

НЕЛИНЕЙНЫЕ
ВОЛНЫ' 2004



Российская академия наук
Институт прикладной физики

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2004

Нижний Новгород
2005

Издано по решению редакционно-издательского совета
Института прикладной физики РАН

УДК 534.222.2
ББК 22.312я4

Н49 **Нелинейные волны' 2004** / Отв. ред. А. В. Гапонов-Грехов,
В. И. Некоркин. — Нижний Новгород: ИПФ РАН, 2005. — 544 с.

ISBN 5-8048-0048-5

В сборник включены обзорные лекции и оригинальные статьи, написанные по материалам лекций, прочитанных на XII школе по нелинейным волнам (Нижний Новгород, 29 февраля — 7 марта 2004 г.). Рассматриваются проблемы пространственно-временного хаоса, структурообразования, синхронизации, волновой динамики, нелинейные явления в плазме, оптике и при взаимодействии сверхсильных полей с веществом, математические проблемы нелинейной динамики и другие аспекты нелинейно-волновой тематики.

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

Ответственные редакторы

академик А. В. ГАПОНОВ-ГРЕХОВ

доктор физико-математических наук В. И. НЕКОРКИН

Рецензенты

доктор физико-математических наук В. Д. Шалфеев

доктор физико-математических наук В. П. Пономаренко

*Проведение школы и издание сборника осуществлены при поддержке
Министерства промышленности, науки и технологий РФ,
Российской академии наук, Российского фонда
фундаментальных исследований, РФЯЦ—ВНИИЭФ, ООО НТП «Поларис»,
Компания «Intel», Spectra-Physics*

ISBN 5-8048-0048-5

© Институт прикладной физики РАН, 2005 г.

СОДЕРЖАНИЕ

<i>Предисловие</i>	5
Нелинейные процессы в физических системах	
<i>А. А. Абрашкин, Е. И. Якубович. Точные решения в гидродинамике</i>	9
<i>А. Б. Езерский. Пространственно-временная динамика солитонов в резонаторах</i>	23
<i>А. А. Куркин, Е. Н. Пелиновский, А. В. Слюняев. Физика волн-убийц в океане</i>	37
<i>Ю. И. Троицкая. Нелинейное резонансное взаимодействие волн с потоками в океане</i>	52
<i>О. В. Руденко. Основы физики интенсивных акустических волн</i>	70
<i>С. Н. Гурбатов, А. И. Саичев. Статистическая нелинейная акустика. Волны и структуры в средах без дисперсии</i>	89
<i>Л. А. Островский. Неклассическая нелинейная акустика</i>	109
<i>А. И. Потапов. Волны деформации в среде с внутренней структурой</i>	125
<i>Ю. Н. Кульчин, Р. В. Ромашко, С. М. Шандаров, Н. И. Буримов, Ю. Ф. Каргин, В. В. Волков, О. Т. Каменев. Применение явления самодифракции световых волн в фоторефрактивных кристаллах для обработки нестационарных спекловых полей</i>	141
<i>М. Ю. Рябикин, А. М. Сергеев. Инжиниринг электронных волновых пакетов в атомах и молекулах для управления генерацией аттосекундных импульсов</i>	149
<i>А. Н. Степанов, Г. М. Фрайман. Рамановское усиление коротких лазерных импульсов в плазме</i>	165
<i>Л. С. Мхитарьян. Z-пинч и термоядерный синтез</i>	179
<i>В. Я. Демиховский. Квантовый хаос в приложениях</i>	188
<i>Вл. В. Кочаровский, В. Ю. Мартьянов. Нелинейность вакуума и перспективы ее наблюдения</i>	202

<i>Д. И. Иудин, В. Ю. Трахтенгерц.</i> Динамическая перколяция	217
<i>Е. Ю. Банникова, К. Ю. Блюх, В. М. Конторович.</i> Эволюция и коллапс самогравитирующего тороидального вихря	243
<i>В. И. Кляцкин.</i> Статистика и реальность в стохастических физических системах	256
<i>Д. И. Трубецков, С. П. Кузнецов, Н. М. Рыскин, А. Е. Храмов.</i> Сложная динамика распределенных нелинейных систем сверхвысокочастотной электроники	287
<i>А. М. Фридман, Р. З. Сагдеев, Е. В. Поляченко.</i> Наблюдаемые проявления хаоса в спиральных галактиках	327

Общие вопросы нелинейной динамики

<i>В. Б. Казанцев, В. И. Некоркин.</i> Фазово-управляемые колебания в нейродинамике	345
<i>В. Г. Яхно, И. В. Нуйдель, А. Е. Иванов.</i> Модельные нейроноподобные системы. Примеры динамических процессов	362
<i>Л. В. Якушевич.</i> Введение в нелинейную физику ДНК	376
<i>Б. П. Безручко, М. Б. Бодров, Т. В. Диканев, А. С. Караваев, В. И. Пономаренко, М. Д. Прохоров, Е. П. Селезнев, И. В. Сысов, Д. А. Смирнов.</i> Некоторые проблемы реконструкции модельных уравнений по временным рядам и пути их решения	381
<i>Н. Г. Макаренко.</i> Реконструкция динамических систем по хаотическим временным рядам	398
<i>Е. М. Лоскутов, Я. И. Мольков, Д. Н. Мухин, А. М. Фейгин.</i> Статистический подход к реконструкции динамических систем	411
<i>Д. В. Тураев, А. Л. Шильников, Л. П. Шильников.</i> Некоторые математические проблемы классической синхронизации	426
<i>С. А. Каценко.</i> Нормальные формы для сингулярно возмущенных систем параболического типа и систем с запаздыванием	451
<i>С. П. Кузнецов, А. С. Пиковский, У. Фойдель.</i> Странный нехаотический аттрактор	484
<i>А. Ю. Лоскутов, А. Б. Рябов.</i> Бильярды	510

ПРЕДИСЛОВИЕ

Два года, прошедшие после предыдущей, одиннадцатой Школы (Нижний Новгород, март 2002), на наш взгляд, показали, что основные направления и тенденции развития нелинейной динамики в целом сохранились. По-прежнему большой интерес исследователей вызывают проблемы пространственно-временного хаоса, структурообразования, синхронизации, волновой динамики, нелинейные явления в плазме, оптике и при взаимодействии сверхсильных полей с веществом, математические проблемы нелинейной динамики и т. д. Получила дальнейшее развитие тематика, которая на прошлой Школе была условно обозначена как «информационные аспекты нелинейной динамики». К этой проблематике тесно примыкают интенсивно развивающиеся в последнее время нейродинамика и динамика фрактальных структур.

В той или иной степени весь этот достаточно широкий класс задач нелинейной динамики нашел отражение в лекциях, прочитанных на двенадцатой Школе «Нелинейные волны — 2004» (Нижний Новгород, март 2004 г.). Настоящий сборник составлен из обзорных лекций и оригинальных статей, написанных по материалам лекций этой Школы. Содержание сборника условно разбито на два раздела. Один из них составляют статьи, посвященные исследованию конкретных физических задач, статьи второго раздела скорее можно отнести к общим вопросам нелинейной динамики. На Школе были прочитаны циклы лекций по теории синхронизации, акустике и гидродинамике, нейродинамике и биофизике, математическим проблемам нелинейной динамики, по нелинейным процессам в плазме и оптике и др. Сложившаяся на предыдущих школах традиция представлять сравнительно новые для нелинейной динамики проблемы была продолжена на Школе 2004 года. В связи с этим можно отметить лекции Л. Ш. Цимринга «Нелинейная динамика сыпучих сред» и Л. В. Якушевича «Нелинейная физика ДНК».

Мы сочли нецелесообразным включать в сборник материалы тех лекций, основное содержание которых ранее было опубликовано в дру-

гих изданиях, хотя они также вызвали большой интерес слушателей. Так, в сборник не вошли материалы лекций А. С. Пиковского (по теории синхронизации) и В. С. Анищенко (по теории статистического резонанса), а также лекций Л. Ш. Цимринга, В. Н. Белых «Кластерная синхронизация связанных хаотических систем», В. Д. Шалфеева «Коммуникации на основе динамического хаоса», В. Л. Воейкова «Роль воды и кислорода в колебательных процессах в биологии», И. И. Мохова «Климатические изменения и нелинейные эффекты», И. С. Аронсона «Самоорганизация и структурообразование в ансамблях проводящих микрочастиц под воздействием сильного электрического поля», Г. Т. Гурии «Структурообразование в крови», А. И. Нейштадта «О разрушении адиабатической инвариантности в системах с быстрыми и медленными движениями», Л. М. Лермана «Динамические системы и структуры в эллиптических уравнениях».

Мы благодарим всех авторов, приславших материалы в данный сборник.

А. В. Гапонов-Грехов

В. И. Некоркин

Нелинейные процессы в физических системах

ТОЧНЫЕ РЕШЕНИЯ В ГИДРОДИНАМИКЕ

А. А. Абрашкин, Е. И. Якубович

Под точными решениями традиционно понимают частные решения нелинейных уравнений, которые в общем случае не интегрируются. Они составляют «золотой фонд» нелинейной науки. Отсюда особый интерес и постоянное внимание к ним.

Своеобразие гидродинамики и, по-видимому, главное ее отличие от остальных областей механики и физики состоит в том, что основное уравнение гидродинамики, описывающее движение жидкости, уравнение Навье — Стокса является существенно нелинейным. Оно было написано французом Навье в 1822 г. и независимо англичанином Стоксом в 1845 г. в следующем виде:

$$\frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V} \nabla) \mathbf{V} = -\frac{\nabla p}{\rho} + \nu \Delta \mathbf{V}. \quad (1)$$

Неизвестной функцией в этом уравнении является векторное поле скорости $\mathbf{V}(x, y, z, t)$, зависящее от трех пространственных координат — x, y, z и времени t . В его правой части стоят сила градиента давления p и вязкая сила (ν — кинематическая вязкость). В невязком пределе, когда $\nu = 0$ (идеальная жидкость), уравнение (1) называют уравнением Эйлера. Нелинейный член в левой части отражает факт движения жидкого объема, он-то и вносит главные трудности в решение уравнений гидродинамики. Уравнение движения жидкости (1) следует дополнить еще условием ее несжимаемости

$$\operatorname{div} \mathbf{V} = 0 \quad (2)$$

и соответствующими граничными условиями. Они имеют простой вид на границе с твердым телом — условия прилипания (вязкая жидкость) или непротекания (идеальная). Для движений жидкости со свободной поверхностью или при склейке разных типов течений они имеют нелинейный характер. Кроме того, в этих случаях подлежит определению и форма границы. Это еще более усложняет анализ течений.

