

# НЕЛИНЕЙНЫЕ ВОЛНЫ

Динамика  
и ЭВОЛЮЦИЯ



НАУКА

АКАДЕМИЯ НАУК СССР  
ИНСТИТУТ ПРИКЛАДНОЙ ФИЗИКИ

# НЕЛИНЕЙНЫЕ ВОЛНЫ

ДИНАМИКА И ЭВОЛЮЦИЯ

СБОРНИК НАУЧНЫХ ТРУДОВ

Ответственные редакторы  
академик А. В. ГАПОНОВ-ГРЕХОВ  
доктор физико-математических наук М. И. РАБИНОВИЧ



Москва «Наука» 1989

УДК [534.2 + 537.86]:532.59

**Нелинейные волны. Динамика и эволюция.**— М.: Наука, 1989.— 400 с. ISBN 5-02-000669-6.

В сборник включены обзорные и оригинальные статьи, написанные по материалам лекций, прочитанных на VIII Всесоюзной школе по нелинейным волнам (Горький, март 1987 г.). Обсуждаются проблемы динамического хаоса, рождения и формирования структур, биологической эволюции и другие проблемы с общих позиций нелинейной динамики.

Книга рассчитана на специалистов, занимающихся исследованиями нелинейных явлений.

Рецензенты: Д. И. ТРУБЕЦКОВ, В. Д. ШАЛФЕЕВ

Н  $\frac{1604030000-226}{055(02)-89}$  120-89, кн. 2 © Издательство «Наука», 1989

ISBN 5-02-000669-6

---

## ПРЕДИСЛОВИЕ

Если говорить о тенденциях развития нелинейной динамики в ее сегодняшнем понимании, то это в первую очередь наука о порядке, хаосе и структурах в нелинейных средах. Формирование этой междисциплинарной науки произошло в последние два десятилетия благодаря чрезвычайно тесному и неординарному взаимодействию трех направлений исследований:

1. Численные эксперименты на современных компьютерах с использованием визуализационной техники, способствующей углубленному пониманию процессов. Подобные эксперименты позволили не только разобраться во многих весьма старых явлениях, но явились в некотором роде подсказкой для построения приближенных аналитических подходов.

2. Развитие мощных аналитических методов математической физики, таких, например, как качественная теория дифференциальных уравнений в трехмерном пространстве, метод обратной задачи рассеяния для определенных классов нелинейных уравнений в частных производных.

3. Экспериментальное доказательство априорных гипотез об универсальности некоторых основных нелинейных моделей применительно к явлениям самой различной природы (физическим, химическим и т. д.).

Замечательные успехи последних лет (построение теории систем, близких к интегрируемым, создание теории детерминированного хаоса и др.) безусловно настраивают на оптимистический лад. Однако теория нелинейных явлений на сегодняшний день выглядит в значительной степени фрагментарно. Предстоит еще много сделать для унификации имеющихся представлений и разработанных близких методов. Кажется очевидным, что центр интересов нелинейной динамики в ближайшие годы еще более сместится в сторону исследования нелинейных полей и хотя и дискретных, но весьма сложных систем типа нелинейных сетей, решеток и т. д. Эти проблемы сейчас волнуют специалистов по искусственному интеллекту, энергетиков (в связи с анализом энергосетей), радиофизиков и радиоастрономов (специальные антенные решетки) и т. д.

Частично эти тенденции развития нелинейной динамики нашли отражение в лекциях прошедшей в 1987 г. в Горьком VIII Всесоюзной школы по нелинейным волнам, материалы которой представлены в данном сборнике.

Так, в сборнике содержится раздел «Эволюция», основу которого составили статьи В. Г. Бабского «Явления самоорганизации у бактерий на клеточном и популяционном уровнях» и Б. Н. Белинцева «„Полевые“ черты в морфогенезе и механохимическая активность клеток». Компьютерная нелинейная физика представлена статьями А. С. Михайлова «Инженерия динамических систем для распознавания образов и обработки информации» и А. В. Рогальского «Численное моделирование двумерных неравновесных сред на многопроцессорном комплексе». Квантовый цикл — обзорами Д. Л. Шепелянского «Квантовая локализация динамического хаоса», Г. Е. Воловика «Нелинейные явления в физике конденсированных сред: Вселенная в капле гелия».

Одна из особенностей VIII школы — в постановке ретроспективного цикла «Нелинейная динамика — от Пуанкаре до наших дней». Помимо представленных в сборнике, в рамках этого цикла были прочитаны лекции Н. Н. Гаврилова «Автоколебания. Андронов и Ван-дер-Поль» и Я. Г. Синая «Случайное в динамике».

В сборник не вошли также материалы лекций А. А. Андропова, А. В. Муравьева, В. Н. Шастина «Новые полупроводниковые активные среды», В. И. Беспалова, А. А. Бетина «Захватывание поперечной пространственной структуры излучения лазеров», А. А. Галеева «Нелинейная физика плазмы и происхождение космических лучей», В. Е. Захарова «Ударные волны на поверхности тяжелой жидкости», опубликованные ранее в других изданиях.

Не успел представить текст лекции «Флуктуационная кинетика реакций» неожиданно скончавшийся Яков Борисович Зельдович. Он принимал участие в работе большинства горьковских школ, и их слушатели испытали на себе обаяние этого выдающегося ученого, быстроту его реакции, его научный темперамент и исключительную заинтересованность в росте научной молодежи.

Мы благодарим всех авторов, приславших материалы в данный сборник.

*А. В. Гапонов-Грехов, М. И. Рабинович*

# НЕЛИНЕЙНАЯ ДИНАМИКА — ОТ ПУАНКАРЕ ДО НАШИХ ДНЕЙ

---

УДК 531+31.19

## НЕЛИНЕЙНАЯ ДИНАМИКА: ПУАНКАРЕ И МАНДЕЛЬШТАМ

Ю. А. ДАНИЛОВ

Нелинейная динамика, в каком бы — узком или широком — смысле мы ее ни понимали, достигла ныне такого этапа в своем развитии, когда уместно оглянуться на пройденное и подвести некоторые итоги. Период штурма и натиска еще продолжается, но для того чтобы дальнейшее продвижение не замедлилось, чтобы не иссяк наступательный порыв, необходимо критически осмыслить достигнутое, подвергнуть тщательному пересмотру основные идеи и понятия, проследить их происхождение, продумать наиболее рациональную схему планомерной «осады узких мест» и достичь ясного понимания того, что сделано теми гигантами духа и мысли, на плечах которых мы, по признанию Ньютона, стоим.

«Для развития науки,— любил подчеркивать Л. И. Мандельштам,— важна не только работа пионеров, создающих новые концепции, в свете которых становится различимым скрывающееся во мраке неизвестное, но и последующий критический анализ этих концепций, очищающий их от случайного и неверного и вносящий в них стройность, ясность и прозрачность, без которых невозможно дальнейшее продвижение» [1, с. 133].

Вклад А. Пуанкаре и Л. И. Мандельштама в создание нелинейной динамики вряд ли можно переоценить. Им мы обязаны созданием этой новой науки, занимающейся изучением систем различной природы (и поэтому с необходимостью вторгающейся на суверенную территорию различных частных наук), выявляющей общие закономерности там, где их, казалось бы, нельзя было и ожидать — среди пестрого разнообразия внешне далеких явлений, описываемых нелинейными теориями, каждая из которых «говорит на своем языке», ставит и решает свои собственные задачи, используя для этого свои индивидуальные методы. Пуанкаре и Мандельштам — истинные творцы нелинейной динамики: первый создал адекватный математический аппарат, второй насытил абстрактные математические схемы ярким физическим содержанием. Разумеется, каждому из них можно было бы посвя-

тить не одну лекцию, и лишь прокрустово ложе школьного расписания вынуждает нас ограничиться самым необходимым.

И Пуанкаре и Мандельштам — каждый в своей области — принадлежали к редкому типу ученых-универсалов, презревших ограниченность узкой специализации, порожденную дифференциацией науки.

Вот что говорит, например, о Пуанкаре легендарный Никола Бурбаки:

«Каждый год математическая наука обогащается массой новых результатов, приобретает все более разнообразное содержание и постоянно дает ответвления в виде теорий, которые беспрестанно видоизменяются, перестраиваются, сопоставляются и комбинируются друг с другом. Ни один математик не в состоянии проследить все это развитие во всех подробностях, даже если он посвятит этому всю свою деятельность. Многие из математиков устраиваются в каком-нибудь закоулке математической науки, откуда они и не стремятся выйти, и не только полностью игнорируют все то, что не касается предмета их исследований, но не в силах даже понять язык и терминологию своих собратьев, специальность которых далека от их. Нет такого математика, даже среди обладающего самой обширной эрудицией, который бы не чувствовал себя чужеземцем в некоторых областях огромного математического мира; что же касается тех, кто, подобно Пуанкаре или Гильберту, оставляет печать своего гения почти во всех его областях, то они составляют даже среди наиболее великих редчайшее исключение» [2, с. 99—100].

А вот на что считает необходимым обратить внимание в творчестве своего учителя А. А. Андронов:

«Если пользоваться известной терминологией В. Оствальда, Л. И. Мандельштам одновременно и классик — по образцовой ясности и законченности опубликованных им работ, по строгости и точности рассуждений, и романтик — по стремлению делиться своими идеями и догадками, по своей любви к преподаванию, по силе своего живого слова, способного вызвать напряженное внимание и радостное возбуждение аудитории.

...Л. И. Мандельштам ощущал все точное естествознание, включая математику и технику, как единое развивающееся целое и не только подчеркивал взаимное влияние математики и физики, физики и техники и т. д., но хотел каждую новую вещь, будь то квантовая механика или теория нелинейных колебаний, понять и усвоить прочно, как необходимую составную часть всей физики, всего точного естествознания.

...Эта несравненная способность к далеко идущим сопоставлениям сочеталась у Л. И. Мандельштама с большой силой и остротой при конкретном исследовании, с умением преодолевать или обходить экспериментальные или вычислительные трудности» [1, с. 100, 102].

Нужно сказать, что и немногочисленные другие математики-универсалы, отобранные по самому строгому «гамбургскому сче-

ту», внесли свою немалую лепту в создание и развитие математического аппарата нелинейной динамики. Каждый из них великолепно владел математикой своего времени, глубоко интересовался проблемами естествознания. Для каждого из них в величественном здании математической науки не было ни закрытых комнат, ни темных закоулков. Каждый мощью своего интеллекта превосходил «многоголового» Никола Бурбаки.

Давид Гильберт (1862—1943) создал геометрию бесконечномерного функционального пространства, разработал прямые методы вариационного исчисления, указал на кинетическую теорию газов как на пример области физики, задачи которой решаются непосредственно с помощью интегральных уравнений и не сводимы к дифференциальным уравнениям.

Герман Вейль (1885—1955) создал аппарат теории представлений групп, все более широко используемой при описании симметрии физических систем, в том числе и систем, изучаемых нелинейной динамикой, получил выдающиеся результаты в области дифференциальной геометрии и решил важную задачу о связи спектра колебаний с формой колеблющейся области.

Джон фон Нейман (1903—1957) получил первоклассные результаты в эргодической теории, математическом обосновании квантовой механики, теории автоматов. Именно ему принадлежит идея об использовании системы «реакция с диффузией» как основы моделирования процессов, происходящих в живых организмах, в частности формообразования и самовоспроизведения.

Вместе с С. Уламом Дж. фон Нейман поставил один из первых численных экспериментов, пытаясь проверить на системе связанных нелинейных осцилляторов один из краеугольных камней статистической механики — гипотезу о равномерном распределении энергии по степеням свободы. Обнаруженный им парадокс — отсутствие тенденции к равномерному распределению — впоследствии привел к открытию солитона в уравнении Кортевега—де Фриза.

Но сколь ни значимы результаты этих последних великих универсалов, они не образуют (в отличие от результатов Пуанкаре) всюду плотного множества в арсенале средств и методов нелинейной динамики, между тем все (или почти все) ее идеи, понятия и методы так или иначе связаны с именем Пуанкаре, хотя и не всегда носят его. Ситуацию здесь довольно точно передает следующий отрывок из доклада Л. И. Мандельштама об оптических работах Ньютона, который мы приводим здесь, лишь слегка изменив текст (у Мандельштама говорится не о Пуанкаре, а о Ньюtone):

«Я чувствую своеобразное затруднение. Когда речь идет о таких открытиях, как открытия Пуанкаре, которые всем нам известны со школьной скамьи, легко очутиться — я знаю это по себе — в положении того любителя литературы, который на вопрос, как ему понравилось «Горе от ума», сказал, что в грибоедовской комедии он в сущности ничего замечательного не видит, так как она сплошь состоит из давно известных поговорок и пословиц.



Чтобы не терять перспективы, мне кажется лучше всего встать на историческую точку зрения. Нужно представить себе, хотя бы в общих чертах, состояние вопроса до Пуанкаре, затем восстановить в памяти то, что сделал Пуанкаре, и, наконец, коротко проследить ту роль, которую его работы сыграли в дальнейшем развитии науки» [1, с. 260].

Последуем совету Мандельштама.

Оценить развитие нелинейной динамики до Пуанкаре не составляет особого труда: нелинейной динамики (тогда еще — нелинейной теории колебаний) как отдельной науки, обладающей своим предметом и методом исследования, не было. В истории нелинейной динамики у Пуанкаре не было предтеч. Существовали отдельные разрозненные результаты, значимость и общность которых никому не были известны. Дифференциальные уравнения, долгое время составлявшие основу математического аппарата нелинейной динамики и поныне не утратившие свои позиции, математики, или, как было принято говорить, геометры, пытались решать путем сведения к более простым. Оценивая в «Аналитическом резюме» свои работы по дифференциальным уравнениям того периода, Пуанкаре заметил:

«Как только принципы исчисления бесконечно малых были установлены, аналитик оказался перед лицом трех проблем:

решение алгебраических уравнений;

интегрирование алгебраических функций;

интегрирование дифференциальных уравнений.

История этих трех проблем одинакова. После длинных и тщетных усилий свести эти проблемы к более простым геометры уступили наконец необходимости изучения проблем самих по себе и были вознаграждены.

Долгое время надеялись, что удастся решить все уравнения в радикалах. От этого пришлось отказаться, и сегодня алгебраические функции нам столь же хорошо известны, как и радикалы, к которым их желали привести. Точно так же и интегралы от алгебраических дифференциалов, которые долго пытались привести к логарифмическим или тригонометрическим функциям, выражаются сегодня посредством новых трансцендентностей.

Примерно то же должно было произойти и с дифференциальными уравнениями. Число уравнений, интегрируемых в квадратурах, крайне ограничено, и постольку, поскольку не решались изучать свойства интегралов самих по себе, вся эта аналитическая область оставалась всего лишь обширной terra incognita, которая казалась навсегда запретной для геометра» [3, с. 580].

О. Л. Коши, И. Л. Фукс, Ш. О. А. Брю и Ж. К. Буке, С. В. Ковалевская проникли в эту ранее не исследованную область. Они установили, что если отказаться от изучения поведения интегралов дифференциальных уравнений, как обыкновенных, так и в частных производных, в целом, т. е. при всех значениях независимой переменной, и сосредоточить усилия на исследовании локальных свойств, т. е. свойств в малом, в окрестности данной

точки, то эти свойства будут существенно отличаться в зависимости от того, будет ли выбранная точка обычной или особой.

Пуанкаре существенно дополнил и расширил результаты своих предшественников, показав, при каких условиях решение в окрестности неособой точки может быть разложено не только по степеням независимой переменной, но и по степеням начальных данных или малого параметра, каким образом эти ряды могут оставаться сходящимися при произвольных значениях независимой переменной.

Но сколь ни важны результаты, полученные Пуанкаре относительно поведения решений дифференциальных уравнений в окрестности обычной точки, свои главные усилия он сосредоточил на выяснении того, что происходит в окрестности особой точки.

Подводя итог этим своим исследованиям, Пуанкаре писал в «Аналитическом резюме»:

«Изучение интегралов дифференциальных уравнений в окрестности данной точки, какова бы ни была его польза с точки зрения числовых вычислений, может рассматриваться лишь как первый шаг. Эти разложения, которые справедливы только в очень ограниченной области... не могут рассматриваться как истинное интегрирование.

Поэтому их следует принять лишь как отправную точку в более глубоком изучении интегралов дифференциальных уравнений, где мы были бы намерены выйти из ограниченных областей, где мы были систематически подготовлены исследовать интегралы по всей плоскости.

Но это изучение может проводиться с двух разных точек зрения.

1. Можно задаться целью выразить интегралы посредством разложений, справедливых всегда и более не ограниченных какой-либо частной областью. При этом приходят к введению в науку новых трансцендентностей; и это введение необходимо, так как старые известные функции позволяют интегрировать лишь небольшое число дифференциальных уравнений.

2. Однако этот способ интегрирования, который дает нам знание свойств уравнения с точки зрения теории функций, один не достаточен, если мы желаем применять дифференциальные уравнения к вопросам механики или физики. Наши разложения не показали бы нам, по крайней мере без значительного труда, будет ли, например, функция постоянно возрастать или колебаться между определенными пределами, или она будет возрастать сверх всякого предела. Другими словами, если функцию рассматривать с точки зрения определения плоской кривой, мы ничего не узнаем об общей форме этой кривой. В некоторых приложениях все эти проблемы имеют такую же важность, как и вычисления, и они составляют новую проблему, которую нам приходится решать» [3, с. 583—584].

Мы видим, что слабое место локального рассмотрения — основной арены, на которой развертываются события, подвластные

классическому анализу, указано Пуанкаре ясно и определенно. Для перехода от рассмотрения в малом к рассмотрению в целом необходимы топологические и теоретико-групповые соображения, и Пуанкаре использует эти соображения, создав топологию и применяя группы Ли.

С волшебной легкостью он переходит от одной области математики к другой, используя технику, наиболее адекватную решаемой задаче, попутно внося в применяемый метод существенные усовершенствования и с щедростью гения разбрасывая новые идеи. Именно Пуанкаре ввел понятие универсальной обертывающей алгебры. Именно ему принадлежит так называемый метод продолжения, суть которого состоит в погружении решаемой задачи в однопараметрическое семейство задач, зависящих от вспомогательного параметра, и в выяснении разрешимости задачи в зависимости от значений этого вспомогательного параметра. Пуанкаре одним из первых стал использовать неподвижную точку и принцип сжатых отображений для доказательства существования решений нелинейных задач и построения эффективных итерационных процедур.

Логика исследования, приведшая в свое время геометров к необходимости исследования дифференциального уравнения самого по себе, без сведения к более простым проинтегрированным ранее, привела Пуанкаре к следующему шагу — к исследованию кривых, определяемых дифференциальными уравнениями. Так началась славная история качественной теории дифференциальных уравнений. Вот как рассказывает об этом сам Пуанкаре в «Аналитическом резюме»:

«Даже когда придут к тому, чтобы то, что было мною сделано для линейных уравнений, проделать для произвольного уравнения, т. е. найти разложения интегралов, справедливые во всей плоскости, это еще не будет основанием для отказа от результатов, которые можно получить другими методами, так как может случиться, что эти методы откроют нам частности, которые разложения не представляли бы нам сразу с очевидностью. Это соображение побудило меня встать на новую точку зрения, и я не мог бы найти лучшего способа дать о ней представление, чем воспроизвести то, что писал в момент, когда начинал эти исследования:

„Итак, необходимо изучать функции, определенные дифференциальными уравнениями, сами по себе, не пытаюсь сводить их к более простым функциям, так же, как это сделано для алгебраических функций, которые пытались сводить к радикалам и которые изучают теперь прямо, и так же, как это сделано для интегралов от алгебраических дифференциалов, которые долго пытались выразить в конечных терминах“.

Исследовать, каковы свойства дифференциальных уравнений, является, таким образом, вопросом, имеющим самый большой интерес. По этому пути уже сделали первый шаг, изучив функцию в окрестности одной точки плоскости. Сегодня речь идет о том, чтобы идти дальше и изучать эту функцию на всем протяжении плос-

кости. В этом исследовании отправной точкой нам будет служить, разумеется, то, что уже известно об изучаемой функции в некоторой области плоскости.

Полное изучение функции состоит из двух частей: 1) качественной (так сказать), или геометрического изучения кривой, определяемой функцией; 2) количественной, или вычисления значений функции.

...Именно с качественной стороны должна начинаться теория всякой функции, и вот почему в первую очередь возникает следующая задача: построить кривые, определяемые дифференциальными уравнениями.

...Это качественное исследование и само по себе будет иметь первостепенный интерес. Различные и чрезвычайно важные вопросы анализа и механики могут быть сведены к нему.

...Таково широкое поле для открытий, которое лежит перед геометрами. Я не притязал на то, чтобы пройти его все, но хотел, по крайней мере, перейти его границы и ограничил себя одним очень частным случаем, тем, который естественно представляется с самого начала, т. е. изучением дифференциальных уравнений первого порядка первой степени» [3, с. 595—597].

Обобщив и специализировав результаты Брио и Буке, а также свои собственные, Пуанкаре обнаружил существование особых точек четырех видов (седел, узлов, фокусов и центров — все названия принадлежат ему), изучил их расположение на плоскости, ввел понятия цикла без контакта и предельного цикла. Тем самым им было выковано оружие, которое через много лет было обнаружено в математическом арсенале учеником Л. И. Мандельштама А. А. Андроновым и стало математическим образом, адекватным автоколебаниям.

Обнаружение сложных — хаотических и стохастических — режимов в детерминированной динамической системе также связано с именем Пуанкаре. Занимаясь изучением так называемой ограниченной задачи трех тел, он открыл существование особых фазовых кривых, отвечающих неустойчивым движениям. Именно они являются тем механизмом, который хаотизирует, запутывает траектории динамической системы. В знаменитых «Новых методах небесной механики» Пуанкаре так описывал гомоклиническую структуру:

«Если попытаться представить себе фигуру, образованную этими двумя кривыми [устойчивая и неустойчивая инвариантные кривые, проходящие через седловую особую точку] и их бесчисленными пересечениями, каждое из которых соответствует двоякоасимптотическому решению, то эти пересечения образуют нечто вроде решетки, ткани, сети с бесконечно тесными петлями; но ни одна из двух кривых никогда не должна пересечь самое себя, но она должна навиваться на самое себя очень сложным образом, чтобы пересечь бесконечно много раз все петли сети.

Поражаешься сложности этой фигуры, которую я даже не пытаюсь изобразить. Ничто не является более подходящим, чтобы

дать нам представление о сложности задачи трех тел и вообще всех задач динамики, в которых нет однозначного интеграла и в которых ряды Болина расходятся» [4, с. 339].

Открытие сложных хаотических режимов позволило не только понять природу неинтегрируемости задач динамики, но и постичь ограниченность так называемого ньютоновского детерминизма, по-новому взглянуть на природу случайного. Экспоненциальное разбегание первоначально близких траекторий, вынужденных оставаться в ограниченной части фазового пространства, приводит к их перепутыванию, т. е. в конечном счете к хаотизации. В одной из своих работ по философии науки («Наука и метод») Пуанкаре говорит о природе случайного так:

«...совершенно ничтожная причина, ускользающая от нас по своей малости, вызывает значительное действие, которого мы не можем предусмотреть, и тогда мы говорим, что это явление представляет собой результат случая.

...Иногда небольшая разница в первоначальном состоянии вызывает большое различие в окончательном явлении. Небольшая погрешность в первом вызвала бы огромную ошибку в последнем. Предсказание становится невозможным, мы имеем перед собой явление случайное» (цит. по [5, с. 323]).

Пуанкаре принадлежит в числе прочих еще одно важное открытие: непрерывный (более того, дифференцируемый) поток в фазовом пространстве, например в окрестности периодической траектории, можно изучать с помощью дискретного отображения, индуцируемого этим потоком на версальном сечении (так называемом сечении Пуанкаре). Тем самым Пуанкаре одним из первых восстановил справедливость, уравнив в правах дискретное отображение сечения в себя и непрерывное (гладкое) отображение фазового пространства в себя. Развернувшиеся впоследствии острые споры о том, что первично — дискретное или непрерывное, несколько напоминают споры «остро-» и «тупоконечников»: в природе непрерывность встречается наряду с дискретностью, и при выборе средства для решения той или иной задачи следует скорее заботиться об его адекватности, нежели отдавать предпочтение определенному подходу только потому, что он дискретен или непрерывен.

С именем Пуанкаре связан и метод нормальных форм, позволяющий избавляться от «лишних» (нерезонансных) членов в правой части уравнений с помощью формальных обратимых замен переменных (существующие теоремы о сходимости рядов, задающих замены, в приложениях, как правило, не используются). Нормальные формы позволяют не только упрощать решаемые уравнения, но и строить разумные базовые модели. Обычно модель выбирается с таким расчетом, чтобы она воспроизводила с большей или меньшей точностью некое множество режимов. Однако при построении модели обычно делаются многочисленные неконтролируемые предположения, не позволяющие в конце анализа однозначно ответить на вопрос, какое отношение к исход-

ным физическим моделям имеет выбранная базовая модель. Приведение к нормальной форме означает разбиение множества исходных моделей на классы эквивалентности с последующим выбором по одному представителю от каждого класса. При таком подходе ничто не мешает после изучения режимов, допускаемых базовой моделью, вернуться к исходной модели без какой бы то ни было потери информации.

В несметных сокровищах наследия Пуанкаре можно найти и многие другие важные понятия и теории, созданные впоследствии на правах «первооткрывателей» теми, кто либо никогда не читал трудов Пуанкаре, либо делал это недостаточно внимательно. В частности, из его работ нетрудно извлечь достаточно подробно проработанные контуры теории бифуркаций, или, как предпочитал называть их сам Пуанкаре, «смен устойчивости».

«К сожалению, — замечает В. И. Арнольд, — бесхитростные тексты Пуанкаре трудны для математиков, воспитанных на теории множеств. [Пуанкаре сказал бы „Петя вымыл руки“ там, где современный математик напишет просто: „Существует  $t < 0$ , такое, что образ точки  $t_1$  при естественном отображении  $t \rightarrow (Петя(t))$  принадлежит множеству грязноруких, и такое  $t_2 \in (t_1, 0]$ , что Петя ( $t_2$ ) принадлежит дополнению вышеупомянутого множества“.] Видимо, поэтому многие его идеи остались незамеченными ближайшими к нему поколениями.

Исключение составляют, пожалуй, лишь Д. Д. Биркгоф и его ученики Г. М. Морс и Х. Уитни. Р. Том в докладе о работах Смейла на Математическом конгрессе в 1966 г. в Москве назвал его чуть ли не единственным математиком, прочитавшим Пуанкаре и Биркгофа» [6, с. 232—233].

Что же касается «наивных» определений Пуанкаре, то попытки обобщения их, как правило, не приводят к новым объектам.

И все же, несмотря на известное высокомерие потомков, труды Пуанкаре не встали мертвым грузом на верхних полках библиотек. Как показала, в частности, конференция по математическому наследию Анри Пуанкаре, состоявшаяся с 7 по 10 апреля 1980 г. в университете штата Индиана, идеи Пуанкаре питают современную математику в гораздо большей мере, чем это может показаться непросвещенному узкому специалисту.

Жизнь другого главного действующего лица нашего повествования — Леонида Исааковича Мандельштама (1879—1944), по словам его ближайшего сотрудника Н. Д. Папалекси, «не отличалась внешним блеском. Он никогда не добивался внешних почестей, не стремился играть какой-либо роли, ему совершенно чужды были честолюбие и славолубие. Но тем полнее и богаче была его внутренняя жизнь. Это была прекрасная жизнь истинного ученого и глубокого мыслителя, искателя научной истины, человека исключительного душевного благородства» [1, с. 5].

Влияние научных идей Л. И. Мандельштама на современную физику в целом, и в особенности на нелинейную динамику, неоспоримо. Отчасти оно прослежено в обзоре [7], посвященном 100-летию

выдающегося ученого. Полученные им результаты по праву считаются классическими. Они вошли в учебники, стали достоянием истории науки и, что гораздо важнее, предметом пристального внимания со стороны тех, кто принимает непосредственное участие в создании нелинейной динамики на ее современном этапе.

Не менее важной, чем собственные научные результаты, была выдвинутая Л. И. Мандельштамом идея выработки нелинейного физического мышления — «создания наглядных физических представлений, имеющих в своей основе адекватные нелинейным физическим объектам математические представления и понятия» [1, с. 107].

Непревзойденный знаток и ценитель линейной теории, Л. И. Мандельштам с присущей ему тонкой физической интуицией и особой, чисто «мандельштамовской» ясностью мышления раньше и лучше других осознал ограниченность линейной теории с ее принципом суперпозиции, теоремами существования и единственности решений. Менее всего склонный принимать новое только потому, что это новое, бережно, чтобы не сказать консервативно, относившийся к старому (в данном случае — к линейной теории), Л. И. Мандельштам видел, сколь широк круг физических явлений, не допускающих описания в рамках линейной идеализации, сколь ненадежной становится «линейная психология», способная скорее вводить в заблуждение, чем служить надежной путеводной нитью исследователю.

Высказанная Л. И. Мандельштамом идея не осталась благим пожеланием. Она была воплощена в плоть и кровь его учениками. Выступая 22 декабря 1944 г. на совместном заседании Московского государственного университета им. М. В. Ломоносова и Академии наук СССР, посвященном памяти Л. И. Мандельштама, А. А. Андронов сказал: «Я перечислю некоторые нелинейные понятия, либо получившие точный физический и математический смысл, либо впервые выдвинутые в этот с 1927 г. период времени.

Я начну с фазового пространства, которое... перестало быть только математической абстракцией и приобрело высокую степень физической наглядности не только потому, что физики с ним свыклись, но и потому, что оказалось возможным приблизить его к нашим органам чувств, наблюдая систематически фазовые траектории на экране осциллографа... Если говорить об автономных системах, то такие физические понятия, как автоколебания, мягкое и жесткое возбуждение автоколебаний, затягивание и т. д., получили теперь твердую математическую основу в виде предельных циклов, теории бифуркаций, областей устойчивости в большом и т. д. Если говорить о неавтономных системах, то такие физические понятия, как феррорезонанс, захватывание разных видов, получили математическую основу в теории периодических решений и их бифуркаций, а ряд других физических понятий, например резонанс второго рода, асинхронное возбуждение и т. д., были вновь выдвинуты, отправляясь от математической теории.

Не все достижения этих лет в направлении выработки нелиней-

ного мышления принадлежат Л. И. Мандельштаму или лицам, так или иначе с ним связанным. Но именно Л. И. Мандельштам вызвал к жизни это новое, опирающееся, с одной стороны, на настоящую математику, с другой стороны, на тонкий радиофизический эксперимент научное направление в теории нелинейных колебаний» [1, с. 120].

С тех пор набор «первичных» нелинейных физических понятий, опирающихся на прочную математическую основу, существенно пополнился. Солитон, различные типы отдельных буфуркаций, цепочки бифуркаций, катастрофы, перемежаемость, диссипативные структуры и т. д. — таков далеко не исчерпывающий их перечень.

Естественно спросить: как мог один человек подняться до столь высокого, чтобы не сказать пророческого, предвидения? Ответ на этот вопрос дал И. Е. Тамм:

«Одна из основных особенностей дарования Л. И., сообщавшая ему особую силу, заключалась, как мне кажется, в редчайшем сочетании в одном человеке ума конкретного, геометрически пластичного и ума абстрактного, логически аналитического. С одной стороны, способность единым взглядом охватить сложное многообразие разнородных явлений, с предельной четкостью усмотреть в них черты сходства и различия и воссоздать все существенное в простой и наглядной модели, с другой стороны, острый интерес к конкретной индивидуальности физического явления, порождавший те чувства непосредственного наслаждения, которые испытывал Л. И. при экспериментировании. В этом истоки и необычайного искусства Л. И. в постановке экспериментов, и его исключительно плодотворной деятельности в области технической физики. И вот с этими свойствами ума «широкого» и «английского», по терминологии Дюгема, в Л. И. сочеталась необычайная сила и тонкость абстрактной логической мысли и необычайная глубина анализа принципиальных основ физической теории, восходящего к основным категориям мышления» [1, с. 131—132].

#### ЛИТЕРАТУРА

1. Академик Л. И. Мандельштам: К 100-летию со дня рождения. М.: Наука, 1979. 312 с.
2. Бурбаки Н. Архитектура математики // Математическое просвещение. М.: Физматгиз, 1960. Вып. 5. С. 99—112.
3. Пуанкаре А. Избранные труды. М.: Наука, 1974. Т. 3. 771 с.
4. Пуанкаре А. Избранные труды. М.: Наука, 1972. Т. 2. 560 с.
5. Пуанкаре А. О науке. М.: Наука, 1983. 560 с.
6. Арнольд В. И. Теория катастроф // Современные проблемы математики. Фундаментальные направления. М.: ВИНТИ, 1986. Т. 5. 235 с.
7. Гапонов-Грегов А. В., Рабинович М. И. Л. И. Мандельштам и современная теория нелинейных колебаний и волн // УФН. 1979. Т. 128, № 4. С. 579—624.



## ОБ АТТРАКТОРАХ

В. С. АФРАЙМОВИЧ

1. Вот уже 60 лет [1] идет работа по созданию адекватного математического образа автоколебательного движения реальной системы. Отметим два важных этапа этого процесса.

А. А. Андроновым и его коллегами и учениками была разработана качественная теория систем дифференциальных уравнений на плоскости и сфере [2, 3]. На огромном числе физических примеров (в частности, из радиотехники) было убедительно показано, что предельный цикл Пуанкаре является вполне адекватным образом автоколебательного периодического движения, а бифуркации предельных циклов описывают, объясняют и предсказывают многие (нелинейные) явления, связанные с периодическими автоколебаниями.

Второй этап начался 15—20 лет назад, когда было осознано, что в сугубо детерминированной системе (в которой по начальному состоянию однозначно определяются все последующие) могут осуществляться нерегулярные, непериодические — стохастические — автоколебания, во многих отношениях неотличимые от реализаций случайного процесса, когда появилось огромное количество конкретных систем, генерирующих стохастические сигналы в широкой области параметров. Начался, но не кончился. Попытки математиков выработать понятие, достаточно адекватно отражающее все интуитивно ясные и наблюдаемые в численных и лабораторных экспериментах свойства «динамической стохастичности», пока к полному успеху не привели. Процесс создания такого понятия продолжается, и цель настоящей работы — представить, хотя бы в общих чертах, мгновенный снимок этого процесса.

2. Математической моделью детерминированной физической системы служит динамическая система (ДС). Предполагается, что существует множество начальных условий  $M$  с заданным на нем расстоянием — фазовое пространство ДС — и что для любого (положительного) значения  $t$  непрерывного или дискретного «времени» имеется непрерывное (дифференцируемое) отображение  $f^t$ , ставящее в соответствие каждой начальной точке  $x$  из  $M$  некоторую другую (или иногда ту же самую) точку  $f^t x$  из  $M$ . При этом для любых значений времени  $t_1$  и  $t_2$  выполнено  $f^{t_1}(f^{t_2}x) = f^{t_2}(f^{t_1}x) = f^{t_1+t_2}x$  для каждого  $x$  из  $M$ . Множество точек  $\{f^t x\}_{t \geq 0}$  называется положительной полутраекторией с началом в  $x$ , точка  $f^t x$  — изображающей точкой на полутраектории, а пара  $(M, \{f^t\})$  — (гладкой) динамической системой.

«Закон»  $f^t$ , по которому начальной точке  $x$  ставится в соответствие точка в фазовом пространстве через время  $t$ , в случае непрерывного времени определяется обычно с помощью системы дифференциальных уравнений  $\dot{x} = X(x)$ . При этом  $f x_0 = \varphi(t, x_0)$  ( $\varphi$  — ре-

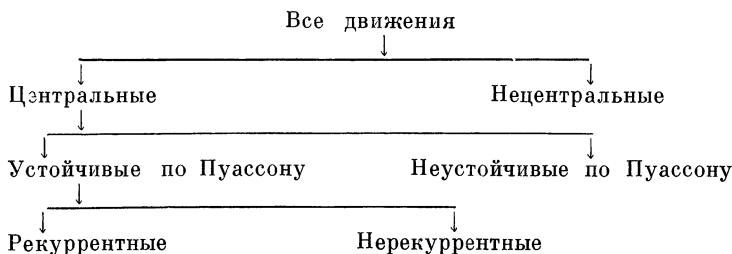
шение этой системы с начальным условием  $x_0$ ), полутраектория  $\{f^t x_0\}_{t \geq 0}$  — кривая в фазовом пространстве  $M$ , а  $M$  может, например, быть либо евклидовым пространством  $\mathbb{R}^n$ , либо (в бесконечномерной ситуации) банаховым.

«Логистическое уравнение» (или «отображение параболы»)  $x_{n+1} = ax_n(1 - x_n)$ , впервые выведенное П. Ф. Верхульстом в 1845 г. при описании роста популяций (см., например, [4]), определяет динамическую систему с дискретным временем  $n \in \mathbb{Z}_+$  (здесь  $M = \mathbb{R}^1$ , полутраектория — последовательность точек на прямой).

Движению реальной системы соответствует движение изображающей точки по полутраектории, а свойства движения отражаются в геометрических и дифференциальных характеристиках полутраекторий.

С этой точки зрения обсудим имеющееся для физических систем с накачкой и диссипацией разделение движений на переходные процессы и установившиеся движения. Основное свойство установившегося движения — в том, что оно «забывает» о начальном состоянии, не зависит от него (в соответствующей области изменения начальных условий). А отсюда следует повторяемость: если какой-либо фрагмент движения проявился, скажем, один раз и больше не появляется, то ясно, что он относится к переходному процессу.

Свойство повторяемости математически отражено в понятиях топологической динамики<sup>1</sup>. Ими активно пользовался (и пропагандировал радиоп физикам тех лет [1]) А. А. Андронов, а сейчас они не очень известны, поэтому приведем часть составленной им таблицы и расшифруем ее:



Изображающая точка, лежащая на устойчивой по Пуассону траектории, возвращается в сколь угодно малую свою окрестность бесконечное число раз (т. е. начальное состояние повторяется со сколь угодно большой точностью), но на значения времени, через которые происходят эти повторения (так называемые времена возврата Пуанкаре), никакие ограничения не накладываются (они могут, например, стремиться к бесконечности и т. п.). Рекур-

<sup>1</sup> Созданная в основном Биркгофом часть теории динамических систем, в которой используется лишь непрерывность отображений сдвига, а не их дифференцируемость.

рентная полутраектория устойчива по Пуассону и, кроме того, обладает тем свойством, что конечный ее кусок аппроксимирует ее всю (с любой степенью точности). Множество предельных точек рекуррентной траектории называется минимальным множеством и состоит только из рекуррентных траекторий<sup>2</sup>. Таким образом, если установившемуся движению отвечает, скажем, некоторая рекуррентная траектория (из какого-то минимального множества), то с абсолютно той же степенью адекватности ему отвечает любая рекуррентная траектория этого минимального множества — по словам А. А. Андропова, «сгусток, состоящий из континуума движений — братьев».

Предположим теперь, что «режиму работы» физической системы отвечает (незамкнутая) устойчивая по Пуассону полутраектория  $L$ . Поскольку начальное состояние с абсолютной точностью неизвестно, то в соответствие этому же режиму мы должны поставить и множество траекторий, предельных к  $L$  — это множество континуально и состоит из устойчивых по Пуассону траекторий, всюду плотных в замыкании  $L$ . Следовательно, если полагать верным, что установившееся движение одно и то же для целого множества (пусть очень близких друг к другу) начальных состояний, то с необходимостью приходится считать математическим образом этого движения целое множество траекторий в фазовом пространстве ДС, являющейся математической моделью процесса. К этому множеству стремятся (притягиваются) полутраектории из «шарика» начальных состояний, поэтому его естественно назвать аттрактором.

Итак, установившемуся движению физической системы отвечает движение изображающей точки по аттрактору (или в малой его окрестности), переходному процессу — движение изображающей точки по полутраектории вне малой окрестности аттрактора.

3. Если определить аттрактор просто как множество предельных точек всех полутраекторий, начинающихся в произвольном фиксированном шарике в фазовом пространстве, то получится очень широкий класс инвариантных множеств, о котором трудно получить какую-либо содержательную информацию. Во всяком случае, этот вопрос не исследован. Начиная с работ С. Смейла (см., например, [5]), предполагается, что область притяжения аттрактора (т. е. множество начальных состояний, через которые проходят полутраектории, стремящиеся к аттрактору) содержит какую-либо окрестность аттрактора: существует такая окрестность  $U$  аттрактора  $A$ , что любая полутраектория с начальной точкой в  $U$  стремится к  $A$  (см., например, [6]).

Иногда это свойство формулируют несколько по-другому: область  $U$  называется поглощающей, если  $f^t U \subset U$ ,  $t > 0$ ,  $\bar{U}$  — замыкание  $U$ ; аттрактор — это пересечение всех сдвигов

<sup>2</sup> Множество же предельных точек устойчивых по Пуассону траекторий — множество центральных движений — может содержать не только устойчивые по Пуассону полутраектории, однако последние всюду плотны в нем.

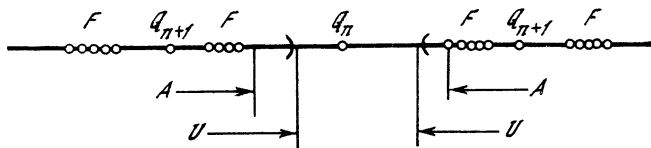


Рис. 1. Полуаттрактор Фейгенбаума  $F$

$U$  — поглощающая область ( $Q_n$ ,  $Q_{n+1}$  — периодические точки периодов  $2^n$  и  $2^{n+1}$ );  $A$  — аттрактор, содержащий  $F$

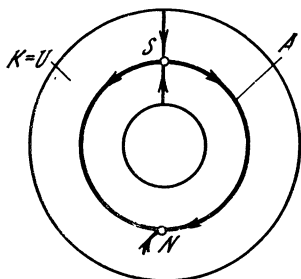


Рис. 2. Аттрактор в поглощающем кольце

поглощающей области  $A = \bigcap_{t>0} f^t U$  (см., например, [7]).

Среди достоинств такого определения отметим следующие: 1) все регулярные аттракторы (устойчивые положения равновесия, устойчивые предельные циклы, притягивающие торы) ему удовлетворяют; 2) если в гладкой динамической системе, непрерывно зависящей от параметра, скажем,  $\mu$ , имеется поглощающая область  $U$  при  $\mu = \mu_*$ , то при значениях  $\mu$ , близких к  $\mu_*$ , тоже имеется близкая к  $U$  поглощающая область (либо сама  $U$  остается поглощающей), поэтому можно говорить о семействе аттракторов в некотором смысле «непрерывно» зависящих от  $\mu$ .

С физической точки зрения кажется естественным требовать, чтобы аттрактор (как объект) не зависел от изменения параметра  $\mu$  (конечно, пока не наступил кризис [7—9]). Поэтому если не подпадающие под это определение аттракторы являются собственными подмножествами объемлющих их аттракторов (построенных с помощью поглощающих областей), то образами установившихся движений следует, по-видимому, считать последние. Именно такой случай имеет место для «полуаттрактора» Фейгенбаума  $F$  отображения параболы в критический момент  $a \approx 3,57$  (последние результаты его исследований см. в [10]). Полуаттрактор  $F$  представляет собой некоторое подобное канторову множество, в любой окрестности которого имеется бесконечное множество неустойчивых периодических точек. Легко сконструировать поглощающую область  $U$ , аттрактором  $A$  в которой будет набор отрезков, содержащих полуаттрактор  $F$ , счетное множество неустойчивых периодических точек и континуальное множество стремящихся к  $F$  полутракторий (рис. 1).

Основной недостаток этого определения в том, что аттрактору могут принадлежать «лишние» траектории, не обладающие свойствами повторяемости. В примере выше — это стремящиеся к  $F$

полутраектории с начальными точками в окрестностях периодических точек. Еще нагляднее это видно на примере рис. 2: поглощающая область — кольцо  $K$ , в котором имеются две неподвижные точки — седло  $S$  и узел  $N$ ; все полутраектории, кроме принадлежащих устойчивой сепаратрисе седла  $S$ , стремятся к  $N$ ; аттрактором же (по этому определению) является неустойчивая сепаратриса  $S$ , объединенная с  $N$  (в то время как естественно считать им узел  $N$ ).

Обобщения этого определения, основанные на понятиях эргодической теории, изложены в [7, 11, 12]. Среди них отметим вероятно предельные множества Милнора и статистически предельные множества Ильяшенко [7].

4. Противоречие, содержащееся в определении аттрактора, основанном на понятии поглощающей области, и в то же время в требовании какой-либо повторяемости для всех траекторий в аттракторе, можно снять (или по крайней мере не ощущать от него неудобства), если от аттракторов перейти к семействам аттракторов.

Итак, пусть гладкая ДС  $(M, \{f_\mu^t\})$  непрерывно зависит от параметра  $\mu$ . Пусть  $\{U_\mu\}$  — непрерывное семейство поглощающих областей (проще всего считать, что границы  $U_\mu$  — гладкие поверхности, непрерывно зависящие от  $\mu$ ), а  $\{A_\mu\}$  — семейство отвечающих им аттракторов. Будем называть  $\{A_\mu\}$  неприводимым при  $\mu \in [\mu_1, \mu_2]$  семейством, если не существует непрерывного семейства поглощающих областей  $U_\mu$ , для которого отвечающие ему аттракторы  $A_\mu$  были бы при каждом  $\mu \in [\mu_1, \mu_2]$  собственными подмножествами  $A_\mu$ . Другими словами, если в  $A_\mu$  при каком-то значении  $\mu$  и есть свой собственный «податтрактор»  $\tilde{A}_\mu$ , то при каком-то значении  $\mu_* \in [\mu_1, \mu_2]$  он перестает быть таковым. При изменении  $\mu$ , скажем, устойчивые по Пуассону траектории могут возникать то в одних, то в других «частях»  $A_\mu$ , могут происходить бифуркации, сопровождающиеся возникновением и исчезновением собственных «податтракторов» в  $A_\mu$ , но это внутренние бифуркации (см. [7—9]), они не сказываются на существовании  $A_\mu$ , не мешают воспринимать семейство  $\{A_\mu\}$  как нечто целостное.

5. Тем не менее в существующие определения аттрактора все же вводят какое-либо условие повторяемости [6, 13—15]. Одно из самых общих математических понятий, отражающих его, — это понятие цепной рекуррентности, основанное на понятии  $\varepsilon$ -траектории, или (в зарубежных источниках) псевдоорбиты Боуэна [16].

**О п р е д е л е н и е** [13, 16—18]. *Будем говорить, что точка  $y$  сцеплена с точкой  $x$  ( $y \Rightarrow x$ ), если существует число  $t(x, y) > 0$  и для каждого  $\varepsilon > 0$  существует натуральное число  $N = N(\varepsilon)$ , набор точек («цепь»)  $x_0 = y, x_1, \dots, x_N = x$  и набор чисел (времена)  $t_1, \dots, t_N, t_i \geq t(x, y)$ , такие, что  $\text{dist}(x_i, f^{t_i}x_{i-1}) < \varepsilon$  для  $i = 1, \dots, N$  (рис. 3). Если время дискретно, то  $t(x, y) \geq 1$ .*

Последовательность отрезков траекторий ДС  $(M, \{f^t\})$  с концами в точках  $x_{i-1}, x_i$  и называется  $\varepsilon$ -траекторией. В случае дискретного времени любая последовательность точек  $\{x_k\}, k = 0, 1, \dots,$

удовлетворяющая неравенству  $\text{dist}(x_i, fx_{i-1}) < \varepsilon$ , является  $\varepsilon$ -траекторией. Заметим, что при вычислениях на ЭВМ мы всегда имеем дело с  $\varepsilon$ -траекторией (а не с истинной траекторией системы).

Если  $x \Rightarrow x$ , то точка  $x$  называется цепнорекуррентной. Устойчивая по Пуассону полутраектория состоит, конечно, из цепнорекуррентных точек. Однако цепнорекуррентная точка может быть окружена точками, никогда не возвращающимися в ее окрестность (см. рис. 4, где окрестность цепнорекуррентной точки  $P$

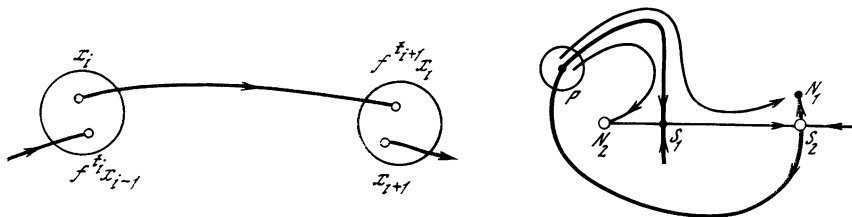


Рис. 3. Отрезок  $\varepsilon$ -траектории

Точки  $x_i, x_{i+1}, f^t x_{i-1}, f^{t+1} x_i$  лежат в шарах радиуса  $\varepsilon$

Рис. 4. Фазовая картинка системы на плоскости, имеющей цепнорекуррентную (но блуждающую) точку  $P$

состоит лишь из точек сепаратрисы седла  $S_1$ , а также из точек, лежащих в областях притяжения узлов  $N_1$  и  $N_2$ ): цепная рекуррентность означает лишь, что приближенное движение может обладать свойством повторяемости.

Пусть  $x$  — цепнорекуррентная точка. Ее класс  $[x]$  называется множеством всех точек  $y$ , таких, что  $y \Rightarrow x$  и  $x \Rightarrow y$  (т. е. соединяемых с  $x$  кусками  $\varepsilon$ -траекторий для произвольно малых  $\varepsilon$ ). Нетрудно понять, что класс любой цепнорекуррентной точки состоит из целых траекторий, т. е. является инвариантным множеством. Класс  $[x]$  — аттрактор (по Конли—Рюэлю), если из  $x \Rightarrow y$  следует  $y \Rightarrow x$  для любого  $y$ . Очевидно, аттрактором по Конли—Рюэлю является любой аттрактор, определяемый с помощью поглощающей области и состоящий лишь из точек, предельных для устойчивых по Пуассону. Полуаттрактор Фейгенбаума является классом любой своей точки и аттрактором по Конли—Рюэлю. В поглощающей области рис. 2 аттрактором является узел  $N$ .

Нетрудно показать, что если  $U$  — поглощающая область и  $A$  — определяемый ею аттрактор, то в  $U$  существует хотя бы один аттрактор по Конли—Рюэлю  $\tilde{A} \subset A$ . Но таких аттракторов может быть любое (и даже бесконечное) количество (см. ниже). Поэтому если  $\{U_\mu\}$  — семейство поглощающих областей и  $\{A_\mu\}$  — отвечающее ему семейство аттракторов, то, во-первых, не ясно, как определять «соответствующее» семейство  $\{\tilde{A}_\mu\}$ , и, во-вторых, если это даже и удастся (в конкретном случае) сделать, то при изменении  $\mu$  семейство  $\{\tilde{A}_\mu\}$  может настолько резко (и непредска-

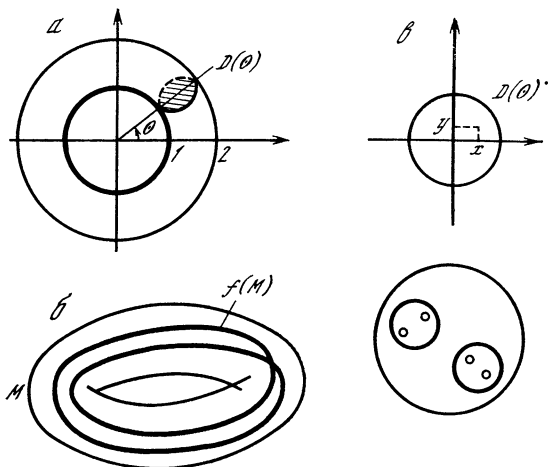


Рис. 5. Построение аттрактора Смейла — Уильямса

$a$  — заполненный тор в  $\mathbb{R}^3$  с введенной системой координат  $\{(\theta, x, y)\}$ ;  $b$  —  $M$  и  $f(M)$ ;  $c$  — следы пересечения  $f(M)$  и  $f(f(M))$  с  $D(\theta)$

зуюмо) меняться (в отличие от  $\{A_\mu\}$ ), что говорить о нем вряд ли имеет смысл. «Плохая» зависимость от параметра — главный недостаток аттракторов по Конли—Рюэлю.

6. Выше мы обсуждали понятие аттрактора вообще. Перейдем к обсуждению понятия странного аттрактора — математического образа стохастического автоколебания. Д. Рюэль, которому (в соавторстве с Ф. Такенсом [14, 19]) принадлежит термин «странный аттрактор», характеристическое свойство стохастических автоколебаний определил как «чувствительную зависимость от начальных условий» [14]. Следовательно, траектории в аттракторе, отвечающем стохастическому автоколебанию, должны быть в каком-то смысле неустойчивыми.

Самая «равномерная» и сохраняющаяся при малых возмущениях неустойчивость траекторий отражена в понятии гиперболичности, введенном Д. В. Аносовым для  $y$ -систем [20] и распространенном С. Смейлом на более широкие классы ДС и инвариантных множеств в них [5], в частности на аттракторы. Точного определения давать здесь не будем, а поясним его на известном примере гиперболического аттрактора — так называемого аттрактора Смейла—Уильямса [5, 12, 14].

В полнотории  $M$ , для которого введены две «амплитудные» координаты  $x, y$  и одна циклическая  $\theta$  (рис. 5), определим ДС с помощью отображения  $f$ :  $\theta_{n+1} = 2\theta_n$ ,  $x_{n+1} = a_1 \cos \theta_n + a_2 x_n$ ,  $y_{n+1} = a_1 \sin \theta_n + a_2 y_n$ , где  $0 < a_2 < a_1 < 1/2$ . Видим, что: 1)  $M$  — поглощающая область:  $f(M) \subset M$ ; 2) диск  $D(\theta)$  при отображении  $f$  равномерно сжимается в диск  $f(D(\theta))$  с коэффициентом сжатия  $a_2$ ; 3) по координате  $\theta$  при каждой итерации происходит растяжение с коэффициентом 2; 4) в диск  $D(\theta)$  при отображении  $f$  переходят

два диска:  $D(\theta/2)$  и  $D(\theta/2 + \pi)$ , и их образы не пересекаются. Отсюда легко вывести, что  $f(M)$  — это заполненный тор, дважды «закрученный» вдоль средней линии ( $x_1 = x_2 = 0$ )  $M$ .

При второй итерации получаем в  $f(M)$  заполненный тор  $f(f(M))$ , дважды «проходящий» вдоль средней линии  $f(M)$  (и четырежды — вдоль средней линии  $M$ ). Продолжая итерирование, в пределе получим аттрактор Смейла—Уильямса, который: 1) состоит из континуального множества «неустойчивых слоев» — незамкнутых кривых, каждая из которых плотна в  $A$  и вдоль которых близкие точки разбегаются с экспоненциальной скоростью; 2) локально (в окрестности каждой точки) устроен одинаково — как произведение канторова множества на отрезок; 3) имеет окрестность, расслоенную на «устойчивые слои» — диски  $D(\theta)$ , по которым точки экспоненциально быстро притягиваются к  $A$ . Оказывается, все перечисленные свойства являются грубыми: они сохраняются при  $C^1$ -малых возмущениях (лишь устойчивые и неустойчивые слои немного «искривятся»). Более того, порожденная  $f$  ДС является грубой в смысле Андронова—Понтрягина: ее  $C^1$ -малое возмущение топологически ей эквивалентно — сводится к ней непрерывной (близкой к тождественной) заменой переменных.

Свойствами, аналогичными 1—3, обладают все гиперболические аттракторы.

7. Но для чувствительной зависимости от начальных условий нужна лишь неустойчивость траекторий в аттракторе и вовсе не обязательно, чтобы эта неустойчивость была одинаковой для всех траекторий — скажем, число неустойчивых направлений для разных траекторий может быть разным, лишь бы они были. Именно таким является знаменитый аттрактор Лоренца—аттрактор, содержащий седловое положение равновесия  $O$  с одномерным неустойчивым и двумерным устойчивым многообразиями и континуальное множество траекторий, лежащих в пересечении двумерных устойчивого и неустойчивого многообразий. Для него свойства, аналогичные свойствам 1 и 3 гиперболического аттрактора, имеют место и являются грубыми, но устроен он геометрически в разных точках по-разному.

В окрестностях точек, не лежащих на неустойчивом многообразии седла  $O$ , аттрактор Лоренца гомеоморфен произведению канторова множества на двумерный диск, вдоль же неустойчивой сепаратрисы это континуальное множество двумерных дисков (кусков неустойчивых слоев) склеено по ней и образует нечто вроде «книги». Об аттракторе Лоренца написано очень много (см., например, [14, 15, 21, 22]). Отметим лишь, что осознание того факта, что в аттракторе Лоренца нет устойчивых траекторий и они не могут появиться при малых возмущениях, и в связи с этим осознание возможности существования динамической стохастичности (или динамического хаоса) и привело к новому этапу развития нелинейной физики, о котором говорилось в п. 1.

8. Если бы во всех ДС, являющихся математическими моделями генерирующих стохастические колебания физических систем,



соответствующие аттракторы имели сплошь неустойчивые траектории, то можно было бы говорить о полной математической корректности основ теории динамического хаоса, оставалось бы изучать свойства этих аттракторов, классифицировать их и т. п. К сожалению, таких аттракторов известно немного: 1) гиперболические, например Смейла—Уильямса [5, 12], Плыкина [23]; 2) квазигиперболические, например Лоренца [4], Лози [24], Белых [25]. Может быть, некоторые упущены, но их число не очень велико. В то же время подавляющее большинство (для сравнения см., например, [26]) автоколебательных стохастических процессов описывается ДС, аттракторы для которых, по-видимому, содержат устойчивые предельные циклы при всех либо почти всех значениях управляющих параметров.

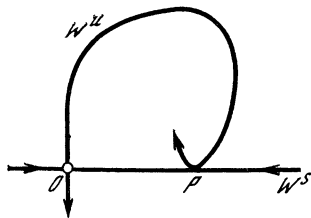


Рис. 6. Гомоклиническая точка  $P$ , принадлежащая касанию устойчивого  $W^s$  и неустойчивого  $W^u$  многообразий седла  $O$

Последнее утверждение основано на следующих математических представлениях и результатах. Странные аттракторы содержат бесконечное множество седловых предельных циклов. Неустойчивая сепаратриса каждого из них также обязана принадлежать аттрактору (иначе вдоль нее произошла бы «утечка» траекторий и аттрактор не был бы таковым). Бифуркации, связанные с касанием неустойчивых и устойчивых сепаратрис предельных циклов, и приводят к появлению устойчивых движений.

Поясним это утверждение на примере отображения Пуанкаре с полутора степенями свободы. Пусть это отображение записывается в виде

$$x_{n+1} = \lambda x_n + f(x_n, y_n, \mu), \quad y_{n+1} = \gamma y_n + g(x_n, y_n, \mu), \quad (1)$$

где  $0 < \lambda < 1 < \gamma$ ,  $f(0, 0, \mu) = g(0, 0, \mu) = 0$ , и седло  $O(0, 0)$  — диссипативная неподвижная точка, т. е.  $\lambda\gamma < 1$ . Предположим, что при  $\mu = 0$  устойчивая и неустойчивая сепаратрисы точки  $O$  касаются одна другой в точке  $P$ , как показано на рис. 6 (при  $\mu < 0$  имеются две точки пересечения  $W^s$  и  $W^u$  в окрестности  $P$ , а при  $\mu > 0$  — ни одной). Тогда теорема Гаврилова—Шильникова [27] утверждает, что для любого  $\mu_* > 0$  существует значение  $\mu_1$  ( $0 < \mu_1 < \mu_*$ ), при котором система (1) имеет устойчивую периодическую траекторию; чем меньше  $\mu_*$ , тем больше ее период и тем уже область ее притяжения. При некоторых дополнительных (но не отражающихся на общности выводов) предположениях Ш. Ньюхаусом [28] установлено, что существует также значение  $\mu_2$ ,  $0 < \mu_2 < \mu_*$ , при котором система (1) имеет бесконечное множество устойчивых периодических траекторий.

Замечательно, однако, что численный эксперимент зачастую не «ловит» эти устойчивые предельные циклы. Показателен следующий пример (один из многочисленных примеров подобного рода).

## В системе Лоренца

$$\dot{x} = -\sigma(x - y), \quad \dot{y} = -xz + rx - y, \quad \dot{z} = xy - bz$$

при  $b = 8/3$ ,  $\sigma = 10,2$ ,  $r = 30,2$  обнаружена бифуркация возникновения контура, составленного из сепаратрис седловых положений равновесия [29]. Из теоретических результатов нелокальной теории бифуркаций [30] с необходимостью следует наличие устойчивых предельных циклов в системе Лоренца при значениях параметров, как угодно близких к приведенным. С помощью ЭВМ тем не менее обнаружить их не удалось. И сколь бы точно ни считали, какая бы мощная и быстродействующая ЭВМ ни была, все равно нашлась бы область параметров, в которой теория гарантирует наличие устойчивых предельных циклов, а машина их не находит.

Приведем слова В. И. Арнольда [31, с. 250] о подобной ситуации (для аттрактора М. Эно [32]): «Наблюдаемое в численном эксперименте приближение к множеству, имеющему локально вид произведения канторова множества на отрезок, не удастся описать в рамках действующих определений гиперболичности<sup>3</sup> (не исключено даже, что в это множество вкраплены области притяжения длинных циклов). Таким образом, математики не признают множество Эно гиперболическим. В то же время с точки зрения экспериментатора движение фазовой точки под действием итераций преобразования  $f$  имеет ярко выраженный стохастический характер...».

9. Итак, мы видим, что свойство физического процесса иметь «чувствительную зависимость от начальных условий» пока не отражено с совершенно полной адекватностью в строгом математическом понятии. Тем не менее аттрактор, построенный с помощью поглощающей области, содержащий бесконечное множество седловых траекторий и имеющий нецелую фрактальную размерность<sup>4</sup>, может служить (и служит) полезным инструментом исследования конкретных систем. Наиболее отчетливо это проявляется при исследовании и описании бифуркаций, отражающих перестройку режимов работы системы при изменении ее параметров. Коснемся только бифуркаций, сопровождающих рост размерности странных аттракторов.

Чтобы в зависящей от параметров системе возник странный аттрактор высокой размерности, необходимо «накопление» неустойчивости при докритическом изменении параметра. Многомерный странный аттрактор не может сразу возникнуть в той области фазового пространства, где до этого не было бифуркаций, где не существовало траекторий с неустойчивыми сепаратрисными

<sup>3</sup> Добавим — и квазигиперболичности.

<sup>4</sup> О размерностных характеристиках странных аттракторов см., например, [18, 22, 33, 34]. Для справки отметим, что ляпуновская размерность аттрактора Лоренца равна примерно 2,06, аттрактора Смейла — Уильямса —  $(1 - \ln 2/\ln a_2)$  (это следует из [33]), а аттрактора Эно, о котором только что говорилось, — примерно 1,26.

многообразиями. Отсюда ясно, что имеются следующие логические возможности возникновения многомерных странных аттракторов.

I. В докритической ситуации существует непритягивающее стохастическое множество и маломерный аттрактор, например состояние равновесия, устойчивый предельный цикл или двумерный тор. В бифуркационный момент маломерный аттрактор исчезает и «бывшее» непритягивающее множество входит в состав родившегося сразу многомерного странного аттрактора.

Именно такой сценарий реализуется в дискретной модели уравнения Гинзбурга—Ландау в области слабых связей:

$$da_j/dt = a_j - (1 + i\beta)|a_j|^2 a_j + e(1 - ic)(a_{j+1} + a_{j-1} - 2a_j), \quad (2)$$

где  $j = 1, \dots, N$ ,  $a_j = a_{j+N}$ ,  $\beta, c > 0$ ,  $\beta c > 1$ , а параметр  $e$  возрастает от нуля.

Согласно [35—37], система (2) имеет решения в виде предельных циклов, описывающих стационарные бегущие волны:

$$a_j = A_n \exp(iw_n t + ij\theta_n), \quad \theta_n = 2\pi n/N, \quad n = 1, \dots, N,$$

$$\omega_n = -\beta + 4e(\beta + c) \sin^2(\theta_n/2), \quad |A_n|^2 = 1 - 4e \sin^2(\theta_n/2),$$

устойчивость которых при  $0 < e \ll 1$  определяется условием  $-\pi/2 < \theta_n < (\pi/2) \bmod \pi$ ,  $\beta c > 1$ . Кроме них, как показано в [37], существует непритягивающее стохастическое множество. По мере увеличения  $e$  эти циклы через бифуркацию Андронова—Хопфа теряют устойчивость. Образовавшиеся в результате этой бифуркации двумерные торы при  $N \gg 1$  также быстро теряют устойчивость. В результате в фазовом пространстве отображения (1) существуют седловые циклы и торы с неустойчивыми многообразиями высокой размерности (от 2 до  $N$ ). Последними теряют устойчивость пара циклов (в силу симметрии по  $\pm j$ ) с  $\theta_n = \pi \pm \pm 2\pi/N$  при  $e = e_* = 0,125 + O(N^{-2})$ .

Образовавшиеся в результате бифуркации Андронова—Хопфа двумерные торы также быстро по параметру  $e$  (в [37] дана оценка интервала существования тора:  $\Delta e \sim \pi^2/16N^2 \ll 1$ ) дестабилизируются, и скачком возникает странный аттрактор с размерностью порядка  $N$ .

II. Другая возможность возникновения многомерных странных аттракторов заключается в том, что при изменении параметра происходит постепенная непрерывная перестройка постоянно существующего аттрактора, при которой его размерность монотонно увеличивается. Здесь следует различать два случая: а) при изменении параметра в аттракторе возникают траектории с большим, чем ранее, числом положительных ляпуновских показателей; б) неустойчивость траекторий в аттракторе усиливается, но число неустойчивых направлений (и, следовательно, число положительных ляпуновских показателей) не растет.

Случай «а» осуществляется в той же системе (2), но при значениях параметра  $e$ , близких к  $e_{**} \gtrsim 0,7$ . В [35, 37] установлено, что при уменьшении  $e$  от больших  $e_{**}$  к меньшим очередной ляпу-

новый показатель  $\lambda_3$  для траекторий в странном аттракторе плавно переходит через нуль, что, естественно, сопровождается увеличением ляпуновской размерности: при  $e = 0,733$  размерность  $D_\lambda = 4,12$ , а при  $e = 0,7$   $D_\lambda = 4,77$ .

Нарастание размерности странного аттрактора, не сопровождаемое возникновением в нем новых неустойчивых направлений (случай «б») осуществляется, например, во всех системах с трехмерным фазовым пространством, имеющим странный аттрактор (в частности, в системе Лоренца). Поскольку ляпуновских показателей всего три и один из них обязан быть нулевым, а еще один — отрицательным в силу диссипативности системы, то остается всего один ляпуновский показатель, увеличение которого и сопровождается нарастанием размерности.

Первая возможность возникновения многомерного аттрактора связана с понятием кризиса, а вторая — внутренней бифуркации. Напомним их. Пусть, как и в п. 4,  $\{U_\mu\}$  — непрерывное семейство поглощающих областей, а  $\{A_\mu\}$  — соответствующее семейство аттракторов,  $\mu \in [\mu_1, \mu_2]$ . Пусть  $\mu_2$  — бифуркационное значение параметра. Если  $A_* = \overline{\lim}_{\mu \rightarrow \mu_2} A_\mu$  — аттрактор системы  $(M, \{f^t\})$ , то  $\mu_2$  называется внутренним бифуркационным значением, если  $A_*$  — не аттрактор, то — кризисным и говорят, что при  $\mu = \mu_2$  происходит кризис семейства аттракторов  $\{A_\mu\}$  (или просто — кризис аттрактора). (Здесь  $\overline{\lim}_{\mu \rightarrow \mu_2} A_\mu$  — это множество всех предельных точек семейства аттракторов  $\{A_\mu\}$ , так называемый верхний топологический предел.) В [7—9] описаны основные механизмы кризисов и внутренних бифуркаций. В частности, оттуда следует, что для предсказания кризисов, и тем более их типов, необходимо при изменении параметров одновременно следить за аттракторами и траекториями, лежащими на границе их областей притяжения. Как следует из приведенных и многих других примеров, это в принципе возможно.

10. Подведем итоги. Из сказанного следует, что понятие аттрактора (в различных вариантах) оказалось плодотворным и хорошо послужило (и служит) как для внутренних нужд теории динамических систем, так и для описания явлений, происходящих в реальных физических процессах, явлений достаточно сложных, как видно из примеров предыдущего пункта. Это позволяет надеяться на то, что дальнейшее развитие в этом направлении теории ДС, может быть с привлечением идей и методов эргодической теории и теории вероятностей (на эту тему см. [11, 13, 18, 22, 33, 34]), приведет к более адекватному отражению феноменов, осуществляющихся в автоколебательных (в широком смысле) физических системах. Хочется надеяться (за последние 20 лет это стало уже традицией), что это путь к построению динамической теории турбулентности.

В заключение автор благодарит М. И. Рабиновича за активное стимулирование к подготовке и написанию работы и полезные замечания.

## ЛИТЕРАТУРА

1. Андронов А. А. Математические проблемы теории колебаний // Первая Всесоюз. конф. по колебаниям. М.; Л.: Гостехтеориздат, 1933. С. 32—71.
2. Андронов А. А., Леонтович Е. А., Гордон И. И., Майер А. Г. Качественная теория динамических систем второго порядка. М.: Наука, 1966. 568 с.
3. Андронов А. А., Леонтович Е. А., Гордон И. И., Майер А. Г. Теория бифуркаций динамических систем на плоскости. М.: Наука, 1967. 487 с.
4. Schuster H. G. Deterministic chaos: An introduction. Weinheim: Physic—Verl., 1984. 456 p.
5. Smale S. Differentiable dynamical systems // Bull. Amer. Math. Soc. 1967. Vol. 73, N 6. P. 747—817.
6. Cosnard M., Demongeot J. Attractors: a deterministic approach // C. r. Acad. Sci. Sér. 1. 1985. T. 300, N 15. P. 551—556.
7. Рюльов В. И., Афраймович В. С., Ильяшенко Ю. С., Шильников Л. П. Теория бифуркаций // Современные проблемы математики. Фундаментальные направления. М.: ВИНТИ, 1986. Т. 5. С. 5—218.
8. Grebogi C., Ott E., Yorke J. A. Crises, sudden changes in chaotic attractors and transient to chaos // Physica D. 1983. Vol. 7, N 2. P. 181—200.
9. Афраймович В. С. Внутренние бифуркации и кризисы аттракторов // Нелинейные волны: Структуры и бифуркации. М.: Наука, 1987. С. 189—213.
10. Вул Е. Б., Синай Я. Г., Ханин К. М. Универсальность Фейгенбаума и термодинамический формализм // УМН. 1984. Т. 39, № 3. С. 3—37.
11. Sinai Ya. G. The stochasticity of dynamical systems // Sel. Math. Sov. 1981. Vol. 1. P. 100—119.
12. Williams R. F. Expanding attractors // Publ. Math. IHES. 1974. Vol. 43. P. 169—203.
13. Ruelle D. Small random perturbation and the definition of attractors // Commun. Math. Phys. 1981. Vol. 82. P. 137—151.
14. Ruelle D. Strange attractors // Math. Intellig. 1980. Vol. 2. P. 126—137.
15. Afraimovich V. S., Shilnikov L. P. Strange attractors and quasiattractors // Nonlinear dynamics and turbulence. Boston; L.; Melbourn: Pitman, 1983. P. 1—34.
16. Bowen R. On axiom A-diffeomorphisms // Reg. Conf. Ser. Math. Providence: Amer. Math. Soc., 1978. Vol. 35. P. 7—28.
17. Conley Ch. Isolated invariant sets and the Morse index // Ibid. Vol. 38. P. 3—90.
18. Eckmann J. P., Ruelle D. Ergodic theory of chaos and strange attractors // Rev. Mod. Phys. 1987. Vol. 57. P. 617—656.
19. Ruelle D., Takens F. On nature of turbulence // Commun. Math. Phys. 1971. Vol. 20. P. 167—192.
20. Аносов Д. В. Геодезические потоки на замкнутых римановых многообразиях отрицательной кривизны. М., 1967. 212 с. (Тр. Мат. ин-та им. В. А. Стеклова; Т. 90).
21. Странные аттракторы / Под ред. Я. Г. Синай, Л. П. Шильникова. М.: Мир, 1979. 253 с.
22. Afraimovich V. S., Pesin Ya. B. Dimension of Lorenz type attractors // Sov. Sci. Rev. C. Math-Phys. 1987. Vol. 6. P. 169—241.
23. Плыкин Р. В. Источники и стоки А-диффеоморфизмов поверхностей // Мат. сб. 1974. Т. 94, № 6. С. 243—264.
24. Lozi R. Un attracteur etrange du type attracteur de Hénon // J. Phys. Coll. C5. 1978. Vol. 179. P. 9—10.
25. Белых В. Н. Модели дискретных систем фазовой синхронизации // Системы фазовой синхронизации. М.: Радио и связь, 1982. С. 161—176.
26. Carrido L., Simó L. Some ideas about strange attractors // Lect. Notes. Phys. 1983. Vol. 179. P. 1—28.
27. Гаерялов Н. К., Шильников Л. П. О трехмерных динамических системах с негрубой гомоклинической кривой // Мат. сб. 1972. Т. 88, № 8. С. 475—492; 1973. Т. 90, № 1. С. 139—156.

28. *Newhouse S. E.*, Asymptotic behavior and homoclinic points in nonlinear systems // *Ann. N. Y. Acad. Sci.* 1980. Vol. 357. P. 292—299.
29. *Афраймович В. С., Быков В. В., Шильников Л. П.* О существовании устойчивых периодических движений в модели // *УМН.* 1980. Т. 35, № 5. С. 164—165.
30. *Быков В. В.* О бифуркациях динамических систем, близких к системам с сепаратрисным контуром, содержащим седло-фокус // *Методы качественной теории дифференциальных уравнений.* Горький: Горьк. гос. ун-т. 1980. С. 44—72.
31. *Арнольд В. И.* Дополнительные главы теории обыкновенных дифференциальных уравнений. М.: Наука, 1978. 304 с.
32. *Hénon M.* A two-dimensional mapping with strange attractor // *Commun. Math. Phys.* 1976. Vol. 50. P. 69—77.
33. *Young L.-S.* Dimension, entropy and Lyapunov exponents // *Ergod. Theor. and Dyn. Syst.* 1982. Vol. 2. P. 109—124.
34. *Бунимович Л. А., Песин Я. Б., Синай Я. Г., Якобсон М. В.* Эргодическая теория гладких динамических систем // *Итоги науки и техники. Современные проблемы математики. Фундаментальные направления.* М.: ВИНТИ. 1985. Т. 2. С. 113—232.
35. *Гапонов-Грехов А. В., Рабинович М. И., Старобинец М. И.* О рождении многомерного хаоса в активных решетках // *ДАН СССР.* 1984. Т. 279, № 3. С. 596—602.
36. *Арансон И. С., Рабинович М. И.* Рождение и динамика спиральных вихрей двумерных моделей Гинзбурга—Ландау // *Изв. вузов. Радиофизика.* 1986. Т. 29, № 12. С. 1514—1517.
37. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* Развитие хаоса в ансамблях динамических структур // *ЖЭТФ.* 1985. Т. 81, № 1. С. 92—94.

УДК 530.182

## ОТ НЕЛИНЕЙНЫХ КОЛЕБАНИЙ — К НЕЛИНЕЙНЫМ ВОЛНАМ

Л. А. ОСТРОВСКИЙ

### ВВЕДЕНИЕ

Название этой статьи не следует, конечно, связывать с хронологией (от колебаний — к волнам). Нелинейные волны стали изучаться отнюдь не позже, чем нелинейные колебания, — достаточно упомянуть работы классиков механики, приведшие к теории ударных волн в газе и нелинейных волн на воде; все это происходило в прошлом веке. Здесь речь пойдет о некоторых идеях теории колебаний в духе Пуанкаре—Мандельштама—Андропова, которые нашли развитие в теории волн, и о тех новых элементах, которые появляются при переходе к волновым системам, имеющим бесконечное число (особенно показателен тут континуум) степеней свободы.

Подходы, выработанные в теории колебаний, начали активно внедряться в теорию волн на рубеже 60-х годов; в значительной степени это было связано с достижениями радиофизики, физики плазмы, лазерной оптики, где с новой степенью подробности началось изучение нелинейных волн в диспергирующих средах. С это-

го времени, пожалуй, все основные «колебательные» понятия — фазовая плоскость, автоколебания, предельный цикл, бифуркация, резонанс, модуляция и т. д. — вошли в обиход теории волн. Можно сказать, что появилась качественная теория волн; естественно, что одновременно с ней активно развивались и приближенные методы. Появились обзоры и книги, в которых теория колебаний и волн излагалась по возможности с единых позиций [1].

Конечно, теория волн — еще и область математической физики, и здесь были достигнуты впечатляющие успехи, связанные с разработкой точных методов решения нелинейных дифференциальных уравнений. Эти яркие достижения на каком-то этапе, я бы сказал, затмили качественную, «колебательную», сторону физики волн. Вскоре они стали использоваться, скажем, для описания динамики солитонов как частиц, построения приближенных методов и т. д. Кроме того, все большую актуальность приобретает изучение неинтегрируемых систем (среды с диссипацией, практически все автоволновые задачи и т. д.). Поэтому попытка обсудить качественную сторону моделей нелинейных волн (в том числе и на материале горьковских школ) и возникающие здесь на современном этапе вопросы представляется своевременной.

## 1. ФАЗОВАЯ ПЛОСКОСТЬ И ФАЗОВОЕ ПРОСТРАНСТВО

Наиболее контрастным фактором, отделившим в свое время теорию колебаний от собственно теории дифференциальных уравнений, можно, по-видимому, считать метод фазовых траекторий, который позволяет построить топологические образы основных типов движений и тем самым получить о них достаточно полное качественное представление. В случае волновых систем конечномерное фазовое пространство может быть построено только для тех или иных частных классов решений. Однако эти классы часто играют первостепенную роль в понимании качественных особенностей волнового процесса.

Метод фазовой плоскости был введен в теорию волн в начале 60-х годов, когда, собственно, и формировалась современная теория нелинейных волн. Р. В. Хохлов [2] исследовал на фазовой плоскости нелинейное взаимодействие гармонической волны с ее второй гармоникой в случаях близких фазовых скоростей (синхронизма), когда между ними происходит эффективный обмен энергией, а Г. И. Фрейдман [3] изучил структуру стационарных ударных электромагнитных волн в линиях с различными дисперсионными характеристиками. Позже использование фазовой плоскости для анализа поведения ударных волн, солитонов, волн огибающих и других типов решений получило широкое распространение.

Среди фазовых траекторий на волновой фазовой плоскости особого внимания (при рассмотрении с принятых здесь позиций)

заслуживают сепаратрисы. На них особенно ясно выявляются различия в роли аналогичных типов решений в случаях колебаний и волн. В теории колебаний сепаратриса — в сущности, нереализуемое решение, граница между областями фазового пространства с топологически различными типами траекторий. В распределенных же системах «сепаратрисные» волны, к которым относятся, вообще говоря, все локализованные стационарные движения — ударные волны, солитоны, ряд автоволн и автоструктур, не только существуют, но и принадлежат к наиболее характерным структурам, которые обнаруживают иногда удивительную устойчивость по отношению к воздействию различного рода возмущений, побуждающую говорить об их сходстве с материальными частицами.

Прежде чем обсуждать причины такой устойчивости, рассмотрим особенности фазовых портретов стационарных бегущих волн, т. е. решений вида  $u = u(\eta)$ , где  $\eta = x - Vt$ ,  $V = \text{const}$ . Эта замена превращает уравнения волн (если они в частных производных) в обыкновенные дифференциальные уравнения, и для автономных систем можно пользоваться методом фазовых траекторий. Однако здесь имеется одна немаловажная особенность: из-за присутствия произвольного параметра  $V$  фазовое пространство оказывается вырожденным. Действительно, все особые траектории (точки равновесия сепаратрисы, предельные циклы) теперь образуют континуум. В качестве элементарных примеров на рис. 1 показаны сепаратрисные поверхности для простейшего уравнения Кортевега—де Фриса—Бюргерса (КдФБ), описывающего солитон и ударную волну:

$$u_t + \alpha u u_x + \beta u_{xxx} = \nu u_{xx}. \quad (1)$$

Эти сепаратрисные поверхности соответствуют решению одной и той же нелинейной краевой задачи вида

$$u(+\infty) = u_1, \quad u(-\infty) = u_2$$

с произвольным параметром  $V$ . Если  $V$  считать собственным значением этой задачи, то существует, вообще говоря, континуум ее решений (хотя, как мы увидим дальше, необязательно все они реализуемы в рамках исходной нестационарной задачи). Наличие такого континуума — одна из главных причин эволюционности нелинейных волн, о которой мы еще упомянем ниже.

Здесь же видны и возможные бифуркации. Так, в уравнении КдФ замкнутая сепаратриса (солитон) существует только при  $V > a$ ; это означает, что в исходных переменных скорость солитона всегда больше скорости линейной длинной волны.

Обсудим немного подробнее другой известный пример, относящийся к автоволнам. Простейшее уравнение возбудимой среды с диффузией имеет вид

$$\mu u_t = D u_{xx} + F(u). \quad (2)$$



Если уравнение  $F(u) = 0$  имеет два корня,  $u_1$  и  $u_2$  (для простоты  $u_1 = 0$ ,  $u_2 = 1$ ), то сепаратрисная поверхность существует на всей оси  $V$ , причем при  $V > V_{\min} = 2\sqrt{DF'(0)}$  это переход типа седло—узел, а при  $V < V_{\min}$  — седло—фокус<sup>1</sup> (рис. 2). Это — известный результат Колмогорова, Петровского и Пискунова [4]. Если же имеются три корня, скажем,  $u_1 = 0$ ,  $u_2 = 1$ ,  $u_3 < 0$ , то появляется вторая сепаратрисная поверхность (рис. 3). Интереснее, однако, то обстоятельство, что существует скорость  $V_\alpha$  ( $V_\alpha = \int_{u_3}^{u_2} F du$ ), при которой возникает сепаратриса, идущая из седла в седло (см. рис. 3, б), при этом вторая сепаратриса идет по-прежнему из седла в фокус (узел).

Отметим следующее важное обстоятельство: если две точки равновесия сливаются в одну ( $u_3 \rightarrow 0$ ), то предельный переход дает решение именно с  $V = V_{\min}$ . В работе [5] численно это было показано для уравнения вида (2) при  $F(u) \sim u^3 - u^2$  (задача решалась применительно к автоволнам турбулентности — движущимся фронтам, разделяющим области ламинарного и турбулентного потоков).

Какие вопросы теории стационарных волн представляются сейчас актуальными? В первую очередь это изучение систем с более чем одной степенью свободы, переход от фазовой плоскости к фазовому пространству. Качественная теория трех- и четырехмерных фазовых пространств развита, конечно, гораздо меньше. Здесь часто приходится использовать ЭВМ. При этом самое интересное — те новые решения, которые принципиально (топологически) невозможны на фазовой плоскости. К таким решениям относятся, в частности, все «обратимые» уединенные решения с  $u(\infty) = u(-\infty)$ , если они меняют знак (таков, например, типичный нервный импульс).

Быть может, самое яркое свойство таких систем — это возможность стохастизации уже в рамках стационарных решений. В случае солитонов такая возможность известна, например, для обобщенного уравнения КдФ [6]

$$u_t + \alpha u u_x + \beta u_{xxx} + \gamma u_{xxxx} = 0. \quad (3)$$

Это уравнение — неинтегрируемое, для него, вообще говоря, существует бесконечное (счетное) число локализованных решений — мультисолитонов — с различным количеством максимумов; все они имеют асимптотику типа  $\exp(-a \pm ib)x$ , т. е. осциллирующие «хвосты» (рис. 4). Такая многостабильность, как мы увидим дальше, создает прямую возможность для стохастизации динамических «ансамблей солитонов». Позже подобные «мультисолитонные» образования были найдены для разных уравнений.

<sup>1</sup> Часто говорят, что  $V > V_{\min}$  есть условие существования переходного решения, однако это оправданно лишь при дополнительном условии знакопостоянства  $u$ .

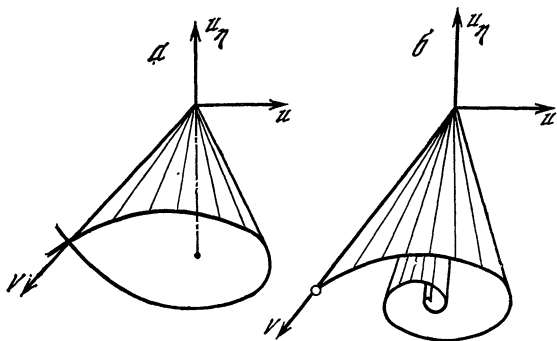


Рис. 1. Вид «сепаратрисной поверхности» для стационарных решений уравнения (1) при  $v=0$  (а) и  $v \neq 0$  (б)

Рис. 2. Структура фазовой плоскости для стационарных решений уравнения (2) при  $V < V_{\min}$  (а) и  $V > V_{\min}$  (б) в случае двух положений равновесия

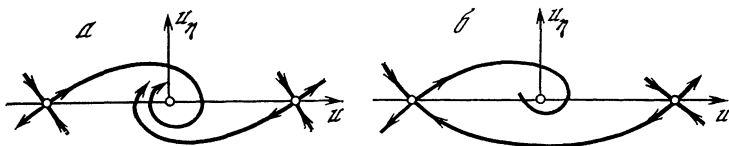
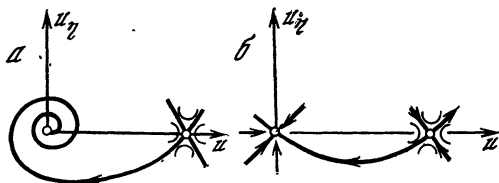


Рис. 3. Структура фазовой плоскости для стационарных решений уравнения (2) при трех положениях равновесия и  $V < V_{\alpha}$  (а),  $V = V_{\alpha}$  (б),  $V = 0$  (в)

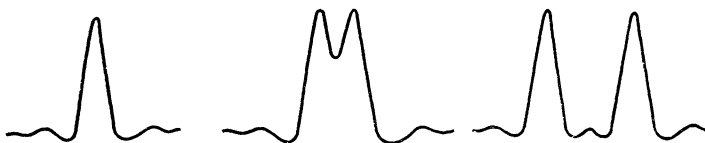
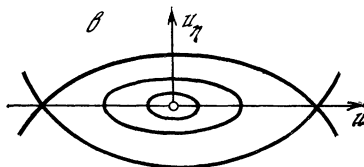


Рис. 4. Некоторые стационарные решения уравнения (3) — мультисолитоны

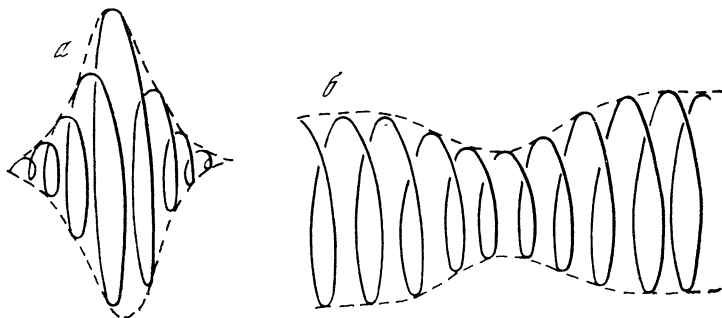


Рис. 5. Возможные уединенные решения уравнения (5)

Что касается активных систем, то возможность стохастизации в них сейчас достаточно хорошо известна, и существование в них режимов типа странного аттрактора представляется естественным. Много изучалась система связанных диффузионных уравнений

$$\tau_1 u_t = D_1 u_{xx} + F_1(u, v), \quad \tau_2 v_t = D_2 v_{xx} + F_2(u, v). \quad (4)$$

Эта система дает, в частности, пример того, как в рамках стационарных решений подчас стираются грани между диссипативными, активными и консервативными системами. Мы уже видели, что при  $V = 0$  стационарный вариант диффузионного уравнения (2) становится консервативным или возможны «смешанные» решения, когда одна сепаратриса идет из седла в седло, а другая — из седла в узел (см. рис. 3). Для консервативного уравнения (3) тоже существуют активные аналоги (типа системы (4)), в которых уравнения стационарных (точнее, статических) решений при  $V = 0$  совпадают с получаемым из (3) при  $V \neq 0$  [7]. При этом и в (4) получается консервативный набор решений типа изображенных на рис. 4. Подчеркнем, однако, что для активной системы это уже не континуум, а одно решение или дискретный набор решений для фиксированных  $V$ . В случае, если они асимптотически устойчивы во времени, их можно рассматривать как аналоги предельного цикла сложной (и притом аperiодической!) формы.

Сравнительно мало внимания пока уделялось вопросу о поляризационных свойствах локализованных волн. Они проявляются, например, в модели «векторного» уравнения КдФ [8]:

$$\mathbf{u}_t + g(\mathbf{u}|\mathbf{u}|^2)_x + p u_{xxx} = 0, \quad (5)$$

где  $\mathbf{u} = (u_y, u_z)$  — вектор в поперечной плоскости. Оно описывает динамику «диспергирующей струны» — цепочки масс, соединенных пружинами и совершающих поперечные колебания. В ней существуют «солитоны вращения», аналогичные солитонам огибающей (в том числе «солитонам затемнения» в виде провалов в периодической структуре), но, возможно, содержащие малое

число периодов закрутки (рис. 5). Уравнение (5), по-видимому, неинтегрируемо, но при замене нелинейного члена на  $u|u|_x^2$  переходит в интегрируемое [9].

## 2. ЭВОЛЮЦИОННОСТЬ

Для построения качественной теории волн первостепенную важность имеет вопрос об условиях существования тех или иных «типовых» структур, в том числе стационарных. Поэтому возникает вопрос об устойчивости или более общий — об эволюционности таких структур, т. е. возможности их формирования из более широкого класса начальных возмущений. Естественно, что фазовое пространство стационарных волн — всего лишь «вершина айсберга», множество меры нуль по отношению ко всему множеству решений исходных нестационарных уравнений. И не только, скажем, устойчивость траектории на фазовой плоскости не гарантирует реальной устойчивости волны, но и неустойчивость ее не означает истинной неустойчивости во времени. Это относится, наконец, к понятию грубости, т. е. сохранения топологии фазового пространства, и в частности данной фазовой траектории, при добавлении к ней произвольных малых членов в исходном уравнении.

Относительно давно и подробно изучалась устойчивость и эволюционность диссипативных периодов ударных волн. Необходимые условия устойчивости для них можно получить, как известно, уже в разрывном приближении, когда рассматриваются крупномасштабные (по сравнению с толщиной ударного перехода) возмущения и ударную волну можно считать резкой границей раздела. Условие устойчивости определяется тогда требованием равенства числа параметров нормальных волн, способных «убегать» от разрыва, числу связывающих их граничных условий. В газодинамическом случае отсюда получается известное условие: устойчивый разрыв должен быть сверхзвуковым по отношению к среде перед ним и дозвуковым по отношению к среде позади него [10].

Однако в общем случае вопрос, насколько свойства стационарных решений позволяют судить о реальной возможности их существования, т. е. об их устойчивости и эволюционности, весьма сложен. Как мы уже видели, прямой связи тут нет, и даже грубость или негрубость траекторий на фазовой плоскости ничего не говорит об их реальной грубости в рамках волновой задачи. Вместе с тем определенные утверждения, подчас даже строгие, сформулировать, оказывается, можно. Это уже было сделано для ударной волны, которая в фазовом пространстве описывается сепаратрисой [3]. Можно показать, что необходимое условие существования и единственности такой сепаратрисы совпадает с упомянутым выше условием устойчивости ударной волны как разрыва — для неустойчивых скачков такого сепаратрисного решения просто не существует. Иными словами, существуют

лишь такие стационарные ударные волны, которые могут возникнуть в результате эволюции плавных возмущений (типа простой волны). Это относится, в частности, к уравнению КдФБ (1).

Для консервативных стационарных волн, в том числе солитонов, с эволюционностью дело обстоит сложнее. В теории колебаний сепаратриса, выходящая из седла и возвращающаяся в него или входящая в другое седло (а именно такая сепаратриса отвечает солитону), — всегда негрубое образование (добавление малых потерь качественно изменяет процесс), однако солитоны в известном смысле вполне эволюционны — во многих уравнениях любое локализованное возмущение в асимптотике распадается на солитоны. Конечно, асимптотическое приближение к выделенному решению не свойственно консервативным системам, для этого нужна диссипация. В данном случае ее роль играет излучение, «сбрасывание» части энергии на бесконечность. Если же в среде, где распространяется «готовый» солитон, добавляются малые потери, то качественных изменений процесса сразу не происходит, солитон лишь плавно эволюционирует по сепаратрисной поверхности.

Поучительные примеры эволюционности или ее отсутствия дает теория автоволн. Интересной (и подчас не совсем верно трактуемой) оказывается в этом смысле уже классическая задача Колмогорова—Петровского—Пискунова, связанная с уравнением (2), точнее, с сепаратрисой, соединяющей две ближайшие точки равновесия (два ближайших нуля  $F(u)$  на оси  $u$ ). Авторы [4] показали, что из всех стационарных решений с  $V \geq V_{\min}$  эволюционно лишь решение с  $V = V_{\min}$ : оно является асимптотикой для начального возмущения в виде ступеньки или перепада, разделяющего области постоянных значений  $u_1 = 0$  и  $u_2 = 1$ . Этот важный результат приводится сейчас повсеместно. Однако необходимо отметить, что и такой эволюционный перепад не может существовать неограниченно долго, поскольку он неустойчив по отношению к малым возмущениям, возникающим в области впереди него. Действительно, линеаризовав уравнение (1) в окрестности точки равновесия  $u_1$  или  $u_2$  и отыскивая возмущение вида  $u' \sim \exp i(\omega t - kx)$ , мы получим дисперсионное уравнение

$$\omega = i(k^2 D - F'_u),$$

где значение  $F'_u$  берется в точке равновесия. Отсюда видно, что область, где  $F'_u = 0$ , всегда устойчива (возмущение затухает), но при  $F'_u > 0$  и  $|k| < k_0 = (F'_u/D)^{1/2}$  возмущение нарастает. Это последнее условие как раз и выполняется в нулевой точке (узел на фазовой плоскости стационарных решений). Поэтому такой автоволновой перепад неустойчив по отношению к длинноволновым возмущениям в разрывном (в том же смысле, что и для ударных волн) приближении (легко видеть, что  $k_0^{-1}$  имеет порядок ширины фронта перепада при  $V = V_{\min}$ ).

По-другому обстоит дело с перепадами, идущими из седла в

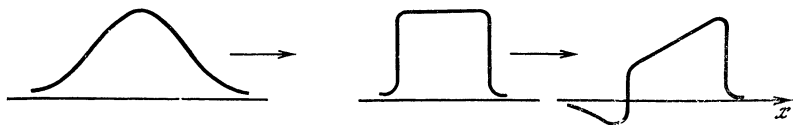


Рис. 6. Эволюция импульса в рамках уравнения (4) при  $\tau_1 \ll \tau_2$

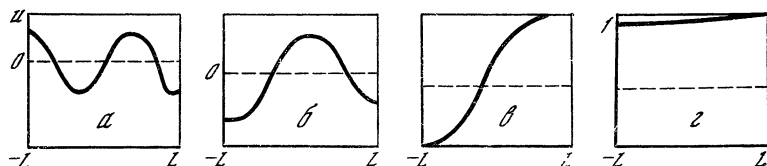
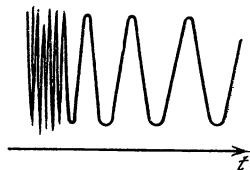


Рис. 7. Эволюция решения уравнения (2) в ограниченной области

$a - t = 0$ ,  $б - 15$ ;  $в - 30$ ;  $г - 60$

Рис. 8. «Регуляризация» хаотического процесса в цепочке активных элементов



седло, в том же уравнении. Здесь обе равновесные области равноправны в том отношении, что в них  $F'_u < 0$  и возмущения затухают. Таким образом, явно «более негрубый» перепад (сепаратриса из седла в седло, единственная скорость) оказывается более «жизнеспособным».

Этот вывод подтверждается анализом более сложной двухкомпонентной системы типа (4) [11, 12]. Если  $a_1 \ll a_2$ , то переменную  $u$  можно считать «быстрой», а  $v$  «медленной» — времена их релаксации сильно различаются. Тогда при каждом  $V = \text{const}$  существует перепад типа седло—седло, который при  $V > V_0$  ( $V_0$  — значение  $V$ , соответствующее неподвижному перепаду) отвечает фронту ( $u_x < 0$ ), а при  $V < V_0$  — спаду ( $u_x > 0$ ). В области вне перепада скорость  $v$  медленно меняется, что приводит к изменению его параметров. В результате формируется стационарный импульс (рис. 6).

Был проделан более детальный анализ устойчивости и бифуркаций такой системы при различных соотношениях ее параметров. Оказалось, что характер решения зависит от отношения величин  $q_{1,2} = \sqrt{D_{1,2}/\delta_{1,2}}$ , имеющих размерность скорости: в зависимости от этого отношения «выживает» либо статическое решение, либо бегущий перепад. (Подробнее об этом см. статью Ю. Д. Калафати и Ю. А. Ржанова в этом сборнике.) Заметим, что неподвижный перепад может начать делиться, генерируя периодическую автоволну (т. е. возникает одномерный ведущий центр).

Заслуживает внимания также численное исследование уравнения (2) (с тремя стационарными точками) в ограниченной области, проведенное И. М. Старобинцем и В. Г. Яхно [13]. Они

отметили еще одно качественное свойство подобных систем: устойчивость данной моды по отношению к более мелкомасштабным и неустойчивость к более крупномасштабным возмущениям. На рис. 7 показан пример эволюции одной из относительно высоких мод системы (такое стационарное решение существует), переходящей в более гладкое решение и в конце концов — в однородное устойчивое состояние.

Один из интересных вопросов — о возможности усложняющейся эволюции волн, приводящей к их стохастизации.

При модовом описании, когда задача сводится к обыкновенным уравнениям, эта возможность очевидна. Возможно, однако, и развитие подобных процессов в пространстве, скажем, для бегущих волн, хотя этот вопрос рассматривался еще относительно мало. Один из ранних примеров стохастизации ансамбля солитонов — в обобщенном уравнении вида (3). Эксперимент в электрической системе (ее можно считать и аналоговой машиной) показал, что периодическое начальное условие приводит к формированию ансамбля солитонов, который при определенных периодах соответствует «замороженной» решетке солитонов, а при других периодах регулярность решетки нарушается и она стохастизируется [14].

Другой пример отвечает, собственно, не волнам (в классическом понимании этого термина), а структурам в струе жидкости: начальное возмущение последовательно усложняется вниз по потоку вплоть до его стохастизации (см., например, [15]).

При численном моделировании наблюдался и обратный процесс — переход хаоса в регулярные колебания (рис. 8).

### 3. МОДУЛЯЦИЯ И РЕЗОНАНСЫ

К вопросу об эволюционности нелинейных волн примыкает и вопрос об их структурной устойчивости по отношению к малым возмущениям в определенных уравнениях (добавление потерь в консервативной системе, неоднородность в пространстве и т. д.). Несколько расширив определение, можно отождествить это свойство с грубостью системы или ее отдельных решений. Наиболее эффективным способом ответа на этот вопрос можно, пожалуй, считать метод возмущений. Вариантов этого метода для волновых задач сейчас известно много, но они все, по существу, строятся по общему принципу [16].

Рассматривается система

$$\mathfrak{E} \hat{M}u = \varepsilon \hat{R}(u), \quad (6)$$

где  $\hat{M}$ ,  $\hat{R}$  — некоторые операторы;  $u$  — набор (вектор) искомых функций, а  $\varepsilon$  — малый параметр. В качестве нулевого приближения берется какое-либо решение уравнения  $\hat{M}u = 0$ ; обозначим его  $u^{(0)}(x, t)$ . Решение отыскивается в виде ряда по степеням  $\varepsilon$ :

$$u = u^{(0)} + \sum_{n=1}^N \varepsilon^n u^{(n)}. \quad (7)$$

Подставляя (7) в (6), в каждом порядке по  $\varepsilon$  мы получим нелинейные неоднородные уравнения

$$\hat{L}u^{(n)} = H^{(n)}, \quad (8)$$

где  $\hat{L}$  — линейный оператор, образованный при вариации оператора  $\hat{M}$ , а  $H^{(n)}$  содержит только функции предыдущих приближений. Отсюда уже можно судить о грубости решений. Если все  $u^{(n)}$  остаются ограниченными (будучи ограниченными в начальный момент), то решение остается близким к  $u^{(n)}$ . Если же  $u^{(1)}$  или какая-либо из  $u^{(n)}$  секулярно растет, то имеет место резонанс и малая правая часть в уравнении (1) обуславливает сильные изменения решения. Эти случаи и представляют наибольший интерес.

Как именно происходит эволюция решения в случае такого резонанса? Часто на этот вопрос можно ответить, если учесть, что во многих случаях резонанс связан с зависимостью  $u^{(0)}$  от одного или нескольких параметров  $A_s$  (амплитуда, фаза, скорость волны):  $u^{(0)} = u^{(0)}(x, t, A_s)$ . Легко видеть, что к числу собственных функций оператора принадлежат все функции  $Y_s = \partial u^{(0)}/\partial A_s$ . Ясно также, что решение следующего приближения будет содержать слагаемые, пропорциональные  $Y_s$ , причем последние, присутствуя в правой части (8), создают резонансное воздействие, и в  $u^{(1)}$  будут секулярно растущие слагаемые, так что (локально во времени)

$$\delta u^{(1)}(t) \sim \delta t \sum_s \alpha_s \frac{\partial u^{(0)}}{\partial A_s} = \sum_s \frac{\partial u^{(0)}}{\partial A_s} \delta A_s,$$

где  $\alpha_s$  — медленно меняющиеся функции, а  $\delta A_s = \alpha_s \delta t$  — приращение параметра  $A$ . Поэтому решение следующего приближения имеет вид

$$u = u^{(0)} + u^{(1)} = u^{(0)} A_s + \sum_s \frac{\partial u^{(0)}}{\partial A_s} \delta A_s = u^{(0)}(A_s + \delta A_s). \quad (9)$$

Отсюда следует важный вывод: резонансная компонента возмущений изменяет (в первом приближении) решение так, что оно остается в семействе  $u^{(0)}(A_s)$ , но при изменяющемся значении  $A_s$ .

Чтобы найти, как именно изменяется  $A_s$ , используют асимптотический метод возмущений: величины  $A_s$  считают медленно меняющимися функциями  $x$  и  $t$ . Изменение параметров  $A_s$  и есть следствие резонанса, а закон этого изменения находится как раз из условия (условие ортогональности), чтобы поправки к  $u^{(0)}$ , если возможно, уже не нарастали. Это условие имеет вид

$$\int H^{(n)} Y_s^* dt = 0, \quad (10)$$

где интеграл берется по периоду для периодических  $u^{(0)}$  или на бесконечном интервале — для локализованных. Величины  $Y^*$  — собственные функции сопряженного оператора  $\hat{L}^*$ . Поскольку



в  $H^{(p)}$  входят производные  $A_{st}$ ,  $A_{sx}$ , то условия (10) эквивалентны дифференциальным уравнениям для  $A_s$ .

К схеме асимптотической теории возмущений сводится (с различными вариациями) огромное число задач как теории колебаний, так и теории волн. Для волн, близких к сумме гармонических, она приводит к уравнениям резонансного взаимодействия при выполнении известных условий синхронизма. Именно так, в сущности, строится теория резонансного триплета, процессов вынужденного рассеяния и др. Так описываются и «волны огибающих» — волновые изменения амплитуды и частоты модулированных волн, и в частности нелинейных волновых пакетов — солитонов огибающих. Мы, однако, обсудим снова уединенные, сепаратрисные волны. Из сказанного ясно, что малые возмущения в уравнениях приводят, вообще говоря, к сильным изменениям решения, но для широкого класса условий, когда справедлива формула (9), эти возмущения «дрейфуют» по сепаратрисной поверхности. Это одна из причин структурной устойчивости солитонов, роднящей их с материальными частицами. В частности, основные типы взаимодействий солитонов могут быть поняты с помощью метода возмущений, когда рассматриваются слабо перекрывающиеся солитоны, взаимодействующие через свои периферийные поля («хвосты»). За последние 10—15 лет на этом пути была построена динамика солитонов как частиц [16, 17].

В зависимости от асимптотики их полей солитоны могут либо отталкиваться, и тогда их сближение сменится удалением (как в обычном уравнении КдФ), либо притягиваться, и тогда образуется связанное состояние, либо (при наличии осциллирующих «хвостов», как в уравнении (2)) демонстрировать сложное поведение, комбинирующее оба этих типа. В частности, возможны устойчивые решетки солитонов [14]; скажем, в уравнении КдФ солитоны благодаря экспоненциальным «хвостам» образуют известную решетку Тода, в которой существуют, в свою очередь, солитоны, из них опять можно составить решетку Тода и т. д. — образуется «иерархия решеток» (рис. 9, а). С другой стороны, есть неустойчивые решетки, которые легко стохастизируются, как в уравнении (2) (рис. 9, б).

Во всех подобных случаях координаты солитонов  $S_i(t)$  удовлетворяют уравнениям в обыкновенных производных:

$$M \frac{d^2 S_i}{dt^2} = \sum_{j=1}^n f(S_i - S_j),$$

где  $f$  определяется асимптотикой солитонов, а  $M$  — масса, т. е. производная полного полевого импульса солитона  $p$  по его скорости  $V$ . Таким образом, здесь мы как бы возвращаемся от волн к колебаниям — от уравнений в частных производных к уравнениям в обыкновенных производных.

Среди современных направлений в классической динамике солитонов мы отметим два. Во-первых, это вопрос о выходе за пре-

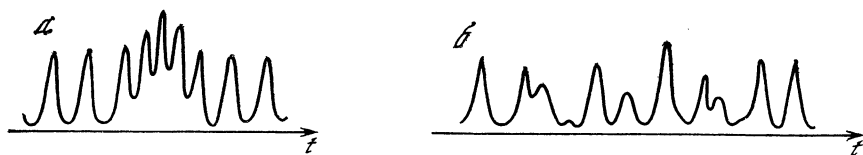


Рис. 9. «Солитон из солитонов» в уравнении КдФ (а) и стохастический ансамбль солитонов в уравнении (3) (б)

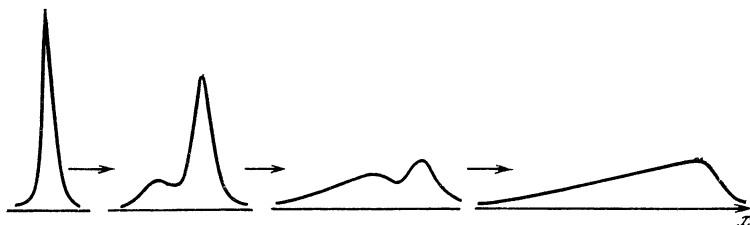


Рис. 10. Затухание солитона в уравнении (1) — адиабатическая стадия и переход к автомодельной волне

делы первого, адиабатического приближения. В следующих приближениях в поле накапливается неадиабатическая, нелокализованная компонента — излучение солитона, отвечающее уже не дискретным резонансным модам, а сплошному спектру возмущений. В этом приближении солитон аналогичен излучающей частице, но его движение вблизи сепаратрисной поверхности возможно теперь лишь в течение ограниченного времени, после чего его структура разрушается и асимптотика решения может быть совсем иной. Например, в уравнении КдФБ с малой вязкостью амплитуда солитона  $A$  сначала затухает по адиабатическому закону

$$A = A_0 (1 + bA_0 t)^{-1}, \quad (11)$$

где  $b = 4\alpha v / 45\beta$ . Постепенно за солитоном вырастает «полочка» — область почти однородного возмущения. В результате солитон превращается в треугольный импульс — автомодельное решение уравнения Бюргера (рис. 10).

Первая стадия подобных решений — вырастание нелокализованного «хвоста» — изучалась В. И. Карпманом и Е. М. Масловым [18] с помощью развитого ими варианта асимптотического метода, основанного на близости системы к точно интегрируемой. К. А. Горшков и В. В. Папко [19] с помощью электромагнитной аналоговой линии воспроизвели весь процесс затухания солитона, все стадии этого процесса.

В этой связи интересны задачи об излучении солитона в подсистему, в которой это излучение является собственной модой. Один из примеров — известные уравнения, описывающие связь высоко- и низкочастотной компонент (с амплитудами  $\psi$  и  $n$  соответственно), например электронной и ионно-звуковой волн в

плазме:

$$i\psi_t + \psi_{xx} = n\psi, \quad n_{tt} - n_{xx} = (|\psi|^2)_{xx}. \quad (12)$$

Для такой системы обсуждалась не только задача об излучении, но и о его обратном влиянии на динамику солитона [20], т. е. о «радиационном трении» для солитона как частицы. Оказалось, что с учетом поправок второго и третьего приближений движение солитона в поле крупномасштабной волны  $n = \tilde{n}(x, t)$  описывается (при достаточно медленных движениях) уравнением

$$\ddot{x}_s = - \frac{2}{(1 + 8c^2/3)} (\tilde{n}_x - c\dot{x}_s), \quad (13)$$

где  $x_s$  — координата солитона;  $c$  — постоянная, имеющая смысл числа квантов в нем, а производная  $\tilde{n}_x$  берется в точке нахождения солитона.

Значительно больше работ, в которых ограничиваются расчеты самого поля излучения (пока оно мало). В последнее время подробно изучалось возмущенное уравнение sin-Гордона

$$u_{tt} - u_{xx} + \sin u = R(u, t), \quad (14)$$

где  $R$  — малая функция. При периодической зависимости  $R(t)$  здесь возможна резонансная генерация поля излучения солитонами и бризерами [21]. В подобных задачах широко используют метод возмущений для обратной задачи рассеяния, о котором мы уже упоминали.

Другое направление можно назвать стохастической динамикой солитонов. Прежде всего это стохастическая динамика солитона как частицы в поле заданной волны или при параметрическом воздействии. Уже довольно давно рассматривалась стохастизация ансамбля солитонов, например в резонансе, возбуждаемом периодической силой [22]. Возможность такой стохастизации (образование газа солитонов) была подтверждена и экспериментально. Делались и попытки статистического описания «газа солитонов» в твердом теле [23] и в плазме [24], причем вводилась и другая, «фононная» подсистема.

Несколько проще описать стохастическое поведение одиночного солитона в поле волны. Такое поведение возникает, например, в рамках системы (12). Много результатов здесь относится к возмущенному уравнению sin-Гордона, например, когда в (14)  $R \sim \sin u \cos nt$ . В подобных случаях динамика солитона близка к поведению частицы, захваченной в поле волны, при движении вблизи сепаратрисы [25].

Для нелинейных волн в диссипативных и активных средах тоже применимы методы адиабатической теории. Правда, здесь в ряде случаев дело обходится без резонансов: движение по сепаратрисной поверхности связано непосредственно с «внешними условиями». Таковы классические задачи об эволюции ударной волны в переменном поле, например в простейшем случае — об эволюции разрыва на фронте треугольного звукового импульса.

Здесь затухание скачка (тоже движение по сепаратрисной поверхности) определяется темпом нелинейной эволюции поля за его фронтом, а структура ударной волны просто «подстраивается» под нее (ситуация, характерная для пограничного слоя). Для автоволн тоже возможно «слежение», скажем, волны переброса за изменением параметров в пространстве. Так, в уравнении (2) изменение в пространстве положений равновесия функции  $F(u)$  приведет к изменению амплитуды перепада, а изменение коэффициента диффузии  $D$  — к изменению его длительности (при этом, конечно, диффузионный член в (2) имеет вид  $(D(x)u_x)_x$ ). Подобные задачи неоднократно рассматривались (см. обзор [26]). К этому же классу относится и уже упомянутая эволюция импульса с фронтом и спадом в рамках уравнений (4), приводящая к формированию стационарного импульса.

Но существуют и ситуации, в которых автоволны определяются не квазистатическими внешними параметрами, а, как и для солитонов, резонансным воздействием возмущений [27]. Пример — уравнение (2) с переменными параметрами, в котором в начальный момент задан неподвижный перепад ( $V = 0$ ), реализуемый, например, при  $F(u) = \alpha^2 u - u^3$ . Если  $D$  и  $\alpha$  — плавно меняющиеся функции  $x$ , то, хотя невозмущенное решение неподвижно в любой точке, неоднородность порождает «резонансный» дрейф перепада, координата которого  $x$  определяется уравнением

$$\dot{x}_s = -\frac{3}{4} D'_x(x_s) - \frac{5}{2} \frac{\alpha'_x(x_s)}{\alpha(x_s)}. \quad (15)$$

Этот результат был подтвержден численными расчетами. При этом для автоволны была применена теория возмущений типа «солитонной».

#### 4. ВОЛНЫ И СТРУКТУРЫ В ТРЕХМЕРНОМ ПРОСТРАНСТВЕ

Насколько «колебательная идеология» остается полезной в дву- и трехмерных задачах? Нужно заметить, что каких-то специфически трехмерных нелинейных решений, не имеющих аналогий или простых предельных переходов к одномерным (или к сосредоточенным системам), известно немного. Зато нет недостатка в разного рода аналогиях. Так, в нелинейной оптике весьма полезной оказалась в свое время пространственно-временная аналогия: поведение одномерных волновых пакетов в нелинейной диспергирующей среде (волн огибающих) и стационарных двумерных пучков описывается аналогичными уравнениями (с заменой  $x, t$  на  $x, y$ ); отсюда, в частности, аналогия между самофокусировкой и самомодуляцией. Что же касается локализованных волн — ударных и солитонов, то здесь хорошо развиты «геометрооптические» (или «геометроакустические») методы, когда фронт распространяется в пространстве по лучам (вообще говоря, зависящим от амплитуды поля).

Обзор таких решений был сделан на одной из прошлых школ; сейчас подобные методы получают более строгое обоснование [28]. Отметим, что позже и здесь были достигнуты определенные успехи в описании излучения искривленных фронтов солитонов, проще говоря, эффекта дифракционных потерь. Еще Дж. Уизем [29] указывал, что рефракция ударных волн может привести к излому фронта (разрыв на разрыве), то же возможно для солитона. В. И. Шрира [30] вывел для возмущений солитонного фронта уравнение Бюргерса, описывающее их нелинейное затухание из-за дифракции; в 1986 г. В. Е. Захаров [31] получил точное решение того же типа для уравнения Кадомцева—Петвиашвили.

Позже начали строиться «геометрические» модели автоволн в виде искривленных фронтов перебора. В сущности так устроен (при сильной релаксационности) знаменитый ревербератор — вращающаяся спиральная волна возбуждения (мне с трудом удалось удержаться от соблазна привести какую-либо эффектную фотографию такого ревербератора в химии или биологии — такие фотографии приводятся во всех статьях и книгах, имеющих хоть какое-то отношение к автоволнам).

В последнее время моделям пространственной динамики автоволновых фронтов уделяется много внимания. Поскольку, как уже говорилось, амплитуда перепада однозначно определяется параметрами среды, теория распространения фронта может иметь в первом приближении кинематический характер. Из кинематических соображений получается уравнение для кривизны фронта  $k(s)$  с учетом условия в начале фронта ( $s$  — координата вдоль фронта). Решение этого уравнения показывает, что в случае постоянной скорости фронта существует спиральная волна, «привязанная», например, к неподвижному центральному кругу. Эта кривая — эвольвента круга с радиусом  $R$ , которая при  $s \gg R$  переходит в архимедову спираль (это было показано еще Винером и Розенблютом в 1946 г.). Если же учесть слабую зависимость скорости от кривизны фронта, то получается и непрерывное решение, включающее центральную область. Мы не описываем подробнее этих решений, поскольку они приведены в трудах предыдущей школы [32].

Существуют и фронты в трехмерном пространстве в виде разного рода «труб» и «свитков» (рис. 14), причем в последнее время делаются попытки определить и классифицировать топологически возможные типы таких волн. В частности, полезную информацию дают законы сохранения топологических констант — индекса Пуанкаре  $m$  (число оборотов волнового вектора при обходе особой точки) и топологического заряда  $n$ , равного числу периодов волн при обходе такой точки [33]. Этот подход интересен для нас тем, что он по существу аналогичен анализу траекторий в фазовом пространстве колебательной системы.

Другой класс задач — о локализованных нелинейных образованиях в пространстве, в том числе о дву- и трехмерных солитонах. Первые исследования трехмерных локализованных решений были

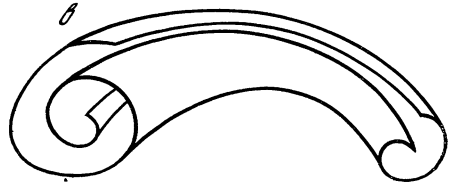
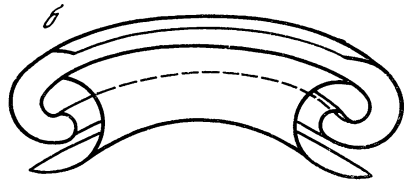
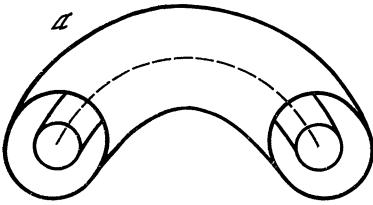
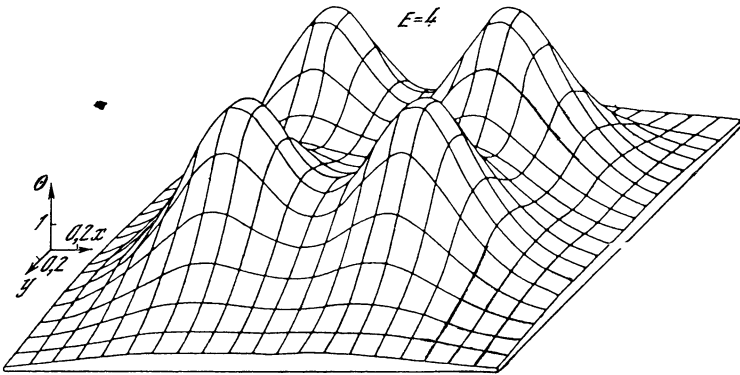


Рис. 11. Возможные структуры автоволновых фронтов в пространстве

$a - (n, l) = (0, 0)$ ;  $b - (2, 0)$ ;  
 $c - (1, K)$

Рис. 12. Двумерные стационарные мультисолитоны



связаны с попытками построить «частицеподобные» решения уравнений нелинейной теории поля (в классическом приближении). Позже внимание теоретиков обратилось на двумерные задачи, и здесь было получено по крайней мере два важных типа двумерных солитонов: солитон Кадомцева—Петвиашвили для квазиплоских дифрагирующих нелинейных пучков и солитон Ларичева—Резника (а затем и другие солитоны) для сильнонелинейных волн Россби. Позже с помощью ЭВМ были найдены различные двумерные солитонные структуры, например в системе, описывающей тепловое самовоздействие электромагнитного пучка в плазме. В докладе К. А. Горшкова, В. А. Миронова и А. М. Сергеева на Горьковской школе 1981 г. [34] описаны решения этой системы в виде одно- и двумерных мультисолитонов (рис. 12); по-видимому, существуют и трехмерные мультисолитоны в сходных системах. Во всех случаях справедлива упомянутая выше простая трактовка таких структур как результата взаимодействия «хвостами» одиночных («одн огорбых») солитонов.

Значительное внимание уделяется взаимодействиям неоднородных солитонов. Так, В. Д. Ларичев и Г. М. Резник [35] изучили динамику пары солитонов Россби, каждый из которых пред-

ставляет собой «вихревую пару» (рис. 13). Оказывается, что такие вихри взаимодействуют почти как солитоны в узком смысле, т. е. возвращаются к исходному состоянию, но не полностью.

Для автоволн дву- и трехмерные локализованные образования, изученные сейчас, — это нередко просто неподвижные стационарные структуры. Можно думать, что есть многочисленные виды трехмерных структур, обобщающих, скажем, известный брусселятор. Вообще разница между сильно нелинейными волнами и

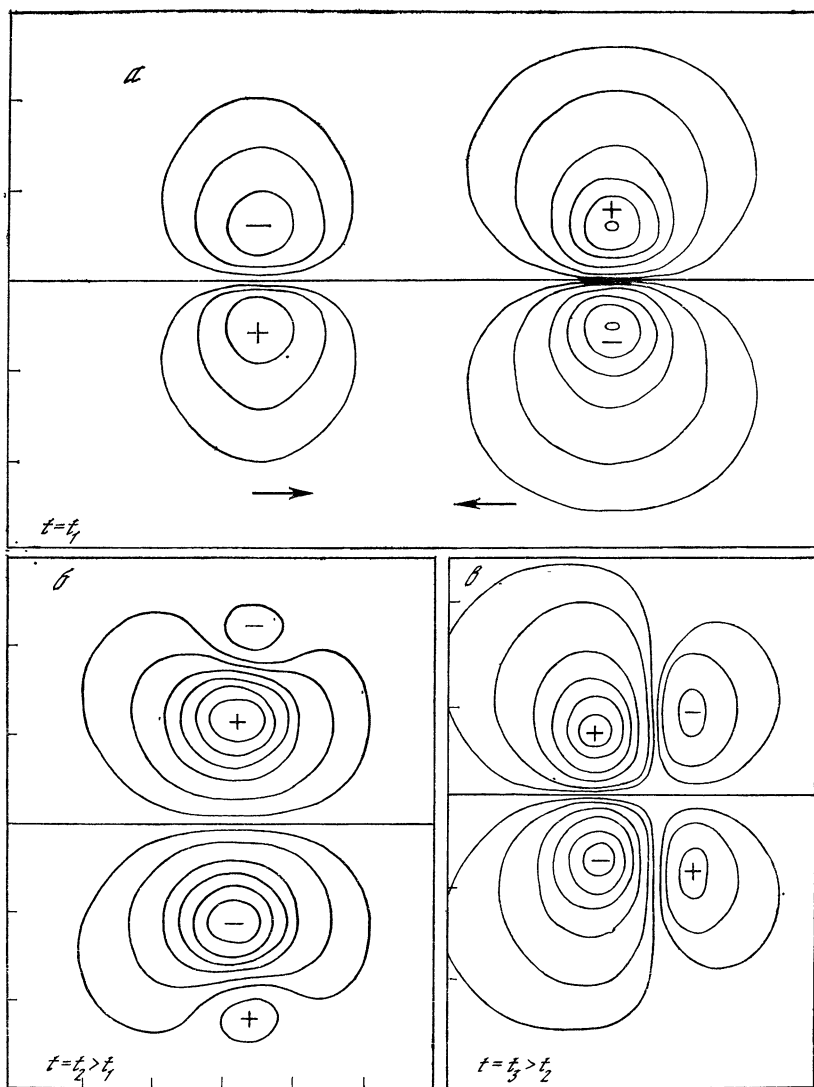


Рис. 13. Взаимодействие локализованных волн Россби в последовательные моменты времени  $t_1$  (а),  $t_2 > t_1$  (б) и  $t_3 > t_2$  (в)

«просто» структурами, особенно между автоволнами и автоструктурами, становится все более условной.

И снова нельзя обойти вопрос о волновой турбулентности. На первом этапе изучалась слабая волновая турбулентность, т. е. ансамбль резонансно взаимодействующих волн в слабонелинейной диспергирующей среде. Отметим, что, кроме такого «хаоса» квазигармонических волн, можно говорить (хотя и несколько условно) о восстановлении «порядка из беспорядка»: когда речь идет об обращении волнового фронта, «испорченные» фронты со сложной, возможно, структурой восстанавливаются после усиления в процессе, например, вынужденного рассеяния [36].

Позже интенсивно изучалась сильная волновая турбулентность, образованная, например, ансамблем солитонов или сложными автоструктурами. Эти эффекты теоретически еще мало изучены, хотя в экспериментах наблюдались неоднократно. Вероятно, весьма общее значение для построения стохастических волновых моделей имеют системы типа Гинзбурга—Ландау:

$$u_t = u^*(1 - \alpha|u|^2) + D\Delta u. \quad (16)$$

Если  $\alpha$  и  $D$  — комплексные параметры, то к частным случаям (16) принадлежат и уравнения автоволн, и нелинейное уравнение Шредингера. Сложные режимы в системах такого рода, бифуркации, стохастические режимы изучаются сейчас очень интенсивно [37—39]. Здесь возникают сложные вопросы о связи «временного хаоса» (маломодовых хаотических движений, соответствующих странному аттрактору) и сложной пространственной динамики. Однако эти вопросы относятся скорее к следующей лекции этого цикла.

Говоря об аналогиях между волнами и частицами, уместно упомянуть и о вихрях как частицах. Динамика двумерных (особенно точечных) гидродинамических вихрей, когда они ведут себя как структурно цельные образования, изучается давно; показано, что система точечных вихрей может стохастизоваться. Недавно была построена такая динамика системы элементарных вихревых диполей [40].

Об интересе к гидродинамическим вихревым структурам говорит уже то, что они рассматривались (с совершенно разных позиций) по меньшей мере в пяти лекциях на предыдущей школе. Среди них такие важные с колебательно-волновой точки зрения вопросы, как генерация и динамика автоволн (или автовихрей) Россби во вращающейся жидкости [41], а также попытка описать движения жидкости как систему замкнутых вихревых нитей (петель), что допускает статистическое описание с помощью уравнения Фоккера—Планка [42]. Возникает естественный вопрос об определении вихрей как солитонов, в том числе о построении для них аналогичной теории возмущений. Здесь, однако, есть одна трудность: обычно вихри нельзя охарактеризовать количеством решений с конечным (или хотя бы с дискретным) набором парамет-



ров; они допускают функциональный произвол. Тем не менее удается задать ряд необходимых условий ортогональности, описывающих движение вихрей.

## ЗАКЛЮЧЕНИЕ

В одной статье невозможно дать обзор всего развития теории нелинейных волн, да это и не нужно (и в этом смысле приведенная короткая библиография — не более чем набор иллюстраций). Мы лишь попытались проследить за эволюцией некоторых понятий и методов при переходе к волнам. Следует признать, что не только для осциллирующих, но и для локализованных, сепаратрисных волн колебательные аналогии оказываются полезными, хотя и ограниченными. Вместе с тем именно такие волны часто служат основой перехода от волн к структурам, а точнее, стирания граней между ними. Вообще переход к структурам (столь популярный сегодня термин!) — это здесь в значительной степени переход к сильнонелинейным (и, в частности, к локализованным) волнам. На этом пути возникают пограничные образования типа «волна—неволна», «волна—частица», «волна—вихрь» и т. д. Здесь же происходит и расплывчатая, не раз обсуждавшаяся граница между волнами и турбулентностью; достаточно упомянуть понятие сильной волновой турбулентности (в том числе ленгмювской — в плазме), иногда почти отождествляемой с солитонами. При этом, конечно, возможны структуры с внутренними колебательными степенями свободы (как, например, солитоны огибающих). Если подойти с физических позиций, то именно в этих классах, по-видимому, лежат экстремальные, наиболее интенсивные и кратковременные виды волн — дельтосекундные лазерные импульсы, самые крутые внутренние волны в океане, волны переброса в биологии и химии...

Такие структуры обычно сочетаются со слабым волновым фоном («излучением»), что важно и с теоретико-полевой точки зрения, и для практических приложений (модели солитонной теплопроводности или переноса возбуждений в белковых молекулах основаны на рассмотрении систем типа «солитоны—фононы», «солитоны—экситоны» и т. п.).

Все это выводит теорию нелинейных волн на весьма интересный, но и трудный уровень, где аналитические методы часто отказывают уже потому, что речь идет в основном о неинтегрируемых системах. Поэтому попытки качественного исследования представляются особенно важными, причем их, по-видимому, целесообразно использовать синергетически, в сочетании с численными методами («теоремное» использование ЭВМ). Можно думать, что на этой основе будет создана новая качественная теория волн, которая получит эффективно используемые «образы» их основных типов.

Трудно сказать, будет это достигнуто в связи с развитием чистой теории или будет стимулироваться прикладными задачами, но это и вряд ли существенно. В любом случае это было бы хоро-

пей иллюстрацией тезиса, высказанного А. Пуанкаре [43]: «Поиски прекрасного приводят нас к тому же выбору, что и поиски полезного...»

#### ЛИТЕРАТУРА

1. Рабинович М. И., Трубецков Д. И. Введение в теорию колебаний и волн. М.: Наука, 1986. 432 с.
2. Хохлов Р. В. О распространении волн в нелинейных диспергирующих линиях // РЭ. 1961. Т. 6. С. 116—1122.
3. Фрейдман Г. И. О зависимости структуры ударных электромагнитных волн в двухпроводных линиях передачи от дисперсионной характеристики системы в линейном приближении // Изв. вузов. Радиофизика. 1963. Т. 6, № 3. С. 536—550.
4. Баренблатт Г. И. Подобие, автомодельность, промежуточные асимптотики. Л.: Гидрометеоиздат, 1978. 208 с.
5. Островский Л. А., Цимринг Л. Ш. Автоволны турбулентности в свдиговых потоках // Изв. АН СССР. ФАО. 1988. Т. 24, № 6. С. 657—662.
6. Gorshkov K. A., Ostrovsky L. A., Pikovskiy A. S., Papko V. V. On the existence of stationary multisolitons // Phys. Lett. A. 1979. Vol. 74. P. 177—179.
7. Pikovskiy A. S., Rabinovich M. I. Stochastic dissipative structures // Thermodynamics and regulation of biological processes. Walter de Cruyter Co, 1984. P. 343—352.
8. Gorbacheva O. B., Ostrovsky L. A. Non-linear vector waves in a mechanical model of a molecular chain // Physica D. 1983. Vol. 8. P. 223—228.
9. Захаров В. Е., Маняков С. В., Новиков С. П., Питаевский Л. П. Теория солитонов. Метод обратной задачи. М.: Наука, 1980. 134 с.
10. Ландау Л. Д., Лифшиц Е. М. Гидродинамика. М.: Наука, 1986. 736 с.
11. Островский Л. А., Яхно В. Г. Формирование импульсов в возбудимой среде // Биофизика. 1975. Т. 20, № 3. С. 489—493.
12. Ortoleva P., Ross J. Spiral waves in reaction-diffraction equations // J. Chem. Phys. 1975. Vol. 63. P. 3398—3431.
13. Starobinets I. M., Yakhno V. G. One-dimensional autowaves. Methods of qualitative description // Self-organization autowaves and structures far from equilibrium. Springer—Verlag, 1986. P. 98—102.
14. Горшков К. А., Папко В. В. Динамические и стохастические колебания решеток солитонов // ЖЭТФ. 1977. Т. 72, № 1(7). С. 178—187.
15. Гапонов-Грехов А. В., Рабинович М. И. Автоструктуры. Хаотическая динамика ансамблей // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 7—44.
16. Gorshkov K. A., Ostrovsky L. A. Interactions of solitons in nonintegrable systems // Physica 3D. 1981, N 1/2. P. 428—438.
17. Горшков К. А., Островский Л. А., Папко В. В. Взаимодействия и связанные состояния солитонов как классических частиц // ЖЭТФ. 1976. Т. 71, № 2(8). С. 585—593.
18. Карпман В. И., Маслов Е. М. Теория возмущений для солитонов // Там же. 1977. Т. 73, № 2(8). С. 537—559.
19. Горшков К. А., Папко В. В. Неадиабатическая стадия затухания солитонов и промежуточные асимптотики // Изв. вузов. Радиофизика. 1977. Т. 20, № 4. С. 360—365.
20. Арансон И. С., Горшков К. А., Рабинович М. И. Возникновение стохастичности при взаимодействии солитона модуляции с низкочастотной волной // ЖЭТФ. 1986. Т. 86, № 3. С. 929—936.
21. Маломед Б. А. Излучение солитона во внешнем резонансном поле // Физика низких температур. 1987. Т. 13, № 5. С. 469—479.
22. Островский Л. А. Ударные волны и солитоны // Изв. вузов. Радиофизика. 1976. Т. 19, № 5/6. С. 661—690.
23. Бишоп А. Солитоны и физические возмущения // Солитоны в действии. М.: Мир, 1981. С. 72—102.
24. Протогенов А. П., Фрайман Г. М. Термодинамика сильной волновой турбулентности // ЖЭТФ. 1987. Т. 92, № 5. С. 1658—1667.

25. *Абдуллаев Ф. Х., Хабибуллаев П. К.* Динамика солитонов в неоднородных конденсированных средах. Ташкент: Фан, 1986. 184 с.
26. *Гуревич А. В., Минц Р. Г.* Локализованные волны в неоднородных средах // УФН. 1984. Т. 142, № 1. С. 61—78.
27. *Альбер А. Я., Островский Л. А.* Эффект дрейфа волны перепада в неоднородной возбудимой среде // Изв. вузов. Радиофизика. 1989. Т. 32, № 3.
28. *Маслов В. П.* Нестандартные характеристики в асимптотических задачах // УМН. 1983. Т. 38, № 6. С. 3—36.
29. *Уизем Дж.* Линейные и нелинейные волны. М.: Мир, 1977. 622 с.
30. *Шрира В. И.* Нелинейная рефракция солитонов // ЖЭТФ. 1980. Т. 79, № 1(7), С. 87—98.
31. *Захаров В. Е.* Ударные волны, распространяющиеся по солитонам на поверхности жидкости // Изв. вузов. Радиофизика. 1986. Т. 29, № 9. С. 1073—1079.
32. *Давыдов В. А., Михайлов А. С.* Спиральные волны в распределенных активных средах // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 261—279.
33. *Гурия Г. Г., Лифшиц М. А.* Топологические основы классификации автоволн: Препр. ИАЭ № 4085. М., 1985. 16 с.
34. *Горишков К. А., Миронов В. А., Сергеев А. М.* Связанные стационарные солитонные образования // Нелинейные волны. Самоорганизация. М.: Наука, 1983. С. 112—128.
35. *Ларичев В. Д., Резник Г. М.* О столкновении двумерных уединенных волн Россби // Океанология. 1983. Т. 23, № 5. С. 725—734.
36. *Беспалов В. И., Пасманик Г. А.* Нелинейная оптика и адаптивные лазерные системы. М.: Наука, 1986. 134 с.
37. *Зельдович Я. Б., Маломед Б. А.* Сложные волновые режимы в распределенных динамических системах // Изв. вузов. Радиофизика. 1982. Т. 25, № 6. С. 591—618.
38. *Маломед Б. А.* Бифуркации и автоволны // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 251—261.
39. *Гапонов-Грехов А. В., Рабинович М. И.* Уравнение Гинзбурга—Ландау и нелинейная динамика неравновесных сред // Изв. вузов. Радиофизика. 1987. Т. 30, № 2. С. 131—143.
40. *Черанов С. Г.* Динамика точечных вихревых движений и спонтанные сингулярности в трехмерных турбулентных потоках // ЖЭТФ. 1987. Т. 93, № 1(7). С. 151—158.
41. *Незлин М. В., Снежкин Е. Н.* Нелинейные вихревые структуры и самоорганизация в неустойчивых течениях на вращающейся мелкой воде // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 67—86.
42. *Мигдал А. А.* Вихревое уравнение Фоккера—Планка // Там же. С. 159—174.
43. *Пуанкаре А.* О науке. М.: Наука, 1983. 560 с.

УДК 530.182

## НЕЛИНЕЙНАЯ ДИНАМИКА И ТУРБУЛЕНТНОСТЬ

М. И. РАБИНОВИЧ

### ВВЕДЕНИЕ

Наверное, ни одно явление не вызвало столь широкого интереса и в то же время столь бурных дискуссий и противоречивых утверждений среди физиков, математиков и инженеров, как явление турбулентности. Это связано не только с исключительной сложностью проблемы, но и с существенно различным ее пониманием инженерами, математиками и физиками. Действительно,

с прикладной точки зрения суть проблемы турбулентности заключается в формулировке некоторых эффективных уравнений (гораздо более простых, чем уравнения Навье—Стокса), которые бы позволили для конкретных турбулентных течений рассчитывать, например, сопротивление обтекания, эффективный теплоперенос и т. п. Написание таких уравнений может опираться и на интуитивные представления, и на разного рода полумпирические гипотезы. Важно лишь, чтобы они позволяли предсказать ответ в реальных экспериментальных ситуациях. Сейчас здесь имеются замечательные результаты (см., например, [1]). Однако с точки зрения физика (механика, математика) один из главных вопросов, почему ламинарное (упорядоченное) течение стало неупорядоченным и как эта неупорядоченность сказывается на средних характеристиках течения, при таком подходе к турбулентности остается в стороне. Главное при этом понять, происходит ли запутывание, перемешивание элементов жидкости в результате, например, действия неконтролируемых пульсаций набегающего на тело потока либо в результате неустойчивостей, присущих самому течению. Шумы и флуктуации, конечно, важны и здесь, но они играют роль лишь механизма, запускающего действие неустойчивости.

Эти вопросы принципиальны для любой турбулентности, независимо от ее «ведомственной» принадлежности — гидродинамической, плазменной, спиновой и проч. Другими словами, речь идет о фундаментальной проблеме: каким способом нелинейное поле (произвольной природы) переходит в неупорядоченное, случайное движение и как это движение описывать?

Эволюция взглядов на природу случайного в теории турбулентности и статистической физике во многом аналогична. Действительно, естественные представления о малых возмущениях, нарастающих в результате развития линейной неустойчивости, высказанные еще Осборном Рейнольдсом, не могли объяснить переход к турбулентности в таких (казавшихся простейшими) течениях, как течение Пуазейля и течение Куэтта между плоскостями. Согласно анализу малых колебаний в рамках уравнений Навье—Стокса, для течения в трубе вообще не существует критического числа Рейнольдса, при котором бы ламинарное течение становилось неустойчивым, в то время как все экспериментаторы прекрасно знают, что при  $Re \gtrsim 5000$  это течение турбулизуется. Подобные «неудачи» линейной теории заставили многих сомневаться в самой возможности применения детерминированных уравнений Навье—Стокса для описания турбулентных течений. Возникло убеждение (Карман, Тейлор), что, подобно движению газов, турбулентность можно понять и истолковать лишь на основе статистического подхода. Эта точка зрения подкреплялась и тем, казалось бы очевидным, обстоятельством, что при достаточно большом значении числа Рейнольдса число степеней свободы, участвующих в движении, столь велико, что описание течения возможно лишь в среднем.

# 1. МОДЕЛЬ ЛАНДАУ—ХОПФА

Первые попытки объяснить неупорядоченное, хаотическое течение воды с чисто динамических позиций были предприняты Л. Д. Ландау (1944) [2] и Е. Хопфом (1948) [3]. Их модели, одинаковые по смыслу, реализуют идею о последовательном усложнении течения за счет развития иерархии неустойчивостей с несоизмеримыми временными масштабами. В результате нелинейной стабилизации этих неустойчивостей устанавливается течение, поле скорости которого оказывается тем более неупорядоченным, чем большее число возмущений с несоизмеримыми масштабами принимает участие в его формировании. Если построить автокорреляционную функцию поля скорости такого течения, то можно заметить, что корреляции достаточно быстро спадают. Обнаружить же определенную регулярность в сформированном таким образом процессе можно, лишь наблюдая его в течение времени, большего, чем время возврата Пуанкаре  $T \sim e^{\alpha N}$  ( $\alpha \approx 1$ ), где  $N$  — число возмущений с независимыми частотами  $\omega_j$  ( $j = 1, 2, \dots, N$ ).

Пример простой динамической системы, где реализуется решение с бесконечным числом таких несоизмеримых частот, построил Хопф. Он предположил, что одномерное течение в замкнутом контуре описывается двумя уравнениями в частных производных ( $u, v, F$  — комплексные функции):

$$\partial u / \partial t = -\hat{M}\{v, u\} - \hat{M}\{v, 1\} + \mu \partial^2 u / \partial x^2, \quad (1)$$

$$\partial v / \partial t = \hat{M}\{v, u\} + \hat{M}\{v, F\} + \mu \partial^2 v / \partial x^2$$

с периодическими граничными условиями  $u(t, x) = u(t, x + 2\pi)$ ,  $v(t, x) = v(t, x + 2\pi)$ . Здесь  $1/\mu$  — аналог числа Рейнольдса,  $F(x)$  — функция, описывающая действие внешнего поля (источник энергии), а оператор  $\hat{M}\{f, g\} = \frac{1}{2\pi} \int_0^{2\pi} f(x-y) \overline{g(y)} dy$  определяет нелинейное взаимодействие между различными возмущениями.

Полагая, что внешнее поле представляется в виде

$$F(x) = \sum_{-\infty}^{\infty} (a_m + ib_m) e^{imx}, \quad (2)$$

где отношение любых  $b_m$  не является рациональным числом, и что среди величин  $a_m/m^2$  нет равных, решение (1) запишем в виде

$$u(x, t) = \sum_{-\infty}^{\infty} u_m(t) e^{imx}. \quad (3)$$

Хопф показал, что построенное таким образом решение при увеличении  $1/\mu$  последовательно усложняется. В фазовом пространстве «течения» (1) этому соответствуют последовательные бифуркации: при  $\mu = \mu_1$  из тривиального состояния равновесия рождается предельный цикл (образ одночастотного периодического течения),

затем при  $\mu_2 < \mu_1$  цикл становится неустойчивым и в результате бифуркации рождается аттрактор в виде незамкнутой намотки на двумерном торе. После  $k$ -й бифуркации (при  $\mu_k < \mu_{k-1}$ ) появляется аттрактор в виде эргодической намотки на  $k$ -мерном торе и т. д.

Гипотеза Ландау—Хопфа сыграла большую роль в становлении идей о динамической природе турбулентности, хотя сами модели (предложенные четыре десятилетия назад) и оказались неадекватными задаче. Дело в том, что аттрактор в виде незамкнутой намотки на многомерном торе (являющийся в рамках этих моделей образом турбулентности) структурно неустойчив, т. е. он разрушается при малом изменении параметров системы. Это означает, что такое неупорядоченное течение (а точнее, течение сложное, с большим числом несоизмеримых частот в спектре Фурье), как правило, реализоваться не может. И тем не менее сама догадка о том, что развитие турбулентности определяется последовательным (по числу Рейнольдса) возбуждением новых степеней свободы течения, как мы увидим, оказалась правильной!

## 2. СТРАННЫЕ АТТРАКТОРЫ

Принципиальное изменение представлений о природе турбулентности (и природе случайного вообще) произошло после открытия и проникновения во все области научного знания феномена динамического хаоса — случайного поведения полностью детерминированных систем. Математическим образом случайного движения динамической системы служит стохастическое множество траекторий в соответствующем фазовом пространстве. Когда речь идет о турбулентности при конечных числах Рейнольдса, т. е. о течении жидкости с конечной вязкостью, нам в первую очередь интересно установившееся стохастическое движение. Образом такого движения в фазовом пространстве является притягивающее стохастическое множество — странный аттрактор [4].

Странный аттрактор — притягивающее множество траекторий, среди которых все (или почти все) являются неустойчивыми — седловыми — может возникнуть после небольшого числа бифуркаций в фазовом пространстве даже весьма простых течений. Наиболее известный пример — конвекция в подогреваемой снизу тороидальной полости, расположенной в вертикальной плоскости [5]. Образом хаотических колебаний (или вращений) жидкости внутри такой полости служит знаменитый аттрактор Лоренца [6].

Как показывают эксперименты с течениями в замкнутых полостях (течение Куэтта между вращающимися цилиндрами, термоконвекция в горизонтальном слое), наиболее распространенные сценарии перехода к хаосу в простых ситуациях — это разрушение квазипериодических движений [7] (механизм, из-за которого не может реализоваться модель перехода Ландау—Хопфа<sup>1</sup>), пере-

<sup>1</sup> Заметим, что усложнение поведения динамической системы, по Ландау — Хопфу, можно тем не менее реализовать в некоторых специальных случаях, например для цепочки однонаправленно связанных ротаторов [8].

межаемость [9] и бесконечная последовательность удвоений периода [10]. В реальных ситуациях могут наблюдаться и более сложные, комбинированные сценарии [11], однако именно обнаружение этих канонических сценариев для конкретных течений показало, что наши представления о динамическом характере процессов в области перехода к турбулентности, по крайней мере для внутренних течений, справедливы (см. [12, рис. 9, 13]).

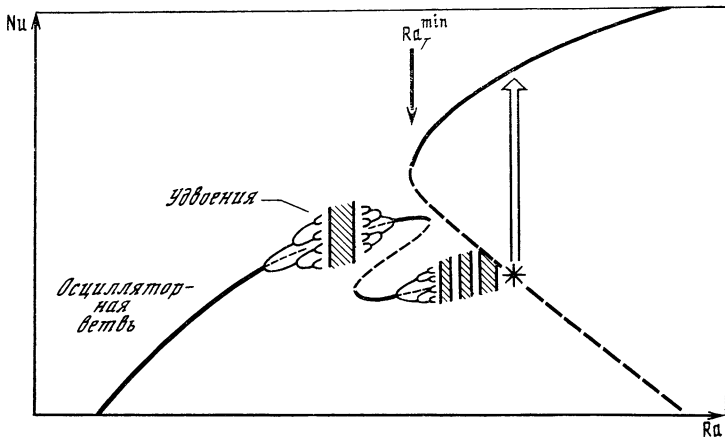
Сейчас о применимости хаотической динамики к описанию возникновения турбулентности можно судить уже не только по характеру бифуркаций течения, предшествующих возникновению беспорядка, но, кроме того, и по тому, что непосредственно по наблюдаемым данным удается восстановить странный аттрактор за точкой перехода! Впервые это было сделано для течения Куэтта—Тейлора и для термоконвекции в замкнутой полости [13, 14].

Теоретические (в том числе и численные) результаты по описанию перехода к турбулентности опираются обычно не на анализ уравнений Навье—Стокса, а на исследование некоторых модельных динамических систем в виде нескольких дифференциальных уравнений в обыкновенных производных — галеркинских мод. Если на первом этапе развития динамической теории турбулентности такое положение представлялось довольно естественным, то в последние годы, когда принципиальная возможность существования нерегулярно пульсирующих решений динамической системы, скажем, пятого порядка, очевидна, возникает необходимость рассмотрения исходных уравнений для гидродинамических полей. Здесь уже имеются некоторые результаты, с одной стороны, подтверждающие возможность использования «конечномерного опыта», а с другой — демонстрирующие потребность в выработке нового, нетрадиционного описания хаотической динамики нелинейных полей (в частности, показана неестественность описания пространственно-временного хаоса с помощью мод линейной краевой задачи<sup>2</sup>). Приведем примеры, иллюстрирующие эти утверждения.

Как известно, полное аналитическое описание перехода к хаосу через последовательность удвоений периода получено для детерминированных систем, описываемых одномерным отображением параболы [15]. Этот сценарий детально исследован и для маломерных моделей с непрерывным временем [16]. В 1983 г. переход к хаосу по сценарию Фейгенбаума был обнаружен в прямом численном эксперименте с полными уравнениями гидродинамики

$$\begin{aligned} \frac{1}{\sigma} \left[ \frac{\partial}{\partial t} \nabla^2 \psi + \frac{\partial}{\partial (x, z)} (\psi, \nabla^2 \psi) \right] &= \text{Ra}_T \theta_x - \text{Ra}_S S_x + \nabla^4 \psi, \\ \partial \theta / \partial t + \partial (\psi, \theta) / \partial (x, z) &= \psi_x + \nabla^2 \theta, \\ \partial S / \partial t + \partial (\psi, S) / \partial (x, z) &= \psi_x + \tau \nabla^2 S, \end{aligned} \quad (4)$$

<sup>2</sup> Сейчас неупорядоченное временное поведение поля с упорядоченной пространственной структурой принято называть хаосом, термин же «турбулентность» используется лишь для пространственно-временного беспорядка (когда неупорядочена и пространственная картина поля).



**Рис. 1.** Изменение теплового потока ( $Nu$  — число Нуссельта) с ростом числа Рэлея  $Ra_T$  для системы (4) при  $Ra_S = 10^4$ ,  $\tau = 10^{-3/2}$  и  $\sigma = 1$  [17]

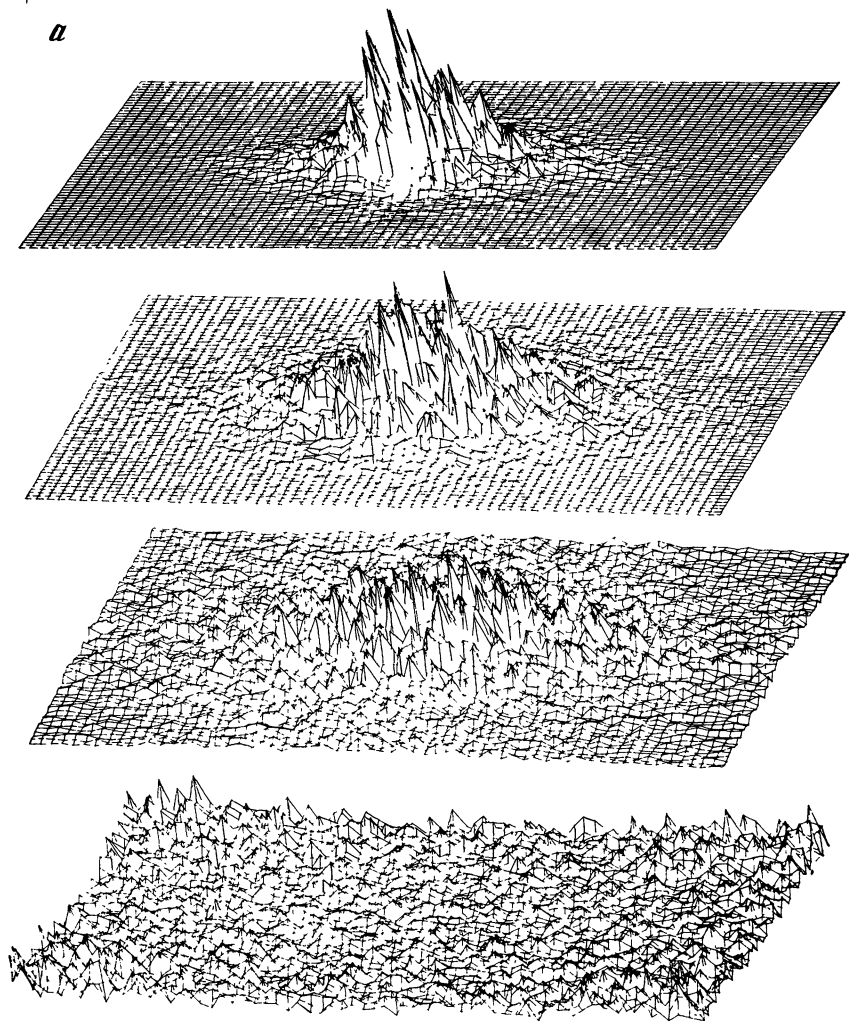
описывающими диффузию температуры и солености в подогреваемом внизу горизонтальном слое жидкости. Здесь  $\sigma$  — число Прандтля;  $0 < \tau < 1$  — отношение диффузии примеси к термодиффузии;  $Ra_T$  и  $Ra_S$  — тепловое и примесное числа Рэлея;  $\psi(x, z)$  — функция тока;  $\theta$  и  $S$  — отклонения температуры и солености (концентрации примеси) от равновесных. Эта система решалась методом конечных разностей при граничных условиях  $\psi = \psi_{zz} = \theta = S = 0$  при  $z = 0, 1$ ;  $\psi = \psi_{xx} = \theta_x = S_x = 0$  при  $x = 0, \lambda$ . По мере увеличения надкритичности  $Ra_T$  были обнаружены бифуркации удвоения периода и рождение хаоса (рис. 1).

