характеризуется сложной структурой, для объяснения которой удобнее обратиться к решению уравнения переноса без учета доплеровских сдвигов. Основной поток излучения в линии $H\alpha$ формируется за ударным разрывом в зоне рекомбинации, и значительная ширина профиля обусловлена высокой температурой излучающего га-

за. Холодный газ, находящийся между наблюдателем и ударным разрывом, поглощает часть излучения, поэтому на широкий эмиссионный профиль накладывается узкий профиль поглощения. Наконец, в зоне радиативного прекурсора вблизи разрыва начинающийся рост ионизации электронным ударом сопровождается радиативными девозбуждениями с переходом электронов на второй уровень. Образующаяся в результате радиативных переходов $3 \rightarrow 2$ эмиссионная компонента частично заполняет профиль поглощения, формирующийся в холодном газе перед ударным разрывом. Интенсивность узкой эмиссионной компоненты чувствительна к плотности невозмущенного газа и становится практически необнаружимой при $\rho_1 < 10^{-11}$ г/см³.

При учете эффектов движущейся среды в коротковолновую сторону сдвигается только широкий эмиссионный профиль, возникающий за ударным разрывом, тогда как остающийся несмещенным узкий эмиссионный компонент является индикатором скорости невозмущенного газа. Таким образом, теоретические расчеты структуры ударных волн позволили не только объяснить происхождение узкого эмиссионного компонента в профиле эмиссионной линии $\text{H}\alpha$ красных гигантов, но также предоставили возможность определения скорости течения газа непосредственно перед ударным разрывом в системе координат, связанной со звездой. В большинстве случаев невозмущенный газ движется от наблюдателя и ударная волна распространяется по газу, падающему по направлению к звезде.

При рассмотрении бальмеровских эмиссионных профилей в потоке излучения ударной волны обнаруживается закономерность, которая заключается в том, что величина доплеровского сдвига составляет около одной трети скорости ударной волны: $\delta V \approx -U_1/3$. Существование такой закономерности связано с двумя следующими причинами. Во-первых, формирование потока излучения в частотах бальмеровских линий происходит в слоях газа, где вследствие торможения газового потока скорость течения составляет около половины скорости ударной волны. Второй причиной, определяющей величину доплеровского смещения, является значительная анизотропия поля излучения ударной волны.

На рис. 4 приведены полярные диаграммы интенсивности излучения в частотах линии $\text{H}\alpha$. Длина радиуса-вектора, соединяющего начало координат с точкой на кривой, пропорциональна интенсивности $I_\nu(\mu)$, тогда как угол между радиусом-вектором и горизонтальной осью $\theta = \cos^{-1}(\mu)$. Для удобства графического представления полярные диаграммы выражены в единицах средней интенсивности излучения J_ν , поэтому в случае изотропного излучения полярная диаграмма

представляет собой окружность единичного радиуса (с учетом симметрии на рис. 4 приведены лишь половины диаграмм).

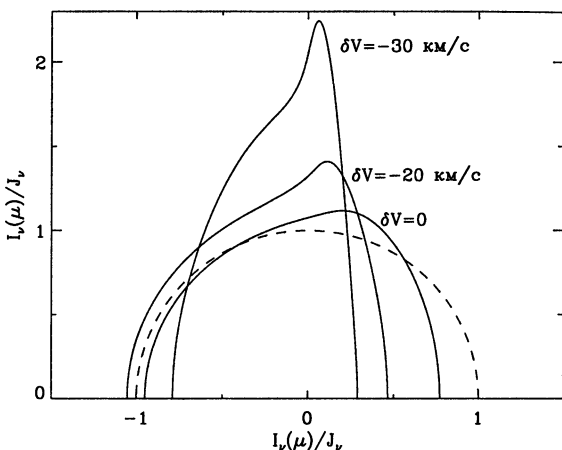


Рис. 4. Полярные диаграммы интенсивности излучения $I_\nu(\mu)$, соответствующие центру линии $H\alpha$ ($\delta V = 0$) и сдвигу в коротковолновую сторону спектра на 20 км/с и 30 км/с. Интенсивность излучения выражена в единицах средней интенсивности J_ν . Штриховой линией показана полярная диаграмма изотропного излучения.

Отклонения от изотропии достаточно малы лишь вблизи центра линии ($\delta V < 10$ км/с), однако быстро возрастают с увеличением δV вследствие меньшей оптической глубины среды в крыльях линии. Поэтому в монохроматический поток излучения

$$F_\nu = \frac{1}{2} \int_0^1 I_\nu(\mu) \mu d\mu \quad (17)$$

существенный вклад вносят лучи полярной диаграммы, составляющие большой угол с нормалью к плоскости ударной волны и, следовательно, характеризующиеся малыми значениями проекции скорости течения газа на нормаль. Вытянутость полярных диаграмм в направлении $\theta = \pm\pi/2$ обусловлена ограниченной геометрической толщиной слоя газа $[x_1, x_n]$, заключающего в себя ударную волну, и прозрачностью среды ($\tau < 1$) в частотах бальмеровских линий. Таким образом, для излучающего газа, движущегося со скоростью U , величина доплеровского сдвига составляет $\delta V \approx -0,7U$.

Литература

1. Зельдович Я. Б., Райзер Ю. П. Физика ударных волн и высокотемпературных гидродинамических явлений. М.: Наука, 1966. 688 с.
2. Fadeyev Yu. A., Gillet D. The structure of radiative shock waves. I. The method of global iterations // *Astron. Astrophys.* 1998. V. 333, № 2. P. 687–701.
3. Fadeyev Yu. A., Gillet D. The structure of radiative shock waves. II. The multilevel hydrogen atom // *Astron. Astrophys.* 2000. V. 354, № 1. P. 349–364.
4. Fadeyev Yu. A., Gillet D. The structure of radiative shock waves. III. The model grid for partially ionized hydrogen gas // *Astron. Astrophys.* 2001. V. 368, № 3. P. 901–911.
5. Fadeyev Yu. A., Le Coroller H., Gillet D. The structure of radiative shock waves. IV. Effects of electron thermal conduction // *Astron. Astrophys.* 2002. V. 392, № 2. P. 735–740.
6. Fadeyev Yu. A., Gillet D. The structure of radiative shock waves. V. Hydrogen emission lines // *Astron. Astrophys.* 2004. V. 420, № 2. P. 423–435.
7. Фадеев Ю. А. Структура ударных волн в атмосферах пульсирующих звезд // *Астрон. журн.* 2001. Т. 78, № 5. С. 421–431.
8. Фадеев Ю. А. Водородный эмиссионный спектр ударной волны в звездной атмосфере // *Письма в Астрон. журн.* 2003. Т. 29, № 11. С. 871–879.
9. Михалас Д. Звездные атмосферы. М.: Мир, 1982. 352 с.
10. Radhakrishnan K., Hindmarsh A. C. Description and use of LSODE, the Livermore solver for ordinary differential equations // LLNL Report UCRL-ID-113855, 1993. 108 p.
11. Mihalas D., Weibel Mihalas B. Foundations of radiation hydrodynamics. N. Y.: Oxford University Press, 1984. 717 p.

Нелинейные явления в окружающей среде

НЕЛИНЕЙНАЯ ДИНАМИКА ВЫСОТНОГО ГРОЗОВОГО РАЗРЯДА

Е. А. Мареев, В. В. Клименко

1. Введение

Электрический разряд в воздухе является классическим объектом нелинейной электродинамики, привлекающим внимание сложностью и разнообразием возникающих при его исследовании проблем.

Когда речь идет о грозовых электрических разрядах, обычно подразумевается наиболее впечатляющая форма такого разряда — молния, т. е. длинная искра в поле грозового облака. Однако в последние годы (начиная с середины 90-х годов XX века) в центре внимания геофизиков оказалось новое явление — разряд в средней атмосфере на высотах от 20 до 100 км, лежащих выше абсолютного большинства облаков. Оказалось, что грозовая активность, обусловленная разделением зарядов и формированием крупномасштабного электрического поля в облаках, сопровождается проникновением поля на достаточно большие высоты, где условия возбуждения разряда облегчаются из-за экспоненциального уменьшения давления воздуха с высотой. Это приводит к возникновению нескольких классов разрядных явлений, различающихся пространственно-временными характеристиками (в том числе, высотой развития разряда), условиями возбуждения и физическими процессами в области разряда. Сложившаяся к настоящему времени классификация подразделяет высотные разряды на четыре основных типа, именуемых эльфами, джетами, спрайтами и гало [1].

Эльфы — наиболее кратковременные события в семействе высотных разрядов. Они представляют собой свечение красного цвета тороидальной формы с горизонтальным масштабом 300—400 км в диапазоне высот 80—100 км, вспыхивающее примерно через 300 мкс после мощной вспышки облако-земля. Продолжительность события — порядка сотни микросекунд. Описанные выше признаки хорошо согласуются с общепринятой в настоящее время гипотезой о том, что эльфы обусловлены импульсом электромагнитного излучения, порождаемого током наиболее мощных молниевых вспышек. Эта гипотеза закрепились и в названии: в англоязычной литературе (в единственном числе) употребляется название Elve — Emission of Light and VLF perturbation from and EMP source (форма множественного числа — Elves). Хотя детали инициации, развития и статистики эльфов, а также их роль в физике нижней ионосферы практически не изучены, данный вид разряда считается наименее «проблематичным» по сравнению с другими видами высотных разрядов.

Джет — весьма редкое и особенно трудное для наблюдений событие. Это свечение голубого цвета в форме узкого конуса, стартующего вблизи

верхней границы облака и распространяющегося до высот около 40 км со скоростью порядка 100 км/с. Тесной корреляции джетов с внутриоблачными вспышками или вспышками облако-земля не отмечено. Кроме обычных голубых джетов, в последнее время выделяют такие классы событий, как «голубые стартеры» и «гигантские джеты». Первые распространяются до высот не более 25 км, а вторые — до высот нижней ионосферы около 70 км. Нужно ли говорить о том, сколь интересны для геофизики эти явления поистине грандиозного масштаба, практически «мгновенно» связывающие тропосферу с ионосферой? Однако наблюдаются гигантские джеты чрезвычайно редко; к настоящему времени зарегистрировано менее десятка событий.

Теория джетов делает первые шаги, в основном по аналогии с теорией внутриоблачного разряда (которая, кстати сказать, также далека от завершения). Критическим фактором для развития теории является весьма ограниченный пока объем экспериментальных данных о джетах.

Наибольшее количество работ в области изучения высотных разрядов в атмосфере посвящено спрайтам. Это связано как с достаточно высокой частотой этих событий (современные оценки дают цифры порядка нескольких тысяч событий в сутки по всему земному шару), так и с чрезвычайно интересной и необычной физикой явления. Спрайт — короткая (порядка миллисекунд) яркая объемная вспышка красного цвета в высотном интервале 50—90 км с поперечным размером порядка десятка километров. В нижней части спрайта часто наблюдаются столбы свечения, где присутствует излучение в голубой области спектра. Оптическая вспышка в мезосфере возникает через несколько миллисекунд (иногда через несколько десятков или даже сотен миллисекунд) после положительного разряда облако-земля, порою на удалении от канала молниевой вспышки (вернее, от нижней части канала) по горизонтали до десяти и более километров. Как известно, положительные вспышки облако-земля, которые переносят положительный заряд на землю, в среднем являются более мощными, чем отрицательные вспышки. Но только этим обстоятельством почти стопроцентный уровень корреляции спрайтов именно с положительными вспышками (что существенно отличает их от эльфов) объяснить трудно.

Наблюдения показывают, что область свечения спрайта в мезосфере имеет ярко выраженную структуру. Большинство спрайтов имеют диффузную (сравнительно однородно светящуюся) верхнюю часть (на высотах более 80 км) с поперечным размером порядка 30 км и структурированную в виде светящихся каналов нижнюю часть (на высотах ниже 70—75 км). Тонкая структура нижней части спрайтов характеризуется множеством светящихся каналов с поперечными размерами от десятков до сотен метров [2] (рис. 1).

05:39:19 UT July 13, 1998

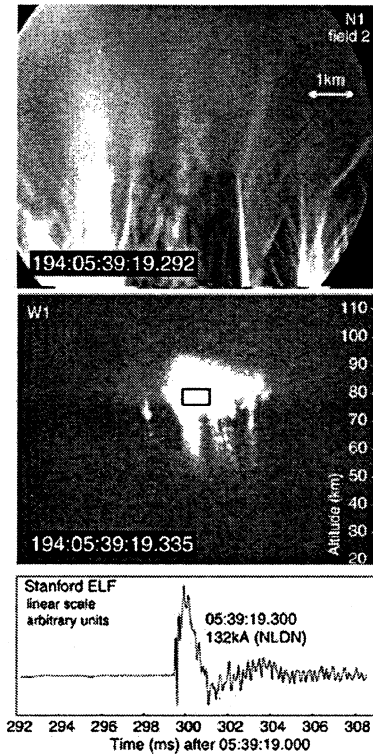


Рис. 1. Особенности тонкой структуры спрайтов [2]

В последнее время как самостоятельный вид высотного разряда выделяются также гало — диффузные свечения на высотах 75—85 км, не обладающие тонкой структурой. В отличие от спрайтов, гало наблюдаются непосредственно над областью вспышки облако-земля. Часто спрайты сопровождаются появлением гало, а гало — появлением спрайтов, но однозначной связи между этими событиями не наблюдается.

Как видно уже из описания морфологических особенностей высотных разрядов в атмосфере, физика этих явлений затрагивает целый ряд проблем нелинейной электродинамики, таких как самосогласованный нагрев и ионизация газа, пороговая динамика, формирование и распространение стримеров, фрактальная структура и динамика, развитие мелкомасштабных неустойчивостей и другие. Учет целого ряда специфических особенностей проблемы (отсутствие сосредоточенных электродов, необходимость учета динамики поля облака на стадии инициации и мол-

нии, принципиально неоднородный и нестационарный характер разряда, наличие релятивистских частиц) делает ее весьма необычной даже с точки зрения физики газового разряда, и без того очень разнообразной. Данная статья посвящена краткому обзору физики спрайтов — пожалуй, наиболее интересного типа высотных разрядов.

2. Методы измерения физических параметров спрайтов

Регистрация высотных разрядов и измерения их физических характеристик осложнены из-за малого времени жизни этого явления, трудностей идентификации сигнала на фоне «обычной» грозовой активности и спорадической природы грозовых разрядов. При проведении оптических наблюдений дополнительные трудности заключаются в высоком поглощении излучения облачной средой, поэтому наземные наблюдения обычно проводятся на горных обсерваториях (2000—4000 метров над уровнем моря) на фоне темного неба, что позволяет при отсутствии облачности на линии визирования регистрировать оптическое излучение высотных разрядов на расстоянии порядка сотен километров от центра грозовой активности. Был организован также ряд кампаний по наблюдению высотных разрядов с борта самолетов, аэростатов и, в самое последнее время, космических аппаратов.

Основная часть информации о физических параметрах спрайтов, накопленной за пятнадцать лет, включая упоминавшиеся выше данные о тонкой структуре, была получена именно оптическими методами, и прежде всего методами видеосъемки [1]. Характерный временной масштаб оптической вспышки 1—10 мс, по существу, соответствует предельному пространственному и временному разрешению используемых видеокамер. Но в последнее время появились публикации изображений спрайтов, сделанные с помощью телескопических объективов, где видна сверхтонкая структура свечения с характерным масштабом менее 100 м [2, 3]. Из-за трудностей абсолютной калибровки видеокамер в статьях обычно дается лишь оценка максимальной яркости, характерное значение которой для всех видов высотных разрядов составляет примерно один мегарэлей (1 рэлей = $10^6/4\pi$ фотон/с·см²·ср).

С помощью щелевого спектрографа, оснащенного усилителем яркости, в спрайтах были зарегистрированы молекулярные полосы излучения N₂1P, по относительным интенсивностям которых удалось оценить температуру электронов в разряде (в предположении бальцовановского распределения по энергиям), оказавшуюся порядка 2 эВ [4, 5]. Была получена также оценка параметра $E/N_m \sim 10^{-15}$ В·см², необходимого для соответствующего нагрева электронов. Это означает, что на высоте порядка 70 км, где концентрация атомов и молекул $N_m \sim 10^{15}$ см⁻³, электрическое поле должно быть порядка $E \sim 100$ В/м. В то же время очевидно, что

функция распределения электронов в разряде не является равновесной, и полученные оценки следует рассматривать как приближенные. Как правило, спектрометры не обладают достаточной светосилой и временным разрешением, чтобы отследить динамику спектра в процессе возникновения и развития спрайта.

Спрайты регистрируются практически на тех же высотах, на которые в высоких широтах проникают энергичные магнитосферные электроны, вызывающие полярные сияния с красным нижним краем — так называемые полярные сияния типа «В» [6, 7]. Так же как и в спрайтах, красный цвет нижней кромки сияния обусловлен излучением системы полос N_2^+1P . В случае полярных сияний интенсивность излучения первой отрицательной системы N_2^+1N ($\lambda = 391,4, 427,8$ и $470,9$ нм) однозначно связана со скоростью ионизации в области свечения [7], так как практически все возбужденные ионы N_2^+ (порог 19 эВ) возникают в процессе ионизации молекул N_2 (порог 15,6 эВ). Известно, в частности, что на 300 актов ионизации излучается 1 фотон с длиной волны 470,9 нм, 4 фотона — 427,8 нм и 12 фотонов — 391,4 нм [7]. Фактически это характеризует поведение функции распределения электронов в области 15—20 эВ при торможении пучка магнитосферных электронов в атмосфере.

Поскольку наличие или отсутствие существенной ионизации в спрайте является важной характеристикой разряда, в ряде работ с помощью фотометров и узкополосных фильтров выполнены измерения яркости эмиссий ионов молекулярного азота N_2^+ . В работе [8] с высоким временным разрешением измерено излучение спрайтов одновременно в синей (N_2^+1N , 427,8 нм; N_2^+2P , 399,8 нм) и красной (N_2^+1P , 580—900 нм) областях спектра. Выявлено, что в спрайтах, состоящих только из верхней диффузной части, излучение N_2^+1N отсутствует или очень слабо. Напротив, в спрайтах с развитой структурой в нижней части это излучение регистрируется и иногда имеет характер импульса в самом начале оптической вспышки, что позволило авторам работы [8] сделать важные выводы: признаки существенной ионизации регистрируются не во всех типах спрайтов, а только в разрядах с развитой нижней частью; признаком существенной ионизации в этом случае служит импульс излучения в полосе N_2^+1N в начальный момент разряда.

Но указанные выводы требуют дополнительной проверки, так как есть принципиальные отличия в процессах формирования функции распределения электронов по энергиям в полярных сияниях и спрайтах. В электрическом разряде из-за резкого падения функции распределения около порогов неупругих процессов заметная ионизация молекул азота и кислорода (пороги 15,6 эВ и 12,1 эВ) может иметь место и при низких интенсивностях излучения ионов молекулярного азота (порог 19 эВ).

С учетом отмеченных принципиальных трудностей оптических наблюдений спрайтов важным направлением исследований является поиск

более универсальных методов регистрации и измерения характеристик спрайтов, включающих использование возможностей радиоволновой диагностики.

В целом ряде работ [1, 9] с помощью наземных датчиков были зарегистрированы СНЧ/ОНЧ-сигналы, запаздывающие на несколько миллисекунд относительно основного сферика, вызванного положительным разрядом облако-земля, и совпадающие с оптической вспышкой спрайта. С учетом параметров сигнала и характеристик волноводного распространения восстановлены значения момента тока в источнике излучения. По характеристикам сигнала трудно определить высоту удаленного источника над поверхностью Земли, однако амплитудно-временная корреляция между найденным моментом тока и яркостью спрайта позволила отнести источник излучения к току, текущему в самом спрайте на высоте порядка 70 км. Полученные в [9] характерные значения моментов тока имеют порядок величины 100 кА·км. При вертикальном размере спрайта порядка 30 км это дает полный ток в спрайте около 3 кА. Если принять поперечный размер спрайта тоже порядка 30 км, то для средней плотности тока в спрайте получим примерно 3 мкА/м^2 .

В ряде работ отмечались также амплитудные и фазовые возмущения сигналов от СДВ-передатчиков ($f \approx 25 \text{ кГц}$), связанные с появлением спрайтов на трассе распространения. Анализ амплитудно-фазовых характеристик таких возмущений одновременно с анализом координат появления спрайтов показал, что в точке приема происходит интерференция двух сигналов — прямого, распространяющегося по дуге большого круга, и сигнала, рассеянного в направлении приемника на неоднородностях ионизации в спрайте. Выяснилось, что рассеяние происходит под большими углами, включая и рассеяние в обратном направлении. Время релаксации возмущения составляет в среднем единицы секунд. Характеристики рассеянных сигналов удовлетворительно объясняются теорией рассеяния на вертикальных цилиндрических неоднородностях с поперечным размером, много меньшим длины волны ($R \sim 1 \text{ км} \ll \lambda \sim 12 \text{ км}$), если электронная концентрация в неоднородностях на высоте порядка 70 км составляет около 10^4 см^{-3} и увеличивается до 10^5 см^{-3} на высоте примерно 50 км [10]. Этим методом получены первые прямые оценки N_e в спрайте.

Итак, в результате дистанционных измерений физических параметров спрайтов были установлены следующие их характеристики:

- характерный объем, занимаемый разрядом: $V_s \sim 30 \times 30 \times 30 \text{ км}^3$;
- пространственные масштабы неоднородностей: внешний порядка 10 км, основных структур порядка 1 км, сверхтонкой структуры — менее 100 м;
- концентрация электронов $N_e \sim 10^4 \text{ см}^{-3}$;
- длительность вспышки $\tau_s \sim 10 \text{ мс}$;
- время релаксации электронной концентрации $\tau_r \sim 1\text{—}10 \text{ с}$;

- температура электронов $T_e \sim 2$ эВ во время вспышки и $T_e \sim 0,02$ эВ в период релаксации;
- полный момент тока в разряде $p_s \sim 100$ кА·км, средняя плотность тока $j = p_s/V_s \sim 3 \cdot 10^{-6}$ А/м²;
- $E/N_m \sim 10^{-15}$ В·см² на высоте 70 км, что соответствует электрическому полю $E \sim 100$ В/м.

3. Физические модели спрайтов

3.1. Физическая причина возникновения спрайтов и гало на достаточно большой высоте весьма проста: пороговое поле пробоя воздуха падает в атмосфере с высотой экспоненциально (в соответствии с экспоненциальным падением давления воздуха), в то время как возмущения электрического поля, возникающие в момент мощных молниевых вспышек облако-земля, уменьшаются с высотой медленнее (по степенному закону) и на высотах примерно 75 км превышают пробойное поле. Однако детали распределения поля с высотой и его динамики, определяемые не только неоднородностью давления, но также профилем проводимости и динамикой вспышки, могут быть очень сложными. Чтобы проиллюстрировать возникающие при этом особенности, рассмотрим вначале задачу о проникновении поля в среду с экспоненциально растущей проводимостью. Как мы увидим, возмущения среды могут быть существенными, и мы учтем их в квазигидродинамическом приближении, а затем уже перейдем к анализу кинетики электронов в электрическом поле.

При моделировании электродинамических процессов в атмосфере выбор уравнений зависит от характера решаемой задачи. Определим характерную частоту процесса ω и проводимость среды σ . Если длина волны $\lambda = 2\pi/\omega(\epsilon_0\mu_0)^{1/2}$ и толщина скин-слоя $2(\omega\mu_0\sigma)^{-1/2}$ велики по сравнению с характерным масштабом задачи, электрическое поле можно рассматривать в квазистатическом приближении: $\text{rot } \mathbf{E} = 0$, следовательно, $\mathbf{E} = -\nabla\phi$. Предполагая, что пространственный масштаб анализируемых задач меньше 100 км, а проводимость меньше 10^{-8} См/м, получаем условие: $\omega \ll 2 \cdot 10^4$ с⁻¹. В рамках квазистатического приближения можно либо пренебречь током смещения для наиболее медленных процессов (например, при описании токов в глобальной электрической цепи), либо учесть его в зависимости от соотношения частоты ω и обратного времени релаксации заряда σ/ϵ_0 . В большинстве случаев ток смещения нужно учитывать. При анализе целого ряда проблем электродинамики атмосферы, разумеется, обе производные по времени в уравнениях Максвелла должны быть сохранены. Типичным примером является излучение молниевой вспышки, приводящее, в частности, к зажиганию разряда на высотах 95—100 км (эльфы).

В пренебрежении магнитным полем, диффузией зарядов и током конвекции уравнение для квазистатического электрического поля может быть записано в виде

$$\operatorname{div} \left(\frac{\partial \mathbf{E}}{\partial t} + \frac{\sigma}{\varepsilon_0} \mathbf{E} \right) = 0. \quad (1)$$

Если проводимость зависит только от вертикальной координаты z , интеграл уравнения (1) имеет вид

$$\frac{\partial E_z}{\partial t} + \frac{1}{\varepsilon_0} j_z = \frac{\partial E_z}{\partial t} + \frac{\sigma}{\varepsilon_0} E_z = f(t), \quad (2)$$

где функция $f(t)$ не зависит от пространственных координат и определяется только граничными условиями. Общее решение уравнения записывается для произвольной функции $\sigma(z)$:

$$E = e^{-\sigma t / \varepsilon_0} \left(C + \int f(t) e^{\sigma t / \varepsilon_0} dt \right). \quad (3)$$

Здесь константа C выбирается в соответствии с начальными условиями. Для экспоненциального профиля проводимости $\sigma = \sigma_0 e^{z/H}$ при мгновенном включении электрического поля E_0 на границе $z = 0$ получаем

$$E = E_0 \left(1 - e^{-z/H} \right) e^{-\sigma t / \varepsilon_0} + E_0 e^{-z/H}. \quad (4)$$

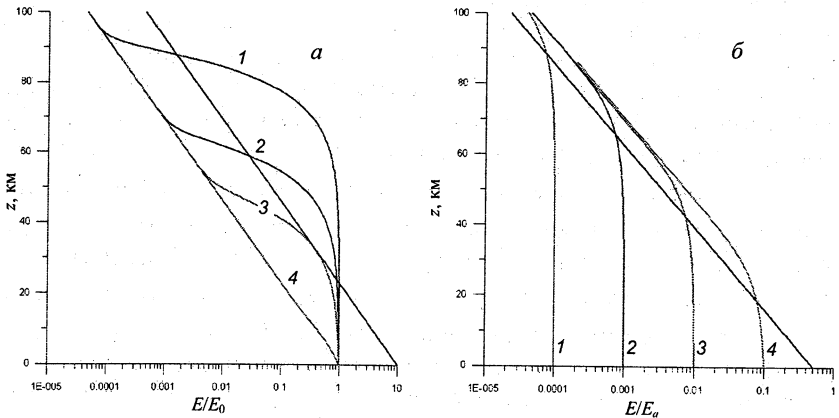


Рис. 2. Переходный процесс: *a* — после мгновенного включения электрического поля на границе (отношение t/t_r равно 0,001 (1); 0,01 (2); 0,04 (3); 1 (4)); *б* — на стадии непрерывного тока (отношение t/t_r равно 0,0001 (1); 0,001 (2); 0,01 (3); 0,1 (4)). Время нормировано на $t_r = \varepsilon_0 / \sigma_0$.

На рис. 2, *a* показаны высотные профили электрического поля (нормированного на величину E_0) в логарифмическом масштабе в различные моменты времени. Время нормировано на время релаксации при $z = 0$:

$\tau_r = \varepsilon_0 / \sigma_0$. Сплошная линия, начинающаяся в точке $E = 10$, соответствует полю пробоя, зависящему экспоненциально от высоты. Принято для простоты, что характерные масштабы изменения плотности нейтрального газа и роста проводимости совпадают, а поле на поверхности в 10 раз меньше пробойного. Из графика видно, что область, занятая полем, превышающим критическое, простирается от определенного уровня H_c , равного в данном случае примерно 20 км, до уровня $H \ln(\varepsilon_0 / \sigma_0 t)$, движущегося вниз со скоростью $V \sim H/t$, обратно пропорциональной времени, до тех пор, пока этот уровень превышает H_c . Для генерации спрайтов, однако, характерна ситуация, когда разряд в мезосфере зажигается на стадии непрерывного тока вспышки облако-земля спустя несколько миллисекунд (иногда сотни миллисекунд) после максимума тока, соответствующего моменту возвратного удара молнии. На стадии непрерывного тока возмущение заряда непрерывно растет: $Q = It$. Если этот заряд распределен равномерно по площади достаточно большого масштаба, связанное с ним возмущение поля можно оценить как $E_0 = aIt / \varepsilon_0 S$, где S — площадь поверхности слоя (так что Q/S — поверхностная плотности заряда), а параметр a порядка единицы учитывает влияние глобальной электрической цепи. Решение уравнения (2) имеет вид

$$E = \frac{Ia}{S\sigma_0} e^{-z/H} \left(1 - e^{-z/H}\right) \left(1 - e^{-\sigma t / \varepsilon_0}\right) + \frac{Ia}{S\varepsilon_0} e^{-z/H}. \quad (5)$$

Рисунок 2, б показывает профили электрического поля (нормированные на $E_a = Ia / \sigma_0 S$ и представленные в логарифмическом масштабе) в различные моменты времени. На нижней границе поле E_0 растет линейно. Прямая, начинающаяся в точке $E_c = 0,4$, соответствует пробойному полю (которое выбрано равным $0,4 E_a$ на нижней границе для иллюстрации), растущему экспоненциально с высотой. Существенная особенность поведения поля, в отличие от случая мгновенного включения поля на границе, состоит в том, что выше точки H_c , определяемой из уравнения $\sigma(H_c)t = \varepsilon_0$, поле $E \approx E_a$ постоянно во времени в течение стадии непрерывного тока и не зависит от полого заряда, перенесенного вспышкой. Если это поле превышает пробойное E_c , область возможного развития разряда простирается от точки H_c (движущейся вниз со скоростью $V \sim H/t$ на стадии непрерывного тока) до бесконечности.

Переходя к вычислению электрического поля для более реалистичного профиля проводимости атмосферы, рассмотрим в качестве примера характерный экспериментальный профиль проводимости ночной атмосферы [1]. В качестве источника также рассмотрим более типичное дипольное распределение заряда. Поле можно найти с помощью следующего приближенного уравнения [11]:

$$\frac{\partial E_z}{\partial t} + \frac{\sigma(E_z)}{\varepsilon_0} E_z = \frac{hI}{\pi \varepsilon_0 z^3}. \quad (6)$$

Здесь h — средняя высота области, в которой распределен заряд, переносимый молниевой вспышкой, I — амплитуда тока. Уравнение (6) на достаточно малых временах дает дипольное поле в окрестности источника. Следует подчеркнуть, однако, что это уравнение не учитывает горизонтальную компоненту электрического поля на высоте, где развивается разряд. Рассматривая задачу о нахождении условий пробоя, можно не учитывать возмущение проводимости из-за ионизации, а ограничиться учетом более быстрого процесса изменения проводимости под действием нагрева электронов. Таким образом, в уравнении (6) учтена лишь зависимость $\sigma(E_z)$, обусловленная изменением частоты столкновений электронов при нагреве.

Чтобы оценить возмущение проводимости $\sigma = e^2 N / m v_e(T_e)$ вследствие нагрева электронов в электрическом поле, можно воспользоваться уравнением баланса температуры [12]:

$$\frac{dT_e}{dt} = \frac{2}{3} \frac{e^2 E^2}{m v_e} - \delta(T_e) v_e(T_e) (T - T_e), \quad (7)$$

где δ — доля энергии, теряемая электроном при столкновении; v_e — частота столкновений электронов; m — масса электрона; T — температура газа. Если характерное время изменения электрического поля велико по сравнению с временем нагрева $(\delta v_e)^{-1}$, (7) сводится к следующему уравнению для квазистационарного распределения температуры:

$$\left(\frac{E}{E_p} \right)^2 \frac{v_e^2(T) \delta(T)}{v_e^2(T_e) \delta(T_e)} = \frac{T_e}{T} - 1. \quad (8)$$

Здесь $E_p = \sqrt{3\delta(T) v_e^2(T) m T / (2e^2)}$ — характерное плазменное поле. Из уравнения (8) очевидно, что при $E \ll E_p$ нагревом электронов можно пренебречь, а в случае $E \gg E_p$ температура электронов существенно возрастает. Используя формулу

$$v_e = 1,84 \cdot 10^9 \cdot (N_m / 10^{17} \text{ см}^{-3}) \cdot (T_e / 10^3 \text{ К})^{5/6} \text{ с}^{-1},$$

справедливую для температуры электронов порядка 1 эВ [13], находим:

$$\begin{aligned} (E/E_p)^{3/4} &= T_e/T, \quad T_e < 10850 \text{ К}; \\ (E/E_p)^{3/4} &= T_e/T \exp[(T_e - 10850)/6500 \text{ К}], \quad T_e > 10850 \text{ К}; \\ \sigma &= \sigma(T) (T/T_e)^{5/6}, \end{aligned} \quad (9)$$

где температура измеряется в градусах Кельвина, а концентрация нейтральных частиц N_m в $[см^{-3}]$. Характерные значения температуры в аргументе экспоненциальной функции получены с помощью экспоненциальной аппроксимации зависимости $\delta(T_e)$ при высоких температурах [11].

Из системы уравнений (9) следует зависимость $\sigma(E)$. Важность учета возмущений проводимости вследствие нагрева иллюстрируется рис. 3, на котором показан профиль проводимости в поле так называемой M -компоненты молнии [14]. Видно, что эффект падения проводимости ярко выражен на высотах 55—75 км благодаря более медленному падению поля с высотой по сравнению с падением плазменного поля. На больших высотах, однако, этот эффект не проявляется из-за быстрого вытеснения поля из области высокой проводимости.

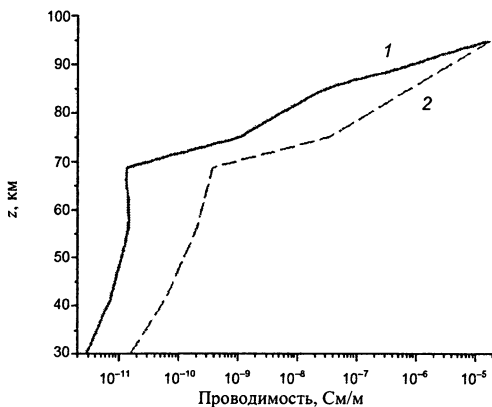


Рис. 3. Профиль возмущения проводимости во внешнем электрическом поле (1) и без внешнего поля (2) на стадии M -компоненты молнии [14].

Наконец, типичный профиль поля для реального профиля проводимости с учетом нагрева приведен на рис. 4 [11], где приведена также кривая изменения порогового поля пробоя. Видно, что условия пробоя, определяемые превышением эффективной частоты ионизации над потерями электронов (в данном случае частотой прилипания), выполняются в некотором интервале высот вблизи высоты 75 км на определенной стадии молниевой вспышки. Для более подробного рассмотрения условий пробоя и сценариев развития разряда в следующем разделе мы рассмотрим основы кинетики нагрева электронов в электрическом поле.

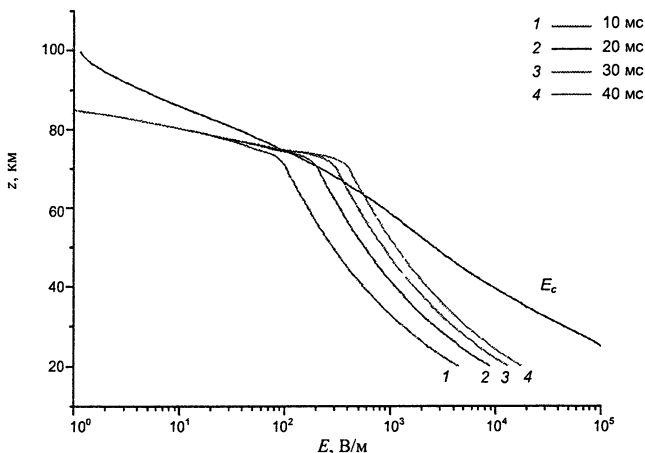


Рис. 4. Профили электрического поля на стадии непрерывного тока при $I = 20$ кА в различные моменты времени [11].

3.2. Наиболее фундаментальная проблема физики спрайтов — развитие разряда в неоднородном слабоионизованном газе. Как уже отмечалось, область свечения спрайта в мезосфере имеет ярко выраженную структуру, причем тонкая структура нижней части спрайтов характеризуется множеством светящихся каналов с поперечными размерами от десятков до сотен метров.

Как известно, в газах низкого давления при не очень больших масштабах промежутка, занятого электрическим полем (когда $pd < 200$ Тор·см), действует таунсендовский (лавиновый) механизм пробоя. Если же давление газа достаточно велико (так что $pd > 10^3$ Тор·см), действует стримерно-лидерный механизм. Естественно предположить, что формирование нижней части спрайтов связано с развитием стримеров в электрическом поле, тогда как в верхней части спрайтов, относящейся к области низких давлений в атмосфере, развитие разряда обусловлено размножением электронных лавин.

Стример представляет собой умеренно ионизованный тонкий канал, который образуется из первичной лавины в достаточно сильном поле и прорастает в одну или в обе стороны. Стример образуется из лавины, если поле ее пространственного заряда достигает величины порядка внешнего. Приближенное равенство

$$E' = eR^{-2} \exp[\alpha(E_0)x] \approx E_0 \quad (10)$$

часто рассматривается как критерий возникновения стримера. С учетом оценки радиуса головки стримера R данный критерий приводит к так на-

зываемому условию Мика: $\alpha(E_0)d \approx 18-20$ [15]. В электроотрицательных газах (к каковым относится воздух) прилипание электронов замедляет процесс ионизации в лавине и приводит к повышению пробивающих полей и граничных значений параметра pd , при которых таунсендовский механизм развития разряда сменяется стримерным. В переходной области значений pd возможно «соревнование» между двумя механизмами [15].

Рассмотрим приближенные решения кинетического уравнения для электронов в слабоионизованной плазме в электрическом поле и на основе этих решений вычислим основные характеристики лавинной стадии размножения электронов.

В общем случае кинетическое уравнение для функции распределения электронов в электрическом поле имеет вид [16]

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} - \frac{e}{m} \left(\mathbf{E} + \frac{1}{c} [\mathbf{v}, \mathbf{H}] \right) \cdot \frac{\partial f}{\partial \mathbf{v}} = S, \quad (11)$$

где S — интеграл столкновений. Пренебрегая влиянием магнитного поля и пространственной неоднородностью, будем искать функцию распределения $f(\mathbf{v}, t)$ в виде разложения по полиномам Лежандра:

$$f(\mathbf{v}, t) = \sum_{k=0}^{\infty} P_k(\cos \theta) \cdot f_k(v, t), \quad (12)$$

где θ — угол между векторами \mathbf{E} и \mathbf{v} . Первые два члена разложения

$$f(v, t) \approx f_0(v, t) + \frac{\mathbf{v}}{v} \cdot \mathbf{f}_1(v, t) \quad (13)$$

описывают изотропную и направленную части функции распределения электронов соответственно. Условие пренебрежения f_2 по сравнению с f_0 , имеющее вид $e^2 E^2 l_e^2 / mv^2 T_e \ll 1$, означает, что на длине свободного пробега электрон набирает энергию меньшую, чем его средняя энергия. В итоге задача сводится к системе двух уравнений для f_0 и f_1 :

$$\frac{\partial f_0}{\partial t} - \frac{eE}{3m} \cdot \frac{1}{v^2} \frac{\partial}{\partial v} (v^2 \cdot f_1) = S_0; \quad \frac{\partial f_1}{\partial t} - \frac{eE}{m} \frac{\partial f_0}{\partial v} = S_1, \quad (14)$$

где S_0 и S_1 — интегралы столкновений для изотропной и направленной частей функции распределения. На высотах мезосферы характерное время релаксации электронов по энергии $(\delta \cdot v_e)^{-1} \sim 10^{-4}$ с, поэтому на временных интервалах, больших 0,1 мс, можно использовать стационарное приближение

$$-\frac{eE}{3m} \cdot \frac{1}{v^2} \frac{\partial}{\partial v} (v^2 \cdot f_1) = S_0; \quad -\frac{eE}{m} \frac{\partial f_0}{\partial v} = S_1. \quad (15)$$

Дальнейшие вычисления зависят от выбора модели для интегралов столкновений. В практически значимом для нас диапазоне энергий $\epsilon \ll 100$ эВ рассеяние электронов на нейтральных молекулах почти изотропно, и каждое столкновение сильно изменяет направление дви-

жения электрона. Кроме того, частота упругих столкновений ν_e существенно больше частоты неупругих столкновений ν_h . Поэтому для направленной части функции распределения в интеграле столкновений имеем $S_1 = -(\nu_e + \nu_h) \cdot f_1 \approx -\nu_e \cdot f_1$. В итоге необходимо решать уравнение

$$-\frac{e^2 E^2}{3m^2} \cdot \frac{1}{v^2} \frac{d}{dv} \left(v^2 \frac{df_0}{dv} \right) = S_0, \quad (16)$$

где интеграл столкновений описывает влияние как упругих, так и неупругих соударений электронов с молекулами: $S_0 = S_{0e} + S_{0h}$.

Имеет смысл принять во внимание следующие приближения:

1. Малые потери энергии в упругих соударениях. При этом S_{0e} имеет форму дивергенции потока электронов $J_e(v)$ в фазовом пространстве скоростей [12]:

$$S_{0e} = \frac{1}{v^2} \frac{d}{dv} \left[v^2 J_e(v) \right], \quad (17)$$

$$J_e(v) = \frac{\delta_e}{2} \cdot \nu_e(v) \cdot v \cdot f_0 + \frac{T_a}{M_a} \nu_e(v) \cdot \frac{\partial f_0}{\partial v} \approx \frac{\delta_e}{2} \cdot \nu_e(v) \cdot v \cdot f_0, \quad (18)$$

где ν_e — транспортная частота упругих столкновений; M_a — масса молекул газа; $\delta_e = 2m/M_a$ — средняя доля энергии, теряемая электроном за время ν_e^{-1} . В окончательном выражении для J_e принято, что средняя энергия электронов много больше средней энергии нейтральных атомов и молекул T_a .

2. Малые потери энергии при неупругих соударениях, пороги которых малы по сравнению с энергией электронов, например, при возбуждении вращательных уровней молекул. Этот случай практически сводится к предыдущему с заменой δ_e на $\delta_e(v) \gg 2m/M_a$. В области $\epsilon \leq 1$ эВ для воздуха можно считать $\delta_e(v) \approx 3 \cdot 10^{-3}$.

3. Большие потери энергии при неупругих соударениях в случае небольшого превышения энергии электронов над порогом возбуждения. В этом случае электрон при столкновении теряет значительную часть своей энергии, исчезает из фазового объема в окрестности порога ν_1 и появляется в элементе фазового объема где-то между $\nu = 0$ и $\nu = \nu_1$. Соответствующую часть интеграла столкновений тогда можно представить в виде

$$S_{0h}(v) = -\nu_h(v) f_0 + Q_0 \cdot \psi(v); \quad Q_0 = 4\pi \int_{\nu_1}^{\infty} \nu_h f_0(v) v^2 dv, \quad (19)$$

где Q_0 — полное число неупругих столкновений с «большой» потерей энергии в единицу времени, а функция $\psi(v)$ описывает характер потерь энергии. Полный интеграл столкновений примет вид

$$S_0 = \frac{1}{2v^2} \frac{d}{dv} \left(\delta_e \cdot \nu_e v^3 f_0 \right) - \nu_h(v) f_0 + Q_0 \cdot \psi(v), \quad (20)$$

где первое слагаемое описывает столкновения с малыми потерями энергии, как упругие, так и неупругие — возбуждение вращательных уровней молекул.

Вид функции $\psi(v)$ можно выбрать исходя из следующих соображений. В составе верхней атмосферы около 78% молекулярного азота, поэтому процессы возбуждения колебательных уровней N_2 с порогами в области $\varepsilon_1 \approx 1-2$ эВ играют роль эффективного стока для энергии электронов. Оптическое излучение спрайтов обусловлено возбуждением электронных переходов N_2 с порогами от $\varepsilon_2 \approx 7,3$ эВ и выше. Пороговые энергии ионизации молекул ε_i равны 12,0 эВ и 15,6 эВ для O_2 и N_2 соответственно. Отсюда следует величина диапазона энергий электронов 1—20 эВ, который нас будет интересовать. Пороговая энергия при возбуждении колебаний N_2 равна примерно $\varepsilon_1 \approx 1,7$ эВ, а потери энергии при возбуждении колебаний кратны $\Delta\varepsilon \approx 0,29$ эВ и с убывающей вероятностью составляют 0,29 эВ, 0,58 эВ, 0,87 эВ и т. д. Поэтому для $\psi(v)$ разумно выбрать форму, описывающую случайные потери энергии с равномерным распределением рассеянных электронов по всему фазовому объему $4\pi v_1^3/3$: $\psi(v) = 3/4\pi v_1^3$.

Применительно к электрическим разрядам в мезосфере будем искать решение уравнения (16) в приближении сильного поля. В области $v < v_1$ учет неупругих столкновений сводится к учету третьего слагаемого в выражении для S_0 , так как здесь $v_h = 0$. После умножения на $4\pi v^2 dv$ и интегрирования от 0 до v уравнение приобретает вид

$$\frac{4\pi}{3} \cdot \frac{e^2 E^2}{m^2 v_e} v^2 \frac{df_0}{dv} - 2\pi \delta_e v_e v^3 f_0 = Q_0 \cdot \frac{v^3}{v_1^3}, \quad (21)$$

решение которого записывается как

$$f_0(v) = e^{-\alpha(E/N) \cdot v^4} \left(\frac{3m^2}{4\pi e^2 E^2 \lambda_e} \frac{Q_0}{v_1^3} \int_v^{v_1} u^2 \exp(\alpha(E/N) \cdot u^4) du + \text{const} \right), \quad (22)$$

где длина свободного пробега $\lambda_e = (\sigma_i N)^{-1} = \text{const}$, $v_e = v/\lambda_e \sim v$ и $\alpha(E/N) = 3m^2 \delta_e / 8e^2 E^2 \lambda_e^2$. Так как E велико, можно положить экспоненты равными единице и записать приближенно

$$f_0(v \leq v_1) \approx \frac{m^2 Q_0}{4\pi e^2 E^2 \lambda_e v_1^3} \left(1 - \frac{v^3}{v_1^3} \right) + C. \quad (23)$$

В области $v \sim v_1$ и выше порога частота неупругих столкновений быстро увеличивается, и в интеграле столкновений можно пренебречь первым слагаемым по сравнению со вторым, поскольку при $v \geq v_1$ выполняется неравенство $\delta_e v_e \ll v_h$. Третье слагаемое в этой области просто равно нулю. В области неупругих столкновений для $f_0(v)$ имеем уравнение

$$\frac{d^2 f_0}{dv^2} + \left(\frac{2}{v} - \frac{1}{v_e} \frac{dv_e}{dv} \right) \frac{df_0}{dv} - \frac{q^2}{v_1^2} f_0 = 0, \quad q^2 = \frac{3m^2 v_e v_h}{e^2 E^2} v_1^2 \approx 2 \frac{v_h}{\delta_e v_e} \frac{\varepsilon_1}{T_e} \gg 1. \quad (24)$$

При условии $\lambda_e = \text{const}$ уравнение записывается в виде

$$\frac{d^2 f_0}{dv^2} + \frac{1}{v} \frac{df_0}{dv} - \frac{q^2}{v_1^2} f_0 = 0. \quad (25)$$

Решение уравнения (25) можно получить в квазиклассическом приближении. Однако в данном случае удобно воспользоваться другим, более наглядным способом. Так как $T_e \sim \varepsilon_1$ и $v_h \gg \delta_e v_e$, то $q^2 \gg 1$. В области $v \geq v_1$ функция распределения $f_0(v)$ быстро убывает, поэтому в уравнении можно пренебречь вторым слагаемым по сравнению с первым и считать $q(v)$ медленно меняющейся функцией. В результате получаем решение

$$f_0(v \geq v_1) = C \cdot \exp\left(-\int_{v_1}^v \frac{q}{v_1} dv\right) = C \cdot \exp\left(-\frac{\sqrt{3} \cdot m}{eE} \int_{v_1}^v \sqrt{v_e v_h} dv\right), \quad (26)$$

где константа C определяется условием нормировки

$$\int_0^{v_1} f_0(v \leq v_1) \cdot 4\pi v^2 dv + \int_{v_1}^{\infty} f_0(v \geq v_1) \cdot 4\pi v^2 dv = 1. \quad (27)$$

При $v \geq v_1$ функция $f_0(v)$ прижата к порогу, поэтому в приграничной области можно положить $v_e(v) \sim v_e(v_1) = \text{const}$. Зависимость частоты неупругих столкновений от скорости вблизи порога аппроксимируем как

$$v_h(v) = N \cdot \sigma_h(v) v_1 \approx v_{h0} (v - v_1) / v_1, \quad (28)$$

где N — концентрация молекул; $\sigma_h(v)$ — сечение неупругих столкновений, которое вблизи порога удовлетворительно аппроксимируется формулой

$$\sigma_h(v) \approx \text{const} \cdot (\varepsilon - \varepsilon_1) \sim (v^2 - v_1^2) \approx 2v_1 (v - v_1). \quad (29)$$

В итоге в области $v \geq v_1$ имеем

$$f_0(v \geq v_1) = C \cdot \exp\left[-q_0 \left(\frac{v}{v_1} - 1\right)^{3/2}\right]; \quad q_0 = \frac{2m\sqrt{v_e(v_1) \cdot v_{h0}}}{\sqrt{3} \cdot eE} v_1 \sim \left(\frac{E}{N}\right)^{-1}. \quad (30)$$

Из условий нормировки f_0 и определения Q_0 получаем

$$C = \frac{3}{4\pi v_1^3} \mu_1 \left(1 - \frac{\mu_1 \mu_2 q_0^2}{4 + \mu_1 \mu_2 q_0^2}\right); \quad Q_0 = 8v_{h0} \frac{\mu_1 \mu_2}{4 + \mu_1 \mu_2 q_0^2};$$

$$\text{где } \mu_1^{-1} = 1 + 2 \left[\frac{1}{q_0^2} + \frac{2\Gamma(4/3)}{q_0^{4/3}} + \frac{\Gamma(2/3)}{q_0^{2/3}} \right], \quad \mu_2 = \frac{\Gamma(8/3)}{q_0^{8/3}} + \frac{2}{q_0^2} + \frac{\Gamma(4/3)}{q_0^{4/3}}.$$

Очевидно, что хвост полученной функции распределения электронов $\exp(-\alpha v^{3/2})$ спадает гораздо медленнее, чем у функции Дрювенштейна $\exp(-\alpha v^4)$ и близок к максвелловскому $\exp(-\alpha v^2)$. На рис. 5 для иллюстра-

ции показано полученное приближенное решение кинетического уравнения при $E/N = 40\text{—}100$ Td ($1 \text{ Td} = 10^{-17} \text{ В}\cdot\text{см}^2 \approx 0,354 \text{ В}/\text{см}\cdot\text{Тор}$), где неупругим процессом является возбуждение колебаний N_2 с $\varepsilon_1 = 1,7$ эВ, $v_1 = 7,7 \cdot 10^7 \text{ см}/\text{с}$ и $v_{i0}/[\text{N}_2] = 1,23 \cdot 10^{-7} \text{ см}^3 \cdot \text{с}^{-1}$, а также функции распределения Максвелла ($T_e = 1,5$ эВ) и Дрювенштейна (100 Td , $\delta_e = 0,15$).

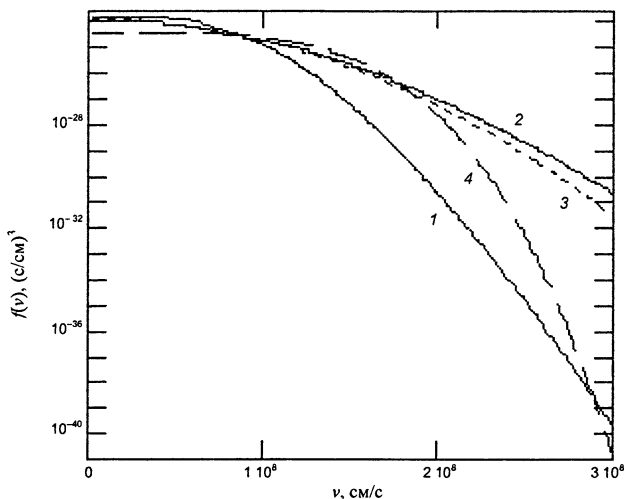


Рис. 5. Функция распределения электронов по скоростям в электрическом поле: приближенное решение кинетического уравнения при $E/N = 40$ Td (1) и при $E/N = 100$ Td (2), распределение Максвелла при $T_e = 1,5$ эВ (3), распределение Дрювенштейна при $E/N = 100$ Td (4) ($\delta_e = 0,15$).

Нашей целью является вычисление на основе функции распределения основных характеристик лавинной стадии разряда в верхней атмосфере: частоты ионизации ν_i , частоты прилипания электронов ν_a и частоты размножения лавины $(\nu_i - \nu_a)$, частот возбуждения оптических переходов ν_r , а также геометрических размеров лавин на пороге стримерной стадии. Процедуры вычисления частот ионизации и возбуждения аналогичны. В частности, для частоты ионизации вычисляем интеграл

$$\nu_i = \int_{v_i}^{\infty} v_{i0} \left(\frac{v}{v_i} - 1 \right) \cdot 4\pi v^2 f_0(v) dv \quad (31)$$

и в результате имеем

$$v_i = \frac{4}{3} v_{i0} \mu_1 \left(1 - \frac{\mu_1 \mu_2 q_0^2}{4 + \mu_1 \mu_2 q_0^2} \right) \frac{\exp \left[-q_0 \left(\frac{v_i}{v_1} - 1 \right)^{3/2} \right]}{q_0^2} \cdot \frac{v_i}{v_i - v_1} \left(1 + \frac{4}{\alpha_i} + \frac{6}{\alpha_i^2} \right), \quad (32)$$

где $\alpha_i = (3q_0 v_i / 2v_1) [(v_i - v_1) / v_1]^{1/2}$. Результаты вычислений суммарной частоты ионизации молекул O_2 и N_2 показаны на рис. 6. Там же приведены частота v_r возбуждения оптических переходов первой положительной системы $N_2 1P$ и частота v_b ионизации молекул N_2 с переходом их в состояние $B^2\Sigma_u$, которое является источником эмиссии $\lambda = 427,8$ нм в синей области спектра. Все сечения неупругих столкновений взяты из работы [17], а при вычислениях использованы их линейные аппроксимации в пороговой области энергий:

$$\begin{aligned} \varepsilon_i(O_2) &= 12,0 \text{ эВ}, & v_i &= 2,0 \cdot 10^8 \text{ см/с}, & v_{i0}[O_2] &= 3,15 \cdot 10^{-8} \text{ см}^3/\text{с}; \\ \varepsilon_i(N_2) &= 15,6 \text{ эВ}, & v_i &= 2,3 \cdot 10^8 \text{ см/с}, & v_{i0}[N_2] &= 6,0 \cdot 10^{-8} \text{ см}^3/\text{с}; \\ \varepsilon_r(N_2) &= 7,3 \text{ эВ}, & v_r &= 1,6 \cdot 10^8 \text{ см/с}, & v_{r0}[N_2] &= 3,2 \cdot 10^{-8} \text{ см}^3/\text{с}; \\ \varepsilon_b(N_2) &= 18,75 \text{ эВ}, & v_b &= 2,8 \cdot 10^8 \text{ см/с}, & v_{b0}[N_2] &= 4,7 \cdot 10^{-10} \text{ см}^3/\text{с}. \end{aligned}$$

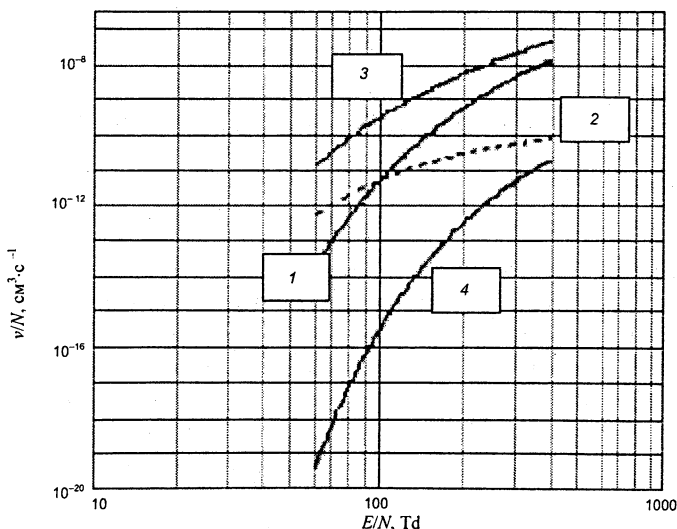


Рис. 6. Нормированные на полную плотность молекул воздуха частоты неупругих соударений электронов в электрическом поле: суммарная частота ионизации молекул O_2 и N_2 (1), частота прилипания электронов в воздухе из-за парных столкновений с O_2 (2), частота возбуждения молекул N_2 на уровень $B^3\Pi_g$ ($N_2 1P$) (3), частота ионизации молекул N_2 с переходом в состояние $B^2\Sigma_u^+$ ($N_2^+ 1N$) (4).

При средней энергии электронов в лавине $\langle \varepsilon \rangle > 1$ эВ основным механизмом потерь электронов является их диссоциативное прилипание к молекулам кислорода в парных столкновениях [18]. Приведенные в [18] экспериментальные данные о коэффициенте двухчастичного прилипания для O_2 при $\langle \varepsilon \rangle$ от 1,5 до 3 эВ хорошо аппроксимируются выражением

$$\frac{v_a}{[O_2]} = \langle \sigma_a v \rangle = A \cdot \exp\left(-\frac{B}{\langle \varepsilon \rangle}\right), \text{ где } A = 4,06 \cdot 10^{-10} \text{ см}^3/\text{с}, B = 8,43 \text{ эВ. При}$$

этом средняя энергия, вычисленная с помощью $f_0(v)$, практически линейно зависит от параметра E/N в диапазоне от 50 до 300 Td: $\langle \varepsilon [\text{эВ}] \rangle = 0,1 + 2,3 \cdot 10^{-2} E/N [\text{Td}]$. Определенная таким способом частота прилипания электронов также приведена на рис. 6. На основе представленных на рис. 6 результатов можно сформулировать следующие краткие выводы:

1. Вычисленные на основе приближенного решения кинетического уравнения частоты ионизации и прилипания хорошо согласуются с данными из других литературных источников, где использовались как соотношения подобия [1], так и детальные численные решения кинетического уравнения методом Монте-Карло [19]. Совпадает и критическое значение параметра $(E/N)^* \approx 100 \text{ Td}$, при котором выполняются необходимые для развития лавины условия $v_i > v_a$.

2. В отличие от ситуации в полярных сияниях, где один фотон в линии 427,8 нм излучается на 75 актов ионизации, здесь один фотон излучается на 10^3 — 10^4 актов ионизации и на 10^4 — 10^5 фотонов, излученных в полосе $N_2 I P$. Значит, в тех случаях, когда с помощью электрофотометров регистрировалось излучение спрайтов в линии 427,8 нм, ионизация и эмиссия в красной области спектра были на несколько порядков больше, чем в полярных сияниях, по крайней мере для тех высот, где еще можно не учитывать гашение возбужденных молекул столкновениями.

3. При $E/N > (E/N)^* \approx 100 \text{ Td}$ нормированная частота $v_i/N > 10^{-9} \text{ см}^3 \cdot \text{с}^{-1}$. Если принять за типичное значение яркости спрайта в красной области спектра 1 мегарелей = 10^{12} фотонов/(с·см²·ср), то для генерации такого свечения на высотах 70—75 км ($N \sim 10^{15} \text{ см}^{-3}$) достаточно иметь на луче зрения 10^6 электронов/см², т. е. при размере разряда порядка 10 км среднюю плотность электронов вдоль линии визирования примерно 1 см^{-3} . Высокие значения яркостей спрайтов могут иметь место и на лавинной стадии разряда.

Геометрические характеристики лавинной стадии разряда определяются:

- 1) достижением головкой лавины критического радиуса, когда ее собственное кулоновское поле сравнивается с внешним приложенным полем, т. е. достижением порога для образования стримера;

2) начальным числом фоновых электронов в единице объема, которое определяет плотность лавин в пространстве и возможность их перекрытия до достижения критического радиуса.

С учетом диффузионного расширения головки лавины можно получить уравнения для критического радиуса и соответствующего времени жизни лавины:

$$R_{cr} = \left[\frac{8 \langle \varepsilon \rangle \cdot u_E}{3 (v_i - v_a) \cdot eE} \ln \frac{E \cdot R_{cr}^2}{e} \right]^{1/2}; \quad t_{cr} = \frac{1}{v_i - v_a} \ln \left(\frac{8 \langle \varepsilon \rangle}{3e^2} u_E t_{cr} \right). \quad (33)$$

Здесь u_E — скорость направленного дрейфа электронов в электрическом поле, которая вычисляется с помощью компоненты функции распределения $f_1(v)$:

$$u_E = \frac{4\pi}{3} \int_0^\infty f_1(v) v^3 dv = \frac{4\pi eE}{3m} \int_0^\infty v^3 \frac{df_0}{dv} dv. \quad (34)$$

В результате путь, пройденный лавиной к моменту t_{cr} , будет равен $l_{cr} = u_E t_{cr}$. Условие перекрытия лавин на этой фазе дает пороговую концентрацию фоновых электронов $n_e^{(0)} > (2R_{cr})^{-3}$. На рис. 7 и 8 показаны зависимости от параметра E/N критических пространственно-временных характеристик лавин и пороговой плотности фоновых электронов соответственно.

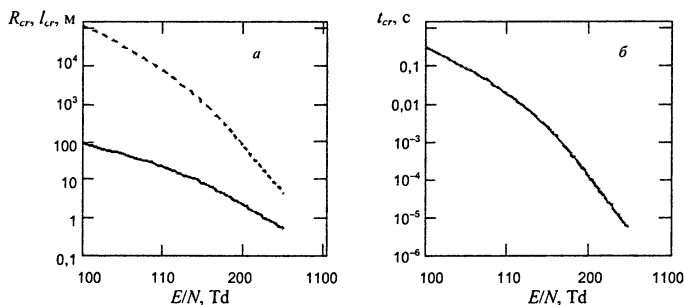


Рис. 7. Зависимость пространственных (а) и временных (б) характеристик электронной лавины от электрического поля к моменту времени, когда ее собственное электрическое поле сравнивается с внешним; R_{cr} (сплошная линия) — критический радиус лавины, l_{cr} (пунктир) — длина пути, пройденного лавиной, t_{cr} — соответствующее время жизни лавины.

В ночных условиях в мезосфере (ионосферная D -область 50—80 км) отсутствуют источники ионизации, и ионизованная компонента состоит из медленно рекомбинирующих положительных и отрицательных ионов. В то же время в результате тройных столкновений постоянно происходит отлипание и прилипание электронов. Малую добавку свободных электро-

нов в стационаре можно оценить для условий термодинамического равновесия из уравнения Саха:

$$\frac{[O_2] \cdot [n_e]}{[O_2^-]} = \frac{g_{O_2} g_e}{g_{O_2^-}} \left(\frac{\mu kT}{2\pi\hbar^2} \right)^{3/2} \exp\left(-\frac{E_1}{kT}\right), \quad (35)$$

где g_i — статистические веса состояний, E_1 — энергия связи (средство) электрона в отрицательном ионе, $\mu = mM/(m + M) \approx m$ — приведенная масса. Для O_2 средство электрона равно 0,6 эВ.

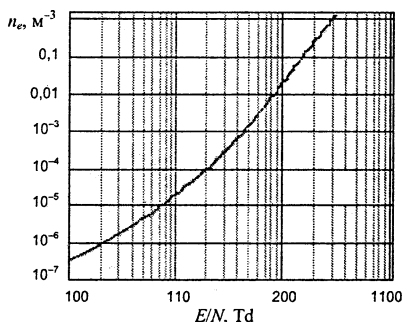


Рис. 8. Кривая на плоскости электрическое поле — фоновая плотность электронов, ниже которой перекрытие лавин в момент t_{cr} не происходит и возможен переход лавины в стримерную стадию. Выше указанной границы лавины перекрываются за время $t < t_{cr}$.

Обозначим $\lambda = n_i^- / n_e = [O_2^-] / n_e$. Тогда для λ имеем

$$\lambda = [O_2] \cdot \frac{g_{O_2^-}}{g_{O_2} g_e} \left(\frac{2\pi\hbar^2}{mkT} \right)^{3/2} \exp\left(\frac{E_1}{kT}\right) = h(T) \cdot [O_2]. \quad (36)$$

Очевидна сильная зависимость коэффициента $h(T)$ от температуры. Так, для трех характерных значений температур для высоты 80 км — 180, 200 и 220 К, имеем $h(T) \approx 1,1 \cdot 10^{-2}$, $2 \cdot 10^{-4}$ и $7,3 \cdot 10^{-6} \text{ см}^3$ соответственно. Примем для последующих оценок $T = 180 \text{ К}$, $[O_2] = 10^{14} \text{ см}^{-3}$ и, следовательно, $\lambda \approx 10^{12}$, что означает при $n_i \sim 10^3 \text{ см}^{-3}$ концентрацию электронов $n_e \sim 10^{-9} \text{ см}^{-3} = 0,001 \text{ м}^{-3}$. Из сопоставления этого значения с графиком на рис. 8 следует, что в данной ситуации при $E/N > 140 \text{ Тд}$ лавины могут непосредственно перерождаться в изолированные стримеры, а при $E/N < 140 \text{ Тд}$ порог для формирования стримеров не достигается, и для дальнейшей эволюции разряда требуется решать дополнительную электродинамическую задачу с учетом изменений электрического поля.

Таким образом, уже приближенное решение кинетического уравнения позволяет в замкнутом виде описать основные характеристики лавинной стадии разряда, определить условия дальнейшего развития разряда по тому или иному сценарию, а также найти реалистичные количественные оценки параметров, необходимые для планирования экспериментов, в том числе плотности разрядного тока, интенсивности электромагнитного излучения и т. д. Следует подчеркнуть, что приведенные на

рис. 7 данные дают представления о тонкой пространственной структуре разряда в диапазоне масштабов 1—10 м.

Что касается дальнейшей картины развития разряда, в частности его тонкой структуры, в наиболее простом и наглядном приближении она может быть описана с привлечением фрактальных моделей. В рамках подобной модели спрайт представляет собой сеть микроразрядов (стримеров) и развивается как самоподдерживающийся перколяционный процесс во внешнем поле. Учитываются различия критических полей для распространения положительных и отрицательных стримеров и релаксация электрического поля в мезосферной плазме. Недавняя модификация модели основана на развитии идей направленной перколяции. Полученная в результате моделирования картина развития разряда удовлетворительно описывает ряд наблюдаемых в экспериментах особенностей тонкой структуры и динамики развития спрайтов [20].

Заключение

За последнее десятилетие изучение высотных разрядов из экзотического занятия, оперирующего результатами разрозненных наблюдений, превратилось в обширное, стремительно развивающееся направление геофизической электродинамики. Огромный интерес к высотным разрядам в атмосфере связан не только с тем, что открытие по-настоящему нового явления — редкое событие в геофизике, но и с тем, что изучение высотных разрядов дало мощный импульс исследованию целого ряда других тесно связанных с ними геофизических проблем:

- свечение ночного неба, полярные сияния;
- электризация грозовых облаков;
- положительные вспышки в атмосфере (разряды типа облако-земля, переносящие положительный разряд на землю);
- высокоэнергичные процессы в атмосфере, роль космических лучей и убегающих электронов;
- мезомасштабная грозовая активность;
- глобальная электрическая цепь;
- химия средней атмосферы;
- турбулентность в верхней мезосфере и роль внутренних гравитационных волн в ее генерации;
- пыль и метеорные следы в верхней мезосфере;
- генерация низкочастотных электромагнитных волн и инфразвука крупномасштабными возмущениями в нижней ионосфере;
- процессы инициации разрядов в атмосфере;
- формирование лидеров и стримеров в различных условиях.

И хотя этап накопления экспериментальных данных, характеризующих морфологию этих явлений, отнюдь не завершен, в последнее время появилась возможность перейти к изучению особенностей структуры и

динамики высотных разрядов, поставить вопрос об их роли в глобальной электрической цепи и балансе малых составляющих атмосферы. Ближайшие перспективы в этой области связаны с проведением координированных наземных и космических экспериментов, а также с разработкой новых методов диагностики параметров высотных разрядов. Решение всех этих задач невозможно без последовательной нелинейной теории развития разряда в неоднородной атмосфере.

Литература

1. Sprites, Elves and Intense Lightning Discharges: NATO Science Series, V. 225 / Ed. by M. Fullekrug, E. Mareev, M. Rycroft. Springer, 2006. 398 p.
2. Gerken E. A., Inan U. S. A survey of streamer and diffuse glow dynamics observed in sprites using telescopic imagery // J. Geophys. Res. 2002. V. 107, № A11, 1344, doi:10.1029/2002JA009248.
3. McHarg M. G., Haaland R. K., Moudry D., Stenbaek-Nielsen H. C. Altitude-time development of sprites // J. Geophys. Res. 2002. V. 107, № A11, 1364, doi:10.1029/2001JA000283.
4. Hampton D. L., Heavner M. J., Wescott E. M., Sentman D. D. Optical spectral characteristics of sprites // Geophys. Res. Lett. 1996. V. 23, № 1. P. 89—92.
5. Green B. D. et al. Molecular excitation in sprites // Geophys. Res. Lett. 1996. V. 23, № 16. P. 2161—2164.
6. Shemansky D. E., Vallance Jones A. Type-B red aurora; The O_2^+ first negative system and the N_2 first positive system // Planet. Space Sci. 1968. V. 16. P. 1115—1130.
7. Омхольт А. Полярные сияния. М.: Мир, 1974. 248 с.
8. Armstrong R. A., Suszcynsky D. M., Lyons W. A., Nelson T. E. Multi-color photometric measurements of ionization and energies in sprites // Geophys. Res. Lett. 2000. V. 27, № 5. P. 653—656.
9. Cummer S. A., Inan U. S., Bell T. F., Barrington-Leigh C. P. ELF radiation produced by electrical currents in sprites // Geophys. Res. Lett. 1998. V. 25. P. 1281—1284.
10. Dowden D. L., Rodger C. J., Nunn D. Minimum sprite plasma density as determined by VLF scattering // IEEE Antennas and Propagation Magazine. 2001. V. 43, № 2. P. 12—24.
11. Mareev E. A., Evtushenko A. A., Yashunin S. A. On the modeling of sprites and sprite-producing clouds in the global electric circuit // Sprites, Elves and Intense Lightning Discharges: NATO: Science Series. V. 225 / Ed. by M. Fullekrug, E. Mareev, M. Rycroft. Springer, 2006. P. 313—340.
12. Гуревич А. В., Шварцбург А. Б. Нелинейная теория распространения радиоволн в ионосфере. М.: Наука, 1973. 272 с.
13. Борисов Н. Д., Гуревич А. В., Милых Г. М. Искусственная ионизированная область в атмосфере. М.: ИЗМИРАН, 1985. 184 с.
14. Yashunin S. A., Mareev E. A., Rakov V. A. Are lightning M-components capable of initiating sprites and sprite halos? // Geophys. Res. Lett. 2007 (accepted).
15. Базелян Э. М., Райзер Ю. П. Искровой разряд. М.: МФТИ, 1997. 320 с.
16. Голант В. Е., Жилинский А. П., Сахаров И. Е. Основы физики плазмы. М.: Атомиздат, 1977. 384 с.
17. Banks P. M., Chappell C. R., Nagy A. F. A new model for the interaction of auroral electrons with the atmosphere: spectral degradation, backscatter, optical emission and ionization // J. Geophys. Res. 1974. V. 79, № 10. P. 1459—1470.
18. Мак-Даниель И. Процессы столкновений в ионизованных газах. М.: Мир, 1967. 832 с.
19. Solov'yev A. A., Terekhin V. A., Tikhonchuk V. T., Altgilbers L. L. Electron kinetic effects in atmosphere breakdown by an intense electromagnetic pulse // Phys. Rev. E. 1999. V. 60, № 6. P. 7360—7368.
20. Hayakawa M., Iudin D. I., Mareev E. A., Trakhtengerts V. Yu. Cellular automation modeling of mesospheric optical emissions, sprites // Phys. Plasmas. 2007 (accepted).

ФРАКТАЛЬНЫЕ ЛАБИРИНТЫ: СТРУКТУРНАЯ ДИНАМИКА

Д. И. Иудин, В. Ю. Трахтенгерц

Фрактальные лабиринты встречаются в природе и технике повсеместно. Это транспортные и коммуникационные сети, системы сбора и распределения ресурсов и информации, речные системы, системы кровоснабжения, молниевые разряды и т. д. Фрактальный характер геометрии этих систем предопределяет особенности их динамического поведения и их транспортных свойств.

1. Блуждания в лабиринтах: аномальная диффузия

Напомним, что физическое содержание понятия размерности, в том числе и фрактальной, определяется изменением массы объекта (или числа составляющих его элементов) $M(L)$ с ростом его линейных размеров L . Если мы рассмотрим малую часть системы с размерами λL ($\lambda < 1$), то для массы фрагмента получим

$$M(\lambda L) = \lambda^d M(L). \quad (1)$$

Решение функционального уравнения (1) имеет простой вид: $M(L) = \text{const} \cdot L^d$. Так, масса длинного провода меняется линейно с λ , т. е. $d = 1$. Для тонкой пластины мы найдем, что $d = 2$, а для бруска $d = 3$. Для фрактальных объектов величина размерности d_f оказывается меньше топологической размерности или размерности объемлющего евклидова пространства $d_f < d$. Таким образом, плотность фрактальных структур уменьшается по степенному закону с ростом их линейных размеров. Благодаря этому транспортные свойства фрактальных структур заметно отличаются от регулярного случая. Мы убедимся в этом, рассматривая проводимость и случайные блуждания в самоподобных лабиринтах.

Проводимость. Начнем с регулярного случая и рассмотрим металлическую решетку размера L^d , к противоположным граням которой приложено напряжение. При изменении линейных размеров L системы ее полное сопротивление изменяется следующим образом:

$$\rho \sim \sigma^{-1} \frac{L}{L^{d-1}}, \quad (2)$$

где σ — проводимость металла. Поскольку σ не зависит от L , полное сопротивление решетки меняется при изменении размеров по степенному закону $\rho \sim L^{2-d} \equiv L^{\tilde{\zeta}}$ с показателем $\tilde{\zeta} = 2 - d$.

Уравнение (2) справедливо для однородных проводников, когда плотность проводящего материала постоянна. Для фрактальных проводников плотность пропорциональна L^{d_f-d} и стремится к нулю при $L \rightarrow \infty$. Это связано с существованием во фрактальных структурах дыр самых различных масштабов вплоть до размера самой системы. Увеличивая L , мы тем самым увеличиваем и размеры непроводящих дыр, что, в свою очередь, уменьшает проводимость. Благодаря самоподобию σ уменьшается на всех масштабах, следуя степенной зависимости

$$\sigma \sim L^{-\tilde{\mu}}. \quad (3)$$

Как следствие (3) и (2), полное сопротивление ведет себя как

$$\rho \sim L^{\tilde{\zeta}}, \quad (4)$$

где критический индекс $\tilde{\zeta} = 2 - d + \tilde{\mu}$ теперь больше, чем аналогичный показатель $2 - d$ для однородного проводника.

Диффузия. Диффузию моделируют обычно процессом случайных блужданий, когда на каждом шаге модельного времени частица смещается на дистанцию a в случайно выбранный соседний узел d -мерной модельной решетки. Предположим, что в момент времени $t = 0$ частица стартовала из начала координат d -мерного решеточного пространства. Тогда после некоторого числа t шагов модельного времени текущая позиция диффундирующей частицы будет описываться вектором смещения

$$r(t) = a \sum_{k=1}^t e_k, \quad (5)$$

где e_k обозначает единичный вектор в направлении прыжка на k -м шагу модельного времени.

Характерная дистанция, на которую частица смещается в среднем в процессе случайных блужданий за t шагов модельного времени, описывается среднеквадратичным смещением $\langle r^2(t) \rangle^{1/2}$, где среднее $\langle \dots \rangle$ понимается как среднее по всевозможным конфигурациям случайных блужданий на решетке. Из (5) мы находим

$$\langle r^2(t) \rangle = a^2 \sum_{k,k'=1}^t \langle e_k \cdot e_{k'} \rangle = a^2 t + \sum_{k \neq k'}^t \langle e_k \cdot e_{k'} \rangle.$$

Поскольку прыжки частицы в различные моменты времени k и k' некоррелированы друг с другом, мы имеем $\langle e_k \cdot e'_k \rangle = \delta_{kk'}$, что приводит нас в результате к закону диффузии Фика:

$$\langle r^2(t) \rangle = a^2 t. \quad (6)$$

Заметим, что (6) не зависит от размерности d решетки. В более общем случае, когда частица имеет вероятность остаться на месте, среднеквадратичное смещение описывается формулой $\langle r^2(t) \rangle = 2dDt$, где D — коэффициент диффузии.

Среднеквадратичное смещение можно обнаружить, рассматривая плотность вероятности $P(r, t)$, которая есть вероятность найти диффундирующую частицу в пределах радиуса r от точки старта посредством соотношения $\langle r^2(t) \rangle = \int dr r^2 P(r, t)$.