Как известно, достаточно общих методов решения нелинейных уравнений не существует. В гидродинамике ситуация усложняется еще тем, что самые простые течения несжимаемой жидкости — это двумерные движения, а потому даже стационарные потоки жидкости описываются уравнениями в частных производных. Тем не менее, в гидродинамике накоплен значительный опыт получения точных результатов. Ниже излагаются три различных подхода к нахождению точных решений системы (1), (2). В целом статья носит обзорно-ознакомительный харак-

тер, но в каждом конкретном случае будут указаны наиболее важные результаты, полученные в последнее время.

1. Плоские стационарные вихревые течения идеальной жидкости

Будем рассматривать плоские ($\frac{\partial}{\partial z} = 0$), стационарные ($\frac{\partial}{\partial t} = 0$) течения в приближении невязкой жидкости. Кроме того, для исключения давления применим к уравнению (1) операцию взятия ротора. Тогда получим следующее уравнение:

$$\text{rot}(\nabla\nabla)\mathbf{V} = 0.$$

Вводя функцию тока ψ по формулам

$$V_x = \frac{\partial\psi}{\partial y}; \quad V_y = -\frac{\partial\psi}{\partial x}; \quad \mathbf{V} = \mathbf{V}\{V_x; V_y\},$$

его можно преобразовать к нелинейному соотношению

$$\frac{D(\Delta\psi, \psi)}{D(x, y)} = 0,$$

эквивалентному уравнению Пуассона

$$\Delta\psi = F(\psi). \tag{3}$$

Здесь F — некоторая, вообще говоря, произвольная функция. Поскольку завихренность плоского течения равна

$$\Omega = \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} = -\Delta\psi,$$

то функция f определяет распределение завихренности по линиям тока. В зависимости от вида этой функции уравнение (3) описывает течения с различным законом изменения завихренности.

Наиболее простой способ задания правой части (3) соответствует однородному распределению завихренности: $f(\psi) = -\Omega = -\Omega_0$. В этом случае общее решение уравнение (3) можно представить с использованием комплексных функций $W = X + iY$, $\bar{W} = X - iY$, $V = V_x + iV_y$ так:

$$V = i\frac{\Omega_0}{2}W + \frac{d\overline{\Phi(W)}}{dW}. \tag{4}$$

Здесь Φ — произвольная аналитическая функция. В случае $\Omega_0 = 0$ функция Φ играет роль потенциала течения. Если, наоборот, $\Phi = 0$,

то решение (4) описывает однородное вращение жидкости. Ограничив область течения круговой областью с некоторым радиусом $|W| = r_0$, получим аналитическое представление для вихревого движения жидкости внутри кругового вихря. Легко проверить, что функция $\Phi(\overline{W}) = i\alpha\overline{W}^2$ определяет течение с софокусными эллиптическими линиями тока и т. д.

А что будет, если функция f не постоянна в области вихревого течения? В этом случае известно лишь ограниченное число точных решений, отвечающих вполне определенному ее виду.

1.1. Вихрь Чаплыгина — Ламба

Зададим функцию f внутри круга $r \leq r_0$ в следующем виде:

$$f(\psi) = -n^2(\psi - \lambda), \quad (5)$$

а вне круга будем считать ее равной нулю (течение потенциально). Тогда решение (3), (5) находится с помощью разделения переменных и в полярных координатах r, θ записывается так [1]:

$$\begin{aligned} \psi &= U\left(r - \frac{r_0^2}{r}\right) \sin \theta; \quad r \geq r_0; \\ \psi &= \frac{2Ur_0}{\beta} J_0\left(\frac{\beta r}{r_0}\right) J_1\left(\frac{\beta r}{r_0}\right) \sin \theta + \lambda \left[1 - \frac{J_0(\beta r/r_0)}{J_0(\beta)}\right]; \quad r \leq r_0, \end{aligned} \quad (6)$$

где U — скорость потока, набегающего на круговую вихревую область с радиусом r_0 ; J_0 — функция Бесселя; λ — произвольная постоянная; $\beta = 3,8317$ — наименьший положительный корень уравнения $J_1(\beta) = 0$; $n = \beta/r_0$. Это решение было получено С. А. Чаплыгиным в 1903 году. Оно описывает стационарный круговой вихрь, обтекаемый однородным на бесконечности потенциальным потоком. Внутренняя структура вихря определяется величиной λ . Если $\lambda = 0$, то вихрь имеет симметричную форму (рис. 1, а) и состоит из двух одинаковых половинок, внутри которых жидкость вращается в разных направлениях. Такое симметричное решение ранее открыл и исследовал Ламб в 1895 году, но он этим случаем и ограничился [2]. Чаплыгин же подробно изучил самую общую ситуацию, когда $\lambda \neq 0$ и вихрь становится несимметричным. Граница, разделяющая его половинки, уже не совпадает с диаметром, а является опирающейся на него дугой (рис. 1, б).

Решение (6) можно также интерпретировать как вихрь, движущийся со скоростью U относительно покоящейся на бесконечности жидкости. Поэтому это вихревое образование называют еще диполем Ламба.

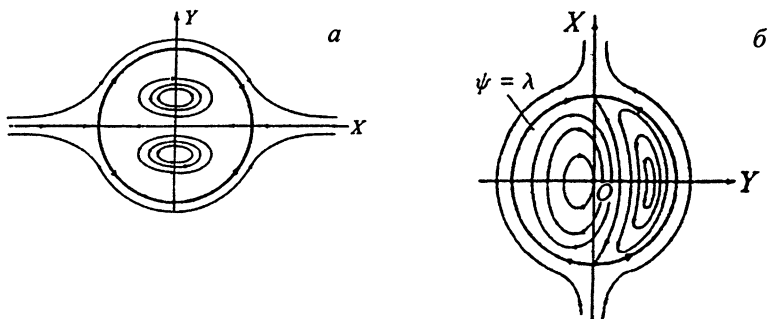


Рис. 1

1.2. Дорожка распределенных вихрей Стюарта

В том случае, когда правая часть уравнения (3) является экспонентой, т. е.

$$\Delta\psi = \alpha \exp(-\lambda\psi); \quad \alpha\lambda > 0, \quad (7)$$

оно называется уравнением Лиувилля, который указал его общее решение, зависящее от произвольной аналитической функции $u + iw$ следующим образом:

$$\exp(-\alpha\psi) = \frac{8}{\alpha\lambda} \frac{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}{(u^2 + w^2 + 1)^2}.$$

Использовать его для описания конкретных течений, однако, оказывается не так просто. Дело в том, что поле скорости искомого движения жидкости должно быть всюду конечным (не содержать особенностей в области течения). Это требование в данном случае является настолько жестким, что удастся указать только одно точное решение, регулярное во всей плоскости. Оно отвечает функции

$$u + iv = (C - \sqrt{C^2 - 1}) \operatorname{th} \frac{1}{2}(x + iy).$$

Здесь C — произвольная постоянная. Решение было получено и исследовано Стюартом в 1967 году [3] и записывается так:

$$\psi = -\frac{k}{4\pi} \ln \left[C \operatorname{ch} \frac{2\pi y}{d} + \sqrt{C^2 - 1} \cos \frac{2\pi x}{d} \right]; \quad (8)$$

$$k = \frac{8\pi}{\lambda}; \quad d^2 = -\frac{\pi k}{\alpha}.$$

Его можно интерпретировать как бесконечную цепочку горизонтально расположенных распределенных вихрей. При удалении по вертикали в $\pm\infty$ течение стремится к однородным потокам, текущим в противоположных направлениях и имеющим скорости $U_{\pm\infty} = \mp \frac{k}{2d}$. Характер течения становится очень простым в двух предельных случаях. При $C = 1$ решение (8) описывает сдвиговый поток с профилем в виде гиперболического косинуса, при $C \gg 1$ оно определяет уже ряд точечных вихрей, отстоящих друг от друга на d . В общем случае при $C > 1$ на каждом периоде течения имеется локализованное вихревое образование, завихренность которого растет при приближении к его центру.

1.3. Обобщенное разделение переменных для уравнения Пуассона

Вид точного решения (8) наталкивает на мысль искать новые варианты решений, используя следующее представление:

$$\psi = G(\varphi(x) + g(y)),$$

где G, φ, g — некоторые функции, причем ни одна из функций φ, g не является многочленом степени два или меньше. Это исследование проделали авторы монографии [4], которые установили, что уравнение (3) допускает разделение переменных такого вида только для следующих типов правых частей $f(\psi)$ с произвольными постоянными α_1, α_2 :

$$\begin{aligned} & \alpha_1 \psi \ln \psi + \alpha_2 \psi, \quad \alpha_1 e^\psi + \alpha_2 e^{-2\psi}, \\ & \alpha_1 \sin \psi + \alpha_2 [\sin \psi \cdot \ln(\operatorname{tg}(1/4\psi)) + 2 \sin(1/2\psi)], \\ & \alpha_1 \sin \psi + \alpha_2 (\operatorname{sh} \psi \cdot \ln(\operatorname{th} 1/4\psi) + 2 \operatorname{sh}(1/2\psi)), \\ & \alpha_1 \operatorname{sh} \psi + \alpha_2 (\operatorname{sh} \psi \cdot \operatorname{arctg}(\exp(1/2\psi)) + \operatorname{ch}(1/2\psi)). \end{aligned}$$

Новые классы точных решений описывают течения с разнообразной топологией линий тока:

- дважды периодические решетки вихрей;
- течения с источником и стоком;
- течения между двумя стенками;
- течение, иллюстрирующее соударение струй.

Исследование [4] подводит важный промежуточный итог в изучении точных решений уравнения Пуассона.

1.4. Обтекание траншеи с образованием присоединенного вихря

Выше изучались свободные вихревые образования, существующие в безграничном пространстве. Но не менее интересно и важно уметь определять и параметры присоединенных вихрей, то есть вихрей, обра-

зующихся за обтекаемыми телами. Все известные результаты здесь получены численными методами. Приведем, однако, пример точного аналитического решения для такого рода задач [5].