В численном эксперименте с полными уравнениями Навье—Стокса удалось обнаружить и другой сценарий перехода к хаосу — через разрушение квазипериодического движения. В работе [18] при расчете обтекания профиля Жуковского ( $Re \approx 10^3 \div 10^4$ ) при увеличении угла атаки был обнаружен переход от стационарного течения (малые углы атаки) к периодическому, затем квазипериодическому и, наконец, хаотическому. Таким образом, сейчас можно утверждать, что переход к хаосу по сценариям, обнаруженным в маломерных динамических системах, реализуется и в моделях, описываемых уравнениями в частных производных<sup>3</sup>.

На следующем примере покажем, что традиционное модовое представление весьма неудобно для описания пространственного беспорядка, типичного для турбулентных течений в не столь три-

<sup>3</sup> Строго говоря, и в этих задачах, где использовался метод сеток, речь идет о конечномерных системах. Однако, поскольку число элементов весьма велико ( $\geq 10^6$ ), представляется вполне допустимым связывать эти результаты с уравнениями Навье — Стокса.



*a*

виальной геометрии, как тороидальная полость. Рассмотрим двумерную турбулентность, описываемую амплитудными уравнениями типа Гинзбурга—Ландау:

$$\begin{aligned} \partial u / \partial t \approx u [1 - (1 + i\beta)|u|^2] + e(1 - ic)\Delta u; \quad u(x, y) = \\ = u(x + L, y + L). \end{aligned} \quad (5)$$

На рис. 2 представлены результаты компьютерного эксперимента с дискретной моделью (5) (исследовалась решетка из  $128 \times 128$  дискретных элементов) [19]. Видно, что с ростом степени неравновесности «среды» (параметр  $1/e = r$ ) пространственный спектр возбуждений сильно уширяется и возникает ощущение, что все возбужденные линейные моды равноправно участвуют



**Рис. 2.** Хаотическая динамика двумерного поля, описываемого дискретным аналогом уравнения (5) [19]

*a* — эволюция пространственного спектра с ростом надкритичности  $r = 1/e$  (сверху вниз); *б* — эволюция пространственных структур при  $e = 20$ ; *в* — картина поля при  $e = 4$

в формировании пространственно-неупорядоченного состояния поля; при этом, как кажется, пространственная структура поля должна быть весьма сложной и запутанной. Однако это впечатление оправдано лишь при достаточно больших  $r$  ( $e \lesssim 5$ ,  $\beta = c = 1,71$ ). При меньших  $r$  ( $e \approx 20 \div 50$ ), несмотря на сравнительно сложный пространственный спектр, картина поля оказывается довольно простой — образуется небольшое число локализованных структур, движение которых хаотично.

### 3. РАЗВИТИЕ ТУРБУЛЕНТНОСТИ. РАЗМЕРНОСТЬ

Эксперименты, убеждающие в реализуемости для определенных видов течений сценариев перехода к хаосу, предсказываемых динамической теорией, и, кроме того, прямое восстановление структуры аттракторов за критической точкой побудили исследователей попытаться определить на основе представлений динамического хаоса природу последующего развития турбулентности по мере увеличения надкритичности. Высказанная еще Ландау и Хопфом идея о том, что с ростом числа Рейнольдса число степеней свободы течения, вовлекаемых в турбулентное движение, должно возрастать, была проверена в лаборатории.

Если действительно реальное турбулентное течение описывается динамической системой, то число степеней свободы этой системы можно оценить, определив в эксперименте размерность <sup>4</sup>

<sup>4</sup> См., например, статью В. С. Афраймовича и А. М. Реймана в наст. сб.

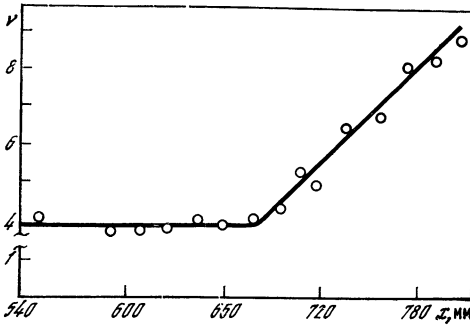


Рис. 3. Изменение корреляционной размерности вдоль пограничного слоя на пластине [21]

Вниз по потоку начинается рост размерности, происходит разрушение двумерных валов

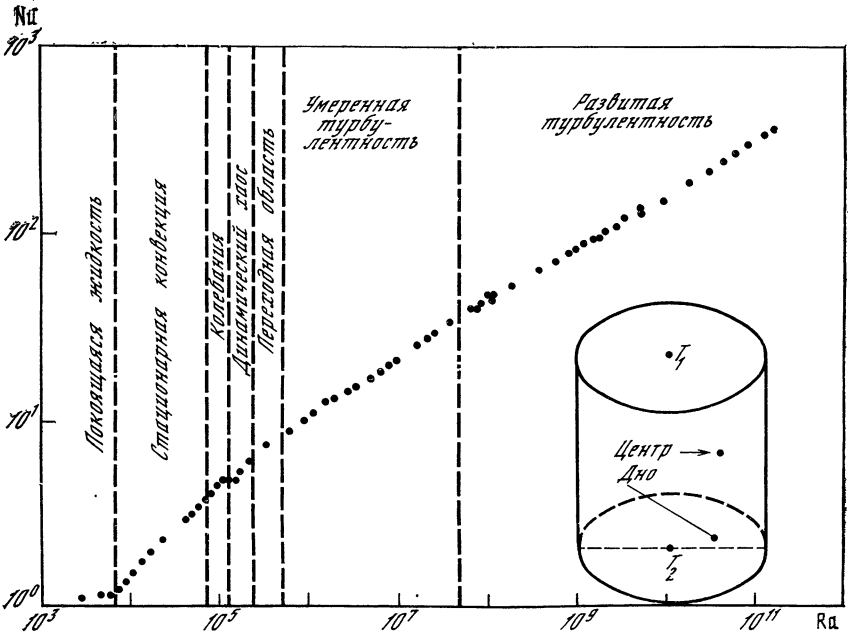


Рис. 4. Возникновение и развитие турбулентности в ячейке, заполненной газообразным гелием

странного аттрактора. В соответствии с высказанными представлениями развитию турбулентности должно отвечать увеличение измеряемой размерности аттрактора при увеличении числа Рейнольдса. Именно этот факт и был установлен в упоминавшихся экспериментах Брандштатера и др. [13, 14].

Существенно, что для определения размерности течения необязательно восстанавливать аттрактор. В 1983 г. Гроссбергер и Прокаччия [20] предложили процедуру измерения размерности движения непосредственно по наблюдаемой. Это позволило затем сделать и следующий шаг — установить природу развития турбулентности в сдвиговых течениях [21]. Непосредственные изме-

рения размерности турбулентного движения как функции координаты показали, в частности, что развитие турбулентности вниз по потоку в пограничном слое (рис. 3) представляет собой постепенное (вдоль течения) вовлечение в турбулентное движение все новых и новых степеней свободы.

Совсем недавно появился эксперимент [22], в котором удалось проследить усложнение динамики течения в ячейке от возникновения регулярной макроскопической конвекции через хаос к развитой турбулентности. Был пройден интервал чисел Рэля от  $10^3$  до  $10^{11}$  (!)<sup>5</sup>. Измерения температуры проводились в двух различных точках ячейки, что, в частности, позволило определить границу возникновения пространственного беспорядка. После возникновения осцилляторной неустойчивости наблюдались стандартные пути возникновения хаоса, затем появился странный аттрактор невысокой размерности. При  $Ra \approx 2,5 \cdot 10^5$  разрушались корреляции между реализациями в разных точках и возникал пространственно-временной хаос — турбулентность. Этому режиму, по-видимому, также отвечает странный аттрактор, но весьма высокой размерности. Результаты этого эксперимента иллюстрирует рис. 4.

## ЗАКЛЮЧЕНИЕ

Еще сравнительно недавно три направления теории турбулентности — статистическое (Рейнольдс), структурное (Толлмин, Таунсенд и др.) и динамическое (Ландау, Лоренц, Рюэль, Арнольд и др.) — развивались параллельно и практически независимо. Результаты, полученные в каждом из них, относились к различным задачам и отвечали на вопросы, возникающие в качественно различных экспериментальных ситуациях. Сейчас кажется удивительным, что такая автономия просуществовала столь долго, и лишь в последние пять лет появилось стремление и возможность (!) к построению единой цельной теории.

Особенно явно сегодня взаимодействие структурного и динамического подходов к турбулентности. Это, в частности, связано со многими экспериментальными успехами, убеждающими исследователей в том, что при умеренных числах Рейнольдса турбулентность представляет собой пространственно-временной хаос взаимодействующих структур. Качественное доказательство этого факта дают эксперименты с визуализацией гидродинамических течений (см., например, [24, 25]), количественное же его обоснование получается при специальной обработке измерений и сравнительно просто добывается в численном эксперименте. Так, при компьютерном исследовании двумерной турбулентности, описы-

---

<sup>5</sup> Эксперименты проводились с газообразным гелием при температуре 4 К в цилиндрическом контейнере; как показано в [23], такое изменение числа  $Ra$  возможно (при практически постоянном числе Прандтля) за счет изменения давления в ячейке.

ваемой уравнениями Навье — Стокса с периодическими граничными условиями [26], было обнаружено, что размерность хаотического множества, вычисленная с помощью определения ляпуновских показателей, обычно много меньше числа элементарных (гармонических) возбуждений, создающих течение. Этот факт допускает наглядную интерпретацию — хаотическая динамика течения определяется лишь небольшим числом независимых (в данном случае крупномасштабных) возбуждений, остальные же с ними жестко связаны, т. е. образуют когерентные структуры. Подобный же результат был получен ранее для турбулентности, описываемой двумерным уравнением Гинзбурга — Ландау [27]. Можно утверждать, что размерность хаотического движения — турбулентности — коррелирует с числом структур (спиральных волн, дефектов волновой решетки и т. д.), взаимодействующих друг с другом.

Чрезвычайный интерес представляет обнаружение связи между размерностью турбулентности при больших числах Рейнольдса и числом мод в инерционном интервале. Пока здесь имеются лишь единичные результаты (см., например, [28]), однако именно при анализе многомерных стохастических множеств, весьма своеобразно меняющихся при увеличении ширины инерционного интервала, возможно, удастся объяснить явление автомодельности турбулентности с точки зрения нелинейной динамики.

#### ЛИТЕРАТУРА

1. Handbook of turbulence / Ed. V. Frisch, and T. M. S. N. Y.: Plenum press, 1977. Vol. 1. Fundamentals and application. 420 p.
2. Ландау Л. Д. К проблеме турбулентности // ДАН СССР. 1944. Т. 44, № 4. С. 329—342.
3. Hopf E. A mathematical example displaying the features of turbulence // Commun. Pure and Appl. Math. 1948. Vol. 1. P. 302—322.
4. Newhouse S., Ruelle D., Takens F. Occurrence of strange attractors near quasiperiodic flow of  $T^m$ ,  $m \leq 3$  // Commun. Math. Phys. 1978. Vol. 64. P. 35.
5. Creveling H. F., de Paz J. F., Baladi J. Y., Schvenhals R. J. Stability characteristics of a single-phase free convection loop // J. Fluid Mech. 1965. Vol. 67, N 5. P. 65—77.
6. Lorenz E. Deterministic nonperiodic flow // J. Atmos. Sci. 1963. Vol. 20, N 1. P. 130—141.
7. Newhouse S., Ruelle D., Takens F. Occurrence of strange attractors near quasiperiodic flow // Commun. Math. Phys. 1978. Vol. 64. P. 35—49.
8. Осипов Г. В. О развитии турбулентности по Ландау в дискретной модели потсковых систем // Изв. вузов. Радиосвязь. 1988. Т. 31, № 5. С. 603—651.
9. Pomeau Y., Manneville P. Intermittent transition to turbulence in dissipative dynamical systems // Commun. Math. Phys. 1980. Vol. 74, P. 189—198.
10. Feigenbaum M. J. Quantitative universality for a class of nonlinear transformations // J. Stat. Phys. 1978. Vol. 19. P. 25—39.
11. Gollub J. P., Benson S. V. Many routes to turbulent convection // J. Fluid Mech. 1980. Vol. 100, N 3. P. 449—470.
12. Гапонов-Грехов А. В., Рабинович М. И. Нелинейная физика. Стохастичность и структуры // Физика XX века. Развитие и перспективы. М.: Наука, 1984. С. 219—280.

13. *Brandstätter A., Swift J., Swinney H. L. et al.* Low-dimensional chaos in a hydrodynamical system // *Phys. Rev. Lett.* 1983. Vol. 51, N 5. P. 1442—1445.
14. *Malraison B., Atten P., Berge P., Dubois M.* An experimental determination for the chaotic regime of two convective systems: Preprint. N. Y.: Springer, 1983. 760 p.
15. *Feigenbaum M. J.* Quantitative universality for a class of nonlinear transformations // *J. Stat. Phys.* 1978. Vol. 19. P. 25.
16. *Klinker T., Meyer-Ilse W., Lauterborn W.* Period doubling and chaotic behaviour in a driven Toda oscillator // *Phys. Lett. A.* 1984. Vol. 101. P. 371—375.
17. *Moore D. R., Toomre J., Kuobloch E., Weiss N. O.* Period doubling and chaos in partial differential equations for thermosolutal convection // *Nature.* 1983. Vol. 303. P. 663—667.
18. *Chia K. N., Osswald G. A., Chia V.* Simulation of self-induced unsteady motion in the near wave of a Joukowski airfoil // *Lect. Not. Eng.* 1986. Vol. 24. P. 118—127.
19. *Арансон И. С., Гапонов Грехов А. В., Рабинович М. И., Сагдеев Р. З.* Ренеточные модели в нелинейной динамике неравновесных сред: Препр. ИПФ АН СССР № 163. Гогький, 1987. 14 с.
20. *Grassberger P., Procaccia I.* Estimating the Kolmogorov entropy from a chaotic signal // *Phys. Rev. A.* 1983. Vol. 28. P. 2591—2600.
21. *Козлов В. В., Рабинович М. И., Рамазанов М. А.* и др. Корреляционная размерность потока и пространственное развитие динамического хаоса в пограничном слое // Письма в ЖТФ. 1987. Т. 15, № 16. С. 986—990.
22. *Libchaber A.* Form chaos to turbulence in Bénard convection // *Proc. Roy. Soc. London.* 1987. Vol. 413, N 1844. P. 63—69.
23. *Threlfall D. C.* Different ways to turbulence in dissipative dynamical system // *J. Fluid Mech.* 1975. Vol. 67. N 2. P. 219—226.
24. *Kobayashi R., Kohama Y.* Spiral vortices in boundary layer transition on a rotating cone // *Laminar-turbulent transition* / Ed. V. V. Kozlov. N. Y.: Springer, 1984. P. 573—576.
25. *Sreenivasan K. R., Meneveau C.* The fractal facets of turbulence // *J. Fluid Mech.* 1986. Vol. 173, N 9. P. 357—369.
26. *Grappin R., Léorat J.* Computation of the dimension of two-dimensional turbulence // *Phys. Rev. Lett.* 1987. Vol. 50, N 10. P. 1100—1103.
27. *Арансон И. С., Гапонов-Грехов А. В., Рабинович И. М.* Развитие хаоса в ансамблях динамических структур // ЖЭТФ. 1985. Т. 89, № 1(7). С. 92—106.
28. *Grappin R., Léorat J., Pouquet A.* Computation of the dimension of a model of fully developed turbulence // *J. Phys.* 1986. Vol. 47, N 11. P. 1127—1135.

УДК 530.182

## РОЖДЕНИЕ И ДИНАМИКА ДВУМЕРНЫХ СТРУКТУР В НЕРАВНОВЕСНЫХ ДИССИПАТИВНЫХ СРЕДАХ

А. В. ГАПОНОВ-ГРЕХОВ, А. С. ЛОМОВ, Г. В. ОСИПОВ,  
М. И. РАБИНОВИЧ

### 1. АТТРАКТОРЫ И ПРОСТРАНСТВЕННЫЕ ФОРМЫ

Одними из основных предметов исследования нелинейной динамики сегодня наряду с динамическим хаосом, солитонами и турбулентностью являются сложные пространственные структуры. Это, в частности, кристаллические и квазикристаллические формы, решетки с разного рода дефектами, фрактальные дендриты и т. д.

Обнаруженное сравнительно недавно единство динамических механизмов их зарождения и развития в неравновесных средах различной физической (и не только физической) природы позволяет сформулировать некоторые базовые уравнения нелинейной теории таких структур.

При не слишком больших надкритичностях движение среды (поля) во многих случаях по одной из пространственных координат можно считать заданным и для описания динамики поля по двум другим пространственным переменным получить так называемые амплитудные уравнения [1—4] или уравнения для параметра порядка [5, 6]. Подобные уравнения естественно разделяются на две группы: потенциальные, имеющие при  $t \rightarrow \infty$  лишь статические решения (так называемые градиентные модели), и такие, которые допускают периодические, квазипериодические или хаотические предельные (при  $t \rightarrow \infty$ ) решения.

Среди потенциальных простейшими, и в этом смысле наиболее каноническими, являются: нелинейное уравнение с диффузией для действительного параметра порядка

$$\partial a / \partial t = \mu a - a^3 + \Delta a, \quad (1)$$

аналогичное уравнение для комплексного параметра порядка

$$\partial a / \partial t = \mu a - |a|^2 a + \Delta a, \quad (2)$$

модель Свифта—Хоенберга [4], учитывающая дисперсию диффузии,

$$\partial a / \partial t = [\mu - (k_2^0 + \nabla^2)^2] a - a^3 \quad (3)$$

и уравнение Хакена [5]

$$\partial a / \partial t = [\mu - (k_0^2 + \nabla^2)^2] a + \beta a^2 - a^3, \quad (4)$$

в котором феноменологически учтена дополнительная квадратичная нелинейность, ответственная, в частности, за формирование многогранных решеток в двумерных средах. Физическая природа этого слагаемого может быть различной. Так, для тепловой конвекции оно связано с зависимостью вязкости от температуры [7] либо поверхностного натяжения от температуры (при конвекции Бенара—Марангони [8]). Уравнения (1) — (4), дополненные граничными условиями, в частности периодическими:

$$a(x, y, t) = a(x + L, y + L, t) \quad (5)$$

или нулевыми:

$$a(x, y)|_{\Gamma} = \nabla a(x, y)|_{\Gamma} = 0, \quad (6)$$

допускают запись в градиентной форме:

$$\partial a / \partial t = -\delta F / \delta a^*. \quad (7)$$

Вариационная производная берется от функционала  $F\{a, a^*\}$ , имеющего смысл свободной энергии. Например, для системы (3),

(6)  $F$  имеет вид

$$F = \int_{\Gamma} \left\{ -\frac{\mu}{2} a^2 + \frac{1}{4} a^4 + \frac{1}{2} [(k_0^2 + \nabla^2) a]^2 \right\} dx dy. \quad (8)$$

Состоянию равновесия (7) отвечает нулевая вариация  $F$ :  $\delta F|_{a=a^0(x,y)} = 0$ , устойчивость же равновесия определяется знаком второй вариации. Поскольку согласно (7), (8)

$$\frac{dF}{dt} = - \int_{\Gamma} \left| \frac{\partial a}{\partial t} \right|^2 dx dy, \quad (9)$$

то для любого нестационарного решения  $dF/dt$  отрицательно, и потому в системах градиентного типа устойчивыми являются лишь статические состояния. Пространственные формы, отвечающие этим состояниям, могут быть чрезвычайно разнообразны и в общем случае образуют континуальное множество, которому в фазовом пространстве отвечает континуум тривиальных аттракторов — состояний равновесия.

По мере увеличения надкритичности временное поведение пространственных структур в неравновесных средах существенно усложняется, и потенциальные модели для их описания уже недостаточны. Приведем здесь примеры «базовых» уравнений, которые описывают при  $t \rightarrow \infty$  не только статические, но и пульсирующие пространственные образования. Это обобщенное уравнение Гинзбурга—Ландау [2, 9, 10]:

$$\partial a / \partial t = \mu a - (1 + i\beta) |a|^2 a + \kappa (1 - ic) \Delta a, \quad (10)$$

подобные же уравнения для сред с параметрической неустойчивостью [11]:

$$\begin{aligned} \partial a_1 / \partial t &= -i\mu a_1^* - (1 + i\beta) |a_1|^2 a_1 + (1 - ic) \Delta a_1, \\ \partial a_2 / \partial t &= -i\mu a_2^* - (1 + i\beta) |a_2|^2 a_2 + (1 - ic) \Delta a_2 \end{aligned} \quad (11)$$

и неавтономное уравнение син-Гордона [12]

$$\partial^2 a / \partial t^2 - \Delta a + \sin a = F(x, y) \sin \omega t - \nu \partial a / \partial t. \quad (12)$$

Фазовое пространство систем (10) — (12) с граничными условиями (5) (или (6)) в общем случае организовано несравненно более сложным образом, чем фазовое пространство потенциальных систем. Оно уже может содержать и предельные циклы, и аттракторы в виде эргодической намотки на двух- или трехмерном торе, и странные аттракторы. Как мы увидим, странные аттракторы могут соответствовать и чисто временному хаосу в ансамбле небольшого числа регулярных пространственных структур. При этом пространственная картина весьма проста и повторяется через неравные отрезки времени (хаотическая динамика здесь только в этом и проявляется). Но возможна и принципиально иная ситуация — сложные пространственные формы, соответствующие решению на странном аттракторе, себя во времени не повторяют,



и при определенных условиях (если можно допустить, что среда безгранична) их следует рассматривать как неупорядоченные и во времени, и в пространстве. Это и есть пространственно-временной хаос.

## 2. МУЛЬТИСТАБИЛЬНОСТЬ. ДЕФЕКТЫ

Простейшие пространственные формы, возникающие в однородных диссипативных средах вблизи порога неустойчивости, — регулярные решетки с квадратными или шестигранными ячейками либо решетки в виде последовательности параллельных валов (роликов). Это коллективные возбуждения среды, которым соответствует минимальное число мод (собственных функций) линейной задачи: одномерная стоячая волна в случае валов, суперпозиция взаимоортогональных стоячих волн одинаковой длины волны в случае квадратов и, наконец, тройка синхронизованных по фазе стоячих волн, развернутых друг относительно друга на  $60^\circ$ , для шестигранников. Такие установившиеся структуры действительно легко описать, однако, когда мы задумываемся над путями их установления, возникают совсем нелегкие вопросы, например: как осуществляется отбор из случайных начальных условий именно данной решетки, почему при наличии ориентационной симметрии установилась решетка, расположенная именно так, почему одинаковые, но по-разному ориентированные решетки в одних случаях сосуществуют, а в других нет, и многие другие.

Обсудим эти проблемы на примере модели (4), которая описывает, в частности, конвекцию Бенара—Марангони в подогреваемом снизу плоском слое жидкости [13]. Граничные условия будем считать периодическими. В рассматриваемой системе предельные режимы всегда статические. Это следует из существования для системы (4), (5) (или (6)) функционала Ляпунова — свободной энергии — в виде

$$F(t) = \int \left\{ -\frac{\mu}{2} a^2 - \frac{\beta}{3} a^3 + \frac{1}{4} a^4 + \frac{1}{2} [(k_0^2 + \nabla^2) a]^2 \right\} dx dy. \quad (13)$$

Как показывает анализ этого функционала, при не слишком малых  $\mu$  он имеет множество локальных минимумов, отвечающих различным устойчивым пространственным формам (рис. 1). Абсолютный минимум реализуется на регулярной решетке из роликов (если  $\beta < \beta^*$ ) или шестигранников (если  $\beta > \beta^*$ ). Локальным минимумам отвечают решетки с различными дефектами (рис. 2, 3). Таким образом, для рассматриваемого класса неравновесных сред характерно явление мультистабильности. Реализуемость при  $t \rightarrow \infty$  той или иной пространственной формы зависит от начальных условий, а переход одной из устойчивых форм в другую определяется конечными флуктуациями.

Одним из основных механизмов установления той или иной решетки является механизм конкуренции структур. В простейшем варианте (регулярных структур) — это хорошо известный эффект

Рис. 1. Схематическое изображение зависимости свободной энергии от вида решения в конфигурационном пространстве для уравнений градиентного типа (1)–(4).

Иллюстрируется явление мультистабильности — в зависимости от начальных условий устанавливается пространственное распределение поля, отвечающее одному из минимумов функционала  $F$

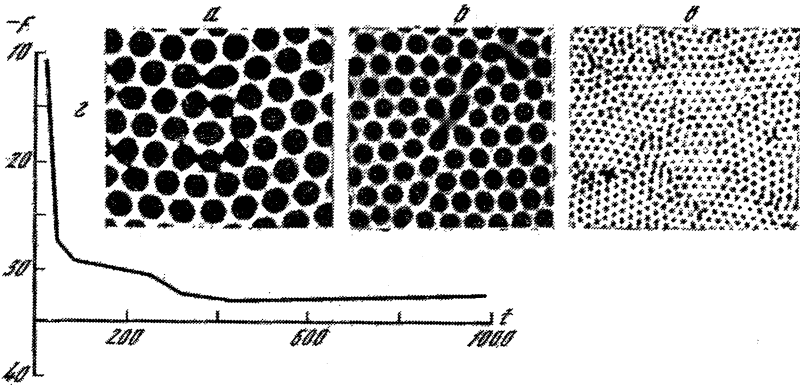
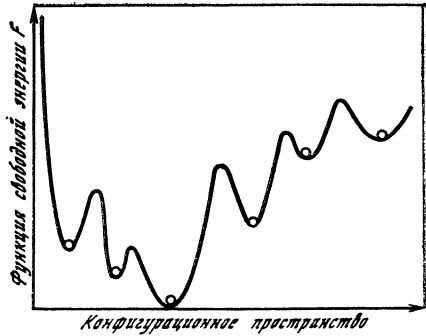


Рис. 2. Примеры стабильных решеток с дефектами, устанавливающихся при разных начальных условиях в системе (4), (5)

$a - e$  — решетки с размером области  $64 \times 64$  ( $a, b$ ) и  $256 \times 256$  ( $e$ );  $z$  — поведение функционала  $F(t)$ , соответствующее формированию из случайных начальных условий структуры  $a$ . Значения параметров:  $\mu = 0,05$ ,  $k_0 = 1$ ,  $\beta = 1$

конкуренции мод [14], описывающий взаимодействие во времени заданных пространственных распределений. Именно с помощью конкуренции пространственно-однородных мод, как кажется, проще всего объяснить установление устойчивой решетки в виде  $n$  параллельных роликов из начально заданной решетки, содержащей, скажем,  $(n + 1)$  ролик. Однако ни в численном эксперименте с системой (4), (5), ни в реальной слабонадкритической конвекции [15] подобного пространственно-однородного подавления одной моды другой не наблюдается. В действительности при взаимодействии близких мод энергетически более выгоден процесс плавной перестройки периода решетки в большей части пространства, сопровождающийся возникновением локального нарушения симметрии — дефекта, после исчезновения которого устанавливается устойчивая периодическая решетка (рис. 4).

Теорию подобных нестационарных дефектов возможно по-

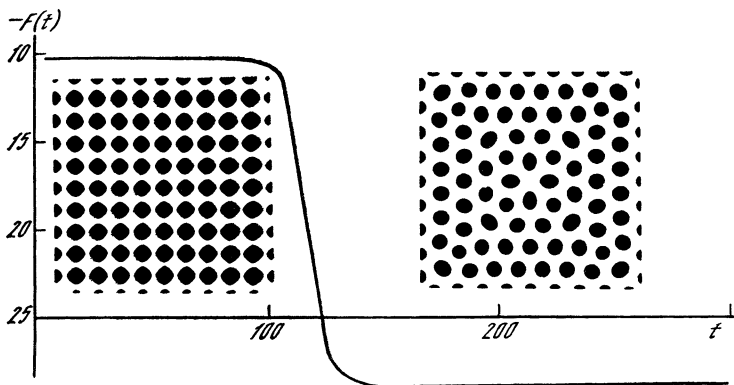


Рис. 3. Скачкообразный переход от метастабильной структуры — четырехгранной решетки — к устойчивой структуре из доменов шестигранной решетки

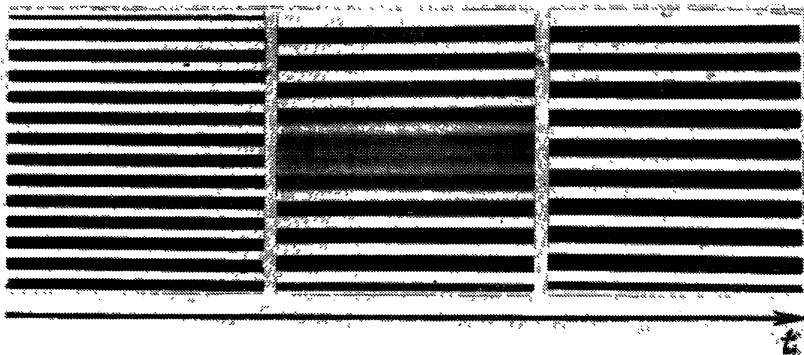


Рис. 4. Эффект пространственно-временной конкуренции роликов  
Численный эксперимент с системой (3), (5) (для устойчивой решетки  $F = -9,5$ )

Строить, если амплитуды взаимодействующих мод считать функциями не только времени, но и пространственных координат. С помощью такого описания возможно, в частности, объяснить и упоминавшийся эффект смены пространственной симметрии за счет рождения и исчезновения дефекта, и эффект сосуществования различных пространственных форм [16]. Заметим здесь, что граница между различными устойчивыми решетками — это протяженный дефект, которому в пространстве стационарных состояний соответствует сепаратриса (см. также [17]).

Пути формирования устойчивой регулярной структуры, очевидно, могут быть разными и зависят от начальных условий. На рис. 5, а представлены результаты численного эксперимента, иллюстрирующие множественность путей формирования шестигранной решетки. Всем этим маршрутам (1—4) соответствует

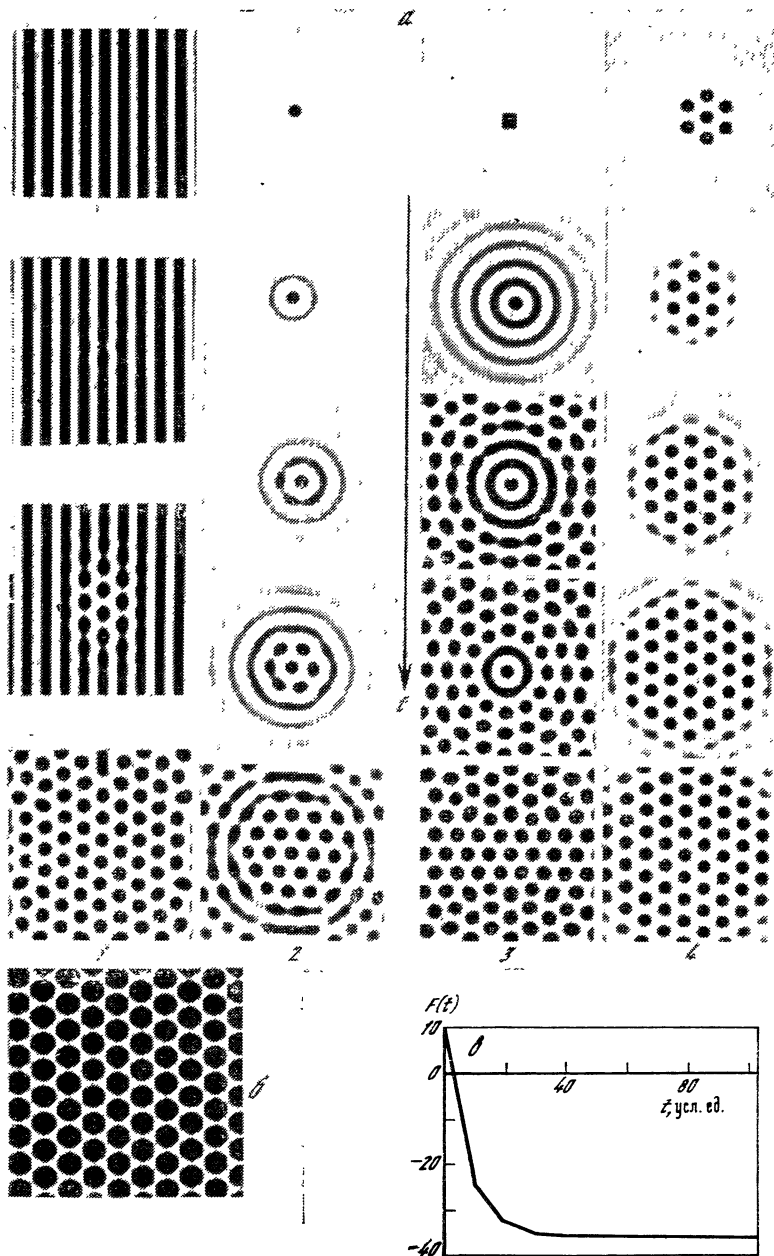


Рис. 5. Многообразие путей установления регулярной шестигранной решетки в модели (4), (5)

$a$  — разные маршруты формирования устойчивой решетки;  $b$  — конечное состояние с  $F_{\min} = -38,1$  для маршрутов 1 — 4;  $c$  — изменение свободной энергии системы при превращении линейной решетки в шестигранную для маршрута 1 на рис. 5,  $a$

переход системы в одно и то же состояние — с наименьшей свободной энергией ( $F_{\min} = -38,1$ ), однако пути достижения этого состояния отличаются. В некоторых случаях, например при возникновении соговой решетки из стоячей волны или близкого к точечному возмущения, формирование решетки шестигранников происходит за счет последовательного развития неустойчивостей — основной и затем модуляционной. Поясним это на примере развития шестигранной решетки из локализованного возмущения цилиндрической формы.

Линеаризация уравнения (4) дает уравнение для возмущений  $\tilde{u}(x, y, t)$  в виде ( $k_0 \equiv 1$ )

$$\tilde{u}_t = [\mu - (\Delta + 1)^2]\tilde{u}. \quad (14)$$

Его решение в цилиндрической системе координат при граничных условиях  $\tilde{u}|_R = 0$  имеет вид цилиндрических волн:

$$\tilde{u}(r, t) = e^{ct} \{A_1 J_0[(1 + \sqrt{\mu - c})^{1/2} r], \quad (15)$$

где  $J_0$  — функция Бесселя нулевого порядка, а постоянная  $c$  определяется из соотношения  $c = \mu - [(\varepsilon_j/R)^2 - 1]^2$ ;  $\varepsilon_j$  —  $j$ -й корень функции Бесселя. Нам интересны лишь нарастающие во времени решения вида (1), т. е. те, которым отвечают положительные  $c$ . Так, для  $R = 32$   $c > 0$  лишь для  $\varepsilon_{10}$  ( $c \approx 0,04$ ). Убедимся, что при достаточно большом  $\beta$  возникшие цилиндрические волны неустойчивы по отношению к азимутальным возмущениям  $\Phi(\varphi, t) \sim \exp[i(\omega t - k\varphi)]$ . Именно эта вторичная неустойчивость и приведет при  $t \rightarrow \infty$  к формированию шестигранной решетки при рассматриваемых начальных условиях.

Поясним задачу сразу в нелинейной постановке. Для этого представим решение уравнения (4) в виде цилиндрической волны, на которой заданы азимутальные возмущения

$$u(r, \varphi, t) = A_1(t)J_0(k_0 r) + A_3(t)J_3(k_3 r) \sin 3\varphi. \quad (16)$$

Для удовлетворения краевых условий необходимо, чтобы  $J_0(k_0 r_0) = 0$ ,  $J_3(k_3 r_0) = 0$ . Вблизи границы неустойчивости цилиндрической волны  $A_1(t)$  и  $A_3(t)$  можно считать медленными функциями времени. Тогда, подставляя выражение (16) в исходное уравнение (4) и пользуясь асимптотическим методом, можно получить для  $A_1$  и  $A_3$  уравнения вида (21) (см. ниже), где квадратичные слагаемые  $A_1^2$  и  $A_1 A_3$  пропорциональны параметру  $\beta$ . Даже при наличии декремента в моде  $A_3$  азимутальные волны при достаточно большом  $\beta$  будут нарастать, и при увеличении  $t$  возникнет шестигранник (см. маршруты 2 и 3 на рис. 5, а).

По своему характеру нарушения симметрии двумерных решеток, т. е. образование дефектов, естественно разделить на две группы: глобальные, приводящие к изменению картины поля во всем пространстве, и точечные (или линейные) — локальные нарушения симметрии (соответственно по обеим или одной из координат) однородных изначально решеток. Первые могут быть ре-

зультатом взаимодействия, скажем, коллективных возбуждений с несоизмеримыми масштабами и допускают описание на языке нелинейной динамики мод (см., например, [18] и рис. 6). Вторые представляют собой топологические особенности двумерного поля и рождаются и исчезают подобно тому, как рождаются и исчезают состояния равновесия (или их континуумы) на фазовой плоскости при локальных бифуркациях (см. рис. 2—4).

Статические, в частности топологические, характеристики дефектов, возникающих на фоне регулярных решеток в различных течениях жидкости (рис. 7), не отличаются от соответствующих характеристик дефектов, например в теории кристаллов, где они подробно изучены [21]. Специфика неравновесных сред здесь состоит в том, что топологические дефекты поля непрерывно движутся, рождаются и умирают, т. е. для полей в неравновесных средах основной интерес представляет нелинейная динамика дефектов. В частности, турбулентность в таких средах может в определенной области надкритичностей рассматриваться как хаотическая динамика дефектов (см. также разд. 5, рис. 7, *в* и рис. 8).

Для потенциальных систем динамика дефектов исследована наиболее подробно. Процесс движения дефектов в таких системах всегда завершается установлением одного из стационарных состояний, которому соответствует локальный минимум функционала  $F$  (дефекты при этом либо останавливаются, либо исчезают). Изменение свободной энергии  $F$  в этом случае можно рассматривать как работу некоторой силы, которую по аналогии с силой, действующей на дислокацию в кристалле при наложении давления, называют силой Peach—Köhler [22]. Для простых дислокаций, в частности представляющих собой переход от роликовой структуры с волновым числом  $2(n+1)\pi/Lx$  к решетке с волновым числом  $2n\pi/L$  (подобно рис. 4, *а*), эту силу легко вычислить [22]. Изменение свободной энергии  $F$  в рассматриваемом случае есть  $dF = \frac{dF}{dk} \frac{2\pi}{Lx} Lx dy$ , т. е. сила, действующая на дислокацию,  $f_{p-k} = dF/dy = 2\pi dF/dk$ . Движение дефекта, очевидно, будет продолжаться до тех пор, пока сила  $f_{p-k}$  не сравняется с действующей на дефект силой трения  $f_D$ , зависящей от величины диффузии в неравновесной среде [22]. Таким образом, движение дефекта подобно движению шарика по бугристому дну ванны, заполненной вязкой жидкостью.

### 3. ЛОКАЛИЗОВАННЫЕ СТРУКТУРЫ

Для иллюстрации принципиальной возможности самозарождения и устойчивого существования нетривиальных локализованных структур в однородных изотропных средах воспользуемся моделью, предложенной в [23]. Эта модель предполагает взаимодействие двух полей (двух компонент) в неравновесной среде и содержит два уравнения<sup>1</sup>:

<sup>1</sup> Можно считать, к примеру, что  $a$  и  $V$  — это поля температуры в двух несмешивающихся горизонтальных слоях жидкости, расположенных один на другом.

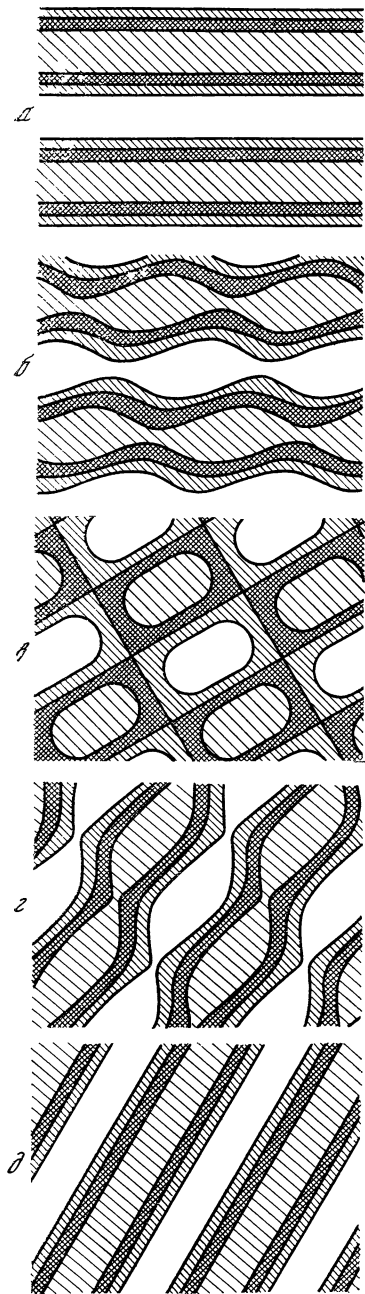


Рис. 6

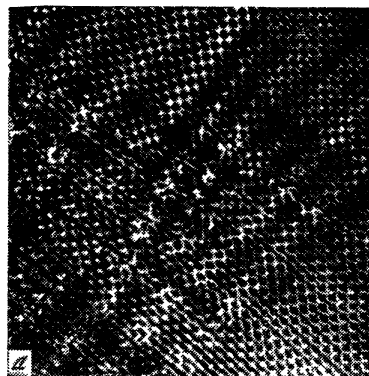


Рис. 7

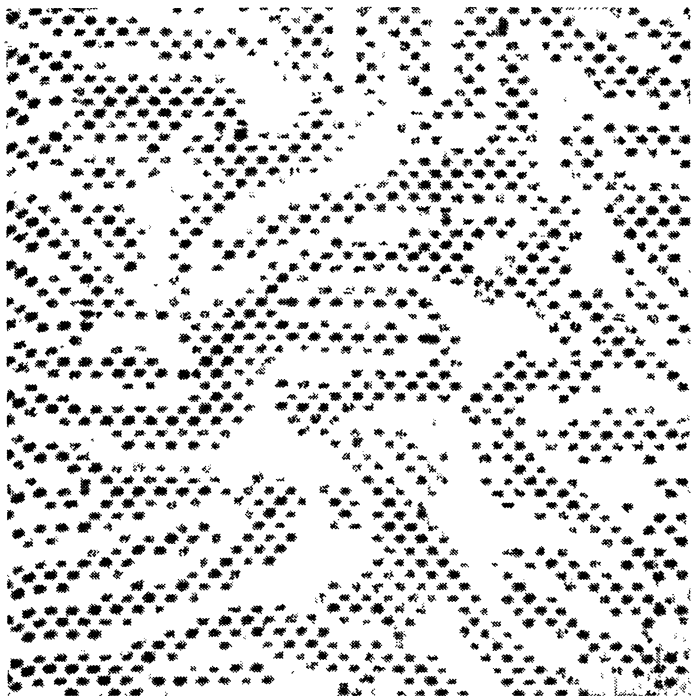


Рис. 8. Дислокационный хаос в дискретной модели уравнения (10)  
Решетка  $64 \times 64$ ,  $\beta = c = 10$ ,  $\kappa = 0,2$

$$\partial a / \partial t = [(V - a) - (1 + \nabla^2)^2]a + \beta a^2 - a^3, \quad (17)$$

$$\partial V / \partial t = (1/\mu)(V - \gamma V^3 + \delta a + D\nabla^2 V), \quad \mu \ll 1. \quad (18)$$

При  $V = 0$  и  $\alpha < 0$  уравнение (17) — это уравнение (4). Второе же уравнение при  $\delta = 0$  совпадает с уравнением (1). Слагаемые  $Va$  в (17) и  $\delta a$  в (18) определяют взаимодействие компонент.

Как показывает анализ [23], система (17), (18) содержит локализованные в пространстве образования при произвольных  $\beta$ , однако нарушение полной азимутальной симметрии происходит лишь при  $\beta > 1$ . При малых же  $\beta$  локализованные структуры имеют форму дисков, характерный размер и стационарная интен-

Рис. 6. Пример взаимных трансформаций периодических пространственных решеток [18]

$a - t=0$ ;  $b - T/16$ ;  $c - 2T/16$ ;  $d - 3T/16$ ;  $e - T/4$

Рис. 7. Дефекты в реальных гидродинамических течениях

$a$  — дислокации в волн медуляции на фоне квадратной решетки ряби; Фарадея [11];  $b$  — дислокационные пары в релаксивной конвекции при периодическом подогреве снизу [19];  $c$  — динамика дефектов в конвективном слое [20]



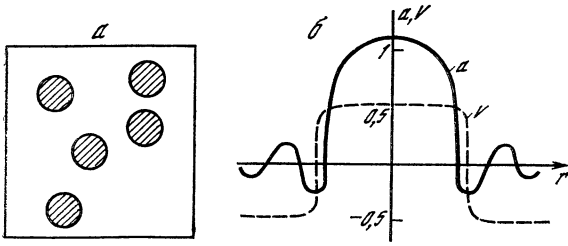


Рис. 9. Локализованные структуры в виде дисков в модели (17), (18) ( $\beta = 0,9$ ,  $\gamma = 4$ ,  $\delta = 0,15$ ,  $D = 0,3$ ) [23] (а) и распределение поля  $a$  и  $V$  для отдельного диска (б)

сивность которых определяются только параметрами среды и не зависят от начальных и граничных условий (рис. 9, а).

Число и взаимное расположение локализованных структур, возникающих в достаточно протяженной двумерной среде, описываемой (17), (18), определяются случайными начальными условиями. Однако они не могут сближаться более чем на масштаб подавления  $l$ , который соответствует характерному размеру пограничной с автоструктурой области, где поле  $a$  отрицательно (см. рис. 9, б). Наличие такой области определяется особенностями диффузии (слагаемые  $2\nabla^2 a + \nabla^4 a$  в (17)). В области подавления неравновесная среда не возбуждается, что и гарантирует устойчивость эффекта самолокализации двумерного поля  $a$ ,  $V$ .

Нетривиальность формы локализованных структур определяется разнообразием линейных возбуждений, которые служат затравкой для последующего нелинейного роста и формирования автоструктуры. Простейшие нетривиальные структуры с центром симметрии — это правильные многогранники. Они получаются в результате взаимодействия лишь двух мод круглой мембраны — радиальной и азимутальной. В частности, автоструктуры в виде уединенных шестигранников и восьмигранников, наблюдавшиеся в [24], можно рассматривать как результат взаимодействия именно таких мод. Согласно [24], определяющую роль в рождении уединенных многогранников играет нелинейность, связанная с зависимостью поверхностного натяжения от температуры. В модели (17), (18) подобная нелинейность учтена слагаемым  $\sim \beta a^2$ . Такая нелинейность обеспечивает устойчивость совместной генерации мод, которые при малых  $\beta$  конкурируют друг с другом.

Высказанные соображения говорят в пользу того, что в однородных неравновесных средах, описываемых уравнениями типа (17), (18), действительно возможно самозарождение и устойчивое существование локализованных многогранников. Численные эксперименты это доказывают.

Для аналитического описания слабонадкритических локализованных структур в ситуации, приближенной к эксперименту [24], воспользуемся уравнениями Буссинеска

$$\partial U / \partial t + (U \nabla) U = \text{Pr} (\theta z_0 + \nabla^2 U),$$

$$\partial \theta / \partial t + (U \nabla) \theta = \text{Ra} (U z_0) + \nabla^2 \theta, \quad \nabla U = 0.$$

Здесь  $\theta(r; t)$  — отклонение температуры от равновесного состояния;  $Pr = \nu/\kappa$  — число Прандтля;  $Ra = (\Delta T g \beta / \nu \kappa) d^3$  — число Рэлея ( $d$  — толщина слоя,  $\Delta T$  — разность температур нижней и верхней поверхностей,  $\nu$  — вязкость,  $\kappa$  — температуропроводность). При больших числах Прандтля ( $Pr \gg 1$ ) и малых надкритичностях ( $|Ra - Ra_c| \ll Ra_c$ ,  $Ra_c$  — значение  $Ra$ , при котором возникает конвективное движение) эти уравнения могут быть приведены к виду (4), в котором переменная  $a$  имеет смысл отклонения температуры  $\theta$ , а квадратичное слагаемое, как отмечалось, связано с зависимостью капиллярного натяжения (или вязкости) от температуры. То же уравнение (4) получается из модели (17), (18), если предположить, что поле  $V$  неоднородно и не меняется во времени, однако в этом случае надкритичность ( $V - \alpha$ ) должна зависеть от  $r$ . Покажем, что в такой модели возможно существование устойчивых пространственно локализованных структур, формы которых не связаны с формой неоднородности  $\alpha'(r) = V - \alpha$  [25].

Будем считать неоднородность  $\alpha'(r)$  радиально симметричной:

$$\alpha'(r) = \begin{cases} \alpha_0, & r \leq r_0, \\ 0, & r > r_0. \end{cases} \quad (19)$$

В результате мы приходим к уравнению

$$\partial a / \partial t = [\alpha'(r) - (1 + \nabla^2)^2]a + \beta a^2 - a^3, \quad (20)$$

где  $\alpha'(r)$  определяется (19). Считая формально нелинейность слабой ( $a^2 \ll 1$ ), будем отыскивать решение (19), (20) в виде суперпозиции собственных функций линеаризованной вблизи  $a = 0$  задачи, решение которой представим в виде

$$a(r, t) = \sum_n F_n(r) \sin(n\varphi) \exp(\lambda_n t),$$

где функции  $F_n(r)$  удовлетворяют условиям  $F_n(r) \rightarrow 0$ ,  $r \rightarrow \infty$ ;  $F_n(r) < \infty$ ,  $r \rightarrow 0$ . Собственные функции  $F_n(r)$  можно найти, сшивая решения уравнения Лапласа на границе  $r = r_0$ . Однако представляют интерес лишь те функции, которым соответствуют нарастающие во времени возмущения, т. е. с  $\lambda_n > 0$ . Вне области положительного  $\alpha'(r > r_0)$  этому условию могут удовлетворять только функции Неймана от комплексного аргумента  $K_n(ikr)$  ( $\bar{k}^2 = -1 + iq$ ), а внутри области  $\alpha' > 0$  — функции Бесселя действительного аргумента  $J_n(kr)$ . Величина  $q$ ,  $k$  в силу действительности  $\lambda_n$  должны удовлетворять соотношению  $\lambda_n = \alpha_0 - (1 - k^2)^2 = q^2 > 0$  и определяются из дисперсионного уравнения  $J_n(kr_0)/kJ'_n(kr_0) = \text{Re}(K_n(ikr_0))/\text{Re}(i\bar{k}k'(i\bar{k}r_0))$ . Это уравнение следует из условий непрерывности  $F_n$  и  $dF_n/dr$  на границе  $r = r_0$ .

Предположим теперь, что параметры  $r_0$ ,  $\alpha_0$  таковы, что несколько показателей  $\lambda_n$  (соответствующих, например, осесимметричному с  $n = 0$  и азимутальному с  $n = 3$  возмущениям) положительны. Тогда, подставляя в (20) решение в виде  $a(r, \varphi, t) =$

$= A_0(t)F_0(r) + A_3(t)F_3(r) \sin(3\varphi)$  и считая амплитуды возмущений медленными функциями времени, из условий ортогональности к собственным функциям сопряженной системы получим для  $A_0$ ,  $A_3$  систему уравнений второго порядка [23]

$$\begin{aligned} dA_0/dt &= \lambda_0 A_0 + C_1 A_0^2 + C_2 A_3^3 - C_3 A_0^3 C_4 A_0 A_3^2, \\ dA_3/dt &= \lambda_3 A_3 + d_1 A_0 A_3 - d_3 A_3^3 - d_3 A_3 A_0^2. \end{aligned} \quad (21)$$

При  $C_1$ ,  $d_1 \neq 0$  (т. е. при наличии квадратичной нелинейности в (20)) эта система имеет устойчивое стационарное решение с  $A_0$ ,  $A_3 \neq 0$ , которому и соответствует уединенная шестигранная ячейка. В тех случаях, когда число таких взаимодействующих мод больше двух, уравнения для амплитуд этих мод могут иметь устойчивые движения с нетривиальной (в том числе и хаотической) динамикой. Такая ситуация, по-видимому, описывает наблюдавшийся в эксперименте [24] режим неперiodического рождения и исчезновения ячеек.

Подобные решения в виде уединенной ячейки могут быть получены и для однородной, но двухкомпонентной среды, описываемой системой (17), (18).

#### 4. РОЖДЕНИЕ СПИРАЛЬНЫХ ВОЛН

Мы убедились, что в слабонеоднородных и даже в однородных средах возможно существование устойчивых локализованных образований универсальной формы — автоструктур. Исследованные автоструктуры обладали определенной «азимутальной» симметрией (инвариантностью по отношению к повороту на заданный угол). В то же время, как известно, в однородных изотропных средах возможны и существенно анизотропные автоструктуры, в частности вращающиеся спиральные вихри. Подобные структуры столь часто встречаются в экспериментах, что возникает ощущение их элементарности и универсальности, не связанной с физической природой неравновесной среды. В чем же природа этой универсальности и каковы механизмы, приводящие к формированию спиральных структур? Можно показать [26], что универсальность спиральных образований связана с весьма общими топологическими свойствами однопараметрических семейств функций  $H_t(x, y) = \text{const}$ , описывающих пространственные образы (мгновенные снимки) поля в двумерных неравновесных средах независимо от их природы. В этом разделе на примере сред, описываемых двумерным уравнением Гинзбурга—Ландау, обсуждается механизм рождения спиралей. Оказывается, такие структуры развиваются при увеличении степени неравновесности среды в результате иерархии пространственных неустойчивостей с различным типом симметрии.

В известных экспериментах (см., например, [20]) доступные прямому наблюдению структуры двумерного поля выделены контрастом либо цветом. В частности, концентрационные структуры,

наблюдаемые в двумерном реакторе, где протекает автокаталитическая химическая реакция, представляют собой именно цветовые «узоры» [27]. Для определенности ограничимся рассмотрением подобных структур, полагая, что кинетика реакции определяется взаимодействием всего двух компонент:

$$\begin{aligned} \frac{\partial n_1}{\partial t} &= n_1 - (n_1^2 + n_2^2)(n_1 - \beta n_2) + \kappa(\nabla^2 n_1 + c\nabla^2 n_2), \\ \frac{\partial n_2}{\partial t} &= n_2 - (n_1^2 + n_2^2)(n_2 + \beta n_1) + \kappa(\nabla^2 n_2 - c\nabla^2 n_1), \end{aligned} \quad (22)$$

где  $n_1(x, y, t)$ ,  $n_2(x, y, t)$  — концентрации этих компонент, а  $\kappa$  характеризует величину диффузии. Система (22) представляет собой вариант известной  $\lambda - \omega$ -системы, используемой для моделирования нелинейной динамики различных неравновесных сред [9].

Распределение цвета в концентрационных структурах будет определяться наложением цветов индивидуальных компонент. Поскольку топология структур (взаиморасположение линий постоянного цвета) должна быть инвариантна относительно пропорционального изменения насыщенности любого из цветов, задача об исследовании структур сводится к анализу линий уровня функции

$$n_1(x, y, t)/n_2(x, y, t) = H_t(x, y) = \text{const.}$$

Нас будут интересовать такие конфигурации поля  $H_t(x, y)$ , которые невозможно перевести непрерывной деформацией в пространственно-однородные. Подобные конфигурации определяются наличием у поля  $H_t(x, y)$  особых точек, в которых поле не определено, «сепаратрис» и т. д., которые и соответствуют элементарным структурам в виде спиралей, вихрей, игл и проч. Бифуркациям — превращениям одних структур в другие — отвечает изменение топологии линий уровня  $H_t(x, y)$ , определяемое системой (22) при изменении либо параметра, либо времени. Найдем такие изменения, используя в качестве параметра величину диффузии  $\kappa$ .

Заменой  $a = n_1 + in_2$  система (22) сводится к уравнению Гинзбурга—Ландау (10). Интересующая нас функция  $H_t(x, y)$  связана с фазой  $\varphi$  комплексного поля  $a = |a| \exp(i\varphi)$  соотношением  $H_t(x, y) = \text{arctg } \varphi(x, y, t)$ . Исследуем полевые структуры, определяемые (10), при периодических граничных условиях (5). Кроме того, учтем, что нетривиальные решения (10) возможны лишь при  $\beta c > 1$ .

Как известно [9, 28], при значениях диффузии  $\kappa > \kappa_0 = (\beta c - 1)L^2/\pi^2(1 + c^2)$  и при  $\beta c > 1$  в рассматриваемой среде устойчив лишь пространственно-однородный режим колебаний:  $a(x, y, t) = \exp(-i\beta t)$ . При  $\kappa < \kappa_0$  пространственно-однородное решение неустойчиво и возмущения  $\zeta$  на его фоне описываются уравнением

$$\begin{aligned} \frac{\partial \zeta}{\partial t} &= -(1 + i\beta)(\zeta + \zeta^* + |\zeta|^2\zeta + 2|\zeta|^2 + \zeta^2) + \kappa(1 - \\ &- i\kappa)\nabla^2 \zeta, \end{aligned} \quad (23)$$

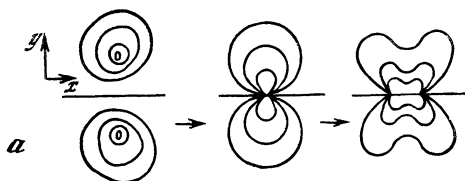
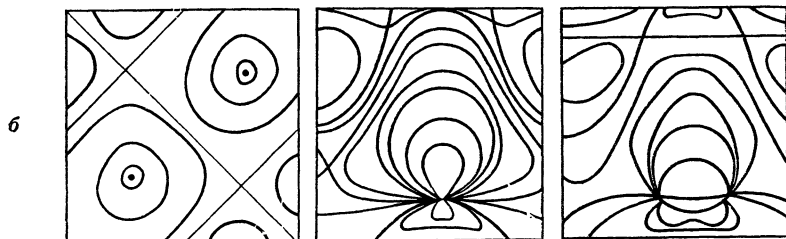


Рис. 10. Рождение спиральных пар по мере увеличения надкритичности (уменьшения диффузии, слева направо) в рамках модели (10)

*a* — изменение линий уровня функции  $H(x, y)$ ; *б* — результаты численного эксперимента [29]



которое связано с (10) заменой  $a(x, y, t) = [1 + \zeta(x, y, t)] \exp(-i\beta t)$ . При  $\kappa < \kappa_0$  равновесное состояние (23)  $\zeta = 0$  неустойчиво по отношению к возмущениям вида

$$\zeta \sim \exp(\pm ik_0 x), \quad \zeta \sim \exp(\pm ik_0 y), \quad (24)$$

где  $k_0 = 2\pi/L$ . В общем случае ляпуновские показатели возмущений вида  $\zeta \sim \exp(ik_n x + ik_m y)$ ,  $k_j = 2\pi j/L$  ( $j = m, n$ ) определяются выражением

$$\lambda_{m,n} = -1 - \kappa(k_m^2 + k_n^2) \pm D_{m,n},$$

$$D_{m,n}^2 = 1 - [\kappa(k_m^2 + k_n^2)]^2 + 2\kappa\beta c(k_m^2 + k_n^2),$$

из которого следует, что при  $\kappa < \kappa_0$  четыре показателя, соответствующие возмущениям вида (24), становятся положительными. Тогда, отыскивая при  $|\kappa - \kappa_0| \ll \kappa_0$  решение (23) в виде

$$\zeta = (1 + iB_1)[C_1 \exp(ik_0 x) + C_2 \exp(ik_0 y) + \text{к. с.}] \quad (25)$$

и считая функции  $C_{1,2}$  медленными, с помощью асимптотического метода получаем уравнения для амплитуд вторичных возбуждений:

$$dC_{1,2}/dt = C_{1,2} [\lambda_1 - 3p(|C_{1,2}|^2 + 2|C_{2,1}|^2)], \quad (26)$$

$\lambda_1 = -1 - \kappa k_0^2 + [1 - (\kappa k_0^2)^2 + 2\kappa\beta c k_0^2]^{1/2}$  — ляпуновский показатель для возмущений вида (24). Величина  $B_1$  в (25) — компонента собственного вектора линеаризованной системы (23);  $B_1^+$  — то же для системы, сопряженной к (23)<sup>2</sup>;  $p = (1 + B_1^2)[1 + \beta(1 + B_1 B_1^+)]/(B_1 - B_1^+)$ . Функция  $H_t(x, y)$  для (26) при  $C_{1,2} \neq 0$  не имеет особенностей, и это решение описывает простейшую структуру в виде периодической решетки «вихрей» (см. рис. 10, *a*).

<sup>2</sup>  $B_1 = (D_1 - 1)/(4\kappa c k_0^2)$ ;  $B_1^+ = -(D_1 + 1)/(4\kappa c k_0^2)$ ;  $D_1^2 = [1 - (\kappa c k_0^2)^2 + 2\kappa\beta c k_0^2]$ .

При дальнейшем увеличении надкритичности (уменьшении  $\kappa$ ), в свою очередь, становится неустойчивым и решение вида (24) — возбуждаются еще четыре волны с волновыми векторами, расположенными под углом  $45^\circ$  к исходным. Это следует из (23): при уменьшении  $\kappa$  еще четыре показателя приближаются к нулю. Соответствующее восьмимодовое решение при этом имеет вид

$$\zeta(x, y, t) = (1 + iB_1)[C_1 \exp(ik_0x) + C_2 \exp(ik_0y) + \text{к.с.}] + \\ + (1 + iB_2)\{C_3 \exp[ik_0(x + y)] + C_4 \exp[ik_0(x - y)] + \\ + \text{к.с.}\},$$

где  $C_{1,2,3,4}$  — медленные комплексные функции времени (по-прежнему  $|\kappa - \kappa_0| \ll \kappa_0$ ), удовлетворяющие системе уравнений

$$dC_{1,2,3,4}/dt = f_{1,2,3,4}(C_1, C_2, C_3, C_4). \quad (27)$$

Можно убедиться, что эта система имеет частное решение вида  $C_1 = C_2 = \text{const}$ ;  $C_3 = C_4 = \text{const}$ , которому в исходных переменных соответствует решение

$$a(x, y, t) = e^{-i\beta t} \{1 + 2C_1(1 + iB_1)[\sin(k_0x) + \sin(k_0y)] + \\ + 4C_3(1 + iB_2)\sin(k_0x)\cos(k_0y)\}. \quad (28)$$

Фазовая функция этого решения имеет вид  $\varphi(x, y, t) = -\beta t + \arctg H_t(x, y)$ , где

$$H_t(x, y) = \frac{2C_1B_1[\sin(k_0x) + \sin(k_0y)] + 2C_3B_2\sin(k_0x)\cos(k_0y)}{1 + C_1[\sin(k_0x) + \sin(k_0y)] + 2C_3\sin(k_0x)\cos(k_0y)}.$$

При выполнении условий  $|B_2/(B_1 - B_2)C_1| \ll 8$ ;  $|B_1/(B_1 - B_2)C_3| \ll 8$  функция  $H_t(x, y)$  в общем случае имеет особенности, т. е. мгновенный снимок, описываемый (28), будет содержать новые по сравнению с (24) элементы структуры.

Вид этих новых элементов определяется поведением линий уровня скалярного поля  $H_t(x, y)$  в малой окрестности особенности. Общие топологические свойства интересующих нас конфигураций изучаются геометрической теорией слоений [30]. Из этой теории, в частности, следует, что в нашем случае (в силу сохранения индекса) ввиду аналитичности функций  $n_{1,2}(x, y, t)$  наиболее типичная бифуркация — это превращение при изменении  $\kappa$  пары вихрей в пару узлов (или спиралей).

Детальный анализ решения (28) показывает, что при уменьшении  $\kappa$  в слоении  $H_t(x, y)$  действительно появляется пара узлов. Они и соответствуют спиральным вихрям. С помощью (28) удастся также проследить за бифуркациями рождения спиралей (см. рис. 10, б).

## 5. СТРУКТУРЫ ВО ВНЕШНИХ ПОЛЯХ

Обсудим здесь некоторые проблемы, связанные с формированием структур в двумерных неравновесных средах с помощью неоднородных внешних полей. Это интересно с разных точек зрения, в частности для управления тепло- и массообменом при термо-

конвекции, средним током в двумерных джоузефсоновских контактах и др. Есть здесь и совершенно необычные задачи, например формирование структур с помощью полей сложной формы — пространственных образов, которые воздействуют на среду лишь ограниченное время  $\tau_0$  (время обучения). При этом в среде может быть сформирована такая функция  $F$ , что данной структуре внешнего образа отвечает локальный минимум  $F$  (аттрактор) и такую среду можно использовать для восстановления исходной картины поля (распознавание). Сюда же примыкают идеи моделирования ассоциативной памяти с помощью неравновесных сред, основанные на эффекте мультстабильности в таких средах.

Сознательное с целью управления воздействие внешним полем на пространственные структуры в эксперименте удается реализовать лишь в весьма ограниченной области параметров, вне которой основную роль играют собственные неустойчивости неравновесной среды. Ситуация, когда устанавливающиеся при  $t \rightarrow \infty$  структуры не повторяют структуру внешнего поля, обнаруживается уже в одномерных задачах — это, в частности, хорошо известный переход соизмеримость—несоизмеримость [30], когда период возникшей структуры не кратен пространственному периоду приложенного поля. В двумерных задачах проявление этого эффекта может быть более разнообразно. Например, ввиду зависимости от начальных условий несоизмеримость может реализоваться лишь по одной из координат. Этот эффект наблюдался, например, при исследовании диссипативного уравнения  $\sin$ -Гордона с периодической правой частью (см. (12)).

Для пространственно-изотропных диссипативных сред естественно ожидать определенной «несимметрии» в поведении поля во времени и пространстве. В частности, пространственная картина поля  $\varphi(x, y, t)$  может не повторять структуру внешнего поля, хотя временное поведение будет полностью синхронизованным. Такой эффект наблюдался для системы (12) при  $\Phi(x, y) = A \sin x \sin y$ , где  $A \approx 10$ . Обнаруживаемые при этом перестройки пространственной картины поля, как правило, представляют собой удвоения (утроения или учетверения) элементарных структур, что демонстрирует рис. 11.

Добавим, что в области параметров, где реализуется режим временной синхронизации, возможен не только переход соизмеримость — несоизмеримость, но и переход несоизмеримость—пространственный хаос и соизмеримость—хаос. Соответствующие режимы удалось, в частности, наблюдать при компьютерном моделировании неавтономного уравнения  $\sin$ -Гордона с периодически неоднородными параметрами:

$$\begin{aligned} \partial^2 \varphi / \partial t + \Lambda \partial \varphi / \partial t - (\partial^2 \varphi / \partial x^2 + \partial^2 \varphi / \partial y^2) + \sin kx \sin ky \sin \varphi = \\ = A \sin \omega t. \end{aligned} \quad (29)$$

Обсуждаемые «несимметричные» режимы, однако, являются весьма экзотическими. Гораздо более типичными оказываются ситуации, для которых пространственный хаос наблюдается в той

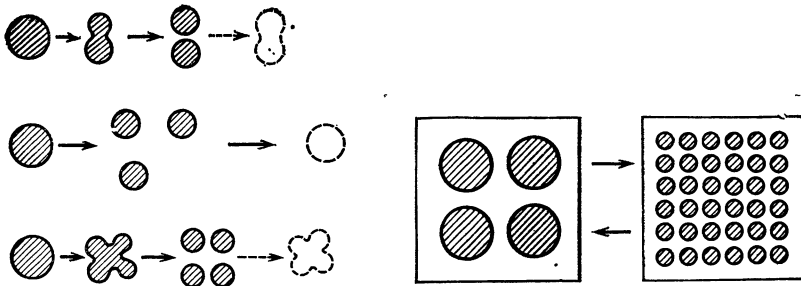


Рис. 11. Удвоение, утроение и учетверение элементарных структур в модели (12), (5) при  $\Phi(x, y) = A \sin x \sin y$

Рис. 12. Регулярная пространственно-временная динамика системы (12), (5) при  $\Phi(x, y) = A \sin x \sin y$  ( $A = 5,0$ )

же области параметров, что и временной, а временная регулярность влечет за собой пространственную (см. пример на рис. 12).

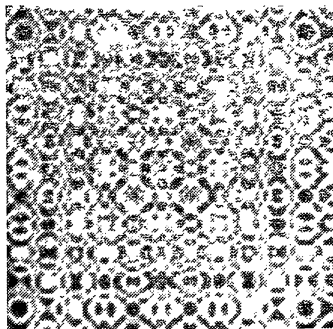
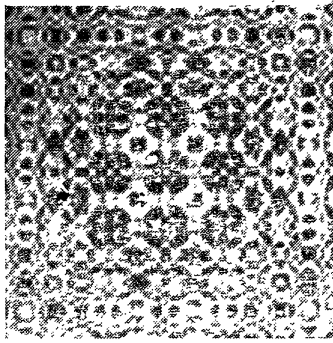
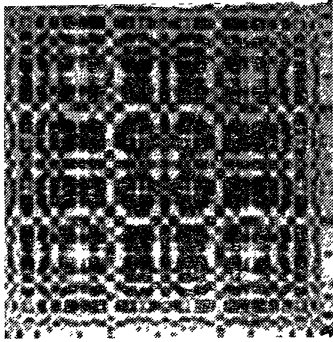
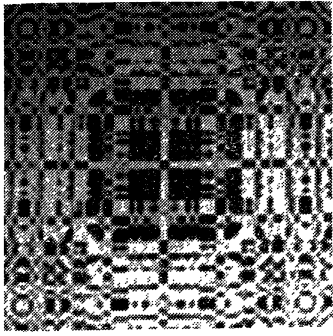
Заметим, что динамику простых пространственных распределений (как и для автономных систем) естественно описывать в рамках модовых моделей типа (21) или (27)<sup>3</sup>. Например, режим периодического во времени с периодом  $T$  удвоения и затем «деления на два» пространственного периода (см. рис. 11) предполагает в простейшей постановке взаимодействие лишь двух мод, резонансно связанных друг с другом внешним полем. В рамках подобных модовых моделей возможно описать большинство наблюдавшихся в компьютерном эксперименте эффектов, в том числе и эффект «структурной перемежаемости» — основную часть периода наблюдается решетка, например из девяти ячеек, сменяющихся на краткое время решеткой, повторяющей структуру внешнего поля.

Картина пространственно-временного хаоса, индуцируемого внешним полем, очевидно, зависит от числа элементарных возбуждений среды. При не слишком большой надкритичности, в частности для модели (29), такой хаос выглядит как случайное блуждание нескольких локализованных структур, слабо взаимодействующих друг с другом. С ростом амплитуды внешнего поля их число возрастает, а размерность соответствующего стохастического множества увеличивается (аналогично см. [28]).

Подчеркнем здесь, что компьютерное исследование многомерного пространственно-временного хаоса представляет весьма сложную проблему, в частности по той причине, что принципиальной может оказаться дискретность численной сетки, когда на первый план выступают неустойчивости, связанные с взаимодействием соседних ее элементов. Хаос в подобных решеточных [31] моделях оказывается весьма специфическим. Поясним это на примере двумерной решетки из затухающих маятников, возбуждаемых

<sup>3</sup> Такое описание оправданно, когда амплитуда невелика и  $\sin \varphi$  можно разложить в ряд, учитывая лишь основную нелинейность  $\sim \varphi^3$ .





внешним полем:

$$\begin{aligned} \partial^2 \varphi_{j,k} / \partial t^2 + \Lambda \partial \varphi_{j,k} / \partial t + \sin \varphi_{j,k} = & \\ = A \sin kx \sin ky \sin \omega t + & \\ + (1/h^2) (\varphi_{j-1,k} + \varphi_{j+1,k} + & \\ + \varphi_{j,k-1} + \varphi_{j,k+1} - 4\varphi_{j,k}), & \quad (30) \\ j, k = 1, 2, \dots, N; & \end{aligned}$$

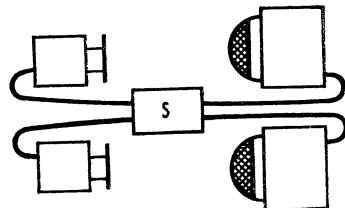
$$\varphi_{j,k} \equiv \varphi_{j+N,k+N}.$$

Как оказалось, при достаточно большой амплитуде внешнего поля ( $A \approx 50$ ) в системе (30) устанавливается мелкомасштабный пространственно-временной хаос, в формировании которого основную роль играют наиболее коротковолновые возбуждения, в частности  $\pi - \pi$  колебания, соответствующие противофазным колебаниям соседних элементов. Размерность хаоса при этом имеет порядок числа элементов в решетке (аналогично см. [28]), а в его пространственной картине отсутствуют какие-либо выделенные элементы — уединенные структуры. Установившийся

Рис. 13. Пространственно-временной хаос в двумерной решетке (30) ( $A = 50,0$ ;  $\Lambda = 0,1$ ;  $h=0,3$ );

Пространственная симметрия картины поля связана с периодическими граничными и симметричными начальными условиями

Рис. 14. Схематическое изображение двухкомпонентного ТВ-аналога, описываемого системой (31)



хаотический режим при этом выглядит [как последовательность никогда не повторяющихся сложных пространственных распределений поля — это иллюстрирует рис. 13.

Сделаем несколько замечаний о моделировании с помощью неавтономных неравновесных сред ассоциативной памяти и использования их для распознавания образов.

Как сейчас представляется, наиболее перспективный путь моделирования ассоциативной памяти (обучение, запоминание и распознавание образов) [32] связан с использованием некоторых абстрактных сред, которые называют спиновыми стеклами. Эти «стекла» представляют собой динамическую систему в виде ограниченной решетки спинов, каждый из которых может находиться в двух состояниях. Связи между спинами формируются в процессе обучения, и делается это так, чтобы каждому запоминаемому образу соответствовал аттрактор в фазовом пространстве системы (т. е. решетки связанных спинов). Чем шире область притяжения такого аттрактора (этим аттрактором может быть просто состояние равновесия), тем надежнее процесс распознавания и воспроизведения образа по его части или при наличии искажений.

Переходу от воспроизведения одного образа к воспроизведению другого соответствует индуцированный внешним воздействием (образом — «претендентом») переход рассматриваемой динамической системы с одного аттрактора на другой. К сожалению, конкретная физическая реализация подобных спиновых стекол до настоящего времени проблематична. Представляется естественным использовать для моделирования ассоциативной памяти двухкомпонентную неравновесную среду, в которой одна из компонент «медленная», а другая — «быстрая». Реализовать подобную среду можно, например, в виде экранов двух мониторов, каждый из которых включен в замкнутый контур монитор—камера—усилитель—обратная связь (рис. 14). Уравнения такой двухкомпонентной среды записываются в виде

$$\begin{aligned} \tau_1 \partial u / \partial t &= F_1(u, v) + D_1 \Delta u + \int v(\xi - \mathbf{r}) u(\xi) d\xi + g(\mathbf{r}) H(\tau_0), \\ \tau_2 \partial v / \partial t &= F_2(u, v) + D_2 \Delta v + \int u(\mathbf{r} - \xi) u(\xi) d\xi, \end{aligned} \quad (31)$$

$$\tau_1 \ll \tau_2 < \tau_0.$$

Здесь поле  $u(\xi)$  непосредственно реагирует на предъявляемый образ  $g(\xi)$  (он действует конечное время  $\tau_0$ ) и воспроизводит его, а поле  $v(\xi)$  играет роль подстраивающейся характеристики среды — при изменении  $v(\xi)$  формируется такой «рельеф» среды (он устойчив благодаря специальному виду  $F_2(u, v)$ ), при котором предъявляемому образу соответствует устойчивое равновесие. В такой распределенной неравновесной системе реализуется принцип параллельного распознавания образов, что, по-видимому, имеет место и в живой природе.

## ЛИТЕРАТУРА

1. *Newell A. C., Whitehead J. A.* Finite bandwidth, finite amplitude convection // *J. Fluid Mech.* 1969. Vol. 38, N 3. P. 279—303.
2. *Stewartson K., Stuart J. T.* A nonlinear instability theory for a wave system in plane Poiseuille flow // *Ibid.* 1971. Vol. 48, N 3. P. 529—545.
3. *DiPrima R. C., Eckhaus W., Segel L. A.* A nonlinear wave-number interaction in near-critical two-dimensional flows // *Ibid.* Vol. 49, N 4. P. 705—744.
4. *Swift J. S., Hohenberg P. C.* Hydrodynamic fluctuation at the convective instability // *Phys. Rev. A.* 1977. Vol. 15, N 1. P. 319—328.
5. *Хакем Г.* Синергетика. Иерархии неустойчивостей в самоорганизующихся системах и устройствах. М.: Мир, 1985. 485 с.
6. *Ма Ш.* Современная теория критических явлений. М.: Мир, 1980. 298 с.
7. *Normand C., Pomeau Y., Velarde M. G.* Convective instability: A physicists approach // *Rev. Mod. Phys.* 1977. Vol. 49. P. 591—612.
8. *Perez-Garcia C.* Stability in thermodynamic systems / Ed. J. Casas-Vazquez, J. Lebon. Berlin.: Springer, 1983. 230 p. (Lect. Notes. Phys.; Vol. 164).
9. *Kuramoto Y.* Chemical oscillations. Waves and turbulence. Berlin: Springer, 1984. 365 p.
10. *Гапонов-Грехов А. В., Рабинович М. И.* Уравнение Гинзбурга—Ландау и нелинейная динамика неравновесных сред // *Изв. вузов. Радиофизика.* 1987. Т. 30, № 2. С. 131—143.
11. *Езерский А. Б., Рабинович М. И., Реутов В. П., Старобинец И. М.* Пространственно-временной хаос капиллярной ряби при параметрическом возбуждении жидкости // *ЖЭТФ.* 1986. Т. 91, № 6(12). С. 2070—2083.
12. *Bishop A. R., Eilbeck J. C., Satiya I., Wysin G.* Pattern selection and low-dimensional chaos in dissipative many degree-of-freedom systems // *Lect. Appl. Math.* 1986. Vol. 23. P. 96—120.
13. *Tesauro G., Cross M. C.* Climbing of dislocations in nonequilibrium patterns // *Phys. Rev. A.* 1986. Vol. 34, N 2. P. 1363—1379.
14. *Рабинович М. И.* Автоколебания распределенных систем // *Изв. вузов. Радиофизика,* 1974. Т. 17, № 4. С. 477—498.
15. *Cellular structures in instabilities* / Ed. J. E. Wesfreid, S. Zaleski. Berlin etc.: Springer, 1984. 390 p. (Lect. Notes. Phys.; Vol. 210).
16. *Haken H.* Pattern formation and chaos in synergetic system // *Phys. scr.* 1985. Vol. T9. P. 111—118.
17. *Coulet P., Elphick C., Gil L., Lega J.* Topological defects of wave patterns / *Phys. Rev. Lett.* 1987. Vol. 59. P. 884—887.
18. *Mermin N. D.* The topological theory of defects in ordered media // *Rev. Mod. Phys.* 1979. Vol. 51, N 3. P. 591—648.
19. *Lowe M., Gollub J. P.* Solution and the commensurate-incommensurate transition in a convecting nematic fluid // *Phys. Rev. A.* 1985. Vol. 31, N 6. P. 3893—3897.
20. *Heutmaker M. S., Gollub J. P.* Wave-vector field of convective flow patterns // *Ibid.* 1987. Vol. 35, N 1. P. 242—259.
21. *Roberts M., Swift J. W., Wagner D. H.* The Hopf bifurcation on a hexagonal lattice // *Contemp. Math.* 1986. Vol. 56. P. 283—318.
22. *Croquette V., Pocheau A.* Wavenumber selection in Rayleigh-Benard convective structure // *Lect. Notes Phys.* 1984. Vol. 210. P. 104—112.
23. *Гапонов-Грехов А. В., Ломов А. С., Рабинович М. И.* Локализованные автоструктуры в двумерных однородных средах // *Письма в ЖЭТФ.* 1986. Т. 44, № 5. С. 224—228.
24. *Вашкевич О. В., Гапонов-Грехов А. В., Езерский А. Б., Рабинович М. И.* Рождение уединенных автоструктур при термоконвекции в слое с неоднородным подогревом // *ДАН СССР.* 1987. Т. 293, № 3. С. 563—567.
25. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* Рождение автоструктур, иерархия неустойчивостей и бифуркации // *Изв. АН СССР. Сер. физ.* 1987. Т. 51, № 6. С. 1133—1150.

26. *Арансон И. С., Рабинович М. И.* Рождение и динамика спиральных вихрей в двумерной модели Гинзбурга—Ландау // Изв. вузов. Радиофизика. 1986. Т. 29. № 12. С. 1514—1517.
27. *Linde H.* Topological similarities in dissipative structures of Marangoni-instability and Belousov-Zhabotinsky-reaction // Self-organization auto-waves and structures far from equilibrium / Ed. V. I. Krinsky. Berlin: Springer, 1984. P. 154—157.
28. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* Развитие хаоса в ансамбле динамических структур // ЖЭТФ. 1985. Т. 81. № 1. С. 92—103.
29. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* и др. Решеточные модели в нелинейной динамике неравновесных сред: Препр. ИПФ АН СССР № 163. Горький, 1987. 24 с.
30. *Тамура И.* Топология слоений. М.: Мир, 1979. 317 с.
31. *Couillet P.* Commensurate-incommensurate transition in nonequilibrium systems // Phys. Rev. Lett. 1986. Vol. 56, N 7. P. 724—727.
32. *Hopfield J. J.* Neural networks and physical systems with emergent collective computational abilities // Proc. Nat. Acad. Sci. USA. 1982. Vol. 79. P. 2554—2558.

УДК 531.32+532.5

## СТОХАСТИЧЕСКАЯ ПАУТИНА И СИММЕТРИЯ СТРУКТУР

Г. М. ЗАСЛАВСКИЙ, Р. З. САГДЕЕВ, Д. А. УСИКОВ,  
А. А. ЧЕРНИКОВ

В гамильтоновских системах область фазового пространства вблизи сепаратрисы играет фундаментальную роль при анализе возникновения хаоса. Сепаратриса представляет собой особую траекторию, проходящую через неустойчивую седловую точку, в которой происходит ее самопересечение. Возмущение системы может быть и слабым, и медленным. Оно может слабо влиять на динамику системы. Однако действие возмущения всегда оказывается сильным, если начальные условия системы принадлежат некоторой области в окрестности сепаратрисы. Именно в этой области зарождается хаос, который радикально изменяет свойства системы и делает ее принципиально неинтегрируемой.

Образование в окрестности сепаратрис гомоклинических и гетероклинических структур описано Пуанкаре [1]. Первая оценка ширины расщепления сепаратрис дана в работе [2]. Существование стохастической динамики в окрестности сепаратрисы и анализ ширины стохастического слоя для различных случаев возмущения были рассмотрены в работах [3—5]. Идея построения отображения в окрестности сепаратрисы, основанная на свойствах траектории системы, высказана в работе [4]. Методика расчета толщины стохастического слоя, использованная в настоящей работе, изложена в [4—8].

### 1. ОБРАЗОВАНИЕ СТОХАСТИЧЕСКОГО СЛОЯ

Нелинейный маятник представляет собой систему, типичную для многих физических задач. Его гамильтониан имеет вид

$$H_0 = 1/2 p^2 - \omega_0^2 \cos x, \quad (1.1)$$

где  $x$  — безразмерная координата;  $p$  — импульс,  $\omega_0$  — частота малых колебаний маятника, а масса положена равной единице. Простейшая форма возмущенной задачи может быть записана так:

$$H = H_0 + V; \quad V \equiv \varepsilon (\omega_0^2/k) \cos (kx - \Delta\omega t), \quad (1.2)$$

где  $\varepsilon$  — безразмерный параметр возмущения;  $k$  и  $\Delta\omega$  — соответственно волновое число и частота возмущения. Гамильтониан (1.2) порождает уравнение движения

$$\ddot{x} + \omega_0^2 \sin x = \varepsilon \omega_0^2 \sin(kx - \Delta\omega t). \quad (1.3)$$

Фазовый портрет невозмущенной задачи (1.3) при  $\varepsilon = 0$  приведен на рис. 1, а. При малом возмущении (рис. 1, б) часть инвариантных траекторий вблизи сепаратрисы разрушается, образуя стохастический слой. В основе оценки его ширины лежат следующие соображения.

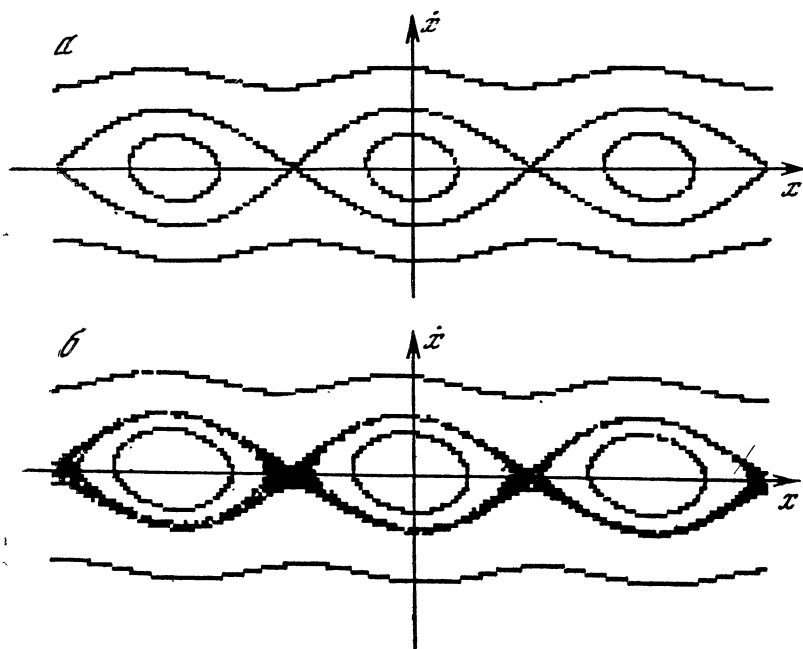


Рис. 1. Фазовый портрет нелинейного маятника в отсутствие (а) и при наличии (б) возмущения

Видны области хаотической динамики в окрестности сепаратрисы

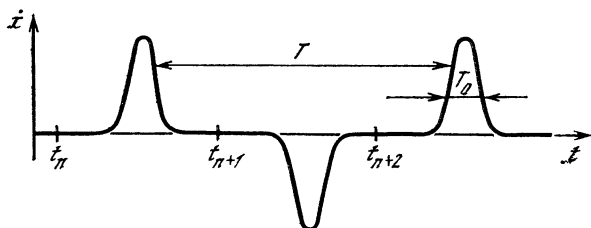


Рис. 2. Зависимость скорости маятника от времени для траектории вблизи сепаратрисы

$T$  — период колебаний на траектории,  $T_0$  — период малых колебаний

При малых амплитудах колебания маятника его можно считать линейным ( $\sin x \approx x$ ). Тогда спектр маятника состоит из одной гармоникой с частотой  $\omega_0$ . В общем случае нелинейная частота маятника  $\omega(E)$  зависит от его энергии  $E = H_0$ . При малых  $E$  частота  $\omega \approx \omega_0$ . Особой траектории — сепаратрисе — соответствует энергия  $E_c = \omega_0^2$ . При  $E \rightarrow E_c$  частота  $\omega(E) \rightarrow 0$ , а период колебаний

$$T = 2\pi/\omega(E) \rightarrow \infty \quad (E \rightarrow E_c = \omega_0^2). \quad (1.4)$$

При этом характерное число гармоник в спектре маятника

$$N \approx \omega_0/\omega(E) = T/T_0 \quad (T_0 = 2\pi/\omega_0) \quad (1.5)$$

стремится к бесконечности.

Из свойства (1.4) следует, что скорость маятника  $p = \dot{x}(t)$  при  $E$  вблизи  $E_c$  имеет вид, изображенный на рис. 2. Такой характер поведения скорости обусловлен тем, что система быстро (за время  $\sim T_0$ ) проходит почти всю ширину потенциальной ямы и надолго (на время  $\sim T \gg T_0$ ) застревает вблизи седловых точек. Поэтому условие  $N \gg 1$  вблизи сепаратрисы дает возможность перейти в возмущенной задаче (1.3) от дифференциального уравнения к конечно-разностному уравнению (отображению).

Оно записывается между двумя последовательными моментами времени  $t_n$  и  $t_{n+1}$  (см. рис. 2), взятыми в окрестности  $n$ -го и  $(n+1)$ -го прохождения вблизи какой-либо из гиперболических точек. Из уравнения

$$\dot{E} = \varepsilon \omega_0^2 \dot{x} \sin(kx - \Delta \omega t), \quad (1.6)$$

которое получается при дифференцировании соотношения  $E = H_0$  и использовании выражений (1.1)–(1.3), видно, что изменение энергии маятника происходит лишь в области, где его скорость  $\dot{x}$  отлична от нуля. Такая область имеет узкий временной интервал  $\sim T_0 = 2\pi/\omega_0$  (см. рис. 2) и эквивалентна действию силового толчка на маятник. Между последовательными толчками проходит интервал времени  $\Delta t_n = t_{n+1} - t_n$  и удовлетворяется условие мгновенности столкновения  $T_0 \ll \Delta t_n$ . Поэтому искомое отображение имеет вид

$$E_{n+1} = E_n + \Delta E_n, \quad t_{n+1} = t_n + (\pi/2)(1/\omega_n + 1/\omega_{n+1}), \quad (1.7)$$

в котором  $\omega_n \equiv \omega(E_n)$ .

Интегрируя (1.6) по времени, для изменения энергии невозмущенного движения  $\Delta E$  получим

$$\Delta E = \int \varepsilon \omega_0^2 \dot{x} \sin(kx - \Delta \omega t) dt. \quad (1.8)$$

При  $\varepsilon = 0$  движение по сепаратрисе описывается соотношением

$$x(t) = 4 \operatorname{arctg} \exp[\omega_0(t - t_n)] - \pi, \quad (1.9)$$

где  $t_n$  — момент  $n$ -го пересечения поверхности  $x = 0$ . Подставляя (1.9) в правую часть (1.8) и интегрируя по времени, для изменения

энергии будем иметь

$$\Delta E = 2\varepsilon\omega_0^3 \int_{-\infty}^{+\infty} \frac{1}{\operatorname{ch} \omega_0 t} \sin [4k \operatorname{arctg} \exp(\omega_0 t) - \Delta\omega t - k\pi - \Delta\omega t_n] dt. \quad (1.10)$$

Запишем правую часть (1.10) в более компактном виде:

$$\Delta E = -\varepsilon\omega_0^2 2^{2k+2} \sin \Delta\omega t_n \int_0^{\infty} \frac{\cos v\tau}{(\operatorname{ch} \tau)^{2k+1}} d\tau, \quad (1.11)$$

где  $v = \Delta\omega/\omega_0 \gg 1$ , и вычислим интеграл:

$$\Delta E = -\frac{2^{4k+1}\varepsilon\omega_0^2}{\Gamma(2k+1)} \Gamma\left(k + \frac{1+iv}{2}\right) \Gamma\left(k + \frac{1-iv}{2}\right) \sin \Delta\omega t_n. \quad (1.12)$$

В частности, если  $k$  целое, то

$$\Delta E = -\pi\varepsilon\omega_0^2 \frac{2^{4k+1}}{(2k)!} \frac{\sin \Delta\omega t_n}{\operatorname{ch}(\pi v/2)} \prod_{r=1}^k \left[ \frac{v^2}{4} + \left( \frac{2r-1}{2} \right)^2 \right]. \quad (1.13)$$

При  $k=1$  из (1.10) следует точная формула

$$\Delta E = -4\pi\varepsilon\omega_0^2 \frac{v^2 \exp(\pi v/2)}{\operatorname{sh}(\pi v)} \sin \Delta\omega t_n. \quad (1.14)$$

Используя это выражение и введя для удобства фазу  $\theta_n$ , получаем вместо (1.7) окончательно

$$\tilde{E}_{n+1} = \tilde{E}_n - \frac{4\pi\varepsilon v^2 \exp(\pi v/2)}{\operatorname{sh}(\pi v)} \sin \theta_n, \quad (1.15)$$

$$\theta_{n+1} = \theta_n + v \ln(32/\tilde{E}_{n+1}),$$

где использованы обозначения

$$v = \Delta\omega/\omega_0; \quad \omega(E) \approx \omega_0 \ln(32/E), \quad (1.16)$$

$$\tilde{E} = -1 + E/\omega_0^2 = (E - E_c)/E_c.$$

Простейшая оценка области стохастичности получается из неравенства [7]

$$|\delta\theta_{n+1}/\delta\theta_n - 1| > 1. \quad (1.17)$$

Отсюда для ширины стохастического слоя по энергии имеем из (1.15)

$$|\tilde{E}_s| < \frac{4\pi\varepsilon v^3 \exp(\pi v/2)}{\operatorname{sh}(\pi v)}. \quad (1.18)$$

Если возмущение в (1.2), (1.3) имеет низкую частоту  $\Delta\omega \ll \omega_0$ , то

$$|\tilde{E}_s| < 4\varepsilon v^2 \quad (1.19)$$

и ширина стохастического слоя в этом случае максимальная. Если, наоборот, возмущение высокочастотное ( $\Delta\omega \gg \omega_0$ ), то

$$|\tilde{E}_s| < 8\pi\varepsilon v^2 \exp(-\pi v/2) \quad (1.20)$$

и ширина стохастического слоя экспоненциально мала.



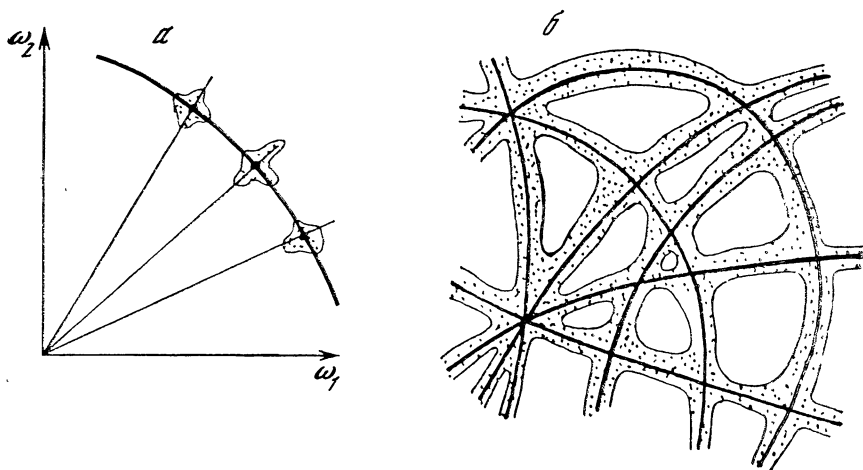


Рис. 3. Области стохастичности при  $N = 2$  (а) и  $N > 2$  (б)

Теория Колмогорова—Арнольда—Мозера (КАМ) [9] определяет условия сохранения инвариантных торов гамильтоновских систем при действии на систему малого возмущения. Если система с  $N$  степенями свободы описывается гамильтонианом

$$H_0 = H_0(I_1, \dots, I_N),$$

зависящим от  $N$  интегралов движения (действий)  $I_k$ , которые независимы и коммутируют, то траектория системы лежит на  $N$ -мерном торе. Сам тор тоже является инвариантом. Возмущенный гамильтониан системы можно представить в виде

$$H = H_0(I_1, \dots, I_N) + \varepsilon V(I_1, \theta_1; \dots; I_N, \theta_N); \quad (1.21)$$

где  $\theta_k$  — фазы, канонически сопряженные действиям  $I_k$ , а  $\varepsilon$  — малый параметр возмущения. При достаточно малых  $\varepsilon$  и при отсутствии вырождения, т. е.

$$\text{Det} \left| \frac{\partial^2 H_0}{\partial I_k \partial I_l} \right| \neq 0, \quad (1.22)$$

большинство инвариантных торов сохраняется в слабо деформированном виде. Часть торов при этом разрушается, однако число сохраняющихся торов так же соотносится к числу разрушенных торов, как множество иррациональных чисел соотносится к множеству рациональных чисел, поэтому разрушенных торов мало и они зажаты между инвариантными торами.

Возможность пересечения стохастических слоев при  $\varepsilon \rightarrow 0$  существует лишь при  $N > 2$ . Это приводит к тому, что фазовое пространство покрывается сетью соединяющихся друг с другом узких каналов со стохастической динамикой частиц внутри. По этой сетке возможен уход частиц сколь угодно далеко — явление, получившее название диффузии Арнольда [10] (рис. 3). При  $N = 2$

условие резонанса в системе имеет вид

$$n_1\omega_1 + n_2\omega_2 = 0, \quad (1.23)$$

где  $n_1, n_2$  — целые числа, а частоты каждой из степеней свободы  $\omega_i$  определяются известными выражениями

$$\omega_i = \partial H_0 / \partial I_i = \omega_i(I_1, \dots, I_N), \quad i = 1, \dots, N. \quad (1.24)$$

При  $N = 2$  условие (1.23) приводит к линейной зависимости между  $\omega_1$  и  $\omega_2$ , поэтому поверхность заданной энергии  $H = E$  пересекает семейство прямых  $\omega_1/\omega_2 = \text{const}$  в точках (см. рис. 3, а). В их окрестности происходит разрушение инвариантов и появляется хаос. Однако малые возмущения создают малые области хаоса, которые не соединяются друг с другом. При  $N > 2$  такое соединение уже возможно (см. рис. 3, б), и в результате возникает сеть, называемая далее стохастической паутиной.

## 2. МИНИМАЛЬНЫЙ ХАОС И СТОХАСТИЧЕСКАЯ ПАУТИНА

Проблема минимального хаоса заключается в определении условий, при которых возникают малые области стохастической динамики при сколь угодно малом возмущении. Существование минимального хаоса означает принципиальную неинтегрируемость системы и возможность перехода при увеличении некоторого параметра возмущения от минимального хаоса к глобальному хаосу, охватывающему значительную область фазового пространства. Два связанных нелинейных осциллятора с гамильтонианом  $H = H_1(I_1) + H_2(I_2) + \varepsilon V(I_1, \theta_1; I_2, \theta_2)$  представляют собой простой пример, в котором в случае общего положения существует минимальный хаос при сколь угодно малых  $\varepsilon$ .

Стохастическая паутина является своеобразной формой проявления минимального хаоса. Фазовое пространство разбивается на отдельные ячейки — структуры, отделенные друг от друга стохастическими слоями. Иными словами, паутина образует в фазовом пространстве некоторую мозаику покрытия его. Характерный размер элементарной мозаичной структуры определяет не только область существования инвариантных торов, но и их максимальные размеры. Поэтому определение условий, при которых существует стохастическая паутина, интересно по крайней мере по следующим причинам: 1) как возможность диффузии по каналам паутины при сколь угодно малых  $\varepsilon$ ; 2) как ограничение на максимальный диаметр инвариантных торов при сколь угодно малых  $\varepsilon$ ; 3) как выяснение внутренней симметрии фазового пространства системы, возникающей при сколь угодно малых  $\varepsilon$ .

Некоторые физические модели имеют довольно универсальную природу. Примером таких моделей является возмущенный нелинейный осциллятор, введенный в уравнениях (1.2), (1.3). Его гамильтониан содержит две плоские волны. На этой модели удастся выяснить фундаментальные и типичные свойства образования

стохастического слоя. С другой стороны, именно такие уравнения, как (1.2), (1.3), возникают во многих физических приложениях, и это позволяет их причислить к категории эталонных уравнений.

Введем еще одно эталонное уравнение:

$$\ddot{x} + \omega_0^2 x = \varepsilon \omega_0^2 \sin(kx - \Delta\omega t), \quad (2.1)$$

которое описывает линейный осциллятор, возмущаемый плоской волной. Гамильтониан, приводящий к (2.1), имеет вид

$$H = 1/2 p^2 + 1/2 \omega_0^2 x^2 + \varepsilon (\omega_0^2/k) \cos(kx - \Delta\omega t). \quad (2.2)$$

Он возникает, в частности, при движении заряженной частицы в постоянном магнитном поле  $B_0$  (которое направлено вдоль  $z$ ) и в поле плоской электростатической волны, распространяющейся вдоль оси  $x$  [11—13]. Задача (2.1) много исследовалась [14—16] в связи с проблемой резонанса волна—частица. Возмущение содержит нелинейность, которая, будучи разложена в ряд Фурье, приводит к большому числу гармоник. Между ними и частотой внешнего возмущения  $\Delta\omega$  возможны резонансы:

$$n\omega_0 = \Delta\omega, \quad (2.3)$$

где  $\omega_0$  имеет смысл ленгмюровской частоты ( $\omega_0 = eB_0/mc$ ). Специфика задачи (2.2) в том, что ее невозмущенная часть является линейной. Поэтому если выполнено условие (2.3), то условие невырожденности (1.28) оказывается невыполненным и теорема КАМ непосредственно неприменима. Это и приводит к образованию стохастической паутины [17].

Нас будет интересовать случай, близкий к точному резонансу (2.3) при некотором значении  $n = n_0$ . Совершим каноническую замену переменных

$$\begin{aligned} x &= (2n_0 I / \omega_0)^{1/2} \cos(\varphi/n_0 - \Delta\omega t); \\ p &= (2n_0 I \omega_0)^{1/2} \sin(\varphi/n_0 - \Delta\omega t), \end{aligned} \quad (2.4)$$

где  $(I, \varphi)$  — переменные действие—угол. Гамильтониан (2.2) преобразуется с помощью замены (2.4) в следующий:

$$H = (n_0 \omega_0 - \Delta\omega) I + \varepsilon \frac{\omega_0^2}{k} \sum_n J_n(k\rho) \cos\left[\frac{n}{n_0} \varphi + \left(\frac{n}{n_0} - 1\right) \Delta\omega t\right], \quad (2.5)$$

где  $J_n(k\rho)$  — функции Бесселя и величина

$$\rho = (2n_0 I / \omega_0)^{1/2} \quad (2.6)$$

имеет смысл ларморовского радиуса, если принять, что  $\omega_0$  — ларморовская частота.

Смысл введения переменных (2.4) в том, чтобы отделить медленное движение от быстрого в окрестности резонанса. При точном

резонансе (2.3) представим (2.5) в виде

$$H = H_0 + V; \quad H_0 = \varepsilon \frac{\omega_0^2}{k} J_{n_0}(k\rho) \cos \varphi; \quad (2.7)$$

$$V = \varepsilon \frac{\omega_0^2}{k} \sum_{n \neq n_0} J_n(k\rho) \cos \left[ \frac{n}{n_0} \varphi + \left( \frac{n}{n_0} - 1 \right) \Delta \omega t \right].$$

Возникшая ситуация уже встречалась нам в предыдущем разделе. Невозмущенное движение описывается гамильтонианом  $H_0$ . Это движение имеет особые траектории — сепаратрисы, которые разрушаются возмущением  $V$ , образуя стохастические слои.

Особые точки находятся из условий

$$\partial H_0 / \partial I = 0, \quad \partial H_0 / \partial \varphi = 0.$$

Гиперболические точки  $(\rho_*, \varphi_*)$  определяются уравнениями

$$J_{n_0}(k\rho_*) = 0, \quad \varphi_* = \pm \pi/2, \quad (2.8)$$

а семейство эллиптических точек  $(\rho_0, \varphi_0)$  — уравнениями

$$J'_{n_0}(k\rho_0) = 0, \quad \varphi_0 = 0, \pi. \quad (2.9)$$

Семейство сепаратрис, проходящих через точки (2.8), имеет структуру паутины (рис. 4). Она состоит из  $2n_0$  лучей и концентрических окружностей с радиусами  $\rho_*^{(s)}$  на плоскости  $(x, p)$ , где  $k\rho_*^{(s)}$  — различные корни функции Бесселя  $J_{n_0}$ . Для понимания физической ситуации достаточно ограничиться случаем больших энергий частиц  $k\rho_0 \gg n_0$ . Тогда

$$H_0 = \pm \varepsilon \frac{\omega_0^2}{k} \left( \frac{2}{\pi k \rho_0} \right)^{1/2} \cos k\tilde{\rho} \cos \varphi, \quad (2.10)$$

где  $\tilde{\rho} = \rho - \rho_0 \ll \rho_0$ , а знак («+» или «-») зависит от фазы эллиптической точки с координатами (2.9).

Период колебаний частиц внутри одной ячейки сепаратрисы на рис. 4 определяется из (2.10) и равен

$$T = \left( \frac{\pi}{2} \right)^{1/2} \frac{4\Delta\omega}{k\varepsilon\omega_0^2 n_0} (k\rho_0)^{3/2} K(\kappa), \quad (2.11)$$

где  $K(\kappa)$  — полный эллиптический интеграл первого рода,

$$\kappa = 1 - H_0^2 k^2 \pi k \rho_0 / 2\varepsilon^2 \omega_0^4. \quad (2.12)$$

Вблизи положения равновесия внутри ячейки паутины частица совершает малые колебания с периодом

$$T_0 = \left( \frac{\pi}{2} \right)^{1/2} \frac{4\Delta\omega}{k\varepsilon\omega_0^2 n_0} (k\rho_0)^{3/2}, \quad (2.13)$$

а для колебаний частиц вблизи сепаратрисы ( $H \rightarrow 0$ ) имеем

$$T = \left( \frac{\pi}{2} \right)^{1/2} \frac{4\Delta\omega}{k\varepsilon\omega_0^2 n_0} (k\rho_0)^{3/2} \ln \left[ \frac{4\varepsilon\omega_0^2}{k^2 H_0} \left( \frac{2}{\pi k \rho_0} \right)^{1/2} \right], \quad (2.14)$$

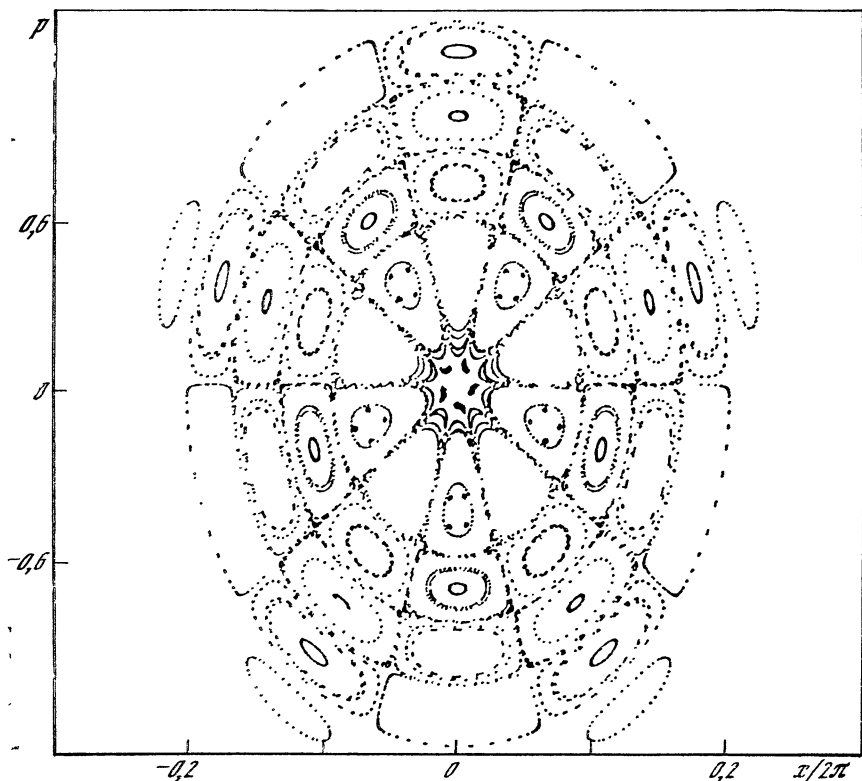
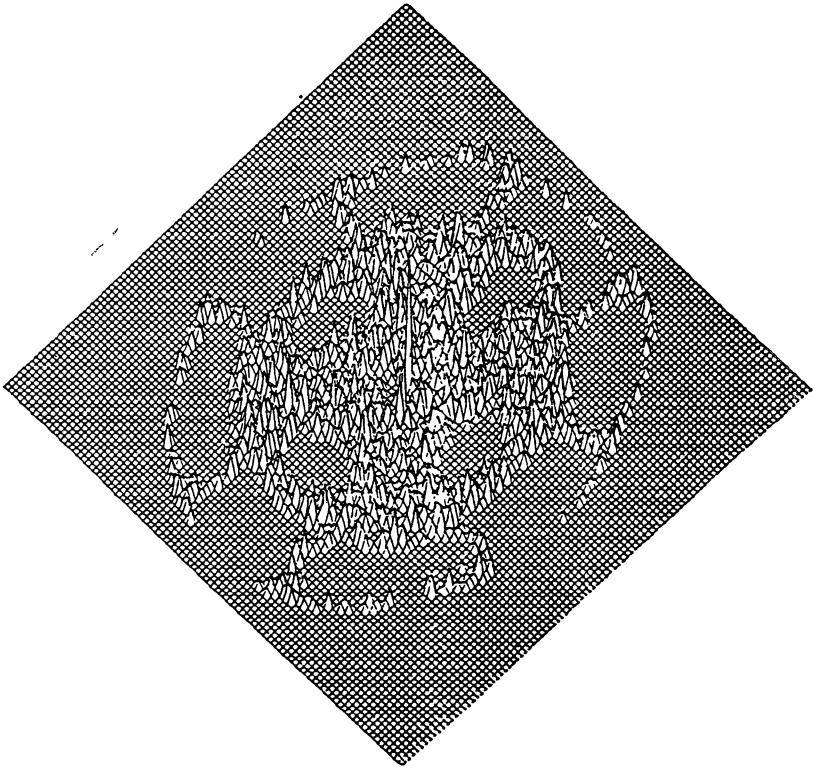


Рис. 4. Сечение Пуанкаре для уравнения (2.1) на плоскости  $(x/2\pi, p)$  при  $n_0 = 5$ ,  $\varepsilon\omega_0^2/k = 1/80$ ,  $k = 15$

и период логарифмически расходится. Это означает, что возмущение  $V$  в (2.7) является высокочастотным ( $\Delta\omega T \gg 1$  при достаточно малых  $H_0$ ). Оно приводит к разрушению сепаратрисной сетки и к образованию стохастического слоя экспоненциальной малой ширины [17]

$$E_s = (2\pi)^{3/2} \frac{4\omega_0^2}{k^3 \varepsilon n_0^2} (k\rho_0)^{3/2} \exp\left[-\frac{(\pi k\rho_0/2)^{3/2}}{kn_0\varepsilon}\right]. \quad (2.15)$$

Таким образом, на фазовой плоскости образуется неограниченная сеть каналов (см. рис. 4), по которым возможен уход частиц. По мере удаления от центра скорость частиц увеличивается. Однако одновременно происходит уменьшение ширины паутины. Согласно формуле (2.15), ширина паутины экспоненциально убывает с ростом энергии частиц, так как величина  $\rho_0 \sim E^{1/2}$  (см. определение (2.6)). Поэтому, хотя формально паутина и является неограниченной, вероятность диффузионного просачивания частиц в область больших энергий экспоненциально мала. Это хорошо отражается на виде функции распределения [17] (рис. 5).



**Рис. 5.** Функция распределения на стохастической паутине для уравнения (2.1) при  $\varepsilon\omega_0^2/k = 0,1$ ,  $k = 15$ ,  $\Delta\omega = 4\omega_0$  и времени счета  $4 \cdot 10^4 \pi/\Delta\omega$

### 3. РАВНОМЕРНАЯ ПАУТИНА

Итак, мы выяснили, что стохастическая паутина существует при минимальной неинтегрируемой размерности динамической системы. Далее одной из важных задач является определение возможных, типичных и исключительных случаев ее существования. К числу последних относится такая паутина, которая имеет примерно одинаковую толщину на всей бесконечной фазовой плоскости. Такой случай возможен [5], он возникает в том же классе задач, содержащих резонансы волна—частица при наличии постоянного магнитного поля. Особый интерес к этому случаю связан с тем, что он является своеобразным генератором нового типа симметрии [18], который будет подробно обсужден ниже.

Рассмотрим снова осциллятор, на который действует не одна плоская волна, как в (2.1), а пакет бесконечного числа плоских волн с одинаковыми амплитудами и с постоянным сдвигом по

частоте:}

$$\ddot{x} + \omega_0^2 x = -\varepsilon \omega_0^2 \sum_{n=-\infty}^{+\infty} \sin(kx - \omega t - n\Delta\omega t). \quad (3.1)$$

Далее удобно положить  $\theta = kx - \omega t$  и представить (3.1) в виде

$$\ddot{x} + \omega_0^2 x = -\varepsilon \omega_0^2 T_0 \sin \theta \sum_{n=-\infty}^{+\infty} \delta(t - nT_0) \quad (3.2)$$

с тем же обозначением  $\Delta\omega = 2\pi/T_0$ . Гамильтониан системы (3.1) или (3.2) имеет следующий вид:

$$H = \frac{1}{2} p^2 + \frac{1}{2} \omega_0^2 x^2 - \varepsilon \frac{\omega_0^2}{k} T_0 \cos(kx - \omega t) \sum_{n=-\infty}^{+\infty} \delta(t - nT_0) \quad (3.3)$$

Его физический смысл легко представить себе, если сопоставить рассматриваемый случай (3.1)—(3.3) с уже рассмотренной задачей (2.2). Вместо плавного периодического возмущения, как в (2.2), на осциллятор действуют периодические  $\delta$ -импульсы, т. е. мгновенные толчки. Кроме того, формула (3.3) соответствует плоскому движению заряженной частицы в постоянном магнитном поле  $\mathbf{B}_0 \parallel Oz$  и в электростатическом поле, направленном вдоль  $Ox$ :

$$E(x, t) = -\varepsilon \omega_0^2 T_0 \sin \theta \sum_{n=-\infty}^{+\infty} \delta(t - nT_0). \quad (3.4)$$

Теперь вырождение может быть особенно сильным при резонансах

$$\alpha = \omega_0 T_0 = 2\pi p/q, \quad (3.5)$$

где  $p$  и  $q$  — целые числа (далее будем считать  $p < q$ ). Условие резонанса (3.5) означает, например при  $p = 1$ , что за время одного оборота частицы в магнитном поле ( $2\pi/\omega_0$ ) она испытывает ровно  $q$  толчков со стороны поля волн.

Уравнение движения (3.2) можно проинтегрировать один раз и получить отображение, связывающее значение  $(p, x)$  через интервал времени  $T_0$ :

$$\hat{T}: \begin{cases} p_{n+1} = [p_n + \varepsilon \omega_0^2 T_0 \sin(n\omega T_0 - kx_n)] \cos \alpha - \omega_0 x \sin \alpha, \\ x_{n+1} = x_n \cos \alpha + \frac{1}{\omega_0} [p_n + \varepsilon \omega_0^2 T_0 \sin(n\omega T_0 - kx_n)] \sin \alpha, \end{cases} \quad (3.6)$$

где  $\alpha = \omega_0 T_0$ . При  $\omega_0 = 0$  отображение (3.6) переходит в стандартное отображение. Наличие постоянной средней скорости  $\omega/k$  у волнового пакета (3.4) приводит к эффекту ускорения частиц. Более симметричный случай возникает при  $\omega = 0$ . Тогда отображение (3.6) переходит в так называемое отображение с подкручи-

ванием на угол  $\alpha$  [5]:

$$\widehat{M}_\alpha: \begin{cases} p_{n+1} = (p_n - \varepsilon\omega_0^2 T_0 \sin kx_n) \cos \alpha - \omega_0 x_n \sin \alpha, \\ x_{n+1} = x_n \cos \alpha + \frac{1}{\omega_0} (p_n - \varepsilon\omega_0^2 T_0 \sin kx_n) \sin \alpha. \end{cases} \quad (3.7)$$

Удобно также ввести безразмерные переменные

$$u = kp/\omega_0, \quad v = -kx \quad (3.8)$$

и представить  $\widehat{M}_\alpha$  в виде

$$\widehat{M}_\alpha: \begin{cases} u_{n+1} = (u_n + K_H \sin v_n) \cos \alpha + v_n \sin \alpha, \\ v_{n+1} = -(u_n + K_H \sin v_n) \sin \alpha + v_n \cos \alpha, \end{cases} \quad (3.9)$$

где

$$K_H = \varepsilon\omega_0 T_0. \quad (3.10)$$

Случай резонанса (3.5) при  $p = 1$  будем обозначать через

$$\alpha_q = 2\pi/q. \quad (3.11)$$

Как мы увидим далее, они играют особо важную роль.

Параметр  $K_H$  характеризует величину взаимодействия частицы с волнами. Если  $K_H = 0$ , то отображение  $\widehat{M}_\alpha$  описывает поворот на угол  $\alpha$  за один шаг. Если параметр  $K_H$  мал, то при резонансном условии (3.11) можно представить гамильтониан (3.3) в более удобной форме. С помощью обозначений (3.5) и (3.8) перепишем его в виде (с точностью до постоянного множителя)

$$H = \frac{1}{2} \alpha_q^2 (u^2 + v^2) - K_H \cos v \sum_{n=-\infty}^{+\infty} \delta(\tau - n), \quad (3.12)$$

где  $\tau = t/T_0$ . Представим  $(u, v)$  в полярных координатах:

$$u = \rho \sin \varphi, \quad v = \rho \cos \varphi,$$

где  $\rho$  — безразмерный ларморовский радиус. С помощью производящей функции

$$F = (\varphi - \alpha_q \tau) I, \quad I = 1/2 \alpha_q \rho^2$$

осуществляется переход в систему координат, вращающуюся с частотой  $\alpha$ . В новых переменных гамильтониан имеет вид

$$\mathcal{H} = H + \frac{\partial F}{\partial \tau} = -K \cos [\rho \cos (\varphi - \alpha_q \tau)] \sum_{n=-\infty}^{+\infty} \delta(\tau - n). \quad (3.13)$$

Преобразуем ряд  $\delta$ -функций:

$$\sum_{n=-\infty}^{+\infty} \delta(\tau - n) = \sum_{j=1}^q \sum_{m=-\infty}^{+\infty} \delta(\tau - (mq + j)).$$



Подставляя это выражение в (3.13) и пользуясь представлением

$$\sum_{m=-\infty}^{+\infty} \delta(\tau - j - mq) = \frac{1}{q} \sum_{m=-\infty}^{+\infty} \exp\left(2\pi im \frac{\tau - j}{q}\right).$$

легко находим

$$\tilde{H} = H_q + V_q, \quad (3.14)$$

$$H_q = -\frac{1}{q} K_H \sum_{j=1}^q \cos \xi_j,$$

$$V_q = -\frac{2}{q} K_H \sum_{j=1}^q \cos \xi_j \sum_{m=1}^{\infty} \cos \left[ \frac{2\pi m}{q} (\tau - j) \right],$$

где

$$\xi_j = \rho \cos \left( \varphi + \frac{2\pi}{q} j \right) = v \cos \left( \frac{2\pi}{q} j \right) + u \sin \left( \frac{2\pi}{q} j \right). \quad (3.15)$$

Величина  $\xi_j = \rho e_j$ , где  $\rho$  — вектор состояния на фазовой плоскости с координатами  $(u, v)$ ;  $e_j$  — единичный вектор, определяющий вершину правильного  $q$ -угольника.

Выражение  $H_q$  будем называть резонансным гамильтонианом. Он определяет некоторую интегрируемую гамильтоновскую систему. Второй член в  $\tilde{H}$  описывает возмущение  $V_q$ , действующее на  $H_q$ . Это возмущение, в частности, разрушает сепаратрисы, образуя узкие области хаотической динамики [18].

Во всех случаях резонанса, в которых  $q \neq 2, 3, 4, 6$ , возникает нетривиальная ситуация. При этом нельзя покрыть плоскость сеткой, которая бы обладала как трансляционной, так и вращательной симметриями одновременно. Такая структура стохастической паутины является очень специфической и относится к категории структур с симметрией так называемого квазикристаллического типа. Примеры паутины для  $q$ , равного 5, 7 и 8, приведены на рис. 6. Она представляет собой множество точек отображения  $\tilde{M}_q$  за фиксированный интервал времени. Структуры на этих рисунках образованы случайным блужданием изображающей точки на фазовой плоскости. Поскольку точка не может выйти за пределы каналов паутины, то за достаточно большое время мы приобретаем достаточно детальную информацию о структуре паутины. Однако эта структура проявляется очень неравномерно во времени.

На рис. 6, *a* приведена паутина, которая достаточно полно изображена за время ее наблюдения (время счета). Совсем другая картина имеет место на рис. 6, *б*. В течение более чем  $10^6$  шагов отображения внутри паутины остается окно в форме фрактальной пятиугольной звезды. Окно зарастает паутиной на временах  $\tau > 1,3 \cdot 10^6$ . Появление окон разных размеров и форм связано с выбором начальных условий. Последовательность формирования

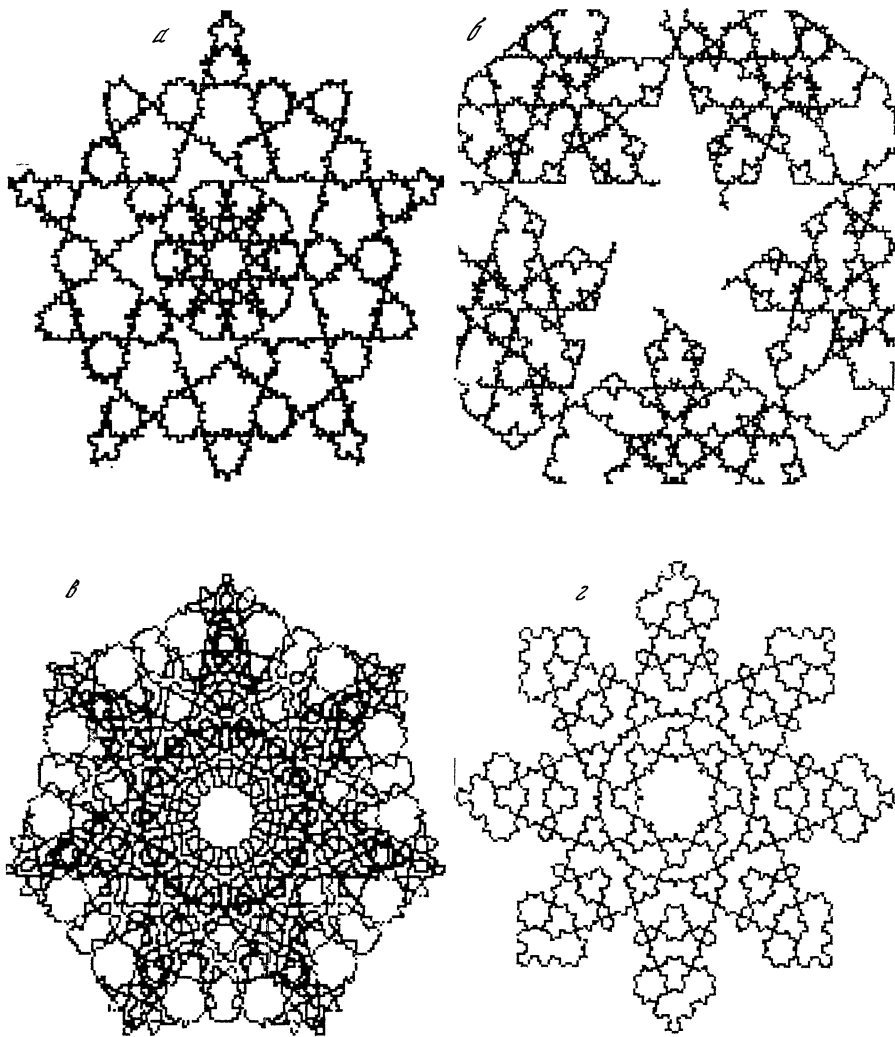


Рис. 6. Стохастическая паутина в случае симметрии 5-го (*а, б*), 7-го (*в*) и 8-го (*г*) порядков

*а, б* —  $K_H = 0,7$ , размер квадрата  $256\pi \times 256\pi$ ; *в* —  $K_H = 0,5$ , размер квадрата  $64\pi \times 64\pi$ ; *г* —  $K_H = 0,6$ , размер квадрата  $128\pi \times 128\pi$

«снежинки» из паутины очень чувствительна к начальным значениям параметров, а это, в свою очередь, связано с существованием канторотов [19], затрудняющих некоторые направления диффузии.

Рост паутины происходит следующим образом. Сначала случайные блуждания точки на плоскости создают некоторую неправильную фигуру типа, например, звезды, у которой не все концы

одинаково выросли. Затем эта звезда достраивается до правильной, и следующим этапом роста паутины начинается формирование звезды большего размера. Конечно, границы этих звезд имеют сложную форму и образуют фрактальные кривые в пределе  $t \rightarrow \infty$ . Поэтому мы будем также называть стохастическую паутину с квазикристаллической симметрией фрактальной паутиной.

Фрактальная паутина имеет центральное незакрытое окно в форме, близкой к правильному  $q$ -угольнику при четных  $q$  и близкой к  $2q$ -угольнику при нечетных  $q$ . Внутри центрального окна также есть сепаратрисы и стохастические слои, которые, однако, не выходят на край окна и не соединяются с основной паутиной. Это связано с тем, что фазовое пространство внутри окна больше всего напоминает по своей структуре фазовый портрет стандартного отображения [8, 14]. Чем меньше  $K_H$ , тем больше размер центрального окна [5]. С уменьшением  $K_H$ , а также с ростом  $q$  паутина становится тоньше, а ее ячейки шире.

Фурье-анализ является важным средством изучения структурных свойств стохастической паутины и рельефов на фазовой плоскости. Кроме того, рентгеновский анализ реальных кристаллов позволяет судить об их симметричных свойствах. Картина фурье-спектра структуры в случае симметрии 5-го порядка (покрытия Пенроуза) появилась в [20] до того, как аналогичная картина была получена в экспериментах [21]. Все эти данные свидетельствуют о том, что фурье-спектр с картины покрытия плоскости достаточно хорошо передает симметричные свойства покрытия, хотя и не дает возможности его однозначно восстановить.

Обозначим через  $S_\Gamma$  множество точек, принадлежащих на фазовой плоскости размером  $\Gamma$  некоторому рисунку (структура). Пусть также

$$\delta(\mathbf{R} - \mathbf{R}_s) = \begin{cases} 1, & s \in S_\Gamma, \\ 0, & s \notin S_\Gamma, \end{cases}$$

где  $\mathbf{R}$  — вектор произвольной точки на плоскости  $\Gamma$ ;  $\mathbf{R}_s$  — вектор фиксированной точки  $s$ . Тогда фурье-образ структуры определяется выражением

$$S(\mathbf{k}) = \lim_{\Gamma \rightarrow \infty} \frac{1}{(2\pi)^2} \int e^{i\mathbf{k}\mathbf{R}} \delta(\mathbf{R} - \mathbf{R}_s) d\mathbf{R}. \quad (3.16)$$

Реально мы всегда имеем дело с конечной областью  $\Gamma$ . Это создает дополнительные граничные эффекты в виде  $S(k)$ . В случае периодических структур  $S$  их иногда легко выделить. Однако в случае неперiodических покрытий плоскости выделить «монокристалл» оказывается невозможным. Здесь все же можно было бы использовать свойство поворотной симметрии и свойство подобия структуры квазикристаллического типа.

На рис. 7 и 8 приведены фурье-спектры соответственно для стохастической паутины, порождаемой отображением  $M_5$ , и для энергетического рельефа той же сглаженной структуры, порождаемой усредненным гамильтонианом  $\bar{H}_5$ . Область  $\Gamma$  выбиралась

в виде круга, что создает достаточно хорошее приближение к паутине на рис. 7. Картины на рис. 7 и 8 близки друг к другу. Это важное следствие, так как оно подтверждает возможность введения симметричного анализа бесконечных паутин динамических систем, подобно тому, как это делается для кристаллов или квазикристаллов. Аналогичный пример приведен на рис. 9 и 10 для  $q = 7$ .

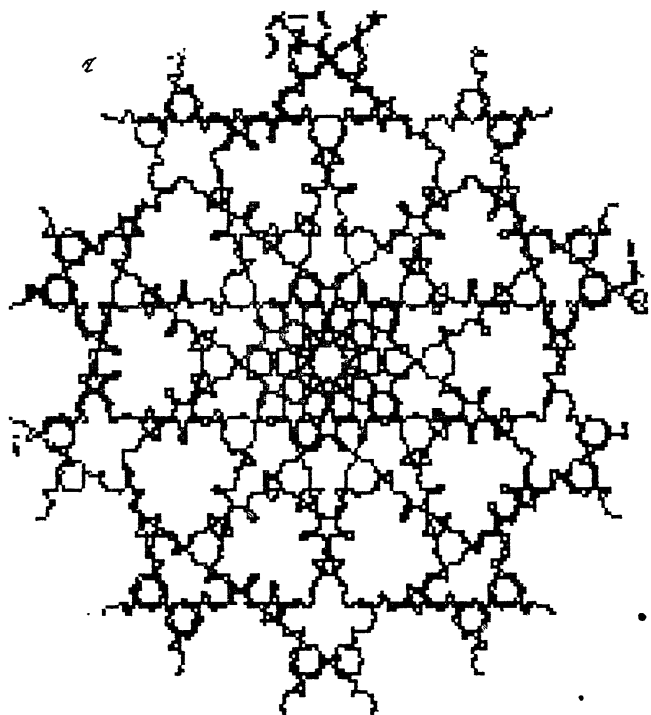
#### 4. О КВАЗИКРИСТАЛЛИЧЕСКОЙ СИММЕТРИИ

Сейчас ясно, что мы столкнулись с новым и еще мало исследованным видом симметрии. Генератор покрытий  $\hat{M}_q$  или усредненный гамильтониан  $\hat{H}_q$  являются возможными средствами выражения этой симметрии. И здесь естествен вопрос, в какой мере симметрия квазикристаллического типа универсальна и где, кроме реальных квазикристаллов, мы можем ее наблюдать.

Возникновение структур в различных гидродинамических течениях является одним из наиболее ранних экспериментальных наблюдений. Среди них наиболее правильная картина связана с конвективными ячейками, которые могут принимать форму одномерных валов или создавать на поверхности слоя квадратную или гексагональную сетку (ячейки Бенара). Общее понимание явления возникновения структур связывает его с различными стадиями изменения состояния среды при переходе ее к турбулентному движению. Чем более правильная и более простая структура, тем выше ее симметрия и тем более упорядоченным является состояние среды.

Турбулентное состояние наиболее однородно (например, вследствие локальной неустойчивости), и симметрия его значительно выше, чем симметрия структурно упорядоченной среды. Максимальной симметрией обладает сильно развитая турбулентность. Путь зарождения турбулентности чрезвычайно сложен. Он определяется не только сложной последовательностью бифуркаций во времени, приводящих в конце концов к хаосу и к непрерывному пространственному спектру, но и не менее сложной последовательностью пространственных перестроек среды, ведущих к возникновению пространственно неупорядоченных структур. Последнее обстоятельство приводит к появлению непрерывного спектра уже по пространственным координатам. Таким образом, картина возникновения турбулентности есть сложный путь рождения пространственно-временного хаоса. Эти идеи все чаще появляются в работах, относящихся к данному вопросу (см., например, [22—25]).

Существует много экспериментальных и численных исследований возможных структурных организаций среды (см., например, сборники обзоров [26—28]), хотя это материалы далеко не всех соведущих на эту тему). Легче всего поддаются анализу различные двумерные структуры, возникающие в задачах тепловой конвекции или электродинамической конвекции, которая может также включать внешнее параметрическое возбуждение. Круг подобных



*б*

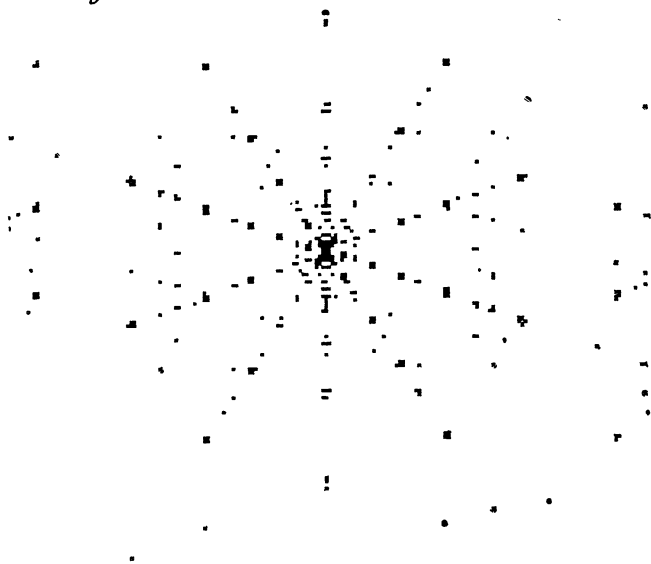
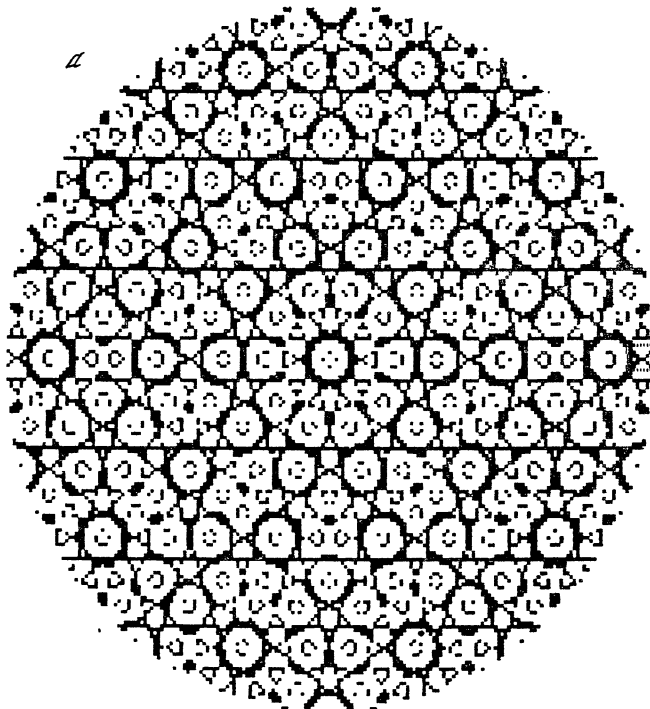


Рис. 7. Элемент паутины для  $q = 5$  (а) (радиус круга  $50\lambda$ ) и его фурье-образ (б)



б

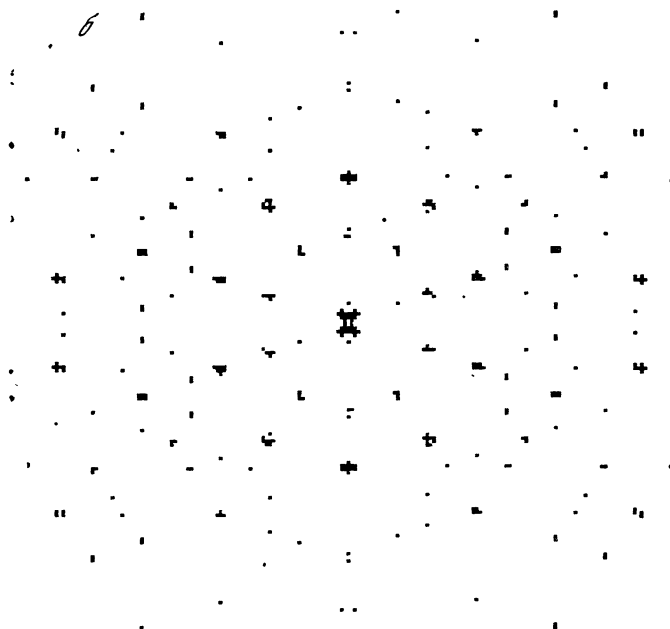
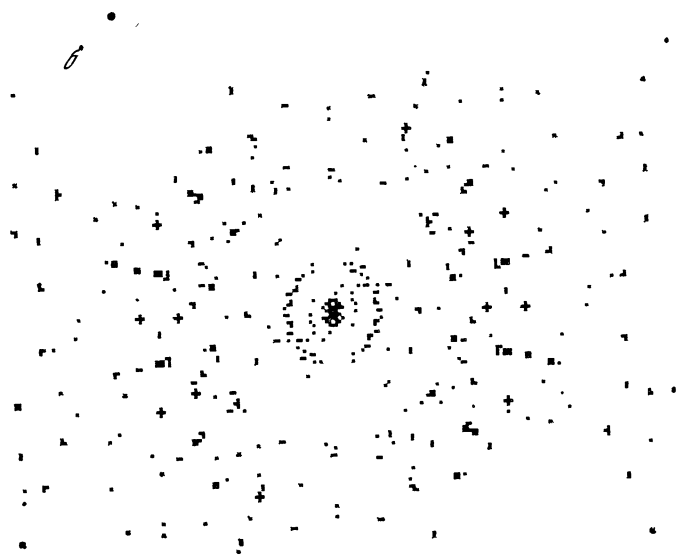
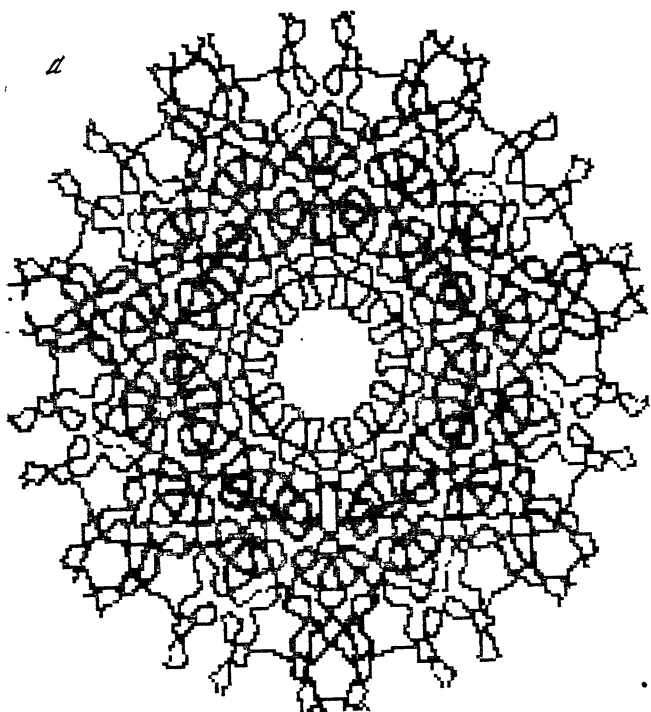
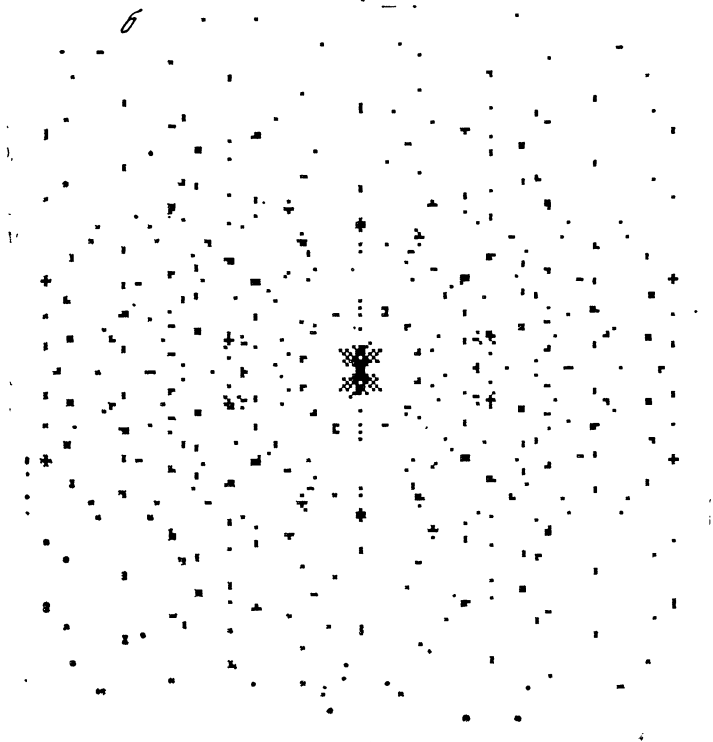
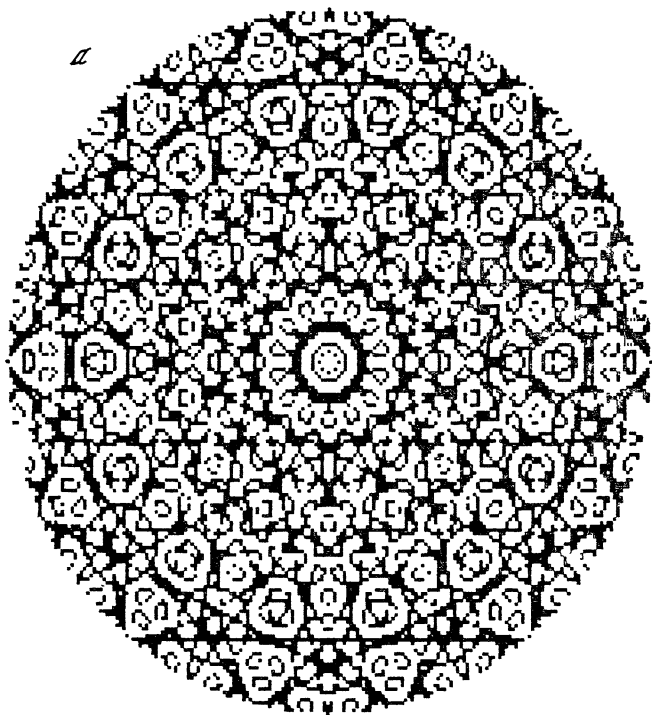


Рис. 8. Элемент энергетического рельефа для  $q = 5$  (а) (радиус круга  $32\pi$ ) и его фурье-образ (б)



**Рис. 9.** Элемент паутины для  $q = 7$  (а) (радиус круга  $24\pi$ ) и его фурье-образ (б)

**Рис. 10.** Элемент энергетического рельефа для  $q = 7$  (а) (радиус круга  $32\pi$ ) и его фурье-образ (б)





задач чрезвычайно широк: строение пены, конвективные ячейки в атмосфере Юпитера, структуры в вихревых решетках, в сдвиговых течениях и др. На этом пути структуры с квазикристаллическим видом симметрии могут сыграть важную роль в гидродинамических средах при анализе возникновения турбулентности [18]. Поясним, почему это возможно.

Обратимся к двумерной гидродинамике несжимаемой жидкости. Уравнение движения имеет вид

$$\frac{\partial}{\partial t} \left( \Delta \psi - \frac{1}{\text{Re}} \Delta^2 \psi \right) + \frac{\partial (\Delta \psi, \psi)}{\partial (x, y)} = 0, \quad (4.1)$$

где  $\psi$  — функция тока ( $v_x = \partial \psi / \partial y$ ,  $v_y = -\partial \psi / \partial x$ );  $\Delta$  — двумерный оператор Лапласа;  $\text{Re}$  — число Рейнольдса.

Рассмотрим теперь  $\psi$  в следующем виде:

$$\psi_k = A(t) \sum_{j=1}^q \cos(k e_j \mathbf{R}). \quad (4.2)$$

Выражение (4.2) для функции  $\psi$ , которое можно назвать «квазоном», обладает следующим свойством:

$$\Delta \psi_k + k^2 \psi_k = 0. \quad (4.3)$$

В силу (4.3) нелинейный член в (4.1) тождественно обращается в нуль, и

$$\dot{A} = - (k^2 / \text{Re}) A,$$

т. е.

$$A(t) = A(0) \exp(-k^2 t / \text{Re}). \quad (4.4)$$

Если, однако, имеется накачка, то квазон может оказаться стационарным решением, и вопрос теперь заключается в определении области устойчивости подобных решений. Она может быть связана с разными неучтенными факторами (сжимаемость, теплопроводность и т. д.). Однако априори нам не следует исключать решений, подобных (4.2), тем более, что они содержат квадратные и гексагональные сетки, как частные случаи.

Примитивные покрытия плоскости квадратной или шестиугольной сетками достаточно распространены в природе. Поэтому нет ничего удивительного в том, что мы их часто встречаем в гидродинамическом эксперименте. Квазикристаллическая симметрия не столь очевидна, хотя и не столь уже редко встречается в природе. Она возникает в результате реального взаимодействия вращательной и трансляционной симметрий, и наличие динамической модели этого взаимодействия позволяет значительно шире взглянуть как на свойства динамических систем, так и на свойства структур. Переход от регулярных структур к пространственному хаосу может сопровождаться последовательностью пространственных бифуркаций, и среди них развитие структур с квазикристаллической симметрией может быть вполне вероятным.

## ЛИТЕРАТУРА

1. Пуанкаре А. Новые методы небесной механики. Избр. тр.: В 2 т. М.: Наука, 1971. Т. 1. 522 с.; 1972. Т. 2. 771 с.
2. Мельников В. К. Об устойчивости центра при периподических во времени возмущениях // Тр. Моск. мат. о-ва. 1963. Т. 12. С. 3—52.
3. Filonenko N. N., Sagdeev R. Z., Zaslavsky G. M. Destruction of magnetic surfaces by magnetic field irregularities // Nucl. Fusion. 1967. Vol. 7. P. 253—266.
4. Заславский Г. М., Филоненко Н. Н. Стохастическая неустойчивость захваченных частиц и условия применимости квазилинейного приближения // ЖЭТФ. 1968. Т. 54, № 5. С. 1590—1602.
5. Заславский Г. М., Захаров М. Ю., Сагдеев Р. З. и др. Стохастическая паутина и диффузия частиц в магнитном поле // Там же. 1986. Т. 91, № 9. С. 500—516.
6. Chernikov A. A., Sagdeev R. Z., Usikov D. A. et al. Minimal chaos and stochastic webs // Nature. 1987. Vol. 326. P. 559—563.
7. Заславский Г. М. Стохастичность динамических систем. М.: Наука, 1984. 272 с.
8. Chirikov B. V. A universal instability of many-dimensional oscillator systems // Phys. Rep. 1979. Vol. 52. P. 265—379.
9. Арнольд В. И. Математические методы классической механики. М.: Наука, 1974. 432 с.
10. Арнольд В. И. О неустойчивости динамических систем со многими степенями свободы // ДАН СССР. 1964. Т. 156, № 1. С. 9—12.
11. Karney C. F. F. Stochastic heating by a lower hybrid wave // Phys. Fluids. 1978. Vol. 21. P. 1584—1599; 1979. Vol. 22. P. 2186—2209.
12. Riyopoulos S., Antonsen T. M., Ott E. Low-amplitude, wave-induced particle energy diffusion in an inhomogeneous magnetic field // Ibid. 1984. Vol. 27. P. 184—191.
13. Fukuyama A., Motota H., Itatani R., Takizuka T. Stochastic acceleration by an electrostatic wave near ion cyclotron harmonics // Phys. Rev. Lett. 1977. Vol. 38. P. 701—704.
14. Лихтенберг А., Либрман М. Регулярная и стохастическая динамика. М.: Мир, 1984. 528 с.
15. Сагдеев Р. З., Шапиро В. Д. Влияние поперечного магнитного поля на затухание Ландау // Письма в ЖЭТФ. 1973. Т. 17, № 7. С. 389—394.
16. Sagdeev R. Z., Zaslavsky G. M. Regular and chaotic dynamics of particles in a magnetic field // Nonlinear phenomena in plasma physics and hydrodynamics / Ed. R. Z. Sagdeev. Moscow: Mir, 1968. P. 65—92.
17. Chernikov A. A., Natanzon M. Ya., Petrovichev B. A. et al. Some peculiarities of stochastic layer and stochastic web formation // Phys. Lett. A. 1987. Vol. 122. P. 39—46.
18. Заславский Г. М., Захаров М. Ю., Сагдеев Р. З. и др. Генерация упорядоченных структур с осью симметрии из гамильтоновской динамики // Письма в ЖЭТФ. 1986. Т. 44, № 7. С. 349—353.
19. Mac Kay R. S., Meiss I. D., Percival I. C. Transport in hamiltonian systems // Physica. D. 1984. Vol. 13. P. 55—81.
20. Mac Kay A. L. Post-classical crystallography // J. Phys. Coll. C3. 1986. Vol. 47, N 7, suppl. P. 153—163.
21. Chechtman D., Blech I., Gratiias D., Cahn J. W. Metallic phase with long-range orientational order and no translational symmetry // Phys. Rev. Lett. 1984. Vol. 53. P. 1951—1953.
22. Sagdeev R. Z., Usikov D. A., Zaslavsky G. M. Nonlinear physics. N. Y.: Harwood Acad. Publ., 1987. 675 с.
23. Newell A. C. Chaos and turbulence: is there a connection? // Perspectives in nonlinear dynamics / Ed. M. F. Shlesinger et al. Singapore: World Sci. 1986. P. 38—70.
24. Езерский А. В., Рабинович М. И., Реутов В. П., Старобинец И. М. Пространственно-временной хаос при параметрическом возбуждении капиллярной ряби // ЖЭТФ. 1986. Т. 91, № 6. С. 2070—2083.