Плотность вероятности для случайных блужданий легко может быть найдена для одномерной цепочки узлов, когда частица может совершать прыжки только в двух направлениях. Прыжки вправо или влево совершаются с вероятностью $p = 1/2$. Для простоты предположим, что $a = 1$.

В процессе случайных блужданий за t шагов модельного времени частица совершит m прыжков вправо и $t - m$ прыжков влево, таким образом, ее текущей координатой будет $x = m - (t - m) = 2m - t$. Вероятность того, что это произойдет, описывается биномиальным распределением:

$$P(m, t) = \binom{t}{m} p^m (1-p)^{t-m} = \binom{t}{m} \left(\frac{1}{2}\right)^t. \quad (7)$$

Для больших t , m , и $t - m$ биномиальные коэффициенты могут быть оценены по формуле Стирлинга: $t! \simeq (2\pi t)^{1/2} (t/e)^t$ и так далее. Поскольку текущая координата частицы x много меньше полного числа ее прыжков вправо или влево, результирующее выражение может быть разложено в ряд Тейлора по степеням величины $x/t \ll 1$, и окончательно мы получим

$$P(x, t) \simeq \frac{1}{(2\pi t)^{1/2}} e^{-x^2/2t}. \quad (8)$$

Соотношение (8) представляет собой распределение Гаусса с шириной t , которая идентична $\langle x^2(t) \rangle$. Поскольку $P(x, t)$ представляет вероятность, то $\int_{-\infty}^{\infty} dx P(x, t) = 1$, что выполняется для (8). В более общем случае d -мерной гиперрешетки (5) переходит в следующее выражение:

$$P(r, t) \simeq \frac{1}{\langle r^2(t) \rangle^{d/2}} e^{-(d/2)r^2/\langle r^2(t) \rangle}, \quad (9)$$

где $\langle r^2(t) \rangle = 2dDt$, и вероятность возвращения частицы в начало координат $P(0, t)$ через t шагов модельного времени пропорциональна $\langle r^2(t) \rangle^{-d/2}$.

Рассмотрим теперь, что же происходит при блужданиях в самоподобных лабиринтах. Очевидно, что благодаря большим дырам, перетяжкам и мертвым концам (тупикам) фрактала движение по нему блуждающей частицы замедляется. Причем поскольку дыры, перетяжки и мертвые концы обнаруживают себя на всех масштабах, то на всех масштабах осуществляется и замедление. Закон Фика (6) утрачивает свою дееспособность, и на смену ему приходит более общая формула для среднеквадратичного смещения:

$$\langle r^2(t) \rangle = t^{2/d_\omega}, \quad (10)$$

где новый критический показатель d_ω (показатель диффузии — *diffusion exponent* или фрактальная размерность случайных блужданий — *fractal dimension of the random walk*) всегда больше 2.

Критический индекс сопротивления $\tilde{\zeta}$ и показатель диффузии d_ω могут быть связаны уравнением Эйнштейна

$$\sigma = \frac{e^2 n}{k_B T} D, \quad (11)$$

которое связывает проводимость постоянного тока σ системы с коэффициентом диффузии D . Здесь e и n обозначают соответственно заряд и концентрацию подвижных частиц.

Для нахождения связи показателя d_ω с величинами $\tilde{\zeta}$ и $\tilde{\mu}$ мы используем обычную скейлинговую аргументацию. Поскольку концентрация частиц n пропорциональна плотности субстрата ($n \sim L^{d_f-d}$), то правая часть в (11) пропорциональна $L^{d_f-d} t^{2/d_\omega-1}$. Левая часть (11) пропорциональна $L^{-\tilde{\mu}}$. Так как время, затрачиваемое диффундирующей частицей на преодоление дистанции длиной L , меняется как L^{d_ω} , мы окончательно находим, что $L^{-\tilde{\mu}} \sim L^{d_f-d+2-d_\omega}$, откуда следует искомое соотношение:

$$d_\omega = d_f - d + 2 + \tilde{\mu} = d_f + \tilde{\zeta}. \quad (12)$$

Для случайных фракталов нахождение величины d_ω далеко не так просто. Исключением являются топологически линейные фракталы с химической размерностью равной единице ($d_l = 1$), которые могут быть рассмотрены как непересекающиеся пути. Вдоль такого пути (в химическом или l -пространстве) диффузия нормальна и, следовательно, $\langle l^2(t) \rangle = t$. Поскольку $l \sim r^{d_f}$, среднеквадратичное смещение

в евклидовом r -пространстве ведет себя как $\langle r^2(t) \rangle = t^{1/d_f}$, что дает нам в этом случае $d_w = 2d_f$.

Для случая перколяционного фрактального кластера величина d_w не может быть вычислена точно, однако ее нижняя и верхняя оценки довольно близки друг к другу в пространствах размерностью $d \geq 3$. Хорошую оценку дает гипотеза Александра — Орбаха [1]: $d_w \simeq 3d_f/2$.

Для фракталов функция распределения $\langle P(r, t) \rangle$, усредненная по всем точкам старта частицы на фрактале, уже не является гауссовой. Следуя аргументации Фишера [2] и Домба [3], предложенной ими для описания самоизбегающих блужданий, можно обнаружить, что $\langle P(r, t) \rangle$ для $r/\langle r^2 \rangle^{1/2} \gg 1$ описывается растянутым гауссовским распределением:

$$\langle P(r, t) \rangle / \langle P(0, t) \rangle \sim \exp[-\text{const} \times (r/\langle r^2 \rangle^{1/2})^u], \quad (13)$$

где показатель u связан с d_w соотношением

$$u = \frac{d_w}{d_w - 1}. \quad (14)$$

Соотношение (14) было проверено численно для огромного числа случайных фракталов, включая перколяционные кластеры. Для топологически линейных фракталов оно выполняется точно. Поскольку на больших временах каждый элемент фрактальной структуры посещается блуждающей частицей примерно с одинаковой вероятностью, возникающая в (13) вероятность возвращения $\langle P(0, t) \rangle$ обратно пропорциональна числу посещенных узлов $S(t)$. Для фракталов $S(t)$ меняется как $\langle r^2(t) \rangle^{d_f/2}$ и, таким образом,

$$\langle P(0, t) \rangle \sim (1/\langle r^2(t) \rangle)^{d_f/2} \sim t^{-d_f/d_w}. \quad (15)$$

2. Фрактальная динамика фильтрационных течений

Содержательным примером динамических явлений в самоподобных лабиринтах служат фильтрационные течения в пористых или дисперсных средах, поровое пространство которых находится вблизи порога перколяционного фазового перехода [4]. Здесь мы обсудим результаты экспериментальной проверки выводов континуальной теории перколяции для фильтрационных течений в дисперсных средах с переменной пористостью. Основное утверждение теории о том, что вблизи порога протекания активная пористость и проницаемость системы

изменяются по скейлинговому закону, было подтверждено в серии экспериментов, где полная пористость модельной среды изменялась под действием одноосного сжатия в цилиндрическом стакане [14].

Пористая среда была представлена неправильными кубиками полиуретана, занимающими объем вертикального цилиндрического стакана между проницаемой перегородкой и проницаемым поршнем. При диаметре цилиндра 0,2 м размеры частичек полиуретана в различных экспериментах составляли 0,7 мм и 5 мм. Перемещением поршня относительно перегородки достигается изменение полной пористости

$$x = 1 - \frac{M}{\rho S h}, \quad (16)$$

где M — полная масса полиуретановых кубиков в стакане, ρ — плотность полиуретана, h — расстояние между проницаемой перегородкой и проницаемым поршнем, S — площадь сечения стакана. Поддон цилиндрического стакана сообщается с резервуаром, поднятым на высоту H относительно уровня стока, так что перепад давления, вызывающий фильтрацию в установке, описывается соотношением

$$\nabla p - \rho_0 \vec{g} = \rho_0 \vec{g} \frac{H}{h}, \quad (17)$$

где p — давление в воде, ρ_0 — плотность воды, \vec{g} — ускорение свободного падения. В одномерной геометрии эксперимента скорость фильтрации определяется расходом θ :

$$u = \frac{\theta}{S}. \quad (18)$$

При каждом фиксированном значении пористости (фиксированное значение h) градиент давления в системе изменялся от нуля до величины порядка 10^5 Па/м. Соответствующие экспериментальные зависимости скорости фильтрации от величины перепада давления для различных значений пористости представлены на рис. 1. Экспериментальные точки хорошо ложатся на теоретические зависимости, полученные с помощью двухчленного закона фильтрации [5]

$$-\frac{k}{\mu} \rho_0 \vec{g} \frac{H}{h} = \vec{u} (1 + \text{Re}), \quad (19)$$

где

$$\text{Re} = \frac{\rho_0 d_0 u}{\mu} \quad (20)$$

— число Рейнольдса для фильтрационного течения; μ — вязкость воды; k — проницаемость; d_0 — параметр с размерностью длины, характеризующий мелкомасштабную структуру порового пространства.

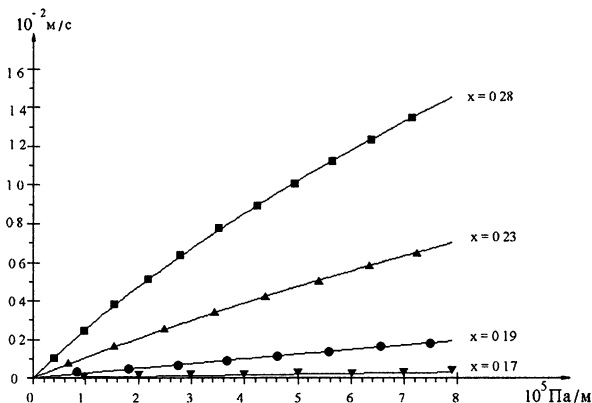


Рис. 1

2.1. Универсальная форма нелинейного закона фильтрации

В экспериментах с монодисперсными ансамблями кубиков с разными характерными размерами (относительные вариации размеров в каждом ансамбле не превышали 20%) система становилась непрозрачной для фильтрационного течения при одном и том же критическом значении полной пористости $x_c = 0,16 \pm 0,1$.

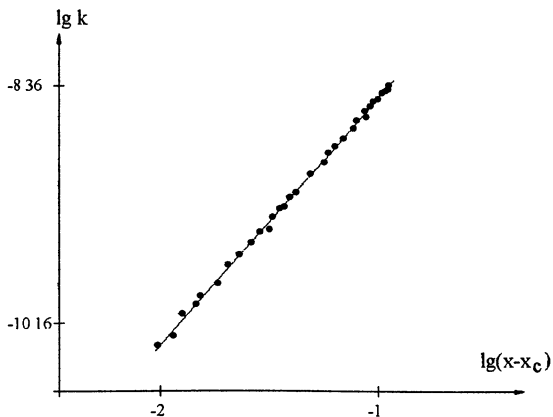


Рис. 2

На рис. 2 проницаемость как функция полной пористости представлена в двойном логарифмическом масштабе. Полученные в экспе-

рименте значения проницаемости хорошо аппроксимируются степенной зависимостью

$$k = k_0(x - x_c)^{1,7 \pm 0,1}, \quad (21)$$

где x_c — критическое значение пористости, $t = 1,7 \pm 0,1$ — критический индекс проницаемости. Величина k_0 пропорциональна квадрату размеров частичек в ансамбле, причем в пределах экспериментальной точности $k_0 = d_0^2$. Величина d_0 не меняется сколько-нибудь заметно при изменении пористости вблизи порога. Это связано с тем, что инерционные эффекты, приводящие к отклонению от линейного закона Дарси, определяются локальными особенностями порового пространства и слабо зависят от его крупномасштабной геометрии. В то же время изменение полной пористости сказывается прежде всего именно на глобальной структуре порового пространства, в меньшей степени затрагивая его локальные геометрические параметры. Поэтому, используя соотношение (21) и вспоминая, что активная пористость m является параметром порядка перколяционного перехода и изменяется вблизи порога по скейлинговому закону, т. е.

$$m \cong (x - x_c)^\beta, \quad (22)$$

где β — критический индекс параметра порядка [6], получаем

$$\text{Re} = \frac{\nu \rho k^{1/2}}{\mu m^{t/2\beta}}. \quad (23)$$

В свое время неоднократно делались попытки выразить число Рейнольдса Re (или характерный размер d_0) через параметры дисперсной системы таким образом, чтобы процесс фильтрации в пористых средах различной структуры описать единой формулой [5]. Результат (23) обеспечивает универсальную форму записи двухчленного закона фильтрации (19) для широкого класса дисперсных сред, объединенных общей геологической историей.

Наблюдаемое в эксперименте критическое значение $x_c = 0,16 \pm 0,1$ является стандартом для трехмерных задач теории протекания, попадающих в один класс универсальности с такими классическими объектами, как задача твердых сфер и трехмерная задача об уровне протекания [6]. В этих задачах предполагается случайно-однородное распределение в пространстве перколирующих элементов. Сильные корреляции в распределении могут значительно изменить величину порога. Достаточно, например, заменить кубики на шарики, для того чтобы в распределении пор возникли сильные корреляции, обусловленные существованием ближнего порядка в ансамбле сферических частиц. При

этом критическое значение понижается до величины порядка 3%, что не сильно отличается от величины теоретического порога для случая регулярной ГЦК упаковки шаров [7, 8].

2.2. Бидисперсный ансамбль

Другим примером резкого понижения порога по сравнению с задачей со случайно-однородной функцией распределения могут служить фрактальные структуры, где медленно спадающие корреляции в ансамбле пор обусловлены самоподобием системы. В данной работе экспериментально был исследован случай, когда систему образовывали два типа кубиков с сильно различающимися размерами (0,7 мм и 0,5 мм). Структура ансамбля контролируется параметром

$$\alpha = \frac{q}{q + Q}, \quad (24)$$

где q, Q — удельные объемы соответственно маленьких и больших кубиков, при этом

$$q + Q = 1 - x. \quad (25)$$

Очевидно, что значения $\alpha = 0$, $\alpha = 1$ соответствуют монодисперсным ансамблям. В пределе сильно различающихся размеров легко получить аналитические выражения для критического значения полной пористости как функции структурного параметра α . Действительно, при значениях $q \ll 1$ ($\alpha \ll 1$) наибольшее количество маленьких кубиков расположено в порах структуры, образованной ансамблем больших кубиков. Очевидно, что в этом случае критическое значение пористости линейно падает с увеличением q : $x_c(q) = x_c(0) - q$, или с использованием (24), (25):

$$x_c(\alpha) = \frac{x_c(0) - \alpha}{1 - \alpha}. \quad (26)$$

Наоборот, когда $Q \ll 1$ ($\alpha < 1$), изменение критического значения полной пористости обусловлено уменьшением размеров системы на величину объема, занимаемого большими кубиками, т. е. $x_c(Q) = (1 - Q)x_c(0)$. Используя опять (24), (25), получаем

$$x_c(\alpha) = \frac{x_c(0) \cdot \alpha}{1 + x_c(0)\alpha - x_c(0)}. \quad (27)$$

Асимптотики (26), (27) представлены на рис. 3 сплошными линиями. Видно, что при $\alpha_c = \frac{x_c(0)}{1 + x_c(0)}$ происходит своеобразный геометрический фазовый переход от режима (26) к режиму (27). Минимальное

значение критического значения полной пористости составляет величину $x_c(\alpha_c) = x_c^2(0)$. Результаты экспериментов представлены квадратами на рис. 3.

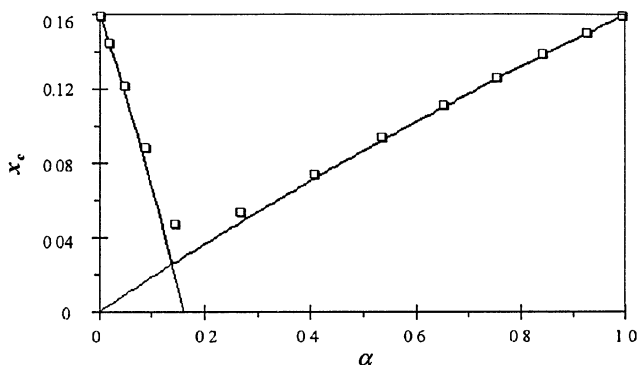


Рис. 3

2.3. Скорость распространения фронта возмущения

В этом разделе представлены результаты экспериментального определения размерности «геодезической» перколяционного кластера D_g .

Дело в том, что размерность геодезической тесно связана с другим критическим показателем, имеющим содержательную аналогию в классической термодинамике. Речь идет о показателе критического замедления. Известно, что критическое замедление является неотъемлемым атрибутом любого фазового перехода. Обсудим смысл этого явления для перколяционного превращения и параллельно выясним связь между показателем критического замедления и величиной D_g .

Пусть имеется некоторое возмущение, распространяющееся следующим образом. Если в момент времени t возбужден узел i , то в момент времени $t + 1$ возбуждаются все соседние с i доступные узлы. Узел, который когда-либо был возбужден, остается возбужденным во все последующие моменты времени. Пусть в момент времени $t = 0$ мы возбудили узел 0. Фронт возбуждения в момент времени t представляет собой совокупность узлов, химическое расстояние от которых до узла 0 равно t .

При $x < x_c$ возмущение в системе не распространяется. При любом $x > x_c$ асимптотическое поведение фронта описывается в пределе

больших времен некоторой, зависящей от x скоростью. При приближении x к x_c сверху эта скорость уменьшается: имеет место критическое замедление. При $x = x_c$ скорость возмущения обращается в 0.

Пусть концентрация доступных узлов равна $x > x_c$. Длина корреляции в такой системе равна $\xi = \tau^{-\nu}$. Во фрактальном режиме время, за которое фронт возмущения переместится на расстояние R , равно $t \sim R^{D_g}$, в частности, расстояние ξ будет преодолено за время $t_\xi \sim \xi^{D_g} \sim \tau^{D_g \nu}$. На больших масштабах поведение системы становится однородным, при этом $t/t_\xi = R/\xi$. Таким образом, за время $t \gg t_\xi$ фронт возмущения распространится на расстояние $R \sim (\xi/t_\xi) t$; эффективная скорость его распространения $v \sim \xi/t_\xi \sim \tau^{(D_g-1)\nu}$. Показатель $\psi = (D_g - 1)\nu$ эффективной скорости определяет величину критического замедления. В трехмерном случае $\psi = 0,31 \pm 0,06$.

В экспериментах с монодисперсными ансамблями измерялась скорость распространения фронта возмущения, вызванного резким включением градиента давления в системе. При приближении к порогу прозрачности эта скорость стремилась к нулю как степень величины $x - x_c$. Соответствующий критический показатель, определяющий величину критического замедления в системе, составил значение $\varphi = 0,31 \pm 0,02$, которое совпадает с результатами компьютерных экспериментов [9].

В заключение рассмотрим простейшее неоднородное обобщение задачи, которое касается известной ситуации отклонения петрофизических данных, получаемых, например, в процессе бурения скважин, от линии «нормального уплотнения», т. е. возникновения зон так называемого «аномального давления», см., например, [12, 13]. Обобщением является зависимость критического параметра от одной пространственной координаты. Именно такая ситуация возникает при фильтрации флюида в пористой среде, испытывающей литостатическое давление. Под действием собственного веса материал изменяет свою пористость, которая на некотором горизонте достигает своего критического значения. Ниже этого горизонта система непроницаема для флюида, причем зоны прозрачности отделены друг от друга флуктуационным фронтом, структура и ширина которого могут быть определены в рамках приближения градиентной перколяции [15–17].

Более интересная для приложений модель возникает тогда, когда полная пористость системы является случайной функцией пространственных координат. При этом целесообразно рассматривать не кластеры пор, а перколяцию зон прозрачности. Присутствие в системе градиентов среднего значения полной пористости обеспечивает здесь формирование иерархии флуктуационных фронтов. Заметим, что аналогичные флуктуационные фронты могут возникнуть из-за пространственной неоднородности критического значения полной пористости.

Более того, когда пористость становится динамической переменной, изменения которой во времени обусловлены геохимическими и геофизическими процессами, в системе возникают эффекты, связанные с нелинейностью пористость — проницаемость вблизи порога перколяции [18, 19]. При этом флуктуационный фронт становится активным элементом распределенного стохастического генератора, например, генератора шума [20].

3. Самоорганизующиеся лабиринты

Динамическим развитием идей перколяции можно считать ситуацию, когда критический параметр перколяционного перехода сам становится внутренней динамической переменной 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, когда система автоматически эволюционирует к состоянию динамического равновесия, близкому к порогу перколяции. В наших предыдущих лекциях [10, 11] мы рассказывали об эффектах динамической перколяции в потенциальном поле на примере молниевых разрядов. Теперь речь пойдет о динамическом скейлинге самоорганизующихся дренажных сетей. Самоорганизующиеся дренажные системы возникают из перколирующих лабиринтов при наличии внешнего поля и по своим структурным свойствам очень напоминают дренажные системы бассейнов крупных рек. Поэтому мы начнем с принятого в гидрологии описания стационарной структуры потоков дренажных систем речных бассейнов, которое опирается на совокупность эмпирических законов Хака, Хортон и Токунаги. Сразу подчеркнем, что существенным отличием самоорганизующихся дренажных сетей от стационарной структуры речных потоков является динамический характер их существования.

3.1. Законы гидрологии

Дренажные системы являются универсальным свойством земных ландшафтов. Маленькие потоки, сливаясь, образуют более крупные потоки, крупные потоки, сливаясь, образуют реки и так далее. Территория, с которой река собирает потоки воды и уносит их в конечном счете в океан, собственно и образует пойму или бассейн реки.

Характеристиками бассейна являются его площадь (или объем — в более общей ситуации) V , длина основного потока l и его габариты L_{\parallel} , L_{\perp} . Известно несколько скейлинговых соотношений — эмпиричес-

ких законов, связывающих между собой эти величины. Наиболее известным из них является закон Хака, который утверждает, что длина основного потока l меняется с увеличением площади бассейна как

$$l \sim V^h, \quad (28)$$

где h именуется обычно показателем Хака. Помимо (28) известен еще целый ряд соотношений, связывающих площадь бассейна V , длину его основного потока l и его поперечные габариты L_{\perp} с характерным продольным размером дренажной системы бассейна $L(=L_{\parallel})$:

$$V \sim L^{\mathfrak{D}}, \quad l \sim L^{\mathfrak{D}}, \quad L_{\perp} \sim L^{\mathfrak{F}}. \quad (29)$$

Кроме того, было обнаружено, что гистограммы площадей бассейнов и длин основных потоков для каждого отдельного ландшафта имеют степенное распределение:

$$\mathfrak{G}(V) \sim V^{-\nu}, \quad \mathfrak{G}(l) \sim l^{-\mu}. \quad (30)$$

Еще задолго до появления фрактальной парадигмы количественное описание упорядоченной системы потоков было введено Хортоном и Стралером. Представленные на стандартной топографической карте потоки, не имеющие притоков в направлении против течения, определяются как потоки первого порядка ($i = 1$). Когда два потока первого порядка соединяются, они образуют поток второго порядка ($i = 2$). Когда два потока второго порядка соединяются, они образуют поток третьего порядка ($i = 3$) и т. д. Кроме того, потоки первого порядка ($i = 1$) могут вливаться в потоки второго ($i = 2$), третьего ($i = 3$) порядка и любые другие потоки более высокого порядка. Аналогично потоки второго порядка ($i = 2$) могут соединяться с потоками третьего порядка и т. д. Суммируя сформулированные выше правила, можно констатировать, что, когда поток порядка i_1 сливается с потоком порядка i_2 , результирующий поток имеет порядок

$$i = \max(i_1, i_2) + \delta_{i_1, i_2}, \quad (31)$$

где δ — символ Кронеккера. Другими словами, порядок увеличивается только при слиянии двух равнозначных потоков.

Хортон определил бифуркационный коэффициент как отношение

$$R_b = \frac{N_i}{N_{i+1}}, \quad (32)$$

где N_i — число потоков порядка i , и коэффициент упорядочения длин как отношение

$$R_r = \frac{r_{i+1}}{r_i}, \quad (33)$$

где r_i — средняя длина потока i -го порядка. Анализ реальных гидрологических данных обнаружил приблизительное постоянство коэффициентов R_b и R_r при изменении порядка потоков в любой дренажной системе. Эти эмпирические факты называют законами Хортонa. Используя итерационное определение фрактальной размерности

$$d_f = \frac{\ln(N_i/N_{i+1})}{\ln(r_{i+1}/r_i)} \quad (34)$$

и подставляя в него соотношения (32) и (33), получаем выражение для фрактальной размерности дренажной системы:

$$d_f = \frac{\ln R_b}{\ln R_r}. \quad (35)$$

Другими словами, справедливость законов Хортонa означает, что дренажные сети являются фрактальными деревьями.

3.2. Динамическое обобщение законов Хака

В основе предлагаемого динамического обобщения законов (28), (29), (30) лежит фундаментальный факт скейлинговой зависимости времени жизни фрагмента дренажной системы от его характерных пространственных размеров. По- существу, речь идет о расширении и перенесении свойства самоаффинности с пространственного континуума на пространственно-временной. Соотношения (28), (29), (30) должны быть дополнены скейлинговой зависимостью времени жизни дренажной системы от продольного размера $L (= L_{\parallel})$:

$$t \sim L^{-\xi}, \quad \mathfrak{G}(t) \sim t^{-\lambda}. \quad (36)$$

По аналогии с гидрологическими параметрами мы можем ввести коэффициент упорядочения времени жизни по формуле

$$R_\tau = \left(\frac{\tau_{i+1}}{\tau_i} \right)^{\delta_\tau}, \quad (37)$$

где τ_i — среднее время жизни потока i -го порядка, а δ_τ — показатель, вообще говоря, отличный от единицы. Анализ показал, что коэффициент упорядочения времени жизни, при соответствующем подборе δ_τ , демонстрирует приблизительное постоянство при изменении порядка потоков.

Выше мы уже подчеркивали, что для построения непротиворечивого алгоритма формирования иерархической системы токов необходимо рассматривать векторную активацию. При этом возникающие

в системе проводящие связи характеризуются еще и направлением текущего по ним тока. С точки зрения теории графов, возникающие в нашей модели проводящие кластеры представляют собой взвешенные ориентированные деревья, т. е. графы, ребра которых имеют определенную ориентацию — направление от одной вершины к другой и характеризуются величиной протекающего вдоль этого направления тока.

Каждой вершине поставим в соответствие пару чисел (a, b) , равных соответственно числу входящих и исходящих ребер. Величина $a+b-1$ совпадает, очевидно, с индексом ветвления. Вся периферия при этом разделится на две группы: вершины типа $(1, 0)$ и типа $(0, 1)$. Возникновению критической пары соответствует рождение пары смежных вершин, одна из которых относится к типу $(1, 0)$, а другая к типу $(0, 1)$. Далее, в случае векторной активации, состояние активной периферийной вершины типа $(1, 0)$ переходит в состояние типа $(1, n)$, где n — число вновь активированных соседей рассматриваемой вершины. Аналогично состояние активной периферийной вершины типа $(0, 1)$ переходит в состояние типа $(n, 1)$. Таким образом, вся совокупность вершин эволюционирующей структуры распределяется между типами $(1, n)$ и $(n, 1)$ с целыми неотрицательными n .

При векторном характере активации исходная критическая пара вершин почти всегда остается в составе ствола формирующегося кластера. Стволом называется неразветвленная центральная часть структуры, представляющая собой ориентированную цепь, состоящую из вершин типа $(1, 1)$ и инцидентных этим вершинам ребер. Ствол заканчивается парой вершин, одна из которых относится к типу $(1, n)$, а другая к типу $(n, 1)$.

Вершины типа $(1, n)$ с $n \geq 2$ и инцидентные им ребра формируют крону ориентированного дерева, а вершины типа $(n, 1)$ и инцидентные им ребра формируют корень. Характерная конфигурация проводящего дерева, толщина ребер которого пропорциональна величине проходящего по ним тока, показана на рис. 4 (моделирование проводилось на кубической решетке $65 \times 65 \times 65$).

Очевидно, что не все структурные элементы формирующегося проводящего кластера равнозначны с точки зрения проходящего по ним тока. Например, ток в стволе кластера существенно превосходит токи на периферии, причем по мере движения от периферийных элементов к стволу растет не только сила тока, но и время его жизни. Существенно и обратное влияние токов на геометрию формирующегося кластера. Самосогласованный учет этого влияния целесообразно проводить в терминах иерархии системы токов. Введем следующие правила построения иерархии.

1. Рассматриваем крону ориентированной структуры (тип $(1, n)$). Всем ребрам кроны присваиваем уровень иерархии равный единице: $h = 1$.

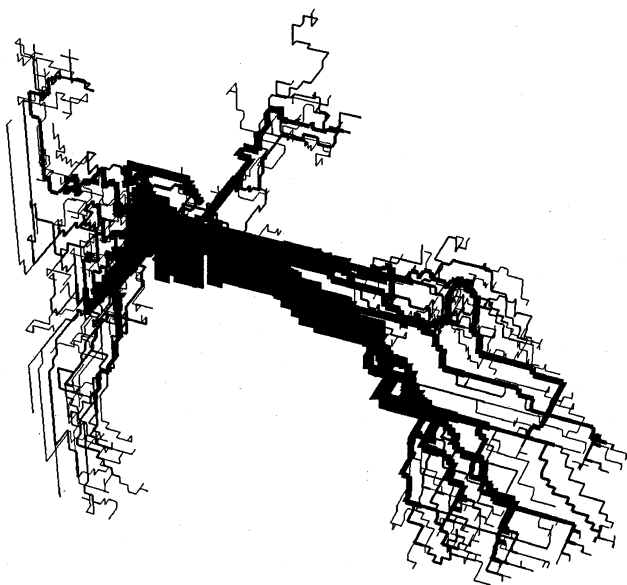


Рис. 4

2. Движемся от выбранной на периферии вершины к стволу. Пропускаем все вершины с $n = 1$. Останавливаемся на вершинах с индексом ветвления $r \geq 2$.

3. Повторяем предыдущий шаг для всех периферийных вершин, принадлежащих кроне (тип (1, 0)).

4. Отсекаем пройденные ребра, а точки остановки объявляем новой периферией (ренормализация). Уровень иерархии оставшихся ребер кроны увеличивается на единицу.

5. Повторяем предыдущие три шага процедуры до тех пор, пока не дойдем до ствола.

6. Запоминаем уровень иерархии ствола $h_t^- = \max(h)$. Знак минус означает здесь, что вычисления относятся к элементам кроны.

7. Повторяем все предшествующие шаги для корня структуры и на последнем этапе запоминаем уровень иерархии ствола $h_t^+ = \max(h)$, вычисленный по элементам корня структуры.

8. Присваиваем стволу уровень иерархии, соответствующий максимальному из значений h_t^- и h_t^+ .

9. Если $h_t^- > h_t^+$, уровни иерархии всех элементов корня увеличиваем на $h_t^- - h_t^+$. При выполнении противоположного условия уровни иерархии всех элементов кроны увеличиваем на $h_t^+ - h_t^-$.

Суть последнего шага процедуры заключается в том, что иерархическая значимость ствола полностью определяется весом максимальной

ной из опирающихся на него конструкций (корня или кроны). Рецессивная часть структуры может попросту отсутствовать (разряд на землю). Этот шаг не выполняется при $h_t^- = h_t^+$, т. е. в случае иерархической симметрии между кроной и корнем.

Уровень иерархии связей является дополнительным параметром, который расширяет возможности рассмотренной выше базовой модели. Изменим прежде всего правило, по которому исчезают проводящие фрагменты кластера. Если периферийная вершина не находит активированных соседей и инцидентное ей ребро имеет минимальный уровень иерархии $h = 1$, то она отмирает вместе с соответствующим ей проводящим ребром, передавая роль периферии ближайшей своей предшественнице, если эта последняя не является точкой ветвления. Если же уровень иерархии инцидентного рассматриваемой вершине ребра $h \geq 2$, активные свойства периферийной вершины (т. е. способность иницировать соседей) сохраняются.

Благодаря описанной модернизации при возникновении асимметрии появляется интересный эффект: элементы рецессивной субструктуры становятся долговителями — они не могут исчезнуть, пока соответствующий им уровень иерархии превышает единицу. Конец этому может положить только деградация доминантной субструктуры до уровня рецессивной.

Заключение

Выше мы определили две проблемы, которые, по нашему мнению, вызывают наибольший интерес в области фрактальной динамики. Это, во-первых, формирование и распространение фронтов возмущений в средах, обладающих фрактальными свойствами. Здесь наибольший интерес вызывают процессы вблизи порога перколяции. В реальных системах, где, как правило, имеют место регулярные градиенты субстанций и крупномасштабные потенциальные поля, большую роль приобретают эффекты градиентной и направленной перколяции. Дальнейшее развитие математических моделей этих процессов поставило бы на количественную основу описание многих реальных процессов, таких как формирование динамических границ между фазами с контрастными физическими свойствами, развитие электрических разрядов или зон трещиноватости в протяженных неоднородных средах и др. Вторая актуальная проблема — это самоорганизующиеся лабиринты, в частности динамические дренажные системы с иерархической нелинейной перестройкой токовых каналов. Здесь применение теории графов представляется наиболее плодотворным.

Литература

1. Alexander S., Orbach R. // J. Phys. Lett. 1982. 43:L625.

2. Fisher M. E. and Burford F. R. J. // Phys. Rev., 156:583, 1967.
3. Domb C. // Adv. Chem. Phys. 1969. 15:229.
4. Федер Е. Фракталы. М.: Мир, 1991.
5. Барентблатт Г. И., Ентов В. М., Рыжик В. М. Движение жидкостей и газов в природных пластах. М.: Недра, 1984. 211 с.
6. Шкловский Б. И., Эфрос А. Л. // УФН. 1975. Т. 117. С. 401.
7. Roberts J. N., Schwartz L. M. Crain consolidation and electrical conductivity in porous media // Phys. Rev. B. 1985. V. 31, № 9. P. 5990.
8. Schwartz L. M., Vanavar J. R. Calculation of electrical transport in continuum system by diffusin simulation // Phys. A. 1989. V. 157, № 9. P. 230.
9. Соколов И. М. Размерности и другие геометрические критические показатели в теории протекания // УФН. 1986. Т. 150, вып. 2. С. 221.
10. Трахтенгерц В. Ю., Иудин Д. И., Григорьев А. Н. О фрактальной динамике активных сред // Нелинейные волны' 2002 / Отв. ред. А. В. Гапонов-Грехов, В. И. Некоркин. Н. Новгород: ИПФ РАН, 2003. С. 287—302.
11. Иудин Д. И., Трахтенгерц В. Ю. Динамическая перколяция в активных средах // Нелинейные волны' 2004 / Отв. ред. А. В. Гапонов-Грехов, В. И. Некоркин. Н. Новгород: ИПФ РАН, 2005. С. 217—242.
12. Добрынин В. М., Серебряков В. А. Методы прогнозирования аномальных высоких пластовых давлений. М.: Недра, 1978. 232 с.
13. Итон Б. Использование получаемых в процессе бурения петрофизических данных для оценки перспектив // Нефтегазовые технологии. 1993. № 3. С. 15—25.
14. Иудин Д. И., Касьянов Д. А., Шалашов Г. М. Фильтрационное течение в среде с переменной пористостью // Доклады РАН. 1999. Т. 2. С. 257—259.
15. Gouet J. F. Invasion noise during drainage in porous media // Physica A. 1990. V. 168. P. 581.
16. Gouet J. F., Sapoval B., Boughaleb Y., Rosso M. Structura of noise generated on diffusion fronts // Physica A. 1990. V. 157. P. 620.
17. Iudin D. I., Korovkin N. V., Molchanov O. A., Surkov V. V., Hayakawa M. Model of earthquake triggering due to gas-fluid «bubble» upward migration I. Physical Rationale // Seismo Electromagnetics. Lithosphere-Atmosphere-Ionosphere Coupling / Ed. by M. Hayakawa and O. A. Molchanov. Tokyo: TERRAPUB, 2002. P. 177—185.
18. Iudin D. I., Korovkin N. V., Molchanov O. A., Surkov V. V., Hayakawa M. Model of earthquake triggering due to gas-fluid «bubble» upward migration II. Finite-automaton model // Seismo Electromagnetics. Lithosphere-Atmosphere-Ionosphere Coupling / Ed. by M. Hayakawa and O. A. Molchanov. Tokyo: TERRAPUB, 2002. P. 177—185.
19. Iudin D. I., Kas'yanov D. A. Percolation model of seismic aktivty // Atmospheric and Ionospheric Electromagnetic Phenomena Associated with Earthquakes / Ed. by M. Hayakawa. Tokyo: Terra Scientific Publishing Company, 1999. P. 911—917.

ВОЛНЫ ЦУНАМИ: ОБЩАЯ ХАРАКТЕРИСТИКА ЯВЛЕНИЯ, МЕХАНИЗМЫ ОБРАЗОВАНИЯ, ПРОГНОЗ

М. А. Носов

1. Общие сведения о волнах цунами

Под волнами цунами обычно понимают поверхностные гравитационные волны с диапазоном периодов 10^2 — 10^4 с. Образование цунами в первую очередь связано с сейсмическими движениями океанического дна (~80% случаев), но существуют и другие механизмы формирования волн. Значительное число событий вызывается оползнями или обвалами (~6%), процессами, связанными с вулканическими извержениями (5%), и метеорологическими причинами (3%). В ряде случаев источник волн остается неизвестным. Последнее время активно обсуждается вопрос о формировании цунами при падении в океан метеоритов. Событие это крайне редкое. Возможно, что за всю историю существования нашей цивилизации оно ни разу не происходило. Но, с другой стороны, нет сомнений, что в геологической истории Земли космогенные цунами оставили свой след.

Следует иметь в виду возможность комбинации различных механизмов формирования волн цунами. Так, например, подводные оползни, спровоцированные землетрясением, могут обеспечить дополнительный вклад в энергию волн, образованных подвижками дна. Рассматривая все причины вкуче, можно утверждать, что побережье любого крупного водоема подвержено опасности цунами.

Наиболее цунамиопасным считается Тихоокеанский регион, в котором известно более 1300 событий. Около 300 цунами имели место в Средиземном море. Имеются сведения о цунами в Атлантическом океане, Карибском, Черном и Каспийском морях. Под пристальным вниманием исследователей теперь находится Индийский океан, хотя и в прошлом, до катастрофы 26 декабря 2004 г., его побережья неоднократно подвергались воздействию цунами.

Цунами являются длинными волнами, в том смысле, что их длина существенно превосходит глубину океана $\lambda \gg H$. Скорость распространения длинных волн в водоеме глубины H определяется простой формулой $V = \sqrt{gH}$, где g — ускорение силы тяжести. При глубине $H \sim 4$ км цунами распространяется со скоростью порядка 200 м/с (720 км/ч), что сопоставимо со скоростью современного реактивного авиалайнера. Амплитуда цунами в открытом океане даже для катастрофических событий обычно измеряется десятками сантиметров и, скорее всего, редко превышает 1 м. Впрочем, в источнике цунами амплитуда смещения поверхно-

сти воды может достигать 10 м и более. Но и в этом случае она много меньше глубины океана.

Малость амплитуды при большом периоде делает волну цунами в открытом океане практически незаметной для наблюдателя, находящегося на борту судна или авиалайнера. Известно, что катастрофическое цунами в Японии (15 июня 1896 г.), которое унесло 28000 жизней, разрушило порт Санрику и прибрежные поселения на протяжении 275 км, не было замечено рыбаками, находившимися в море всего в 40 км от побережья. С математической точки зрения малость амплитуды обеспечивает возможность применения простых линейных моделей для описания распространения волн в открытом океане.

Зависимость скорости распространения цунами от глубины делает эти волны чувствительными к формам рельефа дна. Для цунами свойственны эффекты захвата волновой энергии подводными хребтами и шельфом; неровности дна приводят к рассеиванию волновой энергии.

В реальности скорость распространения гравитационных волн зависит не только от глубины, но и от длины волны. Приведенная выше формула для скорости длинных волн является предельным случаем (при $\lambda \gg H$) более общего выражения $V = \sqrt{g \tanh(kH) / k}$, где $k = 2\pi / \lambda$. Дисперсия волн приводит к преобразованию исходного возмущения в волновой пакет, в котором лидируют наиболее быстрые длинные волны. Отметим, что этот эффект обычно проявляется при распространении волн на достаточно протяженных трассах.

Увеличение амплитуды цунами при подходе к берегу, — что во многом и определяет опасность этих волн, — также связано с рельефом дна. Уменьшение глубины приводит к снижению скорости распространения волн, а следовательно, и к сжатию волнового пакета в пространстве и увеличению амплитуды. Для катастрофических цунами высота заплеска достигает 10—30 м, при этом волна способна проникать вглубь суши (горизонтальный заплеск) на 3—5 км от побережья.

Опасность, которую несут с собой волны цунами, связана со следующими тремя основными факторами: внезапным затоплением части суши, волновым воздействием на сооружения и эрозией. Сильные потоки воды, достигающие скоростей в десятки метров в секунду, способны ломать и перемещать дома, подмывать фундаменты зданий, разрушать мосты, портовые конструкции. Потоки воды часто несут в себе обломки зданий и прочих сооружений, деревья, малые и большие суда, что не оставляет шансов на выживание у человека, попавшего в такой поток. Ущерб, наносимый цунами, может быть также связан с пожарами, загрязнением окружающей среды, эпидемиями, возникшими в результате разрушения береговой инфраструктуры.

С детальной информацией об этом грозном природном явлении можно ознакомиться по монографиям [6, 9, 20] и каталогам цунами [11—14].

2. Сеймотектонические цунами

Наиболее распространенной причиной возникновения волн цунами являются сильные подводные землетрясения. По записям волн на побережье, используя лучевой метод, можно восстановить область очага цунами. Очаг, восстановленный таким образом, как правило, соответствует области проявления афтершоков и области, где произошли заметные остаточные деформации дна. Остаточные деформации обычно являются биполярными, т. е. в одной части источника происходит поднятие дна, а в другой – опускание.

Известны эмпирические зависимости, связывающие средний радиус очага цунами R_{TS} [км] и амплитуду смещения поверхности океана в очаге ξ_0 [м] с магнитудой землетрясения M [2]:

$$\lg R_{TS} = (0,50 \pm 0,07)M - (2,1 \pm 0,6), \quad (1)$$

$$\lg \xi_0 = (0,8 \pm 0,1)M - (5,6 \pm 1,0). \quad (2)$$

Деформация дна происходит не мгновенно и не одновременно во всей активной области. Эволюция деформаций определяется характером распространения разрыва в очаге землетрясения. В качестве меры продолжительности процесса в очаге землетрясения часто используется характеристика T_{hd} [с] (half duration — половина продолжительности). Найти эту временную характеристику можно, например, в Гарвардском сейсмическом каталоге (<http://www.seismology.harvard.edu/>). В результате анализа всех землетрясений с магнитудой $M_w > 7$, содержащихся в этом каталоге, нами была получена следующая регрессионная связь [6]:

$$\lg T_{hd} = (0,42 \pm 0,02)M_w - (1,99 \pm 0,14). \quad (3)$$

Сопоставляя период волн цунами, который можно определить как $T_{TS} = R_{TS} / \sqrt{gH}$, и величину T_{hd} , можно заключить, что продолжительность процессов в очаге землетрясения значительно уступает периоду волны цунами. Поэтому в первом приближении генерацию волн целесообразно рассматривать как мгновенный процесс.

При моделировании цунами обычно полагают, что остаточные деформации дна океана в результате землетрясения возникают мгновенно. Методика расчета остаточных деформаций по параметрам очага землетрясения основана на аналитическом решении статической задачи теории упругости [21]. Далее предполагается, что одновременно с подвижкой дна на поверхности океана образуется возмущение, форма которого полностью аналогична остаточным деформациям дна. Полученное таким образом возмущение водной поверхности используется в качестве начальных условий для решения задачи распространения цунами, которая обычно рассматривается в рамках теории длинных волн.

3. Роль сжимаемости воды и нелинейности

Большинство моделей цунами рассматривают океан как несжимаемую среду. Покажем, что для адекватного описания процесса генерации цунами землетрясением следует принимать во внимание сжимаемость морской воды.

Известно, что жидкость можно считать несжимаемой, если $\Delta\rho/\rho \ll 1$, где ρ — плотность жидкости. Необходимым условием для этого является малость скорости движения жидкости по сравнению со скоростью звука: $v \ll c$. В случае стационарного течения это условие является достаточным. Задача о генерации цунами, очевидно, является нестационарной, поэтому требуется выполнение еще одного дополнительного условия: $T_{hd} \gg (H/c, R_{TS}/c)$, физический смысл которого состоит в том, что время нестационарности должно существенно превосходить время распространения упругой волны вдоль характерного расстояния. Простые расчеты показывают, что второе условие во многих случаях может быть нарушено. Кроме того, в случае реализации бегущей подвижки будет нарушено и первое условие, т. к. скорость вспарывания разрыва превосходит скорость звука в воде.

Заметим, что для задач распространения цунами в открытом океане и наката волн на берег как первое, так и второе условия хорошо выполняются, т. е. водный слой ведет себя как несжимаемая жидкость.

Если рассматривать случай горизонтального абсолютно жесткого дна, то основное различие в поведении сжимаемого океана по сравнению с несжимаемой модельной средой заключается в формировании упругих колебаний водного слоя, характеризующихся дискретным набором нормальных частот $\nu_k = c(1+2k)/4H$, где $k = 0, 1, 2, \dots$. Для типичных условий очага цунами наиболее эффективно возбуждается нормальная частота $\nu_0 = c/4H \sim 0,1$ Гц.

Реальные очаги цунами располагаются, конечно, не на горизонтальном дне, а в области со сложной батиметрией. Но средний уклон океанического дна обычно не превышает величины 0,1. Поэтому поверхность дна можно условно представить как совокупность квазигоризонтальных участков, каждый из которых характеризуется своей собственной глубиной и набором нормальных частот, соответствующим этой глубине. Таким образом, в некоторой заданной точке очага в первую очередь образуются упругие колебания с нормальными частотами, определяемыми глубиной океана в этой точке. Затем спектр упругих колебаний может обогатиться высокими частотами за счет волн, пришедших из соседних мелководных областей. Из-за существования частоты отсечки низкочастотные колебания, сформированные в соседних глубоководных областях, вверх по склону не распространяются.

До недавнего времени упругие низкочастотные колебания водного слоя, образуемые подводными землетрясениями, имели статус теоретически предсказанного эффекта. Недавно, во время сильнейшего землетрясения 2003 г. (Токачи-оки) это явление было впервые зарегистрировано донными датчиками давления [8]. Во избежание недоразумения подчеркнем разницу между таким известным явлением, как T -фаза (гидроакустические волны, излучаемые очагами подводных землетрясений [22]), и упругими низкочастотными колебаниями в очаге цунами. T -фаза не только относится к более высокочастотному диапазону (> 2 Гц), но и регистрируется на значительных расстояниях от источника.

Отметим две особенности эффектов сжимаемости, объясняющие их слабую изученность. Во-первых, упругие низкочастотные колебания водного слоя могут образовываться и распространяться только в открытом океане при достаточно больших глубинах. На мелководье из-за существования частоты отсечки эти колебания не проникают и, следовательно, вблизи побережья наблюдаться не могут. Во-вторых, частотные диапазоны упругих колебаний и волн цунами сильно различаются, следовательно, если рассуждать в рамках линейной теории, эффекты сжимаемости не способны повлиять на волны цунами. В действительности вклад упругих колебаний в волну цунами может быть обеспечен нелинейными механизмами. Часть энергии упругих колебаний может быть передана вниз по спектру к длинным гравитационным волнам. Таким образом, несмотря на то, что упругие колебания не могут непосредственно проявляться у побережья, они все же способны обеспечить дополнительный вклад в цунами вблизи источника [7].

Интересно сопоставить величины энергий упругих колебаний и гравитационных волн (цунами), возбуждаемых деформацией дна. Пусть в некоторый момент времени участок дна площадью S ($S^{1/2} \gg H$) начинает смещаться в вертикальном направлении с постоянной скоростью v . Движение происходит в течение времени τ , после чего дно останавливается. В результате образуется остаточное смещение высотой $v\tau$. Если рассматривать задачу в приближении несжимаемой жидкости, то при условии скоротечности подвижки ($\tau \ll \sqrt{S/(gH)}$) в момент времени $t = \tau$ форма возмущения поверхности воды окажется практически идентичной форме остаточного смещения дна (g – ускорение свободного падения). Поэтому энергию волны цунами можно оценить как потенциальную энергию начального возвышения, имеющего площадь S и высоту $v\tau$: $W_1 = 0,5\rho gS(v\tau)^2$.

В рамках модели сжимаемой жидкости мощность, передаваемая водному слою единицей поверхности подвижного дна, определяется простой формулой ρcv^2 . Зная площадь источника и продолжительность подвиж-

ки, приходим к следующему выражению для энергии упругих колебаний: $W_2 = \rho c v^2 S \tau$. Сопоставляя величины W_1 и W_2 , получаем

$$\frac{W_2}{W_1} = \frac{2c}{g\tau}. \quad (4)$$

Легко убедиться, что в диапазоне значений τ , свойственном реальным сейсмическим событиям (1—100 с), значительная часть энергии, передаваемой от движущегося дна водному слою, существует в виде упругих волн. Этот факт говорит не только об энергетической значимости упругих колебаний, но и обеспечивает возможность заметного вклада нелинейных эффектов в формирование цунами.

В основе математической модели нелинейного механизма формирования цунами лежит предположение о том, что водный слой участвует в двух движениях: медленном (среднем) и быстром (колебательном), т. е. скорость течения, давление и плотность представимы в виде сумм:

$$\mathbf{v} = \langle \mathbf{v} \rangle + \mathbf{v}', \quad p = \langle p \rangle + p', \quad \rho = \langle \rho \rangle + \rho'. \quad (5)$$

Подставим формулы (1) в систему уравнений Эйлера и выполним операцию осреднения по периоду «быстрых» колебаний $4H_{\max}/c$, где H_{\max} — максимальная глубина бассейна, c — скорость звука в воде. Известно, что акустические моды с периодом большим, чем $4H_{\max}/c$, не существуют, следовательно, среднее движение можно описывать как движение несжимаемой жидкости. Пренебрегая квадратичным по средней скорости членом, приходим к линеаризованной системе уравнений для среднего движения:

$$\frac{\partial \langle \mathbf{v} \rangle}{\partial t} = -\frac{\nabla \langle p \rangle}{\langle \rho \rangle} + g - \langle (\mathbf{v}', \nabla) \mathbf{v}' \rangle + \frac{\langle \nabla p'^2 \rangle}{2c^2 \langle \rho \rangle^2}, \quad (6)$$

$$\operatorname{div} \langle \mathbf{v} \rangle = -\frac{1}{\langle \rho \rangle} \operatorname{div} \langle \rho' \mathbf{v}' \rangle. \quad (7)$$

При получении формул (6), (7) использована связь $p' = c^2 \rho'$. От обычных линеаризованных уравнений Эйлера для несжимаемой жидкости уравнения (6), (7) отличаются наличием двух новых членов, которые могут быть интерпретированы как силовое поле и распределенный источник массы. Своим происхождением новые члены обязаны нелинейности исходной системы (уравнений Эйлера). Совместное действие силового поля и распределенного источника массы способно привести к образованию гравитационных волн с амплитудой $A \sim U_{\max}^2 / g$, где U_{\max} — максимальная скорость вертикальной деформации дна [6, 7].

4. Формирование цунами несейсмическими источниками

После цунами сейсмотектонического происхождения наиболее распространенными являются *оползневые цунами*. Этот термин обозначает гравитационные поверхностные волны, образованные подводными оползнями и мутьевыми потоками, обрушением в воду фрагментов крутых берегов, скал и ледников, а иногда и портовых сооружений. В настоящее время только в тихоокеанском регионе известно более 80 событий такого рода.

По сравнению с горизонтальным размером сейсмического источника (10^4 — 10^5 м) береговые и подводные оползни обычно имеют сравнительно небольшие масштабы (10^2 — 10^3 м). Крупнейший из известных в геологической истории оползней Стореджа (Storrega) произошел в позднем четвертичном периоде в районе крутого континентального склона около побережья Норвегии [18]. Его горизонтальный размер оценивается в десятки километров.

Оползневый процесс обычно является результатом длительного накопления осадочного материала в течение десятков и сотен лет. Со временем накопленные на склонах массы осадков теряют устойчивость. Спровоцировать сход оползня может множество факторов [4]: внезапный выброс речных взвесей во время речного паводка; эрозия осадочного слоя на крутом склоне дна; береговое строительство; длительные дожди, приводящие к водонасыщению берегового грунта; обнажение берега во время сильного отлива. Важнейшей причиной оползней и обвалов, конечно, служат землетрясения. Не последнюю роль в инициации оползневых процессов и обвалов играют вулканические извержения.

Интересно, что оползневые цунами могут происходить не только в океанах и морях, но и в крупных реках. Описание одного из таких событий, произошедших на реке Волге в 1597 г., приведено в работе [1]. Нам удалось отыскать упоминание о другом речном цунами, которое имело место на реке Иртыш в 1885 г. [6].

Не вдаваясь в тонкости математических и физических моделей, применяемых для описания оползневых цунами, отметим одну важную физическую особенность процесса. Хорошо известен эффект резонансного возбуждения волн, когда некоторое возмущение (оползневое тело) движется по горизонтали со скоростью, равной скорости длинных волн $c = \sqrt{gH}$. В этом случае генерация длинных волн оказывается особенно эффективной. Интересно, что в действительности скорость движения подводных оползней, как правило, меньше, чем скорость длинных гравитационных волн. Этот факт вытекает из элементарных физических соображений. Если рассматривать движения оползня без учета волнового сопротивления и трения, то его скорость определяется формулой

$$U = \sqrt{2g \frac{\rho_2 - \rho_1}{\rho_2} \Delta H}, \quad (8)$$

где ΔH — изменение вертикального положения центра масс оползня. Предположим, что, спускаясь вниз по склону, оползень достиг некоторой глубины H . Очевидно, что для подводного оползня всегда выполняется неравенство $\Delta H < H$. Сопоставляя скорость оползня U и скорость длинных волн \sqrt{gH} , несложно заключить, что равенство этих величин возможно только при $\rho_2 > 2,0$ г/см³, т. е. для хорошо консолидированных осадков и скальных пород. При учете сил трения требуемая плотность оползневого тела окажется еще больше. Таким образом, «резонанс» возможен только для оползней, состоящих из очень плотных материалов, или когда оползень входит в воду, имея некоторую начальную скорость.

Взрывы вулканических островов (коллапсы кальдер), взрывоподобные извержения подводных вулканов, пирокластические потоки, попадающие в воду, — все эти явления способны вызывать волны, которые по своей разрушительной силе сопоставимы с цунами сейсмотектонического происхождения. Только в Тихоокеанском регионе на настоящее время известно 66 цунами вулканического происхождения, причем в десяти случаях высоты волн достигали 10 м и более (до 55 м).

Основными физическими механизмами возбуждения вулканогенных цунами являются: выброс в воду большого объема вещества (от медленного истечения лавы до взрывного или взрывного извержения), коллапс кальдеры (взрыв вулканического острова), пирокластические потоки, обвалы и т. д., вулканические землетрясения [9].

Одним из ярчайших исторических примеров вулканогенного цунами являются волны, вызванные активностью вулкана Кракатау в августе 1883 г. [16]. 26 августа около 17 ч местного времени произошла серия громких взрывов, и вулкан выбросил облако пепла на высоту до 25 км. При этом образовалось небольшое цунами высотой 1—2 м. Утром 27 августа произошло три колоссальных взрыва. Первый взрыв (5 ч 28 мин) разрушил гору Пербоеватан (Perboewatan) на острове Кракатау, имевшую 130 м высоты. Образовавшаяся при этом кальдера была немедленно заполнена морской водой, что привело к образованию небольшого цунами. В 6 ч 36 мин взорвалась и коллапсировала гора Данан (Danan), имевшая высоту 500 м, что вызвало волну цунами высотой до 10 м. Основной (третий) взрыв произошел в 9 ч 58 мин. Он буквально разорвал на куски оставшуюся часть острова Кракатау (Rakata Island). Вулкан выбросил 9—10 км³ твердой породы и 18—21 км³ пирокластических отложений, которые распределились на площади около 300 км² со средней толщиной 40 м. Пепел покрыл территорию площадью около $2,8 \times 10^6$ км². На месте острова возникла кальдера диаметром 6 км и глубиной 270 м. Звук от третьего

взрыва был сильнейшим из тех, которые когда-либо слышало человечество. Ударные волны в атмосфере семь раз обогнули земной шар. Энергия, высвобожденная в процессе основного извержения вулкана Кракатау, составила $8,4 \times 10^{17}$ Дж. Образовавшиеся в результате третьего сильнейшего взрыва волны цунами достигали высоты 42 м, вглубь суши они проникали на расстояние до 5 км. Средняя высота волн на берегах Зондского пролива (между островами Ява и Суматра) составила около 15 м. Погибло, по крайней мере, 36 тыс. человек. Около 300 поселений было разрушено. Цунами, вызванное взрывом вулкана Кракатау, отмечалось повсеместно. Волны были записаны многими мареографами не только в Индийском, но в Тихом и Атлантическом океанах. Вдали от берегов Индонезии амплитуда волн была сравнительно небольшой.

Длинные волны, аналогичные по своим характеристикам волнам цунами сейсмотектонического происхождения, могут образовываться и в результате воздействия на водный слой различных атмосферных процессов. Эти волны принято называть анемобарическими волнами или метеоцунами [10].

Основными причинами возникновения метеоцунами служат подвижные неоднородности атмосферного давления или напряжение трения ветра. Сразу заметим, что в отличие от иных механизмов генерации цунами в данном случае важную роль играют резонансные эффекты, проявляющиеся, когда скорость распространения атмосферных возмущений и их период оказываются близкими к скорости длинных волн и периоду собственных колебаний акватории.

Аналогично цунами сейсмотектонического происхождения, метеоцунами — явление довольно редкое. Подобно тому, как не каждое землетрясение возбуждает цунами, так и не каждый циклон, атмосферный фронт, цуг внутренних гравитационных волн или иное атмосферное возмущение приводят к формированию метеоцунами. Известно большое количество примеров, когда достаточно сильные атмосферные возмущения не сопровождалась генерацией длинных волн. Тем не менее только в Тихоокеанском регионе зафиксировано 36 событий, классифицируемых как метеоцунами.

Широко известен аномальный характер сейшевых колебаний в бухте Нагасаки (о. Кюсю). Колебания с амплитудой $\sim 0,5$ м и периодом около 30 мин в этой бухте — типичное явление. Оно носит местное название «абики». В ряде случаев волны абики способны достигать значительных амплитуд. Так, например, 31 марта 1979 г. максимальная высота волн, приведших к значительным разрушениям и гибели 3 человек, составила 4,78 м. В работе [17] было показано, что это событие было вызвано прохождением скачка атмосферного давления величиной от 2 до 6 гПа над западной частью Восточно-Китайского моря. Скорость перемещения скачка давления была около 30 м/с. Характерные глубины моря между

областью зарождения возмущения и островом Кюсю составляют от 50 до 150 м. Соответствующие им скорости распространения длинных волн варьируются в пределах 22—38 м/с, что близко к скорости распространения атмосферного возмущения. Период волн, подошедших к бухте Нагасаки, также оказался близким к периоду собственных колебаний. В результате двойного резонанса высота волн возросла более чем в 100 раз. Таким образом, скачок давления всего в несколько гектопаскалей вызвал образование волн абики в бухте Нагасаки высотой в несколько метров.

Другой известный пример — сейшевые колебания с периодом от нескольких минут до нескольких десятков минут, которые регулярно наблюдаются в летнее время у юго-восточного побережья Испании в районе Балеарских островов. В этом отношении наиболее известной является бухта Сьютаделла, расположенная в северо-западной части о. Менорка. Бухта имеет длину порядка 1 км, ширину около 90 м и практически ровное дно при глубине 5 м. В отдельных случаях сейши с типичным периодом около 10 мин достигают здесь высоты 4 м, приводя к серьезным повреждениям судов и береговых построек. Это бедствие получило местное название «фриссага» [19].

По оценкам специалистов потенциальную возможность столкнуться с нашей планетой имеют несколько тысяч крупных объектов (астероиды и кометы) с диаметром более 1 км [25]. Если с Землей столкнется космическое тело размером более 2 км, то это неизбежно приведет к глобальной катастрофе. Объекты относительно небольших размеров регулярно «бомбардируют» Землю, но большинство из них разрушаются и сгорают еще в верхних слоях атмосферы. Критический для каменного метеорного тела размер, при котором оно способно достичь поверхности Земли, составляет около 100 м в диаметре.

Задача описания космогенного цунами является чрезвычайно сложной. Ее оптимальное решение, видимо, состоит в подборе эквивалентного источника и параметризации его характеристик на основе результатов численного моделирования [6]. Падение метеорита в океан создает в воде на начальной стадии полость («кратер»), размеры которой определяются параметрами небесного тела: плотностью, радиусом и скоростью.

На следующем этапе описывают эволюцию волн от начального возмущения (полости), созданного небесным телом. При этом обычно полагают, что в тот момент времени, когда полость сформировалась, скоростью движения частиц воды можно пренебречь. В силу того, что в данной задаче амплитуда и длина волны оказываются сопоставимыми с глубиной океана, применение линейной теории длинных волн оказывается недопустимым. При распространении волн принципиальную роль будут играть нелинейность (вблизи источника) и фазовая дисперсия. Волновой пакет, в котором на больших временах лидируют длинноволновые компоненты, может распространяться со скоростью, более чем в два раза ус-

тупающей скорости длинных волн. Но диспергирующие волны не только распространяются медленнее, но они еще и быстро теряют амплитуду при удалении от области зарождения [6].

В заключение подчеркнем, что вероятность падения на Землю крупного метеорита, который способен вызвать значительные волны цунами, исчезающее мала. Если все же такое событие произойдет, то среди прочих катастрофических последствий волны цунами, несомненно, будут играть не последнюю роль.

5. Прогноз цунами

Существуют два типа прогноза цунами: краткосрочный (оперативный) и долгосрочный. Задача оперативного прогноза состоит в расчете времени добегания волны до защищаемой точки побережья и в определении ее высоты. Долгосрочный прогноз направлен на вероятностную оценку цунамириска данного побережья.

Долгосрочный прогноз цунами в первую очередь опирается на статистический анализ событий, имевших место в прошлом. Очевидно, что по своей силе цунами варьируются в широких пределах: от слабых волн, которые можно зарегистрировать только инструментально, до страшных катастрофических событий, опустошающих побережья на протяжении сотен километров.

Для статистического анализа необходима количественная характеристика, показывающая силу цунами. Определение такой количественной характеристики является весьма нетривиальной задачей, которая на настоящий момент окончательно не решена. Заметим, что с аналогичными трудностями сталкиваются сейсмологи при определении силы землетрясения. Цунами, с одной стороны, характеризуется объективными и количественно измеряемыми параметрами (энергия, амплитуда, период и т. д.), а с другой стороны — субъективными описаниями, отражающими масштабность и степень вызванных волной разрушений или характер ее проявлений на побережье. Как и в случае землетрясений, для оценки цунамиопасности более важными являются именно эти субъективные описания, а не абстрактные физические параметры.

Первая попытка классифицировать цунами была предпринята Зибергом, который ввел 6-балльную шкалу интенсивности цунами по аналогии со шкалой интенсивности землетрясений [24]. Эта шкала не была связана с измерением физических параметров (высоты волн, дальности заплеска и т. д.), она основывалась на описании макроскопических эффектов, показывающих степень разрушений. В дальнейшем шкала Зиберга была несколько модифицирована [15].

Недавно в работе [23] была предложена новая детальная 12-балльная описательная шкала интенсивности цунами. Шкала организована по сле-

дующим трем признакам: (а) воздействию на людей, (б) воздействию на природные и искусственные объекты, включая суда различных размеров, и (в) ущербу, причиняемому зданиям. Поэтому цунами большой амплитуды, обрушившееся на малонаселенное побережье, может иметь невысокую интенсивность по шкале Пападопулоса — Имамуры. И наоборот, цунами с умеренной амплитудой, но поразившее густонаселенное побережье, может характеризоваться достаточно высокой интенсивностью. Ознакомиться со шкалой можно в монографии [6].

Среди количественных характеристик силы цунами, основанных на измеряемых величинах, в настоящее время наиболее широко распространена «шкала интенсивности цунами Соловьева — Имамуры» [26]: $I = 1/2 + \log_2 H$, где H — средняя высота цунами на ближайшем к источнику побережье. Заметим, что такое определение интенсивности не является математически строгим и, следовательно, оставляет много «свободы» для расчета средней высоты цунами. Впрочем, шкала не очень чувствительна к небольшим ошибкам в определении средней высоты волн, т. к. эта величина стоит под знаком логарифма. Важно еще отметить, что для большого числа исторических событий, и тем более для доисторических событий (палеоцунами), единственная информация, имеющаяся в распоряжении, — оценка высот волн в одной или нескольких точках побережья. В этой связи шкала Соловьева — Имамуры достаточно удобна и еще долго будет применяться на практике.

Сама возможность оперативного прогноза цунами основана на том простом факте, что скорость распространения сейсмических волн во много раз больше, чем скорость волны цунами. Тревога объявляется при регистрации подводного землетрясения с магнитудой, превышающей пороговое значение. Заметим, что основанием для такого критерия служит тот факт, что все значимые цунами ($I > 2$) были сформированы в результате землетрясений с магнитудой $M_w > 7$.

Важным успехом, достигнутым в оперативном прогнозе цунами, является возможность быстро (в реальном времени) и с необходимой для практических нужд точностью и надежностью рассчитать время добегаания волны до заданной (защищаемой) точки побережья. Такой расчет может быть выполнен с применением простой лучевой теории. Для него необходимо только знание местоположения источника цунами и распределения глубин в рассматриваемом бассейне.

Гораздо хуже обстоят дела с расчетом высоты заплеска цунами в заданной точке побережья. Необходимые для практических нужд точность и оперативность в решении этой задачи на настоящее время не достигнуты. С одной стороны, это объясняется колоссальным объемом вычислений, который необходимо выполнить для расчета эволюции волны от ее зарождения в источнике до выхода на побережье. С другой стороны, в оперативном режиме невозможно с необходимой точностью рассчитать,

что происходило в источнике цунами. Время, требуемое для надежного определения деформаций дна, произошедших в результате землетрясения, существенно превышает минуты или даже часы, отпущенные для оперативного прогноза.

Существует несколько перспективных путей решения этой проблемы. Один из них реализован в японской системе оперативного прогноза цунами [27]. Методика основана на том, что очаги цунами имеют свойство повторяться. Поэтому задачи, требующие длительных расчетов, решены заранее. Результаты расчетов сведены в специальной базе данных. Когда происходит реальное подводное землетрясение, то в соответствии с его магнитудой и местоположением эпицентра из базы выбираются необходимые данные, по которым с помощью метода интерполяции оперативно рассчитываются возможные высоты заплесков.

Второй путь заключается в использовании глубоководных регистраторов цунами, установленных вдали от побережья (например, DART, JAMSTEC). Сама идея и ее первая реализация связаны с именем С. Л. Соловьева. Одна из возможностей использования таких систем — заблаговременная регистрация волны вдали от берега и корректировка неточностей в расчетах генерации волны в источнике.

Катастрофическое цунами, произошедшее в декабре 2004 г. в Индийском океане, было впервые надежно зарегистрировано спутниковым радиовысотометром [3]. В этой связи следует ожидать развития спутниковых методов мониторинга цунами.

Заметим, что измеритель уровня открытого океана (датчик придонного давления) имеет принципиальное преимущество по сравнению со спутниковым альтиметром. Дело в том, что частотный диапазон волн цунами практически свободен от посторонних сигналов, в то время как соответствующий диапазон длин волн сильно зашумлен (например, вихревыми образованиями). Поэтому волна цунами может быть легко выделена на записи вариаций уровня океана во времени, но не в пространстве.

Определенной перспективой обладают и гидроакустические методы. Анализ материалов записей гидроакустических шумов океана, полученных российской многоцелевой антенной АГАМ в рамках международной программы АТОС (Акустическая термометрия океанского климата) в период 1998—1999 гг., продемонстрировал многообещающие результаты. Система гидрофонов, установленная на Тихоокеанском шельфе Камчатки, зафиксировала гидроакустические сигналы сейсмического происхождения в диапазоне частот 3—70 Гц, которые возникали задолго до появления первого удара землетрясения (от часов до нескольких минут) [5]. Сигналы генерировались микроземлетрясениями, которые происходили в области подготовки сильного землетрясения и свидетельствовали о развитии критической стадии события. Важно отметить, что регистрация таких слабых сейсмических событий возможна только по гидроакустиче-

скому каналу. Дело в том, что коэффициент затухания упругих волн в воде существенно меньше, чем соответствующие характеристики сейсмических волн. Отсюда следует важный вывод о том, что гидроакустический мониторинг процесса подготовки океанического землетрясения может привести к успеху в решении сложной задачи о заблаговременном обнаружении готовящегося близкого землетрясения и своевременном и эффективном предупреждении о возникновении локального цунами.

Работа выполнена при финансовой поддержке РФФИ, проект №04-05-64297, НШ-2104.2003.5.

Литература

1. Диденкулова И. И., Зайцев А. И., Красильщиков А. А., Куркин А. А., Пелиновский Е. Н., Ялчинер А. С. Нижегородское цунами 1597 года на реке Волге: Препринт ИПФ РАН, № 632. Нижний Новгород, 2003. 16 с.
2. Доценко С. Ф., Соловьев С. Л. Математическое моделирование процессов возбуждения цунами подвижками океанского дна // Исследования цунами. М., 1990 № 4. С. 8—20.
3. Куликов Е. А., Медведев П. П., Лаппо С. С. Регистрация из космоса цунами 26 декабря 2004 г. в Индийском океане // ДАН. 2005. Т. 401, № 4. С. 537—542.
4. Куликов Е. А., Рабинович А. Б., Файн И. В., Борнхольд Б. Д., Томсон Р. Е. Генерация цунами оползнями на тихоокеанском побережье Северной Америки и роль приливов в этом процессе // Океанология. 1998. Т. 38, № 1. С. 361—367.
5. Лаппо С. С., Левин Б. В., Сасорова Е. В., Диденкулов И. Н., Карлик Я. С., Морозов В. Е. Гидроакустическая локация области зарождения океанического землетрясения // ДАН. 2003. Т. 388, № 6. С. 805—808.
6. Левин Б. В., Носов М. А. Физика цунами и родственных явлений в океане. М.: Янус-К, 2005. 360 с.
7. Носов М. А., Колесов С. В. Нелинейный механизм формирования цунами в океане в приближении сжимаемой жидкости // Вестн. МГУ. Сер. 3. Физика. Астрономия. 2005. № 3. С. 51—54.
8. Носов М. А., Колесов С. В., Остроухова А. В., Алексеев А. Б., Левин Б. В. Упругие колебания водного слоя в очаге цунами // ДАН. 2005. Т. 404, № 2. С. 255—258.
9. Пелиновский Е. Н. Гидродинамика волн цунами. Н. Новгород: ИПФ РАН, 1996. 276 с.
10. Рабинович А. Б. Длинные гравитационные волны в океане: захват, резонанс, излучение. СПб.: Гидрометеоздат, 1993. 324 с.
11. Соловьев С. Л., Го Ч. Н. Каталог цунами на восточном побережье Тихого океана (1513—1968). М.: Наука, 1975. 203 с.
12. Соловьев С. Л., Го Ч. Н. Каталог цунами на западном побережье Тихого океана. М.: Наука, 1974. 309 с.
13. Соловьев С. Л., Го Ч. Н., Ким Х. С. Каталог цунами в Тихом океане, 1969—1982 г. М.: Изд. МГК АН СССР, 1986. 164 с.
14. Соловьев С. Л., Го Ч. Н., Ким Х. С., Соловьева О. Н., Щетников Н. А. Цунами в Средиземном море 2000 г. до н. э. — 1991 г. М.: Научный мир, 1997. 139 с.
15. Ambraseys N. N. Data for investigation of the seismic sea-waves in the eastern Mediterranean // Bull. Seismol. Soc. Am. 1962. V. 52. P. 895—913.
16. Choi B. H., Pelinovsky E., Kim K. O., Lee J. S. Simulation of the trans-oceanic tsunami propagation due to the 1883 Krakatau volcanic eruption // Natural Hazards and Earth System Sciences. 2003. V. 3. P. 321—332.
17. Hibiya T., Kajiwara K. Origin of Abiki phenomena (a kind of seiches) in Nagasaki Bay // J. Oceanogr. Soc. Japan. 1982. V. 38, № 3. P. 172—182.
18. Jansen E., Befring S., Bugge T. et al. Large submarine slides on the Norwegian continental margin: Sediments, transport, and timing // Marine Geology. 1987. V. 78. P. 77—107.

19. *Montserrat S., Ibbetson A., Thorpe A. J.* Atmospheric gravity waves and the «rissaga» phenomenon // *J. Roy Meteor. Soc.* 1991. V. 117. P. 553—570.
20. *Murty T. S.* Seismic Sea Waves — Tsunamis. 1977 // *Bull. Fish. Res. Board Canada.* № 198. 337 p.
21. *Okada Y* Surface deformation due to shear and tensile faults in a half-space // *Bull. Seis. Soc. Am.* 1985. V. 75, № 4. P. 1135—1154.
22. *Okal E. A., Alasset P.-J., Hyvernaud O., Schindelé F.* The deficient T waves of tsunami earthquakes // *Geophys. J. Int.* 2003. V. 152, № 2. P. 416—432.
23. *Papadopoulos G. A., Imamura F.* A proposal for a new tsunami intensity scale // *ITS 2001 Proc.* P. 569—577.
24. *Sieberg A.* Geologische, physikalische und angewandte Erdbebenkunde. Verlag von Gustav Fischer, Jena. 1927.
25. *Solem J C* Comet and Asteroid Hazards: Threat and Mitigation // *Science of Tsunami Hazards.* 1999. V. 17, № 3. P.141—153.
26. *Soloviev S. L.* Recurrence of tsunamis in the Pacific // *In Tsunamis in the Pacific Ocean / Ed. by W.M. Adams.* 1970. Honolulu: East-West Center Press. P. 149—163.
27. *Tatehata H.* The New Tsunami Warning System of the Japan Meteorological Society // *Science of Tsunami Hazards.* 1998. V. 16, № 1. P. 39—49.

НЕЛИНЕЙНО-ДИСПЕРСИОННАЯ ТЕОРИЯ ВОЛН ЦУНАМИ: ВЗГЛЯД ПОСЛЕ КАТАСТРОФИЧЕСКОГО ЦУНАМИ В ИНДИЙСКОМ ОКЕАНЕ

Е. Н. Пелиновский

1. Введение

Цунами, случившееся 26 декабря 2004 года в Индийском океане после сильнейшего землетрясения в районе Суматры (Индонезия) и Андаманских островов (Индия), оказалось наиболее катастрофическим по своим последствиям в современной истории: около 300 тысяч человек погибло в Индонезии, Таиланде, Индии, Шри Ланка, Мальдивах, Кении и Танзании [1—6]. Волны цунами были инструментально зарегистрированы во всех океанах и даже достигли берегов Канады (как со стороны Тихого океана, так и со стороны Атлантического океана), где их высота достигала 20 см [7, 8]. Это событие планетарного масштаба хорошо известно для бассейна Тихого океана, в частности, волны цунами, зародившиеся у берегов Чили в 1960 году, пересекли Тихий океан и обрушились на берега Камчатки и Японии, имея высоту до 10 м. Аналогичное событие в Индийском океане уже происходило в 1883 году после сильнейшего извержения вулкана Кракатау в Индонезии, когда волны цунами достигли высоты 45 м (погибло около 40 тысяч человек) и распространились также по всему миру, например в проливе Ла-Манш (Английский канал) их высота составила около 15 см [9]. Кадры обрушивающихся на берег гигантских волн цунами во время события 2004 года обошли весь мир (несколько фотографий приведены на рис. 1), они наглядно демонстрировали нелинейный характер волновых процессов, происходящих во время цунами. Среди специалистов по гидромеханике, вычислительной математике, нелинейной физике почти сразу же после цунами развернулась дискуссия о роли и месте нелинейно-дисперсионной теории волн на воде для описания и предсказания цунами. Автор этих строк еще 25 лет назад публиковал книгу «Нелинейная динамика волн цунами» [10], где пытался оценить применимость имеющихся нелинейных теорий к проблеме цунами на основании существовавших в то время крайне грубых оценок параметров волны цунами и немногочисленных случаев их регистрации. Последнее цунами 2004 года дало богатейший материал (более сотни инструментальных записей, многочисленные фото и видео, включая спутниковые изображения), значительно превосходящий по своему объему всю прежнюю историческую информацию, так что стало возможно реально оценить роль и место современных нелинейно-дисперсионных теорий волн на воде. В настоящей статье эта проблема анализируется два года спустя

после события: на часть вопросов уже есть ответы, некоторые пока еще остаются открытыми.