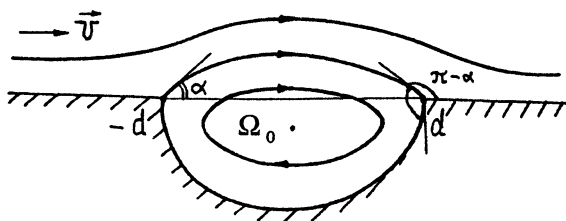


Рис. 2

Рассмотрим обтекание однородным потоком U стационарного вихря, прикрепленного к цилиндрической ямке в плоскости — траншее (рис. 2). Предположим, что верхняя граница вихря — дуга окружности, а угол α , под которым она отходит от плоскости, меньше прямого. Во внешности вихря течение потенциально. Его потенциал представляется формулой

$$\Phi^*(\bar{W}) = \frac{2dU\pi}{\pi - \alpha} \left[1 - \left(\frac{\bar{W} - d}{\bar{W} + d} \right)^{\pi/(\pi - \alpha)} \right]^{-1}.$$

Это выражение описывает потенциальное обтекание однородным потоком возвышения в виде сегмента, опирающегося на отрезок длиной $2d$. В частном случае $\alpha = \pi/2$ он равен функции Жуковского с точностью до несущественного постоянного слагаемого. Поле скорости потенциального течения задается выражением

$$V_{\Pi} = 4d^2\beta^2U \frac{(\bar{W}^2 - d^2)^{\beta-1}}{\left[(\bar{W} + d)^{\beta} - (\bar{W} - d)^{\beta} \right]^2}; \quad \beta = \frac{\pi}{\pi - \alpha}, \quad (9)$$

откуда, в частности, видно, что скорость на бесконечности стремится к постоянной U .

Будем полагать, что завихренность внутри вихревой области распределена однородно и равна Ω_0 . Тогда поле скорости внутри вихря будет иметь вид (4). Кроме того, потребуем, чтобы на границе склейки вихревого и потенциального течений скорость оставалась непрерывной (это условие соответствует обтеканию по схеме Лаврентьева). Положим для простоты, что границе вихря отвечает значение функции тока $\psi_{\Gamma} = 0$, тогда уравнение границы можно представить в виде

$$\Phi(\overline{W}_{\text{гр}}) - \Phi(W_{\text{гр}}) = \left(\frac{\overline{W}_{\text{гр}} + d}{\overline{W}_{\text{гр}} - d} \right)^{\beta} - \left(\frac{W_{\text{гр}} + d}{W_{\text{гр}} - d} \right)^{\beta} = 0.$$

Это равенство определяет связь на границе вихря $W_{\text{гр}}$ и $\overline{W}_{\text{гр}}$, а значит, и вид ее аналитического продолжения — функции $W_{\text{гр}}(\overline{W})$. Зная ее, запишем выражение для поля скорости в вихре:

$$V_{\text{в}} = i \frac{\Omega_0}{2} \left[W - W_{\text{гр}}(\overline{W}) \right] + 4d^2 \beta^2 U \frac{(\overline{W}^2 - d^2)^{\beta-1}}{\left[(\overline{W} + d)^{\beta} - (W - d)^{\beta} \right]^2}. \quad (10)$$

Из него видно, что на границе вихря потенциальное поле скорости (9) и вихревое (10) совпадают, то есть выполнено условие склейки потоков.

По известному полю скорости (10) легко определяется форма траншеи, совпадающей с нижней границей вихря (на ней значение функции тока тоже равно нулю). Полный анализ решения (9), (10) позволяет заключить [6], что вихрь в траншее существует только в том случае, если дуга окружности верхней границы вихря подходит под углом $\alpha < \pi/3$. Причем, при заданном угле α величина $|U(d\Omega_0)^{-1}|$ должна превышать некоторое критическое значение.

2. Птолемеевские течения

В первом разделе мы рассматривали стационарные течения. Уложим несколько ситуацию и обратимся к изучению плоских нестационарных движений идеальной жидкости. Воспользуемся для этого лагранжевыми переменными.

Как известно, уравнения движения жидкости можно представить в двух различных формах в зависимости от того, интересуемся ли мы значениями параметров потока в произвольной точке пространства или стремимся определять «историю» индивидуальной жидкой частицы. Уравнения, которые получаются этими двумя способами, называются соответственно эйлерова и лагранжева формы уравнений гидродинамики, хотя на самом деле обе их предложил Эйлер в 1755 году (за 33 года до Лагранжа) в трактате «Общие принципы движения жидкостей». Уравнения (1), (2) соответствуют эйлеровой формулировке.

Чтобы получить уравнения в форме Лагранжа, запишем уравнение Ньютона для отдельной жидкой частицы

$$\mathbf{r}_{\text{н}} = -\frac{1}{\rho} \nabla p, \quad (11)$$

где $\mathbf{r}(a, b, t)$ — радиус-вектор жидкой частицы; a, b — параметры, нумерующие ее (ими, например, могут быть значения ее координат в начальный момент времени). Для исключения в этом уравнении дифференцирования по неизвестной функции, умножим его последовательно на \mathbf{r}_a и \mathbf{r}_b . Тогда получим

$$\mathbf{r}_a \mathbf{r}_a = X_u X_a + Y_u Y_a = -\frac{1}{\rho} \nabla p \mathbf{r}_a = -\frac{1}{\rho} p_a; \quad (12)$$

$$\mathbf{r}_u \mathbf{r}_b = X_u X_b + Y_u Y_b = -\frac{1}{\rho} \nabla p \mathbf{r}_b = -\frac{1}{\rho} p_b. \quad (13)$$

Вместе с уравнением непрерывности (условием сохранения площади жидкого элемента)

$$\frac{D(X, Y)}{D(a, b)} = \frac{D(X_0, Y_0)}{D(a, b)} = D(a, b), \quad (14)$$

где X_0, Y_0 — начальные координаты жидкой частицы, а D здесь и далее обозначает якобиан преобразования неизвестных функций по лагранжевым переменным, эти уравнения образуют систему уравнений в форме Лагранжа. Если лагранжевые координаты совпадают с X_0, Y_0 , то $D = 1$. В остальных случаях это некоторая не зависящая от времени функция.

Уравнения движения (12), (13) можно свести к одному уравнению, уже не содержащему давление. Продифференцируем первое из них по b , второе по a и вычтем одно из другого. В результате придем к следующему уравнению:

$$X_{ia} X_b - X_{ib} X_a + Y_{ia} Y_b - Y_{ib} Y_a = \Omega D(a, b). \quad (15)$$

Оно выражает теорему Гельмгольца о сохранении завихренности жидких частиц $\Omega(a, b)$ при плоском движении.

Перейдем в уравнениях (14), (15) к комплексным координатам и неизвестным функциям по формулам

$$\chi = a + ib, \quad \bar{\chi} = a - ib; \quad W = X + iY, \quad \bar{W} = X - iY.$$

В новых переменных уравнения движения (12), (13) запишутся так [6]:

$$\frac{D(W, \bar{W})}{D(\chi, \bar{\chi})} = D(a, b); \quad \frac{D(W_t, \bar{W})}{D(\chi, \bar{\chi})} = i \frac{\Omega}{2} D(a, b). \quad (16)$$

Такая компактная форма записи уравнений двумерной гидродинамики позволяет открыть целый класс точных решений. Непосредственной подстановкой в (16) можно убедиться, что выражение

$$W = G(\chi) e^{i\lambda t} + F(\bar{\chi}) e^{i\mu t}, \quad (17)$$

где G, F — аналитические функции, λ, μ — произвольные действительные числа, является его решением. Функции G и F в значительной степени произвольны, так как единственным ограничением на их выбор является требование необращения в нуль якобиана, зависящего от квадратов модулей их производных: $D = |G'|^2 - |F'|^2$.

Траекториями жидких частиц в течениях (17) будут эпициклоиды (гипоциклоиды), т. е. частицы описывают окружность, центр которой, в свою очередь, движется по другой окружности. По таким орбитам вращались планеты в Птолемеевой картине Мира. В связи с этим данный тип течений был назван птолемеевским [6].

Птолемеевские течения — вихревые. Завихренность для них определяется выражением

$$\Omega = 2 \left(\frac{\lambda |G'|^2 - \mu |F'|^2}{|G'|^2 - |F'|^2} \right). \quad (18)$$

В общем случае она неоднородно распределена в плоскости течения.

К числу птолемеевских течений относятся волны Герстнера — гравитационные волны на поверхности жидкости. Выражение, описывающее их, имеет вид

$$W = \chi + iA \exp i(k\bar{\chi} - \omega t),$$

где A — амплитуда волны; k — волновое число; ω — частота волны. Это единственное точное решение для волн на воде, найденное Герстнером еще в 1804 году.

Приведем оригинальные примеры вихрей [6]. Рассмотрим односвязную вихревую область, потенциально обтекаемую снаружи. Пусть выражение

$$W = \alpha v \exp(i\lambda t) + F(\bar{v}); \quad v = \exp(ik\chi), \quad |v| \leq 1, \quad (19)$$

где α — некоторая положительная постоянная, задает птолемеевское течение внутри этой области. Предположим также, что на плоскости лагранжевой переменной v ей соответствует внутренность единичного круга, что эквивалентно условию $b \geq 0$. Выражение (19) является частным случаем птолемеевских течений (17) при $G(\chi) = \exp(ik\chi)$, $\mu = 0$. Форма вихревой области при таком выборе может быть в значительной степени произвольной в зависимости от вида функции F .

Обратимся теперь к определению потенциального течения во внешности вихря. Для этого поступим следующим образом. Запишем комплексную координату некоторой точки внешней области W и комплексную скорость в этой точке V в виде следующих функций комплексного параметра v ($|v| \geq 1$):

$$W = \alpha v \exp(i\lambda t) + F\left(\frac{1}{v}\right), \quad V = \frac{i\lambda\alpha}{v} \exp(i\lambda t). \quad (20)$$

Эти соотношения служат параметрической формой записи искомого потенциального течения. Действительно, из них следует, что V — аналитическая функция от комплексно сопряженной координаты \bar{W} , т. е. течение потенциально. Кроме того, функции (20) совпадают с выражениями для W и V птолемеевского течения на границе вихря. Следовательно, они удовлетворяют условию непрерывности скорости на границе склейки, а следовательно, и условию непрерывности давления.

Переменная v в выражениях (20), подчеркнем, уже не является лагранжевой координатой. Она совпадает с переменной Лагранжа только на границе вихря (при $|v|=1$). Заметим также, что закон спадания скорости на бесконечности такой же, как и для точечного вихря. Именно с целью удовлетворить этому условию выбирались вид G и значение μ .

Завихренность внутри области определяется формулой

$$\Omega = \frac{2\lambda}{1 - \alpha^{-2} |F'|^2}.$$

Для всевозможных функций F , за исключением линейной, она минимальна в центре вихря и возрастает к его границе. Траекториями жидких частиц служат окружности, причем у каждой частицы свой центр вращения, положение которого диктуется конкретным значением функции F .