- 25 Chate H., Manneville P. Transition to turbulence via spatiotemporal intermittency // Phys. Rev. Lett. 1987. Vol. 587. P. 112—115.
26. Cellular structures in instabilities / Ed. T. E. Wesfreid, S. Zaleski. Berlin: Springer. 1984. 389 p.
27. Perspectives in nonlinear dynamics / Ed. M. F. Schlesinger et al. Singapore: World Sci. 1986.
28. Physica D. 1986. Vol. 23. N 1/3.

УДК 536.25

## ДИНАМИКА ДЕФЕКТОВ И ВОЗНИКНОВЕНИЕ ПРОСТРАНСТВЕННОГО ХАОСА В ОДНОМЕРНЫХ СИСТЕМАХ

А. А. НЕПОМНЯЩИЙ

Известно, что диссипативные структуры часто являются пространственно-неупорядоченными. Примером может служить конвекция Рэлея—Бенара в горизонтальном слое, для которой характерны сложные текстуры, включающие области из искривленных валов, разделенные доменными границами, дислокации, дисинклинации и иные виды дефектов [1]. В данной статье рассмотрен ряд гораздо более простых систем, описываемых одномерными уравнениями, для которых динамика дефектов и возникновение пространственного хаоса допускают аналитическое описание.<sup>\*\*\*</sup>

### [1. ОСНОВНАЯ МОДЕЛЬ

Рассмотрим модельную систему, которая описывается вещественным параметром порядка  $A(x, t)$ , подчиняющимся уравнению

$$A_t = A_{xx} + A - A^3; \quad x \rightarrow \pm \infty: |A| < \infty \quad (1)$$

(индекс означает дифференцирование по соответствующей координате). К виду (1) посредством некоторого преобразования масштабов приводится уравнение для амплитудной функции, определяющее нелинейную эволюцию длинноволновой моды конвективной неустойчивости равновесия в горизонтальном цилиндре. В этом случае знак и величина  $A(x)$  описывают направление и интенсивность вращения жидкости в сечении цилиндра, соответствующем данному значению масштабированной продольной координаты  $x$ .

Для уравнения (1) существует функционал, в дальнейшем называемый энергией,

$$U\{A(x)\} = \int [(A^2 - 1)^2/4 + A_x^2/2] dx, \quad (2)$$

для которого

$$A_t = -\delta U/\delta A, \quad U_t = -\int A_t^2 dx \leq 0.$$

Отсюда следует, что при  $t \rightarrow \infty$  система стремится к некоторому стационарному состоянию. Уравнение  $\delta U = 0$ , описывающее

стационарные состояния, формально эквивалентно уравнению одномерного движения частицы в поле с потенциалом  $V(A) = - (A^2 - 1)^2/4$  ( $A$  — координата,  $x$  — время); функционал (2) играет роль действия частицы. Ограниченные при  $x \rightarrow \pm \infty$  решения задачи выражаются через эллиптический синус Якоби и являются периодическими по  $x$ , за исключением решений  $A = 0$  (состояние равновесия),  $A = \pm 1$  (однородное движение жидкости) и сепаратрисных решений

$$A = \pm \operatorname{th} [(x - \xi)/\sqrt{2}]_{\mathfrak{K}}, \quad \xi = \text{const}, \quad (3)$$

которым соответствуют неподвижные дефекты — доменные стенки между областями с противоположным направлением вращения жидкости.

Эквивалентность между «временной» и «пространственной» интерпретацией уравнения  $\delta U = 0$  нарушается при переходе к исследованию устойчивости решений. Устойчивость стационарной структуры определяется спектром задачи для малых возмущений  $\bar{A}(x) \exp \sigma t$ , наложенных на движение  $A(x)$ :

$$\sigma \bar{A} = \bar{A}_{xx} + (1 - 3A^2)\bar{A}; \quad x \rightarrow \pm \infty: |\bar{A}| < \infty. \quad (4)$$

Структура устойчива, если все  $\sigma \leq 0$ , что эквивалентно условию  $\delta^2 U \geq 0$ .

Нетрудно установить неустойчивость «эллиптической» точки  $A = 0$  и устойчивость «гиперболических» точек  $A = \pm 1$ . Устойчивость остальных решений легко определяется на основе известного свойства собственных функций одномерного уравнения Шредингера: волновая функция основного состояния не имеет узлов. Поскольку задача (4) всегда обладает решением  $\bar{A} = A_x$ ,  $\sigma = 0$ , то периодические структуры со знакопеременной функцией  $A_x$  неустойчивы, а решения (3), для которых функция  $A_x$  знакопостоянная, устойчивы. Отметим, однако, что для периодических стационарных решений с периодом  $2L \gg 1$ , описывающих периодическую решетку доменных стенок на расстоянии  $L$  одна от другой, инкремент нарастания моды, вызывающей неустойчивость, экспоненциально мал:  $\sigma \sim \exp(-L\sqrt{2})$ . Поэтому естественно ожидать, что относительно слабое внешнее воздействие может вызвать их стабилизацию.

В качестве примера подобного воздействия рассмотрим пространственно-периодическое изменение локального числа Рэлея вдоль оси цилиндра, которое приводит к следующей модификации уравнения (1):

$$A_t = A_{xx} + [1 + \varepsilon f(x)] A - A^3, \quad f(x + L_0) = f(x). \quad (5)$$

Не зависящие от  $A$  дополнительные члены в амплитудном уравнении, связанные с появлением продольной компоненты градиента температуры, имеют более высокий порядок малости и здесь не учитываются.

Задача описания стационарных конфигураций поля  $A(x)$  эквивалентна задаче о колебаниях нелинейного осциллятора,

параметрически возбуждаемого периодической внешней силой. Конфигурации удобно рассматривать с помощью сечения Пуанкаре в форме дискретного набора точек  $\{u_n = A(x_n), v_n = A_x(x_n), x_n = x_0 + nL_0\}$ , именуемого далее «траекторией». При этом дифференциальное уравнение для  $A(x)$  сводится к некоторому обратимому двумерному отображению  $(u_n, v_n) \rightarrow (u_{n+1}, v_{n+1})$ , сохраняющему меру  $du dv$ , которое, однако, определено не на всей плоскости  $(u, v)$ . Напомним, что сведение к дискретному отображению широко использовалось при анализе стационарных конфигураций дискретных систем (см., например, [2—5]).

Рассматриваемая задача является неинтегрируемой и обладает следующими классами траекторий [6]:

1) квазипериодические траектории, характеризующиеся иррациональным числом вращения  $\omega$ , которым отвечают несоизмеримые структуры, возникающие в результате воздействия внешнего возмущения с периодом  $L_0$  на решетку с несоизмеримым периодом  $L$ . При выполнении условий теоремы Колмогорова—Арнольда—Мозера (КАМ) этот класс траекторий при данном  $\omega$  плотно заполняет некоторую непрерывную замкнутую кривую (гладкий тор КАМ) и может быть непрерывно параметризован, например, углом  $\varphi_0 = \text{Arctg}(v_0/u_0)$ . При достижении некоторого критического значения  $\varepsilon_*(\omega)$  непрерывность кривой нарушается [2];

2) периодические траектории (соизмеримые структуры), для которых  $(u_{n+N}, v_{n+N}) = (u_n, v_n)$  для некоторого целого  $N \geq 1$ . В соответствии с поведением близких траекторий среди них могут быть выделены эллиптические и гиперболические;

3) гомо- и гетероклинические траектории, описывающие структуры, асимптотически переходящие в периодические при  $x \rightarrow \pm \infty$ , но содержащие при конечных  $x$  некоторые «искажения» структуры;

4) хаотические траектории, которым отвечают структуры, содержащие некоторый неупорядоченный набор доменных стенок.

Из континуума решений, параметризуемого величинами  $(u_0, v_0)$ , необходимо отобрать устойчивые решения, удовлетворяющие условию  $\delta^2 U \geq 0$  (функционал  $U$  по сравнению с (2) включает дополнительное слагаемое  $(\varepsilon/2) \int f(A^2 - 1) dx$ ). Удастся сформулировать два общих утверждения, относительно устойчивости решений:

1) неустойчивы несоизмеримые структуры на гладких торах КАМ (доказательство аналогично доказательству неустойчивости решеток при  $\varepsilon = 0$ , только вместо решения  $\tilde{A} = A_x$  следует взять  $\tilde{A} = \partial A / \partial \varphi_0$ );

2) неустойчивы соизмеримые структуры, соответствующие «эллиптическим» периодическим траекториям, для которых непрерывный спектр уравнения Шредингера включает точку  $\sigma = 0$ .

## 2. ДИНАМИКА ДОМЕННЫХ СТЕНОК

Задача существенно упрощается в случае  $\varepsilon \ll 1$ ,  $L_0 \gg 1$  для структур, фазовые траектории которых лежат в окрестности разрушенных сепаратрис точек  $A = \pm 1$ . Именно для таких структур, которые представляют собой набор удаленных друг от друга доменных стенок, естественно ожидать стабилизацию при малых  $\varepsilon$ . В этом случае можно применить «многосолитонную теорию возмущений», аналогичную разработанной для возмущенного уравнения  $\sin$ -Гордона [7, 8]. Нестационарное решение уравнения (5) описывается совокупностью доменных стенок, движение центров которых  $\xi_i(t)$  подчиняется уравнениям движения

$$(2\sqrt{2}/3)\ddot{\xi}_i(t) = -\partial U_w/\partial \xi_i, \quad (6)$$

где энергия  $U_w$  включает собственную энергию доменной стенки, энергию взаимодействия доменной стенки с внешней неоднородностью и энергию взаимодействия между соседними стенками:

$$U_w(\{\xi_i\}) = \sum_i V(\xi_i) + \sum_i W(\xi_i - \xi_{i-1}), \quad (7)$$

где

$$V(\xi_i) = \frac{2\sqrt{2}}{3} + \frac{\varepsilon}{2} \int_{-\infty}^{\infty} f(x) \operatorname{ch}^{-2} \frac{\xi_i - x}{\sqrt{2}} dx,$$

$$W(\xi_i - \xi_{i-1}) = -8\sqrt{2} \exp[-(\xi_i - \xi_{i-1})\sqrt{2}] \quad (\xi_i > \xi_{i-1}).$$

Уравнения (6) становятся неприменимыми при сближении доменных стенок на расстояние порядка единицы и не описывают процесса аннигиляции, однако пригодны для расчета статических конфигураций и анализа их устойчивости.

Модель (7) обладает сходством с известной моделью Френкеля—Конторовой (см., например, [2]). Наиболее существенным отличием является экспоненциальное убывание силы взаимодействия между «частицами», которое приводит к возможности разрывов «цепочек».

Использование модели (7) позволяет рассчитать для заданной плотности доменных стенок  $1/L$  критическое значение параметра  $\varepsilon = \varepsilon_1$ , при котором происходит стабилизация регулярной решетки доменных стенок, захваченных притягивающими неоднородностями, и значение параметра  $\varepsilon = \varepsilon_2$ , при котором становятся возможными устойчивые хаотические конфигурации с «разрывами цепочки». Так, в случае  $f(x) = -\sum_n \delta(x - nL_0)$  (периодическая последовательность локализованных областей с пониженным числом Рэлея)

$$\varepsilon_1 = \varepsilon_2 = 128\sqrt{2} \exp(-L\sqrt{2}).$$

Можно показать также, что структура, содержащая неупорядоченный набор захваченных доменных стенок на произвольных

расстояниях, больших, чем  $L_*(\varepsilon) \approx (1/\sqrt{2}) \ln(\varepsilon/197)$ , является устойчивой (вследствие слабости притяжения между дефектами).

Заметим, что при неоднородности типа (5) наиболее низкоэнергетическими остаются бездефектные состояния с постоянным знаком  $A$ , а при заданной плотности дефектов — периодические структуры. Этот вывод сохраняется и при конечных  $\varepsilon$  и  $L_0$ .

$$A_t = A_{xx} + A - A^3 + \varepsilon g(x) \quad (8)$$

(пространственно-периодический подогрев сбоку). Возможность появления пространственного хаоса в системе (8) обсуждалась в [9].

Пусть  $g(-x) = -g(x)$ , так что внешнее возмущение не делает предпочтительным определенное направление движения жидкости. При малых  $\varepsilon$  результаты качественно не отличаются от случая, рассмотренного выше. Однако с ростом  $\varepsilon$  энергии захваченной доменной стенки при некотором  $\varepsilon = \varepsilon_3$  может стать отрицательной. При  $\varepsilon > \varepsilon_3$  состоянию с наименьшей энергией отвечает некоторая периодическая конфигурация доменных стенок. С ростом  $\varepsilon$  между этими конфигурациями имеет место бесконечная последовательность «фазовых переходов», завершающаяся переходом в оптимальное периодическое состояние, в котором знак функции  $A(x)$  в основном повторяет знак  $g(x)$ . Устойчивые хаотические состояния в этой области представляют собой объединение фрагментов различных структур и обладают более высокой энергией. Как показывают численные расчеты для  $g(x) = \sin kx$ , при некотором  $\varepsilon = \varepsilon_4(k)$  знакопостоянные состояния исчезают. При более высоких  $\varepsilon$  устойчивых хаотических состояний в системе не обнаруживается, а единственной устойчивой конфигурацией является оптимальная периодическая.

В заключение данного раздела заметим, что в случае асимметрии полости относительно вертикальной плоскости амплитудное уравнение (1) в высших порядках содержит слагаемое типа  $\delta A^4$ , которое нарушает симметрию  $A \rightarrow -A$  и энергетическую эквивалентность однородных состояний с  $A > 0$  и  $A < 0$ , что приводит к встречному движению доменных стенок и их аннигиляции с исчезновением «невыгодных» доменов. В этом случае стационарные структуры могут возникать лишь при превышении величиной  $\varepsilon$  некоторого порога пиннинга  $\varepsilon^* = O(\delta)$ . Аналогичное условие имеет место при прокачке жидкости вдоль цилиндра, которой соответствует слагаемое  $-uA_x$  в правой части амплитудного уравнения.

### 3. ДРУГИЕ МОДЕЛИ

Рассмотренная выше модель характеризуется существованием всего двух типов дефектов — доменных стенок (3), которые обязательно чередуются между собой. Несколько более сложная ситуация имеет место для возмущенного уравнения син-Гордона

$$\Phi_{tt} + \alpha\Phi_t - \Phi_{xx} + [1 + \varepsilon f(x)] \sin \Phi = 0, \quad f(x + L_0) = f(x), \quad (9)$$

применяемого для описания распределения магнитного потока в длинных джозефсоновских переходах [7]. Стационарное уравнение (9) при  $\varepsilon = 0$  также обладает двумя типами статических солитонных решений:

$$\Phi = 4 \operatorname{arctg} \exp [\pm (x - \xi)],$$

для которых  $\Phi(+\infty) - \Phi(-\infty) = \pm 2\pi$  (флаксон и антифлаксон). В отличие от уравнения (1) уравнение (9) допускает не только неустойчивые периодические решения с солитонами чередующихся знаков, но и периодические цепочки солитонов одного знака, которые в силу знакопостоянства  $\Phi_x$  устойчивы. При  $\varepsilon \ll 1$ ,  $L_0 \gg 1$  расчет стационарных конфигураций и определение порога хаотизации могут быть произведены на основе уравнений движения вида

$$\ddot{\xi}_i + \alpha \dot{\xi}_i + \partial U / \partial \xi_i = 0,$$

получаемых из уравнений, приводимых в [7, 8], в «нерелятивистском» пределе  $|\dot{\xi}_i| \ll 1$ . Для случая системы микронеоднородностей  $f(x) = \pm \sum_n \delta(x - nL_0)$  порог появления устойчивых хаотических решеток флаксонов одного знака независимо от знака неоднородностей оказывается равным

$$\varepsilon = \varepsilon_1 = 12 \sqrt{3} \exp(-L),$$

где  $L$  — среднее расстояние между флаксонами.

Аналогичную зависимость от  $L$  обнаруживают пороги стабилизации и хаотизации для систем флаксонов разного знака. При наличии стороннего тока (дополнительное слагаемое  $\gamma$  в уравнении (9)) стационарные состояния возможны, лишь если  $\varepsilon$  превышает некоторое пороговое значение  $\varepsilon_*$ . Для соизмеримого состояния со средним расстоянием между флаксонами  $L = (p/q)L_0$   $\varepsilon_* = |\gamma|q \cdot 3\pi \sqrt{3}/4$  (для отталкивающих микронеоднородностей).

В заключение обсудим возможность появления пространственного хаоса в задачах, описываемых уравнением Гинзбурга—Ландау для комплексного параметра порядка. Рассмотрим уравнение

$$A_t = A_{xx} + A - |A|^2 A + \varepsilon e^{ikx}, \quad (10)$$

описывающее динамику огибающей системы конвективных валов при коротковолновой неустойчивости равновесия или параллельного течения в условиях пространственно-периодического внешнего воздействия [10, 11]. При  $\varepsilon = 0$  одномерное уравнение Гинзбурга—Ландау не имеет устойчивых решений, описывающих структурные дефекты, поскольку решения (3) и другие сепаратрисные решения оказываются неустойчивыми [12]. Тем не менее устойчивые солитонные решения появляются при  $\varepsilon \neq 0$  [13], и с ними может быть связано развитие пространственного хаоса [9].

Для двумерного уравнения Гинзбурга—Ландау ( $A_{xx}$  в (10) заменено на двумерный лапласиан  $A_{xx} + A_{yy}$ ), описывающего



трехмерные вторичные конвективные течения [14], уже при  $\varepsilon = 0$  существует устойчивый дефект — вихревая нить, которой соответствует дислокация в системе конвективных валов. Ее динамика для уравнения типа (10) существенно отличается от динамики вихревой нити, описываемой уравнением Гинзбурга—Питаевского ( $-iA_t$  вместо  $A_t$ ); в частности, центр вихря движется не вдоль, а поперек направления «сверхтекучей скорости»  $\nabla \arg A$ . Пиннинг дислокаций на неоднородностях также может быть источником появления пространственного хаоса.

Автор выражает глубокую благодарность Д. А. Киржницу, идеи которого легли в основу данной работы. Автор благодарит также М. А. Закса, Ю. А. Кузнецова, Д. В. Любимова, М. И. Рабиновича и Э. Э. Шноля за полезное обсуждение.

#### ЛИТЕРАТУРА

1. Cellular structure in instabilities. Berlin etc.: Springer, 1984. 390 p. (Lect. Notes Phys.; Vol. 210).
2. Aubry S. The twist map, the extended Frenkel-Kontorova model and the devil's staircase // *Physica D*. 1983. Vol. 7, N 1/3. P. 240—258.
3. Burkov S. E., Pokrovsky V. L., Uimin G. Soliton structures in a discrete chain // *J. Phys. A*. 1982. Vol. 15, N 11. P. L645—L648.
4. Jensen M. H., Bak P. Mean-field theory of the three-dimensional anisotropic Ising model as a four-dimensional mapping // *Phys. Rev. B*. 1983. Vol. 27, N 11. P. 6853—6868.
5. Belobrov P. I., Tret'yakov A. G., Zaslavsky G. M. Methods of non-linear dynamics and equilibrium structures of magnetoelastic chains // *J. Stat. Phys.* 1985. Vol. 38, N 1/2. P. 383—404.
6. Лихтенберг А., Либерман М. Регулярная и стохастическая динамика. М.: Мир, 1984. 528 с.
7. Мак-Лафлин Д., Скотт Э. Многосолитонная теория возмущений // Солитоны в действии. М.: Мир, 1981. С. 210—268.
8. Карпман В. И., Рябова Н. А., Соловьев В. В. Взаимодействие флаксонов в длинных джозефсоновских константах // *ЖЭТФ*. 1981. Т. 81, № 4(10). С. 1327—1336.
9. Couillet P., Elphick C., Repaux D. Nature of spatial chaos // *Phys. Rev. Lett.* 1987. Vol. 58, N 5. P. 431—434.
10. Возовой Л. П., Непомнящий А. А. Конвекция в горизонтальном слое жидкости при наличии пространственной модуляции температуры на границах // *Гидродинамика*. Пермь: Перм. гос. пед. ин-т, 1974. Вып. 7. С. 105—118.
11. Возовой Л. П., Непомнящий А. А. Об устойчивости пространственно-периодических конвективных течений в вертикальном слое с волнистыми границами // *ПММ*. 1979. Т. 43, № 6. С. 998—1007.
12. Василенко Ю. Г., Кузнецов Е. А., Львов В. С. и др. О зарождении вихрей Тейлора в течении Куэтта // *ЖПМТФ*. 1980. № 2. С. 58—64.
13. Couillet P. Commensurate-incommensurate transition in non-equilibrium systems // *Phys. Rev. Lett.* 1986. Vol. 56, N 7. P. 724—727.
14. Непомнящий А. А. О вторичных конвективных движениях в плоском вертикальном слое // *Изв. АН СССР. МЖГ*. 1975. № 4. С. 3—11.

## ЯВЛЕНИЯ САМООРГАНИЗАЦИИ В ЛАЗЕРНОЙ ТЕРМОХИМИИ

Ф. В. БУНКИН, Н. А. КИРИЧЕНКО, Б. С. ЛУКЬЯНЧУК

Исследование многих проблем физики, химии, биологии и т. д. инициировало детальное изучение нелинейных уравнений. Одной из важнейших особенностей поведения нелинейных систем оказалось существование явлений самоорганизации, пространственно-временного упорядочения. С формальной точки зрения самоорганизация отвечает наличию в фазовом пространстве изучаемых систем аттракторов — притягивающих (асимптотически устойчивых) множеств. По характеру выполняемых работ исследования явлений самоорганизации можно условно разбить на два больших класса: 1) формальные исследования отдельных нелинейных задач и их обобщения; 2) изучение динамики процессов в конкретных системах. При этом методы и результаты формальных исследований используются при анализе конкретных систем. В свою очередь, специфические подходы, разрабатываемые применительно к конкретным системам, интерпретация содержания уравнений и возможность экспериментального моделирования оказываются полезными при изучении ряда нелинейных краевых задач.

Адекватный подход к изучению нелинейных уравнений сформулирован А. Пуанкаре в виде программы, суть которой состоит в исследовании качественных изменений свойств системы при изменении описывающих ее параметров. Ключевыми пунктами этой программы являются, в частности, динамическая устойчивость, структурная устойчивость, критические множества.

Большинство работ по самоорганизации посвящено системам пространственно-однородным, т. е. с внешними параметрами, не зависящими от координат. В то же время многие реальные системы являются неоднородными. Например, от точки к точке могут меняться внешняя сила, плотность материала, интенсивность воздействующего на среду излучения и т. д. При анализе таких систем возникают следующие основные вопросы: 1) как меняются структуры, формирующиеся в аналогичной однородной системе, при «включении» неоднородности; 2) каковы пути усложнения поведения системы при изменении параметров, характеризующих неоднородность; 3) какие новые типы структур, отсутствующие в однородном случае, возникают благодаря неоднородности?

Ниже на конкретных примерах мы рассмотрим некоторые особенности самоорганизации неоднородных систем.

### 1. БИФУРКАЦИИ СТАЦИОНАРНЫХ РЕШЕНИЙ

Рассмотрим следующую математическую модель, характерную для задач физики горения [1]:

$$\frac{\partial u}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{\partial^2 u}{\partial z^2}, \quad r \geq 0, \quad z > 0, \quad (1)$$

$$\begin{aligned}
 -\partial u / \partial z &= I(r) + \exp(-1/u) - \eta u, \quad r \geq 0, \quad z = 0; \\
 u(r, z, 0) &= u_0(r, z) \geq 0, \quad u \rightarrow 0 \text{ при } \sqrt{r^2 + z^2} \rightarrow \infty.
 \end{aligned}
 \tag{1}$$

Здесь слагаемое  $\exp(-1/u)$  описывает энерговыделение реакции на поверхности массивного образца,  $(-\eta u)$  — теплопотери, а

$$I(r) = I_0 j(r/r_0) \tag{2}$$

— интенсивность лазерного излучения. В ряде случаев можно принять

$$j(r/r_0) = \exp(-r^2/r_0^2). \tag{2'}$$

В данной задаче пространственная неоднородность вносится источником  $I(r)$ , характеризующимся амплитудой  $I_0$  и эффективной шириной  $2r_0$ .

Мы будем рассматривать условия существования стационарных решений сформулированной краевой задачи.

Один из возможных приближенных способов решения подобных краевых задач состоит в выборе естественных характеристик решения и подходящей аппроксимации нелинейности. В случае задачи (1), (2) поле  $u(r, z)$  характеризуется амплитудой  $u = u(0, 0)$  и эффективным радиусом  $a$ . При  $z = 0$  в окрестности точки  $r = 0$

$$u(r, 0) \approx u(1 - r^2/a^2). \tag{3}$$

С учетом этого нелинейное слагаемое аппроксимируем выражением

$$\exp[-1/u(r, 0)] \approx \exp(-1/u) \exp(-r^2/b^2), \quad b^2 = a^2 u. \tag{4}$$

Приближение (4) эквивалентно известному в физике горения разложению экспоненты по Франк-Каменецкому.

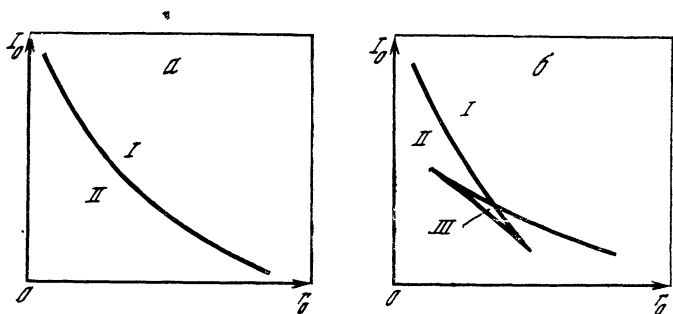
Подставляя (4) в исходную краевую задачу и решая ее относительно  $u(r, z)$ , получим решение в форме  $u(r, z) = f(r, z; I_0, r_0; \eta; u, a)$ , содержащее две неизвестных величины:  $u$  и  $a$ . Теперь, используя исходное приближение (3), получим систему двух трансцендентных уравнений для двух параметров,  $u$  и  $a$  [2]:

$$\begin{aligned}
 u &= f(0, 0; I_0, r_0; \eta; u, a), \\
 -\frac{u}{a^2} &= \frac{1}{2} \frac{\partial^2}{\partial r^2} f(r, 0; I_0, r_0; \eta; u, a) |_{r=0}.
 \end{aligned}
 \tag{5}$$

Оказывается, эта система имеет решения не при всех значениях параметров, и существует так называемая граница устойчивости

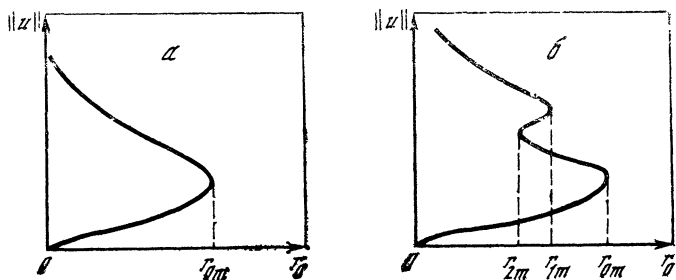
$$I_0 = h(r_0), \tag{6}$$

такая, что при  $I_0 < h(r_0)$  решения имеются, а при  $I_0 > h(r_0)$  решений нет. На рис. 1 показаны два типичных случая разбиения плоскости  $\{I_0, r_0\}$  на области, отличающиеся числом решений. Если  $\eta > \eta_{кр}$ , то плоскость  $\{I_0, r_0\}$  делится на две части, в одной из которых решений нет, а в другой имеются два решения. Если же



**Рис. 1.** Разбиение плоскости параметров  $\{I_0, r_0\}$  на подобласти, отличающиеся числом стационарных решений краевой задачи (1), (2), при  $\eta > \eta_{кр}$  (а) и  $\eta < \eta_{кр}$  (б)

$I$  — решений нет,  $II$  — решений два,  $III$  — решений четыре



**Рис. 2.** Типичные графики ветвления решений при изменении параметра неоднородности  $r_0$  при фиксированном  $I_0$

$\eta < \eta_{кр}$ , то появляется третья подобласть, в которой существуют уже четыре решения системы (5). Тем самым в разных подобластях исходная краевая задача (1), (2) имеет различное число стационарных решений.

На рис. 2 показаны типичные зависимости решения от  $r_0$  при фиксированном  $I_0$ . Как видно из рис. 2, при значениях радиуса локализации источника  $r_0$ , равных  $r_{0m}, r_{1m}, r_{2m}$ , происходит изменение числа стационарных решений краевой задачи (1). Совокупность всех  $r_{im}$  ( $i = 0, 1, 2$ ) для соответствующих значений  $I_0$  и образует бифуркационные границы, показанные на рис. 1.

Из рис. 1 видно, что при изменении  $\eta$  бифуркационная граница  $I_0 = \bar{h}(r_0)$  претерпевает видоизменения, которые характерны для катастрофы типа «ласточкин хвост» (в ряду  $A_n$  касповидных катастроф это  $A_4$ ). Это означает, что построено структурно устойчивое семейство решений, т. е. не слишком большие модификации задачи (1), (2) не приведут к качественным изменениям структуры множества решений — измениться могут лишь количественные значения бифуркационных параметров и характеристик решений.

Таким образом, мы видели, что при изменении параметров источника  $I_0$ ,  $r_0$  меняются как количественные характеристики, так и само число стационарных решений нелинейной краевой задачи, причем в силу структурной устойчивости такое явление характерно и для ряда других задач. Не останавливаясь подробнее на этой задаче, рассмотрим более сложный случай бифуркации решений, когда благодаря неоднородности нелинейная система приобретает способность выходить в режим незатухающих колебаний.

## 2. АВТОКОЛЕБАНИЯ В НЕОДНОРОДНЫХ СИСТЕМАХ

Рассмотрим следующую нелинейную краевую задачу [3]:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \beta_0 n N \exp \left( - \frac{u_0}{u} \right) I(r) - (u - 1), \\ v \frac{\partial (nN)}{\partial t} &= \frac{1}{r} \frac{\partial}{\partial r} \left( r N \frac{\partial n}{\partial r} \right) + \frac{1}{r} \frac{\partial}{\partial r} \left[ r N \alpha \frac{n(1-n)}{u} \frac{\partial u}{\partial r} \right], \\ 0 < r < R; \quad v > 0; \quad \alpha &= \text{const}; \end{aligned} \quad (7)$$

$$N = [u \langle 1/u \rangle]^{-1}, \quad \langle \frac{1}{u} \rangle \equiv \frac{1}{\pi R^2} \int_0^R \frac{1}{u(r, t)} 2\pi r dr;$$

$$\left. \frac{\partial u}{\partial r} \right|_{r=0, R} = 0, \quad \left. \frac{\partial n}{\partial r} \right|_{r=0, R} = 0; \quad u|_{t=0} = 1, \quad n|_{t=0} = n_0.$$

Эта задача описывает в определенных приближениях нагрев двухкомпонентной химически инертной газовой смеси ИК-лазерным излучением и учитывает связь поглощения с концентрацией  $n$  ( $0 < n < 1$ ) поглощающего излучение компонента газа, с температурой  $u$  и с полной плотностью  $N$  газа при граничных условиях, отвечающих тепло- и массоизоляции на границах кюветы. Интенсивность излучения, как и в предыдущем примере, примем в виде (2').

Сформулированная задача содержит закон сохранения

$$\langle nN \rangle \equiv \frac{1}{\pi R^2} \int_0^R 2\pi r n N dr = \langle nN \rangle|_{t=0} = n_0,$$

который устанавливается путем интегрирования второго уравнения в (7) по  $r dr$  и выражает неизменность полного количества каждого из компонентов смеси в ходе процесса.

Если поле излучения пространственно однородно ( $r_0 = \infty$ ), то задача допускает одно или три (в зависимости от значений прочих параметров) однородных стационарных решения, которые находятся из трансцендентного уравнения

$$\beta_0 n_0 \exp(-u_0/u) I_0 = u - 1, \quad n = n_0. \quad (8)$$

Проверка их устойчивости проводится обычным методом: линеаризуя (7) в окрестности некоторого стационарного решения  $\{u = u_c, n = n_0\}$  и полагая  $\{u - u_c, n - n_0\} \sim \exp(pt) J_0(qr)$  ( $J_0(x) -$

функция Бесселя), получим дисперсионное уравнение  $\varphi(p, q) = 0$ . Анализ этого уравнения показывает, что если стационарное однородное решение единственно, то оно устойчиво, если же решений — три, то два устойчивы, а одно (среднее между ними) неустойчиво.

Ситуация качественно меняется, если поле излучения пространственно неоднородно: в определенном диапазоне значений параметров в системе устанавливаются незатухающие колебания переменных. Понять возникновение таких колебаний достаточно просто, если рассмотреть физическое содержание уравнений (7). Прежде всего, второе уравнение в (7) описывает перенос вещества, определяемый тремя механизмами — это диффузия (первое слагаемое), термодиффузия (второе слагаемое) и неоднородное тепловое расширение газа, которое в случае достаточно медленного изменения температуры и, следовательно, в условиях постоянства давления по объему смеси учитывается множителем  $N \neq \text{const}$ . Предположим, что поглощающий излучение компонент смеси тяжелый. Тогда  $\alpha > 0$  и термодиффузионный поток стремится выбросить этот компонент из более нагретых областей среды в менее нагретые. Тепловое расширение неоднородно нагреваемого газа ведет к уменьшению его плотности в более нагретых областях. В результате процесс разыгрывается следующим образом.

Излучение нагревает газ в области с размерами порядка радиуса  $r_0$  пучка. Плотность поглощающего излучение компонента в этой области за счет 2-го и 3-го механизмов массопереноса уменьшается. Это, в свою очередь, ведет к уменьшению поглощательной способности среды, уменьшению количества вводимой в газ энергии и его остыванию. В результате градиенты температуры уменьшаются, плотность газа выравнивается и процесс повторяется. В этом процессе наряду с неоднородностью источника существенной оказывается достаточно сильная температурная зависимость коэффициента поглощения (в модели принято  $\beta(u) = \beta_0 \exp(-u_0/u)$ ), характерная для ряда молекулярных газов в ИК-диапазоне спектра, без которой описанные выше колебания оказываются быстрозатухающими.

Приближенный метод исследования краевой задачи (7), позволяющий установить пороги возникновения автоколебаний, их характерные периоды и частоты, состоит в сведении исходной задачи к системе обыкновенных дифференциальных уравнений для существенных степеней свободы. В качестве таких степеней свободы можно выбрать значения температуры  $u(0, t)$  и плотности поглотителя  $n(0, t)N(0, t)$  в центре кюветы. Запишем приближенно в пределе  $R \rightarrow \infty$

$$\begin{aligned} u(r, t) &\approx \varphi(t) \exp(-r^2/r_0^2) + 1, \\ n(r, t)N(r, t) &\approx [\psi(t) - n_0] \exp(-r^2/r_0^2) + n_0. \end{aligned} \quad (9)$$

Подставляя (9) в (7), умножая оба уравнения на  $\exp(-r^2/r_0^2)$  и интегрируя их почленно по  $rdr/R^2$ , получим систему двух обыкновенных дифференциальных уравнений для функций  $\varphi(t)$  и  $\psi(t)$  [4].

С формальной точки зрения данный метод есть разновидность метода Бубнова—Галеркина, в которой в качестве базиса разложения используются собственные функции соответствующей линейной краевой задачи. В разложении (9) мы ограничились низшим порядком, при котором в (9) содержатся две «галеркинские координаты» — функции  $\varphi(t)$  и  $\psi(t)$ .

Анализ получающейся таким образом системы двух обыкновенных дифференциальных уравнений выполняется известными методами теории нелинейных колебаний. В частности, оказалось, что необходимым условием существования колебаний является  $u_0 > 4$ . Колебания могут возникать как в мягком (т. е. при потере устойчивости стационарным решением), так и в жестком (т. е. под действием возмущений конечной величины) режимах. Для периода колебаний вблизи порога их рождения имеет место оценка  $\tau \sim 5r_0^2/\nu$ , что отвечает характерным временам диффузии на расстояниях порядка радиуса пучка.

Приведем теперь результаты численного анализа краевой задачи (7). Расчеты выполнялись по неявной консервативной разностной схеме, в которой шаги по времени ( $\Delta t = 0,05$ ) и по пространству ( $\Delta r = 0,025$ ) выбирались на основе тестовых расчетов. Зафиксируем параметры задачи:

$$\nu = 3, \quad u_0 = 7, \quad \alpha = 2, \quad n_0 = 0,1, \quad R = 5 \quad (10)$$

и будем менять параметры источника  $r_0$  и  $P = (\beta_0 e^{-u_0}) r_0^2 I_0$ .

Расчеты показали, что плоскость параметров  $\{P, r_0\}$  делится на несколько областей, отличающихся динамикой процесса (рис. 3). Вне области, ограниченной замкнутой кривой, из любых начальных условий происходит выход на стационарное решение, параметры которого зависят от значений  $P$  и  $r_0$ . Для параметров же  $(P, r_0)$  внутри этой области в системе возникают устойчивые колебательные режимы изменения переменных  $u$  и  $n$ . В свою очередь, область существования колебаний делится на три характерные подобласти (I, II, III).

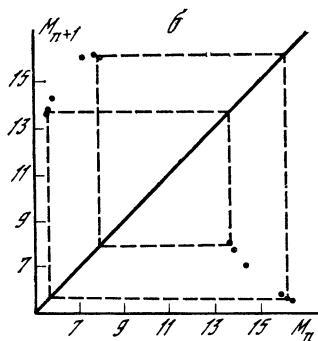
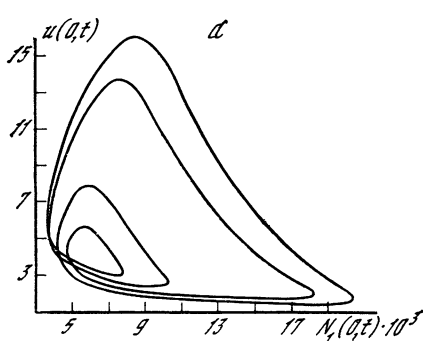
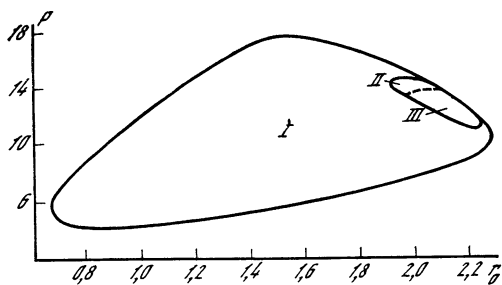
Выберем фазовую плоскость  $\{u(0, t), N_1(0, t)\}$ , где введено обозначение

$$N_1(r, t) = n(r, t)N(r, t).$$

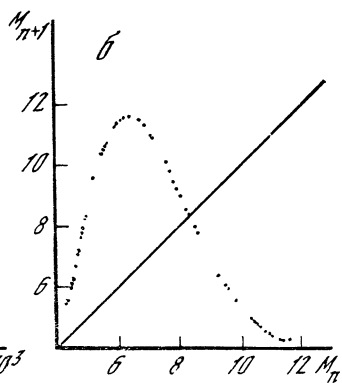
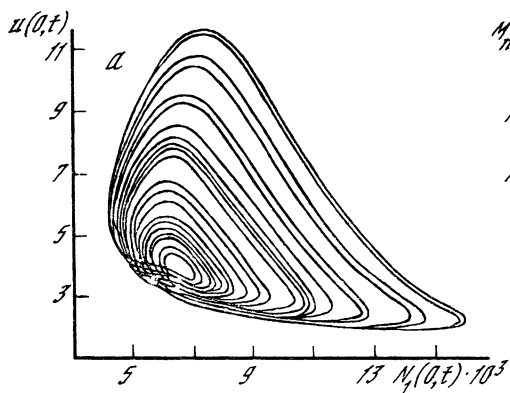
На этой плоскости колебаниям из подобласти I (см. рис. 3) отвечает простая замкнутая кривая без самопересечений — предельный цикл  $S^1$ . Здесь и далее используется обозначение предельных циклов типа  $S^k$ , если на одном периоде колебаний динамическая переменная имеет точно  $k$  максимумов, а на фазовой плоскости такому циклу отвечает замкнутая кривая с  $k$  витками (при  $k > 1$  — с самопересечениями).

При переходе в подобласть II (см. рис. 3) происходит усложнение колебаний. Основной сценарий усложнения — это последовательность бифуркаций удвоения периода:  $S^1 \rightarrow S^2 \rightarrow S^4 \rightarrow S^8 \rightarrow \dots$ . Для примера на рис. 4, а приведен фазовый портрет пре-

**Рис. 3.** Разбиение плоскости параметров  $\{P, r_0\}$  на области, отличающиеся динамикой процесса (качественными особенностями решений краевой задачи (7))



**Рис. 4.** Фазовый портрет (а) и отображение Лоренца (б) для решений краевой задачи (7) в случае, когда происходит выход на предельный цикл  $S^1$  ( $P=15, r_0=1,935$ )



**Рис. 5.** Фазовый портрет (а) и отображение Лоренца (б) для сложного колебательного процесса ( $P=14,8, r_0=2,01$ )



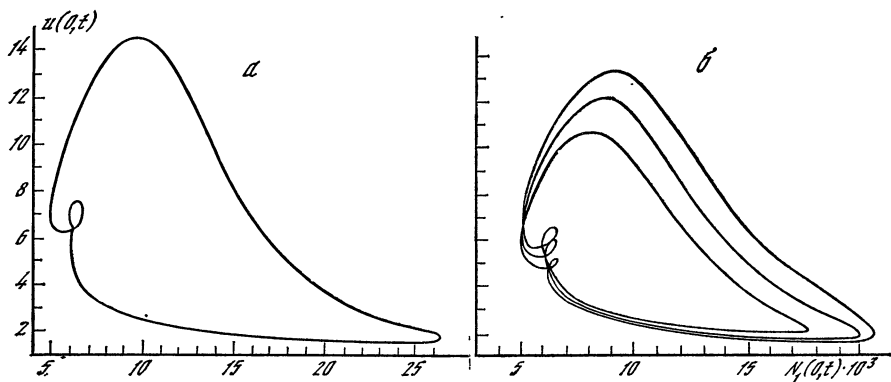


Рис. 6. Фазовые портреты усложненного предельного цикла  $S^1$  (а) и колебательного процесса, содержащего малые витки (б)

а —  $P = 13,9$ ,  $r_0 = 2,12$ ; б —  $P = 14,0$ ,  $r_0 = 2,11$

дельного цикла  $S^4$ , возникшего после первых двух бифуркаций удвоения.

Одним из методов изучения сложных динамических режимов является прием, предложенный Э. Лоренцем. Выберем одну из фазовых переменных, например  $u(0, t)$ , и будем следить за ее последовательными максимумами  $M_n$ ,  $n = 1, 2, \dots$ . Для последовательности  $\{M_n\}$  построим отображение

$$M_{n+1} = f(M_n), \quad n = 1, 2, 3, \dots, \quad (11)$$

т. е. построим график зависимости  $(n + 1)$ -го максимума от  $n$ -го.

На рис. 4, б для режима, выходящего на цикл  $S^4$ , построено отображение (11). Для изучаемой краевой задачи оказалось, что в пределах точности расчета точки отображения ложатся на непрерывную однозначную кривую. Это дает основание предположить, что усложнение динамики описывается теорией М. Фейгенбаума. Согласно этой теории, при изменении некоторого параметра после каскада удвоений периода возникают наряду с иными предельными циклами более сложные, стохастические режимы. Такие режимы были найдены и в задаче (7). На рис. 5 приведены фазовый портрет и отображение Лоренца для сложного режима, содержащего по крайней мере более 300 различных максимумов.

Наконец, в подобласти III (см. рис. 3) наблюдается еще один механизм усложнения динамики, не связанный с бифуркациями циклов. Дело в том, что при близких к области существования колебаний (но вне ее) значениях параметров  $P$ ,  $r_0$  в фазовом пространстве изучаемой системы рождается устойчивая особая точка и фазовые траектории системы начинают отклоняться в сторону зародыша. На рис. 6, а показан фазовый портрет цикла  $S^1$ , усложненного такой деформацией траектории. Как видно из графика, усложнение состоит в появлении дополнительной небольшой петельки, отвечающей маломасштабному биению на фоне регу-

лярного колебательного процесса. На рис. 6, б показан пример более сложного режима, содержащего малые осцилляции.

Таким образом, «включение» пространственной неоднородности привело к качественному усложнению поведения системы: появились режимы, отсутствующие в однородном случае. Заметим, что автоколебания возникают в упрощенных системах, если в (7) либо а) положить  $\alpha = 0$ , т. е. «отключить» термодиффузию, либо б) положить  $N \equiv \text{const} = 1$  [4], но при  $\alpha > 0$ . Следовательно, система (7) описывает взаимодействие двух простых автоколебательных подсистем, возбуждающихся при наличии неоднородного источника. Такое взаимодействие привело к возникновению сложных динамических режимов в комбинированной системе.

### 3. ПРОСТРАНСТВЕННОЕ УДВОЕНИЕ ПЕРИОДА

Наряду с локальным усложнением динамики пространственная неоднородность может вести и к более сложным типам пространственно-временного упорядочения, когда единая когерентная структура по-разному проявляется в различных частях системы.

Рассмотрим следующую краевую задачу [5, 6]:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{c}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + I(r) + \frac{\alpha}{X} \exp\left(-\frac{1}{u}\right) - \\ &- \beta \exp\left(-\frac{\kappa}{u}\right) - \delta u, \\ \frac{\partial X}{\partial t} &= \frac{d}{X} \exp\left(-\frac{1}{u}\right) - \gamma \exp\left(-\frac{\kappa}{u}\right), \quad 0 < r < R, \end{aligned} \quad (12)$$

$$\left. \frac{\partial u}{\partial r} \right|_{r=0, R} = 0, \quad u|_{t=0} = u_0(r), \quad X|_{t=0} = X_0(r).$$

Эта задача была исследована в связи с проблемой нагрева излучением в воздухе металлов типа вольфрама, у которых окисел, образующийся на поверхности, может достаточно эффективно испаряться (сублимировать).

Задавая источник  $I(r)$ , как и раньше, в гауссовском виде (2), будем изучать свойства решений в зависимости от параметров неоднородности ( $I_0, r_0$ ). Прежде всего отметим, что в предельном случае, когда теплопроводностью можно пренебречь, получаем динамическую систему

$$\begin{aligned} \frac{du}{dt} &= I + \frac{\alpha}{X} \exp\left(-\frac{1}{u}\right) - \beta \exp\left(-\frac{\kappa}{u}\right) - \delta u, \\ \frac{dX}{dt} &= \frac{d}{X} \exp\left(-\frac{1}{u}\right) - \gamma \exp\left(-\frac{\kappa}{u}\right). \end{aligned} \quad (13)$$

Исследование этой системы выполняется стандартными методами теории нелинейных колебаний и показывает, что здесь может происходить выход либо на стационарные состояния, либо на простые автоколебания (типа цикла  $S^1$ ) [7].

По аналогии с результатами предыдущего раздела можно ожидать, что при «включении» неоднородности в задаче (12) динамика

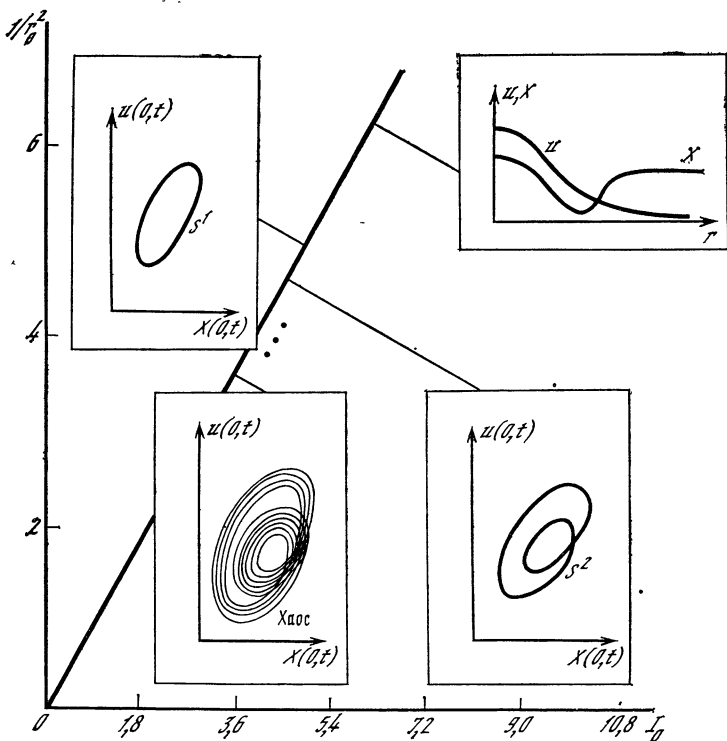


Рис. 7. Зависимость типа режима от параметров источника  $I_0$ ,  $r_0$  при фиксированном значении  $I_0 r_0^2 = 0,9$

Значения констант в краевой задаче (12):  $c = 0,1$ ,  $\alpha = 9,0$ ,  $\beta = 0$ ,  $\kappa = 0,6$ ,  $d = 0,894$ ,  $\gamma = 1,0$ ,  $R = 2,0$ ,  $\delta = 9,0$

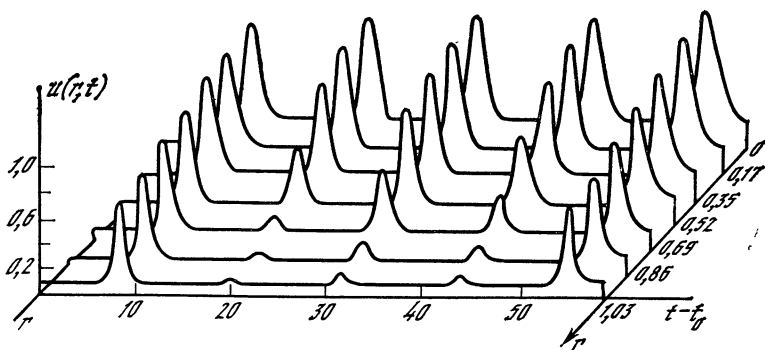


Рис. 8. Зависимость величины  $u$  от времени для разных значений координаты

$I_0 = 1,8$ ,  $r_0 = 1,0$ ,  $R = 1,0$ . Значения остальных параметров те же, что на рис. 7

системы будет усложняться. Численное исследование позволило установить некоторые последовательности бифуркаций при изменении параметров источника. Для примера на рис. 7 показана зависимость типа режима от радиуса локализации источника  $r_0$  при фиксированной величине  $I_0 r_0^2 = 0,9$ . Из диаграммы видно, что при больших значениях  $I_0$  (и соответственно малых  $r_0$ ) система выходит на стационарное состояние с неоднородным распределением величин  $u$  и  $X$  по радиусу. При уменьшении  $I_0$  стационарное состояние теряет устойчивость и рождается предельный цикл  $S^1$ . После каскада бифуркаций удвоения периода возникают сложные нестационарные режимы (в том числе стохастические). При дальнейшем же уменьшении  $I_0$  после некоторой серии бифуркаций остаются возможными только стационарные решения и автоколебания, описываемые моделью (13).

Качественно новый эффект обнаруживается, если проследить за временным поведением переменных в различных точках системы. На рис. 8 показан пример зависимости величины  $u$  от времени для разных значений координаты  $r$ . Видно, что при выбранных значениях параметров в точке  $r = r_1 = 0$  функция  $u(r_1, t)$  совершает регулярные колебания (цикл  $S^1$ ) с периодом  $\tau_1 \approx 12,5$ . Однако при  $r = r_2 = 0,35$  колебания имеют период  $\tau_2 = 2\tau_1$  и на соответствующей фазовой плоскости (например,  $\{u(r_2, t), X(r_2, t)\}$ ) представляются циклом  $S^2$ . В точке же  $r_3 = 0,86$  наблюдаются колебания с периодом  $\tau_3 = 4\tau_1$  (цикл  $S^4$ ). Таким образом, по мере удаления от центра ( $r = r_1$ ) происходят последовательные бифуркации удвоения периода, т. е. имеет место усложнение динамики. Следовательно, в пространственно-неоднородной системе, несмотря на образование некоторой единой когерентной структуры, изучение динамики процесса в малой окрестности какой-либо одной выбранной точки не дает возможности определить динамику в других точках.

#### 4. МЕТАСТАБИЛЬНАЯ САМООРГАНИЗАЦИЯ

Как уже отмечалось, для нелинейных диссипативных систем характерны явления самоорганизации, с математической точки зрения определяемые аттракторами в фазовом пространстве. Однако существует иной тип самоорганизации, возникающий на промежуточной стадии эволюции системы. Оказывается, для многих систем характерно наличие неустойчивых структур, которые тем не менее обладают достаточно широкой областью притяжения (по начальным условиям) и большим временем жизни.

Это явление проиллюстрируем на примере следующей краевой задачи [8, 9]:

$$\frac{\partial u}{\partial t} = -\frac{\partial^2}{\partial x^2} \left[ \frac{\partial^2 u}{\partial x^2} + f(u) \right], \quad 0 < x < R, \quad (14)$$

$$\frac{\partial u}{\partial x} \Big|_{x=0, R} = 0, \quad \frac{\partial^3 u}{\partial x^3} \Big|_{x=0, R} = 0, \quad u(x, 0) = u_0(x),$$

$$f(u) = \beta u + \epsilon u^2 - u^3.$$

Такого типа краевые задачи встречаются при описании расслоения сплавов, эмульсий [10, 11], при исследовании нагрева газовых смесей излучением и т. д.

Задача (14) имеет очевидный закон сохранения

$$\int_0^R u(x, t) dx = \int_0^R u_0(x) dx \equiv \bar{u}_0 R. \quad (15)$$

В дальнейшем примем  $\bar{u}_0 \equiv 0$ . Прежде всего легко видеть, что при этом условии имеется единственное однородное решение:  $u = 0$ . При  $\beta > 0$  это решение оказывается неустойчивым по отношению к малым возмущениям. Действительно, линеаризуя (14) в окрестности  $u = 0$  и полагая  $u - 0 = u_1 \exp(pt + iqx)$ , получим закон дисперсии

$$p = q^2(\beta - q^2), \quad q = n\pi/R, \quad n = \pm 1, \pm 2, \dots \quad (16)$$

При  $R > \pi/\sqrt{\beta}$  в спектре  $p(q)$  присутствуют положительные значения  $p$  и соответствующие возмущения нарастают. В результате в системе формируется некоторая пространственно-неоднородная структура.

Краевая задача (14) имеет функцию Ляпунова

$$V = \int_0^R \left[ \frac{1}{2} \left( \frac{\partial u}{\partial x} \right)^2 - \Phi(u) \right] dx, \quad \Phi(u) = \int_0^u f(u_1) du_1. \quad (17)$$

Действительно, дифференцируя (17) по времени, с учетом (14) находим:

$$\frac{dV}{dt} = - \int_0^R \left\{ \frac{\partial}{\partial x} \left[ \frac{\partial^2 u}{\partial x^2} + f(u) \right] \right\}^2 dx \leq 0, \quad (18)$$

откуда видно, что функция  $V$  со временем не возрастает и ее производная  $dV/dt$  обращается в нуль только на стационарных решениях краевой задачи (14). Из (17) можно показать, что  $V(u)$  ограничена снизу. На основании этих утверждений делается вывод, что из любых начальных условий происходит выход на некоторое стационарное решение краевой задачи (14).

При  $\beta > 0$  все стационарные решения являются пространственно-периодическими с периодами

$$\rho = 2R/m, \quad m = 1, 2, \dots, m_{\max}; \quad \rho_{\min} \leq \rho \leq 2R, \quad (19)$$

т. е. содержат целое число полуволн. Устойчивость каждого из этих решений можно исследовать на основании соответствующей линеаризованной краевой задачи. Этот анализ показывает, что асимптотически устойчивыми являются только те решения, которые имеют максимально возможный период, т. е. решения типа полуволны на всей длине  $R$  [8].

Имея это в виду, рассмотрим теперь динамику системы. Численное интегрирование краевой задачи (14) для достаточно боль-

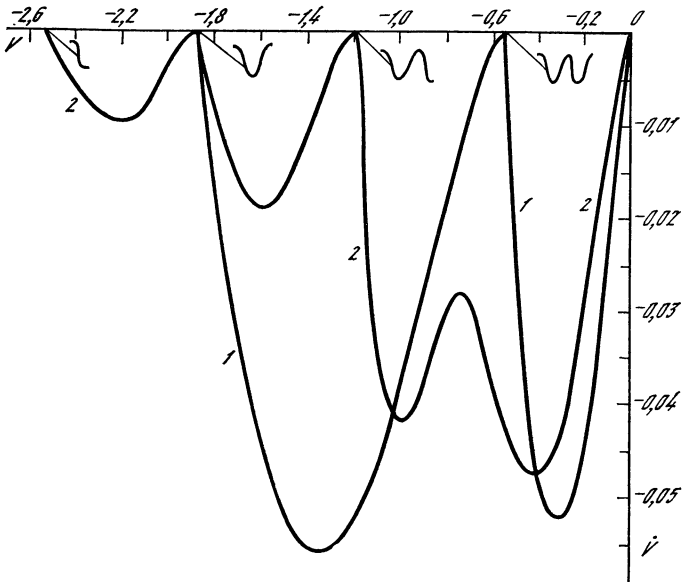


Рис. 9. Фазовый портрет краевой задачи (14) в координатах  $(V, \dot{V})$  для двух различных начальных условий

1 —  $u_0(x) = 0,5 \cos(8\pi x/R)$ , 2 —  $u_0(x) = 0,05 \cos(\pi x/R)$ . Вдоль оси  $V$  приведен качественный вид стационарных решений краевой задачи ( $R = 20$ ,  $\beta = 0,8$ ,  $\varepsilon = 0,2$ )

ших размеров области  $R$  показывает, что развитие возмущений приводит к установлению структуры, близкой к стационарному решению с периодом  $\rho_1 = 2\pi/q_1$ , где  $q_1 = \sqrt{\beta/2}$  — волновой вектор, отвечающий максимуму функции  $p(q)$  в (16). Такая структура ( $u_1 = u(x; \rho_1)$ ) формируется из широкого набора начальных условий и практически не меняется в течение довольно большого промежутка времени. Однако поскольку она неустойчива (при достаточно больших  $R$ ), то малые возмущения разрушают ее и в системе формируется структура  $u_2 = u(x; \rho_2)$ , где  $\rho_2 = 2\pi/q_2$ , а  $q_2$  отвечает максимуму инкремента  $p = p(q; \rho_1)$  неустойчивости решения  $u(x; \rho_1)$ .

Наглядно процесс эволюции можно представить на фазовой плоскости  $\{V, \dot{V}\}$ , имея в виду, что значениям  $\dot{V} = 0$  отвечают стационарные состояния изучаемой системы.

На рис. 9 приведены два примера фазовых траекторий, стартовых из качественно различных начальных условий. Графики показывают, что наиболее сильными притягивающими свойствами обладают стационарные решения, содержащие целое число волн, и особенно первая с периодом  $\rho_1$ . Характерные времена жизни неустойчивых стационарных решений определяются максимумом их инкрементов неустойчивости и могут быть оценены по формуле  $\tau_n \sim 1/p_{n,\max} \sim \rho_n^4$ . Поскольку в процессе эволюции система проходит последовательно в окрестностях стационарных реше-

ний, периоды которых растут с ростом их номера, то при достаточно больших размерах  $R$  благодаря захвату промежуточными (метастабильными) структурами с увеличивающимся временем жизни характерное время выхода на устойчивое состояние может оказаться весьма большим.

Явление метастабильного упорядочения характерно для нелинейных систем. Например, для задачи, рассмотренной в разд. 3, это явление проявляется в том, что устойчивые сложные колебательные структуры формируются в конце некоторой последовательности неустойчивых промежуточных колебательных состояний. Так, если устойчивым является цикл  $S^8$ , то эволюция протекает по цепочке  $S^1 \rightarrow S^2 \rightarrow S^4 \rightarrow S^8$ , причем каждое из промежуточных состояний существует в течение нескольких своих полных периодов [5, 6].

Заметим, наконец, что явление захвата на метастабильное состояние может весьма затруднить численное исследование эволюционных свойств системы.

В заключение следует сказать, что рассмотренные в статье математические модели и их модификации позволяют описать и некоторые иные типы структур: волны переключения, нетривиальные стационарные диссипативные структуры, эффекты обострения и локализации и т. д.

#### ЛИТЕРАТУРА

1. Бункин Ф. В., Галактионов В. А., Кириченко Н. А. и др. Качественные методы в физике лазерного нагрева металлов в воздухе // Изв. АН СССР. Сер. физ. 1985. Т. 49, № 6. С. 1046—1053.
2. Бункин Ф. В., Кириченко Н. А., Лукьянчук Б. С. Неустойчивость поверхностного горения под действием лазерного излучения // Квантовая электрон. 1982. Т. 9, № 10. С. 1959—1967.
3. Бункин Ф. В., Кириченко Н. А., Морозов Ю. Ю. Колебательные процессы при нагреве газовых смесей пучком лазерного излучения // Изв. АН СССР. Сер. физ. 1987. Т. 51, № 6. С. 1162—1169.
4. Бункин Ф. В., Кириченко Н. А., Курдюмов С. П. и др. Автоволновые процессы в газовых средах с термодиффузией, инициируемые лазерным излучением: Препр. ИПМ им. М. В. Келдыша АН СССР № 40. М., 1986. 22 с.
5. Ахромеева Т. С., Бункин Ф. В., Кириченко Н. А. и др. Периодические колебания и диффузионный хаос при нагреве металлов излучением // Изв. АН СССР. Сер. физ. 1987. Т. 51, № 6. С. 1154—1161.
6. Ахромеева Т. С., Бункин Ф. В., Кириченко Н. А. и др. Периодические автоволновые процессы и диффузионный хаос в одной задаче лазерной термохимии: Препр. ИПМ им. М. В. Келдыша АН СССР № 41. М., 1986. 31 с.
7. Бобырев В. А., Бункин Ф. В., Кириченко Н. А. и др. Автоколебательные режимы лазерного горения металлов // Квантовая электрон. 1983. Т. 10, № 4. С. 793—797.
8. Кириченко Н. А. Об асимптотически устойчивых решениях некоторых нелинейных уравнений // ДАН СССР. 1985. Т. 280, № 3. С. 579—583.
9. Кириченко Н. А. Об устойчивости стационарных решений некоторых нелинейных параболических уравнений: Препр. ИОФАН № 109. М., 1986. 43 с.
10. Sahn J. W. The later stages of spinodal decomposition and the beginnings of particle coarsening // Acta met. 1966. Vol. 14, N 12. P. 1685—1692.
11. Hagan P. S., Cohen D. S. Stratified layer formation in particle suspensions // Physica D. 1985. Vol. 17, N 1. P. 54—62.

## ТЕРМОДИФФУЗИОННЫЕ АВТОСОЛИТОНЫ И СТРАТЫ В ПОЛУПРОВОДНИКОВОЙ И ГАЗОВОЙ ПЛАЗМЕ

Б. С. КЕРНЕР, В. В. ОСИПОВ

В целом ряде неравновесных физических, химических и биологических систем могут спонтанно или вынужденно образовываться различного рода диссипативные структуры — автоструктуры [1—3]. Спонтанно автоструктуры возникают при изменении уровня неравновесности системы в результате неустойчивости ее однородного состояния. Вынужденно автоструктуры можно возбудить, воздействуя на систему кратковременным возмущением конечной амплитуды. В устойчивой однородной системе после воздействия на нее локализованного возмущения могут образовываться различного рода локализованные автоструктуры, т. е. уединенные устойчивые собственные состояния — автосолитоны (АС). Нелинейная теория статических и пульсирующих автоструктур большой амплитуды, в том числе АС и одномерных автоструктур — страт, в активных системах с диффузией развита в [4—7]. Ее основные результаты изложены в [8—10].

К системам, близким к активным системам с диффузией [1—3, 8—13], относятся газовая и полупроводниковая плазма, в которой могут образовываться термодиффузионные АС и страты [4, 5, 14, 15]. Существование таких автоструктур связано с сильным влиянием термодиффузии горячих невырожденных носителей заряда на пространственное распределение их концентрации. Термодиффузионные автоструктуры образуются в невырожденной плазме при весьма различных ее параметрах и условиях разогрева [4, 5, 14—19]. Общим необходимым условием существования термодиффузионных автоструктур является  $L \gg l$ , где  $L$  и  $l$  — длины биполярной диффузии и остывания носителей<sup>1</sup>. В данной работе рассматривается физика существования и обсуждаются общие свойства термодиффузионных АС и страт, а также анализируются наиболее характерные условия, при которых они образуются в полупроводниковой и газовой плазме.

### 1. ОСНОВНЫЕ РЕЗУЛЬТАТЫ ТЕОРИИ И ФИЗИКА СУЩЕСТВОВАНИЯ ТЕРМОДИФФУЗИОННЫХ АВТОСОЛИТОНОВ И СТРАТ

В данном разделе для простоты изложения рассматривается «симметричная» электронно-дырочная плазма (ЭДП), для которой параметры электронов и дырок в точности совпадают и их концент-

<sup>1</sup> В диффузионном режиме газового разряда [20—22] при невысокой степени ионизации газа, как правило, выполняется обратное условие  $L \ll l$  и образуются не термодиффузионные, а ионизационные страты. Существование таких страт связано с зависимостью скорости ионизации от концентрации и температуры электронов [20—22]. Нелинейная теория страт и АС большой амплитуды в газовом разряде развита в работах [23—25].



рации ( $n = p$ ) таковы, что выполнены условия

$$\tilde{\tau}_e, \tilde{\tau}_h \ll \tilde{\tau}_{ee} \ll \tau_e, \tau_h. \quad (1)$$

Здесь  $\tilde{\tau}_e$  ( $\tilde{\tau}_h$ ) и  $\tau_e$  ( $\tau_h$ ) — характерные времена релаксации импульса и энергии электронов (дырок) при их столкновении с решеткой;  $\tilde{\tau}_{ee}$  — характерное время межэлектронных соударений. В рассматриваемом случае  $\tilde{\tau}_e = \tilde{\tau}_h = \tilde{\tau}$  и  $\tau_e = \tau_h = \tau$ . При выполнении (1) справедливо «квазигидродинамическое» описание плазмы [22], при котором пространственные распределения концентрации  $n$  и эффективной температуры  $T$  носителей определяются уравнениями

$$\partial n / \partial t = - \operatorname{div} \mathbf{j}_e + G - n / \tau_r, \quad (2)$$

$$3 \partial (nT) / \partial t = - \operatorname{div} \mathbf{j}_e + W - 2n (T - T_l) / \tau, \quad (3)$$

где плотность потока электронов равна

$$\mathbf{j}_e = - (1 + \alpha) D n T^{-1} \nabla T - D \nabla n - \mathbf{j} / 2e = - \nabla (nD) - \mathbf{j} / 2e; \quad (4)$$

$\alpha = \partial \ln \tilde{\tau}_{e,h} / \partial \ln T$ ;  $D$  — коэффициент биполярной диффузии;  $T_l$  — температура решетки полупроводника;  $\tau_r$  и  $G$  — время жизни и скорость генерации носителей;  $\mathbf{j}_e$  — плотность потока энергии носителей;  $W = \mathbf{j} \mathbf{E}$ ;  $\mathbf{j} = \sigma \mathbf{E}$  — плотность тока;  $\mathbf{E}$  — напряженность электрического поля.

Согласно (2), (3), при  $\tau_r = \text{const}$  вольт-амперная характеристика (ВАХ) однородной ЭДП описывается выражениями

$$j^2 = 2n\sigma (T - T_l) \tau^{-1}, \quad E^2 = (T - T_l) [e \mu \tau]^{-1}, \quad (5)$$

где  $\sigma = 2e\mu n$  и  $\mu = e\tilde{\tau}/m$  — проводимость и подвижность носителей. В полупроводниках, как правило, выполняются условия  $s \pm \alpha < 1$  [26, 27], где  $s = \partial \ln \tau / \partial \ln T$ . Поэтому из (5) вытекает, что зависимость температуры однородной ЭДП ( $T = T_h$ ) от величины  $j$  (или  $E$ ) является монотонно возрастающей функцией (рис. 1). Иными словами, перегревная неустойчивость [27] отсутствует, т. е. в рассматриваемой ЭДП при любом  $j$  (или  $E$ ) реализуется единственное однородное состояние  $T = T_h$ ,  $n = n_h$ .

Несмотря на это, однородное состояние ЭДП может стать неустойчивым [15, 28, 29], и в ней могут образовываться поперечные термодиффузионные АС и страты [15], т. е. распределения  $T(x)$ ,  $n(x)$ , неоднородные лишь в направлении линий тока (оси  $x$ ). Такие распределения, согласно (2)–(4), описываются уравнениями

$$\tau_r^0 \partial (\eta \beta) / \partial t = L^2 \partial^2 \eta / \partial x^2 - Q(\eta, \theta), \quad Q = \eta \beta (\tau_r^0 / \tau_r) - 1, \quad (6)$$

$$\frac{3}{2} \tau^0 \eta^{-1} \frac{\partial [\eta \beta T(\theta)]}{\partial t} = l^2 \eta^{-1} \frac{\partial}{\partial x} \left[ \eta \frac{\partial \theta}{\partial x} + \frac{(\beta/2 + \alpha) T}{\beta/2 + \alpha} \frac{\partial \eta}{\partial x} \right] - q(\eta, \theta, J), \quad q = -JT \eta^{-2} + \beta (T - 1) \tau^0 / \tau, \quad (7)$$

где  $\eta = nD(T)$ ;  $\beta = D^0/D$ ;  $J = [\tau^0/4e\mu^0 (n_h^0)^2 T_l] j^2$ ;  $\theta(T) = \int_0^T \frac{(\beta/2 + \alpha) dT}{\beta/2 + \alpha}$  — однозначная функция  $T$ ;  $L = (D^0 \tau_r^0)^{1/2}$ ;  $l = [(5/2 + \alpha) \tau^0 D^0]^{1/2}$ ;  $\eta$  и  $T$  измеряются соответственно в едини-

Рис. 1. Зависимость температуры  $T_h$  однородной ЭДП от уровня ее возбуждения  $J \propto j^2$

Пунктирный участок отвечает области значений  $J_0 < J < J'_0$ , где однородное состояние ЭДП неустойчиво

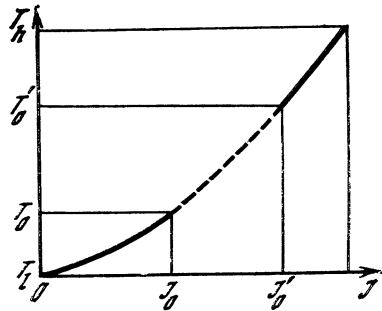
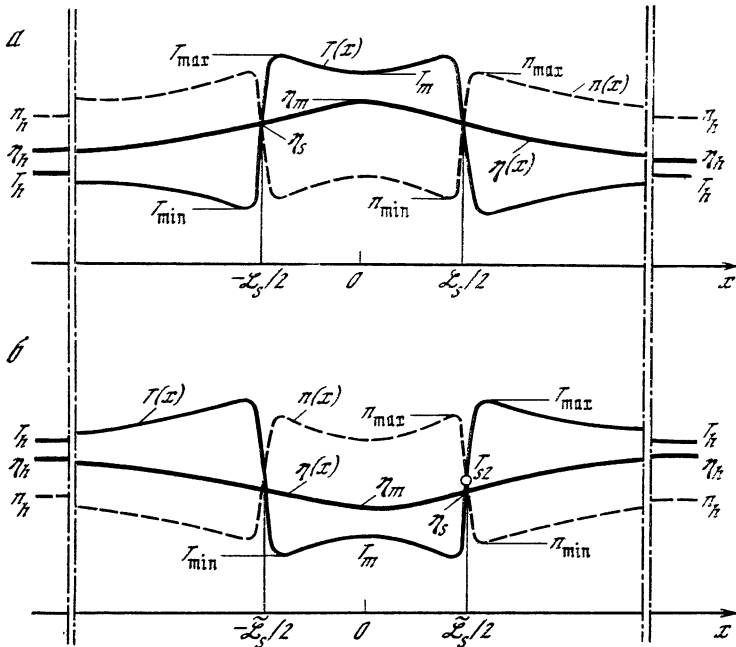


Рис. 2. Распределения концентрации носителей  $n$ , их температуры  $T$  и параметра  $\eta$  в широких горячем (а) и холодном (б) термодиффузионных АС



цах  $n_h^0 D^0$  и  $T_l$ . Верхний индекс ноль у величины означает, что она взята при  $T = T_l$ .

Из уравнения (6) следует, что в горячей ЭДП длина биполярной диффузии  $L$  характеризует масштаб изменения не концентрации носителей  $n$ , а величины  $\eta = nD(T)$ . Характерной же длиной изменения температуры  $T$  является длина остывания горячих носителей  $l \ll L$ . С разным масштабом изменения  $T(x)$  и  $\eta(x)$  в ЭДП и связана физика существования АС. Последний представляет собой область высокой температуры и низкой концентрации носителей (рис. 2, а) [4, 30]. Существование такого горячего АС определяется тем, что благодаря термодиффузии происходит интенсивный выброс горячих носителей из области высокой температуры. В результате концентрация носителей, а следовательно, и проводимость  $\sigma$  в центральной области АС уменьшаются (см.

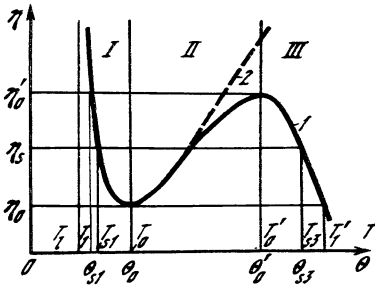


Рис. 3. Кривые И-образной (1) и V-образной (2) локальных связей

эти кривые практически уравниваются термодиффузионным. Последнее, т. е. близость к нулю потока

$$j_D \equiv -\partial\eta/\partial x = -(1 + \alpha)DnT^{-1}\partial T/\partial x - D\partial n/\partial x, \quad (8)$$

следует из уравнения (6), согласно которому  $j_D(x)$  изменяется на длине порядка  $L \gg l$ . Таким образом, в области стенки АС (размера порядка  $l$ ), где  $T(x)$  и  $n(x)$  меняются резко (см. рис. 2, а), вследствие того что  $L \gg l$ , можно считать величину  $\eta(x) \approx \text{const}$ . Резкое изменение  $T(x)$ , описывающее стенку АС, как следует из анализа [4, 5, 30] уравнения (7) при  $\eta = \text{const}$ , реализуется, когда зависимость  $\eta(\theta)$ , вытекающая из уравнения локального баланса энергии носителей

$$q(\eta, \theta, J) = \beta(T(\theta) - 1)\tau^0/\tau - J\eta^{-2}T(\theta) = 0, \quad (9)$$

имеет немонотонный характер (рис. 3) [4, 5, 8–10]. Эта зависимость называется локальной связью (ЛС).

Из уравнений (7) и (9) следует, что немонотонная зависимость  $\eta(\theta)$  в рассматриваемом случае возникает, когда

$$\alpha + s > -1. \quad (10)$$

Это условие является значительно более мягким, чем условие перегретой неустойчивости:  $\alpha + s > 1$  [27]. Оно выполняется при низких температурах, когда рассеяние импульса и энергии носителей происходит соответственно на заряженных центрах и акустических фононах ( $\alpha \approx 3/2$ ,  $s = -1/2$  [26]). В этом случае кривая ЛС имеет И-образный вид (см. рис. 3, кривая 1). Наличие второго падающего участка на ЛС при  $T > T_0'$  обусловлено тем [5], что с ростом  $T$  все большую роль начинает играть рассеяние на оптических фононах.

Условие (10) может выполняться и при высоких температурах  $T \gtrsim \theta_D$  ( $\theta_D$  — температура Дебая), когда рассеяние импульса и энергии носителей определяется их взаимодействием с оптическими фононами или переходами между эквивалентными долинами. В этом случае ЛС может иметь как И-, так и V-образный вид (см. рис. 3). И-образный вид кривой ЛС при  $T \gtrsim \theta_D$  реализуется

рис. 2, а). Это поддерживает высокое значение  $T$  в центре АС, так как приводит к увеличению поступающей к носителям от электрического поля мощности  $W = j^2/\sigma$  и уменьшению мощности  $P = 2n(T - T_1)\tau^{-1}$ , отводимой от носителей к решетке.

Область низкой концентрации в АС не расплывается вследствие того, что диффузионный поток носителей в областях размера порядка  $l$  резко изменяется  $T(x)$  и  $n(x)$  (в стенках АС) практически

за счет возрастания эффективной массы носителей с ростом их энергии или же переходов носителей в более высоко расположенные долины. В частности, условие (10) выполняется при комнатных температурах как для полярных (InSb, PbTe, GaAs), так и для неполярных (Ge, Si) полупроводников (см. также п. 3.1).

### 1.1. Широкие автосолитоны

В устойчивой слабопрогретой ЭДП, для которой кривая ЛС имеет И-образный вид, можно возбудить горячий АС (см. рис. 2, а), размер которого  $\mathcal{L}_s \gg l$ , но  $\mathcal{L}_s \ll L$  [5, 7, 8, 10]. С точностью до величины  $\varepsilon = l/L \ll 1$  в широком АС максимальная и минимальная температура  $T_{\max} = T_{s3}$  и  $T_{\min} = T_{s1}$ , значение  $\eta = \eta_s$  в стенке широкого АС, его ширина —  $\mathcal{L}_s$ , а также значения  $\eta = \eta_m$  и  $T = T_m$  в центре АС определяются из простых, как правило алгебраических, уравнений [5, 8—10, 31]:

$$\int_{\theta_{s1}}^{\theta_{s3}} q(\theta, \eta_s, J) d\theta = 0, \quad q(\theta_{si}, \eta_s, J) = 0, \quad i = 1, 2, 3, \quad (11)$$

$$\mathcal{L}_s = \sqrt{2}L \int_{\eta_m}^{\eta_s} \left[ \int_{\eta_m}^{\eta} Q(\theta_{III}(\eta), \eta) d\eta \right]^{-1/2} d\eta, \quad (12)$$

$$\int_{\eta_s}^{\eta_h} Q(\theta_I(\eta), \eta) d\eta = - \int_{\eta_m}^{\eta_s} Q(\theta_{III}(\eta), \eta) d\eta, \quad q(\eta_m, T_m, J) = 0, \quad (13)$$

где функции  $\theta_{I, III}(\eta)$  отвечают участкам I ( $\theta \leq \theta_0$ , рис. 3) и III ( $\theta \geq \theta'_0$ ) однозначной зависимости  $\theta(\eta)$  на кривой ЛС;  $\eta_h$  — значение  $\eta$  для однородной ЭДП. Величины  $n_{\max}$  и  $n_{\min}$  для АС (см. рис. 3) определяются из соотношения  $D(T)n = \eta_s D^0 n_h^0$ , в которое нужно подставить значения  $\eta_s$  и соответственно  $T_{\min} = T_{s1}$  или  $T_{\max} = T_{s3}$ , найденные из уравнений (11).

Из анализа (11)—(13) следует [5], что ширина горячего АС уменьшается с уменьшением уровня разогрева ЭДП  $J$ , который в данном случае пропорционален квадрату плотности тока  $j^2$ . Из результатов работ [5—8] вытекает, что, когда размер АС достигает величины  $\mathcal{L}_s \sim l \ln(L/l)$ , АС исчезает в точке  $J = J_b$ , где  $d\eta_s/dJ = \infty$  (скачок  $1 \rightarrow 2$  на рис. 4, а). Таким образом, горячий АС (см. рис. 2, а) может существовать при  $J_b < J < J_0$  (кривая АС на рис. 4, а), где  $J_0$  — уровень разогрева ЭДП, при котором ее однородное состояние становится неустойчивым<sup>2</sup>.

<sup>2</sup> В области значений  $J_0 < J < J'_0$ , т. е.  $j_0 < j < j'_0$ , однородное состояние рассматриваемой ЭДП оказывается неустойчивым относительно неоднородных флуктуаций температуры и концентрации носителей с волновым вектором  $k \parallel E$ , причем волновое число критических флуктуаций  $k_0 \approx (lL)^{-1/2}$  [15, 28, 29].

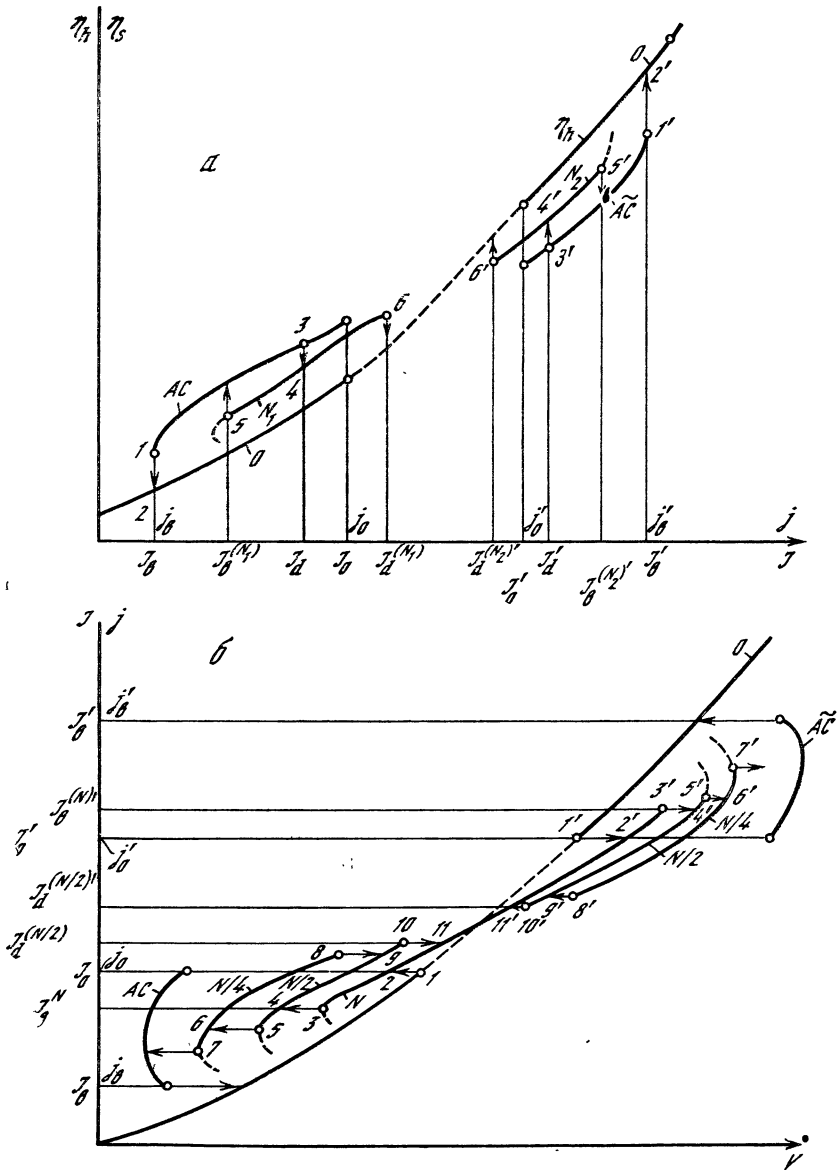


Рис. 4. К пояснению эволюции термодиффузионных АС и страт в плазме с И-образной кривой ЛС

а — зависимости параметров  $\eta_h$  для однородной ЭДП и  $\eta_s$  в стенке АС и страт от уровня возбуждения плазмы  $J \propto j^2$ ; б — вид вольт-амперной характеристики плазмы при образовании в ней поперечных статических АС и страт (в случае генератора тока во внешней цепи образца)

В горячей устойчивой ЭДП можно возбудить АС в виде области низкой температуры и высокой концентрации носителей (см. рис. 2, б). Такой холодный АС может существовать в диапазоне  $J'_0 < J < J'_b$  (кривая  $\widetilde{AC}$  на рис. 4, а), где  $J'_0$  определяет верхнюю границу устойчивости ЭДП (см. рис. 1); в точке  $J = J'_b$  холодный АС (см. рис. 2, б) исчезает (скачок  $1' \rightarrow 2'$  на рис. 4, а). В этой точке  $d\eta_s/dJ = \infty$ , а  $\mathcal{L}_s \sim l \ln(L/l)$ .

При определенных параметрах ЭДП, не доходя до точки  $J = J_0$ , при некотором значении  $J = J_d < J_0$  (точка 3 на рис. 4, а) может произойти спонтанное деление широкого горячего АС сначала на два, а затем и на большее число АС [32]. В результате последовательного деления АС весь образец заполняется стратами (скачок  $3 \rightarrow 4$  на рис. 4, а). Аналогичный эффект может возникать в устойчивой ЭДП и при  $J > J'_0$  в результате деления холодного АС (см. рис. 2, б) при некотором значении  $J = J'_d < J'_b$  (скачок  $3' \rightarrow 4'$  на рис. 4, а). Деление АС в устойчивой ЭДП [32] происходит в результате эффекта «локального пробоя» [4, 5, 8, 10] который обсуждается в п. 1.2.

Картину эволюции АС удобно рассматривать, используя бифуркационную характеристику — зависимость значения  $\eta_s$  в стенке АС от  $J \propto j^2$  (рис. 4, а) или вольт-амперную характеристику (ВАХ) образца (рис. 4, б). На рис. 4 кривые  $O$ ,  $AC$  и  $\widetilde{AC}$  отвечают соответственно однородному состоянию плазмы, горячему и холодному АС. Кривые с номером  $N_1, N_2$  на рис. 4, а и номерами  $N, N/2, N/4$  на рис. 4, б отвечают периодическим стратам с числом страт в образце, равным соответственно  $N_1, N_2, N$  и т. д. Пунктирные участки кривых на рис. 4 отвечают неустойчивым состояниям. Стрелками на бифуркационной характеристике (рис. 4, а) показаны возможные скачки параметра  $\eta_s$  при потере устойчивости АС ( $1 \rightarrow 2, 1' \rightarrow 2'$ ) и страт ( $5 \rightarrow \dots, 5' \rightarrow \dots$ ) или же деления АС ( $3 \rightarrow 4, 3' \rightarrow 4'$ ) и страт ( $6 \rightarrow \dots, 6' \rightarrow \dots$ ). Стрелки на ВАХ (рис. 4, б) показывают возможные скачки величины падения напряжения на образце  $V$  при заданном токе через образец в результате потери устойчивости страт (относительно «перекачки»; скачки  $3 \rightarrow 4, 5 \rightarrow 6, 7 \rightarrow \dots; 3' \rightarrow 4', 5' \rightarrow 6', 7' \rightarrow \dots$ ) или же деления страт (в результате «локального пробоя»: скачки  $8 \rightarrow 9, 10 \rightarrow 11; 8' \rightarrow 9', 10' \rightarrow 11'$ ). (Более подробно об эволюции страт см. п. 1.2.)

## §1.2. Широкие периодические страты

В рассматриваемой ЭДП, для которой  $l \ll L$ , а кривая ЛС имеет И-образный вид (см. рис. 3, кривая 1), при  $J > J_0$  в результате неустойчивости однородного состояния спонтанно скачкообразно возникают [5] периодические горячие широкие страты (рис. 5, а). Эти страты можно рассматривать как последовательность АС, изображенных на рис. 2, а. Значения  $T_{\max}, T_{\min}, \eta = \eta_s$  и  $\mathcal{L}_s$  в стратах (рис. 5, а) также определяются из уравнений (11) и (12),

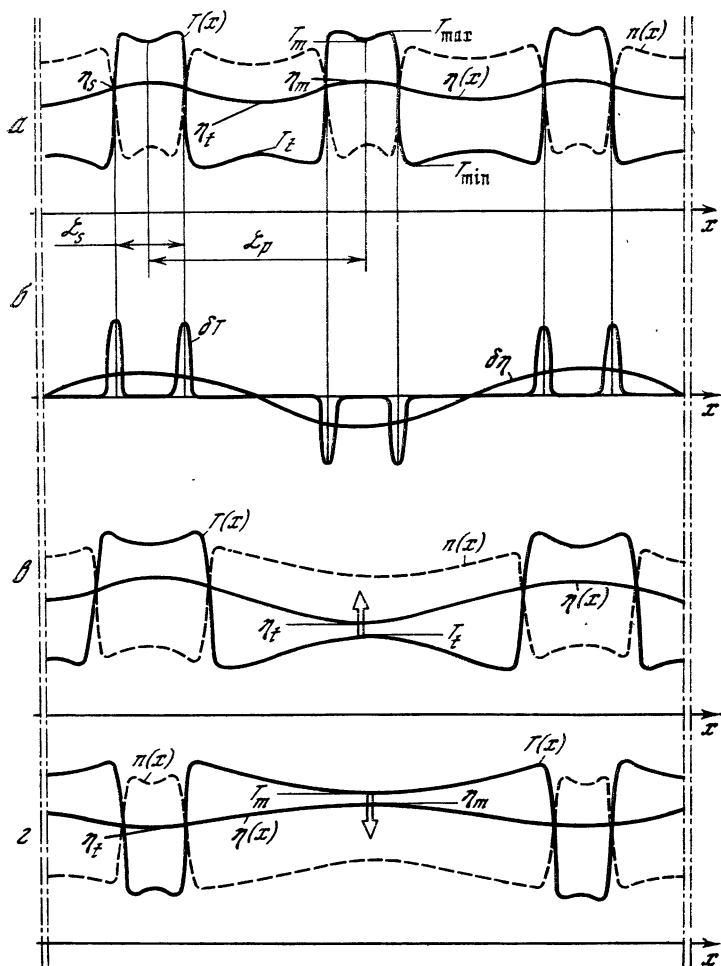


Рис. 5. Широкие периодические термодиффузионные страты периода  $L_p$  (а) и  $2L_p$  (в, г)

б — вид критических для страт а флуктуаций

а значения  $\eta_m$ ,  $T_m$  и  $\eta_t$ ,  $T_t$  в центре горячей страты и между стратами — из уравнений [8]

$$\int_{\eta_s}^{\eta_t} Q(\theta_I(\eta), \eta) d\eta = - \int_{\eta_m}^{\eta_s} Q(\theta_{III}(\eta), \eta) d\eta, \quad q(\eta_m, T_m, J) = 0, \quad (14)$$

$$g(\eta_t, T_t, J) = 0, \quad L_p - L_s = \sqrt{2}L \int_{\eta_s}^{\eta_t} \left[ \int_{\eta_t}^{\eta} Q(\theta_I(\eta), \eta) d\eta \right]^{-1/2} d\eta. \quad (15)$$

**Флуктуационное удвоение периода страт (эффект «перекачки»)**. Период страт  $L_p$  ограничен снизу некоторой величиной  $L_{\min}$ , зависящей от величины  $J$  [4, 5]. Критическая флуктуация температуры  $\delta T$ , нарастание которой приводит к неустойчивости страт с периодом  $L_p = L_{\min}(J)$ , локализована в стенках страт (рис. 5, б) и отвечает увеличению ширины (или амплитуды) одной страты за счет уменьшения ширины (или амплитуды) соседней страты [4, 5]. Развитие такой аperiodической неустойчивости типа «перекачки» приводит к скачкообразному уменьшению числа страт в образце вдвое [5], т. е. к удвоению их периода (рис. 5, в).

**Динамическое деление и уменьшение периода страт (эффект «локального пробоя»)**. В области значений  $J_0 < J < J'_0$ , где однородное состояние ЭДП неустойчиво, период страт  $L_p$  ограничен не только снизу, но и сверху [4, 5] некоторой величиной  $L_{\max}$ , зависящей от  $J$ . Значение  $L_{\max}$  определяется тем, что при больших  $L_p$  величина  $\eta = \eta_t$  между стратами (см. рис. 5, в) достигает значения, равного  $\eta_0$  (см. рис. 3) [4, 5]. В результате происходит «локальный пробой» между стратами, т. е. лавинообразное возрастание  $T$  (отмечено стрелкой на рис. 5, в), приводящее к возникновению в этих областях новых страт (см. рис. 5, а) [5]. Другими словами, при увеличении уровня возбуждения  $J$  число страт в образце скачкообразно увеличивается.

Для пояснения причины возникновения «локального пробоя» в стратах заметим, что связь между величинами  $\eta$  и  $T$  вне стенок страт (вне областей резкого изменения  $T(x)$  на рис. 5, а, в, г) определяется ветвями I и III кривой ЛС (см. рис. 3, кривая I) [5, 6, 8]. Поэтому в случае, когда минимальное значение  $\eta = \eta_t$  в стратах достигнет величины  $\eta_0$  (см. рис. 3), в областях, расположенных между стратами, должен произойти скачок температуры (стрелка на рис. 5, в) от величины  $T_t \approx T_0$  до  $T'_1$  (см. рис. 3) [4, 5]. Этот эффект в какой-то мере напоминает скачок температуры в полупроводниках с S-образной ВАХ [27] при значениях  $E$ , больших пробойного. Нетривиальность эффекта «локального пробоя» состоит в том, что ВАХ рассматриваемой ЭДП однозначна (кривая O — O на рис. 4, б), а локальный пробой между стратами связан с тем, что величина  $\eta$  может существенно изменяться на длине порядка  $L$ , в то время как температура — на длине порядка  $l \ll L$ . Поэтому при скачкообразном увеличении температуры в области между стратами размера порядка  $l$  величина  $\eta$  локально измениться не может. Иными словами, скачок температуры от  $T_t = T_0$  до  $T'_1$  происходит при практически постоянном в области между стратами значении величины  $\eta \approx \eta_0$ . Обусловленное «локальным пробоем» спонтанное увеличение количества страт в образце, т. е. уменьшение периода страт, происходит не в результате потери устойчивости страт, а вследствие динамической перестройки структуры [33].

«Локальный пробой» может происходить и в центре горячей области страты, когда значение  $\eta = \eta_m$  (см. рис. 5, г) при увеличении  $J$  достигает величины  $\eta'_0$  (см. рис. 3) [5]. В этом случае в цент-



ре горячей области происходит лавинообразное локальное уменьшение  $T$  от величины  $T_m \approx T'_0$  до  $T_1$  (см. рис. 3) (символически изображено стрелкой на рис. 5, з). «Локальный пробой» в этом случае приводит к делению страты, т. е. к увеличению количества страт в образце [5, 8].

«Локальный пробой» в стратах в ряде случаев может возникать и в области значений  $J < J_0$  и  $J > J'_0$  (см. рис. 4), где однородное состояние ЭДП устойчиво. Однако при  $J < J_0$  возможен «локальный пробой» лишь в центре горячей области (в окрестности  $\eta = \eta_m$ , рис. 5, а, з) [34], а при  $J > J'_0$  — в области между стратами (в окрестности  $\eta = \eta_t$ , рис. 5, а, в). Последнее связано с тем, что при  $J < J_0$  величина  $\eta = \eta_t \geq \eta_h > \eta_0$  (знак равенства отвечает случаю  $\mathcal{L}_p = \infty$ , т. е. уединенной страте — АС, см. рис. 2, а), а при  $J > J'_0$  величина  $\eta = \eta_m \leq \eta_h < \eta'_0$  [5, 6, 8–10]. Другими словами, в стратах при  $J < J_0$  величина  $\eta_t$  (см. рис. 5, а) не может достигнуть критического значения  $\eta = \eta_0$ , а при  $J > J'_0$  величина  $\eta_m$  (см. рис. 5, а, з) не может достигнуть критического значения  $\eta = \eta'_0$  (см. рис. 3).

**Эволюция АС и страт при изменении тока. Вид ВАХ образца.** Рассмотренные выше эффекты «перекачки» и «локального пробоя» определяют сложную картину эволюции страт в плазме при изменении уровня ее возбуждения  $J$  (в данном случае тока  $j$ ), характеризующуюся многочисленными участками гистерезиса на ее ВАХ (см. рис. 4, б).

При расслаении однородного состояния [15, 28, 29], обусловленного нарастанием критической флуктуации с  $k = k_0 \approx (lL)^{-1/2}$ , при заданном через образец токе  $J = J_0$  с наибольшей вероятностью скачкообразно образуется (скачок 1 → 2 на рис. 4, б) состояние в виде  $N = \mathcal{L}_x / \mathcal{L}_k$  ( $\mathcal{L}_x$  — размер образца,  $\mathcal{L}_k = 2\pi k_0^{-1} \sim (lL)^{1/2}$ ) периодически расположенных широких страт (см. рис. 5, а) [5]. При уменьшении  $J$  широкие горячие страты сужаются и становятся неустойчивыми относительно «перекачки» при некотором  $J = J_b^{(N)}$  (см. рис. 4, б). Нарастание критической флуктуации  $\delta T$  (см. рис. 5, б) может привести к скачкообразному (скачок 3 → 4 на рис. 4, б) уменьшению числа страт в образце вдвое (см. рис. 5, в). При дальнейшем уменьшении  $J$  в результате последовательности таких бифуркаций удвоения периода будет происходить последовательное скачкообразное уменьшение числа страт в образце (скачки 5 → 6, 7 → ... на рис. 4, б). В итоге в образце может спонтанно возникнуть горячий АС (кривая АС на рис. 4), который скачкообразно исчезнет в точке  $J = J_b$ , где  $d\eta_s/dJ = dV/dJ = \infty$ , и образец станет пространственно-однородным (скачок 1 → 2 на рис. 4, а).

Напротив, при увеличении  $J$  страты периода  $\mathcal{L}_k$ , спонтанно образующиеся при  $J = J_0$  (см. рис. 4, б), расширяются, все более приобретая вид холодных страт (см. рис. 5, з). Последние в точке  $J = J_b^{(N)'}$  теряют устойчивость относительно перекачки (скачок

$3' \rightarrow 4'$  на рис. 4, б) [5]. В результате последовательности таких бифуркаций удвоения периода в образце может спонтанно образовываться холодный АС (см. рис. 2, б), который скачкообразно исчезнет в точке  $J = J'_b$ , где  $d\eta_s/dJ = dV/dJ = \infty$  (см. рис. 4).

При изменении  $J$  число страт в образце может изменяться также в результате эффекта «локального пробоя». Возможное при этом удвоение количества страт в образце проиллюстрировано на рис. 4, б скачками  $8 \rightarrow 9$ ,  $10 \rightarrow 11$  на ВАХ, вызванными «локальным пробоем» между стратами (стрелка на рис. 5, в), и скачками  $8' \rightarrow 9'$ ,  $10' \rightarrow 11'$ , обусловленными «локальным пробоем» в центре страт (стрелка на рис. 5, з) [5].

Еще более сложная картина эволюции АС и страт при изменении  $J$  возможна в случаях, когда «локальный пробой» в стратах возникает не только в диапазоне  $J_0 < J < J'_0$ , но также и при  $J < J_0$  и  $J > J'_0$ , где плазма устойчива. В частности, появляется возможность спонтанного заполнения устойчивой ЭДП стратами за счет «локального пробоя» в центре горячего или холодного АС. Эти процессы соответственно приведут к скачкам  $3 \rightarrow 4$  и  $3' \rightarrow 4'$  на рис. 4, а. В результате такого пробоя АС сначала делится на два, а затем большее число АС, которые, трансформируясь в страты, со временем заполняют весь образец [32, 34].

### 1.3. Два типа пичковых автосолиитонов и страт

Выше были рассмотрены АС (см. рис. 2) и страты (см. рис. 5) амплитуда которых  $(T_{\max} - T_{\min}) = (T_{s3} - T_{s1})$  определяется наличием ветви III однозначной зависимости  $\theta(\eta)$  на И-образной кривой ЛС (см. рис. 3, кривая 1). Значения  $T > T'_0$ , отвечающие этой ветви, можно в определенном смысле рассматривать как соответствующую  $\eta = \eta_s$  «горячую устойчивую фазу» ЭДП с температурой  $T \approx T_{s3}$ .

Иная ситуация реализуется в ЭДП, которой отвечает V-образная ЛС (см. рис. 3, кривая 2). В такой ЭДП при  $l \ll L$  могут существовать пичковые АС и страты двух типов: широкие (рис. 6, а, в) [35, 36] и узкие (рис. 6, б, з) [4, 35, 36]. Полуширина пичка  $L_s/2$  для широкого АС и страты независимо от малости  $\epsilon = l/L \ll \ll 1$  составляет величину порядка  $L$  (рис. 6, а), а для узкого АС и страты — порядка  $l$  (рис. 6, б).

Поясним на рассматриваемом примере возникновение в ЭДП широких пичковых АС (рис. 6, а) [35]. Амплитуда  $T_{\max}$  пичковых АС не ограничивается видом кривой ЛС, а определяется диффузионным расплыванием температуры вследствие теплопроводности носителей, а также особенностями интегрального баланса генерации и рекомбинации носителей в АС. Этот баланс в образце размера  $L_s \gg L$  при  $\tau_r = \text{const}$  описывается соотношением

$$\int_0^{L_x} n(x) dx = L_x n_h \equiv L_x G \tau_r. \quad (16)$$

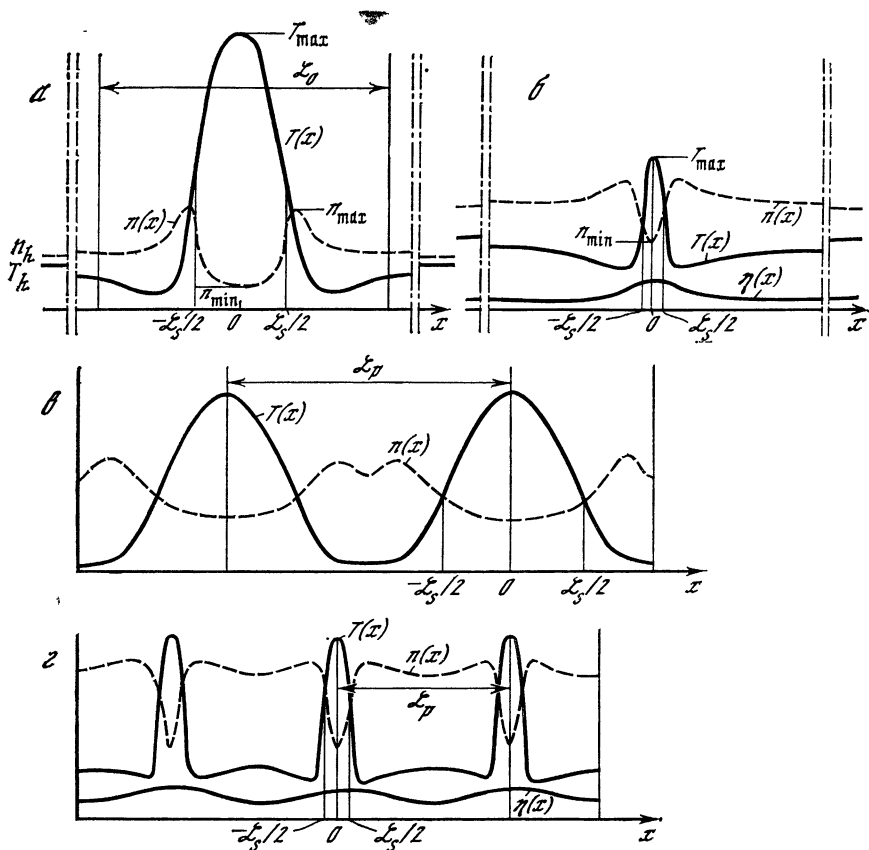


Рис. 6. Пиковые термодиффузионные АС и страты

а, б — соответственно широкий и узкий АС; в, г — широкие и узкие периодические страты

Из (16) приближенно вытекает, что

$$(n_{\max} - n_h)(L_0 - L_s) \approx (n_h - n_{\min})L_s, \quad (17)$$

где  $n_{\max}$  и  $n_{\min}$  — максимальное и минимальное значения концентрации ЭДП в АС;  $L_0$  — размер области локализации АС (см. рис. 6, а). Учитывая, что величина  $n_{\max} - n_h \sim n_h$ , а  $L_0 \sim L$ , из выражения (17) следует, что  $L_s \sim L$ . Поскольку в ряде случаев  $T(x) \propto \exp(-x/l)$ , то легко сделать вывод [35], что при такой ширине пика АС его амплитуда должна иметь гигантскую величину уже при небольших значениях отношения  $L/l > 1$ . Эти качественные соображения подтверждаются численными исследованиями [35], в которых установлено, что уже при  $\varepsilon = l/L = 0,1$  в широком пиковом АС величина  $T_{\max} \approx 100 T_l$ . При уменьшении  $\varepsilon$  значение  $T_{\max}$  резко увеличивается и при  $\varepsilon = 1/15$  составляет примерно  $10^3 T_l$ . Иными словами, при малых  $\varepsilon$  средняя энер-

гия носителей существенно превосходит величину порога ударной ионизации, что необходимо учитывать при нахождении амплитуды АС [35].

В рассматриваемой ЭДП с  $\epsilon \ll 1$  в относительно слабых электрических полях, при которых ее однородное состояние становится неустойчивым, могут спонтанно образовываться страты в виде областей ударной ионизации носителей [36]. Наличие ударной ионизации носителей в пичковых АС или стратах может привести к формированию в ЭДП не широких, а узких пичковых АС и страт (рис. 6, б, г). Об этом свидетельствуют результаты численного моделирования АС и страт, проведенного в [35, 36]. Установлено, что ударная ионизация резко ограничивает амплитуду образующихся узких пичковых АС и страт. Несмотря на интенсивную ударную ионизацию, концентрация носителей в центре АС (страты)  $n_{\text{min}}$  меньше исходной концентрации носителей в однородной ЭДП  $n_h$ . Это связано с самой природой существования устойчивого термодиффузионного АС в ЭДП: АС существует благодаря тому, что термодиффузия горячих носителей из центра АС на его периферию выбрасывает электроны и дырки, образующиеся в результате и межзонной ударной ионизации, и фотогенерации носителей.

## 2. ОБЩИЕ УСЛОВИЯ СУЩЕСТВОВАНИЯ ТЕРМОДИФфуЗИОННЫХ АВТОСОЛИТОНОВ И СТРАТ В ПЛАЗМЕ

Из результатов теории [4—10], кратко изложенных в п. 1.1—1.3, следует, что образование термодиффузионных АС и страт в невырожденной плазме происходит, когда  $L \gg l$ , а кривая локальной связи (ЛС) является немонотонной (см. рис. 3). В большинстве случаев необходимое условие  $L \gg l$  в плазме выполняется благодаря тому, что время рекомбинации носителей  $\tau_r \gg \tau_e, \tau_h$  ( $\tau_e, \tau_h$  — времена релаксации энергии электронов и дырок). Поэтому для выяснения возможности образования АС и страт в плазме достаточно проанализировать кривую ЛС. Из вида кривой ЛС можно непосредственно сделать выводы о форме образующихся АС и страт (см. рис. 2, 5 и 6). При этом в случае И-образной кривой ЛС можно по формулам (11) легко оценить также максимальную и минимальную температуру и концентрацию носителей в АС и стратах (см. п. 1.2).

Уравнение для нахождения кривой ЛС вытекает из уравнения баланса энергии носителей в однородной плазме, если в нем концентрацию горячих носителей  $n$  выразить через их температуру  $T$  и некоторый параметр  $\eta$ . Параметр  $\eta = \eta(n, T)$  определяется как величина, пространственное изменение которой в плазме характеризуется биполярной диффузионной длиной носителей  $L$ . Для нахождения  $\eta$  необходимо, учитывая условие квазинейтральности плазмы, определить суммарный диффузионный и термодиффузионный потоки носителей  $j_D$ . В большом числе случаев выражение

для  $\mathbf{j}_D$  можно записать в виде

$$\mathbf{j}_D = f(n, p, T, T_l) \nabla \eta. \quad (18)$$

Вид функции  $\eta = \eta(n, p, T, T_l)$  зависит от параметров плазмы. Например, в рассмотренном в разд. 1 примере величина  $\eta = nD(T)$ . В плазме с достаточно высокой концентрацией носителей  $\eta = nT$  (см. п. 3.2 и п. 4.1) или же  $\eta = n(T + T_l)$  (см. п. 3.3 и п. 4.2), т. е. параметр  $\eta$  в плотной плазме есть давление газа носителей.

**Поперечные и продольные АС и страты.** Следует различать АС и страты, перпендикулярные и параллельные линиям тока  $\mathbf{j}$ . Перпендикулярные линиям тока АС и страты естественно называть поперечными, а параллельные линиям тока — продольными<sup>3</sup>.

В рассматриваемых ниже случаях квазинейтральная плазма удовлетворяет уравнениям

$$\operatorname{div} \mathbf{j} = 0, \quad \operatorname{rot} \mathbf{E} = 0. \quad (19)$$

Согласно (19), в поперечных стратах меняется электрическое поле  $E$ , а плотность полного тока остается неизменной ( $j = \text{const}$ ). Напротив, в продольных стратах изменяется плотность тока  $j$ , а поле в образце неизменно ( $E = \text{const}$ ). Отсюда следует, что, подставляя в уравнение баланса энергии носителей в однородной ЭДП условие  $j = \text{const}$  и величину  $\eta$ , получим уравнение локального баланса энергии носителей  $q(\eta, T, J) = 0$ , где  $J \propto j^2$ . Если вытекающая из этого уравнения кривая ЛС оказывается И- или V-образной (см. рис. 3), то в рассматриваемой плазме могут возникать поперечные АС и страты.

Если же в уравнение баланса энергии носителей для однородной ЭДП подставить условие  $E = \text{const}$ , то найдем уравнение локального баланса энергии носителей  $q(\eta, T, J) = 0$ , где  $J \propto E^2$ . Вид  $\eta(T)$ , т. е. вид ЛС, вытекающий из этого уравнения, определяет возможность существования в плазме продольных АС и страт, т. е. распределений  $T$  и  $n$ , неоднородных в направлении, перпендикулярном линиям тока. Заметим, что распределение  $T$  и  $n$  в продольных АС и стратах, а также свойства этих автоструктур качественно не отличаются от рассмотренных в разд. 1 на примере ЭДП, в которой возникают поперечные страты (см. рис. 2, 5 и 6).

При рассмотрении распределений  $T(x)$  и  $n(x)$  в продольных АС и стратах необходимо считать, что ось  $x$  на рис. 2, 5 и 6 направлена перпендикулярно линиям тока  $\mathbf{j}$ , причем в этом случае, как отмечалось, значение  $j$  зависит от  $x$ , а величина электрического поля  $E$ , напротив, остается постоянной.

<sup>3</sup> Возможность неустойчивости ЭДП относительно флуктуаций с волновым числом  $k \perp E$ , нарастание которых приводит к спонтанному возникновению продольных страт, выяснена в работах [14, 16].

Поперечные и продольные АС и страты могут возникать при разогреве плазмы как постоянным, так и высокочастотным полем [15, 19, 35, 36]. При рассмотрении высокочастотного поля будем предполагать, что длина его волны много больше размеров образца (плазмы), а частота  $\omega$  удовлетворяет условиям

$$\hbar\omega \ll T, \quad \omega\tau_M \ll 1, \quad \omega\tau_e \gg 1, \quad (20)$$

где  $\tau_M$  — максвелловское время релаксации. Последнее из условий (20) позволяет считать, что температура электронов  $T$  не успевает следить за изменением поля. Отметим, что при высокочастотном разогреве носителей в поперечных стратах  $\bar{j}^2 = \text{const}$ , а изменяется значение  $\bar{E}^2$ . Напротив, в продольных стратах  $\bar{E}^2 = \text{const}$ , но меняется величина  $\bar{j}^2$ . (Величины  $\bar{j}^2$  и  $\bar{E}^2$  есть усредненные за период колебаний поля значения  $j^2$  и  $E^2$ .)

Изложенная в данном разделе программа по определению вида кривой ЛС, а следовательно, свойств АС и страт выполнена ниже для нескольких конкретных случаев, реализующихся в полупроводниковой или газовой плазме при разных физических параметрах и условиях разогрева. Отдельно рассмотрены случаи реализации поперечных (разд. 3) и продольных (разд. 4) АС и страт. Заметим, что в случаях, рассмотренных в разд. 1 и 3, кривая ЛС, отвечающая  $E = \text{const}$  (или  $\bar{E}^2 = \text{const}$ ), является однозначной, т. е. продольные страты в этих случаях не образуются. Напротив, п. 4.1 рассмотрены случаи, когда кривая ЛС оказывается монотонной при  $j = \text{const}$  (или  $\bar{j}^2 = \text{const}$ ), т. е. не образуются поперечные страты.

### 3. ПОПЕРЕЧНЫЕ АВТОСОЛИТОНЫ И СТРАТЫ

#### 3.1. «Неплотная» ЭДП

Рассмотрим полупроводник, в котором за счет фотовозбуждения или инжекции однородно генерируется ЭДП с концентрацией электронов  $n$ , близкой к концентрации дырок  $p$  (величина  $n - p = n_D = \text{const}$ ). Пусть концентрация носителей относительно невысока, а параметры электронов и дырок таковы, что выполнены неравенства (1) и

$$\tau_{eh} \ll \tau_e, \tau_h, \quad (21)$$

где  $\tau_{eh}$  — характерное время рассеяния энергии электронов при их столкновении с дырками. Условия (21) выполняются в полупроводниках, эффективные массы электронов ( $m_e$ ) и дырок ( $m_h$ ), которых отличаются не сильно, или же в полупроводниках, имеющих вырожденную валентную зону. В первом случае  $\tau_{eh} \approx \tilde{\tau}_{eh} (m_e + m_h)^2 (2m_e m_h)^{-1}$ , где  $\tilde{\tau}_{eh}$  — характерное время релаксации импульса электронов при их столкновении с дырками. В полупроводниках с вырожденной валентной зоной эффективность передачи энергии от горячих электронов к тяжелым дыркам, т. е.

малость величины  $\tau_{el}$ , определяется тем, что при столкновении с электроном тяжелая дырка переходит в подзону легких дырок [37].

При выполнении (21) вследствие интенсивных электронно-дырочных соударений носители разогреваются в постоянном или в высокочастотном поле как единая система до некоторой эффективной температуры  $T$ . Когда частота поля такова, что  $\omega\tau_{e,n} \ll 1$ , уравнения, описывающие распределение  $n$  и  $T$  в ЭДП, фактически сводятся к (6), (7), в которых  $j^2 = \bar{j}^2$ . Условие существования АС и страт (см. рис. 2, 5 и 6) в рассматриваемой ЭДП сводится к (10).

В разд. 1 отмечалось, что условие (10) выполняется при комнатных температурах как в полярных, так и в неполярных полупроводниках (Ge, Si). В последних при  $T > \theta_D$  импульс и энергия носителей рассеиваются на неполярных оптических фононах или же при переходах между эквивалентными долинами. В этих случаях  $\alpha \simeq -1/2$ ,  $s \simeq 1/2$  [26, 38], т. е. условие (10) выполнено, а ЭДП расслаивается при температуре  $T_0 \simeq 2T_l$  (см. рис. 3). При  $T > \theta_D$  условия (1), (21) в Ge и Si, по оценкам, выполняются, когда концентрация ЭДП  $n > 10^{15} \div 10^{17} \text{ см}^{-3}$ . В такой ЭДП в достаточно слабых электрических полях  $E \gtrsim 10^8 \div 10^4 \text{ В/см}$  возможно образование пиковых поперечных АС и страт, температура носителей в которых ограничена процессами ударной ионизации (см. рис. 6, б, в). Другими словами, в сравнительно небольших электрических полях в однородном полупроводнике может происходить спонтанное образование статических (или движущихся; см. разд. 5) локальных областей сильной ударной ионизации носителей [35, 36].

В полупроводниках, в которых эффективная масса носителей сильно анизотропна (например, Ge, Si), величина критического значения поля, при котором спонтанно возникают страты, а также форма страт существенно зависят от кристаллографического направления, в котором приложено поле. При  $T > \theta_D$  этот эффект в значительной степени связан с тем, что в разных кристаллографических направлениях зависимость эффективной массы носителей от их энергии может существенно различаться. Если возрастание эффективной массы (непараболичность зоны) заметно лишь при достаточно больших значениях поля, то условие (10) выполняется, но в ограниченном диапазоне поля. В этом случае кривая ЛС имеет И-образный вид (см. рис. 3, кривая 1) и в образце образуются широкие АС и страты (см. рис. 2 и 5). Чем сильнее возрастает эффективная масса носителей с энергией, тем труднее выполнить условие (10). Поэтому при приложении электрического поля в кристаллографических направлениях, в которых эффективная масса оказывается сильно возрастающей функцией энергии носителей, условие (10) может не выполняться ни при каких значениях поля. В этом случае АС и страты в образце не образуются.

Иная ситуация может реализовываться в полупроводниках с  $m_e \ll m_h$  и вырожденной валентной зоной (InSb, InAs и т. п.). В этом случае сильное возрастание эффективной массы электрона с энергией может фактически не повлиять на выполнение условий

образования страт. Это связано с тем, что в таких полупроводниках при выполнении условий (1) и (21) зависимости коэффициента биполярной диффузии  $D$  и времени релаксации энергии ЭДП  $\tau$  от температуры определяются релаксацией соответственно импульса и энергии не легких электронов, а тяжелых дырок на фононах.

Отмеченные особенности энергетического спектра носителей заряда будут определять условия образования и форму страт и для других рассматриваемых ниже случаев, при которых в ЭДП также могут возникать АС и страты.

### 3.2. «Плотная» ЭДП в высокочастотном поле

Поперечные АС и страты образуются в невырожденной ЭДП с такой высокой концентрацией носителей, когда выполняются условия (21) и

$$\tilde{\tau}_{ee} \ll \tilde{\tau}_e, \tilde{\tau}_h \ll \tau_e, \tau_h. \quad (22)$$

В такой плотной ЭДП не только энергия, но и импульс носителей, приобретаемый в поле, рассеивается в основном при электронно-дырочных соударениях, а полный диффузионный поток носителей равен

$$\mathbf{j}_D = -2\mu_e\mu_h(\mu_e + \mu_h)^{-1} \nabla P, \quad (23)$$

где  $P = nT$  — давление электронного газа. Эффективная температура носителей в однородной ЭДП определяется из уравнения баланса энергии носителей

$$\bar{j}^2/\sigma = 2n(T - T_l)/\tau. \quad (24)$$

В рассматриваемом случае  $\mu_e, \mu_h$  — подвижности электронов и дырок, которые определяются их рассеянием на дефектах решетки полупроводника;  $\tau^{-1} = 2^{-1}(\tau_e^{-1} + \tau_h^{-1})$ ; проводимость плазмы

$$\sigma = (e^2 n \tilde{\tau}_{eh}/m^*) (1 + \omega^2 \tilde{\tau}_{eh}^2)^{-1}; \quad m^* = m_e m_h (m_e + m_h)^{-1}; \quad (25)$$

$\tilde{\tau}_{eh} \propto n^{-1} T^{3/2}$  — характерное время электронно-дырочных столкновений.

При частоте поля  $\omega \gg \tilde{\tau}_{eh}^{-1}$  проводимость  $\sigma = (e^2 n/m^* \omega^2) \times \tau_{eh}^{-1} \propto n^2 T^{-3/2}$ . Из (2) и (23) следует, что с характерной длиной порядка  $L$  в «плотной» ЭДП меняется давление газа носителей, т. е.  $\eta = P = nT$ . Вид кривой ЛС, согласно (24), определяется уравнением

$$q(\eta, T, J) = -J\eta^{-2} T^{3/2} + \eta(T - 1)T^{-1}(\tau^0/\tau) = 0, \quad (26)$$

где  $J = [\tau^0 m^* \omega^2 \tilde{\tau}_{eh}^0 / 2e^2 (n_h^0)^2 T_l] \bar{j}^2$ ;  $n$  и  $T$  измеряются в единицах  $n_h^0$  и  $T_l$ . Из (26) вытекает, что условие существования поперечных АС и страт в рассматриваемом случае сводится к неравенству

$$7/2 + s > 0, \quad (27)$$



а характерное значение температуры  $T = T_0$  (см. рис. 3) определяется из уравнения

$$T = [9/2 + s(T)][7/2 + s(T)]^{-1} T_l, \quad (28)$$

Условие (27) выполняется как при низких ( $T \ll \theta_D$ ), так и при относительно высоких температурах ( $T > \theta_D$ ) в большинстве полупроводников. При  $T \ll \theta_D$  энергия носителей в Si, Ge, GaAs и во многих других полупроводниках рассеивается на акустических фононах ( $s = -1/2$ ) [26, 38]. В этом случае из (28) следует, что величина  $T_0 \approx 1,3 T_l$ , т. е. страты, спонтанно возникают уже при слабом однородном разогреве ЭДП. АС можно возбудить в еще более слабо разогретой ЭДП. Максимальная температура в АС и стратах  $T_{\max}$  (см. рис. 2, а и 5, а) ограничена рассеянием энергии носителей на оптических фононах, а необходимая для выполнения условий (22) концентрация ЭДП составляет  $n \approx 10^{14} \div 10^{16} \text{ см}^{-3}$ . Значительно более высокая концентрация носителей необходима для выполнения условий (22) в области температур  $T > \theta_D$ :  $n \gtrsim 10^{17} \div 10^{18} \text{ см}^{-3}$ .

### 3.3. «Плотная» газовая плазма и сильноасимметричная ЭДП в высокочастотном поле

Условия (22) выполняются и в «плотной» ( $s = p$ ) сильноасимметричной ЭДП с  $m_e \ll m_h$ . При этом могут выполняться условия

$$\tau_h \ll \tau_{eh}, \tau_e, \quad (29)$$

при которых в отличие от предыдущего случая (п. 3.2) в поле как единая система с температурой  $T$  разогреваются только электроны, а температура дырок равна  $T_l$ . Полный диффузионный поток носителей в данном случае равен

$$j_D = -\mu_h \nabla [n(T + T_l)], \quad (30)$$

т. е.  $\eta \equiv P = n(T + T_l)$  — давление плазмы. Уравнение баланса энергии электронов в рассматриваемой однородной ЭДП есть

$$\bar{j}^2 / \sigma_e = n(T - T_l)(1/\tau_{eh} + 1/\tau_e). \quad (31)$$

При  $\omega \bar{\tau}_{eh} \gg 1$  из (25) и (31) следует, что

$$q(\eta, T, J) = -J\eta^{-2} T^{3/2} (T + 1)^2 + \eta(T - 1)(T + 1)^{-1} \times \\ \times \{\tau_e^0 [(T + 1)\tau_{eh}^0 T^{3/2} \eta^{-1}]^{-1} + \tau_e^0 / \tau_e\} = 0, \quad (32)$$

где  $J = [\tau_e^0 \bar{\tau}_{eh}^0 m_e \omega^2 / e^2 (n_h^0)^2 T_l] \bar{j}^2$ . Отсюда вытекает, что при низких температурах ( $T \ll \theta_D$ ) кривая ЛС имеет И-образный вид, а при высоких ( $T \gtrsim \theta_D$ ) может иметь как И-, так и V-образный вид. Таким образом, в данном случае могут реализовываться различные типы поперечных АС и страт (см. рис. 2, 5 и 6).

Такие автоструктуры могут реализовываться и в «плотной» газовой плазме. Действительно, в газовой плазме могут

выполняться условия [39, 40]

$$\tilde{\tau}_{ee} \ll \tilde{\tau}_{ea} \ll \tau_{ea}, \quad (33)$$

$$\tau_{ia} \ll \tau_{ei}, \tau_{ea}, \quad (34)$$

эквивалентные соответственно (22) и (29). Здесь  $\tilde{\tau}_{ei}$ ,  $\tilde{\tau}_{ea}$  и  $\tau_{ei}$ ,  $\tau_{ea}$  — характерные времена релаксации соответственно импульса и энергии электронов на ионах и нейтральных атомах газа;  $\tau_{ia}$  — характерное время релаксации энергии ионов на нейтральных молекулах газа. Для газовой плазмы в уравнениях (30)—(32) необходимо заменить  $\mu_h$ ,  $T_l$ ,  $\tau_{eh}$ ,  $\tau_e$  соответственно на  $\mu_i$  (подвижность ионов),  $T_a$  (температура атомов газа),  $\tau_{ei}$ ,  $\tau_{ea}$ .

#### 4. ПРОДОЛЬНЫЕ АВТОСОЛИТОНЫ И СТРАТЫ

##### 4.1. ЭДП в постоянном и высокочастотном поле

Рассмотрим «плотную» ЭДП, разогретую в постоянном ( $\omega = 0$ ) или высокочастотном ( $\omega \ll \tilde{\tau}_{eh}^{-1}$ ) поле, для которой выполнены условия (21) и (22) [14, 17—19]. В этом случае при  $\overline{E^2} = \text{const}$  (см. разд. 2) уравнение баланса энергии в однородной ЭДП имеет вид

$$\sigma \overline{E^2} = 2n(T - T_l)/\tau. \quad (35)$$

Согласно п. 3.2, при выполнении условий (21) и (22) справедливы выражения (23) и (25). Из (25) при  $\omega \ll \tilde{\tau}_{eh}^{-1}$  следует, что проводимость ЭДП  $\sigma \propto T^{3/2}$  и не зависит от концентрации носителей. Учитывая также, что в данном случае, согласно (23),  $\eta = nT$ , из (35) получим (см. разд. 2) уравнение для кривой ЛС:

$$q(\eta, T, J) = -JT^{3/2} + \eta(T - 1)T^{-1}(\tau^0/\tau) = 0, \quad (36)$$

где  $J = [\tau^0 \sigma^0 / 2n_h^0 T_l] \overline{E^2}$ . Отсюда следует, что необходимое условие существования продольных страт — наличие немонотонной кривой ЛС (см. рис. 3 и разд. 2) — сводится к неравенству

$$3/2 + s > 0, \quad (37)$$

а значение  $T = T_0$  определяется из уравнения

$$T = [5/2 + s(T)][3/2 + s(T)]^{-1} T_l. \quad (38)$$

Продольные страты в виде светящихся светло-голубых областей экспериментально наблюдались в GaAs при комнатной температуре в «плотной» ЭДП, разогретой постоянным электрическим полем [41, 42]. Условие (37) для электронов в GaAs при  $T > \theta_D$  может быть выполнено, когда горячие электроны находятся в основном в верхних долинах, где рассеяние их энергии происходит преимущественно в результате междолинных переходов ( $s \approx 1/2$ ) [26]. Такая ситуация реализуется в области статического домена сильного поля, который, как известно, образуется у анодного антизапорного контакта в тонких пленках GaAs. Именно в таких пленках и наблюдались продольные страты [41, 42].

Условия (21), (22) и (37) могут выполняться для других полупроводников при гораздо меньших электрических полях, чем для GaAs. Условия (21), (22) обсуждались в п. 3.2 при рассмотрении плотной плазмы, разогретой высокочастотным полем с  $\omega \tilde{\tau}_{eh} \gg 1$ , когда образуются поперечные страты. При разогреве плазмы в постоянном поле (или высокочастотном поле с  $\omega \tau_{eh} \ll 1$ ), как видно из (25), существенно меняется зависимость проводимости плазмы  $\sigma$  от  $T$  и  $n$ . Вследствие этого в образце возникают не поперечные, а продольные страты. Таким образом, при уменьшении частоты поля  $\omega$  должен наблюдаться переход от продольных к поперечным стратам. Такой переход должен осуществляться и за счет уменьшения концентрации носителей. Последнее связано с тем, что при уменьшении  $n$  условия (22) нарушаются, а справедливыми окажутся условия (1), при которых  $\sigma$  определяется не электронно-дырочными столкновениями, а рассеянием носителей на решетке ( $\sigma \sim n$ ). Как следствие этого, в неоднородной генерируемой по образцу ЭДП переход от продольных к поперечным стратам может осуществляться в пространстве. Такая ситуация, по-видимому, и наблюдалась в GaAs [41].

#### 4.2. Сильноасимметричная ЭДП и газовая плазма

Рассмотрим разогретую в постоянном или высокочастотном поле с  $\omega \lesssim \tilde{\tau}_{eh}^{-1}$ ,  $\tilde{\tau}_{ei}^{-1}$  полупроводниковую или газовую плазму, удовлетворяющую соответственно условиям (22), (29) или (33), (34). При этом, когда  $\tau_{ei} \ll \tau_{ea}$  (или  $\tau_{eh} \ll \tau_e$ ), горячие электроны рассеивают свою энергию на ионах (или дырках). ВАХ такой однородной плазмы имеет неоднозначный и, как правило, S-образный вид, обусловленный тем, что при  $T > T_{кр} = 1,5T_i$  возникает перегревная неустойчивость электронного газа [27, 43]. В работе [14] выяснено, что в случае  $L \gg l$  в такой полупроводниковой или газовой плазме спонтанно возникают продольные страты (см. рис. 5 и 6) при  $T > T_0 \approx 1,3T_i$ , т. е. на участке ВАХ с положительной дифференциальной проводимостью. При  $T < T_0$  в рассматриваемой плазме можно возбудить горячий продольный АС (см. рис. 2 и 6).

Эти результаты, по существу, можно установить из анализа вида кривой ЛС, построенной для  $\bar{E}^2 = \text{const}$  (см. разд. 2). Для газовой плазмы ЛС определяется уравнением

$$q(\eta, T, J) = -JT^{3/2} [1 + \omega^2 (\tilde{\tau}_{ei}^0)^2 (T + 1)^2 T^3 \eta^{-2}]^{-1} + \\ + \eta (T - 1) (T + 1)^{-1} \{ \tau_{ea}^0 [\tau_{ei}^0 (T + 1) T^{3/2} \eta^{-1}]^{-1} \} + \\ + (\tau_{ea}^0 / \tau_{ea}) = 0, \quad (39)$$

где, согласно (2) и (30),  $\eta = n(T + T_a)$ ;  $n$  и  $T$  измеряются в единицах  $n_h^0$  и  $T_a$ ;  $J = (\tau_{ea}^0 e^2 \tilde{\tau}_{ei}^0 / m_e T_a) \bar{E}^2$ . Из (39) видно, что при  $\omega \tilde{\tau}_{ei} \ll 1$  кривая ЛС имеет И- или V-образный вид, причем в случае  $\tau_{ei} \ll \tau_{ea}$  величина  $T_0 \approx 1,3T_a$ . Заметим, что для рассматриваемой плазмы при  $\omega \tilde{\tau}_{ei} \ll 1$  и  $\tau_{ei} \ll \tau_{ea}$  кривая ЛС, построенная

при  $\bar{j}^2 = \text{const}$  (см. разд. 2), оказывается также немонотонной, однако в этом случае характерное значение температуры для кривой ЛС  $T_0 = 3T_a$ , т. е. существенно выше  $T_0 \approx 1,3T_a$ . Выражение (39) справедливо и для асимметричной ЭДП, если в нем заменить  $\eta$ ,  $\tau_{ei}$ ,  $\tilde{\tau}_{ei}$ ,  $\tau_{ea}$ ,  $T_a$  соответственно на  $\eta = n(T + T_l)$ ,  $\tau_{e1}$ ,  $\tilde{\tau}_{e1}$ ,  $\tau_e$ ,  $T_l$ .

## 5. ДВИЖУЩИЕСЯ ПОПЕРЕЧНЫЕ СТРАТЫ

При рассмотрении поперечных страт (см. разд. 1 и 3) для простоты не обсуждались случаи, когда поперечные страты оказываются движущимися. Такая ситуация может возникать при разогреве ЭДП в постоянном электрическом поле вследствие реально существующего различия параметров электронов и дырок. Направление движения страт зависит от соотношения между подвижностями  $\mu_e$  и  $\mu_h$ , концентрациями  $n$  и  $p$  и т. п. При  $n \neq p$ , как известно, появляется амбиполярный снос возмущений плотности квазинейтральной ЭДП, который при  $n > p$  направлен в сторону дрейфа дырок. С другой стороны, в разогретой ЭДП возникает снос возмущений температуры носителей, вызванный потоками их энергий, который при  $n = p$  и  $\mu_e > \mu_h$  направлен в сторону дрейфа электронов. Поэтому, например, при  $n > p$ , но  $\rho = (n - p)/n \ll \ll 1$  в зависимости от величины  $\rho$  и отношения  $\mu_e/\mu_h$  страты могут двигаться как к катоду, так и в противоположном направлении — к аноду. Эти качественные выводы подтверждаются результатами численного моделирования поперечных страт [44].

## 6. О ВЛИЯНИИ НЕОДНОРОДНОСТЕЙ ОБРАЗЦА

Представленная в п. 1.2 эволюция АС и периодических страт (см. рис. 4) относится к исходно идеально однородной плазме. В реальных образцах эволюция АС и страт в большой мере определяется условиями на границах образца, в частности на контактах, и наличием в образце малых неоднородностей. Последние вблизи бифуркационных значений  $J$  ( $J_0$ ,  $J'_0$ ,  $J_b$ ,  $J'_b$ ,  $J_b^{(N)}$ ,  $J_b^{(N)'}$ ,  $J_d$ ,  $J'_d$  и т. д. на рис. 4) сильно влияют на вероятность образования одной из возможных при данном  $J$  структур. Так, например, вблизи точек  $J_b^{(N)}$  или  $J_d^{(N/2)}$  (см. рис. 4, б) число страт в образце может соответственно уменьшаться или увеличиваться не вдвое, а, как правило, на одну страту [8, 10, 33]. При этом, естественно, могут образовываться несимметричные и непериодически расположенные страты [4, 6, 8, 10, 45, 46].

Малая локальная неоднородность, возникающая, в частности, у анодного или катодного контактов образца, может проявляться как затравка для скачкообразного спонтанного возникновения АС в устойчивой плазме при значении уровня разогрева ЭДП  $J$ , близких к критическому значению  $J_0$  (или  $J'_0$ ) (см. рис. 4). Форма образующегося АС (см. рис. 2 и 6) практически не будет зависеть

от вида неоднородности. При наличии сноса (см. разд. 5) образовавшийся в устойчивой ЭДП поперечной АС может оторваться от локальной неоднородности образца и, пробежав по образцу, исчезнуть на контакте, вызвав всплеск тока (или напряжения) во внешней цепи образца. Последовательная генерация таких движущихся поперечных АС может привести к появлению периодических или стохастических высокочастотных колебаний [35].

Явления на контактах могут существенно влиять на свойства поперечных движущихся страт (см. разд. 5). Они могут привести к остановке страт во всем образце или в некоторой его части. При однородной генерации ЭДП в образце  $n$ -типа с антизапорными (омическими)  $n^+$ -контактами вблизи анода может спонтанно образовываться контактная анодная термодиффузионная статическая страта высокой температуры и пониженной концентрации носителей при значениях  $J$ , в десятки раз меньших  $J_0$  (см. рис. 4) [44]. За счет эффекта вытягивания неравновесных носителей вблизи контактной страты, в свою очередь, возникает область повышенного значения поля. В результате в этой области при значениях  $J$ , близких к  $J_0$ , происходит спонтанное образование статических и движущихся к катоду (при  $n > p$ ) страт [44].

## 7. ПУЛЬСИРУЮЩИЕ ТЕРМОДИФУЗИОННЫЕ АВТОСОЛИТОНЫ И СТРАТЫ

Термодиффузионные статические АС и страты могут при некоторых значениях  $J$  превращаться в пульсирующие <sup>4</sup>. Это происходит в тех случаях, когда время жизни носителей  $\tau_r$  есть возрастающая функция их температуры  $T$  или падающая функция концентрации носителей  $n$ . Частота пульсаций на пороге неустойчивости статических автоструктур примерно равна  $\omega_c \approx a_0^{1/2} (\tau_r \tau)^{-1/2}$ , где для широкого АС (см. рис. 2) коэффициент  $a_0 \approx l/L$ , а для широких страт (см. рис. 5)  $a_0 \approx l/\mathcal{L}_p$  [48].

## 8. ТЕРМОДИФУЗИОННЫЕ АВТОСТРУКТУРЫ В СИСТЕМАХ С «ОТРИЦАТЕЛЬНОЙ» ТЕРМОДИФУЗИЕЙ

Выше были рассмотрены термодиффузионные автоструктуры, образующиеся за счет того, что горячие носители вследствие «положительной» термодиффузии выбрасываются из области повышенного значения их температуры. Поэтому в таких АС и стратах температура и концентрация носителей изменяются в противофазе (см. рис. 2, 5 и 6).

Кроме рассмотренных систем с «положительной» термодиффузией, существуют и системы с «отрицательной» термодиффузией. К ним, в частности, относится неравновесная ЭДП, находящаяся в термодинамическом равновесии с решеткой [4], а также двухкомпонентная смесь легкого и тяжелого газов. В такой смеси, чрезвы-

<sup>4</sup> Пульсирующие автоструктуры в активных системах с диффузией были предсказаны в [47, 48], а численно изучены в [49, 50].

чайно слабо разогретой электромагнитным излучением, селективно поглощающимся частицами легкого газа <sup>5</sup>, можно возбудить АС, температура и концентрация легкого газа в котором изменяются в фазе [54]. Существование АС определяется тем, что в области, где высока концентрация легкого газа, происходит наибольшее поглощение электромагнитного излучения, т. е. осуществляется наибольший разогрев, а диффузионному расплыванию концентрации легкого газа препятствует термодиффузионный поток частиц легкого газа, направленный от периферии в центр горячей области.

Форму и свойства АС и страт в системах с «отрицательной» термодиффузией можно непосредственно установить, используя результаты общей теории автоструктур большой амплитуды [4—8, 10]. Форма АС и страт определяется видом кривой ЛС. Уравнение, определяющее кривую ЛС, здесь, как и выше (см. разд. 1 и 2), вытекает из уравнения баланса энергии однородной системы, если в нем концентрацию  $n$  выразить как функцию температуры и параметра  $\eta$ , который плавно изменяется вдоль системы. Параметр  $\eta$  для смеси и тяжелого газа определяется [54] из выражения для потока частиц легкого газа [55]

$$j_D = - T \nabla [nD(T)T^{-1}] \equiv - T \bar{v} \eta, \quad (40)$$

где  $D(T)$  — коэффициент диффузии легкого газа в тяжелом. В системах с «отрицательной» термодиффузией в отличие от рассмотренных выше производная  $\partial \eta / \partial T < 0$ . В результате в них термодиффузионные АС и страты при  $L \gg l$  возникают, когда кривая ЛС имеет N- или  $\Lambda$ -образный вид.

Форма и свойства автоструктур в N- и  $\Lambda$ -системах изучены в работах [4—7]. Из них следует, что распределения температуры в термодиффузионных АС и стратах в системах с «отрицательной» термодиффузией качественно совпадают с изображенными на рис. 2, 5 или 6 соответственно для случаев N- или  $\Lambda$ -образного вида кривой ЛС. Однако существенно, что в данном случае концентрация частиц в АС и стратах изменяется не в противофазе, а в фазе с изменением их температуры.

Другим примером системы с «отрицательной» термодиффузией может быть невырожденная ЭДП, возникающая при однородном поглощении света в тонкой пленке полупроводника [4]. При достаточно высоких температурах ЭДП успевает термализоваться с решеткой, а подводимая к образцу при освещении мощность затрачивается в основном на разогрев решетки [4]. Температура ЭДП и решетки ( $T$ ) описывается усредненным по толщине пленки уравнением теплопроводности

$$\frac{nE_g/\tau_r}{\kappa} = (T - T_s)\kappa(T)l_T^2, \quad (41)$$

<sup>5</sup> В рассматриваемом газе, сильно разогретом электромагнитным излучением, поглощающимся легким компонентом газа, как установлено в [51—53], может возникать неустойчивость, приводящая к расслоению исходно однородного состояния газа. Кинетика такого расслоения численно изучалась в [53].

где  $E_g$  — ширина запрещенной зоны полупроводника;  $\kappa$  — теплопроводность решетки;  $l_T$  — характерная длина изменения температуры;  $T_s$  — температура подложки. В данном случае вместо (8) имеем

$$j_D = -v \nabla (nDv^{-1}) \equiv -v \nabla \eta, \quad (42)$$

где  $v = \exp \left\{ \int^T \delta d \ln T \right\}$ ;  $\delta = \partial \ln \tau_{e,h} / \partial \ln T$ ;  $\tau_{e,h}(\epsilon, T) = f(T) \varphi(\epsilon)$  — время рассеяния импульса носителя с энергией, равной  $\epsilon$ . Если  $\partial \eta / \partial T < 0$  и  $L \gg l_T$ , то в образце могут существовать АС и страты (как правило, узкие пиковые), в которых  $T$  и  $n$  изменяются в фазе [4]. В последнем можно убедиться из анализа уравнения локальной связи  $\eta$  и  $T$ , которое следует из (41), если в него, согласно (42), подставить  $n = \eta v(T) [D(T)]^{-1}$ . Вытекающая из этого уравнения кривая ЛС имеет, как правило,  $\Lambda$ -образный вид.

## 9. «ШАРОВАЯ МОЛНИЯ» В ПОЛУПРОВОДНИКАХ И ГАЗАХ

В газовой и полупроводниковой плазме, разогретой постоянным полем или электромагнитным излучением, кроме рассмотренных выше одномерных АС и страт, могут образовываться двух- и трехмерные термодиффузионные автоструктуры, в том числе АС. АС представляет собой область высокой температуры электронов или частиц газа. В ряде случаев такое сильно неравновесное локализованное самоподдерживающееся состояние можно возбудить в однородных системах, слабо отклоненных от термодинамического равновесия [30, 54]. В этом смысле образование горячего АС в плазме или смеси нейтрального газа напоминает феномен шаровой молнии в атмосфере.

В смеси нейтральных газов АС [54] не расплываются за счет отрицательной термодиффузии, т. е. потока легких частиц газа из периферийных областей в центр АС — горячую область (см. разд. 8). При этом энергия, необходимая для поддержания в АС высокой температуры, может поступать не от электромагнитного поля, а в результате химических реакций. В атмосфере такой реакцией может быть медленное горение водорода, поступающего в центр АС за счет отрицательной термодиффузии.

## ЛИТЕРАТУРА

1. Николис Г., Пригожин И. Самоорганизация в неравновесных системах. М.: Мир, 1979. 512 с.
2. Хакен Г. Синергетика. М.: Мир, 1980. 404 с.
3. Эбелинг В. Образование структур при необратимых процессах. М.: Мир, 1979. 279 с.
4. Кернер Б. С., Осипов В. В. Нелинейная теория стационарных страт в диссипативных системах // ЖЭТФ. 1978. Т. 74, № 5. С. 1675—1697.
5. Кернер Б. С., Осипов В. В. Страты в разогретой электронно-дырочной плазме // Физика и техника полупроводников. 1979. Т. 13, № 4. С. 721.

6. Кернер Б. С., Осипов В. В. Стохастически неоднородные структуры в неравновесных системах // ЖЭТФ. 1980. Т. 79, № 6(12). С. 2218—2237.
7. Кернер Б. С., Осипов В. В. Стационарные и бегущие диссипативные структуры в активных кинетических средах // Микроэлектроника. 1981. Т. 10, № 5. С. 407—432.
8. Кернер Б. С., Осипов В. В. Явления в активных распределенных системах // Там же. 1985. Т. 14, № 5. С. 389—407.
9. Kerner B. S., Osipov V. V. Autosolitons in active systems with diffusion // Selforganization by nonlinear irreversible processes. Berlin etc.: Springer, 1986. P. 118—127.
10. Кернер Б. С., Осипов В. В. Автосолитоны // УФН. 1989. Т. 157, № 2.
11. Васильев В. А., Романовский Ю. М., Яхно В. Г. Автоволновые процессы. М.: Наука, 1987. 240 с.
12. Полак Л. С., Михайлов А. С. Самоорганизация в неравновесных физико-химических системах. М.: Наука, 1983. 282 с.
13. Романовский Ю. М., Степанова Н. В., Чернавский Д. С. Математическая биофизика. М.: Наука, 1984. 304 с.
14. Кернер Б. С., Осипов В. В. Расслоение тока в электронно-дырочной плазме // Физика и техника полупроводников. 1979. Т. 13, № 5. С. 891.
15. Кернер Б. С., Осипов В. В. Слои поля в полупроводниках с положительной дифференциальной проводимостью // ФТТ. 1979. Т. 21, № 3. С. 2342—2346.
16. Кернер Б. С., Осипов В. В. Расслоение электронно-дырочной плазмы при разогреве // ЖЭТФ. 1976. Т. 71, № 4. С. 1542—1554.
17. Гафийчук В. В., Кернер Б. С., Осипов В. В. Расслоение электронно-дырочной плазмы при ее разогреве длинноволновым излучением // ФТТ. 1981. Т. 23, № 8. С. 2305—2310.
18. Гафийчук В. В., Кернер Б. С., Осипов В. В. Условия расслоения тока или поля в электронно-дырочной плазме // Физика и техника полупроводников. 1981. Т. 15, № 11. С. 2171—2178.
19. Гафийчук В. В., Кернер Б. С., Осипов В. В. Образование сгустков электронно-дырочной плазмы, разогретой электрическим полем // Там же. 1985. Т. 19, № 10. С. 1871—1874.
20. Недоспасов А. В. Страты // УФН. 1968. Т. 94, № 3. С. 439—462.
21. Недоспасов А. В., Хаит В. Д. Колебания и неустойчивости низкотемпературной плазмы. М.: Наука, 1979. 168 с.
22. Голант В. Е., Жилинский А. П., Сахаров И. Е. Основы физики плазмы. М.: Атомиздат, 1977. 384 с.
23. Кернер Б. С., Осипов В. В. О нелинейной теории страт в газовом разряде // ДАН СССР. 1981. Т. 257, № 6. С. 1352—1355.
24. Кернер Б. С., Осипов В. В. Нелинейная теория страт в высокочастотном газовом разряде // РЭ. 1982. Т. 27, № 12, С. 2415—2425.
25. Кернер Б. С., Осипов В. В. Нелинейная теория страт в низкотемпературной плазме // Там же. 1983. Т. 28, № 1. С. 132—142.
26. Кочуэлл Э. Кинетические свойства полупроводников в сильных электрических полях. М.: Мир, 1970. 384 с.
27. Волков А. Ф., Коган Ш. М. Физические явления в полупроводниках с отрицательной дифференциальной проводимостью // УФН. 1968. Т. 96, № 4. С. 633—672.
28. Ross D. W. Oscillations of a weakly ionized plasma in a strong electric field // Phys. Rev. 1966. Vol. 146, N 1. P. 176—186.
29. Blötekar K., Weissglas P. Diffusion instabilities in semiconductors // J. Appl. Phys. 1968. Vol. 39, N 3. P. 1645—1653.
30. Кернер Б. С., Осипов В. В. Сильнонеравновесные состояния в системах, слабо отклоненных от термодинамического равновесия // Письма в ЖЭТФ. 1985. Т. 41, № 9. С. 386—388.
31. Кернер Б. С., Осипов В. В. Автосолитоны в горячей плазме полупроводников // ЖЭТФ. 1985. Т. 89, № 2(8). С. 589—607.
32. Гафийчук В. В., Кернер Б. С., Осипов В. В., Южанин А. Г. Самопроизвольное деление автосолитонов в устойчивой электронно-дырочной плазме при изменении уровня ее разогрева // Письма в ЖТФ. 1987. Т. 13, № 21. С. 1299—1303.



33. Кернер Б. С., Осипов В. В. Динамическая перестройка диссипативных структур // ДАН СССР. 1982. Т. 264, № 6. С. 1366—1370.
34. Гафийчук В. В., Кернер Б. С., Осипов В. В., Южанин А. Г. Свойства автосолиитоны в «плотной» электронно-дырочной плазме // Физика и техника полупроводников. 1988. Т. 22, № 11. С. 2051—2060.
35. Дубицкий А. Л., Кернер Б. С., Осипов В. В. Кинетика образования автосолиитона в устойчивой разогретой плазме собственных полупроводников // ФТТ. 1986. Т. 28, № 5. С. 1290—1295.
36. Дубицкий А. Л., Кернер Б. С., Осипов В. В. Спонтанное образование локальных областей ударной ионизации в однородных полупроводниках в слабых электрических полях // Физика и техника полупроводников. 1986. Т. 20, № 7. С. 1195—1198.
37. Дьяконов М. И., Перель В. И., Ясевич И. Н. Эффективный механизм энергетической релаксации горячих электронов в полупроводниках *p*-типа // Там же. 1977. Т. 11, № 7. С. 1364—1370.
38. Зеегер К. Физика полупроводников. М.: Мир, 1977. 615 с.
39. Гуревич А. В., Шварцбург А. Б. Нелинейная теория распространения радиоволны в ионосфере. М.: Наука, 1973. 272 с.
40. Брагинский С. И. Явления переноса в плазме // Вопросы теории плазмы. М.: Атомиздат, 1963. Т. 1. С. 183—272.
41. Кернер Б. С., Синкевич В. Ф. Многошнуровые и многодоменные стационарные состояния в горячей электронно-дырочной плазме // Письма в ЖЭТФ. 1982. Т. 36, № 10. С. 359—362.
42. Кернер Б. С., Осипов В. В., Романко М. Т., Синкевич В. Ф. Расслоение электронно-дырочной плазмы и голубая электролюминесценция в области статического домена GaAs // Там же. 1986. Т. 44, № 2. С. 77—79.
43. Гуревич А. В. О температуре электронов в плазме в переменном электрическом поле // ЖЭТФ. 1958. Т. 35, № 2. С. 392—400.
44. Дубицкий А. Л., Кернер Б. С., Осипов В. В. Математическое моделирование расслоения разогретой электронно-дырочной плазмы в полупроводниковых приборах // Тез. докл. I Всесоюз. конф. по физ. и физ.-хим. основам микроэлектроники. Вильнюс, 1987. С. 431—433.
45. Кернер Б. С., Кузнецова Е. М., Осипов В. В. О стохастически неоднородных состояниях активных сред с диффузией // ДАН СССР. 1984. Т. 277, № 5. С. 1114—1118.
46. Кернер Б. С., Кузнецова Е. М., Осипов В. В. О многообразии неоднородных состояний активных распределенных сред // Микроэлектроника. 1984. Т. 13, № 5. С. 407—426.
47. Кернер Б. С., Осипов В. В. Нелинейная теория неизотермического шнурования тока в транзисторных структурах // Там же. 1977. Т. 6, № 4. С. 337—351.
48. Кернер Б. С., Осипов В. В. Пульсирующие «гетерофазные» области в неравновесных системах // ЖЭТФ. 1982. Т. 83, № 6(12). С. 2201—2214.
49. Гафийчук В. В., Кернер Б. С., Лазурчак И. И., Осипов В. В. Пульсирующие автосолиитоны в активных распределенных средах // Микроэлектроника. 1986. Т. 15, № 2. С. 180—183.
50. Осипов В. В., Лазурчак И. И., Кернер Б. С., Гафийчук В. В. Пульсирующий автосолиитон и ведущий центр в двухкомпонентных системах с диффузией // Там же. 1987. Т. 16, № 1. С. 23—32.
51. Nitzan A., Ross J. Oscillations, multiple steady states and instabilities in illuminated systems // J. Chem. Phys. 1973. Vol. 59, N 4. P. 241—250.
52. Nitzan A., Ortoleva P., Ross J. Symmetry breaking instabilities in illuminated systems // Ibid. 1974. Vol. 60, N 8. P. 3134—3143.
53. Бункин Ф. В., Кириченко Н. А., Лукьянчук Б. С., Морозов Ю. Ю. Диффузионная неустойчивость в поле лазерного излучения // Квантовая электроника. 1983. Т. 10, № 10. С. 2136—2139.
54. Кернер Б. С., Осипов В. В. «Шаровая молния» в смеси нейтральных газов // ДАН СССР. 1987. Т. 292, № 1. С. 82—85.
55. Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. М.: Наука, 1979. 528 с.