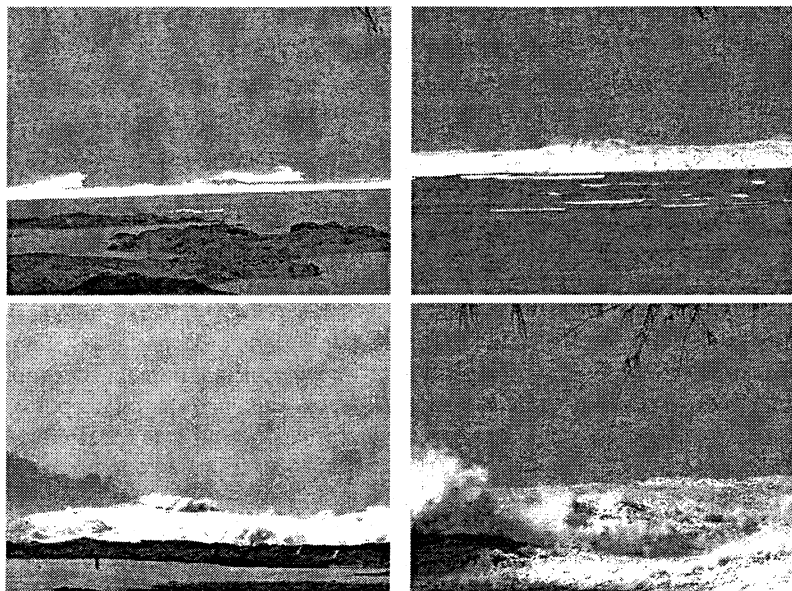


Рис. 1. Фотографии приближающейся волны цунами 26.12.2004

2. Нелинейная теория мелкой воды

Прежде всего дадим краткий обзор существующих моделей волн на воде, которые используются в современной практике прогноза цунами сейсмического происхождения. Грубо характерный размер (диаметр) очага цунами (R) и начальное смещение уровня воды в очаге (H) можно определить по магнитуде землетрясения M , величина которой пропорциональна логарифму энергии, выделившейся при землетрясении; эти эмпирические зависимости представлены на рис. 2 [11]. Как видим, если землетрясение достаточно сильное (магнитуда больше 7 — пороговое значение магнитуды для объявления тревоги цунами на Дальнем Востоке России), то характерный радиус очага составляет десятки или даже сотни километров, значительно превосходя глубину океана. На самом деле, генерация волн цунами при землетрясении — это достаточно сложный процесс [12], зависящий от глубины фокуса землетрясения и направления разлома (вертикальные подвижки при прочих равных условиях генерируют более сильные волны, чем горизонтальные). В настоящее время очаги цунами рассчитываются как остаточные смещения в рамках линейной теории полубесконечного упругого пространства со свободной поверхностью при наличии дельтаобразного движущегося источника в эпицентре

землетрясения [13]. Главная проблема здесь — это выбрать «правильно» параметры движущегося источника. Эта проблема решается в рамках обратных задач сейсмологии на основании сейсмограмм землетрясения, записанных на различных станциях (в частности, сейчас величина магнитуды землетрясения 2004 г. оценена в 9,3). На рис. 3 приведен пример расчета очага цунами 2004 года по сейсмическим данным [5, 14]. Он представляет собой несколько эллипсоидальных сегментов с максимальным подъемом уровня воды до 8 м (в сторону открытого океана) и понижением уровня до 5 м в сторону берега. Протяженность очага вдоль длинной оси около 1000 км, а в поперечном направлении — около 50 км.

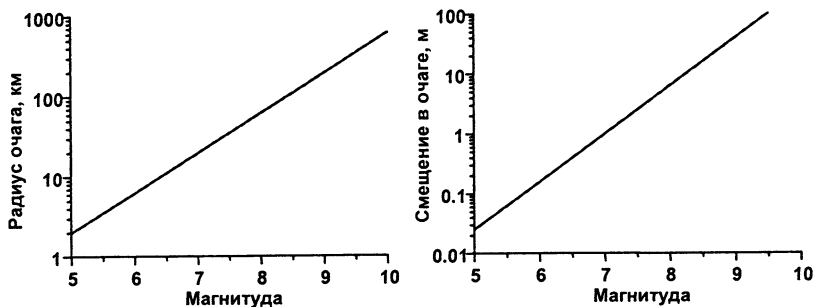


Рис. 2. Грубые оценки параметров очага цунами

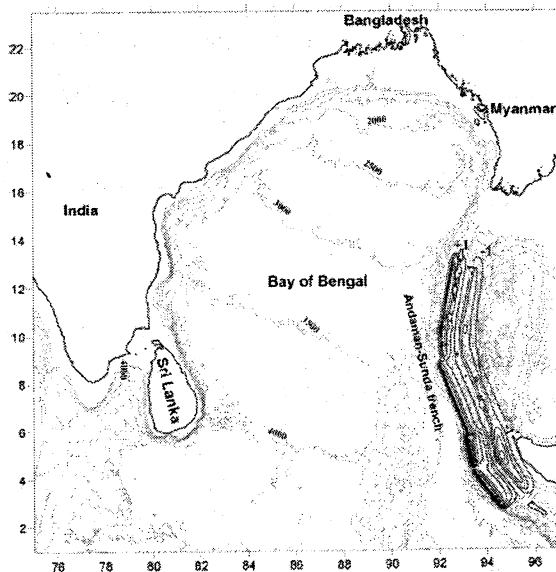


Рис. 3. Очаг цунами 2004 года

Учитывая, что длина волны, выходящей из очага, пропорциональна его размеру и, следовательно, значительно превышает глубину океана. Наиболее подходящей моделью для волн цунами является теория мелкой воды, выводимая строго из уравнений гидродинамики в первом приближении по малому параметру глубина/длина волны. По существу, основным уравнением здесь является классическое волновое уравнение:

$$\frac{\partial^2 \eta}{\partial t^2} - \operatorname{div} [c^2(x, y) \nabla \eta] = 0, \quad c = \sqrt{gh(x, y)}, \quad (1)$$

где $\eta(x, y, t)$ — смещение водной поверхности, $h(x, y)$ — глубина океана и g — ускорение свободного падения. На самом деле для открытого океана приходится учитывать сферичность Земли и ее вращение, так что «рабочей» моделью здесь являются уравнения

$$\begin{aligned} \frac{\partial M}{\partial t} + \frac{gD}{R \cos \theta} \frac{\partial \eta}{\partial \varphi} = fN, \quad \frac{\partial N}{\partial t} + \frac{gD}{R} \frac{\partial \eta}{\partial \theta} = -fM, \\ \frac{\partial \eta}{\partial t} + \frac{1}{R \cos \theta} \left[\frac{\partial M}{\partial \varphi} + \frac{\partial}{\partial \theta} (N \cos \theta) \right] = 0, \end{aligned} \quad (2)$$

где θ и φ — широта и долгота, R — радиус Земли, $f = 2\omega_e \sin \theta$ — параметр Кориолиса, ω_e — частота вращения Земли, M и N — проинтегрированные по глубине компоненты расходы воды, $D = h(x, y) + \eta$ — полная глубина океана. В шельфовых зонах, напротив, вращением Земли и ее сферичностью можно пренебречь, однако эффекты трения о дно в турбулентном придонном потоке становятся существенными, и «рабочей» моделью здесь являются следующие уравнения, где p — коэффициент трения:

$$\begin{aligned} \frac{\partial M}{\partial t} + \frac{\partial}{\partial x} \left(\frac{M^2}{D} \right) + \frac{\partial}{\partial y} \left(\frac{MN}{D} \right) + gD \frac{\partial \eta}{\partial x} + \frac{p}{D^2} M \sqrt{M^2 + N^2} = 0, \\ \frac{\partial N}{\partial t} + \frac{\partial}{\partial x} \left(\frac{MN}{D} \right) + \frac{\partial}{\partial y} \left(\frac{N^2}{D} \right) + gD \frac{\partial \eta}{\partial y} + \frac{p}{D^2} N \sqrt{M^2 + N^2} = 0, \\ \frac{\partial \eta}{\partial t} + \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} = 0. \end{aligned} \quad (3)$$

Подчеркнем, что все эти уравнения нелинейны (хотя нелинейность в открытом океане играет слабую роль), так что процессы нелинейной деформации и обрушения волны цунами они описывают достаточно хорошо [10, 11, 15—17]. Вопросы численного решения уравнения мелкой воды обсуждаются, например, в книгах [18—20]. Перечисленные здесь уравнения реализованы сейчас в ряде вычислительных комплексов, некоторые из них доступны, в том числе комплекс AVI-NAMI, разработанный в нашей группе [21]. Именно эти модели и были использованы после цунами для его моделирования, причем результаты расчетов уже в течение января 2005 г. стали доступны специалистам. Самым первым, выполнив-

шим расчеты распространения цунами как в Индийском океане, так и во всем мире, оказался В. Титов из НОАА [7]. Наша группа оказалась в первой «десятке», представившей результаты для открытого доступа, в настоящее время они опубликованы [22].

3. Дисперсионная теория волн цунами

Эйфорию расчетчиков «взорвал» Е. Куликов из ИОРАН (Москва), поместивший в Интернете результаты сравнения спутниковых данных с оценками скорости распространения волн цунами в рамках точного линейного дисперсионного соотношения (рис. 4). Ясно видно, что оценочная по данным спутниковых изображений скорость распространения вариаций уровня моря меньше скорости длинных волн, используемых в теории мелкой воды, и, следовательно, дисперсия волн на воде должна приниматься во внимание.

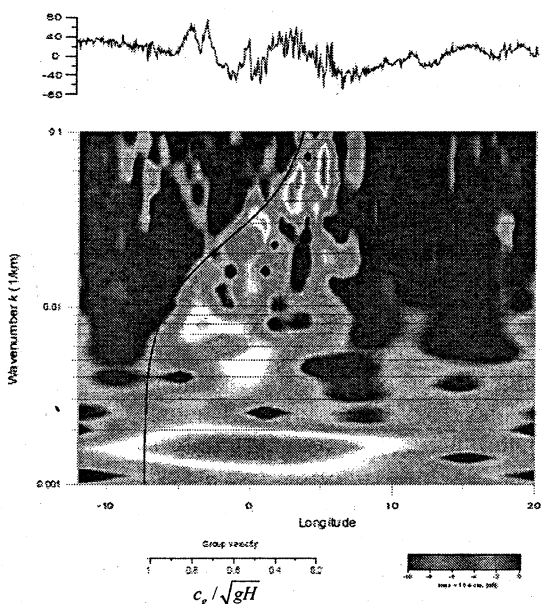


Рис. 4. Сопоставление рассчитанной групповой скорости волн на воде в рамках точного дисперсионного соотношения (сплошная линия) с вэйвлет-анализом спутниковых (Jason 1) данных (профиль уровня моря, зарегистрированный со спутника, представлен сверху).

Совершенно понятно, однако, что меньшая скорость распространения волновых возмущений может быть объяснена и в рамках теории мелкой воды с учетом переменной глубины моря, что приводит к формиро-

ванию многочисленных отраженных и переотраженных волн, приходящих позже головной волны. Эффекты дисперсии, связанные с наличием характерного горизонтального масштаба, хорошо известны в рамках волнового уравнения. Применительно к цунами, они обуславливают цуговый характер волн при однократной подвигке, задержку прихода волны максимальной амплитуды по сравнению с головной, существование захваченных и краевых волн. Поэтому спутниковая запись вполне может быть объяснена и в рамках мелкой воды, что и было немедленно продемонстрировано В. Титовым в Интернете (мы не будем приводить здесь соответствующий рисунок).

И все же аргументы Е. Куликова о «другой» дисперсии, связанной с конечностью вертикального масштаба (глубины), имеют под собой серьезное обоснование. Учитывая длинноволновое разложение дисперсионного уравнения в ряд

$$\omega = ck \left(1 - \frac{k^2 h^2}{6} \right) \quad (4)$$

(ω — частота и k — волновое число монохроматической волны), легко показать, что дисперсия проявляется на расстояниях [10, 11]

$$L_d \sim \frac{\lambda^3}{h^2}, \quad (5)$$

где $\lambda \sim 1/k$ — длина волны. Очаг цунами очень сильно неізотропен (см. рис. 3), и в направлении поперечной оси его масштаб составляет около 50 км, что не так уж сильно превосходит глубину океана (примерно на порядок). Это означает, что на расстоянии около 100 длин волн, или на расстоянии 5000 км, дисперсия должна проявиться в открытом океане. Легко понятно также, что наибольшее излучение от анизотропного источника идет в направлении малой оси, так что в принципе дисперсия может сказаться на расстояниях от Индонезии до Индии, где последствия цунами были катастрофичны и имеется много данных регистрации цунами.

Однако проверка этой гипотезы оказалась непростым делом. Уравнения Буссинеска (в декартовых координатах) хорошо известны в теории нелинейных волн и активно используются при моделировании волновых процессов в прибрежной зоне моря. Фактически существует много версий уравнений Буссинеска, которые в общем виде могут быть записаны как

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \nabla) \mathbf{u} + g \nabla \eta = \mathbf{D}_1, \quad \frac{\partial \eta}{\partial t} + \operatorname{div}[(h + \eta) \mathbf{u}] = D_2, \quad (6)$$

где в правой части содержатся дисперсионные слагаемые. При

$$\mathbf{D}_1 = \frac{h}{2} \frac{\partial}{\partial t} \left[\nabla \operatorname{div}(h \mathbf{u}) - \frac{h}{3} \nabla \operatorname{div}(\mathbf{u}) \right], \quad D_2 = 0 \quad (7)$$

система (6) называется системой Перегринна. Вид дисперсионных функций (даже в линейном приближении) не является однозначным. Наиболее

просто это продемонстрировать на различных вариантах длинноволновой аппроксимации дисперсионного соотношения

$$\omega \approx ck \left(1 - k^2 h^2 / 6\right) \approx \frac{ck}{1 + k^2 h^2 / 6} \approx ck \frac{1 - \alpha k^2 h^2}{1 - (\alpha + 1/6)k^2 h^2} \quad (8)$$

(α — произвольное число), совпадающих в пределе длинных волн, но различающихся в области коротких волн. Этим произволом можно воспользоваться для выбора интегрируемой версии уравнений Буссинеска, на чем мы не будем останавливаться. Более важно, что можно попытаться выбрать аппроксимации дисперсионного соотношения, наиболее точно отражающие дисперсионные свойства в более широкой области волновых чисел, а не только в пределе длинных волн. «Физический» произвол для этих манипуляций связан с неоднозначностью определения скорости течения в (6): это может быть средняя скорость течения по глубине, или поверхностная скорость, или придонная скорость, или скорость на какой-то фиксированной глубине. Сейчас появилось множество работ в этом направлении [11, 23—28], причем не только чисто теоретических, и можно утверждать, что уравнения Буссинеска являются почти точными по нелинейности и дисперсии.

Применительно к проблеме моделирования цунами на больших расстояниях возникают две трудности. Во-первых, необходимо записать уравнения Буссинеска на вращающейся Земле, а сразу не очевидно, как это сделать, если мы пытаемся получить уравнения достаточно полные по нелинейности и дисперсии. Во-вторых, даже уже известные и тестированные уравнения Буссинеска на огромных трассах ведут к большим численным ошибкам, так как необходимо разрешать существенно более малые масштабы (масштаб дисперсии — это глубина воды), чем в рамках теории мелкой воды. Эти трудности еще только преодолеваются, поэтому в ряде коллективов исследователей выполнялись только одномерные, весьма модельные, расчеты. Но и одномерные, и двумерные расчеты показали, что дисперсионные эффекты почти не проявляются на расстояниях от Индонезии до Индии [14, 29, 30]. На рис. 5 показаны результаты расчетов [29] цунами в акватории Индийского океана в рамках теории мелкой воды (слева) и в рамках дисперсионной теории (справа). Аналогичные результаты получены в нашей работе [14] для распределения поля вблизи побережья Бангладеш. Только в небольшом числе точек разность достигает 10%. На больших расстояниях как будто бы влияние дисперсии нарастает; так, разница в расчетах нарастает до 30% вблизи острова Диего — Гарсиа (расположенного на полпути от Индии до Мадагаскара), однако такие расстояния уже на пределе вычислительных возможностей. На больших расстояниях, если мы говорим не о самой первой волне, начинают сказываться также эффекты отражения волн от берегов (обычно моделируемых не поглощающимися вертикальными стенками), так что

расчетное поле начинает отличаться от наблюдаемого в рамках всех моделей. В результате пока еще трудно говорить однозначно о реальной важности дисперсионных эффектов при распространении цунами в открытом океане.

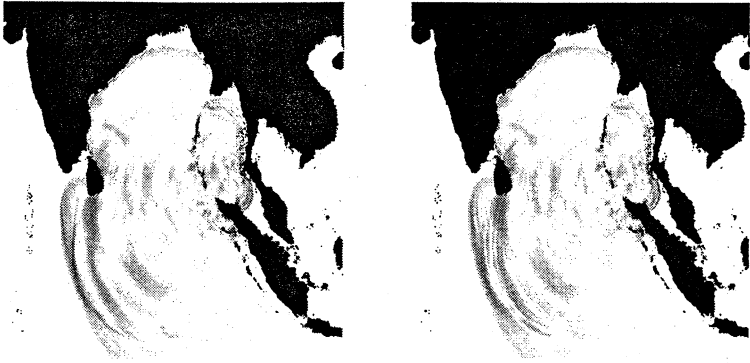


Рис. 5. Результаты расчета цунами 2004 года в рамках теории мелкой воды (слева) и дисперсионной теории (справа).

4. Роль солитонов в проблеме цунами

Итак, по расчетам и наблюдаемым данным цунами 2004 года дисперсия и нелинейность особо не проявились в открытом океане, и здесь линейной теории мелкой воды оказывается достаточно. Поскольку, однако, нелинейность велика вблизи берега, то естественно попытаться искать нелинейные и дисперсионные эффекты в волне, когда она приближается к берегу. Мы уже приводили фотографию обрушивающейся волны на берег (см. рис. 1), и таких доказательств «чистой» нелинейности цунами 2004 года бесчисленное множество. Может ли все-таки дисперсия конкурировать с дисперсией и приводить, например, к формированию солитонов? Оценим, прежде всего, характерные параметры солитонов, используя для этого уравнение Кортевега — де Вриза, справедливое для случая малой нелинейности и дисперсии:

$$\frac{\partial \eta}{\partial t} + \sqrt{gh} \left(1 + \frac{3\eta}{2h} \right) \frac{\partial \eta}{\partial x} + \frac{\sqrt{gh} h^2}{6} \frac{\partial^3 \eta}{\partial x^3} = 0. \quad (9)$$

Аналитическое выражение для солитона в рамках этого уравнения хорошо известно:

$$\eta(x, t) = H \operatorname{sech}^2 \left[\sqrt{\frac{3H}{4h}} \frac{x - \sqrt{gh}(1 + H/2h)t}{h} \right], \quad (10)$$

где H — высота солитона. Характерная длина солитона по уровню $0,5H$ есть

$$\lambda = 2h\sqrt{\frac{4h}{3H}} \ln(1+\sqrt{2}) \cong 2h\sqrt{\frac{h}{H}}. \quad (11)$$

Значения длины солитона для типичных условий даны в таблице. Длина солитона имеет порядок десятка и сотен километров в открытом океане (на глубине 1—5 км) и меньше 1 км на мелкой воде (50—100 м). Как мы уже указывали, длина волн цунами в открытом океане также имеет порядок десятка и сотен километров, что как будто бы указывает на применимость теории солитонов. Однако важен еще масштаб проявления нелинейных и дисперсионных эффектов, а он составляет несколько тысяч километров, см. формулу (5). Именно поэтому существующие вычислительные программы не смогли отловить дисперсионных эффектов, и это задача ближайшего будущего.

Характерная длина солитона [λ , км]

Глубина океана h , м	Высота солитона H , м			
	1	5	10	20
50	0,72	0,32	0,23	0,16
100	2	0,91	0,64	0,46
500	22,8	10,2	7,2	5,1
1000	64,4	28,8	20,4	14,4
5000	720	322	227	161

На мелководье же, как указывалось, нелинейность всегда превалирует. Чтобы дисперсия могла ее компенсировать, шельф должен быть достаточно длинным, чтобы дисперсия успела сказаться. Этот эффект уже ранее анализировался на примере цунами 1983 года в Японском море [11]. Рисунок 6 воспроизводит характерные формы волны цунами на побережье Японии во время этого цунами. Ясно видно большое разнообразие этих форм, в том числе некоторые из них напоминают пару солитонов или развал ударной волны на солитоны. Такие формы особенно характерны для достаточно протяженных шельфов, правда, надо сказать, что и период этого цунами был достаточно мал (до 5 мин, в то время как во время цунами 2004 года период составлял около 30 мин). Тем не менее, и во время цунами 2004 года наблюдались разнообразные формы волн; так, вблизи побережья Таиланда наряду с чисто ударной волной (см. рис. 1) сделаны фотографии волнообразного бора (рис. 7).

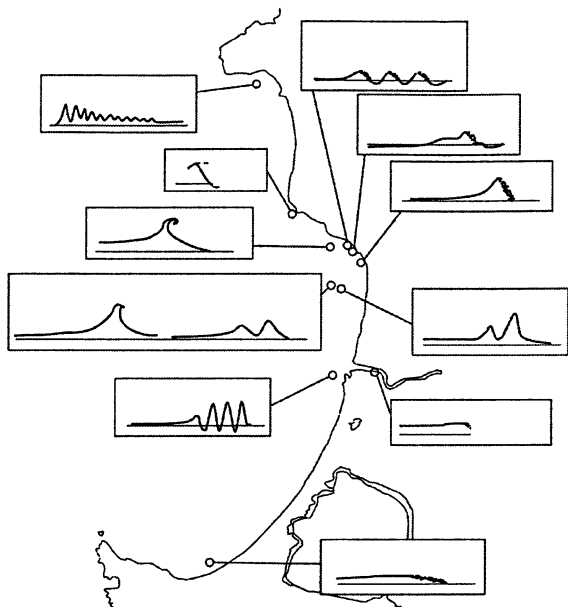


Рис. 6. Формы волн цунами на побережье Японии во время события 1983 года

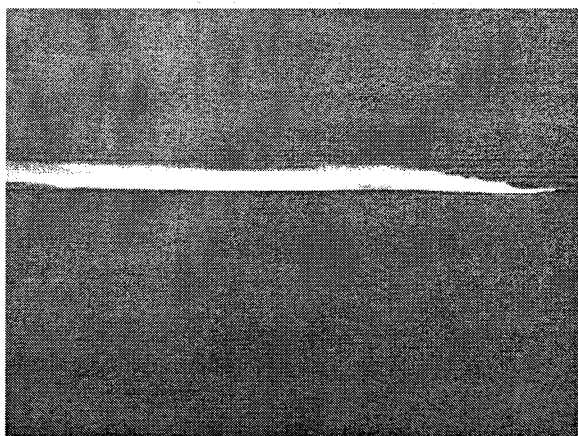


Рис. 7. Фотография волнообразного бора во время цунами 2004 года

Последняя фотография (см. рис. 7) попала к нам сравнительно недавно, поэтому мы вначале обратили внимание на Малаккский пролив (между Индонезией и Малайзией), где мелководный шельф достаточно

протяжен. Именно этот район (рис. 8) был выбран для выяснения роли дисперсионных эффектов [31]. В качестве дисперсионной модели сначала были выбраны уравнения Буссинеска, почти полные по нелинейности, однако мы «запутались» в результатах и вычислительных неустойчивостях, поэтому перешли к более простой модели, основанной на уравнении Кортевега — де Вриза, модифицированном с учетом переменности глубины по пространству. (Параллельно с нами аналогичные расчеты, но в рамках одномерных уравнений Буссинеска, выполнены в [30]). Моделирование волнового поля в эпицентральной области проводилось в рамках теории мелкой воды, а рассчитанная запись цунами на входе в Малаккский пролив (точка 1 на рис. 8) выбиралась в качестве исходной для моделирования в рамках уравнения Кортевега — де Вриза. Результаты расчетов в нескольких точках вдоль пролива (точки показаны на рис. 8) представлены на рис. 9. На каждом из рисунков указано расстояние от первой точки и глубина места. Из рис. 9 видно, что на расстоянии до 370 км (точка 4) обе модели описывают примерно одинаково длинноволновую компоненту волнового профиля. Различие проявляется в описании фронта волны, на котором появляются мелкомасштабные солитонные возмущения в рамках дисперсионной модели. На расстоянии в 450 км на глубине 36 м (точка 5) эти возмущения начинают влиять на форму основной волны.

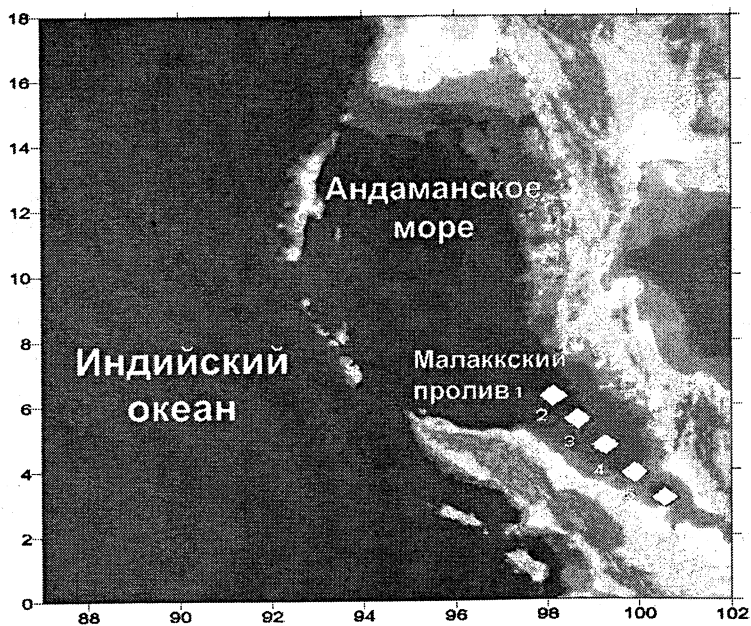


Рис. 8. Малаккский пролив и Андаманское море в Индийском океане

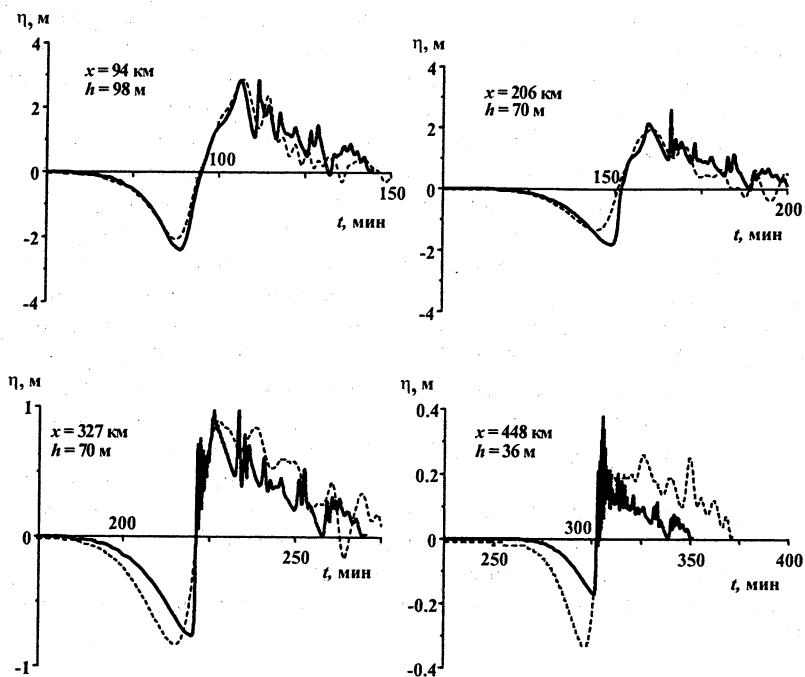


Рис. 9. Результаты расчетов волнового поля в рамках дисперсионной модели (сплошные линии) и теории мелкой воды (штриховые линии) в различных точках.

Более детально фронтальная часть волны в точке 5 показана на рис. 10. Характерная длительность солитонов составляет 1 мин, а их амплитуда приблизительно в два раза больше амплитуды волны, вычисленной в рамках теории мелкой воды. Мы, к сожалению, не могли просчитать динамику волн на больших расстояниях из-за резкого увеличения числа мелкомасштабных осцилляций. Однако ясно, что солитоны на мелководье должны разрушаться и сильно затухать. Не исключено, что именно с этим связан слабый эффект цунами на южном побережье Малайзии. Сейчас мы пытаемся выполнить расчеты в рамках точных уравнений гидродинамики, чтобы понять динамику зарождения и затухания солитонов, однако уже рассмотренный пример убедительно показывает, что солитоны должны быть неотъемлемой частью волн цунами на достаточно протяженных мелководных шельфах, как это и наблюдалось вблизи Таиланда (см. рис. 7).

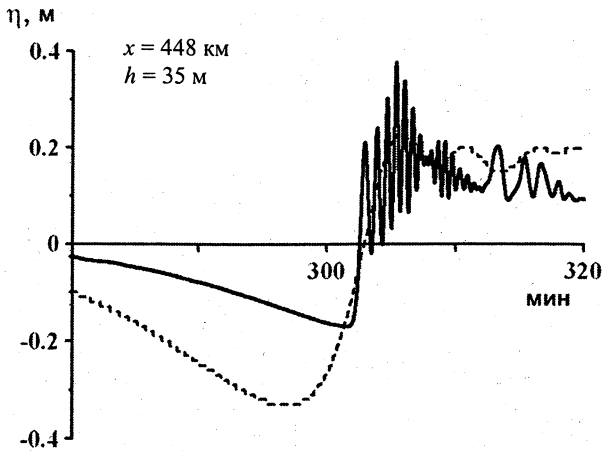


Рис. 10. Рассчитанная фронтальная часть волны в точке 5 (сплошная линия — с учетом дисперсии, штриховая — без учета дисперсии).

5. Заключение

С точки зрения науки, цунами 2004 года в Индийском океане дало уникальный материал для проверки и верификации физико-математических и вычислительных гидродинамических моделей. Оно подтвердило, что базовой моделью для описания сильных цунами сейсмического происхождения может являться нелинейная теория мелкой воды, хорошо реализуемая численно. Между тем, простые физические оценки и весьма трудоемкие расчеты показывают, что на трансокеанических трассах (в десятки тысяч километров) эффекты дисперсии становятся заметными, и они, скорее всего, в ближайшее время по мере развития вычислительных алгоритмов будут подтверждены более аккуратными расчетами. Солитоны, как часть волнового поля, должны быть заметными на длинных и протяженных шельфах, такие районы уже сейчас могут быть выделены на картах. Однако прямого подтверждения существования солитонов в волнах цунами пока еще не получено, хотя тенденция их формирования прослеживается как в расчетах, так и в данных наблюдений. В ближайшие годы, в связи с бурным развитием вычислительных программ, я уверен, солитоны будут выделены в волновом поле.

Работа выполнена при поддержке гранта РФФИ 05-05-64265 и ИНТАС-Сибирь 06-1000013-9236. Автор благодарен А. Зайцеву, А. Козелкову, А. Куркину, Т. Талиповой и А. Ялчинер, с энтузиазмом «вложившимся» в численное моделирование цунами 2004 года.

Литература

1. Никонов А. А. Волна, которая обошла весь мир // Природа. 2005. Вып. 5. С. 23—28.
2. Завьялов А. Д. Землетрясение у берегов Суматры // Природа. 2005. Вып. 5. С. 29—35.
3. Borrero J. C. Field Data and Satellite Imagery of Tsunami Effects in Banda Aceh // Science. 2005. V. 308. P. 1596.
4. Куркин А. А., Пелиновский Е. Н., Полухин Н. В., Зайцев А. И., Кайтренко В. М., Королев Ю. П., Разжигаева Н. Г., Прасетья Г. С., Хидаят Р. Катастрофическое цунами 26 декабря 2004 г., данные полевых исследований северо-восточной части островов Суматра и Сimalур // Изв. АИН РФ. Прикладная математика и механика. 2005. Т. 13. С. 63—75.
5. Lay T., Kanamori H., Ammon C., Nettles M., Ward S., Aster R., Beck S., Bilek S., Brudzinski M., Butler R., DeShon H., Ekstorm G., Satake, Sipkin S. The Great Sumatra-Andaman Earthquake of 26 December 2004 // Science. 2005. V. 308. P. 1127—1133.
6. Choi B. H., Hong S. J., Pelinovsky E. Distribution of runup heights of the December 26, 2004 tsunami in the Indian Ocean // Geophys. Res. Lett. 2006. V. 33, № 13, L13601, doi: 10.1029/2006GL025867.
7. Titov V., Rabinovich A. B., Moffeld H. O., Thomson R. E., Gonzalez F. I. The global reach of the 26 December 2004 Sumatra tsunami // Science. 2005. V. 309. P. 2045—2048.
8. Merrifield M. A., Firing Y. L., Aarup T., Agricole W., Brundrit G., Chang-Seng D., Farre R., Kilonsky B., Knight W., Kong L., Magori C., Manurung P., McCreery C., Mitchell W., Pillay S., Schindele F., Shillington F., Testut L., Wijeratne E. M., Caldwell P., Jardin J., Nakahara S., Porter F.-Y., Turetsky N. Tide gauges observations of the Indian Ocean tsunami, December 26, 2004 // Geophys. Res. Lett. 2005. V. 32, L09603, doi:10.1029/2005GL022610.
9. Pelinovsky E., Choi B. H., Stromkov A., Didenkulova I., Kim H. S. Analysis of tide-gauge records of the 1883 Krakatau tsunami // Tsunamis: case studies and recent developments, Advances in Natural and Technological Hazards Research. Springer, 2005. V. 23. P. 57—77.
10. Пелиновский Е. Н. Нелинейная динамика волн цунами. Горький: ИПФ АН СССР, 1982. 220 с.
11. Пелиновский Е. Н. Гидродинамика волн цунами. Н. Новгород: ИПФ РАН, 1996. 275 с.
12. Левин Б. В., Носов М. А. Физика цунами и родственных явлений в океане. М.: Янус-К, 2005. 360 с.
13. Okada S. Surface displacement due to shear and tensile faults in a half-space // Bull. Seismol. Soc. Am. 1985. V. 75. P. 1135—1154.
14. Ioualalen M., Pelinovsky E., Asavanant J., Lipikorn R., Deschamps A. On the weak impact of the 26 December Indian Ocean tsunami on the Bangladesh coast // Natural Hazards and Earth Science System (submitted).
15. Zahibo N., Pelinovsky E., Talipova T., Kozelkov A., Kurkin A. Analytical and numerical study of nonlinear effects at tsunami modelling // Applied Math. Comp. 2006. V. 174. P. 795—809.
16. Диденкулова И. И., Заибо Н., Куркин А. А., Левин Б. В., Пелиновский Е. Н., Соомере Т. Накат нелинейно деформированных волн на берег // ДАН. 2006. Т. 410, № 5.
17. Zahibo N., Pelinovsky E., Golinko V., Osipenko N. Tsunami Wave Runup on Coasts of Narrow Bays // Int. J. Fluid Mech. Res. 2006. V. 33. P. 106—118.
18. Шокин Ю. И., Чубаров Л. Б., Марчук А. Г., Симонов К. В. Вычислительный эксперимент в проблеме цунами. Новосибирск: Наука, 1988. 385 с.
19. Вольцингер Н. Е., Клеванный К. А., Пелиновский Е. Н. Длинноволновая динамика прибрежной зоны Л.: Гидрометеиздат. 1989. 271 с.
20. Kowalik Z., Murty T.S. Numerical Modeling of Ocean Dynamics. Advanced Series on Ocean Engineering. World Sci. 1993. V. 5. 496 p.
21. Куркин А. А., Зайцев А. И., Ялчинер А., Пелиновский Е. Н. Модифицированный вычислительный комплекс «ЦУНАМИ» для оценки рисков, связанных с цунами // Изв. АИН РФ. Прикладная математика и механика. 2004. Т. 9. С. 88—100.

22. Зайцев А. И., Куркин А. А., Левин Б. В., Пелиновский Е. Н., Ялчинер А., Троицкая Ю. И., Ермаков С. А. Моделирование распространения катастрофического цунами (26 декабря 2004 г.) в Индийском океане // Докл. РАН 2005. Т. 402. С. 388—392.

23. Wei G., Kirby J. T., Grilli S. T., Subramanya R. A fully nonlinear Boussinesq model for free surface waves. Part 1: Highly nonlinear unsteady waves // J. Fluid Mech. 1995. V. 294. P. 71—92.

24. Madsen P. A., Schaffer H. A. Higher-order Boussinesq-type equations for surface gravity waves: derivation and analysis // Phil. Trans. R. Soc. Lond. A. 1998. V. 356. P. 3123—3184.

25. Agnon Y., Madsen P. A., Schaffer H. A. A new approach to high order Boussinesq models // J. Fluid Mech. 1999. V. 399. P. 319—333.

26. Chen Q., Kirby J. T., Dalrymple R. A., Kennedy A. B., Chawla A. Boussinesq modeling of wave transformation, breaking, and run-up // J. Waterway, Port, Coast, and Ocean. Engng. 2000. V. 126. P. 48—56.

27. Madsen P. A., Bingham H. B., Liu H. A new Boussinesq method for fully nonlinear waves from shallow to deep water // J. Fluid Mech. 2002. V. 462. P. 1—30.

28. Madsen P. A., Bingham H. B., Schaffer H. A. Boussinesq-type formulations for fully nonlinear and extremely dispersive water waves: derivation and analysis // Proc. R. Soc. Lond. A. 2003. V. 459. P. 1075—1104.

29. Shigihara Y., Fujima K. Numerical modeling of Indian Ocean tsunami in the Maldives // Sumatra Tsunami on 26 December 2004. Korea, 2005. P. 49—56.

30. Glimsdal S., Pedersen G. K., Atakan K., Harbitz C. B., Langtangen H. P., Lovholt F. Propagation of the Dec. 26, 2004, Indian Ocean tsunami: effects of dispersion and source characteristics // Int. J. Fluid Mech. Res. 2006. V. 33. P. 15—43.

31. Pelinovsky E., Talipova T., Kurkin A., Kozelkov A., Zahibo N. Nonlinear and dispersion effects in the tsunami wave field // Proc. 22nd Int. Tsunami Symposium. Chania, 2005. P. 241—246.

НЕЛИНЕЙНЫЕ МЕХАНИЗМЫ И МОДЕЛИ ФОРМИРОВАНИЯ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ ПРОЦЕССОВ В ВЕРХНЕМ СЛОЕ ОКЕАНА

Ю. И. Троицкая, С. А. Ермаков

1. Введение

Начнем с определения верхнего слоя океана. Средняя глубина океана составляет около 4000 м, и гидрофизические характеристики основной его толщи практически неизменны. Только верхние 300—400 м подвержены сезонной изменчивости, этот слой и называется верхним, или деятельным, слоем океана. Для верхнего слоя океана характерна плотностная стратификация, связанная с тем, что его температура и соленость зависят от глубины. Пример типичных профилей температуры и солености в верхнем слое океана представлен на рис. 1, взятом из монографии [1]. Для стратификации верхнего слоя океана типично наличие перемешанного слоя, толщина которого составляет 60—70 м в умеренных и высоких широтах в летний период и доходит до 250 м в тропическом океане. Взаимодействие с атмосферой, донной топографией, морскими берегами и т. п. вызывает движения в верхнем слое океана, которые могут проявляться на морской поверхности и регистрироваться на изображениях, получаемых с помощью инструментов, размещенных на искусственных спутниках земли. Обобщенный спектр масштабов таких движений был предложен Р. В. Озмидовым [2] на основе обработки данных большого массива наблюдений. В этом спектре выделяются три интервала энергоснабжения: глобальный с масштабом порядка 1000 км, приливной с масштабом в 100 км и ветровой с масштабом в 10—100 м, а также инерционные интервалы между ними (рис. 2).

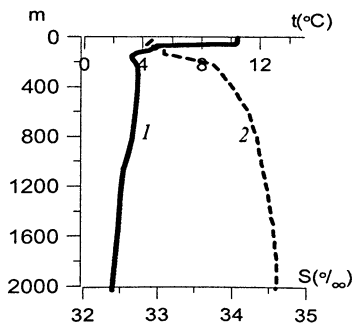


Рис. 1. Пример типичных профилей температуры (1) и солености (2) в верхнем слое океана в северной части Тихого океана [1].

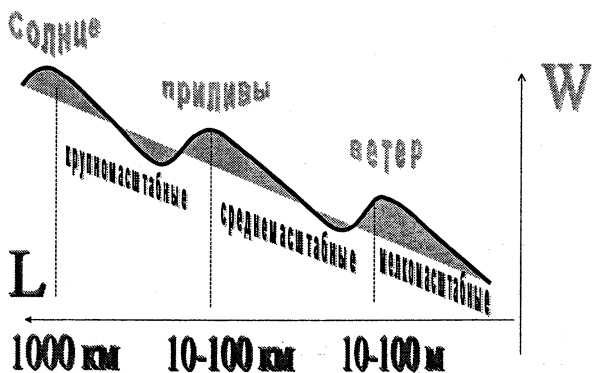


Рис. 2. Обобщенный спектр кинетической энергии движений в океане [2]

Среднемасштабные океанские движения определяют локальные гидрометеороусловия, и в связи с этим знание их параметров важно, например, для решения задач прогноза погоды. При этом, поскольку масштаб таких процессов составляет десятки километров, то оперативная информация о них может быть получена только дистанционными методами. Особое место при этом занимают спутниковые радиолокационные методы, преимуществом которых является сочетание глобальности обзора при высоком пространственном разрешении. Например, радиолокатор с синтезированной апертурой, размещенный на спутнике ENVISAT, обеспечивает радиолокационное изображение в полосе шириной 100 км с разрешением 25 м. Наряду с отмеченными выше, важным преимуществом радиолокационных методов перед, например, оптическими, является их всепогодность, связанная с прозрачностью облаков для электромагнитных волн сантиметрового диапазона, в котором работают спутниковые радиолокаторы. Следует подчеркнуть, однако, важную особенность радиолокационных методов, состоящую в том, что они отображают процессы только на поверхности океана, поскольку в рассматриваемом диапазоне длин радиоволн соленая морская вода является практически идеальным проводником и не проникает в толщу воды. Однако процессы в толще океана могут отображаться на морской поверхности и таким образом проявляться в радиолокационных изображениях поверхности океана.

В качестве примера на рис. 3 представлены радиолокационные изображения из коллекции Гамбургского университета [3], на которых видны проявления на поверхности различных процессов, происходящих в толще океана. Эти изображения получены с помощью радиолокатора бокового обзора с синтезированной апертурой (РСА), в котором полезный сигнал формируется за счет обратного рассеяния радиоволн на поверхностных волнах. Доминирующим механизмом рассеяния радиоволн при углах падения порядка и более 20—30° является резонансный, или брэг-

говский, механизм, когда длина рассеивающей волны λ_{wave} на поверхности воды определяется соотношением

$$\lambda_{wave} = \frac{\lambda}{2 \sin \theta}$$

(λ — длина волны электромагнитного излучения радиолокатора, θ — угол падения).

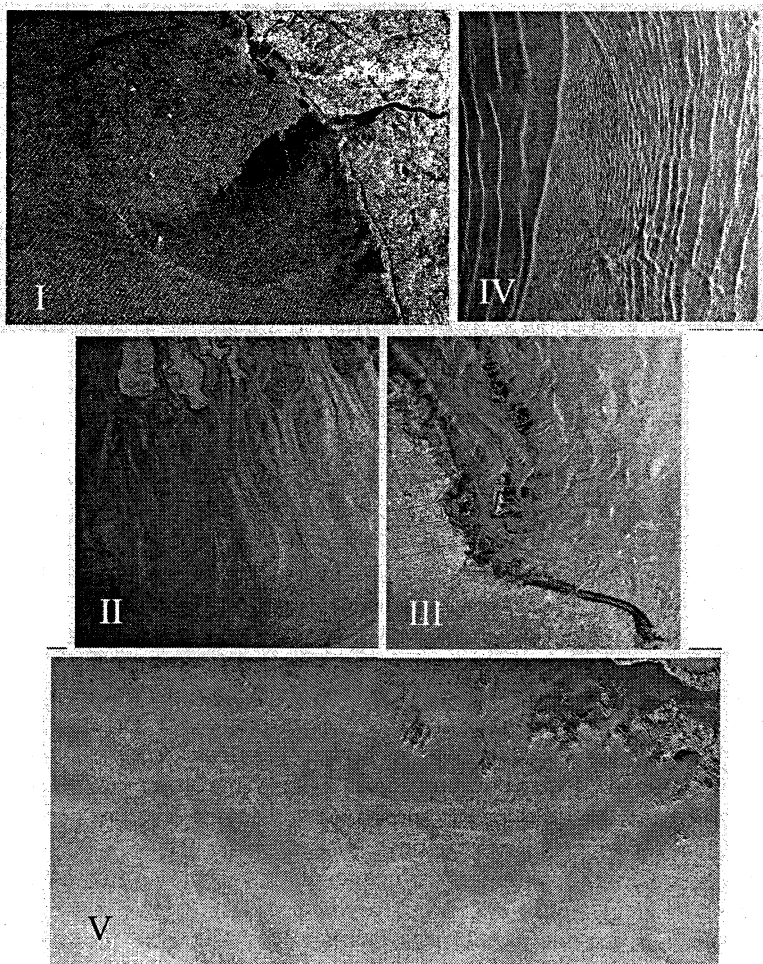


Рис. 3. I — зыбь (португальский шельф), донная топография (II — устье Ганга, III — восточный берег Китая), IV — океанские внутренние волны (Андаманское море), V — береговой фронт (восточный берег Китая).

Для самолетных и судовых радиолокаторов, работающих, как правило, в X-, Ku и Ka-диапазонах под большими углами падения, брэгговские длины волн находятся в диапазоне 4 мм — 2 см. Для спутникового радиолокатора SIR C-X-SAR, работающего в C- и X-диапазонах при умеренных углах падения, резонансные длины волн имеют порядок 2—5 см. Наконец, для PCA, размещенных на спутниках ERS-2, ENVISAT, брэгговская длина волны составляет примерно 7 см. Следует отметить, однако, что для последних из упомянутых PCA углы падения составляют 23°, и в этом случае более корректно описание рассеяния с учетом вкладов как резонансного, так и квазирезонансного механизмов рассеяния. Как показывают расчеты [4], практически следует учитывать рассеяние излучения волнами с длинами до двух-трех брэгговских. Таким образом, при интерпретации радиолокационных изображений морской поверхности необходим анализ модуляции мелкомасштабной ветровой ряби сантиметрового диапазона длин волн.

2. Гидродинамическая модель, основанная на уравнении Хассельмана

2.1. Основные уравнения

Важным элементом количественной теории формирования радиоизображения является гидродинамическая модель, которая описывает воздействие течения на поверхности на приводный ветер и поверхностные волны. При построении такой модели возникают сложности, связанные прежде всего с нелинейностью поля поверхностных волн и их взаимодействия с приводным ветром. Для описания модуляции коротких волн в поле течений используется уравнение Хассельмана [5] для спектральной плотности волнового действия $N = W(\vec{r}, \vec{k}, t) / \omega_0$ (здесь $W(\vec{r}, \vec{k}, t)$ — спектральная плотность энергии поверхностных волн):

$$\frac{\partial N}{\partial t} + \frac{\partial \omega}{\partial \vec{k}} \frac{\partial N}{\partial \vec{r}} - \frac{\partial \omega}{\partial \vec{r}} \frac{\partial N}{\partial \vec{k}} = 2B[N, \vec{U}]N + \text{Int}[N] \equiv R[N, \vec{U}], \quad (1)$$

где $\omega = \omega_0 + k\vec{U}_w(\vec{r}, t)$, $\omega_0 = \sqrt{g|\vec{k}| + T|\vec{k}|^3}$.

Левая часть уравнения (1) представляет собой полную производную от N в пространстве координат и волновых чисел; правая часть — это ветровой инкремент $B[N, \vec{U}]$ и «интеграл столкновений» $\text{Int}[N]$, описывающий энергообмен за счет нелинейного взаимодействия волн. При этом выражение для «интеграла столкновений» имеет вид

$$\begin{aligned} \text{Int}[N] = & 4\pi \iiint T_{0,1,2,3}^2 \delta_{0+1-2-3} \delta_{0+1-2-3}^\omega \times \\ & \times [N_3 N_2 (N_1 + N_0) - N_1 N_0 (N_2 + N_3)] d\vec{k}_1 d\vec{k}_2 d\vec{k}_3. \end{aligned} \quad (2)$$

Выражения для коэффициентов четырехволнового взаимодействия $T_{0,1,2,3}$ приведены, например, в [6]. Они достаточно сложны, и поэтому практически не используются при решении задач о модуляции волн в присутствии неоднородных течений. Для аппроксимации нелинейных слагаемых часто используется релаксационная модель. Рассмотрим подробнее, как она вводится. Сначала рассматривается равновесное состояние, в котором ветровая накачка уравнивается нелинейным взаимодействием волн:

$$2B[N_0, \bar{U}]N_0 + \text{Int}[N_0] \equiv R[N_0, \bar{U}] = 0.$$

Далее рассматривается пространственно однородное возмущение

$$N = N_0 + N_1(\vec{k}) \exp(-\beta_r(\vec{k})t)$$

и находится время релаксации β_r , как решение задачи на собственные значения. Модельное выражение для «интеграла столкновений», полученное в однородной задаче, используется в неоднородной:

$$\frac{\delta R}{\delta N_0} [N_1] = -\beta_r N_1, \quad \beta_r = 2 + 5 \text{Im} \omega.$$

При этом уравнение для пространственно неоднородного возмущения спектральной плотности волнового действия имеет вид

$$\frac{\partial N_1}{\partial t} + \frac{\partial \omega_0}{\partial \vec{k}} \frac{\partial N_1}{\partial \vec{r}} + \beta_r N_1 = k_i \frac{\partial U_{wi}}{\partial x_j} \frac{\partial N_0}{\partial k_j} + 2 \frac{\partial B}{\partial U} U_a N_0. \quad (3)$$

Таким образом, релаксационная модель — квазилокальная как в пространстве координат, так и волновых чисел. Ясно, что эта модель применима:

- 1) в интервале ветровой накачки, т. е. для волн с длиной менее 1 м;
- 2) в случае, когда интеграл столкновений почти «диагональный», т. е. определяющим является нелинейное взаимодействие волн, имеющих близкие волновые числа;
- 3) если масштаб возмущений велик по сравнению с масштабом релаксации волнового поля.

Из уравнения (3) видно, что возмущение спектра ветровых волн в присутствии неоднородного течения вызывается двумя факторами: трансформацией волн на неоднородном течении (первое слагаемое в правой части уравнения (3)) и модуляцией ветрового инкремента (второе слагаемое в правой части).

Первый механизм обсуждается уже более 20 лет. Например, в работах В. И. Таланова с соавторами [7] он интерпретируется как адиабатическое взаимодействие волны с течением. При этом в [8] доказывается подобие кинематики волновых пакетов и движения заряженных частиц в электрическом поле. Уравнения движения для волновых пакетов следуют из кинетического уравнения:

$$\begin{cases} \frac{d\bar{r}}{dt} = \frac{\partial \omega}{\partial \bar{k}}, \\ \frac{d\bar{k}}{dt} = -\frac{\partial \omega}{\partial \bar{r}}. \end{cases}$$

При этом из аналогии кинетического уравнения (2) с уравнением Гамильтона — Якоби следует, что частота подобна гамильтониану:

$$\omega = \omega_0(k) + kU_w(x, t).$$

Наиболее сильное взаимодействие волн с течениями происходит в окрестности точки группового синхронизма, где групповая скорость волны (аналог скорости частиц) близка к скорости течения. Из выражения для гамильтониана вблизи точки группового синхронизма

$$\omega = \omega_0(k_*) + k_*U_w(x, t) + \frac{(k - k_*)^2}{2m_*}$$

(где k_* — волновое число, при котором групповая скорость волны равна средней скорости течения) ясно, что скорость потока аналогична потенциальной энергии, а градиент скорости течения аналогичен потенциальной силе. Это означает, что модуляция волн определяется градиентом скорости течения на поверхности: максимум спектра — в зоне конвергенции, минимум — в зоне дивергенции.

Второй механизм — модуляция инкремента коротких ветровых волн. Он состоит в следующем: неоднородное течение на поверхности воды вызывает модуляцию скорости приводного ветра, что приводит к модуляции инкремента ветровых волн и к модуляции их амплитуды. В отличие от кинематического механизма механизм модуляции инкремента локальный, т. е. он определяется не градиентом скорости течения на поверхности воды, а самой скоростью. При этом соотношение эффективности кинематического механизма и механизма модуляции инкремента определяется соотношением масштаба течения и характерного масштаба нарастания волны в поле ветра.

Оказывается, что на эффект модуляции инкремента принципиально влияет нелинейное взаимодействие волн с ветром. Основным нелинейным эффектом при этом является деформация среднего профиля скорости. Поскольку ветер генерирует волны и, следовательно, передает им некоторое количество движения, то следствием нелинейного взаимодействия волн и ветра должно быть появление отрицательной добавки к средней скорости ветра. С точки зрения теории нелинейных волн — это эффект детектирования волновых возмущений, индуцированных в воздушном потоке волнами на поверхности воды. При этом нелинейная добавка к скорости ветра имеет принципиальное значение для определения

фазы модуляции ветровых волн в присутствии неоднородного течения. Действительно, предположим, что на поверхности воды имеется неоднородное течение и существует приводный ветер. Тогда области на поверхности воды, в которых течение противоположно направлению ветра, находятся в условиях эффективно более сильного ветра, и волны в них генерируются более эффективно. Но это означает, что и отрицательная добавка к скорости ветра в этих областях будет больше. Результат будет зависеть от параметров задачи и должен рассчитываться.

В рамках гидродинамической модели, основанной на уравнении Хасельмана, можно получить явное выражение для вариаций спектра волн в неоднородном поле течений. Если поле скорости на поверхности воды имеет вид гармонической волны

$$U_w = U_{w0} \exp(i(qx - \Omega t)), \quad (4)$$

имеем

$$N_{10} = \frac{iqU_{w0}}{iq\left(\frac{d\omega}{dk} - V_0\right) + \beta_r} k \frac{dN_0}{dk} + \frac{\partial B}{\partial u_*^2} \frac{T_{10}}{iq\left(\frac{d\omega}{dk} - V_0\right) + \beta_r} N_0, \quad (5)$$

Кинематический механизм *Модуляция инкремента*

$$T_{10} = \frac{0,4u_*U_{w0}}{\ln \frac{qz_0}{0,16\eta_0(\beta)} + 0,4u_* \frac{\partial \Delta u_0}{\partial T_1} \Big|_{T_1=0} + i\varphi(\beta)}$$

Здесь $\beta = -(\Delta u_0 + u_0) - \ln \eta_0$, а функции $\eta_0(\beta)$ и $\varphi(\beta)$ получаются на основании решения уравнения, описывающего возмущения скорости ветра, вызванные поверхностным течением (4). В безразмерных переменных оно имеет вид

$$\frac{d}{d\eta}(\eta \Psi_{l\eta\eta}) - i \left[(\ln \eta + \beta) \Psi_{l\eta} - \frac{\Psi_l}{\eta} \right] = 0,$$

а его решения, убывающие на бесконечности, имеют асимптотику

$$\Psi_{l\eta} = A \left(\ln \frac{\eta}{\eta_0} + i\varphi(\beta) \right).$$

Выражение (5) определяет решение прямой задачи: по известному полю течений на поверхности определяется распределение спектра поверхностных волн; при условии применимости формулы (5), очевидно, легко можно найти и решение обратной задачи, т. е. восстановить поле неоднородных течений на поверхности воды по измеренным флуктуациям поверхностного волнения.

2.2. Примеры взаимодействия волн с течением

Рассмотрим примеры количественного описания взаимодействия волн с течением в рамках уравнения Хассельмана.

Модуляция ветровых волн волнами зыби

Рассмотрим длинную поверхностную волну (с длиной более 10 м) для простоты с гармоническим возвышением поверхности $\eta = a \cos(\omega t - kx)$ (где a , k , ω — амплитуда, волновое число и частота длинной волны) и короткие волны на ее фоне. В спектре коротких волн присутствуют, в том числе, «брэгговские» компоненты, которые рассеивают радиоволны. Мощность рассеянных радиоволн P модулирована с пространственными и временными масштабами длинных волн

$$P = P_0 \left(1 + ka \operatorname{Re} \left\{ |MTF| \exp[i(kx - \omega t - \varphi_{MTF})] \right\} \right).$$

Модуляционную передаточную функцию (МПФ, комплексную величину с модулем $|MTF|$ и фазой φ_{MTF}) можно представить в виде суммы компонент, описывающих модуляцию, вызванную различными механизмами [8]

$$MTF = MTF_{ilt} + MTF_{hydro},$$

где MTF_{ilt} — геометрическая МПФ, описывающая модуляцию рассеянных радиоволн уклонами длинных волн; MTF_{hydro} — гидродинамическая МПФ, которая описывает трансформацию спектра брэгговских коротких волн гидродинамическими полями длинной волны.

Для демонстрации эффективности гидродинамической модуляции, приведем рис. 4 из обзорной статьи [9], на котором представлены различные составляющие МПФ как функции частоты длинной волны. Из рис. 4 очевидно, что модуль MTF_{hydro} превосходит модуль MTF_{ilt} для поверхностных волн с периодами более 3—5 секунд. Экспериментально установлено, что модуль гидродинамической $|MTF|$ — величина порядка 5—15, убывающая с ростом частоты волны и скорости ветра, а ее фаза φ_{MTF} близка к нулю (хотя возможны ее существенные вариации), т. е. максимум спектра коротких волн наблюдается вблизи гребней волн зыби.

Для объяснения этих особенностей прежде всего следует учесть эффект модуляции инкремента, чтобы объяснить рост модуля МПФ с увеличением периода волны зыби. Действительно, как уже отмечалось выше, модуляция спектра поверхностных волн за счет эффекта модуляции инкремента возрастает с ростом скорости переменного течения на поверхности воды, а орбитальная скорость в волнах зыби будет возрастать с возрастанием ее периода. Для объяснения фазы МПФ необходимо принять во внимание нелинейное взаимодействие волн с ветром. Как показало сравнение с экспериментом [10], учет конечности амплитуды ветровых волн дает значения модуля и фазы гидродинамической МПФ, удовлетворительно согласующиеся с имеющимися натурными данными.

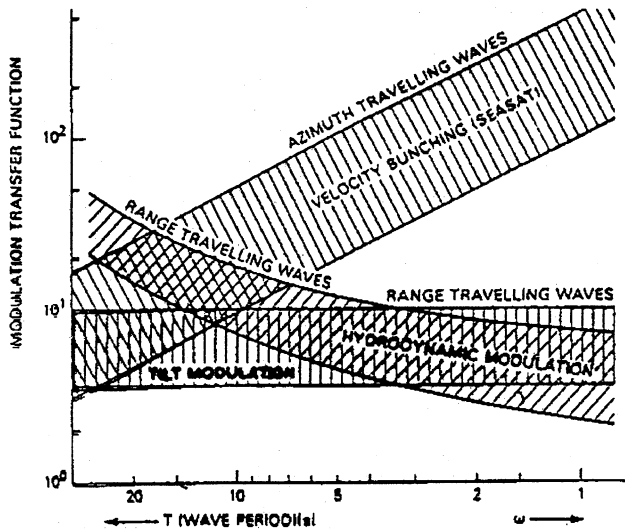


Рис. 4. Компоненты МПФ [9]

Модуляция ветровых волн в присутствии внутренних волн

Рассмотрим проявление кинематического механизма и механизма модуляции инкремента при отображении на морской поверхности океанских внутренних волн. На рис. 5 приведено изображение внутренних волн в Андаманском море, взятое с веб-сайта Гамбургского университета [3]. Вероятнее всего на изображении наблюдается проявление кинематического механизма, когда на переднем склоне внутренних волн (в зоне конвергенции) образуется яркая полоса, соответствующая повышенной интенсивности поверхностного волнения, а на заднем склоне, в зоне дивергенции, — темная полоса [3]. Заметим, что такой типичный вид изображения, на котором наблюдается чередование полос, интенсивность которых выше и ниже фоновой, легко позволяет определить направление распространения внутренних волн (см. рис. 5).

Однако такая картина наблюдается не всегда. На рис. 6 показан пример проявления внутренних волн на поверхности, когда, по-видимому, работает механизм модуляции инкремента. Эти изображения были получены с береговым радиолокатором X-диапазона с рабочей частотой 3 см в ходе эксперимента COPE на Орегонском шельфе в 1995 г. [11]. Были измерены интенсивные внутренние волны приливного происхождения. Одновременно береговым локатором были обнаружены выглаженные области на морской поверхности в фазе с гребнями внутренних волн (а не на склонах).

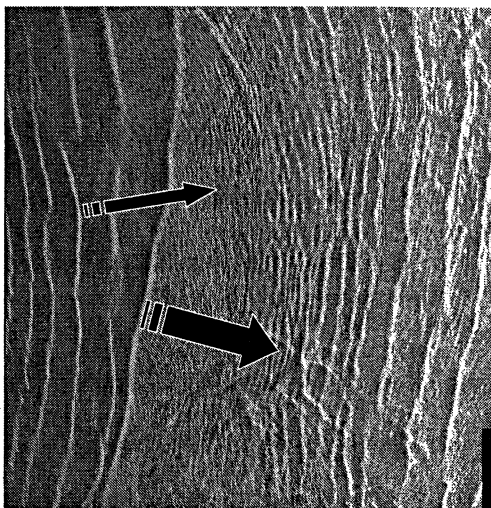


Рис. 5. Внутренние волны в Андаманском море. РСА-изображение получено со спутника ERS-2 (из коллекции изображений [3]). Стрелки указывают направление распространения волн.

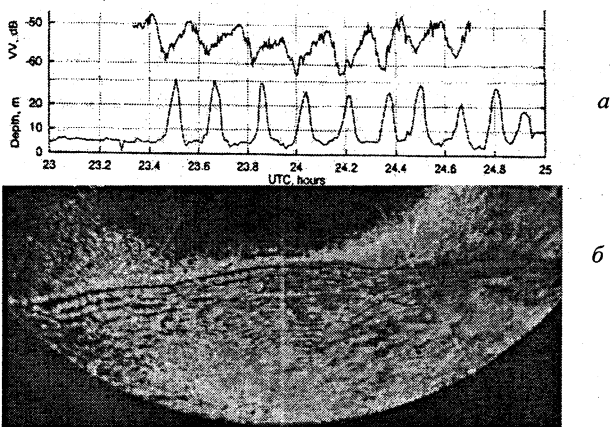


Рис. 6. Эксперимент COPE: *а* — изотермы и сечение рассеяния радиолокационного сигнала, *б* — радиолокационная панорама [11].

Для интерпретации этого эксперимента была предложена модель интенсивных внутренних волн, основанная на решении комбинированного уравнения Кортевега — де Вриза с квадратичной и кубической нелинейностью [12]. Приближенный аналитический метод решения основан на

представлении N -солитонного решения как суперпозиции $2N$ кинков противоположной полярности [12]. В рамках этой модели рассчитывалось поле скорости на поверхности воды и по нему вычислялся модулированный инкремент поверхностных волн. Оказалось, что для условий данного эксперимента скорость ветра над вершиной солитона внутренних волн относительно воды была ниже порога генерации поверхностных волн с длиной 1,5 см, которые являются брэгговскими для радара X -диапазона, используемого в данном эксперименте.

3. Связанные и свободные волны

Следует заметить, что, основываясь на уравнении Хассельмана, не всегда можно количественно описать наблюдаемые явления, в частности невозможно понять, почему на изображении, представленном на рис. 5, видны проявления внутренних волн. Действительно, как отмечалось выше, для того чтобы обеспечить эффективную модуляцию поверхностных волн в поле течений внутренней волны, нужно, чтобы их групповая скорость была близка к фазовой скорости внутренних волн. Для внутренних волн, которые видны на изображениях в Андаманском море (см. рис. 5), фазовая скорость внутренних волн — около 1 м/с, групповая скорость поверхностных волн с длиной волны 7 см (брэгговская длина волны для РСА на европейских спутниках) равна примерно 25 см/с, т. е. условия группового синхронизма не выполняются. Однако волны хорошо видны на изображениях.

Чтобы разрешить этот парадокс, примем во внимание, что уравнение Хассельмана получено в слабонелинейном приближении, т. е. для волн, которые слабо взаимодействуют с ветром и друг с другом, и дисперсионное соотношение которых мало отличается от дисперсионного уравнения для линейных поверхностных волн. Основным источником таких волн является ветер, а слабое взаимодействие их друг с другом приводит к релаксации их спектра к равновесному. Такие волны будем называть свободными.

В то же время, как показывают наблюдения за поверхностью моря или любого достаточно большого водоема, поверхностные волны дециметрового и метрового диапазонов при достаточно сильном ветре становятся сильно нелинейными. Сильная нелинейность приводит к существенному искажению формы волн и возникновению возмущений интересующих нас сантиметрового и дециметрового масштабов, которые, по существу, представляют собой гармоники порождающей волн, поэтому такие гармоники получили в литературе название связанных волн. Фазовая скорость таких волн равна фазовой скорости породивших их длинных волн.

Модуляция длинных волн крупномасштабным течением вызывает эффективную модуляцию дециметровых и метровых волн, которые находятся с этим течением в условиях группового синхронизма, при этом возникает и модуляция связанных волн. Последовательной теории этого процесса в настоящее время не создано, скорее всего, она будет полуэмпирической.

Поскольку связанные и свободные волны имеют сильно различающиеся фазовые скорости, их проявления в радиолокационном сигнале легко можно различить по среднему доплеровскому сдвигу частоты. Когерентный радиолокатор может измерять не только интенсивность рассеянного волнового поля, но и его фазу (а значит, и частоту). Брэгговские волны данной длины могут быть как свободными, так и связанными и, соответственно, распространяться с различными скоростями, вызывая различные доплеровские сдвиги частоты радиолокационного сигнала. В [13] было впервые показано, что наблюдаемый в лабораторном эксперименте доплеровский сдвиг радиолокационных сигналов *X-Ka*-радиолокаторов определяется рассеянием на комбинации свободных и связанных волновых компонент. Этот же вывод иллюстрируется на рис. 7, взятом из работы [14], где в спектре рассеянного радиолокационного сигнала видны два пика, отвечающие доплеровским сдвигам частоты из-за рассеяния на свободных и связанных ветровых волнах.

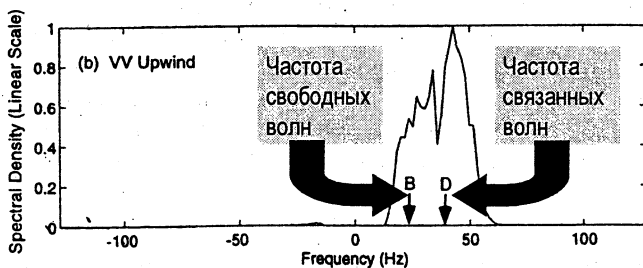


Рис. 7. Пример типичного частотного спектра, измеренного в поле ветровых волн [14].

Рассмотрим подробнее нелинейные механизмы генерации таких волн.

3.1. Паразитные капиллярные волны

Первая из разновидностей вынужденных волн — паразитные капиллярные волны. Они исследовались многими авторами [15—20], в том числе ИПФ РАН [17—19]. Явление состоит в том, что вблизи гребней крутых длинных волн появляются короткие волны (рис. 8).

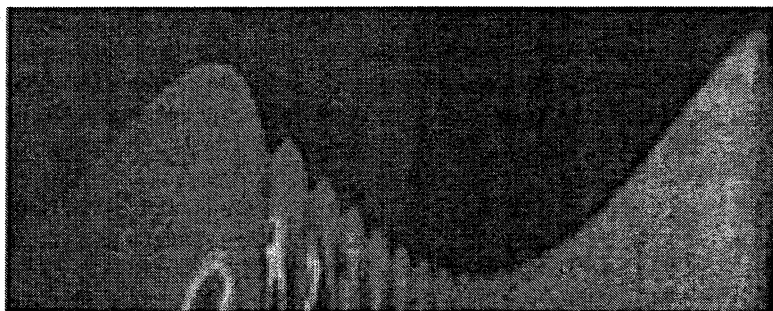


Рис. 8. Пример паразитных капиллярных волн, наблюдавшихся с помощью современных оптических методов [20].

Следуя [16], обсудим кратко физический механизм генерации паразитных капиллярных волн вблизи гребней крутых гравитационных волн. Последние имеют заостренные гребни и пологие впадины. У предельной гравитационной волны угол при вершине составляет 120° , крутизна $AK_{\max} = 0,4432$, что соответствует $h/\lambda \approx 0,07$, а скорость C мало отличается от фазовой скорости линейной волны с тем же пространственным периодом C_0 : $C \approx 1,092C_0$ [20]. В системе отсчета, имеющей скорость C , поле скорости на поверхности воды $U(x)$ стационарно, при этом скорость вблизи гребня волны мала и равна 0 у предельной волны.

Кривизна поверхности воды около гребня имеет резкий пик, при этом капиллярное давление, пропорциональное кривизне, также имеет резкий пик, который является источником для капиллярных волн. Как показано в [20], радиус кривизны поверхности воды вблизи вершины гравитационной волны с крутизной, близкой к предельной, —

$$r_{\min} \approx \frac{24}{K} ((AK)_{\max} - AK). \quad (6)$$

В системе отсчета, бегущей со скоростью, равной фазовой скорости крутой гравитационной волны C , у вершины этой волны имеется покоящийся источник давления, который «обтекается» неоднородным потоком, определяющимся полем орбитальной скорости крутой гравитационной волны (рис. 9, а). Характерный масштаб неоднородности поля скорости вблизи вершины примерно равен r_{\min} (формула (6)). Такой источник будет излучать волны, волновые числа которых определяются из условия, что фазовая скорость этих волн равна скорости потока, в котором находится источник.

Хорошо известно, что зависимость фазовой скорости поверхностных волн c от волнового числа k

$$c = \sqrt{g/k + Tk}$$

(где g — ускорение силы тяжести, T — коэффициент поверхностного натяжения) имеет минимум при $k_* = \sqrt{g/T}$ (длина волны 1,7 см): $c_* = \sqrt[4]{4gT}$ (см. рис. 9, б). Если скорость течения на поверхности воды вблизи гребня волны больше c_* , то будет излучаться два типа волн: гравитационные и капиллярные. При этом вниз по потоку (т. е. влево) будут излучаться более длинные волны, которые имеют групповую скорость меньше фазовой, а против потока (вправо) — более короткие волны, которые имеют групповую скорость больше фазовой. Отсюда ясно, что интересующие нас капиллярные волны возникают на переднем склоне крутой гравитационной волны.

При увеличении крутизны волн скорость вблизи их гребня уменьшается и может оказаться меньше c_* . Тогда излучение волн начнется не от вершины, а от уровня, где скорость потока равна скорости волн. При этом излучаться будут волны с волновым числом k_* ; вправо будут излучаться короткие волны, а влево — длинные. Поскольку волны распространяются на неоднородном потоке, то их длина будет меняться. Вариации волнового вектора можно найти, рассматривая кинематику волн. Примем во внимание, что частота волн, излучаемых в потоке стационарным источником, равна 0, найдем формулу для $k(x)$:

$$U^2(x)k(x) = g + Tk^2(x). \quad (7)$$

При удалении от вершины гравитационно-капиллярной волны скорость $U(x)$ по модулю возрастает. Тогда из формулы (7) и из дисперсионной кривой, изображенной на рис. 9, б, легко видеть, что длина капиллярных волн, распространяющихся вправо, убывает, а длина гравитационных волн, распространяющихся влево, возрастает.

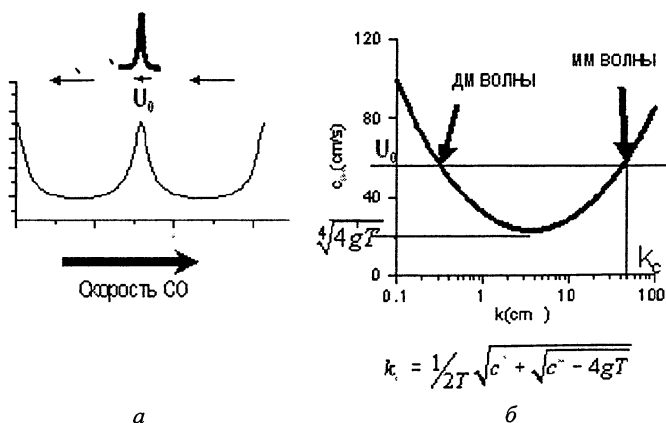


Рис. 9. Профиль в капиллярной волне (а) и дисперсионная кривая для капиллярно-гравитационных поверхностных волн (б).

Из изложенного выше ясно, что для задач дистанционного зондирования взволнованной поверхности воды важен вопрос, каковы будут длина и скорость «паразитных» капиллярных волн. Очевидно, что поскольку это стационарные волны, излучаемые источником, движущимся с фазовой скоростью крутой гравитационной волны (т. е. вершины этой волны), то их фазовая скорость равна этой скорости. При этом при удалении от гребня волны длина капиллярных волн уменьшается от величины 1,7 см, соответствующей минимуму фазовой скорости, а длина гравитационных волн увеличивается. Таким образом, излучаемые вблизи вершины крутой гравитационной волны и распространяющиеся по ее профилю паразитные волны характеризуются целым спектром длин, среди которых может присутствовать и брэгговская длина волны. Однако следует заметить, что, поскольку скорость на поверхности очень быстро возрастает от малого значения вблизи гребня до скорости, слабо отличающейся от фазовой скорости основной волны, на большей части поверхности воды будут присутствовать «паразитные» капиллярные волны с длинами, определяемыми из условия, что фазовая скорость этих волн равна скорости крутой гравитационной волны: $C \approx 1,092 C_0$ [21].

Численные оценки для параметров чистой воды показывают, что гравитационные волны дециметрового диапазона в основном излучают капиллярные волны миллиметрового диапазона (см. таблицу).

Соотношение длин капиллярных и гравитационных волн

Гравитационные волны (длина)	Капиллярные волны (длина в мм)
6 см	5 мм
10 см	3 мм
15 см	2 мм

Амплитуда паразитных капиллярных волн сильно зависит от крутизны несущих волн. Это связано с тем, что источником таких волн является капиллярное давление вблизи гребня, а минимальный радиус кривизны сильно зависит от крутизны волн (см. формулу (6)). При аппроксимации капиллярного давления лоренцевой кривой

$$P_{capillary} = \frac{Tr_{min}}{r_{min}^2 + (3ks)^2}$$

можно получить, что амплитуда излучаемых таким источником волн пропорциональна экспоненте от r_{min} , которая с учетом (6) имеет вид

$$\xi_{capillary} \sim \exp\left(\frac{8}{K}(AK - (AK)_{max})k_{capillary}(U_0)\right).$$

Отсюда видно, что амплитуда капиллярных волн экспоненциально зависит от крутизны несущих волн. Этот факт мы используем потом при интерпретации модуляции поверхностных волн неоднородными течениями.

3.2. Супергармоническая неустойчивость

Рассмотрим еще один процесс, который может приводить к генерации коротких вынужденных волн длинными — обрушение крутых поверхностных волн вблизи гребня. Этот процесс был выделен Лонге — Хиггинсом [22—25] при исследовании устойчивости поверхностных гравитационных волн с крутизной, близкой к предельной. Анализ устойчивости в [22] показал, что при превышении крутизной волны критического значения 0,4292 (близкого к предельному, но ниже его) возникает неустойчивая мода с инкрементом нарастания $0,123 \sqrt{g/r_{\min}}$ и характерным масштабом $0,45r_{\min}$. Эта неустойчивость была названа супергармонической в противоположность субгармонической или модуляционной неустойчивости. Физический механизм этой неустойчивости, по-видимому, аналогичен механизму неустойчивости Кельвина — Гельмгольца в сдвиговом потоке, при котором начальное возмущение нарушает баланс давления, а возникшее возмущение давления приводит к росту первоначального возмущения. Все теоретические результаты изучения супергармонической неустойчивости получены численно. На рис. 10 изображена форма крутых гравитационных волн, на которых развивается супергармоническая неустойчивость (рисунки взяты из обзора [26]). Видно, что в случае длинных волн, когда на развитие супергармонической неустойчивости не оказывают воздействие капиллярные силы, развитие приводит к образованию струи, которая ударяет о поверхность воды. Капиллярные силы «сглаживают» струю, в результате вблизи вершины формируется «утолщение».

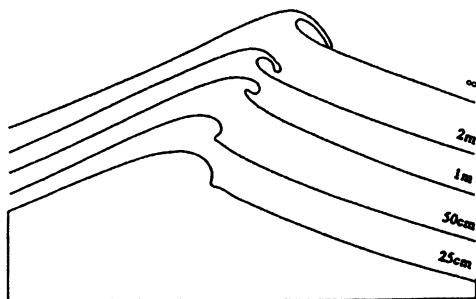


Рис. 10. Форма крутых гравитационных волн, на которых развивается супергармоническая неустойчивость [26]. У кривых указаны длины крутых гравитационных волн.

Если формируется струя, то развитие супергармонической неустойчивости приводит к генерации турбулентности с захватом пузырьков воздуха и образованием капель. В случае «утолщения» эти явления не наблюдаются, этот случай получил в англоязычной литературе название *gentle spilling breaker* («слабое обрушение»). Однако при таком «слабом обрушении» довольно быстро (примерно за 0,1 периода волны) вблизи гребня в лабораторных экспериментах [26, 27] наблюдалось образование области, покрытой случайными поверхностными волнами (волновой турбулентностью). Эксперименты [26, 27] показывают, что: 1) этот процесс сильно пороговый (существует минимальная крутизна несущей волны, необходимая для возникновения неустойчивости, на что указывает и теория [22—25]); 2) возбуждается широкий спектр коротких поверхностных волн, среди гармоник которого присутствуют и брэгговские компоненты сантиметрового диапазона.