В построенном решении функция F должна удовлетворять неравенству

$$|F'| \leq \alpha \quad \text{при} \quad |v| \leq 1$$

для того, чтобы детерминант $D(\chi, \bar{\chi})$ не обращался в нуль внутри вихря. Выполнение этого условия гарантирует, что потенциальное поле скорости во внешности вихря не имеет особых точек.

Известное точное решение для эллиптического вихря Кирхгофа получается из (19), если взять $\alpha = \frac{1}{2}(A+B)$, $F(\bar{v}) = \frac{1}{2}(A-B)\bar{v}$. Очевидно, что границей такого вихря является эллипс A, B , который равномерно вращается с частотой $\lambda/2$ вокруг своего центра. Сами жидкие частицы при этом вращаются по окружности с вдвое большей частотой. Вихрь Кирхгофа является простейшим примером семейства вихрей, которые равномерно вращаются без изменения своей формы. Во вращающейся с угловой скоростью вихря системе отсчета движение в вихре стационарно, линии тока (одна из которых граница вихря) совпадают с траекториями жидких частиц и, следовательно, вихри, границей которых явля-

ются гипоциклоиды, не меняют своей формы. Аналитически они описываются выражением

$$W = \alpha v e^{i\lambda t} + \beta \bar{v}^n, \quad (21)$$

где n — целое неотрицательное число. Оно определяет семейство вихрей неизменной формы, равномерно вращающихся как целое с угловой скоростью $\omega = n\lambda(n+1)^{-1}$. Для $n = 1$ получаем эллиптический вихрь; вихри, соответствующие значениям $n \geq 2$, представляют области гипоциклоидальной формы с числом выступов $n+1$. Условием отсутствия самопересечения границы является неравенство $\beta \leq \frac{\alpha}{n}$, оно обеспечивает однозначность поля скорости потенциального течения. Формула для завихренности течения (21) имеет вид

$$\Omega = \frac{2\lambda}{1 - \left(\frac{\beta n}{\alpha}\right)^2 |v|^{2(n-1)}}.$$

Отсюда ясно, что для эллиптического вихря завихренность постоянна, а для остальных членов семейства она минимальна в центре и возрастает к границе.

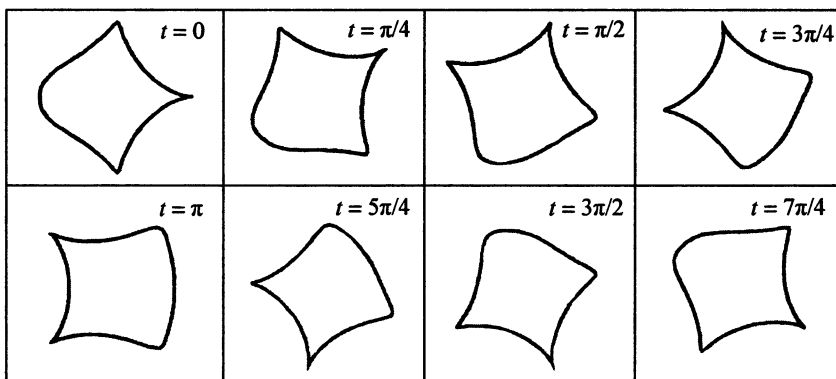


Рис. 3

Для функции F , отличной от степенной, вихревая область помимо вращения деформируется довольно сложным образом. На рис. 3 воспроизведена динамика границы вихревой области течения, описываемого функцией $F(\bar{v}) = 0,2\bar{v}^3 + 0,1\bar{v}^4$. Из него видно, что начальная форма вихря повторяется через период, а ее форма зависит исключительно от вида F .

3. Конические течения

Все вышеприведенные примеры относились к течениям идеальной жидкости, когда вязкость полагается равной нулю. Но завершить рассмотрение нам хочется примерами точных решений для уравнения Навье — Стокса. Их очень немного, они буквально наперечет и, как правило, соответствуют простейшим предположениям о структуре поля скорости. Например, в плоскопараллельных течениях и потоках с круговыми линиями тока нелинейное слагаемое в уравнении (1) обращается в нуль, и уравнения Навье — Стокса легко интегрируются. Но мы хотим выделить особый класс течений, когда инерционные члены как раз оказываются существенными и определяют особенности течений.

Рассмотрим стационарные течения несжимаемой вязкой жидкости. Заметим, что если все компоненты поля скорости будут обратно пропорциональны расстоянию до начала координат R , то инерционные и вязкий члены будут оба обратно пропорциональны R^3 . Давление p можно положить обратно пропорциональным R^2 , тогда градиент давления тоже будет обратно пропорционален кубу радиуса. Таким образом, выбрав оговоренные представления для скорости и давления, можно уменьшить число независимых переменных.

Итак, пусть поле скорости течения обратно пропорционально R и, кроме того, для простоты будем считать, что оно осесимметричное ($\frac{\partial}{\partial \varphi} = 0$) и незакрученное (т. е. $V_\varphi = 0$; R, θ, φ — сферическая система координат). С помощью автомодельной подстановки вида

$$V_R = -\frac{v}{R} f'(\eta), \quad V_\theta = -\frac{v}{R} \frac{f(\eta)}{\sqrt{1-\eta^2}},$$

$$\frac{p-p_\infty}{\rho} = \left(\frac{v}{R}\right)^2 P(\eta), \quad \mathbf{V} = \mathbf{V}(V_R, V_\theta, 0), \quad \eta = \cos \theta,$$
(22)

где $vRf(\eta)$ — функция тока, p_∞ — давление на бесконечности, уравнения Навье — Стокса приводятся к следующему уравнению в обыкновенных производных:

$$-2(1-\eta^2)f' + f^2 - 4\eta f + \alpha\eta^2 + \alpha_1\eta + \alpha_2 = 0. \quad (23)$$

Здесь $\alpha, \alpha_1, \alpha_2$ — некоторые постоянные. Из выражений (22) видно, что при заданном R все неизвестные функции зависят только от меридионального угла θ и принимают одинаковое значение на конической поверхности, соответствующей этому углу. Вследствие этого такие течения называют еще конически симметричными. Уравнение (23) впервые получил Н. А. Слезкин [7]. Независимо его впоследствии переоткрыл Л. Д. Ландау [8].

Общее решение (23) может быть выражено через гипергеометрические функции. Описываемое им поле скорости всегда имеет особенности за исключением единственного случая, когда $\alpha = \alpha_1 = \alpha_2 = 0$. Ему соответствует функция тока вида

$$f = \frac{2 \sin^2 \theta}{C - \cos \theta}, \quad (24)$$

где C — постоянная. Это выражение определяет одностороннее струйное течение. Вследствие выбранной зависимости от R поток импульса через произвольную сферическую поверхность, охватывающую начало координат, постоянен и задается величиной C . Ландау предложил интерпретировать течение (24) как движение, инициируемое находящимся в начале координат источником импульса. Это течение принято называть затопленной струей Ландау [8].

Коническим течениям посвящено множество публикаций. Их полный обзор приведен в [9]. Главный вопрос, который обсуждают авторы этих работ, — возможность интерпретации решений (23) для объяснения свойств реальных течений — например, струй, вытекающих из отверстия в стене или течений внутри конуса. Мы, в свою очередь, предлагаем оригинальный пример течений (22), имеющих физический смысл и допускающих простую интерпретацию [10].

Пусть квадратный трехчлен в (22) имеет вид

$$\alpha \eta^2 + \alpha_1 \eta + \alpha_2 = \alpha(\eta^2 - 1); \quad \alpha > 0.$$

Положим, что $f = -2(1 - \eta^2) \frac{g'}{g}$, тогда уравнение (23) сводится к линейному уравнению $(1 - \eta^2)g'' + \alpha g = 0$, решение которого записывается следующим образом:

$$g = (1 - \eta^2) [C_1 P'_\mu(\eta) + C_2 Q'_\mu(\eta)], \quad \alpha = \mu(\mu + 1).$$

Здесь $P_\mu(\eta), Q_\mu(\eta)$ — функции Лежандра соответственно I и II рода; C_1, C_2 — постоянные. Функция тока рассматриваемого течения равна соответственно

$$f = 2(2\eta - \frac{P'_\mu + CQ'_\mu}{P_\mu + CQ_\mu}); \quad C = \frac{C_2}{C_1}. \quad (25)$$

Задаваемое ею поле скорости зависит от двух произвольных постоянных. Оно регулярно всюду, кроме полуосей $\eta = \pm 1$. Вблизи положительной полуоси поле скорости представляется так:

$$V_\theta = \frac{\sqrt{2\nu\alpha}}{R} \sqrt{x} \ln \alpha x; \quad V_R = -\frac{2\nu\alpha}{R} \ln \alpha x; \quad x = 1 - \eta.$$

Из этих выражений видно, что меридиональная скорость обращается в нуль на оси, а радиальная имеет особенность, причем жидкость вытекает вдоль оси ($V_R > 0$). Точно так же можно показать, что жидкость вытекает и вдоль полуоси $\eta = -1$. Таким образом, решение (25) описывает двустороннюю струю, скорость выплеска которой на осях выплеска обращается в бесконечность.

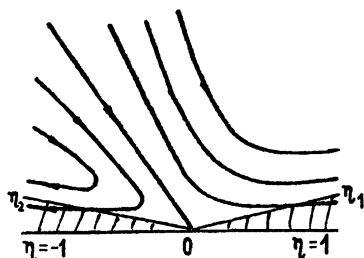


Рис. 4

Особенность радиальной скорости — логарифмическая, «слабая». Ее можно «устранить», если предположить, что на осях выплеска течение турбулизуется. В работе [10] показано, что если рассмотреть течение, которое турбулизовано внутри приосевых конических областей (их границы определяются значениями η_1, η_2 — см. рис. 4), а в пространстве между этими конусами по-прежнему задается формулами (22), (23), то поле скоростей будет всюду конечным. Течение (25) при этом будет предельным для двусторонней осесимметричной струи с турбулизованными приосевыми зонами.

Работа выполнена при финансовой поддержке РФФИ (гранты №02-02-17089, НШ – 1637.2003.2) и президиума РАН по программе «Математические методы в нелинейной динамике».