## ТУРБУЛЕНТНО-ВОЛНОВОЕ ДИНАМО

Б. И. БАРЦ, С. С. МОИСЕЕВ

1. Большой интерес представляют вторичные неустойчивые процессы, приводящие к качественной перестройке системы [1]. Это могут быть либо возбуждение степеней свободы, физически отличающихся от исходных, либо спонтанное нарушение симметрии, смена экспоненциальной эволюции взрывной и т. д. (см., например, [2]). По особенностям развития к вторичным неустойчивостям примыкает турбулентно-волновое динамо (ТВД) — одно-временный рост амплитуды волны и турбулентной энергии за счет пространственно-временной модуляции последней этой волной. Каковы черты такого вторично-неустойчивого процесса в данном случае? Перечислим некоторые из них.

Пусть мы имеем течение в стратифицированной среде (океан, атмосфера), которое находится на пороге устойчивости (либо отрелаксировало к этому состоянию в результате развития первичной неустойчивости). Если длинноволновое возмущение распространяется по течению в слое, где локальное число Ричардсона близко к критическому значению  $Ri_{кр}$ , соответствующему порогу генерации турбулентности, то возможно появление значительной переменной составляющей плотности турбулентной энергии. При этом в турбулентных потоках, воздействующих на волну, определяющими становятся члены, пропорциональные градиентам невозмущенного течения и переменным (с частотой и пространственным периодом волны) коэффициентам турбулентного переноса [3, 4]. Существенно, что параметры стратифицированного течения и волны можно подобрать так, чтобы турбулентные потоки увеличивали энергию волны за счет энергии невозмущенного течения. С ростом же амплитуды волны растет и соответствующая переменная составляющая плотности турбулентной энергии [3]. В результате возникает турбулентно-волновая неустойчивость, причем ее инкремент есть возрастающая функция амплитуды волны [4].

Таким образом, порог линейной (первичной) неустойчивости представляет собой область параметров, для которых могут развиваться вторичные неустойчивые процессы, качественно меняющие эволюцию системы. Заметим, что в [3] была лишь дана верная постановка задачи о вторичной турбулентно-волновой неустойчивости на течении, а сами аргументы носили в основном предварительный и в ряде мест, к сожалению, неверный или гипотетический характер (неверно выражение для инкремента роста амплитуды длинной волны, не конкретизированы выражения для турбулентных потоков, нет критерия развития неустойчивости и др.). Проведенный в [4] и особенно в [5] подробный анализ показал, что отмеченное выше аномальное поведение стратифицированных течений на пороге устойчивости действительно возможно. Более того, из него фактически вытекает, что к классу усиливаемых тур-

булентностью волновых возмущений относятся также собственные волны, генерируемые за счет первичной неустойчивости.

Подчеркнем далее, что наряду с турбулентностью флуктуации параметров (особенно сильные на пороге устойчивости), промодулированные волной, также усиливают волновые возмущения. Такое «агрессивное» поведение шума качественно меняет ситуацию: им нельзя пренебрегать. При этом турбулентность возникает не на последних стадиях развития, а на значительно более ранних. Крупномасштабные моды «обрастают» нарастающей турбулентной «шубой». Это, в частности, существенно меняет структуру вторичных течений; должен также измениться подход к процессу стохастизации течений. В самом деле, учет припорогового шума ускоряет усиление неустойчивого решения, и при этом шум сам усиливается (если только по каким-либо конкретным причинам он не считается заданным). Можно тогда ожидать, что «острова устойчивости», где нет стохастизации решения, перекроются нарастающей шумовой «сеткой».

Как следует из проведенного ниже рассмотрения ТВД, а также нелинейного осциллятора Дюффинга в припороговой области, турбулентность и шум действительно ускоряют развитие неустойчивости, которая обычно принимает характер взрывного процесса. Ускоренный сценарий неустойчивого процесса связан с тем, что корреляционные характеристики шума модулируются нарастающей волной, а также с тем, что шум содержит гармонические составляющие всех частот, в том числе и резонансных с этой волной.

2. Рассмотрим, следуя [4—6], основную схему исследования турбулентно-волнового динамо для стратифицированных течений.

Внутренняя волна описывается уравнениями Рейнольдса для осредненных по турбулентным пульсациям волновых возмущений, дополненными (по сравнению с уравнениями Эйлера) производными турбулентных потоков импульса  $M_{ij} = \langle v_i^T v_j^T \rangle$  и плавучести  $H_j = -g\rho_a^{-1} \langle \rho^T v_j^T \rangle$ , где  $v_i^T$  и  $\rho^T$  — турбулентные пульсации скорости и плотности ( $i, j = x, y, z$ ) [7],  $\rho_a$  — регулярная часть плотности среды,  $\langle \dots \rangle$  — операция усреднения. Эти уравнения записаны в приближении Буссинеска и линеаризованы по малой амплитуде внутренней волны (пренебрежение «собственной» нелинейностью):

$$\begin{aligned} \left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \frac{\partial \Psi}{\partial z} - U' \frac{\partial \Psi}{\partial x} + \frac{1}{\rho_a} \frac{\partial p}{\partial x} &= - \frac{\partial M_{xx}}{\partial x} - \frac{\partial M_{xz}}{\partial z}, \\ - \left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \frac{\partial \Psi}{\partial x} - q + \frac{1}{\rho_a} \frac{\partial p}{\partial z} &= - \frac{\partial M_{zz}}{\partial z} - \frac{\partial M_{zx}}{\partial x}, \\ \left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) q - N^2 \frac{\partial \Psi}{\partial x} &= - \frac{\partial H_x}{\partial x} - \frac{\partial H_z}{\partial z}. \end{aligned} \quad (1)$$

Здесь исходное течение описывается невозмущенной горизонтальной скоростью  $U(z)$  и плотностью  $\rho_0(z)$ ;  $N^2 = -g\rho_a^{-1}\rho'_0(z)$ , где  $\rho_a = \rho_0(0)$ , штрихом обозначена производная по вертикальной координате  $z$ . Волновые движения описываются функцией тока  $\Psi$

и вариацией плавучести  $q = -g\rho_a^{-1}[\rho - \rho_0(z)]$ . Таким образом, полное движение раскладывается на три слагаемых в соответствии с иерархией масштабов: исходное стратифицированное течение, волновые движения, турбулентные пульсации. Граничные условия для (1):  $\Psi = 0$  при  $z = -H$  и  $z = 0$  (на дне и на поверхности).

Функцию тока, описывающую внутреннюю волну, ищем в виде

$$\Psi = [A(t)(\omega k^{-1} - U)F(z) + \Psi_1(t, z)] \exp(i\omega t - ikx)t + \text{к. с.},$$

где  $F(z)$  — распределение вертикальных смещений в невозмущенной волновой моде (решение линейной однородной краевой задачи). Полагая, что возмущающее влияние турбулентных потоков на волну мало меняет ее комплексную амплитуду  $A \equiv |A| \exp(ip)$  за один период, применяем стандартную асимптотическую процедуру метода «медленных» возмущений (двухмасштабное разложение).

Турбулентные потоки импульса и плавучести представляются в виде произведения градиентов соответствующих величин и коэффициентов турбулентного переноса, которые, в свою очередь, согласно гипотезе Колмогорова [7], полагаются пропорциональными корню квадратному из плотности энергий турбулентных пульсаций  $b = 0,5 \sum_{j=x, y, z} \langle v_j^T v_j^T \rangle$  и макромасштабам турбулентности  $L_j$ . Последние принимаются в наших расчетах постоянными, не зависящими от  $b$ .

Предполагается, что в отсутствие волны локальное число Ричардсона  $Ri_0(z) = N^2(U')^{-2}$  на всех глубинах больше критического значения  $Ri_{кр}$ , соответствующего порогу генерации турбулентности, и течение не турбулизовано:  $b(z) \equiv 0$ . Наведенная волной турбулентность возникает в слое ( $h < z < -h + \Delta h$ ), где  $Ri_0$  близко к  $Ri_{кр}$ :

$$(Ri_0 Ri_{кр}^{-1} - 1) \approx A \ll 1, \quad (2)$$

и возмущенное волной число Ричардсона становится меньше  $Ri_{кр}$ . (Если подходящим образом нормировать решение однородной краевой задачи, например положить  $\max |F'| = 1$ , то безразмерная амплитуда  $|A|$  будет выражать порядок малости волнового возмущения.)

Уравнение для плотности энергии турбулентности  $b$  замыкается в соответствии с той же простейшей колмогоровской схемой [7]. Причем для турбулентности, наведенной волной, величина  $\sqrt{b}$  имеет порядок малости амплитуды волны, точнее,

$$\sqrt{b} \approx L_z N |A|. \quad (3)$$

Это обстоятельство позволяет упростить для данного случая уравнение баланса турбулентной энергии. Учитывая его, а также

Вертикальный масштаб  $L_z$  порядка  $\Delta h$  — толщины слоя, в котором исходное течение близко к порогу генерации турбулентности, а  $L_x \sim \pi k^{-1}$ .

неотрицательность  $|A|$ , окончательно получим уравнение для  $\sqrt{\bar{b}}$ :

$$d\sqrt{\bar{b}}/dt = L_z (U')^2 [1 - \kappa_\rho \text{Ri}_0]/2 + |A|S \cos(\Omega t - kx + \rho) \quad (4)$$

при  $t_1 + 2\pi n/\Omega \leq t \leq t_2 + 2\pi n/\Omega$ ,

$$\sqrt{\bar{b}} = 0 \quad \text{при} \quad t_2 + 2\pi n/\Omega \leq t \leq t_1 + 2\pi(n+1)/\Omega,$$

где  $n$  — целое число;  $\Omega(z) = \omega - kU$ ;

$$S(z) = [4\Omega(U')^{-1} - 2\text{Ri}_0 U' \Omega^{-1}]kF + \kappa_\rho \text{Ri}_0 F' + \\ + (U')^{-2} F \frac{d}{dz} [(U')^2 (\kappa_\rho \text{Ri}_0 - 1)], \quad \kappa_\rho^{-1} = \text{Ri}_{\text{кр}}.$$

Используя (4), находим фурье-компоненту  $\sqrt{\bar{b}}$ , и в результате для модуля амплитуды волны  $A$  получим переменную скорость роста, начиная от некоторого порогового значения  $A_n$ . Так, время удвоения амплитуды волны, близкой к пороговой ( $0 < a - 1 \ll \ll 1$ ,  $a = A/A_n$ ), есть убывающая функция амплитуды, как в случае взрывного роста. Далее, увеличение амплитуды еще в 10 раз происходит за время порядка  $30\gamma^{-1}$ , где  $\gamma \approx \Omega \Delta h^3 k^3$ . Анализ показывает, что условие на структуру моды для развития неустойчивости является весьма умеренным; по существу, требуется лишь, чтобы характерные длины изменения волнового возмущения вдоль направления стратификации и перпендикулярно ему были сравнимы. Отсюда, в частности, вытекает (как уже подчеркивалось во введении), что к классу усиливаемых возмущений относятся также собственные волны, генерируемые за счет первичной неустойчивости. Этот вопрос численно анализируется также в [8].

Заметим далее, что в [9] приведены результаты численного решения уравнений, моделирующих турбулентно-волновую неустойчивость (ТВН) в атмосфере, но численный подход не позволил очертить условия проявления ТВН.

3. Рассмотрим теперь влияние шума на развитие неустойчивости в простейшей нелинейной системе — осцилляторе Дюффинга. Уравнения Ланжевена

$$\dot{x} = y, \\ \dot{y} = F(x) - \alpha y + f(t), \quad (5)$$

в которые наряду с детерминированной силой  $F = -\omega_0^2 x + \beta x^3$  ( $\beta > 0$ ) и силой трения ( $\alpha > 0$ ) входит также и случайная сила  $f(t)$  при обычных предположениях о свойствах шума (гауссов, дельта-коррелированный:  $\langle f(t)f(t+\tau) \rangle = 2D\delta(\tau)$ ,  $D > 0$ ) эквивалентны уравнению Эйнштейна—Фоккера для распределения вероятности  $p_t(x, y)$  по координатам  $x$  и скоростям  $y$  (см., например, [10]).

Поскольку для нас важной является область вблизи критической точки  $x_c = \omega_0/\sqrt{\beta}$ , представим силу  $F(x)$  как

$$F(x) = 2\omega_0^2(x - x_c) + \delta F(x) \quad (6)$$

и перейдем к переменным  $z_i (i = 1, 2)$ , являющимся интегралами детерминированного движения, обусловленного линейной частью силы из (6) (см., например, [11]):

$$z_i = [(x - x_c)\mu_i - y]\varphi_i(t) - [(x_0 - x_c)\mu_i - y_0]\varphi_i(0), \quad i = 1, 2, \quad (7)$$

где  $\varphi_i(t) = \exp[(\alpha + \mu_i)t]$ ,  $\mu_{1,2} = (-\alpha \pm \gamma)/2$ ,  $\gamma = \sqrt{\alpha^2 + 8\omega_0^2}$ ,  $x_0, y_0$  — начальные значения.

Перейдем к фурье-образу  $p_i(k)$  функции распределения  $P_t(z)$ :

$$P_t(z) = \frac{1}{(2\pi)^2} \int e^{ikz} p_i(k) dk, \quad (kz) \equiv \sum_i k_i z_i, \quad i = 1, 2. \quad (8)$$

Для функции  $p(k)$  нетрудно получить следующее интегральное уравнение:

$$p_i(k) = p_i^{(0)}(k) \left[ 1 - i \int_0^t \varphi_i(t') (p_i^{(0)}(k))^{-1} \delta \hat{F} k_i p_i^{(0)}(k) dt' \right], \quad (9)$$

где  $\delta \hat{F} = \delta F(x(z_1))|_{z_1=i\partial/\partial k}$ , а  $p_i^{(0)}(k) = \exp[-1/2 b_{ik}(t) k_i k_k]$ ,  $b_{ik}(t) = 2D \int_0^t \varphi_i(t) \varphi_k(t') dt'$ .

В соответствии с (7) имеем для  $x$

$$x(z) = x_t + \frac{1}{\gamma} (z_1 e^{\mu_1 t} - z_2 e^{\mu_2 t}), \quad (10)$$

где  $x_t$  определяется детерминированной «линейной» силой  $F_0$ . Для среднего  $\langle x \rangle$ , с другой стороны, получим

$$\langle x \rangle = x_t + \hat{\delta}_i \chi_t(k)|_{k=0}. \quad (11)$$

Здесь  $\hat{\delta}_i = \frac{i}{\gamma} \left( e^{\mu_1 t} \frac{\partial}{\partial k_1} - e^{\mu_2 t} \frac{\partial}{\partial k_2} \right)$ , а  $\chi_t$  — выражение в квадратных скобках из (9). При отыскании второго слагаемого в правой части (11) будем предполагать, что функция  $\chi_t$  возрастает во времени быстрее, чем  $\varphi_i(t)$  (в дальнейшем это оправдывается).

После этого предположения в результате несложных вычислений нетрудно показать, что

$$\chi_t = \sum_{n=0}^{\infty} \xi^{2n+1} u^n \left[ \frac{4\beta}{(\gamma - \alpha)^2} \right]^n. \quad (12)$$

Здесь  $\xi = \exp\left(\frac{\gamma - \alpha}{2} t\right)$ , а коэффициенты  $u^n$  можно представить в виде суммы неотрицательных слагаемых. Заметим также, что по смыслу сделанного предположения о поведении  $\chi_t$  результат (12) справедлив при достаточно больших временах  $t > 1/\gamma$ , когда  $\xi \gg 1$ , при этом коэффициенты  $u^n$  от  $\xi$  не зависят. Используя неотрицательность слагаемых, из которых состоит  $u^n$ , можно

показать, что

$$\chi_t \geq \sum_p \xi^{6p+1} \left( \frac{4\beta}{(\gamma - \alpha)^2} \right)^{3p} \frac{3v}{2} \left( \frac{\bar{D}^3}{9 \cdot 6^4} \right)^p \frac{(4p)! (2p - 1)!}{(p!)^2 p^{4p}}. \quad (13)$$

Здесь

$$\bar{D} = \frac{2D}{\gamma^2(\gamma - \alpha)}; \quad v = \frac{1}{\gamma} y_0 + (x_0 - x_c) \frac{\gamma + \alpha}{2\gamma}.$$

Подпоследовательность в правой части (13) имеет конечный радиус сходимости

$$\xi_0 = {}^{3/4}e^{2/3} (\gamma - \alpha) / \sqrt{\bar{D}\beta}. \quad (14)$$

Таким образом, по крайней мере к моменту времени  $t_0 \approx \frac{1}{\gamma - \alpha} \ln \frac{\gamma^2(\gamma - \alpha)^3}{D\beta}$ , ряд для  $\chi_t$  расходится, т. е. величина  $\langle x \rangle$  достигает бесконечно больших значений. Можно показать, что при достаточно малых надкритичностях время  $t_0$  существенно меньше времени «выхода на нелинейность» регулярного решения без учета шума. Таким образом, наличие шума значительно ускоряет развитие неустойчивости.

#### ЛИТЕРАТУРА

1. Коган Е. Я., Моисеев С. С., Ораевский В. Н. Гидродинамические модели в применении к исследованию устойчивости замагниченной плазмы // ЖПМТФ. 1965. № 6. С. 41—46.
2. Моисеев С. С., Сагдеев Р. З. Проблемы вторичных неустойчивостей в гидродинамике и плазме // Изв. вузов. Радиофизика. 1986. Т. 29, № 9. С. 1067—1072.
3. Chimonas G. The stability of a coupled wave-turbulence system in a parallel shear flow // Boundary—Layer Meteorol. 1972. Vol. 2. P. 444—452.
4. Моисеев С. С., Пунгин В. Г., Сагдеев Р. З. и др. Некоторые окологороговые и турбулентно-волновые взрывные процессы в гидродинамике // II Всесоюз. съезд океанологов: Тез. докл. Севастополь: МГИ АН УССР, 1982. Вып. 2: Физика и химия океана. С. 15.
5. Моисеев С. С., Суязов Н. В., Эткин В. С. Турбулентно-волновая неустойчивость сдвиговых стратифицированных течений: Препр. ИКИ АН СССР № 905. М., 1984. 18 с.
6. Moiseev S. S., Pungin V. G., Suyazov N. V. Finite-amplitude disturbances and near-threshold instability of hydrodynamic flows // Laminar-turbulent. transition: Symposium, Novosibirsk (USSR). Berlin etc.: Springer, 1984. P. 735—747.
7. Мониш А. С., Яглом А. М. Статистическая гидромеханика. М.: Физматгиз, 1965. Ч. 1. 639 с.
8. Einandi F., Finnigan J., Fud D. Gravity wave turbulence interaction in the presence of a critical level // J. Atmos. Sci. 1984. Vol. 41. P. 661—667.
9. Fud D., Chimonas G., Einandi F., Zeman O. An analysis of wave-turbulence interaction // Ibid. 1982. Vol. 39. P. 2450—2463.
10. Кляцкин В. И. Статистическое описание динамических систем с флуктуирующими параметрами. М.: Наука, 1975. 239 с.
11. Чандрасекар С. Стохастические проблемы в физике и астрономии. М.: Изд-во иностр. лит., 1974. 168 с.

## КРИТЕРИЙ СУЩЕСТВОВАНИЯ БЕГУЩИХ СТРУКТУР В ДВУХКОМПОНЕНТНЫХ РЕАКЦИОННО-ДИФФУЗИОННЫХ СИСТЕМАХ

Ю. Д. КАЛАФАТИ, Ю. А. РЖАНОВ

В настоящее время широко известны примеры неравновесных физических, химических и биологических систем, в которых наблюдаются автоволновые процессы [1—4]. Наиболее яркий пример такой системы — химическая реакция Белоусова—Жаботинского.

Многие характерные особенности обнаруженных автоволновых процессов удается описать с помощью уравнений

$$\tau_P \frac{\partial P}{\partial t} - l_P^2 \frac{\partial^2 P}{\partial z^2} = F_P(P, S), \quad (1)$$

$$\tau_S \frac{\partial S}{\partial t} - l_S^2 \frac{\partial^2 S}{\partial z^2} = F_S(P, S), \quad (2)$$

где  $\tau_P, \tau_S$  — характерные временные, а  $l_P = \sqrt{D_P \tau_P}$ ,  $l_S = \sqrt{D_S \tau_S}$  — пространственные масштабы изменения кинетических переменных, например концентраций химических реагентов  $P$  и  $S$ ;  $D_P$  и  $D_S$  — коэффициенты диффузии;  $F_P(P, S)$ ,  $F_S(P, S)$  — нелинейные функции концентраций. В дальнейшем нас будет интересовать случай, когда нуль-изоклины уравнений (1) и (2)

$$F_P(P, S) = 0, \quad (3)$$

$$F_S(P, S) = 0 \quad (4)$$

имеют вид, показанный на рис. 1 (кривая 1 соответствует уравнению (3) и демонстрирует неоднозначную зависимость  $P$  от  $S$ , кривая 2 соответствует уравнению (4)). Переменную  $P$  принято называть автокаталитической. Кривые 1 и 2 на рис. 1, а пересекаются в трех точках, что соответствует существованию трех корней системы уравнений (3), (4). Это означает, что в системе имеются три стационарных однородных состояния (триггерный режим). Анализ устойчивости этих состояний показывает, что два из них,  $(P_1, S_1)$  и  $(P_3, S_3)$ , устойчивы, а одно  $(P_2, S_2)$  неустойчиво. Кривые на рис. 1, б пересекаются только в одной точке, которая соответствует устойчивому стационарному однородному состоянию системы (ждущий режим).

В ряде работ [4, 5] было показано, что если  $\tau_P/\tau_S \ll 1$  и  $l_S = 0$ , то уравнения (1), (2) содержат решения, описывающие распространение бегущего импульса, который возникает в результате достаточно сильного локального возмущения системы. С другой стороны, известно [1—4, 6], что если  $l_P < l_S$ , то в системе возможно образование неподвижных пространственно-неоднородных структур. Поэтому возникает вопрос: при каком соотношении па-



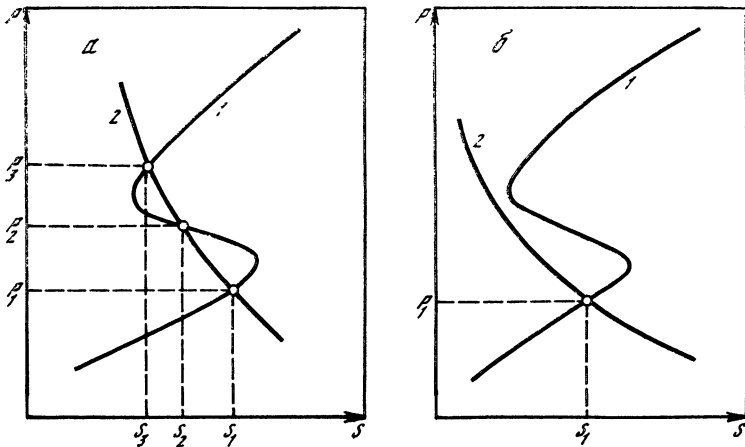


Рис. 1

раметров системы  $\tau_P$ ,  $\tau_S$ ,  $l_P$ ,  $l_S$  существуют бегущие структуры, а при каких неподвижные?

В работе [7] было высказано предположение, что область параметров, при которых существуют бегущие структуры, определяется всего одной универсальной комбинацией  $v_P/v_S$ , где  $v_P = l_P/\tau_P$ , а  $v_S = l_S/\tau_S$ . В ряде последующих сообщений [8, 9] были приведены результаты исследования бегущих структур для случая кусочно-линейных функций  $F_P$  и  $F_S$ . Оказалось, что в этой модели область параметров, при которых существуют бегущие структуры, определяется условием

$$v_P/v_S > 1, \quad (5)$$

что подтверждает выдвинутую в [7] гипотезу.

Цель данной работы — подробный вывод условия (5), исследование устойчивости полученных решений и сравнение теории с результатами численного моделирования.

## 1. ОБСУЖДЕНИЕ МОДЕЛИ

Для исследования стационарной формы волны перейдем в уравнениях (1), (2) к автомодельной переменной  $f = z/l_S - vt/\tau_S$ , тогда

$$-\varepsilon_1 v dP/df - \varepsilon_2^2 d^2P/df^2 = F_P(P, S), \quad (6)$$

$$-v dS/df - d^2S/df^2 = F_S^*(P, S), \quad (7)$$

$$dP/df \rightarrow 0, \quad dS/df \rightarrow 0 \text{ при } f \rightarrow \pm \infty.$$

В дальнейшем мы будем исследовать уравнения (6), (7) в предположении

$$\tau_P/\tau_S = \varepsilon_1 \ll 1, \quad l_P/l_S = \varepsilon_2 \ll 1. \quad (8)$$

Как видно из результатов [4—9], бегущие решения возникают как раз в области параметров, удовлетворяющих условию (8).

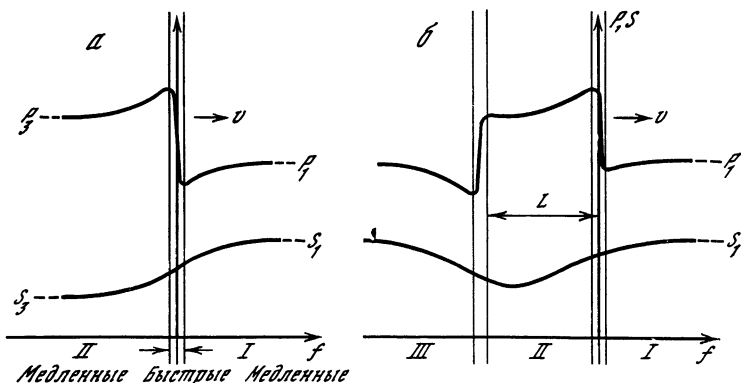


Рис. 2

Наличие малых параметров приводит к тому, что профиль нелинейной волны имеет характерные области: области «быстрого» изменения переменной  $P$  и области «медленного» изменения. Используя методы сингулярной теории возмущений [10—12], получим в нулевом порядке по  $\varepsilon_1$  и  $\varepsilon_2$  приближенные уравнения соответственно для областей «быстрых» и «медленных» изменений. При этом уравнения (6), (7) в области «быстрых» изменений принимают вид

$$-\varepsilon_1 v dP^*/df - \varepsilon_2^2 d^2P^*/df^2 = F_P(P^*, S^*), \quad (9)$$

$$S^* = \text{const}, \quad (10)$$

где звездочкой обозначаются переменные в области «быстрых» изменений. Уравнение (9) удобно записать в символическом виде:  $v = (\varepsilon_2/\varepsilon_1)\varphi(S^*)$ , где  $\varphi(S^*)$  — некоторая функция, которая находится в явном виде, если  $F_P(P, S)$  — кусочно-линейная функция или кубический полином.

В области «медленных» изменений уравнения (6), (7) примут вид

$$F_P(P, S) = 0, \quad (11)$$

$$-vdS/df - d^2S/df^2 = F_S(P, S). \quad (12)$$

Решение уравнений (9), (10) из области «быстрых» изменений необходимо затем «сшить» с решением уравнений (11), (12) из области «медленных» изменений. В задаче о распространении фронта, например, имеются две области «медленных» изменений ( $I$  и  $II$  на рис. 2, а). Вместо того чтобы проводить сшивку решений из областей «быстрых» и «медленных» изменений для «быстрой» переменной  $P$ , мы сформулируем эквивалентную процедуру, но только для «медленной» переменной. В нулевом порядке теории возмущений решения из области  $I$  и из области  $II$  сшиваются с помощью следующих условий:

$$S_I|_{f \rightarrow 0^+} = S_{II}|_{f \rightarrow 0^-}, \quad (13)$$

$$dS_I/df|_{f \rightarrow 0^+} = dS_{II}/df|_{f \rightarrow 0^-}, \quad (14)$$

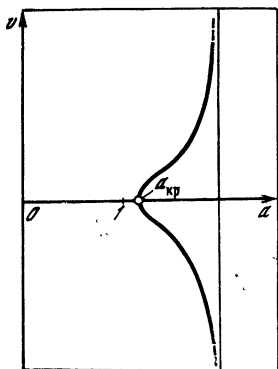


Рис. 3

где  $f \rightarrow 0^-$  означает стремление переменной к нулю слева, а  $f \rightarrow 0^+$  —справа от области «быстрых» изменений.

Уравнения (9)–(12) допускают аналитическое решение, если нелинейные источники  $F_P$  и  $F_S$  являются кусочно-линейными функциями. В качестве этих источников рассмотрим следующие:

$$F_P(P, S) = W\Theta(P - P_0) - P + S, \quad (15)$$

$$F_S(P, S) = [1 - \Theta(P - P_0)]J + (P - S) - (S - S_0), \quad (16)$$

где  $\Theta(P)$  — функция единичного скачка;  $P_0, S_0, W, J$  — параметры. Нуль-изокли-

ны (3), (4), например, имеют три пересечения при

$$0 < a \leq 2, \quad a \leq b \leq 2, \quad (17)$$

где  $a = J/W$ ,  $b = (P_0 - S_0)/W$ . Физический смысл источников и кинетических переменных обсуждался в [7–9].

## 2. РЕШЕНИЕ

Решая уравнения в областях «быстрых» и «медленных» изменений и сшивая решения, получим для скорости фронта

$$v[(v^2 + 4v_P^2)^{-1/2} + (1 - a)(v^2 + 4v_S^2)^{-1/2}] = a + 2(1 - b). \quad (18)$$

Уравнения для одиночного импульса получаются аналогично и устанавливают связь между скоростью распространения импульса и его характерным размером  $L$  (см. рис. 2, б):

$$v[(v^2 + 4v_P^2)^{-1/2} + (1 - a)(v^2 + 4v_S^2)^{-1/2}] = a + 2(1 - b) + (1 - a)[1 - v(v^2 + v_S^2)^{-1/2}] \exp(-\gamma L/l_S), \quad (19)$$

где  $\gamma = + (1/v_S)(v + \sqrt{v^2 + v_S^2})$ , а второе уравнение получается из (19) заменой  $v$  на  $(-v)$ .

Наиболее просто решается уравнение для скорости фронта. Если  $v_P < v_S$ , то уравнение (18) имеет лишь одно решение. Если же

$$v_P/v_S > 1, \quad a > a_{кр} = 1 + v_S/v_P, \quad (20)$$

то в системе возможно существование трех фронтов, движущихся с разной скоростью, причем только два из них устойчивы. Рассмотрим случай, когда уравнение (18) содержит решение, соответствующее неподвижному фронту ( $v = 0$ ). При этом параметры  $a$  и  $b$  должны удовлетворять условию (правило площадей Максвелла)

$$2 - 2b + a = 0. \quad (21)$$

Тогда при  $v_P < v_S$  есть только одно решение  $v = 0$ , а при условиях (20) — три решения (рис. 3):

$$v = 0, \quad v = \pm 2 \sqrt{\frac{v_P^2 (a-1)^2 - v_S^2}{1 - (a-1)^2}}. \quad (22)$$

Кусочно-линейная аппроксимация позволяет проанализировать устойчивость стационарного решения. Оказалось (см. приложение), что инкремент нарастания малых возмущений  $\lambda$  для решения с  $v = 0$  меняет знак при увеличении параметра  $a$ :

$$\lambda = -1 + (a-1)v_P/v_S, \quad (23)$$

и стационарное решение становится неустойчивым. Видно, что условие потери устойчивости стационарного решения совпадает с критерием возникновения бегущего решения (20).

Интересно отметить, что два устойчивых бегущих фронта существуют, только если  $a > 1$ . Последнее означает, что величина  $S$  должна уменьшаться при увеличении  $P$ , т. е. в системе существует отрицательная обратная связь.

Из уравнений для одиночного импульса можно сделать вывод, что стационарное решение

$$v = 0, \quad L = -l_S \ln \left( \frac{a+2-2b}{a-1} \right) \quad (24)$$

существует для всех параметров, где  $a > 2b - 2$ , в области триггерного и ждущего режимов. Полный анализ условий существования бегущих импульсов на основании уравнений (19) затруднителен, однако ясно, что для широких импульсов ( $L/l_S \gg 1$ ) эти уравнения сводятся к уравнениям для скорости фронта. Из самого метода построения решений также ясно, что одиночный импульс можно построить, если уравнение для скорости фронта имеет три решения. Таким образом, мы приходим к выводу, что условия существования бегущих структур определяются выражением (20), при котором неподвижные решения теряют устойчивость.

### 3 ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

Существование движущегося фронта переключения в зависимости от начальных условий исследовалось нами на модели (1), (2), (15), (16) при значениях параметров, соответствующих условию Максвелла (21):

$$W = 6, \quad J = 8, \quad P_0 = 12, \quad S_0 = 2, \quad \varepsilon_1 = 10^{-3}, \quad \varepsilon_2 = 10^{-1},$$

так что  $v_P/v_S = \varepsilon_2/\varepsilon_1 \gg 1$ . При этом реализуется триггерный режим с устойчивыми точками ( $P_1 = 10$ ;  $S_1 = 10$ ) и ( $P_3 = 14$ ;  $S_3 = 8$ ). Для определенности начальные условия выбирались в виде

$$P(z) = P_1 + (P_3 - P_1)\Theta(z_0), \quad (25)$$

$$S(z) = S_1 + (S_3 - S_1)\Theta(z_0 + \Delta z). \quad (26)$$

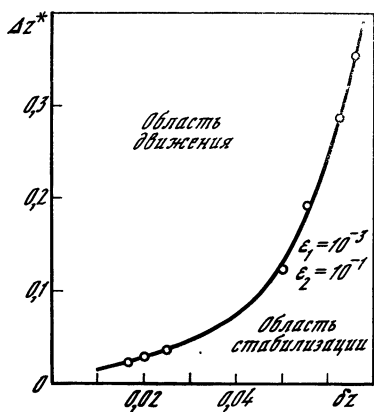


Рис. 4

Вследствие дискретности численной задачи величина  $\Delta z$  кратна расстоянию между соседними узлами равномерной разностной сетки  $\delta z$ . Мы обнаружили, что при каждой величине  $\delta z$  (степени дискретизации) существует некоторое пороговое значение  $\Delta z^*$ . При  $\Delta z \geq \Delta z^*$  начинается устойчивое распространение триггерной волны, при  $\Delta z < \Delta z^*$  устанавливается неподвижная структура. Зависимость  $\Delta z^*$  от  $\delta z$  (рис. 4) монотонна, и  $\Delta z^*$  превышает размер моделируемой системы при  $\delta z \geq \geq 0,07$ . При  $\delta z \geq 0,07$  любые начальные условия, относящиеся к

классу (25), (26), приводили к неподвижному фронту.

С уменьшением шага разностной сетки величина  $\Delta z^*$  достаточно быстро убывает, стремясь к нулю. Это свидетельствует, с одной стороны, в пользу выводов теории о неустойчивости статических решений в континуальном пределе и, с другой стороны, говорит, что причиной стабилизации решений является дискретность системы. Важно отметить, что дискретность может быть связана как с приемами численного моделирования, так и с природой физического объекта.

Оценить влияние дискретизации на поведение решения можно по сопоставлению теоретических и численных результатов. В рассматриваемой задаче степень дискретизации определяется параметром  $\xi = \varepsilon_2/\delta z$ , который можно интерпретировать как число точек разностной сетки, уместающихся в области «быстрого» изменения переменных. Обнаружено, что при  $\xi \geq 5$  наблюдается хорошее согласие теоретических и численных значений скорости распространения фронта. Из этого следует, что аналитические результаты, полученные в нулевом порядке теории возмущений, не только качественно предсказывают поведение решения, но и дают их количественную оценку. При  $\xi < 5$  численно определенные значения скорости значительно отличаются от аналитических, особенно в области параметров, где теория предсказывает существование фронтов,двигающихся с малой скоростью.

**В ы в о д ы.** На примере модели с кусочно-линейной аппроксимацией нелинейных источников мы показали, что бифуркационным параметром, определяющим возникновение бегущих структур в двухкомпонентных системах, действительно является отношение  $v_p/v_s$ . Для рассмотренной модели критическое значение параметра равно единице.

При численном моделировании было обнаружено, что в обла-

сти параметров, где, согласно теории, устойчивы только бегущие структуры, в зависимости от начальных условий существуют как бегущие структуры, так и устойчивые неподвижные. Изменяя шаг сетки и начальные условия, удалось показать, что неподвижные структуры становятся устойчивыми из-за дискретизации при численном моделировании.

## П Р И Л О Ж Е Н И Е.

### ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ НЕПОДВИЖНОГО ФРОНТА

Для исследования устойчивости неподвижного фронта перейдем в уравнениях (1), (2) к переменным  $\tau = t/\tau_S$ ,  $f = z/l_S - vt/\tau_S$  и будем искать решения полученных уравнений в виде

$$P(\tau, f) = P^{st}(f) + e^{\lambda\tau}\delta P(f), \quad (\text{П.1})$$

$$S(\tau, f) = S^{st}(f) + e^{\lambda\tau}\delta S(f), \quad (\text{П.2})$$

где  $P^{st}$ ,  $S^{st}$  — точные решения, соответствующие фронту;  $\lambda$  — инкремент нарастания малых возмущений. Линеаризованные относительно малых отклонений  $\delta P$ ,  $\delta S$  уравнения примут вид

$$\varepsilon_1\lambda\delta P - v\varepsilon_1\frac{d\delta P}{df} - \varepsilon_2^2\frac{d^2\delta P}{df^2} = \left(\frac{\delta F_P}{\delta P}\right)_{\substack{P=P^{st} \\ S=S^{st}}}\delta P + \left(\frac{\delta F_P}{\delta S}\right)_{\substack{P=P^{st} \\ S=S^{st}}}\delta S, \quad (\text{П.3})$$

$$\lambda\delta S - v\frac{d\delta S}{df} - \frac{d^2\delta S}{df^2} = \left(\frac{\delta F_S}{\delta P}\right)_{\substack{P=P^{st} \\ S=S^{st}}}\delta P + \left(\frac{\delta F_S}{\delta S}\right)_{\substack{P=P^{st} \\ S=S^{st}}}\delta S. \quad (\text{П.4})$$

Уравнения (П.3), (П.4), дополненные граничными условиями  $\delta P \rightarrow 0$  и  $\delta S \rightarrow 0$  при  $f \rightarrow \pm\infty$ , являются уравнениями на собственные значения с потенциалом, зависящим от вида решения, соответствующего фронту. Для кусочно-линейной зависимости источников (15), (16) эти потенциалы являются просто дельта-функциями, и уравнения (П.3), (П.4) решаются аналитически. Окончательно уравнение для  $\lambda$  для неподвижного фронта имеет вид

$$\varepsilon_2\alpha_-\alpha_+(\alpha_- + \alpha_+) - [1 - \varepsilon_2(1 - a)](\lambda + 2 + \alpha_-\alpha_+) + a = 0, \quad (\text{П.5})$$

где

$$\sigma_{\pm} = \left\{ \frac{1}{2} \left( \lambda + 2 + \frac{\varepsilon_1\lambda + 1}{\varepsilon_2^2} \right) \pm \left[ \frac{1}{4} \left( \lambda + 2 + \frac{\varepsilon_1\lambda + 1}{\varepsilon_2^2} \right)^2 - \frac{(\varepsilon_1\lambda + 1)(\lambda + 2) - 1}{\varepsilon_2^2} \right]^{1/2} \right\}^{1/2}.$$

Можно убедиться, что с точностью до малых параметров уравнение (П.5) имеет корень

$$\lambda = -1 + (a - 1)v_P/v_S. \quad (\text{П.6})$$

Из (П.6) следует, что неподвижный фронт может потерять устойчивость, если (поскольку  $a \leq 2$ )

$$v_P/v_S > 1, \quad a > 1 + v_S/v_P,$$

что совпадает с условием возникновения б.гущего фронта.

## ЛИТЕРАТУРА

1. *Николис Г., Пригожин И.* Самоорганизация в неравновесных системах. М.: Мир, 1979. 421 с.
2. *Хакен Г.* Синергетика. М.: Мир, 1980. 406 с.
3. *Жаботинский А. М.* Концентрационные автоколебания. М.: Наука, 1974. 311 с.
4. *Васильев В. А., Романовский Ю. М., Яхно В. Г.* Автоволновые процессы в распределенных кинетических системах // УФН. 1979. Т. 128, № 4. С. 625—666.
5. *Островский Л. А., Яхно В. Г.* Формирование импульсов в возбудимой среде // Биофизика. 1975. Т. 20, № 3. С. 489—493.
6. *Белинцев В. Н.* Диссипативные структуры и проблемы биологического структурообразования // УФН. 1983. Т. 141, № 1. С. 55—101.
7. *Калафати Ю. Д., Сербинов И. А., Рябова Л. А.* Об иерархии малых параметров в теории диссипативных структур // ДАН СССР. 1982. Т. 263, № 4. С. 862—864.
8. *Калафати Ю. Д.* Гистерезис скорости распространения фронта переключения в полупроводниках с концентрационной неустойчивостью // Тез. докл. XII совещ. по теории полупроводников. Киев, 1985. Ч. 1. С. 309.
9. *Калафати Ю. Д., Ржанов Ю. А.* Неустойчивость распространения фронта переключения в бистабильных полупроводниках // Тез. докл. Всесоюз. совещ. по самоорганизации в физ., хим. и биол. системах «Синергетика-86». Кишинев: Штиинца, 1986. С. 52.
10. *Заикин А. Н.* Неустойчивость и распространение возбуждения в модели каталитической реакции // Биофизика, 1975. Т. 20. С. 772—777.
11. *Koga S., Kuramoto Y.* Localized patterns in reaction-diffusion systems // Progr. Theor. Phys. 1980. Vol. 63, P. 106—109.
12. *Ortoleva P., Ross J.* Theory of propagation of discontinuities in kinetic systems with multiple time scales: Fronts, front multiplicity, and pulses // J. Chem. Phys. 1975. Vol. 63. P. 3398—3407.

УДК (629.7+519.7) : 57

## ПРОСТРАНСТВЕННО-ВРЕМЕННЫЕ СТРУКТУРЫ В ОПТОЭЛЕКТРОННЫХ УСТРОЙСТВАХ

А. В. МАСТЕРОВ, В. Н. ТОЛКОВ, В. Г. ЯХНО

В настоящее время однородные системы из оптоэлектронных элементов предполагается использовать в ряде устройств, связанных с управлением, параллельной обработкой больших массивов входных сигналов, в том числе и автоматического распознавания изображений [1, 2]. Реализация таких устройств сталкивается с трудностями создания элементной базы для устройств из многих тысяч однородных модулей, организацией удобной и быстрой связи между различными элементами — модулями. В то же время сейчас еще недостаточно развита теория возможных нелинейных динамических процессов в таких распределенных устройствах [2, 3]. Это, в частности, приводит и к тому, что не могут быть реализованы оптимальные схемные решения для устройств, например, распознавания изображений. До сих пор остается еще непонятым, каким образом строить устройства для выполнения эффективных

операций распознавания, с которыми быстро справляется нервная система животных. Поэтому представляется важным исследование коллективной динамики распределенных нейроподобных сред, многие свойства которых моделируются оптоэлектронными устройствами [1, 2].

Целью этой работы является изучение характерных динамических режимов, реализующихся в нейроподобных системах из оптоэлектронных элементов.

Особенность большинства нейроподобных сред связана с существованием нелокальной связи между активными элементами среды. В случае локальной связи информация о возбуждении элемента передается ближайшим соседям и сила этого сигнала монотонно спадает. Для описания процессов в таких средах используются модели так называемого реакционно-диффузионного типа [4—6]. Для нелокальной связи характерно немонотонное изменение силы воздействия сигнала на соседей, информирующего о возбуждении выбранного элемента среды.

Математические модели таких сред представлены интегродифференциальными уравнениями [1, 2, 7, 8]. При этом имеет смысл выделить два крайних случая для вида функции нелокальной связи  $\Phi(\xi, \mathbf{r})$ : а) близкая нелокальная связь, когда немонотонная функция  $\Phi(\xi - \mathbf{r})$  быстро спадает с расстоянием ( $\xi - \mathbf{r}$ ) (например, ближайших активизирует, а чуть более дальних тормозит — так называемое латеральное торможение); б) произвольно дальняя нелокальная связь, определяемая такой функцией  $\Phi(\xi, \mathbf{r})$ , величина которой зависит только от условий задачи и характеризует силу воздействия элемента в точке  $\xi$  на произвольно удаленный элемент в точке  $\mathbf{r}$  (такие связи, например, реализуются в модели Хопфилда [9—11]).

Имеются, конечно, и разнообразные варианты промежуточной нелокальной связи, но при рассмотрении процессов в нейроподобных средах и их оптоэлектронных аналогах эти случаи почти не рассматривались. В нашей работе ограничимся рассмотрением динамики структур в средах с близкой нелокальной связью, соответствующей латеральному торможению. Чтобы выделить особенности поведения структур в таких средах, будем сопоставлять их с известными импульсными решениями в обычных возбудимых средах [4—6].

## 1. МОДЕЛЬ

Оптоэлектронное устройство, моделирующее нейроподобную среду и состоящее из цепей с положительной и отрицательной обратными связями, описывается следующими уравнениями [1, 2, 12]:

$$\begin{aligned} \tau_B \frac{\partial u}{\partial t} = & -u + F_B \left( \int_B [\Phi_{uu}(\xi - \mathbf{x}) u(\xi, t) - \right. \\ & \left. - \Phi_{uv}(\xi - \mathbf{x}) v(\xi, t)] d\xi - T_1 \right) + I_0(\mathbf{x}, t), \end{aligned} \quad (1)$$



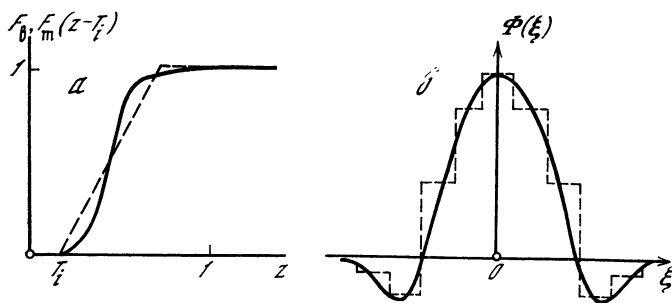


Рис. 1. Характерные параметры оптоэлектронного устройства, моделирующего нейрноподобную среду

*a* — качественный вид нелинейных функций  $F_B$  и  $F_T$  в модели; *b* — функция, характеризующая близкую нелокальную связь в случае латерального торможения. Пунктиром показаны принятые в моделях (8), (9) аппроксимации этих функций

$$\tau_T \frac{\partial v}{\partial t} = -v + F_T \left( \int_D [\Phi_{vu}(\xi - x) u(\xi, t) - \Phi_{vv}(\xi - x) v(\xi, x)] d\xi - T_2 \right). \quad (2)$$

Первое уравнение описывает цепь с положительной обратной связью. В соответствии с принятыми обозначениями для возбудимых сред  $u(x, t)$  — «активаторная» переменная [4—6]. Второе уравнение описывает цепь с отрицательной обратной связью;  $v(x, t)$  — «ингибиторная» переменная (обычно в нейрноподобных средах изменения  $v(x, t)$  более медленные и характеризуют процессы «запоминания» через изменение порогов активации для переменной  $u(x, t)$ );  $F_B, F_T$  — функции ступенчатого вида, описывающие передаточные характеристики усилителей в электронных цепях; величины  $T_i$  ( $i = 1, 2$ ) характеризуют пороги срабатывания усилителей (рис. 1, *a*); функции  $\Phi_{mk}$  описывают пространственные связи между активными элементами, которые воспринимаются  $m$ -й переменной в результате возбуждения  $k$ -й переменной;  $I_0(x, t)$  — внешний сигнал.

В связи с отсутствием регулярных методов аналитического рассмотрения решений нелинейного интегродифференциального уравнения на первом этапе исследования воспользуемся приближенной моделью, записанной в дифференциальной форме. Основные особенности процессов, связанных с близкой нелокальной связью, могут быть выведены из уравнения для активаторной переменной (при  $v = \text{const}$ ) в одномерном случае. Для этого интеграл в уравнении (1) перепишем в виде

$$\int_D \Phi(\xi - x) u(\xi, t) d\xi = \delta_0 u(x, t) + l_1 \frac{\partial u}{\partial x} + l_2^2 \frac{\partial^2 u}{\partial x^2} + l_3^3 \frac{\partial^3 u}{\partial x^3} + \delta_4 l_4^4 \frac{\partial^4 u}{\partial x^4} + \dots, \quad (3)$$

где

$$\delta_0 = \int_{-d}^d \Phi(\eta) d\eta, \quad l_1 = \int_{-d}^d \eta \Phi(\eta) d\eta, \quad l_2^2 = \frac{1}{2} \int_{-d}^d \eta^2 \Phi(\eta) d\eta, \\ l_3^3 = \frac{1}{6} \int_{-d}^d \eta^3 \Phi(\eta) d\eta, \quad \delta_4 l_4^4 = \frac{1}{24} \int_{-d}^d \eta^4 \Phi(\eta) d\eta, \quad (4)$$

параметр  $\delta_4$  равен либо  $+1$ , либо  $-1$ . Такое разложение справедливо для быстроспадающей функции  $\Phi(\eta)$ , т. е. для рассматриваемого здесь случая близкой немонотонности функции пространственной связи.

Приближенная модель, полученная из уравнения (1), имеет вид

$$\tau \frac{\partial u}{\partial t} = -u + F_B(\delta_0 u - T_1 - \beta v) + \alpha_1 \frac{\partial u}{\partial x} + \\ + \alpha_2 \frac{\partial^2 u}{\partial x^2} + \alpha_3 \frac{\partial^3 u}{\partial x^3} + \alpha_4 \frac{\partial^4 u}{\partial x^4}, \quad (5)$$

где

$$\alpha_i = \frac{\partial F}{\partial z} l_i^i (i = 1, 2, 3), \quad \alpha_4 = \frac{\partial F}{\partial z} \delta_4 l_4^4, \quad \beta = \int_{-d}^d \Phi_{uv}(\eta) d\eta.$$

Из формул (4) следует, что для симметричной функции пространственной связи получим  $l_1 = 0$  и  $l_3 = 0$ . Кроме того, видно, что при увеличении немонотонности функции  $\Phi(\xi)$  в случае латерального торможения (рис. 1, б) происходит уменьшение диффузионной длины  $l_2^2$  и характеризующей дисперсию диффузии длины  $l_4^4$ . Сначала при этом происходит изменение знака у  $\delta_4$ , и затем при увеличивающейся длине  $l_4^4$  получаем  $l_2 \rightarrow 0$ . Таким образом, в зависимости от величины немонотонности  $\Phi(\xi)$  имеются два значения параметра  $\delta_4$ :  $\delta_4 > 0$  соответствует условиям обычной возбудимой среды;  $\delta_4 < 0$  — условие, при котором можно ожидать появления новых режимов взаимодействия волн.

## 2. АНАЛИЗ РЕШЕНИЙ ПРИБЛИЖЕННОЙ ОДНОКОМПОНЕНТНОЙ МОДЕЛИ

Одной из основных характеристик обычных автоволновых сред является процесс уничтожения импульсов при столкновении их друг с другом. При столкновении фронтов среда переходит в одно из однородных состояний. Немонотонная связь между активными элементами в пространстве может приводить к новым режимам взаимодействия. Условия, при которых это происходит, можно определить из решений модели (5). В ней нелокальность связей описывается дополнительным слагаемым  $\delta_4 l_4^4 \partial^4 u / \partial x^4$  — дисперсией диффузии. Рассмотрение проводится с помощью аналити-

ческих расчетов при кусочно-линейном виде нелинейной функции

$$-u + F(\delta_0 u - T_1 - \beta v) = \begin{cases} -u & \text{при } u < u_*(T_1, v), \\ 1 - u & \text{при } u > u_*(T_1, v). \end{cases}$$

Новые стационарные решения в такой модели появляются лишь при  $\delta_4 = -1$  и  $\kappa \equiv 2(l_4/l_2)^4 > 0,5$ . Параметры импульсных решений с нулевой скоростью распространения определяются из дисперсионного уравнения

$$\exp\left(-\sqrt{\frac{\sqrt{2\kappa}+1}{2\kappa}}L\right)\left[\cos\sqrt{\frac{\sqrt{2\kappa}-1}{2\kappa}}L + \sin\sqrt{\frac{\sqrt{2\kappa}-1}{2\kappa}}L/\sqrt{2\kappa-1}\right] = 2u_* - 1, \quad (6)$$

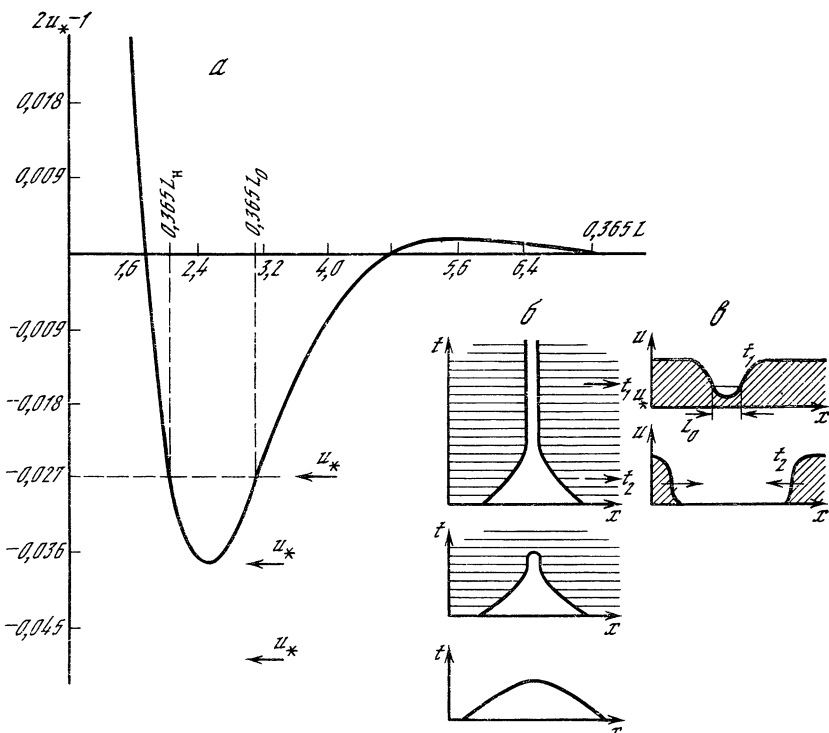
где  $L$  — характерный размер импульса на уровне порога  $u_*$  (рис. 2, а).

Из уравнения (6) видно, что в случае  $\kappa \leq 0,5$  может существовать только один импульс при значениях порога возбуждения среды  $u_* > 0,5$ . Известно, что этот стационарный импульс неустойчив к малым возмущениям и определяет некоторую границу для начальных стимулов, превышение которой необходимо для формирования распространяющихся фронтов возбуждения. В случае  $\kappa > 0,5$  в узком диапазоне значений порога около  $u_* \approx 0,5$  возможно существование дополнительных импульсных решений (рис. 2). При этом важно, что импульсные решения существуют при таких значениях порога возбудимой среды  $u_{* \min} \lesssim u_* \leq 0,5$ , при которых в обычной возбудимой среде отсутствуют импульсные решения, соответствующие сталкивающимся фронтам. По характеру дисперсионной кривой на рис. 2, а и на основе качественных представлений о поведении фронтов естественно предположить, что вновь появившийся импульс с большей длиной  $L_0$  соответствует устойчивому решению, тогда как импульс с меньшей длиной  $L_n$  по-прежнему остается неустойчивым к малым возмущениям<sup>1</sup>. Расчет, сделанный для распространяющихся стационарных фронтов возбуждения в модели (5), показывает, что зависимость скорости фронта

$$a = \frac{l_2}{\tau} (1 - 2u_*) \frac{8(\sqrt{2\kappa}+1)^{3/2}}{3\sqrt{2\kappa}+1}$$

от  $u_*$  (порога возбуждения среды) вблизи значений  $u_* \approx 0,5$  остается такой же, как и в обычной возбудимой среде. Следовательно, при  $\kappa > 0,5$  и  $u_{* \min} \lesssim u_* < 0,5$  возможны такие нестационарные процессы, когда сталкивающиеся фронты либо резко снижают свою скорость перед их исчезновением, либо останавливаются и формируются устойчивый импульс невозбужденного состояния

<sup>1</sup> Существование устойчивых импульсных решений в уравнении, подобном (5), было показано, например, в работе [7].



**Рис. 2.** Решения в однокомпонентной модели (5), записанной через дифференциальные операторы

*a* — зависимость длины неподвижных импульсов  $L$  от величины порога возбудимой среды  $u_*$  при  $\kappa = 40$  ( $l_4 \approx 2,11 l_2$ ); *б* — схемы нестационарных решений модели (5) (возбужденные участки заштрихованы), соответствующие трем порогам возбудимой среды (на рис. 2, *a* они показаны стрелками); *в* — решения  $u(x, t)$  в моменты  $t_1, t_2$  (до и после столкновения фронтов) для процесса, соответствующего верхней схеме рис. 2, *б*

(см. рис. 2, *б*). Такой импульс устойчив к возмущениям его длины с амплитудой, меньшей  $\sim (L_0 - L_H)$ . Если же возмущение сделает длину импульса меньше  $L_H$ , система перейдет к однородному возбужденному состоянию.

Отметим, что в области существования новых стационарных решений ( $\delta_4 = -1$ ) модели (5) стационарные состояния ( $u_0 = 0$  или  $+1$ ) всегда остаются устойчивыми, так же как и в исходном интегродифференциальном уравнении. Следовательно, качественное соотношение между решениями этих уравнений сохраняется. В то же время в области, соответствующей обычной возбудимой среде ( $\delta_4 = 1$ ), высокочастотные пространственные возмущения  $u - u_0 \sim e^{\gamma t} e^{ikx}$  оказываются неустойчивыми ( $\gamma = -1 - l_2^2 k^2 + l_4^2 k^4 > 0$  при  $k = \frac{2\pi}{\lambda} > \frac{l_2}{\sqrt{2} l_4^2} (1 + \sqrt{1 + 2\kappa})^{1/2}$ ). Этот эффект,

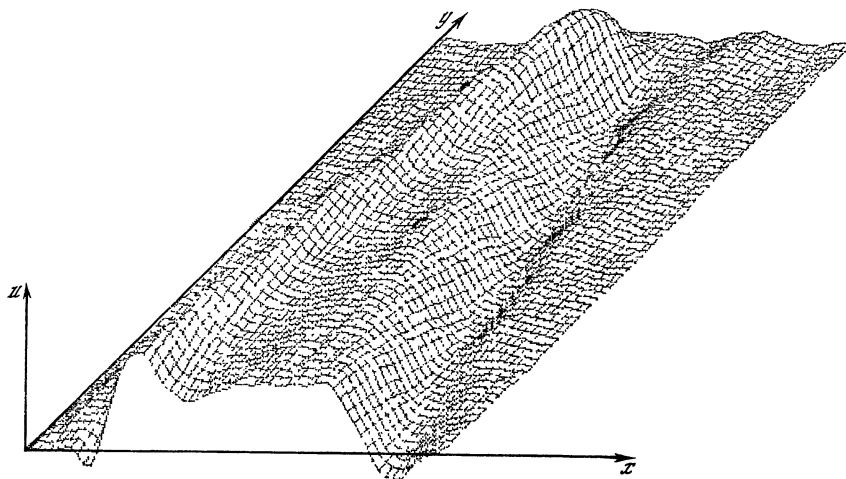
как нетрудно проверить, связан с приближенным характером модели (5). Учет следующих членов в разложении интеграла в ряд (3) убирает такую неустойчивость.

### 3. СТРУКТУРЫ В АНАЛОГОВОЙ МОДЕЛИ ДВУМЕРНОЙ СРЕДЫ С НЕЛОКАЛЬНЫМИ СВЯЗЯМИ

Режимы с одиночными стационарными импульсами, а также различные нестационарные процессы формирования более сложных структур были реализованы в модельных аналоговых устройствах из цепей монитор—камера—обратная связь. Изучение возможных структур проводилось на установке из одной цепи монитор—камера—обратная связь—однокомпонентный ТВ-аналог и на установке из двух цепей — двухкомпонентный ТВ-аналог. Математическая модель, описывающая основные черты работы таких установок, записывается в виде уравнений (1), (2). Нелинейные функции  $F_{в,т}$  определяются параметрами электронных усилительных устройств в камере и мониторе. Функции пространственной связи  $\Phi_{km}(\xi - x)$  определяются настройкой оптического пути от монитора к камере и особенностями частотных характеристик электронных усиливающих устройств.

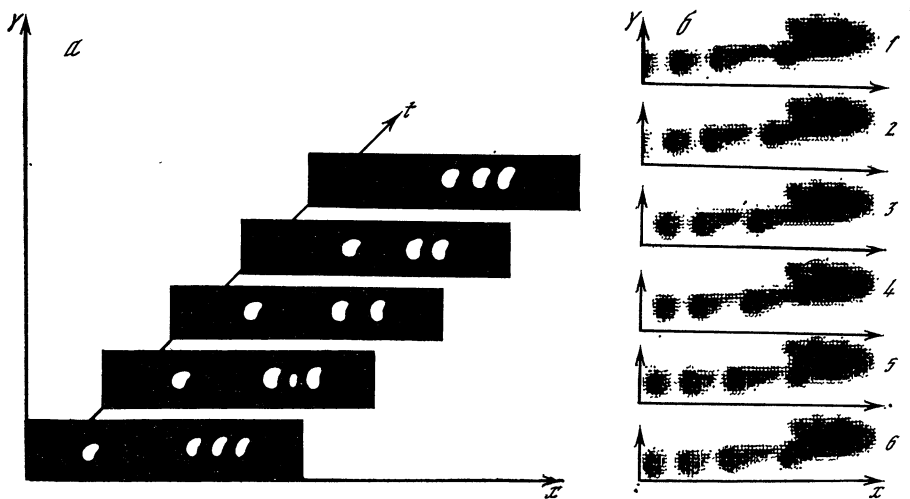
Из-за нелинейных искажений в строчных развертках мониторов, неравномерности частотных передаточных характеристик усиливающих устройств, а также из-за отличия характеристик оптических путей при передаче сигналов от разных точек экрана реальные параметры установки имели пространственные неоднородности: по порогу возбуждения; по несимметричности функций пространственной связи (и следовательно, по величине скорости и направлению сноса возбужденных областей). Тем не менее всегда удавалось выделить определенную область, в пределах которой снос возбуждений отсутствовал либо был слабо выражен.

На однокомпонентной аналоговой модели реализуются распространяющиеся фронты зажигания (при малых порогах возбуждения) или погасания экрана (при больших порогах), а при параметрах, соответствующих малой скорости фронтов и немонотонной функции связи, — уединенные импульсы возбуждения. При монотонной функции связи уединенных импульсов возбуждение получить не удавалось. Эти условия полностью соответствуют аналитическому рассмотрению на модели (5). Вид функции пространственной связи определялся по характеру преобразования, которое происходило с сигналом в виде светящихся отрезков различной длительности в пространстве при одном прохождении этих сигналов через цепь монитор—камера. Чтобы одновременно судить о прохождении отрезков разной длины, на монитор через цифроаналоговый преобразователь подавался сигнал в виде треугольника, интенсивность свечения которого внутри контура задавалась постоянной и максимальной для нашего устройства, а вне контура скачком падала до фонового уровня. Преобразованный сигнал показан на рис. 3.



**Рис. 3.** Амплитуда двумерного сигнала, прошедшего через цепь монитор — камера

Начальная форма сигнала задавалась скачкообразными изменениями интенсивности  $u$  в виде треугольника на плоскости  $(x, y)$



**Рис. 4.** Примеры нестационарных процессов, полученных на однокомпонентном ТВ-аналог

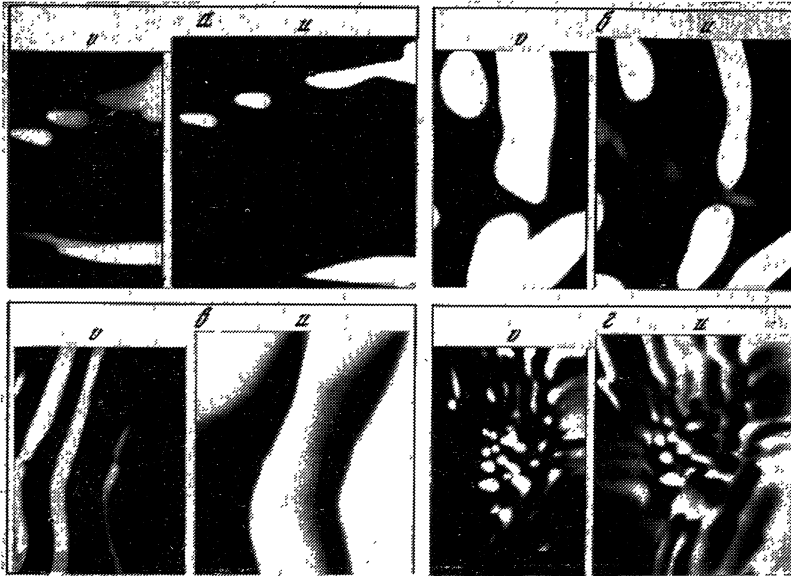
*a* — переходный процесс в цепочке из автоструктур при уничтожении одного из импульсов;  
*б* — генерация дрейфующих импульсов в устройстве с плавной пространственной неоднородностью. Цифрами отмечена временная последовательность кадров, сделанных через 160 мс

Видно, что по краям сигнала появляются участки заторможенных состояний. Это возможно лишь при действии функции пространственной связи с латеральным торможением. В таком устройстве в зависимости от размера начального возмущения формировался один или несколько импульсов, разделенных зонами заторможенного состояния. Внешним воздействием можно было уничтожить каждый из импульсов. Для этого необходимо было уменьшать его размер до определенной критической величины, которая была в 2—3 раза меньше размера стационарного импульса. Если в цепочке импульсов производилось уничтожение одного из импульсов, то за счет слабого дрейфа происходило сближение импульсов на минимально возможное для этих импульсов расстояние. Пример такого переходного процесса показан на рис. 4, а.

В среде с плавной-неоднородной функцией пространственной связи может возникнуть генерация дрейфующих импульсов. Фронт возбуждения распространяется в область пространства, где устойчивыми могут быть лишь уединенные импульсы. Такие импульсы формируются и из-за несимметричности функции пространственной связи продолжают сноситься дальше от места их генерации. Дрейф импульсов описывается в модели (5) прежними решениями, если при  $l_1 \neq 0$  перейти к новой пространственной переменной  $x' = x + (l_1/\tau)t$ . Этот процесс показан на рис. 4, б. Таким образом, имеется хорошее соответствие между наблюдаемыми на аналоговой установке особенностями формирования и взаимодействия импульсов и решениями, полученными из приближенной модели (5).

Режимы «ячеистого» или «пятнистого» возбуждения получались также и на установке «двухкомпонентный ТВ-аналог». Латеральное торможение обеспечивалось совместной работой двух цепей, в каждой из которых диффузионное расплывание сигналов имело характерные размеры  $L_1$  и  $L_2$ . При этом первоначально возбужденные элементы экрана активировали лишь ближайших соседей  $|\xi - x| \leq L_1$ , а следующие за ними элементы в кольце  $L_1 \leq |\xi - x| \leq L_2$  тормозились.

Примеры медленно дрейфующих импульсных структур на двухкомпонентном ТВ-аналоге показаны на рис. 5, а. В такой системе наблюдаются и другие пространственно-временные структуры (рис. 5, б, в). В данном оптоэлектронном устройстве характерные времена изменения сигналов в обоих связанных цепях имели примерно одинаковые значения ( $\tau_b \approx \tau_T$  в модели (1), (2)). Исследование решений модели реакционно-диффузионного типа показало, что при  $\tau_b \approx \tau_T$  и  $L_1 \gg L_2$  происходит переход от решений типа диссипативных структур к решениям в виде распространяющихся импульсов [5, 13]. Поэтому существовала вероятность обнаружить импульсы возбуждения. К сожалению, на используемой нами установке этот переход «замазывался» явлениями сноса областей возбуждения. Тем не менее иногда удавалось наблюдать распространение импульсов возбуждения, обладающих свойствами уничтожать друг друга при столкновении.



**Рис. 5.** Примеры пространственно-временных структур в двухкомпонентном ТВ-аналоге

*a* — генерация дрейфующих импульсов; *б* — локализованные структуры возбуждения; *в* — распространяющиеся волны возбуждения; *г* — периодический режим формирования распространяющихся волн из локализованных структур возбуждения. Периодичность определяется изменением порога возбуждения системы за счет работы АРУ камеры в цепи «активатора». На фотографиях справа — пространственная картина возбуждения «активатора»  $u(x, t)$ , слева — «ингибитора»  $v(x, t)$ ; размер пространственной картины для  $v(x, t)$  уменьшен примерно в 0,79 раза по сравнению с размером картины для  $u(x, t)$

Кроме разнообразных режимов генерации пространственных картинок из автоструктур, установка «двухкомпонентный ТВ-аналог» может осуществлять простые операции преобразования подаваемого на нее начального сигнала. В частности, при соответствующем подборе порогов возбуждения элементов экрана и функции пространственной связи возможно выполнение операций контрастирования, оконтуривания видеообразов, а также некоторые варианты ассоциативного запоминания.

#### 4. ВЗАИМОДЕЙСТВИЕ ИМПУЛЬСОВ В ОДНОРОДНЫХ СРЕДАХ

Особенности взаимодействия импульсов в однородных средах (при симметричной функции пространственной связи в уравнениях (1)–(3)) были исследованы с помощью расчетов упрощенной математической модели оптоэлектронного устройства:

$$\frac{\partial u}{\partial t} = -u + F\left(-T + (1 - \beta v) \int_{-\infty}^{\infty} \Phi(\xi - x) u(\xi, t) d\xi\right), \quad (8)$$

$$\frac{\partial v}{\partial t} = (\tau_B/\tau_T)(u - v). \quad (9)$$



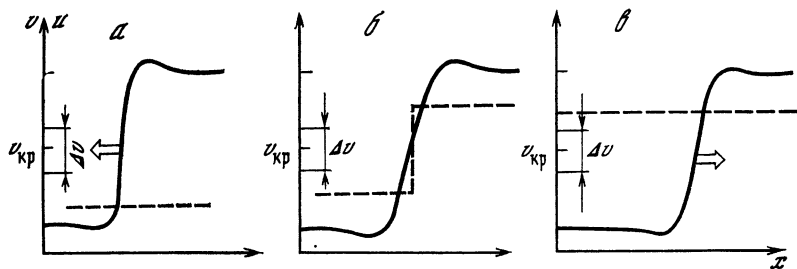


Рис. 6. Направление распространения фронта активации  $u(x - at)$  (сплошная линия) при различных значениях ингибиторной переменной  $v$  (пунктир)  $a - v < v_{кр}$ ;  $b - v = v_{кр}$  или неоднородное распределение  $v(x)$ , пересекающее значение  $v_{кр}$  (см. условия для неподвижного фронта в [4, 6]);  $c - v > v_{кр}$

Эта модель соответствует оптоэлектронному устройству, в котором порог возбуждения активаторной переменной  $u(x, t)$  управляется через отрицательную обратную связь (переменная  $v(x, t)$ ) изменением коэффициента усиления в суммирующем пространственные сигналы устройстве. Расчет проводился для кусочно-линейной аппроксимации нелинейной функции  $F(z)$  и немонотонной функции пространственной связи  $\Phi(\xi)$  (см. рис. 1). Оказалось, что качественное рассмотрение процессов взаимодействия импульсов, описываемых моделью (8), (9), можно проводить при  $\tau_B/\tau_T \ll 1$  (так же, как и для обычной возбудимой среды [4, 6]) с помощью упрощенных характеристик: зависимости скорости фронта возбуждения (перепад активатора  $u(x, t)$ ) от величины медленной ингибиторной переменной  $v(x, t)$ ; нуль-изоклин для соответствующей однородной системы. Дополнительной операцией при рассмотрении решений системы (8), (9) является учет условий, при которых сталкивающиеся один с другим фронты останавливаются.

В соответствии с результатами работ [4, 6, 14] скорость фронта возбуждения зависит от величины переменной  $v(x, t)$  в месте нахождения фронта и меняет свой знак при значении  $v = v_{кр}$ , которое определяется нелинейной функцией  $F(z)$ . При этом фронт имеет нулевую скорость не только при  $v = v_{кр}$ , но и на неоднородности в виде скачка медленной переменной около значения  $v_{кр}$  (рис. 6). Кроме того, как следует из аналитического рассмотрения (см. рис. 2, а), около  $v \approx v_{кр}$  имеется интервал значений  $\Delta v$ :  $\{v_{* \min} \leq v \leq v_{* \max}\}$ , при которых сталкивающиеся фронты останавливаются (для модели (5)  $\Delta v/2$  определяется из уравнения  $u_*(v_*) = 2(u_{* \min} - 0,5)$ ). Воспользуемся теперь этими характеристиками для краткого описания условий реализации полученных на модели (8), (9) режимов взаимодействия импульсов (рис. 7). Свойства сосредоточенных элементов выбраны таким образом, что без внешнего воздействия система находится в стационарном состоянии  $(v_p, u_p)$ , а внешний стимул вызывает генерацию одного импульса и релаксацию системы в то же стационарное состояние.

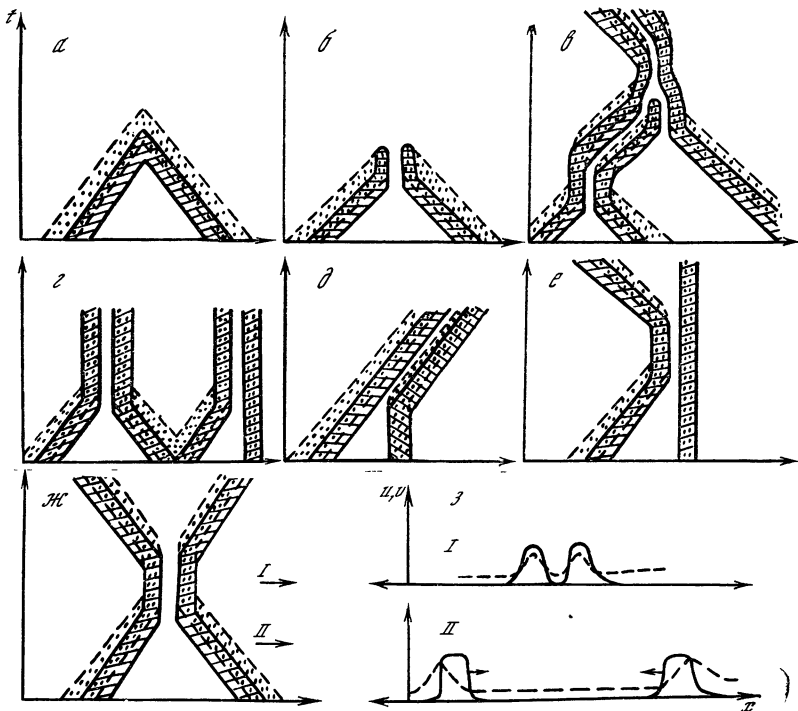


Рис. 7. Схемы режимов взаимодействия импульсов возбуждения на плоскости  $(x, y)$  в среде с латеральным торможением

Области возбуждения (величина «активатора»  $u(x, t)$  больше порога возбуждения) заштрихованы. Области, где величина «ингибитора»  $v(x, t)$  превышает значение  $v_{кр}$  (пунктирные линии), выделены точками.

*a* — уничтожение импульсов при столкновении; *б* — уничтожение импульсов с остановившимися фронтами; *в* — формирование связанных состояний при столкновении двух импульсов; *г* — формирование неподвижных локализованных возбуждений при столкновении двух импульсов или импульса, столкнувшегося с неподвижным возбуждением; *д* — формирование двух распространяющихся импульсов при столкновении импульса с неподвижным возбуждением; *е* — отражение импульса от неподвижного возбуждения; *ж* — отражение двух сталкивающихся импульсов; *з* — пространственное распределение переменных  $u(x, t)$  (сплошная линия) и  $v(x, t)$  (штриховая линия) для времен на рис. 7, *ж*, соответствующих распространению импульсов навстречу друг другу (*II*) и остановке их при столкновении (*I*)

1.  $v_p$  находится вне диапазона  $\Delta v$ . Сталкивающиеся фронты возбуждения не могут остановить друг друга (рис. 7, *a*), и вся динамика взаимодействия импульсов аналогична известной для обычной возбудимой среды [4, 6].

2.  $v_p$  находится внутри диапазона  $\Delta v$ . Однако параметры  $T$  и  $\beta$  такие, что решения уравнения

$$u(x) = F\left(-T + (1 - \beta u) \int_{-\infty}^{\infty} \Phi(\xi - x) u(\xi) d\xi\right) \quad (10)$$

оказываются неустойчивыми в рамках уравнений (8), (9). В этом случае фронты импульсов останавливаются, а подходящие спады импульсов уничтожают это временное связанное состояние (см. рис. 7, б).

3. В уравнении (10) имеется решение, которое является устойчивым в рамках уравнений (8), (9). Если область притяжения к устойчивому импульсу не обладает необходимой шириной, то сталкивающиеся импульсы лишь некоторое время остаются неподвижными, а затем образуют связанное состояние из двух распространяющихся импульсов (рис. 7, в). Направление, в котором будут двигаться импульсы, зависит от того, какой из импульсов первым отразится от места остановки. В более коротком из сталкивающихся импульсов перестройки медленной переменной  $v(x, t)$ , необходимые для отражения, происходят быстрее. Импульс, оказавшийся зажатым между двух других, при таких параметрах системы исчезает (тормозящее влияние окружающих его импульсов настолько снижает амплитуду возбуждения, что выводит ее за пределы области притяжения к устойчивому импульсу) (см. рис. 7, в).

4.  $v_p$  находится внутри диапазона  $\Delta v$ , и существует устойчивая автоструктура — импульс с нулевой скоростью. Имеется конечная область притяжения к такому импульсу. При широкой области притяжения как столкнувшиеся, так и один импульс, подошедший к неподвижной автоструктуре, образуют неподвижный очаг возбуждения (автоструктура из двух импульсов) (рис. 7, г). Если же возмущение, оказываемое распространяющимся импульсом, превышает размер области притяжения, происходит разрушение неподвижного импульса и формирование двух распространяющихся (см. рис. 7, д).

5. В зависимости от длины сталкивающихся импульсов возможна ситуация, когда в области спадов импульсов, остановленных неподвижными фронтами, происходят такие изменения неоднородных распределений  $v(x, t)$  при которых спады превращаются во фронты импульсов возбуждения и происходит отражение этих импульсов<sup>2</sup> (см. рис. 7, е, ж).

Таким образом, в однородных средах с близкой нелокальной пространственной связью в виде латерального торможения в зависимости от параметров среды возможны весьма разнообразные режимы взаимодействия: уничтожение, остановки, отражения, различные связанные состояния импульсов.

Все приведенные на рис. 7 особенности взаимодействия импульсов сохраняются и для процессов в двумерной среде. Численные расчеты подтвердили существование таких взаимодействий. Однако более интересным представляется выявление двумерных эффектов, связанных с самовзаимодействием и взаимодейст-

---

<sup>2</sup> Приведенное качественное описание обосновано пока только численными расчетами. Для анализа условий потери устойчивости неподвижными фронтами и импульсами требуются дополнительные исследования.

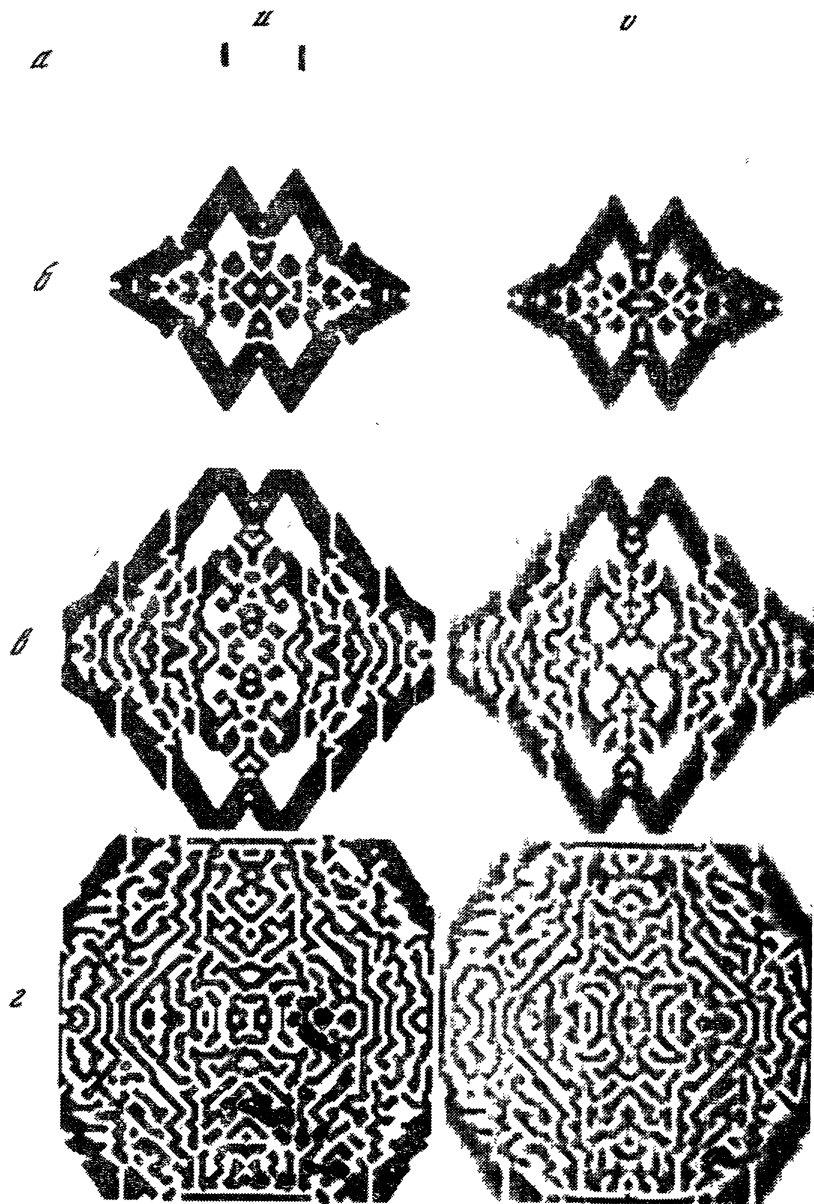


Рис. 8. Формирование двумерной картины возбуждения в среде с анизотропией и латеральным торможением

$a$  — число шагов счета  $t = 0$ ,  $б$  — 115,  $в$  — 200,  $г$  — 300

вием импульсов возбуждения. Оказалось, что новые эффекты такого типа могут быть в среде с анизотропными свойствами.

На рис. 8 приведен пример процесса в среде с анизотропной функцией пространственной связи, которая по-прежнему имеет характер латерального торможения. Импульсы возбуждения распространяются от двух областей, показанных темными линиями при  $t=0$ . Расположение линии фронтов в пространстве достаточно наглядно показывает соотношение скоростей фронтов возбуждения вдоль осей  $x$ ,  $y$  и вдоль диагоналей под углом в  $45^\circ$  к осям. Видно возникновение неустойчивости на распространяющемся фронте. При определенной величине вогнутости линии фронта появляется участок с нулевой скоростью, который затем начинает прорастать вдоль диагоналей или вдоль осей и образовывать структуры заторможенности в областях возбужденного состояния. Аналогичные неустойчивости наблюдаются и на спаде импульса. В этом случае, наоборот, на выпуклых участках спада образуются «языки» возбуждения, прорастающие по областям покоящегося состояния.

Неустойчивость изогнутого фронта вызвана немонотонным характером функции пространственной связи. На вогнутом фронте наибольшее тормозящее воздействие от соседей оказывается на участок, имеющий наибольший изгиб. При этом скорость распространения этого участка падает и изгиб возрастает. Когда скорость становится равной нулю, начинает формироваться двумерная автоструктура. Из-за анизотропных свойств среды автоструктура формируется либо вдоль осей, либо вдоль диагоналей. Наиболее выпуклый участок спада подвергается наименьшему тормозящему воздействию со стороны соседних возбужденных областей, поэтому в нем наиболее легко реализуются условия прорастания структуры активности. Процесс прорастания структур возбуждения по покоящейся области или, наоборот, структур и заторможенных участков по областям возбуждения зависит от расположенных вокруг них других автоструктур. В результате образуется либо стационарный пространственный рисунок, либо взаимодействующие структуры находятся в сложном пульсирующем режиме (повторяющийся процесс образования структуры, ее прорастания и гибели).

Из рис. 8 видна еще одна характерная особенность структурообразования, по-видимому общая для неравновесных сред. Речь идет о последовательности переходов между возможными структурами в среде. Из начального возмущения формируется ближайшая из притягивающих стационарных структур (в данном случае распространяющийся импульс). Затем развивающиеся неустойчивости разрушают первоначальную и формируют более устойчивую структуру (отрезки импульсов), и так до тех пор, пока не останутся лишь наиболее устойчивые для данной среды автоструктуры. Аналогичная направленность переходных процессов известна, например, для простейших возбудимых сред [15]. Отметим также, что с увеличением скорости фронтов возбуждения влия-

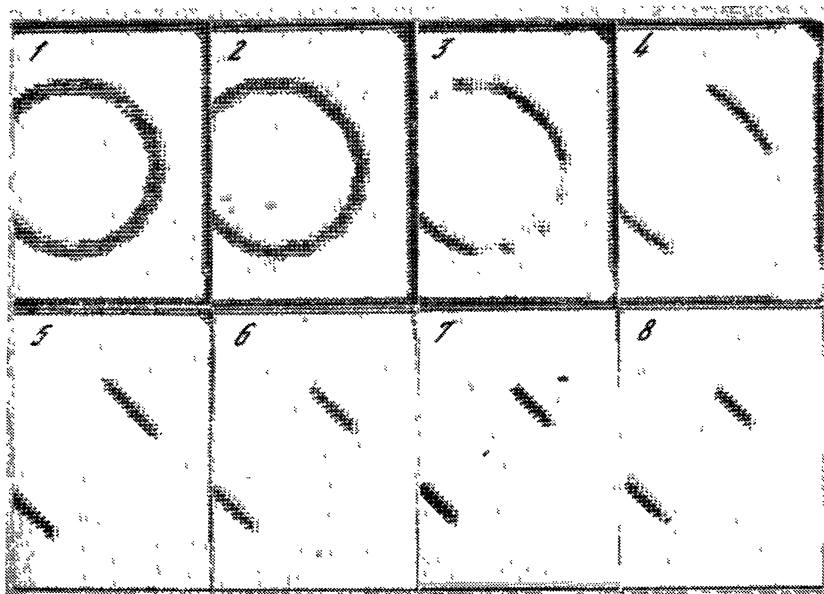


Рис. 9. Формирование линий заданного направления в среде с анизотропной функцией связи (выделенное направление  $\varphi = -\pi/4$  к оси  $x$ )

ние анизотропии среды ослабевает, эффекты неустойчивости и прорастания автоструктур исчезают.

Одним из простых примеров использования анизотропных свойств распределенных сред является выделение с их помощью характерных признаков обрабатываемых видеобразов. Происходит это в параллельном режиме по всему изображению. В качестве примера на рис. 9 приведен динамический процесс выделения линий заданного направления. Параметры среды специально были подобраны таким образом, чтобы формирующиеся в результате обработки начального изображения автоструктуры наибольшее время сохраняли геометрические соотношения исходного видеобраза. Возможны и другие варианты.

Приведенные здесь примеры для структур в двумерных однородных средах показывают, что совместное действие анизотропных и нелокальных свойств функции пространственной связи может приводить к формированию новых автоструктур (неподвижные и прорастающие локализованные области возбуждения или покоя в двумерной среде) и к новым переходным процессам, связанным с неустойчивостями искривленных фронтов.

Таким образом, в распределенных оптоэлектронных устройствах в зависимости от характера нелокальной пространственной связи, анизотропных свойств этой связи и режимов активации отдельных элементов возможны следующие ситуации: а) в двумерной среде могут формироваться новые автоструктуры (непо-

движные и прорастающие локализованные области возбуждения или покоя); б) в процессе взаимодействия происходят остановки, отражения, формирование различных связанных состояний импульсов; в) при определенной кривизне распространяющихся двумерных фронтов возникают неустойчивости, приводящие к разрушению фронтов и формированию устойчивых автоструктур. Достаточно большое разнообразие режимов поведения в таких распределенных средах (по сравнению с динамикой обычных возбудимых сред) показывает большие возможности их использования в разработке универсальных устройств обработки сигналов на однородных активных средах. При этом в качестве элементов таких устройств могут быть использованы не только оптоэлектронные приборы, но и, например, интенсивно изучаемые сейчас биологические материалы с активными свойствами. Представления о структурах в возбудимых средах несомненно будут полезными также и при разработке методов программирования на «мелкозернистых» параллельных вычислительных машинах [3, 16].

Авторы благодарны М. И. Рабиновичу за постоянное внимание к работе и неоценимую помощь в ее проведении.

#### ЛИТЕРАТУРА

1. Дубинин Ф. Д. Оптоэлектронные модели однородных сред. М.: Радио и связь, 1984. 124 с.
2. Левшин В. Л. Биокibernетические оптико-электронные устройства автоматического распознавания изображений. М.: Машиностроение, 1987. 176 с.
3. Ниллис У. Д. Коммутационная машина // В мире науки. 1987. № 8. С. 60—69.
4. Автоволновые процессы в системах с диффузией / Под ред. М. Т. Греховой. Горький: ИПФ АН СССР, 1981. 287 с.
5. Кернер Б. С., Осипов В. В. Явления в активных распределенных системах // Микроэлектроника. 1985. № 5. С. 389—407.
6. Vasilev V. A., Romanovskii Yu. M., Chernovskii D. S., Yakhno V. G. Auto-wave processes in kinetic systems. Berlin: VEB Dt. Verl. Wiss., 1987. 262 p.
7. Белинцев Б. Н., Лившиц М. А., Волькенштейн М. В. Пространственная самоорганизация как результат близкодействующей положительной кооперативности и дальнедействующей отрицательной кооперативности // ДАН СССР. 1981. Т. 257, № 2. С. 487—491.
8. Фролов А. А., Муравьев И. П. Нейронные модели ассоциативной памяти. М.: Наука, 1987. 160 с.
9. Hopfield J. J., Tank D. W. «Neural» computation of decisions in optimization problems // Biol. Cybern. 1985. Vol. 52, N 3. P. 141—152.
10. Веденов А. А., Ежов А. А., Левченко Е. Б. Нелинейные динамические системы с памятью и функции нейронных ансамблей // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 53—67.
11. Кинцель В. Спиновые стекла как модельные системы для нейронных сетей // УФН. 1987. Т. 152, № 1. С. 123—131.
12. Голубев В. Н., Рабинович М. И., Таланов В. И. и др. Критические явления в неоднородных возбудимых средах. Моделирование на «ТВ-аналоге» // Письма в ЖЭТФ. 1985. Т. 42, № 3. С. 84—87.
13. Белинцев Б. Н. Диссипативные структуры и проблема биологического формообразования // УФН. 1983. Т. 141, № 1. С. 55—101.
14. Голубев В. Н., Рабинович М. И., Толков В. Н., Яхно В. Г. Динамические структуры в двухкомпонентных возбудимых средах. Моделирова-

ние на клеточном компьютере // Биомолекулярная электроника. Пушкино, 1987. С. 89—98.

15. *Starobinets I. M., Yakhno V. G.* One-dimensional autowaves, methods of qualitative description, self-organization // *Autowaves and structures far from equilibrium* / Ed. V. L. Krinsky. Berlin etc.: Springer, 1984. P. 98—102.
16. Special issue on parallelism // *Communications of the ASM.* 1986. Vol. 29, N 12.

УДК 530.182

## **ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ДВУМЕРНЫХ НЕРАВНОВЕСНЫХ СРЕД НА МНОГОПРОЦЕССОРНОМ КОМПЛЕКСЕ**

А. В. РОГАЛЬСКИЙ

### **ВВЕДЕНИЕ**

Численное моделирование нелинейных динамических процессов (жизнь во времени) в двумерных и трехмерных неравновесных средах предъявляет экстремальные требования к возможностям вычислительной техники. Подтверждается известный тезис о том, что самые нужные задачи всегда требуют больше ресурсов, чем в состоянии дать мощнейший из компьютеров. Для задач нелинейной динамики это обусловлено рядом факторов.

Во-первых, для большинства задач, например по изучению турбулентности, образованию и взаимодействию структур, развитию пространственно-временного хаоса, необходимо исследовать пространственные области, достаточные не только для формирования и взаимодействия отдельных структур, но и для изучения динамики коллективных движений таких структур. Учет при этом еще и двумерности среды приводит к необходимости проводить исследования в широких диапазонах пространственных переменных.

Во-вторых, различные пространственные возбуждения, наблюдаемые в нелинейных неравновесных средах, характеризуются существенно различными инкрементами, причем многие неустойчивости развиваются лишь на фоне неоднородных состояний. Так, при исследовании режимов развитого пространственно-временного хаоса времена выхода системы на аттрактор, зависящие от пространственных размеров исследуемой области и близости параметров среды к точке бифуркации, могут быть очень большими. Следовательно, и временные масштабы моделирования велики.

Таким образом, при изучении нелинейных процессов в неравновесных средах возникает необходимость исследовать целую иерархию одновременно существующих образований различных пространственных и временных масштабов. Исследование таких режимов, с одной стороны, практически не поддается аналитическому анализу, с другой — при численном моделировании требует



вычислительных средств с реальной производительностью  $F$  не менее нескольких десятков миллионов операций с плавающей точкой в секунду (Мфлоп).

Для обеспечения необходимых вычислительных мощностей можно использовать универсальные ЭВМ с традиционной векторной архитектурой, называемые суперЭВМ. При условии хорошей векторизуемости алгоритмов и монопольном использовании такие суперЭВМ действительно обеспечивают требуемые ресурсы. Однако недостатком таких векторных компьютеров (CDC Cyber 205, Cray Series), помимо их высокой стоимости, является факт резкого снижения реальной (на прикладных задачах) производительности при наличии даже небольшой не векторизуемой доли вычислений [4]. При этом реальная производительность может упасть до единиц мегафлоп при пиковой производительности суперЭВМ сотни мегафлоп.

Альтернативный подход состоит в использовании параллельных вычислительных систем, когда несколько процессоров совместно решают одну задачу. В Институте космических исследований АН СССР совместно с ГХО ИЗОТ НРБ создан многопроцессорный комплекс на базе матричных процессоров (МКМП), основанный на этом подходе.

## 1. СТРУКТУРА И ВОЗМОЖНОСТИ МКМП

Идея использования нескольких процессоров для совместного решения одной задачи не является новой [2, 3]. Тем не менее большинство используемых сегодня многопроцессорных систем работает лишь в мультизадачном режиме, эффективно обрабатывая поток независимых программ. Анализ проблем, существующих в разработке и применении параллельных систем для решения одной задачи, позволяет выделить некоторые причины такого положения.

Разработчики ЭВМ, как правило, пытаются выявить и использовать параллелизм на уровне инструкций существующих программ. Однако отсутствие внутреннего параллелизма в традиционно последовательной структуре большинства программ на этом уровне не позволяет получить здесь существенного выигрыша.

С другой стороны, фундаментальные исследования параллельных систем часто реализуются в принципиально новых подходах к архитектуре, в разработке новых параллельных языков программирования. Помимо предельно сложной (и пока не решенной) проблемы создания базового программного обеспечения (ПО) таких систем, возникает необходимость разработки нового прикладного ПО и отказа от существующих пакетов прикладных программ.

Поэтому не случайно, что параллельные системы из нескольких матричных процессоров, ориентированные на решение одной задачи, появились в ряде научных центров именно среди активных пользователей мощной вычислительной техники: физиков, химиков, геофизиков [4, 5]. Применение матричных процессоров для

ускорения вычислений стимулировало здесь развитие старых и постановку новых, более сложных задач. Соответственно выросли требования к возможностям вычислительных средств. Параллельные системы возникли на основе традиционного серийного оборудования путем его комплексации, развития и добавления необходимого базового ПО. Существенно, что все изменения и дополнения диктовались требованиями реализации прикладных программ.

Многопроцессорный комплекс на базе матричных процессоров в ИКИ АН СССР появился в результате развития вычислительной системы, использующей матричные процессоры (МП). Комплекс состоит из управляющего процессора ЕС-1037 и 10 присоединенных (через каналы ввода—вывода) матричных процессоров ЕС-2706, все оборудование производится серийно ГХО ИЗОТ НРБ [6]. Управляющий процессор ЕС-1037 представляет собой универсальную ЭВМ традиционной архитектуры с производительностью 1,5 млн инструкций в секунду, оперативной памятью 16 Мбайт, 6 каналами ввода—вывода, со скоростью передачи данных по каналу до 2,5 Мбайт/с. Система команд ЭВМ ЕС-1037 полностью совместима с серией IBM/370.

Название «матричный» применительно к процессору ЕС-2706 чисто условное, оно никак не характеризует процессор: он не использует матрицы как операнды и не есть матрица процессоров. Для дальнейшего важно, что МП — универсальный, быстрый вычислитель, устройство для работы с массивами чисел, в том числе и с векторами. Пиковая производительность МП равна 12 Мфлоп, локальная память данных МП до 4 Мбайт. Архитектура МП характеризуется функциональным параллелизмом, который обеспечивает одновременное выполнение таких операций, как сложение и умножение с плавающей точкой, операции целочисленной арифметики, адресации, выборку данных из памяти и регистров. Максимальная производительность МКМП — 120 Мфлоп, максимальная память данных, доступная одной задаче, 50 Мбайт. На реальных задачах, например нелинейной динамики, производительность МКМП составляет 60 Мфлоп.

Разработанное в ИКИ АН СССР базовое ПО комплекса превращает систему с несколькими МП в многопроцессорный комплекс с возможностью совместного решения одной задачи. Состав и структура базового ПО подробно описана в работах [6, 7]. Здесь мы остановимся на возможностях, представляемых пользователю при распараллеливании реальных задач.

Существуют следующие основные возможности работы МКМП: выполнение одинаковых подпрограмм практически одновременно на любых имеющихся МП; выполнение различных подпрограмм на любых непересекающихся подмножествах МП одновременно с работой центрального процессора (ЦП); синхронизация работы любого подмножества МП в нужной точке алгоритма. Возможна такая работа системы, когда сам факт готовности данных по завершению операции инициирует выполнение следующей, т. е. управ-

ление организуется потоком данных, а не инструкций. Таким образом, известные режимы SIMD, MIMD используются в МКМП на другом уровне интеграции. Вместо отдельных инструкций выполняются подпрограммы, вместо элементов массивов в качестве данных — сами массивы. Внутри же МП используется канонический MIMD режим на уровне инструкций.

Практическое использование этих и ряда других возможностей организации работы МКМП мы продемонстрируем на примере задачи численного моделирования двумерных неравновесных сред, описываемых уравнением Гинзбурга—Ландау.

## 2. ДВУМЕРНАЯ РЕШЕТОЧНАЯ МОДЕЛЬ ГИНЗБУРГА—ЛАНДАУ

### Методы, уровни и результаты распараллеливания

Как известно, одной из наиболее ярких и важных задач нелинейной физики является проблема формообразования и развития турбулентности в неравновесных средах. Исследования этих и близких явлений требуют создания единых моделей, понятий и методов, общих для неравновесных сред произвольной природы. При этом особое значение приобретает компьютерное исследование универсальных моделей нелинейной динамики. В работах [8, 9] выделен ряд базовых моделей нелинейных сред. Одной из них является двумерная модель Гинзбурга—Ландау

$$\partial U/\partial t = U[1 - (1 + i\beta) |U|^2] + e(1 - ic)\Delta U. \quad (1)$$

Содержательная теория перехода к пространственно-временному хаосу — турбулентности в двумерных средах, описываемых уравнением Гинзбурга—Ландау, при увеличении надкритичности среды может быть построена на основе решеточной модели:

$$dU_{jk}/dt = U_{jk} - (1 + i\beta) |U_{jk}|^2 U_{jk} + e(1 - ic) \times \\ \times (U_{j+1,k} + U_{j-1,k} + U_{j,k+1} + U_{j,k-1} - 4U_{jk}) \quad (2)$$

или для действительных переменных  $V$  и  $W$  ( $U_{jk} = V_{jk} + iW_{jk}$ ):

$$dV_{jk}/dt = V_{jk} - (V_{jk} - \beta W_{jk})(V_{jk}^2 + W_{jk}^2) + e\Delta V_{jk} + ec\Delta W_{jk}, \quad (3)$$

$$dW_{jk}/dt = W_{jk} - (W_{jk} + \beta V_{jk})(V_{jk}^2 + W_{jk}^2) + e\Delta W_{jk} - ec\Delta V_{jk},$$

где  $\Delta x_{jk} = x_{j+1,k} + x_{j-1,k} + x_{j,k+1} + x_{j,k-1} - 4x_{jk}$ .

Поскольку аналитические результаты при исследовании уравнения (2) удается получить лишь для сравнительно небольших надкритичностей [8], компьютерное моделирование становится здесь основным инструментом исследования. Изучалась решеточная модель, описываемая уравнением (2) с периодическими граничными условиями  $U_{j+N,k} = U_{j,k+N} = U_{jk}$  ( $N \gg 1$ ) и граничными условиями равенства нулю потока:  $U_{N+1,k} = U_{N,k}$ ;  $U_{j,N+1} = U_{j,N}$ . Все параметры ( $e, \beta, c$ ) положительны:  $\beta$  — характеризует

нелинейную дисперсию,  $e$  — диффузионную связь внутри одной компоненты  $V$  или  $W$ , а  $ec$  — перекрестную диффузию между  $V$  и  $W$ . Исследуемый диапазон параметров моделирования представлен ниже:

Размер решетки ( $N \times N$ )	1) $128 \times 128$ 2) $256 \times 256$
Размер системы ОДУ	1) 32 768 2) 131 072
Интервал моделирования по времени $t$	$\approx 10^3$
Диапазон изменения параметра связи	$0 < e < 10^3$
Число различных начальных и граничных условий	$\approx 20$

Моделирование одного варианта с фиксированным набором параметров и характерным временным интервалом, определяемым различными инкрементами возникающих пространственных возмущений, занимает около 40 ч процессорного времени на универсальной ЭВМ ЕС-1037.

Использование одного МП позволило на этой задаче уменьшить время расчета варианта в 40 раз. Сравнительные результаты счета для различных версий программы представлены в табл. 1. Интегрирование проводилось методом Рунге—Кутты—Гилла четвертого порядка с переменным шагом. Максимальная производительность МКМП с одним МП на этой задаче равна 9 Мфлоп. Ускорение получено в результате использования функционального параллелизма МП, т. е. параллелизма на нижнем уровне интеграции — уровне единичных инструкций и данных.

Эффективное применение нескольких МП для совместного решения этой задачи обусловлено использованием решеточной модели. В вычислительной физике решеточные модели используются, во-первых, для разностной аппроксимации непрерывной среды, во-вторых, когда заранее существует информация о преобразовании среды в ансамбль связанных динамических элементов. В последнем случае именно решеточная модель и адекватна физической природе, однако и при пространственной аппроксимации непрерывной среды такая модель зачастую предпочтительней для компьютерного анализа. Одно из ее преимуществ при моделировании нелинейных полей состоит в том, что такой подход не предполагает заранее пространственный и временной характер возмущений, которые могут появляться при развитии различных неустойчивостей. Представление результатов моделирования в этом случае естественно и легко визуализируется. Эквивалентность индивидуальных элементов среды в решеточной модели и локальность их взаимодействия позволяют использовать внутренний параллелизм задачи. Применение МКМП, как мы увидим, дает существенный выигрыш, определяющий саму возможность полного моделирования.

Общая идея алгоритма состоит из следующих этапов:

1) происходит разделение решетки внутри ЦП на  $p$  равных горизонтальных слоев ( $p$  равно числу МП);

**Таблица 1**  
**Времена выполнения программы (мин) для метода**  
**Рунге—Кутты—Гилла**

Конфигурация системы	Астрономическое время	Время ЦП	Время МП	Конфигурация системы	Астрономическое время	Время ЦП	Время МП
Только ЦП	2420	2400	—	ЦП+МП	63	6	55
ЦП+МП	399	9	388	ЦП+10 МП	10	5	9

- 2) каждый из  $p$  слоев помещается в соответствующий МП;
- 3) выполняется интегрирование одновременно во всех МП на один шаг по времени;
- 4) происходит обмен границами слоев через ЦП;
- 5) осуществляется синхронизация событий в МКМП;
- 6) управление возвращается на третий этап до тех пор, пока не будет пройден весь временной интервал.

Несложно получить оценку производительности комплекса с  $p$  матричными процессорами на этой задаче. Общее время прохождения задачи на конфигурации с одним МП ( $p = 1$ ) определяется следующим выражением:  $T_{\Sigma}^1 = T_{\text{М.П}}^1 + T_{\text{ЦП}}^1 + T_{\text{К}}^1$ , здесь  $T_{\text{М.П}}^1$  — полное время работы программы внутри МП;  $T_{\text{ЦП}}^1$  — полное время работы ЦП;  $T_{\text{К}}^1$  — полное время, затраченное на пересылку данных между ЦП и МП. Соответствующее выражение для случая  $p$  матричных процессоров

$$T_{\Sigma}^p = T_{\text{М.П}}^p + T_{\text{ЦП}}^p + T_{\text{К}}^p$$

или

$$T_{\Sigma}^p = (T_{\text{М.П}}^1/p + 8T_{\text{М.П}}^1/N) + t_{in} [2(p - q) + 3] + t_{\text{К}} (2q). \quad (4)$$

где  $N$  — размер решетки;  $t_{in}$  — время инициации операции центральным процессором;  $q$  — количество МП на одном канале;  $t_{\text{К}}$  — время пересылки границы одного слоя между ЦП и МП ( $t_{\text{К}} \approx \alpha N/v_{\text{К}}$ ).

Параметры  $t_{in}$  и  $v_{\text{К}}$  (скорость передачи данных по каналу) являются константами системы МКМП, они не зависят от задачи и соответственно равны  $t_{in} = 5 \cdot 10^3$  с,  $v_{\text{К}} = 1$  Мбайт/с.

Следующее выражение определяет ускорение времени выполнения задачи на конфигурации с  $p$  МП по отношению к времени на системе с одним МП:

$$S(p) = \frac{T_{\Sigma}^1}{T_{\Sigma}^p} = \left\{ \left( \frac{1}{p} + \frac{8}{N} \right) + \frac{t_{in}}{T_{\text{М.П}}^1} [2(p - q) + 3] + \frac{t_{\text{К}}}{T_{\text{М.П}}^1} (2q) \right\}^{-1}. \quad (5)$$

Первые два члена в фигурных скобках определяют эффективность непосредственно параллельного алгоритма, третий и четвертый

учитывают затраты на управление и обмен соответственно. Эффективность алгоритма в нашем случае

$$E(p, N) = \frac{T_{\text{МП}}^1}{T_{\text{МП}}^p} = \frac{1}{1/p - 8/N}. \quad (6)$$

Полному распараллеливанию соответствует  $E(p, N) = p$ . Добавка  $(8/N)$  приводит к существенному снижению эффективности алгоритма уже при сравнительно небольших  $p$ ; так,  $E(8, 256) = 6,4$ . Появление добавки вызвано тем, что при разделении решетки на первом этапе на  $p$  кластеров размер каждого из них равен  $-N(N/p + 8)$ , а не  $N(N/p)$ , так как нужно дополнительно интегрировать по четыре слоя от двух соседних кластеров в силу особенности стандартного метода Рунге—Кутта четвертого порядка. Как следствие необходимо также обмениваться восемью граничными слоями между соседними кластерами на каждом шаге интегрирования.

Учет затрат на инициацию операций МКМП и на обмен границами приводит к следующим значениям относительного ускорения на задаче:

$$S_{N_1}(8) = 2,57, S_{N_2}(8) = 4,65, \text{ где } N_1 = 128, N_2 = 256.$$

Эти не слишком утешительные результаты вызваны прямым перенесением последовательного подхода и стандартного метода Рунге—Кутта на параллельную систему. Использование возможностей независимой работы системы на разных уровнях и адаптация к ней методов интегрирования качественно меняют ситуацию.

Модификация стандартной организации метода Рунге—Кутта четвертого порядка, при которой на каждом из четырех промежуточных шагов интегрирования отбрасывается по два использованных слоя, позволяет получить

$$E(p, N) = \frac{1}{1/p + 3/N}. \quad (6)$$

В этом случае сохраняются расходы на управление и по-прежнему пересылается по восемь слоев на каждом шаге интегрирования.

Другая возможность модификации основана на организации подкачки двух граничных слоев из соседних кластеров на каждом промежуточном шаге интегрирования. Тогда  $E(p, N) = p$  и на каждом целом шаге интегрирования суммарная пересылка также равна восьми слоям, но при этом существенно возрастают расходы на управление. Этот вариант использовался при интегрировании трехмерной модели (относительные расходы на управление здесь на 1,5 порядка меньше) при следующей дополнительной модификации алгоритма.

Обмен границами слоев (четвертый этап) производился во времени на фоне третьего этапа. Это достигается организацией встречного интегрирования в соседних кластерах, когда необходимые на

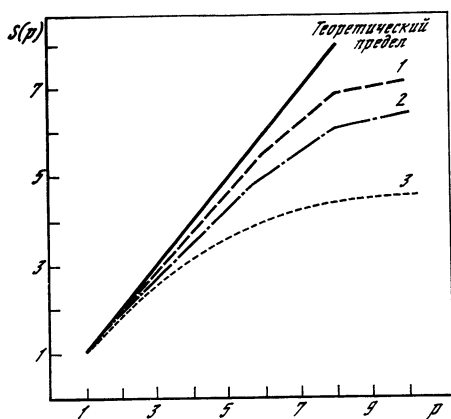


Рис. 1

$$\dot{S}(p) = \left\{ \frac{1}{p} + \frac{t_{in}}{T_{МП}^1} [3(p - q) + 3] + \frac{t_k}{T_{МП}^1} (2q) \right\}^{-1}; \quad (7)$$

функция  $S(p)$  имеет максимальное (при  $p = 16$ ) значение, равное  $S(p) = 10$  для  $N = 256$ .

Экстремальные значения относительного ускорения, полученные на МКМП для этой задачи, представлены на рис. 1 (1 —  $N = 256$ , МП Фортран,  $F = 14,2$  Мфлоп; 2 —  $N = 256$ , МП Ассемблер,  $F = 58,5$  Мфлоп; 3 —  $N = 128$ , МП Ассемблер,  $F = 36,0$  Мфлоп). Видно, что для различных конфигураций системы они находятся в хорошем соответствии с теоретическими оценками. На решетке  $N = 256$  при  $p = 10$  производительность комплекса достигает 58,5 Мфлоп. Время счета одного варианта сокращается при этом до 9 мин, т. е. действительно получаем новое качество, эквивалентное возможностям, предоставляемым вычислительной физике современными суперкомпьютерами.

Положительные результаты применения параллельных систем типа МКМП, по нашему мнению, определяются следующими основными причинами. Во-первых, распараллеливание проводилось на алгоритмическом уровне и на еще более высоком уровне постановки задачи. При этом оказалось, что многие задачи обладают на этом уровне «врожденным» (внутренним) параллелизмом, т. е. параллельны по своей природе [4, 6, 10]. Увидеть и использовать этот уровень можно при решении конкретных задач. Во-вторых, архитектура параллельных систем с присоединенными матричными процессорами и возможности ПО таких систем обладают достаточной гибкостью для реализации крупноблочного распараллеливания, более того, можно сказать, что именно ему такие системы и соответствуют.

Для большинства задач нелинейной физики относительное ускорение на системе зависит от  $T_{МП}^1$  таким образом, что чем более сложен решаемый алгоритм и чем больше данных им нужно обрабо-

следующем шаге интегрирования слою вычисляются первыми и пересылаются в соседние МП (кластеры) одновременно с продолжением счета. Здесь задействован механизм одновременной работы МП, ЦП и обмена данными между МП.

Учет баланса вкладов в управление и обмен значений реальных параметров системы при выборе алгоритма определили новое выражение относительного ускорения:

тать, тем более эффективно используется МКМП. Эта положительная обратная связь стимулирует постановку и дает возможность решения более сложных, например трехмерных, задач. Такая же связь существует в системе и по использованию памяти под задачу: большее число процессоров увеличивает с каждым из них память данных; доступную для трехмерных задач. Например, на трехмерной модели Гинзбурга—Ландау ожидаемая производительность комплекса равна 90 Мфлоп ( $S(p) = 9,8$  при  $p = 10$  и  $N^3 = 128^3$ ).

Различные МП и ЦП могут быть использованы в системе при неизменной идеологии математического обеспечения. Это обеспечивается, в частности, тем, что распараллеливание происходит на уровне не ниже, чем подпрограмма. Представляется маловероятным, что в ближайшее время трансляторы смогут распознавать используемый тип крупноблочного параллелизма, поэтому работа по распараллеливанию не должна быть слишком сложной.

В настоящее время на основе решаемых на МКМП задач создается набор независимых от используемого оборудования моделей параллельных прикладных алгоритмов, реализованных в идеологии МКМП. На базе таких наборов, используя экспертные системы, возможно создание средств автоматического выбора модели и отображения ее на реальную архитектуру. На основе анализа методов решения двумерной модели Гинзбурга—Ландау построена одна из таких базовых моделей — модель параллельного алгоритма решения больших систем ОДУ первого порядка. Интегрирование может проводиться методами Рунге—Кутты четвертого порядка, предиктора-корректора (метод Адамса—Маултона четвертого порядка) и Белирха—Стойера. Определены диапазоны значений параметров модели для оптимального использования на системах типа МКМП. Многие уравнения нелинейной физики в частных и обыкновенных производных сводятся к таким системам и могут решаться на основе этой модели.

### Некоторые результаты численного моделирования

При численном исследовании модели (1) типичными были значения параметров, представленные на с. 187. Наблюдалось исключительное разнообразие режимов структурообразования и хаотической динамики в ансамблях структур по мере уменьшения величины диффузионной связи  $e$ . Часть рассмотренных режимов отражена в табл. 2. Полностью результаты моделирования изложены в работе [9]. Здесь мы остановимся на режимах, при которых влияние элементов «среды» друг на друга мало, т. е. они становятся все более автономными (по мере уменьшения  $e$ ) (см. табл. 2).

Многие из наблюдаемых пространственных структур (в частности, при  $e < 0,25$ ) представляют собой периодические структуры типа кристаллов. Для их описания естественно использовать решения вида  $a_{jk} \approx \sum_{\nu} A_{\nu} \exp(\theta_{\nu j} + \theta_{\nu k})$  с синхронизованными фазами. Существенно упрощающим здесь обстоятельством является  $\tau_0$ ,



Таблица 2

## Режимы структурообразования и хаотической динамики

$e$	Режим
102	Устойчивое однородное равновесие: $ a_{jk} ^2 = 1$
80	Стационарные крупномасштабные структуры: валы, концентрические вихри, спиральные пары
35	Квазистационарные структуры: трансформирующиеся спиральные пары, модулированные концентрические вихри и т. д.
5	Хаотическая динамика структур — структурная турбулентность
0,25	Развитый пространственно-временный хаос высокой размерности — «турбулентность кипения»
$e^0 + 10^{-3}$	Развитый хаос, возникающий из регулярных решеток, иерархия последовательных возбуждений
$e^0 = 1/16$	Инфраструктуры, хаотическая динамика инфраструктур на фоне $\pi - \pi$ -колебаний
$> 1/16$	Устойчивые возбуждения в виде плоских волн, близких к $\pi - \pi$ -колебаниям

что, несмотря на наличие в уравнении (1) нелинейности, вследствие определенного ее вида ( $\sim |a_{jk}|^2$ ), а также дискретности и симметрии «среды» существуют точные решения (1), представляющие собой суперпозицию элементарных возбуждений. В частности, прямой подстановкой можно убедиться, что возможно решение в виде одномерной стоячей волны:

$$a_{jk}^0 = A^0 \exp(iw^0 t) \sin(\pi j/2); \quad |A^0|^2 = 1 - 2e;$$

$$w^0 = -\beta + 2e(\beta + c)$$

(а также решение, зависящее только от  $k$ ), которое является суперпозицией двух встречных волн вида

$$a_{jk}^0 = \exp(iw^0 t) \frac{A^0}{2} \left[ \exp\left(i \frac{\pi}{2} j\right) - \exp\left(-i \frac{\pi}{2} j\right) \right]$$

и двумерной стоячей волны (квадратная решетка)

$$a'_{jk} = A' \exp(iw't) \sin \frac{\pi j}{2} \sin \frac{\pi k}{2}, \quad |A'|^2 = 1 - 4e,$$

$$w' = -\beta + 4e(\beta + c),$$

являющейся суперпозицией четырех волн:

$$a'_{jk} = \exp(iw't) \frac{A}{4} \left\{ \exp\left[i \frac{\pi}{2} (j+k)\right] + \exp\left[i \frac{\pi}{2} (j-k)\right] - \right.$$

$$\left. - \exp\left[-i \frac{\pi}{2} (j+k)\right] - \exp\left[-i \frac{\pi}{2} (j-k)\right] \right\}.$$

Кроме них, имеются еще решения, содержащие и другие волны. Все эти решения неустойчивы (поскольку неустойчива каждая из

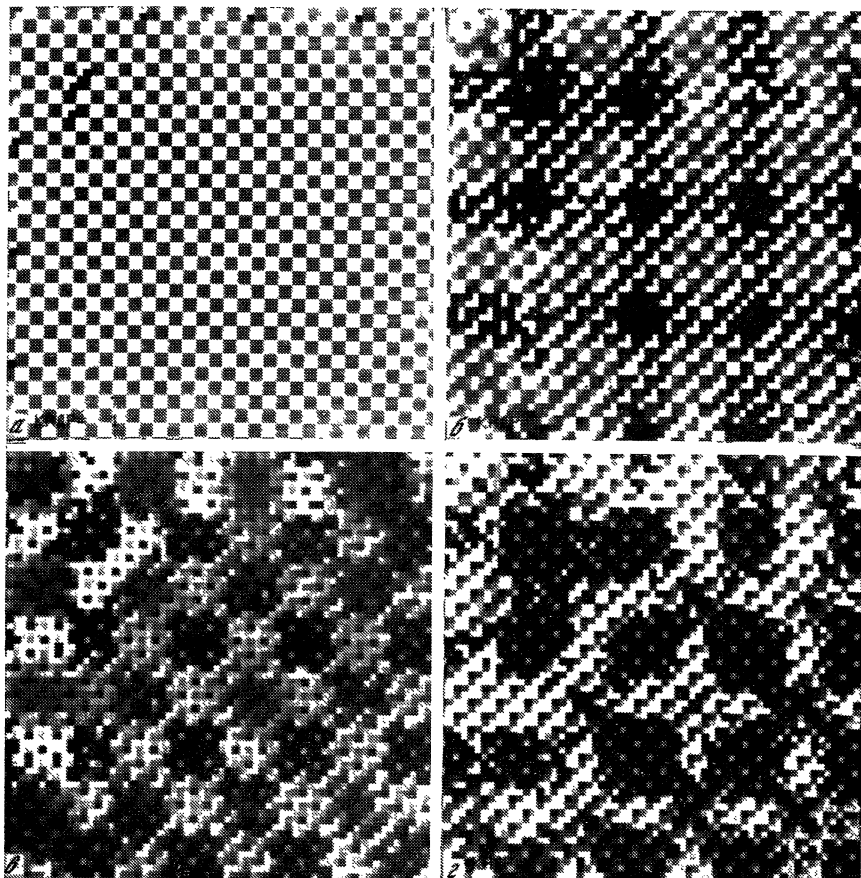


Рис. 2

плоских волн, образующая их), однако разрушение этих образований происходит весьма своеобразно.

При  $0 < e < e^0 = 1/16$  эти решения превращаются под действием слабых возмущений в коротковолновые ( $\theta_m, \theta_n \approx \pi$ ) устойчивые решения. Однако при  $e > e^0$  картина уже существенно иная: вначале на фоне первичной решетки устанавливается длинноволновая продольная модуляция (рис. 2, а), затем к ней добавляется еще и поперечная. В результате этого образуется решетка модуляции — «сверхрешетка», представленная на рис. 2, б. Но и эта структура неустойчива — на фоне ее возникает новая решетка с другим (в общем случае несоизмеримым) пространственным периодом (рис. 2, в). Такой процесс может продолжаться и далее, в результате чего происходит возбуждение всех пространственных масштабов и осуществляется переход к развитому пространственно-временному хаосу (рис. 2, г). Подобный переход к хаосу во времени напоминает переход к турбулентности на поверхности параметриче-

ски возбуждаемой жидкости при изменении надкритичности [11].

При дальнейшем уменьшении  $e$ , начиная со значения  $e^0 = 0,0625$ , ситуация качественно изменилась — пространственная картина стала двухмасштабной. На фоне мелкомасштабного пространственного заполнения в виде  $\pi - \pi$ -колебаний решетки возникли плавные структуры модуляции — инфраструктуры. Замечательно, что весь сценарий развития хаоса, наблюдающийся для самих полей (со стороны больших  $e$ ) в очень узком диапазоне значений  $e$  ( $|e - 0,0625| \approx 10^{-3}$ ), полностью воспроизводится, но уже для инфраструктур.

По мере уменьшения величины связи между индивидуальными элементами решетки все более существенными становятся эффекты, связанные с дискретностью «среды» (автономность элементов при уменьшении  $e$  возрастает). При очень слабой связи  $e$  ( $e < e^0 = 1/16$ ) в уравнении (1) существуют и устойчивы вплоть до  $e = 1/16$  коротковолновые возбуждения с  $\theta_m, \theta_n \approx \pi$ . Очевидно, что таким решениям нет аналогов в непрерывном уравнении Гинзбурга—Ландау, поскольку в данном случае соседние элементы совершают противофазные колебания.

При  $e > 1/16$   $\pi - \pi$ -колебания теряют устойчивость, и в дискретной среде на их фоне, как будет показано, реализуются достаточно плавные структуры. Поскольку дискретное уравнение Гинзбурга—Ландау было первоначально представлено для двумерного ансамбля слабо взаимодействующих структур, то естественно назвать такие плавные вторичные структуры на фоне мелкомасштабного пространственного заполнения (противофазно колеблющихся элементов ансамбля) инфраструктурами.

Для анализа инфраструктур рассмотрим решение, близкое к колебаниям:

$$a_{jk} = A_0 \exp [i\pi (j + k) + i\Omega_0 t] + \exp (i\Omega_0 t) \zeta_{jk}, \quad (8)$$

где  $|A_0|^2 = 1 - 8e$ ,  $\Omega_0 = -\beta + 8e(\beta + c)$ .

Для возмущений  $\zeta_{jk}$  на фоне таких колебаний получаем уравнение

$$\begin{aligned} \zeta_{jk} = & [1 - i\Omega_0 - 2(1 + i\beta) |A_0|^2] \zeta_{jk} - (1 + i\beta) A_0^{2*} \zeta_{jk}^* + \\ & + e(1 - ic) \Delta^2 \zeta_{jk} - (1 + i\beta) [\zeta_{jk} | \zeta_{jk} |^2 + \\ & + \exp [i\pi (j + k)] (2 | \zeta_{jk} |^2 + \zeta_{jk}^2)]. \end{aligned} \quad (9)$$

В уравнении (9)  $\pi - \pi$ -колебаниям отвечает тривиальное равновесие  $\zeta_{jk} = 0$ . Соответствующие ему ляпуновские показатели равны

$$\begin{aligned} \lambda_{mn} = & 8e - 1 + 4e [\cos^2 (\theta_n/2) + \cos^2 (\theta_m/2)] \pm D_{mn}, \\ D_{mn}^2 = & |A_0|^4 - \{4ec [\cos^2 (\theta_n/2) + \cos^2 (\theta_m/2)]\}^2 - \\ & - 8e\beta c [\cos^2 (\theta_n/2) + \cos^2 (\theta_m/2)]. \end{aligned} \quad (10)$$

Из выражения (10) легко получить, что при  $\beta c > 1$  и  $e < 1/16$

$\pi - \pi$ -колебания асимптотически устойчивы. При  $e > 1/16$  пара комплексно-сопряженных корней, отвечающих пространственно-однородному возмущению с  $\theta_m = \theta_n = 0$  проходит через мнимую ось, т. е. реализуется бифуркация Хопфа. При этом на фоне  $\pi - \pi$ -колебаний возбуждаются пространственно-однородные биеения. В фазовом пространстве системы таким квазипериодическим осцилляциям соответствует незамкнутая намотка на двумерном торе. При дальнейшем увеличении  $e$  последовательно возбуждаются уже неоднородные длинноволновые возмущения. При  $e \approx 1/16$  для комплексных амплитуд таких возбуждений можно получить уравнение вида (1). Для этого представим решение (9) в виде

$$\zeta_{jk} = (1 + iB) \exp(i\Omega t) C_{jk} + (1 + iB^*) \exp(-i\Omega t) C_{jk}^* \quad (11)$$

где  $C_{jk}$  будем считать медленной функцией времени и дискретных координат  $j, k$ ;  $B$  — компонента собственного вектора, соответствующая корню  $\lambda_0 = \alpha + i\Omega$ ,  $\theta_m = \theta_n = 0$ . При этом  $\alpha = 16e - 1$ ,  $\Omega^2 = 64(ec)^2 + 16e\beta c - (8e - 1)^2 \approx \beta c + c^2/4 - 1/4$ ,  $e \approx 1/16$ .

Легко проверить, что для  $\pi - \pi$ -колебаний собственные векторы сопряженной системы совпадают с комплексно-сопряженными собственными векторами. Используя это обстоятельство и подставляя (8) в (9), после усреднения получим следующие уравнения для медленной амплитуды:

$$\begin{aligned} \dot{C} &= (16e - 1) C_{jk} - 2(1 + |B|^2) \left[ 1 + i\beta \left( \frac{3}{2} T - \frac{1}{2T} \right) \right] \times \\ &\times C_{jk} |C_{jk}|^2 + e(1 - ic) \Delta C_{jk}. \end{aligned} \quad (12)$$

После замен

$$\begin{aligned} b_{jk} &= \left[ \frac{2(1 + |B|^2)}{16e - 1} \right]^{1/2} C_{jk}, \quad e' = \frac{e}{16e - 1} \gg 1, \\ C' &= C, \quad \beta' = \beta \left( \frac{3}{2} T - \frac{1}{2T} \right); \quad t' = t(16e - 1) \end{aligned}$$

для переменной  $b_{jk}$  получаем исходное уравнение (1), но с масштабно преобразованными параметрами  $e', \beta', t'$ . Как видно, в полученном уравнении для огибающей значение параметра  $e$  уже очень велико. Таким образом, для пространственно-временной модуляции на фоне  $\pi - \pi$ -колебаний действительно справедливо длинноволновое описание, что и объясняет подобие инфраструктур структурам поля при больших  $e$ . Рис. 3 иллюстрирует это — в области очень малых  $e$  мы как бы наблюдаем повторение на фоне  $\pi - \pi$ -колебаний всего сценария развития и усложнения практически тех же пространственных образов, которые развиваются со стороны очень больших  $e$  из пространственно-однородного режима.

Плавные инфраструктуры существуют в очень узкой области изменения параметра  $e$ . По мере увеличения  $e$  через мнимую ось проходят все новые пары собственных значений и последовательно накапливаются бифуркации Хопфа, т. е. возникают торы достаточно высокой размерности. Поскольку (из-за малости  $e$ ) базисные частоты этих торов очень близки, благодаря действию нелинейнос-

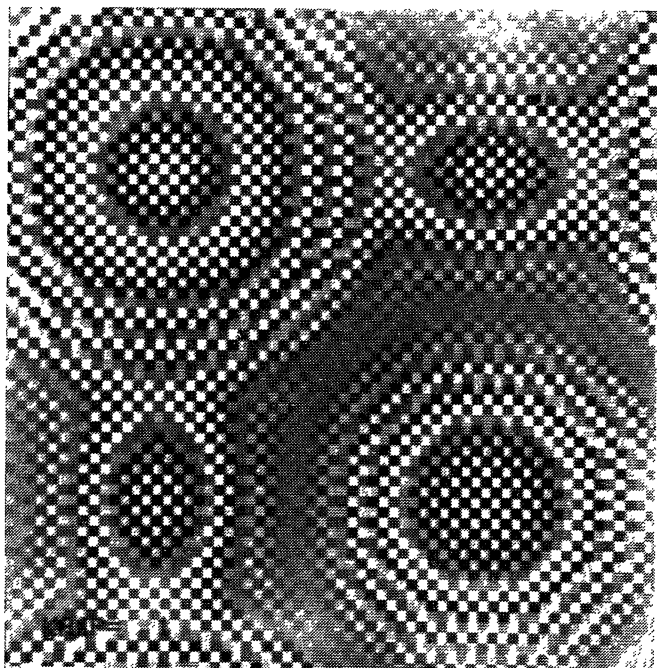


Рис. 3

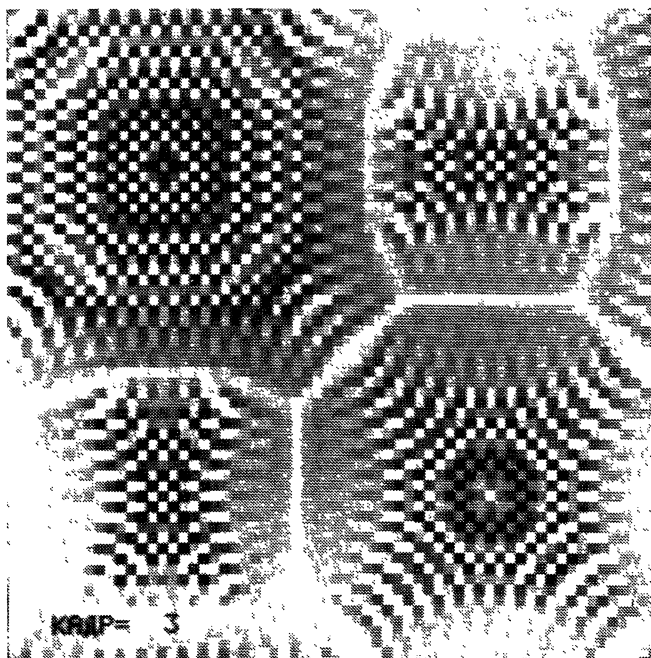


Рис. 4

ти торы разрушаются, что, как демонстрирует численное моделирование, достаточно быстро приводит к возникновению развитого пространственно-временного хаоса очень высокой размерности [9]. Оценкой области существования инфраструктур может служить интервал  $\Delta e$  между двумя последовательными бифуркациями Хопфа:  $\Delta e = e_{12} - e^0$ , где  $e^0$  — бифуркационное значение для возмущения с  $\theta_m = \theta_n = 0$ , а  $e_{12} - \theta_n = 0$ ,  $\theta_m = 2\pi/N$ . Из (10) легко получить, что

$$e_{12} = \frac{1}{12 + 4 \cos(2\pi/N)} \approx e^0 \left( 1 + \frac{\pi^2}{4N^2} \right).$$

Отсюда получаем, что искомый интервал

$$\Delta e = \pi^2/64N^2 \approx o(1/N^2) \ll 1.$$

Очевидно, что в ансамблях с большим числом элементов эта величина очень мала. Например, в ансамбле из  $64 \times 64$  элементов эта область составляет интервал  $4 \cdot 10^{-5}$ , т. е. зарегистрировать инфраструктуры можно лишь в специальном численном эксперименте, который и был поставлен (см. рис. 3 и 4).

### ЗАКЛЮЧЕНИЕ

Проведенный анализ динамики нелинейного поля, описываемого двумерным уравнением Гинзбурга—Ландау (более полный анализ см. в работе [9]), представляет собой синтез аналитического и компьютерного исследований. Подобный подход позволил, с одной стороны, обнаружить при компьютерном моделировании новые, подчас неожиданные явления (например, инфраструктуры на фоне мелкомасштабных решеток и подобие их превращений взаимным трансформациям структур самого поля), с другой стороны, использовать аналитические результаты (в частности, определение точек бифуркаций, опирающихся на вычисления ляпуновских показателей) для осмысления полученных в численном эксперименте эффектов и построить, таким образом, достаточно полную качественную теорию сложного поведения нелинейного двумерного поля.

Параллельный подход в компьютерном моделировании, использующий несколько механизмов (и путей) управления, связанных лишь конечным результатом воздействия, многоуровневая система связей локальной и распределенной памяти, а также подобие применяемых подходов на разных уровнях интеграции соответствуют нелинейной природе решаемых физических задач. В компьютерной науке существует термин «линейное программирование», для параллельных систем, естественно, — «нелинейное программирование», которое дает новое качество при моделировании задач нелинейной физики.

### ЛИТЕРАТУРА

1. Bucher I. Y. The computational speed of supercomputers // Proc. ACM Sigmetrics Conf. Measur. and Model. Comput. Syst. 1983.
2. Hockney R. W. Parallel computers: Architecture, programming and algorithms. Bristol: Hilger, 1981.

3. Wallach Y. Alternating sequential parallel processing // Lect. Notes Comput. Sci. 1982. Vol. 124.
4. Clementy E., Detrich J. Large scale computations on a loosely coupled array of processors // IBM Techn. Rep. KGN-64, 1986. May 28.
5. Велихов Е. П., Лазаров Вл., Марков Ст. и др. Решение задач нелинейной физики на многопроцессорном комплексе ЕС-1037—ЕС-2706: Препр. ИКИ АН СССР № 1169. М., 1986.
6. Многопроцессорный вычислительный комплекс ЕС-1037, ЕС-2706: Опыт системных разработок и численного моделирования нелинейных физических задач ИКИ АН СССР / Под ред. Р. З. Сагдеева, В. И. Шевченко. М., 1987.
7. Chochia G. A., Gribov B. E., D'yachenko A. I. et al. Some results of nonlinear physical problem simulation using the multy array processor system: Prepr. IKI Acad. Sci. USSR. Moscow, 1987.
8. Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И. Развитие хаоса в ансамблях динамических структур // ЖЭТФ. 1986. Т. 89, № 1. С. 92—105.
9. Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И. и др. Решеточные модели в нелинейной динамике неравновесных сред: Препр. ИПФ АН СССР № 163. Горький, 1987.
10. Clementi E., Chin S., Corongiu G. et al. Large-scale computations on a scalar, vector, and parallel «supercomputer» and global AB INITIO simulations with a study on liquid water as an example, IBM DSD, KGN-86. 1986.
11. Езерский А. Б., Рабинович М. И., Реутов В. П., Старобинец И. М. Пространственно-временной хаос капиллярной ряби при параметрическом возбуждении жидкости // ЖЭТФ. 1986. Т. 91, № 6(12). С. 2070—2081.

УДК 537.534

## ЛАЗЕРНО-ИНДУЦИРОВАННЫЕ НЕУСТОЙЧИВОСТИ И ФАЗОВЫЕ ПЕРЕХОДЫ НА ПОВЕРХНОСТИ ТВЕРДЫХ ТЕЛ И ОБРАЗОВАНИЕ УПОРЯДОЧЕННЫХ СТРУКТУР

В. И. ЕМЕЛЬЯНОВ

1. При действии достаточно мощного лазерного излучения с длиной волны  $\lambda_L$  на сильно поглощающие среды на их поверхности образуются сложные упорядоченные структуры. Причиной их возникновения являются лазерно-индуцированные неустойчивости и фазовые переходы на поверхности. Развитие этих неустойчивостей в течение длительности лазерного импульса  $\tau_p$  может радикально изменять процесс взаимодействия излучения с веществом. В настоящей работе дан обзор двух новых классов таких неустойчивостей: интерференционных неустойчивостей (ИН) рельефа поверхности и диффузионно-деформационных неустойчивостей (ДДН).

2. ИН наблюдаются при действии мощных (с интенсивностью  $I \approx 10^7 \div 10^{10}$  Вт/см<sup>2</sup>) коротких ( $\tau_p \approx 10^{-8} \div 10^{-11}$  с) лазерных импульсов на поверхность металлов, полупроводников и диэлектриков. В результате развития ИН после окончания импульса на поверхности остаются поверхностные упорядоченные структуры (ПУС): одномерные, двумерные решетки рельефа поверхности и

более сложные модуляции рельефа с периодом  $d \approx \lambda_L$ . Амплитуды ПУС, существующих в течение  $\tau_p$ , могут значительно превосходить амплитуды этих необратимых ПУС [1].

Геометрия ПУС, образующихся при ИН, характеризуется зависимостью от направления проекции вектора напряженности в падающей волне  $\mathbf{E}_i$  на плоскость поверхности ( $\mathbf{E}_t$ ) и диэлектрической проницаемости среды  $\varepsilon(\omega) = (n + im)^{1/2}$ . В случае  $|\varepsilon| \gg 1$  доминируют решетки с волновым вектором  $\mathbf{q}$ , параллельным  $\mathbf{E}_t$  (продольные ПУС). В случае  $|\varepsilon| \approx 1$  доминируют поперечные ПУС ( $\mathbf{q} \perp \mathbf{E}_t$ ). С ростом числа импульсов доминирующая решетка вследствие нелинейности (см. ниже) стабилизируется, а амплитуда более слабой решетки продолжает нарастать. В результате образуются двумерные решетки (см. обзор [2]). При  $s$ -поляризации падающей волны в случае  $|\varepsilon| \gg 1$  продольная решетка имеет период  $d \approx \lambda_L / \cos \theta$ , а при  $p$ -поляризации образуются две продольные решетки с  $d \approx \lambda_L / (1 \pm \sin \theta)$ , где  $\theta$  — угол падения излучения.

Физический механизм ИН, приводящих к образованию ПУС, состоит из трех этапов. Падающая волна с частотой  $\omega = 2\pi c/\lambda_L$  дифрагирует на затравочной фурье-компоненте рельефа поверхности  $z = z_q(t) \exp(i\mathbf{q}\mathbf{r} + i\lambda''t)$ ,  $\mathbf{r} = \{x, y\}$ , немодулированная плоскость поверхности совпадает с плоскостью  $z = 0$  (ось  $z$  направлена в глубь среды). В результате возникают две приповерхностные дифрагированные волны: стоксова с волновым вектором  $\mathbf{k} = \mathbf{k}_t - \mathbf{q}$  ( $\mathbf{k}_t$  — проекция волнового вектора падающей волны на плоскость поверхности) и частотой  $\omega_s = \omega - \lambda''$  и антистоксова волна с  $\mathbf{k}_a = \mathbf{k}_t + \mathbf{q}$  и  $\omega_a = \omega + \lambda''$ . Благодаря поверхностному электромагнитному резонансу (см. [2]) наибольшие амплитуды среди всех дифрагированных волн имеют волны с  $|\mathbf{k}_{s,a}| \sim |\mathbf{k}| = \omega/c$ . Выделение же направлений векторов  $\mathbf{k}_a$  и  $\mathbf{k}_s$  (и, следовательно, направления  $\mathbf{q}$ ) происходит благодаря специфической зависимости амплитуд дифрагированных волн от направления их распространения по отношению к  $\mathbf{k}_t$ , также благодаря возможности одновременного вхождения в резонанс волн  $\mathbf{k}_s$  и  $\mathbf{k}_a$  (см. [2]).

Дифрагированные волны интерферируют с прошедшей в среду волной накачки и создают интерференционное поле в приповерхностном слое. На этом заканчивается первый (электродинамический) этап. Благодаря оптическому поглощению это поле создает пространственно-периодическое (вдоль поверхности) распределение температуры  $T(\mathbf{r}) \sim \exp(i\mathbf{q}\mathbf{r} + i\lambda''t)$  (второй — тепловой этап). Возникающие вследствие этого силы приводят к нарастанию затравочных модуляций рельефа и развитию ИН. В зависимости от величины лазерной интенсивности обратную связь может замыкать термоупругая сила, приводящая к генерации поверхностных акустических волн (ПАВ), термокапиллярная сила и давление отдачи паров, вызывающие генерацию капиллярных волн в поверхностном расплаве или непосредственный вынос массы вещества при испарении (см. [2]). При быстром остывании модулированной поверхности расплава образуются необратимые ПУС. В линей-



ной теории ИН, включающей в описание все три этапа образования ПУС, были получены выражения инкрементов  $\lambda'$  нарастания амплитуд ( $z_q(t) \sim \exp \lambda' t$ ) в зависимости от  $q$ ,  $I$ ,  $\theta$  и параметров вещества.

Отметим, что значительная часть теоретических работ, посвященных ПУС, ограничена рассмотрением лишь электродинамического этапа [3], при этом геометрия доминирующих решеток определяется по максимумам тепловых интерференционных источников. При  $|\varepsilon| \gg 1$  параметры ПУС, определенные по максимумам  $\lambda' = \lambda'(q)$  [2], совпадают с соответствующими параметрами, получаемыми в электродинамической модели. Это совпадение объясняется тем, что при  $|\varepsilon| \gg 1$  поверхностный электромагнитный резонанс весьма интенсивен и узок и его положение на оси  $q$  определяет положение максимума  $\lambda'$ . Однако в случае  $|\varepsilon| \approx 1$  (например, кварца) предсказания двух этих вариантов рассмотрения качественно различны, причем согласия с экспериментом удастся достичь лишь при использовании выражений для  $\lambda'$  [4].

Выход на стационарный режим генерации ПУС обусловлен нелинейностью дифракции на модуляции рельефа с большой амплитудой [5, 6]. Вычисление амплитуд ПУС в стационарном режиме генерации капиллярных волн проведено в работе [7]. При нормальном падении

$$a \equiv 4z_q = a_{\text{опт}} [2(I/I_{\text{кр}} - 1)]^{1/2}, \quad I_{\text{кр}} \approx 10^6 \text{ Вт/см}^2,$$

$$a_{\text{опт}} = \lambda_L [2n/\pi^2(m^2 + n^2)]^{1/2} \approx \lambda_L/10.$$

Одним из наиболее интересных проявлений ИН является эффект самоиндуцированного возрастания поглощательной способности поверхности до  $A \approx 1$  (от значений  $A \approx 10^{-1} \div 10^{-2}$  для металлов), обусловленный резонансной перекачкой энергии в дифрагированные приповерхностные волны [5, 8]. Максимальное значение  $A$  достигается тогда, когда амплитуда поверхностного рельефа, возникающего при развитии ИН, достигает оптимального значения  $a_{\text{опт}}$ . Как мы видели, такая ситуация вполне вероятна, и при этом должно происходить резкое увеличение энерговклада при лазерном воздействии на поверхность.

3. Ко второму типу неустойчивостей относятся ДДН, которые могут иметь место в широком классе систем с диффузионной переменной (концентрацией неравновесных носителей или температурой в полупроводниках, концентрацией вакансий, дислокаций, пор или аморфной фазы в кристаллизующейся среде), связанной с деформациями упругого континуума. ДДН представляют собой новый класс диссипативных неустойчивостей, при которых доминирующие структуры определяются конкуренцией распределения по модам внешней накачки и диссипации. Геометрия соответствующих ПУС не зависит от поляризации лазерного поля и не определяется непосредственно значением  $\lambda_L$ . Ниже приводятся примеры как эндотермических ДДН, протекающих за счет лазерной энергии, так и экзотермических ДДН, развивающихся благодаря скры-

той теплоте, запасенной в среде. Симметрия возникающих в результате ДДН поверхностных ПС определяется либо симметрией лазерного поля, либо симметрией среды.

Прототипом ДДН является электронно-деформационно-тепловая неустойчивость (ЭДТН) при межзонных переходах в полупроводниках [9—11], развивающаяся следующим образом. Флуктуации поверхностной деформации  $\xi = \text{div } \mathbf{u}$  ( $\mathbf{u}$  — вектор смещения среды), а также концентрация неравновесных носителей  $n_1$  и температуры  $T_1$  от средних значений  $n_0$  и  $T_0$  пространственно модулируют ширину запрещенной зоны:

$$E_g = E_{g0} + \theta\xi - \beta_n n_1 - \beta_T T_1; \quad \beta_n \equiv |\partial E_g / \partial n|; \quad \beta_T \equiv |\partial E_g / \partial T|, \quad (1)$$

где  $\theta$  — потенциал деформации. Это приводит к модуляции мнимой части диэлектрической проницаемости среды:

$$\varepsilon'' = \varepsilon_0'' + \frac{\partial \varepsilon_0''}{\partial \omega} \frac{1}{\hbar} (-\theta\xi + \beta_n n_1 + \beta_T T_1). \quad (2)$$

Здесь  $\partial \varepsilon_0'' / \partial \omega$  — производная, взятая на частоте  $\omega$  лазерного поля

$$\mathbf{E}(\mathbf{r}, z, t) = \mathbf{E} \exp(-\gamma z / 2 + i\omega t) + \text{к. с.}$$

Коэффициент оптического поглощения  $\gamma = \gamma_0 + \gamma_1$ . Модуляция  $\gamma$  вызывает дополнительную модуляцию  $n_1$  и  $T_1$  в соответствии с уравнением

$$\partial Y_j / \partial t + \Delta_j Y_j = (\varepsilon_{j\xi} \xi + \varepsilon_{jT} T_1 + \varepsilon_{jn} n_1) (1 - \gamma_0 z) e^{-\gamma_0 z}, \quad (3)$$

где  $j = n, T$ ;  $Y_T = T_1$ ;  $Y_n = n_1$ ;  $\Delta_j = -\chi_j \Delta + \tau_j^{-1}$ ;  $\Delta$  — оператор Лапласа;  $\chi_T = \chi$  — коэффициент температуропроводности;  $\chi_n = D$  — коэффициент диффузии носителей;  $\tau_T^{-1} = 0$ ;  $\tau_n^{-1} = \tau^{-1} + \gamma_{\text{оже}} n^2$ ;  $\tau$  — время линейной рекомбинации;  $\gamma_{\text{оже}}$  — константа нелинейной оже-рекомбинации. Коэффициенты связи

$$\varepsilon_{j\alpha} = (\partial \varepsilon_0'' / \partial \omega) E^2 \sigma_{j\alpha} / 2\pi \hbar, \quad \sigma_{n\xi} = -\theta / t, \quad \sigma_{nT} = \beta_T / \hbar,$$

$$\sigma_{T\xi} = -\theta \omega / C_V, \quad \sigma_{nn} = \beta_n / \hbar, \quad \sigma_{TT} = \omega \beta_T / C_V, \quad \sigma_{Tn} = \beta_n,$$

где  $C_V$  — теплоемкость единицы объема. Граничные условия имеют вид

$$(\partial Y_j / \partial z)_{z=0} = 0. \quad (4)$$

Флуктуации  $n_1$  и  $T_1$  обуславливают появление сил в уравнении для  $\mathbf{u}$

$$\partial^2 \mathbf{u} / \partial t^2 = c_{it}^2 \Delta \mathbf{u} + (c^2 - c_i^2) \text{grad div } \mathbf{u} + \sum_{j=n, T} f_j \text{grad } Y_j, \quad (5)$$

где  $c_{i,t}$  — продольная и поперечная скорости звука;  $f_n = \theta / \rho$ ;  $f_T = -K\alpha / \rho$ ,  $\rho$  — плотность среды;  $K$  — модуль всестороннего сжатия;  $\alpha$  — коэффициент объемного расширения. В случае одно-

мерной геометрии ПУС граничные условия имеют вид

$$\left(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z}\right)_{z=0} = 0; \left(\sum_{j=n, T} \frac{f_j Y_j}{c_l^2} + \frac{\partial u_z}{\partial z} + (1 - 2\beta) \frac{\partial u_x}{\partial x}\right)_{z=0} = 0, \\ \beta \equiv \frac{c_t^2}{c_l^2}. \quad (5a)$$

Возникающая благодаря действию сил добавочная деформация увеличивает затравочную флуктуацию, что при превышении критической интенсивности  $I_{кр}$  приводит к развитию ЭДТН. При этом с инкрементом  $\lambda'$  нарастают во времени фурье-амплитуды поля упругого смещения среды

$$u_\alpha = N_q a_\alpha \exp(-\kappa_t z) + M_q [b_\alpha \exp(-\kappa_l z) + \lambda^2 \sum_{j=n, T} C_{\alpha j} \Phi_j \exp(-\delta_j z)] \exp(iqx + \lambda t), \quad (6)$$

где  $M_q$  и  $N_q$  — константы;  $a_z = iq$ ;  $b_z = -\kappa_l$ ;  $C_{zj} = -1$ ;  $a_x = \kappa_t$ ;  $b_x = iq$ ;

$$C_{xj} = iq/\delta_j; \quad \kappa_{l,t} = (q^2 + \lambda^2/c_{l,t}^2)^{1/2}; \quad \delta_j = [q^2 + (\lambda + \tau_j^{-1})/\chi_j]^{1/2}; \\ \Phi_j = R_j / \{\chi_j \gamma_0 [\lambda^2 + c_l^2 (q^2 - \delta_j^2)] (1 - \sum_{j=n, T} \varepsilon_{jj} / \chi_j \gamma_0 \delta_j) - \\ - \sum_{j=n, T} R_j c_l^2 (q^2 - \delta_j^2) / \delta_j\}; \quad R_j = -f_j \varepsilon_j \xi / c_l^2,$$

диффузионного поля

$$Y_j = A_j \varepsilon_j \xi \exp(iqx - \delta_j z + \lambda t) \quad (7)$$

и поля поверхностной деформации

$$\xi = A_q \exp(iqx + \lambda t), \quad A_q \sim M_q. \quad (8)$$

В случае аксиальной симметрии лазерного поля в (6) производится замена  $\exp iqx \rightarrow J_m(qr) \cos m\varphi$ ,  $J_m$  — функция Бесселя  $m$ -го порядка (точка  $r = 0$  совпадает с центром лазерного пятна). Тогда при ЭДТН возникают связанные радиально-кольцевые поля деформаций и диффузионной переменной, имеющие конфигурацию концентрических колец в центре и радиальных лучей на периферии («солнце»). Амплитуды этих полей, как и в случае одномерных решеток (6)–(8), затухают в глубь среды на расстояниях порядка  $\delta_j^{-1}$ ,  $\kappa_{l,t}^{-1}$ .