Итак, согласно современным представлениям, связанные волны миллиметрового диапазона представляют собой паразитные капиллярные волны, а механизм генерации связанных волн сантиметрового диапазона связан главным образом с обрушениями, в том числе и со «слабыми обрушениями». Отметим, что, как показали натурные наблюдения [28], именно «слабые обрушения» являются основным механизмом генерации связанных волн в открытом море практически при всех условиях, за исключением ураганного ветра.

3.3. Наблюдения связанных волн в лабораторных экспериментах

В натуральных условиях, когда происходит генерация волн ветром, на поверхности воды присутствуют как связанные, так и свободные короткие поверхностные волны. Отделить связанные волны и изучить их влияние на рассеивающие свойства поверхности воды можно в лабораторном эксперименте. В таких экспериментах крутые гравитационные волны создаются с помощью механического волнопродуктора, совершающего колебания с заданной частотой, соответствующей дециметровому или метровому диапазону длин волн, поэтому свободные волны сантиметрового диапазона исключаются. Серия лабораторных экспериментов, проведенных совместно ИПФ и Университетом Гамбурга [29], продемонстрировала, что с ростом амплитуды длинной волны при переходе через некоторый порог резко изменялся высокочастотный спектр поверхностных волн (рис. 11, *а*). При этом так же резко возрастала рассеянная мощность радиоволн: в частности, при увеличении амплитуды дециметровой волны от 0,15 см до 0,2 см рассеянный сигнал радиолокатора *X*- и *Ka*-диапазонов возрастал на 20 дБ (в 100 раз) (рис. 11, *б*).

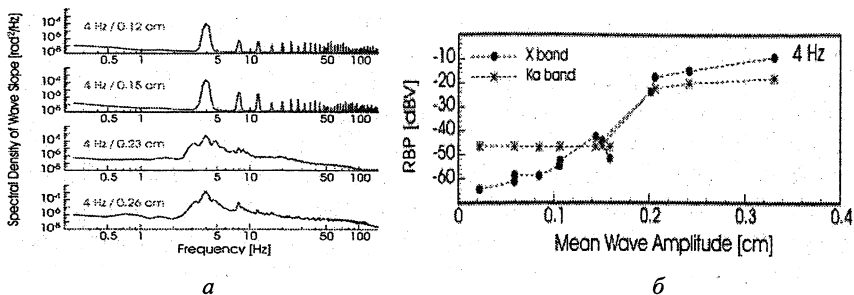


Рис. 11. Изменение спектра поверхностных волн (а) и сечения рассеяния поверхности воды (б) в X- и Ka-диапазонах при увеличении амплитуды дециметровой волны, генерируемой волнопродуктором [13].

Доплеровские спектры рассеянных радиоволн показали, что брэгговские волны в экспериментах [13] были связанными. Брэгговские волны имели длины 4 мм и 1,5 см, а их эффективная фазовая скорость была близка к скорости основной дециметровой волны.

Итак, основной особенностью связанных волн является пороговый резкий характер зависимости их амплитуды от амплитуды несущих поверхностных волн. Это означает, что небольшая модуляция несущих волн будет приводить к существенной модуляции связанных волн (см. рис. 12). Данный эффект был продемонстрирован в лабораторных экспериментах ИПФ РАН [29, 30].

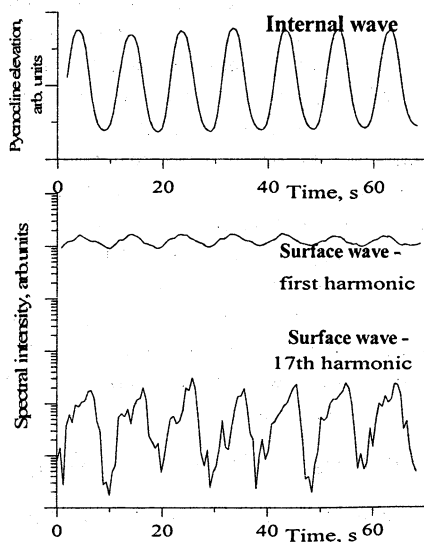


Рис. 12. Смещение изопокны во внутренней волне, модуляция интенсивности основной волны с частотой 4 Гц и ее 17-й гармоники (5-миллиметровой ряби).

Эксперимент состоял в следующем. В бассейне генерировались внутренние волны с периодом 15 с и поверхностные волны с частотой 4 Гц. Групповой синхронизм отсутствовал, поэтому модуляция поверхностных волн была слабая. При этом коэффициент модуляции высших гармоник был больше примерно на порядок. Видно, что при превышении некоторого порога коэффициент модуляции связанных волн существенно выше, чем у основной волны. Итак, в лабораторных условиях для механически генерируемых поверхностных волн показано, что мощность рассеянных радиоволн в основном определяется связанными волнами. При чем модуляция этих волн усиливается за счет порогового характера их генерации. Что будет в случае генерации волн ветром?

Исследования в больших ветро-волновых бассейнах (Франция, США) показывают, что в этом случае важна модуляция как связанных, так и свободных волн. Более того, модель вероятностной структуры уклонов морской поверхности, предложенная Плантом [28], утверждает, что основной сигнал, рассеянный поверхностью, формируется за счет свободных волн, поскольку связанные волны генерируются на очень маленькой площади поверхности (вблизи гребней крутых волн). Возможно, именно по этой причине модель, основанная на решении уравнения Хассельмана для свободных волн, часто хорошо согласуется с экспериментами.

3.4. Модуляция нелинейного сдвига частоты

Обсудим еще один эффект, который не учитывается в уравнении Хассельмана — модуляцию нелинейного сдвига частоты. Этот эффект для спектров слабой турбулентности общего вида рассматривался в работе [31]. Следует заметить, что слагаемые, описывающие модуляцию нелинейного сдвига частоты, входят в кинетическое уравнение в более низком порядке по амплитуде гармоник, чем интеграл столкновений, описывающий энергообмен между спектральными компонентами волнового поля. Однако они не входят в уравнение Хассельмана, поскольку в нем интеграл столкновений получен в предположении статистически пространственно однородного поля поверхностных волн. Для того чтобы учесть модуляцию нелинейного сдвига частоты, надо учесть пространственную неоднородность поля поверхностных волн. Это можно сделать в рамках кинетического уравнения для корреляционной функции двух гармоник спектра поверхностных волн:

$$\zeta(\bar{k}, \bar{k}' + \bar{Q}, t) = \left\langle a^* (\bar{k}, t) a(\bar{k} + \bar{Q}, t) \right\rangle.$$

Если считать, что энергонесущий масштаб спектра поверхностных волн мал по сравнению с масштабом неоднородного течения, то можно определить локальную спектральную плотность волнового действия N :

$$N(\bar{k}, \bar{r}, t) = \int_{q < k_*} \zeta(\bar{k}, \bar{k}' + \bar{Q}, t) \exp(i\bar{r}\bar{Q}) d\bar{Q}.$$

Раскладывая кинетическое уравнение в ряд по малому \bar{Q} , получим кинетическое уравнение для N . В простейшем одномерном случае оно имеет вид

$$\begin{aligned} \frac{\partial N}{\partial t} + \frac{\partial N}{\partial x} \frac{\partial \omega}{\partial k} - \frac{\partial N}{\partial k} \frac{\partial \omega}{\partial x} = -2 \frac{\partial N}{\partial k} \int T(k, k_1, k, k_1) \frac{\partial N(k_1)}{\partial x} dk_1 + \\ + 2 \frac{\partial N}{\partial x} \int 2 \frac{\partial T(k, k_1, k, k_1)}{\partial k} N(k_1) dk_1. \end{aligned}$$

Коэффициент взаимодействия $T(k, k_1, k, k_1)$ — это сложное выражение, но для него можно получить простую приближенную формулу

$$T(k, k_1, k, k_1) = \begin{cases} \frac{1}{30} k k_1^2, & k_1 < k; \\ \frac{1}{30} k^3 + 0,055 (k_1 - 0,8k)^{3/2} k^{3/2}, & k_1 > k. \end{cases}$$

Учет нелинейного сдвига частоты может приводить к эффекту модуляционной неустойчивости в поле поверхностных волн. Можно предположить, что учет этого эффекта позволит разрешить типичную для современных моделей проблему недооценки модуляции коротких поверхностных волн в поле неоднородных течений, которая отмечалась, например, в работе [32]. Модуляционная неустойчивость может привести к существенному усилению модуляции поверхностных волн в поле неоднородных течений по сравнению с моделью, основанной на использовании уравнения Хассельмана.

Выводы

Итак, основным элементом модели формирования изображения на радиолокационных панорамах является модель модуляции брэгговских поверхностных волн в поле неоднородных течений. Эта проблема обсуждается не менее 30 лет и все еще далека от завершения. Причиной этого является очень сложная нелинейная динамика поля коротких поверхностных волн, которые и определяют рассеивающие свойства морской поверхности. В настоящее время выделяют два типа коротких брэгговских волн — свободные и связанные. Свободные волны имеют дисперсионные соотношения, близкие к дисперсионным соотношениям линейных волн, и порождаются за счет генерации ветром или за счет слабонелинейного резонансного взаимодействия гармоник. Связанные волны представляют собой гармоники сильнонелинейных поверхностных волн и имеют фазо-

вые скорости, совпадающие с фазовыми скоростями этих сильнонелинейных волн. Очевидно, связанные волны генерируются и слабонелинейными волнами за счет геометрической нелинейности на поверхности воды. Однако эффективность их генерации резко возрастает при увеличении амплитуды основной волны за счет неустойчивости, которая возникает при превышении амплитудой волны некоторого критического значения. В настоящее время известно два таких механизма неустойчивости, это генерация «паразитной» капиллярной ряби и супергармоническая неустойчивость. Источником капиллярных волн является избыточное капиллярное давление, возникающее вблизи вершины крутой волны. При этом генерация паразитной капиллярной ряби представляет собой черенковское излучение этого источника, находящегося в поле течений длинной волны. В случае волн на воде резонансные условия черенковского излучения выполняются только для капиллярной ряби миллиметрового диапазона. Супергармоническая неустойчивость — это гидродинамическая неустойчивость поверхности воды в крутой гравитационной волне. На нелинейной стадии развития этой неустойчивости генерируется широкий спектр коротких волн, который содержит и брэгговские волны.

Модуляция свободных волн в присутствии неоднородных течений определяется двумя механизмами — кинематическим механизмом и эффектом модуляции инкремента. Этими эффектами очень часто можно объяснить наблюдаемые в природных условиях вариации интенсивности коротких поверхностных волн. Однако в ряде случаев связанные волны учитывать необходимо, причем их вариации в поле неоднородных течений, как правило, значительны. Это обусловлено пороговым характером их генерации по отношению к амплитуде гравитационных волн дециметрового и метрового диапазонов.

В заключение следует заметить, что хотя в настоящее время намечен значительный прогресс в понимании механизмов модуляции коротких поверхностных волн, построение количественной физической модели далеко до завершения.

Работа подготовлена при финансовой поддержке РФФИ (проекты 07-05-00565, 05-05-64137).

Литература

1. Океанология. Физика океана: В 2 т. / Под ред. А.С. Мониной. М.: Наука, 1978.
2. Озмидов П. В. О распределении энергии по разномасштабным движениям в океане // Изв. АН СССР. ФАО. 1965. Т. 1, № 4. С. 439—448.
3. *Werner Alpers, Leonid Mitnik, Lim Hock, Kun Shan Chen.* The tropical and subtropical ocean viewed by ERS SAR <http://www.ifm.uni-hamburg.de>
4. *Holliday D., Gaetan St-Cyr, Nancy C. Woods,* A radar ocean imaging model for small to moderate incidence angles // *Int. J. Remote Sensing.* 1986. V. 7, № 12. P. 1809—1834.

5. *Hasselmann K.* On the non-linear energy transfer in a gravity-wave spectrum. Part 1. General theory // *J. Fluid Mech.* V. 12. 1962. P. 481—500.
6. *Badulin S. I., Pushkarev A. N., Resio D., Zakharov V. E.* Self-similarity of wind-driven seas Nonlinear Processes in Geophysics, 12, 891-945, 2005; SRef-ID: 1607-7946/npg/2005-12-891.
7. *Басович А. Я., Таланов В. И.* О трансформации коротких поверхностных волн на неоднородных течениях // *Изв. АН СССР. ФАО.* 1977. Т. 13, № 7. С. 766—773.
8. *Басович А. Я., Баханов В. В., Таланов В. И.* Влияние интенсивных внутренних волн на ветровое волнение (кинематическая модель) // Воздействие крупномасштабных внутренних волн на морскую поверхность. Горький: ИПФ АН СССР, 1982. С. 8—30.
9. *Hasselmann K., Raney R. K., Plant W. J., Alpers W., Shuchman R. A., Lyzenga D. R., Rufenach C. L., Tucker M. J.* Theory of synthetic aperture radar imaging: a MARSEN view // *J. Geophys. Res.* 1985. V. 90. P. 4659—4686.
10. *Plant W. J., Keller W. C., Cross A.* Parametric dependence of ocean wave-radar modulation transfer functions // *J. Geophys. Res.* 1983. V. 88. P. 9747—9756.
11. *Stanton T. P., Ostrovsky L. A.* Observation of highly nonlinear internal solitons over the continental shelf // *Geophys. Res. Lett.* 1998. V. 25. P. 2695.
12. *Горшков К. А., Долина И. С., Соустова И. А., Троицкая Ю. И.* Модуляция коротких ветровых волн в присутствии интенсивных внутренних волн. Эффект модуляции инкремента // *Изв. РАН. ФАО.* 2003. Т. 39, № 15, С. 661—672.
13. *Gade M., Alpers W., Ermakov S. A., Hühnerfuss H., Lange P. A.* Wind-wave tank measurements of bound and freely propagating short gravity-capillary waves // *J. Geophys. Res.* 1998. V. 103. P. 21697—21710.
14. *Plant W. J., Keller W. C., Hesany V., Hara T., Bock E., Donelan M.* Bound waves and Bragg scattering in a wind wavetank // *J. Geophys. Res.* 1999. V. 104(C2). P. 3243—3263.
15. *Longuet-Higgins M. S.* The generation of capillary waves by steep gravity waves // *J. Fluid Mech.* 1963. V. 16. P. 238—159.
16. *Longuet-Higgins M. S.* Parasitic capillary waves: a direct calculation // *J. Fluid Mech.* 1995. V. 301. P. 79—107.
17. *Рувинский К. Д., Фрейдман Г. И.* О генерации капиллярно-гравитационных волн крупными гравитационными волнами // *Изв. АН СССР. ФАО.* 1981. Т. 17. С. 548—553.
18. *Ruvinky K. D., Feldstein F. I., Freidman G. I.* Numerical simulations of the quasi-stationary stage of ripple excitation by steep gravity-capillary waves // *J. Fluid Mech.* 1991. V. 230. P. 339—353.
19. *Ермаков С. А., Рувинский К. Д., Салашин С. Г.* О локальной связи характеристик ряби на гребнях гравитационно-капиллярных волн с их кривизной // *Изв. АН СССР. ФАО.* 1988. Т. 24, № 7. С. 771—773.
20. *Jiang L., Lin H.-J., Schultz W. W., Perlin M.* Unsteady ripple generation on steep gravity-capillary waves // *J. Fluid Mech.* 1999. V. 386. P. 281—304.
21. *Longuet-Higgins M. S., Fox M. J. H.* Theory of the almost-highest wave. Part 2. Matching and analytic extension // *J. Fluid Mech.* 1978. V. 259. P. 333—344.
22. *Longuet-Higgins M. S., Cleaver R. P.* Crest instabilities of gravity waves. Part 1. The almost-highest wave // *J. Fluid Mech.* 1994. V. 158. P. 115—129.
23. *Longuet-Higgins M. S., Cleaver R. P., Fox M. J. H.* Crest instabilities of gravity waves. Part 2. Matching and asymptotic analysis // *J. Fluid Mech.* 1994. V. 259. P. 333—344.
24. *Longuet-Higgins M. S., Dommermuth D. G.* Crest instabilities of gravity waves. Part 3. Nonlinear development and breaking // *J. Fluid Mech.* 1997. V. 336. P. 33—50.
25. *Longuet-Higgins M. S., Tanaka M.* On the crest instabilities of steep surface waves // *J. Fluid Mech.* 1997. V. 336. P. 51—68.
26. *Duncan J. H.* Spilling breakers // *Annu. Rev. Fluid Mech.* 2001. V. 33. P. 547.

27. *Duncan J. H., Qiao H., Philomin V., Wenz A.* Gentle spilling breakers: crest profile evolution // *J. Fluid Mech.* 1999. V. 379. P. 191—222.
28. *Plant W. J.* A new interpretation of sea-surface slope probability density functions // *J. Geophys. Res.* 2003. V. 108, № C9. P. 3295; doi:10.1029/2003JC001870.
29. *Ермаков С. А., Салашин С. Г.* Об эффекте сильной модуляции капиллярно-гравитационной ряби внутренними волнами // Докл. РАН. 1994. Т. 337, № 1. С. 108—111.
30. *Ermakov S. A., Sergievskaya I. A., Shchegolkov Yu. B., Goldblat V. Yu.* Wave tank study of «cascade» modulation of bound capillary-gravity waves due to internal waves // *Proc. IGARSS IEEE.* Piscataway, N.J., USA, 2000. V. 3. P. 1087—1089.
31. *Львов В. С., Рубенчик А. М.* Пространственно неоднородные сингулярные спектры слабой турбулентности // *ЖЭТФ.* 1977. Т. 78, № 1. С. 127—140.
32. *Romeiser R., Alpers W.* An improved composite surface model for the radar backscattering cross section of the ocean surface. 2. Model response to surface roughness variations and the radar imaging of underwater bottom topography // *J. Geophys. Res.* 1997. V. 102. P. 25251—25267.

Математические проблемы нелинейной динамики

РЕНОРМГРУППОВЫЕ СИММЕТРИИ В КРАЕВЫХ ЗАДАЧАХ

В. Ф. Ковалев, Д. В. Ширков

1. РГ в КТП: возникновение и эволюция

Понятие ренормгрупповой симметрии (РГС) возникло в математической физике в начале 90-х годов XX века [1, 2]¹ как результат объединения понятия *группы Ли преобразований* [6] и понятия *ренормализационной группы* (РГ) (сокращенно — ренормгруппы), возникшего в квантовой теории поля в начале 50-х гг. минувшего столетия в работах Штюкельберга и Петермана. В основу схемы построения РГС были заложены идеи метода ренормализационной группы (МРГ). Этот метод, использующий инфинитезимальный аппарат Софуса Ли для групп непрерывных преобразований, представляет собой алгоритм регулярного уточнения поведения приближенных решений квантовой теории поля в окрестности сингулярности. Он был создан в середине 50-х гг. в работах Боголюбова.

Позже, когда понятия РГ и МРГ были с успехом экспортированы в другие разделы теоретической физики, была установлена близость симметрии РГ-преобразований с симметрией самоподобия (иначе *автоподобности*) и было предложено рассматривать ее (РГ-симметрию) как функциональное обобщение степенного самоподобия, т. е. как функциональное самоподобие.

Простую иллюстрацию использования свойств симметрии решений для уточнения поведения приближений дает задача частичного суммирования двойного ряда, возникающего в разложениях квантовополевой теории возмущений

$$\bar{\alpha}(x, \alpha) = \sum a_{\nu, \mu} \alpha^\mu (\ln x)^\nu, \quad 0 \leq \nu \leq \mu - 1 < \infty \quad (1)$$

и отвечающего безмассовой модели квантовой теории поля с одной константой связи α . Здесь $x = Q^2/\mu^2$ представляет собой отношение квадрата 4-импульса Q к квадрату «нормировочного» импульса μ , а функция $\bar{\alpha}$ именуется эффективной связью (или инвариантным зарядом). Разложение (1) содержит произведение малой величины α на

¹См. также обзоры [3, 4, 5], в которых можно найти дальнейшие ссылки.

параметр $\ln x$, которое оказывается немалым в ультрафиолетовой области $\ln x \gg 1$. Задача эффективного частичного суммирования, превращающего расходящийся ряд (1) в ряд вида

$$\bar{\alpha}(x, \alpha) = \sum_n \alpha^n f_n(\alpha \ln x), \quad (2)$$

где функции f_n растут медленнее степеней логарифма, решается с помощью МРГ следующим образом. Из конечных перенормировочных соотношений Дайсона следует, что должным образом определенная функция $\bar{\alpha}$ инвариантна относительно одновременного одно-параметрического преобразования R_t величин x и α ,

$$R_t : \{x \rightarrow x' = x/t, \quad \alpha \rightarrow \alpha' = \bar{\alpha}(t, \alpha)\}, \quad (3)$$

где первое представляет растяжение (сдвиг) «координаты» x , а второе — более сложное функциональное преобразование характеристики решения. Условие выполнения группового закона композиции преобразований $R_{\tau t} = R_\tau R_t$ приводит к функциональному уравнению для $\bar{\alpha}$

$$\bar{\alpha}(x, \alpha) = \bar{\alpha}(x/t, \bar{\alpha}(t, \alpha)), \quad \bar{\alpha}(1, \alpha) = \alpha, \quad (4)$$

которое допускает также дифференциальную формулировку:

$$X\bar{\alpha} = 0; \quad X = x\partial_x - \beta(\alpha)\partial_\alpha, \quad \text{где } \beta(\alpha) = \partial_\xi \bar{\alpha}(\xi; \alpha) \Big|_{\xi=1}. \quad (5)$$

Здесь X — инфинитезимальный оператор группы Ли преобразований (3), а координата $\beta(\alpha)$ оператора именуется также генератором группы. Согласно определению (5), для вычисления β следует найти коэффициенты $\beta_\nu = a_{1, \nu+2}$ при линейных по $\ln x$ вкладах в возрастающих порядках теории возмущений (1) и отвечающих вкладам неприводимых $(\nu + 1)$ -петлевых диаграмм. При известной функции $\beta(\alpha)$ «эволюция» инвариантной функции связи $\bar{\alpha}$ определяется решением уравнения Ли (уравнения характеристики для (5)):

$$\partial_{\ln x} \bar{\alpha} = \beta(\bar{\alpha}), \quad \bar{\alpha}(1, \alpha) = \alpha. \quad (6)$$

В квантовой электродинамике известны несколько первых членов разложения бета-функции. Ограничиваясь для примера вкладом одно-петлевых диаграмм $\beta(\alpha) = \alpha^2/3\pi$, получаем после подстановки в (6) и интегрирования:

$$\bar{\alpha}_1(x, \alpha) = \frac{\alpha}{1 - \frac{\alpha}{3\pi} \ln x}. \quad (7)$$

Полученный результат отвечает первому вкладу в разложении (2) и по существу решает поставленную задачу улучшения теории возмущений в рассмотренном однопетлевом приближении для $\beta(\alpha)$. Он интересен тем, что точно удовлетворяет функциональному уравнению (4), а при разложении в ряд по α дает бесконечную последовательность слагаемых вида $\alpha(\alpha \ln x)^n$, т. е. старших логарифмических вкладов из каждого порядка теории возмущений.

2. РГ в математической физике: иллюстрация

При переносе РГ-представлений в задачи математической физики [1, 2] в конечном счете ставились те же цели, что и в квантовой теории поля — «исправление» поведения решений в окрестности сингулярности. Однако способ их реализации в математической физике оказался иным. Это связано с тем, что при использовании МРГ в его изначальной формулировке [7] обнаружение группового свойства решения (3) в каждом конкретном случае требует специального, зачастую нестандартного анализа (см., например, обсуждение в [9, 10]), что является алгоритмическим недостатком техники РГ. С целью преодоления этого недостатка для краевых задач математической физики нами совместно с В. В. Пустоваловым был развит (см. [1, 2], а также недавние обзоры [4, с. 232–249], [5]) совершенно иной, отличный от известных ранее, РГ-алгоритм, который объединил РГ-идеологию из квантовой теории поля с регулярным способом построения симметрий решений краевых задач. Именно благодаря этому алгоритму и возникло понятие РГС для решений краевых задач: эти симметрии находятся в результате вычислительной процедуры, подобной используемой в современном групповом анализе, и при этом имеют то же назначение, что и в методе РГ Боголюбова, т. е. получение улучшенного (по сравнению с исходным приближенным) решения.

Проиллюстрируем характерные особенности этого алгоритма на примере построения РГС для решения краевой задачи для уравнения Хопфа:

$$v_z + vv_x = 0, \quad v(0, x) = \epsilon U(x), \quad (8)$$

где U — обратимая функция x . Вводя $u = \epsilon v$, мы вносим амплитуду граничного значения непосредственно в исходное уравнение:

$$u_z + \epsilon uu_x = 0, \quad u(0, x) = U(x). \quad (9)$$

Для малых значений $\epsilon z \ll 1$, т. е. вблизи границы ($z \rightarrow 0$) или для малых амплитуд на границе ($\epsilon \rightarrow 0$) решение задачи (9) по теории возмущений представляется отрезком степенного ряда по ϵz :

$$u = U - \epsilon z U U_x + O((\epsilon z)^2). \quad (10)$$

Очевидно, что это решение не годится при большом удалении от границы, когда $\epsilon z U_x \simeq 1$. Использование РГС позволяет улучшить результат теории возмущений и восстановить правильное поведение решения в окрестности сингулярности (при условии, что такая сингулярность в решении появляется при некоторых конечных значениях z).

Для нахождения этих симметрий *расширим набор переменных*, вовлекаемых в групповые преобразования, путем добавления параметра ϵ к списку независимых переменных. Далее, используя классический вычислительный алгоритм Ли² (см., например, [11]), найдем допускаемую группу симметрии G с генератором

$$X = \xi^z \partial_z + \xi^x \partial_x + \xi^\epsilon \partial_\epsilon + \eta \partial_u, \quad (11)$$

координаты $\{\xi, \eta\}$ которого имеют вид

$$\xi^z = \psi^1, \quad \xi^x = \epsilon u \psi^1 + \psi^2 + x(\psi^3 + \psi^4), \quad \xi^\epsilon = \epsilon \psi^4, \quad \eta = u \psi^3, \quad (12)$$

где ψ^i , $i = 2, 3, 4$, являются произвольными функциями ϵ , u , и $x - \epsilon u z$, а ψ^1 является произвольной функцией всех групповых переменных. Формулы (12) характеризуют бесконечно-мерную алгебру Ли, заданную четырьмя операторами (если амплитуда ϵ не вовлекается в групповые преобразования, то число операторов уменьшается до трех ([12], Т. 1. С. 178):

$$\begin{aligned} X_1 &= \psi^1 (\partial_z + \epsilon u \partial_x), & X_2 &= \psi^2 \partial_x, \\ X_3 &= \psi^3 (x \partial_x + u \partial_u), & X_4 &= \psi^4 (\epsilon \partial_\epsilon + x \partial_x). \end{aligned} \quad (13)$$

Предположим, что известно частное решение краевой задачи (9), заданное неким соотношением $u - W(z, x, \epsilon) = 0$. Поскольку под действием РГ-преобразования решение краевой задачи преобразуется в решение той же краевой задачи, то должно выполняться соответствующее

²Применение аппарата современного группового анализа в РГ-алгоритме подразумевает использование принятой в групповом анализе соответствующей терминологии и обозначений дифференциальной алгебры (подробнее см. [11, 12]). В частности, полагаются известными понятия инфинитезимального оператора группы, точечных симметрий и симметрий Ли — Беклунда (или высших симметрий), инвариантных решений, инвариантных и частично-инвариантных многообразий и т. д.

условие инвариантности (аналогичное уравнению (5)), вычисленное на этом решении:

$$(W - xW_x)\psi^3 - W_x\psi^2 - (\epsilon W_\epsilon + xW_x)\psi^4 = 0. \quad (14)$$

Это условие справедливо для всех z , в том числе и для $z = 0$, когда W заменяется на $U(x)$. В пределе $z \rightarrow 0$ условие (14) дает связь между функциями ψ^i , $i = 2, 3, 4$ (но ограничений на ψ^1 не возникает!), которую, учитывая характер зависимости ψ^i от аргументов, можно легко продолжить на значения $z \neq 0$,

$$\psi^2 = -\chi(\psi^3 + \psi^4) + (u/U_\chi)\psi^3, \quad \chi = x - \epsilon uz, \quad (15)$$

где производная U_χ должна быть выражена в соответствии с краевыми условиями либо в терминах χ или u . Используя (15) в (12), мы приходим к группе меньшей размерности с операторами

$$\begin{aligned} R_1 &= \psi^1 (\partial_z + \epsilon u \partial_x), \quad R_2 = u \psi^3 [(\epsilon z + 1/U_\chi) \partial_x + \partial_u], \\ R_3 &= \epsilon \psi^4 (z u \partial_x + \partial_\epsilon). \end{aligned} \quad (16)$$

Описанная выше процедура, которая сводит (13) к (16), представляет *сужение группы* (11) на *частном решении*, а набор операторов (16) определяет искомую РГС. Это означает, что решение краевой задачи может быть получено с использованием любого из генераторов (16) путем решения соответствующих уравнений Ли (аналогичных (6)) и вычисления конечных преобразований группы. Без ограничения общности можно взять генератор R_3 с $\epsilon \psi^4 = 1$ и получить конечные РГ-преобразования (a — параметр группы)

$$x' = x + azu, \quad \epsilon' = \epsilon + a, \quad z' = z, \quad u' = u, \quad (17)$$

где z и u являются инвариантами РГ-преобразований, а преобразования ϵ и x представляют преобразования сдвига, которые для переменной x зависят также от z и u . При $\epsilon = 0$ с учетом (10) величины x и u связаны соотношением $x = H(u)$, где $H(u)$ — функция, обратная к $U(x)$. Исключая a, z, u из (17) и опуская штрихи у переменных, мы получим искомое решение краевой задачи (9) в неявной форме:

$$x - \epsilon zu = H(u). \quad (18)$$

По сути, это улучшенное решение по теории возмущений (10), которое пригодно не только для малых $\epsilon z \ll 1$, при условии, конечно, что связь (18) определена однозначным образом. В зависимости от $H(u)$

это решение дает либо правильное поведение решения вблизи конечных значений $z \rightarrow z_{sing}$, либо правильное асимптотическое поведение при $z \rightarrow \infty$. Подчеркнем, что для получения решения краевой задачи (18) использовалось *только* знание симметрии решения и соответствующей теории возмущений.

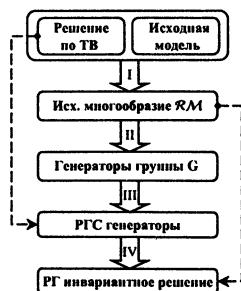
3. Алгоритм построения РГС для краевых задач

Рассмотренный выше пример построения РГС является иллюстрацией общего алгоритма, детальное описание которого применительно к краевым задачам для дифференциальных уравнений можно найти, например, в обзорах в [3, 4], а его обобщение для нелокальных задач — в [5, 13]. Здесь мы только кратко упомянем основные этапы алгоритма и приведем несколько поясняющих замечаний.

В качестве исходной рассматривается система из интегродифференциальных (в том числе дифференциальных и интегральных) уравнений, дополненная соответствующими краевыми (граничными или начальными) условиями. Предполагается также известным некоторое приближенное решение U^α , например представленное отрезком ряда теории возмущений по степеням малого параметра.

Тогда условно схема реализации РГ-алгоритма представима в виде последовательности четырех шагов:

- (I) построения исходного *многообразия* \mathcal{RM} ,
- (II) нахождения допустимой им *группы симметрий* G и
- (III) ее *сужения* на частном решении краевой задачи и нахождения РГС, а также
- (IV) отыскания отвечающего РГС *аналитического решения*.



Для моделей, использующих нелокальные (интегральные) соотношения, исходное многообразие \mathcal{RM} представляется в виде системы из s дифференциальных уравнений k -го порядка и q нелокальных соотношений ($\sigma = 1, \dots, q + s$):

$$F_\sigma(z, u, u_{(1)}, \dots, u_{(r)}, J(u)) = 0, \quad J(u) = \int \mathcal{F}(u(z)) dz, \quad (19)$$

$$u_{(1)} = \{u_i^\alpha\} = D_i(u^\alpha), \quad u_{(2)} = \{u_{ij}^\alpha\} = D_j D_i(u^\alpha), \quad \dots$$

Здесь в z наряду с «традиционными» переменными x включены также параметры задачи p , т. е. $z = \{x, p\}$, а нелокальные переменные вводятся интегральными операциями $J(u)$. Здесь используются принятые в групповом анализе [12] обозначения для независимых переменных $x = \{x^i\}$, $i = 1, \dots, n$ и дифференциальных переменных, $u = \{u^\alpha\}$, $\alpha = 1, \dots, m$, имеющих последовательные производные $u_{(1)}, u_{(2)}, \dots$ и связанных с помощью оператора полного дифференцирования D_i .

Особенностью процедуры построения РГС является многовариантность первого (I) шага алгоритма, назначение которого состоит в том или ином способе вовлечения в групповые преобразования параметров, входящих как в уравнения задачи, так и в краевые условия, от которых зависит частное решение задачи. Выбор конкретной реализации первого шага чаще всего диктуется как видом исходных уравнений и краевых условий к ним, так и видом приближенного решения. Такая многовариантность существует и для приведенного выше примера построения РГС для краевой задачи для уравнения Хопфа [5].

Подобная многовариантность присуща только первому шагу алгоритма и нацелена на охват наиболее широкого круга задач, исследуемых данным методом. Уже следующий шаг схемы, состоящий в вычислении наиболее широкой допускаемой группы симметрии G , выполняется в рамках хорошо разработанных теоретико-групповых методов.

Следующий, второй шаг (II) — вычисление группы симметрии G , допускаемой \mathcal{RM} . Для системы дифференциальных уравнений $F_\sigma = 0$, $\sigma \leq s$ речь идет о локальной группе преобразований в пространстве дифференциальных функций, при которых эта система остается неизменной. Классический алгоритм Ли для нахождения симметрий заключается в построении касательного векторного поля, определяемого оператором

$$X = \xi^i \partial_{z^i} + \eta^\alpha \partial_{u^\alpha} \quad (20)$$

с координатами ξ^i , η^α , которые являются функциями групповых переменных $\{z^i, u^\alpha\}$ и удовлетворяют системе линейных однородных дифференциальных уравнений в частных производных — определяющих уравнений

$$\begin{aligned} X_{(k)} F_\sigma |_{[F_\sigma]} = 0, \quad X_{(k)} = X + \dots + \zeta_{i_1 \dots i_k}^\alpha \partial_{u_{i_1 \dots i_k}^\alpha}, \\ \zeta_{i_1 \dots i_k}^\alpha = D_{i_k} (\zeta_{i_1 \dots i_{k-1}}^\alpha) - u_{j_1 \dots j_k}^\alpha D_{i_k} (\xi^j). \end{aligned} \quad (21)$$

Значок $|_{[F]}$ означает, что результат действия оператора берется на многообразии, задаваемом уравнением $F_\sigma = 0$ и всеми его дифференциальными следствиями.

Решение определяющих уравнений дает набор значений координат ξ^i , η^α , т. е. набор инфинитезимальных операторов (или генераторов группы), которые соответствуют допускаемому векторному полю и образуют алгебру Ли с общим элементом

$$X = \sum_j A^j X_j, \quad \text{где } A^j \text{ — произвольные постоянные.} \quad (22)$$

Отметим, однако, что при попытке использовать РГ-алгоритм в нелокальных задачах, где неприменим классический алгоритм Ли, потребовалась существенная модификация именно второго шага (II) алгоритма [5]. Причиной этого является тот факт, что для многообразия \mathcal{RM} , которое задается системой интегральных уравнений, классический инфинитезимальный алгоритм Ли оказывается неприменимым. Основным препятствием здесь оказывается то, что \mathcal{RM} в этом случае не определено *локально* в пространстве дифференциальных функций, поэтому представление обыкновенного уравнения в виде переопределенной системы уравнений здесь не реализуется. Это препятствие преодолевается, например использованием *канонического* оператора $Y = X - \xi^i D_i$, с помощью которого критерий инвариантности для F_σ относительно допускаемой группы может быть записан в инфинитезимальной форме в виде нелокальных определяющих уравнений

$$Y F_\sigma \Big|_{[F_\sigma]} = 0, \quad \text{где } Y \equiv \int dz \kappa^\alpha(z) \frac{\delta}{\delta u^\alpha(z)}, \quad \kappa^\alpha = \eta^\alpha - \xi^i u_i^\alpha. \quad (23)$$

Поскольку найденная на втором шаге группа G в общем случае является более широкой, чем интересующая нас ренормгруппа, то для получения РГС следует сделать третий шаг (III), заключающийся в *сужении* группы G на многообразии, задаваемом этим частным решением. С математической точки зрения эта процедура состоит в проверке условия инвариантности, близкого по форме к (14), на некотором частном решении краевой задачи $U^\alpha(z)$:

$$\sum_j A^j \kappa_j^\alpha \equiv \sum_j A^j (\eta_j^\alpha - \xi_j^i u_i^\alpha) \Big|_{u^\alpha = U^\alpha(z)} = 0. \quad (24)$$

Алгебраические соотношения (24) «связывают» координаты различных операторов группы X_j , допускаемых \mathcal{RM} , и таким образом «встраивают» краевые условия в РГ-операторы за счет выбора постоянных A_j и/или выбором конкретного вида произвольных функций в координатах ξ^i , η^α . В результате третьего шага и возникает набор операторов

R_i , который определяет группу преобразований, называемую *ренорм-группой*:

$$X \Rightarrow R = \sum_j B^j R_j, \quad R_j = \hat{\xi}_j^i \partial_{x^i} + \hat{\eta}_j^\alpha \partial_{u^\alpha}. \quad (25)$$

Эта группа, имеющая конечную или бесконечную размерность, оказывается в общем случае шире, чем РГ, обсуждаемая в квантовой теории поля, которая обычно дается точечной однопараметрической группой. Кроме того, РГ в математической физике характеризуется большим разнообразием типов РГС, включающих помимо точечных и другие симметрии, например симметрии Ли — Беклунда высокого порядка.

Отметим, что в квантовой теории поля термин «ренормгруппа» используется для обозначения группы преобразований еще до определения конкретной реализации для β -функции (т. е. до проведения операции сужения группы на частном решении). Это обусловлено общим видом РГ-оператора, который одинаков для широкого круга задач квантовой теории поля. В математической физике, ввиду невозможности построения РГС в достаточно общем виде, мы используем термин «ренормгруппа» для обозначения алгебры симметрии, которая получается *после* проведения процедуры сужения группы на решении.

Третий шаг алгоритма завершает процедуру построения РГС, однако для достижения конечной цели РГ-алгоритма — улучшения приближенного решения — необходим еще один, четвертый (IV), заключительный шаг. Этот шаг состоит в использовании операторов РГС для нахождения аналитических выражений для новых, улучшенных (по сравнению с исходными) решений краевой задачи. С математической точки зрения реализация этого шага состоит в использовании условий *ренормгрупповой инвариантности*, задаваемых совместной системой уравнений (19) и условием инвариантности решения краевой задачи под действием РГ-преобразований:

$$\sum_j B^j \left(\hat{\eta}_j^\alpha - \hat{\xi}_j^i u_i^\alpha \right) = 0. \quad (26)$$

В наиболее распространенном случае, когда в роли ренормгруппы выступает *однопараметрическая точечная группа Ли*, условие РГ-инвариантности приводит к *уравнениям в частных производных первого порядка*. Решения соответствующих им характеристических уравнений дают *групповые инварианты* (подобные инвариантным зарядам в квантовой теории поля), через которые выражается искомое решение краевой задачи. В общем случае, в соответствии с известной в современном групповом анализе *теоремой об инвариантном представлении* (см., например, [14, гл. 7]) применительно к решениям краевой

задачи, обладающим свойством функционального самоподобия, РГ-инвариантное решение краевой задачи записывается через функциональные инварианты этой задачи. Это утверждение носит общий характер и сформулировано в [15] в виде Ф-теоремы, являющейся функциональным обобщением известной П-теоремы.

4. РГ-алгоритм для уравнения Хопфа: дополнения

В основе рассмотренного в разделе 2 примера построения РГС для уравнения Хопфа лежал наиболее очевидный вариант: построение РГС из группы точечных симметрий уравнения Хопфа в пространстве групповых переменных, расширенном за счет включения параметра ϵ в число независимых переменных. Такой способ конструирования исходного многообразия \mathcal{RM} не является единственно возможным: как уже было сказано, для первого шага характерна многовариантность. Проиллюстрируем эту особенность РГ-алгоритма на уже использованном в разделе 2 наглядном примере.

Вместо группы точечных преобразований Ли можно также обратиться к симметриям Ли — Беклунда [14] с каноническим оператором $R = \kappa \partial_u$, где κ зависит не только от z , x , ϵ и u , но также и от высших производных от u . Координату κ можно искать в виде степенного ряда по ϵ , а условие инвариантности (14) соответствует обращению в нуль величины κ при $z = 0$. Выбирая то или иное представление для вклада нулевого порядка, мы приходим к выражению для κ либо в виде бесконечного, либо оборванного степенного ряда, например, линейной формы по ϵ ,

$$R = \kappa \partial_u, \quad \kappa = 1 - \frac{u_x}{U_x(u)} - \epsilon z u_x. \quad (27)$$

Этот РГ-оператор (27) эквивалентен генератору точечной группы Ли R_2 в (16) и, следовательно, дает тот же результат.

Другая возможность построения РГС для краевой задачи (9) основана на использовании дополнительной дифференциальной связи, которая совместна как с краевыми условиями, так и с исходными уравнениями. Например, для граничных условий в (9) в виде линейных по своему аргументу функций $U(x) = -x$ можно выбрать дифференциальную связь в виде $u_{xx} = 0$. Это условие совпадает с условием инвариантности исходного уравнения относительно симметрии Ли — Беклунда второго порядка. Вычисление точечной симметрии Ли для объединенной системы, состоящей из этой связи и уравнения Хопфа, дает другой путь вычисления РГС для краевой задачи (9).

Завершая демонстрацию возможностей РГ-алгоритма на примере краевой задачи для уравнения Хопфа, укажем на использование РГС для исследования функционалов от решения. Как уже было показано, решение этой задачи может быть получено применением любого из РГ-операторов (16). Пусть для нас интересно не решение краевой задачи во всем диапазоне изменения переменных и параметров, а некая его характеристика в заданной точке, например значение первой пространственной производной при $x = 0$. Формально эта характеристика может быть задана линейным функционалом от решения вида (19):

$$u_x(z, 0) \equiv u_x^0 = - \int_{-\infty}^{+\infty} dx \delta'(x) u(t, x). \quad (28)$$

Зависимость u_x^0 от переменной z получается продолжением действия РГ-операторов (16) на функционал от решения (28). Опять используем последний из РГ-генераторов (16) в наиболее простом представлении с $\epsilon\psi^4 = 1$. Запишем этот оператор в канонической форме Y и формально продолжим его на нелокальную переменную $u_x^0 = J(u)$:

$$Y + \kappa^J \partial_J \equiv \kappa \partial_u + \kappa^J \partial_J. \quad (29)$$

Интегральное соотношение между κ и κ^J получается применением оператора (29) к уравнению (28), что дает искомую координату κ^J продолженного оператора:

$$\kappa^J = \int \frac{\delta J(u)}{\delta u(z)} \kappa(z) dz = \int \mathcal{F}_u \kappa(z) dz. \quad (30)$$

Рассматривая продолженный оператор в пространстве переменных $\{z, \epsilon, u_x^0\}$, приходим к искомому РГ-оператору для функционала (28). Например, при $U = -x$ этот оператор имеет вид

$$R_4 = \partial_\epsilon - z(u_x^0)^2 \partial_{u_x^0}. \quad (31)$$

Используя инвариант этого генератора $J^0 = \epsilon z - 1/u_x^0 = 1$ с учетом начального условия $u_x^0(z = 0) = -1$, нетрудно восстановить зависимость $u_x^0 = -1/(1 - \epsilon z)$ во всем интервале значений переменной z от начальной точки $z = 0$ до точки, где решение имеет особенность, $z_{sing} = 1/\epsilon$. Подчеркнем, что этот результат получен без построения решения краевой задачи в явном виде, а только при помощи РГС. На первый взгляд, проведенное построение выглядит достаточно громоздко и гораздо проще исходить из очевидного решения (18). Но в более

сложных ситуациях явный вид решения зачастую неизвестен, тогда как РГС удается построить (см., например, [15]).

5. Приложения РГС в математической физике

С практической точки зрения, результаты, посвященные приложениям РГС в математической физике, образуют две большие группы, хотя это деление является в значительной степени условным. К первой группе можно отнести результаты по развитию РГ-алгоритма и демонстрации его возможностей на примерах, имеющих в большей степени методический характер. В частности, уже в первых публикациях [2] было показано, что метод построения РГС и их вид существенным образом зависит от способа формулировки краевой задачи. Например, для модифицированного уравнения Бюргера построение РГС достигается расширением списка групповых переменных за счет параметров, явно введенных в исходное уравнение. Каждый из получаемых при таком построении РГ-операторов восьмимерной алгебры Ли восстанавливает точное решение краевой задачи из приближенного, пертурбативного решения по соответствующему малому параметру. При использовании симметрий Ли — Беклунда (высших симметрий) соответствующая алгебра РГС, как показано на примере линейного параболического уравнения или уравнений геометрической оптики [3], может стать бесконечномерной. На примере этих же математических моделей РГС были построены путем использования информации о краевых данных в форме дополнительной дифференциальной связи. Частным случаем такой связи может служить уравнение погружения, что особенно интересно в приложении к математическим моделям на основе обыкновенных дифференциальных уравнений, для которых задача вычисления симметрии является нетривиальной [2—4].

Ко второй группе можно отнести примеры применения РГ-алгоритма к задачам, для которых использование РГС приводит к новым аналитическим результатам. Здесь, как правило, РГС возникают при сочетании различных подходов, характерных для более простых случаев. Первый пример приложения РГ-алгоритма к решению граничной нелинейной задачи о взаимодействии p -поляризованной электромагнитной волны с неоднородной плазмой [1] является не единственным тому подтверждением. Укажем, например, на РГС краевых задач для широкого круга уравнений, объединенных общим термином «квазичаплыгинские среды», в том числе и для системы уравнений, описывающих в приближении геометрической оптики распространение мощного излучения в нелинейной среде. С помощью этих сим-

метрий найдены новые точные и приближенные решения задачи о самофокусировке плоского (щелевого) лазерного пучка. Было показано, что в данном случае РГС — это формальные симметрии, представленные в виде бесконечных рядов по степеням параметра нелинейности среды. Для краевых условий специального вида эти ряды обрываются и превращаются в конечные суммы. В случае общего положения этого не происходит, и тогда конечная сумма соответствует приближенной РГС для малого параметра нелинейности. В дальнейшем приближенные РГС при наличии сразу двух малых параметров были использованы для нахождения новых решений краевой задачи для системы уравнений, описывающих самофокусировку волнового пучка.

Распространение действия РГ-алгоритма на краевые задачи, использующие интегродифференциальные уравнения, позволило существенно расширить круг анализируемых задач. В частности, в кинетической теории плазмы применением РГС удалось решить ряд задач о динамике плазменных сгустков и кластеров, что отражено в обзорах [5, 15]. Весьма эффективным следствием приложения РГ-алгоритма к нелокальным уравнениям явилась процедура продолжения РГ-оператора на нелокальные переменные с целью сокращенного описания решения на языке его интегральной характеристики — функционала от решения. Существенно при этом, что знание самого решения в явном виде не требуется. Использование этой процедуры оказалось результативным при вычислении функционалов от решений различных задач нелинейной физики.

Более подробное обсуждение вопросов приложения РГС можно найти в обзорах [3–5], [15], [16] и цитированной там литературе.

Обширный список решенных с использованием РГ-алгоритма задач свидетельствует об универсальности метода РГС. Поэтому можно рассчитывать на дальнейшее расширение класса задач, поддающихся исследованию с помощью метода РГС, и на новые объекты, для которых его применение пока не является стандартной процедурой.

Работа выполнена при финансовой поддержке грантов МНТЦ 2289, РФФИ (05-01-00631 и 06-02-16103) и НШ-2339.2003.2.

Литература

1. Ковалев В. Ф., Пустовалов В. В. Функциональная автомодельность в одной из задач теории плазмы с электронной нелинейностью // ТМФ. 1989. Т. 81, № 1(10). С. 69–85.

2. Shirkov D. V. Several topics on renorm-group theory // Renormalization group '91: Proc. 2-nd Int. Conf. / Ed. by D. V. Shirkov, V. B. Priezzhev. WS, Singapore, 1992. P. 1–10;

Kovalev V. F., Krivenko S. V., Pustovalov V. V. The renormalization group method based on group analysis // *ibid.* P. 300—314.

3. *Kovalev V. F., Pustovalov V. V., Shirkov D. V.* Group analysis and renormgroup symmetries // *J. Math. Phys.* 1998. V. 39. P. 1170—1188;

Kovalev V. F., Pustovalov V. V., Shirkov D. V. Preprint hep-th/9706056.

4. *Kovalev V. F., Shirkov D. V.* Bogoliubov Renormalization Group and Symmetry of Solution in Mathematical Physics // *Phys. Rep.* 2001. V. 352(4—6). P. 219; Preprint JINR E2-2000-9. Dubna, 2000; hep-th/0001210.

5. *Kovalev V. F., Shirkov D. V.* Renorm-group symmetry for functionals of boundary value problem solutions // *J. Phys. A.* 2006. V. 39. P. 8061—8073.

6. *Lie S.* Theorie der Transformationsgruppen // *Math. Ann.* 1880. V. 16. P. 441—528; Lie groups: history, frontiers and applications. V. 1: Sophus Lie's 1880 transformation group paper / Ed. by R. Hermann. Brookline, Massachusetts, USA: Math Sci Press, 1975. 563 p.

7. *Bogoliubov N. N., Shirkov D. V.* Charge Renormalization Group in Quantum Field Theory // *Nuovo Cim.* 1956. V. 3. P. 845—863.

8. *Боголюбов Н. Н., Ширков Д. В.* Введение в теорию квантованных полей. 3-е изд. М.: Наука, 1976. 480 с.; *Боголюбов Н. Н., Ширков Д. В.* Квантовые поля. 3-е изд., доп. М.: Физматлит, 2005. 384 с.

9. *Ширков Д. В.* Ренормгруппа Боголюбова // *УМН.* 1994. Т. 49. С. 47—64; Препринт ОИЯИ Р2-94-310; hep-th/9602024.

10. *Chen L. Y., Goldenfeld N., Oono Y.* The Renormalization group and singular perturbations: multiple-scales, boundary layers and reductive perturbation theory // *Phys. Rev.* 1996. V. E54(1). P. 376—394.

11. *Овсянников Л. В.* Групповой анализ дифференциальных уравнений. Москва: Наука, 1978. 400 с.

12. CRC Handbook of Lie Group Analysis of Differential Equations. In V. 3 / Ed. by N. H. Ibragimov. Boca Raton, Florida, USA: CRC Press. V. 1. 1994. 429 p.; V. 2. 1995. 546 p.; V. 3. 1996. 536 p.

13. *Kovalev V. F., Shirkov D. V.* The renormalization group symmetry for solution of integral equations // *Proc. Institute Mathematics of NAS of Ukraine.* 2004. Part 2. V. 50. P. 850—861.

14. *Ibragimov N. H.* Elementary Lie Group Analysis and Ordinary Differential Equations. Chichester: John Wiley & Sons, 1999. 347 p.

15. *Shirkov D. V., Kovalev V. F.* Renormgroup symmetry for solution functionals: Preprint math-ph/0508055. 2005. 28 p.; *Ширков Д. В., Ковалев В. Ф.* Ренормгрупповая симметрия для функционалов от решения: Препринт ОИЯИ, Р-5-12-2004. Дубна, 2004, 58 с.

16. *Shirkov D. V.* Renormalization Group in Different Fields of Theoretical Physics: KEK Report 91-13, Feb. 1992. 85 p.

ВЫРОЖДЕННЫЕ РЕЗОНАНСЫ

А. Д. Морозов

1. Введение

Типичным свойством нелинейных осцилляторов $\ddot{x} + f(x) = 0$ является неизохронность колебаний. К таким осцилляторам относятся, например, уравнения Дюффинга $\ddot{x} + x \pm x^3 = 0$, уравнения с квадратичной нелинейностью $\ddot{x} + x \pm x^2 = 0$, маятниковое уравнение $\ddot{x} + \sin x = 0$. Для этих уравнений удается найти зависимость периода T от значения h интеграла энергии и показать, что $T(h)$ является монотонной (возрастающей или убывающей) функцией на соответствующих интервалах изменения h [1]. Формально нетрудно найти формулу для $T(h)$ при произвольной функции $f(x)$, когда уравнение $\ddot{x} + f(x) = 0$ имеет на фазовой плоскости (x, \dot{x}) область, заполненную замкнутыми фазовыми кривыми. Действительно, используя интеграл энергии $H_0(\dot{x}, x) \equiv \dot{x}^2/2 + U(x) = h = \text{const}$, получаем

$$T(h) = \oint_{L_h} \frac{dx}{\sqrt{2(h - U(x))}}. \quad (1)$$

Здесь $U(x) = \int f(x)dx$, а L_h — замкнутая фазовая кривая, определяемая соотношением $H_0(\dot{x}, x) = h$. Интеграл в (1) для указанных примеров выражается через полный эллиптический интеграл первого рода, откуда и следует монотонность функции $T(h)$. Однако существуют такие $f(x)$, для которых $T(h)$ будет немонотонной функцией. В этом случае $T(h)$ имеет хотя бы одну точку экстремума $h = h_0$. В случае, когда $f(x)$ — полином степени $n > 3$, интеграл в (1) абелев (в частности, гиперэллиптический при $n = 4$). Поэтому исследование зависимости $T(h)$ в общем случае — это непростая задача!

Рассмотрим теперь периодические по времени возмущения уравнений с одной степенью свободы общего вида

$$\begin{aligned} \dot{x} &= \frac{\partial H_0(x, y)}{\partial y} + \varepsilon g(x, y, \nu t), \\ \dot{y} &= -\frac{\partial H_0(x, y)}{\partial x} + \varepsilon r(x, y, \nu t). \end{aligned} \quad (2)$$

Предположим, что невозмущенная система ($\varepsilon = 0$) нелинейная и имеет ячейку $D_0 \subset R^2$, заполненную замкнутыми фазовыми кривыми. Предположим также, что функция $H_0(x, y)$, а также функции g и r , определяющие возмущение, достаточно гладкие (аналитические) по своим

аргументам в области D_0 . Кроме этого предположим, что возмущение непрерывное (гладкое) и периодическое по $\varphi = \nu t$ с периодом 2π , ν — частота возмущения.

Переходя в D_0 от переменных x, y к переменным действие I — угол θ [2], придем к системе:

$$\begin{aligned} \dot{I} &= \varepsilon(rx'_\theta - gy'_\theta) \equiv \varepsilon F(I, \theta, \varphi), \\ \dot{\theta} &= \omega(I) + \varepsilon(-rx'_I + gy'_I) \equiv \omega(I) + \varepsilon G(I, \theta, \varphi), \\ \dot{\varphi} &= \nu, \end{aligned} \quad (3)$$

где ε — малый положительный параметр, функции F, G периодические по θ и φ с периодом 2π , достаточно гладкие в области $\Delta \times T^2$ ($I \in \Delta = (I_{\min}, I_{\max})$, $\theta, \varphi \in T^2$, $T^2 = S^1 \times S^1$ — двумерный тор), $\omega = \partial H_0(I)/\partial I$ (подробности см., например, в [1]). Собственная частота $\omega = 2\pi/T$ зависит от действия I (невозмущенная система нелинейная) и может быть как монотонной, так и немонотонной функцией на рассматриваемом интервале Δ . В данной лекции основное внимание уделим случаю немонотонной функции $\omega(I)$.

Если выполнено условие

$$\omega^{(k)}(I_0) = 0, \quad k = 1, 2, \dots, j-1; \quad \omega^{(j)}(I_0) \neq 0, \quad j > 1, \quad (4)$$

то уровень $I = I_0$ будем называть *вырожденным*. Здесь $\omega^{(k)} = d^k \omega / dI^k$. Величину j будем называть порядком вырождения. Если $\omega^{(1)}(I_0) \neq 0$ ($j = 1$), то уровень $I = I_0$ будем называть *невырожденным*. При изменении частоты возмущения возможно равенство $I_0 = I_{pq}$, где I_{pq} определяется из условия резонанса

$$\omega(I_{pq}) = q\nu/p, \quad (5)$$

p, q — взаимно простые целые числа. В этом случае вырожденный уровень $I = I_0$ совпадает с резонансным уровнем $I = I_{pq}$ и говорят о *вырожденном резонансе*.

Как известно [1, 3], в окрестности

$$I_{pq} = \{(I, \theta) : I_{pq} - C\varepsilon^s < I < I_{pq} + C\varepsilon^s, 0 \leq \theta \leq 2\pi\}, \quad s = \frac{1}{j+1}$$

индивидуального резонансного уровня исходная система (3) приводится к виду

$$\begin{aligned} \dot{u} &= \varepsilon^{1-s} A(v, I_{pq}) + \varepsilon P(v, I_{pq})u + O(\varepsilon^{1+s}), \\ \dot{v} &= \varepsilon^{1-s} b_j u^j + \varepsilon (b_{j+1} u^{j+1} + Q(v, I_{pq})) + O(\varepsilon^{1+s}), \quad j \geq 1, \end{aligned} \quad (6)$$

где

$$A = \frac{1}{2\pi p} \int_0^{2\pi p} F(I_{pq}, v + \frac{q}{p}\phi, \phi) d\phi, \quad P = \frac{1}{2\pi p} \int_0^{2\pi p} \frac{\partial F(I_{pq}, v + \frac{q}{p}\phi, \phi)}{\partial I} d\phi,$$

$$Q = \frac{1}{2\pi p} \int_0^{2\pi p} R(I_{pq}, v + \frac{q}{p}\phi, \phi) d\phi, \quad b_j = \frac{1}{j!} \frac{d^j \omega(I_{pq})}{dI^j}.$$
(7)

Функции $A(v; I_{pq}), P(v; I_{pq}), Q(v; I_{pq})$ периодические по v с наименьшим периодом $2\pi/p$.

Следуя работам [3—7], [9, 10] (см. также [1]) установим возможные структуры вырожденных резонансных зон для любого порядка вырождения и приведем примеры систем с одной степенью свободы, для которых $j \geq 3$. Предварим это исследование рассмотрением невырожденных резонансов.

Далее рассмотрим вырожденные резонансы в системах с двумя степенями свободы, которые в переменных действие — угол можно записать в виде

$$\begin{aligned} \dot{I} &= \varepsilon F(I, J, \theta, \beta), & \dot{J} &= \varepsilon G(I, J, \theta, \beta), \\ \dot{\theta} &= \rho(I, J) + \varepsilon K(I, J, \theta, \beta), & \dot{\beta} &= 1, \end{aligned}$$
(8)

где переменные действия (I, J) — медленные переменные, изменяемые в некоторой односвязной области $\Delta \subset R^2$; $(\theta, \beta) \pmod{2\pi}$ — быстрые угловые переменные; $\varepsilon > 0$ — малый параметр; $\dot{} = d/dt$; ρ — гладкая функция медленных переменных (отношение частот); F, G, K — гладкие функции фазовых переменных, 2π -периодические по углам.

В заключительной части рассмотрим вопрос о вырожденных резонансах в сохраняющих площадь отображениях цилиндра с немонотонным числом вращения.

2. Невырожденные резонансы в системах со степенями свободы $3/2$

Структура индивидуальных резонансных зон описывается системой (6) с $j = 1$. Пренебрегая членами $O(\varepsilon^{3/2})$ и представляя функцию $A(v, I_{pq})$ в виде $A(v, I_{pq}) = \tilde{A}(v, I_{pq}) + B_0(I_{pq})$, приходим к усредненной системе

$$\begin{aligned} \dot{u} &= \varepsilon^{1/2}(\tilde{A}(v, I_{pq}) + B_0(I_{pq})) + \varepsilon P(v, I_{pq})u, \\ \dot{v} &= \varepsilon^{1/2}bu + \varepsilon(b_2u^2 + Q(v, I_{pq})), \end{aligned}$$
(9)

которая определяет структуру невырожденных резонансных зон. Пренебрегая в (9) членами порядка $O(\varepsilon)$, приходим к системе, которую назовем укороченной.

Проходимые, частично проходимые и непроходимые резонансы. Рассмотрим укороченную систему. При $B_0(I_{pq}) \neq 0$ уровень $I = I_{pq}$ назовем *проходимым*, если $\max|A(v)| < |B_0|$, и *частично проходимым*, если $\max|A(v)| > |B_0|$. При $B_0(I_{pq}) = 0$ резонансный уровень $I = I_{pq}$ назовем *непроходимым*. В последнем случае при $B_0(I) \neq 0$ у соответствующей автономной возмущенной системы в окрестности уровня $I = I_{pq}$ существует предельный цикл [1]. Отметим, что $B_0(I_{pq}) = 0$ при консервативных возмущениях. В этом случае дивергенция векторного поля системы (9) тождественно равна нулю и $B_0(I) \equiv 0$.

Фазовые портреты представлены на рис. 1 (по вертикали u , по горизонтали v). На рис. 2 показано поведение фазовых кривых укороченной системы в плоскости исходных переменных x, y . Сепаратрисы седел выделяют резонансную зону.

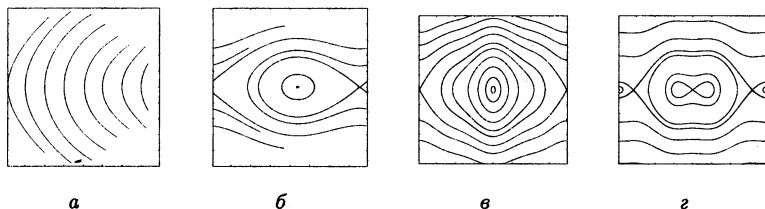


Рис. 1. Фазовые портреты укороченной системы; а: $\max|A(v)| < |B_0|$, б: $\max|A(v)| > |B_0|$, в, з: $B_0(I_{pq}) = 0$.

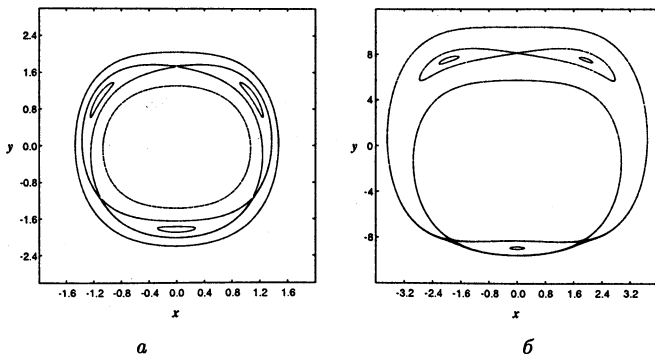


Рис. 2. Соответствующие рис. 1, в, з фазовые портреты системы (9) в плоскости (x, y) при $p = 3$ (а) и $p = 1$ (б); $q = 1$.

Учтем теперь члены порядка $O(\varepsilon)$ в системе (9). Для случая проходимого резонанса поведение фазовых кривых качественно не изменится. В случае частично проходимого резонанса при неконсервативном возмущении состояние равновесия типа «центр» перейдет в состояние равновесия типа «фокус».

В случае непроходимого резонанса система (9) приводится к виду автоколебательного уравнения:

$$\frac{d^2 v}{d\tau^2} - b_1 \tilde{A}(v, I_{pq}) = \sqrt{\varepsilon} \sigma(v, I_{pq}) \frac{dv}{d\tau}, \quad \tau = \sqrt{\varepsilon} t, \quad \sigma = P + Q'_v. \quad (10)$$

Возможные фазовые портреты уравнения (10) приведены на рис. 3. Устойчивому состоянию равновесия уравнения (10) (а также и системы (9)) отвечает в исходной системе устойчивый резонансный режим синхронизации колебаний с периодом $2\pi p/qv$, а устойчивому предельному циклу – двумерный инвариантный тор (подробнее см. в [1]).

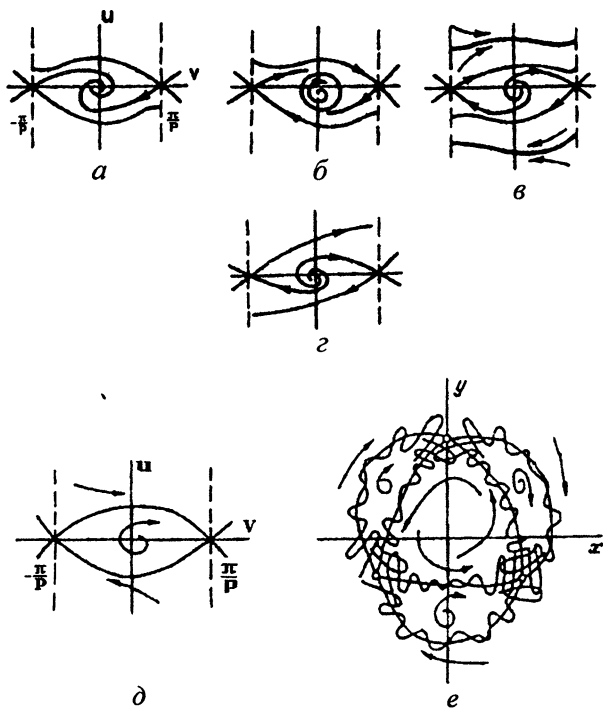


Рис. 3. Возможное поведение фазовых кривых уравнения (10) (а – д) и инвариантных кривых отображения Пуанкаре (е) для случая d и $p = 3$.

3. Вырожденные резонансы в системах со степенями свободы 3/2

Укороченная система, получающаяся из (6) при $j > 1$, имеет вид

$$\begin{aligned} \dot{u} &= \varepsilon^{1-s}(\tilde{A}(v, I_{pq}) + B_0(I_{pq})), \\ \dot{v} &= \varepsilon^{1-s}b_j u^j, \quad s = 1/(j+1), \quad j \geq 2. \end{aligned} \quad (11)$$

Эта система допускает первый интеграл, и поэтому нетрудно построить фазовые портреты. В случае гамильтоновых возмущений имеем $B_0(I_{pq}) = 0$. Справедливо [1, 5] следующее утверждение.

Утверждение 1. Пусть функции g и f в исходной системе являются тригонометрическими многочленами степени N по $\varphi = \nu t$. Тогда функция $\tilde{A}(v; I_{pq})$, определяющая укороченную систему (11), также является тригонометрическим многочленом степени не выше N .

Так как система (11) может иметь только сложные состояния равновесия, то рассмотрим деформации векторного поля этой системы, приводящие к системе с простыми состояниями равновесия:

$$\begin{aligned} \dot{u} &= \varepsilon^{1-s}(\tilde{A}(v, I_{pq}) + B_0(I_{pq})), \\ \dot{v} &= \varepsilon^{1-s}(b_j u^j + \sum_{k=j-1}^1 p_k u^k), \quad s = 1/(j+1), \quad j \geq 2, \end{aligned} \quad (12)$$

где p_k — параметры деформации.

Утверждение 2. Если функции g и f в исходной системе (2) являются тригонометрическими полиномами степени N по $\varphi = \nu t$, тогда система (12) имеет не более чем $2Nj$ простых состояний равновесия.

3.1. Гармонические консервативные возмущения

Рассмотрим случай $N = 1$ и $B_0(I_{pq}) = 0$. Вводя медленное время $\tau = \varepsilon^{1-s}t$, приведем систему (12) к виду

$$\dot{u} = b \sin(v), \quad \dot{v} = u^j + \sum_{k=j-1}^1 p_k u^k, \quad j \geq 2, \quad (13)$$

где постоянная b зависит от I_{pq} (рис. 4).

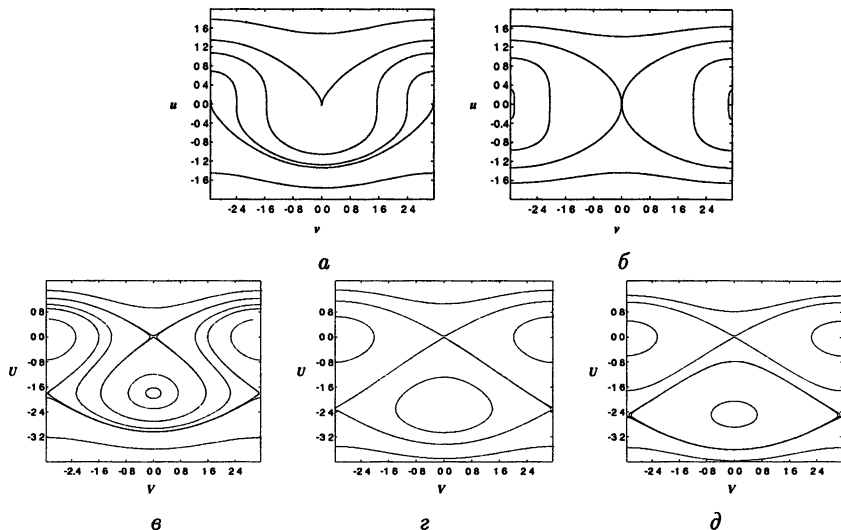


Рис. 4. Фазовые портреты системы (13): $j = 2, p_1 = 0$ (а); $j = 3, p_1 = p_2 = 0$ (б). Перестройки фазового портрета при $b = 1, j = 2$ и изменении параметра p_1 : $p_1 = 1, 8$ (в); $2, 29$ (г); $2, 5$ (д).

3.2. Сепаратрисная сеть

Вырожденный резонанс порядка j распадается на j невырожденных резонансов (см., например, рис. 4, а, в — д). При сближении невырожденных резонансов существует такой момент, когда соответствующие сепаратрисы всех резонансов совпадают (см. рис. 4, г). В этом случае будем говорить о существовании максимальной сепаратрисной сети. Установим условия существования такой сети для системы (13). Пусть уравнение

$$u^j + \sum_{k=j-1}^1 p_k u^k = 0, \quad j \geq 2, \quad (14)$$

определяющее координату u состояний равновесия системы (13), имеет j простых вещественных корней $u_0, u_1, u_2, \dots, u_{j-1}$. Уравнения сепаратрис имеют вид

$$\tilde{H}(u, v) \equiv u^{j+1}/(j+1) + \sum_{k=j-1}^1 p_k u^{k+1}/(k+1) - b \cos v = h_s, \quad (15)$$

где

$$h_s = \frac{u_s^{j+1}}{j+1} + \sum_{k=j-1}^1 \frac{p_k}{k+1} u_s^{k+1} + (-1)^s b, \quad s = 0, 1, \dots, j-1. \quad (16)$$

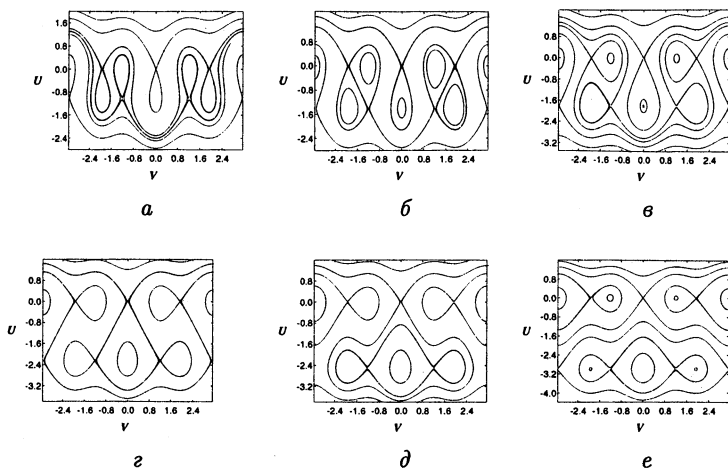


Рис. 5. Траектории системы (18) при $j = 2$, $a = 2$: $p_1 = 1$ (а); $p_1 = 1, 41$ (б); $p_1 = 1, 8$ (в); $p_1 = 2, 245$ (г); $p_1 = 2, 5$ (д); $p_1 = 3$ (е).

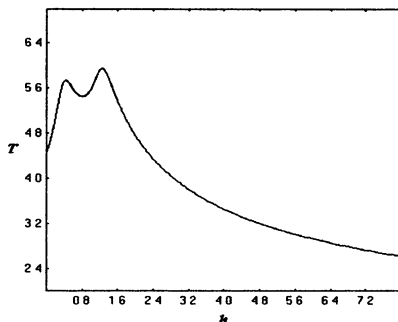


Рис. 6. График функции $T(h)$ при $c_1 = 2$; $c_2 = 1, 2$; $c_3 = -2$; $c_4 = -1, 1$; $c_5 = 1$.

Изменяя параметры c_k нетрудно добиться слияния экстремумов на рис. 6, что дает сложную критическую точку с $j = 4$ ($c_1 = 2$, $c_2 = 0, 5$, $c_3 = -2, 2$, $c_4 = -0, 4$, $c_5 = 1$). Продемонстрируем новые бифуркации для отображения Пуанкаре в окрестности вырождения с порядком $j = 4$, которые не наблюдаются при анализе усредненных систем в консервативном случае. В качестве примера рассмотрим резонанс с $p = 2, q = 1$. Зафиксируем параметры $c_1 = 2$, $c_2 = 0, 5$, $c_3 = -2, 3$, $c_4 = -0, 4$, $c_5 = 1, \varepsilon = 0, 1$ и будем изменять параметр ν .

Перестройки в окрестностях двух резонансов с $p = 2$ представлены на рис. 7. Показано взаимодействие двух резонансных зон с $p = 2$, $q = 1$. Структуры на рис. 7, *б*, *е* мы, следуя [11], называем «вихревой парой» (см. также рис. 5, *б*).

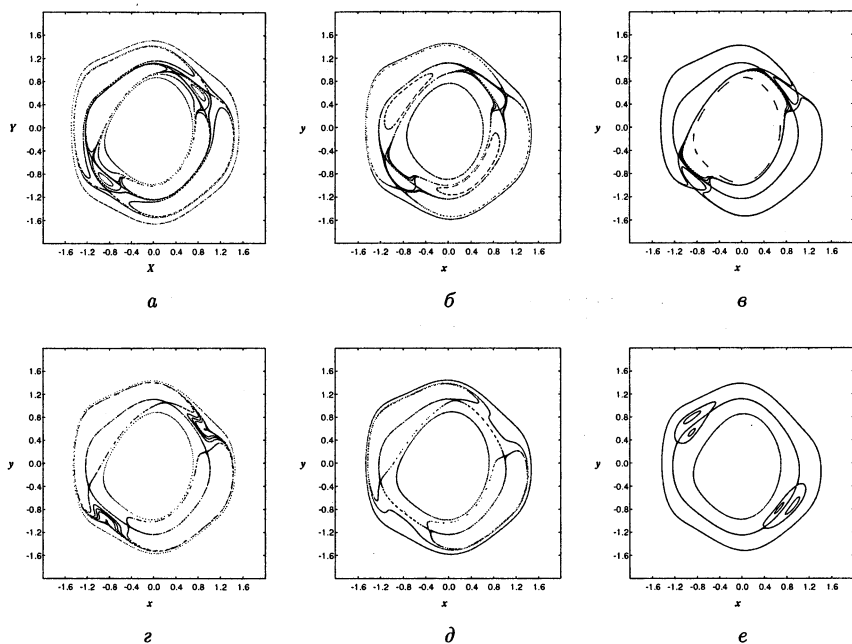


Рис. 7. Отображение Пуанкаре для уравнения (5) при $n = 5$; $c_1 = 2$; $c_2 = 0,5$; $c_3 = -2,3$; $c_4 = -0,4$; $c_5 = 1$; $\varepsilon = 0,1$ и $\nu = 1,93$ (*а*); $\nu = 1,929$ (*б*); $\nu = 1,927$ (*в*); $\nu = 1,92$ (*г*); $\nu = 1,91$ (*д*); $\nu = 1,85$ (*е*).

В заключение этого пункта отметим, что существование вырожденных резонансов второго рода в резонансных зонах первого рода для гамильтоновых систем с $3/2$ степенями свободы является типичным явлением [6].

О вырожденных резонансах в случае неконсервативных возмущений можно прочитать в книге [1] (см. также [6, 9]). Заметим лишь, что диссипация приводит к уменьшению числа не только вырожденных, но также и невырожденных резонансных режимов. При достаточно большой диссипации все резонансы проходимые.

4. Вырожденные резонансы в системах с двумя степенями свободы

Следуя [11], рассмотрим систему (8), к которой приводят, в частности, системы двух связанных осцилляторов. Предположим, что $\nabla\rho \neq 0$.

Определение 1. *Говорят, что в системе (8) имеет место резонанс, если*

$$\rho(I, J) = r = m/n \in \mathbb{Q}. \quad (20)$$

В точках (I, J) , определяемых условием $\nabla\rho(I, J) = 0$, слоение $\rho(I, J) = \text{const}$ имеет особенность. В случае резонансных точек говорят о вырожденном резонансе.