Литература

1. Чаплыгин С. А. Один случай вихревого движения жидкости // Труды физической секции Московского Имперского общества естествоиспытателей. Т. 11. С. 11—14.
2. Ламб Г. Гидродинамика. 2-е изд. М., 1895.
3. Stuart J. T. On finite amplitude oscillations in laminar mixing layer // J. Fluid Mech. 1967. V. 29, № 3. P. 417—440.
4. Андреев В. К., Капцов О. В., Пухначев В. В., Родионов А. А. Применение теоретико-групповых методов в гидродинамике. Новосибирск: Наука, 1994. 318 с.
5. Абрашкин А. А., Якубович Е. И. Обтекание неоднородностей плоской поверхности с образованием присоединенного вихря // ЖПМТФ. 1988. № 5. С. 81—84.
6. Абрашкин А. А., Якубович Е. И. О плоских вихревых течениях идеальной жидкости // ДАН СССР. 1984. Т. 76, № 1. С. 76—78.
7. Слезкин Н. А. Об одном случае интегрируемости полных дифференциальных уравнений вязкой жидкости // Учен. зап. МГУ. 1934. Вып. 2. С. 89—90.
8. Ландау Л. Д. Об одном точном решении уравнений Навье — Стокса // ДАН СССР. 1944. Т. 43, № 7. С. 299—301.
9. Гольдштик М. А., Штерн В. Н., Яворский Н. И. Вязкие течения с парадоксальными свойствами. Новосибирск: Наука, 1989. 366 с.
10. Абрашкин А. А. К теории автомодельных осесимметричных течений вязкой жидкости: Препринт ИПФ АН № 195. Горький, 1988.

ПРОСТРАНСТВЕННО-ВРЕМЕННАЯ ДИНАМИКА СОЛИТОНОВ В РЕЗОНАТОРАХ

А. Б. Езерский

Введение

Возбуждение и взаимодействие солитонов в пространственно неограниченных системах являются предметом исследования очень большого числа теоретических и экспериментальных работ. В ряде волновых задач солитоны возбуждаются в резонаторах. Учет резонаторных эффектов принципиален при возбуждении электромагнитных волн в джозефсоновских переходах [1, 2] или в их механическом аналоге — цепочке маятников, связанных пружинами, в LC-линиях [3], при возбуждении солитонов поверхностных волн или внутренних волн в мелких бассейнах [4], солитонов поперечных упругих волн в отрезках резины [5] (см. рис. 1) и во многих других системах. Отражение на границах нелинейной среды может приводить к качественно новым режимам генерации: солитоны могут возбуждаться на фоне собственных резонаторных мод, что необходимо учитывать для объяснения особенностей режимов, возникающих в этих системах. Кроме этого, в отличие от бесконечно протяженных сред, в резонаторах из-за отражения солитонов от боковых стенок происходят их многократные столкновения, что также существенно усложняет динамику солитонных взаимодействий.

Понятно, что ни одну из нелинейных волн, наблюдающихся в ограниченных системах, с точки зрения строгого математического определения нельзя назвать уединенной волной — солитоном. Однако, как будет показано в статье, при определенных приближениях характеристики нелинейных режимов можно объяснить, если использовать солитон как элементарный объект. В данной статье это будет показано на примере наглядного физического эксперимента по возбуждению солитонов поверхностных волн в 10-метровом лотке гармоническим внешним воздействием.

1. Описание эксперимента

Для изучения пространственно-временной динамики в распределенных системах используются (X - T)-диаграммы: значение переменной (это может быть поле температуры, концентрация примеси, компонента скорости, которые, вообще говоря, зависят от трех координат и времени) регистрируется на интервале $X_1 < X < X_2$ в моменты времени T_1, T_2, T_3, \dots

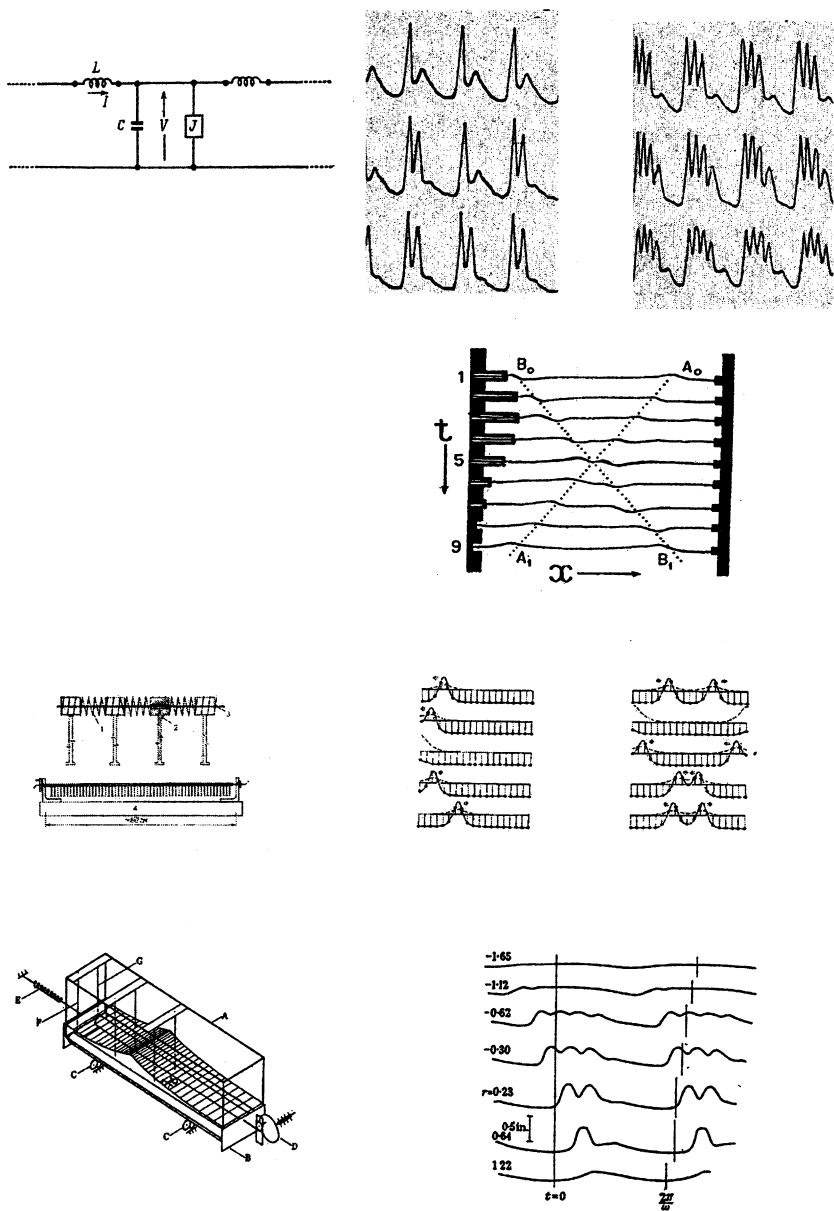


Рис. 1. Солитоны в LC-линии [3] при параметрическом возбуждении волн в полоске резины [4], в цепочке связанных пружинками маятников [1, 2], при параметрическом возбуждении волн в кювете с жидкостью [5].

Определенная таким образом функция двух переменных, (X-T)-диаграмма, строится в ярком виде. Если система одномерная, то диаграмма полностью характеризует ее динамику. В эксперименте для построения таких диаграмм необходимо либо регистрировать сигналы с множества датчиков, либо снимать фильм и обрабатывать мгновенные снимки. При изучении динамики солитонов в резонаторе мы исследовали именно эти диаграммы, которые в наглядном виде представляют информацию о нелинейных волновых взаимодействиях.

1.1. Получение пространственно-временных диаграмм

Солитоны возбуждались в бассейне длиной $L = 10$ метров, и шириной $D = 0,5$ метра. Глубина жидкости была $H = 0,26$ метра. Поверхностные волны возбуждались волнопродуктором, который представлял собой плоскую пластину, периодическое движение которой осуществлялось за счет электромотора. Измерения смещения свободной поверхности воды проводились двумя датчиками сопротивления. Один из датчиков мог перемещаться с помощью координатной системы на расстояние до 5,6 метра в центральной части бассейна, а второй был неподвижным и располагался у стенки, отражающей поверхностные волны. Используя сигналы от двух датчиков, мы могли строить пространственно-временные диаграммы. При этом сигнал от неподвижного датчика использовался как опорный — по нему определялся момент, с которого надо записывать сигнал с датчика, регистрирующего смещение свободной поверхности в центральной части канала. Понятно, что при таком подходе пространственно-временные диаграммы можно построить только для периодических во времени сигналов.

Волнопродуктор позволял возбуждать резонаторные моды с частотами $f_n = \frac{1}{2\pi} \sqrt{gk_n \tanh(k_n H)}$, где $k_n = \frac{\pi n}{L}$, $n = 1, 2, 3, \dots$; g — ускорение свободного падения. В ходе экспериментов частота волнопродуктора выбиралась близкой к частоте моды, длина которой была равна длине резонатора L , то есть $n = 2$, $f_2 \approx \frac{1}{L} \sqrt{gH}$. В условиях эксперимента резонансная частота этой моды составляла $f_2 = 0,1645$ Гц. Измерения показали, что эта мода имеет достаточно высокую добротность. Относительная ширина резонансной кривой (по уровню 0,707) составляла $\gamma = \Delta f_2 / \Delta f \leq 0,02$. Режим нелинейных колебаний в лотке зависит от амплитуды a_{ex} и разности между частотой волнопродуктора и резонансной частотой f_2 возбуждаемой моды.