Используя (6) в граничных условиях (5a), получаем два однородных алгебраических уравнения относительно  $M_q$  и  $N_q$ , откуда следует дисперсионное уравнение, определяющее зависимость  $\lambda$  от  $q$ :

$$4q^2 \kappa_t \kappa_l - (\kappa_t^2 + q^2)^2 = \\ = \lambda^2 \sum_{j=n, T} \frac{R_j [4q^2 \kappa_t \delta_j - (\kappa_t^2 + q^2)^2]}{\chi_j \gamma_0 \delta_j [\lambda^2 + c_l^2 (q^2 - \delta_j^2)] - \lambda^2 \sum_{j=n, T} \varepsilon_{jj} - \sum_{j=n, T} (R_j + \varepsilon_{jj}) (q^2 - \delta_j^2) c_l^2}. \quad (9)$$

В левой части (9) фигурирует так называемый рэлеевский определитель, задающий частоты свободных ПАВ; правая часть (9) описывает действие лазерной накачки ( $R_j \sim E^2$ ).

Это уравнение имеет три типа решений, соответствующих трем качественно различным типам ЭДТН: 1) при дополнительном учете вязкости в (5) дисперсионное уравнение (9) описывает вынужденную генерацию ПАВ:  $\lambda = \lambda' + i\lambda''$ ;  $\lambda', \lambda'' > 0$ , критическая интенсивность  $I_{кр} \approx 10^4$  Вт/см<sup>2</sup> (см. [11, 12]); 2) размягчение частот ПАВ при действии лазерной накачки:  $\lambda = \lambda' + i\lambda''$ ,  $\lambda'' \rightarrow 0$ ,  $\lambda' < 0$  [10]; 3) генерация статических ППС. Положив в (9)  $\lambda'' = 0$  и обозначая  $\lambda' \equiv \lambda$ , получаем для инкремента

$$\lambda = \chi_j \left\{ \left[ \frac{(q + q_{R\epsilon})^2}{4} + qq_{R\epsilon} \frac{2\beta}{1-\beta} \right]^{1/2} + \frac{q_{R\epsilon} - q}{2} \right\}^2 - \chi_j q^2 - \tau_{j0}^{-1}, \quad (10)$$

где  $q_R = R_j / \chi_j \gamma_0$ ;  $q_{R\epsilon} = (R_j + \epsilon_{jj}) / \chi_j \gamma_0$ ;  $\tau_{j0}^{-1} = \tau^{-1} + \gamma_{0\text{же}} n_0^2$ .

Инкремент достигает максимального значения  $\lambda_m$  в точке  $q_m$ , причем

$$q_m = q_{R\epsilon} [(5a^2 + 6a + 1)^{1/2} - (2a + 1)], \quad a = [(2\beta/(1 - \beta))] \times (q_R/q_{R\epsilon}), \quad (11)$$

$$\lambda_m = 1/2 \chi_j q_{R\epsilon}^2 [(5a + 1)(5a^2 + 6a + 1)^{1/2} - 11a^2 - 8a + 1].$$

При  $\epsilon_{jj} \ll R$  максимум достигается при  $q = q_m \approx 0$  и период доминирующей ПУС равен  $d = 2\pi/q_m$ . При  $\epsilon_{jj} \gg R_j$  максимум достигается в точке  $q = 0$ . Однако и в этом случае могут выделяться ПУС с  $q \approx 0$ . Действительно, результирующая деформация к концу импульса ( $t \geq \tau_p$ ) с учетом (8) равна  $\xi(x) = \sum_q A_q \exp(iqx + \lambda\tau_p)$ .

Отсюда для коррелятора деформации имеем  $\langle \xi^2(x) \rangle = \sum_q \langle A_q^2 \rangle \exp(2\lambda\tau_p) \sim \int_0^\infty dq \xi_q^2$ . Затравочные статические деформации имеют радиус корреляции  $r_{кор} \ll d$ , поэтому можно положить  $\langle A_q^2 \rangle = \text{const}$ . Тогда спектральная функция деформаций на поверхности  $\xi_q^2 = q \exp(2\lambda\tau_p)$ . Из (10) следует, что при  $\epsilon_{jj} \gg R_j$  инкремент  $\lambda = \chi_j (q_{R\epsilon}^2 - q^2 - \tau_{j0}^{-1}/\chi_j)$ . С учетом этого получаем, что максимум  $\xi_q^2$  достигается в точке  $q_m = (4\tau_p \chi_j)^{-1/2}$ , при этом период ПУС  $d = 2\pi (4\tau_p \chi_j)^{1/2}$ .

Отметим, что ПУС в виде модуляций рельефа поверхности кремния с периодом  $d \sim \tau_p^{1/2}$ , ориентированные вдоль кристаллографического направления, наблюдались при действии пикосекундных лазерных импульсов в работе [13].

Значение  $I_{кр}$  для ЭДТН определяется из условия  $\lambda = 0$ . Из (9) в случае преобладания либо термоупругого ( $j = T$ , ДТН), либо электронно-деформационного ( $j = n$ , ЭДН) механизмов имеем ( $\lambda \rightarrow 0$ )

$$\lambda = -(2\chi_j q^2 + \tau_{j0}^{-1}) + \frac{2q}{\gamma_0} \left( \frac{R_j}{1-\beta} + \epsilon_{jj} \right). \quad (12)$$

В случае ДТН  $I_{кр} \approx 5 \cdot 10^6$  Вт/см<sup>2</sup>, а при ЭДН  $I_{кр} \approx 5 \cdot 10^7$  Вт/см<sup>2</sup>. Дисперсионное уравнение (9) справедливо и в аксиально-симметричном случае, однако, как видно, все гармоники  $m$  имеют одинаковый инкремент. Это вырождение по  $m$  снимается при учете гауссовской зависимости интенсивности лазерного поля от  $r/E^2$  ( $r \sim \exp[-r^2/r_0^2]$ ). При этом в разных условиях выделяются гармоники с  $m \sim (q_m r_0)$  либо  $m \sim (q_m r_0)^2$ . Результирующие ПУС имеют вид солнца с  $m$  лучами.

Стабилизация ЭДТН происходит благодаря оже-рекомбинации. Зададим при  $z = 0$   $n(0) = n_q \exp iqx$ ,  $\xi(0) = \xi_q \exp iqx$ ,  $T(0) = T_q \exp iqx$ , где фурье-амплитуды  $\xi_q$  и  $T_q$  выражаются через  $n_q$ . При этом  $E_g(0) = E_{g_0} - \Delta E_g \cos gx$ . В стационарном режиме из условия  $\lambda = 0$ , используя (12), где производится замена  $\tau_{n_0}^{-1} \rightarrow \tau^{-1} + \gamma_{оже} |n_q|^2$ , и полагая  $j = n$ , получаем

$$|n_q| = \frac{1}{\gamma_{оже}^{1/2}} \left[ \frac{2q}{\gamma_0} \left( \frac{R_n}{1-\beta} + \varepsilon_{nn} \right) - 2Dq^2 - \tau^{-1} \right]^{1/2}. \quad (13)$$

Зависимость стационарного значения  $n_q$  от  $E^2$  имеет корневой вид, типичный для фазового перехода второго рода (аналогично с помощью (10) получается эта же зависимость и вдали от порога). Зависимость  $n_q$  от  $q$  повторяет зависимость  $\lambda = \lambda(q)$ . Для стационарной амплитуды волны деформации на поверхности имеем  $\xi_q = n_q (D\gamma_0 \delta_n - \varepsilon_{nn}) / n_{\xi}$ , где  $\delta_n = [q^2 + (\tau^{-1} + \gamma_{оже} n_q^2) / D]^{1/2}$ , а для амплитуды температурной решетки  $T_q = \varepsilon_{T\xi} D \delta_n n_q / \varepsilon_{n\xi} \lambda q$ . Используя эти выражения в (1), получаем для амплитуды волны перенормировки ширины запрещенной зоны

$$\Delta E_g = n_q D \delta_n \left[ \frac{2\pi \hbar^2 \gamma_0}{(\partial \varepsilon_0 / \partial \omega) E^2} + \frac{\hbar \omega}{\chi q C_V} \left| \frac{\partial E_q}{\partial T} \right| \right].$$

При  $I = 6 \cdot 10^8$  Вт/см<sup>2</sup> и характерных значениях параметров полупроводников имеет  $\Delta E_q \approx 1$  эВ,  $d \approx 1 \div 5$  мкм,  $\lambda_m \approx 10^{10} \div 10^{11}$  с<sup>-1</sup>. Таким образом, ЭДТН может приводить к фазовому переходу полупроводник—металл с образованием ПУС чередующихся металлической ( $E_g = 0$ ) и диэлектрической фаз. В случае одномерной геометрии лазерного поля на изотропной поверхности должны образовываться одномерные решетки фаз, а в случае аксиальной симметрии поля — ПУС типа солнца. Именно это и наблюдается при действии пикосекундных лазерных импульсов на VO<sub>2</sub> [14].

Как видно из проведенного анализа, ЭДТН протекает за счет энергии лазерного импульса, которая трансформируется непосредственно в энергию упругих деформаций. К аналогичным эндотермическим ДДН относится вакансионно-деформационная неустойчивость (ВДН) и дислокационно-деформационная неустойчивость (ДиДН). При ВДН деформация модулирует энергию образования (или миграции) вакансии аналогично модуляции ширины запрещенной зоны при ЭДТН:  $E_v = E_{v_0} + \theta_v \operatorname{div} \mathbf{u}$ ;  $\theta_v < 0$ . При ДиДН деформация модулирует скорость движения дислокации.

Образование ПУС в виде «солнца» в металлах под действием непрерывного лазерного излучения, которое можно интерпретировать с точки зрения развития ВДН, наблюдалось в работе [15]. В работе [16] наблюдалось образование одномерной решетки концентрации дислокаций, ориентированной вдоль кристаллографической оси на поверхности кремния, под действием миллисекундных лазерных импульсов, которое можно интерпретировать как проявление ДиДН.

В тонких пленках возникновение ДДН может быть обусловлено изгибными деформациями. Примером может служить пористо-деформационная неустойчивость (ПДН) в металлической пленке с порами, находящейся под воздействием сфокусированного лазерного излучения. ПДН описывается уравнением для координаты изгибных деформаций пленки [17], взаимодействующих со сквозными порами в пленке,

$$\partial^2 \zeta / \partial t^2 - c^2 \Delta_{\parallel} \zeta + (c_{\parallel}^2 h^2 / 12) \Delta_{\parallel}^2 \zeta = -GN_1, \quad (14)$$

где  $\Delta_{\parallel}$  — двумерный (вдоль пленки) оператор Лапласа;  $c$  — константа порядка скорости звука;  $c_{\parallel}^2 = \sigma_{\parallel} / \rho$ ;  $\sigma_{\parallel} \approx K^2 \alpha T / c_{\parallel}^2 \rho$  — продольное напряжение, возникающее благодаря неоднородному вдоль  $r$  распределению температуры  $T = T(r)$ , создаваемому лазерным пучком;  $h$  — толщина пленки;  $G = \sigma_{\perp} S_0 / h \rho$ , где  $\sigma_{\perp}$  — перпендикулярное к пленке напряжение, возникающее благодаря несоответствию параметров решеток в пленке и подложке;  $S_0$  — площадь основания сквозной поры высотой  $h$ ;  $N_1$  — флуктуация плотности числа пор (с размерностью  $\text{см}^{-2}$ );  $N = N_0 + N_1$ ,  $N_0$  — средняя плотность числа пор. Уравнение для  $N_1$  имеет вид

$$\frac{\partial N_1}{\partial t} = D \Delta_{\parallel} N_1 - \frac{DN_0 |\theta| h}{kT_0} \Delta_{\parallel}^2 \zeta, \quad (15)$$

где  $D = D_0 a^3 n_0$  — коэффициент диффузии пор;  $D_0$  — коэффициент диффузии вакансий вдоль пленки;  $a$  — параметр решетки в пленке;  $n_0$  — концентрация вакансий;  $\theta = -1/3 E_0$ ;  $E_0$  — энергия миграции вакансий вдоль пленки. Задавая решение (14), (15) в виде  $N_1, \zeta \sim J_0(qr) \exp \lambda t$ , получаем для инкремента  $\lambda = -Dq^2 + aq^2 / (1 + l_0^2 q^2)$ , где  $a = G |\theta| DN_0 h / kT_0 c_{\parallel}^2$ ,  $l_0^2 = (c^2 / c_{\parallel}^2) h^2 / 12$ . Максимальное значение инкремента достигается в точке  $q = q_m$ , причем

$$\lambda_m = 2l_0^{-2} [(a + D)/2 - \sqrt{aD}]; \quad q_m = l_0^{-1} (\sqrt{aD} - 1)^{1/2}$$

и период возникающих кольцевых структур скопления пор и изгибной деформации равен  $d = 2\pi l_0 / (\sqrt{aD} - 1)^{1/2}$ . Образование подобных кольцевых структур сквозных пор в металлических пленках наблюдается при лазерном осаждении их на диэлектрические подложки.

Помимо эндотермических ДДН, рассмотренных выше, возможны также и экзотермические ДДН, которые развиваются за счет запасенной в среде энергии, например скрытой теплоты кристаллизации в аморфной среде. Импульсное лазерное воздействие в этом

случае играет роль первоначальной активации. Мы рассмотрим этот тип ДДН на примере кристаллизационно-деформационно-тепловой неустойчивости (КДТН), протекающей на аморфной поверхности полупроводников.

Пусть аморфная среда занимает полупространство  $z \geq 0$ ;  $a = a(x, y)$  ( $0 \leq a \leq 1$ ) обозначает локальный относительный объем, занимаемый аморфной фазой ( $a = \Delta v_a / \Delta v$ ,  $\Delta v_a + \Delta v_c = \Delta v$ , где  $\Delta v$  — физически малый объем,  $\Delta v_a$  и  $\Delta v_c$  — соответственно объемы аморфной и кристаллической фаз). Переменная  $a$  подчиняется активационному кинетическому уравнению

$$\partial a / \partial t = a \tau_0^{-1} \exp [ - (E_a + \theta_a \operatorname{div} \mathbf{u}) / kT ]. \quad (16)$$

Здесь  $E_a$  — энергия активации;  $\tau_0^{-1}$  — константа скорости кристаллизации. Как видно из (16), деформация в случае КДТН модулирует энергию активации. Уравнение для температуры имеет вид

$$\frac{\partial T}{\partial t} = \chi \Delta T + \frac{E_k a}{C_V \tau_0} \exp \left( - \frac{E_a + \theta_a \operatorname{div} \mathbf{u}}{kT} \right) + \frac{c \gamma E^2}{2\pi C_V} \exp(-\gamma z), \quad (17)$$

где  $E_k$  — скрытая теплота кристаллизации. Уравнения (16) и (17) замыкаются уравнением (5) для  $\mathbf{u}$  с граничными условиями типа (5а).

Положим  $T = T_0 + T_1$ ,  $a = a_0 + a_1$ ,  $\xi = \xi_0 + \xi_1$ , где  $\xi_0 = \operatorname{div} \mathbf{u}_0$ ;  $T_0$ ,  $a_0$  — пространственно-однородные вдоль поверхности решения, а  $\xi_1$ ,  $T_1$ ,  $a_1$  — малые пространственно-неоднородные возмущения. Линеаризуя (16), (17), получаем уравнения нулевого и первого порядка:

$$\frac{\partial a_1}{\partial t} = \tau_0^{-1} \exp \left( - \frac{E_{a0}}{kT_0} \right) \left[ a_0 \left( - \frac{\theta_a}{kT_0} \xi_1 + \frac{E_{a0}}{kT_0^2} T_1 \right) + a_1 \right];$$

$$E_{a0} \equiv E_a + \theta_a \xi_0, \quad (18)$$

$$\frac{\partial T_1}{\partial t} = \chi \Delta T_1 + E_k \tau_0^{-1} C_V^{-1} \exp \left( - \frac{E_{a0}}{kT_0} \right) \left[ a_0 \left( \frac{E_{a0}}{kT_0^2} T_1 - \frac{\theta_a}{kT_0} \xi_1 \right) + a_1 \right]. \quad (19)$$

Разложим далее  $\xi_0(z)$  в ряд:  $\xi_0(z) = \xi_0(0) - |\xi_0'|z$  и используем это разложение в (18) и (19). При условии  $\lambda \tau_0 \exp(E_{a0}/kT_0) \gg 1$  в (18), (19) можно пренебречь членом, пропорциональным  $a_1$ . Тогда (19) приобретает вид

$$\partial T_1 / \partial t = \chi \Delta T_1 + \sigma T_1 e^{-\Gamma z} + g \operatorname{div} \mathbf{u} e^{-\Gamma z}, \quad (20)$$

где обозначено  $\sigma = \tau_0^{-1} \exp(-E_{a0}/kT_0) E_k E_{a0} a_0 / kT_0^2 C_V$ ;  $\Gamma = |\theta_0| \left| \frac{\partial \xi_0}{\partial z} \right| / kT_0$ .

Уравнения (20) и (5) составляют замкнутую систему, аналогичную уравнениям ЭДГН, однако теперь неустойчивость возникает за счет скрытой теплоты. При этом экспоненциально во времени

нарастают фурье-амплитуды поверхностной деформации и температуры, модулирующие, согласно (18), скорость кристаллизации. Поэтому при импульсном лазерном облучении аморфных полупроводников на их поверхности могут образовываться ПУС с чередованием преимущественно аморфной или кристаллической фаз. Эти ПУС могут иметь форму концентрических колец, «солнца», радиально-кольцевых ячеек. Роль активатора может также играть электронный пучок, деформация или просто нагрев в печи. Подобные ПУС действительно наблюдаются как при импульсной лазерной кристаллизации [18], так и при других способах активации процесса кристаллизации [19].

В заключение отметим, что в жидкой фазе возможны неустойчивости, аналогичные рассмотренной выше КДТН. Роль уравнения (5) в этом случае играют уравнения гидродинамики.

#### ЛИТЕРАТУРА

1. Young J. F., Sipe J. E., van Driel H. M. Laser induced periodic surface structure // Phys. Rev. B. 1984. Vol. 30. P. 2002—2015.
2. Ахманов С. А., Емельянов В. И., Коротеев Н. И., Семиногов В. Н. Воздействие мощного лазерного излучения на поверхность полупроводников и металлов: Нелинейно-оптические эффекты и нелинейно-оптическая диагностика // УФН. 1985. Т. 147, № 4. С. 675—745.
3. Siegmann A. E., Fauchet P. M. Stimulated Wood's anomalies on laser illuminated surfaces // IEEE J. Quant. Electron. 1986. Vol. 22. P. 1384—1403.
4. Емельянов В. И., Конов В. И., Семиногов В. Н., Токарев В. Н. Образование поверхностных периодических структур при воздействии импульсного излучения CO<sub>2</sub>-лазера на плавленый кварц: Препр. ИОФАН № 89. М., 1987.
5. Емельянов В. И., Семиногов В. Н. Самоиндуцированный резонанс локального поля и аномально высокая поглощательная способность на поверхности конденсированных сред // Изв. АН СССР. Сер. физ. 1986. Т. 50. С. 2273—2279.
6. Емельянов В. И., Семиногов В. Н., Соколов В. И. Дифракция света на поверхности с большой амплитудой модуляции рельефа и поверхностные нелинейно-оптические эффекты // Квантовая электрон. 1987. Т. 14. С. 33—46.
7. Емельянов В. И., Семиногов В. Н. Нелинейный режим лазерной генерации капиллярных волн и образование упорядоченных поверхностных структур: Препр. НИЦТЛ АН СССР № 15. Троицк, 1986. 23 с.
8. Емельянов В. И., Семиногов В. Н. Аномально высокая поглощательная способность и аномально быстрый нагрев шероховатой поверхности конденсированных сред // Квантовая электрон. 1987. Т. 14. С. 47—54.
9. Емельянов В. И. Нелинейно-оптическая деформация акустической подсистемы полупроводника и сверхбыстрое плавление поверхности мощными короткими лазерными импульсами: Препр. МГУ № 5. М., 1985.
10. Емельянов В. И., Уварова И. Ф. Электронно-деформационно-тепловая неустойчивость и фазовый переход полупроводник—металл под действием лазерного излучения с образованием сверхструктур // Изв. АН СССР. Сер. физ. 1986. Т. 50. С. 1214—1219.
11. Емельянов В. И., Уварова И. Ф. Самовозбуждение поверхностной акустической волны при межзонных переходах в полупроводниках под действием света // Акуст. журн. 1985. Т. 31. С. 481—487.
12. Shmidt M., Dransfeld K. Excitation of surface acoustical wave under laser illumination of GaAs // Appl. Phys. A. 1982. Vol. 28. P. 211—219.



13. Jost D., Lüthy W., Webber H. P. Laser pulse dependent surface ripples on silicon // Appl. Phys. Lett. 1986. Vol. 49. P. 625—627.
14. Бугаев А. А., Клочков А. В. Необратимые изменения в пленке двуокиси ванадия при пикосекундном лазерном воздействии // ФТТ. 1984. Т. 26. С. 3487—3489.
15. Гусев В. С., Козлова Е. К., Портнягин А. И. О роли термоградиентных явлений в лазерной электрохимии // Квантовая электрон. 1987. Т. 14. С. 323—331.
16. Вейко В. П., Имас Я. А., Либенсон М. Н. и др. Формирование регулярных структур на поверхности кремния под действием миллисекундного импульса неодимового лазера // Изв. АН СССР. Сер. физ. 1985. Т. 49. С. 1236—1239.
17. Ландау Л. Д., Лифшиц Е. М. Теория упругости. М.: Наука, 1987.
18. Баскин Б. Л., Поляков А. А., Трушин В. Н., Яровецкий И. Д. Воздействие импульсного лазерного излучения пикосекундной длительности на поверхность германия // Письма в ЖТФ. 1985. Т. 11. С. 1251—1257.
19. Александров Л. Н. Кинетика кристаллизации и перекристаллизации полупроводниковых пленок. Новосибирск: Наука, 1985. 224 с.

УДК 535.21 : 517.9

## ВЫНУЖДЕННОЕ РАССЕЯНИЕ И ПОВЕРХНОСТНЫЕ СТРУКТУРЫ

В. К. ГАВРИКОВ, А. В. КАЦ, В. М. КОНТОРОВИЧ,  
И. С. СПЕВАК

Открытое вскоре после изобретения лазеров вынужденное рассеяние (ВР) представляет собой столь же общее явление, что и нелинейный сдвиг частоты осциллятора в поле чужой сильной волны. По существу, ВР — это неустойчивость, которую проявляют практически все низкочастотные моды в поле достаточно мощного (лазерного) излучения, т. е. каждому спонтанному рассеянию соответствует, как давно установлено, его вынужденный аналог. Характеризующий ВР инкремент — это не что иное, как положительная мнимая часть сдвига частоты, что соответствует притоку энергии в данную степень свободы, а не ее утечке (как при нелинейном затухании). Инкремент ВР пропорционален интенсивности лазерного поля.

Вынужденный аналог спонтанного рассеяния на поверхностных волнах, рассмотренного Л. И. Мандельштамом [1] и А. А. Андроновым и М. А. Леонтовичем [2], обсуждался впервые в [3—5] для прозрачных, а в [6] — для бесконечно проводящих сред. Был выяснен ряд характерных особенностей этого вида ВР, в частности показано, что максимум инкремента соответствует появлению одной или нескольких скользящих рассеянных волн [5]. Однако на пути наблюдения вынужденного рассеяния на поверхностных волнах (ВРПВ) в недиссипативных условиях практически непреодолимым препятствием<sup>1</sup> стало даже малое поглощение. Приблиз-

<sup>1</sup> Тем не менее ВРПВ, по-видимому, наблюдалось (см. [7—9]).

тельно в то же время (и даже несколько ранее, если говорить о первых наблюдениях [10]) экспериментально было обнаружено, что при лазерном облучении металлов, полупроводников и диэлектриков, сопровождающемся значительными тепловыми эффектами, в определенном диапазоне интенсивностей и длительностей облучения на их поверхности образуются периодические структуры (ПС) — решетки (см. обзоры [11, 12] и цитируемую там литературу).

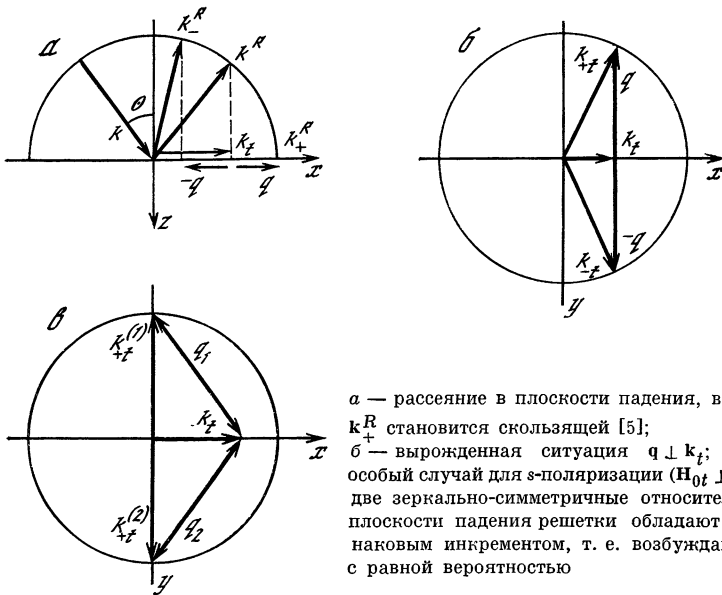
Период решеток зависит от угла падения  $\theta$  и по порядку величины равен длине волны  $\lambda$  воздействующего света, а ориентация ПС определяется направлением поляризации и углом падения. Согласно сложившимся к настоящему времени представлениям, образование ПС обусловлено рассеянием электромагнитной волны на затравочных неровностях поверхности материала с последующей интерференцией падающего и рассеянного излучений и обратным влиянием интерференционной части интенсивности на медленные движения среды (границы) совершенно так же, как это происходит в процессах ВР. По существу, возникновение ПС есть проявление ВРПВ за счет тепловых механизмов (в недиссипативных средах механизм раскачки был обусловлен световым давлением).

Замечательно, что параметры ПС не столь детально связаны с конкретным механизмом диссипативного взаимодействия, а зависят в основном от геометрии рассеяния, определяющей структуру теплового источника, который находится из рассмотрения электродинамической части задачи.

Рассеянные поля первого порядка <sup>2</sup>, отраженные ( $R$ ) и преломленные ( $T$ ) на неоднородностях поверхности  $z = \zeta(\mathbf{r}, t) = \text{Re } \zeta \times \exp i(\mathbf{q}\mathbf{r} - \Omega t)$ , билинейны по падающему полю  $\text{Re } \mathbf{E}_0 \exp i(\mathbf{k}\mathbf{r} - \omega t)$  и возвышению  $\zeta$ :  $\mathbf{E}^R = \hat{R}\mathbf{E}_0\zeta$ , где  $\hat{R}$  — обобщенные френелевские коэффициенты [2, 13]; им соответствуют комбинационные частоты  $\omega + j \text{Re } \Omega$  и волновые векторы  $\mathbf{k}_{jt} = \mathbf{k}_t + j\mathbf{q}$  ( $j = \pm$ ) в плоскости раздела (рис. 1). Вследствие этого возникают биения вектора Пойнтинга (которые запишем в виде  $\text{Re}(Qk\zeta)$ ) с той же частотой и длиной волны, что и возвышение  $\zeta$ , породившее рассеянную волну. Поглощение энергии через тот или иной механизм обратного влияния [11] приводит в уравнении для  $\zeta$  к появлению секулярного вынуждающего члена  $\sim Q\zeta$ , что и порождает неустойчивость с инкрементом  $\gamma \sim \text{Re}[q^\mu Qe^{i\beta}]$ , пропорциональным мощности теплового источника <sup>3</sup>, но, вообще говоря, сдвинутым относительно него по фазе. Параметры  $\mu$  и  $\beta$  зависят от конкретного ме-

<sup>2</sup> В отдельных случаях, например для почти идеально проводящей среды, резонансный характер скользящей рассеянной волны приводит к столь значительному росту ее амплитуды, что необходимо рассмотрение существенно нелинейной по возвышению задачи с привлечением дифракционных спектров порядка выше первого [6].

<sup>3</sup> Мы отвлекаемся от очевидной необходимости учитывать вклад затухания в уравнении для  $\zeta$ , например, считая, что инкремент существенно превышает соответствующий декремент.



а — рассеяние в плоскости падения, волна  $k_t^R$  становится скользящей [5];  
 б — вырожденная ситуация  $q \perp k_t$ ; в — особый случай для s-поляризации ( $H_{0t} \perp k_t$ ) две зеркально-симметричные относительно плоскости падения решетки обладают одинаковым инкрементом, т. е. возбуждаются с равной вероятностью

Рис. 1. Геометрия рассеяния электромагнитной волны на гофрированной поверхности

ханизма [11]: для испарительного механизма  $\beta = 0$ ,  $\mu = 0$ , термокапиллярный эффект в низкочастотной области дает  $\beta = \pi$ ,  $\mu = -2$  (другие варианты см. в [12, 13]). Таким образом, информация о возможных типах ПС содержится в основном в выражении для интерференционного потока  $Q$ .

С точностью до членов первого порядка по  $\zeta$  нормальная к поверхности компонента потока излучения  $S = c |E, H|/4\pi$  на частоте биений при  $z = \zeta(r, t)$  может быть представлена в виде

$$\frac{c |E_0|^2}{16\pi} Q \frac{k\zeta}{2} \exp i(qr - \Omega t) + \text{к. с.} \quad (1)$$

Безразмерный поток  $Q$  для произвольно поляризованного излучения удобно представить в виде билинейной формы s- и p-компонент падающего поля:

$$Q = Q_s \sin^2 \psi + Q_p \cos^2 \psi + Q_{sp} \sin \psi \cos \psi, \quad (2)$$

где угол поляризации  $\psi$  определен согласно выражению

$$E_0 = (E_0 \sin \psi) e_s + (E_0 \cos \psi e^{-i\delta}) e_p, \quad (3)$$

$e_s \equiv e_y$ ;  $e_p = [e_y k]/k$ ;  $E_0 = |E_0|$ ,  $\delta$  — сдвиг фазы между s- и p-компонентами полей. В парциальных потоках удобно выделить вклады s- и p-компонент рассеянных полей (относительно плоскости рассеяния):

$$Q_s = i |T_s|^2 \left\{ -k_x k_{+x} / k^2 + b_{ss}^+ R_+^{ss} + b_{sp}^+ R_+^{ps} - \left( \begin{smallmatrix} \text{к. с.} \\ + \rightarrow - \end{smallmatrix} \right) \right\}, \quad (4)$$

$$Q_p = i |T_p|^2 \left\{ -k_x k_{+x} / k^2 + b_{pp}^+ R_+^{pp} - b_{ps}^+ R_+^{sp} - \left\{ \begin{array}{c} \text{к. с.} \\ + \rightarrow - \end{array} \right\} \right\}, \quad (5)$$

$$Q_{sp} = i \left\{ T_s T_p^* e^{i\delta} [-k_x k_{+y} k_z^T / k^3 + b_{pp}^+ R_+^{ps} - b_{ps}^+ R_+^{ss}] + \right. \\ \left. + T_s^* T_p e^{-i\delta} [k_x k_{+y} k_z^T / \varepsilon k^3 + b_{sp}^+ R_+^{pp} + b_{ss}^+ R_+^{sp}] - \left( \begin{array}{c} \text{к. с.} \\ + \rightarrow - \end{array} \right) \right\}. \quad (6)$$

Здесь  $R_j^{\alpha\beta}$  — коэффициенты Френеля для дифракционных спектров первого порядка [13];  $T_s$ ,  $T_p$  — френелевские коэффициенты прохождения по полю;

$$b_{sp}^j = \frac{k_{jy}}{k_{jt}} \left( 1 + \frac{k_z^{T*} k_{jz}^R}{k^2} \right), \quad b_{ss}^j = \frac{k_{jx}}{k k_{jt}} (k_{jz}^R + k_z^{T*}), \quad j = \pm, \quad (7)$$

коэффициенты  $b_{ps}$  и  $b_{pp}$  получаются соответственно из  $b_{sp}$  и  $b_{ss}$  заменой  $k_z^T \rightarrow k_z^T / \varepsilon$ ,  $(k_{jz}^R)^2 = k^2 - (\mathbf{k}_t + j\mathbf{q})^2$ ,  $(k_z^T)^2 = \varepsilon k^2 - k_t^2$ .

При  $|\varepsilon| \gg 1$  формулы существенно упрощаются, что позволяет провести анализ инкремента ВР для ряда интересных случаев аналитически. В резонансном приближении (соответствующем вкладу скольльзящей рассеянной  $p$ -волны) выражение для потока имеет вид

$$Q = 4i (F_+ f_+ - F_-^* f_-^*), \quad |\varepsilon| \gg 1. \quad (8)$$

Первый множитель

$$F_j \equiv F(\mathbf{k}_{jt}) = [\mathbf{k}_{jt} - \mathbf{k}_t, \mathbf{H}_{0t}|_z [\mathbf{k}_{jt} \mathbf{H}_{0t}|_z^* / k^2 E_0^2] \quad (9)$$

описывает медленную угловую зависимость, определяющую ориентацию ПС<sup>4</sup>, а множитель

$$f_j \equiv f(-k_{jz}^R / k) = \frac{2 \operatorname{Re} \sqrt{\varepsilon}}{|\varepsilon| (1/\sqrt{\varepsilon} - k_{jz}^R / k)}; \quad \operatorname{Im}, \operatorname{Re} k_{jz}^R \leq 0, \quad (10)$$

описывает быструю («резонансную») зависимость, определяющую период ПС. Максимум  $|f|$  достигается при  $\operatorname{Im} k_{jz}^R = -\operatorname{Im} \sqrt{\varepsilon} k / |\varepsilon| \ll k$ , ширина максимума порядка  $\operatorname{Re} \sqrt{\varepsilon} / |\varepsilon| \ll 1$ . Условие резонанса в (10) близко к условию скользкого распространения  $k_{jz}^R \rightarrow 0$  и выделяет однопараметрическое множество возможных П (характеризуемых волновыми векторами  $\mathbf{q}$ ). При этом, однако, остается неизвестной ориентация  $\mathbf{q}$ . Последнюю можно определить, исследуя на экстремум медленную зависимость (9).

В случае вещей резонансных  $F_j$  (произвольная линейная поляризация) положения экстремумов по медленной и быстрой зависимостям находятся независимо, так как выражение для инкремента факторизуется:

$$\operatorname{Re} (Q e^{i\beta}) = -4F_+ \operatorname{Im} (f_+ e^{i\beta}), \quad |f_+| \gg |f_-|, \quad (11)$$

<sup>4</sup> Из представления (9) следует, что ориентацию решеток удобно определять относительно вектора тангенциальной составляющей магнитного поля  $\mathbf{H}_{0t}$ , что обусловлено  $H$ -поляризацией поверхностной волны.

в окрестности одного (+) резонанса. Существенно, что положение экстремумов малочувствительно к  $\beta$ , т. е. к механизму. Для двойного резонанса ( $\mathbf{q} \perp \mathbf{k}_t$ ,  $k_{jx} = k_x$ ,  $f_+ = f$ ), когда одновременно становятся скользящими как стоксова, так и антистоксова рассеянные волны <sup>5</sup>, параметр  $\beta$ , напротив, важнее:

$$\operatorname{Re}(Qe^{i\beta}) = -4[(F_+ - F_-) \sin \beta \operatorname{Re} f_+ + (F_+ + F_-) \cos \beta \operatorname{Im} f_+], \quad (12)$$

и инкремент не факторизуется, за исключением нормального падения:

$$\operatorname{Re}(Qe^{i\beta}) = -8F_+ \cos \beta \operatorname{Im} f_+, \quad \theta = 0, \quad (13)$$

причем (13) справедливо и в общем случае эллиптической поляризации <sup>6</sup>. Рассмотрим теперь экстремумы величины  $\operatorname{Re}(Qe^{i\beta})$  для различных поляризацій и углов падения (роль фактора  $q^\mu$  легко может быть учтена [13]).

#### Нормальное падение эллиптически поляризованного пучка.

Поскольку, согласно (9), в этом случае  $F_\pm = [|\mathbf{q}, \mathbf{H}_0|]^2/k^2 E_0^2$ , то, учитывая (13), видим, что максимуму инкремента отвечает направление большей оси эллипса поляризации [15]. В предельном случае линейной поляризации [11, 16—18] штрихи решетки перпендикулярны вектору  $\mathbf{E}_0$  ( $\mathbf{q} \parallel \mathbf{E}_0$ ), период решетки  $\Lambda = \lambda + o(1/\sqrt{|\varepsilon|})$ ,  $\lambda$  — длина волны света, причем малое отличие от  $\lambda$  зависит от механизма взаимодействия.

**Неполяризованный свет.** Усредняя по ориентациям  $\mathbf{H}_0$  и учитывая, что  $\langle H_{0\alpha} H_{0\beta}^* \rangle = E_0^2 (\delta_{\alpha\beta} - k_\alpha k_\beta / k^2) / 2$ , получим

$$\langle F_j \rangle = (\cos^2 \theta + \sin^2 \theta k_{jx}^2 / k^2 - \sin \theta k_{jx} / k) / 2, \quad |\mathbf{k}_{jt}| \approx k, \quad (14)$$

где  $\theta$  — угол падения. Из (14) следует, что  $\langle F_j \rangle$  достигает максимума при рассеянии «назад»:  $k_{jx} = -k$ ,  $\langle F_j \rangle_{\max} = (1 + \sin \theta) / 2$ . Сравнивая соответствующий максимум инкремента с его значением в двойном резонансе ( $\mathbf{q} \perp \mathbf{k}_t$ ), находим, что в простейшем случае  $\beta = 0$  для малых углов падения ( $1 + \sin \theta < 2 \cos^4 \theta$ ,  $\theta < \theta_0 \approx 24^\circ$ ) максимальному значению  $\operatorname{Re} Q$  соответствует вырожденная геометрия, в соответствии с чем при угле падения  $\theta_0$  должна происходить переориентация решеток [19].

**Произвольная линейная поляризация.** Из представления (9) видно, что экстремумы сомножителей  $F_j$  достигаются при  $\mathbf{q} \perp \mathbf{H}_{0t}$  и  $\mathbf{k}_{jt} \perp \mathbf{H}_{0t}$ . Поскольку длина вектора  $\mathbf{k}_{jt}$  фиксирована ( $k_{jt} = k$  в резонансе), в то время как длина  $\mathbf{q}$  меняется, то второй экстремум более существен для определения максимума инкремента

<sup>5</sup> Эти и другие особые и вырожденные геометрии рассеяния см. в [14]. Существенная роль вырожденного случая  $\mathbf{q} \perp \mathbf{k}_t$  при формировании ПС обсуждается в [11]. Отметим также особую геометрию  $\mathbf{k}_{+t} \perp \mathbf{H}_{0t}$  для  $s$ -поляризации (см. рис. 1).

<sup>6</sup> Обращение в нуль инкремента при  $\beta = \pi/2$  соответствует отсутствию ВР для нормального падения при недиссипативном механизме воздействия [5]. Поскольку при нормальном падении формирование ПС наблюдается, то это означает, что ответствен за это механизм, при котором  $\cos \beta \neq 0$ .

та. При линейной поляризации комплексное сопряжение в (9) можно опустить. Вводя углы  $\chi_j = \widehat{\mathbf{k}_{jt}\mathbf{H}_{0t}}$  и  $\psi_1 = \widehat{\mathbf{k}_t\mathbf{H}_{0t}}$ , получим, что главный максимум  $(F_j)_{\max} = A(1+a)$  достигается при  $\chi_j = 3\pi/2$ , где  $A = |\mathbf{H}_{0t}|^2/E_0^2$ ,  $a = \sin\theta \sin\psi_1$  (рис. 2). Остается сравнить значение  $\text{Re}(Qe^{i\beta})$  в этом максимуме с соответствующим выражением для вырожденной геометрии. Результат существенно зависит от сдвига фаз  $\beta$ . В частном случае  $\beta = 0$  угол падения, при котором происходит переориентация решеток, определяется углом  $\psi_1$  [13], причем если излучение близко к  $p$ -поляризованному ( $-1 < \cos 2\psi_1 < 0$ ), то преобладающей является решетка  $\mathbf{k}_{jt} \perp \perp \mathbf{H}_{0t}$ .

**Наклонное падение циркулярно поляризованного пучка** ( $\psi = \pi/4$ ,  $\delta = \pm \pi/2$ ). В этом случае из-за комплексности  $F_j$  медленную зависимость нельзя исследовать независимо от быстрой и положение экстремумов зависит от сдвига фаз  $\beta$ . Для определенности ниже считаем, что формирование ПС определяется  $\text{Re } Q$  (т. е.  $\beta = 0, \pi$ ):

$$\text{Re } Q/4 = -\text{Re } F \text{Im } f - \text{Im } F \text{Re } f, \quad (15)$$

где у  $F_j, f_j$  опущены индексы. Учитывая явное выражение (10) для  $f$  ( $-k_{jz}^R/k$ ), найдем вначале экстремумы  $\text{Re } Q$  по  $k_{jz}^R$ . Проще всего анализ выполняется для сред с  $\varepsilon' < 0$ ,  $\varepsilon'' \ll |\varepsilon'|$  (полупроводников). Наибольший экстремум

$$(\text{Re } Q/4)_{\text{extr}} = -\text{Im } F \pm |F| \text{sgn } \text{Re } F \quad (16)$$

достигается при

$$-\text{Im } k_{jz}^R/k = \frac{\text{Im } \sqrt{\varepsilon}}{|\varepsilon|} + \frac{\text{Re } \sqrt{\varepsilon}}{|\varepsilon|} \frac{\text{Im } F + |F| \text{sgn } \text{Re } F}{\text{Re } F}. \quad (17)$$

Выражая  $F_j$  через углы между плоскостью падения и плоскостями рассеяния  $\varphi_j = \widehat{\mathbf{k}_t\mathbf{k}_{jt}}$ :

$$F_j = \sin^2 \psi \cos^2 \theta \sin^2 \varphi_j + \cos^2 \psi \cos \varphi_j (\cos \varphi_j - \sin \theta) + \\ + \sin \psi \cos \psi \cos \theta \sin \varphi_j [\cos \varphi_j e^{i\delta} + (\cos \varphi_j - \sin \theta) e^{-i\delta}], \quad (18)$$

зависимость (16) легко исследовать численно. Результат в случае  $\beta = 0$  (ему соответствует верхний знак в (16), (17)) приведен на рис. 3, причем учтено, что для вырожденной геометрии  $\text{Re } Q = 4 \cos^4 \theta$  ( $-\text{Im } f)_{\max} = 4 \cos^4 \theta$ . Для левого вращения ( $\delta = \pi/2$ ) график  $\text{Re } Q$  получается путем отражения относительно прямой  $\varphi_+ = \pi$ , что соответствует зеркальной симметрии картины рефлексов относительно плоскости падения при замене направления вращения. Такое поведение подтверждается экспериментом [20]. При малых углах падения ( $\theta < 22^\circ$ ) реализуется вырожденная геометрия ( $\mathbf{q} \perp \mathbf{k}_t$ ). При  $\theta \rightarrow 0$  максимумы  $\text{Re } Q$  становятся слабо выраженными (их угловая ширина  $\Delta\varphi \sim 2/\sqrt{|\varepsilon|} \sin 2\theta$ ). Это находится в согласии с тем, что при нормальном падении рассеяние изотропно.

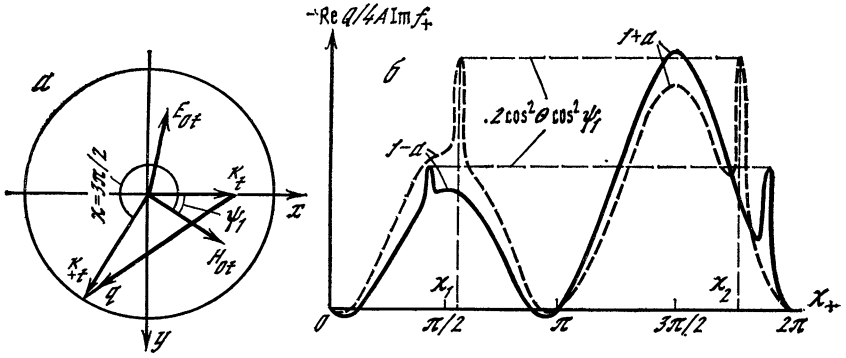


Рис. 2. Периодические структуры для произвольной линейной поляризации  $a$  — схема рассеяния излучения (пучок близок к  $s$ -поляризованному,  $0 < \cos 2\psi_1 < 1$ ; так как векторы  $\mathbf{H}_{0t}$  и  $\mathbf{E}_{0t}$  взаимно перпендикулярны (за исключением чистых  $s$ - и  $p$ -поляризаций), то из условия  $\mathbf{k}_{jt} \perp \mathbf{H}_{0t}$ , вообще говоря, не следует  $\mathbf{k}_{jt} \parallel (\mathbf{E}_{0t})$ ;  $b$  — зависимость инкремента от угла  $\chi_+ = \widehat{\mathbf{k}_{+t}\mathbf{H}_{0t}}$  для  $\theta = 45^\circ$  (сплошная линия) и  $\theta = 20^\circ$  (штриховая линия) при  $\psi_1 = 30^\circ$  (во втором случае инкремент максимален для вырожденной геометрии)

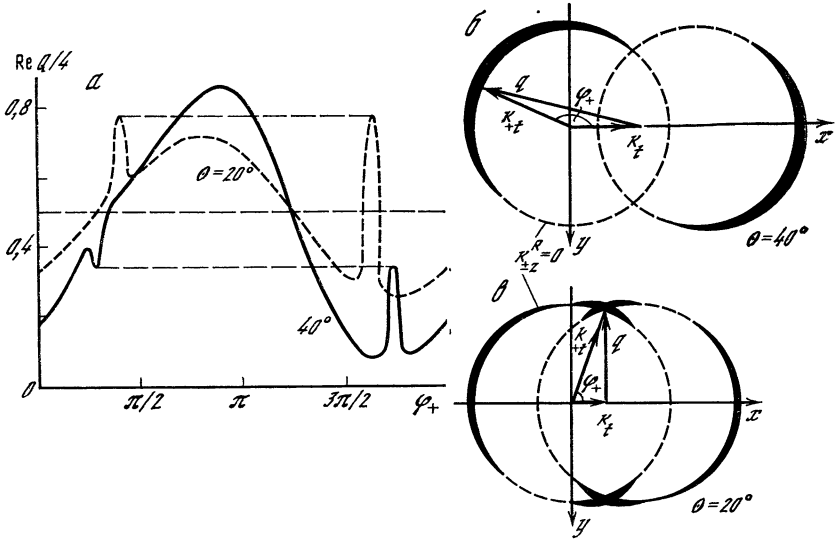


Рис. 3. Формирование ПС циркулярно поляризованным светом

$a$  — интерференционный поток как функция угла  $\varphi_+ = \widehat{\mathbf{k}_t\mathbf{k}_{+t}}$  для излучения с правой поляризацией;  $b, \epsilon$  — картина рефлексов при освещении ПС более коротковолновым излучением, падающим нормально к поверхности. Толщина луночек пропорциональна превышению интерференционным потоком порогового уровня. Диаграмма  $b$  согласуется с имеющимися наблюдениями [20]

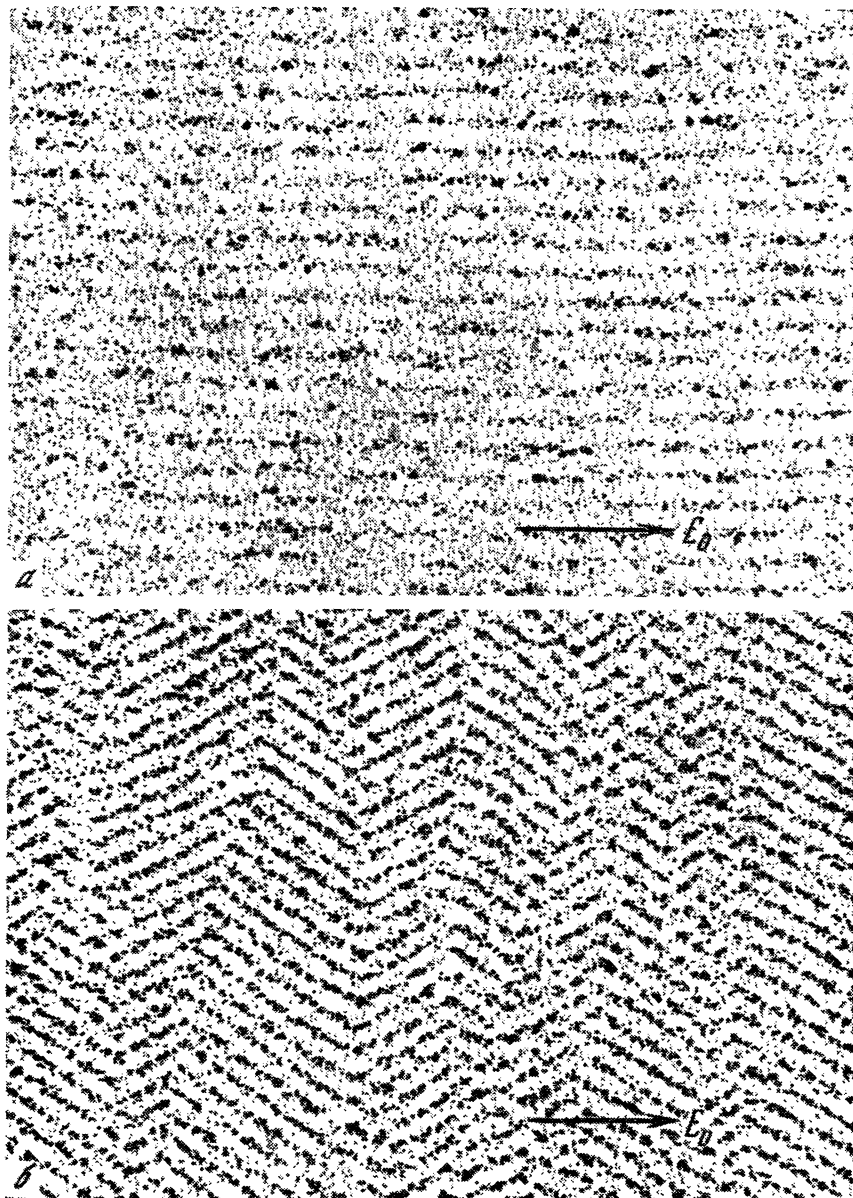


Рис. 4. Микрофотографии ПС, фотоиндуцированных  $p$ -поляризованным излучением в пленке  $\text{AgCl} - \text{Ag}$  на стеклянной подложке [21–23]

Угол падения  $\theta = 70^\circ$ , рассеянные волны ТЕ-типа:  $a$  — толщина пленки  $h = 30$  нм,  $\mathbf{q} \perp \mathbf{E}_0$ ,  $\mathbf{k}_t$ , вырожденная геометрия рассеяния;  $b$  —  $h = 100$  нм, паркетная структура, соответствующая зеркально-симметричным вариантам рассеяния относительно плоскости падения (ср. рис. 1,  $\delta$ )



В заключении отметим следующее. При рассеянии падающего излучения на заданной ПС в общем случае возбуждается гибридная волна (отличны от нуля  $s$ - и  $p$ -компоненты полей). В предельном случае  $|\varepsilon| \gg 1$  в резонансных условиях френелевские коэффициенты удовлетворяют неравенству  $|R_j^{s\alpha}| \ll |R_j^{p\alpha}|$ ,  $\alpha = p, s$ , т. е. рассеянная волна является волной  $p$ -типа независимо от поляризации падающего излучения. Отсюда следует, что максимуму интерференционного потока  $[E^R, \mathbf{H}_{01}]_z$  соответствует геометрия рассеяния  $\mathbf{k}_{jt} \perp \mathbf{H}_{01}$ .

ПС на поверхности твердых тел, обусловленные тепловыми механизмами, требуют больших потоков излучения, близких к порогу разрушения материала ( $\sim 1$  МВт/см<sup>2</sup> при длительностях импульсов 10—100 нс). В то же время фотоструктурные механизмы, ответственные за образование ПС в тонких пленках AgCl—Ag [21—23], реализуются в непрерывном режиме при значительно меньших уровнях мощности ( $\sim 10$  мВт). При наличии на поверхности образца пленки (или естественного окисного слоя [12]) резонанс осуществляется на волноводных модах, которые могут быть как ТМ-, так и ТЕ-типа. В тех случаях, когда резонансная рассеянная волна (аналог скользящей волны в предыдущем изложении) является модой ТМ-типа, картина ПС близка к описанной выше, в частности  $\mathbf{q} \parallel \mathbf{E}_0$  при нормальном падении [21—23]. Если же резонанс осуществляется с волной ТЕ-типа, то возникающие в пленке ПС ориентированы «ортогонально» предыдущим (при той же поляризации воздействующего излучения). Например, при нормальном падении ( $\mathbf{q} \perp \mathbf{E}_0$ ) паркетные структуры (рис. 4) возбуждаются не  $s$ -, а  $p$ -поляризацией [21—23] и т. п.

Мы признательны авторам [21—23] за предоставленные фотографии и полезное обсуждение.

#### ЛИТЕРАТУРА

1. *Мандельштам Л. И.* О шероховатой свободной поверхности жидкости // Полн. собр. тр. М.: Изд-во АН СССР, 1948. Т. 1. С. 246—260; — *Idem* // *Ann. Phys.* 1913. Bd. 41, N 8. S. 609—624.
2. *Андронов А. А., Леонтович М. А.* К теории молекулярного рассеяния света на поверхности жидкости // Андронов А. А. Собр. тр. М.: Изд-во АН СССР, 1956. Т. 1. С. 5—18; — *Idem* // *Ztschr. Phys.* 1926. Bd. 38. S. 485—499.
3. *Бункин Ф. В., Самохин А. А., Федоров М. В.* Вынужденное рассеяние света на поверхности жидкости // Письма в ЖЭТФ. 1968. Т. 7, № 11. С. 431—433.
4. *Бункин Ф. В., Самохин А. А., Федоров М. В.* К теории вынужденного рассеяния света на поверхности жидкости // ЖЭТФ. 1969. Т. 56, № 3. С. 1057—1063.
5. *Гавриков В. К., Кац А. В., Конторович В. М.* Вынужденное рассеяние на поверхностных волнах // ДАН СССР. 1969. Т. 186, № 5. С. 1052—1054; — То же // ЖЭТФ. 1970. Т. 58, № 4. С. 1318—1331.
6. *Кац А. В., Маслов В. В.* Вынужденное рассеяние электромагнитных волн на поверхности хорошо проводящей жидкости // ЖЭТФ. 1972. Т. 62, № 2. С. 496—504.
7. *Ashkin A., Dziedzic J. M.* Radiation pressure on a free liquid surface // *Phys. Rev. Lett.* 1973. Vol. 30, N 4. P. 139—142.

8. *Иванова О. Н., Чернов С. П., Шепелев А. В.* Экспериментальное исследование взаимодействия мощного лазерного излучения со свободной поверхностью жидкости при полном внутреннем отражении // *Квантовая электрон.* 1974. Т. 1, № 9. С. 2077—2080.
9. *Козлов Л. Н., Маляровский А. И., Семенов Ю. Г.* и др. Исследование вынужденного рассеяния света на поверхности жидкости // *Тр. МФТИ. Сер. общ. и молекуляр. физики.* 1976. № 8. С. 100—107.
10. *Birnbaum M.* Semiconductor surface damage produced by ruby lasers // *J. Appl. Phys.* 1965. Vol. 36, N 11. P. 3688—3689.
11. *Ахманов С. А., Емельянов В. И., Коротеев Н. И., Семиногов В. Н.* Воздействие мощного лазерного излучения на поверхность полупроводников и металлов: Нелинейно-оптические эффекты и нелинейно-оптическая диагностика // *УФН.* 1985. Т. 147, № 4. С. 673—745.
12. *Бонч-Бруевич А. М., Коченгина М. К., Либенсон М. Н.* и др. Возбуждение поверхностных и волноводных мод интенсивным лазерным излучением и их влияние на характер поверхностного разрушения конденсированных сред // *Изв. АН СССР. Сер. физ.* 1982. Т. 46, № 6. С. 1186—1193.
13. *Слепак И. С., Конторович В. М., Кац А. В., Гавриков В. К.* О формировании поверхностных структур при тепловом воздействии когерентного излучения // *ЖЭТФ.* 1987. Т. 93, № 1. С. 104—117.
14. *Гавриков В. К., Кац А. В.* О вынужденном рассеянии света на поверхностных волнах при полном внутреннем отражении // *Квантовая электрон.* 1976. Т. 3, № 5. С. 1112—1120.
15. *Keilman F., Bai Y. H.* Periodic surface structure frozen into CO<sub>2</sub> laser-induced quartz // *Appl. Phys.* 1982. Vol. A29. P. 9—18.
16. *Sipe J. E., Young J. F., Preston J. S., van Driel H. M.* Laser-induced periodic surface structure. I. Theory // *Phys. Rev. B.* 1983. Vol. 27, N 2. P. 1141—1154; II. Experiment on Ge, Si, Al and brass // *Ibid.* P. 1155—1172; III. Fluence regimes, the role of feedback, and the details of the induced topography in germanium // *Ibid.* 1984. Vol. 30, N 4. P. 2001—2015.
17. *Прохоров А. М., Свахин А. С., Сычугов В. А.* и др. Возбуждение и резонансные преобразования поверхностной электромагнитной волны при облучении твердого тела мощным лазерным излучением // *Квантовая электрон.* 1983. Т. 10, № 5. С. 906—912.
18. *Анисимов В. Н., Баранов В. Ю., Большой Л. А.* и др. О развитии периодических структур на поверхности металлов и полупроводников под действием излучения импульсного CO<sub>2</sub>-лазера // *Поверхность.* 1983. Т. 7. С. 138—144.
19. *Аксенов В. П., Журкин Б. Г.* Образование периодической структуры при воздействии мощного когерентного излучения на поверхность полупроводников // *ДАН СССР.* 1982. Т. 265, № 6. С. 1365—1366.
20. *Van Driel H. M., Sipe J. E., Young J. F.* Laser-induced periodic surface structure on solids: a universal phenomenon // *Phys. Rev. Lett.* 1982. Vol. 49, N 26. P. 1955—1958.
21. *Агеев Л. А., Блоха В. Б., Милославский В. К.* Особенности периодических структур, индуцированных в тонких пленках AgCl—Ag при наклонном падении поляризованного лазерного излучения // *ЖТФ.* 1985. Т. 55, № 10. С. 1967—1972.
22. *Агеев Л. А., Блоха В. Б., Милославский В. К.* Фотоиндуцированные периодические структуры на металлической поверхности // *Поверхность.* 1986. Т. 8. С. 124—130.
23. *Агеев Л. А., Блоха В. Б., Милославский В. К.* Периодические структуры, индуцированные светом в пленках AgCl—Ag // *Изв. АН СССР. Сер. физ.* 1986. Т. 50, № 8. С. 1605—1608.

## РАСПРЕДЕЛЕННЫЙ ВОЛНОВОЙ КОЛЛАПС В МОДЕЛИ НЕЛИНЕЙНОГО УРАВНЕНИЯ ШРЕДИНГЕРА

С. Н. ВЛАСОВ, В. И. ТАЛАНОВ

Проблема коллапса волн, описываемого нелинейным уравнением Шредингера (НУШ)

$$i\psi_t + \Delta_D \psi + |\psi|^{2\sigma} \psi = 0, \quad (1)$$

возникла впервые в теории самофокусировки (размерность пространства  $D = 2$ ) в средах с кубичной нелинейностью ( $\sigma = 1$ ) [1, 2] и была достаточно подробно изучена в работах [3—5]. Дальнейшее развитие эта проблема получила в теории ленгмюровского коллапса [6], для которого уравнение (1) играет роль модельного. На модели коллапсирующих каверн строится теория сильной турбулентности плазмы и поглощения волн в нелинейных средах [6]. Об актуальности проблемы коллапса волн свидетельствует и ряд работ, опубликованных в самое последнее время [7, 8].

Возникновение сингулярностей в ходе эволюции решений нелинейных динамических уравнений — явление, характерное для целого ряда физических задач, рассматриваемых без учета механизмов нелинейной или высокочастотной диссипации. Упомянутая выше самофокусировка волновых пучков и самосжатие волновых пакетов в нелинейных средах, самофокусировка электромагнитных волн в плазме, коллапс ленгмюровских волн — примеры таких явлений. Как правило, коллапсом завершается нелинейная стадия модуляционной неустойчивости в консервативных системах высокой размерности. Примером сингулярного поведения диссипативной структуры может служить коллапс решений уравнений со взаимной нелинейной диффузией, в частности хемотаксический коллапс.

Общей чертой для всех подобных систем является конкуренция механизмов нелинейной фокусировки пространственно-временных траекторий квазичастиц (частиц) и их волновой дисперсии (диффузии). Коллапс возникает в тех случаях, когда нелинейная рефракция преобладает над дисперсией на всех стадиях вплоть до образования за конечное время сингулярности. Поскольку конкуренция указанных механизмов зависит от размерности системы, то коллапс наглядно иллюстрирует различный пространственно-временной характер нелинейных процессов в пространстве разной размерности.

В настоящее время существует много терминов, обозначающих явление взрывного возникновения сингулярности: схлопывание, коллапс, нелинейная фокусировка, в иностранной литературе появился термин «blow up» (взрыв) [8]. В процессе образования особенностей формируются и характерные динамические структуры, столь же фундаментальные, как, например, солитоны. Эти структуры, учитывая их профиль, локализованный характер и

особенности динамики, можно назвать взрывающимися солитонами (в многомерном случае — кавитонами).

В настоящей работе обращается внимание на то, что преимущественное поглощение энергии волн коллапсирующими кавернами происходит не на традиционно рассматриваемой стадии формирования особенности  $t < t_0$ , а на следующей за моментом возникновения особенности стадии  $t \geq t_0$ , которую будем называть фокальной. При этом наблюдается качественное различие в пространственно-временной структуре коллапсирующих каверн разной размерности.

## 1. УСТОЙЧИВОСТЬ СОЛИТОНОВ И ДОСТАТОЧНЫЕ УСЛОВИЯ КОЛЛАПСА

Фундаментальными решениями нелинейного уравнения Шредингера являются солитоны — локализованные ( $\|\psi\|_2 < \infty$ ) образования, характеризующиеся равновесием между нелинейной рефракцией и дисперсией. Условия существования и устойчивости (в пределах той же размерности) солитонов ограничивают на плоскости параметров  $(D, \sigma)$  некоторую область. В той части, где солитоны неустойчивы, возможен коллапс. Для общего анализа коллапсирующих решений уравнения (1) разобьем, следуя [8], плоскость параметров  $(D, \sigma)$  на области: I — докритическую  $0 < \sigma < 2/D$ ; II — критическую,  $2/D = \sigma$ ; III — закритическую,  $2/D < \sigma < \infty$  при  $D < 2$  и  $2/D < \sigma < 2/(D - 2)$  при  $D > 2$  [8].

В каждой области могут быть определены однопараметрические солитонные решения (кавитоны) размерности  $D$ :  $\psi_0 = \lambda^{1/2\sigma} \times R_0(\lambda^{1/2}x) e^{i\lambda t}$  с интегралом энергии (числа частиц)  $N = \int |\psi|^2 d^D x$ , равным  $N_0(\lambda) = \lambda^{1/\sigma - D/2} N_D$ , где  $N_D = \int R_0^2(x) d^D x$ . Согласно критерию Вахитова—Колоколова [9], эти решения устойчивы в области I ( $dN_0/d\lambda > 0$ ). В области II они неустойчивы по отношению к конечным возмущениям ( $dN_0/d\lambda = 0$ ), в области III — экспоненциально неустойчивы ( $dN_0/d\lambda < 0$ ). Верхнее значение  $\sigma = 2/(D - 2)$  при  $D > 2$  определяет границу существования солитонных решений. Обычно рассматриваемому случаю кубичной нелинейности ( $\sigma = 1$ ) при  $D = 1$  соответствует область I, при  $D = 2$  — область II, при  $D = 3$  — область III.

Зависимость  $N_0(\lambda)$  можно представить как зависимость  $N$  от характерного размера солитона  $a_0 = \lambda^{-1/2}$ . В докритической области  $dN_0/da_0 < 0$ , в критической области  $dN_0/da_0 = 0$  и, наконец, в закритической области  $dN_0/da_0 > 0$ . Уже отсюда видно, что эволюция локализованных волновых сгустков в разных областях будет различна. Пусть начальное распределение имеет характер солитона с энергией  $N > N_0(a_0)|_{t=0}$ , где  $a_0$  — размер солитона, тогда нелинейная рефракция преобладает над дисперсией и при  $t > 0$  идет сжатие сгустка. В докритической области с уменьшением  $a_0$  сжатие прекратится, так как неминуемо наступит момент, когда  $N$  станет меньше  $N_0(a_0)$ . Этим и обеспечивается устойчивость солитонов в этой области. В критическом случае, когда  $dN_0/da_0 = 0$ ,

начальное преобладание  $N$  над  $N_0$  сохраняется и в процессе сжатия, что приводит к неограниченному уменьшению размеров сгустка (коллапсу). В закритической области превышение  $N$  над  $N_0$  в процессе сжатия только усиливается, т. е. сжатие идет убыстряющимся темпом. Подобные качественные соображения показывают, что в области I коллапс невозможен, здесь фундаментальными решениями нелинейного уравнения Шредингера являются устойчивые солитоны, а задача Коши для нелинейного уравнения Шредингера имеет решение для всех  $t > 0$ . В областях II и III всегда можно указать начальные условия, приводящие к образованию сингулярности за конечное время.

## 2. ТЕОРЕМА ВИРИАЛА И ДОСТАТОЧНЫЕ УСЛОВИЯ КОЛЛАПСА ВОЛН

Формулировка достаточных условий коллапса основывается на теореме вириала (метода моментов [10]). Определим положительную величину, аналогичную моменту инерции относительно центра масс системы тел с распределенной в пространстве плотностью  $\rho = |\psi|^2$ :

$$I(t) = \int (\mathbf{x} - \langle \mathbf{x} \rangle)^2 |\psi|^2 d^D x,$$

где  $\langle \mathbf{x} \rangle = \frac{1}{N} \int \mathbf{x} |\psi|^2 d^D x$  — координата центра «масс» (плотности энергии) волнового сгустка. Согласно теореме вириала (в формулировке работы [8]),

$$\ddot{I} = 4 \left[ 2H + \frac{D}{1+\sigma} \left( \frac{2}{D} - \sigma \right) \int |\psi|^{2\sigma+2} d^D x \right], \quad (2)$$

где  $H = \int \left( |\nabla \psi|^2 - \frac{1}{\sigma+1} |\psi|^{2\sigma+2} \right) d^D x$  — гамильтониан уравнения (1), являющийся интегралом движения.

Согласно (2), при  $D > 2/\sigma$  (критическая и закритическая области) достаточным условием коллапса является отрицательность гамильтониана;  $H < 0$ ; при этом  $I < 0$  и волновой пучок стягивается к центру масс.

Для критического случая  $\ddot{I} = 8H$  и уравнение (2) интегрируется:  $I = 4Ht^2 + Bt + C$ , где  $B$  и  $C$  — константы, определяемые начальной структурой волнового сгустка [10]. В частности, для волнового сгустка типа кавитона  $\psi = AR_0(x)$  имеем  $H < 0$  при  $|A| > 1$  ( $N\{\psi\} > N\{R_0\} = N_0$ ). Следовательно, любой кавитон с  $N > N_0$  (в теории самофокусировки это соответствует мощности пучка больше критической) схлопывается за конечное время. Это время обычно меньше времени глобального коллапса, определяемого условием  $I = 0$ . Таким образом, теорема вириала формулирует лишь достаточные условия возникновения особенности, которая может возникнуть и при  $H > 0$ . Теорема вириала ничего не говорит о фактическом времени возникновения особенности и энергии (мощности), которая в ней концентрируется. Однако уже в рамках

первых численных экспериментов по самофокусировке [11] было замечено, что в окрестности особенности концентрируется мощность порядка критической ( $N = N_0$ ). Более точное подтверждение этому было получено в численном эксперименте, описанном в работе [12].

В закритической области III ( $\sigma > 2/D$ ) второй член в (2) отрицателен и возникновение коллапса облегчается. Условие  $H < 0$  и в этом случае сохраняется как достаточное.

В докритическом случае ( $\sigma < 2/D$ ) даже при  $H < 0$  никаких выводов о существовании коллапса на основании теоремы вириала сделать нельзя. Имеются доказательства, что коллапс в этом случае отсутствует и задача Коши для нелинейного уравнения Шредингера имеет решение для всех  $t > 0$  [13].

Таким образом, плоскость ( $D, \sigma$ ) может быть разбита на область I, где существуют решения нелинейного уравнения Шредингера в виде устойчивых солитонов, а коллапс невозможен, и области II, III, где всегда можно создать начальные условия, приводящие к возникновению особенности в решении нелинейного уравнения Шредингера.

### 3. СТРУКТУРА ПОЛЯ ВБЛИЗИ ОСОБЕННОСТИ

Какова же структура возникающей особенности? Под структурой будем понимать пространственно-временной характер поля в окрестности особой точки. По исходной физической постановке, как правило, условия применимости НУШ в непосредственной окрестности особенности нарушаются, и изучение собственно сингулярного поведения его решений становится, строго говоря, неправомерным. Однако, как показывают конкретные оценки, выход на асимптотику поля около особенности возможен и в условиях применимости НУШ, поэтому асимптотические формулы могут быть использованы, например, в оценках разного рода вторичных нелинейных эффектов в окрестности особенности. Кроме того, исследование особых решений нелинейного уравнения Шредингера принципиально интересно для понимания всей структуры решения. Сингулярные решения играют здесь примерно ту же роль, что и солитоны в докритической области: на них в конечном итоге разбивается начальное волновое поле.

Обычно принято обсуждать структуру особенности в момент времени  $t < t_0$ , непосредственно предшествующий моменту возникновения сингулярности. Такой подход оправдан в известной степени лишь в критическом случае ( $D\sigma = 2$ ), когда коллапс носит мгновенный характер, и при  $t > t_0$  поле вновь регулярно. В закритической области ( $D\sigma > 2$ ) коллапс может приобрести распределенный во времени характер, т. е. сингулярность будет существовать конечное время за моментом  $t = t_0$ . Это делает принципиально интересным и изучение чисто сингулярных решений НУШ.

В настоящее время структура особенности достаточно строго обоснована лишь в критическом случае, к которому, в частности, относится теория самофокусировки волн в кубичной среде [4, 10, 14].

Начиная с работы [14], построение асимптотики поля в окрестности фокальной точки  $t_0$  производится с помощью так называемого обобщенного преобразования фокусировки, которое для симметричного случая  $\psi = \psi(r, t)$  запишем при произвольных значениях размерности и нелинейности в виде

$$\psi = g(s) V(\eta, s) \exp\left(-\frac{i}{2D} \frac{g'_s}{g} \eta^2\right), \quad (3)$$

где  $\eta = rg^\sigma(s)$ ;  $s_i = g^{2\sigma}(t)$ . В области  $t < t_0$ , где поле регулярно, его амплитуда в центре  $r = 0$  удовлетворяет соотношению, вытекающему из уравнения (1):

$$|\psi(0, t)|'_t = -D\varphi''_{rr} |\psi(0, t)|, \quad (4)$$

где  $\varphi = \arg \psi(r, t)$ .

Из (3) и (4) следует, что  $|V(0, s)| = \text{const}$ , а  $[\arg V(0, s)]''_{\eta\eta} = 0$ . Таким образом, все изменение амплитуды поля в точке  $r = 0$  в дофокальной области описывается функцией  $g(s)$ , которая, по условию, должна стремиться к бесконечности при  $t \rightarrow t_0$ . Из определения зависимости  $s(t) = \int_t^t g^{2\sigma} dt$  следует, что при достаточно быстрой расходимости поля ( $g^{2\sigma} > (t_0 - t)^{-\alpha}$ ,  $\alpha \geq 1$ ) преобразование (5) относит фокальную точку  $t = t_0$  на бесконечность.

Подстановка (3) в (1) приводит к уравнению для  $V(\eta, s)$ :

$$iV_s + \Delta_{\eta\eta} V + |V|^{2\sigma} V + ib(s) \eta \partial V / \partial \eta + a(s) \eta^2 V = 0, \quad (5)$$

где  $b(s) = (g'_s/g)(\sigma - 2/D)$ ;  $a(s) = g''_{ss}/2Dg + (2\sigma D - D - 2)/12D^2g^2$ .

При  $\sigma = 2/D$  (область II) коэффициент  $b(s)$  в (5) обращается в нуль. Можно потребовать, чтобы  $a(s) \rightarrow 0$  при  $s \rightarrow \infty$ . Когда  $a(s)$  достаточно мало, в области  $a(s) \eta^2 \ll 1$  уравнение (5) сводится к виду (1) и имеет решение  $R_0(\eta) e^{is}$ . В области  $a(s) \eta^2 \gg |V|^{2\sigma}$  нелинейность несущественна и решение (5) носит излучательный характер:

$$V \approx g^{-1} \psi_c \exp\left(\frac{i}{2D} \eta^2 \frac{g'_s}{g} + i \int \frac{|\psi_c|^{2\sigma}}{g^{2\sigma}} dt\right), \quad (6)$$

где  $\psi_c = \text{const}$ . В этой области амплитуда  $V$  стремится к нулю как  $g^{-1}$ . Сшивая решение (6) с  $R_0(\eta) e^{is}$  в некоторой промежуточной точке  $\bar{\eta}(s)$ , получим уравнение для определения функции  $a(s)$ , посредством которого определяется  $g(s)$ :

$$\exp(-1/\sqrt{a(s)}) g(s) = \psi_c \quad (7)$$

с асимптотическим при  $s \rightarrow \infty$  решением  $a(s) = c_0/s$ ,  $g(s) = \sqrt[4]{s} e^{\sqrt{s}}$ , где  $c_0$  — постоянная. При этом  $g(t) = \sqrt{\frac{|\ln(t_0 - t)|}{t_0 - t}}$ . При  $\sigma > 2/D$  (область III) в (7) можно сделать замену

$$V = V_0 \exp\left[-\frac{ig'_s}{4g} \left(\sigma - \frac{2}{D}\right) \eta^2\right],$$

при которой (7) перейдет в

$$iV_0' s + \Delta_{\eta\eta} V_0 + |V_0|^{2\sigma} V_0 - (Db/2)V_0 + \tilde{a}\eta^2 V_0 = 0, \quad (8)$$

где

$$\tilde{a}(s) = \frac{g''_{ss}}{4g} \sigma + \frac{g'^2_s}{4g^2} (\sigma^2 - \sigma).$$

Решение (8) вблизи оси не совпадает с  $R_0$ , и поэтому рассуждения, аналогичные данным выше для случая  $\sigma D = 2$ , провести не удастся. Однако можно заметить, что стационарные режимы (8) при  $\sigma D > 2$  могут существовать только при  $b$  и  $\tilde{a}$ , слабо зависящих от  $s$ . Только в этом случае возможно уравновешивание усиления, описываемого членом, содержащим  $b(s)$ , затуханием из-за излучения, возникающего из-за наличия члена  $\tilde{a}\eta^2$ . При  $b = \text{const}$  и  $\tilde{a} = \text{const}$  имеем  $g = e^s$  и  $g(t) = 1/\sqrt{t_0 - t}$ . Решение уравнения (5) при такой зависимости  $g(t)$  является автомодельным вблизи  $\eta = 0$  [1]. Вдали от этой точки оно должно перейти в решение (6), не являющееся автомодельным, что может привести к поправкам к  $g(t)$ .

#### 4. СЛАБЫЙ И СИЛЬНЫЙ КОЛЛАПС

Коллапс с конечным значением первого интеграла  $N = N_0$  (число частиц, мощность, энергия) получил название сильного [7]. Его лучевая структура схематически представлена на рис. 1, а. Коллапсирующая особенность носит мгновенный (безабберационный) характер и представляется на временной оси отдельной точкой. В критическом случае при  $N > N_0$  и колоколообразном профиле начального волнового сгустка коллапс имеет многофокусный характер с захватом в каждую особенность значения интеграла  $N = N_0$  [4]. Энергия первоначального сгустка в конечном счете поглощается (или рассеивается) в виде отдельных вспышек, каждая из которых представляет собой взрывающийся ( $g \rightarrow \infty$ ) кавитон. Коллапс, при котором число захваченных частиц (энергии)  $N$  стремится к нулю при  $t \rightarrow t_0$ , был назван слабым в отличие от критического случая [7]. На первый взгляд имеется принципиальное отличие между коллапсом в критической области (сильным) и в закритической области (слабым): в первом случае в особенность захватывается конечное число частиц  $N_0$ , определяемое только параметрами уравнения ( $D, \sigma$ )<sup>1</sup>, во втором — захваченное к моменту  $t \rightarrow t_0$  число частиц (энергия) равно нулю.

Отсутствие формально к моменту  $t \rightarrow t_0$  захвата в схлопывающуюся каверну конечной энергии волнового поля с самого начала затрудняло построение теории сильной турбулентности и поглощения волн на основе модели коллапсирующих сгустков. Было предложено несколько путей обхода этой трудности.

<sup>1</sup> В теории самофокусировки волн в кубичной среде это так называемая критическая мощность, зависящая от параметров среды и частоты волны.



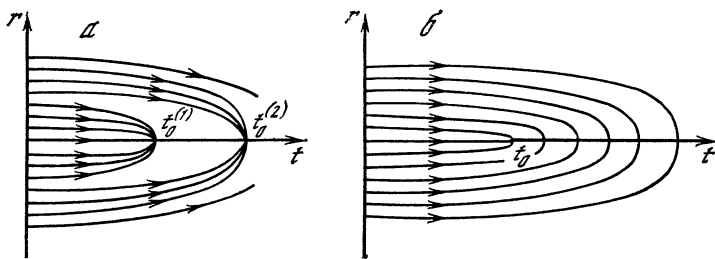


Рис. 1. Пространственно-временная фокусировка в критической (а) и закритической (б) областях

Один из них основан на том, что при слабом коллапсе фактическое поглощение энергии (первого интеграла) определяется тем значением амплитуды поля  $\psi_0$ , при котором уравнение (1) становится неприменимым, а в одном акте коллапса это количество примерно равно ( $D=3$ )

$$\Delta N = \psi_0^2 \xi_0^3 \sim 1/\psi_0,$$

где  $\xi_0 \sim 1/\psi_0$  — характерный размер области поглощения [7]. В такой модели количество поглощенной энергии в одном акте тем меньше, чем выше порог нелинейного поглощения  $\psi_0$ . Другой рецепт устранения слабости трехмерного коллапса предложен в [15]. Предполагается, что коллапс носит квазиклассический характер (т. е. описывается уравнениями геометрической оптики) вплоть до области сильного поглощения. Если считать, что число частиц сохраняется, то при  $t \rightarrow t_0$  условия применимости квазиклассики только улучшаются'

$$|\psi_0|^2 a^2 \sim 1/a \rightarrow \infty$$

при размере  $a$ , стремящемся к нулю. В этом случае в особенность могла бы пойти вся энергия, запасенная в первоначальном сгустке. Заметим, что аналогичная ситуация должна бы иметь место и в двумерном квазиклассическом коллапсе (теории самофокусировки [7]), при котором начальные условия  $|\psi|^2 a^2 \gg 1$  сохраняются для всех значений  $t$  вплоть до  $t_0$ . Тем не менее количество поглощенной энергии точно равно  $N_0$ .

Кажущееся противоречие разрешается, если учесть, что квазиклассический коллапс неустойчив по отношению к коротковолновым возмущениям, причем неустойчивость носит взрывной характер. В двумерном случае эта неустойчивость приводит к возникновению особенности слабого типа. В связи с этим высказывалось предположение, что трехмерный коллапс в рамках нелинейного уравнения Шредингера принципиально несимметричен и сопровождается стохастическим разбрызгиванием коллапсирующей каверны на более мелкие, каждая из которых будет в процессе коллапса так же стохастически дробиться. При этом поглощение первоначально-

го сгустка осуществляется всем ансамблем каверн. Как отмечалось в [7], модель множественного дробления коллапсирующих каверн пока еще не получила подтверждений в численном эксперименте.

## 5. ТРЕХМЕРНЫЙ РАСПРЕДЕЛЕННЫЙ КОЛЛАПС

Мы хотели бы обратить внимание на то что с точки зрения поглощения в коллапсирующей каверне энергии (числа частиц  $N$ ) первоначального волнового сгустка принципиального различия между режимом сильного и слабого коллапса нет, по крайней мере в трехмерном случае.

Дело в том, что при  $D = 3$  слабому коллапсу с формально нулевой поглощаемой в особенности энергией соответствует лишь начальная дофокальная стадия  $t < t_0$ , собственно поглощение в особенности происходит на более поздней стадии ( $t > t_0$ ) и носит распределенный во времени характер. Это наглядно иллюстрируется картиной пространственно-временных лучей на рис. 1, б. Лучи при  $r \neq 0$  остаются ориентированными к центру каверны и после возникновения в ней особенности поля. Это значит, что поток энергии в особенность сохраняется и за точкой начального фокуса  $t = t_0$  (фокальная «нить» на рис. 1, б).

Таким образом, в закритическом случае слабый при  $t < t_0$  коллапс переходит в сильный при  $t > t_0$ . Время жизни особенности зависит от начального профиля волнового сгустка, его энергии и положения точки  $(D, \sigma)$  в области III. В критическом случае ( $D\sigma = 2$ ) вся стадия поглощения сжимается к точке  $t = t_0$  — коллапс становится мгновенным. При значительных превышениях  $N$  над  $N_0$  лучевая картина носит многофокусный характер с поглощаемой в каждом фокусе энергией, равной  $N_0$ . При незначительном отклонении от критического случая ( $D\sigma \gtrsim 2$ ) можно ожидать возникновения в центре каверны серии поглощающих особенностей конечной длительности («многонитевая» структура особенности), которые сольются во времени при уходе точки  $(D, \sigma)$  в глубь закритической области. Что касается поглощенной энергии первоначального сгустка, принципиальной разницы между критической и закритической областями нет. В обоих случаях поглощение происходит на стадии сильного коллапса  $t \gtrsim t_0$ ; в критическом случае — мгновенно порциями  $N_0$ , в закритическом — непрерывно или серией вспышек конечной длительности.

В наиболее интересном случае трехмерного симметричного коллапса в кубичной среде ( $D = 3, \sigma = 1$ ), описываемого уравнением

$$i\psi_t + \psi_{rr} + (2/r)\psi_r + |\psi|^2\psi = 0, \quad (9)$$

при постоянстве потока энергии, подтекающего к особенности закон изменения по  $r$  амплитуды и фазы поля  $\psi = \psi_0 \exp(+i\phi - iht)$  в окрестности особенности универсален:

$$\psi_0 = 1/r |\ln r|^{1/2}; \quad \psi_0^2 r^2 = C,$$

где константа  $C$  определяется величиной потока энергии. Из сказанного ясно, что при изучении поглощения энергии коллапсирующей каверной наибольший интерес представляет фокальная стадия коллапса  $t \geq t_0$ . Однако именно здесь из-за наличия особенности поля при  $r = 0$  трудно применить традиционные при  $t < t_0$  численные методы. Чтобы избежать указанной трудности, предположим что в окрестности  $r = 0$  сосредоточено сильно локализованное многофотонное поглощение. Для этого вместо (9) будем интегрировать уравнение

$$i\psi_t + \psi_{rr} + (2/r)\psi_r + |\psi|^2\psi + i\alpha|\psi|^2\psi = 0 \quad (10)$$

при достаточно малом значении многофотонного поглощения  $\alpha$ . На рис. 2 показаны изменения амплитуды поля в точке  $r = 0$  в зависимости от времени  $t$  при  $n = 8$ ,  $\alpha_0 \approx 3,9 \cdot 10^{-11}$  и начальном условии  $\psi = 7e^{-1,125r^2}$ . Для сравнения там же приведена аналогичная зависимость для двумерного коллапса при том же начальном условии. Видно, что трехмерный коллапс имеет распределенный характер, а двумерный — типично многофокусный характер.

Это подтверждается и зависимостью от времени полной энергии  $N(t)$  (рис. 3). В отличие от ступенчатого изменения полной энергии при двумерном коллапсе энергия трехмерного сгустка монотонно убывает за точкой  $t = t_0$  вследствие поглощения в окрестности  $r = 0$ .

Таким образом, в закритической области III трехмерный коллапс при  $\sigma = 1$  носит принципиально распределенный во времени характер. Диссипация энергии начального волнового сгустка происходит при этом не на дофокальной стадии, а после возникновения особенности, в том числе и при наличии сильного нелинейного поглощения. Можно ожидать, что и при несимметричных начальных условиях образующиеся коллапсирующие каверны будут выгорать таким же образом, преимущественно на фокальной ( $t > t_0$ ) стадии.

В заключение заметим, что эффект распределенного во времени коллапса в рамках НУШ с локальной нелинейностью лишь

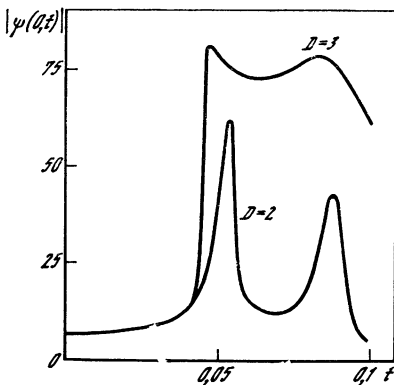
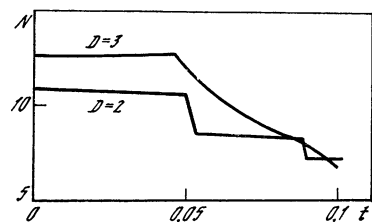


Рис. 2. Зависимость амплитуды поля в центре каверны от времени

Рис. 3. Зависимость полной энергии каверны от времени



внешне похож на «эффект воронки», описанный для коллапса ленгмюровских волн в [16]; последний требует более сильной особенности поля в окрестности  $r = 0$  и не реализуется в распределенном коллапсе, несимметричные возмущения каверн при распределенном коллапсе в особенность не захватываются.

#### ЛИТЕРАТУРА

1. Таланов В. И. О самофокусировке волновых пучков в нелинейных средах // Письма в ЖЭТФ. 1965. Т. 2, № 5. С. 218—222.
2. Kelly P. L. Self-focusing of optical beams // Phys. Rev. Lett. 1965. Vol. 15, N 26. P. 1005—1008.
3. Ахманов С. А., Сухоруков А. П., Хохлов Р. В. Самофокусировка и дифракция света в нелинейной среде // УФН. 1967. Т. 93, № 1. С. 19—70.
4. Луговой В. Н., Прохоров А. М. Теория распространения мощного лазерного излучения в нелинейной среде // Там же. 1973. Т. 111, № 62. С. 203—247.
5. Progress in quantum electronics. Vol. 4 / Ed. J. H. Sanders, S. Stenholm. L.: Pergamon press, 1975. 110 p.
6. Захаров В. Е. Коллапс ленгмюровских волн // ЖЭТФ. 1972. Т. 62, № 5. С. 1745—1750.
7. Захаров В. Е., Кузнецов Е. А. Квазиклассическая теория трехмерного волнового коллапса // Там же. 1986. Т. 91, № 4 (10). С. 1310—1324.
8. Rypdal K., Rasmussen J. J. Blow up in nonlinear Schrödinger equations. I, II // Phys. scr. 1986. Vol. 33, N 6. P. 481—504.
9. Вахитов Н. Г., Колоколов А. А. Стационарные решения волнового уравнения в среде с насыщением нелинейности // Изв. вузов. Радиофизика. 1973. Т. 16, № 7. С. 1020—1028.
10. Власов С. Н., Петрищев В. А., Таланов В. И. Усредненное описание волновых пучков в линейных и нелинейных средах (метод моментов) // Там же. 1971. Т. 14, № 9. С. 1353—1363.
11. Гольдберг В. Н., Таланов В. И., Эрм Р. Э. Самофокусировка аксиально-симметричных волновых пучков // Там же. 1967. Т. 10, № 5. С. 674—686.
12. Дышко А. Л., Луговой В. Н., Прохоров А. М. Самофокусировка интенсивных световых пучков // Письма в ЖЭТФ. 1967. Т. 6, № 5. С. 655—657.
13. Weinstein M. I. Nonlinear Schrödinger equations and sharp interpolation estimates // Commun. Math. Phys., 1982. Vol. 87, N 4. P. 567—576.
14. Власов С. Н., Пискунова Л. В., Таланов В. И. Структура поля вблизи особенности при самофокусировке // ЖЭТФ. 1978. Т. 75, № 5(11). С. 1602—1607.
15. Захаров В. Е., Кузнецов Е. А., Мушер С. Л. О квазиклассическом решении трехмерного волнового коллапса // Письма в ЖЭТФ. 1985. Т. 41, № 3. С. 125—127.
16. Захаров В. Е., Шур Л. Н. Об автомодельных режимах волновых коллапсов // ЖЭТФ. 1981. Т. 81, № 6(12). С. 2019—2031.

## НЕУСТОЙЧИВОСТИ И СТРУКТУРЫ В КОГЕРЕНТНЫХ НЕЛИНЕЙНО-ОПТИЧЕСКИХ СИСТЕМАХ, ОХВАЧЕННЫХ ДВУМЕРНОЙ ОБРАТНОЙ СВЯЗЬЮ

С. А. АХМАНОВ, М. А. ВОРОНЦОВ

При исследованиях сильных нелинейных эффектов в оптике основное внимание уделяется системам с внешней обратной связью. Простейшая система такого типа — резонатор Фабри—Перо, заполненный практически безынерционной слабонелинейной средой с кубической нелинейностью; даже слабая нелинейность показателя преломления приводит здесь к бистабильности. Помимо этого ставшего теперь уже тривиальным поведения, оптический резонатор, заполненный нелинейной средой, обнаруживает и более интересные эффекты.

Икеда [1] показал, что при достаточно сильных полях, таких, при которых нелинейный фазовый сдвиг в резонаторе становится сравнимым с  $2\pi$ , в системе возникают неустойчивости и хаос. Хотя саму оптическую бистабильность удается теперь наблюдать в весьма разнообразных условиях, в том числе с помощью непрерывных и квазинепрерывных лазеров малой мощности, число экспериментов по оптическим неустойчивостям и хаосу пока еще невелико.

Исследуемые в подавляющем большинстве работ оптические резонаторы, заполненные средой с кубической нелинейностью, и в особенности стандартные гибридные системы с электрооптическими кристаллами, далеко не в полной мере позволяют понять удивительно многообразную динамику нелинейных оптических систем. Действительно, в них фактически не используется трехмерный характер оптического поля, а наблюдаемые эффекты во многом аналогичны эффектам в нелинейных системах с сосредоточенными постоянными (единственное исключение — так называемая поперечная оптическая бистабильность [2]).

В предлагаемой работе главный акцент делается на специфике оптических задач. Мы ограничимся в основном системами, где нелинейность определяется прежде всего зависимостью действительной части показателя преломления от интенсивности света; в них неустойчивости и хаос теснейшим образом связаны с четырехфотонными параметрическими процессами.

Оптический сигнал по сравнению, например, с электромагнитным, распространяющимся в электрических цепях, имеет существенное отличие: оптическое поле распределено в пространстве. Комплексная амплитуда вектора напряженности электрического поля световой волны  $A(\rho, t)$  зависит не только от времени  $t$ , но и от пространственных переменных  $\rho = \{x, y, z\}$ . В каждом фиксированном сечении  $z = \text{const}$  комплексная амплитуда поля  $A$  определяется двумерным вектором  $\mathbf{r} = \{x, y\}$ , лежащим в плоскости,

перпендикулярной направлению распространения волны  $OZ$ . В этом смысле можно говорить о двумерном характере сигнала в оптической системе.

Естественным обобщением идеи обратной связи в оптических системах может являться двумерная оптическая обратная связь. На рис. 1 показана структурная схема некоторой оптической системы с двумерной обратной связью ( $G$  — некоторый оператор преобразования светового поля в контуре обратной связи). Если операторы  $G$  и  $W$  реализуют поточечное преобразование поля, то такого рода система с двумерной обратной связью фактически распадается на континуум независимых систем с одномерной обратной связью и может быть проанализирована точно так же, как и обычные радиофизические системы. Примером такой организации обратной связи являются бистабильные оптические устройства [3].

Для оптики гораздо более характерны невырожденные преобразования координат. Например, при распространении света в системе из-за дифракции происходит эффективное «перемешивание» координат. В результате поточечное рассмотрение часто вообще теряет смысл [2]. Такую систему следует описывать дифференциальными уравнениями в частных производных.

Заметим, что даже сравнительно простые преобразования поля (изменение масштаба светового пучка в контуре обратной связи, сдвиг и т. д.) не являются вырожденными.

Нелинейные оптические системы с пространственным преобразованием поля в контуре обратной связи интересны в нескольких отношениях. Динамические процессы в таких системах отличаются сложностью и многообразием [4]. Являясь обобщением известных нелинейных систем, они представляют несомненный интерес с точки зрения развития физики нелинейных явлений. Оптические системы с двумерной обратной связью привлекательны и в связи с развитием оптических компьютеров, адаптивных систем коррекции и формирования световых пучков. Они идеально подходят для реализации многочисленных итеративных схем решения многомерных задач [5].

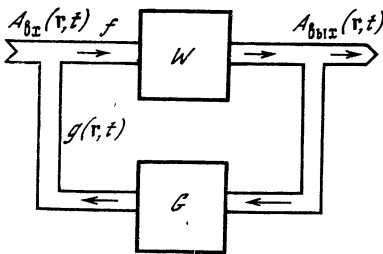


Рис. 1. Структурная схема системы с распределенной обратной связью

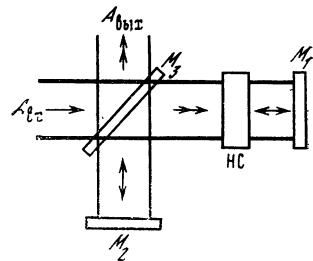


Рис. 2. Нелинейный интерферометр Физо

$M_1, M_2$  — зеркала,  $M_3$  — светоделитель, HC — тонкий слой нелинейной среды

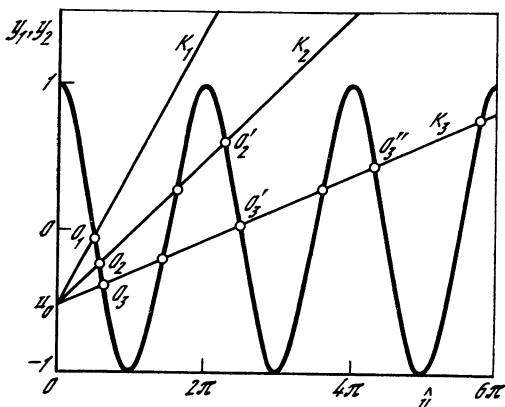


Рис. 3. Графическое решение уравнения (4)

$y_1 = \cos u$ ,  $y_2 = (u - u_0)/K$ .  
 $K = K_1$  — единственная стационарная точка,  $K = K_2$  — два устойчивых состояния  $O_2, O_2'$  (оптическая бистабильность),  $K = K_3$  — устойчивые точки  $O_3, O_3', O_3''$  (мультистабильность);  
 $K_3 > K_2 > K_1$

Пожалуй, наиболее изученной системой с оптической обратной связью является нелинейный интерферометр Фабри—Перо [3]. В такой системе для различных типов нелинейности среды зависимость интенсивности выходного поля  $I_{\text{вых}}$  от интенсивности поля на входе в систему  $I_{\text{вх}}$  носит гистерезисный характер. Обратная связь в такой системе является внутренней и реализуется при отражении поля от зеркал интерферометра. Если не учитывать дифракционные эффекты (что и делалось в первых работах), то здесь мы имеем дело с простейшим типом вырожденного преобразования поля в контуре обратной связи. Система как бы распадается на множество идентичных одномерных систем с обратной связью. Учет дифракции поля или пространственной диффузии нелинейной среды приводит к нелокальной связи полей. В результате происходит изменение поперечной структуры излучения (поперечная бистабильность).

Возьмем за основу простейшую модель нелинейной системы, усложняя которую перейдем к исследованию достаточно общего случая двумерной обратной связи с невырожденным преобразованием поля. На рис. 2 показана схема нелинейного интерферометра Физо. Входная волна  $A_{\text{вх}}$  предполагается плоской, монохроматической и линейно поляризованной. Отраженная от зеркал  $M_1$  и  $M_2$  волна интерферирует с входным полем  $A_{\text{вх}}$ . Будем полагать, что в пространство между зеркалами интерферометра помещен тонкий слой нелинейной среды НС. Изменение показателя преломления среды под действием поля будем описывать релаксационным уравнением

$$\tau \frac{dn}{dt} + n = n_0 + n_2 I_c, \quad (1)$$

где  $I_c$  — интенсивность поля в среде. Характерное время релаксации показателя преломления  $\tau$  полагаем много больше времени распространения волны в системе, что позволяет не учитывать запаздывания волны.

Модуляция показателя преломления среды создается волной, интенсивность которой определяется соотношением

$$I_c = (I_1 + I_2) (1 + 2\gamma \cos u), \quad (2)$$

где  $\gamma = (I_1 I_2)^{1/2} / I_0$  — видность интерференционной картины;  $I_1 = (1 - R)I_{\text{вх}}$ ;  $I_2$  — интенсивности интерферирующих волн;  $I_0 = I_1 + I_2$ ;  $u = 2dkn$  — набег фазы волны при распространении в среде;  $d$  — толщина слоя нелинейной среды;  $k = 2\pi/\lambda$  — волновое число;  $R$  — коэффициент отражения зеркал.

С точностью до постоянного множителя интенсивность выходного поля  $I_{\text{вых}}$  совпадает с  $I_c$ . Из (1), (2) легко получить уравнение, определяющее динамику изменения фазы интерферирующих волн:

$$\tau du/dt + u = u_0 + K \cos u, \quad (3)$$

где  $u_0 = 2kd(n_0 + n_2 I_c)$ ;  $K = 4kdn_2 \gamma I_0$  — коэффициент усиления в контуре обратной связи. Стационарные состояния системы  $\hat{u}$  определяются из уравнения

$$(\hat{u} - u_0)/K = \cos \hat{u}. \quad (4)$$

Графическая интерпретация этого уравнения показана на рис. 3. Так же как и в интерферометре Фабри—Перо, в рассматриваемой системе при сильных нелинейностях (больших коэффициентах усиления  $K$ ) существует множество устойчивых стационарных состояний поля, между которыми возможны переходы, сопровождающиеся скачкообразным изменением фазы световой волны.

Собственно говоря, сама по себе описанная система не представляет сколь-нибудь значительного практического интереса. Она интересна в другом отношении. На ее базе сравнительно просто организовать различные преобразования поля в контуре обратной связи. Заменяв зеркало  $M_2$  в схеме, изображенной на рис. 2, призмой П реализуем простейший тип преобразования поля (зеркальное отражение относительно оси, проходящей через центр пучка). В рамках теории плоских волн такое преобразование является вырожденным. В системе оказываются попарно связанными точки, расположенные на равном расстоянии относительно оси поворота поля (на рис. 4, *a* точки 1 и 2). Аналогичные системы исследовались в работах [6, 7].

Обозначим через  $u_1$  и  $u_2$  фазовые набег в нелинейной среде в точках 1 и 2 поперечника пучка. По аналогии с (3) легко записать соответствующие уравнения для фазовых функций  $u_1(t)$  и  $u_2(t)$ :

$$\tau du_1/dt + u_1 = \varphi_1 + K_1 \cos u_2, \quad (5)$$

$$\tau du_2/dt + u_2 = \varphi_2 + K_2 \cos u_1,$$

где через  $\varphi_1$  и  $\varphi_2$  обозначены стационарные сдвиги фазы волны в точках 1 и 2;  $K_1, K_2$  — соответствующие коэффициенты усиления.

Положим для простоты  $\varphi_1 = \varphi_2 = 2\pi m$ ,  $m$  — целое,  $K_1 = K_2 = K$ . Стационарные состояния  $\hat{u}_1, \hat{u}_2$  определяются из решения уравнений

$$\hat{u}_1 = K \cos \hat{u}_2, \quad \hat{u}_2 = K \cos \hat{u}_1,$$

или

$$\hat{u}_1 = K \cos [K \cos \hat{u}_1]. \quad (6)$$



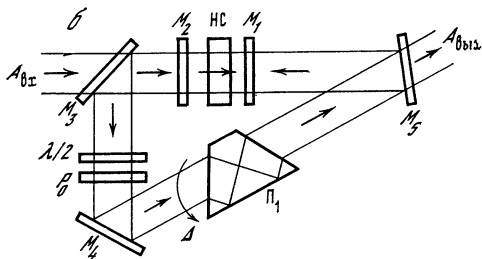
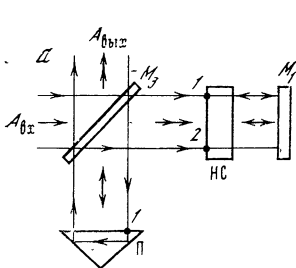


Рис. 4. Пространственное преобразование поля в контуре обратной связи нелинейного интерферометра Физо

а — зеркальное отображение; б — поворот поля на угол  $\Delta$

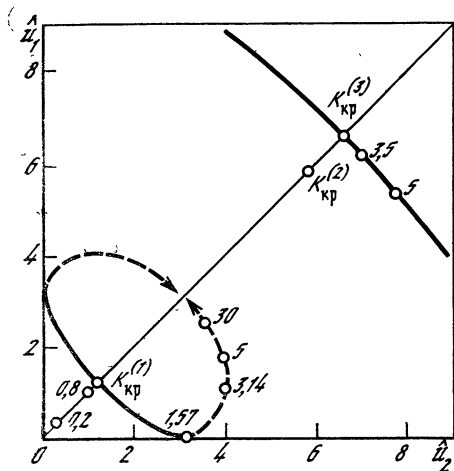


Рис. 5. Стационарные состояния системы (5)

Сплошная кривая — точки типа «устойчивый узел», пунктирная — «устойчивый фокус»



Рис. 6. Пространственная структура светового пучка в контуре обратной связи системы при зеркальном отображении изображения с помощью крышеобразной призмы (см. рис. 4, а)

а — исходный пучок; б — пространственная бистабильность, в — оптическая мультистабильность

На рис. 5 изображены кривые, состоящие из стационарных точек, соответствующих различным коэффициентам  $K$ . При  $K < K_{кр}^{(1)}$  стационарные фазовые сдвиги равны. Тем самым интенсивность выходного поля остается пространственно однородной. При  $K > K_{кр}^{(1)}$  интенсивности в точках, расположенных симметрично относительно оси зеркального отображения поля в контуре

обратной связи, различны. При  $K = \pi/2$  достигается максимальное различие интенсивностей. На рис. 6 показаны распределения интенсивности выходного поля, полученные в экспериментах с гибридной системой, которая представляет собой модель описанной системы.

При  $K > K_{\text{кр}}^{(2)}$  существует несколько устойчивых точек (см. рис. 5), число которых возрастает с увеличением коэффициента усиления  $K$ . При больших значениях  $K$  в системе возникает неустойчивость пространственной структуры светового пучка. На рис. 6, в приведено распределение интенсивности, соответствующее этому случаю. Конечно, для описания пространственных структур, изображенных на рис. 6, в, геометрическое описание непригодно. Необходимо учитывать дифракцию светового пучка, которая приводит к пространственной связи точек. Вырождение преобразования поля в контуре обратной связи снимается.

Учесть дифракционные эффекты можно и чисто феноменологически, вводя в уравнения (3), (5) поперечную диффузию фазы:

$$\tau \partial u(\mathbf{r}, t) / \partial t + u(\mathbf{r}, t) = d \Delta_{\perp} u(\mathbf{r}, t) + K(x, -y) \cos [u(x, -y, t) + \varphi_0],$$

$$\Delta_{\perp} = \partial^2 / \partial x^2 + \partial^2 / \partial y^2. \quad (7)$$

Наличие знака минус у аргумента функций  $K(\mathbf{r})$  и  $u(\mathbf{r}, t)$  в правой части уравнения (7) связано с учетом поворота поля вокруг оси  $OX$  на  $180^\circ$ . Коэффициент диффузии  $d$  можно связать с характерной шириной дифракционного расплывания  $(\lambda L)^{1/2}$  ( $L$  — оптическая длина пути в системе). В ряде случаев «перемешивание» точек связано с пространственной диффузией нелинейной среды [2].

Рассмотрим более сложный тип преобразования поля в контуре обратной связи. Поместим между зеркалами  $M_4$  и  $M_5$  нелинейного интерферометра Физо призму Дове, с помощью которой повернем поле в контуре обратной связи на угол  $\Delta$  (см. рис. 4, б). Однонаправленное распространение волны в оптической схеме реализуется с помощью пластинки  $\lambda/2$  и соответствующим образом ориентированного поляризатора  $P_0$ . При  $\Delta = 2\pi/m$ , где  $m$  — целое число, имеем вырожденное преобразование (имеется в виду описание на языке плоских волн). Связанными оказываются  $m$  точек. При учете диффузии обратная связь перестает быть вырожденной. Легко показать, что уравнение примет вид

$$\tau \partial u / \partial t + u = \partial \Delta_{\perp} u + K(r, \theta) \cos [u(r, \theta - \Delta, t) + \varphi_0], \quad (8)$$

где  $u = u(r, \theta, t)$ ;  $\Delta_{\perp} = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}$  — оператор Лапласа в цилиндрических координатах. Система описывается дифференциальным уравнением в частных производных со сдвинутым аргументом. Эффекты, связанные с поворотом изображения

в системе с телевизионной обратной связью, обсуждаются в работе [5].

Заметим, что если характерное время релаксации нелинейного отклика среды  $\tau$  имеет тот же порядок, что и время распространения света в оптическом тракте системы, то необходимо учитывать запаздывание (сдвиг аргумента по времени). В простейшей системе с поточечным преобразованием поля (аналогичной рассмотренной выше) запаздывание вызывало возникновение стохастической неустойчивости выходной интенсивности [8]. Поворот поля приводит к сдвигу пространственного аргумента.

Перейдем к анализу частного случая уравнения (8). Ограничимся одномерной задачей:  $u = u(\theta, t)$ ,  $K = \text{const}$ ,  $\varphi_0 = \pi/2$ ,

$$\tau \partial u / \partial t + u = d \partial^2 u / \partial \theta^2 + K \sin u(\theta - \Delta, t) \quad (9)$$

с периодическими граничными условиями  $u(0, t) = u(2\pi, t)$ ,  $\partial u(0, t) / \partial \theta = \partial u(2\pi, t) / \partial \theta$ . Ситуацию, описываемую этим уравнением, можно реализовать в системе, изображенной на рис. 4, б, если на одно из зеркал интерферометра наклеить маску в виде сравнительно узкого кольца.

В системе (9) возможно возникновение движущихся периодических структур. Пространственный период и скорость движения зависят от величины сдвига  $\Delta$  и коэффициента усиления  $K$ . В линейном приближении угловая скорость перемещения структур описывается соотношением

$$\omega_m = K \sin(m\Delta), \quad (10)$$

где  $m = 1, 2, 3, \dots$  — числа, характеризующие пространственный период. Численные исследования показали, что оценка скорости движения (10) имеет место и в общем случае. При угле поворота  $\Delta$ , таком, что  $\sin(m\Delta) = 0$ , структуры неподвижны.

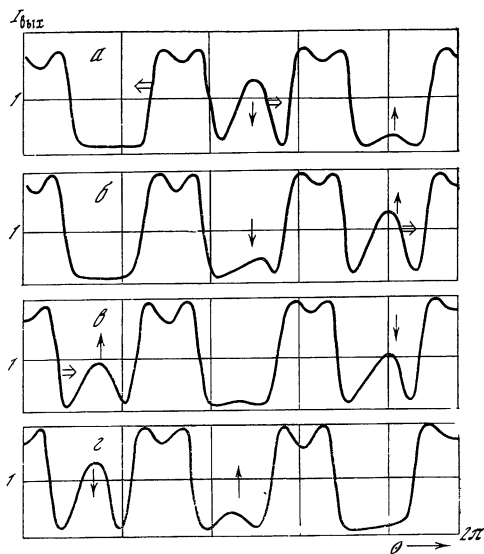
На рис. 7 показана зависимость  $I_{\text{вых}}(\theta) = 1 + \cos u(\theta)$  в различные моменты времени  $t$ . Наряду с ярко выраженными движущимися пиками существует образование, перемещающееся навстречу с существенно большей скоростью и практически не взаимодействующее с основной структурой. Следует отметить, что изображенная на рис. 4, б система с поворотом поля является хорошей моделью для исследования сложных автоволновых явлений. Круговая симметрия задачи обеспечивает возможность наблюдать автоволны практически неограниченное время.

Для экспериментального исследования процессов в нелинейных системах с двумерной оптической обратной связью использовалась структура фотопроводник—диэлектрическое зеркало—слой жидкого кристалла [9]. Схема установки показана на рис. 8.

Модуляция показателя преломления слоя жидкого кристалла осуществлялась с помощью управляющего света, падающего на фотопроводник. Сопротивление фотопроводника изменялось пропорционально распределению интенсивности управляющего света. В результате напряжение, приложенное к тонкому ( $\sim 10$  мкм) слою жидкого кристалла, оказывалось пространственно-проду-

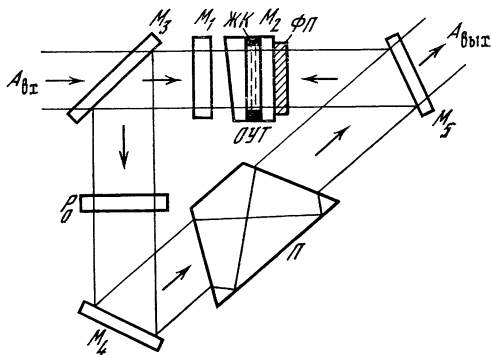
**Рис. 7.** Распределение интенсивности  $I_{\text{вых}}(\theta)$  в системе с поворотом поля на угол  $\Delta = 2,16$  рад в различные моменты времени при  $K = 6$  (численное моделирование)

$a - t = 8\tau$ ;  $b - 8,5\tau$ ;  $c - 9\tau$ ;  
 $z - 9,5\tau$



**Рис. 8.** Схема экспериментальной установки

$M_1, M_2$  — зеркала интерферометра Физо,  $M_3$  — светоделитель,  $M_4, M_5$  — поворотные зеркала в контуре обратной связи, ОУТ — оптически управляемый транспарант, ЖК — слой жидкого кристалла, ФП — фотопроводник,  $P_0$  — поляризатор (регулировка коэффициента усиления в контуре обратной связи)

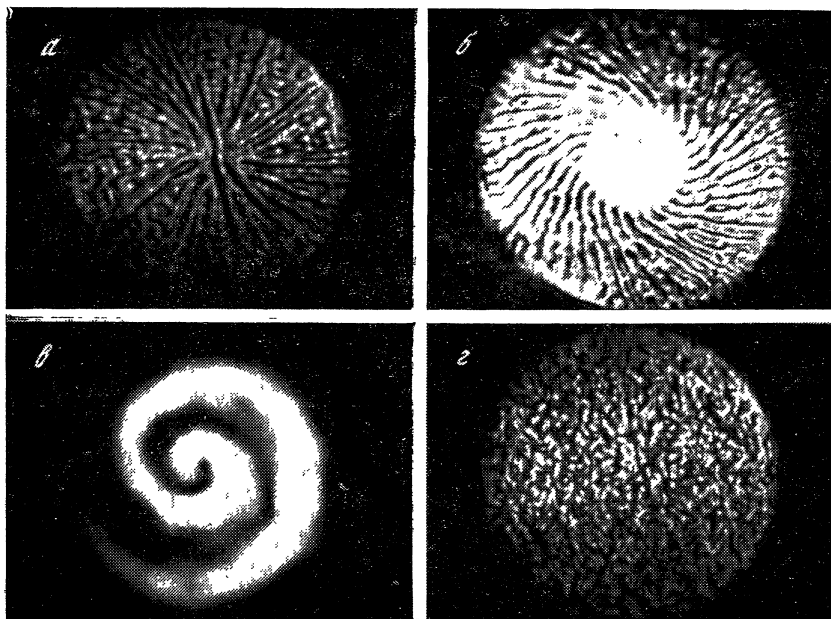


лированным в соответствии с распределением интенсивности в контуре обратной связи, что, в свою очередь, вызывало модуляцию показателя преломления жидкого кристалла.

Входное поле, отражаясь от диэлектрического зеркала, не достигало фотопроводника и практически не оказывало влияния на изменение показателя преломления.

Уравнения (3), (7), (9), записанные для системы с чисто оптической обратной связью, в равной степени применимы и для описания процессов в рассматриваемой гибридной системе. Диффузионный член в уравнениях в этом случае связан в основном с диффузией заряда в фотопроводнике.

Исследовались различные преобразования поля в контуре обратной связи: сдвиг, поворот, изменение масштаба. На рис. 9 показаны возникающие при этом структуры: все они, как правило, движущиеся. При большом коэффициенте усиления (интенсивно-



**Рис. 9.** Пространственные структуры светового пучка в контуре обратной связи

*a* — поворот изображения на угол  $\Delta = 180^\circ$ ; *б* — ротационная неустойчивость ( $\Delta = 30^\circ$ ); *в* — поворот изображения с изменением масштаба, *г* — стохастизация пространственной структуры поля

сти света  $I_2$  в контуре обратной связи) регулярные структуры разрушаются, происходит стохастизация поля.

В заключение подчеркнем, что сильные нелинейные самовоздействия в системах с двумерной обратной связью приводят к генерации световых полей, не имеющих даже отдаленных аналогов в линейной оптике. Особенно яркий пример в этом смысле — вращающиеся поля, ревербераторы. Вместе с тем ряд наблюдавшихся нами волновых структур может быть непосредственно сопоставлен с волновыми структурами, исследуемыми в гидродинамике, химии, биологии [10—13].

Приведенные данные показывают, что нелинейно-оптические системы с двумерной обратной связью открывают новые возможности в изучении динамики таких структур. Оптика представляет уникальные возможности управления параметрами нелинейной системы, изучения неустановившихся процессов, взаимодействия структур.

Описанную в настоящей статье гибридную систему с двумерной обратной связью можно считать эффективным аналоговым компьютером, способным решать широкий класс задач волновой динамики сильнонелинейных сред.

## ЛИТЕРАТУРА

1. *Ikeda K.* Optical turbulence: Chaos in optical bistability // *J. Phys. C.* 1983. Vol. 2, N 2. P. 183—192.
2. *Розанов Н. Н.* Гистерезисные явления в распределенных оптических системах // *ЖЭТФ.* 1981. Т. 20, № 1. С. 96.
3. *Gibbs H. M.* Optical bistability: Controlling light with light. N. Y. etc.: Acad. press, 1985.
4. *Ахманов С. А., Воронцов М. А., Пруидзе Д. В., Шмальгаузен В. И.* Пространственная неустойчивость светового поля в нелинейной системе с двумерной обратной связью: Препр. МГУ № 33. М., 1986. 4 с.
5. *Ferraro G., Hausler G.* TV optical feedback system // *Opt. Eng.* 1980. Vol. 19, N 4. P. 442—450.
6. *Kompanets I. N., Parfenov A. V., Popov Yu. M.* Bistable properties of spatial light modulator with internal feedback // *Opt. Commun.* 1981. Vol. 36, N 5. P. 415.
7. *Smith P. W., Tomlinson W. J.* Bistable optical devices promise subpicosecond switching // *IEEE Spectrum.* 1981. June. P. 24.
8. *Ikeda K., Daido H., Akimoto O.* Optical turbulence: Chaotic behaviour of transmitted light from a ring cavity // *Phys. Rev. Lett.* 1980. Vol. 45. P. 705.
9. *Васильев А. А., Компанец И. Н., Парфенов А. В.* Достижения в области разработки и применений оптически управляемых пространственных жидкокристаллических модуляторов света // *Квантовая электрон.* 1983. Т. 10, № 6. С. 1079.
10. *Ахманов С. А., Воронцов М. А., Иванов В. Ю.* Крупномасштабные поперечные нелинейные взаимодействия в лазерных пучках. Новые типы нелинейных волн, возникновение «оптической турбулентности» // *Письма в ЖЭТФ.* 1988. Т. 47, вып. 12. С. 611—614.
11. *Кринский В. И., Жаботинский А. М.* Автоволновые структуры и перспективы их исследования // *Автоволновые процессы в системах с диффузией.* Горький: ИПФ АН СССР, 1981. С. 6—32.
12. *Арансон И. С., Гапонов-Грехов А. В., Рабинович М. И.* Развитие хаоса в ансамблях динамических структур // *ЖЭТФ.* 1985. Т. 89, № 1(7). С. 92—105.
13. *Гапонов-Грехов А. В., Рабинович М. И.* Уравнение Гинзбурга—Ландау и нелинейная динамика неравновесных сред // *Изв. вузов. Радиофизика.* 1987. Т. 30, № 2. С. 131.

УДК 530.16+517.9

## РАЗМЕРНОСТЬ И ЭНТРОПИЯ В МНОГОМЕРНЫХ СИСТЕМАХ

В. С. АФРАЙМОВИЧ, А. М. РЕЙМАН

Странному аттрактору, являющемуся математическим образом автоколебательного движения реальной системы, присущи самые разнообразные признаки и свойства, каждому из которых можно сопоставить какую-либо характеристику реального процесса. Эти свойства могут быть количественно описаны соответствующей математической величиной — числом, функцией и т. д. Нам представляется, что наиболее важные (и употребимые сейчас) характеристики странного аттрактора, помимо уже привычного спектра Фурье, — это размерность аттрактора и энтропия динамической системы на нем. В настоящей статье речь пойдет именно о размерности и энтропии (а это с необходимостью заставит нас коснуться и других понятий и свойств, связанных со стохастичностью в динамических системах).

### §1. ОПРЕДЕЛЕНИЕ РАЗМЕРНОСТИ

Пусть  $A$  — аттрактор для системы  $n$  обыкновенных дифференциальных уравнений (в более общем случае  $A$  — подмножество метрического пространства  $E$  с заданной функцией расстояния  $\text{dist}$  между любыми двумя своими точками). Обозначим через  $N(\varepsilon)$  минимальное число шаров<sup>1</sup> диаметром  $\varepsilon$ , совокупность которых покрывает аттрактор  $A$ . Тогда число

$$c(A) = \overline{\lim}_{\varepsilon \rightarrow 0} \ln N(\varepsilon) / (-\ln \varepsilon)$$

называется (верхней) предельной емкостью или фрактальной размерностью множества  $A$ .

**Пример 1.** Пусть  $A$  — отрезок  $[0, 1]$  на прямой  $R^1 (= E)$ . Очевидно, чтобы покрыть его интервалами длины  $\varepsilon$ , достаточно  $\varepsilon^{-1} + O(1) = N(\varepsilon)$  множеств, поэтому  $c(A) = 1$ .

Нетрудно также установить, что емкость единичного куба в  $n$ -мерном евклидовом пространстве равна  $n$ . И вообще, любое

<sup>1</sup> Шар диаметром  $\varepsilon$  с центром в точке  $x$  — это множество  $B = \{y \in E \mid \text{dist}(x, y) < \varepsilon/2\}$ .

гладкое  $n$ -мерное многообразие (множество, определенным образом склеенное из  $n$ -мерных евклидовых кубов) имеет емкость, равную  $n$ . Таким образом, неравенство размерности множества целому числу говорит о его сложном геометрическом устройстве.

**Пример 2.** Пусть  $A$  — канторово множество на отрезке  $[0, 1] = E$ , т. е. множество точек  $\{x\}$ , записываемых в виде  $x = \sum_{i=1}^{\infty} (a_i/3^i)$ , где  $a_i$  принимает значение 0 либо 2. Напомним, что это множество может быть построено следующим образом. Из отрезка  $[0, 1]$  изымается средняя треть — открытый интервал, затем из оставшихся двух отрезков также убираются средние трети и т. д. То, что остается на единичном отрезке после бесконечной последовательности изъятий, и является канторовым множеством  $A$ .

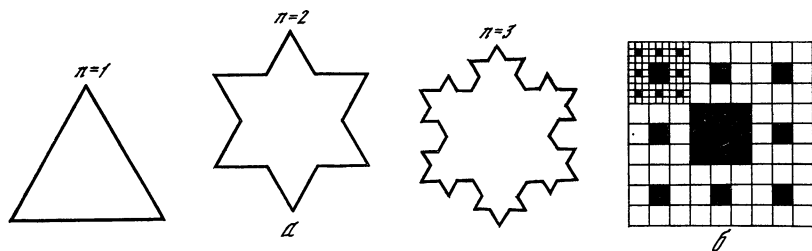
Пусть  $\epsilon = \epsilon_n = 3^{-n}$ . Покроем множество  $A$  интервалами длины  $\epsilon_n$ . Поскольку на  $n$ -м шаге построения канторова множества остается  $2^n$  замкнутых отрезков длиной  $3^{-n}$  каждый, то нетрудно видеть, что  $N(\epsilon_n) > 2^n$  и  $c(A) \geq \ln 2/\ln 3$ . С другой стороны, если для любого сколь угодно малого  $\delta > 0$  взять  $\tilde{\epsilon}_n = (3 - \delta)^{-n}$ , то  $N(\tilde{\epsilon}_n) \leq 2^n$ , поэтому  $c(A) \leq \ln 2/\ln(3 - \delta)$ . Из произвольности  $\delta$  следует, что  $c(A) = \ln 2/\ln 3 \approx 0,6309$ .

При вычислении размерности канторова множества мы использовали, во-первых, его самоподобие и, во-вторых, то обстоятельство, что и сам процесс его построения самоподобен: на последующем шаге делается то же, что и на предыдущем, но с элементами, уменьшенными в постоянное число раз. Для всех множеств, обладающих этим свойством, имеется эмпирическое правило вычисления размерности, в ряде случаев строго доказанное: если множество  $A$  строится с помощью самоподобного процесса, причем на любом шаге линейный элемент длины  $l$  заменяется  $p$  линейными элементами длиной  $q^{-1}l$  каждый ( $q > 1$ ), то предельная емкость  $c(A)$  равна  $\ln p/\ln q$  [1].

**Пример 3** (рис. 1).

**А.** Кривая фон Коха. Здесь  $p = 4$ ,  $q = 3$ , поэтому  $c(A) = \ln 4/\ln 3 \approx 1,2618$ .

**Б.** Ковер Серпинского. Здесь  $p = 8$ ,  $q = 3$ ,  $c(A) = \ln 8/\ln 3 \approx 1,8927$ .



**Рис. 1.** Процессы построения кривой фон Коха (а) и ковра Серпинского (б)



Но, конечно, чаще встречаются «неоднородные» множества, которые получаются либо вообще не по самоподобным правилам, либо в разных местах этих множеств самоподобие различно. И здесь важно знать не только предельное значение отношения  $\ln N(\varepsilon)/(-\ln \varepsilon)$ , но и «скорость» его стремления к пределу. Некоторую информацию о ней дает метаразмерность [2].

## 2. МЕТАРАЗМЕРНОСТЬ

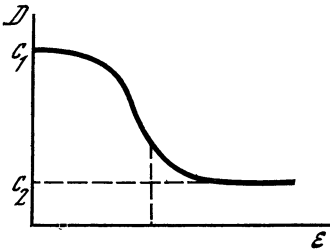
Приведем простейший пример неравномерного стремления к пределу отношения  $D(\varepsilon) = \ln N(\varepsilon)/(-\ln \varepsilon)$ .

**Пример 4.** Пусть множество  $A$  представимо в виде суммы двух непересекающихся множеств:  $A = A_1 \cup A_2$ , причем  $c(A_1) = c_1$ ,  $c(A_2) = c_2$  и  $c_1 > c_2$ . Очевидно начиная с некоторого  $\varepsilon$   $N(\varepsilon) = N_1(\varepsilon) + N_2(\varepsilon)$ , где  $N_i(\varepsilon)$  — число шаров диаметром  $\varepsilon$ , в совокупности покрывающих  $A_i$ ,  $i = 1, 2$ . Предположим, что  $N_1(\varepsilon) = \varepsilon^{-c_1}(1 - 2\varepsilon^\delta)$ ,  $N_2(\varepsilon) = \varepsilon^{-c_2}$  с точностью до членов более высокого порядка малости, причем  $0 < \delta < 1$ . Тогда

$$\frac{\ln N(\varepsilon)}{-\ln \varepsilon} = \frac{\ln N_1(\varepsilon) + \ln(1 + N_2(\varepsilon)/N_1(\varepsilon))}{-\ln \varepsilon} \rightarrow c_1 \text{ при } \varepsilon \rightarrow 0;$$

$$\frac{\ln N(\varepsilon)}{-\ln \varepsilon} = \frac{\ln(\varepsilon^{-c_2} + \varepsilon^{-c_1} - 2\varepsilon^{-c_1+\delta})}{-\ln \varepsilon} < c_2 + \frac{\ln 2}{-\ln \varepsilon} \text{ при } 1 - 2\varepsilon^\delta < \varepsilon^{c_1-c_2},$$

т. е. график  $D(\varepsilon)$  имеет вид, изображенный на рис. 2. Еще раз отметим, что если  $\delta \ll 1$ , то для очень малых значений  $\varepsilon$   $D(\varepsilon)$



будет порядка  $c_2$ , несмотря на то, что  $\lim_{\varepsilon \rightarrow 0} D(\varepsilon) = c_1$ . Число  $\delta$  и есть метаразмерность.

Дадим строгое определение. Пусть существует  $\lim_{\varepsilon \rightarrow 0} [\ln N(\varepsilon)/(-\ln \varepsilon)] = c$ . Предположим также, что существует  $\lim_{\varepsilon \rightarrow 0} N(\varepsilon)\varepsilon^c = \mu_0$ .

Положим

$$\delta = \lim_{\varepsilon \rightarrow 0} \frac{\ln |N(\varepsilon)\varepsilon^c - \mu_0|}{\ln \varepsilon}$$

Рис. 2. График зависимости отношения  $D$  от  $\varepsilon$

и назовем число  $\delta$  метаразмерностью (если предела нет, заменяем его верхним пределом). Из определения следует, что для некоторого положительного числа  $a$  с точностью до членов более высокого порядка малости  $N(\varepsilon) = \varepsilon^c (\mu_0 \pm a\varepsilon^\delta)$ . Отсюда видно, что если  $0 < \delta < 1$ , то допредельное значение  $D(\varepsilon)$  может существенно отличаться от  $c$  на большом интервале значений  $\varepsilon$ , тем более близких к нулю, чем меньше  $\delta$ .

Может оказаться, что емкость  $c(A)$  равна натуральному числу, и тогда фрактальность множества  $A$  определяется метаразмерностью.

Пример 5 [2]. Будем строить канторово подобное множество  $A$  так же, как в примере 2, но выбрасывая на  $n$ -м шаге не среднюю треть, а «дырку» размером  $\varepsilon_n$ , где  $\varepsilon_n$  стремится к нулю быстрее, чем экспоненциально; именно пусть  $\varepsilon_n = \frac{1}{2^n} \prod_{k=1}^n (1 - a^k)$ ,

$0 < a < 1$ . Тогда, очевидно, для любого  $\alpha > 0$  длина  $l_n$  любого остающегося отрезка на  $n$ -м шаге построения удовлетворяет неравенству  $l_n \geq (2 + \alpha)^{-n}$  для всех номеров  $n$ , больших некоторого. Из [1] следует, что в этом случае  $c(A) \geq \ln 2 / \ln(2 + \alpha)$ . Из произвольности  $\alpha$  вытекает, что  $c(A) = 1$ . С другой стороны, это множество «дырявое» — между любыми его точками на отрезке находится бесконечно много «дырок», выбрасываемых в процессе построения. Такие множества, емкость которых является натуральным числом и совпадает с обычной размерностью объемлющего «хорошего» пространства, а сами они сложно геометрически устроены (скажем, как в этом примере, обладают канторовской структурой), называются толстыми или жирными фракталами.

В [2] вычислена метаразмерность множества  $A$ :  $\delta = \ln(1/a) / \ln 2$ . Можно сказать, что достаточным признаком жирного фрактала наряду с целой емкостью является конечная метаразмерность<sup>2</sup>.

### 3. $(n, \varepsilon)$ -РАЗДЕЛЕННОСТЬ

В предыдущих пунктах были введены характеристики геометрической сложности объекта — размерность и метаразмерность — без учета того обстоятельства, что объект — это множество траекторий динамической системы. Если вспомнить, что рассматриваемое подмножество фазового пространства инвариантно, т. е. состоит из целых траекторий, то понятию размерности можно придать «динамическое» содержание.

Предположим (для простоты), что время в рассматриваемой динамической системе дискретно, и пусть  $A$  — ее инвариантное подмножество в фазовом пространстве  $E$  (например,  $A$  — аттрактор). Напомним, что здесь полутраектория — это множество точек  $\{f^k x\}_{k=0}^{\infty}$ , где  $f$  — отображение, порождающее динамическую систему, а отрезок траектории временной длины  $n$  — это множество точек  $\Gamma = \{f^k x\}_{k=0}^n$ .

Назовем  $(n, \varepsilon)$ -разделенными два отрезка,  $\Gamma_1$  и  $\Gamma_2$ , траекторий временной длины  $n$ , если  $\Gamma_1 = \{f^k x\}_{k=0}^n$ ,  $\Gamma_2 = \{f^k y\}_{k=0}^n$  и существует  $k_0$ ,  $0 \leq k_0 \leq n$ , такое, что  $\text{dist}(f^{k_0} x, f^{k_0} y) \geq \varepsilon$  (рис. 3) [4].

Обозначим через  $A_{n, \varepsilon}$  ансамбль отрезков траекторий временной длины  $n$  ( $A_{n, \varepsilon} = \{\Gamma(1 \leq k \leq n)\}$ ), удовлетворяющих следующим условиям:

- 1) любые два отрезка из этого ансамбля  $(n, \varepsilon)$ -разделены.
- 2) в множестве  $A$  не существует отрезка траектории временной

<sup>2</sup> О толстых фракталах см., например, [3].

длины  $n$ , который был бы  $(n, \varepsilon)$ -разделен с любым отрезком из  $A_{n, \varepsilon}$ .

Условие 2 означает, что, какой бы отрезок временной длины  $n$  в множестве  $A$  мы ни взяли, он целиком лежит в  $\varepsilon$ -окрестности некоторого отрезка из ансамбля  $A_{n, \varepsilon}$ . Это позволяет следующим образом интерпретировать приведенное определение. Пусть у нас есть ЭВМ, которая может различать отрезки траекторий временной длины  $n$  лишь с точностью  $\varepsilon$ . Тогда  $A_{n, \varepsilon}$  — это всевозможные отрезки длины  $n$ , хранящиеся в памяти нашей ЭВМ. Пусть  $C_{n, \varepsilon}$  —

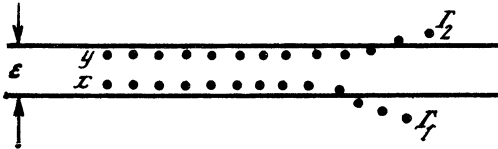


Рис. 3. К определению  $(n, \varepsilon)$ -разделенности ( $n = 13$ ,  $k_0 = 10$ )

количество элементов в множестве  $A_{n, \varepsilon}$ . Ясно, что и при уменьшении  $\varepsilon$ , и при увеличении  $n$  число  $C_{n, \varepsilon}$  должно возрасти. Уменьшение  $\varepsilon$  означает повышение чувствительности (точности) нашей ЭВМ и возможность разрешения все более тонкой геометрической структуры множества  $A$ , поэтому можно предположить, что величина  $C_{n, \varepsilon}$  как-то связана с предельной емкостью множества  $A$ . Это действительно так: справедлива формула [5]:

$$c(A) = \lim_{n \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \frac{\ln C_{n, \varepsilon}}{-\ln \varepsilon}.$$

С другой стороны, увеличение  $n$  при фиксированном значении  $\varepsilon$  может добавлять элементы в  $A_{n, \varepsilon}$  лишь в результате различного временного поведения траекторий в  $A$ , что связано, во-первых, с их неустойчивостью и, во-вторых, с неодинаковостью их поведения в  $A$ , т. е. со сложностью разбиения множества  $A$  на траектории. Эту неустойчивость и эту сложность характеризует специальная математическая величина, называемая топологической энтропией  $h_{\text{топ}}$  [6]. Как установлено в [4],

$$h_{\text{топ}} = \lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} \frac{\ln C_{n, \varepsilon}}{n}.$$

Топологическая энтропия и емкость характеризуют предельное поведение величины  $C_{n, \varepsilon}$ , рассматриваемой как функция двух переменных. Можно ввести величину

$$T_{n, \varepsilon} = \ln C_{n, \varepsilon} / (n - \ln \varepsilon),$$

которую мы назовем отношением Такенса [7]; тогда из приведенных выше формул вытекают равенства

$$c(A) = \lim_{n \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} T_{n, \varepsilon}; \quad h_{\text{топ}} = \lim_{\varepsilon \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} T_{n, \varepsilon}.$$

Пусть наша ЭВМ имеет две следующие ручки управления: вращая одну из них, мы увеличиваем точность ( $\varepsilon \rightarrow 0$ ), вращая

другую, увеличиваем временной интервал; если сначала до упора повернуть первую ручку, а затем вторую, то получим размерность, если же сначала вторую, затем первую — то энтропию. Можно одновременно вращать обе ручки — устремлять к бесконечности и  $n$ , и  $-\ln \varepsilon$  так, чтобы  $n/(-\ln \varepsilon) = \text{const}$ , тогда тоже будем получать некие числа, которые в [7] названы динамическими инвариантами.

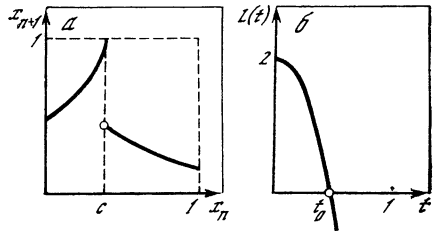
Можно теперь заключить: «динамический смысл» размерности в том, что она характеризует асимптотику числа  $\varepsilon$ -различных отрезков траекторий в аттракторе при  $\varepsilon \rightarrow 0$ .

#### 4. НИДИНГ-РЯДЫ.

##### ТОПОЛОГИЧЕСКАЯ ЭНТРОПИЯ ОДНОМЕРНЫХ ОТОБРАЖЕНИЙ

Исходя из определения, данного в предыдущем разделе (или в [6]), можно составить алгоритм счета топологической энтропии на ЭВМ и, казалось бы, ее сосчитать. Однако, как показал опыт, сделать это трудно даже для современных машин. Лишь для динамических систем, порождаемых одномерными отображениями,

Рис. 4. Разрывное отображение отрезка, моделирующее отображение Пуанкаре системы дифференциальных уравнений (а), и функция  $L(t)$ , корень которой определяет топологическую энтропию (б)



существует развитая математическая теория, позволяющая, в частности, определять и топологическую энтропию. Итак, пусть динамическая система задается отображением  $f$  типа изображенного на рис. 4, т. е.  $x_{n+1} = f(x_n)$ . Введем в рассмотрение функцию

$$\theta_i(x) = \begin{cases} +1, & \text{если функция } f^i \text{ возрастает в точке } x, \\ -1, & \text{если функция } f^i \text{ убывает в точке } x. \end{cases}$$

(Напомним, что  $f^i$  означает  $i$ -ю итерацию отображения  $f$ .) Тогда каждой точке  $x$  (которая при итерировании никогда не попадет в точку  $c$ ) мы можем сопоставить так называемый нидинг-ряд [8, 9]:

$$K(x) = \theta_0(x) + \theta_1(x)t + \dots + \theta_k(x)t^k + \dots$$

В множестве нидинг-рядов можно ввести лексикографический порядок и, например, установить, что если  $x < y$ , то и  $K(x) \leq K(y)$ .

Нидинг-ряды обладают многими замечательными свойствами, связанными с тем, что эти ряды фактически являются некоторыми универсальными координатами — они инвариантны относительно

любых непрерывных замен переменных. Но основное из этих свойств заключается в том, что динамика системы определяется лишь конечным числом нидинг-рядов (в ситуации, представленной на рис. 4, например, двумя:

$$K^+ = \lim_{x \rightarrow c+0} K(x); \quad K^- = \lim_{x \rightarrow c-0} K(x).$$

В частности, и для топологической энтропии существует простая формула. Пусть  $L(t) = K^+ - K^-$  (рис. 4, б) и пусть  $t_0$  — ближайший к началу координат положительный корень функции  $L(t)$ , тогда [8, 10]

$$h_{\text{топ}} = -\ln t_0.$$

Заметим, что отображение необязательно должно быть разрывным; нужно лишь, чтобы оно было кусочно-монотонным, тогда для энтропии существуют простые формулы, подобные приведенной.

А. Л. Железняк проверил на ЭВМ эффективность алгоритма нахождения топологической энтропии отображения «параболы»  $x_{n+1} = 4\lambda x_n(1-x_n)$  для нескольких значений параметра  $\lambda$  (в частности,  $\lambda = 0,75; 0,89; 0,95$ )<sup>3</sup>. Оказалось, что степень полинома, приближающего функцию  $L(t)$ , не превосходит 25 для нахождения энтропии с точностью  $10^{-2}$  и корень находится достаточно быстро.

В заключение раздела об энтропии отметим, что  $h_{\text{топ}}$  является оценкой сверху (точной верхней гранью) для метрических энтропий — энтропий Колмогорова—Синая — по всем инвариантным мерам, допускаемым динамической системой (о них речь пойдет ниже).

## 5. ДЕТЕРМИНИРОВАННО ПОРОЖДЕННЫЕ И «СЛУЧАЙНЫЕ» НАБЛЮДАЕМЫЕ

Наблюдаемая (сигнал, реализация и т. п.) — это функция от времени  $x = f(t)$  (если время дискретно, тогда наблюдаемая — это последовательность  $\{a_i\}_{i=0}^{\infty}$ ), по которой мы судим о процессе в исследуемой физической системе. Если рассматриваемая система является динамической (в строгом смысле), то наблюдаемая — функция от точки на ее траектории. Имея дело с природой (или лабораторным экспериментом), мы, конечно, заранее не знаем, можно ли адекватно описать исследуемый процесс динамической системой — у нас имеется лишь «голая» наблюдаемая.

Изложим некую теорию (следуя в основном работе [7]), оставаясь в рамках которой, можно в принципе отличить динамический процесс от «случайного», недетерминированного. Приведем определение из [7] (считая для простоты, что время дискретно).

Итак, наблюдаемая  $\underline{a} = \{a_i\}_{i=0}^{\infty}$  называется детерминированно порожденной, если:

- 1) существуют динамическая система (т. е. отображение  $f$ )

<sup>3</sup> За это авторы приносят ему искреннюю благодарность.

с конечномерным (в смысле емкости или в обычном, «топологическом» понимании) фазовым пространством  $M = \{x\}$ , начальная точка  $x_0$  и функция  $\varphi$ , такие, что  $\varphi(f^i(x_0)) = a_i$ ,  $i = 0, 1, \dots$ ;

2) траектории этой динамической системы не могут разбежаться быстрее, чем экспоненциально, т. е. при некоторых  $c, \lambda > 0$

$$\text{dist}(f^t x, f^t y) \leq \text{const} \cdot e^{\lambda t} \cdot \text{dist}(x, y);$$

3) функция  $\varphi$  липшиц-непрерывна, т. е.

$$|\varphi(x) - \varphi(y)| \leq \text{const} \cdot \text{dist}(x, y)$$

для любых точек  $x, y$  в фазовом пространстве.

Введем пространство наблюдаемых  $B$  как множество всевозможных последовательностей  $\underline{a} = (a_0, a_1, \dots)$ , каждая из кото-

рых удовлетворяет неравенству  $\sum_{i=0}^{\infty} |a_i|/2^i < \infty$ . Легко убедиться,

что: 1) если сложение наблюдаемых и умножение на скаляр определить покомпонентно, то  $B$  — линейное пространство; 2) число

$\sum_{i=0}^{\infty} |a_i|/2^i$ , ставящееся в соответствие каждой наблюдаемой  $\underline{a}$ ,

удовлетворяет всем свойствам нормы; 3) пространство  $B$ , снабженное так введенной нормой, является полным, т. е. любая сходящаяся по этой норме последовательность имеет своим пределом элемент из  $B$ . Следовательно,  $B$  — банахово пространство. Далее определим в  $B$  динамическую систему, задав отображение сдвига  $\underline{a} \rightarrow f(\underline{a})$ , где  $f(\underline{a}) = (a_1, a_2, \dots)$ , если  $\underline{a} = (a_0, a_1, a_2, \dots)$ . Это и есть универсальная динамическая система, порождающая любую ограниченную (или возрастающую медленнее экспоненты) наблюдаемую с помощью проекции на первую координату  $\pi_0$ :

$\underline{a} \rightarrow a_0$ . (Другими словами, любому элементу из  $B$  нужно поставить в соответствие значение  $a_0$ .) Возьмем теперь произвольную наблюдаемую и интерпретируем ее как начальное состояние нашей динамической системы в пространстве  $B$ . Выпустим из  $\underline{a}$

полутраекторию  $\{f^k \underline{a}\}_{k=0}^{\infty}$  и рассмотрим ее предельное множество  $\text{clos}(\{f^k \underline{a}\}_{k=0}^{\infty})$ , которое обозначим через  $A\underline{a}$  (можно показать, что  $A\underline{a}$  компактно). Предельную емкость  $c(A\underline{a})$  естественно назвать размерностью наблюдаемой.

Предположим, что  $c(A\underline{a}) < \infty$ ;  $n$  — натуральное число, такое, что  $n \geq 2c(A\underline{a}) + 1$  и является наименьшим из натуральных чисел, удовлетворяющих неравенству;  $M = \{(a_0, \dots, a_{n-1})\}$  —  $n$ -мерное подпространство пространства  $B$ ;  $\pi_n$  — естественная проекция из  $B$  на  $M$ , т. е. отображение, ставящее в соответствие любой наблюдаемой вектор из ее первых  $n$  координат. По теореме Манье [11] среди всех проекций  $B \rightarrow M$  типичными являются взаимно однозначные (и, очевидно, взаимно непрерывные) на множестве  $A\underline{a}$ . Предположим, что наша естественная проекция  $\pi_n$  типична. Тогда на подмножестве  $E = \pi_n(A\underline{a})$   $n$ -мерного пространства  $M$  определена динамическая система (порождаемая отображением  $\tilde{f} = \pi_n \circ f \circ \pi_n^{-1}$ ), а значение функции  $\varphi$ , определяющей

исходную наблюдаемую, — это просто значение первой координаты  $n$ -мерного вектора. Поскольку расстояние между двумя  $n$ -мерными векторами, определяемое по формуле  $\text{dist}(x, y) = \sum_{i=0}^{n-1} |x_i - y_i|/2^i$ , эквивалентно обычному евклидову расстоянию, то динамическую систему  $(f, M): (a_0 a_1 \dots a_{n-1}) \rightarrow (a_1 a_2 \dots a_n) \rightarrow (a_2 a_3 \dots a_{n+1}) \rightarrow \dots$  можно представлять себе в обычном  $n$ -мерном евклидовом пространстве, как и делается в численном эксперименте (см. ниже).

Таким образом, если размерность наблюдаемой конечна, то мы можем предъявить конечномерную динамическую систему (с точностью до предположения о типичности проекции  $\pi_n$ , которое с физической точки зрения можно считать всегда выполненным), продуцирующую данную наблюдаемую. Легко видеть, что и функция  $\varphi$  — проекция на первую координату — липшиц-непрерывна с константой Липшица, равной единице. Для детерминированной порожденности необходимо еще выполнение условия 2. Проверить это условие можно, используя понятие максимального ляпуновского показателя [12] и алгоритм его нахождения из наблюдаемой (см., например, [13]). Если максимальный ляпуновский показатель конечен, то условие 2 выполнено <sup>4</sup>. Следовательно, наблюдаемая детерминированно порождена и, как следует из [7], отношение Такенса ограничено:  $T_{n,\varepsilon} \leq \text{const} < \infty$ ; топологическая энтропия при этом конечна.

Нетрудно понять, что если значения  $a_i$  для каждого фиксированного номера  $i$  порождаются своим случайным процессом, никак не скоррелированным с процессами, порождающими значения  $a_j$ ,  $j \neq i$ , то наблюдаемая  $a$  не может быть детерминированно порожденной. В [7] установлено, что большинство наблюдаемых не являются детерминированно порожденными. Именно в [7] показано, что в пространстве наблюдаемых существует такое множество  $\bar{B}$  полной меры <sup>5</sup>, что для каждой наблюдаемой  $a \in \bar{B}$   $C_{n,\varepsilon}(a) \approx (1/2\varepsilon)^n$ . Отсюда сразу следует, что отношение Такенса не ограничено и размерность бесконечна.

Итак, основной вывод этого пункта: конечность размерности наблюдаемой и максимального ляпуновского показателя влечет детерминированную порожденность; бесконечность размерности или(и) энтропии говорит о «случайности» наблюдаемой (и процесса).

## 6. АЛГОРИТМ ТАКЕНСА

Остановимся более подробно на алгоритме обработки наблюдаемой с целью подчеркнуть то обстоятельство, что размерность

<sup>4</sup> При нахождении максимального ляпуновского показателя из наблюдаемой встречаются довольно существенные трудности вычислительного характера, о которых здесь говорить не будем, отсылая читателя к работе [13].

<sup>5</sup> Мера в пространстве наблюдаемых вводится как прямое произведение обычных мер Лебега на прямой.

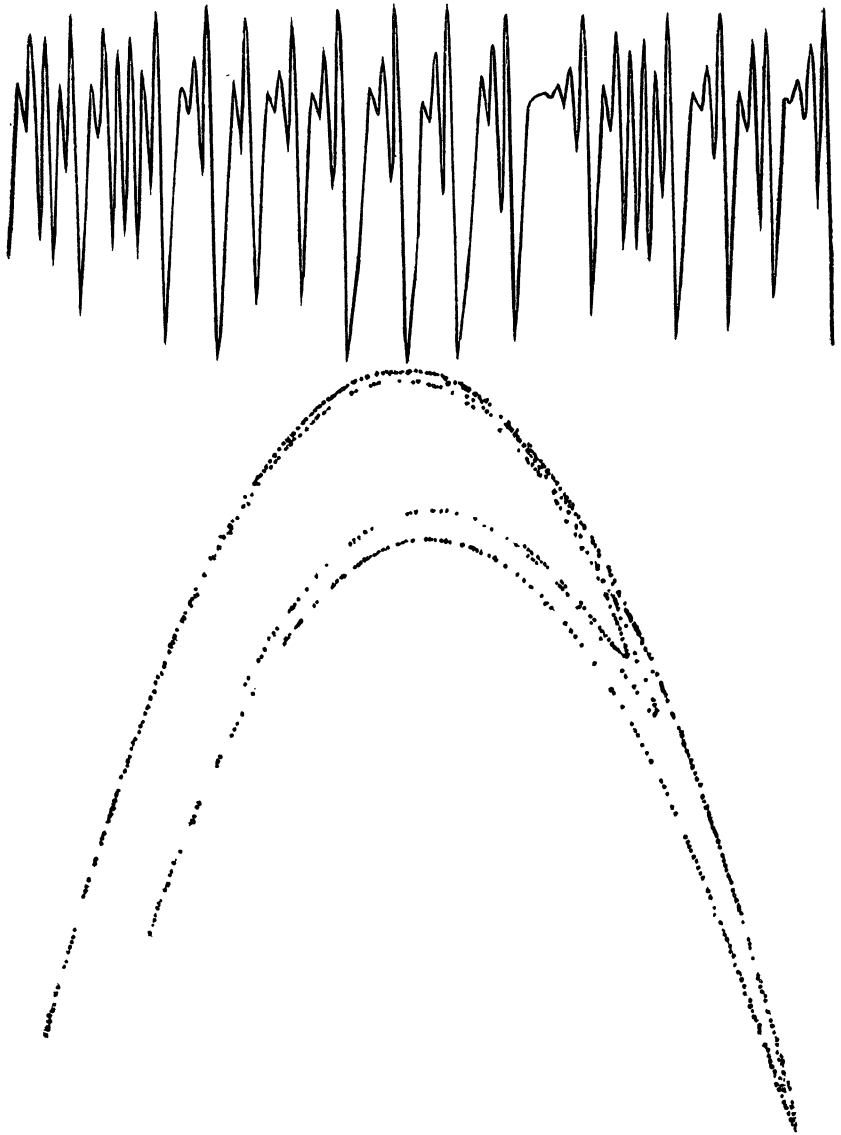


Рис. 5. Восстановление структуры фазового пространства по наблюдаемой (для отображения Эно)

(и энтропия) может служить эффективным инструментом получения информации о процессе.