Определение 2. *Будем говорить, что в точке (I_*, J_*) имеет место вырождение k -го порядка, если в этой точке все производные от функции ρ по переменным I, J до порядка k включительно равны нулю и имеется по крайней мере одна ненулевая производная порядка $k + 1$.*

Если при этом

$$\rho_* = \rho(I_*, J_*) = r = m/n \in \mathbb{Q} \quad (21)$$

с некоторыми взаимно простыми целыми m, n , то будем говорить, что в точке (I_, J_*) имеет место вырожденный резонанс $m : n$ с порядком вырождения k .*

Пусть точка (I_*, J_*) — изолированная критическая точка функции ρ , для которой имеем вырожденный резонанс $m : n$ с порядком вырождения k . Тогда система (8) вблизи этой точки преобразуется к виду

$$\begin{aligned} \dot{x} &= \mu^{k+1} f_*(\varphi) + O(\mu^{k+2}), \\ \dot{y} &= \mu^{k+1} g_*(\varphi) + O(\mu^{k+2}), \\ \dot{\varphi} &= \mu^{k+1} Q_*^{(k+1)}(x, y) + O(\mu^{k+2}), \\ \dot{\beta} &= 1, \quad \mu = \varepsilon^{(k+2)^{-1}}, \end{aligned} \quad (22)$$

где правые части $2\pi/n$ -периодические по φ . Эта система приводится в случае вырождения первого порядка к следующей укороченной системе (нормальной форме):

$$\begin{aligned} \xi' &= f(\varphi), \\ \eta' &= g(\varphi), \\ \varphi' &= \xi^2 + \delta\eta^2, \end{aligned} \quad (23)$$

где штрих означает производную по медленному времени $\tau = \mu^{k+1}t$, $\delta = \pm 1$ или $\delta = 0$.

В настоящее время исследованы лишь некоторые примеры систем вида (23) [9].

5. Отображения цилиндра с немонотонным числом вращения

Рассмотрим симплектическое отображение

$$T : \begin{cases} \bar{x} = x + K \sin \theta, \\ \bar{\theta} = \theta + p_1 \bar{x} + p_2 \bar{x}^2, \end{cases} \quad (24)$$

где K, p_1, p_2 — параметры ($p_2 \neq 0$); при $p_2 = 0$ имеем стандартное отображение Чирикова. Отображение (24) называют стандартным неэквивариантным отображением (SNTM, см. [11]). Это связано с тем, что функция $f(x) = p_1 x + p_2 x^2$ определяет число вращения и является немонотонной.

Для отображения (24) резонансные уровни $x = x_{pq}$ в области вращательных движений определяются из условия $f(x) = 2\pi q/p$, где p, q — взаимно простые целые числа. Поэтому $x_{pq}^{\pm} = (-p_1 \pm \sqrt{p_1^2 + 8\pi p_2 q/p})/2p_2$.

Вырожденный уровень $x = x_d$ с первым порядком вырождения ($j = 2$) определяется из условия $f'(x_d) = 0$. Отсюда получаем $x_d = -p_1/2p_2$. При выполнении условия $x_d = x_{pq}$ говорят о вырожденном резонансе.

При $p = 1, q = 1$ на резонансном уровне расположена пара неподвижных точек (гиперболическая и эллиптическая), для которых $\theta = 0; \pi$. При $p_1 = 1, p_2 = -0,038$ имеем $x_{11}^+ \simeq 15,9; x_d \simeq 13,2$. На рис. 8, а показаны траектории отображения (24) вблизи вырожденного резонанса ($p = 1, q = 1$). При изменении параметра p_2 и $p_1 = 1$ получаем картины, аналогичные рис. 3, а, $\epsilon - \delta$. На рис. 8, б показаны «вихревые пары» ($p = 2, q = 1$), а также нерегулярные траектории. В этом случае на резонансном уровне $x = x_{21}$ расположены периодические периода 2 точки (подробнее см. в [1, 4]).

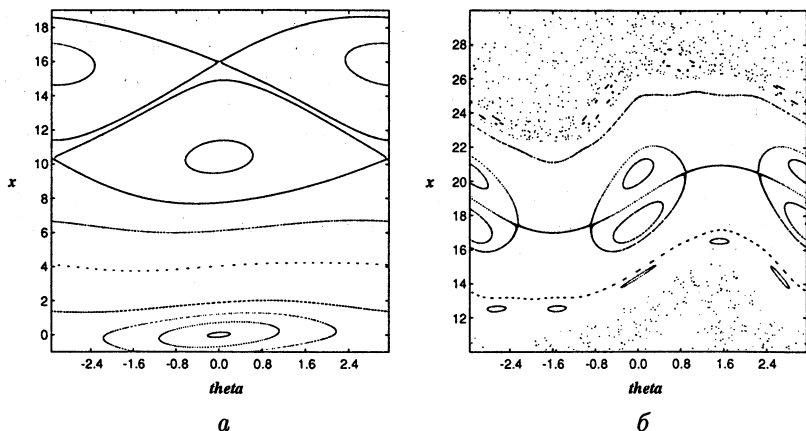


Рис. 8. Траектории отображения (24): при $p_1 = 1$, $p_2 = -0,038$, $K = 0,5$ (а); при $p_1 = 1$, $p_2 = -0,02635$, $K = 4$ (б).

Работа по исследованию резонансов частично поддерживается грантами РФФИ (№06-01-00270), НШ (№9686.2006.1).

Литература

1. Морозов А. Д. Резонансы, циклы и хаос в квазиконсервативных системах. М.; Ижевск: РХД, 2005. 422 с.
2. Арнольд В. И. Математические методы классической механики. М.: Наука, 1974. 431 с.
3. Морозов А. Д., Шильников Л. П. О неконсервативных периодических системах, близких к двумерным гамильтоновым // ПММ. 1983. Т. 47, вып. 3. С. 385–394.
4. Morozov A. D. Degenerate resonances in Hamiltonian systems with 3/2 degrees of freedom // Chaos. 2002. V. 12, № 3. P. 539–548.
5. Морозов А. Д. О вырожденных резонансах в двумерных периодических системах // Вестн. ННГУ. Сер. Математика. 2003. Вып. 1(1). С. 33–44.
6. Морозов А. Д., Ашаева Е. В. К вопросу о существовании вырожденных резонансов в гамильтоновых системах // Вестн. ННГУ. Сер. Математика. 2004. Вып. 1(2). С. 126–136.
7. Morozov A. D. On degenerate resonances in nearly Hamiltonian systems // Regular and Chaotic Dynamics. 2004. V. 9, № 3. P. 337–350.
8. Soskin S. M., Luchinsky D. G., Mannela R., Neiman A. B., McClintock P. V. E. Zero-dispersion nonlinear resonances // Int. J. Bifurcation and Chaos. 1997. V. 7, № 4. P. 923–936.

9. Морозов А. Д. О резонансах в уравнениях с немонотонным вращением // Вестн. ННГУ. Сер. Математика. 2006. Вып. 1(4). С. 126—136.

10. Morozov A. D., Boykova S. A. On investigation of the degenerate resonances // Regular and Chaotic Dynamics. 1999. V. 4, № 1. P. 70—82.

11. Howard J. E., Humpherys J. Nonmonotonic twist maps // Physica D. 1995. V. 80, № 3. P. 256—276.

12. Карабанов А. А., Морозов А. Д. Об усреднении вблизи вырожденного резонанса в четырехмерной двухчастотной задаче // Труды СВМО. 2004. Т. 6, № 1. С. 292—357.

ЗАТЯГИВАНИЕ ПОТЕРИ УСТОЙЧИВОСТИ ПРИ ДИНАМИЧЕСКИХ БИФУРКАЦИЯХ

А. И. Нейштадт

В классической теории бифуркаций поведение системы, зависящей от параметра, рассматривается для значений параметра, близких к некоторому критическому бифуркационному значению. Это статическая теория в том смысле, что значение параметра не зависит от времени. В теории динамических бифуркаций параметр медленно изменяется во времени и проходит через значение, которое было бы критическим в классической статической теории. Некоторые явления, которые здесь возникают, совершенно отличны от, на первый взгляд, естественных, предсказаний, которые могут быть сделаны на основе статической теории. Одно из таких явлений — затягивание потери устойчивости.

1. Общее описание явления затягивания потери устойчивости

Рассмотрим систему дифференциальных уравнений вида

$$\begin{aligned}\dot{x} &= f(x, \tau), \\ \dot{\tau} &= \varepsilon.\end{aligned}\tag{1}$$

Здесь x — вектор, который будем называть вектором фазовых переменных, а τ — скалярная переменная, которую будем называть параметром, $0 < \varepsilon \ll 1$, точка обозначает производную по времени t . Для простоты изложения вектор x будем считать двумерным. Поскольку параметр τ изменяется медленно, целесообразно сначала рассмотреть уравнение для x отдельно при каждом фиксированном значении τ . Будем предполагать, что это уравнение при каждом τ из некоторого отрезка $[a, b]$ имеет невырожденное положение равновесия $x = X(\tau)$ с комплексно-сопряженными собственными числами $\lambda_{1,2}(\tau)$, $\text{Im } \lambda_{1,2}(\tau) \neq 0$. Пусть имеется критическое значение параметра $\tau = \tau_*$ такое, что при $\tau < \tau_*$ собственные числа лежат в левой полуплоскости, а при $\tau > \tau_*$ — в правой полуплоскости, $\text{Im } \lambda_{1,2}(\tau_*) \neq 0$. Таким образом, при переходе параметра через критическое значение равновесие теряет устойчивость, оставаясь невырожденным. В пространстве x, τ имеется кривая равновесий, ее часть с $\tau < \tau_*$ составлена из асимптотически устойчивых равновесий, а часть с $\tau > \tau_*$ — из неустойчивых равновесий (рис. 1). Эти части называют соответственно устойчивой и неустойчивой частью кривой равновесий. В случае общего положения в рассматриваемой задаче имеет место бифуркация Пуанкаре — Андронова — Хопфа [1]: при $\tau = \tau_*$ от кривой равновесий ответвляется семейство предельных циклов. На рис. 1 показан случай, когда циклы устойчивы и лежат в области $\tau > \tau_*$; в этом случае бифуркация называ-

ется мягкой [1]. Пусть теперь параметр τ медленно изменяется со временем: $\dot{\tau} = \varepsilon > 0$. Пусть в начальный момент времени $\tau < \tau_*$, и фазовая точка находится недалеко от кривой равновесий. Тогда она притянется к кривой равновесий и будет двигаться вблизи нее, по меньшей мере, до момента $\tau = \tau_*$.

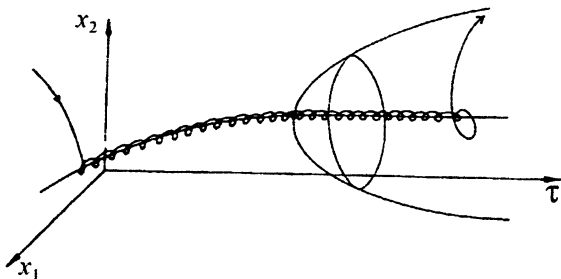


Рис. 1

Явление затягивания потери устойчивости состоит в следующем. При $\tau = \tau_*$ ничего не происходит, и фазовая точка продолжает двигаться уже вблизи неустойчивой части кривой равновесий еще долго, в течение времени $\sim 1/\varepsilon$, как бы мало ни было значение ε . За это время параметр τ успевает измениться на величину порядка 1. Только потом происходит быстрый уход от неустойчивой части кривой равновесий (говорят о срыве с кривой равновесий). После срыва фазовая точка в пространстве x, τ , возможно, притянется к поверхности, составленной из циклов, ответвившихся в результате бифуркации от кривой равновесий (см. рис. 1). Но может быть и так, что при соответствующих значениях τ эти циклы уже не существуют, и фазовая точка, например, притянется к другому устойчивому режиму или уйдет на бесконечность.

Выше говорилось о равновесиях, но это равновесия для системы с замороженным параметром. Для системы с изменяющимся параметром значение $X(\tau)$, вообще говоря, изменяется со временем, так что это не равновесие. Если $X(\tau) = X_* \equiv \text{const}$, то X_* — действительно равновесие, и мы имеем тривиальный пример затягивания потери устойчивости. В области устойчивости ($\tau < \tau_*$) фазовая точка так близко притягивается к равновесию, что требуется большое время, чтобы в области неустойчивости уйти от равновесия на расстояние порядка 1. Но этот случай бесконечно вырожденный: определенная функция от τ должна оказаться тождественно постоянной. Мы этим случаем заниматься не будем. Однако, хотя в общем случае функция $X(\cdot)$ — не постоянная, явление затягивания потери устойчивости, в некотором смысле, очень общее.

Явление (не тривиального) затягивания потери устойчивости было открыто в выполненной под руководством Л. С. Понтрягина работе М. А. Шишковой [2] для следующей системы:

$$\begin{aligned} \dot{z} &= (\tau - i)(z - \tau) + \gamma(z - \tau)|z - \tau|^2, \\ \dot{\tau} &= \varepsilon. \end{aligned} \quad (2)$$

Здесь z — комплексная переменная, $z = u + iv$, так что вместо комплексного уравнения для z можно написать систему двух вещественных уравнений для u и v ; $\gamma = \text{const}$. Для $\varepsilon = 0$ ($\tau = \text{const}$) система имеет равновесие $z = \tau$ с собственными числами $\lambda_{1,2} = \tau \pm i$. Это равновесие асимптотически устойчиво при $\tau < 0$ и неустойчиво при $\tau > 0$. Выписанное уравнение является уравнением в нормальной форме для бифуркации Пуанкаре — Андронова — Хопфа, мягкой при $\gamma < 0$ и жесткой при $\gamma > 0$. В работе [2] показано, что если параметр τ медленно изменяется со временем ($\varepsilon > 0$), то имеет место затягивание потери устойчивости: фазовые точки, притянувшиеся к кривой (здесь — прямой) равновесий при $\tau < -1$ сорвутся с нее при $\tau \approx 1$. Так что потеря устойчивости при любом γ выглядит как жесткая. В [2] введен и основной метод исследования затягивания потери устойчивости: рассмотрение динамики в комплексном времени.

В работе [3] результаты [2] распространены на некоторые частные классы систем. Существование (не тривиального) затягивания потери устойчивости можно вывести также из результатов [4].

Оказывается, в классе аналитических систем явление затягивания потери устойчивости — не просто общее, а неизбежное. А именно, в ситуации, описанной в начале этого раздела, если система аналитична, то обязательно имеет место затягивание потери устойчивости [5, 6].

Оказывается также, что это явление связано именно с аналитичностью. Есть примеры систем конечной гладкости, и даже систем класса C^∞ , но не аналитических, для которых потеря устойчивости не затягивается: срыв с потерявшего устойчивость равновесия происходит в интервале значений параметра, стягивающемся при $\varepsilon \rightarrow 0$ к бифуркационному значению [6].

При затягивании потери устойчивости происходит следующее. Система (1) в области $a \leq \tau \leq \tau_*$ имеет интегральную кривую Γ_- , близкую к устойчивой части кривой равновесий. В области $\tau_* \leq \tau \leq b$ эта система имеет интегральную кривую Γ_+ , близкую к неустойчивой части кривой равновесий. Пусть d — расстояние между точками кривых Γ_- и Γ_+ в сечении $\tau = \tau_*$. Кривую Γ_- можно продолжить в область $\tau > \tau_*$, а кривую Γ_+ — в область $\tau < \tau_*$. Пусть $\rho(\tau)$ — расстояние между точками этих кривых в зависимости от значения параметра τ . Ясно, что в области неустойчивости это расстояние растет по закону $\dot{\rho} \sim (\tau - \tau_*)\rho$ и, значит, $\ln(\rho/d) \sim$

$\sim (\tau - \tau_*)^2/2\varepsilon$. Если $\tau_* + \tau_d$ — значение параметра τ , при котором впервые станет $\rho \sim 1$, то $\tau_d \sim \sqrt{\varepsilon |\ln d|}$. Значение τ_d/ε можно назвать временем затягивания.

Оказывается, если система аналитична, то величина d экспоненциально мала: $d \lesssim e^{-\text{const}/\varepsilon}$ [6]. Тогда $\tau_d \gtrsim 1$, а это и означает затягивание потери устойчивости.

Можно привести пример системы гладкости C^V , в котором $d \sim \varepsilon^V$. В этом случае $\tau_d \sim \sqrt{V\varepsilon |\ln \varepsilon|} \rightarrow 0$ при $\varepsilon \rightarrow 0$, и затягивания нет. Для любой функции $\beta(\cdot)$, $\beta(\varepsilon) \rightarrow \infty$ при $\varepsilon \rightarrow 0$, можно привести пример системы гладкости C^∞ , в котором $d \geq e^{-\beta(\varepsilon)|\ln \varepsilon|}$. В этом случае $\tau_d \sim \sqrt{\beta(\varepsilon)\varepsilon |\ln \varepsilon|}$, и можно выбрать функцию $\beta(\cdot)$ так, что $\tau_d \rightarrow 0$ при $\varepsilon \rightarrow 0$, и затягивания нет.

Из этого описания ясно, что экспоненциально малый шум должен разрушать затягивание. Тем не менее затягивание наблюдается при численных и даже при реальных экспериментах (некоторые из этих экспериментов были выполнены до аналитического описания явления затягивания), см., например, [7 — 10].

Естественный класс систем для явления затягивания потери устойчивости — быстро-медленные системы (называемые также системами с быстрыми и медленными переменными или системами с быстрыми и медленными движениями) [11]. Это системы вида

$$\begin{aligned} \dot{x} &= f(x, y, \varepsilon), & x \in R^n, \\ \dot{y} &= \varepsilon g(x, y, \varepsilon), & y \in R^m. \end{aligned} \quad (3)$$

Частным случаем таких систем являются системы вида (1). В системе (3) переменные x называют быстрыми, а переменные y — медленными. Первое из уравнений системы (3) при $y = \text{const}$ и $\varepsilon = 0$ называют быстрым уравнением (или быстрой системой). Пусть быстрое уравнение для всех значений y из некоторой области имеет невырожденное положение равновесия $x = X(y)$. Рассмотрим поверхность равновесий, называемую также медленной поверхностью $\{x, y: x = X(y)\}$. Будем считать, что у поверхности равновесий есть устойчивая и неустойчивая части. Для устойчивой части все собственные числа равновесий лежат в левой полуплоскости комплексной плоскости. Для неустойчивой части у равновесий есть собственные числа в правой полуплоскости. Устойчивая и неустойчивая части поверхности равновесий разделены поверхностью на единицу меньшей размерности, составленной из равновесий, у которых некоторые из собственных чисел — чисто мнимые, а остальные собственные числа лежат в левой полуплоскости.

Уравнение

$$\dot{y} = \varepsilon G(y), \quad G(y) = g(X(y), y, 0)$$

называется медленным уравнением. Оно определяет фазовый поток на поверхности равновесий, называемый медленным потоком. Пусть у медленного потока есть траектории, пересекающие границу между устойчивой и неустойчивой частями поверхности равновесий. Если система аналитична, то потеря устойчивости неизбежно затягивается. Фазовая точка быстро-медленной системы (3), начавшая движение вблизи отрезка такой траектории, лежащего на устойчивой части поверхности равновесий, притягивается к этой траектории и движется вблизи нее еще время порядка $1/\varepsilon$ после пересечения границы области устойчивости [6]. Так что срыв с неустойчивой части поверхности равновесий происходит при значениях медленных переменных, находящихся на расстоянии порядка 1 от границы области устойчивости. Аналогично случаю системы (1), здесь происходит следующее. Быстро-медленная система имеет интегральные многообразия Γ_- и Γ_+ , лежащие соответственно вблизи устойчивой и неустойчивой частей поверхности равновесий. При значениях медленных переменных, принадлежащих границе области устойчивости, точки многообразий Γ_- и Γ_+ в аналитическом случае экспоненциально ($Oe^{-\text{const}/\varepsilon}$) близки. Поэтому поверхности Γ_- и Γ_+ можно продолжить в области неустойчивости и устойчивости соответственно на расстояние порядка 1 от границы области устойчивости.

Аналогичное затягивание потери устойчивости сопровождается в аналитических быстро-медленных системах (3) потерю устойчивости цикла быстрой системы. Пусть при каждом значении медленной переменной y из некоторой области быстрая система имеет невырожденный цикл L_y . В фазовом пространстве системы (3) возникает поверхность циклов $\{x, y: x \in L_y\}$. Будем считать, что у поверхности циклов есть устойчивая и неустойчивая части. Для устойчивой части все мультипликаторы циклов лежат внутри единичного круга. Для неустойчивой части у циклов есть мультипликаторы вне единичного круга. Устойчивая и неустойчивая части поверхности циклов разделены поверхностью на единицу меньшей размерности, составленной из циклов, у которых есть мультипликаторы на единичной окружности (не в точке 1), а остальные мультипликаторы лежат внутри единичного круга. Уравнение

$$\dot{y} = \varepsilon G(y),$$

где $G(y)$ — среднее значение функции $g(x, y, 0)$ на цикле L_y (усреднение проводится по времени), называется медленным уравнением [12]. Оно определяет поток на проекции поверхности циклов на пространство медленных переменных (медленный поток). Пусть у медленного потока есть траектории, пересекающие проекцию на пространство медленных переменных границы между устойчивой и неустойчивой частями поверхности

циклов. Если система аналитична, то потеря устойчивости затягивается. Именно, рассмотрим двумерный цилиндр, составленный из циклов, проектирующихся на одну из указанных траекторий медленного потока. Фазовая точка быстро-медленной системы (3), начавшая движение вблизи части этого цилиндра, составленной из устойчивых циклов, притягивается к этому цилиндру и движется вблизи него еще время $\sim 1/\varepsilon$ после пересечения границы области устойчивости [6].

Аналогичное затягивание потери устойчивости сопровождается в аналитических быстро-медленных отображениях потерю устойчивости неподвижной точки быстрого отображения. Быстро-медленное отображение — это отображение вида

$$\begin{aligned} x &\mapsto f(x, y, \varepsilon), \quad x \in R^n, \\ y &\mapsto y + \varepsilon g(x, y, \varepsilon), \quad y \in R^m. \end{aligned} \tag{4}$$

Для отображения (4) переменные x называют быстрыми, а переменные y — медленными. Первое из соотношений (4) при $y = \text{const}$, $\varepsilon = 0$ определяет отображение, называемое быстрым отображением. Пусть быстрое отображение для всех значений y из некоторой области имеет невырожденную неподвижную точку $x = X(y)$. Рассмотрим поверхность неподвижных точек, называемую также медленной поверхностью $\{x, y: x = X(y)\}$. Будем считать, что у поверхности неподвижных точек есть устойчивая и неустойчивая части. Для устойчивой части все мультипликаторы неподвижных точек лежат внутри единичного круга. Для неустойчивой части у неподвижных точек есть мультипликаторы вне единичного круга. Устойчивая и неустойчивая части поверхности неподвижных точек разделены поверхностью на единицу меньшей размерности, составленной из неподвижных точек, у которых есть мультипликаторы на единичной окружности (не в точке 1), а остальные мультипликаторы лежат внутри единичного круга. Отображение

$$y \mapsto y + \varepsilon G(y), \quad G(y) = g(X(y), y, 0)$$

называется медленным отображением. Поведение фазовой точки при итерациях этого отображения приближенно описывается сдвигом за единицу времени вдоль фазового потока, определяемого уравнением

$$\dot{y} = \varepsilon G(y)$$

(медленного потока). Пусть у медленного потока есть траектории, пересекающие границу между устойчивой и неустойчивой частями поверхности неподвижных точек. Если система аналитична, то потеря устойчивости неизбежно затягивается [6].

2. Существование затягивания потери устойчивости

Ниже излагается простой способ установления неизбежности затягивания потери устойчивости в аналитических системах. Будем опять рас-

смагивать систему (1). Введем вместо x новую переменную $\xi = x - X(\tau)$. Тогда

$$\dot{\xi} = f(X(\tau) + \xi, \tau) - \varepsilon \partial X / \partial \tau = A(\tau) \xi + O(|\xi|^2) + \varepsilon h(\tau). \quad (5)$$

Матрица $A(\tau)$ — невырожденная, при $\tau < \tau_*$ ее собственные числа лежат в левой полуплоскости, а при $\tau > \tau_*$ есть собственные числа в правой полуплоскости.

Пусть $\xi_\varepsilon(\tau, \varepsilon)$ — положение равновесия для уравнения (5) при $\tau = \text{const}$: $\xi_\varepsilon(\tau, \varepsilon) = -\varepsilon A^{-1} h + O(\varepsilon^2)$. Введем новую переменную $\xi_1 = \xi - \xi_\varepsilon$. Получим

$$\dot{\xi}_1 = A_1(\tau, \varepsilon) \xi_1 + O(|\xi_1|^2) + \varepsilon^2 h_1(\tau, \varepsilon). \quad (6)$$

Уравнение (6) имеет тот же вид, что и уравнение (5), но член, не обращающийся в 0 в начале координат, есть $O(\varepsilon^2)$. Повторяя аналогичные действия N раз, получим уравнение

$$\dot{\xi}_N = A_N(\tau, \varepsilon) \xi_N + O(|\xi_N|^2) + \varepsilon^{N+1} h_N(\tau, \varepsilon). \quad (7)$$

Оказывается, функция h_N допускает оценку: $|h_N| < a^N N!$, $a = \text{const} > 0$. Дело в том, что определение функции h_r включает дифференцирование по τ функции h_{r-1} . Так что определение функции h_N включает $(N+1)$ -кратное дифференцирование по τ функций из правой части уравнения (5). А производные аналитической функции растут с ростом порядка r производной не быстрее, чем $(\text{const})^r r!$ Это объяснение явления. Фактически, для получения оценки функции $|h_N|$ проще использовать оценки Коши (оценки максимума модуля производной аналитической функции в области через максимум модуля самой функции в более широкой области [13]), см. [6]. Теперь надо выбрать значение N , при котором величина $\varepsilon^N a^N N!$ будет иметь наименьшее значение (заметим, что с ростом N эта величина сначала убывает, а затем быстро растет). По формуле Стирлинга $\varepsilon^N a^N N! \lesssim (\varepsilon a N / e)^N$. Рассматривая N как непрерывную переменную, находим для точки минимума значение $N = 1/(\varepsilon a)$. Беря в качестве N целую часть этой величины, находим, что величина h_N экспоненциально мала: $|h_N| < c e^{-1/(\varepsilon a)}$, $c = \text{const} > 0$. Далее, $A_N = A + O(\varepsilon)$, так что при $\tau < \tau_* - \text{const} \varepsilon$ собственные числа матрицы A_N лежат в левой полуплоскости. Вне экспоненциально малой окрестности начала координат последний член в правой части уравнения (7) много меньше первого члена. Поэтому фазовая точка, начавшая движение при $\tau < \tau_*$, $\tau_* - \tau \sim 1$, не слишком далеко от начала координат, двигаясь в течение времени $\sim 1/\varepsilon$, обязана подойти экспоненциально близко, на расстояние $e^{-c_1/\varepsilon}$, $c_1 = \text{const} > 0$, к началу координат. Ей потребуется затем время, по меньшей мере, по-

рядка $1/\varepsilon$, чтобы отойти на расстояние ~ 1 от начала координат. А это и означает затягивание потери устойчивости.

Описанный способ доказательства наличия затягивания потери устойчивости легко распространяется на случай потери устойчивости равновесия или цикла в быстро-медленных системах, неподвижной точки в быстро-медленных отображениях. Но он не позволяет вычислить асимптотику времени затягивания потери устойчивости.

Для установления этой асимптотики полезно, следуя [2], рассматривать динамику системы в комплексном времени.

3. Время затягивания потери устойчивости, максимальное затягивание

Будем рассматривать систему (1), вектор фазовых переменных x для простоты формулировок будем считать двумерным. Будем называть время, измеряемое в единицах параметра τ , медленным временем. В пределе, при $\varepsilon \rightarrow 0$, притяжение к положению равновесия и срыв с него происходят мгновенно. Соответствующие предельные значения медленного времени называются моментами падения и срыва [11]. Функция $\Pi(\cdot)$, задающая зависимость момента срыва от момента падения, называется функцией входа-выхода [14]. Оказывается, при достаточно общих предположениях существуют значения τ_*^- и τ_*^+ параметра τ такие, что при $\tau < \tau_*^-$ будет $\Pi(\tau) = \tau_*^+$ (конечно, $\tau_*^- < \tau_* < \tau_*^+$). То есть фазовые точки, притянувшиеся к равновесию в разные моменты медленного времени, предшествующие моменту τ_*^- , сорвутся с равновесия приблизительно одновременно, в момент τ_*^+ . Это явление называется наличием максимального затягивания, а точка τ_*^+ — буферной точкой или стенкой [15]. Конечно, это явление отсутствует для тривиальных примеров затягивания, упомянутых в п. 2. Для $\tau \in (\tau_*^-, \tau_*)$ будет $\Pi(\tau) \in (\tau_*, \tau_*^+)$, и функция Π монотонна: в этом интервале чем раньше фазовая точка притянулась к равновесию, тем позже она с него сорвется (как и в случаях «тривиального» затягивания). Для момента падения τ_*^- значение функции входа-выхода не определено: фазовые точки, притянувшиеся к равновесию при $\tau \approx \tau_*^-$, срываются с равновесия в различные моменты времени, большие τ_*^+ . (Это становится очевидным, если обратить направление времени: тогда фазовые точки, притянувшиеся к равновесию в разные моменты медленного времени, большие τ_*^+ , должны из-за явления максимального затягивания срывать с равновесия при $\tau \approx \tau_*^-$).

С вещественной точки зрения наличие максимального затягивания объясняется следующим образом. Вся динамика рассматривается при $\tau \in [a, b]$ (см. п. 2). Рассмотрим фазовую точку, находящуюся вблизи кривой равновесий при $\tau = a$. Обозначим Γ_- траекторию этой фазовой точки, а τ_*^+ — соответствующий момент срыва (предполагаем, что $\tau_*^+ < b$). Аналогично, изменив направление времени, рассмотрим фазовую точку, находящуюся вблизи кривой равновесий при $\tau = b$, ее траекторию Γ_+ , и момент срыва $\tau_*^- > a$. Фазовая точка $p(\tau)$, находящаяся вблизи равновесия (и, значит, вблизи Γ_-) при $\tau \in (a, \tau_*^-)$, в дальнейшем, до момента τ_* , успевает притянуться к Γ_- ближе, чем фазовая точка, находящаяся при $\tau = \tau_*^-$ на Γ_+ . Поэтому $p(\tau)$ должна находиться вблизи Γ_- больше, чем точки Γ_+ и, следовательно, оставаясь вблизи Γ_- , уходит от Γ_+ (и, значит, от кривой равновесий) при $\tau \approx \tau_*^+$. Фазовая точка $q(\tau)$, притянувшаяся к равновесию (и, значит, к Γ_-) при $\tau_* \in (\tau_*^-, \tau_*)$, в дальнейшем, до момента τ_* , успевает притянуться к Γ_- меньше, чем точки кривой Γ_+ . Поэтому $q(\tau)$ должна находиться вблизи Γ_- меньше времени, чем точки Γ_+ , и должна уйти от Γ_- (и, значит, от кривой равновесий) до момента τ_*^+ . Чтобы определить значения τ_*^\pm и построить функцию входа-выхода, полезно рассматривать динамику в комплексном времени.

При движении вдоль оси вещественного времени фазовые точки с разными начальными условиями сначала, до момента τ_* , экспоненциально сближаются, а потом расходятся. Приходится следить за компенсацией сжатия и растяжения. На плоскости комплексного времени есть пути, соединяющие точки отрезков оси вещественного времени до и после момента потери устойчивости и такие, что вдоль них одно из собственных чисел равновесия становится чисто мнимым. Вдоль комплексно-сопряженных путей становится чисто мнимым другое собственное число. При движении по таким путям проекция на соответствующий собственный вектор отклонения фазовой точки от равновесия имеет приблизительно постоянную длину (если начальное отклонение достаточно мало). В начале и в конце такого пути расстояние фазовой точки от равновесия приблизительно одно и то же. В силу аналитичности системы результат не должен зависеть от выбора пути: можно было бы двигаться и по вещественной оси. А это и означает затягивание потери устойчивости. Если τ^- и τ^+ — левый и правый концы одного из рассматриваемых путей, то $\tau^+ = \Pi(\tau^-)$. Аккуратные формулировки и доказательства содержатся в [6].

Вблизи точки τ_* рассматриваемое семейство путей является гладким. Но при его продолжении могут возникать особенности. Эти особенности могут быть связаны с наличием особенностей у правых частей системы, с обращением в нуль одного из собственных чисел, с наличием кратных собственных чисел. В частности, если при некотором комплексном значении τ_c медленного времени собственное число равновесия обращается в нуль, то концы проходящего через τ_c пути — это и есть, при достаточно общих предположениях, τ^- и τ_* .

Пример (Л. С. Понтрягин — М. А. Шишкова [2]).

Рассмотрим уравнение

$$\dot{z} = (\tau - i)(z - \tau), \quad \tau = \varepsilon t, \quad (8)$$

получающееся из уравнения (2) при $\gamma = 0$. Введем переменную $\xi = z - \tau$. Получим линейное неоднородное уравнение

$$\dot{\xi} = (\tau - i)\xi - \varepsilon. \quad (9)$$

Общее решение соответствующего однородного уравнения имеет вид

$$\eta = e^{\Psi(\tau)/\varepsilon} c, \quad \Psi(\tau) = (\tau - i)^2 / 2, \quad c = \text{const.}$$

На плоскости комплексных значений τ вдоль любого пути $\{\tau: \text{Re } \Psi(\tau) = \text{const}\}$ будет $|\eta| = \text{const}$. Эти пути — равнобочные гиперболы и пара прямых, их семейство показано на рис. 2.

Рассмотрим отрезок одной из указанных гипербол, соединяющий точки вещественной оси. Пусть σ — вещественный параметр вдоль такого пути, $\tau = \tau(\sigma)$. Обозначим $\vartheta = \sigma / \varepsilon$ время вдоль этого пути. Уравнение (9) перепишется в виде

$$d\xi / d\vartheta = i\omega(\sigma)\xi + \varepsilon h(\sigma),$$

где ω — вещественная, не обращающаяся в нуль функция. Введем $\xi_1 = \xi + i\varepsilon h(\sigma) / \omega(\sigma)$. Получим

$$d\xi_1 / d\vartheta = i\omega(\sigma)\xi_1 + O(\varepsilon^2).$$

Для решения этого уравнения на временах $\vartheta \sim 1/\varepsilon$ величина $|\xi_1|$ постоянна с точностью $O(\varepsilon)$. Поэтому на концах рассматриваемого пути значения $|\xi_1|$ отличаются лишь на $O(\varepsilon)$. Значит, для фазовой точки, моментом падения которой является левый конец пути, моментом срыва будет правый конец пути.

В точке $\tau = i$ собственное число обращается в нуль, и семейство путей имеет особенность. Рассмотрим путь, состоящий из

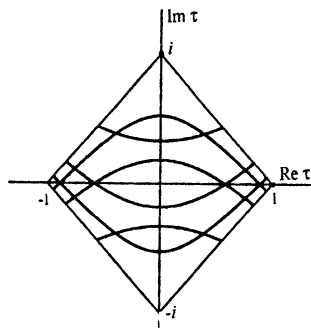


Рис. 2

отрезка, соединяющего точки -1 и i , и отрезка, соединяющего точки i и 1 . Вдоль первого из этих отрезков введем вещественный параметр σ : $\tau = i + \sigma(1 + i)$. Обозначим $\vartheta = \sigma/\varepsilon$. Уравнение (9) переписывается в виде

$$d\xi/d\vartheta = 2i\sigma\xi - \varepsilon(1+i).$$

Для решения этого уравнения величина $|\xi|$ вдали от точки $\sigma = 0$ испытывает лишь колебания порядка ε , но в окрестности точки $\sigma = 0$ меняется на величину порядка $\sqrt{\varepsilon}$. Решение, которое при $\sigma = -1$ (т. е. $\tau = -1$) принимает значение $O(\varepsilon)$, примет при $\sigma = 0$ (т. е. $\tau = i$) значение

$$\begin{aligned} \xi_{-}(i) &= -\int_{-\infty}^0 e^{-i\sigma^2/\varepsilon} (1+i) d\sigma + O(\varepsilon) = \\ &= -\sqrt{\varepsilon} \int_{-\infty}^0 (\cos u^2 - i \sin u^2) (1+i) du + O(\varepsilon) = -\sqrt{\frac{\varepsilon\pi}{2}} + O(\varepsilon). \end{aligned}$$

Аналогично, рассматривая уравнение (9) вдоль отрезка, соединяющего точки i и 1 , получаем, что решение, которое при $\tau = 1$ принимает значение $O(\varepsilon)$, примет при $\tau = i$ значение $\xi_{+}(i) = \sqrt{\varepsilon\pi/2} + O(\varepsilon)$. Таким образом, исходное уравнение (9) имеет два решения, $\xi_{-}(\tau)$ и $\xi_{+}(\tau)$, такие, что $\xi_{\mp}(\mp 1) = O(\varepsilon)$ и $\xi_{\mp}(i) = \mp\sqrt{\varepsilon\pi/2}$. Разность $\xi_{+}(\tau) - \xi_{-}(\tau)$ является решением однородного уравнения, ее модуль постояен вдоль каждого из этих отрезков и равен $|\xi_{+}(i) - \xi_{-}(i)| = \sqrt{2\pi\varepsilon} + O(\varepsilon)$. Поэтому $|\xi_{-}(1)| = \sqrt{2\pi\varepsilon} + O(\varepsilon)$, что гарантирует быстрый рост величины $|\xi_{-}(\tau)|$ в вещественном времени при $\tau > 1$. Поэтому $\tau_{*}^{-} = -1$, $\tau_{*}^{+} = 1$.

Существенное упрощение в примере (8) по сравнению с исходным примером из [2] (см. (2)) состоит в том, что в уравнение (8) не входит величина z . Для анализа уравнения (2) можно переписать его в виде системы уравнений

$$\begin{aligned} \dot{z} &= (\tau - i)(z - \tau) + \gamma(z - \tau)^2(w - \tau), \\ \dot{w} &= (\tau + i)(w - \tau) + \gamma(w - \tau)^2(z - \tau) \end{aligned}$$

с начальным условием $z(\tau_0)$, $w(\tau_0) = \bar{z}(\tau_0)$, где τ_0 вещественно, $\tau_0 < 0$, $z(\tau_0) = \tau_0 + O(\varepsilon)$. Тогда при комплексных τ должно быть $w(\tau) = \bar{z}(\bar{\tau})$. Решение системы надо рассматривать при значениях τ , принадлежащих квадрату с вершинами -1 , i , $+1$, $-i$ (см. рис. 2). Для оценки значения z в точке внутри квадрата можно двигаться сначала из точки τ_0 по дуге гиперболы на рис. 2, а затем вниз по прямой $\operatorname{Re} \tau = \operatorname{const}$. При движении по дуге гиперболы величина $|z - \tau|$ остается приблизительно постоянной, а

при движении по прямой убывает (т. к. при линеаризации правой части уравнения для z около значения $z = \tau$ получается линейное уравнение, для которого собственное число при движении по дуге гиперболы чисто мнимое, а для движения по прямой вниз имеет отрицательную вещественную часть). Значение w оценивается с помощью соотношения $w(\tau) = \bar{z}(\tau)$. В общем случае системы (1) около точки $\tau = \tau_c$, где одно из собственных чисел обращается в нуль, систему можно переписать в форме, где одно из уравнений имеет вид

$$\dot{z} = az^2 + b(\tau - \tau_c) + \dots,$$

а многоточие обозначает члены более высокого порядка. Для положения равновесия имеем $z = Z(\tau) = \sqrt{-(b/a)(\tau - \tau_c)} + \dots$, так что в точке τ_c координаты положения равновесия как функции параметра имеют ветвление. Для уравнения (2) $a = b = 0$, и координаты равновесия являются гладкими функциями τ . Случай $a \neq 0, b \neq 0$ рассмотрен в [17].

Для задачи о затягивании потери устойчивости периодической траектории системы с медленно изменяющимся параметром имеются неулучшаемые оценки максимального затягивания [18]. Вплоть до предполагаемого момента максимального затягивания вычислена функция входа-выхода [18]. Некоторые аспекты затягивания потери устойчивости в этой задаче рассмотрены в [19]. Для задачи о затягивании потери устойчивости неподвижной точки отображения с медленно изменяющимся параметром в некоторых случаях функция входа-выхода вычислена в [20].

4. Численный пример: затягивание потери устойчивости в системе Циглера [21]

Система Циглера [22] — один из модельных примеров в теории неконсервативных механических систем, нагруженных следящими силами. Система изображена на рис. 3.

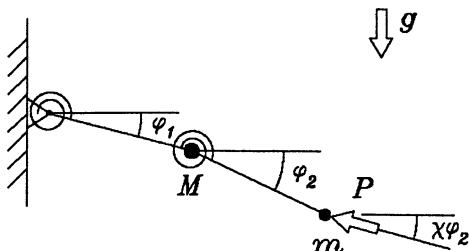


Рис. 3

Она состоит из двух стержней, соединенных друг с другом и со стеной шарнирами так, как показано на рис. 3. Имеется поле тяжести. Система находится в вертикальной плоскости. Стержни считаются невесомыми, но несут на концах массы. Шарниры снабжены пружинами, учитывается трение в шарнирах. К свободному концу второго (не соединенного со стеной) стержня приложена следящая сила, угол которой с горизонталью пропорционален углу отклонения этого стержня от горизонтали. Если величина P этой силы меньше некоторого порога P_* , то система находится в асимптотически устойчивом равновесии. Параметры системы подобраны так, что при увеличении величины следящей силы в системе происходит мягкая бифуркация Пуанкаре — Андронова — Хопфа: пара сопряженных собственных чисел пересекает мнимую ось, равновесие теряет устойчивость, и от него ответвляется устойчивый предельный цикл.

Пусть теперь величина P следящей силы медленно линейно растет со временем, $P = P_0 + \epsilon t$, и проходит через пороговое значение P_* . Поведение угла φ_2 отклонения второго стержня от горизонтали показано на рис. 4.

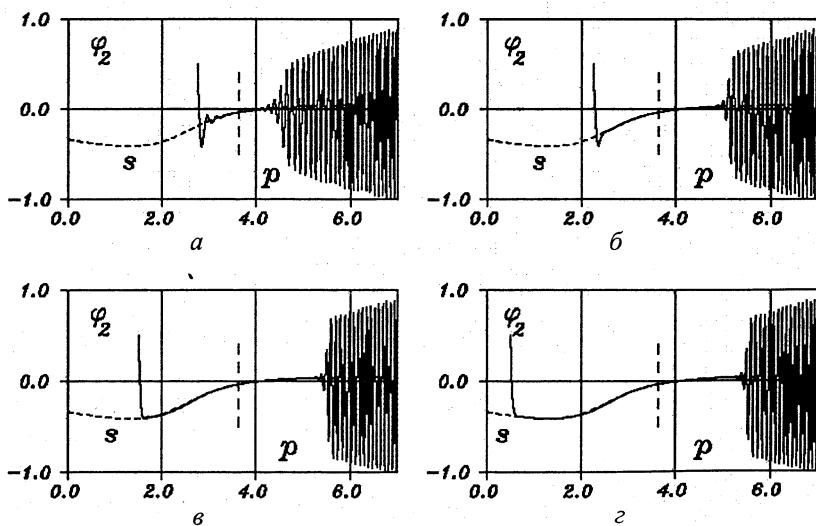


Рис. 4

По вертикальной оси отложено значение φ_2 , по горизонтальной — величина силы P . Пунктирная кривая S отвечает значению φ_2 в положении равновесия. Вертикальная штриховая линия отмечает пороговое значение

следящей силы. Разные рисунки отвечают движениям, начинающимся при разных значениях следящей силы. Фазовая точка притягивается к устойчивому равновесию, отслеживает его, а затем срывается с равновесия, и начинаются колебания, отвечающие движению вблизи предельного цикла. Видно, что потеря устойчивости затягивается: колебания начинаются при значениях следящей силы существенно бóльших, чем «статическое» пороговое значение. На рис. 4, *a*, *b* чем дольше система находилась вблизи устойчивого равновесия, тем позже начинаются колебания. Рисунки *в*, *г* демонстрируют наличие максимального затягивания: время нахождения вблизи устойчивого равновесия увеличивается, а колебания начинаются при одном и том же значении следящей силы.

Рисунок 5 дает объяснение наблюдаемым явлениям с точки зрения теории, обсуждавшейся в п. 3.

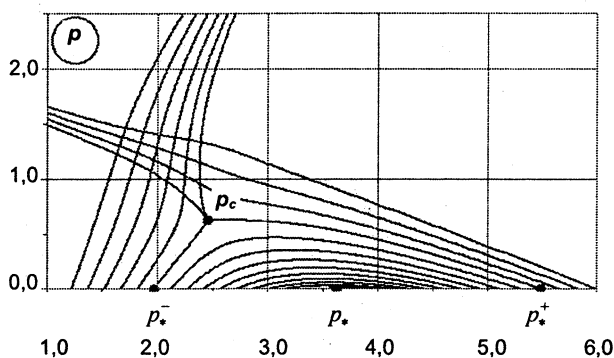


Рис. 5

На рис. 5 изображены линии уровня функции $\text{Re } \Psi(P)$, где $\Psi(P) = \int_{P_c}^P \lambda(\tau) d\tau$, и $\lambda(\tau)$ — собственное число равновесия, отвечающее за поте-

рю устойчивости. Эти линии уровня определяют функцию входа-выхода. Для линии уровня, пересекающей вещественную ось в двух точках, если притяжение к равновесию происходит при значениях силы вблизи левой из этих точек, то срыв с равновесия — вблизи правой. Семейство линий уровня имеет особенность, связанную с тем, что в комплексной точке P_c собственное число λ обращается в нуль. Точки пересечения с вещественной осью особой линии уровня, проходящей через точку P_c , определяют величину максимального затягивания. Если притяжение к равновесию

произошло при значениях силы, меньших значения в левой из этих точек, то срыв с равновесия произойдет при значениях силы, близких к значению в правой из этих точек.

Работа выполнена при поддержке РФФИ (06-01-00117, НШ-13.12.2006.1).

Литература

1. Арнольд В. И. Геометрические методы в теории обыкновенных дифференциальных уравнений. Изд. 2-е. Ижевск: НИЦ РХД, 2000. 400 с.
2. Шишкова М. А. Рассмотрение одной системы дифференциальных уравнений с малым параметром при высших производных // Докл. АН СССР. 1978. Т. 209, № 3. С. 576—579.
3. Каримов С. Асимптотическое поведение решений определенных классов дифференциальных уравнений с малым параметром перед производными в случае смены устойчивости стационарной точки на плоскости «быстрых движений» // Дифференц. уравнения. 1985. Т. 21, № 10. С. 1698—1701.
4. Sibuya Y. Sur réduction analytique d' un système d' equations différentiales ordinares linéaires content un paramètre // J. Fac. Sci. Univ. Tokio. Sect. 1. 1958. V. 10. P. 527—540.
5. Нейштадт А. И. Асимптотическое исследование потери устойчивости равновесия при медленном прохождении пары собственных чисел через мнимую ось // Успехи мат. наук. 1985. Т. 40, вып. 5. С. 300—301.
6. Нейштадт А. И. О затягивании потери устойчивости при динамических бифуркациях, I, II // Дифференц. уравнения. 1987. Т. 23, № 12. С. 2060—2067; 1988. Т. 24, № 2. С. 226—233.
7. Erneux T., Mandel P. Stationary, harmonic and pulsed operations of an optically bistable laser with saturable absorber. II // Phys. Rev. A. 1984. V. 30, № 4. P. 1902—1909.
8. Tsois T. T., Sane R. C., Lindstrom T. H. The bifurcation behaviour of catalytic reaction system due to slowly varying control parameter // AIChE. 1987. V. 34. P. 383—388.
9. Baer S. M., Erneux T., Rinzel J. The slow passage through a Hopf bifurcation: delay, memory effects and resonance // SIAM J. Appl. Math. 1989. V. 49, № 1. P. 55—71.
10. Benoit E. (ed.) Dynamic bifurcations: Lecture Notes in Mathematics. V. 1493. Berlin: Springer, 1991. 219 p.
11. Мищенко Е. Ф., Розов Н. Х. Дифференциальные уравнения с малым параметром и релаксационные колебания. М.: Наука, 1975. 247 с.
12. Понтрягин Л. С., Родыгин Л. В. Приближенное решение одной системы обыкновенных дифференциальных уравнений с малым параметром при производных // Докл. АН СССР. 1960. Т. 131, № 2. С. 255—258.
13. Арнольд В. И. Малые знаменатели и проблемы устойчивости в классической и небесной механике // Успехи мат. наук. 1963. Т. 18, вып. 6. С. 91—112.
14. Звонкин А. К., Шубин М. А. Нестандартный анализ и сингулярные возмущения обыкновенных дифференциальных уравнений // Успехи мат. наук. 1984. Т. 39, вып. 2. С. 77—127.
15. Diener F., Diener M. Maximal delay // Benoit E. (ed.). Dynamic bifurcations: Lecture Notes in Mathematics. V. 1493. Berlin: Springer, 1991. P. 71—86.
16. Neishtadt A. I. On calculation of stability loss delay time for dynamical bifurcations // XI Int. Congress on Mathematical Physics (Paris, 1994). Cambridge, MA: Int. Press, 1995. P. 280—287.

17. *Neishtadt A. I., Simó C., Treschev D. V.* On stability loss delay for a periodic trajectory // *Progress in Nonlinear Differential Equations and Their Applications*. 1996. V. 19. P. 253—278.
18. *Su J.* Effects of periodic forcing on delayed bifurcations // *J. Dynam. Differential Equations*. 1997. V. 9, № 4. P. 561—625.
19. *Fruchard A.* Canards et râteaux // *Annales de l'institut Fourier*. 1992. V. 42, № 4. P. 825—855.
20. *Нейштатдт А. И., Сидоренко В. В.* Запаздывание потери устойчивости в системе Циглера // *Прикл. мат. и мех.* 1997. Т. 61, вып. 1. С. 18—29.
21. *Ziegler H.* Die Stabilitätskriterien der Elastomechanik // *Ing. Archiv*. 1952. V. 20, № 1. S. 49—56.

Нелинейные явления в живых системах

БАКТЕРИАЛЬНАЯ КЛЕТКА С ФИЗИЧЕСКОЙ ТОЧКИ ЗРЕНИЯ

А. А. Веденов

Бактериальная клетка рассматривается как совокупность полимеризующих молекулярных машин: ДНК-полимераз, РНК-полимераз и рибосом. Условие синхронизации машин позволяет оценить число РНК-полимераз и рибосом в фазе экспоненциального роста. Экспериментальные данные о проникновении антибиотика в клетку и гипотеза о существовании только одной мишени у антибиотика позволяют оценить декремент экспоненциального затухания числа бактерий в популяции в растворе с антибиотиком и минимальную ингибирующую концентрацию (МИС) для бактерий дикого типа и мутантов.

Строение бактериальной клетки

Примерные размеры клетки бактерии *Escherichia coli* $1 \times 1 \times 2$ мкм. От внешней среды ее отделяет клеточная стенка — белковая оболочка; внутри находится цитоплазма — водный раствор тысяч низкомолекулярных веществ и полимеров — нуклеиновых кислот, белков и их комплексов. В стенке закреплены насосы, закачивающие в цитоплазму вещества, необходимые для размножения, и насосы, откачивающие вредные или ненужные вещества. В стенке также закреплены моторы (с выступающими наружу длинными винтовыми жгутиками), приводящие бактериальную клетку в движение. У некоторых бактерий (микоплазм, хламидий) клеточной стенки нет и цитоплазма отделена от внешней среды мембраной; они могут плавать или ползать по поверхности [1]. Пример бактериальной клетки приведен на рис. 1 [2].

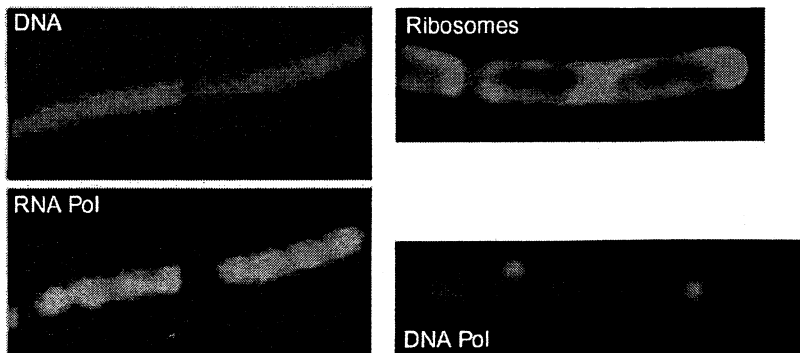


Рис. 1. Бактериальная клетка [2]

Полностью первые геномы бактерий были расшифрованы в середине 1990-х гг. Самый короткий расшифрованный геном у бактерии *Mycoplasma genitalium* (рис. 2) составляет 470 генов [3]. Самый длинный из расшифрованных бактериальный геном — 7 825 генов — у почвенной нитевидной бактерии *Streptomyces coelicolor*, производящей значительную часть натуральных антибиотиков, используемых в медицине и ветеринарии [4].

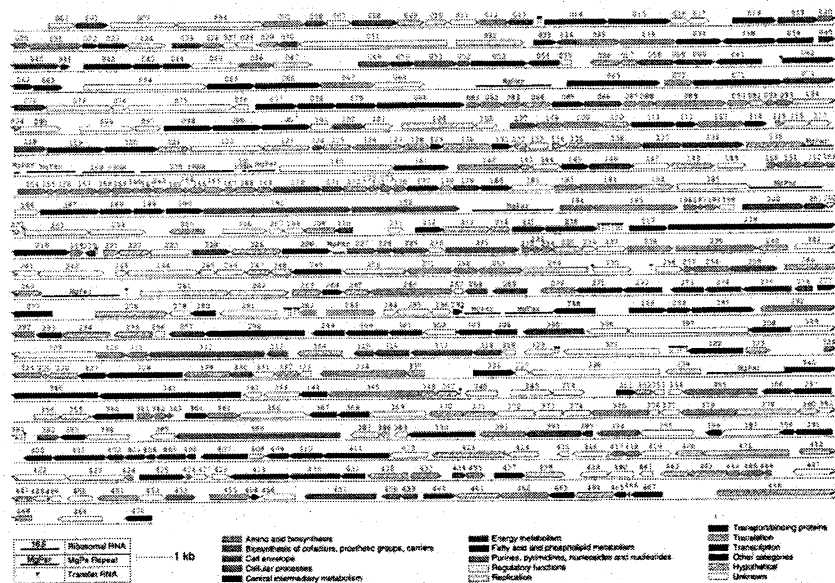


Рис. 2. Расположение генов ДНК-полимеразы, РНК-полимеразы, рибосомальной РНК и белков в геноме *Mycoplasma genitalium* [3].

Число молекул различных белков в бактериальной клетке меняется в очень широких пределах — от единиц до сотен тысяч, 99% массы бактериальной клетки составляют наиболее распространенные 100—150 белков.

Бактерии крайне сложны и разнообразны. Основное, что объединяет бактерии — это способность размножаться делением одной материнской клетки (после роста, продолжающегося некоторое время — часы или сутки) на две одинаковые дочерние клетки.

Во всех бактериях есть линейные полимеры (ДНК, РНК, белки) и молекулярные машины, осуществляющие полимеризацию, необходимую для роста и размножения бактерий. Это ДНК-полимеразы, РНК-полимеразы, рибосомы. Нуклеиновые кислоты полимеризуются из нуклеоти-

дов, белки — из аминокислот. В течение последнего десятилетия в экспериментах с отдельными молекулярными машинами были найдены их скорости при различных значениях расхода АТФ. Например, для *E. coli* скорость ДНК-полимеразы составляет $v_1 = 800$ нуклеотидов в секунду, скорость РНК-полимеразы $v_2 = 50$ нуклеотидов в секунду, скорость рибосомы $v_3 = 16$ аминокислот в секунду.

В жизни бактериальных популяций наблюдаются различные состояния. Мы рассмотрим основные — фазу размножения (экспоненциального роста популяции) и фазу гибели (экспоненциального спада числа жизнеспособных бактерий).

Экспоненциальный рост популяции

Для каждой питательной среды, согласно экспериментальным данным, у каждого штамма бактерий имеется определенное количество молекулярных машин каждого сорта. Например, для богатой питательной среды в каждой бактериальной клетке *E. coli* имеются несколько ДНК-полимераз, тысячи РНК-полимераз, десятки тысяч рибосом.

Эти количества таковы, что исходное число машин (построенных из белков и рибонуклеиновой кислоты) каждого сорта удваивается за одно и то же время:

$$N = N_0 \cdot 2^{(t/\tau)},$$

где τ — период удвоения бактериальной культуры, находящейся в фазе экспоненциального роста (рис. 3 [5]), t — время.

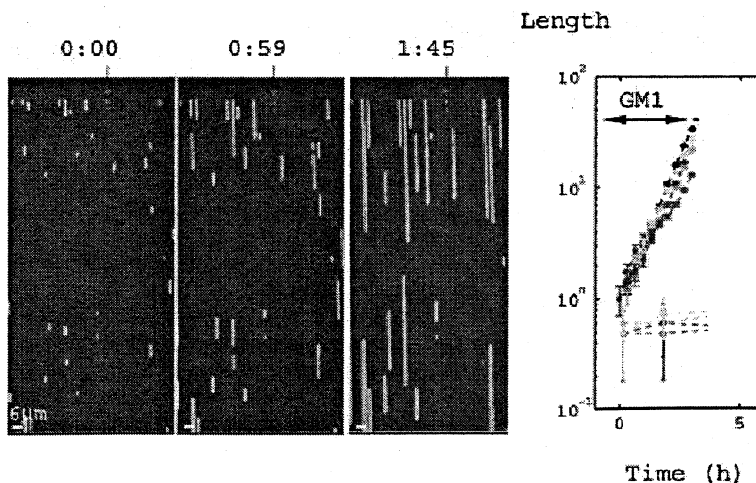


Рис. 3. Экспоненциальный рост популяции бактерий в одномерных желобках шириной в несколько микрометров [5].

Удвоение бактериальной клетки (после которого происходит деление) включает идущие одновременно процессы репликации, транскрипции и трансляции.

Геном у бактерий — это молекула ДНК, которая состоит из пар комплементарных (дополнительных) нуклеиновых оснований АТ (аденин-тимин) и ГЦ (гуанин-цитозин) и образует двухнитевую спираль диаметром 10 ангстрем и длиной в сотни микрон.

При репликации ДНК-полимераза пропускает через себя одну нить ДНК (называемую матричной) и полимеризует из диффундирующих из окружающего раствора оснований А, Т, Г, Ц комплементарную нить, т. е. такую, что против этих нуклеотидов выстраиваются соответственно Т, А, Ц, Г. ДНК-полимеразу с одной входящей и двумя выходящими нитями ДНК называют репликативной вилкой.

Транскрипция — это полимеризация транскриптов (РНК-копий одиночных генов или нескольких идущих подряд генов) с помощью РНК-полимеразы. Каждый транскрипт является копией матрицы — одной из двух нитей ДНК, составляющих данный ген. В сумме транскрипты составляют так называемую матричную РНК, или мРНК.

Трансляция — это создание белковой копии транскрипта. В этом процессе каждый белок полимеризуется рибосомой из аминокислот, приносимых к рибосоме транспортными РНК. Это происходит в процессе копирования каждого транскрипта мРНК по определенным правилам: каждой тройке нуклеотидов соответствует определенная аминокислота.

На каждом гене находятся и движутся вдоль двухнитевой ДНК несколько РНК-полимераз, от каждой из них отходит один растущий однонитевой РНК-транскрипт, вдоль которого от внешнего конца движутся несколько рибосом, и от каждой из них отходит нить белка (рис. 4).

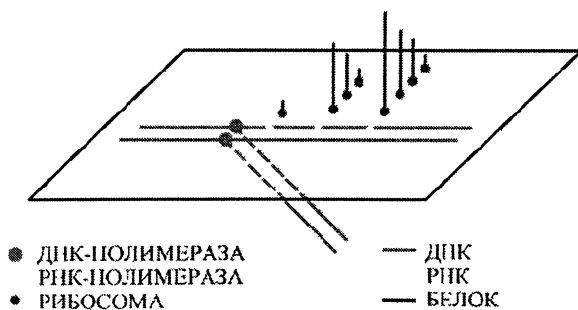


Рис. 4. Схема расположения полимеров и полимеризующих машин в растущей клетке

В действительности синтез ДНК, РНК, белка и других веществ в растущей бактериальной клетке неравномерен во времени. В частности, име-

ет место модуляция скорости полимеризации, связанная с влиянием последовательно появляющихся регуляторных белков и РНК.

В целом рост бактерии и отделение дочерней клетки от материнской внешне похожи на вытекание последовательности жидких капель из крана — их размеры и время между отделением капель зависит от свойств жидкости (в частности, от химического состава, определяющего плотность и поверхностное натяжение), и сам процесс при изменении условий может превратиться из регулярного в хаотический.

К моменту отделения дочерней клетки все полимерные молекулы, за исключением ДНК и ДНК-полимераз, присутствуют в цитоплазме в виде большого количества (от сотен до сотен тысяч) отдельных молекул — полимерных клубков или комплексов из нескольких молекул (рибосомы, РНК-полимеразы). Поэтому примерно половина молекул и комплексов каждого сорта оказывается в отделяющейся дочерней клетке, обеспечивая ее идентичность материнской бактериальной клетке.

В экспериментах *in vitro* процессы репликации, транскрипции и трансляции могут идти отдельно (см., например, [6, 7]).

Оценка числа основных полимеризующих машин в *E. Coli*

Проведем оценку числа РНК-полимераз и рибосом в клетке *E. coli*.

Каждая рибосома состоит из двух субъединиц, и каждая субъединица состоит из рибосомальной РНК (рРНК) и нескольких десятков белков. Ген рРНК длиной $l_{31} = 3000$ нуклеотидов в бактерии *E. coli* представлен 7 копиями. Рассмотрим экспоненциально растущую культуру бактерий в такой среде, когда время удвоения равно $\tau = 40$ мин и бактерии имеют $n = 2$ генома на клетку [8].

Число x_{31} транскриптов генов рРНК (первый индекс означает: 1 — ДНК, 2 — РНК, 3 — белок и рибосомы) есть произведение числа этих генов $7 \times n$ на число РНК-полимераз на этом гене l_{31}/d_2 и на число циклов прохождения гена этими полимеразами за время удвоения $\tau/(l_{31}/v_2)$

$$x_{31} = 7 \cdot n \cdot (l_{31}/d_2) \cdot \tau / (l_{31}/v_2).$$

При скорости $v_2 = 50$ нуклеотидов в секунду, расстоянии между центрами соседних РНК-полимераз (на матричной ДНК) порядка их диаметра $d_2 = 80$ нуклеотидов и $\tau = 40$ мин = 2400 с получим число транскриптов рРНК:

$$x_{31} = 7 \cdot n \cdot v_2 \cdot \tau / d_2 = 7 \cdot 2 \cdot 50 \cdot 2400 / 80 = 21000. \quad (1)$$

20 белков малой рибосомальной субъединицы 30 S кодируются (файл 1GIX из PDB-банка) суммарной последовательностью 7200 нуклеотидов, т. е. средняя длина одного транскрипта $l_{32} = 7200/20 = 360$ нуклеотидов. Число копий каждого из этих транскриптов равно

$$n \cdot v_2 \cdot \tau / d_2 = 2 \cdot 50 \cdot 2400 / 80 = 3000.$$

Пусть расстояние между центрами соседних рибосом на транскрипте порядка диаметра рибосомы $d_3 = 80$ нуклеотидов. Если трансляция с данного транскрипта продолжается только пока транскрипт связан с РНК-полимеразой, то с каждого транскрипта делается $l_{32}/d_3 = 4,5$ белковых копии. При этом число молекул каждого белка рибосомальной субъединицы (и, следовательно, число рибосом) равно

$$x_{32} = n \cdot v_2 \cdot \tau / d_2 \cdot l_{32} / d_3 = 3000 \cdot 4,5 = 13500, \quad (2)$$

т. е. несколько меньше, чем число транскриптов рРНК (1). Экспериментально найдено $x_3 = 26000$.

Число РНК-полимераз на кодирующей рРНК части ДНК равно

$$x_{21} = 7 \cdot n \cdot (l_{31} / d_2) = 7 \cdot 2 \cdot 3000 / 80 = 525.$$

Число РНК-полимераз на кодирующей рибосомальные белки части ДНК равно

$$x_{22} = n \cdot 7200 / 80 = 180. \quad (3)$$

Экспериментально найдено $x_{21} + x_{22} = 1000$.

Если две репликативные вилки одновременно продвигаются от исходной точки репликации, то время синтеза ДНК составляет

$$(4 \ 700 \ 000 \text{ нуклеотидов} / 2) / 800 \text{ нуклеотидов} / \text{с} = 2 \ 800 \text{ с} = 47 \text{ мин}. \quad (4)$$

Для почвенной бактерии *S. coelicolor* при длине генома 8 000 000 на богатой среде $\tau = (8 \ 000 \ 000 / 2) / 800 = 5 \ 000 \text{ с}$. При $d_2 = 100$ нуклеотидов, 6 копиях гена рРНК и $n = 1,5$ генома на клетку получаем $x_{31} = 6 \cdot n \cdot v_2 \cdot \tau / d_2 = 6 \cdot 1,5 \cdot 50 \cdot 10 \ 000 / 100 = 45 \ 000$ рибосом при экспериментальном значении 30 000.

Фаза гибели популяции

Молекулы антибиотика, попавшие в бактериальную клетку из внешней среды, связываются с полимеризующими молекулярными машинами и замедляют или останавливают их работу. В результате бактерии размножаются медленнее или начинают погибать (рис. 5).

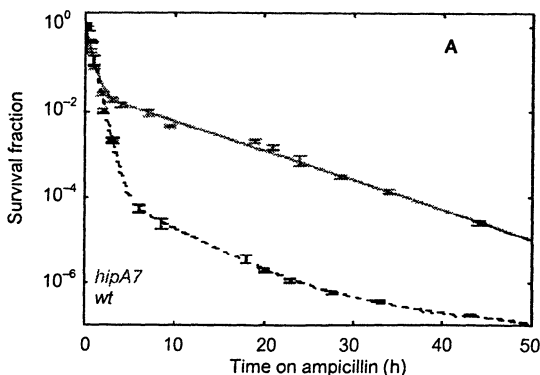


Рис. 5. Экспоненциальное падение численности популяции во времени в среде с антибиотиком [5].

В качестве примера рассмотрим антибиотик рифампицин. Считается, что его молекулы селективно ингибируют РНК-полимеразу, нековалентно связываясь с ней.

Измерения процесса проникновения рифампицина из раствора в бактериальную клетку *E. coli* показали, что эта аккумуляция протекает за $\tau_{rif} = 40$ с и насыщается при концентрации рифампицина $c = 2$ мкг/мл (рис. 6). При этом концентрация рифампицина в клетке составляет $c_{rif} = 3$ микрограмм на грамм сухого веса [9].

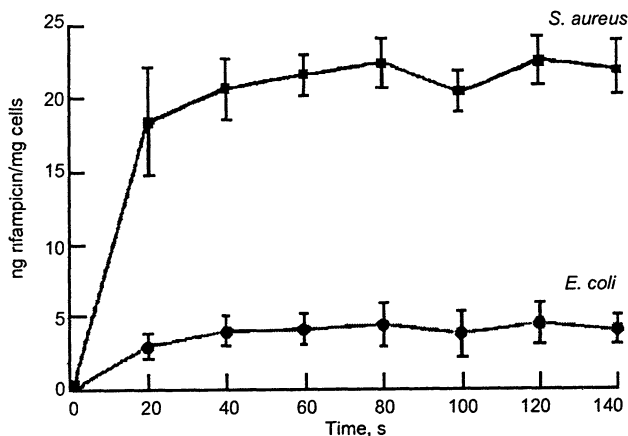


Рис. 6. Накопление рифампицина в бактериальных клетках [9]

Принимая

сухой вес равным 0,2 полного веса,

плотность бактерии равной 1 г/см^3 ,

молекулярный вес рифампицина равным 600 Дальтон,

объем бактериальной клетки $v_{cell} = 2 \text{ мкм}^3$,

получаем:

масса рифампицина в клетке равна $3 \cdot 10^{-6} \cdot 0,2 \cdot 2 \cdot 10^{-12} \text{ г} = 12 \cdot 10^{-19} \text{ г}$,

масса молекулы рифампицина равна $600 \cdot 1,6 \cdot 10^{-24} \text{ г} = 10^{-21} \text{ г}$,

т. е. в клетке в объеме 2 мкм^3 находится $N_{rif} = 1200$ молекул рифампицина.

При этом в растворе находится $c = 2 \text{ мкг/см}^3$ рифампицина, т. е. в объеме 2 мкм^3 раствора находится $2 \cdot 10^{-12} \cdot 2 \cdot 10^{-6} / 10^{-21} = 4 \text{ 000}$ молекул рифампицина, так что отношение концентраций антибиотика в клетке и в растворе равно $\eta = 0,3$.

Можно ожидать, что взаимодействие $x_{21} + x_{22}$ РНК-полимераз, движущихся вдоль нити ДНК, с N_{rif} молекулами рифампицина будет приводить с некоторой вероятностью к торможению полимеризации транскриптов и прекращению процесса роста и деления для каждой бактерии. С увеличением N_{rif} вероятность должна возрастать и при $x_{21} + x_{22} = N_{rif}$

должна быть порядка единицы; для оценки примем, что эта вероятность линейно зависит от N_{rif} .

Пусть полное число размножающихся бактерий равно n . Тогда число бактерий, перестающих размножаться за 1 с, можно оценить так:

$$dn/dt = -\gamma n, \quad (5)$$

где $\gamma = (1/\tau_{rif}) N_{rif}/(x_{21} + x_{22})$, или

$$\gamma = (1/\tau_{rif}) (\eta \cdot c \cdot v_{cell}) / (x_{21} + x_{22}). \quad (6)$$

С учетом размножения бактерий

$$dn/dt = (-\gamma + 1/\tau)n,$$

где τ — время удвоения в отсутствие антибиотика. Эксперимент (рис. 7) качественно правильно описывается формулами (5), (6).

Концентрация антибиотика, при которой рост популяции сменяется падением, называется минимальной ингибирующей концентрацией; она определяется условием $-\gamma + 1/\tau = 0$, т. е.

$$c_{MIC} = (\tau_{rif}/\tau) \cdot (x_{21} + x_{22}) / (\eta \cdot v_{cell}),$$

что качественно согласуется с экспериментальным результатом для рифампицина на рис. 7.

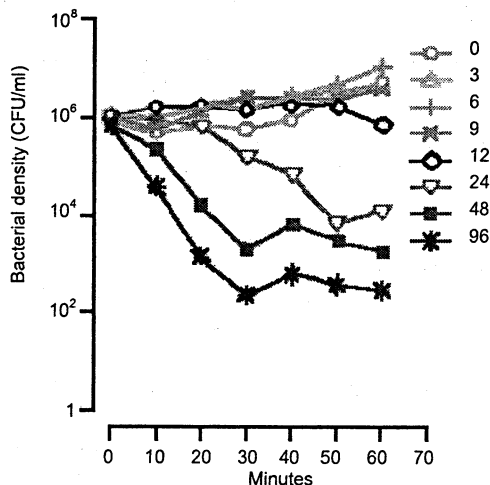


Рис. 7. Падение численности популяции *E. coli* во времени при разных концентрациях рифампицина, мкг/мл [10].

Мутации

Одна из проблем современной медицины — это возникающая и увеличивающаяся устойчивость бактерий к антибиотикам, применяемым для

лечения. Она возникает из-за мутаций (т. е. химических замен в молекуле ДНК) у исходных бактерий (так называемого «дикого типа»).

Мутации в ДНК приводят к тому, что у потомков размножающейся бактериальной клетки молекулярные машины слегка изменяют свой химический состав и форму, поэтому молекула антибиотика слабее связывается с такой машиной и меньше мешает ее работе. В результате мутировавшие бактерии продолжают размножаться (т. е. они оказываются устойчивыми к антибиотику), и чтобы антибиотик действовал, приходится увеличивать его концентрацию. Например, для *E. coli* MIC некоторых мутантов (с заменой аминокислоты в РНК-полимеразе) составляют [11]:

МИС, мкг/мл	Замена аминокислоты	Изменение объема аминокислоты	Изменение типа аминокислоты
0—12,5	дикий тип		
100—200	I572L	124 → 124	гидрофобная → гидрофобная
25—50	L511Q	131 → 114	гидрофобная → гидрофильная
100—200	H526L	118 → 124	заряженная(+) → гидрофобная
200—400	H526Y	118 → 141	заряженная(+) → гидрофобная
50—400	Q148L	114 → 124	гидрофильная → гидрофобная
3000—4000	P564L	90 → 124	гидрофобная → гидрофобная
100—200	D516G	91 → 48	заряженная (-) → гидрофобная

При замене аминокислоты изменяется ее объем и тип, что вызывает деформацию РНК-полимеразы. Это приводит к двум следствиям. Во-первых, на некоторую величину $\Delta E > 0$ уменьшается энергия связи РНК-полимеразы с антибиотиком. Вероятность образования комплекса РНК-полимеразы с антибиотиком пропорциональна бoльцмановскому фактору

$$\xi = \exp(-\Delta E / T), \quad (7)$$

поэтому в формуле (6) нужно ввести поправочный множитель ξ , учитывающий уменьшение этой вероятности

$$\gamma = (1/\tau_{rif}) (\xi \cdot \eta \cdot c \cdot v_{cell}) / (x_{21} + x_{22}). \quad (8)$$

Поскольку МИС определяется равенством $\gamma = 1/\tau$, из (7), (8) получаем

$$\Delta E = T \cdot \ln [(\tau/\tau_{rif}) (\eta \cdot c_{MIC} \cdot v_{cell}) / (x_{21} + x_{22})]. \quad (9)$$

Если величина ΔE мала, то она пропорциональна деформации, так что логарифм c_{MIC} линейно зависит от деформации РНК-полимеразы у мутанта.

Во-вторых, в отсутствие антибиотика РНК-полимеразы мутанта, как правило, будет работать хуже, чем у дикого типа. Из физических экспериментов с отдельными РНК-полимеразами известно, что скорость РНК-полимеразы падает с увеличением силы F , препятствующей протягиванию полимерной нити. Скорость обращается в нуль, если сила F превы-

шает величину E_0/δ , где E_0 — энергия АТФ, которая тратится на полимеризацию одного звена цепи, а δ — происходящее при этом перемещение цепи относительно РНК-полимеразы (эта оценка справедлива, если эффективность работы машины велика) [12] (рис. 8).

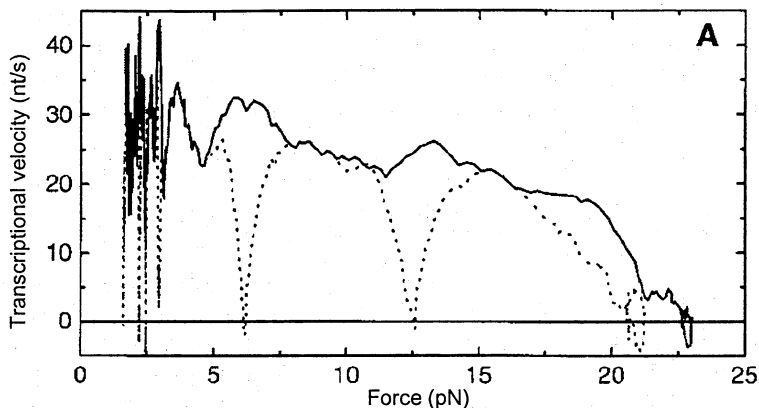


Рис. 8. Зависимость скорость — сила для РНК-полимеразы [12]

При малых же значениях F скорость определяется скоростью диффузии поступающих из раствора звеньев полимерной цепи. Согласно [13] скорость РНК-полимеразы равна $v = v_0 \cdot n_r / (n_r + n_0)$, где n_r — концентрация рибонуклеотидов в растворе, $v_0 = 22$ нуклеотида/с, $n_0 = 250$ мкмоль/л.

Вследствие деформации РНК-полимеразы у мутанта может либо уменьшиться скорость диффузии новых звеньев полимеризующейся цепи, либо уменьшиться энергия их связи с РНК-полимеразой. В результате должна упасть скорость роста бактерии и должно увеличиться время удвоения популяции. Этот вывод согласуется с опытом, где наблюдается линейное падение инкремента роста популяции $1/\tau$ с ростом $\ln c_{\text{МИС}}$ [14]:

$$1/\tau = (1/\tau_0) \cdot (1 - 0,08 \cdot \log(c_{\text{МИС}}/0,01)),$$

τ_0 — время удвоения дикого типа (для него $c_{\text{МИС}} = 0,01$ мкг/мл).

Для преодоления лекарственной устойчивости мутантов предлагается применить гибкие молекулы (рис. 9), которые смогут адаптироваться к форме РНК-полимеразы [15], или использовать лекарства, имеющие одновременно много мишеней в клетке [16]. Возможно, однако, что при этом будет теряться селективность действия антибиотика, которую необходимо сохранять для уменьшения побочных действий лекарства.