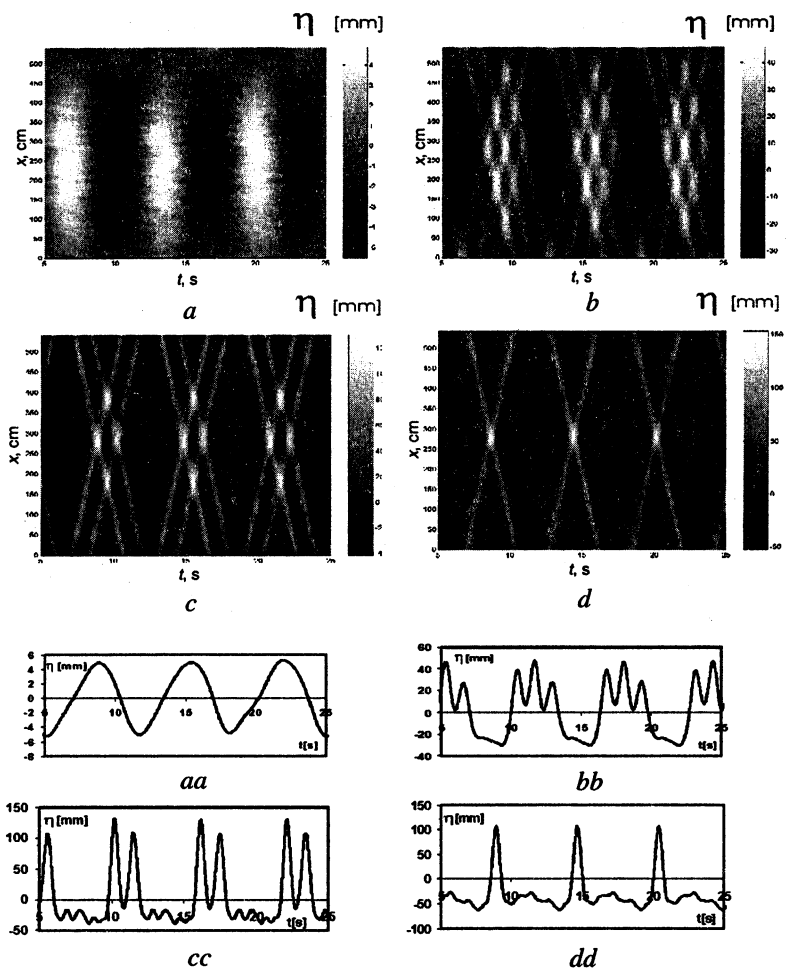


Рис. 2. Пространственно-временные диаграммы и временные ряды с неподвижного датчика: *a, aa* — $a_{ex} = 2,3$ см, $f_{ex} = 0,151$ Гц, $\Delta = (f_{ex} - f_2)/f_2 = -0,087$; *b, bb* — $a_{ex} = 8,2$ см, $f_{ex} = 0,158$ Гц, $\Delta = -0,045$; *c, cc* — $a_{ex} = 8,2$ см, $f_{ex} = 0,170$ Гц, $\Delta = 0,028$; *d, dd* — $a_{ex} = 8,2$ см, $f_{ex} = 0,177$ Гц, $\Delta = 0,07$.

На рис. 2 представлены характерные диаграммы для различных параметров внешнего воздействия. На этом же рисунке представлены временные ряды, полученные при помощи неподвижного датчика, расположенного у стенки. Как видно из этих диаграмм, при малой амплитуде внешней силы в резонаторе возбуждаются стоячие поверхностные волны (рис. 2, *a* и соответствующий временной ряд (*aa*)). При больших амплитудах в системе возникают распространяющиеся импульсы (рис. 2, *b—d*)

и временные ряды (bb)—(dd). Импульсы на пространственно-временных диаграммах видны как яркие полосы. Наклон этих полос соответствует скорости распространения импульсов и в условиях нашего эксперимента близок к скорости поверхностных волн на мелкой воде (подробное обсуждение скорости распространения импульсов будет проведено в дальнейшем). Число импульсов зависит от частоты возбуждения. Под числом импульсов мы понимаем следующее. Представим стоячую волну как суперпозицию двух бегущих волн, распространяющихся в противоположных направлениях. Синхронно с каждой из этих волн распространяются импульсы — в большинстве наших экспериментов от 1 до 3. Понятно, что количество импульсов, зарегистрированных датчиком, будет разным в зависимости от точки резонатора. Легко понять, что только в центре резонатора и у его боковых границ количество импульсов, регистрируемых датчиками, совпадает с количеством импульсов в каждой из бегущих волн. Эксперименты показали, что для частоты меньшей, чем резонансная частота моды, возбуждаются три импульса на периоде волны. Две пересекающиеся тройки таких треков представлены на рис. 2, *b*. А при частоте возбуждения большей, чем резонансная, — один: только два пересекающихся трека представлены на рис. 2, *d*.

Разбиение плоскости параметров на области с качественно различной динамикой показано на рис. 3. Внутри областей, отмеченных цифрами, наблюдается возбуждение соответственно 1, 2, 3 импульсов на периоде волн. При достаточно большой амплитуде есть небольшая область, где наблюдается возбуждение 4 импульсов. При управляющих параметрах, попадающих в зоны, отмеченные пунктирной линией, наблюдалась мультистабильность: при одних и тех же величинах частоты и амплитуды колебаний волнопродуктора, но различных начальных условиях могли наблюдаться различные режимы.

Например, в зоне 2—3 наблюдались возбуждение двух, трех или перенная генерация двух и трех импульсов на периоде. Последний из режимов соответствует удвоению периода — нелинейные волны, возбуждаемые в системе, имеют период вдвое больший, чем период колебаний волнопродуктора.

Пространственно-временные диаграммы показывают, что импульсы, распространяющиеся в противоположных направлениях, взаимодействуют. На диаграммах можно найти, например, области, где импульсы сливаются. Вместо двух максимумов в результате слияния в определенной области диаграммы возникает один максимум. Размер области, где происходит слияние, зависит от амплитуды импульсов. Если мы определим скорость движения импульса как скорость точки, соответствующей максимуму, то взаимодействие импульсов приводит к существенному изменению их скорости.

a_{ex} [cm]

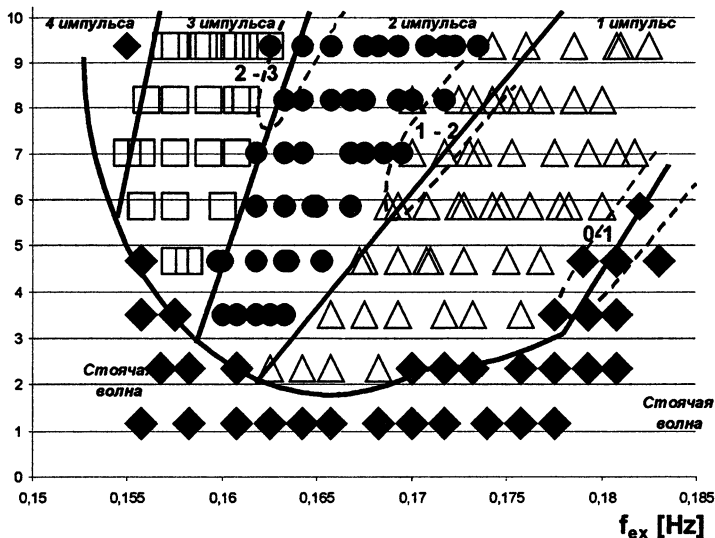


Рис. 3. Разбиение плоскости (f_{ex} , a_{ex}) на области с существенно различными режимами генерации нелинейных волн [6]. Области, выделенные пунктирными линиями, соответствуют параметрам, при которых режимы генерации зависят от начальных условий.

1.2. Выделение солитонов

Детальный анализ экспериментальных данных показал, что для определенного интервала управляющих параметров импульсы, возбуждаемые в резонаторе, близки по своим характеристикам к солитонам. Чтобы доказать это, мы предположили, что в резонаторе солитоны возбуждаются на фоне гармонической волны. Вопрос состоит в том, как разделить импульсы и гармонические колебания. Очевидно, что это нельзя сделать при помощи линейных фильтров: частота повторения импульсов в точности равна частоте гармонических колебаний. Поэтому при помощи пакета LabView были составлены программы, позволяющие срезать импульсы и сглаживать полученные временные ряды. Подробно процедура обработки сигнала обсуждалась в [6]. Было установлено, что форма импульсов близка к классическим солитонам $\cosh^{-2}(t/\tau_s)$ всюду за исключением областей, где амплитуда поля мала. Длительность импульса τ_s уменьшается с увеличением амплитуды импульса A_s . Экспериментальные данные представлены на рис. 4. На этом же графике пока-

зана теоретическая зависимость длительности классического солитона от его амплитуды.

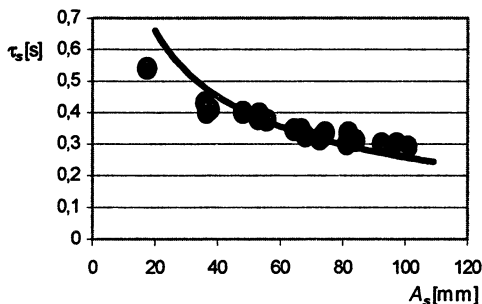


Рис. 4. Экспериментально полученная зависимость длительности солитона от амплитуды (точки). Для сравнения представлена теоретическая зависимость

$$\tau_s = \frac{1}{V_s} \sqrt{\frac{8H^3}{3A_s}}, \quad V_s = V_0 \left(1 + \frac{A_s}{2H}\right) \text{ (кривая)}.$$

Мы назвали наблюдающиеся в эксперименте импульсы солитонами, чтобы подчеркнуть, что они локализованы в определенной области пространства, а их характеристики близки к характеристикам классического солитона.

2. Обсуждение результатов

2.1. Представление временных рядов поверхностных волн

Выделение из временных рядов солитонов и гармонической волны позволяет сравнить свойства нелинейных волн, наблюдающихся в эксперименте, с хорошо известными решениями уравнения Кортевега — де Вриза (см., например, [2]). Смещение свободной поверхности воды подчиняется уравнению

$$\frac{\partial \eta}{\partial t} + V_0 \frac{\partial \eta}{\partial x} + \frac{3}{2} \frac{V_0}{H} \eta \frac{\partial \eta}{\partial x} + \frac{1}{6} V_0 H^2 \frac{\partial^3 \eta}{\partial x^3} = 0, \quad (1)$$

которое имеет решение в виде солитона:

$$\eta_s(x - V_s t) = A_s \cosh^{-2} \left(\sqrt{\frac{3A_s}{8H^3}} (x - V_s t) \right), \quad V_s = V_0 \left(1 + \frac{A_s}{2H}\right), \quad (2)$$

где $V_0 = \sqrt{gH}$ — скорость поверхностных волн бесконечно малой амплитуды в мелкой воде. Понятно, что длительность солитона пропор-

циональна $A_s^{-0.5}$, скорость солитона зависит от его амплитуды — нелинейная добавка к скорости пропорциональна A_s .

Мы предполагаем, что нелинейные волны, возбуждаемые в резонаторе, состоят из двух солитонов, распространяющихся навстречу друг другу, причем каждый из солитонов распространяется на фоне гармонической волны. В нулевом приближении мы пренебрежем взаимодействием встречных волн и представим нелинейное поле смещения поверхности в виде четырех компонент:

$$\eta_s(x - V_s t) + \eta_0 \sin(\omega t - kx - \varphi_s) + \eta_s(x + V_s t) + \eta_0 \sin(\omega t + kx - \varphi_s), \quad (3)$$

где ω, k — частота и волновое число гармонических волн; η_0 — амплитуда волн; φ_s — фазовый сдвиг между солитоном и гармонической волной. Отметим, что рядом с боковой стенкой, где расположен неподвижный датчик, смещение свободной поверхности может быть представлено как $\eta = 2\eta_0 \sin(\omega t - \varphi_s) + 2\eta_s(t)$.