Пусть  $\underline{a}$  — произвольная наблюдаемая из  $B$ ,  $\underline{a} = \{a_i\}_{i=0}^{\infty}$ .  
 $\sum_{i=0}^{\infty} |a_i|/2^i < \infty$ . Поочередно будем фиксировать  $n = 1, 2, 3, \dots$ ,  
и для каждого  $n$  осуществим следующую процедуру. Фиксиру-



ем  $\varepsilon$  и определяем (по индукции) множество индексов  $J(n, \varepsilon)$ : 1)  $0 \in J(n, \varepsilon)$ ; 2)  $i \in J(n, \varepsilon)$  тогда и только тогда, когда для всех  $j$ ,  $0 \leq j < i$ , таких, что  $j \in J(n, \varepsilon)$ , справедливо неравенство  $\max \{|a_i - a_j|, |a_{i+1} - a_{j+1}|, \dots, |a_{i+n} - a_{j+n}|\} \leq \varepsilon$ . Пусть  $\tilde{C}_{n, \varepsilon}$  — число элементов в множестве  $J(n, \varepsilon)$ . Положим  $c_n = \overline{\lim}_{\varepsilon \rightarrow 0} \ln \tilde{C}_{n, \varepsilon} / (-\ln \varepsilon)$ ,  $c = \lim_{n \rightarrow \infty} c_n$ .

Ясно, что  $c_n = c(\pi_n(A\underline{a}))$  (см. разд. 5), и если  $c(A\underline{a}) < \infty$ ,  $n > 2c(A\underline{a}) + 1$ ,  $\pi_n$  — типичная проекция [11], то  $c_n = c(A\underline{a})$ . В [5, 7] установлено, что если наблюдаемая  $\underline{a}$  порождается динамической системой (гладкой или липшиц-непрерывной), т. е. является гладкой или липшиц-непрерывной функцией  $\varphi$  от точки на траектории этой системы, принадлежащей аттрактору  $A$ , то (в типичной ситуации)

$$c(A) = \lim_{n \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \ln \tilde{C}_{n, \varepsilon} / (-\ln \varepsilon), \text{ т. е. } c(A) = c(A\underline{a}),$$

$$h_{\text{топ}} = \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \ln \tilde{C}_{n, \varepsilon} / n$$

(для липшиц-непрерывного случая эти равенства заменяются на неравенства « $\geq$ »; см. [7]).

Приведенный алгоритм носит название алгоритма Такенса, а значение  $n_0$ , начиная с которого  $c_n$  стабилизируется, называется размерностью вложения. Из теории следует, что в типичном случае  $n_0 \leq \tilde{n}_0$ , где  $\tilde{n}_0$  — ближайшее справа к  $2c(A\underline{a}) + 1$  целое число (конечно,  $n_0 \geq c(A\underline{a})$ ). Если  $c_n = c_{n+1}$  при  $n \geq n_0$ , то в силу изложенной теории есть основания считать, что искомый процесс (с точностью до указанных выше предположений о типичности) описывает система из  $n_0 + 1$  обыкновенных дифференциальных уравнений. Нужно лишь правильно подобрать динамические переменные. Несмотря на фактически полную пока неясность в решении последней проблемы, знание даже просто конечномерности задачи принципиально важно (см. об этом также разд. 11).

Определив  $n_0$ , можно в  $n_0$ -мерном евклидовом пространстве восстановить полутраекторию дискретной динамической системы, моделирующей процесс:  $(a_0 a_1 \dots a_{n_0-1}) \rightarrow (a_1 a_2 \dots a_{n_0}) \rightarrow \dots$ , и получить визуальное представление об искомом странном аттракторе (рис. 5).

## 7. ИНВАРИАНТНАЯ МЕРА

Важной характеристикой систем с аттракторами является инвариантная мера. Дело в том, что если для гамильтоновой системы фазовый объем сохраняется, то для систем с диссипацией и подкачкой энергии он уменьшается с течением времени. Поэтому начальное распределение со временем изрезается, эволюционируя к некоторому предельному (если, конечно, предел существует). Это предельное распределение, задающее инвариантную меру на аттракторе, безусловно содержит значительную

информацию о нем. Дадим строгое определение инвариантной меры.

Пусть  $\{f^t\}_{t=0}^{\infty}$  — отображение сдвига рассматриваемой динамической системы. Тогда мера  $\rho$  называется инвариантной, если для любого измеримого подмножества  $D$  фазового пространства  $\rho((f^t)^{-1}D) = \rho(D)$ , т. е. мера не меняется при сдвиге множества назад на время  $t$ . Например, обычная мера, порождаемая длиной, на отрезке  $[0, 1]$  инвариантна для отображения  $x \rightarrow 2x \bmod 1$ : при обратном отображении множество распадается на два, каждое половинной меры. Для отображения параболы (см. разд. 4) при  $x \in [0, 1]$ ,  $\lambda = 1$  инвариантна мера с плотностью  $d\rho = (\pi \sqrt{x(1-x)})^{-1} dx$ .

Инвариантная мера  $\rho$ : 1) сосредоточена на аттракторе  $A$ , если мера любого подмножества, не содержащего точек  $A$ , равна нулю; 2) вероятностная, если  $\rho(A) = 1$ ; 3) эргодическая, если в  $A$  не существует инвариантных подмножеств промежуточной (строго между нулем и единицей) меры.

Инвариантные меры могут быть, в частности, построены следующим способом. Именно, зафиксируем произвольную точку  $x$ , принадлежащую, скажем, странному аттрактору, и выпустим из нее полутраекторию  $\{f^t x\}_{t \geq 0}$ . Для любого измеримого подмножества  $D$  фазового пространства и числа  $T_k > 0$  определим  $\tau_k$  — время, проводимое траекторией  $\{f^t x\}$  на отрезке от нуля до  $T_k$  в множестве  $D$ . Тогда мерой  $\rho(D)$  множества  $D$  можно назвать  $\lim_{k \rightarrow \infty} \tau_k / T_k$ , если этот предел существует<sup>6</sup>, и показать, что

она удовлетворяет всем перечисленным выше свойствам. Такую меру можно строить по наблюдаемой  $a$ : для каждого  $n = 1, 2, \dots$  поочередно строить плотность распределения на аттракторе; скажем, для  $n = 1$  это обычная гистограмма.

Существуют и другие способы построения инвариантных мер; различных инвариантных мер может быть много, и каждая из них отражает те или иные свойства динамической системы на аттракторе. О том, какие из этих мер считать физическими, см., например, обзор [12].

## 8. ПОТОЧЕЧНАЯ РАЗМЕРНОСТЬ И КОРРЕЛЯЦИОННЫЙ ИНТЕГРАЛ

Как уже говорилось, определять размерность исходя из алгоритма Такенса вычислительно сложно. В [14, 15] был предложен другой алгоритм, позволяющий по реализации найти некоторое число, названное авторами этих работ корреляционной размерностью. В данном разделе рассказывается об этом алгоритме и об установленной сравнительно недавно связи корреляционной размерности с предельной емкостью.

<sup>6</sup> Известно, если для любого замкнутого множества  $F$   $\overline{\lim} \mu(F) \leq \rho(F)$ , то полученная мера  $\rho$  является мерой Боголюбова — Крылова (через  $\mu_k$  обозначается допредельная мера, определяемая  $\tau_k / T_k$ ).

Пусть  $x$  — произвольная точка в аттракторе  $A$ . Обозначим через  $B(x, \delta)$  шар радиуса  $\delta$  с центром в точке  $x$ . Пусть  $\rho$  — инвариантная, вероятностная эргодическая мера, сосредоточенная на  $A$ . Положим

$$\bar{d}_\rho(x) = \overline{\lim}_{\delta \rightarrow 0} \ln \rho(B(x, \delta)) / \ln \delta;$$

$$\underline{d}_\rho(x) = \underline{\lim}_{\delta \rightarrow 0} \ln \rho(B(x, \delta)) / \ln \delta$$

и назовем эти величины верхней и нижней поточечными размерностями соответственно.

В [16] доказано, что если для  $\rho$ -почти всех точек  $x$  (т. е. для множества полной  $\rho$ -меры)  $\underline{d}_\rho(x) = \bar{d}_\rho(x) = d$ , то и емкость множества полной меры в  $A$  равна  $d$ .

Пусть теперь нам дана наблюдаемая  $\underline{a} = \{a_i\}_{i=0}^\infty$ . Зафиксируем достаточно большое натуральное число  $N$  и натуральное  $m < N$ . Обозначим через  $x(i)$  вектор  $(a_i, a_{i+1}, \dots, a_{i+m})$ . Выражение  $C^m(\varepsilon) = \overline{\lim}_{N \rightarrow \infty} N^{-2} \cdot \{\text{число упорядоченных пар } [x(i), x(j)], \text{ таких, что } \text{dist}[x(i), x(j)] < \varepsilon\}$  называется корреляционным интегралом. Предел  $\lim_{\varepsilon \rightarrow 0} \ln C^m(\varepsilon) / \ln \varepsilon$  (если он существует) называется корреляционной размерностью.

Пусть  $C_i^m(\varepsilon) = \overline{\lim}_{N \rightarrow \infty} N^{-1} \cdot \{\text{число индексов } 0 \leq j \leq N, \text{ таких, что } \text{dist}[x(i), x(j)] < \varepsilon\}$ . Легко заметить, что  $C_i^m$  — это средняя частота попадания точки на траектории восстановленной по алгоритму Такенса динамической системы в  $m$ -мерный шар радиусом  $\varepsilon$  с центром в точке  $x(i)$ , а верхний (или нижний) предел выражения  $\ln C_i^m(\varepsilon) / \ln \varepsilon$  при  $\varepsilon \rightarrow 0$  — это верхняя (или нижняя) поточечная размерность относительно меры  $\rho$ , определяемой этой средней частотой. Если мера  $\rho$  эргодична, то для  $\rho$ -почти всех точек в  $\text{clos}\{x(i)\}_{i=0}^\infty$  верхняя (и нижняя) поточечная размерности будут одинаковы. Поэтому в данном случае, учитывая, что

$$C^n(\varepsilon) = \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^{N-1} C_i^n(\varepsilon),$$

получаем, что корреляционная размерность совпадает с верхней поточечной размерностью. А если верхняя и нижняя поточечные размерности совпадают, то корреляционная размерность совпадает с предельной емкостью множества полной меры в  $A_{\underline{a}}$ . (Конечно, здесь уже предполагается, что  $n \geq n_0$  ( $n_0$  — размерность вложения), в противном случае смысл получаемых величин неясен.)

Следует отметить, что корреляционный интеграл несет, вообще говоря, большую информацию, чем просто корреляционная размерность. «Изломы» в нем, подобные изгибам на рис. 2, свидетельствуют о структурированности сигнала (см. ниже п. 10).

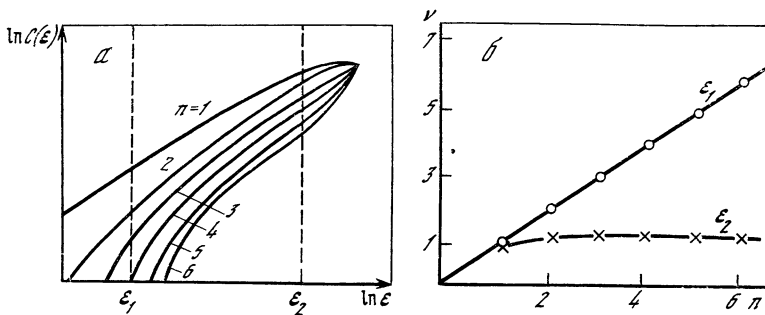


Рис. 6. Корреляционный интеграл (а) и локальные наклоны при различных  $\varepsilon$  (б) для отображения Эно с шумом

Например, есть возможность отделить динамический сигнал от белого шума малой интенсивности [17] (рис. 6).

При реальном вычислении размерности экспериментатор не имеет, конечно, возможности переходить к пределу как при  $N \rightarrow \infty$ , так и при  $\varepsilon \rightarrow 0$ . Поэтому, чтобы вычисленная им величина верно аппроксимировала размерность аттрактора в фазовом пространстве исследуемой (через наблюдаемую) системы, необходимо выполнение некоторых математических условий. Сформулируем их в виде следующего

К р е д о  и д е а л ь н о г о  ( с  т о ч к и  з р е н и я  м а т е м а т и к а )  
 э к с п е р и м е н т а т о р а :

1. Существует конечномерный аттрактор, являющийся математическим образом исследуемого процесса.

2. Существует единственная мера  $\rho$ , определяемая средней частотой попаданий итерированной точки.

3. Мера  $\rho$  эргодична.

4.  $d_\rho(x) = \bar{d}_\rho(x)$ .

5.  $x(0) = (a_0, \dots, a_{n-1})$  — типичная точка для меры  $\rho$ .

Тогда  $\lim_{\varepsilon \rightarrow 0} \ln C^n(\varepsilon) / (-\ln \varepsilon) = d_\rho$  и совпадает, как уже гово-

рилось, с емкостью множества полной меры в аттракторе. Следовательно, если  $N$  достаточно велико, а  $\varepsilon$  достаточно мало, найденная экспериментатором величина хорошо приближает предельную емкость истинного аттрактора <sup>7</sup>.

Несмотря на внешнюю ограничительность условий 2—5, экспериментальный опыт и тестовые исследования счета размерности «впрямую» и с помощью наблюдаемой показывают, что на практике об этих требованиях можно не вспоминать — главное, чтобы процесс описывался конечномерной системой.

<sup>7</sup> Предполагается, что емкость аттрактора совпадает с емкостью множества полной меры в нем. Это выполнено для «хорошей», физической, меры, какой мы считаем меру  $\rho$ .

## 9. РАЗМЕРНОСТЬ В ЭКСПЕРИМЕНТАХ. ПРЕДВАРИТЕЛЬНАЯ РАБОТА С НАБЛЮДАЕМЫМИ

В этом разделе мы приведем список (разумеется, далеко не полный) экспериментов — вычислительных и лабораторных, в которых измерялась размерность с помощью изложенных (или близких к ним) алгоритмов. Некоторые особенности исследования конкретных задач будут обсуждены ниже<sup>8</sup>; здесь просто перечислим накопленные результаты, что, как нам кажется, необходимо сделать, поскольку они пока еще не очень хорошо известны. Кроме того, обсудим независимые от конкретных алгоритмов вопросы предварительной обработки наблюдаемых.

Перечислим вначале тестовые, математические системы, исследование которых позволяет очертить рамки возможностей измерения размерности: система Лоренца [19], отображение Эно [14], отображение Заславского [15], система Рабиновича—Фабриканта [15, 20], система уравнений численного моделирования течения в гидродинамическом канале [21].

Модельные системы могут служить для оптимизации алгоритмов, отладки программ и т. д. Дело в том, что при конкретной реализации алгоритма Такенса (или алгоритмов Грассбергера—Прокачиа и т. д.) сразу возникают вопросы, не существующие с математической точки зрения. Первый из них следующий. Пусть наблюдаемая  $a(t)$  порождается системой с непрерывным временем. Чтобы применить описанные выше алгоритмы, нужно перейти к дискретному времени — рассмотреть значения  $a$  в моменты времени  $i\tau$ ,  $i=0, 1, \dots$ . Как выбрать шаг дискретизации  $\tau$ ? Математика здесь ответа не дает. Известно лишь [5], что для любого (типичного) значения  $\tau$  размерность, восстанавливаемая по наблюдаемой  $a(i\tau)$ , одна и та же.

С другой стороны, нетрудно понять, что для любого фиксированного  $n$  существуют настолько малые значения  $\tau$ , что близким начальным точкам в последовательности  $x(i\tau)$ , где  $x(i\tau) = (a(i\tau), a[(i+1)\tau], \dots, a[(i+n)\tau])$ , соответствуют близкие куски траекторий большой длины. А это означает, что восстанавливаемый аттрактор вытягивается вдоль главной диагонали  $n$ -мерного пространства  $a(i\tau) = a[(i+1)\tau] = \dots = a[(i+n)\tau]$  и расположен в ее малой окрестности, причем диаметр пересечения этой окрестности с нормальной к диагонали плоскостью тем меньше, чем меньше  $\tau$ . Следовательно, при малых  $\tau$  на большом интервале масштабов  $\varepsilon$  мы будем наблюдать в корреляционном интеграле размерность (наклон кривой), равную 1.

В [19], где численно исследуется вопрос о влиянии времени дискретизации на получаемое значение размерности, предлагается отталкиваться от  $T_R$  — среднего времени возврата траектории. В [19] на примере системы Лоренца показано, что при  $\tau \ll T_R$  размерность равна 1. При увеличении  $\tau$  значение размер-

<sup>8</sup> Конкретный алгоритм исследования размерности с указанием некоторых особенностей «технологии» вычислительного эксперимента см. в [18].

ности монотонно увеличивается; в области  $\tau \in [0, 2T_R; 0,5 T_R]$  это значение стабилизируется и становится равным истинной размерности аттрактора Лоренца ( $\approx 2,06$ ). То значение  $\tau$ , при котором значение размерности стабилизируется, по-видимому, и можно считать оптимальным.

Остановимся теперь на приведении наблюдаемой к удобной для дальнейшей обработки форме. Но сначала приведем пример «неудобной» наблюдаемой.

**Пример 6<sup>9</sup>.** Пусть  $a(t)$  — наблюдаемая, имеющая вид последовательности импульсов, такой, что время, занимаемое их фронтами, мало по сравнению с длительностями вершин, расположенных в малых окрестностях значений  $a = \alpha$  и  $a = \beta$  (рис. 7). Здесь, какое бы  $\tau$  мы ни взяли, для любого  $n$  подавляющая часть точек восстанавливаемой траектории будет расположена вблизи вершин  $n$ -мерного куба со стороной  $|\beta - \alpha|$ , и для огромного интервала масштабов  $\varepsilon$  будет выполняться неверная (соответствующая нулевой размерности) асимптотика.

Для того чтобы верно (и быстро) получить значение размерности, необходимо от наблюдаемой  $a(t)$  перейти к другой наблюдаемой  $b(t)$ , которая обеспечивает достаточное число градаций амплитуды сигнала; здесь, например, можно за  $b(t)$  взять текущую длину импульсов, скажем, в виде  $b(t) = \int_0^t a(t) dt$ .

Дефектность наблюдаемой может быть связана и со следующим обстоятельством. Фактически наблюдаемая — это проекция радиус-вектора изображающей точки в (конечно- или бесконечномерном) фазовом пространстве на некоторую прямую. Как мы знаем, из теорем Манье и Такенса следует, что в типичной ситуации информация о размерности при проектировании не теряется — размерность может быть восстановлена по наблюдаемой. Но при этом аттрактор может так сплюснуться по определенным направлениям, что восстановить истинное значение практически очень трудно или даже невозможно.

**Пример 7 [22].** Для двумерного отображения вида

$$x_{n+1} = \alpha x_n + b [2y_n], \quad y_{n+1} = (2y_n) \bmod 1$$

в качестве наблюдаемой выбрана величина  $z_n = \alpha x_n + (1 - \alpha) y_n$ . Если  $\alpha \ll 1$ , то размерность, вычисляемая с помощью  $z_n$ , будет совпадать с размерностью аттрактора для отображения  $y_{n+1} = (2y_n) \bmod 1$ , т. е. с единицей.

<sup>9</sup> Ситуация, описания в этом примере, встречается довольно часто при обработке импульсных сигналов с хаотической скважностью (биоэлектрические сигналы, в том числе электрокардиограммы; сигнал, полученный в эксперименте по рассеянию на поверхностных волнах в жидкости (рис. 7, а) — здесь узкий луч света нормально падает на взволнованную поверхность жидкости, а рассеянный свет принимается направленным приемником и т. д.).

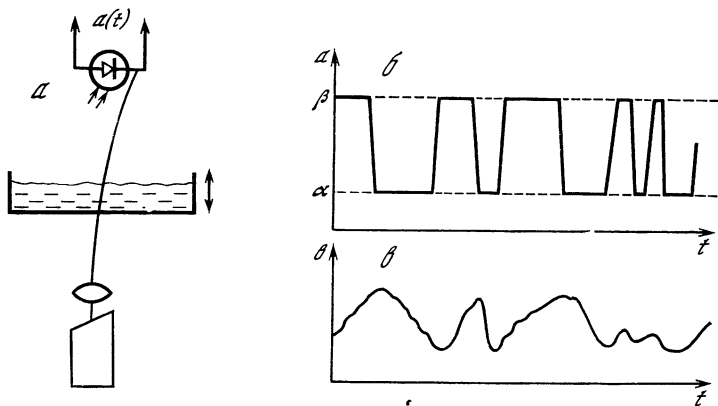


Рис. 7. «Неудобная» наблюдаемая в эксперименте

$a$  — схема эксперимента;  $b$  — исходный сигнал;  $c$  — профильтрованный сигнал

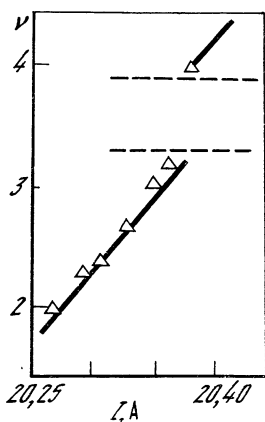


Рис. 8. Изменение размерности  $\nu$  с ростом управляющего параметра (силы тока) в эксперименте с течением проводящей жидкости между цилиндрами под действием электрического тока в магнитном поле [30]

Таким образом, из приведенных примеров вытекает, что: 1) перед «запуском» наблюдаемой в обработку желательно привести ее к удобной для соответствующего алгоритма форме; 2) для уверенности в результате необходимо менять саму наблюдаемую (конечно, если есть такая возможность).

В натурном или лабораторном эксперименте, где наблюдаемую выбирать не приходится, следует помнить, что она может описывать свойство не всей системы, а лишь некоторой подсистемы (определяемой проектированием). Однако в тех случаях, когда размерность измеряется не для фиксированных значений управляющих параметров, а для целой их области, можно с достаточной долей уверенности предполагать, что там, где она плавно меняется, она измерена верно. Перечень (конечно, тоже далеко не полный) физических систем, в которых размерность измерялась при различных значениях управляющих параметров, дан в таблице. Заметим, что во всех этих экспериментах размерность сравнительно невелика (не превышает 8).

Объект исследования	Параметр	Литература
Течение Куэтта	Число Рэлея	[23]
Пограничный слой на пластине	Координата	[24]
Сдвиговое течение с обратной связью	Число Рейнольдса	[25]
Электроно-дырочная плазма в полупроводнике	Напряженность поля	[26]
Ферроэлектрик (ниобат бария—натрия)	Плотность тока	[27]
Электроэнцефалограмма	Фаза сна	[28]

Изолированные скачки размерности могут служить указанием на кризисы [29] соответствующих странных аттракторов (см. рис. 8 и п. 11).

#### 10. АЛГОРИТМ ВЫЧИСЛЕНИЯ РАЗМЕРНОСТИ И СТРУКТУРА КОРРЕЛЯЦИОННОГО ИНТЕГРАЛА

В этом пункте мы обсудим некоторые вопросы, возникающие при конкретном вычислении корреляционной размерности. Сначала выпишем приближенное выражение для корреляционного интеграла:

$$C^n(\varepsilon) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \Theta(\|r_{ij}\| - \varepsilon) \equiv \frac{1}{N} \sum_{i=1}^N \frac{N_i(\varepsilon)}{N},$$

где  $\Theta$  — функция Хевисайда ( $\Theta(x) = 1$  при  $x > 0$ ,  $\Theta(x) = 0$  при  $x \leq 0$ );  $N_i(\varepsilon)$  — число точек из отрезка траектории временной длины  $N$ , попадающих в  $\varepsilon$ -окрестность  $i$ -й точки;  $N$  — достаточно большое натуральное число. (Напомним, что  $N_i(\varepsilon)/N$  — приближенная мера, определяемая средней частотой, для шара радиусом  $\varepsilon$  с центром в  $i$ -й точке.) Чтобы вычислить корреляционную размерность, мы определим наклон кривой — графика  $\ln C^n$  ( $\ln \varepsilon$ ). Но при попытке это сделать для любого конкретного случая мы сразу же столкнемся с тем, что график имеет разные производные для разных интервалов  $\varepsilon$ . Какую из них считать «истинной» корреляционной размерностью? Чтобы ответить на этот вопрос, необходимо знать причины появления изломов на графике. Выяснением этих причин мы и займемся. Предварительно отметим, что все предположения, изложенные в «Кредо» (см. разд. 8), мы считаем выполненными, и всякий раз оговаривать этот вопрос не будем. (Основные — единственность меры, определяемой средней частотой, равенство верхней и нижней поточечных размерностей.)

Рассмотрим причины изломов в корреляционном интеграле.

**Неоднородность аттрактора.** Мы уже видели в разд. 2, что изломы в соответствующем графике могут появляться из-за неоднородности исследуемого множества. Для корреляционной



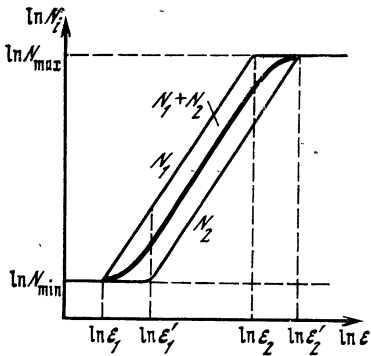


Рис. 9. Сокращение интервала степенной асимптотики для неоднородного аттрактора [31]

$$N_1(\varepsilon) = \begin{cases} N_{\min}, & \varepsilon < \varepsilon_1, \\ a_1 \varepsilon^\nu, & \varepsilon_1 < \varepsilon < \varepsilon_2, \\ N_{\max}, & \varepsilon > \varepsilon_2, \end{cases} \quad N_2(\varepsilon) = \begin{cases} N_{\min}, & \varepsilon < \varepsilon'_1, \\ a_2 \varepsilon^\nu, & \varepsilon'_1 < \varepsilon < \varepsilon'_2, \\ N_{\max}, & \varepsilon > \varepsilon'_2, \end{cases}$$

где  $N_{\min}$  и  $N_{\max}$  — максимальное и минимальное число точек, которые могут попасть в  $\varepsilon$ -окрестность данной точки. (Обычно  $N_{\min} = 1$ , т. е. в  $\varepsilon$ -окрестности точки содержится только она сама;  $N_{\max} = N$  — длина реализации.) Видим, что если интервалы  $(\varepsilon_1, \varepsilon_2)$  и  $(\varepsilon'_1, \varepsilon'_2)$  пересекаются, то скейлинг  $\sim \varepsilon^\nu$  для суммы  $N_1 + N_2$  наблюдается лишь на этом пересечении; при других значениях  $\varepsilon$  форма интеграла другая (рис. 9).

**Недостаточная длина реализации.** Ясно, что если число  $N$  невелико, то частные суммы  $N_i(\varepsilon)$  могут вести себя совершенно по-разному для разных  $i$ , а это, как только что было показано, приводит к появлению дополнительных изломов.

Изломы, обусловленные неоднородностью аттрактора и недостаточной длиной реализации, исчезают при увеличении  $N$ ; точнее, область истинного скейлинга ( $\sim \varepsilon^\nu$ ) при увеличении длины реализации  $N$  расширяется, «съедая» эти несущественные изломы. Это обстоятельство может служить для их диагностирования и борьбы с ними.

Следует отметить, что длина реализации связана с измеряемой размерностью — очевидно, чем больше размерность, тем большую длину реализации нужно иметь, чтобы обеспечить достаточное заполнение аттрактора. В [25] для алгоритма, аналогичного здесь приведенному, дана следующая эмпирическая оценка на  $N$ :  $\ln N \sim \nu \ln(\varepsilon_{\max}/\varepsilon_{\min})$ , где  $(\varepsilon_{\min}, \varepsilon_{\max})$  — интервал  $\varepsilon$ , на котором  $C^n(\varepsilon) \sim \varepsilon^\nu$ ,  $\nu$  — корреляционная размерность.

**Структурированность наблюдаемой.** Изломы в корреляционном интеграле могут быть обусловлены тем, что наблюдаемая является смесью сигналов, порожденных динамическими системами с разными размерностями соответствующих аттракторов.

размерности так тоже может быть. В корреляционном интеграле частные суммы  $N_i(\varepsilon)$  имеют тот же скейлинг (то же асимптотическое представление), что и  $C^n(\varepsilon)$ :  $N_i(\varepsilon) \sim a_i \varepsilon^\nu$ , но числа  $a_i$  — плотности точек — могут быть различными. Из-за этого искомый скейлинг для  $C^n(\varepsilon)$  может осуществиться в более узком диапазоне  $\varepsilon$ .

**Пример 8 [31].** Рассмотрим две частные суммы в общей сумме  $C^n(\varepsilon)$ :  $N_1(\varepsilon)$  и  $N_2(\varepsilon)$ , причем пусть

В простейшем таком случае  $a(t) = a_0(t) + \delta_1 a_1(t) + \dots + \delta_k a_k(t)$ , где  $a_i(t)$  порождается динамической системой, размерность аттрактора у которой равна  $\nu_i$ , и  $\delta_1 > \delta_2 > \dots > \delta_k$ ,  $\nu_0 < \nu_1 < \dots < \nu_k$ . Тогда при последовательном уменьшении  $\varepsilon$  в корреляционном интеграле будут проявляться составляющие все с меньшими амплитудами  $\delta_i$ , т. е. мы увидим интервалы  $(\varepsilon'_i, \varepsilon''_i)$ , на которых тангенс угла наклона в графике корреляционного интеграла будет равен приблизительно  $\nu_i$ . Число  $(k-1)$  в этом случае можно назвать глубиной структурированности. Здесь наличие излома несет полезную информацию об исследуемом процессе. Отметим, что увеличение длины реализации не уничтожит эти изломы, а сделает их лишь более отчетливыми.

До сих пор в этом разделе мы рассуждали так, как будто бы имели дело с сигналом чисто динамического происхождения. На самом деле любая наблюдаемая, порожденная реальным процессом, содержит в себе шум; продуцирующая в точности эту наблюдаемую динамическая система бесконечномерна, и поэтому корреляционный интеграл обязан иметь излом, отделяющий масштаб динамической составляющей от малых масштабов  $\varepsilon$ , где существует шум. Мы видели (см. рис. 6), что исследование локальных наклонов позволило отделить хаотический сигнал динамического происхождения от аддитивного белого шума. Итак, четвертой причиной, порождающей «изломы», является наличие шумов в наблюдаемой. Эти изломы также не устраняются увеличением длины реализации и проявляются при малых  $\varepsilon$ .

## 11. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

1. Размерность несет важную информацию при построении феноменологической модели исследуемого процесса: она позволяет определить число ее степеней свободы. Из теории следует, что если  $\nu^0$  — размерность, то  $n (\geq 2\nu^0 + 1)$  функций достаточно для описания системы<sup>10</sup>. Если сигнал структурирован с глубиной структурированности, скажем,  $k$ , то можно представить себе иерархическую систему из  $k$  динамических систем с последовательно возрастающими размерностями фазового пространства, каждая из которых все менее огрубленно — все более тонко — описывает исследуемый процесс.

Приведем следующий пример. При исследовании перехода к турбулентности в пограничном слое на плоской пластине в аэродинамической трубе [24] применялись два уровня описания динамики процесса. В грубом приближении на начальном участке

<sup>10</sup> На самом деле число степеней свободы может быть и меньше. Напомним, что алгоритм нахождения корреляционной размерности устроен так же, как и алгоритм Такенса нахождения предельной емкости (см. разд. 6), т. е. строится зависящее от  $n$  семейство корреляционных интегралов и определяется последовательность  $\nu(n)$  корреляционных размерностей. Значение  $\nu^0$ , после которого  $\nu(n) = \nu(n+1)$ , будет так называемой размерностью вложения. В принципе  $\nu^0$  может оказаться меньше, чем  $2\nu^0 + 1$  (но, конечно, всегда больше  $\nu^0$ ).

пластины учитывалась лишь волна Толлмина—Шлихтинга, вводящая возмущения в поток; ей соответствует предельный цикл в некотором фазовом пространстве; размерность  $\nu_1 = 1$ . Более точное (с меньшими значениями  $\epsilon$ ) измерение показывает  $\nu_2 = 4$ ; это соответствует наличию еще трех независимых движений — собственных колебаний механических частей установки и электрической наводки на управляющий элемент. Можно следить за зависимостью от параметра отдельно размерностей  $\nu_1$  и  $\nu_2$ , т. е. размерностей, определяемых по разным интервалам  $\epsilon$ . В рассматриваемом примере параметром служит значение координаты «вниз по потоку». При его изменении на большом интервале значение  $\nu_2$  сохраняется, а размерность грубого описания  $\nu_1$  растет до 3,5. На этапе трехмерных возмущений  $\nu_2$  начинает расти; растет и  $\nu_1$ , не превышая, однако,  $\nu_2$ .

Если известна дополнительная (кроме размерности) информация об исследуемой системе (например, информация об особенностях спектра мощности сигнала, кратностях пиков и т. д.), то в некоторых случаях этого уже достаточно для построения феноменологической модели.

**П р и м е р 9** [25]. В эксперименте на аэродинамической трубе исследовался переход к турбулентности в сдвиговом течении с акустической обратной связью: регистрировались пульсации скорости, измерялись спектр мощности и корреляционная размерность. В частности, было установлено, что при скорости потока 10,2 м/с в спектре мощности имеются два четко выраженных пика на частотах  $\omega_1$  и  $2\omega_1 + \Delta$  (их уровень превышал уровень сплошного спектра на 20 дБ), а корреляционная размерность  $\nu \approx 2,6$ . Естественно было предположить, что стохастические пульсации в этом режиме могут быть приближенно описаны системой с двумя степенями свободы, распадающейся (при отсурствии взаимодействия) на две системы с одной степенью свободы каждая, одна из которых отвечает автоколебаниям на частоте  $\omega_1$ , а другая — на  $\omega_2 \approx 2\omega_1$ . Из физических соображений естественно выбрать одну из наиболее простых таких систем:

$$\begin{aligned} u_{tt} - \alpha_1 (1 - \gamma_1 u^2) u_t + \omega_1^2 u + \beta_1 u^3 &= \beta_1 u w, \\ w_{tt} - \alpha_2 (1 - \gamma_2 w^2) w_t + \omega_2^2 w + \beta_2 w^3 &= \beta_2 u^2. \end{aligned} \quad (1)$$

Как показывает численный счет [32], в широкой области параметров в фазовом пространстве системы (1) существует странный аттрактор, размерность которого в зависимости от коэффициентов диссипации  $\alpha_{1,2}$  непрерывно меняется от 2 до 4. Спектр мощности реализации, полученной из системы (1), очень похож на спектр исходного сигнала (в частности, и при  $\nu = 2,6$ ). Поэтому систему (1) можно считать хорошим кандидатом на модель, адекватно описывающую исследуемое течение.

Более простая ситуация, когда формулы, определяющие динамическую систему (модель рассматриваемого процесса), уже известны из каких-либо априорных соображений, неизвестны

лишь числовые коэффициенты, вводящие в них. Здесь, меняя коэффициенты до тех пор, пока измеряемые размерность и энтропия не совпадут с заранее заданными, получаем их правильные значения.

В этом духе проводились исследования и в работе [33], в которой измерения энтропии позволили очистить динамически порожденный сигнал от внешнего шума, а затем, используя линейные авторегрессионные модели, вычислить оптимальный набор коэффициентов и функций, входящих в предполагаемое отображение  $x_{n+1} = F(x_n)$ .

Эта же методика была применена для исследования динамики реальных процессов в экспериментальной установке, реализующей генератор Ван-дер-Поля с аддитивным шумом [33]. При этом длина реализации составила всего несколько тысяч точек.

2. Размерность и энтропия — числовые характеристики, усредненно описывающие отдельные стороны стохастичности. А это значит, что по ним можно сравнивать процессы, говорить о более сильной и более слабой стохастичности, находить значения параметров, при которых они претерпевают резкие, скачкообразные изменения, и определять бифуркации, порождающие эти скачки. Так был обнаружен в [34] кризис аттрактора для параметрически возбуждаемого осциллятора, описываемого уравнением  $\ddot{x} + k\dot{x} + (1 - q \cos \omega t)x + x^3 = 0$ . Именно в [34] было численно установлено, что при  $q = 50$ ,  $\omega = 2,04$ ,  $k = 0,4576$  размерность аттрактора резко увеличивается, а затем была найдена бифуркация, объясняющая скачок.

Здесь возникает новая, чисто математическая задача об асимптотике зависимости размерности (или энтропии) от параметра для различных бифуркаций, приводящих к возникновению странных аттракторов. Один из результатов на эту тему, полученный М. А. Шерешевским и В. С. Афраймовичем, описывает зависимость дробной части размерности от надкритичности при рождении странного аттрактора в результате касательной бифуркации (см., например, [29]), другими словами, при возникновении перемежаемости. Пусть  $d(\varepsilon)$  — дробная часть размерности рождающегося при  $\varepsilon = 0$  в результате касательной бифуркации странного аттрактора, тогда имеет место одна из следующих асимптотических формул [35]:

$$d(\varepsilon) = a\sqrt{\varepsilon} + o(\sqrt{\varepsilon}); \quad d(\varepsilon) = a\sqrt{\varepsilon} \ln 1/\sqrt{\varepsilon} + b\sqrt{\varepsilon} \ln \ln 1/\sqrt{\varepsilon} + c\sqrt{\varepsilon} + o(\sqrt{\varepsilon}),$$

где  $a$ ,  $b$ ,  $c$  — константы,  $\varepsilon$  — надкритичность. Применяется первая или вторая формула в зависимости от того, содержит или не содержит в критический момент неустойчивая сепаратриса седло-узловой предельного цикла гетероклинические траектории другого седлового цикла <sup>11</sup>.

<sup>11</sup> Напомним, что при касательной бифуркации устойчивый и седловой предельные циклы сливаются, образуя седло-узловой цикл, и затем исчезают.

3. Как уже обсуждалось в разд. 5, любая ограниченная наблюдаемая порождена некоторой универсальной динамической системой. В случае непрерывного времени это динамическая система сдвигов, определяемая следующим образом: каждой функции  $f(x)$ , определенной на бесконечном интервале значений  $x$ , и значению времени  $t$  ставится в соответствие функция  $f(x+t)$ ; другими словами, на множестве, скажем, непрерывных функций определен оператор сдвига  $T^t$ . Расстояние в пространстве функций можно ввести, например, по формуле

$$\text{dist}(f, g) = \int_{-\infty}^{\infty} |f - g| e^{-|x|} dx$$

(если функции определены лишь при  $x \geq 0$ , то нижний предел в интеграле следует заменить нулем).

Итак, любую функцию из рассматриваемого класса можно считать наблюдаемой этой динамической системы, в том числе и мгновенный снимок (например, значение при  $t = t^*$  решения  $A(x, t)$  уравнения  $A'_t = \mu A(1 - 2 \cos qx) - A^3 + A''_{xx}$ ). Следовательно, каждому мгновенному снимку мы можем поставить в соответствие размерность аттрактора универсальной динамической системы, определяемой этим снимком; получаем естественную числовую характеристику пространственного хаоса. Отметим, что если мгновенный снимок периодичен по  $x$ , то ему соответствует периодическая траектория динамической системы сдвигов — размерность равна единице; если квазипериодичен, ему соответствует обмотка на торе — размерность равна целому числу, большему единицы; дробная размерность отвечает хаотическому пространственному поведению.

Если исследуемый процесс имеет две или три пространственные координаты, то ему можно поставить в соответствие так называемую динамическую систему с двумерным или трехмерным временем. Отображение сдвига для нее на пару  $(t_1, t_2)$  или тройку  $(t_1, t_2, t_3)$  выглядит следующим образом:  $f(x, y) \rightarrow f(x + t_1, y + t_2)$  или  $f(x, y, z) \rightarrow f(x + t_1, y + t_2, z + t_3)$ . Мгновенному снимку (который в этом случае выглядит как некий рельеф) тоже можно поставить в соответствие предельное множество универсальной динамической системы с двумерным (или трехмерным) временем и его размерность считать характеристикой пространственного хаоса. При изменении управляющего параметра (например, вкачиваемой в систему энергии) размерность может меняться, пространственные картины могут усложняться — можно говорить о пространственных бифуркациях и т. д. Как нам кажется, такой динамический подход при исследовании пространственных свойств процесса может оказаться плодотворным.

В заключение авторы выражают искреннюю благодарность М. И. Рабиновичу за стимулирующие работу полезные обсуждения.

## ЛИТЕРАТУРА

1. *Mandelbrot B.* Fractals: Form, chance and dimension. San-Francisco: Freeman, 1977. 365 p.
2. *Umberger D. K., Mayer-Kress G., Jen E.* Hausdorff dimension for sets with broken scaling symmetry // Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. P. 42—53.
3. *Umberger D. K., Farmer J. D.* Fat fractals on the energy surface // Phys. Rev. Lett. 1985. Vol. 55, N 7. P. 661—664.
4. *Bowen R.* Entropy of group diffeomorphisms and homogeneous spaces // Trans. Amer. Mat. Soc. 1971. Vol. 153. P. 401—414.
5. *Takens F.* Detecting strange attractors in turbulence // Lect. Notes Math. 1980. Vol. 898. P. 336—382.
6. *Adler R., Konheim G., McAndrew M. H.* Topological entropy // Trans. Amer. Mat. Soc. 1965. Vol. 114. P. 309—319.
7. *Takens F.* Distinguishing deterministic and random systems // Nonlinear dynamics and turbulence / Ed. G. I. Barenblatt, G. Iooss, D. D. Joseph. N. Y.: Pitman, 1983. P. 314—333.
8. *Milnor J., Thurston W.* On iterated maps of the interval: Preprint. Princeton, 1977. 8 p.
9. *Малкин М. И.* О топологической классификации кусочно-монотонных отображений интервала // Методы качественной теории дифференциальных уравнений. Горький: ГГУ, 1980. С. 186—191.
10. *Малкин М. И.* О непрерывности энтропии разрывных отображений отрезка // Методы качественной теории дифференциальных уравнений. Горький: ГГУ, 1982. С. 35—47.
11. *Mañe R.* On the dimension of the compact invariant sets of certain nonlinear maps // Lect. Notes Math. 1981. Vol. 898. P. 230—242.
12. *Eckmann J. P., Ruelle D.* Ergodic theory of chaos and strange attractors // Rev. Mod. Phys. 1985. Vol. 57, N 3. P. 617—656.
13. *Eckmann J. P., Kamphorst S. O., Ruelle D., Cilibeto S.* Liapunov exponents from time series: Preprint. 1986. 12 p.
14. *Grassberger P., Procaccia I.* Characterization of strange attractors // Phys. Rev. Lett. 1983. Vol. 50, N 5. P. 346—349.
15. *Grassberger P., Procaccia I.* Measuring the strangeness of strange attractors // Physica D. 1983. Vol. 9. P. 189—208.
16. *Young L.* Dimension, entropy and Liapunov exponents // Ergod. Theory and Dyn. Syst. 1982. N 2. P. 109—124.
17. *Ben-Mizrachi A., Grassberger P., Procaccia I.* Measuring the dimension of experimental (noisy) attractors // Phys. Rev. A. 1984. Vol. 29, N 2. P. 975.
18. *Арансон И. С., Рейман А. М., Шехов В. Г.* Методы измерения корреляционной размерности в эксперименте // Наст. сб. С. 262—266.
19. *Caputo J. P., Malraison B., Atten P.* Determination of attractor dimension and entropy for various flows. An experimentalist's viewpoint // Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. P. 180—190.
20. *Рабинович М. И., Фабрикант А. Л.* Стохастическая автомодуляция волн в неравновесных средах // ЖЭТФ. 1979. Т. 74, № 4. С. 617—621.
21. *Brandstätter A., Swinney H. L., Chapman G. T.* Characterizing turbulent channel flow // Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. P. 150—157.
22. *Badii R., Politi A.* On the fractal dimension of filtered chaotic signals // Ibid. P. 67—73.
23. *Brandstätter A., Swift J., Swinney H. L. et al.* Low-dimensional chaos in a hydrodynamic system // Phys. Rev. Lett. 1983. Vol. 51. P. 1442—1444.
24. *Козлов В. В., Рабинович М. И., Рамазанов М. П. и др.* Корреляционная размерность потока и пространственное развитие динамического хаоса в пограничном слое // Письма в ЖТФ. 1987. Т. 13, № 16. С. 896—901.

25. Громов П. Р., Зобнин А. Б., Рабинович М. И. и др. Конечномерные аттракторы в сдвиговых течениях с обратной связью // ДАН СССР. 1987. Т. 292, № 2. С. 284—287.
26. Held G. A., Jeffries C. D. Characterization of chaotic instabilities in an electron — hole plasma in germanium // Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. P. 158—170.
27. Martin S., Martienssen W. Transition from quasiperiodicity into chaos in the periodically driven conductivity of BSN crystal // Ibid. P. 191—197.
28. Babloyantz A. Evidence of chaotic dynamics of brain activity during the sleep cycle // Ibid. P. 241—256.
29. Афраймович В. С. Внутренние бифуркации и кризисы аттракторов // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 189—213.
30. Tabeling P. Sudden increase of the fractal dimension in a hydrodynamic system // Phys. Rev. A. 1985. Vol. 31, N 5. P. 3460—3462.
31. Holzjuss J., Mayer-Kress G. An approach to error-estimation in the application of dimension algorithms // Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. P. 114—122.
32. Aranson I. S., Rabinovich M. I., Starobinets I. M. Transitions in chaos // Nonlinear and turbulent processes in physics. N. Y., 1984. Vol. 3. P. 1139.
33. Crutchfield J. P., McNameara B. S. Equations of motion from a data series // J. Complex Syst. 1987. N 1. P. 1—14. Preprint.
34. Афраймович В. С., Рабинович М. И., Угодников А. Д. Критические точки и фазовые переходы в параметрически возбуждаемом ангармоническом осцилляторе со стохастическим поведением // Письма в ЖЭТФ. 1983. Т. 38, № 2. С. 64—67.
35. Шерешевский М. А. Об асимптотике хаусдорфовой размерности базисного множества, рождающегося при исчезновении состояния равновесия типа седло—седло // Функцион. анализ и его прил. 1987. Т. 21, № 1. С. 88—89.

УДК 519.254

## МЕТОДЫ ИЗМЕРЕНИЯ КОРРЕЛЯЦИОННОЙ РАЗМЕРНОСТИ В ЭКСПЕРИМЕНТЕ

И. С. АРАНСОН, А. М. РЕЙМАН, В. Г. ШЕХОВ

1. Несмотря на то что методы теории динамических систем используются для анализа и диагностики разнообразных случайных реализаций менее пяти лет, полученные здесь результаты очень интересны и в определенном смысле уникальны [1]. Дело в том, что применение традиционных способов анализа (корреляционного, спектрального, регрессионного и т. д.) во многих случаях в принципе не может дать желаемой информации. Например, наличие сплошного спектра временной реализации может быть обусловлено, с одной стороны, внешними шумами, с другой — сложной собственной динамикой нелинейной системы [2].

Наиболее эффективным методом, с помощью которого в эксперименте удается установить детерминистическую природу стохастического сигнала, является построение корреляционного интеграла и определение по нему корреляционной размерности и энтропии [3].

Процедура построения корреляционного интеграла состоит в следующем. Сформируем из временной реализации  $x(t)$  квантованием через некоторое время  $\tau$  набор чисел  $\{x_n\}_{n=1}^N$ , где  $n = 1, \dots, N$ ;  $x_n = x(n\tau)$ . Из полученного набора построим  $d$ -мерные векторы  $\mathbf{x}_i = (x_i, \dots, x_{i+d-1})$ . Тогда корреляционным интегралом порядка  $d$  будем называть функцию  $C^d(l)$ , равную вероятности того, что расстояние между двумя векторами меньше  $l$  (здесь мы считаем, что для векторов  $\mathbf{x}_i, \mathbf{x}_j$  тем или иным способом введена функция расстояния  $\rho(\mathbf{x}_i, \mathbf{x}_j)$ )<sup>1</sup>.

Установлено, что для типичных процессов, порождаемых динамическими системами, существует область масштабов  $l$ , для которой, начиная с некоторых достаточно больших  $d$ ,  $C^d(l) \sim l^{\nu_s}$ , где  $\nu$  — корреляционная размерность. Ограниченность величины  $\nu$ , в частности, означает, что при соответствующем огрублении можно считать, что сигнал порождается динамической системой размерности не более  $2\nu + 1$  [4].

Нахождение корреляционной размерности реальных сигналов обычно связано с большими вычислительными сложностями: необходима обработка временных реализаций длиной  $N$  порядка  $10^4$  и более точек, причем для построения корреляционного интеграла требуется произвести  $N^2$  сравнений векторов, что и приводит к большим затратам машинного времени. Более того, с ростом  $\nu$  неизбежно возрастает величина  $N$ , при которой наблюдается сходимость корреляционных интегралов  $C^d(l)$  при  $N \rightarrow \infty$  к промежуточной асимптотике  $l^\nu$  [5].

В настоящей работе предложен, во-первых, эффективный алгоритм построения корреляционного интеграла для экспериментальных сигналов, во-вторых, принцип его нахождения, в определенном смысле обобщающий оригинальный [6].

2. Обычно входной сигнал  $x(t)$  вводится в ЭВМ через АЦП и хранится в виде последовательности чисел в двоичном коде  $\{x_n\}$ . Интервал квантования  $\tau$  выбирается из условия  $\tau \leq 1/2F_{\max}$ , где  $F_{\max}$  — эффективная граница спектра сигнала. На практике это условие удобно контролировать по виду зависимости  $x_{n+1}(x_n)$ : при слишком малой величине  $\tau$  все точки оказываются выстроенными вдоль диагонали квадрата, длина стороны которого равна размаху сигнала  $x(t)$ ; увеличением  $\tau$  добиваются наилучшего заполнения квадрата.

Обычно динамический диапазон сигнала не превышает 40—50 дБ, однако максимальное значение  $x_{\max} = \max_{\forall t} |x(t)|$  может сильно меняться от реализации к реализации. Для удобства вычислений необходимо либо применять АЦП большой разрядности (12—15 разрядов), либо, при использовании 8—10-разрядных

<sup>1</sup> Для типичных сигналов корреляционная размерность и энтропия не зависят от способа определения расстояния [3].

<sup>2</sup> Оценка размерности  $2\nu + 1$ , как правило, является завышенной. В большинстве же случаев аттрактор удастся вложить в евклидово пространство размерности  $\nu + 1$  [1].



АЦП, регулировать усиление, приводя  $x_{\max}$  к максимальному значению разрядной сетки преобразователя.

Основой для построения  $C^d(l)$  является измерение расстояний между  $d$ -мерными векторами  $\mathbf{x}_i = (x_i, x_{i+1}, \dots, x_{i+d-1})$  и  $\mathbf{x}_j = (x_j, x_{j+1}, \dots, x_{j+d-1})$ . Для ускорения обработки на ЭВМ эффективно использование двух вариантов вычисления расстояния:

$$\rho(\mathbf{x}_i, \mathbf{x}_j) = \begin{cases} \frac{1}{d} \sum_{k=1}^d |x_{i+k-1} - x_{j+k-1}|, & (1a) \\ \max_{k=1 \dots d} |x_{i+k-1} - x_{j+k-1}|. & (1b) \end{cases}$$

Оба варианта соответствуют кубической форме ячеек, на которые разбивается фазовое пространство при подсчете вероятностей<sup>3</sup>. Такой способ определения расстояний позволяет использовать операции с целыми числами, которые, особенно на малых ЭВМ, выполняются существенно быстрее, чем операции с плавающей запятой. Удобнее применять вариант (1б), не требующий деления и не выводящий значения  $\rho(\mathbf{x}_i, \mathbf{x}_j)$  за пределы целочисленной разрядной сетки ЭВМ.

Будем строить распределение вероятностей для величины  $\rho(\mathbf{x}_i, \mathbf{x}_j)$  в интервале  $(0, x_{\max})$ , введя разбиение этого интервала на  $n$  ячеек. Очевидно, что  $\rho_1 = |\mathbf{x}_i - \mathbf{x}_j|$  есть расстояние между двумя одномерными векторами  $\mathbf{x}_i, \mathbf{x}_j$ . Поскольку множество значений  $\rho$  состоит из целых чисел, то его элементом можно сопоставить номера ячеек (номер ячейки  $m_1$ , соответствующий значению  $\rho_1$ , равен  $m_1 = \rho_1 n / x_{\max}$ ). Аналогично определяются номера ячеек  $m_d$  и для всех других значений  $d$ .

Пусть программа обработки организована так, что в каждом элементе вспомогательного двумерного массива  $C[n, d_{\max}]$  накапливается информация о числе попаданий  $d$ -мерного вектора в интервал  $\left[ \frac{m-1}{n} x_{\max}, \frac{m}{n} x_{\max} \right]$ . Тогда для корреляционного интеграла получаем выражение

$$C^d(l) = \sum_{k=1}^{m^*} C(k, d), \quad \text{где } l = \frac{m^*}{n} x_{\max}.$$

Заметим, что можно организовать процесс вычисления расстояний между  $d$ -мерными векторами таким образом, чтобы на промежуточных стадиях определялись все предшествующие расстояния ( $d^* < d$ ). Дополнительный выигрыш по времени дает замена линейной сетки ячеек на логарифмическую, заранее вычисленную и занесенную в таблицу.

<sup>3</sup> Способ определения расстояний, основанный на евклидовой норме:

$$\rho(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{d} \left[ \sum_{k=1}^d (x_{i+k-1} - x_{j+k-1})^2 \right]^{1/2}, \quad \text{соответствует сферической форме ячеек.}$$

3. В настоящее время корреляционные интегралы определяются путем вычисления вероятности для расстояний от набора фиксированных векторов для всех остальных, сформированных по входному сигналу. Уже на этом этапе предполагается, что сигнал обладает свойством эргодичности, так как информация получается путем анализа достаточно длинной реализации. Однако само вычисление корреляционного интеграла основано на фазовом усреднении, поскольку вероятность определяется по всем комбинациям  $\{x_i, x_j\}$  или по всему фазовому пространству. Покажем, что условие эргодичности позволяет вычислить корреляционный интеграл путем только временного усреднения.

Рассмотрим сигнал  $x(t)$  и сдвинутый на время  $\tau_0$  сигнал  $x(t - \tau_0)$ , где  $\tau_0 \gg \tau_k$ ,  $\tau_k$  — время корреляции процесса  $x(t)$ . Очевидно, что, несмотря на статистическую независимость этих сигналов, обе реализации эквивалентны (в отношении статистических характеристик). Требование ко времени  $\tau_0$  должно учитывать и время перемешивания системы, но для реальных сигналов условие  $\tau_0 \gg \tau_k$  оказывается достаточным.

Будем строить корреляционный интеграл следующим образом: в качестве пар векторов выберем  $x_i = \{x(i\tau), x[(i+1)\tau], \dots, x[(i+d-1)\tau]\}$ ,  $x_j = \{x(i\tau - \tau_0), x[(i+1)\tau - \tau_0], \dots, x[(i+d-1)\tau - \tau_0]\}$ , т. е.  $j = i - \tau_0/\tau$ . Распределение вероятности величины  $\rho(x_i, x_j)$  будем определять прямым измерением времени выполнения условия  $\rho < l$ . Так как предполагается, что анализируемый сигнал  $x(t)$  обладает свойством эргодичности, то при  $N \rightarrow \infty$  соответствующие величины будут сходиться к вероятностям, полученным традиционным способом (на основе усреднения по фазовому пространству).

Как показали численные эксперименты, такое построение корреляционных интегралов зачастую оказывается более эффективным по скорости сходимости к промежуточной асимптотике  $l^N$  по сравнению с методами фазового усреднения. Так, в системе Лоренца [5] получим достоверное значение  $\nu = 2,03$  при  $N = 16\ 000$ , что эквивалентно по затратам времени полной обработке реализации всего из 100 точек, для которой такая сходимость не может быть достигнута. Достоинством такого алгоритма является также то, что он позволяет вести обработку сигнала в реальном масштабе времени, так как не требуется запоминание реализации процесса.

В то же время наличие постоянного сдвига между парами векторов эквивалентно некоторому прореживанию сигнала, в результате чего могут оказаться заполненными не все ячейки сетки вероятностей. Такая ситуация возникает при обработке сигналов, содержащих мощные дискретные спектральные компоненты (см., например, [7, 8]). При этом описанная методика в ряде случаев может давать существенную ошибку вычисления  $\nu$ , и даже значительное увеличение длины реализации не позволяет восстановить истинное распределение вероятностей.

Этого недостатка лишен метод редуцированного корреляционного интеграла: вычисление вероятности производится не по всем парам  $\{x_i, x_j\}$ , а только по ограниченному их числу. При этом  $M$  точек в фазовом пространстве являются опорными (относительно этих точек строятся гиперкубы сетки вероятностей), и каждая из них сравнивается с  $K$  точками ( $K, M \leq N$ ). При этом максимальная достоверно измеряемая размерность определяется логарифмом отношения полного числа пар  $MK$  к числу опорных точек; степень же изрезанности  $C^d(l)$ , т. е. качество усреднения частных сумм, определяется значением  $M$  [3, 8]. Очевидно, выгодно выбирать  $K = N$ , а  $M$ , как показывает эксперимент, может составлять величину порядка нескольких десятков. Такое сокращение вычислений позволяет поднять скорость обработки почти до уровня алгоритма временного усреднения. Для микро-ЭВМ типа LSI-11/23 время обработки одной реализации длиной 90 K точек составляет несколько минут, что позволяет обработать данные непосредственно в ходе эксперимента.

Описанные алгоритмы позволяют существенно поднять скорость обработки сигналов в эксперименте и открывают возможность построения нового класса анализирующих систем.

Авторы признательны М. И. Рабиновичу за постоянный интерес к работе и обсуждение результатов.

#### ЛИТЕРАТУРА

1. Schuster H. G. Deterministic chaos: An introduction. Weinheim: Physik Verl., 1984. 224 p.
2. Гапонов-Грегов А. В., Рабинович М. И. Нелинейная физика. Стохастичность и структуры // Физика XX века. Развитие и перспективы. М.: Наука, 1984. С. 219—280.
3. Афраймович В. С., Рейман А. М. Размерность и энтропия в многомерных системах // Наст. сб. С. 238—262.
4. Mane R. On dimension of the compact invariant sets of the certain nonlinear maps // Lect. Notes Math. 1981. Vol. 898. P. 230—242.
5. Dimensions and entropies in chaotic systems / Ed. G. Mayer-Kress. Berlin; Heidelberg: Springer, 1986. 343 p.
6. Grassberger P., Procaccia I. Characterization of strange attractors // Phys. Rev. Lett. 1983. Vol. 50, N 5. P. 346—349.
7. Козлов В. В., Рабинович М. И., Рамазанов П. И. и др. Корреляционная размерность потока и пространственное развитие динамического хаоса в пограничном слое // Письма в ЖТФ. 1987. Т. 18, № 16. С. 986—991.
8. Громов П. Р., Зобнин А. Б., Рабинович М. И. и др. Конечномерные аттракторы в сдвиговых течениях с обратной связью // ДАН СССР. 1987. Т. 292, № 2. С. 284—287.

## КВАНТОВАЯ ЛОКАЛИЗАЦИЯ ДИНАМИЧЕСКОГО ХАОСА

Д. Л. ШЕПЕЛЯНСКИЙ

К настоящему времени явление динамического хаоса в классических системах достаточно хорошо изучено и в основном понято [1—3]. Вместе с тем исследование простых моделей показало, что динамика квантовых систем, стохастических в классическом пределе, обладает рядом особенностей (см., например, [2, 4, 5]), из которых наиболее интересным является эффект квантового ограничения диффузии [4—6]. Примером физической системы, в которой указанные выше эффекты играют существенную роль, служит высоковозбужденный атом водорода в микроволновом поле [7]. При напряженности поля, превышающей некоторое критическое значение, движение классического электрона становится хаотическим, что приводит к диффузионной ионизации атома [8—11]. Такая картина процесса согласуется с экспериментальными данными [7] и была более детально подтверждена в эксперименте [12]. Вместе с тем при определенных условиях в этой системе происходит квантовое ограничение диффузии, что ведет к резкому падению вероятности ионизации по сравнению с классическим значением [13—15]. Ниже будет показано, что квантовая локализация хаоса проявляется и при диффузионном возбуждении молекул (см. [16, 17]).

Прекращение диффузии в квантовой системе является следствием локализации собственных функций квазиэнергии (СФКЭ), которая аналогична локализации Андерсона в твердом теле [18]. Эта аналогия позволяет рассматривать возбуждение многоуровневой системы в периодическом поле как задачу о локализации в твердом теле и использовать полученные там результаты [19]. При этом номер невозмущенного уровня играет роль пространственной координаты (номер узла в решетке). Следует, однако, подчеркнуть, что, несмотря на указанную аналогию, явление квантовой локализации хаоса принципиально отличается от локализации Андерсона тем, что в системе нет никаких случайных параметров и соответствующий твердотельный потенциал является детерминированной функцией номера узла решетки.

## 1. МОДЕЛЬ КВАНТОВОГО РОТОРА

Для анализа динамики квантовых систем, стохастических в классическом пределе, рассмотрим модель квантового ротатора с гамильтонианом

$$H = \hat{n}^2/2 + k \cos \theta \delta_T(t), \quad (1)$$

где  $\hat{n} = -i\partial/\partial\theta$ ;  $\delta_T(t)$  — дельта-функция с безразмерным периодом  $T$ ;  $\theta$  — фаза в интервале  $[0, 2\pi]$ ;  $V(\theta) = k \cos \theta$  — внешнее возмущение;  $\hbar = 1$  [4—6, 18]. Здесь  $H_0(n) = n^2/2$  задает энергии

невозмущенных уровней  $n$ . Динамика соответствующей классической системы определяется уравнениями движения с гамильтонианом (1), где  $n, \theta$  — классические сопряженные переменные действие—фаза.

После интегрирования на период  $T$  получаем стандартное отображение [1, 3]

$$\bar{p} = p + K \sin \theta, \quad \bar{\theta} = \theta + \bar{p}, \quad (2)$$

где черта означает новые значения переменных через период, а  $p = Tn$ ,  $K = kT$ . При  $K \leq K_{\text{кр}} = 0,9716 \dots$  [20] изменение  $n$  ограничено:  $|\Delta n| \lesssim \sqrt{k/T}$ , а при  $K > K_{\text{кр}}$  величина  $(\Delta n)^2$  растет по диффузионному закону со скоростью  $D = (\Delta n)^2/\tau = D_0(K)/T^2$ , где  $D_0(K)$  — скорость диффузии по  $p$  в стандартном отображении (2), а  $\tau$  — число периодов возмущения. Условие квазиклассичности имеет вид  $k \gg 1$ ,  $T \ll 1$ ,  $K = \text{const}$ .

В единой хаотической компоненте зависимость скорости диффузии от параметра стохастичности  $K$  приближенно описывается следующим выражением:

$$D_0 \approx \begin{cases} \frac{K^2}{2} [1 + 2J_2(K) + 2J_2^2(K)]_x & K \geq 4,5, \\ 0,30(\Delta K)^3, & K < 4,5_x \end{cases} \quad (3)$$

где  $J_2(K)$  — функция Бесселя;  $\Delta K = K - K_{\text{кр}}$ . При  $K \geq 4$ , 5 зависимость  $D_0(K)$  имеет вид осцилляций, затухающих с ростом  $K$  [3, 21]. Предельное значение  $D_0 = K^2/2$  соответствует квазилинейному приближению, когда фазы  $\theta$  ( $\tau$ ) в (2) случайны и независимы. При  $K \rightarrow K_{\text{кр}}$  в (3) использована эмпирическая формула, полученная из численных экспериментов [6]. Значение показателя  $\eta \approx 3$  в степенной зависимости  $D_0 \propto (\Delta K)^\eta$  близко к значению, принятому в [22].

Численные эксперименты с квантовым стандартным отображением (1) показали [4—6, 18, 23, 24], что с течением времени диффузионный рост  $\langle n^2 \rangle$  останавливается, т. е. под действием поля эффективно возбуждается конечное число уровней ( $\Delta n \sim l$ ). Естественная интерпретация такого квантового ограничения диффузии связана с локализацией СФКЭ, которая аналогична локализации Андерсона в одномерной решетке [6, 18, 23, 24]. В работах [5, 6] была получена теоретическая оценка для числа возбужденных уровней и длины локализации СФКЭ:

$$l = \alpha D \sim \Delta n \sim \tau_D, \quad (4)$$

где  $\alpha$  — некоторая числовая константа.

Вывод соотношения (4) основан на следующих соображениях. Пусть одно невозмущенное состояние разлагается на  $l$  СФКЭ с квазиэнергиями  $v_i$ . Так как эти  $v_i$  распределены на интервале  $[0, 2\pi]$ , то среднее расстояние между ними  $\Delta v \sim 1/l$  (рассматривается случай, когда невозмущенные уровни равномерно распределены в интервале  $2\pi$ ). Если первоначально возбуждено одно

невозмущенное состояние, то диффузия будет продолжаться только в течение конечного времени  $t \sim \tau_D$ , пока не проявится дискретность спектра. Согласно соотношению неопределенности,  $\tau_D \sim 1/\Delta v \sim l$ . За это время диффузионно возбудится число уровней  $\Delta n \sim (D\tau_D)^{1/2} \sim l$ . Откуда и получаем соотношение (4). При этом условие его применимости —  $D \gg 1$ .

## 2. ЛОКАЛИЗАЦИЯ СФКЭ

Рассмотрим уравнение, которому удовлетворяет СФКЭ модели (1) с квазиэнергией  $v$  [18]:

$$u_n^- = \exp\{i[v - TH_0(n)]\} u_n^+, \quad u^+(\theta) = \exp[-iV(\theta)] u^-(\theta). \quad (5)$$

Здесь  $u^\mp(\theta)$  — значения функции до и после действия возмущения;  $u_n^\pm$  — фурье-компоненты  $u^\pm(\theta)$ . После простых преобразований (5) сводится к виду

$$\hat{H}_{ss}u = \left[ \cos \frac{\hat{V}}{2} \operatorname{tg} \left( \frac{v}{2} - \frac{T}{2} \hat{H}_0 \right) \cos \frac{V}{2} - \frac{1}{2} \sin \hat{V} \right] u = 0, \quad (6)$$

где  $u = \exp(\pm iV/2)u^\pm$ . После фурье-преобразования  $\exp(-iV/2) = \sum_r W_r^- \exp[i(r\theta + \varphi_r)]$  видно, что гамильтониан  $H_{ss}$  соответствует некоторой одномерной решетке с взаимодействующими соседями и энергией  $E = -\sum_r W_r^- W_{-r}^- \sin \varphi_r \cos \varphi_{-r}$ . При

этом квазиэнергия  $v$  как бы определяет потенциал взаимодействия, а собственное значение энергии  $E$  рассматривается как параметр. Кроме того, невозмущенный номер уровня  $n$  в модели (1) соответствует дискретной пространственной координате в решетке.

Поскольку в одномерной случайной цепочке все собственные функции локализованы [19], то естественно ожидать экспоненциальной локализации СФКЭ в (1). В том случае, когда  $\cos(V/2) \neq 0$ , можно ввести  $\bar{u} = u \cos(V/2)$  и, разделив уравнение (6) на  $\cos(V/2)$ , свести задачу к случаю диагонального беспорядка с  $\bar{H}_{ss} = \operatorname{tg} \left( \frac{v}{2} - \frac{T}{2} \hat{H}_0 \right) - \operatorname{tg} \frac{\hat{V}}{2}$ . Такой метод был неявно использован в [18]. Следует, однако, подчеркнуть, что такой путь приводит к возникновению нефизической особенности, которая не позволяет рассматривать широкий класс потенциалов с  $|V| \geq \pi$ .

Форма уравнения (6) удобна для проведения аналогии с задачами твердого тела. Вместе с тем для численных экспериментов удобней переписать (5), (6) следующим образом. Введем  $\bar{u} = \exp(\pm iV/2)u^\pm/g$ , где  $g$  — произвольная действительная функция  $\theta$ , и ограничимся случаем, когда  $g$  и  $V$  четные. Тогда из (5) получаем

$$\sum_r \bar{u}_{n+r} W_r \sin(\chi_n + \varphi_r) = 0, \quad (7)$$

где  $\exp(-iV/2)g = \sum_r W_r \exp[i(r\theta + \varphi_r)]$ ;  $\chi_n = [v - TH_0(n)]/2$ .

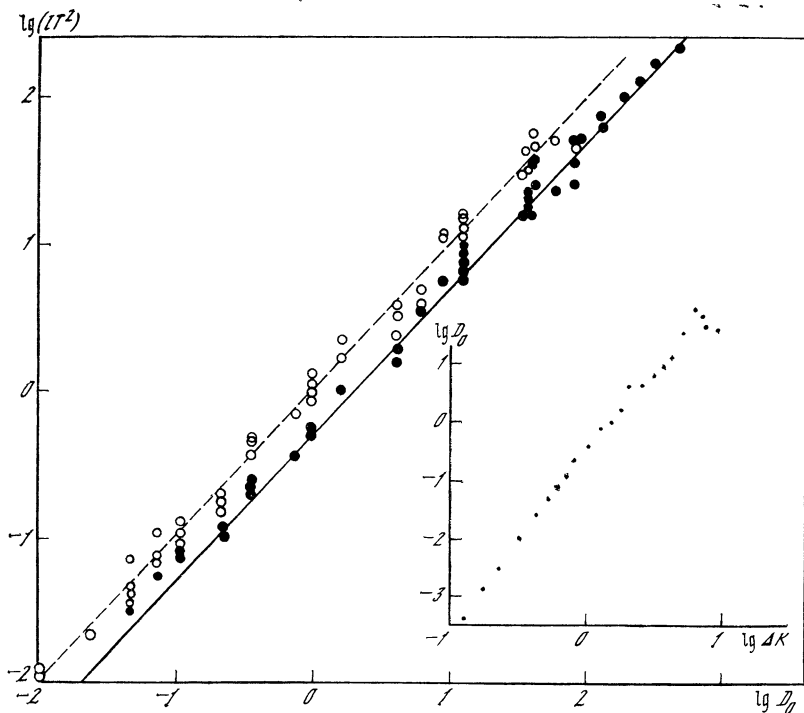


Рис. 1

В том случае, когда отличны от нуля только  $W_r$  с  $|r| \leq N$ , столбец из  $2N$  известных значений  $\overline{u_n}$  определяет значения СФКЭ при любом  $n$ .

Рекуррентное вычисление  $\overline{u_n}$  из (7) можно рассматривать как некоторую динамическую систему в дискретном времени  $n$ . Динамика  $2N$ -компонентного вектора определяется матрицей переноса  $M_n$ , вид которой легко определяется из (7). Можно показать, что динамика по  $n$  является гамильтоновой и в системе имеются  $N$  положительных и  $N$  отрицательных показателей Ляпунова (ПЛ), причем  $\gamma_i^+ = -\gamma_i^-$  (см., например, [3]). Асимптотически длина локализации СФКЭ определяется минимальным положительным ПЛ  $\gamma_1 = 1/l$  [23, 24]. Отличие  $\gamma_1$  от нуля приводит к тому, что все СФКЭ экспоненциально локализованы:  $u_n \propto \exp(-|n|/l)$ , а спектр квазиэнергий является чисто точечным. Техника вычисления *всех* ПЛ подробно описана в [3].

Выбрав  $g = 1$ , получим, что для (1)  $W_r = J_r(k/2)$ ,  $\varphi_r = -\pi r/2$ ,  $\chi_n = (v - Tn^2/2)/2$ , где  $J_r(k/2)$  — функция Бесселя. Из-за быстрого падения  $W_r$  при  $|r| > k/2$  можно ограничиться конечным числом соседей  $N \sim k/2$ . Результаты численных экспериментов [23], полученных методом ПЛ (точки на рис. 1), показывают, что в квазиклассической области ( $T \leq 1$ ,  $5 \leq k \leq 75$ ,  $1,5 \leq K \leq$

$\leq 29$ ,  $T/4\pi$  — типичное иррациональное число) длина локализации СФКЭ удовлетворительно описывается соотношением (прямая)

$$l = D_0 (K)/2T^2, \quad (8)$$

которое сохраняется не только в области развитого хаоса с  $K \gg \gg 4,5$ , где мера островков устойчивости пренебрежимо мала [1], но и при  $\Delta K = K - K_{кр} \ll 1$ , когда скорость диффузии определяется сложной критической структурой [22] и устойчивая компонента занимает около 50% всей фазовой плоскости [1]. При этом скорость диффузии в стандартном отображении  $D_0$  меняется на четыре порядка (рис. 1, на вставке дана зависимость  $D_0$  от  $\Delta K$  [6]). Численное значение  $\alpha = 1/2$  получено путем сравнения (4) с точно решаемой моделью Ллойда [23].

На рис. 1 приведены также численные данные [6] для длины локализации стационарного распределения вероятностей по невозмущенным уровням, устанавливающегося в процессе эволюции начального состояния с  $n = n_0$ :  $\bar{f}(n) \propto \exp(-2 |n - n_0| / l)$ . Среднее значение  $\langle \alpha \rangle = 1,04$  (пунктирная прямая) близко к единице и существенно отличается от значения  $\alpha = 1/2$  для СФКЭ. Причина этого заключается в сильных флуктуациях СФКЭ [6, 25].

### 3. ДИФФУЗИОННЫЙ ФОТОЭФФЕКТ В АТОМЕ ВОДОРОДА

Рассмотрим процесс ионизации высоковозбужденного атома водорода в монохроматическом линейно-поляризованном поле. Для состояний, вытянутых вдоль поля (магнитное квантовое число  $m = 0$ , а параболические числа  $n_1 \gg n_2$ ), динамику возбуждения можно описывать одномерным гамильтонианом [13—15]

$$H = -\frac{1}{2n^2} + \varepsilon n^2 \cos \omega t \left[ \frac{3}{2} - 2 \sum_{s=1}^{\infty} \frac{J'_s(s)}{s} \cos(s\lambda) \right], \quad (9)$$

где  $n, \lambda$  — сопряженные действие — фаза;  $\varepsilon, \omega$  — напряженность и частота поля и использованы атомные единицы.

От непрерывных уравнений движения удобно перейти к отображению за один оборот электрона вокруг ядра. Чтобы найти производящую функцию  $G$  для этого отображения запишем координату  $z$  в невозмущенных переменных и проинтегрируем уравнения Гамильтона на одном обороте. В результате получим, что

$$G(\bar{N}, \varphi) = \bar{N}\varphi + 2\pi (-2\omega \bar{N})^{-1/2} + kA(\bar{\kappa}) \cos \varphi.$$

Здесь  $N = E/\omega = -1/(2\omega n^2)$ ;  $\varphi$  — сопряженная с  $N$  переменная, равная значению  $\omega t$  в момент прохождения перигелия;  $k = 0,822\pi\varepsilon/\omega^{5/2}$ , а  $A(\kappa) = \frac{\kappa^{2/3}}{0,411} J'_\kappa(\kappa)$ , где  $J_\kappa(\kappa)$  — функция Ангера (при целых  $\kappa$  она совпадает с функцией Бесселя, при  $\kappa \rightarrow 0$  производная  $J'_\kappa(\kappa) = 1/2\kappa$ ,  $\kappa = \omega (-2\omega \bar{N})^{-3/2}$ ).



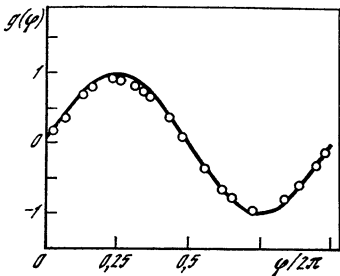


Рис. 2

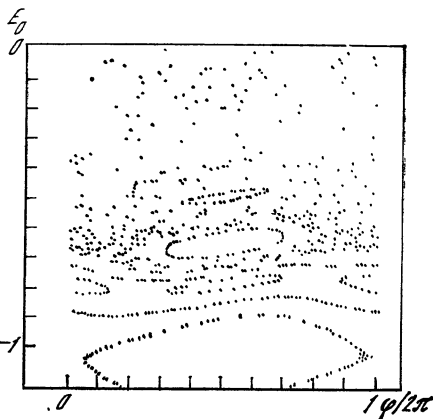
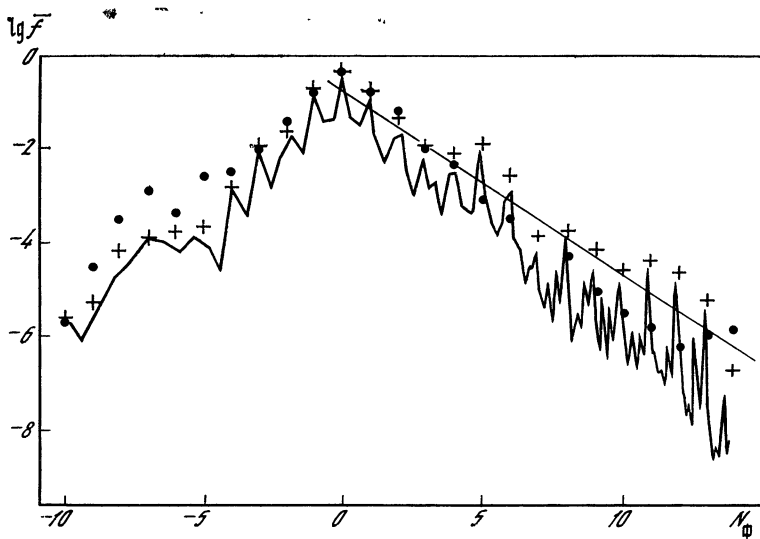


Рис. 3

Рис. 4



При  $\kappa \gg 1$  функция  $A \rightarrow 1$  и отображение принимает особенно простой вид

$$\bar{N} = N + k \sin \varphi, \quad \bar{\varphi} = \varphi + 2\pi\omega (-2\omega \bar{V})^{-1/2}, \quad (10)$$

где черта означает новые значения переменных после одного оборота. Для проверки (10) при численном решении непрерывных уравнений (9) фиксировалась фаза  $\omega t = \varphi$  в момент прохождения перигелия. По трем полученным значениям  $\varphi$  находилась функция  $g(\varphi) = (\bar{N} - N)/k$ . Сравнение теоретической зависимости  $g = \sin \varphi$  (линия) с численными данными (точки) представлено на рис. 2 для начальных условий с  $\varepsilon_0 = \varepsilon n_0^4 = 0,04$ ,  $\omega_0 = \omega n_0^3 = 1,5$ . Таким образом, (10) вполне удовлетворительно описывает динамику вплоть до  $\omega_0 \approx 1$ .

Линеаризацией по  $N$  отображение (10) локально может быть сведено к стандартному отображению (2) с параметрами  $k$ ,  $T = 6\lambda\omega^2 n_0^5$  и  $K = kT = \varepsilon_0/\varepsilon_c$ , где  $\varepsilon_c \approx 1/(49\omega_0^{1/3})$ . Глобальная диффузия в системе возникает при  $K > 1$ , т. е.  $\varepsilon_0 > \varepsilon_c$ , что согласуется с [11]. В процессе диффузии значение  $K$  растет, фазы становятся случайными и независимыми, а скорость диффузии по  $N$  равна  $D = k^2/2$ .

Пример фазовой плоскости в переменных  $E_0 = N\omega n_0^2$ ,  $\varphi$  при  $\varepsilon_0 = 0,04$ ,  $\omega_0 = 3$  приведен на рис. 3. Показаны шесть устойчивых траекторий и одна хаотическая. Отметим, что ионизация в (10) происходит в результате одного «последнего» толчка, который переводит  $N < 0$  в  $N > 0$ , после чего траектория уходит на бесконечность. Время ионизации, выраженное в числе оборотов, равно  $t_I = N_I^2/D$ , где  $N_I = 1/(2\omega n_0^2)$ .

Теперь перейдем к квантованию (10) [26]. Из обычного соотношения  $\hat{E} = -i\partial/\partial t$  следует, что  $\hat{N} = -i\partial/\partial\varphi$  ( $-\infty < \varphi < \infty$ ). Поскольку возмущение периодически по  $\varphi$ , то в системе (10) наряду с квазиэнергией сохраняется квазиимпульс. Для начального уровня  $n_0$  квазиимпульс равен дробной части  $N_0 = -N_I$ . Внешнее возмущение приводит к возбуждению высоких гармоник квазиимпульса. Поскольку  $\varphi \propto t$ , то этот квазиимпульс соответствует обычной квазиэнергии в исходной системе (9). Таким образом, квантовое отображение (10) описывает процесс диффузии по квазигармоникам в (9). Номер квазигармоники соответствует числу поглощенных фотонов и равен  $N_\Phi = N - N_0$ . Квантование дает  $\hat{N}_\Phi = -i\partial/\partial\varphi$  ( $0 \leq \varphi \leq 2\pi$ ). В результате получаем отображение для волновой функции

$$\bar{\psi} = \exp(-i\hat{H}_0) \hat{P} \exp(-ik\cos\varphi) \psi, \quad (11)$$

где  $\hat{H}_0 = 2\pi[-2\omega(N_0 + \hat{N}_\Phi)]^{-1/2}$ , а  $\hat{P}$  — оператор проекции на дискретные состояния ( $N < 0$ ).

Как и в модели (1), следствием квантовых эффектов является локализация диффузии. Однородность диффузии приводит к однородной экспоненциальной локализации стационарного распределения с длиной  $l_\Phi = D$ . В результате устанавливается распределение вида [6, 26]

$$\bar{f}_{N_\Phi} = \frac{1}{2l_\Phi} (1+x) e^{-x}, \quad x = \frac{2|N_\Phi|}{l_\Phi}, \quad l_\Phi = 3,33 \frac{\varepsilon^2}{\omega^{10/3}}. \quad (12)$$

Пример стационарного распределения показан на рис. 4 для  $n_0 = 100$ ,  $\varepsilon_0 = 0,04$ ,  $\omega_0 = 3$  (сплошная линия — данные квантового моделирования (9), крестики — суммарная вероятность в интервале  $[N_\Phi - 1/2, N_\Phi + 1/2]$ , точки — результат итерирования отображения (11), прямая — среднеквадратичная подгонка по максимумам в распределении  $\bar{f}$  при  $N_\Phi > 0$ ). Здесь экспериментальное значение  $l_\Phi$  в 1,6 раза больше теоретического (12).

В том случае, когда длина локализации сравнима с числом фотонов, требуемым для ионизации  $l_\Phi \approx N_I = -N_0$ , в системе происходит делокализация и процесс возбуждения близок к классическому. Полученные результаты находятся в согласии с [14, 15]. Вместе с тем они позволяют описать процесс возбуждения на высоких уровнях и объяснить наблюдавшиеся в [14] многофотонные пики в распределении по уровням.

Из (11) получаем, что скорость однофотонной ионизации (в числе оборотов) равна  $\gamma_{1\Phi} = (k/2)^2$ , что согласуется со стандартным результатом теории возмущений ( $k \ll 1$ ). При  $k \gg 1$  и  $N_I \ll k$  распределение ионизованных электронов пропорционально  $J_{N_\Phi}^2(k)$ . В области локализации, когда  $N_I > l_\Phi > k > 1$ , скорость ионизации равна  $\gamma_\Phi \sim \sum_{N=N_I-k}^{N_I} \bar{f}_N \sim k \bar{f}_{N_I}$  (вероятность, теряемая за один оборот). В физическом времени получаем

$$\Gamma_\Phi \sim k f_{N_I} (k\omega)^{3/2} \sim \frac{\omega^{5/4}}{\sqrt{n_0}} \left( \frac{\varepsilon_q}{\varepsilon_0} \right)^{3/2} \exp \left[ -2 \left( \frac{\varepsilon_q}{\varepsilon_0} \right)^2 \right], \quad (13)$$

где  $\varepsilon_q = \omega_0^{3/4} / \sqrt{6,66n_0}$  — граница делокализации. При  $\varepsilon_0 > \varepsilon_q$  скорость ионизации определяется классической диффузией и при  $\omega_0 \approx 1$  оказывается значительно больше, чем скорость однофотонной ( $\omega_0 = n_0/2$ ) ионизации при той же напряженности поля  $t_I \gamma_{1\Phi} \sim (n_0/2)^{-1/3} \ll 1$  (см. [15, рис. 1]).

В заключение отметим, что длина локализации (12) (выраженная в числе поглощенных фотонов) может быть записана в виде  $l_\Phi = 2\pi^2 \mu^2 \varepsilon^2 \rho^2$ , где  $\mu = 0,411/\omega^{5/3} n^3$  и  $\rho = n^3$  — соответственно однофотонный матричный элемент и плотность состояний в атоме водорода (9) (подробнее см. [25, 26]). В таком виде эта формула применима и к более общей задаче о возбуждении системы с заданной плотностью уровней монохроматическим полем, которая может рассматриваться как модель бесстолкновительной диссоциации молекул [17].

#### ЛИТЕРАТУРА

1. Chirikov B. V. A universal instability of many-dimensional oscillato $\Gamma$  systems // Phys. Rep. 1979. Vol. 52, N 5. P. 263—380.
2. Заславский Г. М. Стохастичность динамических систем. М.: Наука, 1984. 271 с.
3. Лихтенберг А., Либерман М. Регулярная и стохастическая динамика. М.: Мир, 1984. 528 с.
4. Casati G., Chirikov B. V., Ford J., Izrailev F. M. Stochastic behavior of a quantum pendulum under a periodic perturbation // Lect. Notes Phys. 1979. Vol. 93. P. 334—352.
5. Chirikov B. V., Izrailev F. M., Shepelyansky D. L. Dynamical stochasticity in classical and quantum mechanics // Sov. Sci. Rev. C. 1981. Vol. 2. P. 209—267.
6. Чуриков Б. В., Шепелянский Д. Л. Локализация динамического хаоса в квантовых системах // Радиопизика. 1986. Т. 29, № 9. С. 1041—1049.
7. Bayfield J. E., Koch P. M. Multiphoton ionization of highly excited hydrogen atoms // Phys. Rev. Lett. 1974. Vol. 33, N 5. P. 258—261.

- ξ . Делоне Н. Б., Крайнов В. П., Зон Б. А. Диффузионный механизм ионизации высоковозбужденных атомов в переменном электромагнитном поле // ЖЭТФ. 1978. Т. 75, № 2(6). С. 445—453.
9. Leopold J. G., Percival I. C. Microwave ionization and excitation of Rydberg atoms // Phys. Rev. Lett. 1978. Vol. 41, N 14. P. 944—947.
  10. Меерсон Б. Н., Окс Е. А., Сасоров П. В. Стохастическая неустойчивость осциллятора и ионизация высоковозбужденных атомов под действием электромагнитного излучения // Письма в ЖЭТФ. 1979. Т. 29, № 1. С. 79—82.
  11. Делоне Н. Б., Крайнов В. П., Шепелянский Д. Л. Высоковозбужденный атом в электромагнитном поле // УФН. 1983. Т. 140, № 3. С. 355—392.
  12. Van Leeuwen K. A. H., Oppen G. V., Renwick S. et al. Microwave ionization of hydrogen atoms: Experiment versus classical dynamics // Phys. Rev. Lett. 1985. Vol. 55, N 21. P. 2231—2234.
  13. Shepelyansky D. L. Quantum diffusion limitation at excitation of Rydberg atom in variable field // Proc. Intern. Conf. Quantum Chaos, Como 1983, Plenum 1985. P. 187—204.
  14. Casati G., Chirikov B. V., Shepelyansky D. L. Quantum limitations for chaotic excitation of the hydrogen atom in a monochromatic field // Phys. Rev. Lett. 1984. Vol. 53, N 27. P. 2525—2528.
  15. Casati G., Chirikov B. V., Guarneri I., Shepelyansky D. L. New photoelectric ionization peak in the hydrogen atom // Ibid. 1986. Vol. 57, N 7. P. 823—826.
  16. Шурык Э. В. Нелинейный резонанс в квантовых системах // ЖЭТФ. 1976. Т. 71, № 6(12). С. 2039—2056.
  17. Акулин В. М., Дыхне А. М. Динамика возбуждения многоуровневых систем зонного типа в лазерном поле // Там же. 1977. Т. 73, № 6(12). С. 2098—2105.
  18. Fishman S., Grepel D. R., Prange R. E. Quantum dynamics of a nonintegrable system // Phys. Rev. A. 1984. Vol. 29, N 4. P. 1639—1647.
  19. Лифшиц И. М., Гредескула С. А., Пастур Л. А. Введение в теорию неупорядоченных систем. М.: Наука, 1982. 358 с.
  20. Greene J. M. A method for determining a stochastic transition // J. Math. Phys. 1979. Vol. 20, N 6. P. 1183—1201.
  21. Rechester A. B., White R. B. Calculation of turbulent diffusion for the Chirikov-Taylor model // Phys. Rev. Lett. 1980. Vol. 44, N 24. P. 1586—1589.
  22. MacKay R. S., Meiss J. D., Percival I. C. Transport in Hamiltonian systems // Physica D. 1984. Vol. 13, N 1/2. P. 55—81.
  23. Shepelyansky D. L. Localization of quasienergy eigenfunction in action space // Phys. Rev. Lett. 1986. Vol. 56, N 7. P. 677—680.
  24. Blümel R., Fishman S., Grinisti M., Smilansky U. Localization in the quantum description of the periodically perturbed rotor // Lect. Notes Phys. 1987. Vol. 263. P. 212—228.
  25. Шепелянский Д. Л. Локализация диффузионного возбуждения в многоуровневых системах: Препр. ИЯФ № 86-187. Новосибирск, 1987. 25 с.
  26. Casati G., Guarneri I., Shepelyansky D. L. Exponential photonic localization for the hydrogen atom in a monochromatic field: Prepr. INP N. 87-29. Novosibirsk, 1987. 13 p.

# СЛУЧАЙНОСТЬ И ПРЕДСКАЗУЕМОСТЬ ДИНАМИЧЕСКОГО ХАОСА

Ю. А. КРАВЦОВ

## 1. СОГЛАШЕНИЯ О СЛУЧАЙНОСТИ

В современной теории вероятностей, основанной на понятии вероятностной меры, к случайным относят величины, задаваемые своими вероятностными распределениями. Той же традиции придерживается и теоретическая физика. В частном случае спектрально-корреляционной теории, оперирующей только с моментами первого и второго порядка, признаками случайного, хаотического поведения процессов служат спадающие корреляции и сплошной характер спектра.

Наряду с теоретико-множественным определением случайности в физике встречаются другие трактовки, опирающиеся на те или иные специфические аспекты поведения процессов и не тождественные теоретико-множественному подходу. Так, в экспериментальной физике к случайным относят такие ситуации, которые характеризуются понятиями неконтролируемости, невозпроизводимости, неповторяемости, непредсказуемости. Несмотря на распространенность такого взгляда, эти качественные понятия до сих пор еще не получили адекватного формализованного описания. Кроме того, в последнее время в качестве синонимов случайности часто выступают такие характеристики физических систем, как фрактальная структура (фрактальная размерность) и гиперболичность (экспоненциальное разбегание траекторий в фазовом пространстве).

В математике тоже были предложены альтернативные трактовки случайности: достаточно сослаться на алгоритмический подход, развиваемый А. Н. Колмогоровым и его последователями [1—4]. Алгоритмический подход отождествляет случайность с алгоритмической сложностью последовательности или процесса и служит примером еще одного соглашения о случайности. Этот подход получил распространение в теории динамических систем.

Перечисленные выше примеры определений и признаков случайности свидетельствуют о том, что физики и математики фактически имеют дело не с одним, а с несколькими соглашениями о случайности. Все эти соглашения исходят из того, чтобы рассматривать случайность как отсутствие закономерности; однако они отличаются толкованием отсутствия закономерностей, которое выступает как отсутствие простых алгоритмов в алгоритмической теории вероятностей, как быстрое расщепление корреляций при анализе спадающей корреляции, как несоответствие прогнозу с позиций предсказуемости—непредсказуемости и т. д.

Применительно ко многим физическим системам различные

соглашения о случайности более или менее согласуются друг с другом. Однако в ряде важных случаев один и тот же процесс (последовательность, поле), случайный в рамках одного соглашения, может оказаться детерминированным в рамках другого. Например, динамический хаос выступает как случайный процесс с позиций алгоритмического подхода, поскольку хаос алгоритмически сложен, и одновременно как неслучайный, детерминированный процесс с точки зрения его предсказуемости на ограниченных интервалах времени; возможность предсказания связана с тем, что динамический хаос описывается детерминированными уравнениями.

## 2. СЛУЧАЙНОСТЬ КАК НЕПРЕДСКАЗУЕМОСТЬ

Сосуществование нескольких соглашений о случайности приводит к известным трудностям в терминологии, которые, однако, можно преодолеть, если уточнить, каким соглашением мы руководствуемся.

В данной работе мы сделаем попытку формализовать соглашение, в соответствии с которым случайность отождествляется с непредсказуемостью. Несмотря на чрезвычайно широкое распространение этого критерия в физике вообще и в экспериментальной физике в особенности, соглашение о случайности как непредсказуемости вплоть до недавнего времени еще не получило адекватного математического описания.

Это соглашение предполагает, что случайность и детерминированность не противопоставляются, а рассматриваются как полюса единого свойства, а именно свойства частичной детерминированности. Первоначально концепция частично детерминированных (т. е. частично предсказуемых) процессов была разработана для описания предсказуемости динамического хаоса [5—8], а затем перенесена на поля [9, 10]. Рассмотрим основные черты этого подхода, имеющего непосредственное отношение к интерпретации сложного поведения динамических систем. Сначала опишем три типа процессов, с которыми приходится иметь дело в задачах о предсказуемости.

Пусть  $x(t)$  — *реальный* процесс, происходящий в некоторой физической системе. Динамическое уравнение для многокомпонентной величины  $x(t)$  запишем в символической форме:

$$M(d/dt, x, f) = 0, \quad (1)$$

где  $f(t)$  — разнообразные флуктуационные факторы, как внешние, так и внутренние, относительно которых мы не располагаем детальной информацией и которые мы можем охарактеризовать лишь статистическими характеристиками.

*Наблюдаемый* процесс  $y(t)$  всегда отличается от реального процесса  $x(t)$ , во-первых, потому что наблюдение осуществляется при помощи измерительных приборов, подверженных действию шумов  $v(t)$  (инструментальный или измерительный

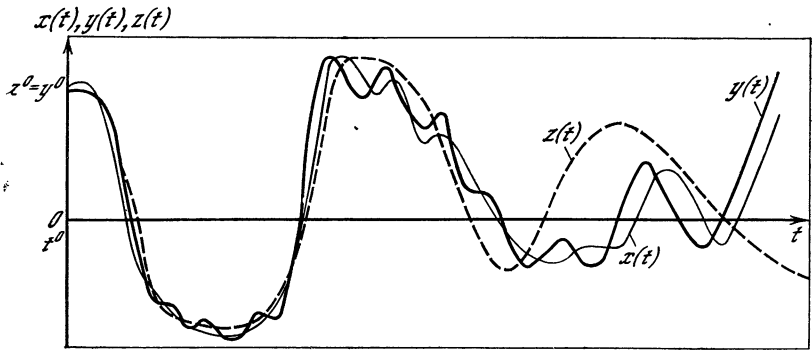


Рис. 1. Три типа процессов, рассматриваемых в задаче о предсказуемости  $x(t)$ ,  $y(t)$ ,  $z(t)$  — соответственно реальный, наблюдаемый и модельный (прогнозируемый) процессы

шум), а во-вторых, потому что приборы не в состоянии зарегистрировать все без исключения компоненты реального процесса  $x(t)$ . В простейшем, но далеко не всегда реализуемом случае, когда шум аддитивен и когда используется линейный, безынерционный прибор с числом входов, совпадающим с числом компонент реального процесса  $x(t)$ , имеем

$$y(t) = x(t) + v(t). \quad (2)$$

Предсказание всегда осуществляется на основе той или иной модели (алгоритма), и поэтому говорить о детерминированности—случайности в смысле предсказуемости—непредсказуемости наблюдаемого процесса  $y(t)$  можно только по отношению к *модельному* процессу (прогнозу)  $z(t)$ . Прогноз опирается на упрощенные, идеализированные представления о реальном процессе  $x(t)$ , и поэтому модельное уравнение для прогноза  $z(t)$  всегда проще, чем уравнение (1) для  $x(t)$ :

$$M_1(d/dt, z, 0) = 0. \quad (3)$$

Модельный оператор  $M_1$ , вообще говоря, отличается от исходного оператора  $M$ ; в частности, он, как правило, имеет меньший порядок относительно  $d/dt$  и игнорирует флуктуационные воздействия  $f(t)$ , полагая их равными нулю. Кроме того, размерность модельного процесса  $z(t)$  может быть меньше размерности  $x(t)$ .

Начальными условиями для  $z(t)$  мы можем распорядиться по своему усмотрению — это тоже один из элементов модельного описания, лежащего в основе любого прогнозирования. Для определенности мы примем здесь простейшие допущения: если измеренное значение наблюдаемого процесса  $y(t)$  в начальный момент времени  $t = t^0$  равно  $y^0 \equiv y(t^0)$ , то мы полагаем  $z^0 \equiv z(t^0) = y^0$ . Более сложный способ задания начальных условий, основанный на предварительной фильтрации и усреднении  $y(t)$ , описан в работе [5], тогда как в работе [11] использован

адаптивный подход. Все три процесса условно представлены на рис. 1, как если бы все они были однокомпонентными. Существенно, что модельный процесс  $z(t)$  сам может быть случайным, скажем, в смысле алгоритмической сложности. Например, это может быть процесс, генерируемый «бесшумовым» странным аттрактором.

### 3. СТЕПЕНЬ ДЕТЕРМИНИРОВАННОСТИ КАК КОЛИЧЕСТВЕННАЯ МЕРА ПРЕДСКАЗУЕМОСТИ

В отличие от стандартного подхода, при котором мерой качества прогноза служит средний квадрат  $\sigma_\eta^2 = \langle \eta^2 \rangle = \langle |y - z|^2 \rangle$  ошибки  $\eta = |y - z|$ , мы предпочитаем использовать иной показатель — степень детерминированности  $D$ , которую определим как нормированный коэффициент корреляции между моделью  $z(t)$  и наблюдением  $y(t)$ :

$$D(\tau) = \langle y(t)z(t) \rangle / \sqrt{\langle y^2(t) \rangle \langle z^2(t) \rangle}, \quad t = t^0 + \tau. \quad (4)$$

Угловые скобки здесь означают статистическое усреднение. При теоретическом рассмотрении их можно трактовать как статистическое усреднение с двумерной условной плотностью вероятности «наблюдение—прогноз»

$$w_2 = w_2(y, z, t | y^0, z^0, t^0), \quad (5)$$

которая в силу принятого условия  $z^0 = y^0$  при  $t \rightarrow t^0$  ведет себя как  $\delta(y - z)w_1(y, t)$ . В итоге

$$\langle y(t)z(t) \rangle_{theor} = \iint yzw_2(y, z, t | y^0, z^0, t^0) dydz. \quad (6)$$

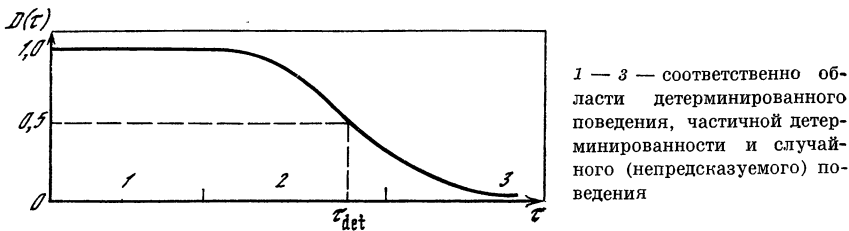
Эмпирический же подход предполагает многократный «запуск» исследуемой системы в моменты  $t_1^0, t_2^0, \dots, t_N^0$  и последующее усреднение по реализациям. Если  $y_j(t) = y(t_j^0 + \tau)$  и  $z_j(t) = z(t_j^0 + \tau; y_j^0, t_j^0)$  —  $j$ -е реализации процессов  $y(t)$  и  $z(t)$ , то в предположении статистической стационарности

$$\langle y(t)z(t) \rangle_{empiric} = \frac{1}{N} \sum_{j=1}^N y(t_j^0 + \tau) z(t_j^0 + \tau; y_j^0, t_j^0). \quad (7)$$

Общий характер зависимости  $D(\tau)$  показан на рис. 2. Равенство степени детерминированности  $D$  единице при  $\tau = 0$  отвечает принятому здесь начальному условию  $z^0 = y^0$ . Степень детерминированности, хотя и выражается через  $\sigma_\eta^2$  ( $D \simeq 1 - \sigma_\eta^2 / 2 \langle y^2 \rangle$ ), но имеет перед  $\sigma_\eta^2$  важное преимущество. Дело в том, что при помощи зависимости  $D(\tau)$  естественно вводится время детерминированного поведения  $\tau_{det}$  — это время, в течение которого степень детерминированности не опускается ниже некоторого уровня. Здесь мы примем этот уровень равным  $1/2$ :  $D(\tau_{det}) = 1/2$ .

Частичная детерминированность наблюдаемого процесса  $y(t)$  по отношению к модельному процессу  $z(t)$  означает в соответствии





1 — 3 — соответственно области детерминированного поведения, частичной детерминированности и случайного (непредсказуемого) поведения

Рис. 2. Типичный вид зависимости степени детерминированности от времени

с рис. 1, что процесс  $y(t)$  предсказуем на малых временах  $\tau \ll \tau_{det}$  (степень детерминированности близка к единице,  $|1 - D| \ll 1$ ), непредсказуем на больших временах (при  $\tau \gg \tau_{det}$  степень детерминированности принимает малые значения,  $|D| \ll 1$ ) и частично предсказуем на временах порядка времени детерминированного поведения (здесь  $0 \lesssim |D| \lesssim 1$ ).

С позиций теории частичной детерминированности термины типа «детерминированный хаос» оправданны лишь на малых временах ( $\tau \ll \tau_{det}$ ). В общем же случае естественно говорить о частично детерминированном хаосе.

#### 4. ВРЕМЯ ДЕТЕРМИНИРОВАННОГО ПОВЕДЕНИЯ ДИНАМИЧЕСКОГО ХАОСА

Это время зависит от многих факторов, которые мы разобьем на три группы: воздействие флуктуационных сил  $f(t)$ , воздействие инструментальных шумов  $v(t)$  и неточность модели, т. е. различие между операторами  $M$  и  $M_1$ . Приведем лишь один пример оценки  $\tau_{det}$  для систем гиперболического типа. Если придерживаться принятого здесь соглашения о задании начальных условий по правилу  $z^0 = y^0$ , то в таких системах «расстояние»  $\eta = |y - z|$  между наблюдением и прогнозом увеличивается по экспоненциальному закону<sup>1</sup>:

$$\langle \eta^2(\tau) \rangle \sim (\sigma_f^2 + \sigma_v^2 + \sigma_{\Delta M}^2) e^{2h\tau},$$

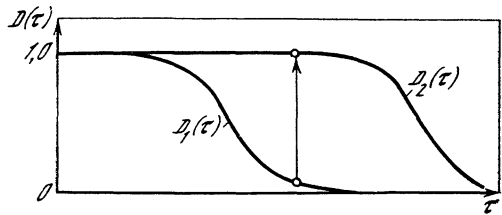
где  $h$  — наибольший ляпуновский показатель, а величины  $\sigma_f^2$ ,  $\sigma_v^2$  и  $\sigma_{\Delta M}^2$  отражают действие указанных выше трех факторов. Учитывая связь между  $D$  и  $\langle \eta^2 \rangle$  (см. разд. 3), из условия  $D(\tau_{det}) = 1/2$  получаем оценку

$$\tau_{det} \approx \frac{1}{2h} \ln \frac{\langle y^2 \rangle}{\sigma_v^2 + \sigma_f^2 + \sigma_{\Delta M}^2}.$$

Из этой оценки следует, что максимальное время детерминированного поведения, достигаемое при предельно точных изме-

<sup>1</sup> При другом соглашении, когда начальные данные постоянно обновляются, ошибка  $|\eta|$  ограничена:  $|\eta| \leq \text{const}$  [11].

Рис. 3. Переход от случайности к детерминированности ( $D_1 \rightarrow D_2$ ) при использовании более удачной модели  $z_2(t)$  вместо  $z_1(t)$



рениях ( $\sigma_v^2 \ll \sigma_f^2$ ) и для удовлетворительной модели ( $\sigma_{\Delta M}^2 \ll \sigma_f^2$ ), определяется лишь флуктуационными воздействиями:

$$\tau_{\text{det max}} \approx \frac{1}{2h} \ln \frac{\langle y^2 \rangle}{\sigma_f^2}.$$

Отсюда следует важный вывод, сделанный в работах [5—8], что случайность, отождествляемая с непредсказуемостью, вызвана в конечном счете действием флуктуационных факторов. Время  $\tau_{\text{det max}}$  естественно отождествить с понятием горизонта предсказуемости, введенным Лайтхиллом в работе [12]. Величина  $\tau_{\text{det max}}$  характеризует время забывания начальных условий, которое в работах [5—8] интерпретируется как время динамической памяти. Значение  $\tau_{\text{det max}}$  сопоставимо также со временем информационной памяти [13—15] и со временем обратимого поведения физических систем.

Говоря о времени детерминированного поведения, важно отметить, что случайность (непредсказуемость) процесса  $y(t)$  по отношению к модели  $z(t)$  не противоречит тому, что тот же самый процесс на том же «расстоянии»  $\tau$  от начала отсчета окажется детерминированным по отношению к другой, более удачной модели  $z_2(t)$ . Соответствующий перевод процесса из случайного в детерминированный показан на рис. 3, где изображены степени детерминированности  $D_1(\tau)$  и  $D_2(\tau)$ , отвечающие моделям  $z_1(t)$  и  $z_2(t)$ . Разумеется, такой перевод возможен только при условии  $\tau < \tau_{\text{det max}}$ .

Таким образом, в рамках теории частичной детерминированности нет смысла говорить об истинной, абсолютной случайности хотя бы потому, что в нашем распоряжении нет количественных критериев такой истинности. Случайность здесь выступает как относительное свойство, проявляющееся при сравнении с моделью. Вообще, не располагая моделью или критерием, невозможно отличить детерминированное от случайного.

В отличие от сложностного подхода, который выносит суждение о случайности только после проведения всех мыслимых тестов на отсутствие закономерностей, концепция частичной детерминированности ограничивается проверкой только некоторых гипотез в рамках возможности, которыми располагает экспериментатор. В итоге деление процессов на случайные и детерминированные в рамках теории частичной детерминированности производится

по схеме проверки гипотез, на которую опираются, в сущности, все естественные науки.

В связи с обсуждаемой здесь относительностью понятия случайности отметим, что исследователь, располагающий мощной ЭВМ, имеет определенные преимущества по переводу процессов из разряда случайных в детерминированные по сравнению с исследователем, лишенным ЭВМ, поскольку компьютер облегчает перебор различных моделей, отличающихся, скажем, значениями нескольких параметров. При таком переборе неточность модели, которую мы вправе классифицировать как незнание, выступает здесь на тех же правах, что и различные флуктуационные факторы: степень детерминированности может возрасти как за счет улучшения модели (уточнение параметров, учет еще одной степени свободы и т. д.), так и за счет снижения флуктуационных воздействий. В итоге граница детерминированность—случайность, характеризуемая временем  $\tau_{\text{det}}$ , оказывается подвижной; она зависит как от уровня шумов, так и от способности исследователя предлагать удачные гипотезы. С этой точки зрения при заданном уровне шумов границу динамической стохастичности следует определить как границу незнания в рамках принятой системы гипотез, а поскольку любая система гипотез не может быть полной (это некоторая модификация утверждения Геделя), всегда существует возможность хотя бы слегка расширить область детерминированного поведения за счет сужения области незнания.

### 5. СООТНОШЕНИЕ МЕЖДУ ЧАСТИЧНОЙ ДЕТЕРМИНИРОВАННОСТЬЮ И ЧАСТИЧНОЙ КОГЕРЕНТНОСТЬЮ

Если мы не располагаем динамической моделью наблюдаемого процесса  $y(t)$  и в качестве прогноза  $z(t)$  берем значение  $y$  в предыдущий момент  $t-\tau$ , т. е.  $z(t) = y(t-\tau)$ , то степень детерминированности  $D(\tau)$  совпадает со степенью когерентности наблюдаемого процесса  $\gamma(\tau)$ :

$$\gamma(\tau) = \langle y(t) y(t-\tau) \rangle / \langle y^2 \rangle, \quad (8)$$

а время детерминированного поведения — со временем когерентности  $\tau$ . Иными словами, степень когерентности характеризует наихудшую степень детерминированности.

Время детерминированного поведения  $\tau_{\text{det}}$  может заметно превышать время когерентности  $\tau_c$ . Различие между  $\tau_{\text{det}}$  и  $\tau_c$  представляется естественным, поскольку эти времена характеризуют разные стороны наблюдаемых явлений: если время  $\tau_c$  связано с шириной спектра процесса:  $\tau_c \sim 1/\Delta\omega$ , то интервал  $\tau_{\text{det}}$  определяется флуктуациями и неточностью модели.

Важно отметить, что возможность детерминированного (модельного) описания физических систем на умеренных интервалах времени ( $\tau \lesssim \tau_{\text{det}}$ ) не противоречит возможности статистического

(вероятностного) описания, которое применимо, грубо говоря, начиная со времен порядка  $\tau_c$ , отвечающих расщеплению корреляций. Поэтому при надлежащем обеспечении экспериментов можно надеяться на предсказание турбулентных потоков, ветрового волнения и других гидродинамических систем на времена, превышающие время корреляции.

## 6. ПРИМЕРЫ

Для иллюстрации общих положений были проделаны расчеты, которые выявили исключительно важную роль шумов и не менее важную роль модели в вопросе о предсказуемости систем, демонстрирующих сложное поведение. На рис. 4 представлены результаты вычисления степени детерминированности  $D$  и коэффициента корреляции  $\gamma$  для системы, которая описывается одномерным кусочно-линейным отображением отрезка  $0,1$  самого в себя [8]:

$$x_n = F(x_{n-1}) + f_{n-1}, \quad F(x) = 2x \pmod{1}, \quad (9)$$

где  $n$  — дискретное время (на рис. 4 дискретное время  $\mu$  отсчитывается от начального момента  $n^0$ , так что  $\mu = n - n^0$ ; эта величина аналогична интервалу  $\tau = t - t^0$ );  $f_n$  — флуктуационное воздействие с равномерной плотностью вероятности в интервале  $-10^{-p} < f < 10^{-p}$ ;  $p$  — показатель, характеризующий силу флуктуаций ( $p = 2 \div 6$ ). В данном случае мы пренебрегли измерительным шумом, положив  $y = x$ , а для модели  $z$  взяли уравнение  $z_n = F(z_{n-1})$ .

Из рис. 4 следует, что при значениях показателя  $p$ , лежащих в интервале между  $p = 2$  и  $p = 6$ , т. е. при уровне шумов от  $10^{-6}$  до  $10^{-2}$ , степень детерминированности  $D$  близка к единице в течение 5 — 10 тактов, а затем довольно быстро спадает до нуля. Время детерминированного поведения оказывается сильно зависящим от уровня шумов.

В то же время коэффициент корреляции  $\gamma$  спадает до уровня  $\sim 1/2$  уже на первом такте, при этом поведение  $\gamma$  обнаруживает слабую чувствительность к уровню шумов.

Эти закономерности проявляются и для других отображений, в том числе для отображения, даваемого логистической кривой. Оказалось, что при переходе к хаосу через фэйгенбаумовские бифуркации удвоения периода степень детерминированности  $D$  претерпевает существенные изменения, свидетельствующие об ухудшении предсказуемости. Однако время детерминированного поведения в рассмотренных нами вместе с Л. Ш. Ильковой задачах всегда было больше, чем время корреляции.

Более реалистичный пример — предсказуемость в автостохастической системе Кияшко, Пиковского и Рабиновича — был рассмотрен в работе [16]. Здесь также было обнаружено, что предсказуемость простирается на времена, превышающие время корреляции, и что введение дополнительного шума почти не

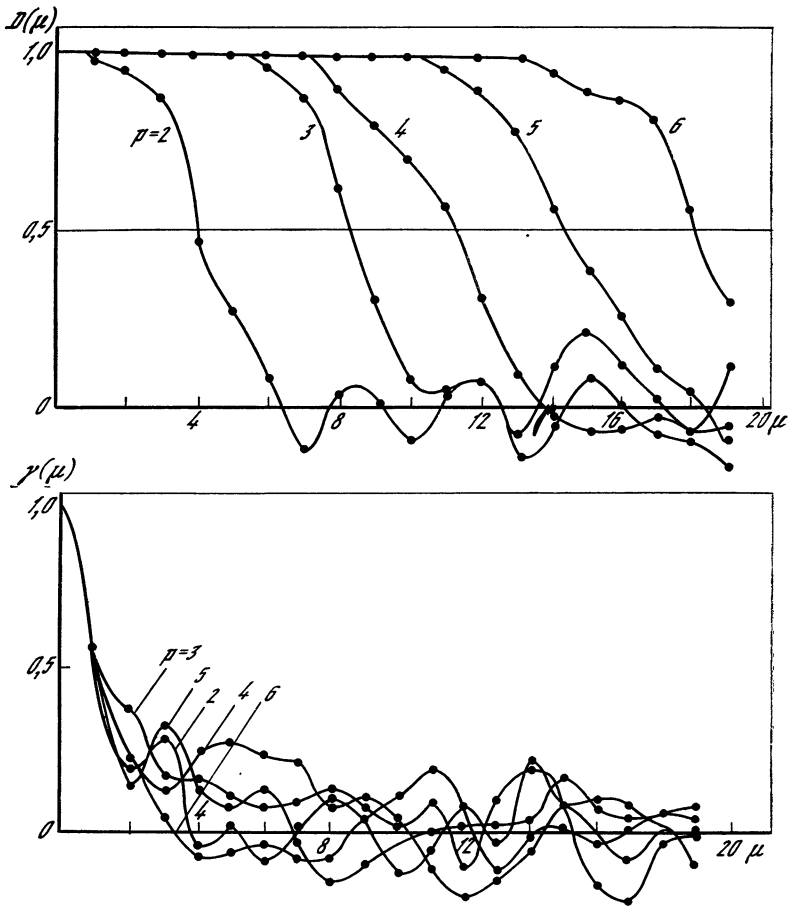


Рис. 4. Зависимость степени детерминированности  $D$  и степени когерентности  $\gamma$  от дискретного времени  $\mu$  для отображения (9) при наличии шума. Параметр  $p$  характеризует интенсивность шума  $I \approx 10^{-p}$

отражается на коэффициенте корреляции, но существенно влияет на степень детерминированности.

В отношении систем, проходящих через точку бифуркации, можно предвидеть одну интересную особенность, заключающуюся в следующем. При медленном изменении бифуркационного параметра можно ожидать, что степень детерминированности резко уменьшится ровно в 2 раза, тогда как при относительно быстром прохождении бифуркационной точки степень детерминированности сохранится на уровне, близком к единице. Переход от одного типа поведения к другому существенно зависит от уровня шумов.

## 7. ЧАСТИЧНО ДЕТЕРМИНИРОВАННЫЕ ПОЛЯ

Представление о частично детерминированных процессах естественным образом обобщается на поля, в том числе на волновые поля [9]. В этом случае наблюдаемое поле  $Y(\mathbf{r}, t)$  следует сопоставить с модельным полем  $Z(\mathbf{r}, t)$  в некоторой пространственно-временной области. При таком подходе становится ясным, что при определенных условиях можно добиться удовлетворительной предсказуемости, например спекл-неоднородных полей, характеризующихся коротким радиусом корреляции. Идея предсказуемости довольно проста: по экспериментальным данным, полученным, скажем, на отрезке  $-l < x < l$ , можно восстановить источники поля, а по ним уже легко экстраполировать поле вне отрезка  $-l < x < l$ . Важно, что расстояние предсказуемого поведения может заметно превысить радиус корреляции спекл-неоднородного поля [10].

Численные расчеты, касающиеся хаотического поведения распределенных систем, нами не проводились. Мы убеждены, однако, что при надлежащем обеспечении экспериментов (самое сложное здесь — получить надежные начальные данные) можно надеяться на предсказание турбулентных гидродинамических потоков, ветрового волнения и других гидродинамических систем на времена, превышающие время корреляции.

В этой связи отметим, что возможность детерминированного (модельного) описания физических систем на умеренных интервалах времени ( $\tau \leq \tau_{\text{det}}$ ) и в ограниченных областях пространства ( $l < l_{\text{det}}$ ) не противоречит возможности статистического (кинетического) описания, которое применимо, грубо говоря, начиная со времен порядка  $\tau_c$ , отвечающих расцеплению корреляций. Иными словами, кинетический и динамический методы описания могут сосуществовать в интервале  $\tau_c < \tau < \tau_{\text{det}}$ . Мы имеем здесь дело просто с разными сторонами одного и того же явления: если кинетический подход ориентируется на получение средних характеристик, то модельное описание дает предсказание мгновенных (локальных) значений поля.

## 8. МОДЕЛЬНЫЕ ВЕРОЯТНОСТНЫЕ ПРЕДСТАВЛЕНИЯ ПРИ ОБРАБОТКЕ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

На наш взгляд, вероятности, извлекаемые из эксперимента на базе частотного подхода или же задаваемые на основе теоретических соображений, всегда носят модельный, гипотетический характер. Ценность этих моделей определяется тем, насколько хорошо они описывают свойства реальных систем на коротких, умеренных и длинных интервалах времени. Геделевские соображения о неполноте любой системы гипотез и в этом случае заставляют нас с большой осторожностью относиться к экстраполяции статистических характеристик далеко вперед.

Здесь мы хотели бы отметить, что даже саму вероятность можно рассматривать как степень детерминированности по от-

ношению к той или иной модели. Для этого знак усреднения  $\langle \cdot \rangle$  в (4) следует переопределить как среднее значение суммы индикаторных функций. Обозначив такое среднее фигурными скобками, получим

$$\langle yz \rangle \rightarrow \{yz\} = \frac{1}{N} \sum_{i=1}^N \chi_{\varepsilon}(y_i, z_i), \quad (10)$$

где  $\chi_{\varepsilon}(y, z) = 1$  при  $|y - z| \leq \varepsilon/2$  и  $\chi_{\varepsilon}(y, z) = 0$  при  $|y - z| > \varepsilon/2$  ( $\varepsilon$  — интервал дискретизации функций  $y$  и  $z$ ). Учитывая, что  $\chi_{\varepsilon}(y_i, y_i) = \chi_{\varepsilon}(z_i, z_i) = 1$  и, как следствие,  $\{y^2\} = \{z^2\} = 1$ , имеем

$$D_y(z_i) = \frac{1}{N} \sum_{i=1}^N \chi_{\varepsilon}(y_i, z_i) = \frac{n_{y,z}(\varepsilon)}{N}, \quad (11)$$

где  $n_{y,z}(\varepsilon)$  — число совпадений между  $y$  и  $z$  при дискретности разбиения  $\varepsilon$ .

Если зафиксировать уровень  $z = z_0 = \text{const}$ , то степень детерминированности (11) дает нам обычную вероятность попадания  $y(t)$  в интервал  $(z_0 - \varepsilon/2, z_0 + \varepsilon/2)$ :

$$P(z_0 - \varepsilon/2 < y < z_0 + \varepsilon/2) \cong n_{y,z_0}(\varepsilon)/N = D_y(z_0).$$

В то же время в условиях, когда все члены модельной последовательности  $z_i$  совпадают с  $y_i$ , мы имеем  $D_y(z_i) = 1$ . В этом последнем случае можно говорить о том, что экспериментатору удалось построить удовлетворительную модель  $z_i$  для наблюдаемой последовательности  $y_i$ . Обычный вероятностный подсчет дал бы для такого совпадения очень малую вероятность, а алгоритмическая теория вероятности квалифицировала бы совпадение  $y_i$  и  $z_i$  как событие с конструктивно нулевой мерой [2].

## ЗАКЛЮЧЕНИЕ

Анализ хаотического поведения динамических систем с позиций теории частично детерминированных процессов, которая основана на отождествлении случайности с непредсказуемостью, позволяет сделать следующие выводы:

1. В рамках представлений о частичной детерминированности динамический хаос следует классифицировать как частично-детерминированный процесс с ограниченным временем детерминированного (предсказуемого) поведения  $\tau_{\text{дет}}$ . С этой точки зрения термин «детерминированный хаос» правильно было бы заменить на «частично-детерминированный» хаос.

2. Степень детерминированности  $D(\tau)$  существенно зависит от уровня шумов в системе и от выбранной модели, тогда как степень когерентности (коэффициент автокорреляции) от уровня шумов практически не зависит. Время детерминированного поведения может быть увеличено как за счет снижения шумов,

так и за счет улучшения предсказательной модели, или, иными словами, за счет уменьшения «шумов незнания».

3. Степень когерентности (коэффициент автокорреляции) динамического хаоса выступает как наихудшая оценка степени детерминированности, реализуемая при условии, что в качестве модели используется значение наблюдаемого процесса в предшествующий момент времени. Вследствие этого время детерминированного поведения, как правило, заметно превышает время автокорреляции:  $\tau_{\text{det}} < \tau_c$ .

4. Детерминированное описание динамических систем на ограниченных интервалах времени  $\tau < \tau_{\text{det}}$  не противоречит возможности статистического (кинетического) описания. Это позволяет говорить о сосуществовании детерминированного и кинетического описаний в интервале  $\tau_c < \tau < \tau_{\text{det}}$  и о принципиальной возможности предсказания поведения ряда систем, которые традиционно считаются стохастическими.

#### ЛИТЕРАТУРА

1. Колмогоров А. Н. Три подхода к количественному определению информации // Проблемы передачи информации. 1965. Т. 1. С. 1—22.— То же // Колмогоров А. Н. Теория вероятностей и математическая статистика. М.: Наука, 1986. С. 467—471.
2. Шень А. А. Частотный подход к определению случайной последовательности // Семиотика и информатика. М.: ВИНТИ, 1982. № 18. С. 14—42.
3. Martin-Löf P. The definition of random sequences // Inform. Control. 1966. Vol. 9, N 7. P. 602.
4. Звонкин А. К., Левин Л. А. Сложность конечных объектов и развитие концепции информации и случайности при помощи теории алгоритмов // УМН. 1970. Т. 25, № 1. С. 85—101.
5. Кравцов Ю. А., Эткин В. С. К вопросу о роли флуктуационных сил в динамике автостохастических систем, ограничение времени предсказуемости и разрушение слабых периодических режимов // Изв. вузов. Радиофизика. 1981. Т. 24, № 8. С. 992—999.
6. Kravtsov Yu. A., Etkin V. S. Dynamic correlation and predictability in stochastic and dynamic systems in the presence of fluctuations // Nonlinear and turbulent processes in Physics / Ed. R. Z. Sagdeev. N. Y.: Harwood Acad. Publ. 1984. Vol. 3. P. 1411—1417.
7. Кравцов Ю. А., Эткин В. С. Степень динамической корреляции и проблема выявления динамической природы случайных процессов // РЭ. 1984. Т. 29, № 12. С. 2358—2364.
8. Илькова Л. Ш., Кравцов Ю. А., Мергелян О. С., Эткин В. С. Мера частичной детерминированности динамического хаоса // Изв. вузов. Радиофизика. 1985. Т. 28, № 7. С. 929—932.
9. Кравцов Ю. А., Петников В. Г. Частично детерминированные волновые поля // ДАН СССР. 1985. Т. 285, № 4. С. 871—873.
10. Kravtsov Yu. A., Petnikov V. G. Partially deterministic wave-fields in optics // J. Opt. Soc. Amer. A. 1987. Vol. 4, N 1. P. 11—16.
11. Вул Е. Б., Синай Я. Г., Ханин К. М. Универсальность Фейгенбаума и термодинамический формализм // УМН. 1984. Т. 39, № 3. С. 1—39.
12. Lighthill J. The recently recognized failure of predictability in Newtonian dynamics // Proc. Roy. Soc. A. 1986. Vol. 407, N 1832. P. 35—50.
13. Shaw R. Strange attractors, chaotic behavior and information flow // Ztschr. Naturforsch. A. 1981. Bd. 36. S. 80—102.



14. *Crutchfield J. P., Packard N. H.* Symbolic dynamics of one-dimensional maps: Entropies, finite precision and noise // Intern. J. Theor. Phys. 1982. Vol. 21. P. 433—466.
15. *Crutchfield J. P., Packard N. H.* Symbolic dynamics of noisy chaos // Physica. D. 1983. Vol. 7, N 2. P. 202—223.
16. *Илькова Л. Ш., Кравцов Ю. А., Пиковский А. С.* Определение степени детерминированности экспериментально наблюдаемых хаотических режимов // Крат. сообщ. ФИАН по физике. 1986. № 11. С. 7—9.

УДК 539.194+539.191

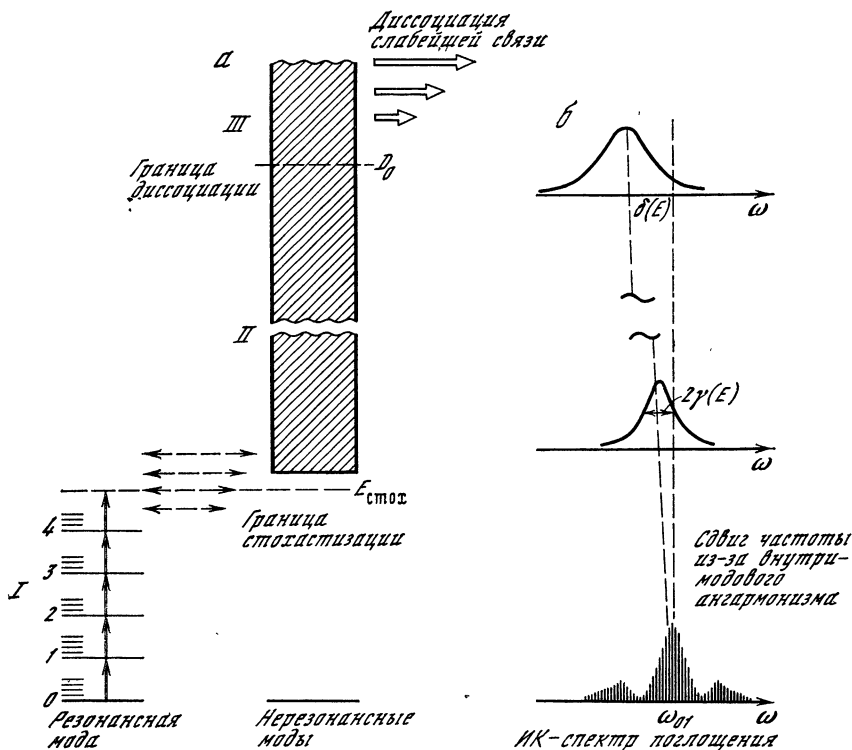
## СТОХАСТИЧЕСКАЯ КОЛЕБАТЕЛЬНАЯ ДИНАМИКА МНОГОАТОМНЫХ МОЛЕКУЛ

В. С. ЛЕТОХОВ, А. А. МАКАРОВ, Е. А. РЯБОВ

### ВВЕДЕНИЕ

С точки зрения классической механики многоатомная молекула — это система осцилляторов, связанных относительно слабым нелинейным взаимодействием. Общим свойством таких классических систем является стохастичность движения при достаточно большой колебательной энергии [1—4]. Хотя многоатомная молекула, строго говоря, является квантовым объектом, все же нам представляется, что те реально наблюдаемые свойства высоковозбужденных колебательных состояний многоатомных молекул, о которых пойдет речь, на практике неотличимы от свойств стохастического колебательного движения в классическом пределе (см. также работу [5]).

Стохастизация колебательного движения очень важна для понимания многофотонного возбуждения (МФВ) и диссоциации (МФД) многоатомных молекул под действием интенсивного ИК лазерного излучения, частота которого настроена в резонанс с какой-либо активной в ИК-поглощении колебательной модой (см. [6]). На рис. 1, *a* упрощенно показана модель ИК МФВ и МФД многоатомной молекулы. Первые несколько ИК-фотонов молекула поглощает на квантовых переходах между нижними дискретными колебательно-вращательными уровнями (область *I* на рис. 1, *a*). Внутримодовый ангармонизм резонансной колебательной моды в этой области дает красный сдвиг спектра ИК-полосы поглощения (рис. 1, *b*). Взаимодействие ИК-поля с колебаниями молекулы имеет резонансный или мод-селективный характер. По мере роста колебательной энергии молекулы становится существенным межмодовый ангармонизм, который приводит к взаимодействию резонансной моды с остальными нерезонансными по отношению к ИК-излучению модами. Для колебательно-возбужденной молекулы возникает много межмодовых резонансов или резонансов Ферми, перекрытие которых в соответствии с критерием Чирикова [1, 5] приводит к стохастизации колебательного движения (область *II* на рис. 1, *a*).



**Рис. 1.** Модель ИК многофотонного возбуждения многоатомной молекулы. Схематически показаны изменения структуры колебательных состояний (а) и эволюция спектра полосы ИК-поглощения с ростом энергии молекулы (б).

а: I — область резонансного мод-селективного взаимодействия с дискретными уровнями резонансной моды; II — область стохастичности колебательного движения, в которой ИК-возбуждение не является мод-неселективным; III — область нестационарного колебательного перевозбуждения

В области стохастичности радикально меняется форма полосы поглощения молекулы в окрестности частоты резонансной моды (рис. 1, б). Из-за межмодового взаимодействия образуется одно-одно-уширенная полоса поглощения, ширина  $2\gamma$  которой зависит от скорости внутримолекулярного перераспределения колебательной энергии молекулы, а центральная частота смещена в длинноволновую область на величину  $\delta(E)$ , зависящую от среднего межмодового ангармонизма. И ширина  $2\gamma$ , и сдвиг  $\delta(E)$  колебательной полосы растут по мере увеличения колебательной энергии, но, как правило, сдвиг не превышает заметно ширины. Это связано с тем, что далеко не все колебательные моды молекулы дают вклад в межмодовый ангармонизм. Именно такое соотношение между шириной и сдвигом ИК фундаментальной полосы многоатомной молекулы обеспечивает возможность

поглощения большого числа ИК-фотонов вплоть до границы диссоциации слабейшей связи молекулы и даже выше нее (область III на рис. 1, а).

Явление стохастизации колебательного движения в возбужденной многоатомной молекуле в последние годы стало предметом экспериментальных и теоретических исследований, результаты которых кратко представлены ниже.

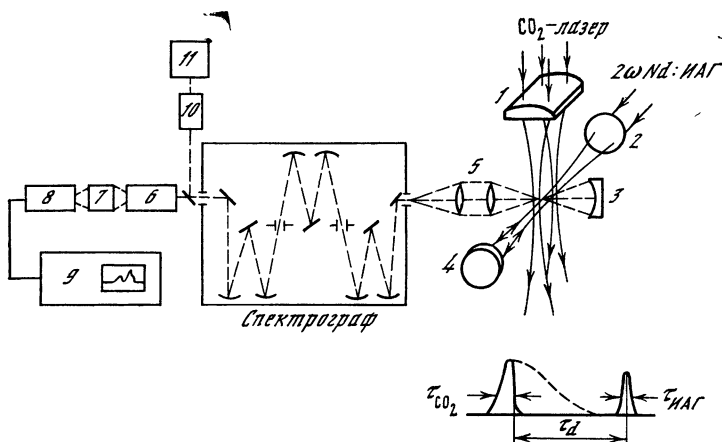
## 1. ЭКСПЕРИМЕНТ И ЕГО РЕЗУЛЬТАТЫ

Для прямых экспериментов по спектроскопическому зондированию колебательных распределений при возбуждении молекул импульсом  $\text{CO}_2$ -лазера наиболее адекватен метод спектроскопии антистоксова спонтанного комбинационного рассеяния (КР). Легко показать, что в гармоническом приближении интегральный сигнал антистоксова КР в колебательной полосе пропорционален энергии в соответствующей моде и что для этой величины различные эффекты ангармонизма носят характер лишь малых несущественных поправок. Уже в первом эксперименте [7] с молекулами  $\text{SF}_6$  и  $\text{CF}_3\text{I}$  были зарегистрированы сигналы от мод, нерезонансных по отношению к возбуждающему ИК-излучению.

В дальнейшем эксперимент был усовершенствован с целью получения более полной количественной информации. Схема показана на рис. 2. Импульс  $\text{CO}_2$ -лазера с длительностью 20 нс возбуждает исследуемые молекулы. Импульс второй гармоники Nd : ИАГ-лазера с длительностью 10 нс осуществляет зондирование. Задержка зондирующего импульса относительно возбуждающего регулируется с точностью 10 нс. Давление газа соответствует отсутствию столкновений молекул в газе от начала возбуждения до окончания зондирования. Рассеянный свет проходит через тройной спектрограф и регистрируется в режиме счета фотонов. Детали эксперимента наиболее полно описаны в публикациях [8, 9].

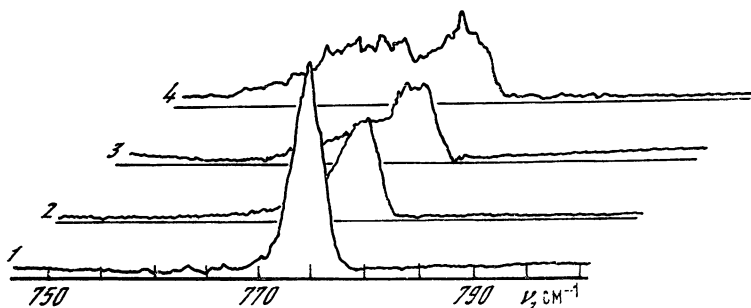
В первую очередь измеряются интегральные сигналы антистоксова КР во всех разрешенных и достаточно сильных колебательных полосах. Спектрографом вырезается существенный участок спектра, и измерения проводятся с помощью ФЭУ. Как было отмечено выше, интегральные сигналы прямо дают значения энергий в модах. Главное качественное наблюдение состоит в том, что сигнал присутствует во всех полосах. Это уже само по себе говорит о возбуждении всех колебательных мод молекулы (результаты нашей ранней работы [10] были позднее воспроизведены в работе [11] группы ученых из США, где был сделан тот же качественный вывод). Естественное объяснение — наличие внутримолекулярных механизмов перераспределения колебательной энергии.

Однако при каких уровнях возбуждения по полной колебательной энергии молекул работают эти механизмы? Знание только энергий в модах еще не позволяет этот уровень возбуждения оце-



**Рис. 2.** Схема эксперимента по зондированию методом спектроскопии КР колебательного распределения, создаваемого в результате ИК многофотонного возбуждения

1 — 5 — линзы и зеркала, 6 — усилитель изображения, 7 — объектив, 8 — многоканальная система регистрации SIT-500, 9 — система обработки спектров WP-4, 10 — фотоэлектронный умножитель, 11 — счетчик импульсов



**Рис. 3.** Спектры стоксова КР на частоте полносимметричной моды  $\nu_1$  при возбуждении молекулы  $\text{SF}_6$  импульсами  $\text{CO}_2$ -лазера, резонансными трехкратно-вырожденной моды  $\nu_3$  [8]

Плотность энергии  $\Phi$ , Дж/см<sup>2</sup>, равна: 1 — 0, 2 — 0,02, 3 — 0,036, 4 — 0,07

нить даже грубо. Действительно, в принципе возможны два крайних случая: а) в результате возбуждения все молекулы имеют приблизительно одинаковую полную колебательную энергию; б) в результате возбуждения малая (неизвестная) доля молекул  $q$  приобретает значительно большую энергию, а остальные молекулы,  $1-q$ , остаются на нижних колебательных уровнях. Ранее проведенные эксперименты указывали на то, что скорее реализуется второй случай [12, 13]. Измерения стоксовых спектров КР, проводимые в наших экспериментах с помощью многоканальной системы регистрации, позволяют это легко выявить. На рис. 3 отчетливо видны после возбуждающего импульса две

группы молекул (из-за ангармонического сдвига они дают сигналы на заметно отличающихся частотах). В ряде случаев обработка спектров позволяет определить величину  $q$ , а значит, и уровень возбуждения  $\varepsilon_q$  верхней группы молекул достаточно точно.

Другим способом измерения величин  $q$  и  $\varepsilon_q$  при известных значениях энергий в модах является измерение сигналов антистоксова КР в обертоновых и составных полосах молекул. Эти сигналы квадратичны по энергиям в модах в верхней группе молекул [9, 14] и поэтому дают недостающие данные. В ряде случаев удается найти полосы, измерения в которых хотя и трудоемки, но реализуемы.

Для каждой из исследованных молекул  $\text{SF}_6$ ,  $\text{CF}_3\text{I}$ ,  $\text{CF}_2\text{HCl}$ ,  $\text{CF}_3\text{Br}$ ,  $\text{CF}_2\text{Cl}_2$  проводился комплекс доступных измерений. Для различных плотностей энергии  $\Phi$  импульса  $\text{CO}_2$ -лазера определялись характеристики создаваемых распределений. В результате воссоздается следующая довольно простая, общая для всех исследованных молекул картина (рис. 4). Часть молекул, которую выше мы обозначили через  $q$ , возбуждается относительно высоко. Энергия в этой группе распределена по модам статистически-равновесно. Равновесность означает, что для молекул с  $\varepsilon_q > E_{\text{гр}}$  энергии  $\varepsilon_{iq}$  в модах  $\nu_i$  хорошо описываются одной эффективной температурой  $T_{\varepsilon\Phi}$ :

$$\varepsilon_{iq} = \hbar\nu_i [\exp(\hbar\nu_i/kT_{\varepsilon\Phi}) - 1]^{-1}, \quad \varepsilon_q = \sum_i \varepsilon_{iq}. \quad (1)$$

Часть молекул находится на низких возбужденных уровнях резонансной моды (как правило, в резонансной моде по сравнению с остальными имеется заметный избыток энергии). Часть молекул остается в исходных состояниях. С ростом  $\Phi$  растет как доля молекул  $q$  в верхней равновесной группе, так и их уровень возбуждения  $\varepsilon_q$ . Однако при уменьшении  $\Phi$  величина  $\varepsilon_q$  стремится к некоему постоянному значению  $E_{\text{гр}}$  (рис. 5). Никаких изменений сигналов на бесстолкновительной шкале времени не происходит. Поэтому если говорить о времени перераспределения энергии по модам в верхней группе молекул, то наши эксперименты могут дать для него лишь оценку сверху ( $\lesssim 10$  нс.)

Ниже приведены значения энергетических границ стохастизации колебательного движения для различных молекул, измеренные в экспериментах по ИК МФВ с помощью спектроскопии КР:

Молекула	$\text{SF}_6$	$\text{CF}_3\text{I}$	$\text{CF}_2\text{HCl}$	$\text{CF}_3\text{Br}$	$\text{CF}_2\text{Cl}_2$
$E_{\text{гр}} = E_{\text{стох}}, \text{ см}^{-1}$	5000	6000	$\lesssim 5000$	7500	$\lesssim 7800$

То, что в двух случаях дана лишь оценка сверху, связано с недостаточной чувствительностью, чтобы уверенно измерить выход на постоянное значение. Однако существенное уменьшение наклона зависимости  $\varepsilon_q(\Phi)$  при указанных значениях уже наблюдалось.

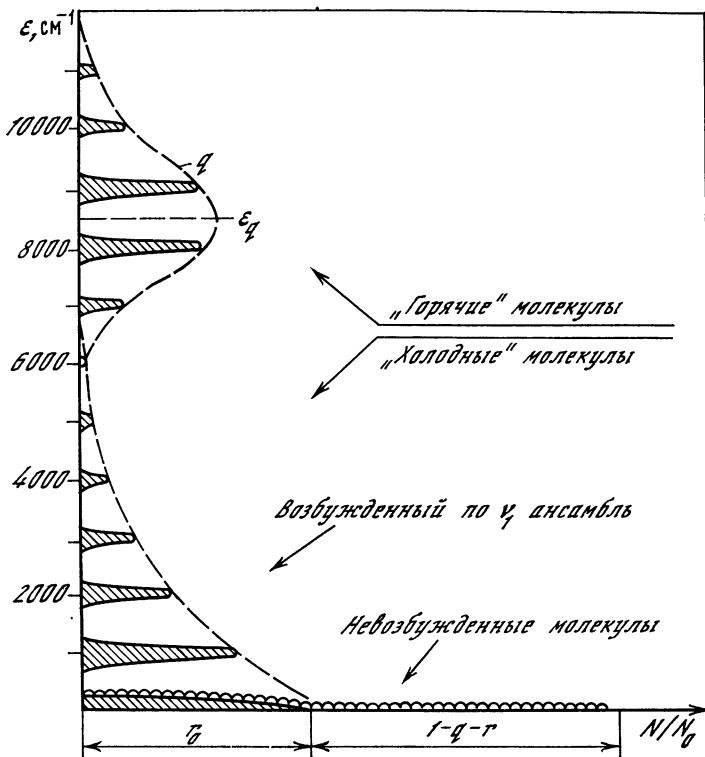


Рис. 4. Качественная картина колебательного распределения в  $\text{CF}_3\text{Br}$  в результате ИК многофотонного возбуждения

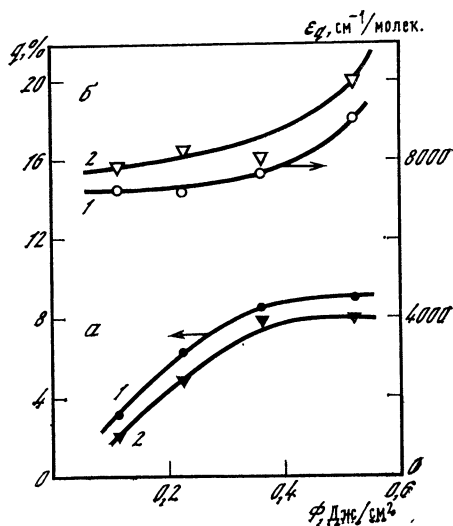


Рис. 5. Зависимости доли  $q$  молекул  $\text{CF}_3\text{Br}$  в верхнем ансамбле (а) и их средней энергии  $\epsilon_q$  (б) от плотности энергии  $\Phi$  импульса  $\text{CO}_2$ -лазера [9]

Приведены два варианта определения  $q$  и  $\epsilon_q$  из экспериментальных данных: 1 — из данных по сигналам КР в полосах  $\nu_3$  и  $2\nu_3$ ; 2 — из данных по  $\nu_2$ ,  $\nu_3$  и  $\nu_2 + \nu_3$

Тот факт, что ниже  $E_{гр}$  молекулы с равновесным распределением полной колебательной энергии по модам просто не наблюдаются (!), говорит в пользу того, что  $E_{гр}$  представляет собой физическую границу в молекуле, разделяющую области с существенно различной структурой квантовых состояний:  $E_{гр} \approx E_{стох}$ .

## 2. ТЕОРИЯ: КОНЦЕПЦИИ И МОДЕЛИ

При создании теории мономолекулярного распада многоатомных молекул выше границы диссоциации было постулировано (см., например, монографии [15, 16]) наличие механизмов, обеспечивающих быстрый обмен энергией между колебательными модами, так что при усреднении во времени энергия распределена по модам статистически равновесно. Это неизбежно должно означать, что равновесность энергий в модах является свойством истинных собственных состояний  $|g\rangle$  колебательного гамильтониана молекулы. Только тогда временная эволюция первоначально созданного преимущественного возбуждения одной из мод приведет к равновесию. Эксперименты, о которых было рассказано выше, вероятно, проводят границу между областью колебательных энергий, где прекрасно «работает» квазигармоническое приближение, составляющее основу молекулярной спектроскопии, и областью колебательных энергий, где «работает» статистический подход.

Статистическое описание предполагает малость относительных флуктуаций. В ансамбле близких гармонических состояний  $|h\rangle$  флуктуаций энергий в модах порядка самих энергий (справедлива больцмановская статистика). Поэтому истинные состояния  $|g\rangle$ , которые бы обладали малыми флуктуациями, могут быть образованы только как суперпозиции большого числа гармонических состояний. (Такие состояния иногда называют глобальными [17, 18].) Главный механизм перемешивания гармонических состояний молекулы — межмодовый ангармонизм. Возьмем какое-нибудь гармоническое состояние  $|h_1\rangle$  с энергией  $E_1$ . Оно связано со всеми близкими гармоническими состояниями  $|h_n\rangle$  той же колебательной симметрии недиагональными матричными элементами ангармонического взаимодействия  $V_{1n}$ . Если величины  $V_{1n}$  малы по сравнению с дефектами энергий  $E_1 - E_n$ , то состояние  $|h_1\rangle$  мало возмущено. В противоположном случае, если найдется такое состояние  $|h_2\rangle$ , для которого выполнено условие

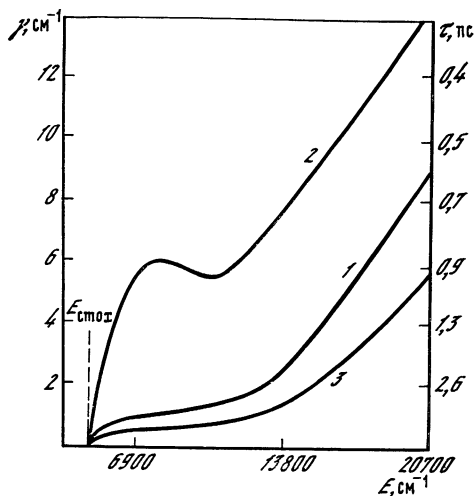
$$|V_{12}| \gtrsim |E_1 - E_2|, \quad (2)$$

то состояния  $|h_1\rangle$  и  $|h_2\rangle$  окажутся сильно перемешанными. Если же в начальный момент времени молекула находится в состоянии  $|h_1\rangle$ , то в дальнейшем она будет осциллировать между состояниями  $|h_1\rangle$  и  $|h_2\rangle$ .

Дефекты энергий  $E_1 - E_n$  задаются всеми возможными относительно малыми алгебраическими суммами нормальных частот,

**Рис. 6.** Зависимость ширины переходов  $\gamma$  и времени внутримолекулярной релаксации  $\tau$  от энергии молекулы [19]

Расчет для мод  $\nu_1$ ,  $\nu_3$  и  $\nu_5$  (кривые 1, 2, 3 соответственно) модельной молекулы  $\text{CFCl}_2\text{Br}$ . Константы междоуровневого ангармонизма третьего порядка  $X_{ijk} = 2,4 \text{ см}^{-1}$ . Расчетное значение  $E_{\text{стох}} = 4629 \text{ см}^{-1}$



такими, как  $\nu_i + \nu_j - \nu_k$  (резонансы третьего порядка),  $\nu_i + \nu_j + \nu_k - \nu_l$  и  $\nu_i + \nu_j - \nu_k - \nu_l$  (резонансы четвертого порядка) и т. д. Матричные элементы взаимодействия  $V_{ln}$  определяются соответствующими константами в разложении потенциальной энергии по нормальным координатам (в частности, при членах  $q_i q_j q_k$  для резонансов третьего порядка) и колебательными квантовыми числами участвующих в резонансах мод в состояниях  $|h_1\rangle$  и  $|h_n\rangle$ . Значения матричных элементов  $V_{ln}$  увеличиваются с ростом колебательных квантовых чисел, следовательно, в среднем увеличиваются с ростом энергии.

Далее, пусть выполнение условия (2) — это не простая случайность для каких-то отдельных пар уровней, а закономерность, и оно обеспечивается не одним резонансом, а несколькими, включающими в себя все моды молекулы (с ростом энергии такая ситуация неизбежно должна возникнуть). В классической механике это не что иное, как критерий стохастичности Чирикова. В нашей же схеме рассуждений на языке квантовой механики это критерий того, что цепочка резонансов перемешает множество близких гармонических состояний, отвечающих самым различным возможным для данной полной энергии наборам колебательных квантовых чисел. Как следствие этого: а) в динамике энергии в модах должны стремиться к равновесию, б) истинные стационарные состояния должны обладать теми статистическими свойствами, которые мы обсуждали выше.

Теоретические расчеты, которые могли бы подтвердить полученные в эксперименте значения  $E_{\text{гр}} = E_{\text{стох}}$  для конкретных молекул, пока вряд ли осуществимы в силу чрезвычайно большой трудоемкости численного решения уравнения Шредингера для систем многих связанных осцилляторов и отсутствия существенных данных об ангармонических константах. Поэтому представляют интерес модельные подходы. Так, в работе [19] обна-



ружена резкая граница между двумя областями с качественно отличной структурой состояний, что находится в согласии с одним из наиболее важных экспериментальных наблюдений. В этой работе каждая мода молекулы описывается зависящими от энергии частотой и шириной. Ширина в духе идей квантовой кинетики отвечает скорости релаксации энергии в моде к равновесию. Исходя из данной структуры межмодовых резонансов с помощью температурной техники квантовых функций Грина была получена самосогласованная система уравнений для нахождения ширины. Было показано, что существует некая граничная энергия, ниже которой для ширины существует только нулевое решение, а выше существуют два решения — одно из них (нулевое) неустойчиво, а устойчивое решение соответствует ширинам, отличным от нуля и растущим с ростом энергии (рис. 6).

По-видимому, точные результаты в будущем должны продемонстрировать некую условность границы между двумя областями колебательной энергии. Однако имеющиеся на сегодня экспериментальные и теоретические результаты указывают на то, что область перехода довольно узкая.