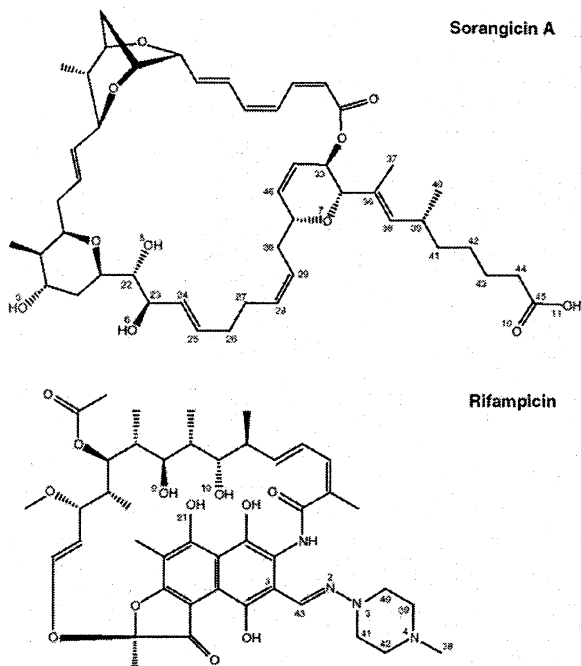


Рис. 9. Структура гибкого и жесткого ингибиторов РНК-полимеразы [15]

Заключение

Геномы, полимеразы и рибосомы разных бактерий различны. Число молекулярных машин и время удвоения популяции зависят от среды обитания бактерии. Машины в бактериальной клетке взаимодействуют между собой [17]. Лекарства, как правило, имеют в бактериальной клетке не одну мишень [14, 16]. Поэтому приведенные оценки носят лишь качественный характер.

В заключение хочу поблагодарить за полезные обсуждения И. Б. Стечкину и А. Л. Чернякова.

Литература

1. *Wolgemuth C. W., Igoshin O., Oster G.* The Motility of Mollicutes // *Biophys. J.* 2003. V. 85. P. 828.
2. *Errington J.* Dynamic proteins and a cytoskeleton in bacteria // *Nature cell biology.* 2003. V. 5. P. 175.
3. *Fraser C. M. et al.* The minimal gene complement of *Mycoplasma genitalium* // *Science.* 1995 V. 270. P. 397.

4. *Bentley S. D. et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2) // *Nature*. 2002. V. 417. P. 141.
5. *Miller C. et al.* SOS Response Induction by β -Lactams and Bacterial Defense Against Antibiotic Lethality // *Science*. 2004. V. 305. P. 1629.
6. *Braun D., Goddard N. L., Libchaber A.* Exponential DNA Replication by Laminar Convection // *PRL*. 2003. V. 91. P. 158103-1.
7. *Bulleid N. J., Freedman R. B.* The transcription and translation in vitro of individual cereal storage-protein genes from wheat // *Biochem. J.* 1988. V. 254. P. 805.
8. *Bremer H., Dennis P. P.* Modulation of chemical composition and other parameters of the cell by growth rate // *Escherichia coli* and *Salmonella*: cellular and molecular biology. 2nd ed. V. 2. Washington, D.C.: ASM Press, 1996. P. 1553—1569.
9. *Williams K. J., Piddock L. J. V.* Accumulation of rifampicin by *Escherichia coli* and *Staphylococcus aureus* // *J. antimicrobial chemotherapy*. 1998. V. 42. P. 597.
10. *Regoes R. R. et al.* Pharmacodynamic Functions: a Multiparameter Approach to the Design of Antibiotic Treatment Regimens // *Antimicrobial agents and chemotherapy*. 2004. V. 48, № 10. P. 3670—3676.
11. *Reynolds M. G.* Compensatory Evolution in Rifampin-Resistant *Escherichia coli* // *Genetics*. 2000. V. 156. P. 1471.
12. *Wang M. D. et al.* Force and Velocity Measured for Single Molecules of RNA Polymerase // *Science*. 1998. V. 282. P. 902.
13. *Harada Y. et al.* Direct observation of DNA rotation during transcription By *Escherichia coli* RNA polymerase // *Nature*. 2001. V. 409. P. 113.
14. *O'Neill A. J. et al.* Molecular Genetic and Structural Modeling Studies of *Staphylococcus aureus* RNA Polymerase and the Fitness of Rifampin Resistance Genotypes in Relation to Clinical Prevalence // *Antimicrobial agents and chemotherapy*. 2006. V. 50, № 1. P. 298—309.
15. *Campbell E. A. et al.* Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase // *The EMBO J.* 2005. V. 24. P. 674—682.
16. *Csermely P., Agoston V., Pongor S.* The efficiency of multi-target drugs: the network approach might help drug design // *Trends in pharmacological sciences*. 2005. V. 26. P. 178—182; www.arxiv.org/q-bio.MN/0412045.
17. *Epshtein V., Nudler E.* Cooperation Between RNA Polymerase Molecules in Transcription Elongation // *Science*. 2003. V. 300. P. 801.

ГЕНОМ КАК ФРАКТАЛЬНОЕ МНОЖЕСТВО, ПОРОЖДАЕМОЕ ИНВЕРСИОННО-ДИФФУЗИОННОЙ ДИНАМИКОЙ

*С. А. Ларионов, А. Ю. Лоскутов,
С. Д. Рыбалко, Е. В. Рядченко*

1. Введение

Как известно, ДНК представляет собой макромолекулярное образование в форме двойной спирали, состоящей из двух цепей нуклеотидов, соединенных между собой водородными связями. Нуклеотиды — это низкомолекулярные соединения, включающие гетероциклическое кольцо из атомов углерода и азота (азотистое основание), пятиуглеродное сахарное кольцо (пентозу) и фосфатную группу. Азотистые основания присутствуют в двух типах — пурины и пиримидины. Пиримидины имеют одно шестичленное кольцо, а пурины — два конденсированных кольца: одно из них пятичленное, другое — шестичленное.

В ДНК присутствуют два пурина — аденин (*A*) и гуанин (*G*) — и два пиримидина — цитозин (*C*) и тимин (*T*). Азотистые основания соединены с пентозным кольцом гликозидной связью. В пентозном кольце положения углеводов нумеруются некоторым стандартным способом. В нуклеотидной цепи атом в положении 5' одного пентозного кольца соединен с атомом в положении 3' следующего пентозного кольца через фосфатную группу, образуя сахарофосфатный остов ДНК, состоящий из 5' — 3' связей. На концах цепи, следовательно, присутствуют 5' свободная группа и 3' свободная группа, что обычно употребляется для указания направления цепи. Таким образом, нуклеотиды соединяются друг с другом в длинную полинуклеотидную цепь, которую обычно представляют в виде последовательности *ATTGCCAA...*

Посредством водородных связей ветви спирали ДНК соединяются между собой так, что более крупное пуриновое основание — аденин — в одной цепи имеет обычно своим партнером в другой цепи меньшее по размерам пиримидиновое основание — тимин. Соответственно, больший по размерам гуанин соединяется с цитозином. При этом пара *A — T* имеет две водородные связи, а *G — C* три. Стало быть, суммарное число пуринов в одной ветви равно сумме пиримидинов в другой. Состав цепей ДНК характеризуется отношением $(G + C)/(A + T)$, варьирующимся от 26 до 74%, характерным для каждого вида. Соответствие *A — T* и *G — C* называют правилом комплементарности,

а сами цепи — комплементарными. При такой организации, вследствие возможной нерегулярности последовательности, одна из цепей может быть более легкой и, кроме того, различные участки цепи могут иметь различную силу связи с комплементарной цепью за счет нерегулярного количества водородных связей. Это может отражаться на функциональных свойствах.

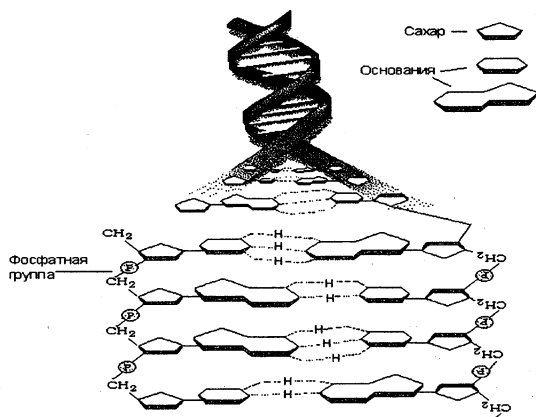


Рис. 1. Двойная спираль ДНК

Среди факторов, существенно влияющих на функциональность ДНК, также обычно учитывают энергию стэкинга, т. е. вертикальные гидрофобные взаимодействия между парами нуклеотидов, уложенных в спирали один над другим в виде стопки, и взаимодействия с раствором. При изменении внешних условий (температуры и уровня pH) данные подсистемы существенно определяют конформацию спирали. Фосфатные группы находятся на внешней стороне спирали, а азотистые основания — внутри, так что их плоскости перпендикулярны оси спирали (рис. 1). Спирали в ДНК представляются несколькими формами (*A*, *B*, *Z*), допускающими дополнительную внутреннюю классификацию, образуя, таким образом, семейства. Они задаются двумя основными параметрами — числом пар нуклеотидов на виток и расстоянием между соседними повторяющимися элементами. По числу пар на шаг спирали ДНК варьируется от 10,2 в *B*-форме до 12 в *Z*-форме. Стандартная форма спирали (*B*-форма) наиболее часто присутствует при организации ДНК. При этом возможны динамические переходы в *A* и *Z* формы в зависимости от внешних условий.

Кроме вирусов, где ДНК находится лишь в покрывающей ее белковой оболочке, обычным местом присутствия ДНК являются клетки

живых организмов. Клетки по характеру структурной организации разделяют на два основных класса — прокариоты и эукариоты. В прокариотах — по сути клетках бактерий — одна цепь ДНК в кольцевой организации находится в сложном образовании — нуклеоиде, которое не отделено мембраной от остальных составляющих элементов клетки. В эукариотах набор цепей ДНК находится соответственно в целом множестве иерархически организованных сложных комплексов — хромосомах, которые, в свою очередь, локализованы в ядре клетки. Кроме того, ДНК в одноцепочечном виде присутствует в митохондриях и хлоропластах (аналог митохондрий для растительных клеток) — одном из видов органелл эукариотов, отвечающих за преобразование энергии в необходимые формы для работы клетки.

Термин *геном* применяется к полной последовательности либо ко всей совокупности последовательностей ДНК с учетом их локализации и функций. Так, для эукариотов различают геномы митохондрий и ядерный геном. Ядерный геном состоит из набора ДНК, заключенного хромосомах. По скорости эволюции геномы ядра и митохондрий существенно отличаются, но по функциям в клетках они пересекаются в зависимости от организма. Полная длина ДНК геномов колеблется от $5 \cdot 10^3$ — $5 \cdot 10^5$ пар нуклеотидов для вирусов и до 2 — $3 \cdot 10^9$ пар нуклеотидов для млекопитающих.

Как известно, основным свойством ДНК является хранение, обработка и воспроизводство информации, заложенной в ее последовательности, а также эволюционная адаптация к изменяемой внешней среде. При всем при этом данные процессы должны происходить на основе информации, хранящейся в той же последовательности, которая накладывает специфические ограничения на организацию таких последовательностей.

В ДНК весьма условно можно различить три класса последовательностей: высоко повторяющуюся, умеренно повторяющуюся и уникальную.

К первому классу часто относят сателлитную ДНК, представляющую собой непрерывную последовательность тандемных повторов размером от двух до нескольких тысяч и повторяющуюся соответственно от нескольких раз для длинных повторов до нескольких тысяч для коротких. Эти последовательности часто однозначно локализованы в определенных частях хромосом, в частности в теломерах и центромерах — областях, отвечающих за структурную организацию хромосомы в целом. К классу умеренно повторяющихся можно отнести различные типы тандемных повторов, диспергированно распределенные по ДНК несколькими видами. Здесь также весьма условно можно выделить

частые, редкие, смешанные, сгруппированные и смешанные, гнездовые. Данные типы последовательностей образуют семейства. Обычно при этом к коротким последовательностям относят 100—500 п. н. (пар нуклеотидов) и к длинным — от $5 \cdot 10^3$ и более. В отдельных геномах наблюдается до 100 семейств размером от 100 до 10^4 элементов, распределенных по геномной ДНК. В ряде случаев они образуют целые кластеры из десятков семейств, где могут находиться одно внутри другого в одной последовательности, общая длина которой может достигать порядка $15 \cdot 10^3$ п. н. Рассеянные повторы могут находиться как в интронах, так и между генами и внутри сателлитных последовательностей. Среди них наблюдаются также истинные палиндромы, т. е. участки, которые «читаются» в любую сторону одинаково, и комплементарные палиндромы, т. е. имеющие продолжение в виде комплементарной цепи. Семейства повторов в сумме могут занимать более 20% всей геномной последовательности. Гены и псевдогены (дефектные, некодирующие гены) также могут иметь тандемные повторы, образовывать кластеры и иметь длины непрерывных последовательностей, достигающие $2-6 \cdot 10^4$ п. н. В данной работе представлен метод, который позволяет эффективно отображать *всю хромосому*, даже если в ней содержатся десятки и сотни миллионов нуклеотидов, легко находить похожие и подобные участки, идентифицировать функциональные и структурные элементы, а также определять фрактальные свойства некоторых частей последовательности ДНК. Он основан на представлении ДНК как блуждания частицы на плоскости [1]. Данное представление восходит к работе [2] (см. также [3—6]). Метод позволяет также по-новому подойти к описанию эволюции и моделированию генома (см. ниже).

2. ДНК как двумерное блуждание

Представим нуклеотидную цепь A, T, G, C на квадратной решетке следующим образом: если встретим аденин (A), то сделаем шаг вправо, если тимин (T) — то влево, если гуанин (G), то вверх, и если цитозин (C), то вниз. Обозначим эти координаты $AGTC$ (двигаясь против часовой стрелки от оси X). Тогда исходная последовательность нуклеотидов будет соответствовать определенной *траектории блуждания* на плоскости $AGTC$. Очевидно, такое отображение ДНК складывается из двух несводимых друг к другу последовательностей $A - T$ и $G - C$, которые можно рассматривать отдельно друг от друга.

Легко видеть, что существует большое количество последовательностей, которые можно рассматривать как блуждание на плоскости,

и изучать возникающие при этом фрактальные структуры и т. д. (см., например, [7, 8] и приведенные там ссылки). Однако применительно к ДНК, начиная только с достаточно большого числа элементов, появляется возможность выявлять функциональные и структурные участки хромосомы по характерному «рисунку» блуждания. При этом число элементов в последовательности нуклеотидов может быть порядка длины последовательности всей хромосомы.

Современные компьютерные методы позволяют автоматизировать обработку данным способом и сравнивать как участки хромосом различных организмов, так и хромосомы в целом. Такое сравнение, в принципе, можно выполнить и методом выравнивания. И хотя этот метод является достаточно эффективным, выделение схожих участков часто оказывается невозможным при анализе последовательностей, состоящих из большого (порядка 10^5 — 10^6) числа элементов.

На сегодняшний день полностью расшифрованы последовательности ДНК хромосом уже достаточно большого числа живых организмов. К ним относятся дрожжевые клетки *Saccharomyces cerevisiae*, на примере которых мы рассмотрим данное представление ДНК.

На рис. 2 изображены карты *AGTC* седьмой (вверху) и восьмой (внизу) хромосом, содержащих 1 090 948 и 562 644 нуклеотидов соответственно. На рис. 3 показано представление девятой хромосомы той же клетки, состоящей из 439 887 нуклеотидов. Видно, что эти рисунки имеют много общего: каждая хромосома в таком представлении имеет три характерные особенности. Прежде всего, это длинные почти прямолинейные участки. Кроме того, явно видны области, в пределах которых траектория словно блуждает случайным образом. Наконец, четко прослеживаются фрагменты вытянутой формы.

Очевидно, что на длинных прямолинейных участках преобладают определенные нуклеотиды в последовательности ДНК. Это хорошо видно на рис. 3. По наклонам данных участков, ввиду различия концентрации пуринов и пиримидинов, допустимо предположить неравенство веса (размера) комплементарных ветвей спирали ДНК и говорить об их анизотропии. Поэтому интересно исследовать распределение по длинам и направлениям прямолинейных участков траектории. Это можно сделать следующим образом.

Зафиксируем рамку (окно) размером N нуклеотидов с учетом интересующего нас масштаба и будем рассматривать движение этой рамки вдоль траектории блуждания на плоскости *AGTC* с шагом в один нуклеотид. Построим в координатах *AGTC* диаграмму такого движения следующим образом. На каждом шаге перемещения рамки зафиксируем радиус-вектор, соединяющий ее начало и конец. Этот радиус-

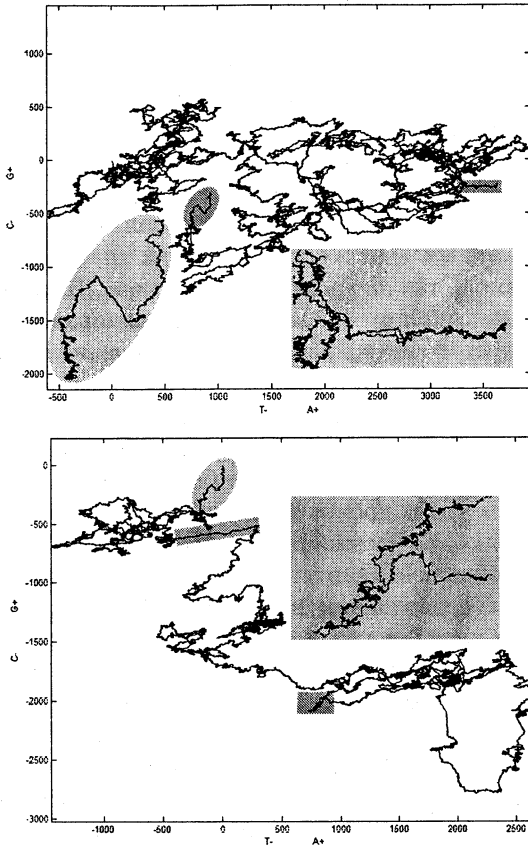


Рис. 2. Отображение *AGTC* 7-й (вверху) и 8-й (внизу) хромосомы *S. cerevisiae*.

вектор будет характеризоваться определенной длиной и направлением. На диаграмме отметим конец этого радиус-вектора. Сдвинем рамку на один шаг. Тогда получится некоторый новый радиус-вектор. Аналогичным образом отметим и его на диаграмме. И так далее. В результате получим диаграмму на плоскости (рис. 4), каждая точка i ($0 < i < M - N$, M — количество нуклеотидов в последовательности, N — размер рамки) которой обозначает, насколько далеко и в каком направлении изображающая точка отошла от начала i -го отрезка за N шагов. Очевидно, для каждой последовательности существует свой характерный размер рамки.

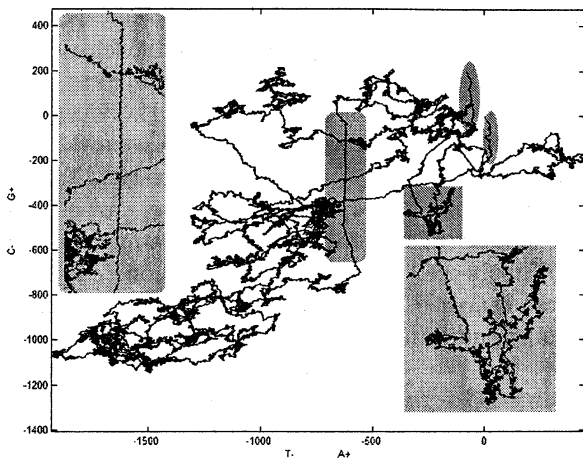


Рис. 3. Карта 9-й хромосомы *S. cerevisiae*

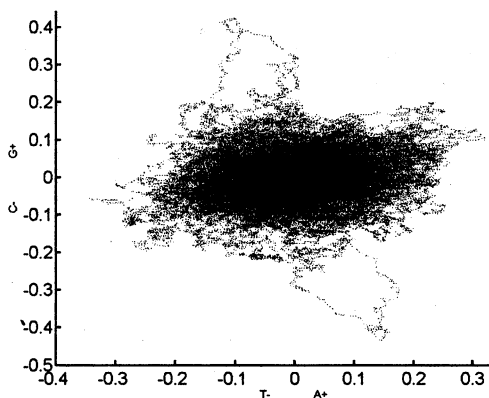


Рис. 4. Распределение по длинам 12-й хромосомы *S. cerevisiae* (см. далее рис. 5 (вверху)).

Данное построение позволяет легко выделить участки с преобладанием определенных нуклеотидов. В частности, выбросы (см. рис. 4) характеризуют направление и длину наблюдаемых прямолинейных участков.

Представление последовательности нуклеотидов в виде карты *AGTC* позволяет рассматривать функционально-структурную организацию хромосом на различных масштабах. Так, легко определяют

ся практически идентичные (и любых размеров) участки хромосом, в частности теломеры, которые на рис. 2, 3 выделены серыми эллипсами. Визуально, не проводя статистического анализа, легко обнаружить подобные регионы в разных хромосомах. Как следует из построения, движение по этим участкам происходит в разных направлениях, т. е. данные участки являются комплементарными. Также легко заметить участки длинных повторов. На рис. 3 в центре выделен длинный повтор (сглаженный прямоугольник), кодирующий белок Flo11, участвующий в процессе споруляции. В восьмой (см. рис. 2, внизу) хромосоме также выделен прямолинейный участок (прямоугольник с наклоном), кодирующий миозин — белок, необходимый для движения клетки.

Легко найти также палиндромы со значительной долей псевдослучайных включений. Напомним, что палиндромом называется последовательность, которая в обоих направлениях читается одинаково. В молекулярной биологии и биоинформатике под палиндромом в большинстве случаев, если это не оговаривается отдельно, подразумевают комплементарный палиндром, т. е. последовательность, продолженную в той же цепи своей *комплементарной* последовательностью. На рис. 2 (внизу) выделен диффузный палиндром (серый прямоугольник справа), кодирующий несколько белков, участвующих в процессе митоза. На рис. 3, представляющем девятую хромосому дрожжей, в левом нижнем углу выделен кластер инвертированных палиндромов с сильным дрейфом. В данном субтеломерном участке кодируется кластер, представленный альфа-глюкозидазой, АТФ-зависимой геликазой, серипауперином-1/14 и мембранным гликопротеином. На рис. 2 в седьмой хромосоме дрожжей можно видеть горизонтальный палиндром, который представляет собой пример мобильного элемента генома семейства ретротранспозонов TYA Gag и TYB Pol.

Отметим, что, по-видимому, палиндромные структуры сыграли очень важную роль в формировании генома и геномного кода [9].

3. Фрактальные свойства представления

Применение описанного представления ДНК как блуждания на плоскости гораздо шире, нежели только для идентификации и сравнения участков. В комплексе с другими современными методами, такими как фрактальный и фурье-анализ [10], вейвлет-преобразование [11] и др., оно дает возможность проводить *подробную* сегментацию и, вследствие достаточной наглядности, становится инструментом для

формулировки и проверки гипотез по организации и свойствам геномов в целом [9].

В данном разделе кратко рассмотрим фрактальные свойства описанного представления ДНК на примере тех же клеток *S. cerevisiae*. На рис. 5 в одном масштабе показаны все 16 хромосом этого организма. Видно, что все карты имеют структуру, характерную для статистики Леви.

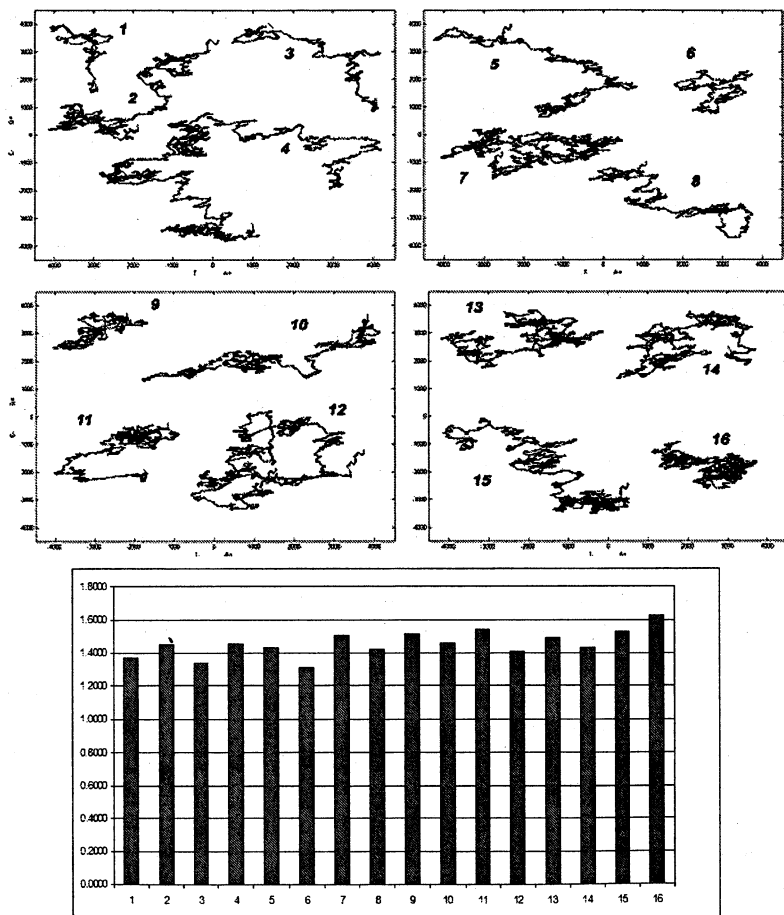


Рис. 5. Карты *AGTC* всех 16 хромосом *S. cerevisiae* (вверху) и соответствующие значения фрактальной размерности (внизу).

В нижней части рис. 5 представлены значения фрактальной размерности каждой из карт хромосом *S. cerevisiae*. Расчет фрактальной размерности проводился следующим образом.

Выберем случайный нуклеотид в последовательности *AGTC*. Пусть (x, y) — координаты этого нуклеотида на плоскости *AGTC*. Построим окружность случайного радиуса L с центром в точке (x, y) . Пусть N — количество точек последовательности, оказавшихся внутри такой окружности. Построим график, где по оси X отложено значение $\log L$, а по оси Y — $\log N$. Как известно, тангенс угла наклона прямолинейного участка этого графика будет равен фрактальной размерности.

Максимальное значение фрактальной размерности ($d_f = 1,62$) соответствует карте 16-й хромосомы. Это и понятно. Траектория блуждания для нее является наиболее плотной (см. рис. 5 (вверху)). Наименьшее значение имеет карта 6-й хромосомы. Это связано с тем, что ее длина (270 150 нуклеотидов) достаточно мала для образования инверсионных кластеров и диффузных палиндромов. При этом для первой хромосомы (230 210 нуклеотидов), сравнимой по длине с 6-й, имеет место включение мобильного элемента генома — туретротранспозона, имеющего в своей структуре диффузный палиндром. Это и приводит к увеличению значений фрактальной размерности.

В целом, значения фрактальной размерности всех карт хромосом *S. cerevisiae* находятся в пределах $1,31 \div 1,61$.

4. Представление хромосом человека

В данной части мы приведем в качестве примера представления ДНК как двумерного блуждания некоторые хромосомы человека. Такое представление было выполнено впервые в работе [12].

На рис. 6, 7 показаны карты хромосом с первой по четвертую. Прежде всего, бросается в глаза явно выраженный общий диагональный наклон траектории. Это означает, что в исходной последовательности число компонент $A-T$ преобладает над числом компонент $G-C$.

Кроме того, явно выражено огромное количество гигантских палиндромов. Обнаружить такие палиндромы достаточно сложно из-за их больших размеров (до десятков миллионов нуклеотидов) и сильного мутационного дрейфа, усложненного другими видами реорганизации. Как уже отмечалось, по-видимому, палиндромы сыграли ключевую роль в эволюционном процессе [9].

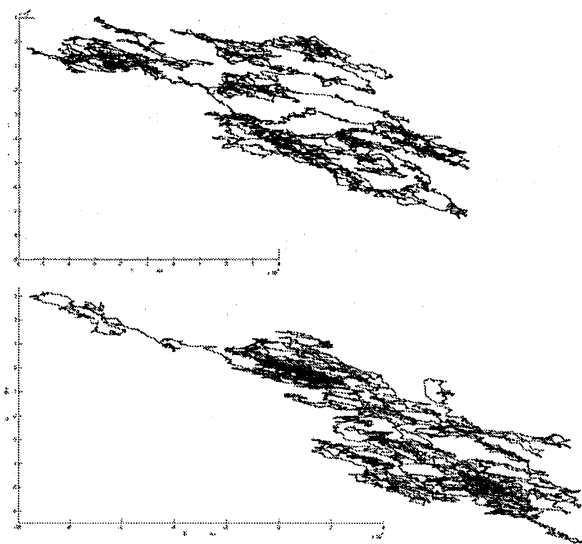


Рис. 6. Первая (247 249 719 нуклеотидов) и вторая (242 951 149 нуклеотидов) хромосомы человека.

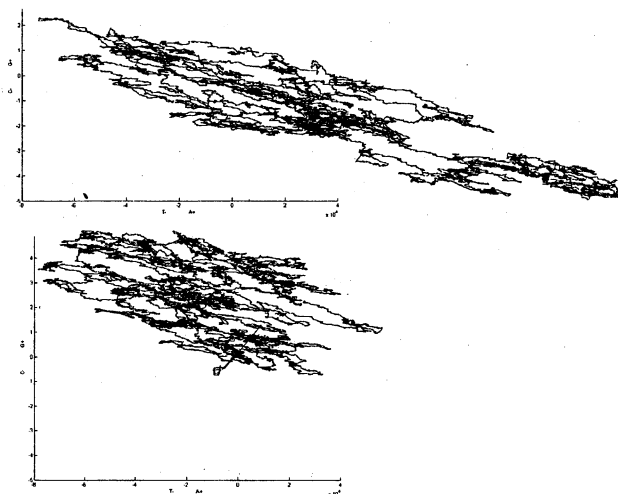


Рис. 7. Третья (199 501; 827 нуклеотидов) и четвертая (191 273 063 нуклеотидов) хромосомы человека.

5. Модель ДНК как двумерного блуждания

Описанное представление последовательности нуклеотидов позволяет построить математическую модель, достаточно детально описывающую карту ДНК на плоскости. Это дополняет анализ физических принципов строения и организации хромосом [1]. Используя методы символической динамики, можно получить последовательности, состоящие из символов словаря разбиения. В качестве отображения кодирующего ДНК мы использовали отображение единичного интервала, способного проявлять сложную динамику, известное как отображение Манневилля.

5.1. Моделирование с использованием случайных процессов

Отображение Манневилля было впервые введено в работе [13] как пример дискретной динамической системы с перемежаемостью:

$$x \mapsto f(x, a) = x + x^a \pmod{1}, \quad x \in [0; 1], \quad a > 1. \quad (1)$$

Эта работа привлекла большое внимание, поскольку динамика, схожая с динамикой данного отображения, была обнаружена во многих других областях, относящихся к моделированию последовательностей ДНК [14], физике ионных каналов [15], неравновесной термодинамике [16] и т. д.

Известно, что в зависимости от значений параметра a отображение (1) проявляет три качественно различных динамических режима. При $1 \leq a < 3/2$ динамика является нормальной в том смысле, что флуктуации случайной величины, порождаемой отображением, распределены по закону Гаусса. При $3/2 < a < 2$ поведение системы является переходным с аномальными отклонениями. Если $a > 2$, то динамика становится аномальной со статистикой Леви с показателем $\alpha = 1/(a - 1)$.

Для описания последовательности *AGTC* посредством символического представления, перепишем отображение Манневилля в явном виде:

$$x_{n+1} = f(x_n, a) = x_n + x_n^a \pmod{1}, \quad x \in [0; 1], \quad a = 1,6. \quad (2)$$

Разобьем отрезок $[0, 1]$ на подынтервалы как $[0; x_A; x_G; x_T; 1]$, $0 \leq x_A \leq x_G \leq x_T \leq 1$, и сопоставим каждому символу A , G , T и C определенный подынтервал следующим образом: $A = [0; x_A]$, $G = [x_A; x_G]$, $T = [x_G; x_T]$, $C = [x_T; 1]$. Теперь, в соответствии с тем, в какой из

этих подынтервалов попадает изображающая точка преобразования (2), отобразим всю траекторию на карту $AGTC$ по правилу, аналогичному описанному в п. 2. Именно, если изображающая точка находится в подынтервале $[0; x_A]$, сделаем шаг вправо, если $G = [x_A; x_G]$, то вверх, если $[x_G; x_T]$, то влево, и если $C = [x_T; 1]$, то вниз. Тогда порожденная последовательность будет иметь достаточно сложную структуру с регулярными и «хаотическими» доменами (рис. 8).

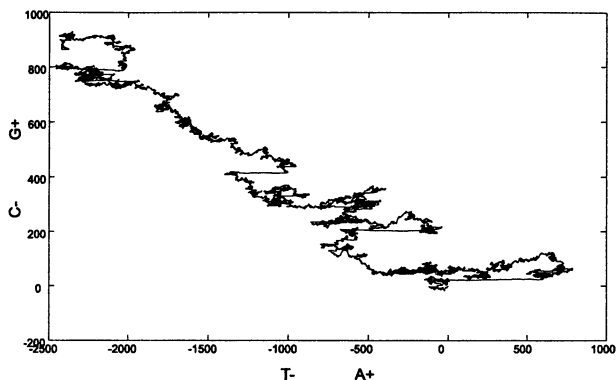


Рис. 8. Пример символического представления динамики отображения (2) при $x_A = 1/8$; $x_G = 1/4$; $x_T = 3/4$.

Целью нашего моделирования является изучение процесса возникновения подпоследовательностей квазипалиндромного типа и их мутационного дрейфа с помощью методов символической динамики. Палиндромные и квазипалиндромные последовательности составляют значительную часть геномов для многих видов [12].

Обобщением данного типа последовательностей, с учетом комлементарной симметрии составных элементов (нуклеотидов), может быть инверсионный градиент. Случайное блуждание, порождаемое специальной параметризацией отображения Манневилля, имеет свойства, близкие к тем, которые наблюдаются в геномных последовательностях. Ниже описано несколько характерных примеров.

Сначала рассмотрим пример диффузии (мутационного дрейфа) уже созданных палиндромов. Пусть отображение Манневилля задано в виде (2). Введем *делеции* следующим образом:

$$\begin{aligned}
 x_n & \mapsto \mathbf{G}, \\
 x_{n+1} = f(x_n, a) & \mapsto C, \\
 x_{n+2} = f(x_{n+1}, a) + \varepsilon\xi(x) & \mapsto \mathbf{A}.
 \end{aligned} \tag{3}$$

Здесь символ C должен быть удален из порожденной последовательности. Предполагается, что значение $x = x_{n+2}$, следующее за делецией, возмущено шумом $\xi(x)$ очень малой амплитуды ε , где $\xi(x)$ — нормально распределенная на отрезке $[-1; 1]$ случайная величина. Смысл добавления слабого шума состоит в том, что любая мутация вносит некоторое возмущение в порожденную последовательность. Благодаря хаотическому характеру отображения это может привести к изменению последующих элементов последовательности после очень долгого времени итерирования n . Можно предположить, что конкретный вид распределения существенно не изменит результат. Процесс делеций схематически представлен на рис. 9, а.

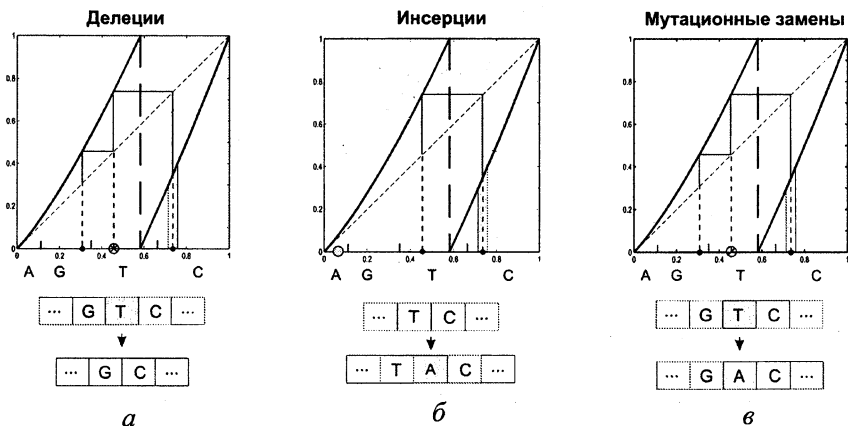


Рис. 9. Представление делеций (а), инсерций (б) и мутаций (в)

Инсерции могут быть представлены следующей системой уравнений:

$$\begin{aligned}
 x_n & \mapsto C, \\
 & \mapsto G, \\
 x_{n+1} = f(x_n, a) + \varepsilon\xi(x) & \mapsto A.
 \end{aligned} \tag{4}$$

Здесь символ G вставляется между C и A без шага итерации. Так же как и в процессе делеции, инсерция вызывает малые возмущения последовательности (аддитивный шум в выражении для x_{n+1}). Процесс инсерции схематически представлен на рис. 9, б.

Процесс *мутационных замен* можно определить следующим образом:

$$\begin{aligned}
 x_n & \mapsto G, \\
 x_{n+1} = f(x_n, a) & \mapsto T, \\
 x_{n+2} = f(x_{n+1}, a) + \varepsilon\xi(x) & \mapsto A.
 \end{aligned} \tag{5}$$

В этом случае символ T подставляется вместо символа C , который появляется из x_{n+1} путем символического представления в процессе итерирования (рис. 9, в). Под мутациями понимают совокупность инсерций, делеций и мутационных замещений, происходящих в последовательности с некоторой вероятностью. Процесс, называемый мутационной заменой, схематически представлен на рис. 10. На этом рисунке построены две одинаково порожденные (одни и те же начальные условия) символические последовательности: одна — с мутациями (серая ломаная), а другая — без них (черная ломаная).

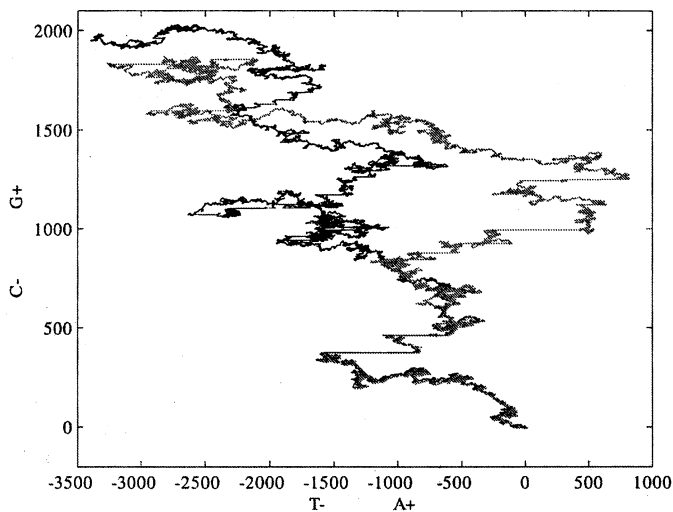


Рис. 10. Примеры моделирования ДНК с мутациями и без них.

В описанной модели образования последовательности ДНК можно контролировать частоту появления мутаций. Это позволяет получать обратные подпоследовательности палиндромов максимально близкими, насколько это необходимо, к исходным (прямым) подпоследовательностям. Данное моделирование может быть интересно с точки зрения параметризации мутационного дрейфа и оценки характерных времен диффузии палиндромов [12]. В процессе, показанном на рис. 10, вероятности появления делеций, инсерций и мутаций равны $p_d = 1/32$, $p_i = 1/64$ и $p_m = 1/64$ соответственно. Это значит, что за один шаг в отображении Манневилля делеция может произойти с вероятностью p_d . Такие же правила действуют для инсерций с вероятностью p_i и мутаций с вероятностью p_m .

5.2. Моделирование на основе хаотического отображения

В данном разделе мы рассмотрим алгоритм, порождающий квазипалиндромы. Он основан на использовании символической динамики для *модифицированного* отображения Манневилля. В качестве порождающего символическую последовательность отображения рассмотрим преобразование отрезка $[0; 1]$ в себя следующего вида:

$$x_{n+1} = f(x_n, a, b) = x_n + bx_n^a(1/2 - x_n)(1 - x_n)^a \pmod{1}, \quad x \in [0; 1]. \quad (6)$$

В окрестности точек $x = 0$ и $x = 1$ оно имеет характер исходного отображения Манневилля. Поэтому при соответствующих параметрах оно проявляет перемежаемость с двумя «ламинарными» участками: вблизи точек 0 и 1. Параметры, гарантирующие такой режим перемежаемости, были выбраны следующими: $a = 3$ и $b = 300$ (рис. 11, *a*). Для получения символической последовательности из символов A, G, T и C отрезок $[0; 1]$ был разбит на 64 равные части, каждой из которых ставился в соответствие однозначно триплет (s_1, s_2, s_3) , где $s_i \in A, G, T, C$.

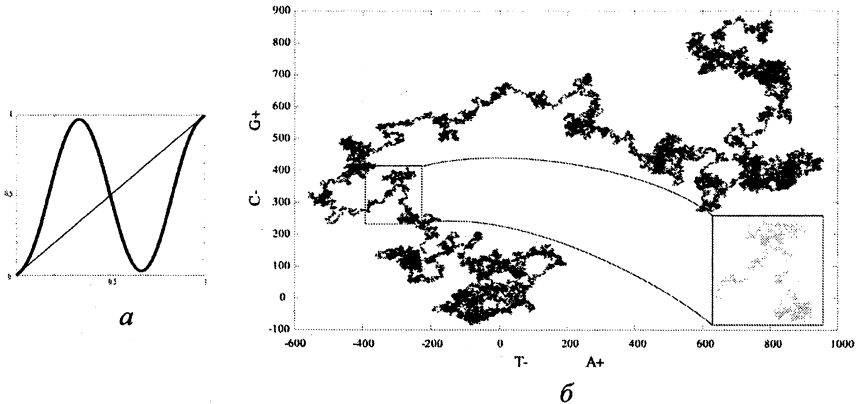


Рис. 11. Последовательность, порождаемая симметричным отображением Манневилля.

На рис. 11, *б* представлена последовательность искусственной ДНК как двумерного блуждания, порожденная отображением (6) длиной 600 000 символов. Хорошо видно, что данная последовательность имеет сложный характер и внешне напоминает последовательность реальной ДНК (см. рис. 2, 3 и 5). На врезке показан отрезок полученной последовательности, похожий на диффундировавший палиндром. Отметим, что при таком способе получения последовательности

ДНК нигде не использовались случайные функции. Вся сложность последовательности обеспечивается только слабо хаотическим характером (перемежаемость со статистикой Леви) измененного отображения Манневилля (6).

6. Заключительные замечания

Природа фрактального множества хромосом в их представлении на карте *AGTC*, по нашему мнению, состоит в свойствах комплементарности нуклеотидов и общем рекомбинационном процессе эволюции. Соответствующая наблюдаемая во многих геномах *G – C*-позиционная симметрия, постулированная на уровне состава ДНК Чарграффом в виде второго правила (о том, что в составе одной из цепей ДНК число молекул тимина равно числу молекул аденина, а число молекул гуанина равно числу молекул цитозина), предположительно отражает результат релаксации обобщенного инверсионно-диффузионного процесса в течение миллионов лет.

Построенная в последнем пункте модель, описывающая ДНК именно как случайное блуждание, позволяет формализовать процесс инверсионно-диффузионной динамики. Это дает возможность наглядно отобразить адекватность постановки вопроса о природе генома.

Литература

1. Ларионов С. А., Лоскутов А. Ю., Рядченко Е. В. Геном как двумерное блуждание // Докл. РАН. 2005. Т. 405, № 6. С. 755–759.
2. Голомб С. В. Код ДНК // Математические методы в биологии / Под ред. Р. Беллмана. М.: Наука, 1966. С. 251–271.
3. Gates M. A. Simple DNA sequence representations // Nature. 1985. V. 316. P. 219.
4. Mizraji E., Ninio J. Graphical coding of nucleic acid sequences // Biochimie. 1985. V. 67. P. 445.
5. Berthelsen Ch. L., Glazier J. A., Skolnik M. H. Global fractal dimension of human DNA sequence threaded as pseudorandom walks // Phys. Rev. A. 1992. V. 45. P. 8902–8913.
6. Vincens P., Buff L., Andre C. et al. A strategy for finding regions of similarity in complete genome sequences // Bioinformatics. 1998. V. 14, № 8. P. 715–725.
7. Abramson G., Alemany P. A., Cerdeira H. A. Noisy Levy walk analog of two-dimensional DNA walks for chromosomes of *S. cerevisiae* // Phys. Rev. E. 1998. V. 58. P. 914–918.
8. Rosas A., Nogueira Jr. E., Fontanari J. F. Multifractal analysis of DNA walks and trails // arXiv:cond-mat/0209396. 2002.

9. *Larionov S. A., Loskutov A., Ryadchenko E. V.* // (Готовится к публикации).

10. *Лобзин В. В., Четветкин В. П.* Порядок и корреляции в геномных последовательностях // УФН. 2000. Т. 170. С. 57–81.

11. *Loskutov A., Postnikov E. B., Larionov S. A., Ryadchenko E.* Analysis of DNA structure as a 2d walk by complex wavelet transform // Proc. Int. Conf. «Wavelets and Applications – 2006» (July 10–14, 2006, EPFL, Lausanne, Switzerland).

12. *Larionov S. A., Loskutov A., Ryadchenko E. V., Rybalko S.* Visual genomics: Gigantic palindrome diffusion and certain features of genomes evolution // Proc. Int. Symp. Evolution of Biomolecular Structure (University of Vienna, Austria, 25–27 May 2006).

13. *Manneville P.* Intermittency, self-similarity and $1/f$ spectrum in dissipative dynamical systems // J. Physique. 1980. V. 41. P. 1235–1243.

14. *Allegrini P., Barbi M., Grigolini P., West B. J.* Dynamical model for DNA sequences // Phys. Rev. E. 1992. V. 52. P. 5281–5297.

15. *Toth T. I., Liebovitch L. S.* Models of ion channels kinetics with chaotic subthreshold behaviour // Z. Angew. Math. Mech. 1996. V. 76. P. 523–524.

16. *Bugatti M., Grigolini P., Palatella L.* Nonextensive approach to the entropy of symbolic sequences // Physica A. 1999. V. 268. P. 214–221.

17. *Gaspard P., Wang X.-J.* Sporadicity: between periodic and chaotic dynamical behaviours // Proc. Natl. Acad. Sci. (USA). 1988. V. 85. P. 4591–4602.

РОЛЬ НЕУСТОЙЧИВОСТЕЙ, НЕЛИНЕЙНЫХ КОЛЕБАНИЙ И ВОЛН В РЕГУЛЯЦИИ ТУБУЛИНОВОГО ЦИТОСКЕЛЕТА

Е. А. Катруха, Г. Т. Гурия

1. Введение

Колебания и волны играют ключевую роль в процессах передачи информации в пространственно протяженных биологических системах. После того как Герман Гельмгольц установил, что скорость распространения нервного импульса у лягушки составляет 26,2 м/с [1], эстафетные механизмы передачи сигнала в биологических, химических и физических системах стали предметом интенсивного экспериментального и теоретического изучения [2—5]. Наиболее значительные «подвижки» во взглядах на природу механизмов генерации и распространения нервных импульсов произошли после знаменитой работы А. Л. Ходжкина и А. Ф. Хаксли [6]. Оказалось, что распространение возбуждения в нерве опосредуется трансмембранной электрической деполяризацией, распространяющейся по нервному волокну в соответствии с так называемыми телеграфными уравнениями [7]. Формально распространение нервного импульса описывается параболическими уравнениями подобно фронтам горения [4] и концентрационным волнам в реакционно-диффузионных системах [8, 9]. За работы по развитию основ неравновесной термодинамики реакционных систем с диффундирующими реагентами И. Р. Пригожину в 1977 году была присуждена Нобелевская премия по химии [10].

Анализ широкого класса моделей активных сред показал, что в системах, далеких от термодинамического равновесия, могут спонтанно возникать устойчивые пространственные, временные (колебания) и пространственно-временные структуры (волны) [11, 12]. Помимо актуальных проблем в индивидуальном развитии [13] и морфогенезе [14—16], реакционно-диффузионные модели стали широко использоваться в клеточной биологии для интерпретации колебательных процессов в гликолизе [17], в регуляции цитоплазматических колебаний ионов Ca^{2+} [18—19].

После выхода в свет книги Г. Хакена [20] стало ясно, что достаточно широкий круг явлений самоорганизации в природе может трактоваться с единых идейных позиций. В некотором смысле можно сказать, что Хакену выпала честь указать на правомерность распространения методов классической теории колебаний и волн [21, 22] за рамки радиотехники и физики в целом.

С математической точки зрения в основе ключевых процедур синергетики лежат асимптотические методы (методы многих масштабов), перенесенные с систем обыкновенных дифференциальных уравнений [23], на уравнения в частных производных, описывающих эволюцию про-

странственно протяженных реакционно-диффузионных систем, в духе, принятом в механике «сплошных сред» [24, 25].

Применительно к «биологическим» системам «математическое» свойство «сплошности» принято толковать как отсутствие сингулярностей, внутренних особенностей в рассматриваемом объекте [26—28]. В свете вышеизложенного внутриклеточные процессы не вполне корректно описывать в рамках реакционно-диффузионного приближения, т. к. наряду с биохимическими, её динамике свойственны и механо-химические степени свободы [29]. Последние обуславливаются наличием в клетке цитоскелета. Армирующий внутриклеточное пространство цитоскелет представляет собой ажурную трехмерную сеть, сформированную полимерными микротрубочками, состоящими из молекул тубулина и актина [31].

Цитоскелет обладает чертами лабильности и способен претерпевать значительные трансформации в ходе клеточного цикла и при воздействии на клетку внешними агентами [31—33]. Пронизывающее внутриклеточное пространство полимерные волокна способны удлиняться (или укорачиваться) в результате реакций сорбции (или же десорбции) мономеров на их активные участки. Последние выступают при этом как центры поглощения — своеобразные сингулярные стоки (источники) мономеров из жидкой фазы при полимеризационном росте (распаде) элементов цитоскелета.

О «сингулярности» цитоскелетных структур на фоне достаточно плавных концентрационных структур, имеющих место в жидкой фазе цитоплазмы, представляется уместным говорить в силу масштабных различий. Поперечные размеры у актиновых и тубулиновых нитей составляют 10 и 25 нм соответственно, в то время как характерные концентрационные структуры в клетках, формирующиеся за счет микродиффузии, обычно соизмеримы по масштабу с размерами самих клеток, т. е. превышают микрометры [34, 35].

С математической точки зрения при анализе неравновесных концентрационных структур в клетке мы имеем дело со структурами, существующими не всюду, а «почти всюду» в рассматриваемой области [30]. Размерность множества, занимаемого цитоскелетом, формально может полагаться при таком подходе равной нулю. Существование в клетке нитевидных цитостатических (цитоскелетных) структур с концентрационными, как несложно понять, взаимообусловлено.

Выяснение того, какую роль способны играть нестабильности, колебания и волны в динамических механизмах трансформации тубулинового цитоскелета, является предметом данной работы.

2. Строение и функции цитоскелета

Цитоскелет представляет собой клеточную органеллу, состоящую из трех основных компонентов: актиновых волокон, промежуточных фила-

ментов и тубулиновых микротрубочек. Тонкие актиновые волокна концентрируются на периферии клетки у внешней мембраны и участвуют в сократительной деятельности клетки и ее движении как целого. Промежуточные филаменты расположены кольцом вокруг ядра клетки. Микротрубочки расходятся преимущественно радиально от ядра к периферии клетки. Именно из них образуется так называемое «веретено деления», отчетливо наблюдающееся при видеомикроскопии элементарных стадий клеточного деления (рис. 1) [31, 33].

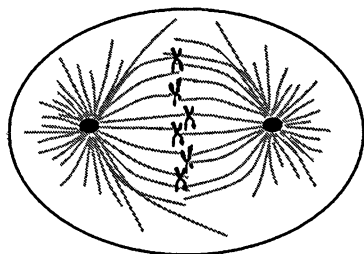


Рис. 1. Схематическое изображение делющейся клетки (стадия метафазы). Микротрубочки (показаны серым цветом), формирующие веретено деления, радиально расходятся от центросом. В центральной области клетки их свободные концы «захватывают» хромосомы.

Микротрубочки представляют собой полые цилиндрические супрамолекулярные структуры с внешним диаметром 24 нм, внутренним диаметром 13 нм и длиной до 100 мкм, состоящие из молекул белка тубулина (рис. 2) [36]. Микротрубочки в норме представляют собой 13-заходную правозакрученную спиральную структуру. В силу химической анизотропии, свойственной отдельным молекулам тубулина и стерическим ориентационным эффектам при их полимеризации, у каждой микротрубочки принято различать «плюс» и «минус» концы. Минус-конец обычно закрепляется вблизи ядра, на центросоме, а плюс-конец — у внешней мембраны. Помимо участия в перераспределении хромосом в процессе деления, микротрубочки выполняют в клетке ряд других физиологических функций. Вместе с актиновыми филаментами тубулиновые микротрубочки принимают участие в движении клетки как целого [31].

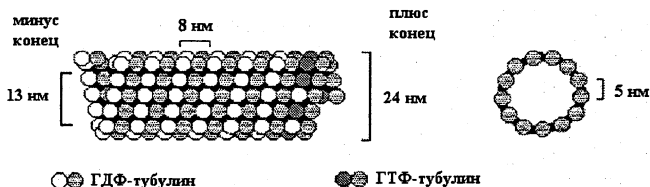


Рис. 2. Структура и характерные размеры микротрубочки и составляющих ее элементов

Кроме того, сеть из микротрубочек, проникая во все области клетки, служит «рельсами» для активного внутриклеточного транспорта (молекулярных моторов), а также сигнальной системой, реагирующей на изменения окружения [31]. Микротрубочки — достаточно жесткие по полимерным меркам структуры, их персистентная длина¹ в зависимости от условий может достигать 5 мкм [37]. Поскольку представляющие для нас интерес эукариотические клетки имеют характерные размеры от 5 мкм до 100 мкм, то преимущественное радиальное расположение микротубулиновых волокон позволяет им исполнять роль каркаса, обеспечивая устойчивость формы клетки [38]. Способность клеток к значительным трансформациям структуры тубулинового цитоскелета связана, в частности, с необходимостью адаптивно реагировать на внешние воздействия. Средства современной флуоресцентной микроскопии позволяют визуализировать микротубулиновый цитоскелет (рис. 3) и следить за динамикой его трансформаций.

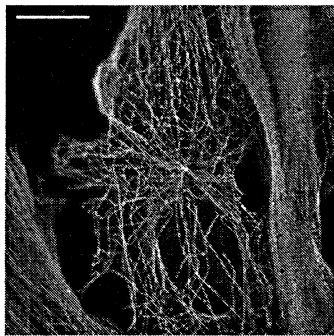


Рис. 3. Радиальная система флуоресцирующих микротрубочек в клетках Vero (эпителий почки зеленой африканской макаки)². Размер светлого прямоугольника в левом верхнем углу соответствует 10 мкм.

3. Феноменологические черты динамики микротрубочек

Наблюдение за поведением плюс-концов *in vivo*, показывает, что на всех стадиях клеточного цикла, за исключением митоза, они обычно расположены на периферии клетки. То есть, как правило, наблюдаются квазистатистические распределения микротрубочек по длинам, в которых преобладают протяженные экземпляры. При этом местоположение плюс-концов незначительно колеблется за счет локальной сорбции/десорбции тубулина, однако амплитуда невелика в сравнении с характерной длиной микротрубочек [38]. В процессе митоза поведение микротрубочек коренным образом изменяется. Амплитуда перемещений плюс-концов резко возрастает. Периоды достаточно медленного их роста чередуются со ста-

¹ Персистентная длина — принятый в физике полимеров термин, используемый для обозначения расстояния вдоль полимерной цепи, на котором утрачивается ориентационная корреляция между звеньями [50].

² Фотография публикуется с разрешения О. В. Коваленко, лаборатория молекулярной биологии цитоскелета, НИИ физико-химической биологии, Институт белка, Пушино.

диями стремительной деполимеризации [46]. Переключения между стадиями роста и деполимеризации микротрубочек происходит внешне случайным, нерегулярным образом.

По-видимому, подобного рода нестационарная динамика микротрубочек имеет важное физиологическое значение. За счет «трепетаний» плюс-концов осуществляется эффективный пространственный поиск хромосом [39]. Как только плюс-конец микротрубочки достигает центромеры¹ (кинетохора) хромосомы, он крепится к ней и прекращает движение. На следующей стадии митоза происходит растаскивание хромосом к центрам дочерних клеток.

Сеть тубулиновых волокон в живой клетке находится под управлением ряда биохимических факторов [31, 35, 40]. Механизмы, приводящие к переключению цитоскелета из квазистационарного режима в динамический, скорее всего, носят пороговый характер [41] и в настоящее время активно изучаются [33, 34].

Выяснилось, что соизмеримые с размером микротрубочек «трепетания» могут иметь место не только в живых клетках (т. е. в системах *in vivo*), но и в искусственных, так называемых реконструированных биохимических системах², т. е. системах *in vitro* [42]. По-видимому, впервые в системе *in vitro*, представляющей собой водный раствор молекул тубулина, магния и гуанозинфосфата, колебания оптической плотности были зарегистрированы в работе [43] (рис. 4).

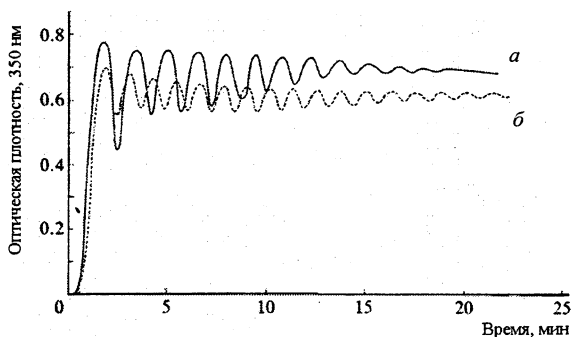


Рис. 4. Кривые изменения во времени общего количества полимеризованного тубулина, определяемого по оптической плотности (длина волны 350 нм), в системе *in vitro* [43]. Концентрация тубулина в растворе равна: а — 150 мкМ, б — 100 мкМ (1 М = 1 моль/л).

¹ Центромера (кинетохор) — макромолекулярная белковая структура, расположенная в механическом центре хромосомы. Образует перетяжку, разделяющую хромосому на два плеча (две части).

² Реконструированными системами *in vitro* в современной биологии и биохимии называются искусственные системы, содержащие строго определенный набор элементов реальных систем.

Обнаружилось, что при определенных начальных концентрациях ГТФ (гуанозинтрифосфат) и тубулина в системе имеют место колебания, которые постепенно затухают, что связывают с исчерпанием молекул ГТФ и постепенным накоплением в растворе молекул ГДФ (гуанозиндифосфат). Добавление в раствор биохимической системы регенерации ГТФ из ГДФ значительно увеличивает продолжительность колебаний [44].

Удалось выяснить, что описанные выше объемные колебания связаны с нестационарным изменением длины отдельных микротрубочек, содержащихся в растворе (рис. 5) [42, 45]. Явление, в ходе которого отдельные микротрубочки проходят через сменяющие друг друга фазы роста и деполимеризации, получило в литературе название динамической неустойчивости (dynamic instability) [42].

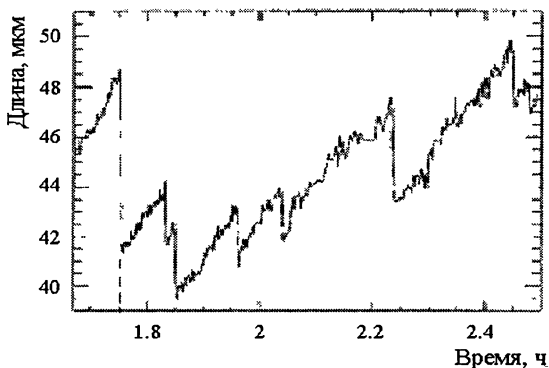


Рис. 5. Динамическая неустойчивость микротрубочки [45]. По оси абсцисс отложено время, по оси координат — размер индивидуальной микротрубочки.

В биологической литературе момент внезапной смены роста микротрубочки на ее стремительную деполимеризацию принято называть катастрофой или моментом катастрофы (catastrophe event). Обратное событие в динамике микротрубочки носит название «спасения» (rescue event). Природа механизмов, лежащих в основании «катастроф» и «спасений», на протяжении двадцати лет широко обсуждается [47]. Большое число работ посвящено обсуждению гипотезы о существовании «ГТФ-крышки».

Суть гипотезы сводится к следующему. Молекулы ГТФ и ГДФ в растворе способны связываться с молекулами тубулина с приблизительно равными константами скоростей. Однако рост микротрубочек происходит только за счет присоединения к их плюс-концу молекул тубулина, связанных с ГТФ (тубулин-ГТФ). После присоединения молекулы тубулин-ГТФ к полимеру, связанная с тубулином молекула ГТФ способна гидролизироваться, превращаясь в ГДФ. В соответствии с обсуждаемой гипоте-

зой полагается, что микротрубочка, состоящая только из молекул тубулина-ГДФ, является нестабильной. Микротрубочку же от деполимеризации предохраняет своеобразная «крышка» из молекул тубулина-ГТФ, образующаяся на растущем конце благодаря тому, что гидролиз происходит с некоторым запаздыванием после присоединения очередной молекулы. За двадцать лет было предложено значительное количество моделей, в основе которых лежит гипотеза ГТФ-крышки [42, 46, 47].

Помимо этого круга моделей, явно основанных на гипотезе о существовании ГТФ-крышки, для описания колебаний было предложено несколько моделей феноменологического характера [48, 56]. Исторически одной из первых была модель М. Доджером и С. Лейблера [48]. В ней аксиоматически постулируется, что микротрубочки могут находиться только в двух состояниях — растущем и деполимеризующемся. Частота переключений между ними полагается изменяющейся стохастически. Делается предположение лишь о виде зависимости частоты переключений от концентрации тубулина-ГТФ в растворе. Скорость же процесса полимеризации полагается независимой от концентрации тубулина, связанного с ГТФ в растворе. На основании изложенных выше соображений авторы составили систему кинетических уравнений, описывающих динамику плюс-концов растущих и деполимеризующихся микротрубочек [48]. Решения системы находились методами Монте-Карло и среднего поля в неограниченной по пространству области.

В результате среди решений удалось выделить два класса, соответствующих ограниченному и неограниченному росту микротрубочки. Первый характеризуется стохастическими пульсациями длины микротрубочки в области центра нуклеации (центросомы), а второй — бегущей с постоянной скоростью в бесконечность волной плотности плюс-концов микротрубочек. Выяснилось, что смена типа поведения микротрубочки должна происходить пороговым образом, при выполнении специальных условий на некоторые из параметров системы, характеризующих скорость роста и деполимеризации, а также частоту переключения между двумя состояниями [48].

При всей привлекательности данного подхода, за его рамками остается целый ряд существенных обстоятельств. Прежде всего, не раскрывается природа пульсаций микротрубочек, так как их нестабильность просто аксиоматически постулируется. Наблюдающееся в решениях нестационарное поведение априори присуще системе, а не является результатом процессов взаимодействия ее частей. Постулирование всего двух состояний (роста и деполимеризации) не объясняет присутствие в ансамбле микротрубочек, не растущих и не деполимеризующихся, хотя таковые наблюдаются экспериментально и могут составлять заметную долю от общего числа [46, 49].

Концентрация тубулина-ГТФ в растворе полагается однородной и постоянной во времени, что также не соответствует экспериментальным фактам [49, 57]. По сути, игнорируется роль пространственных градиентов концентраций веществ, участвующих в полимеризации. Как показали последние исследования, в клетках существует диффузионный механизм авторегуляции, лимитирующий и направляющий рост ансамбля микротрубочек [34, 35]. Он активно используется клеточной системой регуляции сети тубулиновых полимеров при переходе от стационарной к ярко выраженной неравновесной динамике, наблюдаемой на стадии митоза. Повисают в «воздухе» и такие, казалось бы, «законные» вопросы, как: «Каким образом и за счет чего происходит внезапное переключение динамических режимов? Существует ли возможность управления переключением?»

Размышления о возможной роли диффузионных факторов в нетривиальной динамике микротрубочек начинаются с, казалось бы, естественных аналогий. Внешне пространственные смещения плюс-концов микротрубочек, происходящие за счет элементарных актов агрегации-дисагрегации молекул тубулина, во многом напоминают процессы, наблюдаемые при так называемой лимитированной диффузией агрегации (DLA — Diffusion Limited Aggregation) [51]. В ряде случаев при лимитированности процессов агрегации диффузией имеет место образование достаточно сложных структур, к которым применимы методы фрактальной геометрии. Однако в системах, к которым применима концепция DLA, таким как ветвящиеся полимеры, образование трещин, сорбция «мономерных звеньев», как правило, идет необратимо, в то время как при росте микротрубочек имеют место как прямые, так и обратные процессы десорбции. Поэтому перенос развитых методов теории DLA на интересующую нас систему «в лоб» не проходит. Возникает вопрос, существует ли в принципе в динамике микротрубочек какое-либо ограничение, может быть вызванное и не диффузией, а обусловленное чем-то иным? Если да, то какова его роль и может ли оно приводить к нестабильностям и нестационарному поведению?

Кроме диффузии, рост микротрубочек непосредственно зависит также от кинетических особенностей распределения в системе запасенной в ГТФ энергии. Стоит заметить, что ГТФ играет в биологии роль химического носителя энергии, как и хорошо всем известный АТФ (аденозинтрифосфат). Энергия выделяется в реакциях гидролиза, при которых ГТФ превращается в ГДФ. Без запаса энергии, либо без постоянного ее поступления извне полимеризационный рост микротрубочек невозможен. Поэтому возникает еще один крайне интересный вопрос: может ли агрегация и образование структур в ансамбле микротрубочек каким-либо образом лимитироваться или же модулироваться энергетическими факторами? Если да, то каким образом такого рода лимитация связана с диффузионными и прочими реакционными процессами?

Совокупность приведенных выше соображений и вопросов подтолкнула авторов к поиску новых путей описания динамики микротрубочек. С физической точки зрения система микротрубочек представляет собой нитевидную распределенную систему, демонстрирующую черты колебательной и стохастической динамики. Общие механизмы формирования неоднородных пространственных распределений, проявляющиеся в виде концентрационных волн и диссипативных структур, в реакционно-диффузионных системах (т. е. «сплошных», по сути, системах) хорошо изучены применительно к целому ряду физико-химических и биологических систем [8, 52—54]. Поэтому нам представлялось актуальным проанализировать динамику роста и деполимеризации микротрубочек в свете идей теории неравновесных процессов.

В данной работе мы исходили из того, что микротрубочки и реакционные компоненты раствора, в котором они находятся, представляют собой взаимосвязанную пространственно распределенную «почти всюду сплошную» систему. Необходимым условием существования автоколебаний и автоволн в далеких от равновесия системах является, как известно, наличие в кинетическом механизме *запаздывающей отрицательной обратной связи* [21, 55]. Последняя, в принципе, может реализовываться как за счет процессов, происходящих в самой микротрубочке (к примеру, за счет задержки гидролиза ГТФ в ГТФ-крышке), так и быть результатом взаимодействия микротрубочек и компонентов окружающего их раствора, т. е. вызываться внешними, по сути, кинетическими причинами. Действительно, из общих соображений ясно, что при росте микротрубочки за счет сорбции молекул тубулина-ГТФ в растворе должно происходить локальное понижение концентрации в окрестности ее плюс-конца. Очевидно, что если такое локальное обеднение раствора не вполне компенсируется в каждый момент времени за счет диффузии, то скорость сорбции тубулина-ГТФ из раствора должна замедляться. Иными словами, наличие элементов отрицательной запаздывающей обратной связи в регуляции роста микротрубочек наверняка должно иметь место в тех случаях, когда диффузия молекул тубулина-ГТФ в растворе не является «мгновенной»¹.

В пользу того, что рост трубочек на самом деле лимитируется процессами диффузии и реакциями присоединения молекул тубулина-ГТФ к плюс-концам, говорит ряд экспериментальных данных. Прежде всего, твердо установлено, что скорость присоединения молекул тубулина-ГТФ из раствора к плюс-концам микротрубочек действительно определяется локальной концентрацией этих молекул [42]. Соответственно рост микротрубочек непосредственно регулируется концентрационными градиентами веществ, участвующих в реакциях полимеризации [34, 35]. Сверх того, было установлено, что рост микротрубочек сам по себе приводит к образованию в окружающем микротрубочку растворе пространственных гра-

¹ Корректней, видимо, говорить «происходящей быстрее характерной скорости роста микротрубочки».

диентов концентраций [57]. Круг замкнулся. Из всего вышеизложенного ясно, что скорость роста микротрубочек на самом деле регулируется по типу отрицательной обратной связи за счет относительно медленных диффузионно лимитируемых процессов перераспределения молекул тубулина-ГТФ в жидкой фазе [57].

Несложно убедиться, что рост микротрубочек лимитируется энергетическими факторами по типу отрицательной обратной связи. Действительно, количество энергоресурсов, в данном случае молекул ГТФ, в клетке не может быть неограниченным. Поскольку для процесса полимеризации необходимо потребление энергоресурсов, то рост тубулиновых волокон будет приводить к общему уменьшению количества молекул ГТФ, что, в свою очередь, приведет к замедлению роста. Можно сказать, что таким образом в соответствии с принципом Ле-Шателье система реагирует на рост микротрубочек.

Понятно, что энергорегуляция динамики микротрубочек определяется, с одной стороны, диффузионными градиентами энергетически богатых молекул тубулина-ГТФ в окрестности плюс-концов, а с другой, градиентами десорбировавшихся из микротрубочек энергетически бедных молекул тубулина-ГДФ. Интенсивность протекания процессов доставки вещества, наделенного соответствующим энергетическим потенциалом, так же как и интенсивность процессов «отведения отработанного материала», в принципе может быть нестационарной. Наличие в рассматриваемой системе регуляции элементов отрицательной запаздывающей обратной связи наводит на мысль, что при определенных условиях в ней может иметь место потеря устойчивости стационарного состояния по А. А. Андронову с последующим рождением предельного цикла! Иными словами, из достаточно общих «нелинейных» соображений следует, что фазы интенсивного роста микротрубочек могут чередоваться с фазами их деполимеризации. Представляет большой интерес нахождение области параметров, в которых динамика микротрубочек будет иметь такого рода колебательные или же хаотические черты.

Понятно, что построение параметрических диаграмм состояния невозможно без построения физико-математической модели рассматриваемой системы. В силу сложности системы еще раз перечислим те ее элементы, которые будут нами постоянно держаться в поле зрения. Итак, мы будем рассматривать в простейшем одномерном приближении реакционную систему особого типа, у которой имеется:

а) концентрационная подсистема с диффундирующими в растворе компонентами;

б) один из компонентов раствора (тубулин-ГТФ) способен к динамической преципитации¹ («выпадению в осадок»), т. е. переходу из жидкой фазы в форму жесткой полимерной микротрубочки;

¹ Преципитация — образование твердой фазы (нерастворимого осадка) в жидкости (растворе) в результате химической реакции между растворенными компонентами.

в) рост микротрубочек есть следствие реакций присоединения к ним молекул тубулина-ГТФ из раствора, сокращение — реакций десорбции;

г) рост микротрубочки имеет место лишь в том случае, если в окрестности плюс-конца локально выполняются условия «пересыщенности» жидкой фазы, т. е. процессы сорбции доминируют над процессами десорбции;

д) энергообеспеченность роста микротрубочек обуславливается диффузионными процессами и реакциями с участием молекул тубулина, связанных с ГТФ и ГДФ, в растворе.

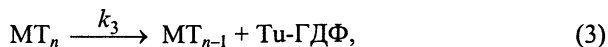
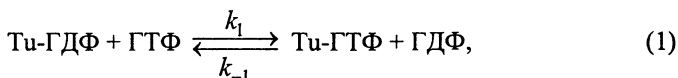
Поскольку рост и деполимеризация микротрубочек определяются локальными значениями концентрации молекул тубулина в окрестности плюс-концов, то динамика микротрубочек должна непосредственно быть связанной с сопутствующими концентрационными волнами в растворе. Для систем без преципитации (так называемых реакционно-диффузионных) механизмы возникновения и распространения концентрационных волн подробно изучены [8, 9]. Известно, что появление автоволн и пространственно неоднородных стационарных структур в неравновесных реакционно-диффузионных системах может быть следствием потери устойчивости стационарного состояния [15]. (К слову, и системы с выпадением в осадок, так же как и системы, в которых наблюдаются явления колебательной поверхностной кристаллизации, к настоящему времени изучены неплохо [58, 59]).

В свете вышеизложенного представлялось целесообразным понять, в какой мере масштабные трансформации в динамике микротрубочек, проявляющиеся в форме своеобразных «трепетаний тубулиновых нитей», могут быть следствием потери устойчивости в регуляции рассматриваемой реакционно-диффузионной-преципитационной системы.

Построенная ниже простая математическая модель позволила качественно изучить проблему параметрической устойчивости микротрубочек. Удалось построить диаграмму состояний и выяснить условия перехода от стационарных к динамическим режимам поведения микротрубочек.

4. Математическая модель

В данной работе мы будем полагать, что ключевую роль в динамике микротрубочек играют следующие реакции [67]:



где MT_n обозначает микротрубочку, содержащую n молекул тубулина, Tu -ГТФ и Tu -ГДФ обозначают растворенный тубулин, связанный соответственно с молекулами ГТФ и ГДФ. В ходе реакции с константой скорости k_1 в растворе происходит обмен связанной с тубулином молекулы ГДФ на молекулу ГТФ. Обратная реакция (с константой скорости k_{-1}) отображает обмен связанной с тубулином молекулы ГТФ на ГДФ, также происходящий в растворе. Реакция (2) описывает удлинение микротрубочки при полимеризации тубулина. Последняя реакция соответствует деполимеризации, при которой из микротрубочек высвобождается тубулин, связанный с молекулой ГДФ. Предполагается, что молекулы тубулина способны диффундировать в растворе.

Перейдем к построению кинетических уравнений, описывающих динамику микротрубочек в растворе. Будем полагать, что миграция плюсконцов, в силу жесткости микротрубочки, происходит радиально, от центра к периферии клетки. Система кинетических уравнений, учитывающая реакции (1)—(3), в общем случае имеет вид

$$\frac{\partial p(\vec{r}, t)}{\partial t} = \beta \operatorname{div}(k_3 p - k_2 u p), \quad (4)$$

$$\frac{\partial u(\vec{r}, t)}{\partial t} = D \Delta u - k_2 u p + k_1 [\text{ГТФ}] v - k_{-1} [\text{ГДФ}] u, \quad (5)$$

$$\frac{\partial v(\vec{r}, t)}{\partial t} = D \Delta v + k_3 p - k_1 [\text{ГТФ}] v + k_{-1} [\text{ГДФ}] u, \quad (6)$$

где $u(\vec{r}, t)$ соответствует значению концентрации молекул Tu -ГТФ в точке с радиус-вектором \vec{r} в момент времени t , $v(\vec{r}, t)$ — концентрации молекул Tu -ГДФ, $p(\vec{r}, t)$ — плотности вероятности нахождения плюсконца микротрубочки, деленной на число Авогадро N_A ; k_1 [$M^{-1} \cdot c^{-1}$], k_{-1} [$M^{-1} \cdot c^{-1}$], k_2 [$M^{-1} \cdot c^{-1}$], k_3 [c^{-1}] обозначают константы скоростей соответствующих реакций (1)—(3), β [м] — стерический фактор, равный приращению длины микротрубочки при присоединении к ней одной молекулы тубулина [36]. Концентрации молекул ГТФ и ГДФ выступают в качестве параметров.

Также необходимо выполнение дополнительных условий:

$$N_A \cdot \int_V p(\vec{r}, t) dV = 1, \quad (7)$$

$$\int_V \left[u(\vec{r}, t) + v(\vec{r}, t) + \frac{|\vec{r}|}{\beta} p(\vec{r}, t) \right] dV = \frac{N_{tot}}{N_A}, \quad (8)$$

где N_{tot} обозначает общее количество молекул тубулина в системе, а V — общий объем системы. Уравнение (7) отображает полноту рассматриваемых

мой системы событий. Соотношение (8) отражает сохранение в ходе реакций (1)—(3) общего количества молекул тубулина. Инициация полимеризации предполагается происходящей на затравке, расположенной в начале координат, т. е. при $\bar{r} = 0$. Рост микротрубочек происходит в радиальном направлении. Оказывается удобным для дальнейшего рассмотрения ввести также дополнительный параметр $Tu = N_{tot} / VN_A$, равный общему количеству молекул тубулина, деленному на число Авогадро и объем системы, и по смыслу соответствующий «общей молярной концентрации» молекул тубулина в растворе.