Представление (3) пренебрегает взаимодействием встречных гармонических волн и солитонов. В частности, хорошо известный эффект увеличения амплитуды солитона при отражении от стенки и появления фазового сдвига не может быть описан в рамках представления (3). В предположении, что модель (3) справедлива, были измерены параметры солитонов. Используя процедуру разделения сигнала на гармоническую волну и солитон (п. 1.2) по временным рядам, полученным с неподвижного датчика, мы выяснили зависимость характеристик солитона от управляющих параметров, кроме зависимости длительности солитона на уровне $0,4199A_s$ ($0,4199 = \cosh^{-2}(1)$) от его амплитуды (рис. 4). На этом же рисунке был измерен также фазовый сдвиг φ_s . При малой амплитуде внешней силы фазовый сдвиг был близок к нулю. С увеличением амплитуды фазовый сдвиг увеличивался до 40 — 60 градусов. Амплитуда солитона зависела от частоты и амплитуды волнопродуктора. При одной и той же амплитуде внешнего воздействия амплитуда солитона и фаза солитона зависели от частоты внешнего воздействия.

Измеряя значения времени, при которых достигается максимум амплитуды солитона для различных положений датчика смещения свободной поверхности, можно определить траекторию движения солитона. Пример таких измерений показан на рис. 5.

Если исключить область, где происходит слияние двух максимумов, то можно определить скорость движения солитона на различных участках траектории. Линейная аппроксимация показала, что скорость солитона имеет приблизительно постоянную величину на всех участках $V_s \approx 2\pi f_{ex} L$. Солитоны распространяются синхронно с гармонической

волной. В центральной части резонатора, где происходит столкновение двух встречных солитонов, возникает область, где солитоны сливаются — вместо двух максимумов существует один. Размер области слияния L_m зависит от амплитуды солитона. Длина слияния, так же как и размер солитона, уменьшается при увеличении амплитуды солитона.

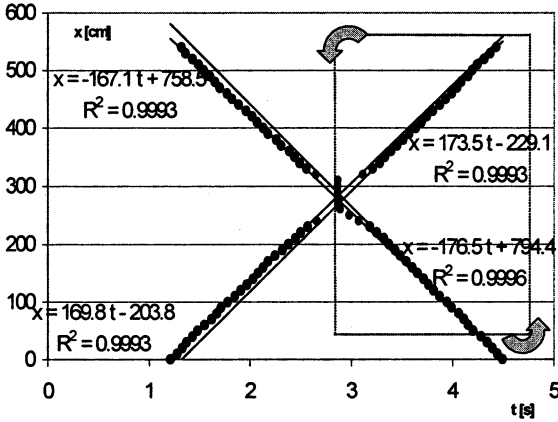


Рис. 5. Траектории солитона в резонаторе, полученные при $a_{ex} = 8,2$ см, $f_{ex} = 0,177$ Гц. Если развернуть выделенную область на 90° , то получатся траектории, изображенные далее на рис. 7.

2.2. Сравнение с результатами Максворти

Полученные нами результаты по взаимодействию солитонов в резонаторах естественно сравнить с результатами по динамике солитонов в безграничных средах, полученными в [7]. Типичная траектория для случая столкновения двух солитонов, показанная на рис. 5, состоит из 4 участков. Стрелки показывают направление движения солитонов. Наилучшая линейная аппроксимация и корреляция точек показаны для каждого из участков траекторий. Укажем некоторые особенности траекторий солитонов. Как показывает линейная аппроксимация траекторий, скорости солитонов на 4 участках несколько отличаются. Хотя это различие небольшое ($\pm 3\%$), оно достаточно хорошо проявляется, так как точки на каждом из участков траектории имеют высокую (0,999) степень корреляции. Имеется также небольшой фазовый сдвиг — задержка солитонов, вызванная их взаимодействием.

Траектории солитонов, взаимодействующих в резонаторе, можно сравнить с траекториями, приведенными в работе Максворти (рис. 6) [7]. Как было показано в этой работе, столкновение двух встречных солито-

нов поверхностных волн одинаковой амплитуды в безграничной среде эквивалентно отражению одного солитона от вертикальной стенки. Действительно, при отражении от вертикальной стенки горизонтальная скорость возмущений равна нулю, а при столкновении двух солитонов одинаковой амплитуды можно выделить сечение (ось симметрии), в котором горизонтальная скорость будет равна нулю. Это обеспечивает эквивалентность двух задач. Максворти обнаружил, что существует определенное время, в течение которого максимум находится в одной точке у боковой стенки. Это время τ , называют временем остановки.

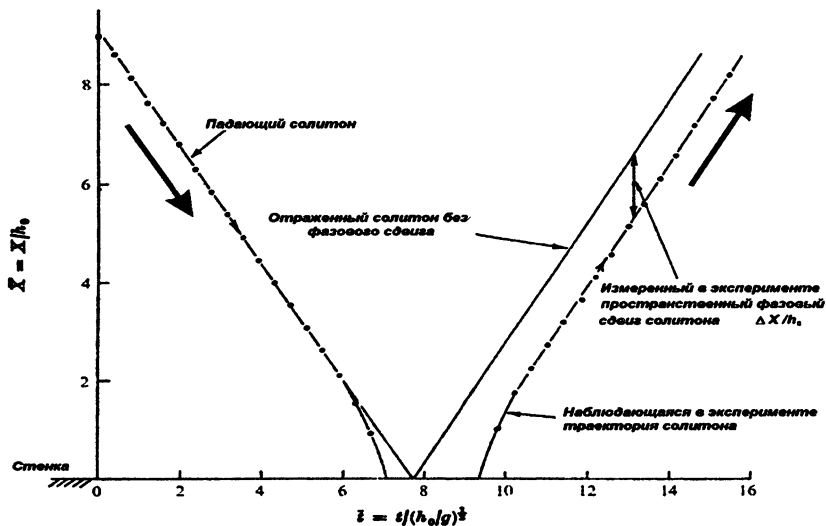


Рис. 6. Траектория солитона, полученная Максворти [7] (по горизонтальной оси отложено безразмерное время, по вертикальной — безразмерная координата).

Вместо длины слияния L_m , которая наблюдается при взаимодействии солитонов в нашем эксперименте, существует время остановки τ , при отражении единичного солитона от вертикальной стенки в экспериментах Максворти. Длина слияния и время остановки уменьшаются при увеличении амплитуды солитонов.

Различие траекторий обусловлено тем, что по-разному обрабатываются данные. В экспериментах [7] анализировались мгновенные снимки — кадры кинофильма. На каждом из мгновенных снимков положение максимума амплитуды определялось при фиксированном значении времени. В нашем эксперименте анализировались временные ряды в различных точках и определялось время, при котором достигается мак-

симум амплитуды при различном значении координаты вдоль резонатора. Подробный анализ того, как способ обработки данных влияет на форму траекторий солитонов, представлен в [6].

Как было отмечено Максворти, амплитуда солитона, отраженного от стенки, несколько меньше, чем до отражения. Это обусловлено потерями. В случае возбуждения солитонов в резонаторе, кроме потерь есть еще взаимодействие солитонов с монохроматической модой, которое может приводить и к увеличению амплитуд солитонов. Поэтому для экспериментов в резонаторе нельзя утверждать, что амплитуда солитонов до взаимодействия должна быть больше, чем после взаимодействия. Это и демонстрируют полученные нами данные. До столкновения скорость солитонов 167,1 см/с и 169,8 см/с, а после столкновения — 173,5 см/с и 176,5 см/с.

2.3. Уравнения для амплитуды и фазы солитона

Сравним характеристики солитонов поверхностных волн, измеренные в эксперименте, с результатами полуфеноменологической теории.

Для описания солитонов электромагнитных волн в [3] были выведены уравнения для амплитуды и фазы солитона, распространяющегося в резонаторе на фоне монохроматической моды. Аналогичные уравнения для поверхностных волн на мелкой воде выглядят следующим образом:

$$\frac{dE_s}{dt} = \frac{3}{2} \frac{\omega \eta_0}{H} E_s \cos \varphi_s - \delta E_s, \quad (4)$$

$$\frac{d\varphi_s}{dt} = \frac{\sqrt{gH}}{2} \frac{A_s k}{H} - \frac{3}{2} \eta_0 k \sqrt{\frac{g}{H}} \sin \varphi_s - \Delta. \quad (5)$$

Здесь под E_s понимается энергия солитона $E_s = \int_{-\infty}^{\infty} \eta_s dx \sim A_s^{3/2}$, а под

фазой солитона φ_s понимается безразмерное время между максимумом солитона и нулем монохроматической моды, $\Delta = 2\pi(f_{ex} - f_2)$ — это расстройка, а δ — коэффициент, описывающий экспоненциальное затухание солитона. Уравнение (4) — это энергетический баланс. Изменения амплитуды солитона обусловлены передачей энергии от гармонической волны (первое слагаемое в (4)) и диссипацией энергии (второе слагаемое в (4)). Следует подчеркнуть, что коэффициент δ описывает диссипацию, которая не зависит от масштаба возмущений. Такая диссипация действительно наблюдается для поверхностных волн, однако этот параметр, в отличие от всех других, входящих в систему (4), (5), мы не определили при проведении нашего эксперимента и в дальнейшем выберем его из

условия, чтобы теоретически предсказанные параметры солитона были наиболее близки к экспериментальным данным. Уравнение (5) — это кинематическое условие. Скорость солитона в системе отсчета, движущейся со скоростью гармонической волны ($V_{ph} = \frac{\omega}{k}$), зависит от амплитуды солитона A_s (первое слагаемое в (5)).

Гармоническая волна переносит солитон как частицу с соответствующей скоростью, зависящей от фазы. Этот перенос описывается вторым слагаемым в правой части (5). Изменения фазы обусловлены также расстройкой между частотой внешней силы и резонансной частотой (третье слагаемое). Следует подчеркнуть, что система (4), (5) получена для солитона и монохроматической волны, распространяющихся в одном направлении. Для солитона и волны в противоположном направлении будут те же уравнения, а взаимодействие встречных возмущений не принимается во внимание.