## ЗАКЛЮЧЕНИЕ

В экспериментах, описанных выше, исследовалось колебательное распределение, создаваемое в результате ввода энергии в многоатомную молекулу через одну из ее нормальных мод. В этих экспериментах отчетливо видна энергетическая граница  $E_{гр}$ , при переходе через которую наблюдается резкое качественное изменение распределения энергии по модам. Ниже этой границы поглощенная энергия концентрируется в резонансной моде. Если полная колебательная энергия превышает  $E_{гр}$ , то наблюдается возбуждение не только резонансной, но и всех остальных мод. Более того, вся совокупность экспериментальных результатов хорошо количественно объясняется в предположении, что полная энергия при превышении над границей распределена по модам статистически равномерно.

Для детального понимания механизмов, приводящих к перераспределению колебательной энергии в каждом конкретном случае, первостепенный интерес представляют времена процесса. Временная шкала наших экспериментов ( $\sim 10$  нс) оказалась слишком длинной для прямого наблюдения динамики перераспределения. Поэтому следующим шагом должны быть эксперименты с более короткими лазерными импульсами. Пока колебательная динамика прямо наблюдалась только в электронно-возбужденных состояниях сложных полициклических молекул по эволюции спектров видимой флуоресценции (см. обзор [20]), однако энергии в модах количественно не измерялись.

Для изучения динамики возможен также спектроскопический подход. Времена перераспределения энергии между модами многоатомной молекулы должны проявляться в уширении спектров

колебательных полос (см., например, [21, 22]). Экспериментально уже опробована как спектроскопия обертонов высокочастотных колебаний [23, 24], так и спектроскопия фундаментальных полос молекул, приготовленных в высоковозбужденных колебательных состояниях [25, 26]. Однако следует отметить, что трактовка спектроскопических измерений такого рода не является тривиальной задачей. В общем случае в уширение полосы дают вклад не только релаксация энергии в моде, но и эффект чисто фазовой внутримолекулярной релаксации [27, 28], а также неоднородность спектра. Вероятно, в дальнейшем для устранения неоднородного уширения, кроме техники охлажденных сверхзвуковых струй, будут также использоваться некоторые стандартные методы лазерной спектроскопии, например фотонное эхо.

Совсем по-другому внутримолекулярная релаксация сказывается на линиях в микроволновом спектре, т. е. для чисто вращательных переходов. Взамен привычного релаксационного уширения линий теория здесь предсказывает, что быстрое перераспределение энергии между колебательными модами должно приводить к сужению неоднородной структуры переходов из ансамбля близко расположенных уровней [29]. Возможно, этот эффект также интересно использовать для определения  $E_{\text{стох}}$ .

#### ЛИТЕРАТУРА

1. Заславский Г. М., Чириков Б. В. Стохастическая неустойчивость нелинейных колебаний // УФН. 1975. Т. 105, № 1. С. 3—39.
2. Балеску Р. Равновесная и неравновесная статистическая механика. М.: Мир, 1978. Т. 2.
3. Заславский Г. М. Стохастичность динамических систем. М.: Наука, 1984. 271 с.
4. Лихтенберг А., Либерман М. Регулярная и стохастическая динамика. М.: Мир, 1984. 528 с.
5. Шуряк Э. В. Нелинейный резонанс в квантовых системах // ЖЭТФ. 1976. Т. 71, № 6. С. 2039—2056.
6. Bagratashvili V. N., Letokhov V. S., Makarov A. A., Ryabov E. A. Multiple photon infrared laser photophysics and photochemistry. Chur etc.: Harwood Acad. Publ., 1985. 512 p.
7. Баграташвили В. Н., Вайнер Ю. Г., Должииков В. С. и др. Прямое наблюдение методом спектроскопии КРС эффекта стохастизации колебательной энергии в молекулах при взаимодействии с сильным лазерным ИК полем // Письма в ЖЭТФ. 1979. Т. 30, № 8. С. 502—506.
8. Doljikov V. S., Doljikov Yu. S., Letokhov V. S. et al. Intramolecular and intermolecular redistribution of vibrational energy in MP IR excitation:  $\text{CF}_2\text{HCl}$  molecule // Chem. Phys. 1986. Vol. 102, N 1/2. P. 155—163.
9. Должииков Ю. С., Летохов В. С., Макаров А. А. и др. Меж- и внутримолекулярное распределение колебательной энергии при ИК многофотонном возбуждении. Молекула  $\text{CF}_3\text{Br}$  // ЖЭТФ. 1986. Т. 90, № 6. С. 1982—1994.
10. Баграташвили В. Н., Вайнер Ю. Г., Должииков В. С. и др. Меж- и внутримолекулярное распределение колебательной энергии молекул при многофотонном возбуждении ИК лазерным излучением // Там же. 1981. Т. 80, № 3. С. 1008—1025.
11. Mazur E., Burak I., Bloembergen N. Collisionless vibrational energy redistribution between infrared and Raman active modes in  $\text{SF}_6$  // Chem. Phys. Lett. 1984. Vol. 105. N 3. P. 258—262.

12. *Ambartzumian R. V., Makarov G. N., Puretzky A. A.* Investigation of multiple photon excitation of  $\text{OsO}_4$  by dissociation yield saturation // *Opt. Commun.* 1978. Vol. 27, N 1. P. 79—84.
13. *Баграташвили В. Н., Должигов В. С., Летохов В. С.* Кинетика спектров ИК поглощения молекул  $\text{SF}_6$ , колебательно возбуждаемых мощным импульсом  $\text{CO}_2$ -лазера // *ЖЭТФ*. 1979. Т. 76, № 1. С. 18—25.
14. *Doljikov Yu. S., Letokhov V. S., Makarov A. A.* et al. Raman probing of overtone and combination bands to study the vibrational energy distribution produced by multiple-photon excitation // *Chem. Phys. Lett.* 1986. Vol. 124, N 4. P. 304—308.
15. *Кондратьев В. Н., Никитин Е. Е.* Кинетика и механизм газофазных реакций. М.: Наука, 1974. 558 с.
16. *Робинсон П., Холбрук К.* Мономолекулярные реакции. М.: Мир, 1975. 380 с.
17. *Rice S. A.* An overview of the dynamics of intramolecular transfer of vibrational energy // *Photoselective chemistry. Pt 1 / Ed. J. Jortner, R. D. Levine, S. A. Rice. N. Y. etc.: Wiley, 1981. P. 117—200. (Adv. Chem. Phys.; Vol. 47).*
18. *Зеваил А.* Селективная лазерная химия — возможно ли это? // *Физика за рубежом*. М.: Мир, 1982. С. 63—89.
19. *Кузьмин М. В., Летохов В. С., Стучебрюхов А. А.* Пороговая энергетическая зависимость скорости внутримолекулярной колебательной релаксации изолированной многоатомной молекулы // *ЖЭТФ*. 1986. Т. 90, № 2. С. 458—470.
20. *Bloembergen N., Zewail A. H.* Energy redistribution in isolated molecules and the question of mode-selective laser chemistry revisited // *J. Phys. Chem.* 1984. Vol. 88, N 23. P. 5459—5465.
21. *Дыкман М. И., Кривоглаз М. А.* Квантовая теория нелинейных осцилляторов, взаимодействующих со средой // *ЖЭТФ*. 1973. Т. 64, № 3. С. 993—1006.
22. *Bagratashvili V. N., Kuzmin M. V., Letokhov V. S., Stuchebrukhov A. A.* Theory of multiple-photon IR excitation of polyatomic molecules in the model of active and passive modes of a vibrational reservoir // *Chem. Phys.* 1985. Vol. 97, N 1. P. 13—29.
23. *Reddy K. V., Heller D. F., Berry M. J.* Highly vibrationally excited benzene: Overtone spectroscopy and intramolecular dynamics of  $\text{C}_6\text{H}_6$ ,  $\text{C}_6\text{D}_6$ , and partially deuterated or substituted benzenes // *J. Chem. Phys.* 1982. Vol. 76, N 6. P. 2814—2837.
24. *Baggott J. E., Chuang Mei-Chen, Zare R. N.* et al. Structure and dynamics of the excited  $\text{CH}$ -chromophore in  $(\text{CF}_3)_3\text{CH}$  // *Ibid.* 1985. Vol. 82, N 3. P. 1186—1194.
25. *Evseev A. V., Krivitsun V. M., Kuritsyn Yu. A.* et al. IR absorption spectrum of  $\text{CrO}_5\text{Cl}_2$  molecules for high-lying states of the vibrational quasi-continuum // *Chem. Phys.* 1986. Vol. 106, N 1. P. 131—149.
26. *Баграташвили В. Н., Ионов С. И., Летохов В. С.* и др. Однородный спектр колебательных переходов и время внутримолекулярной релаксации сильновозбужденной многоатомной молекулы // *Письма в ЖЭТФ*. 1986. Т. 44, № 10. С. 450—453.
27. *Nitzan A., Persson B. N. J.* Vibrational dephasing by the exchange mechanism: some new results // *J. Chem. Phys.* 1985. Vol. 83, N 11. P. 5610—5618.
28. *Макаров А. А., Тяжт В. В.* Чисто фазовая внутримолекулярная релаксация и ее влияние на спектры переходов между высоковозбужденными колебательными состояниями молекул // *ЖЭТФ*. 1987. Т. 93, № 1. С. 17—30.
29. *Макаров А. А.* Микроволновая спектроскопия как метод исследования стохастизации колебательного движения в многоатомной молекуле. Оптика и спектроскопия. 1987. Т. 62, № 5. С. 1183—1185.

УДК 536 182

## ЯВЛЕНИЯ САМООРГАНИЗАЦИИ У БАКТЕРИЙ НА КЛЕТОЧНОМ И ПОПУЛЯЦИОННОМ УРОВНЯХ

В. Г. БАБСКИЙ

Бактерии представляют собой, как правило, подвижные микроорганизмы, что позволяет им активно искать лучшие условия обитания. Возникающие при этом явления, примеры которых будут приведены в статье, характеризуются тесно связанными между собой процессами, происходящими на молекулярном, организменном (или клеточном, поскольку речь идет об одноклеточных) и популяционном уровнях.

### 1. АРХИТЕКТОНИКА БАКТЕРИАЛЬНЫХ КОЛОНИЙ

В лабораторных условиях бактерии часто выращивают на стандартных твердых или полутвердых питательных средах (обычно на базе агар-агара) [1]. При этом бактериальной суспензией покрывают поверхность среды, и, вообще говоря, каждая бактерия дает начало бактериальной колонии. В процессе роста колонии еще много загадочного: клетки делятся, нагромождаясь одна на другую, и их взаимное расположение определяется адгезионными и другими свойствами поверхности клеток. Колония растет, как правило, в виде полусферической шапочки на поверхности среды; необходимые для роста клеток питательные вещества поступают из толщи среды посредством диффузии; также посредством диффузии отводятся продукты метаболизма. Предельный размер колонки достигается, когда в зоне, прилегающей к колонии, наступает истощение питательной среды и (или) отравление продуктами метаболизма. Поверхность колонии может быть гладкой или «шероховатой», да и сама колония, помимо простой полусферической формы, может принимать весьма причудливые (например, спиральные) очертания [2]. Эти отличия отражают физиологические и генетические изменения, происходящие в клетках.

Оказывается, что в сечении колония не является однородной, а обладает развитой структурой [2]: по мере удаления от поверхности колонии поры становятся более крупными, что облегчает процесс диффузии. Это результат дифференцировки клеток, при-

чем не только по форме и размерам, но и по физиологической активности [3]. Таким образом, колонию можно рассматривать как простейший многоклеточный организм, а переход популяции бактерий к такому способу существования — как адаптацию к прогрессирующей экстремальной ситуации — дефициту питательных веществ и (или) накоплению продуктов метаболизма, причем эта адаптация индуцирована диффузией. Дело в том, что недостаток или избыток тех или иных веществ в колонии, созданный процессами метаболизма и диффузии, приводит к физиологическим и морфологическим изменениям делящихся клеток и, как следствие этого, к возникновению пространственной структуры, облегчающей диффузию.

Здесь пока еще не используется более мощный с точки зрения эволюции процесс адаптации — присущая клеткам подвижность. Это связано с тем, что органы подвижности бактерий приспособлены для перемещения в жидкости, в то время как поверхность твердых агаризованных сред покрывает тонкая жидкая пленка, толщина которой (сравнимая с характерным размером бактерий) зависит от окружающей влажности и концентрации агар-агара в питательной среде. Тем не менее существует изложенный ниже пример, когда популяция бактерий перемещается и в такой поверхностной пленке жидкости, переходя для этого в особую социальную фазу.

## 2. РОЕНИЕ БАКТЕРИЙ

Явление роения характерно для бактерий рода *Proteus* [2, 4]. Здесь рост колонии не прекращается на стадии полусферической шапочки предельного размера; после достижения этого размера на периферии колонии в пленке на поверхности среды появляются в результате блокировки деления клеток особые формы бактерий (швермеры), во много раз более длинные, чем обычные бактерии, и с существенно большим числом жгутиков (органы подвижности бактерий). Сцепленные между собой в единую подвижную структуру (пленку роения), они осуществляют коллективное радиальное перемещение (собственно роение). По истечении определенного промежутка времени (периода роения) швермеры начинают делиться, давая начало новой колонии нормально делящихся клеток. В результате вся поверхность агаризованной среды оказывается покрытой кольцевыми колониями, концентрическими относительно исходной колонии (рис. 1, а).

Существующие экспериментальные факты позволяют трактовать это как проявление генетического триггера (см. [5]). Предполагается, что этот триггер переключает внутриклеточный метаболизм с синтеза белка флагеллина  $P_1$ , из которого состоят жгутики (швермеры называют фабрикой флагеллина), на синтез некоторого (пока не идентифицированного) белка  $P_3$ , необходимого для завершения процесса деления бактерий. Гипотетическая схема такого триггерного переключения состоит в том, что

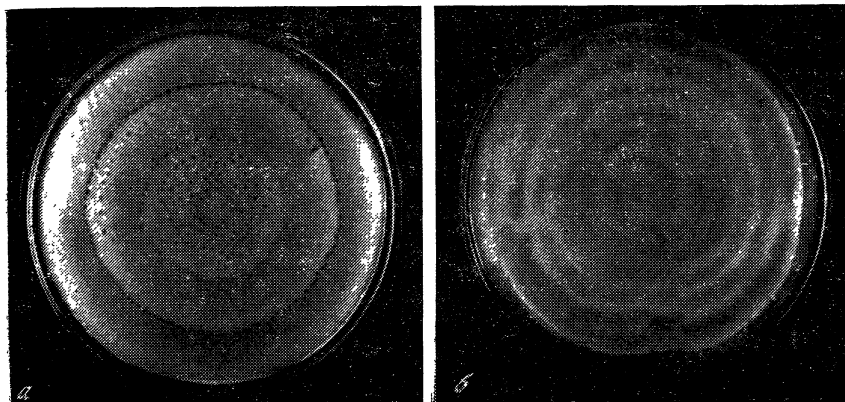


Рис. 1. Картина роения *Proteus vulgaris* на мясо-пептонном агаре с концентрацией 0,9 (а) и 1,1% (б) [4]

Диаметр чашки Петри 90 мм

некоторый летучий продукт метаболизма бактерий  $P_2$  интенсифицирует синтез флагеллина  $P_1$  и при последующей конкуренции за общий аминокислотный пул (субстрат  $S$ ) синтез  $P_3$  падает — именно это и приводит к блокировке деления и образованию покрытых жгутиками швермеров.

Замечательно, что переключение триггера во второе устойчивое состояние (на нормальный синтез  $P_3$ ) практически не зависит от внешних условий, а характеризует инерционность внутриклеточных процессов [5]. Другими словами, период роения не зависит от внешних условий; экспериментально [4] это выражается в том, что, например, при увеличении концентрации агар-агара в питательной среде уменьшается толщина пленки жидкости, растет ее сопротивление движению швермеров и, как следствие этого, уменьшается расстояние между кольцевыми колониями (рис. 1, б), однако время формирования одной кольцевой зоны остается практически неизменным, т. е. период роения (на рис. 1 около 3,5 ч) является собственной характеристической триггера.

Коллективное движение швермеров при более подробном рассмотрении носит весьма сложный характер (наблюдавшие его под микроскопом пишут об этом явлении с восхищением [2]). Наблюдаемая в эксперименте [4] постоянная скорость движения границы роения позволяет предложить в качестве «имитационной» модель с нелинейной диффузией:

$$\partial P_1 / \partial t = f_1(P_1, P_2, S) + \text{grad}(D(P_1)\text{grad}P_1), \quad D \sim P_1^n \quad (n > 1), \quad (1)$$

$$\partial P_2 / \partial t = f_2(P_1, P_2, S), \quad \partial S / \partial t = f_3(P_1, P_2, S), \quad P_2 \sim P_3,$$

при этом движение (роя) швермеров рассматривается как движение флагеллина  $P_1$  с коэффициентом диффузии  $D$ , зависящим от

$P_1$  (зависимость  $D \sim P_1^n$  отражает как коллективность движения, так и конечную скорость распространения границы роения).

Необычность модели (1) по сравнению с другими уравнениями реакции—диффузии [5] заключается в том, что диффузия происходит при остановившейся кинетике:  $f_1 = f_2 = f_3 = 0$ ; действительно, деление клеток остановилось, образовались швермеры, и они движутся. С другой стороны, переключение триггера на синтез  $P_3$  и рост колоний происходит при пренебрежимо малой подвижности клеток ( $D = 0$ ). С биологической точки зрения чередование коллективной подвижности швермеров со статическим ростом колоний представляет собой своеобразную смену разных социальных фаз развития. Решение же модели (1) можно, не уточняя вида функций  $f_1, f_2, f_3$ , трактовать как движение фронта флаголлина, амплитуда и скорость распространения которого испытывают релаксационные колебания во времени. Роение у бактерий является лишь одной из альтернатив быстрого перемещения к свежим порциям питательной среды. Большая группа других явлений, приводящих к направленной миграции клеточных популяций, носит название таксиса.

### 3. ХЕМОТАКСИС

При культивировании бактерий в жидкой или полужидкой (например, с добавлением небольших концентраций агар-агара) питательной среде их рост происходит во всем объеме субстрата. При этом большинство бактерий в норме обладает способностью воспринимать градиенты концентрации различных веществ. И хотя отдельные микроорганизмы, вращая жгутики, мигрируют случайным образом (см. обзор в [4]), вся популяция имеет тенденцию к перемещению в направлении увеличения концентрации (положительный хемотаксис) или ее уменьшения (отрицательный хемотаксис). Базовая математическая модель, описывающая хемотаксис у бактерий, имеет вид [6]

$$\partial b / \partial t = \text{grad} (m(s) \text{grad} b - b\beta(s) \text{grad} s) + F_1(s, b),$$

$$\partial s / \partial t = D\Delta s + F_2(s, b), \quad (2)$$

где  $b$  — концентрация микроорганизмов;  $s$  — концентрация вещества, индуцирующего хемотаксис (аттрактанта в случае положительного хемотаксиса и репеллента в случае отрицательного);  $m(s)$  — коэффициент диффузии (симметричного случайного блуждания) микроорганизмов;  $D$  — коэффициент диффузии  $s$ ;  $F_1(s, b)$ ,  $F_2(s, b)$  — функции, описывающие производство и деградацию микроорганизмов и вещества  $s$ , которое может метаболизироваться микроорганизмами или быть продуктом метаболизма;  $\beta(s)$  — коэффициент хемотаксической подвижности или несимметричного случайного блуждания ( $\beta(s) > 0$  при положительном и  $\beta(s) < 0$  при отрицательном хемотаксисе).

Модель (2), в которой, кроме обычной диффузии, учтена отри-

цательная перекрестная диффузия, отражающая хемотаксическое поведение, может описывать большинство явлений, наблюдающихся при экспериментальном исследовании хемотаксиса у бактерий: это бегущие импульсы, напоминающие фронт распространения пламени, когда, например, зона бактерий, потребляющая аттрактант, создает крутой градиент его концентрации и перемещается вместе с ним; стационарные неоднородные распределения микроорганизмов. Последние могут возникать в фиксированных градиентах вещества  $s$ ; скопления бактерий могут быть довольно существенными — теоретически даже говорят о «хемотаксическом коллапсе» (это и есть очевидное проявление отрицательной диффузии) [7]. Заметим, что стационарные неоднородные состояния соответствуют стационарным решениям системы (2), отвечающим (мягко или жестко) от пространственно-однородного распределения  $b$  и  $s$ .

В отличие от процесса роста колоний и роения, процесс хемотаксиса у бактерий детально изучен с позиций биохимии, генетики, механики. Тем не менее уже простая модель реакции—диффузии (но нелинейной диффузии!) дает вполне удовлетворительное описание. Дальнейшее развитие моделей нелинейной диффузии для описания явлений самоорганизации в биологии представляется весьма перспективным.

#### ЛИТЕРАТУРА

1. *Перт С. Д.* Основы культивирования микроорганизмов и клеток. М.: Мир, 1978. 331 с.
2. *Пешков М. А.* Цитология бактерий. М.; Л.: Изд-во АН СССР, 1955. 220 с.
3. *Wimpenny J. W. T., Parr J. A.* Biochemical differentiation in large colonies of enterobacter cloacae // *J. Gen. Microbiol.* 1979. Vol. 114, N 2. P. 487—489.
4. Микроорганизмы в космическом полете / Под общ. ред. К. М. Сытника. Киев: Наук. думка, 1983. 156 с.
5. *Романовский Ю. М., Степанова Н. В., Чернавский Д. С.* Математическая биофизика. М.: Наука, 1984. 304 с.
6. *Keller E. F., Segel L. A.* Initiation of slime mold aggregation viewed as an instability // *J. Theor. Biol.* 1970. Vol. 26, N 3. P. 399—415.
7. *Childress S., Percus J. K.* Nonlinear aspects of chemotaxis // *Math. Biosci.* 1981. Vol. 56, N 1. P. 217—237.

УДК 577.3.001.57

### «ПОЛЕВЫЕ» ЧЕРТЫ В МОРФОГЕНЕЗЕ И МЕХАНОХИМИЧЕСКАЯ АКТИВНОСТЬ КЛЕТОК

Б. Н. БЕЛИНЦЕВ

Ни одна из конструкций, созданных человеком, не способна конкурировать по сложности решаемых задач с простейшими из живых созданий. Одной из таких задач является производство себе подобных. Каким образом яйцеклетка и спермий, объединяясь, дают зародыш, а из него, в свою очередь, развивается взрослый организм?



Ключевой идеей современного подхода к обоснованию зародышевого развития является идея о самоорганизации. Эта идея сформулирована в результате осознания принципиальных трудностей, с которыми сталкивается чисто причинностная трактовка развития, или так называемая механика развития.

Общая точка зрения последней выражается в предположении, что план строения организма, т. е. план упорядоченного пространственного распределения различных типов специализированных клеток, в той или иной форме существует уже в раннем зародыше. Неправомерность этой точки зрения вытекает из экспериментов, вскрывших существование так называемых эмбриональных (дришевских) регуляций [1, 2]. В этих экспериментах при перемешивании, добавлении или удалении иногда значительной доли клеточного материала раннего зародыша наблюдалось развитие «правильного целого». Например, половина бластулы морского ежа может формировать уменьшенное по размерам, но вполне развитое животное со всеми положенными структурами.

Подобные результаты приводят к выводу, что план строения, определяющий разную судьбу в развитии разных частей зародыша, не задан изначально, но формируется в результате пространственной саморегуляции развития [3]. В развивающемся зародыше активность клеток (включение нужных генов, синтез нужных белков и т. д.) проявляется в свое время, в своем месте. Этот макроскопический порядок не задан заранее, но именно порождается как продукт группового поведения клеток. Задача состоит в том, чтобы обосновать такую самоорганизацию зародышевого материала исходя из его более элементарных свойств.

### 1. РАЗМЕТКА МОРФОГЕНЕЗА И МОРФОГЕНЕТИЧЕСКОЕ ПОЛЕ

В развитии разных зародышевых зачатков у зародышей разных видов удается выделить общую компоненту. А именно многие данные указывают на то, что прежде активации определенных, специфических для данного зачатка генов и соответствующих биосинтезов осуществляется разметка [4]. Ею устанавливается группа клеток, в которой эта активация произойдет на следующем этапе развития. В названной последовательности событий все, что определяет уникальность того или иного органа, приходится на вторую стадию. Разметка же от соответствующей специфичности не зависит и в этом смысле универсальна [4].

Разметку можно охарактеризовать рядом типичных, повторяющихся черт, которыми определяется так называемый полевой аспект развития. Данный аспект выражается в тех экспериментально регистрируемых ситуациях, когда локальное воздействие на зародыш приводит к переопределению судьбы в развитии целых областей [5].

Характерные признаки «полевого» поведения таковы:

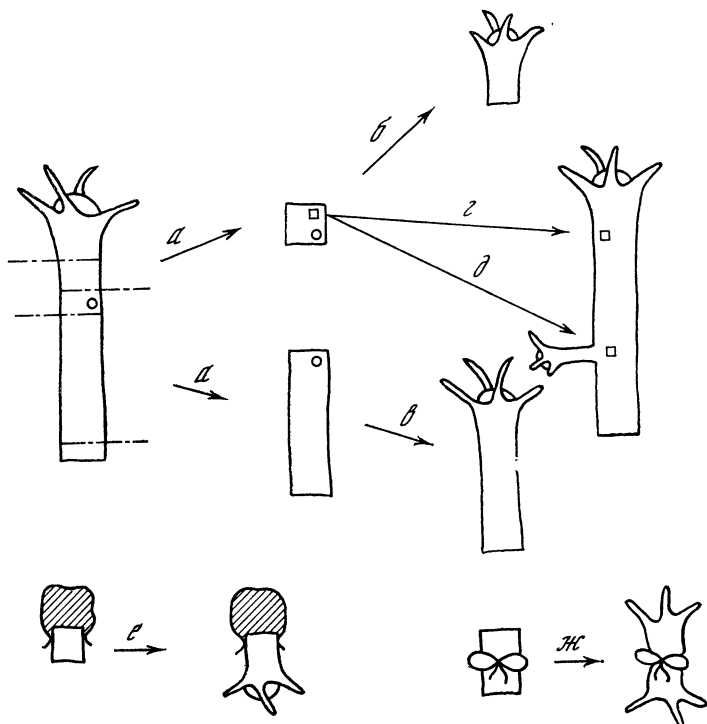


Рис. 1. «Полевые» аспекты развития, проявляющиеся при регенерации фрагментов гидроидного полипа [15]

1. Эффект подавления. Как правило, тот или иной эмбриональный зачаток формируется лишь частью клеточного материала, способного его формировать. При этом зарождающийся побег является источником действия, подавляющего способность к зарождению другого такого же в окружающих областях. Характерная величина радиуса распространения подавляющего действия  $R \approx \lesssim 1$  мм.

На рис. 1 представлены «полевые» аспекты развития, проявляющиеся при регенерации фрагментов гидроидного полипа [15]. Любая часть туловища способна регенерировать головной отдел, но эта способность проявляется в зависимости от того, как выделен сегмент (сравните судьбу клеток, отмеченных кружком, при двух способах выделения сегмента — б и в). Существующий головной отдел или формирующийся зачаток подавляет развитие нового; этот эффект падает с удалением от имеющегося зачатка (г и д). Эффект подавления на нижнем конце сегмента устраняется, если блокировать развитие головного зачатка на верхнем конце (е) или изолировать нижний конец путем перетяжки (ж).

2. Регуляция пропорций (размерная инвариантность) проявляется на ранних этапах развития, когда весь клеточный материал

зародыша оказывается в пределах указанного радиуса. Свойство размерной инвариантности выражается в том, что изменение общего объема зародыша (путем устранения или добавления клеточного материала) оставляет неизменными пропорции развивающихся структур. Пример — упоминавшиеся дришевские регуляции.

3. Разметка осуществляется путем самоорганизации. Начало формирования побега в компетентной области, как правило, связано с некоторой изначальной неоднородностью или внешним стимулом. Однако после внешней инициации дальнейший процесс разметки осуществляется автономно. Так, в некоторых экспериментах удавалось наблюдать развитие побега вообще не в том месте, где это развитие было инициировано [6] (см. далее рис. 6, в).

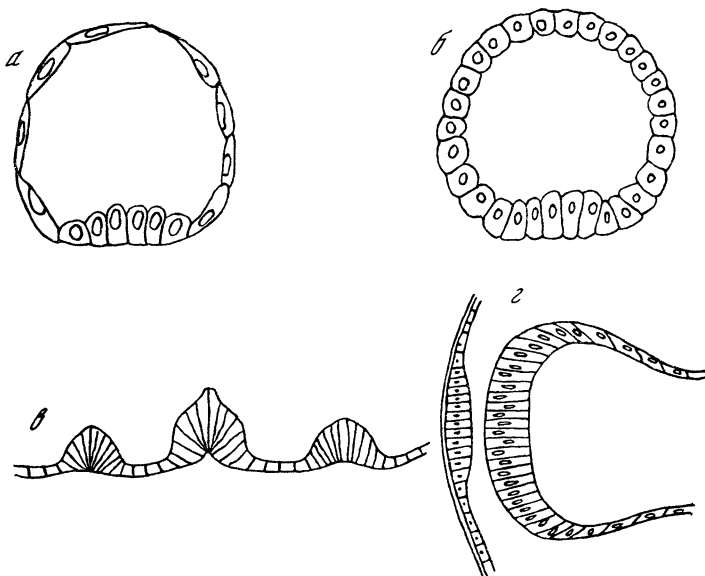
4. Макроскопические паттерны. Радиус действия поля подавления может оказаться меньше размера области зародыша, в которой может развиваться порождающий его зачаток. В таком случае возможно возникновение более или менее регулярной решетки повторяющихся зачатков. Это происходит, в частности, при морфогенезах зачатков различных структур кожного покрова — перьев, волосков, щетинок, чешуек, желез. Геометрия этой решетки может быть различной — для перьевых зачатков ее ячейки представляют собой правильные шестиугольники, волоски или щетинки часто организованы либо периодически повторяющимися рядами, либо вообще нерегулярно.

Перечисленные признаки поведения, связанного с разметкой морфогенеза, дают представление о феноменологическом содержании понятия морфогенетического поля. Именно с морфогенетическим полем связывают элементарное звено самоорганизации надклеточного порядка в зародышевом развитии. Очевидно, понятие морфогенетического поля, или, более определенно, набор связанных с ним проявлений, нуждается в интерпретации в терминах измеряемых параметров, характеризующих клетки и их взаимодействия. Для этого необходимо охарактеризовать соответствующие регистрируемые формы активности клеток.

## 2. РЕГИСТРАЦИЯ РАЗМЕТКИ. ЭЛЕМЕНТАРНЫЕ СВОЙСТВА КЛЕТОЧНОГО МАТЕРИАЛА

Наблюдения тонких срезов эмбриональных тканей показывают, что закладка многих, едва ли не всех, структур в зародыше осуществляется путем локальных деформаций пластов сцепленных клеток — эпителиев. Единственным регистрируемым признаком предварительной разметки во всех случаях является вычленение доменов столбчатых клеток [7]. (На рис. 2 показаны некоторые из соответствующих многочисленных примеров.) В следующий промежуток времени эпителиальный пласт деформируется благодаря активным движениям выделенной группы клеток.

На клеточном уровне такой разметке можно соотнести следующие свойства клеток.



**Рис. 2.** Первичная организация эпителиальных морфогенезов на примерах раннего зародыша *Eriperipatus trinidadensis* (а), поздней бластулы морского ежа (б), глазных рудиментов зародыша *Limulus* (в) и рудиментов линзы и глаза у зародышей позвоночных (г)

1. Индивидуальные клетки обладают механохимической активностью. За счет метаболических источников энергии (гидролиза АТФ) они способны самопроизвольно деформироваться, обладая двумя дискретными локально-устойчивыми формами: изотропной и морфологически поляризованной.

2. Активные деформации клеток порождают упругие напряжения в зародышевом материале. Скорость релаксации этих напряжений существенно превосходит скорость активных деформаций.

3. Упругие напряжения модулируют процесс активной деформации клеток, так, что касательные натяжения в пласте подавляют активное вытяжение клеток в нормальном направлении. Релаксация упругих напряжений в эмбриональном материале приводит к утрате черт дальнего порядка: появляется более или менее хаотическое распределение столбчатых клеток вместо стандартного расчленения, показанного на рис. 2.

4. Имеет место эффект контактной поляризации клеток [8, 9] — транслокации механохимической активности между контактирующими клетками.

Подчеркнем, что названные свойства не являются произвольными постулатами. Они выделены [10, 11] на основе надежного фактического материала, полученного в экспериментальных исследованиях [3, 8, 12—15].

Выделенные свойства клеточного материала сами по себе не элементарны. Обуславливающие их субклеточные механизмы, очевидно, достаточно сложны. Однако при постановке вопроса об организации групповой активности клеток в эмбриональной ткани эти свойства могут рассматриваться как элементарные. Именно на их основе с помощью математической модели далее интерпретируются черты «полевого поведения» в ходе самоорганизации разветки морфогенеза.

### 3. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Выделенные постулаты определяют наличие двух типов обратной связи и тем самым обеспечивают регуляцию клеточной активности. Благодаря эффекту «контактной поляризации» локальный домен морфологически поляризованных клеток стремится расшириться — положительная обратная связь. С другой стороны, в результате этого расширения нарастают касательные натяжения, подавляющие поляризацию новых клеток, — отрицательная обратная связь.

Приведенные качественные соображения не дают полной картины группового поведения клеток. Для этой цели необходимо более строгое математическое рассмотрение.

Идеализированный эпителиальный пласт будем представлять в виде плоского клеточного монослоя. Отдельные клетки представляются как упругие элементы, релаксированным состоянием которых является изотропное. Клетки пласта, помимо того, рассматриваются как механически активные элементы, поскольку обладают способностью переходить в альтернативное устойчивое состояние, соответствующее колонообразной форме. Монослой прикреплен к эластичному субстрату, в котором могут возникать только упругие, т. е. чисто пассивные силы.

Моделью не постулируется какой-либо определенный механизм активных деформаций клеток. Механическая активность представляется в общей форме следующим образом. В качестве характеристики внутриклеточной системы регуляции механохимической активности вводится скалярный динамический параметр  $P$  [10, 11]. По предположению, он равен нулю в состоянии выключенной механохимической активности и положителен для клеток, вытянутых вследствие такой активности нормально к плоскости пласта.

Параметр  $P$  в нашем подходе играет такую же роль, как параметр порядка в феноменологической теории фазовых переходов [16]. Определяемый как самая медленная макроскопическая степень свободы параметр порядка позволяет формулировать замкнутое описание перехода с изменением макроскопической фазы, несмотря на отсутствие микроскопической теории. Аналогично введенный параметр  $P$  характеризует некоторую «обобщенную координату», которой отслеживается лимитирующая темп стадия

сложной совокупности координированных субклеточных процессов, ведущих к морфологической поляризации клетки.

Основываясь на этом общем определении, можно переформулировать в терминах замкнутой математической модели регистрируемые макроскопические черты процесса клеточной поляризации (см. свойства 1—4). Такая формулировка представляется в виде уравнения

$$\partial P / \partial t = (1/\tau) f(P) + D \nabla^2 P - \alpha \sigma_{ii}. \quad (1)$$

Здесь распределение клеточной поляризации в пласте определено скалярным, непрерывным полем  $P(\mathbf{r}, t)$ ;  $\mathbf{r} = (x, y)$ .

Будучи скрытым параметром,  $P$  не может быть измерен непосредственно в эксперименте. Однако при гистологических наблюдениях всегда удается дискриминировать поляризованные и неполяризованные клетки. Поэтому доступны измерению и такие характеристики, как скорость распространения фронта клеточной поляризации<sup>1</sup> и геометрические параметры паттернов поляризованных клеток. На основании этих данных можно получить оценки параметров  $\tau$  и  $D$  уравнения (1). Соответствующие оценки будут приведены в последующих публикациях.

Функция  $f(P)$  в уравнении (1) отражает авторегулируемую внутриклеточную динамику, обуславливающую активные деформации. Наличие двух дискретных стабильных морфологических состояний клеток моделируется И-образным видом зависимости  $f(P)$  (рис. 3). Вторым слагаемым в уравнении (1) представлен эффект контактной поляризации клеток, третьим — влияние касательных напряжений на процесс активного сжатия в плоскости пласта ( $\sigma_{ii}$  — след тензора касательных напряжений). (Более подробные разъяснения см. в работе [17].)

Касательные напряжения включают в себя активную (связанную с использованием метаболической энергии) и пассивную (связанную с эластичностью клеток) составляющие:

$$\sigma_{ik} = \sigma_{ik}^{\text{акт}} + \sigma_{ik}^{\text{пас}}. \quad (2)$$

Первая возникает благодаря активному механизму внутриклеточных перестроек. Величина  $\sigma_{ik}^{\text{акт}}$  предполагается линейно связанной с параметром  $P$ , отслеживающим эти перестройки:

$$\sigma_{ik}^{\text{акт}} = \alpha \delta_{ik} P. \quad (3)$$

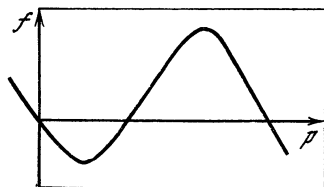


Рис. 3. Характер функции  $f(P)$  в уравнении (1), отражающий бистабильность активной морфологической динамики клеток

<sup>1</sup> Согласно наблюдениям [13], процесс распространения поляризации в пласте имеет характер триггерной волны.

Здесь символ Кронекера  $\delta_{ik}$  отражает изотропный (в плоскости  $\{x, y\}$ ) характер натяжений, генерируемых нормально удлиняющимися клетками.

Упругие напряжения, фигурирующие в выражении (2), порождаются активными клеточными деформациями. Для формального представления такой взаимосвязи вполне можно использовать условие механического равновесия. Приемлемость этого условия оправдывается наблюдаемой быстротой механической релаксации по сравнению с процессом клеточной поляризации [12].

Условие баланса касательных сил выражается в виде

$$-qu_i + \partial\sigma_{ik}/\partial x_k = 0. \quad (4)$$

Здесь первый член соответствует упругой реакции подложки;  $u_i$  — вектор касательных смещений;  $q$  — коэффициент сдвиговой жесткости подложки; второй член представляет силу, возникающую вследствие неоднородной деформации самого пласта.

Уравнение (4) с использованием стандартного выражения тензора упругих напряжений через тензор деформаций [18] позволяет восстановить поле деформаций  $u_{ik}$  по полю активных напряжений  $\sigma_{ik}^{\text{акт}}$  (т. е. распределению  $P(\mathbf{r}, t)$ ) [19]. Решая это уравнение, находим величину  $\sigma_{ii}$ , которая, в свою очередь, контролирует динамику клеточной поляризации (см. (1)). Результат таков:

$$\sigma_{ii} = \frac{\alpha(1+\sigma)}{R_e^2} \int K_0 \left( \frac{|\mathbf{r}-\mathbf{r}'|}{R_e} \right) P(\mathbf{r}', t) d\mathbf{r}' + \alpha(2-\sigma)P, \quad (5)$$

где величина  $R_e$  выражается через параметры задачи:  $R_e^2 = E/q(1-\sigma^2)$ ;  $E$  — модуль Юнга пласта;  $\sigma$  — коэффициент Пуассона;  $K_0$  — функция Мак-Дональда (функция источника для оператора  $R_e^2\nabla^2 - 1$ ). Выражение (5) вместе с (1) дает замкнутое описание коллективной динамики клеточной поляризации. Последнее имеет вид уравнения относительно  $P(\mathbf{r}, t)$ :

$$\frac{\partial P}{\partial t} = \frac{1}{\tau} \hat{f}(P) + D\nabla^2 P - \frac{\eta}{R_e^2} \int K_0 \left( \frac{|\mathbf{r}-\mathbf{r}'|}{R_e} \right) P(\mathbf{r}', t) d\mathbf{r}'. \quad (6)$$

Здесь  $\hat{f}(P) = f(P) + \alpha\alpha(\sigma - 2)P$ ;  $\eta = \alpha\alpha(1 + \sigma)$ .

Существенной чертой самого процесса и его описания (6) является эффективная нелокальность. Локальная поляризация клеток генерирует упругие напряжения, распространяющиеся на прилегающие неполяризованные области (интегральный член в уравнении (6)). Этот эффект существен до расстояний порядка  $R_e$ , растет по мере расширения области поляризации, препятствуя этому расширению. Тем самым обеспечивается нелокальная отрицательная обратная связь в процессе клеточной поляризации. Наряду с  $R_e$  модель имеет второй характерный пространственный масштаб. Он определяет длину когерентности контактной поляризации и оценивается как  $\tilde{R}_p = \sqrt{D\tau}$ .

Следует выделить специальный случай, когда пласт активных клеток не контактирует с пассивной подложкой. В этом случае

первый член в уравнении (4) исчезает, и в результате только один из двух характерных масштабов, а именно  $R_p$ , остается в задаче (формально  $R_e \rightarrow \infty$ ). Формально бесконечный радиус упругой реакции делает эффект граничных условий существенным для поведения всего пласта. В важном частном случае, когда пласт имеет неизменную общую протяженность,

$$\sigma_{ii} = \alpha \langle P \rangle (1 + \sigma) + (2 - \sigma) \alpha P. \quad (7)$$

Соответственно уравнение для  $P$  имеет вид

$$\partial P / \partial t = (1/\tau) \bar{f}(P) + D \nabla^2 P - \alpha a \langle P \rangle (1 + \sigma). \quad (8)$$

Здесь угловые скобки соответствуют усреднению по всему пласту.

Наличие такой когерентной составляющей в динамике существенно для придания черт глобальной регуляции клеточной активности и, в частности, обеспечивает свойство размерной инвариантности формирующихся паттернов.

#### 4. АНАЛИЗ ПОВЕДЕНИЯ В РАМКАХ МОДЕЛИ (6)

Исходное состояние — пласт — до включения морфогенетической активности соответствует стационарному решению уравнения (6):

$$P \equiv 0 \quad (\bar{f}(0) = 0). \quad (9)$$

Анализ его устойчивости по отношению к нормальным (гармоническим) модам возмущений в линейном приближении указывает на существование формальной предпосылки неустойчивости тьюрингова типа (колоколообразная дисперсионная кривая (рис. 4)) [20]. Порог устойчивости достигается при

$$\sqrt{\eta} R_e > R_p, \quad \mu_{кр} = 0, \quad \mu = \sqrt{\beta + 2\eta} - \sqrt{\eta} - R_p/R_e, \quad \beta = f'_p(0). \quad (10)$$

Волновое число возбуждаемой за порогом (при  $\mu > 0$ ) неустойчивой моды выражается в виде

$$k_{кр}^2 = \sqrt{\eta}/R_p R_e - 1/R_e^2. \quad (11)$$

Таким образом, физические посылки, заложенные в основу модели, обеспечивают механизм самоорганизации коллективной механохимической активности клеток через неустойчивость тьюрингова типа.

Характер бифуркации (закритический или докритический) пространственно-неоднородных решений уравнения с диффузионным оператором и нелокальной пространственной связью типа (6) исследован в работе [21]. Отсылая к ней за более точными оценками, укажем лишь на существенную зависимость его от отношения выделенных пространственных масштабов  $r = R_p/R_e$ .

При  $r \ll 1$  устойчивые пространственно-неоднородные распределения  $P(r)$  могут формироваться уже до порога тьюринго-



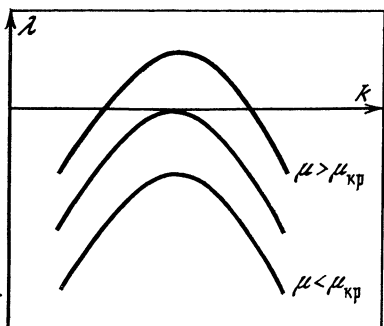
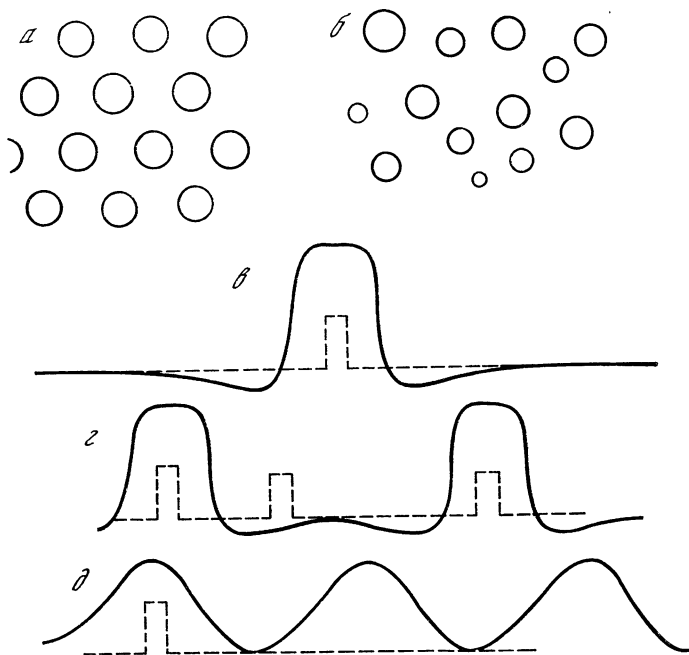


Рис. 4. Форма дисперсионной зависимости  $\lambda(k)$  для линейризованного уравнения (6) при трех значениях бифуркационного параметра  $\mu$

Рис. 5. Вид стационарных решений уравнения (6) в плане для двумерных распределений  $P(r)$  (а, б) и в профиль для одномерных распределений  $P(x)$  (в — д)



вой неустойчивости. Тьюрингова мода для таких макросостояний, вообще говоря, не является доминирующей; дальнего порядка они лишены. При этом, однако, наличествует элемент ближнего порядка, проявляющегося в том, что домены поляризованных клеток (значение  $P$  в окрестности правой нулевой точки функции  $\tilde{f}(P)$ ) разделены промежутками не уже радиуса упругой реакции  $R_e$ . Характерно также отсутствие однозначного асимптотического состояния, т. е. зависимость такового от начальных условий.

На рис. 5 представлены стационарные решения уравнения (6) для двумерных  $P(r)$  и одномерных  $P(x)$  распределений параметра  $P$ . Паттерн  $P(r)$ , генерируемый тьюринговой модой (бифуркация при  $r \approx 1$ ), обладает дальним порядком (а и в). При этом локальный триггер индуцирует коллективную, пространственно-

периодическую моду клеточной активности ( $\partial$ ). При  $r \ll 1$  тьюрингова мода не является доминирующей; поэтому дальний порядок отсутствует ( $\delta - \varepsilon$ ). При докритическом зарождении негомогенных состояний («жестким» начальным возмущением — пунктир — асимптотически устойчивого исходного однородного состояния) последние имеют вид локализованных доменов, формирующихся по месту начального возмущения ( $\theta$ ). В зоне, окружающей домен поляризованных клеток, развитие нового подавляется ( $\varepsilon$ ). Формальный анализ таких «нетьюринговых», контрастных структур проведен в работах [24, 25].

При сопоставимых характерных длинах  $R_p \approx R_e$  тьюрингова мода  $\cos k_{кр} r$  вблизи порога неустойчивости исходного состояния оказывается доминирующей. Ветвление одномерных периодических решений с длиной волны  $\lambda = 2\pi/k_{кр}$  происходит заkritически [22]. Двумерные (ячеистые) пространственно-регулярные распределения  $P(r)$  зарождаются с конечной амплитудой до порога тьюринговой неустойчивости [21]. Для случая  $r \approx 1$  характерно формирование однозначного паттерна  $P(r)$  независимо от начальных условий (это справедливо лишь в окрестности порога устойчивости исходного состояния).

Мы ограничимся этим беглым перечнем результатов анализа математической модели (6), поскольку чисто формально данная модель является частным случаем более общего класса моделей «реакция—диффузия» [23], достаточно освещенных в литературе [23—27].

В целом можно утверждать, что математическая модель, к которой привели постулированные на основе эксперимента механохимические свойства клеточного материала, способна охватить широкое многообразие типов макроскопической надклеточной организации. Это делает рассмотренный механизм весьма вероятным источником первичной пространственной организации (разметки) эпителиальных морфогенезов.

В качестве иллюстрации приведем результаты расчетов, выполненных на основании модели с целью интерпретации конкретных экспериментальных данных.

Наглядную демонстрацию проявления полевых черт морфогенеза дают эксперименты по внешней стимуляции почкования у гидроидов [6]. При индуцированном почковании первичным откликом на внешний стимул (например, укол иглой) является выделение во внешнем эпителиальном слое клеток ствола домена столбчатых клеток. При двухцентральной стимуляции в эксперименте регистрируется нетривиальная, «интерференционная» картина, в которой локализация внешнего стимула к развитию боковой почки может не совпадать с локализацией морфогенетического ответа (рис. 6,  $\theta$ ). Результаты серии подобных экспериментов, схематически представленные на рис. 6, указывают на существенную автономность морфогенетического отклика по отношению к индуцирующему внешнему сигналу (разметка путем самоорганизации групповой клеточной активности). В них также явно выражено

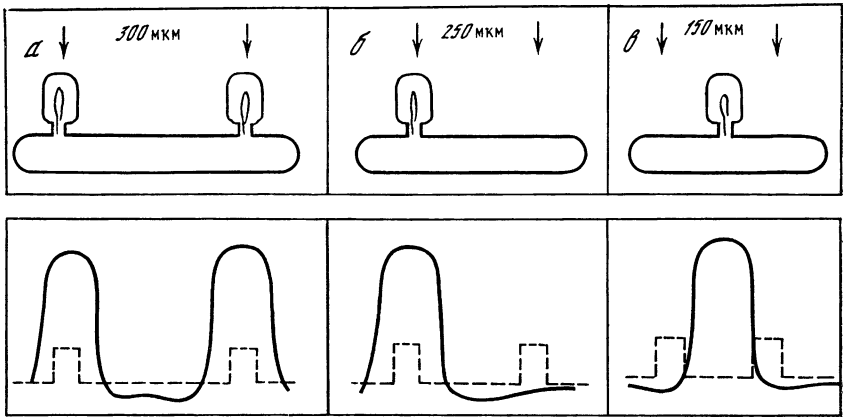


Рис. 6. Схема опытов Пликерта [6] по индукции бокового почкования у гидроидов

Вверху *a* — каждый из двух достаточно удаленных стимулов индуцирует отдельную почку; *b* — при меньшем удалении стимулов только один из них индуцирует формирование почки; *c* — близко расположенные стимулы индуцируют один общий рудимент, расположенный между точками внешнего воздействия. Внизу — объяснение экспериментов Пликерта в рамках математической модели (6) (показаны рассчитанные асимптотические профили  $P(x)$ , порождаемые различными начальными возмущениями  $P(x, 0)$  (пунктир); при расчете использованы значения параметров в безразмерных единицах:  $\tau = 5 \cdot 10^{-2}$ ,  $D_p = 5 \cdot 10^{-4}$ ,  $\alpha_0 = 50$ ,  $R_e = 0,1$ ,  $f = 0,5 P + 15 P^2 - 10 P^3$ )

«поле подавления» в окрестности формирующегося зачатка бокового побега. Как видно из показанных расчетных кривых (см. рис. 6, внизу), приведенные экспериментальные данные вполне удовлетворительно интерпретируются моделью.

Еще одна типичная черта «полевого» поведения — инвариантность пропорций зародыша при варьировании в эксперименте его размеров, как следует из более детальных наблюдений [3], также связана с актом первичной разметки морфогенеза. Соответствующая регуляция в ходе разметки после удаления части клеточного материала, как выяснилось, осуществляется за счет пропорционального изменения размера зоны столбчатых клеток в бластуле. Объяснение этого эффекта легко получить, анализируя поведение пласта согласно модели (8).

При условии  $R_p \ll L$  стационарное решение уравнения (8) имеет вид доменного распределения со значениями параметра  $P$  в областях плато, соответствующими двум корням уравнения  $(1/\tau)\dot{f}(P) - (1 + \sigma)\alpha_0 P = 0$ . Таких распределений может быть множество (с разным числом и размерами доменов), но для них всех величина  $\langle P \rangle$  одина и не зависит от размеров пласта [9]. Существенно, что из данного множества стационарных решений единственным асимптотически устойчивым является состояние лишь с одним доменом. Для последнего размерная инвариантность величины  $\langle P \rangle$  означает одновременно масштабную инвариантность геометрических пропорций паттерна  $P(r)$  [10]. Именно таким об-

разом, мы считаем, интерпретируется одно из характерных выражений дришевских регуляций.

Можно привести еще ряд примеров, иллюстрирующих описательную силу предложенной модели самоорганизации разметки в эмбриональных эпителиях (см. [11]). В целом можно считать, что важный аспект пространственной регуляции морфогенеза — те черты самоорганизации надклеточного порядка в зародышевом развитии, которые приписывают «морфогенетическому полю», — удается трактовать на базе более элементарных свойств клеточного материала. Главным из этих свойств является механохимическая активность клеток, выражающаяся в их способности к активным деформациям.

#### ЛИТЕРАТУРА

1. Дриш Г. Витализм, его история и система. М., 1915.
2. Белоусов Л. В. Введение в общую эмбриологию. М.: Изд-во МГУ, 1980.
3. Белоусов Л. В., Богдановский С. Б. Клеточные механизмы эмбриональных регуляций у морских ежей // Онтогенез. 1980. Т. 11. С. 467—475.
4. Wolpert L. Positional information and the spatial pattern of cellular differentiation // J. Theor. Biol. 1969. Vol. 25. P. 1—45.
5. Зусман М. Биология развития. М.: Мир, 1977.
6. Plickert G. Mechanically induced stolon branching in *E. Viricula* // Development and cell biology of coelenterates / Ed. P. Tardent, R. Tardent. Amsterdam: Elsevier; North-Holland, 1980.
7. Wessels N. K. A catalogue of processes responsible for metazoan morphogenesis // Evolution and development / Ed. J. T. Bonner. Berlin; Heidelberg: Springer, 1982.
8. Черданцев В. Г. Пространственные развертки морфогенетических клеточных движений как элементы ротового поля бесхвостых амфибий // Онтогенез. 1977. Т. 8. С. 335—360.
9. Лучинская Л. В., Белоусов Л. В. Электронно-микроскопическое исследование межклеточных взаимодействий в эксплантатах эмбриональных тканей // Цитология. 1977. Т. 25. С. 939—944.
10. Белинцев Б. Н., Белоусов Л. В., Зарайский А. Г. Модель эпителиальных морфогенезов на основе упругих сил и контактной поляризации клеток. I // Онтогенез. 1985. Т. 16. С. 5—14.
11. Белинцев Б. Н., Белоусов Л. В., Зарайский А. Г. Модель эпителиальных морфогенезов на основе упругих сил и контактной поляризации клеток. II. Биологические следствия // Там же. С. 437—449.
12. Белоусов Л. В., Дорфман Я. Г., Черданцев В. Г. Быстрые изменения формы и клеточной архитектуры изолированных фрагментов эмбриональной ткани как экспериментальная модель морфогенеза // Там же. 1974. Т. 5. С. 323—333.
13. Петров К. В., Белоусов Л. В. Кинетика контактной поляризации клеток в индуцированных тканях // Там же. 1984. Т. 15. С. 643—648.
14. Kolega J. Effects of mechanical tensions on protrusive activity and microfilament and intermediate filament organization in an epidermal epithelium moving in culture // J. Cell Biol. 1986. Vol. 102. P. 1400—1409.
15. Belintsev B. N., Belousov L. V., Zaraisky A. G. Model of pattern formation in epithelial morphogenesis // J. Theor. Biol. 1987.
16. Ландау Л. Д., Лифшиц Е. М. Статистическая физика. М.: Наука, 1976.
17. Белинцев Б. Н., Савин Д. Физический механизм, определяющий пространственную организацию при эпителиальном формообразовании // Биофизика. 1985. Т. 30. С. 284—291.
18. Ландау Л. Д., Лифшиц Е. М. Теория упругости. М.: Наука, 1965.
19. Белинцев Б. Н. Элементарные процессы формирования надклеточной организации при морфогенезе: Автореф. дис. ... д-ра биол. наук. М., 1986.

20. *Белинцев Б. Н.* Диссипативные структуры и проблема биологического формообразования // УФН. 1983. Т. 141. С. 55—101.
21. *Лившиц М. А.* Возникновение диссипативных структур по типу фазового перехода 1-го рода при вырожденных бифуркациях // Биофизика. 1984. Т. 29. С. 699—705.
22. *Белинцев Б. Н., Лившиц М. А., Волькенштейн М. В.* Пространственная самоорганизация как результат близкойдействующей положительной кооперативности и дальнедействующей отрицательной кооперативности // ДАН СССР. 1981. Т. 257. С. 487—491.
23. *Belintsev B. N., Livshitz M. A., Volkenstein M. V.* Pattern formation in systems with nonlocal interactions // Ztschr. Phys. B. 1981. Vol. 44. P. 345—351.
24. *Кернер Б. С., Осипов В. В.* Нелинейная теория стационарных страт в диссипативных системах // ЖЭТФ. 1978. Т. 74. С. 1675—1686.
25. *Кернер Б. С., Осипов В. В.* Стохастически неоднородные структуры в неравновесных системах // Там же. 1980. Т. 79. С. 2218—2238.
26. *Nitzan A., Ortoleva P.* Scaling and Ginzburg criteria for critical bifurcations in nonequilibrium reacting systems // Phys. Rev. A. 1980. Vol. 21. P. 1735—1755.
27. *Kuramoto Y., Tsuzuki T.* On the formation of dissipative structures in reaction-diffusion systems // Progr. Theor. Phys. 1975. Vol. 54. P. 687—689.

УДК 530.1+517.9

## ДРЕЙФ РЕВЕРБЕРАТОРА В АКТИВНОЙ СРЕДЕ ПРИ ВЗАИМОДЕЙСТВИИ С ГРАНИЦАМИ

В. Н. БИКТАШЕВ

### ВВЕДЕНИЕ

Одно из характерных свойств, присущих активным средам, в которых могут распространяться автоволны (например, сердечной ткани или среде с реакцией Белоусова—Жаботинского), — это способность таких сред содержать в себе и источники автоволн. Типичными для двумерных сред являются источники вращающихся спиральных волн, называемые также вихрями или ревербераторами.

Существование ревербератора может быть не связано с неоднородностями среды, но обеспечено лишь начальными условиями [1—5]. Такой источник ведет себя как выделенный долгоживущий объект, навязывающий окружающей среде присущий ему ритм и обладающий известной устойчивостью по отношению к внешним воздействиям [6—8]. Небольшие внешние воздействия лишь немного изменяют характер автономной эволюции ревербератора. Степень этого воздействия быстро убывает с удалением «точки приложения» внешних сил от специфической области вблизи центра вращения спиральной волны, называемой ядром ревербератора.

Рассмотрим случай простейшей эволюции ревербератора, когда в однородной бесконечной среде ревербератор излучает вращающуюся с постоянной скоростью вокруг неподвижного центра

спиральную волну. Тогда влияние границ среды, небольших неоднородностей или других источников автоволн обычно приводит к изменению частоты и смещению положения ревербератора с течением времени (т. е. к дрейфу положения и фазы).

Представления о механизмах такого воздействия в настоящий момент недостаточны и носят феноменологический характер. Для построения теории полезно исследование простейших случаев, допускающих аналитические подходы. В данной работе для задачи о дрейфе ревербератора развивается формальный подход «поле — частица», при котором ревербератор рассматривается как точечная частица, взаимодействующая с окружающей средой, а эволюция окружающей среды описывается «полевым» уравнением, в котором устранены быстрые осцилляции.

## 1. МОДЕЛЬ. СВОБОДНЫЙ РЕВЕРБЕРАТОР

Выбрана одна из простейших моделей автоволновых сред, относящаяся к так называемым  $\Lambda - \Omega$ -системам [9], описывающим среды из диффузионно связанных нелинейных осцилляторов:

$$\dot{u} = (1 - i\Omega)u - (1 - i\alpha)u|u|^2 + \nabla^2 u; \quad (1)$$

$$u = u(\mathbf{r}, t) \in \mathbb{C}, \quad \alpha, \Omega \in \mathbb{R}, \quad \mathbf{r} \equiv (x, y), \quad \nabla \equiv (\partial/\partial x, \partial/\partial y), \\ \dot{\phantom{u}} \equiv \partial/\partial t.$$

Это уравнение является также частным случаем обобщенного уравнения Гинзбурга—Ландау [10]. Переменная  $u$  имеет смысл комплексной амплитуды колебаний: скорость изменения фазы этой переменной определяет частоту колебаний, а градиент фазы соответствует локальному волновому числу. Замена  $\alpha \rightarrow -\alpha$  эквивалентна замене  $\Omega \rightarrow -\Omega$ ,  $u < u^*$  (звездочка означает комплексное сопряжение), для определенности пусть  $\alpha > 0$ .

В рамках этой модели Хаганом [3] был особо исследован предельный случай  $\alpha \rightarrow 0$ , который оказался достаточно простым, чтобы получить асимптотическое представление для решений типа стационарно вращающегося ревербератора:

$$u^0 = \exp(im\Phi - i\omega^0 t) U(\rho). \quad (2)$$

Здесь  $\rho, \Phi$  — полярные координаты (считаем, что центр вращения находится в начале координат);  $U$  — комплексная функция;  $m \neq 0$  — «топологический заряд» (число рукавов спирали), для определенности  $m > 0$ . Это представление удовлетворяет условию

$$d \arg(U)/d\rho \rightarrow k^0, \quad \rho \rightarrow \infty; \quad \omega^0 = \Omega - \alpha + \alpha(k^0)^2, \quad (3)$$

т. е. при больших радиусах линии равной фазы близки к архимедовым спиральям, причем

$$k^0 \approx (2/\alpha) \exp[-\pi/2m\alpha + C'(m)] \quad (4)$$

однозначно определено при каждом  $\alpha$ , т. е. является собственным числом задачи; здесь  $C'(m)$  — константы, определяемые численно. Это решение будем называть свободным ревербератором.

## 2. ПОЛЕВОЕ УРАВНЕНИЕ

Представим уравнение (1) в терминах амплитуды и фазы  $a$ ,  $\varphi$ ,  $u = a \exp(i\varphi)$ :

$$\dot{a} = [1 - a^2 - (\nabla\varphi)^2]a + \nabla^2 a;$$

$$\dot{\varphi} = -\Omega + \alpha a^2 + [\nabla(a^2)\nabla\varphi]/a^2 + \nabla^2\varphi.$$

Если производные  $\nabla a$  и  $\nabla^2\varphi$  и параметр  $\alpha$  малы, то переменная  $a$  является быстрой по сравнению с  $\varphi$ , ее адиабатическое исключение приводит к замкнутому уравнению для  $\varphi$ :

$$\dot{\varphi}_* = -\Omega + \alpha - \alpha(\nabla\varphi)^2 - \frac{\nabla\varphi\nabla(\nabla\varphi)^2}{1 - (\nabla\varphi)^2} + \nabla^2\varphi.$$

Если оператор  $|\nabla\varphi|$  также мал (что действительно имеет место в данной модели для свободного ревербератора при  $\alpha \rightarrow 0$ ), то четвертым членом в правой части можно пренебречь. Остающееся уравнение заменой

$$\varphi = -(1/\alpha)\ln\psi - (\Omega - \alpha)t$$

приводится к линейному полювому уравнению

$$\dot{\psi} = \nabla^2\psi \tag{5}$$

(этот вывод предложен в работе [11]). Нас интересуют не совсем обычные (неограниченно растущие) решения этого уравнения. Например, решению уравнения (1) в виде плоской волны с волновым числом  $k$  отвечает функция  $\psi$  вида  $\psi = \exp(\alpha^2 k^2 t - \alpha k x)$ . Кроме того, при наличии в среде замкнутых контуров, при обходе которых  $\varphi$  меняется на  $2\pi m$ ,  $m \neq 0$ , функция  $\psi$  также будет неоднозначной — при обходе таких контуров она будет увеличиваться (уменьшаться) в  $\exp(2\pi m \alpha)$  раз.

Взаимодействие волн с непроницаемой границей можно описывать в рамках уравнения (5). Так, падающая из точки  $x \rightarrow +\infty$  волна при условии  $\partial u/\partial x = 0$ ,  $x = 0$  описывается функцией  $\psi = \exp(\alpha^2 k^2 t) \operatorname{ch}(\alpha k x)$ .

## 3. АДИАБАТИЧЕСКОЕ ПРИБЛИЖЕНИЕ

Пусть теперь задана автоволновая среда в ограниченной (для простоты — конечной) области с условиями непроницаемости на границах:

$$(\mathbf{n}\nabla)u|_{\Gamma} = 0, \tag{6}$$

где  $\mathbf{n}$  — нормаль к границе  $\Gamma$ . Для описания дрейфа рассмотрим систему отсчета, движущуюся по траектории  $\mathbf{r} = \mathbf{r}^0(t)$  относи-

тельно лабораторной. В этой системе отсчета вместо (1), (6) имеем

$$\dot{u} = (1 - i\Omega)u - (1 - i\alpha)u|u|^2 + \nabla^2 u + \left(\frac{dr^0}{dt}\nabla\right)u, \quad (7)$$

$$(\mathbf{n}\nabla)u|_{(\mathbf{r}+\mathbf{r}^0(t))\in\Gamma} = 0 \quad (8)$$

$$\text{или } (\mathbf{n}'(t)\nabla)u|_{\mathbf{r}\in\Gamma'(t)} = 0 \quad (8')$$

— краевую задачу с движущимися границами.

Пусть выбранная система отсчета движется вместе с ревербератором, т. е. решение задачи (7), (8) почти периодически и в некоторой окрестности начала координат все время близко к свободному ревербератору (вблизи границ такой близости быть не может — свободный ревербератор не может удовлетворять условию (8)).

Из численных экспериментов [6] видно, что скорость дрейфа быстро убывает с увеличением расстояния между ревербератором и границей. Оценка, сделанная в данной работе, согласуется с этим наблюдением — скорость убывает экспоненциально.

С другой стороны, естественно ожидать, что с увеличением размеров области время установления какого-либо устойчивого стационарного режима в ней будет расти не быстрее, чем по степенному закону (см., например, (5): известно, что для уравнения теплопроводности это время пропорционально квадрату размеров среды).

В этом случае при достаточно больших расстояниях между ревербератором и границей граница  $\Gamma'$  в выражении (8') не успевает сдвинуться сколь-нибудь существенно за время установления стационарного режима, соответствующего текущему положению границы  $\Gamma'$ .

Для этого стационарного режима имеем задачу

$$(1 - i\Omega)u - (1 - i\alpha)u|u|^2 + \nabla^2 u + (\mathbf{c}\nabla)u = -i\omega u, \quad (9)$$

$$(\mathbf{n}\nabla)u|_{\Gamma'} = 0, \quad (10)$$

или в тригонометрической форме:

$$\nabla^2 a + (1 - a^2 - (\nabla\varphi)^2)a = -(\mathbf{c}\nabla)a, \quad (9')$$

$$a\nabla^2\varphi + 2(\nabla a\nabla\varphi) = a(\Omega - \omega - \alpha a^2 - (\mathbf{c}\nabla)\varphi),$$

$$(\mathbf{n}\nabla)a|_{\Gamma'} = 0, \quad (\mathbf{n}\nabla)\varphi|_{\Gamma'} = 0. \quad (10')$$

Вектор скорости  $\mathbf{c}$  и частота  $\omega$  оказываются собственными числами: при заданных  $\alpha$  и  $\Gamma'$  они однозначно определены. Согласно [3, 10], свободный ревербератор устойчив при малых  $\alpha$  при  $m = 1$ ; будем предполагать, что стационарный режим (9), (10) также устойчив. Наоборот, стационарные ревербераторы при  $m > 1$  неустойчивы, так что достаточно рассмотреть случай  $m = 1$ ; однако обобщение для  $m > 1$  не вызывает трудностей. Далее ограничимся лишь анализом задачи (9), (10).

Такое приближение в известной степени аналогично использованию уравнений электростатики в задаче о движении системы



электрических зарядов, когда скорости зарядов много меньше скорости света. В соответствии с этой аналогией условие применимости приближения сведется к условию

$$|c| \ll \alpha k^0,$$

так как волновому числу  $k$  в данной модели соответствует скорость передачи информации  $2\alpha k$ .

#### 4. МЕТОД РЕШЕНИЯ

Решение задачи (9), (10) находится методом спивания асимптотик, если расстояние до границ среды достаточно велико. Будем предполагать, что при не слишком больших  $|r|$  решение близко к свободному ревербератору, а при не слишком малых описывается приближением почти плоских волн. Сравнив соответствующие асимптотические представления в области, где они оба применимы, получим оценки для частоты и скорости дрейфа ревербератора. За недостатком места приводятся лишь основные этапы вычислений (более подробно см. [12]).

**Внутренняя область.** В сделанных предположениях естественным способом решения (9) при небольших  $|r|$  является теория возмущений со свободным ревербератором в качестве нулевого приближения и с малыми параметрами  $c$  и  $\omega^0$ . Используем при этом тот факт, что при небольших  $\rho$  свободный ревербератор (2) — (4), в свою очередь, близок к решению типа (2) уравнения (1) с  $\alpha = 0$  и сам получается из него при использовании теории возмущений по  $\alpha$  [3]. Итак, при небольших  $\rho$  будем искать решение (9') в виде  $a = a^0 + a'$ ,  $\varphi = \varphi^0 + \varphi'$ ,  $a' \ll a^0$ ,  $\varphi' \ll \varphi^0$ , рассматривая  $\alpha$ ,  $c$  и  $(\omega - \Omega)$  как малые параметры и ограничиваясь линейным приближением по каждому из этих параметров.

Невозмущенное решение описывается [3] функциями  $\varphi^0 = m\Phi$ ,  $a^0 = a^0(\rho) = P(\rho, m)$ , где  $P(\rho, m)$  — решение уравнения Гринберга [13]:

$$\frac{d^2 a^0}{d\rho^2} + \frac{1}{\rho} \frac{da^0}{d\rho} + a^0 \left( 1 - a^{02} - \frac{m^2}{\rho^2} \right) = 0, \quad (11)$$

причем

$$P(\rho, m) \approx 1 - m^2/2\rho^2, \quad \rho \rightarrow \infty; \quad P(\rho, m) \approx \varepsilon(m)\rho^m, \quad \rho \rightarrow 0,$$

Исследование уравнений для поправок  $a'$ ,  $\varphi'$  в предположении, что  $a'$ ,  $\varphi'$  растут не слишком быстро при  $\rho \rightarrow \infty$  (что необходимо, чтобы полученное представление можно было сшить с внешним), приводит к следующему. При

$$1 \ll \rho \ll \exp(\sqrt{2/\alpha m^2})$$

имеет место

$$\begin{aligned} \varphi(\rho, \Phi) \approx m\Phi + \alpha m^2 \ln \rho [^{1/2} \ln \rho + C(m)] + \\ + c m \rho [^{1/2} \ln \rho + B(m)] \sin(\Phi - \theta) - \sum_{n=2}^{\infty} 2\rho^n \operatorname{Im}(C_n e^{in\Phi}), \quad (12) \end{aligned}$$

и аналогично для  $a(\rho, \Phi)$  с точностью до аддитивных добавок, имеющих смысл произвольных малых сдвигов невозмущенного решения в пространстве и времени. Здесь  $c = (c \cos \theta, c \sin \theta)$ ;  $C_n$  — произвольные (не слишком большие) коэффициенты;  $B(m)$ ,  $C(m)$  — константы, которые могут быть найдены численно. Согласно [3],

$$C(1) \approx -0,098; \quad C(2) \approx -0,998; \quad C(3) \approx -1,3; \dots \quad (13)$$

Автором были оценены также и константы  $B(m)$ :

$$B(1) \approx -0,31, \quad B(2) \approx -0,78, \quad B(3) \approx -1,01, \dots \quad (14)$$

Внешняя область. Задача (9), (10) приводит к уравнению для «полевой» переменной:

$$\nabla^2 \psi + (c\nabla)\psi - \alpha^2 k^2 \psi = 0. \quad (15)$$

Здесь  $k$  — волновое число, соответствующее частоте  $\omega = \Omega - \alpha + \alpha k^2$ . С учетом неоднозначности  $\psi$  общее решение (15) имеет вид

$$\begin{aligned} \psi = \exp[-(c\rho)/2 - \alpha m \Phi] \sum_{n=-\infty}^{+\infty} \exp(in\Phi) \times \\ \times [G^n K_{n+i\alpha m}(\alpha k \rho) + F^n I_{n+i\alpha m}(\alpha k \rho)] \end{aligned} \quad (16)$$

с произвольными  $F^n$ ,  $G^n$ . Здесь  $k^2 = k^2 - c^2/4\alpha$ ;  $K_{n+i\alpha m}$ ,  $I_{n+i\alpha m}$  — линейно независимые решения модифицированного уравнения Бесселя с комплексным индексом  $n + i\alpha m$ , причем

$$\begin{aligned} K_\mu(z) \approx \frac{1}{\sqrt{z}} e^{-z}, \quad I_\mu(z) \approx \frac{1}{\sqrt{z}} e^z, \quad z \rightarrow \infty, \quad I_\mu(z) = O(1), \quad z \rightarrow 0; \\ K_\mu \equiv K_{-\mu}, \quad I_\mu \equiv I_{-\mu}, \quad (K_\mu^*)^* \equiv K_\mu, \quad (I_\mu^*)^* \equiv I_\mu. \end{aligned}$$

Вещественность ряда (16) требует

$$G^{-n} \equiv (G^n)^*; \quad F^{-n} \equiv (F^n)^*.$$

Для свободного ревербератора все  $F^n$ ,  $G^n$  равны нулю, кроме  $G^0$ , который можно положить равным единице. Чтобы при небольших  $\rho$  движущийся ревербератор слабо отличался от свободного, необходимо выполнение неравенств  $G^n \ll 1$ ,  $n \neq 0$ . В остальном  $G^n$ ,  $F^n$  следует выбрать исходя из граничных условий

$$(n\nabla)\psi|_{\Gamma'} = 0. \quad (17)$$

Проще всего это сделать, если граница  $\Gamma'$  задается уравнением  $\rho = R(\Phi)$ , причем  $R$  удовлетворяет неравенствам

$$\alpha k R \gg 1, \quad \frac{1}{R} \frac{dR}{d\Phi} \ll \alpha k R.$$

Тогда после несложных выкладок получаем, что

$$F^n = \frac{1}{2\pi} \int_0^{2\pi} \exp(-in\Phi) \exp[-2\alpha k R(\Phi)] / d\Phi. \quad (18)$$

Ниже мы увидим, что  $F^n$  играют роль сил, действующих на ревербератор и приводящих его к дрейфу. То, что эти «силы» определяются в основном уже в нулевом порядке по  $s$  (отличие  $k$  от  $k^0$  несущественно), делает предположение об адиабатичности поля ревербератора самосогласованным.

**Сшивки внутреннего и внешнего представлений.** Сопоставим в области (11) два представления для решения (9'), (10') в виде (12) и (16), переписав (16) для  $\varphi$ . Сшивку можно производить отдельно по каждой угловой моде.

Сопоставление по нулевой моде показывает, что соответствие между внутренним и внешним представлениями может быть достигнуто при условии

$$k \approx k^0 (1 + \pi F^0) \quad (19)$$

или для частоты

$$\omega \approx \omega^0 + 2\pi\alpha k^2 F^0,$$

где  $k^0$  — волновое число свободного ревербератора ([3], ср. (4)):

$$k^0 \approx \frac{2}{\alpha} \exp \left[ -\frac{\pi}{2m\alpha} + C(m) - \gamma \right],$$

$\gamma \approx 0,5771\dots$  — постоянная Эйлера—Маскерони.

Сопоставление по первой моде, в свою очередь, дает необходимое условие

$$ce^{-i\theta} \approx -2\pi i m \alpha^2 k \{1 + im\alpha[C(m) - 1 - 2B(m)]\} F^1. \quad (20)$$

Остается убедиться, что сопоставление по остальным угловым модам приводит к определению произвольных с точки зрения внутреннего представления постоянных  $C_n$  через граничные условия.

Полученные уравнения движения (19), (20) решают поставленную задачу: определяют дрейф фазы и положения ревербератора в зависимости от его расположения по отношению к границам среды. Эта зависимость входит в них через посредство обобщенных сил  $F^0$ ,  $F^1$ , которые могут быть найдены из решения краевой задачи для полевого уравнения (15), (17), в простых случаях — по формулам (18).

**Пример.** Пусть граница  $\Gamma^1$  представляет собой окружность радиусом  $S \gg (\alpha k)^{-1}$ , а ядро ревербератора смещено относительно центра окружности на расстояние  $s$  под углом  $\sigma$  к оси  $x$ , причем  $\alpha ks \lesssim 1$ . Так как  $s \ll S$ , приближенно имеем

$$R(\Phi) \approx S - s \cos(\Phi - \sigma).$$

Подставляя это в формулы (18)—(20), получим, что частота ревербератора в круговой области

$$\omega \approx \omega^0 + 2\pi\alpha k^2 \exp(-2\alpha ks) I_0(2\alpha ks),$$

а его дрейф можно описать уравнениями движения

$$sd\sigma/dt \approx 2\pi m\alpha^2 k \exp(-2\alpha kS) I_1(2\alpha ks),$$

$$ds/dt \approx -2\pi m^2\alpha^3 k B'(m) \exp(-2\alpha kS) I_1(2\alpha ks).$$

Здесь  $I_0, I_1$  — модифицированные функции Бесселя индексов 0 и 1;  $\omega_1^0$  — частота свободного ревербератора, а коэффициенты  $B'(m)$  выражаются через  $C(m), B(m)$  (см. (13), (14)) по формуле

$$B'(m) = 1 + 2B(m) - C(m)$$

и, по крайней мере для первых  $m$ , положительны:

$$B'(1) \approx 0,48, \quad B'(2) \approx 0,44, \quad B'(3) \approx 0,3.$$

Итак, в рассматриваемой модели ревербератор отталкивается от границы, однако скорость отталкивания много меньше скорости дрейфа вдоль границы (отношение порядка  $\alpha$ ).

**Обсуждение.** Получены оценки параметров взаимодействия ревербератора с границами среды в простой модели.

Основным обстоятельством, обеспечивающим возможность применения асимптотических методов, является различие в характере процессов, протекающих в окрестности ядра ревербератора и в остальной среде.

В выбранной модели это различие особенно наглядно. Функция  $\psi$  может быть истолкована как потенциал короткодействующего поля, источником которого является область вблизи ядра ревербератора и которое, отразившись от границ, оказывает обратное влияние на ревербератор и обуславливает его дрейф (однако потенциалом эта функция может быть названа лишь условно, поскольку она многолистка).

Результатом является описание движения ревербератора в терминах системы уравнений «поле—частица». При этом поле  $\psi$  описывается линейным уравнением (5), а роль сил, действующих на частицу-ревербератор и определяющих его пространственный и временной дрейф, выполняют коэффициенты  $F^0, F^1$ , описывающие деформацию поля  $\psi$  в окрестности ядра (несмотря на линейность полевого уравнения, принцип суперпозиции к нему применим с оговорками в силу многолистности  $\psi$ ; по этой же причине силы не могут быть выражены через обычные производные внешнего поля, а определяются более сложным образом).

К сожалению, простые оценки показывают, что непосредственная проверка полученных уравнений движения численным экспериментом должна быть весьма непроста вследствие экспоненциальных зависимостей от малого параметра  $\alpha$ . Так, для справедливости сделанных в работе предположений размеры модельной среды должны быть не меньше, чем характерное расстояние, на котором ослабевает влияние границ. Однако уже при  $\alpha \approx 0,2$  это расстояние, согласно сделанным оценкам, должно примерно в  $10^3$ — $10^4$  раз превышать размеры ядра ревербератора. По-видимому, реальным путем обхода этих трудностей может явиться обоб-

щение теории на случай не слишком больших расстояний до границы или на случай конечных  $\alpha$ .

Автор благодарен Б. А. Маломеду и Э. Э. Шнолю за полезные обсуждения и постоянное внимание к работе.

#### ЛИТЕРАТУРА

1. Перцов А. М., Панфилов А. В. Спиральные волны в активных средах. Ревербератор в модели Фитц Хью—Нагумо // Автоволновые процессы в системах с диффузией. Горький: ИПФ АН СССР, 1981. С. 77—91.
2. Cohen D. S., New J. C., Rosales R. R. Rotating spiral wave solution of reactions-diffusion equations // SIAM J. Appl. Math. 1978. Vol. 35, N 3. P. 536—547.
3. Hagan P. S. Spiral waves in reaction-diffusion equations // Ibid. 1982. Vol. 42, N 4. P. 762—786.
4. Kuramoto Y., Koga S. Turbulized rotating chemical wave // Progr. Theor. Phys. 1981. Vol. 66, N 3. P. 1081—1083.
5. Mikhailov A. S., Krinsky V. I. Rotating spiral waves in excitable media: The analytical results // Physica. 1983. Vol. 90. P. 346—371.
6. Ермакова Е. А., Перцов А. М. Взаимодействие вращающихся спиральных волн с границей // Биофизика. 1986. Т. 31, вып. 5. С. 855—861.
7. Руденко А. Н., Панфилов А. В. Дрейф и взаимодействие вихрей в двумерной неоднородной активной среде // Stud. biophys. 1983. Vol. 98, N 3. P. 183—186.
8. Фаст В. Г. Исследование автоволновых источников в ткани сердца методом компьютерной реконструкции: Дис. ... канд. физ.-мат. наук. М.: МФТИ, 1986.
9. Kopell N., Howard L. N. Plane wave solutions to reaction-diffusion equations // Stud. Appl. Math. 1973. Vol. 52, N 4. P. 291—328.
10. Kuramoto Y., Tsuzuki T. On the formation of dissipative structures in reaction-diffusion systems. Reduction perturbation approach // Progr. Theor. Phys. 1975. Vol. 54, N 3. P. 687—699.
11. Kuramoto Y., Tsuzuki T. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium // Ibid. 1976. Vol. 55, N 2. P. 356—369.
12. Бикташев В. Н. Дрейф ревербератора в активной среде при взаимодействии с границами: Препр. НЦБИ АН СССР. Пущино, 1987.
13. Greenberg J. M. Axi-symmetric, time-periodic solutions for lambda-omega system // SIAM J. Appl. Math. 1978. Vol. 35, N 1. P. 1—16.

УДК 577.3+772.2+773.3

### ФОТОХРОМНЫЕ И НЕЛИНЕЙНЫЕ ПРОЦЕССЫ В ПЛЕНКЕ «БИОХРОМ-БР»

Н. Н. ВСЕВОЛОДОВ

Создание технических элементов и систем на основе биологических материалов — одно из перспективных направлений развития современной науки и техники. Микроустройства, изготовленные на основе биоматериалов природного происхождения, могут играть роль управляющих элементов, обрабатывать и преобразовывать информацию, служить приемными и исполнительными устройствами [1]. Среди белковых систем живой природы важное

место занимают биологические светочувствительные комплексы как преобразователи световой энергии.

Попытки использовать их в качестве фоторегистраторов предпринимались ранее. Например, создать фоточувствительную пленку из зрительного родопсина или растительного или бактериально-хлорофилла в принципе возможно, но сложность, дороговизна, а главное, неустойчивость таких фотоматериалов к внешним воздействиям не привели к их внедрению в современную фототехнологию [2, 3].

Положение изменилось с появлением в 1973 г. первых работ по исследованию фотохимических процессов в бактериородопсине (БР) — ретинальбелковом комплексе, выделенном из галофильных микроорганизмов. В 1979 г. в Институте биологической физики АН СССР была создана первая фотохромная пленка, параметры которой обладали некоторым преимуществом перед традиционными. Позже новый фотохромный материал был назван «Биохром-БР»<sup>1</sup>.

Оказалось, что «Биохром» обладает способностью записывать оптическую информацию непосредственно в процессе экспозиции в постоянном и динамическом (реверсивном) режимах с высоким разрешением и различными способами (побитовым, голографическим, поляризационным). Он может применяться в современных оптико-механических системах записи, хранения и обработки информации. Недавно показано, что «Биохром-БР» может оказаться перспективной нелинейной средой для использования его в системах нелинейной оптики.

Основные результаты по исследованию нелинейных процессов получены на пленках из БР, поэтому кратко опишем структуру и свойства этого белка.

Около 15 лет назад в мембранной оболочке бактерий, живущих в соленых озерах планеты, были обнаружены участки пурпурного цвета. Эти участки, поглощая солнечный свет, преобразуют его энергию в энергию жизнедеятельности клетки. Они относительно легко выделяются из биомассы таких микроорганизмов и ресуспендируются в воде или высушиваются для получения порошка, который может храниться годами при нормальных условиях.

Каждый участок (фрагмент) состоит из нескольких десятков тысяч молекул бактериородопсина, упакованных в кристаллический монослой, и фактически представляет собой двумерный белковый кристалл. Такие фрагменты названы пурпурными мембранами (ПМ).

На рис. 1 показан схематический разрез бактерии. Почти все работы проводятся с фрагментами ПМ, так как при разрушении кристаллической структуры молекулы БР денатурируют. Однако

<sup>1</sup> Биохромным материалом предложено называть любой фотохромный материал из белковых молекул или комплексов природного происхождения. Добавка БР означает, что основа данного материала — бактериородопсин; для хлорофилла будет «Биохром-Хл» и т. д.

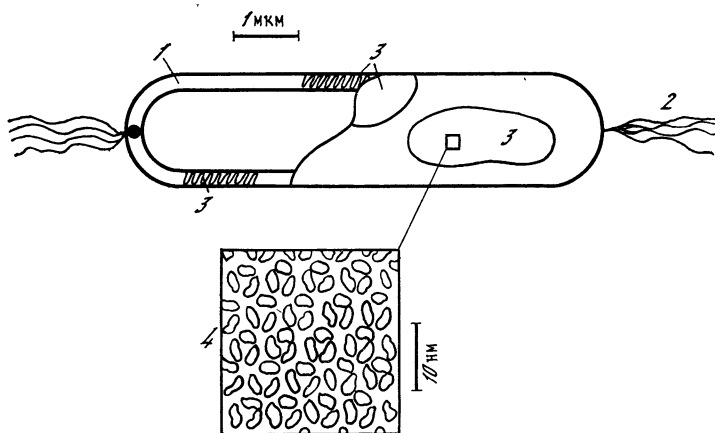


Рис. 1. Упрощенная схема-разрез соленлюбивой бактерии

1 — оболочка-мембрана, 2 — флагеллы (для перемещения), 3 — пурпурные мембраны (ПМ), 4 — увеличенный участок ПМ (хорошо видна гексагональная упаковка молекул бактериородопсина)

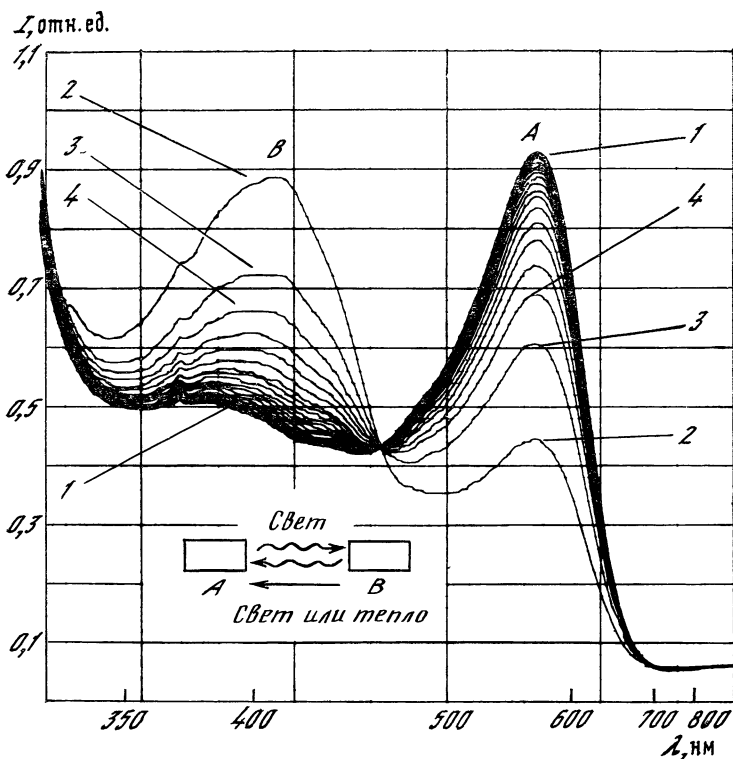
в обычных условиях ПМ необычайно устойчивы, сохраняют свои физико-химические, фотохимические и другие свойства вне клетки годами.

Одно из этих свойств — фотохромизм и сопутствующая ему фотоанизотропия. Именно эти свойства использованы в пленках «Биохром», которые представляют собой нанесенный на подложку практически из любого материала (стекло, лавсан и т. д.) ПМ, смешанный с водорастворимым полимером, прозрачным в видимой области спектра (желатин, поливиниловый спирт и др.). Спектр поглощения такой пленки приведен на рис. 2. Под действием света исходный спектр смещается в синюю область, а после окончания действия света самопроизвольно возвращается в исходное состояние. Время возвращения определяется составом полимера и предысторией ПМ. Его можно ускорить синим светом, т. е. стереть изображение или точку, записанную желтым светом. Это чисто фотохромный процесс.

При 50° С процесс возврата в исходную форму для БР прекращается, и изображение хранится бесконечно долго. Кинетика переходов из формы А в В, и наоборот, приведена на рис. 3.

Исследования показали, что «Биохром-БР» как фотохромный материал обладает некоторыми преимуществами перед известными ранее традиционными фотохромами (см. таблицу). Основные из них — высокая цикличность и близкая к теоретической чувствительность.

Необходимо отметить, что существуют биохимические и биотехнологические способы модификации БР, в результате которых исходный спектр смещается, но фотохимические свойства не изменяются и способность к фотохимии остается. Таким способом

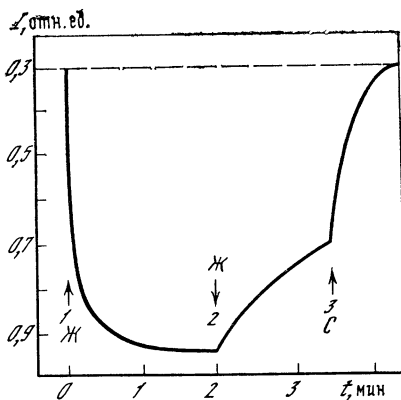


**Рис. 2.** Спектры поглощения пленки «Биохром-БР»

1 — исходный (до засветки), 2 — после 1 с желтого света (в темноте), 3, 4 и т. д. — постепенное восстановление исходного спектра. Время восстановления зависит от типа «Биохрома» и времени экспозиции. А и В — исходная и фотоиндуцированная спектральные формы соответственно

**Рис. 3.** Кинетика образования и исчезновения формы В

Измерение в области 410 нм. 1 — включение желтого света, 2 — его выключение, 3 — включение синего света



можно создавать «Биохромы» различного цвета. Времена записи и стирания ограничены и для «Биохрома-БР» составляют единицы микросекунд в области комнатных температур, но при охлаждении они могут достигать единиц пикосекунд и менее.

Механизм фотохромии в молекулах БР не совсем ясен, хотя различных теорий на этот счет существует достаточно. За короткий срок БР и ПМ исследованы довольно подробно большим количе-



**Сравнительные характеристики обратимых несеребряных регистрирующих фотоматериалов (фотохромы) пленочного типа и пленки «Биохром»**

Характеристика «идеального» фотохрома	Спиропираны	Датизонаты Ртути	Самцилден-анилин	Тионингоиды	Феноксинафтаценхинон	Фотополимеры	Термопласты	Биохромы
Поглощение в видимой области оптического диапазона 400—700 нм	—	—	—	+	—	—	—	+
Светочувствительность, близкая к теоретической, равной $10^{-3}$ Дж/см <sup>2</sup>	—	—	—	—	—	—	+	+
Разрешающая способность не хуже 5000 мм <sup>-1</sup>	+	+	+	+	+	+	—	+
Число циклов (запись — стирание) не менее 100 000	—	—	+	—	—	+	—	+
Время хранения оперативной информации — часы и больше	—	+	—	—	+	+	+	±
Дифракционная эффективность (для голографии) не менее 10%	+	—	+	—	—	—	+	≤5%
Использование распротраненных лазеров	—	—	—	+	—	—	+	+
Работа и хранение в нормальных условиях не менее 10 лет	+	—	+	+	?	—	—	+
Способность к поляризационной записи	+	—	—	+	—	—	—	+
Возможность к записи цвета	—	—	—	—	—	—	—	+

Пр и м е ч а н и е. «+» — обладает указанным свойством, «—» — не обладает.

ством различных биологических, физико-химических и других методов. Имеются подробные обзоры [4, 5]. Сегодня клетки для получения реальных количеств ПМ культивируются и выращиваются в специальных люмостатах, а методика выделения из них ПМ хорошо налажена во многих лабораториях.

Перейдем к рассмотрению оптических свойств «Биохрома-БР». Известно, что в каждом фрагменте ПМ диполи молекул БР лежат практически в плоскости мембран, а их расположение хаотично, т. е. не имеет избранного направления. С другой стороны, сами мембраны также лежат в основном в одной плоскости. Таким образом, исходно пленка изотропна по отношению к падающему излучению неполяризованного луча. Однако если луч линейно-поляризован и его длина волны попадает в область поглощения исходной формы А, то в пленке наводится оптическая ось, параллельная этой поляризации, т. е. появляется оптическая анизотро-

пия. В этот момент любой другой луч взаимодействует с ней, как с анизотропным материалом.

В одной из первых работ по исследованию нелинейного эффекта Вейгерта в пленках «Биохром-БР» наблюдали зависимость фотоанизотропии от длины волны и ее интенсивности [6]. Наведение фотоанизотропии производилось линейно-поляризованным лазерным излучением последовательно на пяти длинах волн (422, 458, 488, 515, 633 нм), а измерение наведенного дихроизма и двулучепреломления — в области 440—680 нм. Наведенный дихроизм принимает максимальное значение в области 570 нм, а двулучепреломление в этой области падает до нуля при любых значениях длин волн возбуждения.

После выключения возбуждающего света «Биохром-БР» возвращается в изотропное состояние за время порядка времени перехода формы В в А и, таким образом, готов к новому циклу фотоанизотропного наведения. Заметим, что в классическом эффекте Вейгерта наведенная анизотропия не исчезает.

Описанный выше эффект позволяет использовать «Биохром-БР» для записи интерференционной решетки двумя ортогональными по поляризации лучами, т. е. осуществлять поляризационную голографию [7].

Интенсивность света для поляризационной записи в «Биохроме-БР» на два порядка меньше, чем для красителей.

Предложение использовать «Биохром-БР» как нелинейную среду для обращения волнового фронта впервые появилось в работе [8]. В этой работе использовалась схема попутного четырехпучкового взаимодействия. Позже теми же авторами в схеме встречного четырехпучкового взаимодействия было реализовано высококачественное пространственно-поляризационное обращение волнового фронта, эффективность которого для циркулярно-поляризованных вращающихся в разных направлениях двух волн накачки составляла 99% [9].

Как уже отмечалось выше, «Биохром-БР» обладает высокой энергетической чувствительностью, достигающей в некоторых образцах  $5 \cdot 10^{-4}$  Дж/см<sup>2</sup>. При этом в нем отсутствуют необратимые фотохимические процессы даже при очень высоких плотностях энергии падающего излучения, достигающего сотен ватт на квадратный сантиметр. Нелинейная среда с такими характеристиками может быть использована для наблюдения оптической бистабильности и других эффектов самовоздействия в нелинейном интерферометре Фабри—Перо [9]. Для пленок толщиной порядка 100 мкм удельный коэффициент самовращения достигает огромной величины — 2,5 см/мВт.

Помещение «Биохрома-БР» в плечо трехзеркального резонатора аргонового лазера приводило к возникновению оптической бистабильности и возможности управлять генерацией лазера от внешнего маломощного источника света со скоростью переключения 50 мкс [10]. Скорость переключения можно увеличить, если работать при низких температурах.

Способность «Биохромов» к записи динамических голограмм можно использовать в интерферометрии вибрирующих объектов. Один из недостатков голографической интерферометрии заключается в отсутствии адаптации измерительной схемы к паразитным низкочастотным колебаниям. Для устранения этого недостатка используется схема с фоторефрактивными средами. Как правило, это фоторефрактивные кристаллы типа ниобата лития и т. д. Однако такие кристаллы обладают малой чувствительностью и, следовательно, большими временами записи.

Недавно экспериментально показана возможность голографической обработки сигнала фазовой модуляции волоконно-оптического интерферометра с использованием в качестве динамической фоторефрактивной среды пленок «Биохром-БР» [11]. Адаптивность такого интерферометра к низкочастотным флуктуациям значительно повышается, а требуемая для его работы световая мощность на два порядка ниже.

В принципе пленку «Биохром» можно рассматривать как двумерный кристалл или квазикристалл, состоящий из кристаллических доменов — ПМ с нелинейными свойствами. В работе [12] показано влияние фотохромных процессов в молекулах БР на нелинейные свойства пленки «Биохром-БР», и в частности на генерацию второй гармоники (ВГ) импульсного лазера с излучением на длине волны 1064 нм. Генерация отраженной и прошедшей ВГ получена в пленках толщиной 100 мкм. Обнаружено, что при одновременной подсветке пленки в области генерации гармоники слабым ( $0,1-100$  мВт/см<sup>2</sup>) излучением перестраиваемого лазера интенсивность ВГ зависит как от интенсивности, так и от длины волны подсветки. При отсутствии подсветки интенсивность ВГ максимальна. При увеличении интенсивности подсветки в области исходной формы (т. е. в области 570 нм) происходит переход формы А в форму В и интенсивность ВГ падает, при обратном переходе В в А (под действием подсветки в области 412 нм) она возрастает. Происходит переключение системы генерации ВГ. С другой стороны, излучение накачки при 1064 нм не попадает в полосу поглощения БР и не влияет на переключение. Таким образом, при считывании информации, заложенной в ВГ, не происходит ее искажения или стирания.

Такой нелинейно-оптический метод считывания может оказаться перспективным для элементов памяти, если использовать быстрые пикосекундные переходы в БР, которые проявляются при низких и сверхнизких температурах. Такие переходы последовательно проявляются при постепенном нагревании «Биохрома-БР» от гелиевых температур до комнатных [5].

Таким образом, обнаруженные и исследованные в пленках «Биохром-БР» нелинейные эффекты могут быть использованы в различных устройствах как современной техники, так и в перспективной молекулярной биоэлектронике.

## ЛИТЕРАТУРА

1. *Иваницкий Г. Р.* Биологические микроустройства // Вестн. АН СССР. 1984. № 3. С. 118—128.
2. Пат. № 1423991 (Англия).
3. Пат. № 4084969 (США).
4. *Methods in Enzymology.* 1982. Vol. 88. 772 p.
5. *Балашов С. П., Литвин Ф. Ф.* Фотохимические превращения бактериородопсина М.: Изд-во МГУ, 1985. 183 с.
6. *Всеволодов Н. Н., Дюкова Т. В., Корчемская Е. Я.* и др. Нелинейный вейгерт-эффект в пленках Биохром на основе бактериородопсина // Укр. физ. журн. 1984. Т. 24, № 27. С. 1107—1108.
7. *Всеволодов Н. Н., Дюкова Т. В., Корчемская Е. Я.* и др. Светоиндуцированная анизотропия пленок «Биохром» // Светочувствительные биокомплексы и оптическая регистрация информации. Пушино, 1985. С. 209.
8. *Всеволодов Н. Н., Дюкова Т. В., Жердиченко В. В.* и др. Попутное четырехлучковое взаимодействие в пленках Биохром // Там же.
9. *Всеволодов Н. Н., Иваницкий Г. Р., Соскин М. С., Тараненко В. Б.* Пленки Биохром — реверсивная среда для оптической записи // Автометрия. 1986. № 2. С. 41—48.
10. *Баженов В. Ю., Тараненко В. Б., Соскин М. С.* Гигантская нелинейность и оптическая бистабильность на основе обратимых фотопревращений бактериородопсина в пленках Биохром // Тез. XII Вессоюз. конф. по когерентной и нелинейной оптике. М.: Изд-во МГУ, 1985. Ч. 1. С. 325.
11. *Барменков Ю. О., Зосимов В. В., Кожевников И. М.* и др. Регистрация сигнала фазовой модуляции волоконно-оптического интерферометра с помощью динамической голограммы в бактериородопсине // Акуст. журн. 1987. Т. 23, № 3. С. 568—569.
12. *Акципетров О. А., Ахмедиев Н. Н., Всеволодов Н. Н.* и др. Фотохромизм в нелинейной оптике: Фотоуправляемая генерация второй гармоники молекулами бактериородопсина // ДАН СССР. 1987. Т. 293, № 3. С. 592—594.

УДК 577.3

## ИНЖЕНЕРИЯ ДИНАМИЧЕСКИХ СИСТЕМ ДЛЯ РАСПОЗНАВАНИЯ ОБРАЗОВ И ОБРАБОТКИ ИНФОРМАЦИИ

А. С. МИХАЙЛОВ

Последовательный характер выполнения операций в традиционных компьютерах накладывает жесткие ограничения на скорость обработки информации, что препятствует их возможному использованию в задачах искусственного интеллекта. Очевидно, что деятельность мозга связана с параллельной обработкой информации (например, мозг решает примерно за 100 шагов задачи, требующие от традиционного компьютера выполнения нескольких миллионов операций [1]). Однако существует фундаментальное различие между мозгом и любым параллельным компьютером. Мозг не производит никаких вычислений.

Бессмысленно допускать, что мозг кошки, настаивающей в прыжке мышью, решает в считанные доли секунды сложные диф-

ференциальные уравнения, описывающие траекторию. Эволюционное происхождение человеческого мозга предполагает, что его основные механизмы обработки информации должны быть теми же, что и у более примитивных животных, и они не могут сводиться к логическим операциям с математическими величинами (подробнее см. [2]).

Альтернативой вычисления служит аналоговая обработка информации. В своей простейшей форме она представляет собой прямую имитацию. Техника электронной аналоговой имитации широко использовалась в середине нашего столетия, до изобретения цифровых ЭВМ. Чтобы решить, например, задачу об определении траектории снаряда, можно попытаться построить специальную электрическую схему с лампами, сопротивлениями, конденсаторами и т. п., так что динамика токов в ней будет подчиняться тем же самым дифференциальным уравнениям, как и для движения снаряда. Тогда, подобрав подходящие начальные условия и включив эту электрическую схему, по ее динамике можно предсказать движение снаряда.

Новое поколение аналоговых устройств для обработки информации, которое, как можно ожидать, будет особенно эффективно в задачах искусственного интеллекта, должно основываться на иной, гораздо более изощренной элементной базе. Чтобы использовать в полной мере преимущества параллельной обработки, это устройство должно представлять собой сеть из связанных между собой простых динамических систем. Подобные системы могут иметь очень малые (даже молекулярные [3]) размеры. Кроме того, работа таких устройств не будет сводиться к простой имитации внешних явлений — в ней будут использоваться некоторые из механизмов, свойственных мозгу.

## 1. ПРИНЦИПЫ АНАЛОГОВОЙ ОБРАБОТКИ ИНФОРМАЦИИ

Ясно, что мозг не в состоянии полностью моделировать все явления в окружающем его мире. Вместо этого он использует в своей работе построенную в нем самую картину внешнего мира, которая включает только свойства, являющиеся существенными для жизнедеятельности человека или животного. Для этого необходима первичная обработка поступающего из внешнего мира информационного потока, с тем чтобы выделить из него существенные характеристики. Подобная первичная обработка сопровождается поэтому сжатием информации.

Сжатие информации означает проведение классификации: различные состояния окружения необходимо отнести к одному и тому же классу, если они обладают одними и теми же свойствами. Ясно, что классификация может быть успешной, лишь если она основывается на определенных инвариантах внешнего мира. Распознавание этих инвариантных характеристик является одной из наиболее важных функций мозга.

Например, все буквы «а» независимо от своего написания должны распознаваться как единый объект; все треугольники с разными ориентациями и отношениями сторон должны рассматриваться как объекты, отличные от всех квадратов или окружностей. Мы умеем узнавать людей, т. е. способны определить, что перед нами один и тот же человек, независимо от его возраста, одежды, положения тела, выражения лица и т. п. Более того, мозг способен распознавать также временные последовательности событий. В частности, нам нетрудно установить идентичность мелодии независимо от того, исполняется ли она какими-то музыкальными инструментами, человеческим голосом, хором или же целым симфоническим оркестром.

Распознавание образов предполагает, что мозг хранит в своей памяти наборы некоторых типичных картин и может сравнивать с ними любую поступающую на вход картину, основываясь в сравнении на определенных критериях. Традиционные компьютеры крайне неэффективны в задачах распознавания образов, поскольку они способны производить сравнение картин только последовательно, по очереди, одну за другой. Кроме того, для них картины необходимо вначале преобразовать в последовательные списки — описания.

Высокой эффективности можно достигнуть, если осуществлять сравнение одновременно со всеми имеющимися картинами. Это нетрудно сделать с помощью распределенных динамических систем (см. разд. 2).

Обучение представляет собой тесно связанную с распознаванием задачу. Типичные картины, хранящиеся в мозге, являются либо генетически унаследованными, либо индивидуально приобретенными. Различие между этими двумя ситуациями не очень глубокое. В обоих случаях такие типичные картины каким-то путем извлекаются из опыта, и разница лишь в характерных временах.

В абстрактной форме процесс обучения можно описать следующим образом. Имеется набор картин, о которых известно, что они принадлежат к одному и тому же классу (например, приводят к одному и тому же эффекту). Требуется найти такую картину-прототип, которая будет достаточно близкой ко всем картинам из этого класса и далекой от картин из других классов согласно определенному критерию. Чтобы решить эту задачу, достаточно случайно генерировать большое число различных картин и произвести сравнение каждой из них со всеми имеющимися в наличии картинами из данного набора. В качестве картины-прототипа тогда можно выбрать ту картину, для которой совпадение будет наиболее полным. Очевидно, что такой механизм обучения можно успешно реализовать, лишь располагая устройством, способным к параллельному сравнению картин, поскольку в противном случае процесс обучения занимал бы слишком много времени.

Если располагать мощным аналоговым устройством для параллельной обработки картин, это могло бы также существенно упро-

стить процесс *решения задач*, являющийся одним из наиболее важных свойств интеллекта. Решить некоторую задачу — значит найти оптимальную картину (траекторию на карте, последовательность событий, путь по дереву решений некоторой игры). Поэтому если наше устройство способно генерировать много различных картин и «запускать» их одновременно, так что они будут конкурировать друг с другом и в результате «выживет» лишь наиболее эффективная, то это позволит находить решение, не прибегая к логическим операциям с математическими величинами.

Эти аргументы свидетельствуют, что организация параллельной обработки картин обеспечивает серьезные преимущества и, скорее всего, именно таким образом осуществляется работа мозга. Искусственное устройство, наделенное такими способностями, можно рассматривать как сложную динамическую систему, выполняющую в процессе своей эволюции функции генерации, преобразования и сравнения различных структур. В результате приходим к фундаментальной проблеме целенаправленного создания сложных динамических систем с нужными свойствами. До определенной степени это напоминает генную инженерию, которая ставит своей целью получение новых биологических объектов с заранее заданными характеристиками.

## 2. ДВЕ АЛЬТЕРНАТИВЫ В АНАЛОГОВОМ РАСПОЗНАНИИ ОБРАЗОВ

Один из возможных путей реализации аналогового распознавания образов состоит в том, чтобы построить распределенную динамическую систему (или дискретную сеть), аттракторами которой в ее конфигурационном пространстве являлись бы типичные картины-образы. Тогда каждая такая картина будет обладать своей областью притяжения и всякое начальное условие, представляющее собой какую-то допустимую картину, обязано попасть в одну из областей притяжения. С течением времени, в ходе эволюции, эта начальная структура трансформируется в наиболее близкую из хранящихся в памяти структур-аттракторов, а именно в ту, области притяжения которой она принадлежала. Следовательно, подавая на вход в качестве начального условия для такой распределенной системы некоторую структуру, мы, очевидно, будем проводить ее автоматическое (т. е. аналоговое) распознавание, которое будет параллельным, поскольку оно осуществляется в течение однократной эволюции данной динамической системы.

Указанную ситуацию можно реализовать, воспользовавшись распределенными системами с реакцией и диффузией, которые описываются при наличии двух реагирующих компонент уравнениями

$$\begin{aligned} \dot{u} &= f(u, v) + D_u \Delta u, \\ \dot{v} &= g(u, v) + D_v \Delta v, \end{aligned} \quad (1)$$

Хорошо известно (см., например, [4]), что при определенных условиях системы с реакцией и диффузией могут обладать большим количеством разных диссипативных структур, представляющих собой устойчивые неоднородные пространственные распределения реагентов  $u$  и  $v$ . Фактически каждая такая диссипативная структура представляет собой аттрактор в конфигурационном пространстве всех возможных распределений реагентов. Поэтому системы с реакцией и диффузией можно в принципе использовать для аналогового распознавания образов [5—10].

Однако использование систем с реакцией и диффузией в этих целях имеет серьезные ограничения. Прежде всего неизвестно, как построить систему с реакцией и диффузией, аттракторами которой являлись бы заданные структуры. Кроме того, невозможно предсказать заранее, какими будут размеры и форма областей притяжения для таких структур. В результате оказывается невозможным указать критерий, согласно которому будет осуществляться классификация картин.

Другой подход был предложен в работах Хопфилда [11—13] и сейчас широко обсуждается в литературе (см. [13—15]). Он формулируется на языке дискретных сетей и позволяет указать явный рецепт для построения динамических систем с заранее заданными аттракторами.

Рассмотрим динамическую систему, которая представляет сеть из  $N$  связанных между собой элементов, каждый из которых описывается вещественной переменной  $S_i$ ,  $i = 1, 2, \dots, N$ . Пусть эволюция этой системы подчиняется динамическим уравнениям

$$\dot{S}_i = -\partial \mathcal{H} / \partial S_i, \quad (2)$$

в которых функция  $\mathcal{H}$  играет роль энергии и дается выражением

$$\mathcal{H} = -\frac{1}{2} \sum_{i \neq j} J_{ij} S_i S_j - \alpha \sum_i \left( \frac{1}{2} S_i^2 - \frac{1}{4} S_i^4 \right). \quad (3)$$

Если положительный параметр  $\alpha$  достаточно велик, то в устойчивых стационарных состояниях такой системы переменные  $S_i$  принимают лишь значения, близкие к  $+1$  или  $-1$ . Следовательно, в пределе  $\alpha \rightarrow +\infty$  всякий аттрактор системы (2), (3) будет представлять собой некоторый набор  $\{\xi_i\}$  со значениями величин  $\xi_i$ , равными  $+1$  или  $-1$ .

При произвольной матрице связей  $J_{ij}$  аттракторами динамической системы (2), (3) будут, вообще говоря, некоторые случайные наборы. Однако можно так выбрать элементы этой матрицы, чтобы система обладала в качестве аттракторов заданными наборами  $\{\xi_i^{(m)}\}$ . А именно, если взять  $J_{ij}$  в виде

$$J_{ij} = \frac{1}{M} \sum_{m=1}^M \xi_i^{(m)} \xi_j^{(m)}, \quad (4)$$



то каждый из взаимно ортогональных наборов  $\{\xi_i^{(m)}\}$ , удовлетворяющих условиям

$$\frac{1}{N} \sum_{i=1}^N \xi_i^{(m)} \xi_i^{(m')} = \Delta_{mm'}, \quad (5)$$

будет являться аттрактором этой динамической системы (пока их суммарное число  $M$  не слишком велико).

Главный недостаток модели Хопфилда, как и имеющихся ее обобщений, заключается в том, что взаимодействия в ней не являются локальными: каждый элемент сети взаимодействует со всеми другими. Поскольку распознаваемые такой системой образы есть различные двоичные последовательности  $\{\xi_i\}$ , число элементов в системе  $N$  определяется длиной этих последовательностей. С увеличением  $N$  число связей между элементами растет как  $N^2$ . Необходимость создания столь большого количества связей серьезно ограничивает возможности реализации модели Хопфилда на микроскопическом уровне — в устройствах с элементами молекулярных размеров. Кроме того, как и для систем с реакцией и диффузией, в модели Хопфилда заранее неизвестны области притяжения аттракторов, а поэтому распознавание происходит согласно неустановленным критериям.

В описанных выше случаях распознаваемый образ подается в качестве начального условия, а динамическая система имеет столько аттракторов, каково число запомненных образов. Возможен, однако, и альтернативный подход. Рассмотрим такую динамическую систему, которая описывается некоторым набором параметров  $\{q_k\}$ , причем при заданном наборе этих параметров она обладает всего одним аттрактором. Пусть, однако, эта динамическая система довольно чувствительна к изменению своих параметров, так что при изменении величин  $\{q_k\}$  один аттрактор сменяется другим. Ясно тогда, что каждому из возможных аттракторов будет отвечать своя область в пространстве параметров, образованном всевозможными наборами  $\{q_k\}$ . Если теперь выбрать в качестве набора  $\{q_k\}$  распознаваемый образ и проследить, притяжением к какому из аттракторов будет завершаться эволюция системы, это также позволит осуществить аналоговое распознавание.

В качестве примера рассмотрим динамическую систему с конкуренцией мод, предложенную в работах [16—18]. Эта система характеризуется следующей совокупностью дифференциальных уравнений:

$$\dot{A}_m = (\beta R \sum_{k=1}^K \chi_k^{(m)} q_k - \gamma) A_m, \quad m = 1, 2, \dots, M, \quad (6)$$

в которых

$$\sum_{m=1}^M A_m + R = R_0 = \text{const}, \quad \sum_{k=1}^K q_k^2 = 1. \quad (7)$$

Указанные уравнения описывают кинетику автокаталитических реакций со схемой



при условии, что эффективная скорость размножения  $w_m$  реагента  $A_m$  равна

$$w_m = \sum_{k=1}^K \chi_k^{(m)} q_k. \quad (9)$$

Более подробно о возможной практической реализации такой схемы рассказано в [18].

Кинетические уравнения (6) описывают конкуренцию  $M$  различных мод, каждая из которых отвечает определенному образу. Простой анализ показывает, что в общем случае эти уравнения имеют всего одно устойчивое стационарное решение, являющееся единственным аттрактором данной системы. Если  $\max_m w_m < \gamma/\beta R_0$ , то с течением времени концентрации всех реагентов  $A_m$  стремятся к нулю. Если же  $\max_m w_m > \gamma/\beta R_0$ , то выживает лишь та мода  $m^*$ , для которой величина  $w_m$  максимальна; в пределе  $t \rightarrow \infty$  концентрация соответствующего реагента стремится к значению

$$A_{m^*} = R_0 - \gamma/\beta w_{m^*}, \quad (10)$$

а концентрации остальных реагентов обращаются в нуль. Иными словами, в ходе конкуренции побеждает та мода, для которой  $w_m$  наибольшая.

Допустим теперь, что каждый из наборов  $\{\chi_k^{(m)}\}$  характеризует один из  $M$  хранящихся в памяти системы образов, а набор  $\{q_k\}$  представляет собой образ, предъявленный системе для распознавания. Как следует из определения (9), величина  $w_m$  будет максимальна для того хранящегося в памяти образа  $\{\chi_k^{(m)}\}$ , который наиболее близок к предъявленному.

Следовательно, исход эволюции для динамической системы (6) прямо указывает, к какому из хранящихся образов наиболее близок распознаваемый. При этом заранее известен критерий близости — величина скалярного произведения (9). Результат распознавания не зависит от начальных условий для такой динамической системы (нужно лишь, чтобы в начальный момент времени были отличны от нуля концентрации всех реагентов  $A_m$ ).

Разумеется, описанная выше система — это всего лишь простейший пример использования альтернативного подхода к построению аналоговых распознающих устройств. Тем не менее она демонстрирует возможность так проводить распознавание, чтобы распознаваемый образ кодировался в наборе параметров динамической системы, а не формировал бы ее начальное условие.

### 3. СЕМАНТИЧЕСКИЕ СТРУКТУРЫ

Работа мозга по распознаванию образов носит многоступенчатый, иерархический характер. Очевидно, что анализ зрительной картины или звукового сигнала нельзя осуществить путем однократного применения простейшей процедуры описанного выше типа. Рассмотрим, например, как могла бы выглядеть последовательность шагов, если задача заключается в том, чтобы отличить треугольник от всех других многоугольников.

Первый шаг состоит в выделении линий и проверке того, что их достаточно большие участки близки к отрезкам прямых. Затем необходимо выделить вершины, т. е. точки, где сходятся две из таких линий. Далее необходимо построить как бы промежуточное, сжатое изображение картины — ее граф. На этом графе фиксируется лишь наличие связей между вершинами, независимо от длины или ориентации линий, осуществляющих такие связи. Теперь, когда каждый граф отвечает целому классу, например всем различным треугольникам, следует осуществить распознавание графов, что дает окончательный ответ на поставленную задачу.

Видно, что даже решение такой простой задачи требует ее перевода на некоторый внутренний язык и построения фактически определенных семантических структур — графов. Элементы такой структуры — это уже в сущности абстрактные объекты, или понятия (в описанном выше случае — вершина и связи, осуществляемые отрезками прямых линий). Каждый граф, в свою очередь, также может рассматриваться в качестве некоторого понятия (например, граф треугольника) и служить элементом, входящим в состав семантической структуры следующей степени.

Основной вид деятельности мозга у высших животных и человека — это операции с семантическими структурами: их распознавание, генерация, передача, преобразование и сравнение. С этой точки зрения появление у человека языка не было чем-то совершенно новым. Как писал известный специалист в области искусственного интеллекта М. Минский, «задолго до того, как наши предки научились говорить, у них уже возникли специальные механизмы мозга для представления объектов, различий и причин; эти механизмы позднее легли в основу нашего языка (и грамматики в том числе)» [19].

Все необходимые операции с семантическими структурами должны осуществляться в мозге аналоговым образом, как некоторые динамические процессы в сложной распределенной нелинейной системе. Мозг как бы является средой, где семантические структуры «живут своей жизнью»: эволюционируют, взаимодействуют и конкурируют между собой.

Современные работы по инженерии динамических систем для аналоговой обработки информации — это лишь начало обширной программы работ по созданию искусственных устройств, близких по принципам своей деятельности (но не по физическому строению) к человеческому мозгу. Здесь пока делаются самые первые шаги, решаются отдельные модельные задачи.

Ниже мне хотелось бы рассказать об одной такой задаче, связанной с обработкой семантических структур, а именно об использовании динамических систем для распознавания и передачи графов.

Для простоты ограничимся графами, которые образованы однотипными элементами со связями одного вида. Если перенумеровать все элементы в некоторой последовательности, граф можно задать матрицей его связей  $T_{ij}$ , в которой  $T_{ij} = 1$ , когда элементы с номерами  $i$  и  $j$  связаны в графе между собой, и  $T_{ij} = 0$ , если связь между ними отсутствует. Кроме того, будем считать, что  $T_{ii} = 0$ . Связи предполагаются симметричными, а поэтому  $T_{ij} = T_{ji}$ .

Задание графа посредством матрицы  $T_{ij}$  неоднозначно, поскольку получающаяся матрица зависит от выбора последовательности нумерации элементов. Попытаемся поэтому построить для каждого графа такую физическую систему, динамика которой определялась бы только картиной связей, а не нумерацией элементов в графе.

Предположим, что каждому элементу графа сопоставлена некоторая массивная частица. Все эти частицы расположены вдоль одной прямой, причем частицы с номерами  $i$  и  $j$  упруго связаны друг с другом, если связаны между собой соответствующие элементы графа. Такая физическая модель будет описываться системой дифференциальных уравнений

$$\ddot{x}_i = - \sum_{j=1}^N T_{ij} (x_i - x_j), \quad (11)$$

где  $x_i$  — координата  $i$ -й частицы; массы всех частиц и упругости всех струн одинаковы и положены равными единице.

В результате каждому графу будет сопоставлена как бы некоторая одномерная молекула. Эта молекула обладает определенным спектром колебаний, т. е. набором собственных частот  $\{\omega_\alpha\}$ ,  $\alpha = 1, 2, \dots, N$ , таких, что  $\lambda_\alpha = -\omega_\alpha^2$  есть собственные значения вещественной симметричной матрицы:

$$R_{ij} = -T_{ij} + \delta_{ij} \sum_{k=1}^N T_{ik}. \quad (12)$$

Очень важно, что спектр колебаний почти однозначно определяет такую одномерную молекулу, а следовательно, и породивший ее граф.

Вообще говоря, тем же набором собственных значений будет обладать любая матрица  $R'$ , полученная из  $R$  посредством невырожденного линейного преобразования  $U$ :

$$R'_{ij} = U_{ik} R_{kl} U_{lj}^{-1}; \quad (13)$$

Однако, чтобы сопоставить матрице  $R'$  некоторую молекулу рассматриваемого вида, необходимо выполнить два условия. Во-первых, недиагональные элементы матрицы  $R'$  должны быть только

нулями или минус единицами. Во-вторых, как следует из (12), всякий диагональный элемент этой матрицы должен быть равен взятой с обратным знаком сумме остальных элементов, стоящих в ней в том же столбце (или строке, что эквивалентно из-за симметрии матрицы), т. е.

$$R'_{ii} = - \sum_{j \neq i} R'_{ij}. \quad (14)$$

За редкими исключениями, при таких двух ограничениях все получаемые матрицы  $R'$  будут отличаться от  $R$  лишь изменением порядка нумерации строк и столбцов, а следовательно, отвечать одной и той же молекуле.

Итак, каждый граф почти всегда можно однозначно характеризовать набором частот  $\{\omega_\alpha\}$ , представляющих собой собственные частоты колебаний соответствующей одномерной молекулы. Это позволяет свести задачу распознавания графов к распознаванию генерируемых ими мелодий.

Попытаемся построить такую динамическую систему-детектор, которая автоматически восстанавливала бы граф по его мелодии, т. е. набору частот  $\{\omega_\alpha\}$ . В качестве такой гипотетической системы вновь рассмотрим одномерную молекулу, матрица связей которой  $\tau_{ij}$  уже не фиксирована, а может случайно меняться со временем. Пусть, кроме того, колебания этой молекулы совершаются с небольшим затуханием, а к одной из частиц приложена вынуждающая сила  $f(t)$ .

При заданных связях  $\tau_{ij}$  такая система описывается уравнениями

$$\ddot{x}_i = - \sum_{j=1}^N \tau_{ij} (x_i - x_j) - \gamma \dot{x}_i + f(t) \delta_{i1}. \quad (15)$$

Допустим, что вынуждающая сила  $f(t)$  есть мелодия, генерируемая некоторым графом, и она характеризуется некоторым набором частот  $\{\omega_\alpha\}$ . Очевидно, что при заданной  $f(t)$  и достаточно высокой добротности молекулы-детектора сумма средних квадратов колебаний образующих ее частиц будет тем больше, чем ближе совпадают частоты собственных колебаний  $\{\omega'_\alpha\}$  молекулы-детектора с набором частот  $\{\omega_\alpha\}$ , т. е. чем ближе мы находимся к ситуации полного резонанса.

Получаемая путем усреднения по интервалам времени, содержащим много периодов колебаний, величина

$$F = \sum_i \overline{x_i^2(t)} \quad (16)$$

в установившемся режиме дается выражением вида

$$F = \sum_{\alpha, \beta} \frac{C_{\alpha\beta}^2}{(\omega_\alpha - \omega'_\beta)^2 + \gamma^2} \quad (17)$$

и достигает абсолютного максимума при полном совпадении наборов  $\{\omega_\alpha\}$  и  $\{\omega'_\alpha\}$ . Это, конечно, предполагает отсутствие перекрытия отдельных резонансов, т. е. выполнение условия  $\gamma \ll \omega_\alpha$ .

Пусть теперь связи  $\tau_{ij}$  могут случайно меняться с течением времени, но, однако, эти случайные мутации происходят достаточно редко: интервал между двумя последовательными мутациями в среднем гораздо больше времени релаксации молекулы-детектора  $t_{\text{рел}} = 1/\gamma$ . Допустим для простоты, что вероятность  $w$  случайного изменения (разрыва или рождения) любой связи  $\tau_{ij}$  одинакова, но тем меньше, чем большее значение приобрела величина  $F$ , т. е. сумма средних квадратов смещений частиц в молекуле-детекторе. Удобно выбрать вероятность  $w$  в виде

$$w = Ge^{-F/\theta}, \quad (18)$$

где параметр  $\theta$  играет роль температуры.

При выполнении определенных условий построенная динамическая система будет проводить основное время в состоянии с картиной связей, идентичной графу подаваемой мелодии.

Действительно, поскольку изменения связей происходят медленно, оправдано адиабатическое приближение. Тогда можно считать, что величина  $F$  в (18) принимает установившееся значение (17), отвечающее мгновенному набору частот  $\{\omega'_\alpha\}$  в молекуле-детекторе, т. е. мгновенной картине связей:  $F = F(\{\tau_{ij}\})$ . В результате описанной выше случайной динамики установится следующее равновесное распределение вероятности для картин связей:

$$P(\{\tau_{ij}\}) \sim \exp\{F(\{\tau_{ij}\})/\theta\}, \quad (19)$$

и наиболее вероятная картина будет отвечать максимуму  $F$ , т. е. совпадению наборов частот  $\{\omega_\alpha\}$  в подаваемой на вход мелодии и набора собственных частот  $\{\omega'_\alpha\}$  молекулы-детектора. Поскольку, как уже отмечалось выше, каждый набор  $\{\omega_\alpha\}$  однозначно определяет некоторый граф, это означает, что наиболее вероятная картина связей  $\tau_{ij}$  будет идентична графу поступающей мелодии. Таким образом, такая гипотетическая динамическая система обладает способностью восстанавливать граф по его мелодии.

Приведенный пример демонстрирует принципиальную возможность кодирования и передачи семантических структур посредством меняющегося со временем сигнала. Мы видим, что имеется возможность производить также и аналоговую дешифровку сигнала, т. е. восстанавливать по нему исходный граф. Более подробное обсуждение этих вопросов будет проведено в отдельной публикации. Здесь хотелось бы лишь заметить, что в практических приложениях, вероятно, будет удобнее использовать стохастическую динамику Метрополиса [20] вместо равных вероятностей мутации (18) и применять процедуру «имитации отжига» (см., например, [21]).

Разумеется, нельзя утверждать, что именно такая схема используется в работе мозга. Учитывая, однако, исключительную роль

которую играют семантические структуры в деятельности мозга, следует ожидать, что он использует какие-то способы аналогового кодирования для графов при передаче их по каналам связи [22, 23].

#### ЛИТЕРАТУРА

1. Фелдман Дж. А. Сетевые модели // Реальность и прогнозы искусственного интеллекта. М.: Мир, 1987. С. 137—147.
2. Михайлов А. С. Физики задумываются над механизмом работы мозга. // Природа. 1987. № 3. С. 15—26.
3. Рамбиди Н. Г., Замалин В. М. Молекулярная микроэлектроника: Истоки и надежды. М.: Знание, 1985. 64 с.
4. Полак Л. С., Михайлов А. С. Самоорганизация в неравновесных физико-химических системах. М.: Наука, 1983. 351 с.
5. Балкарей Ю. И., Никулин И. Г., Елинсон М. И. Континуальные среды для обработки информации: (Автоволновые среды) // Проблемы современной радиотехники и электроники. М.: Наука, 1980. С. 431—449.
6. Евтихов М. Г., Балкарей Ю. И., Никулин М. Г., Елинсон М. И. Математическое моделирование динамической фазовой памяти автоволновой среды // Микроэлектроника. 1978. Т. 7. С. 421—430.
7. Гуляев Ю. В., Балкарей Ю. И., Ржанов Ю. А. и др. О возможности моделирования некоторых биологических процессов в активной полупроводниковой среде // ДАН СССР. 1981. Т. 250. С. 82—85.
8. Балкарей Ю. И., Евтихов М. Г., Елинсон М. И. Исследование явления самодостройки в активной диффузионной среде // Микроэлектроника. 1978. Т. 10. С. 78—83.
9. Крицкий В. И., Михайлов А. С. Автоволны. М.: Знание, 1984. 64 с.
10. Ebeling W. Pattern dynamics and optimization by reaction-diffusion systems // J. Stat. Phys. 1986. Vol. 45. P. 891—903.
11. Hopfield J. J. Neural networks and physical systems with emergent collective computation abilities // Proc. Nat. Acad. Sci. USA. 1982. Vol. 79. P. 2554—2559.
12. Hopfield J. J. Neurons with gradual response have collective computational properties like those of two-state neurons // Ibid. 1984. Vol. 81. P. 3088—3092.
13. Hopfield J. J., Tank D. W. Collective computation with continuous variables // Disordered systems and biological organization / Ed. A. Bienenstock et al. Berlin. etc.: Springer, 1986. P. 155—170.
14. Веденов А. А., Левченко Б. Б. Об одном классе нелинейных систем с памятью // Письма в ЖЭТФ. 1985. Т. 41, № 8. С. 328—330.
15. Веденов А. А. Моделирование элементов мышления. М.: Наука, 1988. 311 с.
16. Герзон С. А. Опознающие синергетические устройства: Препр. Ин-та теорет. и эксперим. физики № 7. М., 1985. 17 с.
17. Герзон С. А. Динамические модели для описания взаимодействия слов языка: Препр. Ин-та теорет. и эксперим. физики № 139. М., 1987. 7 с.
18. Герзон С. А., Михайлов А. С. Возможный механизм анализа изображений неравновесными биохимическими системами // ДАН СССР. 1986. Т. 291. С. 228—230.
19. Минский М. Общение с вземным разумом // Реальность и прогнозы искусственного интеллекта. М.: Мир, 1987. С. 231—244.
20. Metropolis N., Rosenbluth A. W., Rosenbluth M. N. et al. Equation of state calculations by fast computing machines // J. Chem. Phys. 1953. Vol. 6. P. 1087—1092.
21. Kirkpatrick S., Gelatt C. D., Vecchi M. P. Optimization by simulated annealing // Science. 1985. Vol. 21. P. 1087—1098.
22. Hogg T., Huberman B. A. Artificial intelligence and large scale computation: a physics perspective // Phys. Rep. 1987. Vol. 156. P. 227—310.
23. Mikhailov A. S. Graph identification by simulated annealing // J. Phys. A. 1988. Vol. 21. P. L487—L490.

УДК 630.1.5 61+530.182

## НЕЛИНЕЙНЫЕ ЯВЛЕНИЯ В ФИЗИКЕ КОНДЕНСИРОВАННЫХ СРЕД: ВСЕЛЕННАЯ В КАПЛЕ ГЕЛИЯ

Г. Е. ВОЛОВИК

Одно из наиболее интересных нелинейных явлений в физике — спонтанное образование упорядоченных структур. Структуры возникают в двух на первый взгляд существенно различных типах систем. Это, во-первых, диссипативные структуры в далеких от равновесия динамических системах (ячейки Бенара в турбулентности, автоволновые процессы в химических и биологических системах типа реакции Белоусова—Жаботинского, нестационарный эффект Джозефсона в сверхпроводимости, когерентная оптическая генерация в лазерах и т. д.). Во-вторых, это когерентные состояния в системах, находящихся в равновесии (упорядочение атомов в кристаллах, магнитных моментов в ферро- и антиферромагнетиках, дипольных моментов в сегнетоэлектриках, молекулярных осей в жидких кристаллах, образование сверхпроводящего состояния в металлах и сверхтекучего состояния в жидких изотопах гелия и в нейтронных звездах и т. д.). Не исключено, что и физический вакуум представляет собой спонтанно образовавшуюся упорядоченную структуру.

Свойства упорядоченных структур во многом сходны, поскольку имеют общее происхождение — спонтанное нарушение симметрии. Последнее означает, что в первоначально однородной и изотропной системе спонтанно появляются выделенные направления или положения либо в обычном пространстве-времени, либо в изотопическом пространстве. Каждая структура характеризуется своим типом нарушения симметрии, и структуры, обладающие одинаковым нарушением симметрии, имеют одинаковые свойства, описываются одинаковым параметром порядка. Это позволяет предсказывать поведение одной системы путем изучения более доступного для наблюдения поведения другой системы.

В этом отношении наиболее интересно исследование систем с максимально сильным нарушением симметрии, которые тем самым сочетают в себе свойства большинства остальных систем. Среди имеющихся в настоящее время в лабораторных условиях веществ наиболее мощное нарушение симметрии осуществляется в сверх-



текущих фазах жидкого  ${}^3\text{He}$ . Эти фазы сочетают в себе свойства упорядоченных магнетиков, жидких кристаллов и сверхтекучего  ${}^4\text{He}$ , а в состоянии вращения также свойства кристаллов и сегнетоэлектриков. Причем сочетание этих свойств отнюдь не механическое, а приводит к ряду новых фундаментальных явлений, изучение которых полезно не только для физики конденсированных сред и нелинейных динамических структур, но также может пролить свет и на свойства вакуума физики элементарных частиц. Среди трех экспериментально обнаруженных сверхтекучих фаз гелия-3 ( ${}^3\text{He-A}$ ,  ${}^3\text{He-B}$  и  ${}^3\text{He-A}_1$ ) и еще десятка теоретически возможных структура сверхтекучего  ${}^3\text{He-A}$  наиболее близко соответствует структуре физического вакуума, поэтому мы рассмотрим именно эту фазу.

## 1. КОГЕРЕНТНОЕ СВЕРХТЕКУЧЕЕ СОСТОЯНИЕ

Сверхтекучесть жидкого  ${}^3\text{He}$  существенно отлична от открытой полвека назад сверхтекучести более тяжелого и гораздо более доступного изотопа гелия-4. Однако происхождение когерентного сверхтекучего состояния у них одинаковое: спонтанное нарушение калибровочной симметрии. Поэтому мы начнем с намного более простого случая сверхтекучести в  ${}^4\text{He}$ . Его атомы не обладают спиновым моментом и поэтому являются бозонами. Наиболее интересное свойство ансамбля бозонов заключается в том, что при охлаждении до некоторой температуры  $T_c$  (для  ${}^4\text{He}$  она порядка 1 К) образуется высококогерентное состояние (бозе-конденсат) макроскопическая часть атомов конденсируется в одном и том же квантовом состоянии с наименьшей энергией. Более того, не только квантовые состояния всех атомов идентичны, но и фазы волновых функций атомов бозе-конденсата совпадают между собой. Следовательно, для данного когерентного состояния появляется выделенное значение  $\varphi$  фазы, т. е. с образованием конденсата нарушается калибровочная симметрия.

Благодаря когерентности движение бозе-конденсата может быть описано волновой функцией единственного атома  $\psi = |\psi|e^{i\varphi}$ , однако модуль  $|\psi|$  должен быть нормирован на полное число атомов  $N$  бозе-конденсата:  $\int |\psi|^2 d^3r = N$ . Поскольку волновая функция  $\psi$  приобретает таким образом макроскопический характер, уравнение, которому она подчиняется, является существенно нелинейным. Это вполне аналогично образованию когерентного состояния в лазерах, где внутри оптического резонатора возникает одна макроскопическая мода электромагнитного поля, комплексная амплитуда которой эквивалентна параметру порядка  $\psi$  в сверхтекучем  ${}^4\text{He}$ . В обоих случаях спонтанно возникает когерентность по фазе: в одном случае у фотонов, в другом случае у атомов  ${}^4\text{He}$ .

В статическом случае уравнение для  $\psi$  получается варьированием функционала свободной энергии Гинзбурга—Ландау

$$\delta F/\delta\psi^* = 0; F = \int \{-\alpha(1 - T/T_c)|\psi|^2 + \beta|\psi|^4 + j|\nabla\psi|^2\}d^3r \quad (1)$$

и представляет собой вариант нелинейного уравнения Шредингера. Отметим, однако, что динамическое уравнение для  $\psi$  имеет, как правило, существенно более сложный вид, и неизвестно, есть ли область параметров, где оно совпадало бы с нелинейным уравнением Шредингера. С другой стороны, можно указать ситуации, когда  $\psi$  подчиняется нелинейному уравнению диффузии.

Явление сверхтекучести, т. е. отсутствие трения при протекании жидкости сквозь узкие каналы, — непосредственный результат когерентности фаз волновых функций атомов  ${}^4\text{He}$  в бозе-конденсате: шероховатости на стенках могут рассеивать отдельные атомы, но бессильны перед коллективной системой бозе-конденсата. В результате в жидкости имеются два вида движения: движение сверхтекучей компоненты, состоящей из атомов бозе-конденсата, не испытывающего трения, и движение нормальной компоненты, состоящей из не вошедших в конденсат некогерентных атомов, не отличающееся от течения обычной вязкой жидкости.

## 2. ${}^3\text{He-A}$ — СВЕРХТЕКУЧИЙ АНТИФЕРРОМАГНИТНЫЙ ЖИДКИЙ КРИСТАЛЛ

В жидком  ${}^3\text{He}$  сверхтекучесть возникает более экзотическим образом. Поскольку атомы  ${}^3\text{He}$  обладают ядерным спином  $1/2$ , они являются фермионами и поэтому не могут образовывать бозе-конденсат — в одном квантовом состоянии может находиться не более одного фермиона. Переход в сверхтекучее состояние в системе фермионов происходит по следующему сценарию: при охлаждении до некоторой температуры  $T_c$  (в жидком  ${}^3\text{He}$  это примерно  $10^{-3}$  К, а в сверхпроводниках может достигать 100 К) фермионы спариваются в своеобразные молекулы (куперовские пары), которые уже являются бозонами и могут выпадать в бозе-конденсат.

Существенно, что в  ${}^3\text{He-A}$ , а возможно, и в некоторых сверхпроводниках эти куперовские молекулы отнюдь не имеют сферической формы. Они обладают сразу двумя осями анизотропии: осью пространственной анизотропии  $\mathbf{l}$  и осью магнитной (спиновой) анизотропии  $\mathbf{d}$ . В высококогерентном состоянии бозе-конденсата оси  $\mathbf{l}$  и  $\mathbf{d}$  всех куперовских пар упорядочены, в результате, помимо сверхтекучих свойств, вызванных фазовой когерентностью,  ${}^3\text{He-A}$  обладает как жидкокристаллическими, так и антиферромагнитными свойствами.

Благодаря этим дополнительным степеням свободы параметр порядка (волновая функция куперовской пары, одинаковая для всех пар в бозе-конденсате) содержит не одну, а девять комплексных амплитуд, которые обычно записываются в виде  $3 \times 3$  комплексной матрицы  $A_{ik}$ . Векторы  $\mathbf{l}$  и  $\mathbf{d}$  являются определенными комбинациями компонент этой матрицы. Функционал Гинзбурга—Ландау для  $A_{ik}$  содержит пять членов 4-го порядка и три

градиентных члена:

$$\begin{aligned}
 F = \int d^3r [ & -\alpha(1 - T/T_c) A_{ik}^* A_{ik} + \beta_1 A_{ki}^* A_{ki}^* A_{mn}^* A_{mn} + \\
 & + \beta_2 A_{ki}^* A_{ki} A_{mn}^* A_{mn} + \beta_3 A_{ki}^* A_{mi}^* A_{kn}^* A_{mn} + \beta_4 A_{ki}^* A_{mi} A_{kn}^* A_{mn} + \\
 & + \beta_5 A_{ki}^* A_{mi} A_{kn}^* A_{mn}^* + \gamma_1 \nabla_1 A_{km} \nabla_i A_{km}^* + \gamma_2 \nabla_i A_{ki} \nabla_m A_{km}^* + \\
 & + \gamma_3 \nabla_i A_{km} \nabla_m A_{ki}^* ].
 \end{aligned} \tag{2}$$

Это приводит к колоссальному числу различных решений уравнения

$$\delta F / \delta A_{ik}^* = 0, \tag{3}$$

для исследования и классификации которых используются методы теории групп и топологии. Соответственно усложняется и динамика, которая должна учитывать не только многокомпонентное когерентное движение параметра порядка, но и существенно более сложное по сравнению с <sup>4</sup>He движение некогерентной подсистемы жидкости.

### 3. <sup>3</sup>He-A И ФИЗИКА ЭЛЕМЕНТАРНЫХ ЧАСТИЦ

К настоящему времени наибольшее продвижение достигнуто в исследовании поведения <sup>3</sup>He-A вблизи равновесного состояния, соответствующего решению уравнения (3) с минимальной свободной энергией (2), в том числе вблизи основного состояния (вакуума сверхтекучего <sup>3</sup>He-A), представляющего собой равновесное состояние при нулевой температуре. В самом равновесном состоянии в идеальном случае бесконечной системы параметр порядка  $A_{ik}$  не зависит от координат, и, следовательно, векторы  $\mathbf{l}$  и  $\mathbf{d}$  постоянны в пространстве, хотя и могут быть ориентированы в произвольных направлениях. При наличии границ вакуум, вообще говоря, становится неоднородным в пространстве из-за граничных условий на вектор  $\mathbf{l}$ .

Динамика <sup>3</sup>He-A вблизи вакуумного состояния включает в себя динамику параметра порядка и динамику элементарных возбуждений на фоне вакуума, составляющих некогерентную подсистему жидкости. Последняя включает в себя кванты колебаний векторов  $\mathbf{l}$  и  $\mathbf{d}$  и других компонент матрицы  $A_{ik}$ , а также набор фермионных возбуждений, получающихся при разрыве куперовских пар. Такая двухжидкостная динамика <sup>3</sup>He-A вблизи вакуума напоминает квантовую теорию поля в физике элементарных частиц. Элементарным возбуждениям в <sup>3</sup>He-A соответствуют элементарные частицы: фермионы (электроны и нейтрино) и бозоны (фотоны, гравитоны,  $W$ -бозоны). С другой стороны, параметр порядка  $A_{ik}$  меняет локально свойства жидкого <sup>3</sup>He, причем пространственные и временные изменения параметра порядка  $A_i$  действуют на элементарные возбуждения, как калибровочные поля на элементарные частицы. Таким образом, матрица параметра порядка  $A_{ik}$ , которая в <sup>3</sup>He описывает сверхтекучее движение, в физике эле-

ментарных частиц соответствуют калибровочные поля: электромагнитное, гравитационное и др.

Самым поразительным в этой аналогии является то, что в ряде случаев соответствие является точным, а именно явления описываются одинаковыми формулами и приводят к одинаковым эффектам, хотя и выраженным различными физическими языками. Ниже приводятся некоторые выдержки из словаря, позволяющего переводить с языка  ${}^3\text{He-A}$  на язык физики элементарных частиц.

### Физика ${}^3\text{He-A}$

Основное состояние  
 Элементарные возбуждения  
 Поле орбитального вектора  $\mathbf{l}$   
 Орбитальные волны (распространяющиеся колебания вектора  $\mathbf{l}$ )  
 Кванты орбитальных волн  
 Скорость Ферми  
 Поле тех компонент матрицы  $A_{ik}$ , которые создают локальную геометрию для элементарных возбуждений  
 Клепшинг-моды (распространяющиеся колебания этих компонент)  
 Кванты спин-орбитальных волн  
 Неспаренные фермионы  
 Спин  
 Боголюбовский изоспин в пространстве частица — дырка  
 Ось магнитной анизотропии  $\mathbf{d}$   
 Уравнение Боголюбова для фермионов  
 Динамические уравнения Гинзбурга—Ландау  
 Логарифмическая расходимость градиентной энергии  
 Импульс  
 Обмен импульсом между сверхтекучей и нормальной подсистемами  
 Квантовый вихрь, оканчивающийся на еж в поле вектора  $\mathbf{l}$   
 Еж в поле вектора  $\mathbf{d}$   
 Частицеподобные солитоны

### Физика элементарных частиц

Вакуум  
 Элементарные частицы  
 Электромагнитное поле  
 Электромагнитные волны  
 Фотоны  
 Скорость света  
 Гравитационное поле  
 Гравитационные волны  
 W-бозоны  
 Лептоны (электроны и нейтрино)  
 Изотопический спин  
 Спин  
 Ось квантования изоспина  
 Уравнения Дирака для фермионов  
 Уравнения Максвелла и Эйнштейна  
 Явление нуль-заряда (логарифмическая экранировка электрического заряда из-за поляризации вакуума)  
 Киральный заряд  
 Киральная аномалия (рождение кирального заряда из вакуума)  
 Магнитный монополю Дирака  
 Магнитный монополю  $g'$  Хоофта—Полякова  
 Адроны?

Это соответствие полезно для обеих сторон. В  ${}^3\text{He-A}$  оно позволяет с единой точки зрения калибровочных полей описать ряд явлений, детально исследованных в квантовой теории поля. Для физики элементарных частиц сверхтекучий  ${}^3\text{He}$  представляет собой модель, которая показывает один из возможных механизмов происхождения электромагнитного, слабого и гравитационного полей, нарушения четности в слабых взаимодействиях (фермионы

в  ${}^3\text{He-A}$ , соответствующие электронам, имеют только левую поляризацию), массы  $W$ -бозона и т. д. Существенно, что квантовая теория поля в  ${}^3\text{He-A}$  является полностью замкнутой и не содержит расходимостей, которые характерны для квантовой теории поля в физике элементарных частиц и требуют введения параметров обрезания. Это следствие того, что структура вакуума в  ${}^3\text{He-A}$  известна в отличие от вакуума в физике элементарных частиц. Подробнее об аналогии между этими двумя квантовыми теориями поля см. в обзоре [1].

#### 4. СТРУКТУРЫ С ТОПОЛОГИЧЕСКИМ ЗАРЯДОМ

Знание структуры основного вакуумного состояния позволяет находить устойчивые решения уравнения (3), описывающие неоднородные в пространстве состояния, у которых локально вакуум совпадает с вакуумом сверхтекучего  ${}^3\text{He-A}$ , но ориентация вакуума, т. е. векторов  $\mathbf{l}$  и  $\mathbf{d}$  и других компонент матрицы  $A_{ik}$ , медленно меняется от точки к точке. Устойчивость таких неоднородных вакуумов жестко связана с топологией. Только те состояния устойчивы по отношению как к малым, так и к большим возмущениям, которые обладают нетривиальной топологической структурой, т. е. такой структурой, которую невозможно непрерывным образом преобразовать в однородное состояние.

Центральным понятием здесь является топологический заряд; структуры, которые могут быть непрерывно продеформированы друг в друга, обладают одним и тем же топологическим зарядом, причем топологический заряд основного состояния, как правило, равен нулю. Имеет место закон сохранения топологического заряда: при слиянии различных структур происходит сложение (групповое) зарядов. Топологический заряд имеет несколько компонент: одна из них описывает топологию структур в виде линий (вихри, дислокации, дисклинации), другая — структуры в виде точек (ежи, монополи), третья — в виде поверхностей (доменные стенки) и т. д. Топологический заряд различает сингулярные структуры (дефекты, обладающие кором, внутри которого упорядочение нарушается, т. е. вектор  $\mathbf{l}$  или  $\mathbf{d}$  невозможно определить) и несингулярные структуры (солитоны). Благодаря сохранению топологического заряда такие структуры устойчивы и в статике, и в динамике, им не страшны ни диссипация, ни ее отсутствие: в большинстве случаев единственная возможность устранения подобной структуры — это ее аннигиляция со структурой, имеющей противоположный топологический заряд (так что при сложении зарядов получается нуль).

В качестве примера рассмотрим топологически нетривиальные структуры в  ${}^3\text{He-A}$ , имеющие форму линий. Они характеризуются двумя числами,  $N$  и  $m$ , причем  $N$  принимает лишь четыре значения, которые обычно выбираются как  $0$ ,  $1/2$ ,  $-1/2$  и  $1$  со сложением по модулю  $1$ , т. е.  $1 + 1 = 0$ , а  $m$  принимает любые целые и полуцелые значения с обычным арифметическим законом сложения,

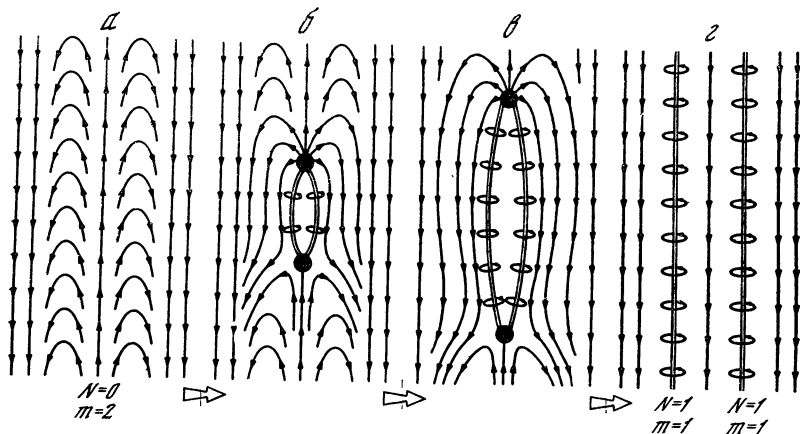


Рис. 1

т. е.  $1 + 1 = 2$ . Объект с  $N = 0$  и  $m = 2$  (так называемый несингулярный вихрь) изображен на рис. 1, а. Стрелками показано поле вектора  $\mathbf{l}$  в сечении плоскостью, проходящей через ось вихря. Обращение в нуль топологического заряда  $N$  означает, что структура несингулярна: вектор  $\mathbf{l}$  определен в каждой точке пространства. Однако непрерывный переход в однородное состояние запрещен из-за отличия от нуля второго топологического заряда  $m$ . Условие  $m = 2$  означает, что течение сверхтекучей компоненты вокруг вихря квантуется и соответствует двум квантам циркуляции скорости течения вокруг вихря (квант циркуляции равен  $h/M$ , где  $M$  — масса куперовской пары,  $h$  — постоянная Планка). Поле фазы бозе-конденсата, ответственное за сверхтекучее течение вокруг вихря, на рисунке не изображено.