Наиболее простым представляется псевдоодномерное приближение, рассматриваемое в области, параллельной оси абсцисс, с сечением S и длины L . Рост микротрубочек происходит в положительном направлении вдоль оси абсцисс. Правая граница $x = L$ соответствует местоположению клеточной стенки, которая считается непроницаемой. Система кинетических уравнений (4)—(8) в этом случае принимает вид

$$\frac{\partial p(x,t)}{\partial t} = \beta \frac{\partial}{\partial x} (k_3 p - k_2 u p), \quad (9)$$

$$\frac{\partial u(x,t)}{\partial t} = D \frac{\partial^2 u}{\partial x^2} - k_2 u p + k_1 [\Gamma T \Phi] v - k_{-1} [\Gamma D \Phi] u, \quad (10)$$

$$\frac{\partial v(x,t)}{\partial t} = D \frac{\partial^2 v}{\partial x^2} + k_3 p - k_1 [\Gamma T \Phi] v + k_{-1} [\Gamma D \Phi] u, \quad (11)$$

$$N_A S \cdot \int_0^L p(x,t) dx = 1, \quad (12)$$

$$\int_0^L \left[u(x,t) + v(x,t) + \frac{x}{\beta} p(x,t) \right] dx = \frac{N_{tot}}{N_A}. \quad (13)$$

5. Диаграмма состояния

Анализ системы уравнений (9)—(13) проводился численно [68]. Результаты расчетов представлены на диаграмме, изображенной на рис. 6. По оси абсцисс отложена безразмерная общая концентрация молекул тубулина в системе:

$$Tu^0 = \frac{k_2}{k_3} Tu, \quad (14)$$

а по оси ординат — величина

$$\gamma = \frac{k_{-1} [\Gamma D \Phi]}{k_1 [\Gamma T \Phi]}. \quad (15)$$

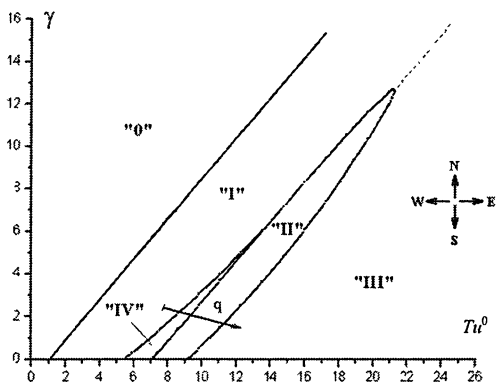


Рис. 6. Диаграмма состояний, построенная в соответствии с системой уравнений (9)—(13). Линии границ разбивают параметрическую плоскость на пять областей: зоны стационарных режимов «0», «I», «II», зону нестационарных режимов «III», а также метастабильную зону «IV».

Выяснилось, что параметрическая плоскость (Tu^0, γ) (см. рис. 6) разбивается семейством найденных нами кривых на пять областей, каждой из которых соответствует свой динамический режим. Области значений параметров, обозначенной на диаграмме символом «0», соответствует наличие в системе единственного решения в виде унимодального устойчивого распределения микротрубочек по длинам с выраженным максимумом на левой границе. Фактически это соответствует существованию в системе только затравки (центросомы) и полному отсутствию полимеризовавшихся микротрубочек.

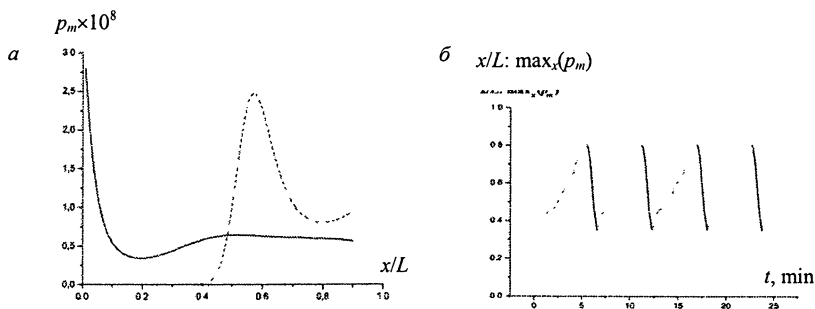


Рис. 7. Решения, соответствующие областям «I» (сплошная линия) и «II» (пунктирная линия) на диаграмме состояния (а). Изменение во времени положения максимума нестационарного распределения плотности плюс-концов микротрубочек, характерное для режимов, соответствующих области «II» (б): периоды медленной полимеризации (пунктирная линия) чередуются с волнами быстрой деполимеризации (сплошная линия).

Областям решений «I» и «III» соответствуют стационарные устойчивые решения (рис. 7, а). В области значений параметров «I» распределение микротрубочек по размерам характеризуется наличием отчетливо выраженного пика на левом крае (часть коротких микротрубочек у центросомы) и локального максимума, расположенного на существенном удалении от левого края (остальная часть, на удалении от центра клетки) (см. рис. 7, а). В области параметров «III» характерные распределения микротрубочек по размерам, как видно из рис. 7, а, имеют минимум на левой границе ($x = 0$), экстремум примерно в центре рассматриваемой области и локальный максимум на границе $x = L$. Это говорит о том, что при значениях параметров из данной области в системе должны доминировать протяженные микротрубочки, сопоставимые по размерам с размером клетки. Фактически данное решение отвечает распределениям микротрубочек, наблюдающимся на всех стадиях клеточного цикла, за исключением митоза.

Область «II» соответствует нестационарным режимам, проявляющимся в виде периодически бегущих справа налево волн деполимеризации микротрубочек, чередующихся с фазами их медленного полимеризационного роста (рис. 7, б). На границе области «II» имеет место параметрическая неустойчивость стационарных решений.

Представляется, что отвечающие данной области параметров режимы наблюдаются в динамике реальных микротрубочек как *in vitro* [43, 46], так и *in vivo* на определенных стадиях клеточного деления [60, 61].

Для параметров из области «II» рассмотрим поведение изображающей систему точки на плоскости, где по оси абсцисс отложен «центр масс» плотности вероятности нахождения плюс-концов, а по оси ординат — концентрации тубулина-ГТФ (рис. 8). Видно, что на указанной плоскости имеет место предельный цикл Пуанкаре. То есть в системе явным образом реализуются автоколебания! Более тщательный анализ показывает, что по мере роста микротрубочек (движение по нижней ветви цикла) «центр масс» концентрации тубулина-ГТФ начинает смещаться к центру нуклеации (центросоме). Это соответствует обеднению концентрации способного к полимеризации тубулина в области правой границы, образованию там «провала» за счет диффузионного размыва. Поэтому в момент, когда микротрубочки достигают правой границы, локальное пресыщение в области плюс-конца приводит к тому, что запускается процесс деполимеризации (верхняя ветвь на рис. 8).

В области «IV» рассматриваемая система имеет два типа решений. Наряду со стационарными метастабильными решениями, имеют место и нестационарные решения, схожие с наблюдаемыми в области «II». В зависимости от истории эволюции значений параметров в системе может установиться либо стационарный, либо нестационарный режим. Если параметры (Tu^0 , γ) будут медленно изменяться так, что на диаграмме состояния (см. рис. 6) изображающая точка будет плавно перемещаться из

области «I» в область «IV», то система будет оставаться в стационарном метастабильном состоянии до тех пор, пока параметры не пересекут границу между областями «IV» и «II». Пересечение последней повлечет за собой «жесткое» возбуждение пространственных осцилляций и волн, характерных для всей области «II». Примечательно, что при изменении параметров в обратном направлении, т. е. движении изображающей точки из зоны «II» в зону «IV», система сохранит нестационарное поведение вплоть до границы между зоной «IV» и зоной «I». Иными словами, в области значений параметров «IV» система демонстрирует отчетливо выраженное гистерезисное поведение.

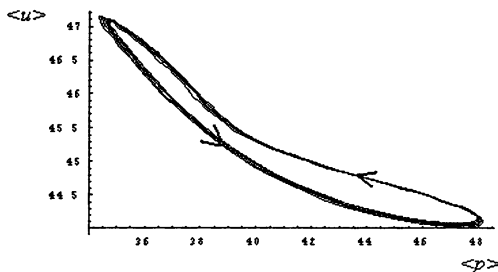


Рис. 8. Зависимость от времени «центров масс» концентрации тубулина-ГТФ и плотности плюс-концов при величинах параметров, соответствующих зоне «II».

Качественно поведение системы показано на рис. 9, а. Вдоль оси абсцисс отложено изменение параметров Tu^0 и γ вдоль вектора \mathbf{q} , проходящего через четыре разные области на диаграмме состояний (см. рис. 6). По оси ординат отложен средний радиус кривизны $\langle R \rangle$ решения дискретного аналога системы (9)—(13) $p_m(t)$, $u_m(t)$, $v_m(t)$, $m \in \{1, M\}$, представляющую однопараметрическую кривую в $3M$ -мерном пространстве [62]. Для режимов, характерных для областей параметров «I» и «III» на диаграмме состояний, справедливо $\langle R \rangle = 0$. Было обнаружено, что при проходе из области «III» в область параметров «II» в системе мягко возбуждаются автоколебания. О величине амплитуды соответствующего предельного цикла можно судить по среднему значению радиуса кривизны $\langle R \rangle$ за период. При антипараллельном вектору \mathbf{q} движению из зоны «III» в зону «II» амплитуда $\langle R \rangle$ растет как квадратичный корень величины за критичности (расстояния от $\mathbf{q} = \mathbf{q}^*$, т. е. границы областей «II» и «III», до текущей изображающей точки на диаграмме состояний). Иными словами, при переходе системы через границу областей «III» и «II» имеет место критическая бифуркация Андронова — Хопфа [20].

При антипараллельном вектору \mathbf{q} изменении параметров в точке $\mathbf{q} = \mathbf{q}_1$ происходит бифуркация удвоения цикла (рис. 9, б). Однако перехода к каскаду удвоений Фейгенбаума не наблюдается [63], так как далее,

при $q = q_2$, наблюдается бифуркация схлопывания «удвоившегося ранее цикла» к однопериодичному решению (уменьшение периода вдвое), которое продолжает существовать при дальнейшем изменении параметров вплоть до границы областей «IV» и «I».

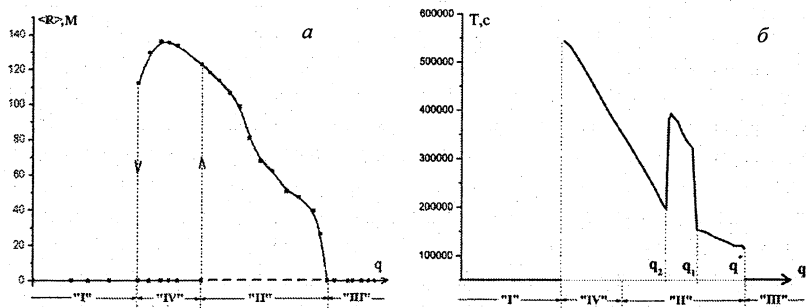


Рис. 9. Зависимости: *a* — среднего радиуса кривизны $\langle R \rangle$ траектории решения системы (9)—(13); *б* — периода предельного цикла T от изменения параметров Tu^0 и γ вдоль вектора q (см. рис. 6).

Из рис. 9, *a* видно, что в области «I» величина $\langle R \rangle$ равняется нулю, что соответствует устойчивому стационарному решению. Дальнейшее медленное изменение параметров вдоль вектора q оставляет решение системы на нижней ветви $\langle R \rangle = 0$ вплоть до границы с областью «II». Как уже отмечалось, пересечение этой границы приводит к жесткому возбуждению колебаний, что выражается в скачкообразном увеличении величины $\langle R \rangle$. Система переходит на верхнюю ветвь (см. рис. 9, *a*), отвечающую ненулевым значениям $\langle R \rangle$. Отметим, что во всей области значений параметров «II» стационарное решение системы (9)—(13) существует, но оно неустойчиво (обозначено горизонтальной пунктирной линией на рис. 9, *a*).

При изменении параметров из области «II» в противоположном вектору q направлении система остается на верхней ветви до тех пор, пока изображающая точка не пересечет границу областей «IV» и «I». При пересечении этой границы происходит потеря устойчивости нестационарного решения, и система возвращается на нижнюю ветвь. Таким образом, история изменения параметров определяет ветвь, на которой будет находиться система при значении параметров из области «IV».

Наличие целой области метастабильных состояний у системы регуляции микротубулинового цитоскелета может играть важную биологическую роль. При определенных условиях микротубулиновый цитоскелет находится в покоящемся, но «взведенном» состоянии [41]. Небольшого

внешнего воздействия, иногда укола клетки иглой, достаточно для мгновенного перехода цитоскелета к масштабной перестройке.

6. Классификация цитостатических воздействий

Построенная диаграмма состояния может быть использована для качественного анализа механизмов влияния различных биохимических агентов на динамику цитоскелета. С кинетической точки зрения их действие на рассматриваемую систему наиболее часто проявляется в изменении эффективных констант скоростей реакций [32, 64, 65]. В рамках развиваемого подхода изменение констант, в свою очередь, влечет за собой изменение параметров модели и сдвиг изображающей систему точки на диаграмме состояний. В табл. 1 представлены сведения об изменениях ключевых параметров модели Tu^0 и γ при воздействии на систему ряда биохимических регуляторных факторов. Стрелка вверх соответствует увеличению значения параметра при действии соответствующего фактора, вниз — уменьшению.

Таблица 1

Название фактора	k_1 [ГТФ]	k_2	k_3	k_{-1} [ГДФ]	Tu	$\frac{[ГТФ]}{[ГДФ]}$	Tu^0	γ	Класс
Таксол, доцетаксел	—	↑	↓	—	—	—	↑	—	E
Тау, MAP2, MAP4	—	↑	↓	—	—	—	↑	—	E
Колхицин, винбластин	—	↓	↑	—	—	—	↓	—	W
Ор18/статмин	—	—	↑	—	↓	—	↓	—	W
Митохондриальные факторы	↑	—	—	↓	—	↑	—	↓	S
Факторы ГТФ-ГДФ-обмена (GEFs)	↑	—	—	↓	—	↑	—	↓	S
Малые ГТФазы (rho, rap, ras)	↓	—	—	↑	—	↓	—	↑	N
Активирующие ГТФазы белки (GAPs)	↓	—	—	↑	—	↓	—	↑	N

Из табл. 1 видно, что рассмотренные нами факторы могут быть разбиты на две группы. Одни оказывают влияние на параметр Tu^0 , а дру-

гие — на γ . Каждая из групп, в свою очередь, разбивается на два класса в зависимости от того, к увеличению или уменьшению соответствующего параметра приводит действие входящих в данный класс агентов. В рамках модели действие агентов сводится к смещению изображающей точки на диаграмме состояний в направлении, параллельном осям, поэтому удобно соотнести обозначения указанных четырех классов с направлениями, принятыми для указания сторон света (см. рис. 6). К первому классу, обозначаемому буквой E, относятся вещества (таксол, доцетаксел, см. табл. 1), вызывающие смещение изображающей точки на восток. Аналогичным образом колхицин и винбластин могут быть отнесены к классу W, а митохондриальные факторы и факторы ГТФ-ГДФ-обмена — к классу S и так далее.

Известно, что вещества, оказывающие цитостатическое действие (цитостатики), нарушают нормальное прохождение клеткой митоза [32, 66]. Эти агенты подавляют высокую динамическую лабильность микротрубочек, необходимую для пространственного разделения хромосом. Таким образом, цитостатики блокируют ту фазу митоза, в которой клетка находилась бы в области параметров, отвечающей нестационарным режимам, т. е. области «II».

В рамках предложенного подхода зона «II» соотносится с состояниями, в которых тубулиновый цитоскелет демонстрирует макроскопические трансформации. Выше отмечалось, что клетка на определенных стадиях митоза должна проходить через области нестационарного поведения. Вследствие этого понятно, что более эффективными следует считать те цитостатические препараты (либо их сочетания), которые при прочих равных условиях вызывают большее смещение изображающей точки на диаграмме состояния из областей нестационарного поведения микротрубочек. Кратчайшим путем является смещение вдоль нормали к оси области «II» (см. рис. 6). Например, за счет одновременного смещения вверх и влево (что характерно для совместного действия веществ из классов W и N), или же вниз и вправо (при сочетании веществ S- и E-классов). Это означает, что при использовании указанных веществ в приведенных сочетаниях требуемый цитостатический эффект будет достигаться при меньших дозировках. Откуда непосредственно следует, что сопутствующая токсичность противоопухолевых цитостатических препаратов при одновременном их применении также будет снижена. В табл. 2 представлены данные о совместном действии веществ (сочетаемость), принадлежащих различным из перечисленных классов (знаком + обозначены сочетания, при которых имеет место эффект взаимного усиления воздействия препаратов. Прочерк обозначает отсутствие эффекта).

Взаимное усиление (синергизм) действия различных цитостатических агентов достигается лишь в тех случаях, когда имеет место влияние одного агента на «тубулиновую» составляющую (т. е. на изменение ско-

рости реакций полимеризации/деполимеризации), а другого — на «энергетическую», т. е. на баланс между концентрациями молекул ГТФ и ГДФ.

Таблица 2

	N	S	E	W
N	+	-	-	+
S	-	+	+	-
E	-	+	+	-
W	+	-	-	+

7. Дискуссия

В связи с биологическими и медицинскими приложениями проблема регуляции динамических трансформаций микротубулинового цитоскелета представляет большой научный и практический интерес [33, 41]. К настоящему времени исследован широкий класс биохимических веществ, обладающих цитостатическим регуляторным действием [32, 34]. Разработанный нами подход позволил расклассифицировать регуляторные воздействия на четыре основных типа. При этом открылась возможность, опираясь на диаграмму состояния микротубулинового цитоскелета, судить об эффективности совместного действия разных регуляторов (см. табл. 2).

Микротубулиновый цитоскелет клетки высокочувствителен к условиям проведения эксперимента. Наибольшие экспериментальные затруднения возникают при отыскании условий, при которых проявляется динамическая нестабильность микротрубочек [42, 49]. В свете проведенного анализа эти наблюдения становятся понятными. Дело в том, что области, отвечающие нестационарным режимам на построенной нами диаграмме состояния, представляют собой узкие полосы (см. рис. 6). Поэтому случайное попадание параметров в эти обладающие малой площадью области маловероятно.

Намного проще направленная постановка опытов по регистрации нестабильностей в динамике микротрубочек в системах *in vitro*. Достаточно задать высокую начальную концентрацию тубулина и ГТФ и дожидаться возбуждения колебаний в ходе естественного понижения уровня ГТФ [44]. Отметим, что снижение уровня ГТФ со временем на диаграмме состояния (см. рис. 6) соответствует движению изображающей точки снизу вверх. При этом, как несложно видеть, система микротрубочек на границе между зонами «III» и «II» должна претерпевать переход из стационарного состояния в нестационарное. Таким образом, ясно, что в системах *in vitro* появление нестабильностей в динамике микротрубочек обусловлено рождением предельного цикла (см. рис. 8). В терминах теории катастроф, при

переходе системы через границу областей «III» и «II» имеет место за критическая бифуркация Андронова — Хопфа [21].

Природа нестационарных пульсаций микротрубочек, при которых стадии полимеризационного роста внезапно сменяются стадиями деполимеризации, интенсивно обсуждается более 25 лет [42, 46]. Наиболее широко дебатруется гипотеза о существовании в окрестности плюс-конца микротрубочки так называемой ГТФ-крышки (GTP-cap). Согласно данной гипотезе, гидролиз присоединенной к тубулину молекулы ГТФ происходит не сразу после присоединения очередного димера к микротрубочке, а спустя некоторое время [46]. Образующийся слой молекул тубулина-ГТФ в окрестности плюс-конца предотвращает деполимеризацию и стабилизирует всю микротрубочку [46]. Переход от роста микротрубочки к ее деполимеризации связан с исчезновением ГТФ-крышки.

Нельзя исключить, что в определенных экспериментальных условиях задержка гидролиза ГТФ, в принципе, способна вызвать нестабильность в динамике микротрубочек [42, 46]. Однако проведенный выше анализ показал, что, и не прибегая к этой гипотезе (опираясь только на реакционно-диффузионные представления в рамках модели (9)—(13)), удается получить решения, соответствующие динамическим нестабильностям тубулиновых микротрубочек. Попутно заметим, что в рамках нашего подхода сам собой снимается активно обсуждаемый в альтернативных подходах вопрос о виде частотного спектра случайных событий «переключения» микротрубочек с роста на деполимеризацию и обратно [56].

В рамках развитого нами подхода динамика плюс-конца микротрубочки определяется исключительно реакциями сорбции/десорбции молекул тубулина. При этом скорость реакции сорбции зависит от локальной концентрации молекул тубулина, связанных с ГТФ. В этом смысле изменение размера микротрубочки регулируется реакционно-диффузионными процессами в окружающей системе. Проведенный анализ показал, что рост микротрубочки связан с распространением концентрационных автоволн молекул тубулина в растворе. Бегущая слева направо автоволна ГТФ вызывает рост микротрубочек, а распространяющаяся справа налево волна ГДФ сопровождает деполимеризацию. По аналогии с упоминавшимися ранее диффузионно-лимитированными процессами (DLA [51]) можно сказать, что критические явления в динамике микротрубочек относятся к энергетически лимитированным процессам агрегации-дисагрегации (ELAD).

Вопрос о роли концентрационных автоволн в динамике микротрубочек, насколько нам известно, не рассматривался ранее теоретически. Однако в пользу влияния концентрационных градиентов на динамику микротрубочек *in vivo* свидетельствуют недавно опубликованные экспериментальные данные [35, 57]. Это позволяет надеяться, что полученные нами теоретические результаты могут оказаться небезытересными специалистам в области клеточной биологии.

Стоит заметить, что в предпринятой в данной работе попытке описания критических явлений в динамике микротрубочек развиваются общие физико-химические положения теории диссипативных структур [52]. Распределение реагентов в растворе, в котором формируются микротрубочки, описывается в рамках реакционно-диффузионного приближения. А сами микротубулиновые нити трактуются как «преципитат» из жидкой фазы, т. е. как выпавшее в осадок вещество. В рамках такого подхода проблема роста микротрубочек по своей физико-химической сути предстает как проблема конденсации молекул тубулина из одной фазы в другую. Несмотря на то что построенная нами кинетическая реакционно-диффузионно-преципитационная (РДП) модель, вне всякого сомнения, является сильно упрощенной и ухватывает лишь часть из наиболее характерных черт реальной системы, на ее основе удалось построить диаграмму состояния рассматриваемой системы.

В физике переход вещества из одной формы пространственного упорядочения в другую принято рассматривать в рамках теории фазовых переходов с термодинамических позиций [69]. Традиционный взгляд на явления смены агрегатного состояния связан с выяснением *условий термодинамической устойчивости* фаз [73]. Сосуществование двух фаз возможно лишь в случае, если на границе раздела фаз химические потенциалы равны. О спектре агрегатных состояний системы и возможных переходах между ними при изменении термодинамических параметров принято судить по виду «диаграммы состояния» [73]. Поэтому построение таких диаграмм является одной из ключевых задач в физике конденсированного состояния. Надо сказать, что диаграммы состояния широко используются и в физической химии, например, при изучении термодинамических условий выпадения в осадок растворенных в жидкой фазе солей [74]. Применительно к рассмотренной нами ELAD-системе уместно говорить о неравновесных переходах, поскольку в ходе формирования и роста микротрубочек имеет место энергопотребление — расходуются молекулы ГТФ.

Ввиду того что тубулиновые микротрубочки способны формироваться и в реконструированных биохимических системах, нам представлялось естественным рассматривать их с позиций современной неравновесной физической химии [72]. С этой точки зрения реконструированная система может рассматриваться как состоящая из двух взаимодействующих между собой подсистем (рис. 10). В «концентрационной» подсистеме молекулы тубулина, ассоциированные с ГТФ или ГДФ, способны индивидуально перемещаться в пространстве. Напротив, в «преципитационной» подсистеме отдельные молекулы тубулина плотно упакованы в цилиндрическую 13-заходную спиральную структуру (см. рис. 2). В этом смысле тубулиновая микротрубочка представляет собой элемент конденсированной фазы. При этом условная «граница раздела фаз» представляет собой

кольцо с внешним размером 24 нм и внутренним 13 нм, лежащее в торцевом сечении микротрубочки, в зоне местоположения ее плюс-конца. Понятно, что указанная «граница раздела фаз» должна быть в равновесной ситуации неподвижной (испытывать только флуктуационное дрожание), так как при этом скорости прямых и обратных процессов (сорбции/десорбции) должны быть равны.

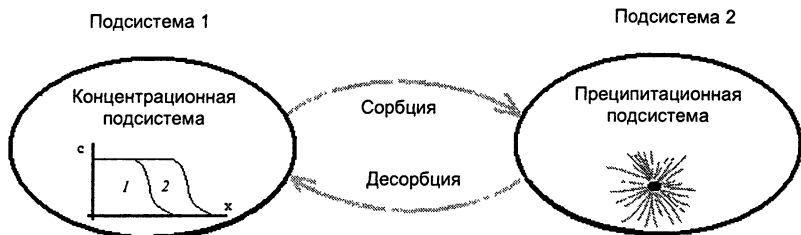


Рис. 10. Принципиальная схема рассмотренной реакционно-диффузионно-преципитационной системы.

Однако взаимное действие указанных выше подсистем в неравновесных условиях может быть, вообще говоря, и нескомпенсированным. Например, в условиях, когда доминируют процессы сорбции молекул тубулина, будет иметь место рост микротрубочки. Как следствие, будет смещаться граница раздела фаз. При доминировании процессов десорбции, трубочки будут укорачиваться, а концентрационная фаза насыщаться молекулами тубулина-ГДФ. Действие подсистем друг на друга, как выяснилось, не всегда бывает компенсаторно сбалансированным в каждый момент времени. Наряду с режимами, которым отвечают стационарные распределения микротрубочек по длинам, имеющими место в зонах «0», «I» и «III» (см. рис. 6); возможны и нестационарные режимы, которым отвечают областям «II» и «IV» на параметрической диаграмме состояния (см. рис. 6). Как мы видели, в последнем случае поведение системы определяется наличием в фазовом пространстве предельного цикла А. Пуанкаре. Представляется важным подчеркнуть, что при этом масштабные периодические изменения длины микротрубочек оказываются сопряженными с концентрационными автоволнами в растворе. В физике широко известно, что при приближении к критическим состояниям имеет место нарастание крупномасштабных флуктуаций и увеличение радиуса корреляции (второго корреляционного момента) [75, 76]. В рассмотренной нами неравновесной системе увеличение пространственно-временной корреляции в критических условиях обеспечивается за счет автоволновых процессов и сопряженных с ними процессов полимеризации-деполимеризации.

В определенной области параметров (см. «II», «IV», рис. 6) имеет место своеобразная взаимная синхронизация в поведении подсистем, кото-

рая проявляется в виде взаимообусловленной пространственно-временной ритмики. Рождение предельного цикла и связанных с ним пространственно-временных форм трансформации элементов тубулинового цитоскелета идет в полном соответствии с открытым А. А. Андроновым механизмом (см. рис. 9, а).

Выражаем благодарность профессору И. А. Воробьеву за ценные советы и замечания.

Работа выполнена при поддержке РФФИ, грант № 04-04-49423а.

Литература

1. *Helmholtz H.* Ueber die Fortpflanzungsgeschwindigkeit der Nervenreizung // *Annalen der Physik und Chemie.* 1850. V. 79. P. 329—330.
2. *Hertz H.* Allgemeine Theorie der Elektrizität und des Magnetismus // *Die Fortschritte der Physik.* 1880. V. 36. P. 765—767.
3. *Колмогоров А. Н., Петровский И. Г., Пискунов Н. С.* Исследование уравнения диффузии, соединенной с возрастанием количества вещества, и его применение к одной биологической проблеме // *Бюлл. МГУ. Сер. А.* 1937. Вып. 6. С. 1—16.
4. *Зельдович Я. Б., Франк-Каменецкий Д. А.* Теория теплового распространения пламени // *Журн. физической химии.* 1938. Т. 12, вып. 1. С. 100—105.
5. *Wiener N., Rosenblueth A.* The Mathematical Formulation of the Problem of Conduction of Impulses in a Network of Connected Excitable Elements, Specifically in Cardiac Muscle // *Arch. Inst. Cardiol Mex.* 1946. V. 16. P. 205—265 [Русский перевод: *Винер Н., Розенблют А.* Проведение импульсов в сердечной мышце: Математическая формулировка проблемы проведения импульсов в сети связанных возбудимых элементов, в частности в сердечной мышце // *Кибернетический сборник.* 1961. Вып. 3. С. 7—86].
6. *Hodgkin A. L., Huxley A. F.* A quantitative description of membrane current and its application to conduction and excitation in nerve // *J. Physiology.* 1952. V. 117. P. 500—544.
7. *Scott A.* Active and Nonlinear Wave Propagation in Electronics. N. Y.: Wiley, 1970 [Русский перевод: *Скотт Э.* Волны в активных и нелинейных средах в приложении к электронике. М.: Советское радио, 1977. 368 с.].
8. *Жаботинский А. М.* Концентрационные колебания. М.: Наука, 1974.
9. *Колебания и бегущие волны в химических системах / Под ред. Р. Филд, М. Бургер.* М.: Мир, 1988. 720 с.
10. *Пригожин И.* Время, структуры и флуктуации // *УФН.* 1980. Т. 131, вып. 2. С. 185—207.
11. *Автоволновые процессы в системах с диффузией / Под ред. М. Т. Греховой.* Горький: ИПФ АН СССР, 1981.
12. *Васильев В. А., Романовский Ю. М., Яхно В. Г.* Автоволновые процессы. М.: Наука, 1987. 240 с.
13. *Wolpert L.* Principles of development. Oxford: University Press, 1998.
14. *Koch A J, Meinhardt H.* Biological pattern formation — from basic mechanisms to complex structures // *Rev. Modern Physics.* 1994. V. 66. P. 1481—1507.
15. *Turing A. M.* The chemical basis of morphogenesis // *Phil. Trans. R. Soc. London.* 1952. V. 237. P. 37—72.
16. *Thom R.* Mathematical models of morphogenesis. N. Y.: Halsted Press, 1982 [Русский перевод: *Том Р.* Математические модели морфогенеза. М.: РХД, 2006. 136 с.].
17. *Иваницкий Г. П., Кринский Б. И., Сельков Е. Е.* Математическая биофизика клетки. М.: Наука, 1978. 308 с.
18. *Leichtleier J., Girard S., Peralta E., Clapham D.* Spiral calcium wave propagation and annihilation in *Xenopus laevis* oocytes // *Science.* 1991. V. 252. P. 123—126.

19. *Gorbunova Y. V., Spitzer N. C.* Dynamic interactions of cyclic AMP transients and spontaneous Ca²⁺ spikes // *Nature*. 2002. V. 418. P. 93—96.
20. *Haken H.* Synergetics, an introduction. Nonequilibrium Phase-Transitions and Self-Organization in Physics, Chemistry and Biology. Springer, 1977 [Русский перевод: *Хаген Г.* Синергетика. М.: Мир, 1980].
21. *Андронов А. А., Витт А. А., Хайкин С. Э.* Теория колебаний. 2-е изд. М.: Физматгиз, 1959. 926 с.
22. *Горелик Г. С.* Колебания и волны. М.: Физматгиз, 1959.
23. *Боголюбов Н. Н., Митропольский Ю. А.* Асимптотические методы в теории нелинейных колебаний. М.: Наука, 1974.
24. *Вайнберг М. М., Треногин В. А.* Теория ветвления решений нелинейных уравнений. М.: Наука, 1969. 529 с.
25. *Sattinger D. H.* Topics in stability and bifurcation theory. Lecture notes in Mathematics, N309. N. Y.: Springer-Verlag, 1973.
26. *Данилов Ю. А.* Лекции по нелинейной динамике. Элементарное введение. М.: Постмаркет. 2001. 184 с.
27. *Mandelbrot B. B.* The fractal geometry of nature. San Francisco: Freeman, 1982.
28. *Де Жен П.* Идеи «скейлинга» в физике полимеров. М.: Мир, 1987. 368 с.
29. *Wang N., Naruse K., Stamenovic D., Fredberg J. J., Mijailovich S. M., Tolic-Norrelykke M., Polte T., Mannix R., Ingber D. E.* Mechanical behavior in living cells consistent with the tensegrity model // *PNAS*. 2001. V. 98. P. 7765—7770.
30. *Халмош П.* Теория меры. М.: Факториал. 2003. 256 с.
31. *Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P.* *Molecular Biology of Cell* N. Y.: Garland Science, 2002.
32. *Wilson L., Panda D., Jordan M.* Modulation of microtubule dynamics by drugs: a paradigm for the actions of cellular regulators // *Cell structure and function*. 1999. V. 24. P. 328—335.
33. *Potapova T. A., Daum J. R., Pittman B. D., Hudson J. R., Jones T. N., Sattinover D. L., Stukenberg P. T., Gorbsky G. J.* The reversibility of mitotic exit in vertebrate cells // *Nature*. 2006. V. 440. P. 954—958.
34. *Niethammer P., Bastiaens P., Karsenti E.* Stathmin-tubulin interaction gradients in motile and mitotic cells // *Science*. 2004. V. 303. P. 1862—1866.
35. *Caudron M., Bunt G., Bastiaens P., Karsenti E.* Spatial coordination of spindle assembly by chromosome-mediated signaling gradients // *Science*. 2005. V. 309. P. 1373—1376.
36. *Li H., DeRosier D. J., Nicholson W. V., Nogales E., Downing K. H.* Microtubule structure at 8 Å resolution // *Structure*. 2002. V. 10. P. 1317—1328.
37. *Janson M. E., Dogterom M.* A bending mode analysis for growing MT // *Biophys. J.* 2004. V. 87. P. 2723—2736.
38. *Воробьев И. А., Григорьев И. С., Бориси Г. Г.* Динамика микротрубочек в культивируемых клетках // *Онтогенез*. 2000. Т. 31, № 6. С. 420—428.
39. *Holy T. E., Leibler S.* Dynamic instability of microtubules as an efficient way to search in space // *PNAS*. 1994. V. 91. P. 5682—5685.
40. *Grigoriev I., Borisy G., Vorobjev I.* Regulation of microtubule dynamics in 3T3 fibroblasts by Rho family GTPases // *Cell Motility and the Cytoskeleton*. 2006. V. 63. P. 29—40.
41. *Karsenti E., Vernos I.* The mitotic spindle: a self-made machine // *Science*. 2001. V. 294. P. 543—547.
42. *Mitchison T. J., Kirschner M. W.* Dynamic instability of microtubule growth // *Nature*. 1984. V. 312. P. 237—242.
43. *Carlier M. F., Melki R., Pantaloni D., Hill T. L., Chen Y.* Synchronous oscillations in microtubule polymerization // *PNAS*. 1987. V. 84. P. 5257—5261.
44. *Melki R., Carlier M. F., Pantaloni D.* Oscillations in microtubule polymerization: the rate of GTP regeneration on tubulin controls the period // *EMBO J.* 1988. V. 7. P. 2653—2659.
45. *Fygenon D. K., Braun E., Libchaber A.* Phase diagram of microtubules // *Phys. Rev. E.* 1994. V. 50 P. 1579—1588.

46. *Desai A., Mitchison T. J.* Microtubule polymerization dynamics // *Annu. Rev. Cell Dev. Biol.* 1997. V. 13. P. 83—117.
47. *Janosi I. M., Chretien D., Flyvbjerg H.* Structural microtubule cap: stability, catastrophe, rescue, and third state // *Biophys J.* 2002. V. 83. P. 1317—1330.
48. *Dogterom M., Leibler S.* Physical aspects of the growth and regulation of microtubule structures // *Phys. Rev. Lett.* 1993. V. 70. P. 1347—1350.
49. *Caudron N., Valiron O., Usson Y., Valiron P., Job D.* A reassessment of the factors affecting microtubule assembly and disassembly in vitro // *J. Mol. Biol.* 2000. V. 297. P. 211—220.
50. *Волькенштейн М. В.* Биофизика. М.: Наука, 1988. 592 с.
51. *Vicsek T.* *Fractal Growth Phenomena.* World Scientific, 1992. 400 p.
52. *Николс Г., Пригожин И.* Самоорганизация в неравновесных системах. М.: Мир, 1979.
53. *Эбелинг В.* Образование структур при необратимых процессах. Введение в теорию диссипативных структур. М.: РХД, 2004. 256 с.
54. *Koch A. J., Meinhardt H.* Biological pattern formation — from basic mechanisms to complex structures // *Rev. Modern Phys.* 1994. V. 66. P. 1481—1507.
55. *Guria G. T., Livshits M. A.* Nonequilibrium spatio-temporal selforganization due to delayed negative feedback // *Z. Phys. B. Condensed Matter.* 1982. V. 47. P. 71—77.
56. *Houchmandzadeh B., Vallade M.* Collective oscillations in microtubule growth // *Phys. Rev. E.* 1996. V. 53. P. 6320—6324.
57. *Pearson C. G., Gardner M. K., Paliulis L. V., Salmon E. D., Odde D. J., Bloom K.* Measuring nanometer scale gradients in spindle microtubule dynamics using model convolution microscopy // *Mol. Biol. Cell.* 2006. V. 17. P. 4069—4079.
58. *Ross J., Muller S., Vidal C.* Chemical waves // *Science.* 1988. V. 240. P. 460—465.
59. *Буравцев В. Н.* Периодический фазовый переход в растворе аммиака // *Журн. физической химии.* 1983. Т. 57. С. 1822—1824.
60. *Sammak P. J., Borisy G. G.* Direct observation of microtubule dynamics in living cells // *Nature.* 1988. V. 332. P. 724—726.
61. *Grill S. W., Gonczy P., Stelzer E. H., Hyman A. A.* Polarity controls forces governing asymmetric spindle positioning in the *C. elegans* embryo // *Nature.* 2001. V. 490. P. 630—633.
62. *Гельфанд И. М.* Лекции по линейной алгебре. М.: Наука, 1970. 273 с.
63. *Feigenbaum M. J.* Quantitative universality for a class of nonlinear transformations // *J. Stat. Phys.* 1978. V. 19, № 1. P. 25—52.
64. *Andersen S. S. L.* Spindle assembly and the art of regulating microtubule dynamics by MAPs and Stathmin/Op18 // *Trends Cell Biol.* 2000. V. 10. P. 261—267.
65. *Rubin C. I., Atweh G. F.* The role of stathmin in the regulation of the cell cycle // *J. Cell Biochem.* 2004. V. 93. P. 242—250.
66. *Honore S., Pasquier E., Braguer D.* Understanding microtubule dynamics for improved cancer therapy // *Cell Mol. Life Sci.* 2005. V. 62. P. 3039—3056.
67. *Nucleation theory and applications / Ed. by Schmelzer J.W.P., Ropke G., Priezhev V.R.* Dubna: JINR, 1999.
68. *Катруха Е. А., Гурия Г. Т.* Динамические нестабильности тубулинового цитоскелета. Диаграмма состояния // *Биофизика.* 2006. Т. 51, вып. 5. С. 885—893.
69. *Ландау Л. Д., Лившиц Е. М.* Статистическая физика. М.: Наука, 1964. 568 с.
70. *Геккель Э.* Красота форм в природе. СПб, 1907.
71. *Вейль Г.* Симметрия. М.: Наука, 1968. 192 с.
72. *Пригожин И., Конденди Д.* Современная термодинамика. От тепловых двигателей до диссипативных структур. М.: Мир, 2002. 464 с.
73. *Gibbs J. W.* *The Collected Works. Thermodynamics.* N. Y.: Longmans, Green and Co., 1928. V. 1.
74. *Мелвин-Хьюз Э.-А.* Физическая химия: В 2 т. М.: Изд-во иностранной литературы. 1962.
75. *Ма Ш.* Современная теория критических явлений. М.: Мир, 1980. 299 с.
76. *Балеску Р.* Равновесная и неравновесная статистическая механика: В 2 т. М.: Мир, 1978.

АЭРОИОНЫ ЧИЖЕВСКОГО И ЯВЛЕНИЕ СВЕЧЕНИЯ ВЛАЖНОГО ВОЗДУХА

В. Л. Воейков, С. Ф. Чалкин, Р. Р. Асфарамов

1. Введение

Александр Леонидович Чижевский сообщил о роли отрицательно заряженных ионов воздуха в поддержании жизненно важных функций организма более 80 лет тому назад [1]. Предложенная им искусственная аэроионизация воздуха успешно применяется как в промышленности, так и в быту [2, 3]. Надежно подтверждены полученные А. Л. Чижевским свидетельства благотворного действия отрицательных аэроионов на здоровье человека, животных и растений [4—7]. Однако, несмотря на бесспорные доказательства того, что присутствующие в природном воздухе аэроионы абсолютно необходимы для нормальной жизнедеятельности, что их дефицит во вдыхаемом человеком воздухе служит одной из существенных причин возникновения и развития хронических и инфекционных заболеваний, масштабы использования аэроионотерапии в медицине далеко не соответствуют ее эффективности.

Что же сдерживает применение аэроионов для профилактики и лечения заболеваний? Нередко представители академической биологии и медицины, которые сами аэроионизации не используют, относятся к утверждениям о важной биологической функции отрицательно заряженных ионов воздуха скептически. Во-первых, исходя из обычных представлений биохимии и физиологии, из-за ничтожности содержания отрицательно заряженных ионов в «свежем» воздухе (сотни — тысячи частиц ионов в 1 см^3 , тогда как содержание кислорода в том же объеме воздуха порядка 10^{18} молекул) рассматривать их в качестве действующего фактора не представляется возможным. А с точки зрения физики, время жизни частиц, полученных при ионизации молекул газов, столь мало, что трудно себе представить, как ионы, генерируемые ионизаторами, находящимися на расстоянии метров от объекта воздействия, могут оказывать на него хоть какое-то действие.

Со всеми этими возражениями встретился еще А. Л. Чижевский. Естественно, он задумывался и о механизме действия аэроионов, генерируемых его «электроэфлювиальными люстрами», и о том, почему они достигают своих мишеней. Чижевский пришел к мысли, что отрицательно заряженный кислород (действующее начало ионизированного воздуха) обладает высокой биологической активностью не за счет электрического заряда, а поскольку представляет собой некую активированную форму кислорода, которая запускает некие лавинообразные процессы, реализующиеся в физиологические эффекты. Предположение, что биологическое действие ионизированного воздуха как-то связано со свободными

радикалами, высказали позднее Krueger и Reed [8], а идентичность отрицательных аэроионов и супероксид-анион-радикалов была установлена только в 1992 г. [9]. Вскоре было обнаружено, что в водных средах, подвергнутых действию аэроионов, растет содержание перекиси водорода [10]. По мнению М. Н. Кондрашовой и соавт., благотворное медико-биологическое действие аэроионов обусловлено активацией ими пероксидазных окислительных реакций в диапазоне физиологической нормы [11].

На основании исследований свободнорадикального окисления кислородом биомолекул в водных системах, анализа результатов перекисно-водородной и озонотерапии, мы выдвинули гипотезу, что целительное действие аэроионов Чижевского обусловлено тем, что они выступают в роли своеобразных «искр», запускающих и поддерживающих в организме процессы с участием активных форм кислорода (АФК). Оказалось, что существенная часть потребляемого живыми организмами кислорода непосредственно восстанавливается, превращаясь в АФК — богатые энергией и метастабильные частицы, которые в норме не накапливаются, так как тут же рекомбинируют друг с другом или устраняются другими способами. В ходе таких процессов генерируется наиболее ценная форма энергии — энергия электронного возбуждения, играющая важнейшую биоэнергетическую и биоинформационную роль [12, 13]. Такие реакции склонны к самозатуханию, и тогда в организме даже при избытке молекулярного кислорода начинают накапливаться продукты неполного сгорания продуктов обмена веществ. Для поддержания нормального режима «горения» поступающий в организм кислород должен содержать хотя бы малую примесь АФК [14].

Если ответ на вопрос, почему ничтожное количество аэроионов — свободнорадикальных АФК — обладает столь высокой биологической активностью, хотя бы в первом приближении получен, второй вопрос — каким образом ионизированные частицы газов не только обладают длительным временем жизни в воздухе, но и могут переноситься от ионизаторов на значительные расстояния, практически не обсуждается. В то же время именно он послужил в 1928 г. препятствием для публикации А. Л. Чижевским статьи об аэроионизации в «Британском журнале актиотерапии и физиотерапии». Как пишет сам Чижевский, его статья была отвергнута, поскольку английские физики подвергли сомнению саму возможность того, что столь короткоживущие ионы могут достигать своих мишеней. А. Л. Чижевский в своем ответе писал: «/рецензенты/... упустили из виду один важный факт. Дело в том, что человек < > окружен «рубашкой» мельчайших частиц влаги... Как только легкий, короткоживущий ион попадает в эту область и оседает на микрочастице влаги, он становится долгоживущим...» [15]. Другими словами, он признал, что влажность воздуха играет важнейшую роль в действии аэроионов. Другие исследователи впоследствии также обращали внимание на то, что как при

очень низкой, так и при очень высокой влажности воздуха отрицательные аэроионы либо исчезают, либо теряют свою эффективность. Сейчас уже нет сомнений в том, что многие атмосферные явления, отрицательно влияющие на здоровье, связаны с исчезновением из воздуха ионизированного кислорода. К таким явлениям относятся сухие и горячие ветры, такие как «фён» в Австрии и Швейцарии, «мистраль» во Франции, «шарав» в Израиле [8]. Таким образом, биологические эффекты активированного кислорода, возникающего в газовой фазе, так или иначе связаны с влажностью воздуха, хотя в доступной нам литературе специальных исследований этого вопроса мы не обнаружили.

В настоящей работе мы сообщаем о новом явлении: об инициации сверхслабым источником УФ-фотонов процессов, в ходе которых во влажном воздухе продуцируются АФК, развиваются реакции, протекающие по принципу разветвленно-цепных, генерируется энергия электронного возбуждения, распространяющаяся по влажному воздуху на макроскопические расстояния от источника УФ-фотонов. Это явление, по видимому, не только объясняет возможность воздействия различного рода аэроионизаторов, как искусственных, так и природных, на удаленные объекты, но и ставит новые вопросы, связанные с различного рода атмосферными явлениями. Это явление было обнаружено при изучении особенностей работы нового детектора одиночных фотонов, «Биотокс 7с», предназначенного для одновременной регистрации фотонов отдельно в видимой и УФ областях спектра [16].

2. Материалы и методы

На рис. 1 представлена блок-схема прибора «Биотокс 7с». Прибор оснащен двумя фотозлектронными умножителями: ФЭУ-101, максимально чувствительным в области 380—490 (Вид-ФЭУ) нм и практически нечувствительным к фотонам с $\lambda < 300$ нм, и ФЭУ-142, предназначенным для исследований в вакуумной ультрафиолетовой (до $\lambda = 112$ нм) и ультрафиолетовой областях спектра (УФ-ФЭУ). Максимальная спектральная чувствительность (~10%) ФЭУ-142 лежит в диапазоне 200—250 нм, а чувствительность к фотонам с $\lambda \geq 400$ нм менее 0,1%. Окна ФЭУ направлены друг на друга и расположены на расстоянии 5 см. Окно ФЭУ-101 выполнено из стекла (пропускание $\lambda > 320$ нм), а ФЭУ-142 — из MgF_2 (пропускание $\lambda > 110$ нм). Диаметр окон обоих ФЭУ равен 1 см. Рабочее напряжение на ФЭУ-142 составляло 1700 В, а на ФЭУ-101 — 800 В. Фотокатоды обоих ФЭУ заземлены.

Камера для образцов представляет собой выполненную из дюралюминия цилиндрическую гильзу с двумя отверстиями на противоположных стенках (см. рис. 1). В режиме регистрации излучения от образца, помещенного во внутреннюю полость гильзы, она повернута отверстиями

к окнам ФЭУ; для смены образца гильзу поворачивают на 90° так, что окна ФЭУ закрываются сплошной стенкой гильзы для предотвращения их засветки. Поворот гильзы осуществляется вращением светонепроницаемой крышки камеры для образцов.

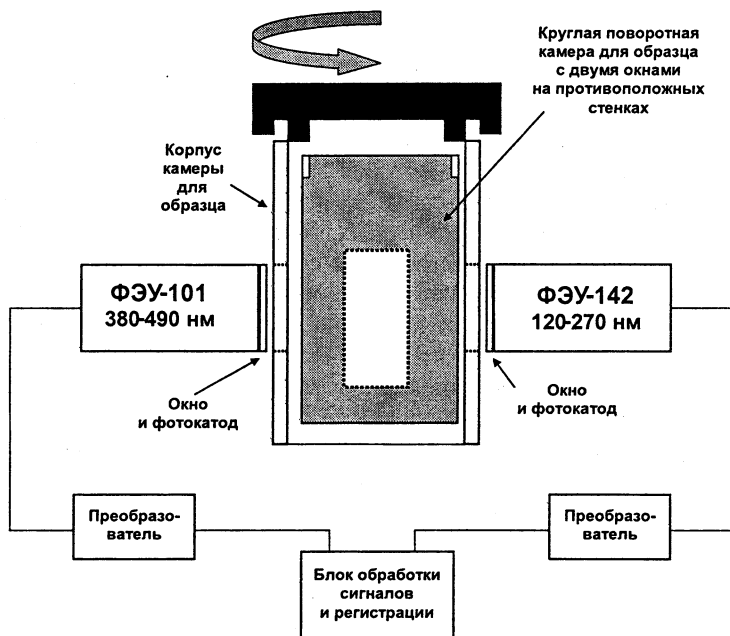


Рис. 1. Блок-схема прибора «Биотокс 7с» (масштаб не соблюден)

Влажность и температуру воздуха в камере для образцов измеряли с помощью комбинированного зонда, присоединенного к сетевому микропроцессору ИВТМ-7 МК-С (ОАО «Практик-НЦ», Москва, Зеленоград). Зонд был встроен в крышку камеры. В ней также имелись отверстия для шлангов для продувания камеры газами, которые либо предварительно осушали пропусканием через колонну, заполненную сухим силикагелем, либо увлажняли при пропускании сквозь заполненную водой емкость. В ряде экспериментов на дно камеры вне поля видимости ФЭУ устанавливали емкость с дистиллированной водой. В качестве газов использовали лабораторный воздух, баллонные газы: азот, аргон или гелий (категории не ниже «высокой чистоты») по отдельности или в смеси с лабораторным воздухом.

3. Результаты

Как видно из рис. 2, при включении УФ-ФЭУ скорость счета фотонов (ССФ) Вид-ФЭУ возрастает более чем в 3 раза по сравнению с таковой до включения УФ-ФЭУ при влажности воздуха в камере около 20%. При пропускании через камеру сухого воздуха одновременно снижается ССФ как Вид-ФЭУ, так и УФ-ФЭУ, причем ССФ УФ-ФЭУ снижается быстрее. При выключении тока сухого воздуха (отсчет времени 1500 с), когда в камеру начал диффундировать более влажный лабораторный воздух, ССФ обоими ФЭУ начинает возрастать.

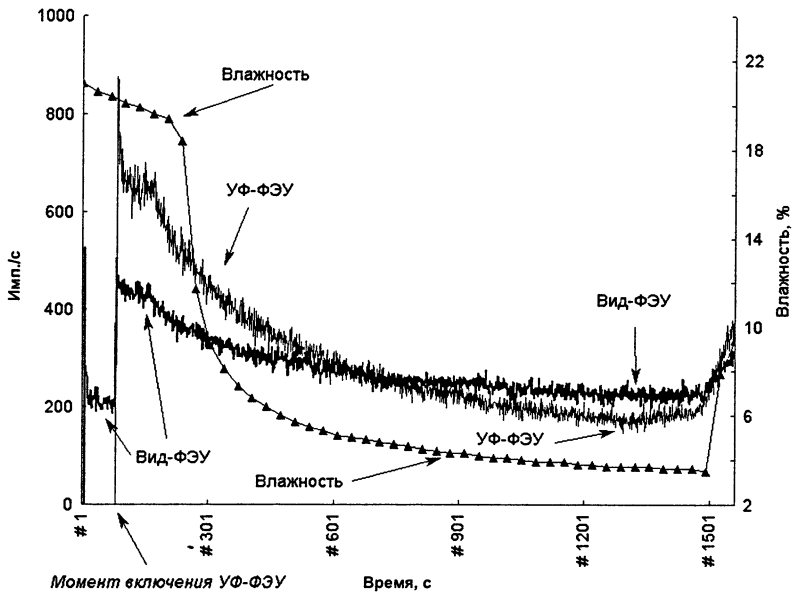


Рис. 2. Влияние УФ-ФЭУ на показания Вид-ФЭУ при наличии между ними оптического контакта в зависимости от влажности воздуха в камере для образцов. Абсцисса — время с начала эксперимента [с], левая ордината — скорость счета фотонов [имп./с], правая ордината — влажность [%].

При пропускании через камеру воздуха со 100%-ной влажностью показания обоих ФЭУ при наличии контакта между ними с увеличением влажности возрастают. При этом ССФ Вид-ФЭУ возрастает быстрее, чем ССФ УФ-ФЭУ и многократно превышает последнюю. При влажности воздуха свыше 50% ССФ Вид-ФЭУ приобретает ярко выраженный колебательный характер (рис. 3).

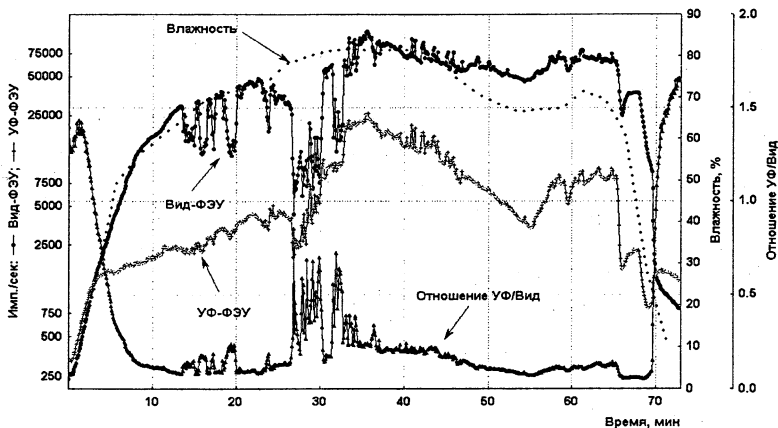


Рис. 3. Влияние повышения влажности воздуха в камере, в которую обращены окна Вид-ФЭУ и УФ-ФЭУ, на ССФ в видимой и УФ областях спектра — левая ордината и на отношение ССФ в УФ и видимой областях спектра — вторая правая ордината.

При увеличении концентрации во влажном воздухе инертных газов (гелия, аргона) ССФ обоими ФЭУ возрастает, но скорость счета УФ-фотонов возрастает особенно резко и превышает таковую при продувке камеры обычным влажным воздухом (рис. 4). В последнем секторе № 2 скорость продувки камеры влажным воздухом была снижена, что сопровождалось снижением влажности в камере и уменьшением ССФ как Вид-ФЭУ, так и УФ-ФЭУ.

Чрезвычайно интенсивные колебания ССФ УФ-ФЭУ наблюдались при значительном обогащении влажного воздуха в камере гелием. Амплитуды таких колебаний достигали сотен тысяч импульсов за 1 с, а колебания совершались с периодичностью в несколько секунд.

При выключенном УФ-ФЭУ влияния влажности воздуха на показания Вид-ФЭУ зарегистрировано не было. Это видно из графика на рис. 5. В этом опыте на дно просушенной камеры устанавливали сосуд с водой и регистрировали ССФ Вид-ФЭУ при выключенном УФ-ФЭУ в течение длительного времени, когда влажность воздуха увеличивалась за счет испарения воды. Как видно на рис. 5, на фоне установившегося в камере примерно стационарного уровня влажности воздуха около 62% сразу после включения УФ-ФЭУ наблюдается резкий скачок ССФ Вид-ФЭУ, сопровождаемый интенсивными колебаниями ССФ как Вид-ФЭУ, так и УФ-ФЭУ. Включение УФ-ФЭУ сопровождается еще двумя интересными явлениями. Во-первых, сразу наблюдается постепенный рост температуры (почти на 2 градуса за последующие 3 часа), а также рост стационарного уровня влажности, по-видимому, связанный с повышением температуры. В ходе всего эксперимента температура в помещении была постоянной и составляла 25 °С.

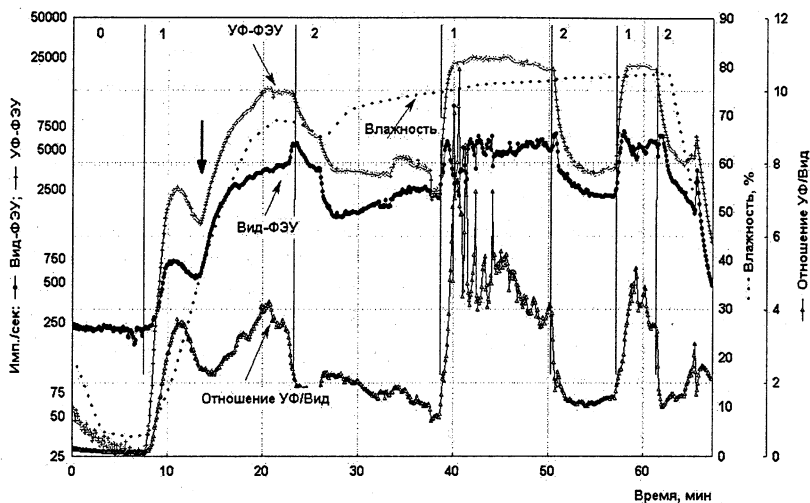


Рис. 4. Влияние последовательной продувки камеры влажным воздухом, обогащенным аргоном (сектора №1) и влажным воздухом (сектора №2) на ССФ. Сектор №0 — продувка сухим воздухом, толстая стрелка в первом секторе №1 — увеличение скорости продувки камеры. В последнем секторе №2 скорость продувки камеры влажным воздухом была снижена, что сопровождалось снижением влажности в камере.

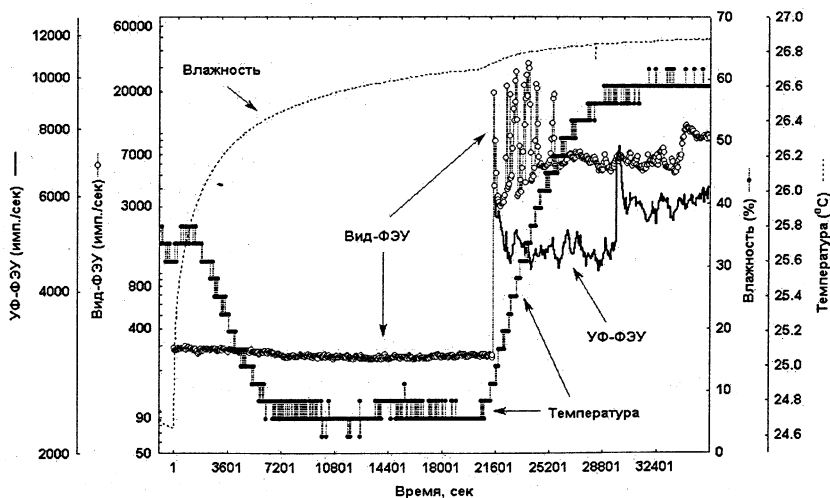


Рис. 5. Отсутствие влияния влажности на показания Вид-ФЗУ при выключенном УФ-ФЗУ (до отметки времени 21600 с). Влияние работы УФ-ФЗУ на температуру и влажность воздуха в камере. Все графики представлены после их сглаживания по 60 точкам.

Следует отметить, что на первом этапе эксперимента, когда работал только Вид-ФЭУ, температура в камере несколько снижалась, что было обусловлено, по-видимому, затратами энергии на испарение воды. Учитывая, что камера выполнена из металла и не теплоизолирована, столь сильное повышение температуры воздуха при включении УФ-ФЭУ свидетельствует о значительном освобождении энергии в ходе процессов, протекающих в воздухе при работе УФ-ФЭУ.

Как видно из рис. 2—5, при работе обоих ФЭУ наблюдалось развитие колебательных режимов ССФ. При длительной регистрации ССФ в условиях постепенного нарастания влажности воздуха в камере (подобно условиям эксперимента на рис. 5) колебания ССФ Вид-ФЭУ постепенно становились строго упорядоченными, и такой их характер сохранялся в течение десятков минут.

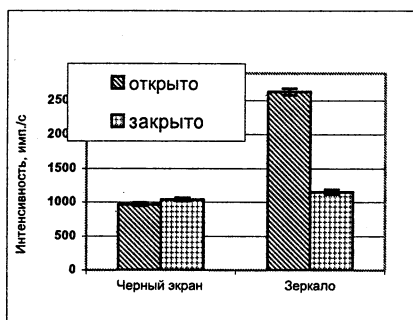


Рис. 6. Влияние характера экрана, установленного перед УФ-ФЭУ, на скорость счета им фотонов. Слева — перед окном УФ-ФЭУ установлен черный экран с шероховатой поверхностью, справа — установлен зеркальный экран. Темная диагональная штриховка — шторка между экраном и окном ФЭУ открыта, светлая сетка — шторка закрыта.

Если перед окном УФ-ФЭУ на расстоянии не более 1 см было установлено зеркало, то уже при влажности воздуха более 5—10% ССФ УФ-ФЭУ значительно возрастала (рис. 6). Если между зеркалом и окном ФЭУ была закрыта шторка или находилось другое светонепроницаемое препятствие, эффект не наблюдался. Установка против окна УФ-ФЭУ черного шероховатого (не глянцевого) экрана не влияла на скорость счета им фотонов. Установка зеркала против окна Вид-ФЭУ не влияла на скорость счета им фотонов.

4. Обсуждение

4.1. Возможный механизм инициирования свечения влажного воздуха

При работе с детектором одиночных фотонов, оснащенным двумя ФЭУ с разной спектральной чувствительностью (в видимой и УФ областях спектра), были обнаружены два эффекта, зависящие от влажности воздуха, заполняющего пространство между двумя ФЭУ. Во-первых, при наличии оптического контакта между окнами Вид-ФЭУ и УФ-ФЭУ при включении последнего повышалась скорость счета фотонов в видимой

области спектра. Во-вторых, влажность воздуха оказывала влияние на скорость счета фотонов УФ-ФЭУ и на стабильность его показаний вне зависимости от наличия его контакта с Вид-ФЭУ. При минимальной влажности воздуха первый эффект был слабо выражен, а второй практически отсутствовал, и степень выраженности обоих эффектов возрастала с увеличением влажности воздуха. В доступной нам литературе сведений о таких явлениях обнаружить не удалось.

Обнаруженные эффекты не могут быть объяснены наводками в электронных схемах, влиянием влажности воздуха на работу электронных схем и другими артефактами. Объяснение им следует искать исходя из того, что УФ-ФЭУ, находящийся под рабочим напряжением, воздействует на газовую фазу, с которой контактирует его окно. В то же время, поскольку показания Вид-ФЭУ при отсутствии его оптического контакта с включенным УФ-ФЭУ не зависят от влажности воздуха, его можно рассматривать как пассивный регистратор событий, происходящих в газовой фазе.

Для выяснения механизмов обнаруженных эффектов следует рассмотреть некоторые общие свойства фотоэлектронных умножителей, особенности конструкции ФЭУ-142 и проанализировать явления, при которых возможно возникновение сверхслабого излучения в наших экспериментальных условиях.

1. Известно, что в фотоумножителях за счет электронных ударов фотоэлектронами, разогнанными в электрическом поле до энергий, превышающих 100 эВ, возникает люминесценция остаточных газов и поверхности динодов. В спектре люминесценции представлены УФ-фотоны с энергией в несколько электрон-вольт, часть которых может направляться и в сторону фотокатода.

2. Контактующее с воздухом окно ФЭУ-142 выполнено из кристалла MgF_2 , который прозрачен вплоть до 110 нм и не является препятствием для фотонов, возникающих при работе ФЭУ. MgF_2 гигроскопичен, и при контакте с влажным воздухом на поверхности окна ФЭУ-142 образуется водяная пленка. Если поверхность, на которой сорбируется вода, представляет собой мозаику из частично положительных и частично отрицательных зарядов, то адсорбированная ею вода упорядочена. Поскольку молекулы воды представляют собой диполи, то, согласно модели Гилберта Линга, в первом адсорбционном слое каждая молекула воды окружена со всех четырех сторон молекулами, дипольные моменты которых направлены относительно нее антипараллельно. На этом слое адсорбируется новый слой воды, который, в свою очередь, адсорбирует следующий и т. д. [17]. В результате поверхность может удерживать множество прилежащих друг к другу поляризованных монослоев воды [18, 19], и эта трехмерная вода отлична как от обычной жидкой воды, так и ото льда. Возможны и другие модели укладки воды на поверхности гигроскопичного кристалла, например в виде клатратов (газгидратов) [20]. Так или иначе, весьма вероятно, что водяная пленка на поверхности полированно-

го кристалла MgF_2 , также хотя бы частично многослойна и поляризована. Высокоэнергетические фотоны, возникшие при работе УФ-ФЭУ, неизбежно встречают ее на своем пути.

3. Под действием коротковолновой радиации $100 < \lambda < 190$ нм (12,3—6,5 эВ) молекулы воды, как известно, диссоциируют на атом водорода и гидроксильный радикал ($H-O-H \rightarrow H\bullet + \bullet OH$). Недавно выполненные расчеты свидетельствуют, что микрокапли и кристаллы чистой воды эффективно поглощают ультрафиолет с еще более низкой энергией — до 4 эВ [21], что также может сопровождаться гомолитичной диссоциацией молекул воды.

4. Повышение ССФ в УФ-области спектра с возрастанием влажности воздуха в камере указывает, что в пленке воды развиваются процессы, при которых освобождается энергия в форме электронного возбуждения, что характерно для разветвленно-цепных реакций с участием кислорода. Если в пленке, поглотившей УФ-фотон, возникают свободные радикалы $H\bullet$ и $\bullet OH$, они могут реагировать с кислородом с образованием перекисных радикалов, обеспечивающих рост и разветвление цепей. Вклад в развитие цепных процессов может вносить и недавно обнаруженная реакция окисления воды электронно-возбужденным (синглетным) кислородом до перекиси водорода с образованием малоустойчивых более высоких перекисей (HO_2O_2 , HO_4O_2 , HO_6O_2), которые могут порождать новые активные центры. Эту реакцию катализирует определенным образом организованная вода [22]. К тому же, если активные центры возникают в организованной среде, то освобождающаяся в ходе их реакций энергия электронного возбуждения может в ней накапливаться и даже «складываться» до более высоких уровней, соответствующих УФ-области спектра, что способствует возникновению новых активных центров.

Указанные выше особенности работы ФЭУ, в частности ФЭУ-142, свойства разветвленно-цепных процессов с участием кислорода позволяют с единых позиций объяснить все результаты, полученные при анализе влияния влажности воздуха на ССФ УФ-ФЭУ. В частности, возникновение колебательных режимов излучения может объясняться тем, что при высокой влажности воздуха толщина пленки достигает таких размеров, что инициированные в ней окислительные процессы сопровождаются столь интенсивным освобождением энергии, что пленка разрушается. Процесс затухает до формирования новой пленки.

Распространение возбуждения от тонкой пленки воды, адсорбированной на окне УФ-ФЭУ, на расстоянии 5 см, и к тому же с усилением, является новым и неожиданным наблюдением, говорящим, что во влажном воздухе в мягких условиях может быть инициировано образование и размножение активных центров, протекание окислительных реакций разветвленно-цепного типа с участием кислорода. Возникновение колебательных режимов счета фотонов Вид-ФЭУ может быть обусловлено волнообразным переносом возбуждения по влажному воздуху, подобным

тому, который наблюдается в реакциях Белоусова — Жаботинского в конденсированных фазах.

4.2. Возможные следствия явления «свечения влажного воздуха», инициированного УФ-фотонами

Весьма вероятно, что описанное здесь явление «свечения влажного воздуха», точнее — распространения по воздуху на макроскопические расстояния самоподдерживающихся за счет окислительно-восстановительных реакций волн возбуждения, реализуется во многих природных и технических системах, где существуют упорядоченные пленки воды.

Так, хорошо известно, что атмосфера соляных пещер и шахт, вырытых в месторождениях сильвинита (хлориды калия и натрия), благотворно действует на здоровье человека. Оказалось, что воздух там постоянно высоко насыщен легкими отрицательно заряженными аэроионами [23], которые, как отмечалось выше, являются индикаторами его насыщенности разнообразными активными формами кислорода. На поверхности гигроскопичных кристаллов сильвинита образуется водяная пленка. Хлорид калия в сильвините содержит природный радиоактивный изотоп ^{40}K , составляющий 0,012% от всего калия. Период полураспада ^{40}K — $1,4 \times 10^9$ лет, а энергия его β -распада (основной тип распада ^{40}K) — 1,35 МэВ. Каким образом столь гигантская энергия может преобразоваться в энергию электронного возбуждения, лежащую в диапазоне менее 10 эВ? Известно, что при попадании в воду, показатель преломления которой $n = 1,33$, бета-частицы с энергией выше 263 кэВ возникает черенковское излучение, максимум энергии квантов которого лежит в районе 5 эВ и которое к тому же обладает свойствами когерентности. Если же показатель преломления среды еще больше, что характерно для водных растворов и коллоидов, то пороговое значение энергии бета-частиц, вызывающих черенковское излучение, снижается [24]. Следовательно, распад каждого атома ^{40}K , растворенного в водяной пленке или находящегося вблизи нее, должен сопровождаться черенковским излучением. Последнее инициирует в пленке воды, а затем и в окружающей атмосфере процессы, аналогичные тем, что наблюдаются в камере использованного нами прибора. Не связан ли терапевтический эффект спелеотерапии с пребыванием человека в атмосфере, в которой непрерывно «гуляют» волны электронного возбуждения и непрерывно генерируются АФК?

В связи с вышесказанным следует напомнить, что значительная часть внутриклеточной воды, воды на поверхностях раздела клетка — среда, на поверхностях раздела живая ткань — воздух структурирована и многослойна [25], а все клетки характеризуются существенно повышенным относительно среды содержанием ионов калия. Ежесекундно в организме человека (масса 70 кг) происходит 4400 распадов изотопа ^{40}K . Этот фактор может вносить определенный вклад как в генерацию энергии электронного возбуждения, так и в распространение возбуждения и во

внутренней среде организма, и с его поверхности в окружающую среду. При поглощении фотонов Черенковского излучения молекулы воды, входящие в ее организованные структуры, могут распадаться, а при наличии молекулярного кислорода возможно развитие разветвленно-цепных процессов, способствующих существенно усилению и распространению исходного возбуждения. Помимо радиоактивных распадов, с неизбежностью происходящих как в клетке, так и во внеклеточной среде, кванты энергии, достаточные для расщепления молекул воды, регулярно генерируются в ходе метаболических процессов в виде так называемого митогенетического излучения [26]. Митогенетическое излучение, точнее — генерация в ходе метаболических процессов энергии электронного возбуждения, достаточной для расщепления молекул воды, также может обеспечивать общее возбужденное состояние внутренней и внешней среды организма, необходимое для реализации биоэнергетических и биоинформационных процессов.

Что касается аэроионизаторов, то хорошо известно, какое значение А. Л. Чижевский придавал не только напряжению, подаваемому на иглы электроэффлювиальных люстр, но и числу этих игл, и их взаимному расположению, отчего, с нашей точки зрения, может зависеть характер распространения волны возбуждения по воздуху к мишеням, на которые действуют АФК (или само возбуждение).

Можно также предположить, что инициированное УФ-фотонами незатухающее возбуждение влажного воздуха, обладающее при определенных условиях даже способностью к автоусилению, может вносить вклад и в некоторые атмосферные процессы. Однако более глубокое обсуждение этого вопроса требует заинтересованного участия специалистов по физике и химии атмосферы.

В любом случае, обнаруженное явление может обеспечивать организованное распространение возбуждения от локальных источников на значительные расстояния как в окружающей живые организмы среде, так и в их жидкой фазе, внося свой вклад в энергоинформационную целостность живых систем.

Литература

1. *Чижевский А. Л.* Действие отрицательных ионов воздуха на некоторые физиологические функции животных [Доклад, читанный в Зоологическом музее МГУ 17 марта 1922 г., литогр. в 1-й типографии калужского губисполкома]. На правах рукописи. Калуга, 1922. 22 с.
2. *Чижевский А. Л.* Аэроионификация в народном хозяйстве. М.: Госпланиздат, 1960.
3. *Чижевский А. Л.* Аэроионы и жизнь. М.: «Мысль», 1999. 716 с.
4. *Soyka F., Edmonds A.* The Ion Effect. New York: Bantam, 1978.
5. *Tikhonov V. P., Temnov A. A., Kushnir V. A., Sirota T. V., Litvinova E. G., Zakharchenko M. V., Kondrashova M. N.* Complex herpeutical effect of ionized air: stimulation of immune system and decrease of excessive serotonin. H₂O₂ as a link between the two counterparts // IEEE Trans. Plasma Sci 2004. V. 32. № 4. P. 1661—1667.

6. Goel N., Terman M., Terman J. S., Macchi M. M., Stewart J. W. Controlled trial of bright light and negative air ions for chronic depression // *Psychol. Med.* 2005. V. 35. P. 945—955.
7. Скунцов В. П. Аэроионы и жизнь. Саранск: «Красный Октябрь», 2005. 135 с.
8. Krueger A. P., Reed E. J. Biological impact of small negative ions // *Science*. 1976. V. 193. P. 1209—1233.
9. Goldstein N. I., Goldstein R. N., Merzlyak M. N. Negative air ions as a source of superoxide // *Int. J. Biometeorol.* 1992. V. 36. P. 118—122.
10. Kosenko E. A., Kaminsky Yu. G., Starovskaya I. G., Sirota T. V., Kondrashova M. N. The stimulatory effects of negative air ions and hydrogen peroxide on the activity of superoxide dismutase // *FEBS Lett.* 1997. V. 410. P. 309—312.
11. Kondrashova M. N., Grigorenko E. V., Tichonov A. N., Sirota T. V., Temnov A. V., Starovskaya I. G., Kosyakova N. I., Lange N. V., Tichonov V. P. The primary physico chemical mechanism of the beneficial biological/medical effects of negative air aions // *IEEE Tranact. Plasma Sci.* 2000. V. 28. P. 230—237.
12. Voeikov V. Reactive Oxygen Species, Water, Photons, and Life // *Rivista di Biologia / Biology Forum*. 2001. V. 94. P. 193—214.
13. Воейков В. Л. Регуляторные функции активных форм кислорода в крови и в модельных водных системах: Дис. ... д.б.н. М.: МГУ, 2003.
14. Воейков В. Л. Возможные физико-химические механизмы биологического действия аэроионов Чижевского (супероксида) // *Космос и мировая история: Материалы Международной научной конференции к 105-летию со дня рождения основоположника космической экологии и социологии А. Л. Чижевского*. Отв. ред. Л. В. Голованов. М.: ИИКС «Эльф-3» РФБ «Гелиос». 2002. С. 211—224.
15. Чижевский А. Л. Тайна живого воздуха. Калуга, 2004. С. 100—101.
16. Воейков В. Л., Чалкин С. Ф. Способ газового анализа и газоанализатор для его осуществления: Патент РФ на изобретение № 2235311, зарегистрирован 27 августа 2004 г.
17. Ling G. N. A new theoretical foundation for the polarized-oriented multilayer theory of cell water and for inanimate systems demonstrating long-range dynamic structuring of water molecules // *Physiol. Chem. Phys. & Med. NMR*. 2003. V. 35. P. 91—130.
18. Pashley R. M., Kitchener J. A. Surface forces in adsorbed multilayers of water on quartz // *J. Colloid and Interface Sci.* 1979. V. 71. P. 4910—500.
19. Fisher J. R., Gamble R. A., Middlchurer J. The Kelvin equation and the condensation of water // *Nature*. 1981. V. 290. P. 575—576.
20. Кисловский Л. Д. Роль воды в лабильности поверхностных структур // *Активная поверхность твердых тел / Под ред. Г. И. Дистлера, П. Ю. Бутягина*. М.: ВИНТИ, 1976. С. 267—279.
21. Новаковская Ю. В. Ультрафиолетовое излучение и роль воды в озоновом слое земли // МГУ им. М. В. Ломоносова, химический факультет. Ломоносовские чтения, 2002.
22. Xu X., Muller R. P., Goddard W. A. 3rd. The gas phase reaction of singlet dioxygen with water: a water-catalyzed mechanism // *Proc. Natl Acad. Sci. USA*. 2002. V. 99. P. 3376—3381.
23. Вавилов В. А., Хлюстов В. Н., Ковалева Г. П., Емельянов А. Д., Ефимова И. Е., Обухов Е. А., Карамов С. К. Применение спелеотерапии у больных с бронхолегочной патологией // *Кремлевская медицина. Клинический вестник* 1999. № 4, <http://grcc.pmc.ru/Vestnik/V99-4/14.html>.
24. Ross H. H. Recent applications of Cerenkov radiation // *Liquid scintillation counting. Recent applications and development*. Vol. 1. New York; London: AP, 1980. P. 375—385.
25. Pollack G. H. Cells, gels and the engines of life. A new unifying approach to cell function. Ebner & Sons, Seattle, WA, USA. 2001. 320 p.
26. Гурвич А. Г., Гурвич Л. Д. Митогенетическое излучение, физико-химические основы и приложения в биологии и медицине. М.: Медгиз, 1945. 283 с.

БИОЛОГИЧЕСКИЙ ВИД ВОЛОКОННЫХ СВЕТОВОДОВ

*Ю. Н. Кульчин, О. А. Букин, С. С. Вознесенский, А. Н. Галкина,
С. В. Гнеденков, А. Л. Дроздов, В. Г. Курявый, Т. Л. Мальцева,
С. Л. Синебрюхов, А. И. Чередниченко*

Введение

Биологические системы в процессе эволюции развили уникальные способы решения проблем выживания: от средств устрашения хищников и обеспечения механической жесткости, до сложных процессов фотосинтеза. Скелеты различных морских организмов построены на основе разнообразных минералов: кальцит, арагонит, апатит и кремнезем [1]. Особое внимание привлекают к себе такие морские организмы, как диатомовые, радиолярии и губки, скелеты которых сформированы на основе кремнезема [2]. Это обусловлено тем, что кремний и его оксиды являются одними из основных материалов, используемых в микро- и оптоэлектронике для решения широкого круга задач.