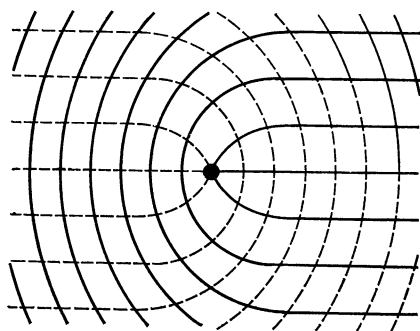


Рис. 2

Два вихря с топологическими зарядами  $N = 1$  и  $m = 1$  каждый схематически изображены на рис. 1, в. Каждый из вихрей сингулярен: фаза бозе-конденсата не определена внутри кора, обозначенного двойной линией. Течение вокруг каждого из вихрей, обозначенное круговыми линиями, соответствует одному кванту циркуляции. Благодаря законам сложения топологических зарядов для  $N$  и  $m$  ( $1 + 1 = 0$  и  $1 + 1 = 2$  соответственно) два сингулярных вихря могут слиться с образованием несингулярного вихря. На рис. 1 показан обратный процесс расщепления несингулярного вихря на два сингулярных. Вначале (рис. 1, б) образуются два экзотических объекта, имеющие сходство с маг-

нитным монополю Дирака. Это ежи (точечные топологические дефекты в поле вектора  $\mathbf{l}$ ; вектор  $\mathbf{l}$  не определен в центре каждого из ежей), имеющие хвосты в виде сингулярных вихревых линий (дираковские струны). Монополи расходятся (рис. 1, *в*), оставляя после себя сингулярные вихри (рис. 1, *г*).

Линейный дефект с  $N = 1/2$  и  $m = 1/2$  изображен на рис. 2. На этот раз изображены поля в сечении вихря поперечной плоскостью. Вектор  $\mathbf{l}$  в этой структуре постоянен и не изображен. После вектора  $\mathbf{d}$  изображено сплошными линиями, а поле фазы  $\varphi$  бозе-конденсата изображено с помощью вспомогательного векторного поля  $\mathbf{u}$  штриховыми линиями (угол вектора  $\mathbf{u}$  в плоскости соответствует фазе  $\varphi$ ). Этот дефект представляет собой, как видно из рисунка, комбинацию из линейного дефекта в поле  $\mathbf{d}$  (спиновая дисклинация) и такого же дефекта в поле вектора  $\mathbf{u}$ , представляющего собой вихрь с половинным квантом циркуляции. Существование подобного гибрида вихря с дисклинацией является результатом необычного сочетания магнитных и сверхтекучих свойств в  $^3\text{He-A}$ . Одно из самых поучительных свойств такого гибрида заключается в том, что по отдельности каждый из дефектов существовать не может. Это напоминает ситуацию с кварками, которые по отдельности не наблюдаемы. Если мы будем пытаться отделить два дефекта один от другого, то при их разделении энергия притяжения между ними растет пропорционально расстоянию между ними, как и при разделении кварков. Это всего лишь один из нескольких примеров конфайнмента, встречающихся в сверхтекучем  $^3\text{He}$ .

Имеются в сверхтекучем  $^3\text{He}$  и топологические объекты, имеющие большое сходство с частицами. Это частицеподобные солитоны, которые по своей топологической структуре совпадают с адронами в некоторых моделях сильных взаимодействий.

Разнообразие структур в сверхтекучем  $^3\text{He}$  увеличивается благодаря тому, что внутри данного топологического класса структур возможны структуры с различной симметрией и, следовательно, с разными физическими свойствами. Например, для несингулярного вихря найдено два решения с различным нарушением четности: в одном из вихрей вдоль оси течет спонтанный поток массы, в то время как в другом — спонтанный поток спина. Таким образом, помимо топологического заряда, структуры характеризуются также и элементами группы симметрии.

Теоретическое исследование неоднородных структур осуществляется в несколько этапов. На первом этапе проводится грубое разбиение структур по топологическим классам (каждый класс характеризуется определенным топологическим зарядом). На втором этапе внутри каждого класса находятся подклассы, характеризующиеся различными элементами симметрии. На третьем этапе для каждого подкласса симметрии находятся решения уравнений Гинзбурга—Ландау, сравниваются энергии различных структур, их относительная устойчивость и делаются выводы о фазовых переходах между структурами (бифуркациях) при изме-

нении внешних условий. И наконец, наиболее сложный этап — исследование динамики структур: их движения, взаимопревращения, слияния и аннигиляции. Именно на этом этапе осуществляется наиболее тесный контакт с динамическими системами, являющимся объектом изучения настоящей школы по нелинейным волнам. (Подробнее о структурах в сверхтекучем  $^3\text{He}$  см. в обзоре [2].)

#### ЛИТЕРАТУРА

1. *Volovik J. E.* Singular behavior of the superfluid  $^3\text{He-A}$  at  $T=0$  and quantum field theory // *J. Low-Temp. Phys.* 1987. Vol. 67, N 3/4. P. 301—312.
2. *Volovik J. E., Salomaa M. M.* Quantized vortices in superfluid  $^3\text{He}$  // *Ibid.* Vol. 59, N 3. P. 533—613.

УДК 551.46

## НЕЛИНЕЙНАЯ ТЕОРИЯ НАБЕГАНИЯ МОРСКИХ ВОЛН НА БЕРЕГ

Е. Н. ПЕЛИНОВСКИЙ

Наверное, нет необходимости говорить о практической значимости этой теории, ведь ущерб от наводнений во всем мире стоит на первом месте в статистике стихийных бедствий. На протяженность зоны затопления побережья морскими волнами влияет много физических и географических факторов: скорость ветра, размеры акваторий, подводный склон, очертания берега и т. д.; поэтому существует много эмпирических формул для описания высоты и длины наката волн на берег, привязанных к конкретным географическим пунктам. Целью настоящей статьи является описание современного состояния проблемы наката морских волн на берег с общезначимых позиций.

В общем случае выделяют три характерных процесса трансформации волны в прибрежной зоне (рис. 1, *a*). В зоне *I* (зона усиления) происходит разгон морских волн под действием ветра и их усиление из-за уменьшения полной глубины бассейна. Для описания характеристик волнения на этом этапе применяют спектрально-лучевые модели с различными аппроксимациями интеграла столкновений и источников энергии ветровых волн [1]. В зоне *II* (зоне прибоя) происходит разрушение волны, причем возможно неоднократное разрушение. Различают несколько типов разрушения гребня волны [2]. Вдали от берега при сильном ветре обычно происходит образование «барашков», возникает скользкий бурун, перемещающийся от вершины волны к ее подножию (рис. 1, *b*). Вблизи берега на сравнительно крутых откосах происходит разрушение волны по типу ныряющего буруна, когда гребень волны заворачивается и обрушивается; этот тип разрушения запечатлен на известной картине японского художника



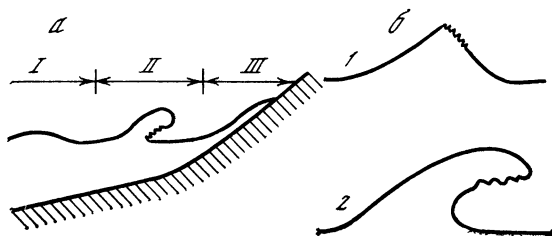


Рис. 1. Трансформация волн в прибрежной зоне (а) и типы и разрушений морских волн (б)  
1 — скользящий бурун,  
2 — ныряющий бурун

Хокусаи. В зоне III (зона наката) волновой поток накатывается на берег, его параметры, собственно, и определяют разрушительную силу морских волн (описание этого процесса см. далее).

На практике разделение прибрежной области на три зоны сделать не всегда удастся. Так, если ветер слабый, то волна не успевает разогнаться в прибрежной зоне и ее разрушения не происходит вообще либо волна обрушивается непосредственно на берегу. Другой важный пример — волны цунами и приливные волны, размеры которых огромны и вероятность их обрушения мала [3]. В этих случаях тип наката определяют как вздымающийся бурун — это подтопление берега (наводнение) без разрушения гребня волны. Очевидно, что вздымающийся бурун представляет собой наиболее простейший тип волнового движения с физической точки зрения, ведь волновой профиль непрерывен всюду. Экспериментальные факты свидетельствуют о том, что такой тип наката реализуется в случае, когда волна в прибрежной зоне проходит

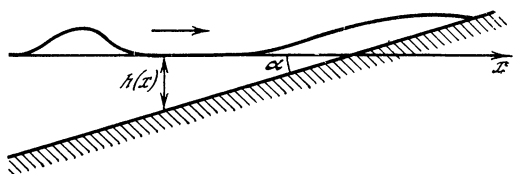


Рис. 2. Геометрия задачи о накате волны на плоский откос

расстояние, сравнимое со своей длиной (эти данные легко интерпретируются в теории, и соответствующие оценки представлены ниже), поэтому можно ожидать, что наиболее важным фактором здесь является только нелинейность волны.

Если определить параметр нелинейности как отношение высоты волны к невозмущенной глубине бассейна, то вблизи уреза этот параметр стремится к бесконечности. В то же время дисперсия и диссипация остаются малыми, так как волна проходит относительно малое расстояние. Это означает, что базовой теорией для описания наката морских волн без обрушения является нелинейная теория мелкой воды, «подправленная» с учетом малой дисперсии и диссипации. Здесь мы ограничимся описанием только точных результатов в рамках нелинейной теории мелкой воды, более общее изложение проблемы содержится в работах [4, 5].

Приведем известные уравнения теории мелкой воды в случае наката волны на плоский откос (рис. 2):

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + g \frac{\partial \eta}{\partial x} = 0, \quad \frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} [(-\alpha x + \eta) u] = 0. \quad (1)$$

Здесь  $u$  — скорость потока;  $\eta$  — уровень воды;  $h = -\alpha x$  — невозмущенная глубина бассейна и  $g$  — ускорение свободного падения.

Получим сначала критерии подобия для системы (1). Введем безразмерные переменные:

$$\tilde{t} = \omega t, \quad \tilde{x} = \alpha x/R, \quad \tilde{\eta} = \eta/R, \quad \tilde{u} = \alpha u/\omega R, \quad (2)$$

где  $R$  и  $\omega$  — характерные высота и частота набегающей волны. В этих переменных система (1) принимает вид

$$\frac{\partial \tilde{u}}{\partial \tilde{t}} + \tilde{u} \frac{\partial \tilde{u}}{\partial \tilde{x}} + \frac{1}{\text{Fr}} \frac{\partial \tilde{\eta}}{\partial \tilde{x}} = 0, \quad \frac{\partial \tilde{\eta}}{\partial \tilde{t}} + \frac{\partial}{\partial \tilde{x}} [(-\tilde{x} + \tilde{\eta}) \tilde{u}] = 0, \quad (3)$$

и она определяется единственным параметром

$$\text{Fr} = \omega^2 R/g\alpha^2, \quad (4)$$

имеющим фундаментальное значение (его смысл будет пояснен ниже).

Вернемся снова к исходной системе (1). В переменных  $u$  и  $D = -\alpha x + \eta$  она приводится к системе уравнений с постоянными коэффициентами и в силу гиперболичности имеет сохраняющиеся римановы инварианты

$$\mathcal{Y}_{\pm} = u \pm 2\sqrt{gD} + g\alpha t. \quad (5)$$

Но тогда эффективным приемом решения таких уравнений является преобразование годографа, что возможно при ненулевом якобиане  $\Delta = \partial(x, t)/\partial(\mathcal{Y}_+, \mathcal{Y}_-)$  (его величина будет подсчитана далее). В результате получаем систему уравнений [6]

$$\partial x/\partial \mathcal{Y}_- - (3/4 \mathcal{Y}_+ + 1/4 \mathcal{Y}_- - g\alpha t) \partial t/\partial \mathcal{Y}_- = 0, \quad (6)$$

$$\partial x/\partial \mathcal{Y}_+ - (3/4 \mathcal{Y}_- + 1/4 \mathcal{Y}_+ - g\alpha t) \partial t/\partial \mathcal{Y}_+ = 0.$$

Хотя эта система и нелинейна, она легко приводится к линейной исключением  $x$ :

$$\frac{\partial^2 t}{\partial \mathcal{Y}_+ \partial \mathcal{Y}_-} + \frac{3}{2(\mathcal{Y}_+ - \mathcal{Y}_-)} \left( \frac{\partial t}{\partial \mathcal{Y}_-} - \frac{\partial t}{\partial \mathcal{Y}_+} \right) = 0. \quad (7)$$

Удобно ввести новые переменные

$$\lambda = 1/2 (\mathcal{Y}_+ + \mathcal{Y}_-) = u + g\alpha t, \quad \sigma = 1/2 (\mathcal{Y}_- - \mathcal{Y}_+) = 2\sqrt{gD} \quad (8)$$

и привести (7) к виду

$$\frac{\partial^2 \Phi}{\partial \lambda^2} - \frac{\partial^2 \Phi}{\partial \sigma^2} - \frac{1}{\sigma} \frac{\partial \Phi}{\partial \sigma} = 0; \quad (9)$$

тогда через  $\Phi$  определяются все физические переменные (подробности см. в [5, 6]):

$$\begin{aligned} \eta &= \frac{1}{2g} \left( \frac{\partial \Phi}{\partial \lambda} - u^2 \right), \quad u = \frac{1}{\sigma} \frac{\partial \Phi}{\partial \sigma}, \\ x &= \frac{1}{2g\lambda} \left( \frac{\partial \Phi}{\partial \lambda} - u^2 - \frac{\sigma^2}{2} \right), \quad t = \frac{\lambda - u}{g\alpha}. \end{aligned} \quad (10)$$

Итак, исходная система нелинейных уравнений мелкой воды свелась к линейному волновому уравнению (9). Основное преимущество такой записи заключается в том, что линии подвижного уреза соответствует  $\sigma = 0$ , так как полная глубина на урезе  $D = 0$ , и, следовательно, уравнение (9) решается в полупространстве с фиксированной границей в отличие от исходных уравнений.

Представленные здесь формулы в принципе позволяют по заданной  $\Phi(\sigma, \lambda)$  однозначно восстановить волновое поле всюду. Фактически это, однако, трудно сделать из-за неявности формул преобразования, что привело к пессимистическому высказыванию, сделанному в работе [7]: «Основной вклад [6] не в вычислении заплеска, а скорее в показе того, что существуют волны возвышения в нелинейном длинноволновом приближении, которые распространяются по берегу постоянного уклона без обрушения». Между тем существует весьма важная практическая ситуация, когда преобразования (10) особенно эффективны. Речь идет о накате волны, которую вдали от побережья можно считать линейной, причем если интересоваться только величинами максимальных уровней подъема воды на побережье, то расчетные формулы становятся явными.

Рассмотрим, например, решение (9), отвечающее накату монохроматической волны:

$$\Phi(\sigma, \lambda) = A J_0(l\sigma) \cos l\lambda, \quad (11)$$

где  $A$  и  $l$  — произвольные постоянные;  $J_0$  — функция Бесселя. На бесконечности ( $\sigma \rightarrow \infty$ ) поле линейно и связь между всеми величинами становится явной, что позволяет однозначно выразить  $A$  и  $l$  через высоту подходящей к берегу волны  $H$  (на расстоянии  $L = h/\alpha$  от берега) и ее частоту  $\omega$ . Затем, рассматривая (10) и (11) при  $\sigma = 0$  (на подвижном урезе), легко найти экстремумы функций  $\eta$  и  $u$ , характеризующие максимальную высоту наката  $R$  и скорость наката  $U$ :

$$\frac{R}{H} = \left( \frac{\pi\omega}{\alpha} \sqrt{\frac{h}{g}} \right)^{1/2}, \quad U = \frac{\omega R}{\alpha}. \quad (12)$$

Формулы (12), как можно показать, справедливы при  $R > 2H$ .

Мы уже говорили, что преобразования (10) имеют смысл при  $\Delta \neq 0$ . Элементарный расчет показывает, что  $\Delta \neq 0$  только при [8]

$$\text{Gr} = \omega^2 R / g\alpha^2 \leq 1. \quad (13)$$

Отсюда вытекает физический смысл параметра  $\text{Gr}$ , выведенного ранее только из соображений размерности. Если  $\text{Gr} < 1$ , то ре-

шение остается непрерывным всюду и волны выходят на берег без обрушения — режим спокойного подтопления берега (наводнение). В случае  $B\gamma > 1$  в рамках уравнений мелкой воды обязательно наступает градиентная катастрофа и волна должна разрушиться. На самом деле из-за влияния диссипации и дисперсии процесс формирования ударной волны может замедлиться, что приведет к увеличению критического значения  $B\gamma$ . И действительно, анализ экспериментальных данных о разрушении морских волн [9] приводит к величине критического числа  $B\gamma \approx 1,5$ . Этот факт свидетельствует в пользу развитой здесь теории.

Отметим, что из условия (13) следует, что при прочих равных условиях короткие волны обрушиваются раньше, чем длинные, — этот факт физически очевиден. В частности, волны с высотами до 10 м и периодами более 10 мин (волны цунами, приливные волны, штормовые нагоны), как правило, не обрушиваются на типичных шельфах ( $\alpha \lesssim 10^{-2}$ ), в то время как для ветровых волн ( $T \approx 6$  с) характерны значения  $B\gamma > 1$ .

Используя (12) и (13), можно оценить сверху максимальную величину коэффициента усиления волны на шельфе:

$$\max \frac{R}{H} = \frac{2\pi\sqrt{2}}{(8\pi\sqrt[3]{2})^{1/5}} \left(\frac{h}{H}\right)^{1/5} \simeq 2,75 \left(\frac{h}{H}\right)^{1/5}. \quad (14)$$

Так как  $B\gamma > 1$ , то при обрушении из-за диссипации энергии высота наката будет уменьшаться. Подчеркнем, что максимальная высота относительно слабо зависит от параметра нелинейности падающей волны и для типичных условий  $R/H$  не превышает 5—10.

Рассмотрим теперь накат волны произвольной формы. Если, как и ранее, считать, что падающая волна задана далеко от берега, где она линейна, а интересоваться только динамикой подвижного уреза ( $\sigma = 0$ ), то может быть предложена простая схема расчета характеристик наката [10]. Действительно, из выражения (10) при  $\sigma = 0$  вытекает

$$\begin{aligned} u(t) &= F(t + u/g\alpha), \\ \eta(t) &= \psi\left(t + \frac{u}{g\alpha}\right) - \frac{1}{2g} u^2(t), \\ F(\lambda) &= \frac{d\psi}{d\lambda}, \quad \psi(\lambda) = \frac{1}{2g} \frac{\partial \Phi}{\partial \lambda}(\lambda, 0). \end{aligned} \quad (15)$$

Если наряду с нелинейной системой (1) рассмотреть линейную систему

$$\partial u / \partial t + g\partial \eta / \partial x = 0, \quad \partial \eta / \partial t - \alpha \partial (xu) / \partial x = 0 \quad (16)$$

и применить к ней линейный вариант преобразований (10), легко выяснить физический смысл функции  $\psi$  — она описывает колебания уровня воды на неподвижном уреze в линейной постановке:  $\psi(t) = \eta_{\text{лин}}(t, x = 0)$  [11]. Следовательно, решение нелинейной системы (1) выражается через решения линейной системы (16)

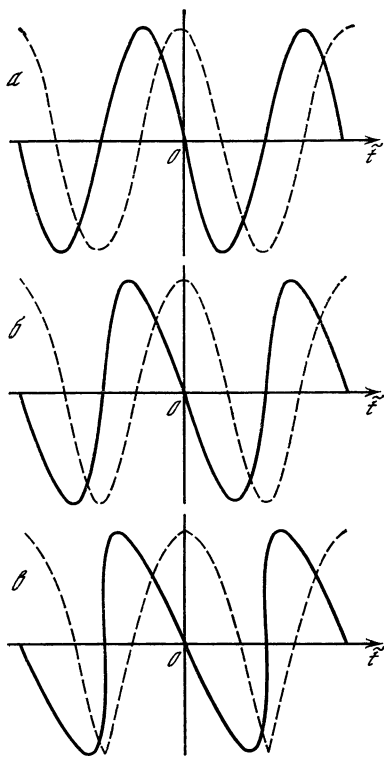


Рис. 3. Осциллограммы колебаний уреза (штриховая линия) и его скорости движения (сплошная линия) при значениях  $B\gamma$ , разных 0 (а), 0,5 (б) и 1 (в)

ределаяет величину вертикального ускорения частиц воды, оно должно быть меньше  $g\alpha^2$ .

С помощью (16) условию (17) может быть придана следующая эквивалентная форма:

$$\max \left. \frac{\partial \eta}{\partial x} \right|_{x=0} \leq \alpha, \quad (18)$$

т. е. крутизна волны на урезе должна быть меньше угла откоса. Последнее условие исходя из наглядных геометрических представлений (при выполнении (18) можно ожидать пересечения уровня воды с береговой линией) уже постулировалось ранее [13], так что здесь мы его строго обосновали [5, 14].

Итак, в случае наката импульсной волны имеют место два процесса: трансформация волны на шельфе (этот процесс может быть полностью описан в рамках линейной теории) и искажение ее формы под действием нелинейности (последнее рассчитывается

с трансформированием масштаба времени, как в римановой волне.

Из (15) также следует, что экстремумы функций  $u$  и  $\eta$  совпадают с экстремумами  $F$  и  $\psi$ ; следовательно, расчет максимальных характеристик наката может быть проведен в рамках линейной теории. Этот вывод [3, 10—12] имеет очень важное значение и позволяет использовать решения классических линейных задач для волн над плоским откосом. Нелинейность влияет на асимметричность формы колебаний уреза, приводя к увеличению крутизны переднего склона волны скорости и формированию скачка производной в осциллограмме уровня при его максимальном понижении (рис. 3). Более того, из (15) можно определить условия однозначности решения — формирования разрыва при  $\partial u / \partial t \rightarrow \infty$ :

$$\frac{\max d^2\psi/dt^2}{g\alpha^2} \leq 1; \quad (17)$$

в случае монохроматической волны условие (17) сводится к  $B\gamma \leq 1$ . Величина  $d^2\psi/dt^2$  оп-

римановым преобразованием координат). Примеры исследования наката импульсных волн приведены в работах [5, 15].

Выше описан процесс наката морских волн в рамках нелинейной теории мелкой воды. Учет влияния дисперсии на этот процесс является весьма сложным делом, и аналитических результатов здесь не известно. Даже в линейной постановке такие задачи являются трудными, и точные формулы получены только для углов, кратных  $\pi/2$  (см., например, [13]). Поскольку вдали от уреза волна всегда является короткой (при  $\alpha = \text{const}$ ), то усиление волны происходит на небольших расстояниях в приустьевой зоне, это приводит к ограничению величины коэффициента усиления волны [13] (близкая формула получается и из приближенной дифференциальной модели [16]):

$$R/H = 2 \sqrt{\pi/2\alpha} \quad (19)$$

(здесь  $\alpha$  есть угол заложения откоса в радианах). Эта формула, в сущности, согласуется с (12), так как в рамках теории мелкой воды  $\omega \sqrt{h/g} \lesssim 1$ , и является верхней границей отношения  $R/H$ .

В этом кратком обзоре суммированы только аналитические результаты (далеко не все). Ясно, что для описания наката морских волн на берег с учетом реальных очертаний береговой линии и неоднородности  $h(x, y)$  необходимо использовать численные методы; соответствующие результаты приведены в работах [4, 5].

#### ЛИТЕРАТУРА

1. Давидан И. Н., Лопатухин Л. И., Рожков В. А. Ветровое волнение в Мировом океане. Л.: Гидрометеоздат, 1985. 255 с.
2. Galvin C. J. Breaker type classification on three laboratory beaches // J. Geophys. Res. 1968. Vol. 73, N 12. P. 3651—3660.
3. Пелиновский Е. Н. Нелинейная динамика волн цунами. Горький: ИПФ АН СССР, 1982. 226 с.
4. Накат цунами на берег / Под ред. Е. Н. Пелиновского. Горький: ИПФ АН СССР, 1985. 215 с.
5. Вольцингер Н. Е., Клеванный К. А., Пелиновский Е. Н. Длинноволновая динамика прибрежной зоны. Л.: Гидрометеоздат, 1988. 415 с.
6. Carrier G. F., Greenspan H. P. Water waves of finite amplitude on a sloping beach // J. Fluid Mech. 1958. Vol. 4, pt 1. P. 97—109.
7. Le Mehaute B., Koh C., Hwang L. S. A synthesis of wave runup // J. Waterways and Harbors Div. Proc. Amer. Soc. Civ. Eng. 1968. Vol. 94, N 1. P. 77—92.
8. Кайстренко В. М., Пелиновский Е. Н., Симонов К. В. Накат и трансформация волн цунами на мелководье // Метеорология и гидрология. 1985. № 10. С. 68—75.
9. Balsillie J. H. Redefinition of shore breaker classification as a numerical continuum and a design shore-breaker // J. Coast. Res. 1985. Vol. 1, N 3. P. 247—254.
10. Мазова Р. Х., Осипенко Н. Н., Пелиновский Е. Н. Влияние нелинейности на характеристики наката длинных волн // Изв. АН СССР. Физика атмосферы и океана. 1987. Т. 23, № 9. С. 950—955.
11. Мазова Р. Х., Пелиновский Е. Н., Шаврацкий С. X. Одномерная теория наката необрушивающихся волн цунами на берег // Процессы возбуждения и распространения цунами. М.: ИОАН, 1983. С. 98—103.

12. Zelt J. A. Tsunamis: The response of harbours with sloping boundaries to long wave excitation // W. M. Keck Laboratory of hydrolics and water resources: Report N KH-R-47. Pasadena, 1986. 218 p.
13. Стоккер Дж. Волны на воде. М.: Изд-во иностр. лит. 1959. 618 с.
14. Shuto N. Standing waves in front of a sloping dike // Coast. Eng. Japan. 1972. Vol. 15. P. 14—21.
15. Иванов В. В., Пономарев Н. А. О накате импульсной волны на берег // Теоретические и экспериментальные исследования длинноволновых процессов. Владивосток, 1985. С. 73—77.
16. Пелиновский Е. Н. Дифференциальная модель волн на воде // ДАН СССР. 1988. Т. 300, № 5. С. 1231—1234.

УДК 539 184(048)

## СВЕРХИЗЛУЧЕНИЕ. МЕТОД ЭЛЕКТРОДИНАМИКИ СПЛОШНЫХ АКТИВНЫХ СРЕД

В. В. ЖЕЛЕЗНЯКОВ, В. В. КОЧАРОВСКИЙ, Вл. В. КОЧАРОВСКИЙ

### ВВЕДЕНИЕ

Явление коллективного спонтанного излучения — сверхизлучения (СИ) — наблюдается в открытых макроскопических образцах возбужденной среды в случае достаточно высокой концентрации  $N$  активных молекул. На возможность его указал Дике в 1954 г. [1]. В простейшем варианте рассматривается достаточно короткий образец двухуровневых молекул, которые в начальный момент времени  $t = 0$  были инвертированы, т. е. переведены из нижнего состояния  $\psi_1$  в верхнее возбужденное состояние  $\psi_2$ . При этом внешнее по отношению к образцу электромагнитное поле на частоте  $\omega_0$  перехода  $2-1$  предполагается отсутствующим, т. е. дальнейшее излучение образца происходит спонтанно.

Явление СИ заключается в том, что распад возбужденного состояния системы молекул, взаимодействующих между собой посредством собственного резонансного излучения, происходит коллективно за время, много меньшее характерного времени спонтанного распада возбужденного состояния изолированной молекулы  $T_1 = 3\hbar c^3/4\omega_0^3 d^2$  ( $d$  — дипольный момент молекулярного перехода,  $\hbar$  — постоянная Планка,  $c$  — скорость света). В результате вся запасенная в образце внутренняя энергия молекул высвечивается в виде короткого импульса квазимонохроматического электромагнитного излучения, схематически показанного на рис. 1. Его мощность  $Q_{\text{СИ}}$  на много порядков превосходит мощность некогерентного спонтанного излучения  $Q_{\text{спонт}}$  такого же числа изолированных молекул и в вытянутых образцах сосредоточена в узком телесном угле вдоль длинной оси образца. Время задержки  $t_d$  импульса СИ обычно превышает его длительность  $\tau$  в 10—20 раз и в известных экспериментах лежит в пределах от единиц до сотен наносекунд.

Исследование СИ Дике началось в СВЧ-диапазоне в неравновесных спиновых системах (см. [2]). Первые эксперименты по оптическому СИ были осуществлены в середине 70-х годов в разреженных газах возбужденных молекул ( $N \sim 10^{10} \div 10^{15} \text{ см}^{-3}$ ) [3—5]. В них использовались вращательные, колебательно-вращательные и электронно-колебательные переходы для получения СИ в далеком ИК, ближнем ИК и видимом диапазонах соответственно. Исходная инверсия активных молекулярных переходов 2—1 достигалась при помощи накачки нерезонансным им лазерным импульсом с использованием какого-либо третьего энергетического уровня молекул. Недавно было получено также СИ на переходах между ридберговскими (высоковозбужденными) состояниями атомов в парах металлов в миллиметровом диапазоне [6] и на экситонных переходах в твердотельных образцах с примесными активными центрами в видимом диапазоне [7, 8].

Трудности реализации СИ по сравнению с созданием обычных лазеров и мазеров связаны с необходимостью быстрого достижения большой плотности инвертированных молекул в условиях достаточно слабой релаксации их высокочастотной поляризации (см. ниже неравенство (13)). Чем привлекательно преодоление этих трудностей? Прежде всего возможностью исключить из схемы сверхизлучательной генерации высокодобротные резонаторы и специальные зеркала, без которых обычные лазеры и мазеры работать не могут [9, 10]. Кроме того, реализация сверхизлучательного режима позволяет получить ценную информацию по спектроскопии и кинетике молекулярных возбуждений [7, 11].

## 1. ПОДХОД ЭЛЕКТРОДИНАМИКИ СПЛЮШНЫХ СРЕД

Очевидно, что большая мощность СИ по сравнению с мощностью некогерентного спонтанного излучения обусловлена эффективной взаимной фазировкой дипольных колебаний молекул, обеспечивающей когерентность их излучения к моменту высвечивания  $t_d$ . Как описать этот процесс? На самом начальном этапе, при  $t \ll t_d$ , процесс самофазировки молекул, предоставленных самим себе, имеет квантостатистический характер [5, 12]. Поэтому для исчерпывающего описания свойств СИ, прежде всего статистических, в принципе необходимо основываться на последовательной квантовой электродинамике системы молекул и поля в вакууме.

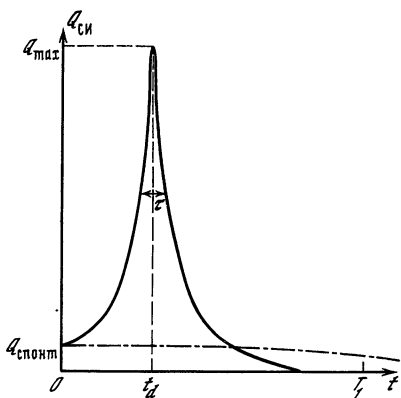


Рис. 1



Начала соответствующей теории СИ были заложены еще в первой работе Дике [1] для модели крупинки, размеры которой малы по сравнению с длиной излучаемой волны:  $L_R \ll \lambda$ , а затем развиты в работах [13—19] и других для протяженных образцов с учетом эффектов запаздывания и переизлучения волн (в частности, с использованием метода обратной задачи рассеяния). Однако ввиду сложности решения операторных уравнений квантовой электродинамики указанный подход реально не дает возможности сколь угодно полно проследить пространственно-временную эволюцию импульса СИ и описать его флуктуационные свойства, а главное — скрывает физический механизм явления.

Вместе с тем с появлением большого числа фотонов процесс СИ приобретает классический характер, так что необходимость прибегать к квантовой электродинамике пропадает и достаточно полуклассического приближения [20, 21]:

$$\operatorname{rot} \mathbf{E} = -c^{-1} \dot{\mathbf{B}}, \quad \operatorname{rot} \mathbf{B} = c^{-1} (\dot{\mathbf{E}} + 4\pi \dot{\mathbf{P}}) + 4\pi \sigma c^{-1} \mathbf{E}, \quad (1)$$

$$\ddot{\mathbf{P}} + 2T_2^{-1} \dot{\mathbf{P}} + (\omega_0^2 + T_2^{-2}) \mathbf{P} = \mathbf{e}(\mathbf{eE}) \omega_c^2 / 4\pi, \quad (2)$$

$$\Delta \dot{N} = -2T_1^{-1} N_2 + 2\mathbf{E} \dot{\mathbf{P}} / \hbar \omega_0. \quad (3)$$

(точка обозначает  $\partial/\partial t$ ). В этом приближении электрическое  $\mathbf{E}$  и магнитное  $\mathbf{B}$  поля описываются классическими уравнениями Максвелла и на поверхности образца удовлетворяют граничным условиям непрерывности тангенциальных составляющих. В (1) учитывается также возможность омической диссипации за счет электропроводности  $\sigma$  «фоновой» среды (для открытых образцов аналогичным фактором служат дифракционные потери излучения через боковую поверхность).

Активная среда для простоты моделируется системой одинаково ориентированных в направлении  $\mathbf{e}$  и равномерно распределенных в образце двухуровневых молекул с однородным уширением спектральной линии ( $T_2^{-1} \gg T_1^{-1}$ ). Коэффициент связи макроскопической поляризации единицы объема  $\mathbf{P}$  с полем  $\mathbf{E}$  в квантовомеханическом уравнении (2) (в пренебрежении поправкой Лоренца к действующему полю) определяется разностью населенностей  $\Delta N = N_2 - N_1$  энергетических уровней молекул в единице объема, точнее, квадратом так называемой кооперативной частоты

$$\omega_c^2 = -8\pi d^2 \omega_0 \hbar^{-1} \Delta N. \quad (4)$$

В инвертированной среде  $\Delta N > 0$  и значение  $\omega_c^2 < 0$ . Поскольку обычно дипольный момент перехода  $d \lesssim 10^{-18}$  ед. СГС и инверсия  $|\Delta N| \ll 10^{20}$  см $^{-3}$ , то  $|\omega_c| \ll \omega_0$  даже в СВЧ-диапазоне. В процессе СИ, когда внешняя накачка отсутствует, значение  $\Delta N$  и с ним  $\omega_c^2$  постепенно меняются вследствие работы поля  $\mathbf{E}$  над током  $\dot{\mathbf{P}}$  (см. (3)). В результате меняется диэлектрическая проницаемость среды

$$\varepsilon(\omega) = 1 + i \frac{4\pi\sigma}{\omega} - \frac{\omega_c^2}{(\omega + iT_2^{-1})^2 - \omega_0^2}, \quad (5)$$

которая определяется в линейном приближении для полей вида  $\mathbf{E} = (\mathbf{E}_0/2) \exp(-i\omega t) + \text{к.с.}$  при  $\mathbf{E} \parallel \mathbf{P} \parallel \mathbf{e}$  с использованием среднего по высокочастотному периоду значения  $\Delta N(t)$ .

Применение для анализа СИ уравнений (1)–(3) и выражения (5) для диэлектрической проницаемости, в которых уже учтена квантовая (двухуровневая) природа молекулярных осцилляторов, отвечает переходу к классической электродинамике сплошных сред (при  $N^{-1/3} \ll \lambda$ ). На этом пути, как было показано в [21], может быть достигнуто ясное понимание физического механизма СИ Дике, а также установлено существование аналогичных процессов в классической электронике. Если наряду с указанным подходом привлечь процедуру феноменологического квантования неустойчивых нормальных волн в активной среде, то удастся упростить и развить описание макроскопических проявлений коллективных квантовых флуктуаций в СИ [22–24].

## 2. МОДЕЛЬ ДИКЕ — СВЕРХИЗЛУЧЕНИЕ КРУПИНКИ

Эффект СИ проще всего объяснить [21] в пределе крупинчатого образца объемом  $V \ll \lambda^3 = (2\pi c/\omega_0)^3$ . В этом случае, считая для простоты крупинку шарообразной и полагая  $\sigma = 0$ , можно воспользоваться решениями двух известных электродинамических задач: об излучении в вакууме точечного высокочастотного диполя  $\mathbf{P}\dot{V} = (\mathbf{P}_0 V/2) \exp(-i\omega_0 t) + \text{к.с.}$  и о поляризации диэлектрического шарика во внешнем квазистатическом поле. Согласно первой, на молекулы крупинки действует поле реакции их собственного излучения  $\mathbf{E}_{\text{изл}} = 2\mathbf{P}\dot{V}/3c^3$ . Согласно второй, существует резонансное значение диэлектрической проницаемости  $\epsilon_p = -2$ , при котором в шарике возникает конечная однородная поляризация  $\mathbf{P}_0$  и соответствующее внутреннее поле  $\mathbf{E}_0 = 4\pi\mathbf{P}_0/(\epsilon_p - 1)$  под действием сколь угодно малого поля  $\mathbf{E}_{\text{изл}}$ . При этом энергия высокочастотного поля в инвертированной крупинке определяется известной электродинамической формулой и оказывается отрицательной за счет большого отрицательного вклада энергии колебаний поляризации:

$$W = V \frac{|E_0^2|}{16\pi} \frac{d[\omega\epsilon(\omega)]}{d\omega} \approx V \frac{|E_0^2|}{8\pi} \frac{9\omega_0^2}{\omega_c^2} < 0. \quad (6)$$

Вместе с тем мощность радиационных потерь положительна:

$$Q_{\text{изл}} = \omega_0^4 |P_0^2| V^2 / 3c^3 > 0.$$

Поэтому вычисление по формуле  $\omega'' = -Q_{\text{изл}}/2W$ , отвечающей закону изменения энергии  $\dot{W} = -Q_{\text{изл}}$ , приводит к положительному значению инкремента:

$$\omega'' = -\omega_c^2 \omega_0^2 V / 12\pi c^3 > 0. \quad (7)$$

Выражение (7) получено в предположении  $\omega'' \gg T_2^{-1}$ , т. е. без учета некогерентной релаксации поляризации. Ее учет приводит

к дополнительному отрицательному слагаемому  $-T_2^{-1}$  в правой части (7). Отметим, что в газе в условиях случайного теплового движения молекул благодаря их диполь-дипольному взаимодействию значение  $T_2^{-1} \gtrsim \omega_0 |\omega_c / 2\omega_0|^2$  [25], и, следовательно, в газовой крупинке объемом  $V \lesssim 3\pi c^3 / \omega_0^3 \approx (\lambda/3)^3$  СИ невозможно.

Итак, результат (7) означает, что поляризация (и поле) в крупинке возрастает экспоненциально:  $P_0(t) = P_0(0) \exp(\omega''t)$ . Если  $\omega'' \gg T_2^{-1}$ , т. е. плотность инвертированных молекул достаточно высока, то коллективное взаимодействие молекул через поле реакции излучения будет превалировать над некогерентной релаксацией поляризации и произойдет быстрое когерентное высвечивание энергии инвертированных молекул из крупинки за время  $t_d \ll T_2$ . В итоге ясно, что СИ Дике в случае крупинки является следствием диссипативной неустойчивости ее дипольной моды, обладающей отрицательной энергией и поэтому нарастающей за счет потерь энергии на излучение в окружающее пространство: потери на излучение  $Q_{\text{изл}} \sim |E_0^2| \rightarrow$  уменьшение  $W \rightarrow$  рост  $|W| \sim |E_0^2| \rightarrow$  рост  $Q_{\text{изл}} \dots$

В адиабатическом приближении развитие этой диссипативной неустойчивости на временах  $t \ll T_{1,2}$  описывают уравнения [21]

$$\dot{Q}_{\text{изл}} = 2\omega''(t) Q_{\text{изл}}, \quad \Delta \dot{N} = -2Q_{\text{изл}} / V\hbar\omega_0, \quad (8)$$

которые совместно с (4), (7) учитывают изменение внутренней энергии молекул  $\hbar\omega_0\Delta N(t)V/2$  в процессе СИ. Решение уравнений (8) при начальном значении  $Q_{\text{изл}}(0) = Q_{\text{спонт}} = \hbar\omega_0\Delta N_0V/T_1$  приводит к известной форме импульса СИ крупинки [5, 12, 20]:

$$Q_{\text{изл}} = \frac{\hbar\omega_0\Delta N_0V}{4\tau} \text{ch}^{-2}\left(\frac{t-t_d}{2\tau}\right), \quad \Delta N = -\Delta N_0 \text{th}\left(\frac{t-t_d}{2\tau}\right). \quad (9)$$

Длительность импульса СИ  $\tau = [2\omega''(0)]^{-1}$  и время его задержки  $t_d$  (см. рис. 1) фактически задаются начальным числом инвертируемых молекул:

$$\tau = T_1 / \Delta N_0 V, \quad t_d = \tau \ln(\Delta N_0 V). \quad (10)$$

Оно же определяет и превышение максимальной мощности СИ над мощностью некогерентного спонтанного излучения:  $Q_{\text{изл}}(t_d) / Q_{\text{спонт}} = \Delta N_0 V / 4$ . Мощность СИ оказывается прямо пропорциональной квадрату концентрации молекул. Последнее свидетельствует о том, что в процессе СИ все активные молекулярные диполи в крупинке оказываются сфазированными и ведут себя как один большой диполь. Поэтому, в частности, к моменту исчезновения инверсии ( $\Delta N(t_d) = 0$ ), когда в обычных импульсных лазерах и мазерах генерация заведомо прекращалась бы, здесь излучение достигает максимума  $Q_{\text{изл}}(t_d)$  и продолжается до полного девозбуждения системы:  $\Delta N \rightarrow -\Delta N_0$ .

### 3. МОДЕЛЬ ОГРАНИЧЕННОГО ОДНОМЕРНОГО ОБРАЗЦА

При анализе СИ протяженных образцов с длиной  $L \gg \lambda$  необходимо дополнительно учитывать эффекты распространения когерентного излучения в активной среде. Имея в виду одномерную постановку задачи, рассмотрим плоский однородный слой активной среды, занимающий область  $0 < z < L$ , и ограничимся исследованием плоских волн, распространяющихся перпендикулярно к слою, т. е. вдоль оси  $z$  (рис. 2). Такая модель адекватно описывает СИ в длинном открытом цилиндрическом образце с малым поперечным сечением  $S \ll L\lambda$  (иначе может быть существенна сложная структура поля в поперечном сечении). При этом в силу малости числа Френеля  $F = S/L\lambda \ll 1$  следует учитывать дифракционное излучение через боковую поверхность цилиндра, добавляя в (1) к омической проводимости  $\sigma$  эффективную величину дифракционной диссипации поля [12]  $c/6\pi FL$ . Впрочем, в условиях фиксированной структуры поля в поперечном сечении образца дифракция может быть и неважна, как, например, в задаче о СИ в одномодовом световоде, активированном примесными молекулярными центрами.

Решение поставленной задачи в определенных условиях можно представить как суперпозицию симметричного набора плоских неоднородных волн вида  $E = (E_{\perp}/2) \exp(-i\omega t + ikz) + \text{к. с.}$ , комплексные частоты  $\omega = \omega' + i\omega''$  и волновые числа  $k = k' + ik''$  которых удовлетворяют внутри образца дисперсионному уравнению  $\omega^2 \epsilon(\omega) = c^2 k^2$ , получающемуся из (1) и (2).

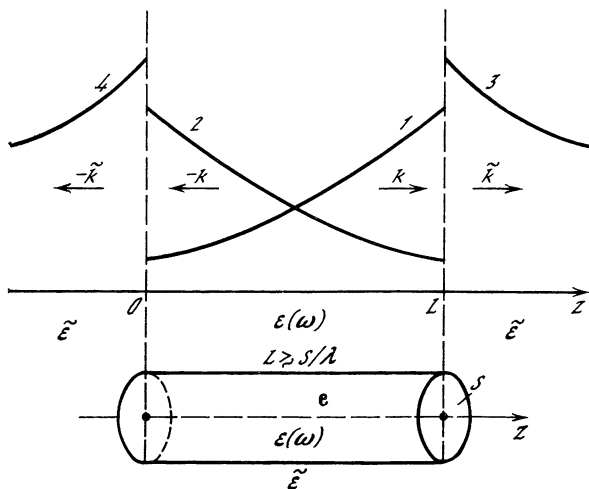


Рис. 2. Пространственная структура поля моды ограниченного одномерного образца длиной  $L \geq S/\lambda$

$$1 - E_{\perp}^{(1)} \exp(ikz), \quad 2 - E_{\perp}^{(2)} \exp(-ikz), \quad 3 - E_{\perp}^{(3)} \exp(i\tilde{k}z), \quad 4 - E_{\perp}^{(4)} \exp(-i\tilde{k}z)$$

#### 4. НОРМАЛЬНЫЕ ВОЛНЫ В ДВУХУРОВНЕВОЙ СРЕДЕ

Выясним сначала дисперсионные свойства однородных ( $k = \text{Re } k$ ) волн в безграничной двухуровневой среде при фиксированной инверсии  $\Delta N$ . Согласно (5), существуют волны двух типов -- электромагнитная волна и волна поляризации<sup>1</sup> [21]:

$$\omega_{e,p} = \omega_0 - iT_2^{-1} + \frac{1}{2} [ck - \omega_0 + i(T_2^{-1} - 2\pi\sigma)] \times \\ \times \left\{ 1 \pm \left[ 1 + \frac{\omega_c^2}{[ck - \omega_0 + i(T_2^{-1} - 2\pi\sigma)]^2} \right]^{1/2} \right\}. \quad (11)$$

В отсутствие инверсии спектр волн (11) известен как поляритонный; он изучался в физике твердых тел [26, 27], а также газов [28]. Для инвертированной среды спектр принимает вид, изображенный на рис. 3, где для конкретности выбраны определенные численные значения параметров. Неустойчивость волн возникает при  $-\omega_c^2 > 8\pi\sigma T_2^{-1}$  и наиболее эффективна в центре линии при  $ck = (\omega_0^2 + T_2^{-2})^{1/2}$ , где инкременты максимальны:

$$\omega_{e,p}'' = -T_2^{-1} + 1/2(T_2^{-1} - 2\pi\sigma) \{1 \pm [1 - \omega_c^2/(T_2^{-1} - 2\pi\sigma)^2]^{1/2}\}. \quad (12)$$

Существенно, что неустойчивой может быть только одна из двух волн -- электромагнитная или волна поляризации -- в зависимости от соотношения между диссипативными параметрами  $2\pi\sigma \leq T_2^{-1}$  (на рис. 3 в каждом из этих случаев неустойчивая волна выделена жирной линией).

Традиционные представления о мазерной (лазерной) неустойчивости [2, 20] относятся к электромагнитной волне  $\omega_e(k)$  в условиях сильной релаксации поляризации и слабой диссипации поля ( $T_2^{-1} \gg |\omega_c| \gg 2\pi\sigma$ ), когда плотность инвертированных молекул сравнительно невелика и  $\omega_e' \approx ck$ . Тогда резонансное значение  $\omega_e' \approx -\omega_c^2 T_2/4 - 2\pi\sigma$ , а наличие волны поляризации с  $\omega_p'' \approx -T_2^{-1}$  несущественно; и в результате при  $\omega_e'' > 0$  реализуется обычный режим *суперлюминесценции*, т. е. вынужденного усиления спонтанного электромагнитного излучения несфазированных возбужденных молекул.

В этом случае фотон, спонтанно излученный на одном конце образца, по мере распространения в активной среде вызывает (индуцирует) излучение фотонов остальными возбужденными мо-

<sup>1</sup> Рассматриваем только интересующую нас область резонанса  $|\omega_0 - ck| \ll \ll \omega_0$ , считая, что  $T_2^{-1}, 2\pi\sigma, |\omega_c| \ll \omega_0$ , и для определенности ограничиваясь обычно реализуемым случаем  $T_2^{-1} \gg \omega_0 |\omega_c/2\omega_0|^2$ , когда поляризуемость активной среды заведомо мала, т. е.  $|\varepsilon - 1| \ll 1$ , и амплитуда поляризации в волнах  $|\dot{P}_\perp| \ll |E_\perp| \approx |B_\perp|$ . Название «волна поляризации» для одной из дисперсионных ветвей (11) связано с тем, что в ней амплитуда поляризации активной среды больше (обычно много больше), чем в электромагнитной волне при той же амплитуде поля  $E_\perp$ .

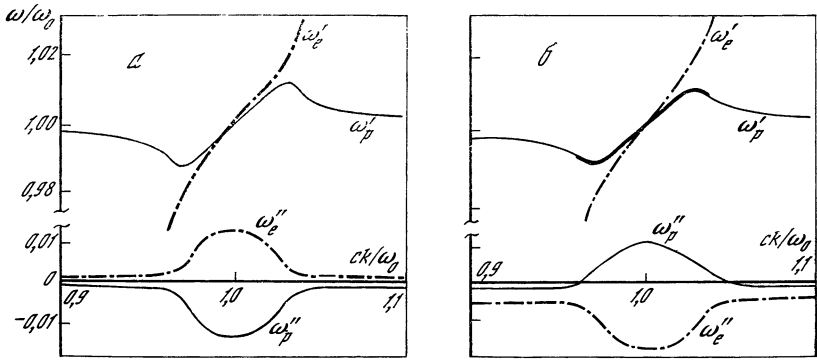


Рис. 3. Поляритонный спектр нормальных волн в инвертированной среде при  $-\omega_c^2 = 10^{-3} \omega_0^2 \gg T_2^{-2} = 10^{-6} \omega_0^2$

а — в отсутствие диссипации ( $\sigma = 0$ ); б — для диссипации  $\sigma = 5 T_2^{-1} / 2\pi$

лекулами, причем каждый из вновь родившихся фотонов сам становится причиной появления новых фотонов. В итоге интенсивность излучения на другом конце образца может значительно превышать суммарную интенсивность некогерентного спонтанного излучения такого же числа отдельных, изолированно излучающих молекул. Однако в отличие от СИ процесс суперлюминесценции развивается медленно в масштабе времени  $T_2$  некогерентной релаксации дипольных колебаний отдельных молекул, и поэтому возникающая в образце поляризация оказывается сравнительно небольшой, а связанное с ней коллективное поведение молекул проявляется слабо. Его описание не представляет затруднений, поскольку благодаря большой ширине спектральной линии  $T_2^{-1} \gg \gg \omega_e''$  систему (1)–(3) можно свести к балансным уравнениям переноса спектральной интенсивности излучения. При этом дифференциальное уравнение (2) заменяется алгебраическим соотношением, например  $P_{\perp} = -iT_2 d^2 \hbar^{-1} \Delta N E_{\perp}$  для медленных амплитуд бегущих электромагнитных волн на частотах вблизи резонанса  $\omega_0$ .

Напротив, эффект СИ характерен для образцов с большой концентрацией активных молекул, когда

$$T_2^{-1} \ll |\omega_c| \quad (13)$$

и в силу быстротечности (когерентности) процесса в масштабе  $T_2$  необходим учет собственной пространственно-временной динамики колебаний поляризации среды. В этом случае поле в образце уже нельзя представить в виде набора некогерентных, спектрально неперекрывающихся компонент (спектрально некоррелирующих фотонов), поскольку в силу существенной нестационарности процесса эффективная ширина спектральных компонент излучения (порядка  $\max \omega_{e,p}$ ) превышает ширину спектральной линии сре-

ды  $T_2^{-1}$ . Согласно (12), в интересующих нас условиях эффективного отбора энергии от инвертированного образца за счет диссипации поля ( $2\pi\sigma > T_2^{-1}$ ) решающую роль играет развитие неустойчивости волны поляризации. Резонансное значение ее инкремента имеет вид

$$\omega_p'' \approx -\omega_c^2/8\pi\sigma - T_2^{-1} \quad (14)$$

при  $|\omega_c| \ll 2\pi\sigma$ , когда вместо уравнений Максвелла (1) фактически достаточно ограничиться соотношением  $\dot{\mathbf{P}} = -\sigma\mathbf{E}$  (ср. (7)), и достигает максимальной величины  $\omega_p'' \approx |\omega_c|/2$  при очень большой плотности инвертированных атомов, когда  $|\omega_c| \gg 2\pi\sigma$ .

Подчеркнем, что всюду, говоря о СИ как о спонтанном процессе, мы имеем в виду лишь соответствующую постановку задачи с отсутствующим внешним полем (см. введение), тогда как по отношению к каждой молекуле речь идет, конечно, об индуцированном излучении под действием коллективного (самосогласованного) поля остальных молекул. Поясним сказанное, используя соотношение энергетического баланса

$$2\omega''|E_{\perp}^2|/8\pi = \hbar\omega_0\Delta N\rho - \sigma|E_{\perp}^2|/2.$$

Согласно ему, скорость роста энергии поля определяется конкуренцией омических потерь и индуцированного высвечивания двухуровневых молекул.

Вероятность  $\rho$  индуцированного перехода одной молекулы с верхнего уровня на нижний в единицу времени находится из уравнений (3) и (2):

$$\rho = \text{Im}(-\omega E_{\perp}^* P_{\perp})/2\hbar\omega_0\Delta N = -\omega_c^2|E_{\perp}^2|/16\pi\omega''\hbar\omega_0\Delta N.$$

Здесь для простоты пренебрегается некогерентной релаксацией поляризации ( $T_2^{-1} \ll \omega''$ ) и рассматривается резонансный случай:  $k = \omega_0/c$ ,  $\omega' = \omega_0$ . Как видим, вероятность индуцированного высвечивания молекул определяется спектральной плотностью излучения, обратно пропорциональной инкременту  $\omega''$ . Из соотношения энергетического баланса после подстановки в него величины  $\rho$  имеем  $\omega'' = -\omega_c^2/4\omega'' - 2\pi\sigma$ . Отсюда при  $|\omega_c| \ll 2\pi\sigma$  приходим к выражению (14) (без  $T_2^{-1}$ ), а при  $|\omega_c| \gg 2\pi\sigma$  получается предельное значение инкремента  $\omega_p'' \approx |\omega_c|/2$ . Отличие этого результата от соответствующего результата для суперлюминесценции, известного в теории лазеров, объясняется тем, что ширина спектра сверхизлучения  $\Delta\omega \approx \omega_p''$  превышает релаксационную ширину перехода  $T_2^{-1}$ . Последнее исключает возможность применения стандартного выражения для вероятности  $\rho$  индуцированного перехода молекулы, получающегося по теории возмущений для случая монохроматического поля [2, 20]. Отметим, что в типичных оптических экспериментах по СИ значение  $|\omega_c| \approx 10^{10} \text{ с}^{-1}$ .

## 5. ФИЗИЧЕСКИЙ СМЫСЛ НЕУСТОЙЧИВОСТИ ВОЛНЫ ПОЛЯРИЗАЦИИ

Покажем, что неустойчивость волны поляризации представляет собой диссипативную неустойчивость волны отрицательной энергии. Волновые неустойчивости такого рода хорошо известны, например, в физике плазмы [29], электродинамике [30, 31] и гидродинамике [32]. В рассматриваемом случае характер неустойчивости легко выясняется на основе закона изменения длины вектора Блоха  $s_B = (P \sqrt{1 + (\omega_0 T_2)^{-2}}, -\dot{P}/\omega_0, \Delta N d)$  и закона изменения плотности энергии среды и поля, которые следуют из (2), (3) и (1), (3) соответственно:}

$$\begin{aligned} \frac{\partial}{\partial t} \left[ (\Delta N d)^2 + \left( \frac{\dot{P}}{\omega_0} \right)^2 + \left( 1 + \frac{1}{\omega_0^2 T_2^2} \right) P^2 \right] = \\ = - \frac{4(\dot{P})^2}{T_2 \omega_0^2} + \frac{\omega_c^2 \hbar N_2}{2\pi T_1 \omega_0}; \end{aligned} \quad (15)$$

$$\frac{\partial}{\partial t} \left( \frac{E^2 + B^2}{8\pi} + \frac{\hbar \omega_0 \Delta N}{2} \right) = - \frac{\hbar \omega_0 N_2}{T_1} - \sigma E^2 - \frac{c}{4\pi} \operatorname{div} [\mathbf{E}, \mathbf{B}]. \quad (16)$$

Чтобы определить знаки плотности энергии  $w$  и плотности мощности потерь  $Q$  для линейных волн в среде с фиксированным значением  $\Delta N = \Delta N_0$ , следует сделать в (15) замену  $\partial (\Delta N^2)/\partial t \approx \approx 2\Delta N_0 \partial (\Delta N)/\partial t$ , найти производную  $\partial (\Delta N)/\partial t$  и подставить ее в (16). Тогда после усреднения по высокочастотному периоду закон изменения плотности энергии для медленных амплитуд поля и поляризации плоской волны, вообще говоря, неоднородной вдоль оси  $z$ , примет вид

$$\dot{w} = -Q - (c/8\pi) \partial \operatorname{Re}[E_{\perp} B_{\perp}^* \exp(2\omega''t - 2k''z)]/\partial z, \quad (17)$$

$$\begin{aligned} w = [(|E_{\perp}^1|^2 + |B_{\perp}^1|^2)/16\pi + \pi \omega_c^{-2} (|\omega^2| + \omega_0^2 + T_2^{-2}) |P_{\perp}^2|] \times \\ \times \exp(2\omega''t - 2k''z), \end{aligned} \quad (18)$$

$$Q = (\sigma |E_{\perp}^2|^2/2 + 4\pi \omega_c^{-2} T_2^{-1} | \omega^2 P_{\perp}^2 |) \exp(2\omega''t - 2k''z). \quad (19)$$

Найденная волновая энергия (18) отвечает приближенному значению усредненной суммы  $(E^2 + B^2)/8\pi + \hbar \omega_0 \Delta N/2$  (см. (16)) в линейной задаче и включает как энергию электромагнитного поля, так и энергию колебаний поляризации молекул. В свою очередь, потери этой энергии определяются омической (или дифракционной) диссипацией электрического поля и столкновительной (или спонтанной) релаксацией поляризации, а также неоднородностью потока энергии вдоль направления распространения волны. Используя решение (11) с  $k = k' + ik''$ , можно убедиться, что в инвертированной среде, где  $\omega_c^2 < 0$ , плотность энергии волны поляризации отрицательна ( $w_p < 0$ ), а электромагнитной волны — положительна ( $w_e > 0$ ). Поэтому, согласно (17), при отборе энергии из образца вследствие диссипации или излу-



чения поля <sup>2</sup> амплитуда волны поляризации будет увеличиваться, т. е. будет происходить фазировка и нарастание дипольных колебаний инвертированных молекул [21]. Указанный процесс подобен рассмотренному в разд. 3 для дипольной моды инвертированной крупинки.

Отрицательный знак энергии волны поляризации свидетельствует о том, что при ее возбуждении у среды отбирается часть энергии. Это связано с тем, что только при переводе первоначально инвертированных молекул в суперпозиционное состояние  $\psi = C_1\psi_1 + C_2\psi_2$  с  $C_{1,2} \neq 0$ , т. е. только при заселении нижнего энергетического уровня  $\psi_1$ , у молекул можно создать высокочастотный дипольный момент  $\int (C_1\psi_1) \text{er} (C_2\psi_2)^* dV + \text{к. с.}$ . Поэтому если наряду с появлением поляризации при указанном отборе энергии молекул возникает не слишком большая энергия электромагнитного поля, то в целом совершается отрицательная работа и, следовательно, среда приобретает *отрицательную волновую энергию*. Согласно (18), именно такая ситуация имеет место для волны поляризации.

## 6. СВЕРХИЗЛУЧЕНИЕ МОД С ДИСКРЕТНЫМ СПЕКТРОМ

Рассмотрим теперь неустойчивость волн поляризации в модели ограниченного одномерного образца (см. разд. 3, рис. 2). В этом случае наличие частичных, даже малых отражений электромагнитного поля от торцов цилиндрического образца приводит к тому, что из непрерывного спектра самовозбуждающихся волн выделяются отдельные моды с дискретным спектром [21]. Последние являются собственными решениями уравнений (1), (2) при  $\Delta N = \text{const}$  и внутри однородного активного слоя ( $0 < z < L$ ) представляют собой суперпозицию двух встречных линейно-поляризованных волн вида

$$E = 1/2 E_{\perp}^{(1)} \exp(-i\omega t + ikz) + 1/2 E_{\perp}^{(2)} \exp(-i\omega t - ikz) + \text{к. с.}, \quad (20)$$

где  $k = \sqrt{\varepsilon(\omega)}\omega/c$  (см. (5)). На границах слоя указанное электрическое поле сшивается с полями волн  $(E_{\perp}^{(3),(4)}/2)\exp(-i\omega t \pm i\tilde{k}z) + \text{к. с.}$ , уходящих от слоя во внешнюю среду с диэлектрической проницаемостью  $\tilde{\varepsilon} = \text{const}$ , где  $\tilde{k} = \sqrt{\tilde{\varepsilon}}\omega/c$ . Аналогичным условиям непрерывности подчиняется магнитное поле.

Эти граничные условия приводят к следующим соотношениям между амплитудами волн:

$$\begin{aligned} E_{\perp}^{(1),(2)} &= (\sqrt{\varepsilon(\omega)} \mp \sqrt{\tilde{\varepsilon}}) E_{\perp}^{(4)}/2\sqrt{\varepsilon(\omega)}, \\ E_{\perp}^{(3)} &= E_{\perp}^{(1)} \exp[i\omega Lc^{-1}(\sqrt{\varepsilon(\omega)} - \sqrt{\tilde{\varepsilon}})] + \\ &+ E_{\perp}^{(2)} \exp[i\omega Lc^{-1}(\sqrt{\varepsilon(\omega)} + \sqrt{\tilde{\varepsilon}})] \end{aligned} \quad (21)$$

<sup>2</sup> Отбор энергии в случае однородной волны поляризации с  $k'' = 0$  означает, что соответствующие потери (19)  $Q_p > 0$ . Это условие для волн резонансной частоты  $\omega_0$  выполняется при  $T_2^{-1} < 2\text{лс} < -\omega_c^2 T_2/4$ .

и к характеристическому уравнению для дискретного спектра мод  $\omega_n$ :

$$\omega_n \sqrt{\varepsilon(\omega_n)} = \frac{\pi n c}{L} - \frac{ic}{L} \ln \left[ \frac{\sqrt{\varepsilon(\omega_n)} - \sqrt{\bar{\varepsilon}}}{\sqrt{\varepsilon(\omega_n)} + \sqrt{\bar{\varepsilon}}} \right], \quad (22)$$

где  $n$  — целое число.

Если в левую часть (22) подставить выражение (5) для диэлектрической проницаемости, то станет ясно, что неполное отражение поля от торцов образца, описываемое последним слагаемым в (22), играет роль, аналогичную диссипации поля в образце. В результате оказывается, что в открытом слое существуют уже известные нам два типа мод с неоднородной пространственной структурой (20) ( $\text{Im}k \neq 0$ ) — электромагнитные и поляритонные. Их дисперсионные свойства в резонансном приближении  $|\omega_0 - \pi n c/L| \ll \omega_0$  (см. сноску 1) аналогичны свойствам электромагнитных волн и волн поляризации безграничной среды и описываются формулой (11) с дискретными значениями  $k_n = \pi n/L$ , если заменить в ней  $\sigma$  на

$$\sigma_{\text{эф}} = \sigma + \sigma_{\text{изл}}, \quad 2\sigma_{\text{изл}} = (c/2L) \ln R^{-1}, \\ R = |(\sqrt{\varepsilon(\omega_n)} - \sqrt{\bar{\varepsilon}})/(\sqrt{\varepsilon(\omega_n)} + \sqrt{\bar{\varepsilon}})|^2. \quad (23)$$

(Для простоты пренебрегаем малым смещением частот мод из-за наличия мнимой части коэффициента отражения  $R$ .)

Величина  $\sigma_{\text{изл}}$  характеризует диссипацию поля мод за счет излучения через торцы образца и поэтому входит аддитивно с величиной  $\sigma$ , определяющей омические и дифракционные потери энергии. Интегрируя закон изменения энергии (17) по объему образца, легко показать, что в инвертированном образце энергия поляритонных мод отрицательна. Следовательно, СИ является результатом развития их диссипативной неустойчивости, возникающей при наличии положительных потерь энергии, когда  $T_2^{-1} < < 2\pi\sigma_{\text{эф}} < -\omega_c^2(0)T_2/4$ . Согласно оценкам, выделение и преобладание СИ мод с дискретным спектром над СИ волн с непрерывным спектром происходит уже при малом коэффициенте отражения  $R \gtrsim (NLS)^{-1/4}$ , обычно при  $R \approx 10^{-2} \div 10^{-4}$ .

Подчеркнем, что здесь речь идет о не слишком длинных и не слишком коротких образцах. Ограничение сверху имеет вид  $L \lesssim c/|\omega_c(0)|$ . Оно обеспечивает получение импульса СИ простой формы и возможность пренебрежения запаздыванием излучения, приходящего из разных частей образца за время развития неустойчивости. Иначе, т. е. в более длинных образцах, коллективное спонтанное излучение молекул на различных участках происходит независимо, так что к моменту эффективного снятия инверсии образуется случайная последовательность бегущих импульсов, а пространственная структура дискретных мод не успевает установиться. Ограничение снизу следует из требования существования сильной неустойчивости мод ( $\omega_p \gg T_2^{-1}$ , см. (14) и (23)) и при

$\sigma = 0$  имеет вид <sup>3</sup>

$$2\pi\sigma_{\text{изл}} \equiv (c/2L) \ln R^{-1} \ll -\omega_c^2(0)T_2/4. \quad (24)$$

В указанной ситуации при условии эффективного излучения через торцы образца ( $2\pi\sigma_{\text{изл}} \gg |\omega_c(0)|$ ) определяющую роль обычно играет одна диссипативно неустойчивая мода с номером  $n \approx 2L/\lambda \gg 1$  и инкрементом  $\omega_p'' \approx -\omega_c^2/8\pi\sigma_{\text{изл}} \propto |\Delta N|$ , а процесс СИ мало чем отличается от рассмотренного в разд. 2 для крупинки. В известном приближении «среднего поля» [12, 21], использующем усредненную по объему образца разность населенностей  $\overline{\Delta N}(t)$  для определения мгновенного значения инкремента поляритонной моды  $\omega_p''(t) \propto \overline{\Delta N}$ , вновь справедливы адиабатические уравнения вида (8) с заменой  $\Delta N \rightarrow \overline{\Delta N}$ ,  $Q_{\text{изл}}/V \rightarrow Q$ , где  $Q(t)$  — мощность излучения из образца в расчете на единицу его объема. В результате, если в момент инверсии молекул мощность излучения была задана на флуктуационном уровне  $Q(0) \ll \ll Q_{\text{max}} = \hbar\omega_0\overline{\Delta N}_0\omega_p''(0)/2$ , то спустя время задержки  $t_d = = \tau \ln [Q_{\text{max}}/Q(0)] \gg L/c$ , необходимое для коллективной спонтанной фазировки диполей, возникает мощный короткий импульс СИ (9) с длительностью  $\tau = [2\omega_p''(0)]^{-1} = \hbar c \ln(R^{-1})/8\pi L\omega_0 d^2 \overline{\Delta N}_0$ , определяемой первоначальной инверсией  $\overline{\Delta N}_0$ . Аналогичная картина имеет место и для трехмерных образцов, например в форме шара [21].

Отметим, что в слое с хорошо отражающими граничными зеркалами ( $R \approx 1$ ) или в слое с периодическими граничными условиями (модель активного световода, замкнутого в кольцо), где пространственная структура мод однородна ( $\text{Im } k \approx 0$ ), быстрая коллективная релаксация инвертированных молекул посредством диссипативной неустойчивости не может быть обеспечена излучением через торцы ( $\sigma_{\text{изл}} \approx 0$ ). Однако в подобном резонаторе она будет возможна при наличии омической (или дифракционной) диссипации поля. Его осциллограмма тогда примет вид либо указанного на рис. 1 одноимпульсного процесса с длительностью  $\tau \approx 4\pi\sigma/|\omega_c^2(0)|$  (при  $2\pi\sigma \gtrsim |\omega_c(0)| \gg T_2^{-1}$ ), либо медленно релаксирующего осцилляторного процесса — так называемой оптической нутации вектора Блоха с периодом  $T \sim 4\pi/|\omega_c(0)|$  (при  $|\omega_c(0)| \gg 2\pi\sigma$ ,  $T_2^{-1}$ ) [33]. При определяющей роли омических потерь соответствующий процесс перехода энергии активной среды в импульс омического нагрева аналогичен СИ и может быть назван *сверхпоглощением*. Можно показать, что осциллирующий характер процессов сверхпоглощения и СИ при условии  $|\omega_c(0)| \gg 2\pi\sigma_{\text{эф}}$  связан с неадиабатическим взаимодействием поляритонных и электромагнитных мод в образце с быстро меняющейся инверсией  $\Delta N(t)$ .

<sup>3</sup> При нарушении условия (24), т. е. для более коротких образцов с  $R \ll 1$ , СИ формируется волнами непрерывного спектра (см. разд. 7 ниже).

## 7. ОДНОНАПРАВЛЕННОЕ СВЕХИЗЛУЧЕНИЕ ВОЛН НЕПРЕРЫВНОГО СПЕКТРА

Рассмотрим теперь случай предельно малого коэффициента отражения ( $R \rightarrow 0$ ), когда за время развития СИ  $t_d$  не успевает произойти выделение мод с дискретным спектром и ими, следовательно, можно пренебречь.

Оговоримся, что даже в отсутствие отражений от торцов образца одновременное существование встречных волн может приводить к изменению инверсии среды  $\Delta N(t, z)$  в пространстве и во времени, вызывающему взаимное влияние встречных волн друг на друга. Однако в реальных условиях это обстоятельство обычно не очень существенно, поскольку встречные волны начинают снижать инверсию с противоположных концов образца, причем их СИ зажигается в различные (случайные) моменты времени.

Поэтому СИ встречных волн можно описывать независимо на значительных промежутках времени вплоть до момента столкновения двух встречных фронтов снятия инверсии, после которого СИ фактически прекращается.

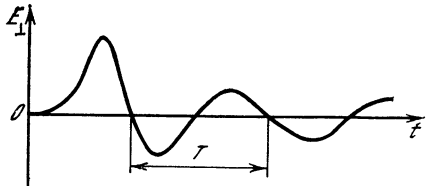


Рис. 4

В результате возникает задача об однонаправленном СИ, в которой рассматриваются волны, бегущие только в одном направлении  $+z$  вдоль оси цилиндрического образца, начиная с точки  $z = 0$  (см. рис. 2). Она сводится к уравнению-синус Гордона с диссипацией.

Отсылая за подробностями к работам [12—17, 24, 34—36], мы приведем только качественный результат исследования процесса однонаправленного СИ на временах  $t \ll T_{1,2}$  в случае  $|\omega_c(0)| \ll c/L$ ,  $2\pi\sigma$ . Оказывается, что развивающаяся неустойчивость волн имеет абсолютный характер и приводит к формированию плавно меняющихся самосогласованных профилей разности населенностей  $\Delta N(t, z)$  и амплитуд поля  $E_{\perp}(t, z)$  и поляризации  $P_{\perp}(t, z)$ . Их временные профили (и соответствующая осциллограмма дифракционного излучения из цилиндрического образца) на малом расстоянии от края образца  $z \ll (c/2\pi\sigma) \ln \varphi_0^{-1}$  имеют вид релаксирующих осцилляций (рис. 4) с периодом порядка  $T = 4\pi c \ln(\varphi_0^{-1}/z) |\omega_c^2(0)|$ , задаваемым начальным значением кооперативной частоты (4), а на большем расстоянии приобретают одноимпульсный характер с типичными значениями длительности  $\tau \sim [2\omega_p''(0)]^{-1} = 4\pi\sigma / |\omega_c^2(0)|$  и времени задержки  $t_d \sim \tau \ln \varphi_0^{-2}$ .

Выше введен начальный полярный угол вектора Блоха  $\varphi_0 \sim (NLS)^{-1/2} \ll 1$ , определяющий флуктуационный уровень поляризации в образце ( $|P_{\perp}|_{t=0} = d\Delta N_0 \sin \varphi_0$ ). Импульс СИ простой

формы <sup>4</sup> здесь, как и в разд. 6, получается именно в результате диссипативной неустойчивости волн поляризации, обусловленной эффективными омическими или дифракционными потерями  $2\pi\sigma \gg |\omega_c(0)|$ . При нарушении последнего условия, а также в случае очень длинных образцов  $L \gg c/|\omega_c(0)|$  СИ имеет сложный осциллирующий характер.

## 8. ЦИКЛОТРОННОЕ СВЕРХИЗЛУЧЕНИЕ — КЛАССИЧЕСКИЙ АНАЛОГ СВЕРХИЗЛУЧЕНИЕ ДИКЕ

Трактовка СИ квантовой системы двухуровневых молекул как диссипативной неустойчивости волн поляризации с отрицательной энергией позволяет предположить, что явление СИ возможно и в классических системах, где существуют волны с аналогичными свойствами. Примером такой классической системы служит симметричная одномерная модель цилиндрической магнитной ловушки длиной  $L$  (геометрию см. на рис. 2), в которой имеются два встречных моноскоростных потока электронов (с концентрацией  $N_e$  и энергией  $\mathcal{E}$ ), летящих вдоль однородного магнитного поля  $B_0$  параллельно оси  $z$  в условиях аномального эффекта Доплера. Последнее означает, что скорость электронов  $v_{\parallel}$  превышает фазовую скорость электромагнитных волн в среде  $c_0 = c/\sqrt{\epsilon_0}$ , роль которой может играть замедляющая электродинамическая система или диэлектрик с проницаемостью  $\epsilon_0 > 1$ .

В этой ситуации возможен эффект циклотронного СИ [38, 39] — поток классических электронных осцилляторов в магнитном поле может спонтанно излучать короткие когерентные импульсы в модах с дискретным спектром, выделяющихся вследствие отражения поля на торцах ловушки. Максимальная мощность импульсов, как и в случае СИ Дике, пропорциональна квадрату электронной концентрации:  $Q_{\max} \propto N_e^2$ . Механизмом циклотронного СИ является диссипативная неустойчивость так называемой медленной циклотронной волны отрицательной энергии, развивающаяся за счет потери энергии электромагнитного поля — омической или дифракционной, обусловленной высвечиванием излучения через боковую поверхность ловушки.

Соответствующий спектр нормальных волн — электромагнитной  $\omega_I(k) = \omega'_I + i\omega''_I$  и медленной циклотронной  $\omega_{II}(k) = \omega'_{II} + i\omega''_{II}$  — описывается дисперсионным уравнением, подобным таковому для двухуровневой среды, и схематически изображен на рис. 5 при  $\omega_B \gg \omega_c$ , где  $\omega_B = eB_0/mc$  — гирочастота электронов,  $\omega_c = \sqrt{4\pi e^2 N_e}/m$  — плазменная частота ( $e$  и  $m$  — величина заряда и релятивистская масса электрона).

<sup>4</sup> Следует отметить также возможность управления формой импульса СИ и инициирования самого процесса СИ посредством введения в образец в начальный момент времени  $t = 0$  слабого затравочного импульса резонансного излучения [4, 37].

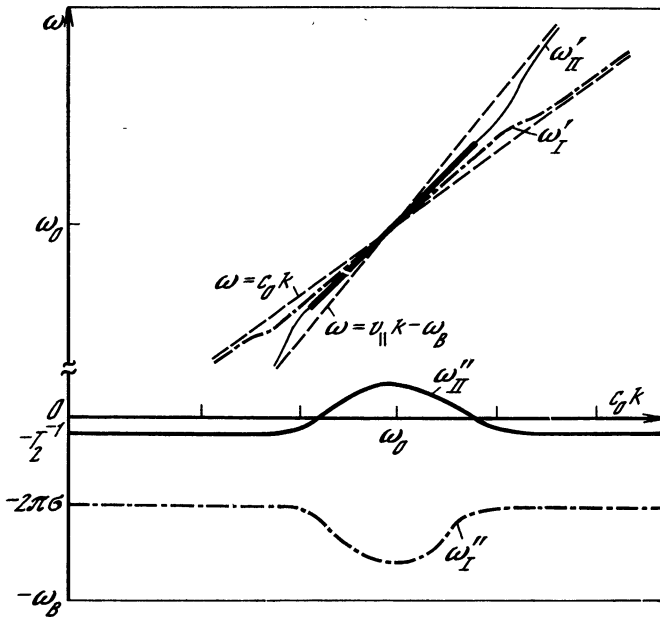


Рис. 5

Характерная форма импульса циклотронного СИ в случае неустойчивости одной <sup>5</sup> медленной циклотронной моды ловушки с коэффициентом отражения  $R = 1$  представлена на рис. 6. Максимальная мощность циклотронного СИ в расчете на единицу объема ловушки равна  $Q_{\max} = \mathcal{E} N_e \omega_{\perp}^2 (v_{\parallel} - c_0) / \omega_0 (\epsilon_0 - 1) \epsilon_0 c_0$ , где  $\omega_0 = \omega_{BC_0} / (v_{\parallel} - c_0)$  — частота излучения. Длительность импульса порядка нескольких  $\tau = 2\pi\sigma c_0 / \omega_{\perp}^2 (v_{\parallel} - c_0)$ , время задержки  $t_d = \tau \ln [Q_{\max} / Q(0)]$ . Обе величины  $\tau$  и  $t_d$  малы по сравнению с временем некогерентной релаксации циклотронных колебаний электронов  $T_2$ . Более подробный анализ нелинейного этапа диссипативной неустойчивости в классической системе осцилляторов, аналогичного нелинейной стадии СИ Дике, см. в работе [38].

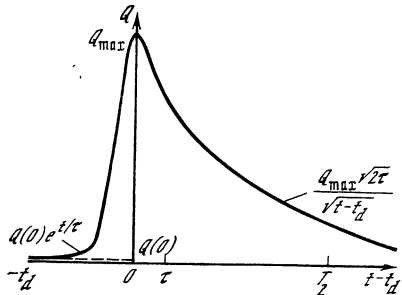


Рис. 6

Наряду с общими свойствами циклотронного СИ и СИ Дике существует, конечно, ряд различий этих двух процессов. Наиболее

<sup>5</sup> В случае одновременной неустойчивости нескольких медленных циклотронных мод СИ происходит в режиме «эстафеты мод», отвечающем равнозамедленному движению электронного потока и распространению мощности излучения вверх по частотному спектру мод ловушки [38].

значительное различие связано с типом нелинейности, определяющей форму импульса СИ. Для квантовой системы двухуровневых молекул — это нелинейность типа насыщения, обусловленная изменением разности населенностей их энергетических уровней и одинаково влияющая на инкременты поляритонных мод с различными волновыми числами. Для классической системы движущихся электронов в магнитном поле — это нелинейность резонансного типа, вызванная доплеровским изменением частоты колебаний электронов при их торможении под действием поля и оказывающая выборочное влияние на инкременты медленных циклотронных мод с различными волновыми числами.

Следует подчеркнуть, что установленная конкретная аналогия между СИ Дике в системе квантовых молекулярных осцилляторов и циклотронным СИ в системе классических электронных осцилляторов, движущихся в магнитном поле в условиях аномального эффекта Доплера, отнюдь не исчерпывает, а фактически лишь ставит проблему СИ (коллективного спонтанного излучения) в классической электронике и физике плазмы. В частности, согласно [40], возможность реализации коллективного спонтанного излучения движущегося сгустка электронов имеется также в лазерах на свободных электронах. При этом представляет интерес излучение электронным сгустком как дискретных мод, так и волн непрерывного спектра, в том числе в режиме однонаправленного высвечивания импульса СИ.

### ЗАКЛЮЧЕНИЕ

На наш взгляд, развитые выше представления электродинамики сплошных активных сред в отношении явления СИ как диссипативной неустойчивости волн отрицательной энергии в рассмотренных и других случаях позволяют не только установить физический механизм этого явления, но и значительно продвинуться в его анализе.

Согласно изложенному, в ограниченных активных образцах с конечным, пусть небольшим коэффициентом отражения процесс СИ, ведущий к возникновению мощных коротких импульсов простой формы, обусловлен диссипативной неустойчивостью поляритонных мод с дискретным спектром. Последняя обязана отрицательной энергии таких мод и ее эффективной диссипации за счет потерь на излучение из образца. При достаточно малом коэффициенте отражения ( $R \rightarrow 0$ ) определяющей становится генерация волн непрерывного спектра. В этом случае запасенная в образце энергия молекул обычно высвечивается в режиме однонаправленного СИ. Этот режим в зависимости от величины дифракционных (или омических) потерь в образце реализуется либо на электромагнитных волнах (при  $2\lambda\sigma < T_2^{-1}$ ), либо на волнах поляризации (при  $2\lambda\sigma > T_2^{-1}$ ).

Предложенный подход позволяет корректно учесть влияние на СИ отражений от границ образца, геометрические факторы,

роль нелинейного взаимодействия мод в формировании профиля импульса СИ и др. Кроме того, если квантовать не осцилляторы поля в вакууме, как это обычно делается в квантовой электродинамике, а непосредственно моды активного образца, то удастся существенно упростить решение и квантовых задач, например установление статистики времени задержки импульса СИ, определяемой квантовыми флуктуациями в начале развития процесса СИ [22—24].

#### ЛИТЕРАТУРА

1. Dicke R. H. Coherence in spontaneous radiation processes // Phys. Rev. 1954. Vol. 93, N 1. P. 99—110.
2. Ярус А. Квантовая электроника. М.: Сов. радио, 1980. 488 с.
3. Skribanowitz N., Hermann I. P., Mac Gillivray J. C., Feld M. S. Observation of Dicke superradiance in optically pumped HF gas // Phys. Rev. Lett. 1973. Vol. 30, N 8. P. 309—312.
4. Schuurmans M. F. H., Vreken Q. H. F., Polder D. Superfluorescence // Advances in atomic and molecular physics / Ed. D. Dates, B. Bederson. N. Y.; L.: Acad. press, 1981. Vol. 17. P. 167—228.
5. Андреев А. В., Емельянов В. И., Ильинский Ю. А. Коллективное спонтанное излучение (сверхизлучение Дике) // УФН. 1980. Т. 131, № 4. С. 653—694.
6. Moi L, Goy P., Gross M. et al. Rydberg-atom masers // Phys. Rev. A. 1983. Vol. 27, N 4. P. 2043—2081.
7. Пабойкин Ю. В., Самарцев В. В., Зиновьев П. В., Силаева Н. В. Когерентная спектроскопия молекулярных кристаллов. Киев: Наук. думка, 1986. 204 с.
8. Florian R., Schwan L. O., Schmid D. Time-resolving experiments on Dicke superfluorescence of  $O_2^-$  centers in KCl. Two-color superfluorescence // Phys. Rev. A. 1984. Vol. 29, N 5. P. 2709—2715.
9. Андреев А. В., Ильинский Ю. А., Хохлов Р. В. О роли коллективных и индуцированных процессов при генерации мессбауэровского  $\gamma$ -излучения // ЖЭТФ. 1977. Т. 73, № 4. С. 1296—1300.
10. Baldwin G. C., Feld M. S. Kinetics of nuclear superradiance // J. Appl. Phys. 1986. Vol. 59, N 11. P. 3665—3671.
11. Carlson N. W., Jackson D. J., Schawlow A. L. Superradiance triggering spectroscopy // Opt. Commun. 1980. Vol. 32, N 2. P. 350—354.
12. Gross M., Haroche S. Superradiance: An essay on the theory of collective spontaneous emission // Phys. Rep. 1982. Vol. 93, N 5. P. 301—396.
13. Haake F., King H., Schroder G. et al. Fluctuations in superfluorescence // Phys. Rev. A. 1979. Vol. 20, N 5. P. 2047—2063.
14. Haake F., King H., Schroder G. et al. Delay-time statistics of superfluorescence pulses // Ibid. 1981. Vol. 23, N 3. P. 1322—1333.
15. Prasad S., Glauber K. J. Initiation of superfluorescence in a large sphere // Ibid. 1985. Vol. 31, N 3. P. 1583—1597.
16. Mostowski J., Sobolewska B. Initiation of superfluorescence from a sphere // Ibid. 1983. Vol. 28, N 5. P. 2943—2952.
17. Mostowski J., Sobolewska B. Three-dimensional theory of initiation of superfluorescence // Ibid. 1984. Vol. 30, N 3. P. 1392—1400.
18. Рунасов В. И. К теории сверхизлучения Дике: точное решение квазиодномерной квантовой модели // ЖЭТФ. 1982. Т. 83, № 5. С. 1711—1724.
19. Худсон В. И. К динамике интегрируемых квантовых систем // Там же. 1985. Т. 88, № 5. С. 1757—1770.
20. Файн В. М. Фотоны и нелинейные среды. М.: Сов. радио, 1972. 472 с.
21. Железняков В. В., Кочаровский В. В., Кочаровский Вл. В. Эффект сверхизлучения и диссипативная неустойчивость в инвертированной двухуровневой среде // ЖЭТФ. 1984. Т. 87, № 5. С. 1565—1581.
22. Kocharovsky V. V., Kocharovsky Vl. V. Superradiance statistics for three-dimensional samples // Opt. Commun. 1985. Vol. 53, N 5. P. 345—348.



23. *Кочаровский В. В., Кочаровский Вл. В.* Квантовая динамика диссипативной неустойчивости поляритонных мод: Препр. ИПФ АН СССР № 110. Горький, 1984. 35 с.
24. *Железняков В. В., Кочаровский В. В., Кочаровский Вл. В.* Волны поляризации и сверхизлучение в активных средах // УФН. 1989. Т. 159.
25. *Климонтович Ю. Л.* Кинетическая теория электромагнитных процессов. М.: Наука, 1980. 376 с.
26. *Агранович В. М., Гинзбург В. Л.* Кристаллооптика с учетом пространственной дисперсии и теория экситонов. М.: Наука, 1979. 432 с.
27. Кристаллы / Под ред. Б. И. Веркина, А. Ф. Прихотько. Киев: Наук. думка, 1983. 526 с.
28. *Железняков В. В., Кочаровский В. В., Кочаровский Вл. В.* Пространственная дисперсия в газе двухуровневых молекул: Препр. ИПФ АН СССР № 126. Горький, 1985. 67 с.
29. *Арцимович Л. А., Сагдеев Р. З.* Физика плазмы для физиков. М.: Атомиздат, 1979. 320 с.
30. *Люиселл У.* Связанные и параметрические колебания в электронике. М.: Изд-во иностр. лит., 1963. 352 с.
31. *Незлин М. В.* Динамика пучков в плазме. М.: Энергоиздат, 1982. 264 с.
32. *Островский Л. А., Рыбак С. А., Цимринг Л. Ш.* Волны с отрицательной энергией в гидродинамике // УФН. 1986. Т. 150, № 3. С. 417—437.
33. *Кочаровский В. В., Кочаровский Вл. В.* К теории самосогласованной оптической нутации // Изв. вузов. Радиофизика. 1985. Т. 28, № 9. С. 1099—1111.
34. *Mac Gillivray J. C., Feld M. S.* Limits of superradiance as a process for achieving short pulses of high energy // Phys. Rev. A. 1981. Vol. 23, N 3. P. 1334—1348.
35. *Mac Gillivray J. C., Feld M. S.* Theory of superradiance in an extended, optically thick medium // Ibid. 1976. Vol. 14, N 3. P. 1169—1189.
36. *Габутов И. Р., Захаров В. Е., Михайлов А. В.* Нелинейная теория суперфлуоресценции // ЖЭТФ. 1984. Т. 86, № 4. С. 1204—1216.
37. *Варнаевский О. П., Куркин А. Н., Леонтович А. М.* и др. Когерентное усиление ультракоротких импульсов в активированных кристаллах // Там же. С. 1227—1239.
38. *Железняков В. В., Кочаровский В. В., Кочаровский Вл. В.* Циклотронное сверхизлучение — классический аналог сверхизлучения Дике // Изв. вузов. Радиофизика. 1986. Т. 29, № 9. С. 1095—1116.
39. *Zheleznyakov V. V., Kocharovskiy V. V., Kocharovskiy Vl. V.* The cyclotron superradiance in plasma physics and electronics: A classical analog of Dicke's superradiance // Proc. Contrib. Pap. Intern. Conf. Plasma Phys. Kiev: Nauk. dumka, 1987. Vol. 4. P. 111—114.
40. *Bonifacio R., Casagrande F.* The superradiant regime of a free electron laser // Nucl. Instrum. and Meth. Phys. Res. A. 1985. Vol. 239, N 1. P. 36—42.

УДК 523.4

## ПРОИСХОЖДЕНИЕ И ЭВОЛЮЦИОННАЯ ДИНАМИКА КОЛЕЦ УРАНА

Н. Н. ГОРЬКАВЫЙ, А. М. ФРИДМАН

Узкие, эллиптические кольца Урана — один из самых загадочных астрономических объектов. В настоящем виде их можно считать аттракторами. Но как они образовались в околопланетном диске Урана на его протостадии?

В данной работе подводятся итоги десятилетней дискуссии о природе колец. Обсуждаются проблемы образования и динамики колец Урана: связь с линдбладовскими резонансами от внешних спутников; возникновение эксцентриситета и устойчивость наблюдаемых структур.

## 1. ГИПОТЕЗЫ О ПРИРОДЕ КОЛЕЦ УРАНА

**Свойства колец Урана.** Кольца Урана открыты 10 марта 1977 г. по затмению звезды (рис. 1) [1, 2]. Узкие и плотные колечки, разделенные широкими, практически пустыми зонами, совсем не похожи на классические кольца Сатурна. У большинства урановских колечек обнаружен эксцентриситет  $e$ : частицы, образующие кольцо, расположены вдоль эллиптической орбиты с  $e \approx 10^{-(2+3)}$ . Ширина колечек зависит от азимутального угла и в апоцентре может быть в несколько раз больше, чем в перицентре. Ширина у восьми колец 1—10 км, а у кольца  $\epsilon$  от 20 км (в перицентре) до 100 км (в апоцентре). Края колец очень резкие; некоторые кольца имеют постоянную внутреннюю структуру [3, 4].

Многие свойства колец казались непонятными. Во-первых, узкое кольцо вследствие обмена угловым моментом между сталкивающимися частицами должно быстро расплываться (за считанные десятилетия), увеличивая свою ширину и уменьшая резкость краев [5]. Во-вторых, несферичность поля Урана вызывает прецессию эллиптических орбит, скорость которой зависит от величины большой полуоси. Дифференциальная прецессия частиц на внешнем и внутреннем краях кольца за несколько сотен лет должна разрушить эллиптическое узкое кольцо (и образовать круговое и более широкое) [5]. Тем не менее кольца не расплываются и сохраняют эллиптичность, прецессируя как единое целое.

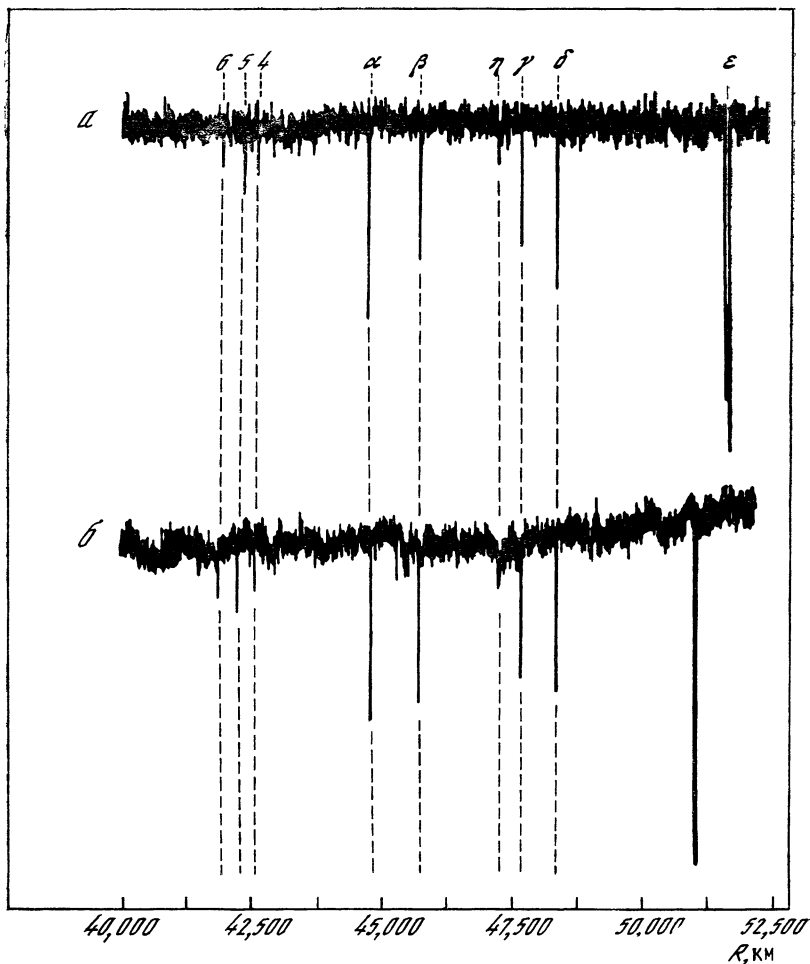
Наиболее важными вопросами происхождения и динамики колец можно считать следующие:

1. Как образовались кольца Урана? Что собрало околопланетное вещество в узкие кольца, далеко отстоящие одно от другого?

2. Как возник эксцентриситет у колец?

3. Почему кольца устойчивы?

**Первые гипотезы о связи с известными спутниками.** В статье [1] высказана идея, что расстояния между кольцами объясняются резонансами с известными спутниками Урана. В работе Дермотта и Голда [6] положения пяти колец (идентифицированных в 1977 г. и обозначенных  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ) сопоставлялись с серией трехчастотных резонансов от Ариэля—Титании и Ариэля—Оберона. Предполагалось, что частицы «застревают» на резонансных орбитах при движении к планете. Позже Акнес [7] и Голдрайх и Никольсон [8] показали, что в зоне колец более значительны трехчастотные резонансы, связанные с Мирандой, а не Ариэлем; при этом даже самые сильные трехчастотные резонансы могут контроли-



**Рис. 1.** Изменение светимости звезды при затмении ее кольцами Урана [3] *a* — до покрытия звезды диском Урана; *b* — после покрытия. Слабые кольца 4, 5, 6 и  $\eta$  идентифицированы на данных регистрограммах позже. По оси абсцисс — расстояние от центра планеты (в км), по оси ординат — интенсивность прошедшего света звезды. Профили эллиптического кольца  $\epsilon$  смещены вследствие наблюдений под разными азимутальными углами (со стороны Урана)

рывать движение частиц в очень узкой зоне (в несколько десятков метров), меньшей, чем ширина колец. Но даже критики работы [6] (см. [7, 8]) не отказываются от идеи резонансной природы колец.

Акснес [7] высказывает мысль, что только определенные типы резонансов могут тормозить вещество, движущееся к планете, а также упоминает замечание Колумбо о приблизительных ре-

зональных соотношениях между самими кольцами. В статьях [8, 9] высказывается гипотеза, что кольца Урана — это сильно-нелинейные спиральные волны, резонансно возбуждаемые в оптически тонком диске. Стейгман [10, 11] модифицирует гипотезу Дермотта—Голда, связывая расположение колец с трехчастотными резонансами от Миранды—Ариэля и Миранды—UVI — не открытого спутника на орбите с радиусом 105 221 км. Но в 1978 г. обнаружены еще четыре кольца Урана ( $\eta$ , 4, 5, 6), и сопоставлять положение всех девяти колец с трехчастотными резонансами от внешних спутников стало трудно. С учетом критических замечаний [7, 8] это серьезно подорвало позиции резонансных гипотез (в том числе и модели Дермотта—Голда).

**Гипотезы о неизвестных «спутниках-пастухах».** В 1979 г. появилась серия гипотез, предполагающих наличие в зоне колец нескольких неоткрытых спутников. Авторы [12, 13] считали, что в каждом кольце расположено по спутнику, при этом кольца — или постоянно возобновляемые газообразные «шлейфы» невидимых спутников [12], или скопления частиц на сложных, «банановых» орбитах вблизи спутника [13]. Голдрайх и Тримейн [5] предположили, что каждое кольцо расположено между двумя «спутниками-пастухами», не дающими частицам кольца расплываться. Введение «пастухов» давало также возможность объяснить эксцентриситет колец [14]. Устойчивость кольца к дифференциальной прецессии объяснялась силами самогравитации [15]. В ноябре 1980 г. «Вояджер-1» открыл возле узкого, эллиптического кольца F Сатурна два «спутника-пастуха» (Пандора и Прометей), после чего идея «спутников-пастухов» получила самое широкое распространение.

**Гипотеза о резонансной природе колец Урана, определяемой серией неоткрытых спутников.** В работе [16] (см. также [17]) развивается концепция, по которой существование колец связано с интенсивным столкновительным разрушением частиц вблизи планеты. Близость планеты определяет значительную величину относительных сдвиговых скоростей частиц в дифференциально-вращающемся диске ( $\Omega a_0$ , где  $\Omega = \sqrt{GM_p/R^3}$  — угловая скорость вращения,  $a_0$  — размер частицы). Внешняя граница  $R_r$  колец определяется равенством  $\Omega a_0 \sim \sqrt{Gm/a_0}$  [16] (здесь  $\sqrt{Gm/a_0}$  — скорость убегания с поверхности частиц [18]), что дает  $R_r \sim (M_p/\rho)^{1/3}$ . Внутри этой границы конденсационный рост частиц ограничен несколькими метрами [16], а вне ее не останавливается вплоть до образования спутников. Таким образом, кольца из множества макрочастиц образуются там, где формирование спутников запрещено. Сосуществование колец и спутников «вперемешку» возможно только в достаточно узкой зоне между областями колец и спутников (пример — кольцо F с «пастухами» на краю колец Сатурна).

Модель «спутников-пастухов», предполагающая существование по всей зоне колец до 18 спутников, явно противоречит данной концепции образования колец. С другой стороны, в начале 80-х го-

дов в кольцах Сатурна открыто несколько узких (иногда с заметным эксцентриситетом) колец, связанных не с «пастухами», а с резонансами от внешних спутников (см. [19, 20], а также рис. 1 из [17]). Это ставит под сомнение обязательность модели «пастухов» и для колец Урана.

В [21] выдвинута гипотеза, по которой положения колец Урана соответствуют линдбладовским резонансам низшего порядка (1 : 2, 2 : 3, 3 : 4, ...) от серии неоткрытых спутников за внешней границей колец. Серьезным доводом в пользу данной гипотезы явилось обнаружение удивительной закономерности в расположении колец Урана: вблизи внешней границы колец существует несколько орбит, каждая из которых резонансна сразу с парой колец. Отметим, что именно возле внешней границы колец Юпитера и Сатурна открыты серии мелких спутников. Естественно ожидать, что аналогичная серия спутников существует и возле колец Урана, хотя бы в силу достаточно равномерного распределения вещества в протодиске. По расположению колец в [21] были вычислены радиусы орбит тех невидимых спутников, каждый из которых определяет расположение двух колец сразу. Для объяснения особенностей кольца  $\epsilon$  было предположено, что вблизи его внешнего края находится «спутник-пастух».

## 2. ПРОЛЕТ «ВОЯДЖЕРА-2» ВБЛИЗИ УРАНА В ЯНВАРЕ 1986 г.

Американский аппарат «Вояджер-2» открыл 10 новых спутников и ряд более слабых колец (табл. 1) [22—24], и тем самым оказалось возможным подвести итоги дискуссии о природе колец. Внутри зоны колец спутников не обнаружено, за исключением одного, самого мелкого, вблизи внешней границы зоны колец (аналогичного внутреннему «пастуху» кольца F у Сатурна). Таким образом, только одно из девяти главных колец оказалось окруженным «спутниками-пастухами». С другой стороны, «Вояджер» подтвердил гипотезу о резонансной природе колец Урана (орбиты части открытых спутников хорошо совпали с орбитами, вычисленными ранее по расположению колец [17, 25, 26]; см. табл. 1).

Рис. 2 демонстрирует корреляцию между резонансами низшего порядка от новых спутников Урана и кольцами. Вверху совмещены диаграмма, характеризующая распределение колец (заштрихованные столбцы), и диаграмма резонансного влияния (сплошная линия). Высота заштрихованных столбцов для каждого 1000-километрового интервала  $H_{\text{рез}} = 1,2 (N + N'/\alpha_r)$ , где  $N$  — число классических колец (плюс кольцо 1986U1R);  $N'$  — число слабых колец, открытых «Вояджером»;  $\alpha_r$  — отношение значимости (вклада в диаграмму) между этими двумя классами колец. Высота диаграммы резонансного влияния зависит от числа и порядка резонансов:  $H_{\text{рез}} = \sum_j 3/(n_j + m_j)$ , где  $j$  — число резонансов

Таблица 1  
Характеристики спутников и колец Урана

Наименование	Радиусы орбит спутников, км		Резонансы от спутников в зоне колец (36—53)·10 <sup>8</sup> км	Наименование	Радиусы орбит, км [4, 22]	e·10 <sup>8</sup> [4]
	открытых [22, 24]	предсказанных [24, 25]				
Спутники						
1985 U 1	86 000 ± 26		—	ε	51 456,3 ± 4,9	4,924 ± 0,023
1986 U 5	75 258 ± 30		[1 : 2] *	Кольцо 1	50 660 ± 30	?
1986 U 4	69 942 ± 27		[1 : 2; 3 : 5]	1986 U 1 R	50 030 ± 30	?
1986 U 1	66 086 ± 33	66 450	[1 : 2; 3 : 5; 2 : 3]	δ	48 306,1 ± 4,7	(0,020 ± 0,140)
1986 U 2	64 352 ± 27		1 : 2 [3 : 5; 2 : 3]	γ	47 632,3 ± 5,1	(0,121 ± 0,201)
1986 U 6	62 676 ± 25	62 470	1 : 2; [3 : 5; 2 : 3; 3 : 4]	η	47 183,6 ± 5,2	(0,014 ± 0,025)
1986 U 3	61 777 ± 27	61 860	1 : 2; [3 : 5; 2 : 3; 3 : 4]	β	45 668,8 ± 5,0	0,438 ± 0,022
1986 U 9	59 173 ± 27	58 600	1 : 2; [3 : 5; 2 : 3; 3 : 4; 4 : 5]; 5 : 6	α	44 726,6 ± 4,9	0,759 ± 0,026
1986 U 8	53 798 ± 49	55 380	3 : 5; [2 : 3; 3 : 4; 4 : 5; 5 : 6]; 6 : 7 . . . ]	4	42 579,3 ± 4,8	1,065 ± 0,029
1986 U 7	49 771 ± 18	51 580	2 : 3; [3 : 4; 4 : 5; 5 : 6]; 6 : 7 . . . ]	5	42 243,4 ± 4,9	1,900 ± 0,029
				6	41 845,9 ± 4,6	1,001 ± 0,024
				1-й сегмент кольца	41 760 ± 30	?
				2-й сегмент кольца	41 470 ± 30	?
				3-й сегмент кольца	38 430 ± 50	?
				Кольцо 2	38 280 ± 50	?
				1986 U 2 R	37 000—39 500	—

\* Квадратными скобками выделены спутниковые резонансы, попадающие в зону (41—52)·10<sup>8</sup> км.

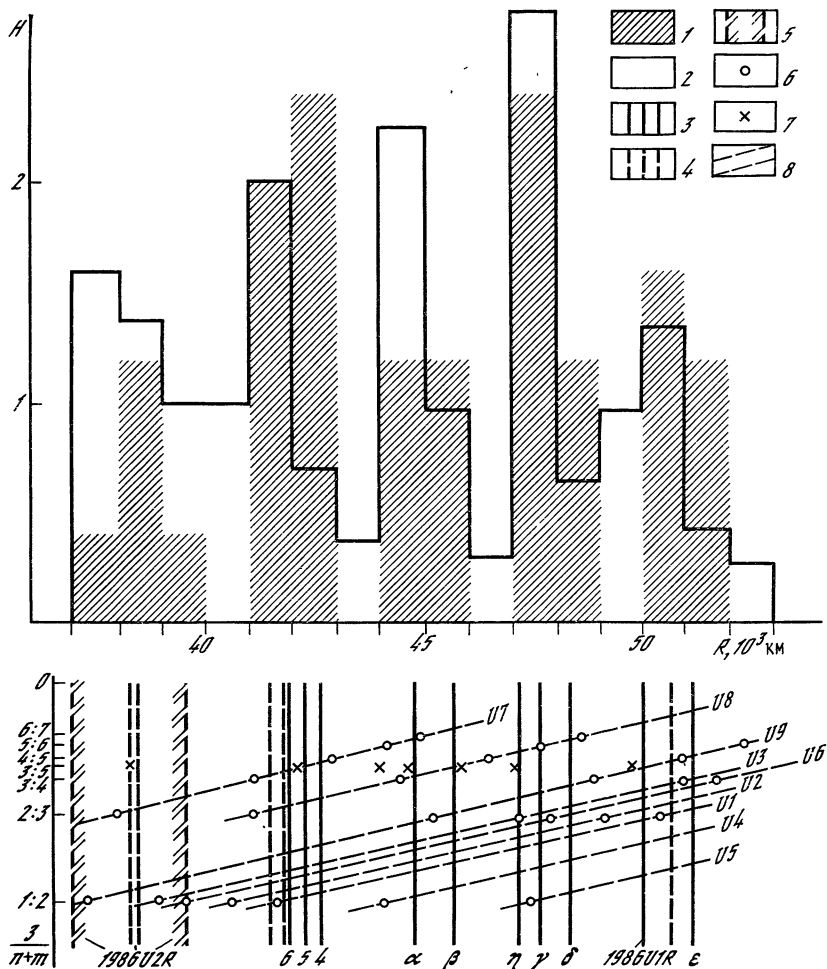


Рис. 2. Связь расположения резонансов и колец

1 — гистограмма распределения колец, 2 — гистограмма распределения резонансов, 3 — классические плотные кольца, 4 — слабые кольца, открытые «Вояджером» в 1986 г., 5 — граница широкого пылевого кольца 1986U2R, 6 — линдбладовские резонансы низшего порядка — 1 : 2, 2 : 3, . . . , 6 : 7, 7 — резонанс 3 : 5, 8 — линии, соединяющие линдбладовские резонансы отдельных спутников

(типа  $n : m$ ) в данном интервале. Внизу показано расположение колец и рассматриваемых резонансов от девяти новых спутников Урана. Боковая ось отражает вклад каждого резонанса в диаграмму (в наименовании спутников опущен год открытия — 1986).

Вычисленный по стандартной методике (см. [27]) коэффициент корреляции между двумя диаграммами на рисунке равен 0,6. Более детальный анализ показывает, что корреляция между кольцами и резонансами определяется главным образом резонансами

Таблица 2

Коэффициент корреляции между кольцами Урана в зоне (36—53)·10<sup>3</sup> км и резонансами низкого порядка от внешних спутников

n	R, 10 <sup>3</sup> км	$\alpha_r$		
		3	2	1
9	36—53	0,59 ± 0,16	0,61 ± 0,15	0,60 ± 0,15
7	36—53	0,71 ± 0,12	0,75 ± 0,11	0,76 ± 0,10
7	40—53	0,78 ± 0,11	0,80 ± 0,10	0,80 ± 0,10

Обозначения. n — число учитываемых спутников, R — ширина рассматриваемой зоны колец,  $\alpha_r$  — отношение вкладов от основных и слабых колец в диаграмму распределения.

типа 1 : 2, 2 : 3, 3 : 4, 3 : 5 и в меньшей степени 4 : 5 и 1 : 3. Резонансы высших порядков на коэффициент корреляции влияют незначительно. Кольца, открытые «Вояджером-2», хорошо соответствуют модели резонансной природы колец Урана: при учете «новых» колец коэффициент корреляции возрастает. Наибольший вклад в коэффициент корреляции вносят резонансы от спутников 1986U1, 1986U9, 1986U6, 1986U3, 1986U7, 1986U8, 1986U5 (спутники расположены по степени уменьшения вклада). Спутники 1986U2 и 1986U4, видимо, не участвовали в формировании колец (возможно, из-за более позднего образования); при исключении резонансов от этих спутников коэффициент корреляции значительно повышается (табл. 2). Особенно велик коэффициент корреляции в зоне наиболее плотных колец (40—53 тыс. км; см. табл. 2). Проведенный выше анализ переводит гипотезу о резонансной природе колец Урана в разряд наблюдательных фактов. Естественно возникает проблема выяснения механизма образования колец вблизи резонансов.

### 3. МОДЕЛЬ РЕЗОНАНСНОГО ПРОИСХОЖДЕНИЯ КОЛЕЦ УРАНА

Очевидно, что на протостадии вокруг Урана существовали не отдельные узкие кольца, а достаточно однородный диск макрочастиц. Поверхностную плотность его можно оценить, считая, что она близка к средней плотности в соседней зоне мелких спутников: 10—20 г/см<sup>2</sup> (это сравнимо с плотностью кольца С Сатурна). Рассмотрим влияние резонансных внешних спутников на эволюцию протоколец Урана.

**Спиральные волны и возникновение резонансных щелей.** На первый взгляд два факта противоречат гипотезе резонансного накопления колец:

а) зона резонансного влияния спутника очень узка: естественная ширина резонанса [9, 28]

$$\Delta L \sim R_L (M_s/M_p)^{1/2} \approx \begin{cases} 30 \text{ км для спутника Сатурна Мимаса,} \\ 4 \text{ км для спутника Урана с } a_s = 50 \text{ км;} \end{cases} \quad (1)$$



здесь  $R_L$  — резонансный радиус,  $M_s/M_p$  — отношение масс спутника и планеты;

б) локальное резонансное воздействие крупного тела на скопление мелких тел, которое ранее ассоциировалось с уменьшением плотности: щель Кассини — следствие резонанса 1 : 2 от Мимаса, деления Кирквуда — следствие резонансов от Юпитера в поясе астероидов.

Однако представления о локальности резонансного возмущения неверны для среды макрочастиц с коллективными свойствами (давление, самогравитация и т. д.). В таких средах резонансное возмущение распространяется от точки резонанса в виде спиральной волны на многие сотни километров [9, 28]. Например, щель Кассини, образовавшаяся на месте спиральной волны, простирается на 4,5 тыс. км, что на два порядка превышает ширину резонанса (1) [9]. Спиральная волна, унося угловой момент частиц диска, вызывает дрейф частиц к планете и образование щели при массе спутника больше критической [29]:

$$\frac{M_s}{M_p} \gtrsim \frac{1}{n} \frac{c}{\Omega_s R_s} \sqrt{\frac{\tau}{\tau^2 + 1}}. \quad (2)$$

Здесь  $\tau$  — оптическая толщина колец;  $c$  — скорость хаотического движения частиц;  $\Omega_s$  и  $R_s$  — угловая частота и орбитальный радиус спутника;  $n$  характеризует порядок линдбладовского резонанса типа  $n/(n + 1)$ .

При меньших массах спутника щель не образуется — она закрывается из-за диффузии сталкивающихся частиц быстрее, чем освобождается при резонансном выметании. Если масса спутника достаточно мала:

$$M_s/M_p \lesssim (c/\Omega_L R_L)^2 \quad (3)$$

(индекс  $L$  обозначает характеристики резонансной орбиты), то резонансным воздействием на диск можно пренебречь, так как амплитуда резонансных движений (1) сравнима с толщиной колец  $h \approx c/\Omega$  ( $c$  — амплитудой тепловых флуктуаций движения частиц). Из (2), (3), можно показать, что ближайšie к планете спутники Сатурна и Урана с радиусом 1 км  $\lesssim a_s \lesssim 200$  км вызывают спиральные волны, не образующие щели (фактически только Мимас с  $a_s \approx 200$  км образовал щель Кассини, генерируя спиральную волну плотности). Такие спиральные волны в большом количестве обнаружены в кольцах Сатурна [28]. Отметим, что спиральные волны плотности распространяются наружу от резонансной орбиты в плоскости диска, а акустические волны и спиральные волны, изгибающие плоскость колец (и соответствующие резонансу спутника с частотой вертикальных осцилляций частиц [28]), движутся от точки резонанса к планете.

**Спиральные волны и накопление колец.** Спиральные волны от новых спутников Урана с радиусом 10—50 км не могли вызвать заметного дрейфа макрочастиц в протокольцах Урана. Следова-

тельно, накопление узких колечек не может быть связано с данным резонансным перемещением частиц. Но в протокольцах Урана существовал другой важнейший фактор переноса вещества: сильный поток пыли к планете, унесший основную массу протодиска. Самым существенным фактором торможения и падения пыли является аэродинамическое трение, приводящее к выметанию микронных частиц за 100—1000 лет [30]. Как взаимодействует мощный планетоцентрический поток пыли со спиральными волнами?

В зоне спиральной волны могут возникнуть два встречных потока пыли. Как показывают современные теоретические исследования [31], в области линдбладовских резонансов рождаются четыре типа волн: 1) спиральные волны плотности, распространяющиеся наружу — от планеты; 2) акустические волны, распространяющиеся к планете; 3) стоячие волны вследствие вязкости; 4) изгибные волны,двигающиеся к планете. Первые два типа волн способны захватывать и переносить мелкие частицы, «выгружая» их в радиальной области уменьшения амплитуды, т. е. в радиусе нескольких сот километров от резонансной орбиты. Какая из этих волн преобладает по амплитуде, зависит от параметров околопланетного диска.

Заметим, что при наличии планетоцентрического потока пыли спиральная волна плотности не только создает встречный ему поток мелких частиц, двигаясь навстречу, но и тормозит основной поток пыли к планете. Отметим также, что зоны волн плотности в кольцах Сатурна выделяются повышенной оптической толщей, связанной, во-первых, с повышенной поверхностной плотностью [28] и, во-вторых, с повышенным содержанием мелкой компоненты, являющимся следствием усиленного разрушения рыхлых частиц в возмущенной зоне [32]. Таким образом, резонансные спиральные волны, вызывая изменения в планетоцентрическом потоке пыли, могут приводить к образованию кольцевых уплотнений вблизи резонанса — на расстоянии, не превышающем характерной длины распространения спиральной волны.

Оценим максимальное характерное время нарастания такого колечка. Пусть радиальная часть протодиска, из которого образовались кольца, занимает область  $(r_1, r_2)$ . К настоящему времени практически все вещество аккрецировало на планету. Количество пыли, аккрецировавшее на планету за время существования протодиска  $t_L$ , есть  $2\pi r_1 \sigma_d v_d t_L$  ( $\sigma_d, v_d$  — поверхностная плотность и радиальная скорость пыли в области внутреннего радиуса  $r_1$ ). Очевидно, это количество пыли не может быть меньше массы протодиска  $\sigma_0 \pi (r_2^2 - r_1^2) \approx \sigma_0 \pi r_2^2$  ( $\sigma_0$  — средняя поверхностная плотность протодиска), поскольку, кроме вещества протодиска, на планету аккрецирует межпланетная пыль за счет, например, эффекта Пойнтинга—Робертсона. Итак,  $2\pi r_1 \sigma_d v_d t_L \geq \sigma_0 \pi r_2^2$ , откуда  $\sigma_d v_d \geq \sigma_0 r_2^2 / 2r_1 t_L$ .

Оценим максимальное время  $t_{\max}$ , за которое поток пыли к планете  $\sigma_d v_d$  сможет создать кольцо с контрастом плотности  $\Delta\sigma$  и

пириной  $\Delta r$ , воспользовавшись равенством  $2\pi r \Delta r \Delta \sigma = 2\pi r v_d \sigma_d t_{\max}$ , откуда

$$t_{\max} = \frac{\Delta r \Delta \sigma}{v_d \sigma_d} \leq \frac{\Delta r \Delta \sigma}{\sigma_0 r_2^2} 2r_1 t_L \approx \frac{\Delta r}{r_2} \frac{\Delta \sigma}{\sigma_0} t_L \approx 2 \cdot 10^{3+5} \text{ лет.} \quad (4)$$

Здесь были использованы следующие числовые значения:  $\Delta \sigma / \sigma_0 \approx \approx 1$ ,  $\Delta r \approx 10$  км,  $r_2 = 5 \cdot 10^4$  км,  $r_1 = 2,6 \cdot 10^4$  км,  $t_L \approx 10^{7+9}$  лет.

В [17, 33] показано, что кольцевые флуктуации в диске с недиффузионными потоками пыли неустойчивы — нарастают (аккреционная неустойчивость). Это относится как к спонтанным крупномасштабным флуктуациям, так и к индуцированным мелкомасштабным уплотнениям, в которых диффузия подавлена (вопрос об устойчивости тонких колец по отношению к диффузионному распылению будет рассмотрен позже).

#### 4. ПОЧЕМУ КОЛЬЦА ЭЛЛИПТИЧНЫ?

**Эксцентриситет колец и внешние силы.** Превращение кругового кольца в эллиптическое соответствует неосесимметричному возмущению с  $m_\varphi = 1$  (возмущение характеристик кольца с периодом  $2\pi$  вида  $\sim \exp(im_\varphi \varphi)$ )<sup>1</sup>. Наблюдаемые характерные амплитуды отклонения кольца от круговой орбиты 40 — 400 км ( $e \approx 10^{-(2+3)}$ ). Амплитуда резонансных возмущений колец Урана на расстоянии

$\Delta r$  от точки Лидбладова резонанса равна [34]  $\Delta r \sim \frac{M_s}{M_p} \frac{R_L^2}{\Delta r} \approx \approx (0,1 \text{ км } (e \approx 2 \cdot 10^{-6}))$ , если  $\Delta r \approx 100$  км. Это показывает, что эксцентриситет колец не связан с возмущениями от внешних спутников.

**Эксцентриситет колец и саморазвитие.** В классической задаче Лапласа—Максвелла круговое абсолютно твердое кольцо вокруг планеты неустойчиво относительно возмущения типа  $m_\varphi = 1$  — кольцо смещается и падает на планету. Устойчиво ли круговое уплотнение в дифференциально-вращающемся диске неупругих частиц относительно возмущений  $m_\varphi = 1$ ? В качестве первого шага к решению данной проблемы рассмотрим устойчивость неосесимметричных возмущений такого диска в линейном приближении. Запишем систему уравнений переноса для плоского вращающегося диска неупругих частиц [35]:

$$\begin{aligned} \frac{\partial \sigma}{\partial t} + \frac{\partial}{r \partial r} (r \sigma v_r) + \frac{\partial}{r \partial \varphi} (\sigma v_\varphi) &= 0, \\ \frac{\partial v_r}{\partial t} + v_r \frac{\partial}{\partial r} v_r + v_\varphi \frac{\partial}{r \partial \varphi} v_r - \frac{v_\varphi^2}{r} &= -\frac{1}{\sigma} \frac{\partial P}{\partial r} - \frac{\partial \Psi}{\partial r} + \\ + \frac{4}{3} \nu \frac{\partial}{\partial r} \frac{1}{r} \frac{\partial}{\partial r} r v_r + \frac{4}{3\sigma} \frac{\partial v_\sigma}{\partial r} \left( \frac{\partial v_r}{\partial r} - \frac{v_r}{2r} \right) &- \end{aligned}$$

<sup>1</sup> Формально возмущение  $m_\varphi = 1$  сдвигает центр тяжести круговой орбиты, что с точностью до членов порядка  $e^2$  ( $e$  — эксцентриситет эллипса) соответствует эллиптической орбите.

$$\begin{aligned}
& -\frac{2}{3\sigma} \frac{1}{r} \frac{\partial}{\partial r} \left( \nu \sigma \frac{\partial v_\varphi}{\partial \varphi} \right) - \frac{4}{3} \nu \frac{1}{r^2} \frac{\partial v_\varphi}{\partial \varphi} + \\
& + \frac{1}{r\sigma} \frac{\partial}{\partial \varphi} \left[ \nu \sigma \left( \frac{1}{r} \frac{\partial v_r}{\partial \varphi} + r \frac{\partial}{\partial r} \frac{v_\varphi}{r} \right) \right], \\
\frac{\partial v_\varphi}{\partial t} + v_r \frac{\partial v_\varphi}{\partial r} + \frac{v_r v_\varphi}{r} + v_\varphi \frac{\partial}{r \partial \varphi} v_\varphi = & -\frac{\partial P}{\sigma r \partial \varphi} - \frac{\partial \Psi}{r \partial \varphi} + \\
& + \frac{1}{\sigma r^2} \frac{\partial}{\partial r} r^3 \nu \sigma \frac{\partial}{\partial r} \frac{v_\varphi}{r} + \frac{1}{\sigma r} \frac{\partial}{\partial r} \left( \nu \sigma \frac{\partial}{\partial \varphi} v_r \right) + \nu \frac{\partial}{r^2 \partial \varphi} v_r + \\
& + \frac{1}{\sigma r} \frac{\partial}{\partial \varphi} \left[ \frac{4}{3} \nu \sigma \left( \frac{1}{r} \frac{\partial v_\varphi}{\partial \varphi} + \frac{v_r}{r} - \frac{1}{2} \frac{\partial v_r}{\partial r} \right) \right], \quad (5) \\
\frac{3}{2} \left( \frac{\partial T}{\partial t} + v_r \frac{\partial}{\partial r} T + v_\varphi \frac{\partial}{r \partial \varphi} T \right) + T \left( \frac{1}{r} \frac{\partial}{\partial r} r v_r + \frac{1}{r} \frac{\partial v_\varphi}{\partial \varphi} \right) = \\
= \frac{1}{\sigma r} \frac{\partial}{\partial r} \left( \chi r \frac{\partial T}{\partial r} \right) + \frac{1}{\sigma r^2} \frac{\partial}{\partial \varphi} \left( \chi \frac{\partial T}{\partial \varphi} \right) + \nu \left[ r \frac{\partial}{\partial r} \left( \frac{v_\varphi}{r} \right) \right]^2 + \\
+ 2\nu \left[ \left( \frac{\partial v_r}{\partial r} \right)^2 + \left( \frac{v_r}{r} \right)^2 \right] + 2\nu \left[ \left( \frac{1}{r} \frac{\partial v_\varphi}{\partial \varphi} \right)^2 + 2 \frac{v_r}{r^2} \frac{\partial v_\varphi}{\partial \varphi} + \right. \\
+ \frac{1}{2} \left( \frac{\partial v_r}{r \partial \varphi} \right)^2 + \left. \frac{\partial v_r}{\partial \varphi} \left( \frac{\partial}{\partial r} \frac{v_\varphi}{r} \right) \right] - \frac{2}{3} \nu \left[ \left( \frac{1}{r} \frac{\partial}{\partial r} r v_r \right)^2 + \right. \\
+ \left. 2 \frac{1}{r^2} \left( \frac{\partial}{\partial r} r v_r \right) \frac{\partial v_\varphi}{\partial \varphi} + \left( \frac{1}{r} \frac{\partial v_\varphi}{\partial \varphi} \right)^2 \right] - \sigma E^-.
\end{aligned}$$

Здесь  $\nu$  — коэффициент кинематической вязкости;  $\chi$  — коэффициент теплопроводности;  $P = \sigma T$ ;  $T = v^2/3$ ;  $\Psi$  — гравитационный потенциал. Коэффициенты переноса проинтегрированы по  $z$ . Отсюда для возмущений  $A \sim \exp(-i\omega t + ikr + im_\varphi \varphi)$  получим систему алгебраических уравнений ( $kr \gg 1$ ):

$$\begin{aligned}
(\gamma + im_\varphi \Omega) \hat{\sigma} + ik\sigma_0 \hat{v}_r &= 0, \\
(\gamma + im_\varphi \Omega) \hat{v}_r - 2\Omega \hat{v}_\varphi &= i \frac{2\pi G \sigma_0 - kc^2}{\sigma_0} \hat{\sigma} - ik\hat{T} - \frac{4}{3} \nu k^2 \hat{v}_r, \\
(\gamma + im_\varphi \Omega) \hat{v}_\varphi + \frac{\kappa^2}{2\Omega} \hat{v}_r &= -\nu k^2 \hat{v}_\varphi - ik\alpha \hat{T} - ik\beta \hat{\sigma}, \\
\frac{3}{2} (\gamma + im_\varphi \Omega) \hat{T} + ikc^2 \hat{v}_r &= \chi k^2 \hat{T} - \Delta E_\sigma \hat{\sigma} - \Delta E_T \hat{T} - ik\mu \hat{v}_\varphi,
\end{aligned} \quad (6)$$

где

$$\begin{aligned}
\alpha &= \frac{\partial \nu}{\partial T_0} (-r\Omega'); & \beta &= \frac{1}{\sigma_0} \frac{\partial \nu \sigma_0}{\partial \sigma_0} (-r\Omega'); & \mu &= 2\nu (-r\Omega'); \\
\Delta E_\sigma &= \frac{1}{\sigma_0} \left( \frac{\partial E^- \sigma_0}{\partial \sigma_0} - \frac{\partial E^+ \sigma_0}{\partial \sigma_0} \right); & \Delta E_T &= \frac{\partial E^-}{\partial T_0} - \frac{\partial E^+}{\partial T_0}; \\
\gamma &\equiv -i\omega; & c^2 &\equiv T_0.
\end{aligned}$$

В диске выполняется энергетический баланс  $E^+ = E^-$ , где  $E^+ = \nu (r\Omega')^2$  — разогрев дифференциально-вращающегося диска

Таблица 3

## Корни бездиссипативного дисперсионного уравнения

Решение	$m_\varphi$			
	0	1	2	3
$\omega_1$	$-\Omega$	0	$\Omega$	$2\Omega$
$\omega_2$	0	$\Omega$	$2\Omega$	$3\Omega$
$\omega_3$	0	$\Omega$	$2\Omega$	$3\Omega$
$\omega_4$	$\Omega$	$2\Omega$	$3\Omega$	$4\Omega$

вследствие вязкости;  $E^-$  — остывание диска из-за неупругих соударений частиц. Из равенства нулю детерминанта системы (6) получаем дисперсионное уравнение четвертого порядка:  $\gamma^4 + A\gamma^3 + B\gamma^2 + C\gamma + D = 0$ . Если пренебречь всеми столкновительными процессами и эффектами давления и самогравитации, то для сферически-симметричного гравитационного поля дисперсионное уравнение имеет следующие решения  $\omega_i$  (см. табл. 3 для первых значений  $m_\varphi$ ). Естественно ожидать, что низкочастотные неустойчивые решения, связанные с неупругими столкновениями частиц, будут появляться на месте нулевых корней бездиссипативного уравнения. Действительно, при радиальных колебаниях ( $m_\varphi = 0$ ) учет диссипативных процессов превращает одну из нулевых частот в частоту колебаний температуры (что может приводить к тепловой неустойчивости). Другой нулевой корень в разных моделях приводит к диффузионной [36, 37] и аккреционной [17, 33] неустойчивостям.

Рассмотрим не исследованный ранее случай  $m_\varphi = 1$ . Из упрощенного дисперсионного уравнения для  $\omega \ll \Omega$  и  $kr \ll 1$  получим низкочастотное решение:

$$\omega = \frac{1}{2\Omega} (s'k^2c^2 - 2\pi G\sigma_0k - 2\Omega\omega_p) - i \frac{\Delta E_T}{3\Omega^2} (sk^2c^2 - 2\pi G\sigma_0k - 2\Omega\omega_p); \quad (7)$$

где

$$s' = \frac{5}{3} + \frac{14}{9} \frac{\nu \Delta E_T}{c^2} + \frac{2}{3} \frac{\alpha\mu}{c^2} - \frac{4}{3} \frac{\sigma_0}{\Omega c^2} (\Delta E_T \beta - \Delta E_\sigma \alpha);$$

$$s = 1 + [\Omega (2\alpha c^2 + 3\beta \sigma_0) - \sigma_0 \Delta E_\sigma - 5\Omega^2 \nu] / \Delta E_T c^2;$$

а  $\omega_p$  — скорость прецессии из-за несферичности поля Урана (с гармоникой  $J_2$ ):

$$\omega_p = \frac{\Omega^2 - \kappa^2}{2\Omega} \simeq \Omega - \kappa = \frac{3}{2} J_2 \left( \frac{R_p}{r} \right)^2 \Omega.$$

Условие неустойчивости записывается в виде ( $s \approx 1$ )

$$sk^2c^2 - 2\pi G\sigma_0 k - 2\Omega\omega_p < 0. \quad (8)$$

Наиболее неустойчивы длины волн  $\lambda \sim c^2/G\sigma \approx 0,5$  км ( $c \approx 0,2$  см/с;  $\sigma \approx 10$  г/см<sup>2</sup>) с инкрементом  $\gamma \sim vk^2$  (характерное время роста  $t_x \sim \gamma^{-1} \lesssim 0,1$  года). Можно полагать, что именно данная неустойчивость неосесимметричных возмущений с  $m_\varphi = 1$  (эллипс-неустойчивость) ответственна за появление эксцентриситета у колец Урана.

## 5. ПОЧЕМУ КОЛЬЦА АНОМАЛЬНО УСТОЙЧИВЫ?

Почему дифференциальная прецессия не разрушает эллиптические кольца? Голдрайх и Тримейн [15] показали, что устойчивость кольца относительно дифференциальной прецессии обеспечивается самогравитацией. Если узкое кольцо с эксцентриситетом обладает достаточной плотностью (кольцо  $\epsilon$  с  $\sigma \approx 32,9$  г/см<sup>2</sup>; кольца  $\alpha$  и  $\beta$  с  $\sigma \approx 1,5 \div 2$  г/см<sup>2</sup> [4]), то оно будет прецессировать как единое целое. Максимальная ширина эллиптических колец Сатурна и Урана не превышает 100 км (для стабилизации более широких колец требуется очень большая плотность).

Почему кольца не расплываются из-за диффузии? Почему резки края колец? Эта проблема стала острее обсуждаться после пролета «Вояджера-2». Возник вопрос: почему частицы колец просто не упали на планету? По оценкам аэродинамического трения даже метровые частицы должны упасть на Уран за время  $10^{6-8}$  лет [30]. Устойчивость колец можно понять, если учесть, что процессы расплывания кругового и эллиптического колец качественно различаются. Действительно, расплывание кругового кольца — это переход сталкивающихся (или дрейфующих) частиц на соседние квазикруговые орбиты, не пересекающиеся с основным кольцом. При расплывании эллиптического кольца такие соседние орбиты тоже будут эллиптическими с близким эксцентриситетом. В несферическом поле кольцо прецессирует со скоростью, отличной от скорости прецессии орбиты оторвавшейся частицы (из-за отличия в больших полуосях). Вследствие этого орбита частицы неизбежно пересечет кольцо, и частица, израсходовав энергию относительного движения на неупругие соударения, застревает в основной массе частиц.

Оценим, как далеко «отпускает» кольцо частицы, прежде чем захватить. Это расстояние должно соответствовать резкости края кольца. Рассмотрим два вложенных непересекающихся эллипса (с совпадающим фокусом и почти совпадающими направлениями на апоцентры) с большими полуосями  $a$  и  $a + \delta a$  и эксцентриситетами  $e$  и  $e + \delta e$  ( $\delta a \ll a$ ,  $\delta e \ll e$ ). Найдем угол  $\Delta\varphi_{\min}$  — минимальный угол между линиями аспид, при котором эллипсы соприкасаются:

$$\Delta\varphi_{\min} \approx \arccos \left[ - \left( \frac{\delta a}{2ae} + \frac{a\delta e}{\delta a} \right) \right] - \arccos \left( \frac{\delta a}{2ae} + \frac{a\delta e}{\delta a} \right). \quad (9)$$

Учитывая, что  $\delta e$  связано с соударениями и сравнимо с тепловыми эксцентриситетами ( $\delta e \approx h/a$ ), получим условие «догона»  $h \lesssim \lesssim \delta a \approx 2ae$ . Для сильного неравенства  $\Delta\varphi_{\min} \approx \frac{\delta a}{2ae} + \frac{a\delta e}{\delta a}$ . Край кольца догоняет частицу из-за дифференциальной прецессии  $\Delta\omega_p$  за характерное время

$$t_p \approx \frac{\Delta\varphi_{\min}}{\Delta\omega_p} \approx \frac{\delta a/2ae + a\delta e/\delta a}{\frac{21}{4} J_2 \Omega \left(\frac{R_p}{a}\right)^2 \frac{\delta a}{a}}. \quad (10)$$

Сравнивая (10) и время диффузионного расплывания края  $t_d \approx (\delta a)^2/\nu$ , найдем равновесное  $\delta a$ . Например, для кольца  $\epsilon$  резкость края  $\delta a \approx 400 \div 750$  м для времени свободного пробега частиц  $t_c \approx (1 \div 10) \Omega^{-1}$ , что хорошо соответствует наблюдаемым данным [22]. Точно так же кольцо будет улавливать и мелкие частицы, падающие на планету вследствие аэродинамического торможения. Отметим, что даже очень малый эксцентриситет ( $e \approx 10^{-(5 \div 6)}$ ) должен резко повышать диффузионную стабильность кольца.

### ЗАКЛЮЧЕНИЕ

Узкие кольца Урана образовались вблизи резонансов от внешних спутников. Первоначально круговые колечки, взаимодействуя с окружающим диском, превратились в эллиптические. Узкие, эллиптические кольца оказались устойчивыми к диффузионному расплыванию и дифференциальной прецессии и сохранились до наших дней.

Общий сценарий, изложенный в данной работе, требует дальнейшей детализации и количественных моделей отдельных этапов эволюции. Например, неизвестен ответ на вопрос: можно ли из уравнений переноса получить нелинейное локализованное устойчивое решение, соответствующее узкому эллиптическому кольцу? Но уже сейчас ясно, что кольца Урана имеют резонансное происхождение и являются ярким примером самоорганизации.

### ЛИТЕРАТУРА

1. Elliot J. L., Dunham E., Mink D. The rings of Uranus // Nature. 1977. Vol. 267, N 5609. P. 328—330.
2. Millis R. L., Wasserman L. H., Birch P. V. Detection of rings around Uranus // Ibid. P. 330—331.
3. Elliot J. L., Nicholson P. D. The rings of Uranus // Planetary rings / Ed. R. Greenberg, A. Brahic. Tucson(Ariz.): Univ. Arizona Press, 1984. P. 25—72.
4. French R. G., Elliot J. L., Levine S. E. Structure of the Uranian rings. II. Ring orbits and widths // Icarus. 1986. Vol. 67, N 1. P. 134—163.
5. Goldreich P., Tremaine S. Towards a theory for Uranian rings // Nature. 1979. Vol. 277, N 5692. P. 97—99.
6. Dermott S. F., Gold T. The rings of Uranus: Theory // Ibid. 1977. Vol. 267, N 5612, P. 590—593.
7. Aksnes K. Quantitative analysis of the Dermott—Gold theory for Uranus's rings // Ibid. Vol. 269, N 5631. P. 783.

8. Goldreich P., Nicholson P. Revenge of tiny Miranda // *Ibid.* P. 783—785.
9. Goldreich P., Tremaine S. The formation of the Cassini division in Saturn's rings // *Icarus*, 1978. Vol. 34, N 2. P. 240—253.
10. Steigmann G. A. Resonance orbits and the rings of Uranus // *Nature*. 1978. Vol. 274, N5670. P. 454—455.
11. Steigmann G. A. Concerning the mass of the suggested sixth satellite of Uranus // *Mon. Notes. Roy. Astron. Soc.* 1979. Vol. 186, N. 2. P. 19p—22p.
12. Van Flandern T. C. The rings of Uranus: Invisible and impossible // *Science*. 1979. Vol. 204. N 4398. P. 1076—1077.
13. Dermott S. F., Gold T., Sinclair A. T. The rings of Uranus: Nature and origin // *Astron. J.* 1979. Vol. 84, N 8. P. 1225—1234.
14. Goldreich P., Tremaine S. The origin of the eccentricities of the rings of Uranus // *Astrophys. J.* 1981. Vol. 243, N 3(1). P. 1062—1075.
15. Goldreich P., Tremaine S. Precession of the  $\epsilon$ -ring of Uranus // *Astron. J.* 1979. Vol. 84, N 10. P. 1639—1644.
16. Горькавый Н. Н., Фридман А. М. О причине существования границы между планетными кольцами и спутниками планет и о размерах и структуре частиц колец // *Письма в «Астрон. журн.»* 1985. Т. 11, № 8. С. 628—633.
17. Горькавый Н. Н., Фридман А. М. Структура планетных колец // *Нелинейные волны. Структуры и бифуркации* / Под ред. А. В. Гапонова-Грехова, М. И. Рабиновича. М.: Наука, 1987. С. 86—104.
18. Горькавый Н. Н. К вопросу об устойчивости колец Сатурна // *Письма в «Астрон. журн.»* 1985. Т. 11, № 6. С. 469—474.
19. Holberg J. B., Forrester W. T., Lissauer J. J. Identification of resonance feature within the rings of Saturn // *Nature*. 1982. Vol. 297, N 5861. P. 115—120.
20. Porco C., Nicholson P. D., Borderies N. et al. The eccentric Saturnian ringlets at  $1.29R_s$  and  $1.45R_s$  // *Icarus*. 1984. Vol. 60, N 1. P. 1—16.
21. Горькавый Н. Н., Фридман А. М. О резонансной природе колец Урана, определяемой его неоткрытыми спутниками // *Астрон. циркуляр.* 1985, № 1391. С. 1—2. — То же // *Письма в «Астрон. журн.»* 1985. Т. 11, № 9. С. 717—720.
22. Smith B. A., Soderblom L. A., Beebe R. et al. Voyager-2 in the Uranian system: Imaging science results // *Science*. 1986. Vol. 233, N 4759. P. 43—64.
23. Lane A. L., Hord C. W., West R. A. et al. Photometry from Voyager-2: Initial results from the Uranian atmosphere, satellites and rings // *Ibid.* P. 65—70.
24. Owen W. M., Jr., Synnott S. P. Orbits of the ten small satellites of Uranus // *Astron. J.* 1987. Vol. 93, N 5. P. 1268—1271.
25. Горькавый Н. Н., Фридман А. М. Резонансная природа колец Урана и предсказание его новых спутников // *УФН.* 1986. Т. 150, № 4. С. 628—632.
26. Горькавый Н. Н., Фридман А. М. Об открытии «Вояджером-2» предсказанных спутников, определяющих резонансную природу колец Урана // *Письма в «Астрон. журн.»* 1987. Т. 13. № 3. С. 237—244.
27. Шиголов Б. М. Математическая обработка наблюдений. М.: Наука, 1969. 344 с.
28. Cuzzi J. N., Lissauer J. J., Esposito L. W. et al. Saturn's rings: Properties and processes // *Planetary rings* / Ed. R. Greenberg, A. Brahic. Tucson (Ariz.): Univ. of Arizona Press, 1984. P. 73—199.
29. Goldreich P., Tremaine S. Disk-satellite interactions // *Astrophys. J.* 1980. Vol. 241, N 1. P. 425—441.
30. Broadfoot A. L., Herbert F., Holberg J. B. et al. Ultraviolet spectrometer observations of Uranus // *Science*. 1986. Vol. 233, N 4759. P. 74—79.
31. Meyer—Vernet N., Sicardy B. On the physics of resonant disk-satellite interaction // *Icarus*. 1987. Vol. 69, N 1. P. 157—175.
32. Esposito L. W. Structure and evolution of Saturn's rings // *Ibid.* 1986. Vol. 67, N 3. P. 345—357.



33. Горькавый *Н. Н.* К динамике планетных колец: Дис. ... канд. физ.-мат. наук. М., 1986. 136 с.
34. Goldreich *P.*, Tremaine *S.* The dynamics of planetary rings // *Annu. Rev. Astron. and Astrophys.* 1982. Vol. 20. P. 249—283.
35. Горькавый *Н. Н.* К теории линейных колебаний вращающихся столкновительных сред // *Науч. информ. Астрон. совета АН СССР.* 1986. Вып. 61. С. 132—147.
36. Ward *W. R.* On the radial structure of Saturn's rings // *Geophys. Res. Lett.* 1981. Vol. 8, N 6. P. 641—643.
37. Lin *D. N. C.*, Bodenheimer *P.* On the stability of Saturn's rings // *Astro-phys. J. Lett.* 1981. Vol. 248, N 2. P. L83—L86.

## СОДЕРЖАНИЕ

Предисловие . . . . .	3
НЕЛИНЕЙНАЯ ДИНАМИКА — ОТ ПУАНКАРЕ ДО НАШИХ ДНЕЙ	
<i>Данилов Ю. А.</i> Нелинейная динамика: Пуанкаре и Мандельштам . . . . .	5
<i>Афраймович В. С.</i> Об аттракторах . . . . .	16
<i>Островский Л. А.</i> От нелинейных колебаний — к нелинейным волнам . . . . .	29
<i>Рабинович М. И.</i> Нелинейная динамика и турбулентность . . . . .	50
<i>Гапонов-Грехов А. В., Ломов А. С., Осипов Г. В., Рабинович М. И.</i> Рождение и динамика двумерных структур в неравновесных диссипативных средах . . . . .	61
СТРУКТУРЫ	
<i>Заславский Г. М., Сагдеев Р. З., Усиков Д. А., Черников А. А.</i> Стохастическая паутина и симметрия структур . . . . .	84
<i>Непомнящий А. А.</i> Динамика дефектов и возникновение пространственного хаоса в одномерных системах . . . . .	106
<i>Бункин Ф. В., Кириченко Н. А., Лукьянчук Б. С.</i> Явления самоорганизации в лазерной термохимии . . . . .	113
<i>Кернер Б. С., Осипов В. В.</i> Термодиффузионные автосолитоны и страты в полупроводни- ковой и газовой плазме . . . . .	127
<i>Барц Б. И., Моисеев С. С.</i> Турбулентно-волновое динамо . . . . .	153
<i>Калафати Ю. Д., Ржанов Ю. А.</i> Критерий существования бегущих структур в двухкомпонент- ных реакционно-диффузионных системах . . . . .	159
<i>Мастеров А. В., Толков В. Н., Яхно В. Г.</i> Пространственно-временные структуры в оптоэлектронных устройствах . . . . .	166

<i>Рогальский А. В.</i> Численное моделирование двумерных неравновесных сред на многопроцессорном комплексе . . . . .	183
<i>Емельянов В. И.</i> Лазерно-индуцированные неустойчивости и фазовые переходы на поверхности твердых тел и образование упорядоченных структур . . . . .	198
<i>Гавриков В. К., Кац А. В., Конторович В. М., Сневак И. С.</i> Вынужденное рассеяние и поверхностные структуры . . . . .	208
<i>Власов С. Н., Таланов В. И.</i> Распределенный волновой коллапс в модели нелинейного уравнения Шредингера . . . . .	218
<i>Ахманов С. А., Воронцов М. А.</i> Неустойчивости и структуры в когерентных нелинейно-оптических системах, охваченных двумерной обратной связью . . . . .	223

### ДИНАМИЧЕСКИЙ ХАОС

<i>Афраймович В. С., Рейман А. М.</i> Размерность и энтропия в многомерных системах . . . . .	238
<i>Арансон И. С., Рейман А. М., Шехов В. Г.</i> Методы измерения корреляционной размерности в эксперименте . . . . .	262
<i>Шепелянский Д. Л.</i> Квантовая локализация динамического хаоса . . . . .	267
<i>Кравцов Ю. А.</i> Случайность и предсказуемость динамического хаоса . . . . .	276
<i>Летохов В. С., Макаров А. А., Рябов Е. А.</i> Стохастическая колебательная динамика многоатомных молекул . . . . .	288

### ЭВОЛЮЦИЯ

<i>Бабский В. Г.</i> Явления самоорганизации у бактерий на клеточном и популяционном уровнях . . . . .	299
<u>Белинцев Б. Н.</u> «Полевые» черты в морфогенезе и механохимическая активность клеток . . . . .	303
<i>Бикташев В. Н.</i> Дрейф ревербератора в активной среде при взаимодействии с границами . . . . .	316
<i>Всеволодов Н. Н.</i> Фотохромные и нелинейные процессы в пленке «Биохром-БР» . . . . .	324
<i>Михайлов А. С.</i> Инженерия динамических систем для распознавания образов и обработки информации . . . . .	331

## ПРИЛОЖЕНИЯ НЕЛИНЕЙНОЙ ДИНАМИКИ

*Воловик Г. Е.*

Нелинейные явления в физике конденсированных сред: Вселенная в капле гелия . . . . . 343

*Пелиновский Е. Н.*

Нелинейная теория набегания морских волн на берег . . . . . 351

*Железняков В. В., Кочаровский В. В., Кочаровский Вл. В.*

Сверхизлучение. Метод электродинамики сплошных активных сред . . . . . 358

*Горькавый Н. Н., Фридман А. М.*

Происхождение и эволюционная динамика колец Урана . . . . . 376

---

---

## CONTENTS

Introduction . . . . .	3
------------------------	---

### NONLINEAR DYNAMICS: FROM POINCARÉ TO THE PRESENT DAY

<i>Danilov Yu. A.</i> Nonlinear dynamics: Poincaré and Mandelstam . . . . .	5
<i>Afraimovich V. S.</i> Attractors . . . . .	16
<i>Ostrovsky L. A.</i> From nonlinear oscillations to nonlinear waves . . . . .	29
<i>Rabinovich M. I.</i> Nonlinear dynamics and turbulence . . . . .	50
<i>Gaponov-Grekhov A. V., Lomov A. S., Osipov G. V., Radinovich M. I., vich M. I.</i> Pattern formation and dynamics of two-dimensional structures in nonequilibrium' dissipative media . . . . .	61

### STRUCTURES

<i>Zaslavsky G. M., Sagdeev R. Z., Usikow D. A., Chernikov A. A.</i> Stochastic spider web and structural symmetry . . . . .	84
<i>Nepomnyashchy A. A.</i> Dynamics of defects and onset of spatial chaos in one-dimensional systems . . . . .	106
<i>Bunkin F. V., Kirichenko N. A., Luk'yanchuk B. S.</i> Self-organization phenomena in laser thermochemistry . . . . .	113
<i>Kerner B. S., Osipov V. V.</i> Thermal-diffusion autosolitons and strats in semiconductor an gas plasmas . . . . .	127
<i>Barts B. I., Moiseev S. S.</i> The wave turbulent dynamo . . . . .	153
<i>Kalafati Y. D., Rzhanov Y. A.</i> Criteria of the existence of moving structures in two-component reaction-diffusion systems . . . . .	159

<i>Masterov A. V., Tolkov V. N., Yakhno V. G.</i>	
Spatio-temporal structures in opto-electronic devices . . . . .	166
<i>Rogal'skii A. V.</i>	
Multiple array processors for two-dimensional nonequilibrium media computer simulation . . . . .	183
<i>Emel'yanov V. I.</i>	
Laser-induced instabilities and phase transitions on the surfaces of solids with the formation of the ordered structures . . . . .	198
<i>Gavrikov V. K., Kats A. V., Kontorovich V. M., Spevak I. S.</i>	
Stimulated scattering and surface structures . . . . .	208
<i>Talanov V. I., Vlasov S. N.</i>	
Distributed wave collapse in the nonlinear Schrodinger equation	218
<i>Akhmanov S. A., Vorontsov M. A.</i>	
Instabilities and structures in coherent nonlinear optical systems	228

## DYNAMIC CHAOS

<i>Afraimovich V. S., Reiman A. M.</i>	
Dimensions and entropies in multidimensional systems . . . . .	238
<i>Aranson I. S., Reiman A. M., Shekhov V. G.</i>	
Measurement methods for correlation dimensions . . . . .	262
<i>Shepelyansky D. L.</i>	
Quantum localization of dynamic chaos . . . . .	267
<i>Kravtsov Yu. A.</i>	
Randomness and predictability in dynamic chaos . . . . .	276
<i>Letokhov V. S., Makarov A. A., Ryabov E. A.</i>	
Stochastic vibrational dynamics of polyatomic molecules . . . . .	288

## EVOLUTION

<i>Babski V. G.</i>	
The phenomena of self-organization in bacterial cells and populations . . . . .	299
<i>Belintsev B. N.</i>	
Mechano-chemical activity of cells and self-organization of embryonic patterns . . . . .	303
<i>Biktashev V. N.</i>	
Drift of a reverberator in an active medium due to the interaction with boundaries . . . . .	316
<i>Vsevolodov, N. N.</i>	
Photochromic and nonlinear processes in biochrom — BR films	324
<i>Mikhailov A. S.</i>	
Engineering of dynamic systems for pattern recognition and information processing . . . . .	331

APPLICATIONS OF NONLINEAR DYNAMICS

*Volovik G. E.*

Nonlinear phenomena in condensed matter: Universe in a helium droplet . . . . . 343

*Pelinovsky E. N.*

Nonlinear theory of sea wave run-up . . . . . 351

*Zheleznyakov V. V., Kocharovskiy V. V., Kocharovskiy Vl. V.*

Superradiance: Approach to electrodynamics of continuous active media . . . . . 358

*Gor'kavyj N. N., Fridman A. M.*

The origin and evolutionary dynamics of uranian rings . . . . 376

Научное издание

**НЕЛИНЕЙНЫЕ ВОЛНЫ**  
**Динамика и эволюция**

Утверждено к печати  
Институтом прикладной физики АН СССР

Редактор издательства *Л. Е. Кононенко*  
Художественный редактор *М. Л. Храмцов*  
Технический редактор *Н. Н. Плохова*  
Корректор *Н. Б. Габасова*

ИБ № 39745

Сдано в набор 27.01.89

Подписано к печати 4.05.89

Т-00160 Формат 60×90<sup>1/16</sup>

Бумага типографская № 1

Гарнитура обыкновенная

Печать высокая

Усл. печ. л. 25,0. Усл. кр. отт. 25,0. Уч.-изд. л. 27,3

Тираж 1950 экз. Тип. зак. 2574

Цена 4 р 70 к.

Ордена Трудового Красного Знамени  
издательство «Наука»

117864, ГСП-7, Москва, В-485, Профсоюзная ул., 90

2-я типография издательства «Наука»  
121099, Москва, Г-99, Шубинский пер., 6



ИЗДАТЕЛЬСТВО «НАУКА»  
ГОТОВИТСЯ К ПЕЧАТИ КНИГА:

**ВОЗДЕЙСТВИЕ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ  
НА МАТЕРИАЛЫ**

*Р. В. Арутюнян, В. Ю. Баранов, Л. А. Бобышов и др.*

—25 л. 3 р. 40 к.

Книга посвящена одной из наиболее актуальных проблем современной физики, связанной с ускорением внедрения в промышленность лазерной технологии. Авторы — ведущие советские специалисты в этой области науки. В книге изложены результаты многочисленных экспериментов по обработке материалов излучением импульсно-периодических лазеров, рассмотрены теоретические модели протекающих при этом физических процессов. Представлены результаты численного моделирования различных режимов лазерного воздействия на материалы.

Для научных работников и инженеров, специализирующихся в области лазерной техники и технологии.

Для получения книг почтой  
заказы просим направлять по адресу:

- 117192 Москва, Мичуринский проспект, 12, магазин «Книга — почтой»  
Центральной конторы «Академкнига»;  
197345 Ленинград, Петрозаводская ул., 7, магазин «Книга — почтой»  
Северо-Западной конторы «Академкнига»  
или в ближайший магазин «Академкнига», имеющий отдел  
«Книга — почтой».