Оптическая связь, основанная на использовании кварцевых волоконных световодов сегодня прочно вошла в нашу повседневную жизнь. Однако, как выяснилось, идея создания и использования кварцевых волокон была реализована Природой задолго до появления человека на Земле. Стекланные губки класса Hexactinellida (Porifera) представляют собой группу древнейших многоклеточных животных, которые обитают преимущественно в глубоководных зонах всех океанов, закрепившись в донных осадках с помощью особых кремнийсодержащих элементов, называемых спикулами. Как показывают результаты проведенных нами исследований, особенности структурной организации спикул стекланных губок по своим механическим и оптическим свойствам позволяют отнести их к уникальным фотонным биоматериалам природного происхождения. Геометрические размеры спикул, в зависимости от их вида и возраста, могут изменяться в достаточно широких пределах: диаметр от 0,01 до 30 мкм, а длина от единиц миллиметров до более чем 5 м. Минеральный состав спикул включает в себя аморфную двуокись кремния. В настоящей работе представлены результаты структурного, оптического и спектрального исследований спикул стекланных губок, обитающих в бассейне морей Тихого океана.

1. Морские губки как объект исследования

Морские губки относятся к типу беспозвоночных, образуя тупиковую ветвь в основании филогенетического древа многоклеточных [2]. Эти уникальные животные широко распространены в природе. На рис. 1 приведены фотографии живых губок, показывающие лишь малую часть их

природного разнообразия. К типу губок (Spongia, Porifera) относятся морские и пресноводные организмы, которые сохранили многие признаки колонии простейших. Сквозь тело губок процеживается вода, из которой губки получают питательные вещества. Всего известно более 3 тысяч видов губок.

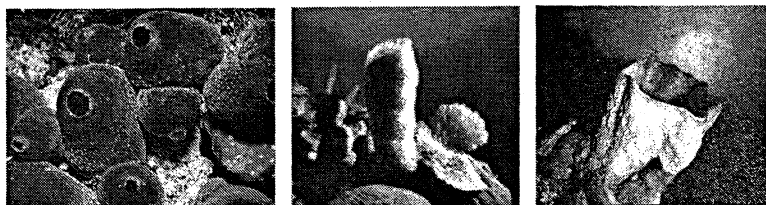


Рис. 1. Губки в живой природе

Губки различаются по типу скелетных образований. Скелет — основной признак, по которому губки разделяются на классы. Основу скелета губок составляют специфические образования, называемые спикулами, имеющие большое разнообразие форм и различающиеся по функциональному назначению в скелете губки. Спикулы традиционно подразделяются по размеру на две категории: мегасклеры и микросклеры (рис. 2). Мегасклеры обычно формируют основу, скелет губки, тогда как микросклеры располагаются по телу губки.

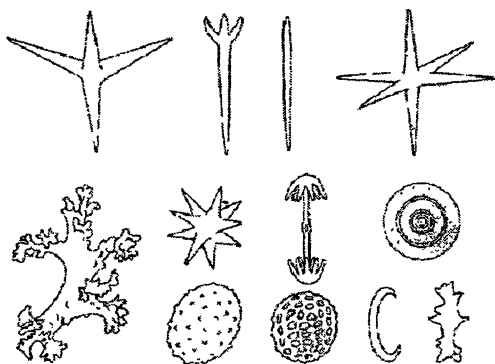


Рис. 2. Спикулы губок: крупные макросклеры (вверху) и мелкие микросклеры (внизу). Масштаб на рисунке не соблюден.

Обычно выделяют три класса губок.

1) *Известковые губки* (Calcispongia, или Calcareae): морские животные, обычно не длиннее 15 см, скелет которых образован одно-, трех- или четырехлучевыми спикулами, состоящими из углекислого кальция.

2) *Обыкновенные губки* (Demospongiae): имеют весьма разнообразные скелеты, составленные из одно- или четырехлучевых кремнеземных спикул или вовсе без них. К этому классу относятся как пресноводные, так и морские организмы.

3) *Стеклянные, или шестилучевые, губки* (Hexactinellida): скелет составлен часто сливающимися кремнеземными шестилучевыми спикулами, поэтому выглядит как бы сотканным из стеклянных нитей. В этот класс входят только морские, чаще всего глубоководные, губки длиной до 90 см, которые бывают как одиночными, так и колониальными.

На наш взгляд, наиболее привлекательным объектом исследования являются морские губки, скелет которых образуют протяженные спикулы длиной не менее единиц сантиметров, состоящие из аморфного гидратированного кварца, депонированного вокруг белковой основы. К такого рода губкам можно отнести обитающие на глубинах от 35 до 5000 метров в холодной воде морей Тихого океана стеклянные губки родов: *Euplectellidae* (рис. 3), *Hyalonema sieboldi* (рис. 4) и *Ferronema sp.* (рис. 5). На сегодня наиболее изученными являются структурные и оптические свойства базальных спикул стеклянной губки *Euplectella aspergillum* [2, 3]. Поэтому в нашей работе остановимся на результатах исследований спикул двух родов стеклянных морских губок, которые для краткости будем нумеровать по тексту номерами I и II соответственно.

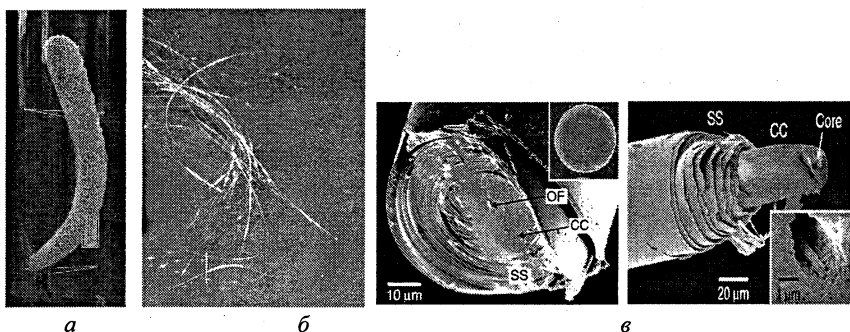


Рис. 3. Стеклянная губка рода *Euplectellidae* (а), образцы спикул (б), СЭМ-изображения разломленного поперечного участка базальной спикулы губки *Euplectella sp.* до и после обработки раствором NaOCl (в).

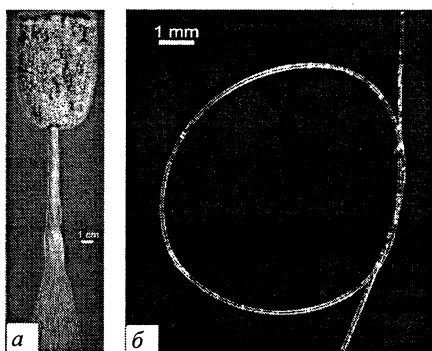


Рис. 4. Стеклянная губка *Hyalonema sieboldi* (а) и ее базальная спикула (б).



Рис. 5. Стекло́нная губка рода *Ferronema* sp. и образцы их спикул

2. Особенности строения материала спикул стеклянных морских губок

В мире существует более 500 видов морских стеклянных губок класса *Hexactinellida*, скелет которых составляют шестилучевые спикулы, состоящие из аморфного гидратированного кварца. Установлено, что спикулы, в зависимости от их назначения и расположения в губке, имеют различия в строении, однако сами принципы организации спикул сохраняются [2].

Так, в губках различают скелетные и базальные (якорные) спикулы. Базальные спикулы представляют собой тонкие нити (волоски), состоящие из центральной органической нити и слоев кремнезема вокруг него. Основное назначение этих спикул — обеспечение связи с окружающей средой (и реакция на изменения условий жизни) путем регистрации и излучения спикулой светового сигнала (гипотеза).

В отличие от базальных, скелетообразующие спикулы — это более сложная механическая структура, предназначенная для придания формы и обеспечения жесткости каркаса губки.

В наших работах для исследования выбирались спикулы морских губок I и II рода длиной от 0,01 до 0,05 м и с диаметрами от 40 до 150 мкм.

Все кремнийорганические спикулы имеют центральное ядро, называемое аксиальной нитью. Основная задача аксиальной нити — контроль за полимеризацией кремнезема [4]. Аксиальная нить размещается в центральном канале спикулы, что отчетливо видно на фотографиях (рис. 6, а и рис. 7, а), полученных при 1150-кратном оптическом увеличении изображений торцов спикул I и II рода. Кроме того, хорошее представление о структуре базальной спикулы дают СЭМ-изображения разломленного поперечного участка базальной спикулы губки *Euplectella* sp. на рис. 3, в из работы [5].

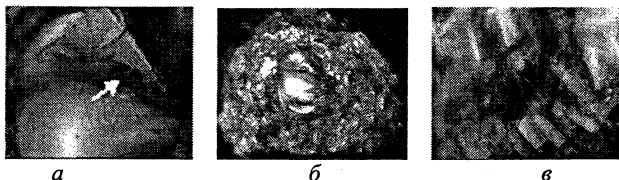


Рис. 6. Спикула I: а — фотография центрального цилиндра спикулы с каналом для белковой нити (стрелка указывает на канал); б — фотография сколото́го неполированного торца спикулы (увеличение в 297 раз); в — фотография фрагмента кольцевой структуры спикулы (увеличение в 1157 раз).

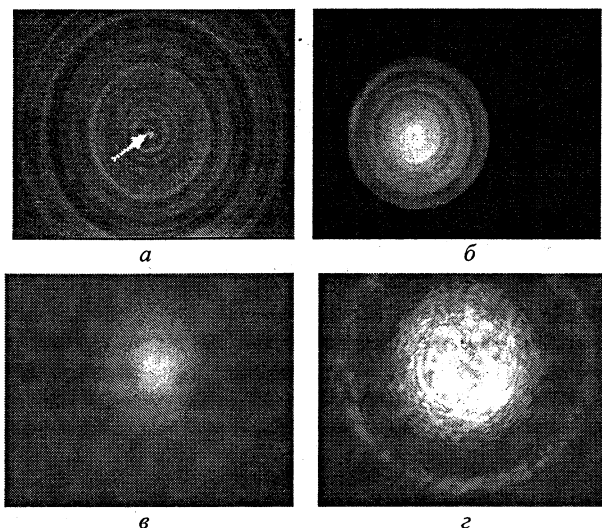


Рис. 7. Спикула II: *а* — фотография центрального цилиндра спикулы с каналом для белковой нити (стрелка указывает на канал, увелич. в 1157 раз); *б* — фотография выходного торца спикулы при возбуждении белым светом (увеличение 438 раз); *в* и *г* — фотографии распределения поля излучения ($\lambda = 0,63$ мкм) из спикулы и кварцевого волоконного световода с одинаковыми диаметрами сердцевины соответственно.

Механизм, который регулирует концентрацию и полимеризацию кремнезема, свидетельствует о том, что основным белком аксиальной нити является силикатеин. Это уникальный фермент, катализирующий гидролиз и поликонденсацию кремнийферроалголят, он является матрицей (шаблоном) при формировании высокоупорядоченной структуры спикулы. Формирование спикул происходит вокруг аксиальной нити [5] из наночастиц кремнезема размером 100—120 нм, которые далее объединяются в микрочастицы, образуя плотный слой двуокиси кремния (рис. 6, в), которые на фотографии видны в виде отдельных блоков. Наночастицы кремнезема, окруженные протеиновым матриксом, образуют внешнюю зону аксиальной нити, что обуславливает слоистость поперечной структуры спикулы и, как следствие, ее механическую прочность (рис. 6, б и рис. 7, а). Как только процесс силификации заканчивается, спонгоциты выделяют фибриллы белка спонгина вокруг спикулы. На следующих стадиях происходит чередование процессов силификации и выделения спонгина. О композитности структуры спикул стеклянных морских губок говорят и результаты их рентгеновской фотоэлектронной спектроскопии (рис. 8), которые подтверждают наличие в их составе больших концентраций кварца и белковых соединений. Завершающая стадия роста спикулы — покрытие ее аморфным кремнеземом (для придания ей жесткости, гладкости или для сцепления с другими спикулами).

Детальное исследование методом динамической ультрамикротвердометрии на поперечных шлифах с использованием пирамиды Берковского таких важных эксплуатационных характеристик материала, как микротвердость и модуль Юнга, показали (рис. 9), что для спикул морских губок их значения близки к соответствующим значениям для плавного кварца [6]. Но в то же время распределения микротвердости и модуля Юнга неоднородны по поперечному сечению спикулы и имеют тенденцию спада от центра к периферии, что, как видно из рис. 6, в и рис. 7, а, объясняется слоистостью структуры спикулы, в которой слои из кварца чередуются со слоями белковых соединений, а также утончением размеров слоев кварца с их приближением к периферии (см. рис. 7, а). Именно это качество композитного по своей природе материала спикул морских губок обеспечивает их чрезвычайно высокую стойкость к изгибу, позволяющую, в отличие от обычных кварцевых волоконных световодов, завязывать их практически в узел (рис. 4, б).

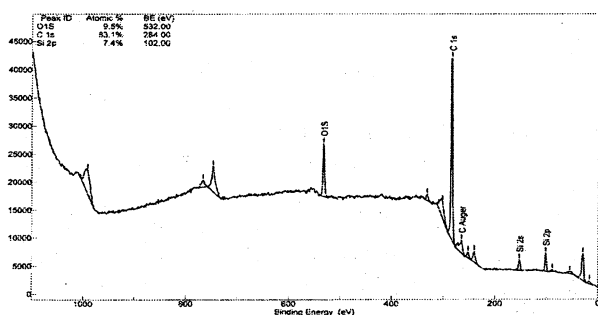


Рис. 8. Фотоэлектронный спектр для образцов спикул II

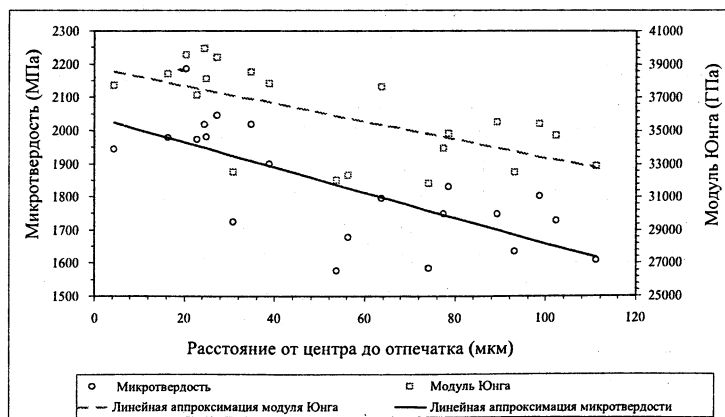


Рис. 9. Распределение микротвердости и модуля Юнга материала базальной спикулы I по поперечному сечению образца.

3. Оптические характеристики кремнийорганических спикул стеклянных морских губок

Как показано в разделе 2, спикулы стеклянных морских губок состоят из однородного центрального матрикса (кремнийорганического состава) и слоистой оболочки (чередование слоев кремнезема и органики белковой природы). При этом в [2, 7] установлено, что осевая нить базальной спикулы (1—2 мкм в диаметре и с максимальным содержанием Na) имеет показатель преломления 1,45—1,48, что очень близко к значению показателя преломления аморфного кварца 1,458. Вокруг нее находится цилиндрическая область (область гладкого цилиндра 15—40 мкм в диаметре, она хорошо видна на рис. 3, в, 6, б и рис. 7, б), которая имеет более низкий показатель преломления — 1,425 и для которой уменьшение показателя преломления объясняется высоким содержанием органических веществ и/или низкой степенью уплотнения кварца. Во внешней части спикулы наблюдается постепенное увеличение показателя преломления (от 1,433 до 1,438), в соответствии с постепенным изменением состава и увеличением диаметра блоков кварца в слоистой области (см. рис. 8, а). Различия показателей преломления во всех этих областях коррелируют с присутствием органических слоев между слоями кремнезема, что видно из измеренных зависимостей для распределения микротвердости по сечению спикул (рис. 9).

Как видно, по своим оптическим характеристикам спикулы близки к обычным кварцевым волоконным световодам, при этом натуральное оптическое волокно имеет ряд преимуществ, позволяющих по некоторым параметрам превосходить искусственное. Преимуществами являются: отсутствие двойного лучепреломления и необходимости выполнения сложных технологических процессов, связанных с внедрением в материал специальных добавок (например, ионов натрия); улучшенные механические свойства спикул, а также их низкотемпературный синтез. Вследствие достаточно большой разности показателя преломления сердцевины и окружающей среды (воздух или морская вода), спикулы стеклянных губок являются достаточно хорошими световодами. Например, для спикул I рода с диаметром сердцевины $D \sim 140$ мкм на длине волны излучения He-Ne-лазера $\lambda = 632$ нм потери излучения оказываются не более 0,1 дБ/м (рис. 10). При этом обращает на себя внимание низкая угловая расходимость выходящего из спикул излучения в сравнении с расходимостью излучения из близких по диаметру кварцевых волоконных световодов, что иллюстрируется рис. 7, в, г.

Результаты экспериментов по прохождению белого света (галогеновая лампа, спектральный диапазон 400—1600 нм) через базальную спикулу губки *Hyalonema sieboldi* [4] показали, что спикулы хорошо пропускают свет в видимой части спектра, что иллюстрируется полученными за-

висимостями, приведенными на рис. 11 ([4]). Как видно, в диапазоне длин волн 400—615 нм спикула практически не пропускает свет, и аналогичная картина наблюдается в ближней инфракрасной области (от 1310 нм и выше). В диапазоне же длин волн от 600 до 1310 нм спикула имеет достаточно хорошее пропускание, и лишь в окрестностях областей длин волн 770, 960 и 1150 нм наблюдаются три небольших пика поглощения. Экспериментальные исследования пропускания белого света сквозь спикулы продемонстрировали наличие концентрации излучения вблизи осевой области гладкого цилиндра (см. рис. 7, б) и сильный цветовой градиент от белого к красному по длине спикулы. Полученные результаты позволяют предположить, что наблюдаемая в спикулах стеклянных губок слоистая кремнийорганическая структура (рис. 6, рис. 7 и рис. 12) приводит к возникновению периодической пространственной модуляции диэлектрической проницаемости с периодом, сравнимым с длиной волны светового излучения, что делает их схожими с известными микроструктурированными волоконными световодами.

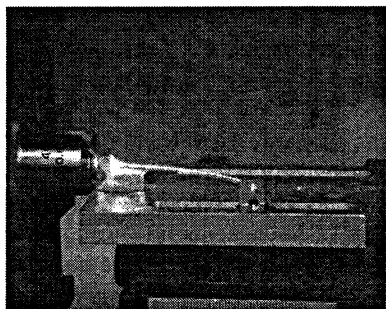


Рис. 10. Фотография спикулы I, возбужденной излучением He-Ne-лазера ($\lambda = 0,63$ мкм).

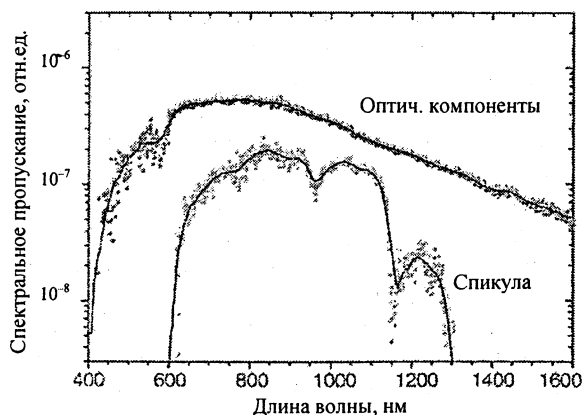


Рис. 11. Спектральное пропускание спикул стеклянных морских губок рода I

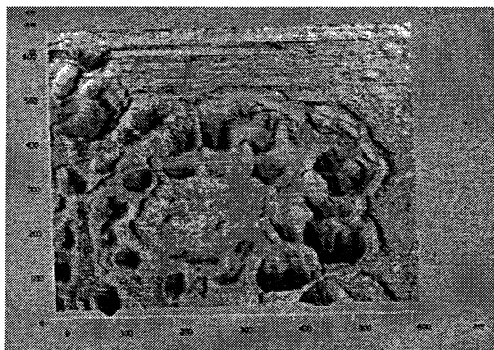


Рис. 12. Фрагмент изображения полированного торца спикулы II, полученного с использованием атомного силового микроскопа.

Наличие трансляционной симметрии в поперечном сечении световедущей области не только позволяет объяснить околоосевую концентрацию излучения и дифференциальное ослабление спикулой различных областей спектра излучения, но и сделать предположение о том, что спикулы стеклянных губок являются одномерными природными фотонными кристаллами.

При возбуждении спикул импульсами лазерного возбуждения второй гармоники Nd:YAG-лазера $\lambda = 523$ нм с длительностью импульсов 12 нс, частотой следования 10 Гц и с импульсной энергией порядка 30 мДж наблюдается значительный рост интенсивности флуоресценции в длинноволновую область. На рис. 13, *а* приведены экспериментальные зависимости спектров флуоресценции излучения из спикул рода II, полученные при различных интенсивностях излучения накачки. Видно, что спектры флуоресценции затянуты в длинноволновую область с максимумом на длине волны 770 нм, положение которого не зависит от величины интенсивности на входе. Характер зависимости отличается от соответствующей зависимости спектра флуоресценции, измеренного для обычного многомодового кварцевого волоконного световода (рис. 13, *б*). Столь разное поведение экспериментальных зависимостей спектров флуоресценции, по-видимому, связано с содержанием в спикулах морских губок крупных органических комплексов. При измерении зависимости интенсивности излучения флуоресценции на длине волны максимума ($\lambda = 770$ нм, рис. 13, *в*) наблюдается ее насыщение, что также говорит о нелинейности процесса преобразования энергии, связанной с наличием высокой концентрации неоднородно распределенных органических комплексов (белок спонгин), что находится в соответствии с результатами [8]. Это делает спикулы морских губок перспективным материалом для создания активных оптоэлектронных приборов, таких как, например, волоконные усилители и лазеры.

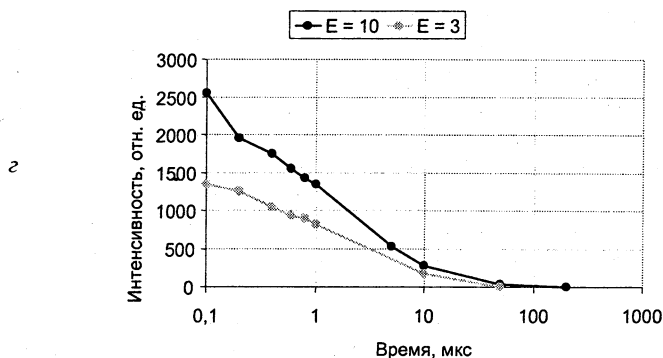
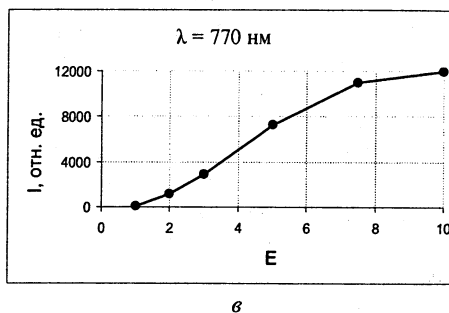
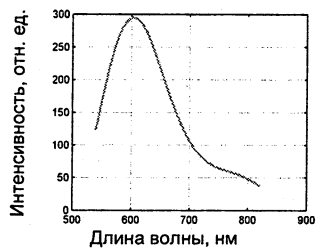
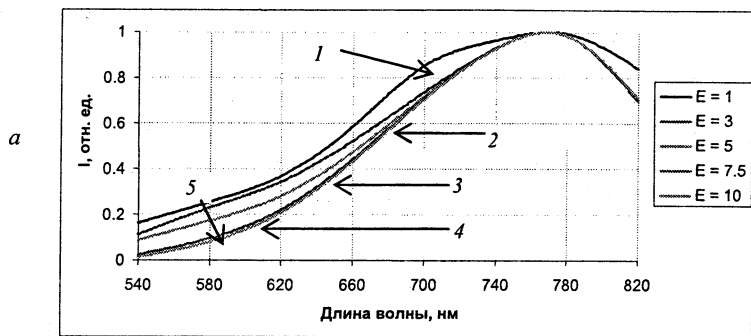


Рис. 13. Спектры флуоресценции спикулы при разных относительных значениях интенсивности возбуждающего излучения ($\lambda = 523$ нм) на входе: 1 — 1; 2 — 2,5; 3 — 3; 4 — 7,5; 5 — 10 (*a*); спектр флуоресценции многомодового кварцевого световода (*б*); зависимость насыщения флуоресценции спикулы от интенсивности излучения накачки (*в*); зависимость времени жизни флуоресценции излучения из спикулы при разных значениях интенсивности излучения накачки (*г*).

Особенностью спикул стеклянных губок как объекта флуометрии является время жизни флуоресценции. На рис. 13, *з* приведены результаты экспериментального исследования длительности флуоресценции излучения из спикул, полученные для разных значений мощности импульсов индуцирующего излучения. Полученные данные демонстрируют большие значения времени жизни флуоресценции для этих объектов (десятки микросекунд) в сравнении с таким объектом, как морская зоопланктон или кварцевые волоконные световоды ($\sim 10^{-10}$ — 10^{-9} с).

Заключение

Исследования последних лет продемонстрировали важную роль наноструктур в различных областях науки и техники. При этом, в связи с углублением знаний о строении и функционировании природных объектов и живых организмов на молекулярном уровне, исследователи пытаются разработать общий подход к получению и использованию искусственных материалов с наноразмерной структурой. В сущности, очевидно, что все природные материалы построены из наноструктур или нанообъектов. Управляя размерами и формой наноструктур, таким материалам можно придать совершенно новые функциональные характеристики, резко отличающиеся от характеристик обычных материалов. К числу уже известных наноструктур относятся углеродные трубки, белки, ДНК и т. д. Понимание и освоение этого процесса в перспективе делает возможным получение искусственных структур с заданными физическими, химическими и биологическими свойствами.

Как показали результаты настоящей работы, осуществляемая в природе комбинация белка спонгина и морской воды позволяет при низких температурах получить перспективный для оптоэлектроники композитный наноматериал, сочетающий эластичность и прочность белка с упругостью и прочностью кремнезема. В своих исследованиях мы находимся только в начале пути изучения и практического освоения этих новых и интересных по своим оптическим и механическим характеристикам объектов. Понимание процессов синтеза таких материалов открывает перспективы целенаправленного их изготовления, что, по-видимому, благодаря уникальности их свойств, позволит в дальнейшем внести существенные изменения в технику и технологию оптических систем связи и оптических вычислений.

Литература

1. Дроздов А. Л. Биология для физиков и химиков. Владивосток: Дальневост. ун-т., 2005. 414 с.; Joanna Aizenberg, Vikram C. Sundar, Andrew D. Yablon, James C. Weaver, and Gang Chen. Biological glass fibers: Correlation between optical and structural properties // *Proc. Natl. Acad. Sci. USA*. 2004 March 9; 101(10). P. 3358—3363.

2. *Werner E.G. Müller, Klaus Wendt, Christopher Geppert, Matthias Wiens, Andreas Reiber, Heinz C. Schröder.* Novel photoreception system in sponges? Unique transmission properties of the stalk spicules from the hexactinellid *Hyalonema sieboldi*. *Biosensors and Bioelectronics*. 2006. V. 21. P. 1149—1155.
3. *Joanna Aizenberg, James C. Weaver, Monica S. Thanawala, Vikram C. Sundar, Daniel E. Morse, Peter Fratzl.* Skeleton of *Euplectella* sp.: Structural Hierarchy from the Nanoscale to the Macroscale // *Science*. 2003. V. 309. P. 275—278.
4. *Mari'a-J. Uriz, Xavier Turon, Mikel A. Becerro, Gemma Agell.* Siliceous Spicules and Skeleton Frameworks in Sponges: Origin, Diversity, Ultrastructural Patterns, and Biological Functions // *Microscopy Research and Technique*. 2003. V. 62. P. 279—299.
5. *Vikram C. Sundar, Andrew D. Yablon, John L. Grazul, Michla Ilan, Joanna Aizenberg.* Fiber-optical features of a glass sponge // *Nature* 2003. V. 424. P. 899—900.
6. *Самсонов Г.В. и др.* Физико-химические свойства окислов. М.: Металлургия, 1978. 472 с.
7. *Эрлих Г., Ересковский А. В., Дроздов А. Л., Крылова Д. Д., Ханке Т., Майснер Х., Хайнеман С., Ворх Х.* Современный подход к деминерализации спикул стеклянных губок (*Porifera: Hexactinellida*) с целью извлечения и исследования протеинового матрикса // *Биология моря*. 2006. Т. 32, № 3. С. 18—23.
8. *Маслов Д. В., Остроумов Е. Е., Фадеев В. В.* Флуометрия насыщенных сложных органических соединений с высокой локальной концентрацией флуорофоров // *Квантовая электроника*. 2006. Т. 36, № 2. С. 163—168.

Периодические орбиты, локализация в пространстве нормальных мод и проблема Ферми, Паста и Улама. С.Г. Флах // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. – С. 9–23.

Ожидается, что длинноволновая мода в ангармонической атомной цепочке быстро отдаст свою энергию другим модам. 50 лет тому назад Ферми, Паста и Улам (ФПУ) обнаружили, что в действительности наблюдается противоположный результат. Обсуждается продолжение нормальных мод, т. е. периодических орбит (ПО) гармонической цепочки в ангармонический режим. Полученные решения экспоненциально локализованы в пространстве нормальных мод. Траектория ФПУ является слабым возмущением точной ПО и остается в ее окрестности в течение долгого времени. Результаты обобщаются на случаи двух- и трехмерных систем, а также на случай макроскопической решетки, с некоторыми неожиданными последствиями.

Симметрии и их нарушения в системах во внешних переменных полях. С.Г. Флах // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 24–38.

Проводится анализ симметрии нелинейных динамических систем под воздействием внешних переменных полей. Выпрямление транспорта в современной литературе часто сопровождается понятием рэчточного эффекта. Обсуждается выпрямление потока невзаимодействующих частиц, двигающихся в пространственно-периодическом потенциале. Экспериментально этот эффект был обнаружен в направленной диффузии холодных атомов в оптических решетках. Принцип симметрий применен к потоку энергии в нелинейных полевых уравнениях. Результирующий направленный транспорт солитонов был обнаружен в кольцевых Джозефсоновских переходах. Симметричный анализ проведен для квантовых спинов во внешних магнитных полях. Предсказана намагниченность с компонентой, перпендикулярной к приложенным магнитным полям.

Фазовые компактоны. А.С. Пиковский, Ф. Розенау // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 39–54.

Рассматривается фазовая динамика в цепочке дисперсионно связанных автогенераторов. В квазинепрерывном пределе основная модель сводится к уравнению типа Кортевега – де Вриза, но с нелинейной дисперсией. Решениями являются компактоны (удлиненные волны с компактным носителем) и коватоны (компактные пары перепад-антиперепад, распространяющиеся с общей скоростью и находящиеся на произвольном расстоянии друг от друга).

Пространственно-временная динамика топологических особенностей волновых полей: источники, стоки, дырки, дефекты. А.Б. Езерский // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 55–67.

При исследовании пространственно-временной динамики периодических структур в неравновесных средах широко используется подход, при котором основное внимание уделяется не всему полю параметра порядка, а топологическим особенностям этих полей. Основная идея – показать, что в пространственно-периодических системах, возникших благодаря неустойчивости гидродинамического типа, возможно идентифицировать такие топологические особенности и провести сравнение их характеристик с решениями эталонных уравнений.

О реализации некоторых классических моделей и феноменов нелинейной динамики на основе связанных неавтономных осцилляторов. С.П. Кузнецов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 68–84.

Обсуждается идея реализации ряда абстрактных моделей и феноменов нелинейной динамики на базе специального класса систем, которые могут быть построены как устройства радиотехники и электроники. Это системы из связанных автогенераторов, которые благода-

ря принудительной периодической модуляции параметров по очереди выходят за порог генерации, передавая друг другу возбуждение. Сформулированы модели в виде систем неавтономных дифференциальных уравнений с периодическими коэффициентами и приведены компьютерные результаты, свидетельствующие о присутствии ожидаемых феноменов.

Фазовые кластеры и краткосрочная память в сети нейронов с последеполаризацией. В.В. Клиньшов, В.И. Некоркин // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 85–101.

Предложена и исследована колебательная модель системы рабочей памяти мозга в виде многомерной динамической системы, описывающей сеть пирамидальных нейронов с последеполаризацией под действием подпорогового периодического сигнала. Показано, что в такой сети возможно запоминание информации в виде кластеров периодической активности. Наиболее интересным свойством данной системы является способность запоминания нескольких образов одновременно, что является базовым свойством рабочей памяти мозга. Хранение нескольких образов соответствует существованию в системе нескольких последовательно возбуждающихся кластеров периодической активности. Каждый кластер представляет один образ и возбуждается в своей фазе внешнего сигнала. Построена зависимость максимального числа образов, которое система может хранить одновременно, от параметров связи.

Непрерывный вейвлетный анализ и некоторые его приложения к задачам нелинейной динамики. А.А. Короновский, А.Е. Храмов // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 102–117.

Изложены основные понятия, связанные с новым направлением прикладной математики – вейвлетным анализом, рассмотрены различные аспекты непрерывного вейвлетного преобразования и возможности его применения для исследования сложных режимов, порожденных нелинейными динамическими системами. В качестве примеров рассматриваются методы эффективного выделения ламинарных и турбулентных фаз во временных рядах различной природы, включая биологические ряды электроэнцефалограмм, и диагностики хаотической синхронизации с помощью введения непрерывного множества фаз хаотического сигнала.

Системы связи с хаотическими носителями: проблемы, достижения и перспективы. А.С. Дмитриев // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 118–136.

Рассматривается технология связи, использующая в качестве носителя информации динамический хаос. Показывается, что к настоящему времени она переходит из стадии фундаментальных исследований в стадию практического применения.

Математические модели, точные решения, методы анализа. О.В. Руденко // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 139–150.

Обсуждается ряд моделей нелинейной акустики, представляющих интерес и для других разделов физики нелинейных волн. Приводятся точные решения, имеющие физический смысл. Описываются некоторые способы их нахождения или угадывания. На примерах показана связь «нестрогих» подходов с корректным математическим анализом нелинейных моделей на основе их симметрий.

Актуальные проблемы, связанные с нелинейной акустикой. О.В. Руденко // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 151–169.

Рассмотрены направления, привлекившие наибольшее внимание в последние годы: 1. Звуковой удар от тел, движущихся со сверхзвуковой скоростью. Явления при формировании *N*-волны в атмосфере. Выбросы, их экологические последствия. Удар и перспективы

сверхзвуковой пассажирской авиации. 2. Акустический термомод – проблема реализации термомодерного синтеза в коллапсирующем пузырьке, при кавитации. Излучение более коротковолновое, чем при сонолюминесценции одиночного пузырька. Явление неограниченной кумуляции. 3. Свойство скелетных мышц демпфировать удары. Нелинейная динамика мышечного сокращения. Данные по измерениям сдвиговой упругости мышц, новые модели.

О «неклассических» проявлениях микроструктурно-обусловленной акустической нелинейности на примере контактодержащих сред. В.Ю. Зайцев // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 170–190.

Обсуждаются причины возникновения «неклассических» нелинейных свойств микро неоднородных сред. Приводятся примеры использования нелинейных эффектов для получения информации о микроструктуре среды. Рассмотрены примеры «неклассических» нелинейных эффектов, связанных не только с нелинейно-упругим поведением, но и с амплитудно-зависимым поглощением (в частности, наблюдения акустического аналога Люксембург-Горьковского эффекта), проявления медленной динамики за счет термоупругих эффектов и т. д.

Динамика вихрей в слоистых сверхпроводниках с внутренним эффектом Джозефсона. В.В. Курин, А.В. Чигинев // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 191–205.

Дается представление о микроскопическом устройстве высокотемпературных купратных сверхпроводников и способах описания различных взаимодействующих степеней свободы сверхпроводника. Дается понятие об эффекте Джозефсона и его проявлениях в слоистых высокотемпературных сверхпроводниках. Рассматриваются различные механизмы связи между джозефсоновскими слоями в высокотемпературном сверхпроводнике и обсуждаются их влияние на структуру вихревых решеток. Особое внимание уделяется вопросу о влиянии оптических фононов на устойчивость синфазного режима движения джозефсоновских вихрей в высокотемпературных сверхпроводниках.

Вынужденное рассеяние Мандельштама – Бриллюэна лазерных мод Гаусса – Лагерра: новые явления. Ф.А. Стариков // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 206–221.

Показано, что обращение волнового фронта при вынужденном рассеянии Мандельштама – Бриллюэна (ВРМБ) сфокусированных вихревых пучков Гаусса – Лагерра не реализуется как явление, так как обращенная стоковая мода и аналогичная мода противоположной спиральности имеют одинаковые коэффициенты усиления ВРМБ. Поэтому (например, для тороидальной лазерной моды LG_0^1) стоков пучок состоит из случайной суперпозиции этих мод. Для довольно широкого класса вихревых пучков (например, для лазерной моды LG_1^1) предсказывается явление, которое можно назвать *преобразованием* волнового фронта при ВРМБ. Суть его в том, что в стоковом пучке происходит селекция несопряженной ортогональной моды, которая является основной гауссовой модой.

Генерация быстрых электронов при электрон-ионных столкновениях в сильных полях. Г.М. Фрайман, А.А. Балакин // *Нелинейные волны–2006*. – Нижний Новгород: ИПФ РАН, 2007. С. 222–236.

Рассмотрены электрон-ионные столкновения в сильных электромагнитных полях в плазме. Предложен метод их аналитического и численного описания на основе точечного отображения. Из анализа точечного отображения получены основные особенности рассеяния в сильных полях: появление многопоточности, стохастичности движения частиц и «сингулярности» концентрации перед «жестким» ударом. Получены оценки для эффективной частоты столкновений и распределения частиц по энергиям в результате столкновений в нерелятивистском и ультрарелятивистском пределах. Показано, что с ростом поля все энер-

гетические эффекты, обусловленные столкновениями (джоулев нагрев, генерация быстрых электронов, излучение на высоких гармониках), возрастают. Приводится сравнение полученных результатов с экспериментальными.

Гравитация и супергравитация. Ю.М. Зиновьев // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 239–247.

Конструктивный подход рассматривает гравитацию как "обычную" теорию поля, описывающую безмассовые частицы со спином 2 так же, как электродинамика описывает безмассовые частицы со спином 1 – фотоны. Таким образом все взаимодействия в гравитации (самодействие и взаимодействие с материей) могут быть воспроизведены и в этом смысле поняты. Но квантование ведет к неперенормируемой теории. Наиболее естественный путь к лучшей квантовой теории – это суперсимметричное расширение гравитации – супергравитация. Но это приводит к введению суперчастиц – целому поколению новых элементарных частиц. И мы надеемся, что рождение таких новых частиц будет возможно на LHC (Большой адронный коллайдер в CERN, Женева), который должен заработать в 2007 году.

Физика частиц на ускорителях и гравитация. Ю.М. Зиновьев // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 248–253.

Обычная формулировка гравитации содержит только один масштаб – массу Планка ~ 10^{19} ГэВ. Поэтому многие годы казалось, что гравитационные взаимодействия слишком слабы, чтобы быть измеримыми в ускорительных экспериментах. Но исследования в супергравитации и теории струн возобновили интерес к старому вопросу – какова в действительности размерность нашего пространства? Среди многих возможностей, которые рассматривались за последние 30 лет, есть интересный сценарий, в котором реальный масштаб гравитации может быть существенно ниже, например 10^3 – 10^4 ГэВ. Это означает, что эксперименты на LHC могут открыть новую физику, связанную с гравитацией!

«Центральная машина» в активных галактических ядрах. В.С. Бескин // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 254–267.

Многие компактные астрофизические источники имеют струйные выбросы. Это говорит в пользу единого механизма извлечения энергии, коллимации и ускорения частиц. Согласно существующей точке зрения «центральная машина» должна быть связана с быстровращающейся молодой звездой, нейтронной звездой или черной дырой, работающей как униполярный индуктор в присутствии сильного магнитного поля. С другой стороны, долгое время не было ясно, можно ли непосредственно использовать стандартный подход к магнитосфере черной дыры, где релятивистские эффекты играют ведущую роль. Мы показываем, что черная дыра, вращающаяся во внешнем магнитном поле, может работать как униполярный индуктор несмотря на то, что сам механизм извлечения энергии не связан с электрическими токами, текущими по горизонту.

Физика релятивистских джетов. Е.В. Деришев, В.В. Кочаровский, Вл.В. Кочаровский, В.Ю. Мартыанов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 268–288.

Дано краткое описание основных наблюдательных фактов и теоретических представлений о релятивистских струях плазмы (джетах) в активных ядрах галактик, микроквазарах и источниках космических гамма-всплесков. Обсуждаются физические особенности релятивистской неравновесной плазмы, ответственные за динамику джетов, ускорение частиц и их излучение, а также за формирование магнитного поля.

Самоорганизация плазмы и магнитного поля в неравновесных астрофизических объектах. И.Н. Топтыгин // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 289–303.

Рассматриваются проблемы, связанные с самоорганизацией плазмы и магнитного поля: генерация бесстолкновительных ударных волн в неравновесных астрофизических объектах; нелинейные эффекты при ускорении частиц на сильно ударном фронте в турбулентной среде; перестройка крупномасштабного магнитного поля ускоренными частицами вблизи ударного фронта; нерезонансная раскачка крупномасштабных альфвеновских волн ускоренными частицами.

Пульсации звезд. Ю.А. Фадеев // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 304–319.

При рассмотрении радиальных звездных пульсаций предполагается, что сферически-симметричные движения вещества обусловлены действием силы тяжести и градиента давления, эффекты вязкости пренебрежимо малы, а перенос энергии описывается уравнением лучистой теплопроводности. Решение линеаризованных уравнений представляет набор собственных функций, соответствующих колебаниям типа стоячей волны со свободной внешней границей. Возбуждение неустойчивости и установление автоколебаний предельного цикла связаны с модуляцией потока излучения в слоях частично ионизованного гелия. Периодические ударные волны, возникающие при нелинейных пульсациях большой амплитуды, являются причиной истечения вещества из звезды. Основная нерешенная проблема теории звездных пульсаций заключается в отсутствии корректного описания переноса энергии турбулентной конвекцией.

Ударные волны в атмосферах пульсирующих звезд. Ю.А. Фадеев // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 320–334.

Протяженность ударной волны в атмосфере пульсирующей звезды значительно меньше шкалы высот, поэтому предположения стационарного потока и плоской геометрии оказываются хорошим приближением при рассмотрении структуры волны. Уравнения газовой динамики и атомной кинетики сводятся к системе жестких обыкновенных дифференциальных уравнений, тогда как проблема переноса лучистой энергии представляет собой краевую задачу. Самосогласованное решение всех уравнений может быть получено итерационным способом. Взаимосвязь газодинамического потока и поля излучения, порождаемого этим потоком, особенно велика в частично ионизованном водородном газе вследствие малой длины пробега лаймановских фотонов и образования области лучистого прогрева перед ударным разрывом. Основная доля излучения ударной волны возникает на ее хвосте в зоне рекомбинации атомов водорода.

Нелинейная динамика высотного грозового разряда. Е.А. Мареев, В.В. Клименко // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 337–359.

Дан краткий обзор физических характеристик спрайтов – пожалуй, наиболее интересного типа высотных разрядов, которые наблюдаются в мезосфере и тесно коррелируют с положительными разрядами молнии – и методов их диагностики. Спрайты рассматриваются как часть динамического процесса, включающего формирование поля и заряда в облаке и молниевый разряд облако-земля, приводящий к инициации разряда в мезосфере. В динамике спрайта важнейшую роль играет крупномасштабная неоднородность среды, приводящая к развитию таунсендовского разряда вблизи мезопаузы и формированию стримеров в нижней части мезосферы. Рассмотрены приближенные решения кинетического уравнения для электронов в слабоионизованной плазме в электрическом поле и на основе этих решений даны оценки основных характеристик лавинной стадии размножения электронов применительно к условиям мезосферы.

Фрактальные лабиринты: структурная динамика. Д.И. Иудин, В.Ю. Трахтенгерц // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 360–377.

Фрактальные лабиринты встречаются в природе и технике повсеместно. Это транспортные и коммуникационные сети, системы сбора и распределения ресурсов и информации, речные системы, системы кровоснабжения, молниевые разряды и т. д. Фрактальный характер геометрии этих систем предопределяет особенности их динамического поведения и транспортных свойств.

Волны цунами: общая характеристика явления, механизмы образования, прогноз. М.А. Носов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 378–392.

Изложена общая информация о явлении цунами. Описываются механизмы генерации волн подводными землетрясениями, оползнями, вулканическими извержениями, падением метеоритов и подвижными атмосферными фронтами. Обсуждаются роль сжимаемости воды и нелинейности и их возможный вклад в амплитуду цунами. Описываются возможности и перспективы развития методов прогноза цунами.

Нелинейно-дисперсионная теория волн цунами: взгляд после катастрофического цунами в Индийском океане. Е.Н. Пелиновский // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 393–407.

Обсуждаются современное состояние нелинейно-дисперсионной теории волн на воде и ее применимость к описанию цунами в океане. Сначала описывается нелинейная теория мелкой воды, в основе которой лежит двумерное волновое уравнение. В рамках этой модели рассчитано большинство исторических цунами. Затем на физическом уровне обсуждаются дисперсионные эффекты, связанные с сверхдальним распространением цунами. В последнее время разработаны сильно нелинейные и дисперсионные модели Буссинесковского типа, однако расчеты на большие расстояния пока весьма трудоемки и не гарантируют высокой точности. В заключение обсуждается роль солитонов в проблеме цунами. Развиваемые сейчас теории иллюстрируются на примере наиболее разрушительного цунами, случившегося в Индийском океане 26/12/2004 г.

Нелинейные механизмы и модели формирования спутниковых изображений процессов в верхнем слое океана. Ю.И. Троицкая, С.А. Ермаков // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 408–430.

Представлен обзор развитых в последнее время моделей поверхностного проявления процессов в верхнем слое океана, в которых существенную роль играет нелинейность волновых процессов. Обсуждаются основные нелинейные механизмы, ответственные за модуляцию поверхностных волн в присутствии неоднородных течений. В частности, обсуждается механизм генерации мелкомасштабных поверхностных волн гравитационными волнами большой амплитуды и его роль при формировании гидродинамического и радиолокационного контраста в присутствии неоднородных течений на морской поверхности. Описан механизм модуляции ветрового инкремента поверхностных волн с учетом нелинейности взаимодействия волн с ветром. Представлены результаты проверки развитых моделей в лабораторных и натуральных экспериментах, а также примеры интерпретации спутниковых изображений.

Ренормгрупповые симметрии в краевых задачах. В.Ф. Ковалев, Д.В. Ширков // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 433–446.

Дается обзор исследований, связанных с использованием и расширением понятий функциональной автоматодельности и ренормгруппы Боголюбова в краевых задачах математической физики. Наиболее существенным достижением при переносе ренормгрупповых представлений из квантовой теории поля в математическую физику явилось создание алгоритма нахождения симметрий ренормгруппового типа с помощью современной теории

групп преобразований. Предложен общий алгоритм построения ренормгрупповых симметрий для математических моделей, использующих дифференциальные уравнения. Дается обобщение ренормгруппового алгоритма и связанного с ним понятия ренормгрупповой симметрии для задач с нелокальными (интегральными) уравнениями. Обсуждается приложение этого обобщенного алгоритма к моделям, содержащим линейные функционалы от решений.

Вырожденные резонансы. А.Д. Морозов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 447–460.

Рассматриваются резонансы в двухчастотных системах, близких к гамилтоновым. Дается определение резонанса и вырожденного резонанса. Для систем с $3/2$ степенями свободы приводится двумерная усредненная система, описывающая поведение решений в индивидуальной резонансной зоне. Основное внимание уделяется вырожденным резонансам. В случае вырожденных резонансов в системах с двумя степенями свободы усредненная система трехмерная. Приводятся нормальные формы усредненных систем в эллиптическом, гиперболическом и параболическом случаях. Рассматриваются сохраняющие площадь отображения цилиндра с немонотонным числом вращения. Обращается внимание на два типа бифуркаций перехода от невырожденных резонансов к вырожденному резонансу: а) с образованием «вихревых пар».

Затягивание потери устойчивости при динамических бифуркациях. А.И. Нейштадт // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 461–476.

В классической теории бифуркаций поведение системы, зависящей от параметра, рассматривается для значений параметра, близких к некоторому критическому бифуркационному значению. В теории динамических бифуркаций параметр медленно изменяется во времени и проходит через значение, которое было бы критическим в классической статической теории. Пусть при бифуркационном значении параметра положение равновесия или предельный цикл теряют асимптотическую устойчивость в линейном приближении, оставаясь невырожденными. В аналитических системах потеря устойчивости неизбежно затягивается: фазовые точки остаются вблизи неустойчивого равновесия ещё в течение долгого времени после момента бифуркации; за это время параметр изменяется на величину порядка 1. Такое затягивание потери устойчивости отсутствует в неаналитических системах.

Бактериальная клетка с физической точки зрения. А.А. Веденов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 479–490.

Бактериальная клетка рассматривается как совокупность полимеризующих молекулярных машин: ДНК-полимераз, РНК-полимераз и рибосом. Условие синхронизации машин позволяет оценить число РНК-полимераз и рибосом в фазе экспоненциального роста. Экспериментальные данные о проникновении антибиотика в клетку и гипотеза о существовании только одной мишени у антибиотика позволяют оценить декремент экспоненциального затухания числа бактерий в популяции в растворе с антибиотиком и минимальную ингибирующую концентрацию для бактерий дикого типа и мутантов.

Геном как фрактальное множество, порождаемое инверсионно-диффузионной динамикой. С.А. Ларионов, А.Ю. Лоскутов, С.Д. Рыбалко, Е.В. Рядченко // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 491–508.

На основе представления последовательности нуклеотидов ДНК как двумерного блуждания на плоскости предлагается новый способ изучения и сравнения различных участков хромосом. В частности, посредством данного представления впервые обнаружены некоторые новые особенности ДНК. Проанализированы также свойство самоподобия и некоторые известные функциональные и структурные элементы. Впервые представлены некоторые из полностью расшифрованных хромосом человека.

Роль неустойчивостей, нелинейных колебаний и волн в регуляции тубулинового цитоскелета. Е.А. Катруха, Г.Т. Гурия // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 509–534.

Динамические неустойчивости и критические явления, свойственные элементам клеточного цитоскелета – тубулиновым микротрубочкам, – рассматриваются с позиций теории диссипативных структур. Построена феноменологическая модель, описывающая ключевые эффекты, свойственные динамике тубулиновых волокон.

Аэроионы Чижевского и явление свечения влажного воздуха. В.Л. Воейков, С.Ф. Чалкин, Р.Р. Асфарамов // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 535–547.

Одной из проблем, связанных с механизмом биологического действия отрицательно заряженных аэроионов (активных форм кислорода), является их перенос на значительные расстояния, несмотря на их малое время жизни. Мы обнаружили, что при поглощении отдельных УФ-фотонов пленкой воды, смачивающей гигроскопичную поверхность, происходит вспышка излучения фотонов, которая инициирует в контактирующем с пленкой влажном воздухе окислительные реакции, сопровождаемые генерацией электронного возбуждения. Возбуждение распространяется по воздуху на макроскопические расстояния. При определенной влажности возбуждение воздуха приобретает колебательно-волновой характер. Данное явление может вносить свой вклад как в механизм действия аэроионов, получаемых при искусственной ионизации воздуха, так и в природные процессы, где существует структурированная вода.

Биологический вид волоконных световодов. Ю.Н. Кульчин, О.А. Букин, С.С. Вознесенский, А.Н. Галкина, С.В. Гнеденков, А.Л. Дроздов, В.Г. Куравый, Т.Л. Мальцева, С.Л. Синябрухов, А.И. Чередниченко // *Нелинейные волны–2006.* – Нижний Новгород: ИПФ РАН, 2007. С. 548–559.

Идея создания волоконных световодов была реализована природой задолго до появления человека на Земле. Стекланные губки представляют собой группу древнейших многоклеточных животных, которые обитают преимущественно в глубоководных зонах океана, закрепившись в донных осадках с помощью кремнийсодержащих спикул. Скелет глубоководных губок состоит из индивидуальных кремнеземных элементов, называемых спикулами. Геометрические размеры спикул, в зависимости от вида и возраста, могут изменяться в достаточно широких пределах. Представлены результаты структурных, физико-химических исследований спикул стекланных губок. Результаты исследований показали, что спикулы по своим механическим и оптическим свойствам являются уникальными биоматериалами, аналогами фотонных кристаллов и могут в перспективе использоваться в качестве новых материалов для оптоэлектроники.

Periodic orbits, localization in space of normal modes, and the Fermi-Pasta-Ulam problem. S.G. Flach.

The excitation of a long wavelength mode in a finite anharmonic atomic chain is expected to quickly share its energy with other normal modes due to anharmonicity. Fermi, Pasta and Ulam (FPU) discovered 50 years ago that this must not be the case. It is shown that normal modes of the harmonic chain can be continued into the anharmonic case as exact periodic orbits (PO). These orbits are characterized by exponential localization in normal mode space. The FPU trajectory is a slight perturbation off the exact PO, and remains close to it for long times. The concept of these POs can be extended to two-, three-dimensional and infinite lattices, with surprising consequences.

Symmetries and their violation in systems in external variable fields. S.G. Flach.

The symmetry analysis approach to the response of nonlinear dynamical systems to external time-periodic fields is discussed. In the simplest case of a single particle in a one-dimensional periodic potential the symmetries and ways of their violation in the case of the ratchet effect, i.e. when the motion of the particle is rectified, are considered. The results have been recently confirmed in experiments on directed diffusion of cold atoms in optical lattices. The symmetry approach is extended to the case of nonlinear partial differential equations and necessary conditions for appearance of a directed energy flow are obtained. This flow is realized with topological solitons and was observed in Josephson ring junctions. Finally, the symmetry analysis is applied to the case of quantum spins in ac magnetic fields. Ways of achieving a dc magnetization component which is perpendicular to the applied fields are considered.

Phase compactons. A.S. Pikovsky, Ph. Rosenau.

Phase dynamics in a chain of dispersively coupled autogenerators is considered. In the quasicontinuous limit, a Korteweg-de Vries equation with nonlinear dispersion is the basic model. Its solutions are compactons – solitary waves with compact carrier, and covatons – compact, arbitrarily spaced apart kink-antikink pairs propagating with the same velocity.

Spatio-temporal dynamics of topological features of wave fields in nonequilibrium media: sources, sinks, holes, defects. A.B. Ezersky.

Investigations of spatio-temporal dynamics of periodic structures in nonequilibrium media are frequently based on topological features of the fields of order parameter such as temperature, velocity, impurity concentration, and others, rather than on the order parameter field as a whole. The main idea of the lecture is to demonstrate that such topological features may be identified in spatially periodic systems produced by hydrodynamic instability and compared with solutions of reference equations.

On realization of some classic models and phenomena of nonlinear dynamics in coupled non-autonomous oscillators. S.P. Kuznetsov.

The lecture is devoted to a certain general idea of realization of abstract models and phenomena of nonlinear dynamics on the base of a special class of systems that may be implemented, e.g., as electronic devices. These systems are composed of two or more coupled self-oscillators, which become active turn by turn due to periodic forcing (modulation of the parameters) with a transfer of the excitation from each subsystem to its partner as it becomes active. The models under consideration are represented by ordinary differential equations with periodic coefficients. Computer results illustrate the expected dynamical phenomena.

Phase clusters and short-term memory in a neuronal network with afterdepolarization. V.V. Klinshov, V.I. Nekorkin.

An oscillatory model of the system of working memory of the brain in the form of a multidimensional dynamic system describing a network of pyramidal neurons with afterdepolarization under the action of subthreshold periodic signal is considered. It is shown that in such a system

information can be stored in the form of clusters of periodic activity. The most interesting property of the system is its capability to remember several images simultaneously, which is the basic feature of the working memory of the brain. Storage of several images corresponds to existence in the system of several successively excited clusters of periodic activity. Each cluster represents one image and is excited in the phase of its external signal. Maximal number of images that can be stored in the system simultaneously is plotted as a function of coupling parameter.

Continuous wavelet analysis and some applications to problems of nonlinear dynamics.

A.A. Koronovsky, A.E. Khramov.

The lecture deals with the basic notions of wavelet analysis that is a new field of modern science. Different aspects of the continuous wavelet transform and its applications to analysis of complex signals generated by nonlinear dynamical systems are considered. The methods of effective selection of laminar and turbulent phases from time series of different systems (including biological ones) as well as diagnostics of the chaotic synchronization regime by means of a continuous set of chaotic signal phases are described.

Communication systems with chaotic carriers: problems, advances and prospects.

A.S. Dmitriev.

We consider the communication technology that uses dynamical chaos as information carrier. It is shown that it is currently moving from the stage of fundamental research to the stage of practical applications.

Mathematical models, exact solutions, methods of analysis. O.V. Rudenko.

Some models of nonlinear acoustics that are interesting to other branches of the physics of nonlinear waves are considered. Physically meaningful exact solutions are presented. Methods of seeking or guessing such solutions are described. The connection between "nonrigorous" approaches and correct mathematical analysis of nonlinear models on the basis of their symmetries is demonstrated on concrete examples.

Topical problems related to nonlinear acoustics. O.V. Rudenko.

The lecture concerns the trends that have attracted primary attention in the recent years: 1. Acoustic shock from bodies moving with supersonic speed. Phenomena occurring at formation of N-wave in an atmosphere. Emissions and their ecological implications. Shock and perspectives of supersonic civil aviation. 2. Sonofusion – the problem of implementing nuclear fusion in a collapsing bubble at cavitation. Radiation more shortwave than at sonoluminescence of a single bubble. The phenomenon of unlimited cumulation. 3. The property of skeletal muscles to damp shocks. Nonlinear dynamics of muscle contraction. Data of measurements of shear elasticity of muscles.

On «non-classical» manifestations of structurally induced acoustic nonlinearity on an example of contact-containing media. V.Yu. Zaitsev.

The physical origin of «non-classical» nonlinear properties of microinhomogeneous media is discussed. Examples of using nonlinear effects for acquiring information about medium microstructure are presented. Also given are examples of «non-classical» nonlinear effects associated not only with nonlinearly elastic behavior but also with amplitude-dependent absorption (in particular, observations of the acoustic analog of the Luxembourg-Gorky effect), manifestations of slow dynamics due to thermoelastic effects, etc.

Vortex dynamics in layered superconductors with internal Josephson effect. V.V. Kurnn, A.V. Chiginev.

A microscopic device of high-temperature cuprate superconductors and ways of describing different interacting degrees of freedom of a superconductor are considered. The Josephson effect and its manifestations in layered high-temperature superconductors are briefly discussed. Different mechanisms of coupling between Josephson layers in high-temperature superconductor are surveyed and their influence on vortex layer structure is analysed. Special attention is focused on the influence of optical phonons on stability of in-phase motion of Josephson vortices in high-temperature superconductors.

Stimulated Brillouin scattering of Laguerre-Gaussian laser modes: new phenomena.

F.A. Starikov.

It is shown that the phase conjugation at stimulated Brillouin scattering (SBS) of focused vortex Laguerre-Gaussian beams is not realized as a phenomenon, as SBS gain coefficients of conjugated Stokes mode and analogous mode of opposite helicity are the same. Therefore, for example, for doughnut LG_0^1 laser mode, the Stokes beam consists of random superposition of these modes. For a rather wide class of vortex beams (for example, for LG_1^1 laser mode), another phenomenon is predicted that can be referred to as *phase transformation* at SBS. It is essentially the following: in the Stokes beam a nonconjugated mode is selected, that is a principal Gaussian mode orthogonal to the laser mode.

Generation of fast electrons at electron-ion collisions in strong fields.

G.M. Fraiman,
A.A. Balakin.

Electron-ion collisions in strong electromagnetic fields in plasma are considered. A method of their analytical and numerical description on the basis of point mapping is proposed. Point map analysis reveals basic features of scattering in strong fields: arising of multiframe stochastic motion of particles and «singularity» of concentration prior to «hard» impact. Effective frequency of collisions and energy distribution of particles as a result of collisions in nonrelativistic and ultrarelativistic limits are estimated. It is shown that all the energy effects due to collisions (Joule heating, fast electron generation, radiation at high harmonics) increase with increasing field. The obtained results are compared with data of experiments.

Gravity and supergravity.

Yu.M. Zinoviev.

A constructive approach considers gravity as a "normal" field theory describing massless spin-2 particles exactly in the same way as electrodynamics describes massless spin-1 particles – photons. In this way all the nonlinear interactions of gravity (self-interaction and interaction with matter) could be reproduced and in this sense understood. But quantization leads to a non-renormalizable theory. One of the most natural ways to better the quantum theory is to consider supersymmetric extension of gravity – supergravity. But this results in the introduction of superparticles, a whole generation of new elementary particles. And we hope that production of such new particles will be possible on LHC (Large Hadron Collider in CERN, Geneva) which is to be started in 2007.

Gravity and accelerator physics.

Yu.M. Zinoviev.

Common formulation of the gravity theory contains one natural scale only – the so-called Planck mass $\sim 10^{19}$ GeV. For many years it seemed that gravitational interactions are too weak to be measurable in any accelerator experiments. But investigations in supergravity and superstring theories renewed an interest in the old question: how many dimensions does our space-time really have? Among many possibilities that have been considered in the last 30 years there is a very interesting scenario which makes possible that real scale of gravity could be much lower, say $10^3 \sim 10^4$ GeV. This means that new experiments on LHC could reveal new physics related with gravity!

«Central engine» in active galactic nuclei.

V.S. Beskin.

Many compact astrophysical sources have jets. That says in favour of the unique mechanism of energy extraction, collimation, and particle acceleration. According to the present point of view the «central engine» is to be connected with a rapidly rotating young star, neutron star or black hole working as unipolar inductor in the presence of strong magnetic field. On the other hand, it was not clear that one can use directly the standard approach to the black hole magnetosphere where the general relativistic effects play the leading role. We show that a black hole rotating in the external magnetic field can work as a unipolar inductor, although the mechanism of energy extraction is not connect with the electric currents flowing along the horizon.

The physics of relativistic jets. E.V. Derishev, V.V. Kocharovskiy, V.I. Kocharovskiy, V.Yu. Martyanov.

Basic observations and theoretical treatment of relativistic plasma outflows (jets) in active galactic nuclei, microquasars, and cosmic gamma-ray bursts are considered. Peculiarities in the physics of relativistic non-equilibrium plasma responsible for the jet dynamics, particle acceleration and emission, and magnetic field formation are discussed.

Self-organization of plasma and magnetic field in nonequilibrium astrophysical objects. I.N. Toptygin.

The lecture concerns problems of self-organization of plasma and magnetic field such as generation of collisionless shock waves in nonequilibrium astrophysical objects; nonlinear effects accompanying particle acceleration at strong shock front in a turbulent medium; restructuring of large-scale magnetic field by accelerated particles near shock front; nonresonant build-up of large-scale Alfvén waves by accelerated particles.

Stellar pulsations. Yu.A. Fadeyev.

The radial stellar pulsations are assumed to be spherically symmetric motions of gas due to gravity and the pressure gradient with viscosity neglected and the energy transfer treated as radiation heat conduction. The eigenfunctions of the linearized equations correspond to standing waves with free outer boundary. Both excitation of the instability and attainment of the limiting cycle are due to the radiation flux modulation in the layers of the partially ionized helium. Periodic shock waves accompanying the large-amplitude nonlinear pulsations are responsible for gas outflows from the star. The energy transfer by turbulent convection remains the main unsolved problem of the stellar pulsation theory.

Shock waves in atmospheres of pulsating stars. Yu.A. Fadeyev.

The thickness of shocks in atmospheres of pulsating stars is much smaller than the scale height, so that assumptions of the steady-state flow and the planar geometry allow us to consider shock wave structure to a good accuracy. Equations of gas dynamics and atomic kinetics are stiff ordinary differential equations, whereas the radiation transfer is treated as a two-point boundary value problem. The self-consistent solution is obtained using the iteration procedure. The coupling between the gas flow and the radiation field created by this flow is strongest in the partially ionized hydrogen gas because of the small mean free path of Lyman photons and the existence of the radiative precursor ahead of the discontinuity. The major portion of radiation is created in the tail of the shock wave in the hydrogen recombination zone.

Nonlinear dynamics of high-altitude storm discharge. E.A. Mareev, V.V. Klimenko.

The lecture is concerned with the physics of sprites, one of the most interesting types of high-altitude discharges observed in the mesosphere that strongly correlate with positive strikes. Physical characteristics of sprites and methods of their diagnostics are briefly reviewed. Sprites are considered as part of the dynamical process that includes field and charge formation in the cloud and the cloud-earth strike initiating discharge in the mesosphere. Large-scale medium inhomogeneity giving rise to development of Townsend discharge near mesopause and to streamer formation in the lower mesosphere plays an important role in sprite dynamics. Approximate solutions of kinetic equation for electrons in weakly ionized plasma in electric field are considered; basic characteristics of avalanche multiplication of electrons in application to the conditions of the mesosphere are accessed.

Fractal labyrinths: structural dynamics. D.I. Iudin, V.Yu. Trakhtengerts.

Fractal labyrinths are encountered everywhere in the nature and in technics. These are transport and communication networks, systems for collection and distribution of resources and information, river systems, blood supply systems, striking, and so on. The fractal character of geometry of these systems specifies features of their dynamic behavior and transport properties.

Tsunamis: general characteristics of the phenomenon, mechanisms of formation, prognosis. M.A. Nosov.

This paper is intended to outline the state-of-the-art knowledge on the tsunamis. First, we provide general information on the tsunami phenomenon. Then, we describe up to date models of tsunami generation by submarine earthquakes, landslides, volcanic eruptions, meteorites impacts, and moving atmospheric pressure fronts. The role of water compressibility and non-linearity is considered in the context of their contribution to tsunami wave amplitude. Special attention is paid to the existing and perspective methods of tsunami forecast.

Nonlinear-dispersive theory of tsunami waves: outlook after the hazardous Indian Ocean tsunami. E.N. Pelinovsky.

State-of-the-art of the nonlinear-dispersive theory of water waves and its application to tsunami are discussed. First, the nonlinear shallow-water theory based on 2D wave equation is described. Most of the historic tsunamis have been modeled in the framework of this model. Second, dispersion effects for tsunami waves at long distances are considered at the physical level. Strongly nonlinear and dispersive Boussinesq systems have been developed by now, but long distance computing requires high resources and does not guarantee high accuracy. In conclusion, the role of solitary waves (solitons) in the tsunami problem is discussed. The currently developed theories are illustrated on an example of the hazardous Indian Ocean tsunami of 26/12/2004.

Nonlinear mechanisms and models of satellite imaging of processes in the upper ocean. Yu.I. Troitskaya, S.A. Ermakov.

The recently developed models of surface manifestations of processes occurring in the upper ocean, where nonlinearity plays a significant role, are surveyed. Basic nonlinear mechanisms responsible for surface wave modulation in the presence of inhomogeneous flows are discussed. In particular, the mechanism of generation of small-scale surface waves by large-amplitude gravity waves and its role in formation of hydrodynamic and radar contrast in the presence of inhomogeneous flows on sea surface are considered. The mechanism of modulation of surface wave wind increment with allowance for nonlinearity of wave-wind interaction is described. Results of testing the models in laboratory and full-scale experiments, as well as examples of interpreting satellite images are presented.

Renormgroup symmetry for solutions of boundary value problems. V.F. Kovalev, D.V. Shirkov.

We review the evolution of the renorm group algorithm and the related renorm group symmetry introduced in mathematical physics for solution of boundary value problems described by differential equations. Different methods of calculating these symmetries based on modern group analysis are described. We discuss the essential progress made recently in the application of this algorithm to models with integral equations. Several physical illustrations from nonlinear physics demonstrate the potentialities of the algorithm for both, models with differential equations and models with nonlocal terms in the form of integral equations and linear solution functionals.

Degenerate resonances. A.D. Morozov.

Resonances in two-frequency systems close to the Hamiltonian one are considered. Resonance and degenerate resonance are defined. A two-dimensional averaged system describing behaviour of solutions in an individual resonance zone is presented for systems with $3/2$ degrees of freedom. Primary attention is focused on degenerate resonances. In the case of degenerate resonances in systems with two degrees of freedom an averaged system is three-dimensional. Normal forms of averaged systems are presented for elliptic, hyperbolic, and parabolic cases. Area-preserving maps of a cylinder with non-monotonic twist maps are considered. The attention is drawn to two types of bifurcations of the transition from non-degenerate resonances to degenerate resonance: a) with formation of loops, b) with formation of «vortical pairs».

Stability loss delay for dynamical bifurcations. A.I. Neishtadt.

In the classical bifurcation theory behaviour of systems is considered, depending on parameter, for values of the parameter close to some critical, bifurcational one. In the theory of dynamical bifurcations the parameter is changing slowly in time and passes through the value that would be bifurcational in the classical static theory. Let at a bifurcational value of the parameter the equilibrium or the limit cycle lose asymptotic linear stability but remain nondegenerate. In analytical systems the stability loss is inevitably delayed: the phase points remain near the unstable equilibrium for a long time after the bifurcation; during this time the parameter changes by a magnitude of order 1. Such delay is generally not found in non-analytical systems.

Bacterial cell in terms of physics. A.A. Vedenov.

A bacterial cell is considered as an ensemble of polymerizing molecular machines: DNA-polymerases, RNA-polymerases, and ribosomes. The condition of machine synchronization allows estimation of the number of RNA polymerases and ribosomes at the phase of exponential growth. Available experimental data on antibiotic invasion into a cell and the hypothesis that an antibiotic has only one target enable one to determine the exponential decrement of the amount of bacteria in the population with antibiotic and minimal inhibitory concentration for wild or mutant bacteria.

Genome as a fractal set generated by inversion-diffusion dynamics. S.A. Larionov, A.Yu. Loskutov, S.D. Rybalko, E.V. Ryadchenko.

A new method of investigation and comparison of different sections of chromosomes is proposed based on representing a sequence of DNA nucleotides as two-dimensional wandering on a plane. Such a representation, in particular, revealed some new features of DNA. The properties of similarity and some known functional and structural elements were analyzed. Several fully identified human chromosomes are presented for the first time.

The role of instabilities, nonlinear oscillations and waves in tubulin cytoskeleton. E.A. Katrukha, G.T. Guria.

Dynamic instabilities and critical phenomena of cytoskeleton filaments called microtubules are considered within the framework of the self-organization theory. An extended phenomenological model reflecting key features of microtubule dynamics is formulated.

Chizhevsky air ions and the phenomenon of humid air luminance. V.L. Voeikov, S.F. Chalkin, R.R. Asfaramov.

One of the problems related to the mechanism of biological action of light negative air ions (reactive oxygen species) is transmission of these short-living chemical species at long distances. We discovered the phenomenon that may probably resolve this contradiction. When a thin layer of water hydrating a hygroscopic surface absorbs rare UV-photons capable of splitting water molecules, a flash of photon emission in UV- and visible regions of the spectrum is observed. This flash initiates oxidative reactions accompanied by generation of electronic excitation in humid air contacting water film. Excitation propagates through the air at macroscopic distances. Depending on air humidity, it may gain the oscillatory-wave character. This phenomenon may endow into some natural processes where ordered water films may form.

Biological species of fiber-optic waveguides. Yu.N. Kulchin, O.A. Bukin, S.S. Voznesensky, A.N. Galkina, S.V. Gnednikov, A.L. Drozdov, V.G. Kuryavy, T.L. Maltseva, S.L. Sinebruchov, A.I. Therednichenko.

The idea of Nature as an engineer of fiber-optic waveguides is quite old. The deep-sea are the oldest metazoan phylum still extant today that have basalja (anchorage) fibrous that permit successful colonization on soft sediments. Most sponges produce silica-made skeletons consisting of individual elements (spicules). Depending on the type, age and morphology of sponges, the proportions of their spicules vary from several centimeters to 2 meters in length and from 0.1 to 100 mm in diameter. Results of structural, optical, and spectral properties of biosilica spicules are presented. The analysis indicates that biosilica spicules are unique biomaterials promising for photonics and microelectronics applications.

CONTENTS

<i>Preface</i>	5
----------------------	---

Nonlinear dynamics of complex systems

<i>S.G. Flach</i> . Periodic orbits, localization in space of normal modes, and the Fermi-Pasta-Ulam problem	9
<i>S.G. Flach</i> . Symmetries and their violation in systems in external variable fields	24
<i>A.S. Pikovsky, Ph. Rosenau</i> . Phase compactons.....	39
<i>A.B. Ezersky</i> . Spatio-temporal dynamics of topological features of wave fields in nonequilibrium media: sources, sinks, holes, defects	55
<i>S.P. Kuznetsov</i> . On realization of some classic models and phenomena of nonlinear dynamics in coupled non-autonomous oscillators	68
<i>V.V. Klinshov, V.I. Nekorkin</i> . Phase clusters and short-term memory in a neuronal network with afterdepolarization.....	85
<i>A.A. Koronovsky, A.E. Khramov</i> . Continuous wavelet analysis and some applications to problems of nonlinear dynamics.....	102
<i>A.S. Dmitriev</i> . Communication systems with chaotic carriers: problems, advances and prospects.....	118

Nonlinear oscillations in structured systems

<i>O.V. Rudenko</i> . Mathematical models, exact solutions, methods of analysis.....	139
<i>O.V. Rudenko</i> . Topical problems related to nonlinear acoustics	151
<i>V.Yu. Zaitsev</i> . On «non-classical» manifestations of structurally induced acoustic nonlinearity on an example of contact containing media.....	170
<i>V.V. Kurin, A.V. Chiginev</i> . Vortex dynamics in layered superconductors with internal Josephson effect.....	191
<i>F.A. Starikov</i> . Stimulated Brillouin scattering of Laguerre-Gaussian laser modes: new phenomena	206
<i>G.M. Fraiman, A.A. Balakin</i> . Generation of fast electrons at electron-ion collisions in strong fields	222

Nonlinear effects in gravity and astrophysical plasma

<i>Yu.M. Zinoviev</i> . Gravity and supergravity.....	239
<i>Yu.M. Zinoviev</i> . Gravity and accelerator physics	248

<i>V.S. Beskin.</i> «Central engine» in active galactic nuclei.....	254
<i>E.V. Derishev, V.V. Kocharovskiy, V.I. Kocharovskiy, V.Yu. Martyanov.</i> The physics of relativistic jets.....	268
<i>I.N. Toptygin.</i> Self-organization of plasma and magnetic field in nonequilibrium astrophysical objects.....	289
<i>Yu.A. Fadeyev.</i> Stellar pulsations.....	304
<i>Yu.A. Fadeyev.</i> Shock waves in atmospheres of pulsating stars.....	320

Nonlinear phenomena in the environment

<i>E.A. Mareev, V.V. Klimenko.</i> Nonlinear dynamics of high-altitude storm discharge	337
<i>D.I. Iudin, V.Yu. Trakhtengerts.</i> Fractal labyrinths: structural dynamics.....	360
<i>M.A. Nosov.</i> Tsunamis: general characteristics of the phenomenon, mechanisms of formation, prognosis.....	378
<i>E.N. Pelinovsky.</i> Nonlinear-dispersive theory of tsunami waves: outlook after the hazardous Indian Ocean tsunami	393
<i>Yu.I. Troitskaya, S.A. Ermakov.</i> Nonlinear mechanisms and models of satellite imaging of processes in the upper ocean.....	408

Mathematical problems of nonlinear dynamics

<i>V.F. Kovalev, D.V. Shirkov.</i> Renormgroup symmetry for solutions of boundary value problems.....	433
<i>A.D. Morozov.</i> Degenerate resonances	447
<i>A.I. Neishtadt.</i> Stability loss delay for dynamical bifurcations	461

Nonlinear phenomena in living systems

<i>A.A. Vedenov.</i> Bacterial cell in terms of physics.....	479
<i>S.A. Larionov, A.Yu. Loskutov, S.D. Rybalko, E.V. Ryadchenko.</i> Genome as a fractal set generated by inversion-diffusion dynamics.....	491
<i>E.A. Katrukha, G.T. Guria.</i> The role of instabilities, nonlinear oscillations and waves in tubulin cytoskeleton	509
<i>V.L. Voeikov, S.F. Chalkin, R.R. Asfaramov.</i> Chizhevsky air ions and the phenomenon of humid air luminance	535
<i>Yu.N. Kulchin, O.A. Bukin, S.S. Voznesensky, A.N. Galkina, S.V. Gnedenkov, A.L. Drozdov, V.G. Kuryavy, T.L. Maltseva, S.L. Sinebruchov, A.I. Therednichenko.</i> Biological species of fiber-optic waveguides.....	548

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2006

Сборник научных трудов

Утверждено к печати Институтом прикладной физики РАН,
603950 Нижний Новгород, ул. Ульянова, 46

Редакторы *Н. Н. Кралина, Л. О. Кузнецова*
Верстка *М. В. Башевой, С. Н. Новиковой*
Корректор *Д. П. Семенова*

Подписано к печати 15.05.2007 г.
Формат 60 × 90 ¹/₁₆. Бумага офсетная № 1.
Усл. печ. л. 36. Уч.-изд. л. 30. Темплан 2007 г. Поз. 1. Тираж 350 экз.

Отпечатано в полном соответствии с качеством
предоставленных диапозитивов в ГУП ППП «Типография “Наука”», АИЦ РАН
121099 Москва, Шубинский пер., 6

Заказ № 958