Качественный анализ решений системы (4), (5) показал (см., например, [3]), что устойчивые стационарные состояния для фазы φ_s и амплитуды A_s удовлетворяют соотношениям

$$\varphi_s = \arccos\left(\frac{2H\delta}{3\omega\eta_0}\right), \quad (6)$$

$$A_s = \frac{2\Delta}{k} + \frac{3\omega\eta_0}{Hk} \sqrt{1 - \left(\frac{2H\delta}{3\omega\eta_0}\right)^2}. \quad (7)$$

Используя процедуру выделения фазы амплитуды и фазы солитона, описанную в п. 1.2 мы можем сравнить экспериментальные данные с теоретическими предсказаниями. Из представленной теоретической модели следует, что фаза солитона не зависит от расстройки. Поэтому мы представили значения фазы φ_s как функцию гармонической волны η_0 для всех частот возбуждения, при которых в эксперименте наблюдался один солитон (рис. 7). Зависимость амплитуды солитона A_s от амплитуды гармонической моды представлена на рис. 8 для двух различных частот возбуждения. При построении теоретических кривых, изображенных на рис. 7, 8 непрерывными линиями, значение затухания было выбрано равным $\delta = 0,0015$ 1/с. Как видно из представленных графиков, есть достаточно хорошее качественное соответствие между теоретическими и экспериментальными зависимостями. Отметим, что использованная при теоретических расчетах величина коэффициента δ хорошо коррелирует с измерениями добротности гармонической моды с частотой f_2 по ширине резонансной кривой. Непосредственные измерения показали, что

$Q^{-1} = \Delta f_2 / f_2 \approx 0,02$. Если предположить, что затухание моды с частотой f_2 такое же, как у солитона, то в результате пересчета получим близкую величину добротности: $Q^{-1} = \Delta f_2 / f_2 \approx 0,023$.

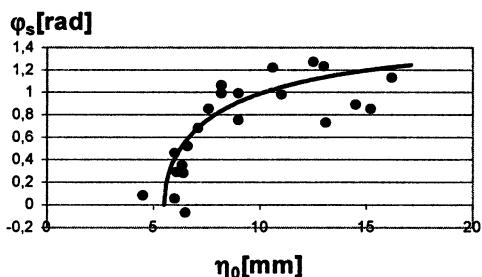


Рис. 7. Зависимость фазы ϕ_s солитона от амплитуды η_0 гармонической волны (точки) и теоретическая зависимость (6) для $\Delta = 0,063$ (кривая).

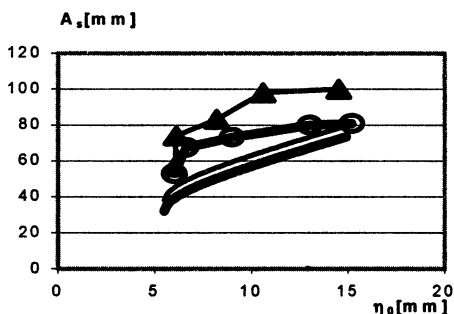


Рис. 8. Зависимость амплитуды солитона A_s от амплитуды гармонической волны η_0 : точки, соединенные ломаными кривыми — экспериментальные данные, гладкие кривые — теоретическая зависимость (7). Толстые линии соответствуют расстройке $\Delta = 0,063$ [рад], тонкие — $\Delta = 0,075$ [рад].

Заключение

В данной лекции показано, что солитоны на поверхности мелкой воды, так же как солитоны в LC-линиях, генерируются на фоне крупномасштабной резонаторной моды. Для поверхностных волн, также как и для электромагнитных волн в [3], обнаружены мультистабильность и удвоение периода возбуждаемых в резонаторе нелинейных колебаний. Построение пространственно-временных диаграмм позволило обнару-

жить связанные состояния солитонов, в которых солитоны распространяются парами и тройками. Проведенное в работе сравнение пространственно-временной динамики нелинейных полей при столкновении двух солитонов в резонаторе и в безграничной системе показало, что взаимодействие с резонаторными модами может существенно влиять на скорости взаимодействующих солитонов.

Выяснено, что для теоретического описания генерации солитонов поверхностных волн могут быть использованы дифференциальные уравнения для амплитуды и фазы солитона, распространяющегося на фоне гармонической волны. Стационарные решения этих уравнений позволяют объяснить характеристики солитонов, которые наблюдались в эксперименте.

Литература

1. *Скотт Э.* Волны в активных и нелинейных средах в приложении к электродинамике. Пер. с англ. / Под ред. Л. А. Островского и М. И. Рабиновича. М.: Сов. радио, 1970. 368 с.
2. Солитоны в действии / Под ред. К. Лонгрена, Э. Скотта, Нью-Йорк, 1977. Пер. с англ. / Под ред. А. В. Гапонова-Грехова, Л. А. Островского. М.: Мир, 1981. 312 с.
3. *Горшков К. А., Островский Л. А., Папко В. В.* Параметрическое усиление и генерация импульсов в нелинейных распределенных системах // Изв. вузов. Радиофизика. 1973. Т. 16, № 8. С. 1195—1204.
4. *Chester W., Borns J. A.* Resonant oscillations of water waves. II. Experiment // Proc. Roy. Soc. A. 1968. Sec. A **306**. P. 23—39.
5. *Potapov A. I., Vesnitsky A. I.* Interaction of solitary waves under head-on collisions. Experimental investigation // Wave motion. 1994. V. 19. P. 29—35.
6. *Езерский А. Б., Полухина О. Е., Броссар Ж., Маран Ф., Мутбази И.* Динамика солитонов, возбуждаемых в резонаторах на поверхности мелкой воды // Изв. вузов. Прикладная нелинейная динамика. 2004. № 2 (принято к печати).
7. *Maxworthy T.* Experiments on collision between solitary waves // J. Fluid Mech. 1976. V. 76. P. 177—185.

ФИЗИКА ВОЛН-УБИЙЦ В ОКЕАНЕ

А. А. Куркин, Е. Н. Пелиновский, А. В. Слюняев

Сейчас под аномально высокими, или «волнами-убийцами», понимаются волны большой амплитуды, неожиданно появляющиеся на морской поверхности как бы из ниоткуда и так же быстро исчезающие. На английском языке для их обозначения обычно используют термины *freak, rogue, giant waves*. Долгое время волны-убийцы являлись предметом морского фольклора и приключенческой литературы. В рассказах бывалых моряков о волнах-убийцах их форма представляется разной: иногда говорят о «стене воды», или о «дыре в море», или о нескольких больших волнах («трех сестрах»). В инженерных приложениях часто под волной-убийцей понимается большой гребень, который сопровождается неглубокой, но протяженной впадиной, находящейся впереди или позади огромной волны. За последние примерно 30—50 лет волны-убийцы стали предметом научных исследований, и их существование сейчас (после получения инструментальных данных) можно считать доказанным. Изучение таких огромных и необычных волн принципиально как для кораблестроения в связи с необходимостью обеспечения живучести кораблей при сильном волнении, так и для морской гидротехники, занимающейся проектированием нефтяных и газовых платформ на морских шельфах, а также портовых сооружений. Одним из принятых критериев, выделяющих волны-убийцы, является отношение их высоты (от подножия до вершины) к так называемой значительной высоте ветрового волнения (среднее значение одной трети больших волн): оно должно превышать двойку ($H_{\max}/H_s > 2$). При этом принимается, что в остальном волна-убийца имеет характерные параметры ветровых волн (т. е. ее длительность примерно 6—15 с), так что волны цунами, достигающие очень больших высот, «отсеиваются» этим определением.

Сейчас накоплен большой архив фотографий аномальных волн, их можно найти в Интернете [1] (см. также обзор [2] и книгу [3]). Такие волны встречались в различных районах Мирового океана. Судя по этим фотографиям и свидетельствам очевидцев волн-убийц, они могут возникать на фоне почти спокойного моря и являть собой одиночные импульсы, либо группы волн. Часто такие волны на фотографиях имеют крутые фронты и могут быть асимметричными, что свидетельствует о нелинейном характере волн-убийц. Подчеркнем также, что, судя по фотографиям, они представляют собой почти двумерные волны со слабой вариацией амплитуды вдоль фронта. По всей очевидности, волна-убийца может появиться как в открытом море, так и вблизи берега, а также и на самом берегу (мощные заплески).

Особый интерес представляют инструментальные записи волн-убийц, позволяющие дать количественный анализ явления. Приведем здесь запись ветрового волнения, сделанную на нефтяной платформе в Северном море 1 января 1995 г., она получила название «новогодней волны» (рис. 1). Характерная длина волн в записи составляет 220 м, глубина места 70 м, и потому условия соответствуют глубокой воде. Значительная высота волн в записи около 11 м, средняя крутизна $ka \sim 0,25$, амплитуда волны-убийцы 18,5 м, а высота 26 м. Такая высота аномальной волны при указанной длине должна, вообще говоря, вести к обрушению волны, аномальная волна является сильнонелинейной. Отношение H_{\max}/H_s здесь более 2,2. Другие записи волн-убийц собраны в книге [3], там же можно найти регистрации аномальных волн из космоса, полученные радиолокатором с синтезированной апертурой.

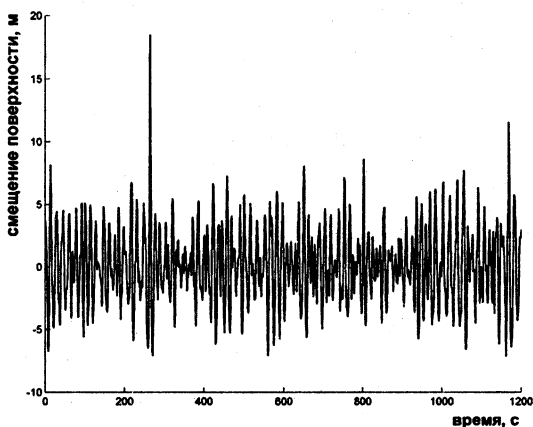


Рис. 1. «Новогодняя волна»

В настоящей статье мы обсудим основные физические механизмы генерации аномально высоких волн.

1. Дисперсионное сжатие волновых пакетов

В линейной теории поле ветровых волн может быть представлено как сумма независимых спектральных компонент, имеющих различные волновые числа (частоты) и направления распространения. В статистическом описании фазы всех спектральных компонент случайны и распределены равномерно в интервале $0—2\pi$, обеспечивая стационарный гауссовый процесс в силу центральной предельной теоремы. Существование редких экстремальных событий можно интерпретировать как локальную суперпозицию большого числа монохроматических волн с под-

ходящими фазами и направлениями, приводящую к формированию пространственно-временных каустик в волновом поле. Для одномерного распространения (двумерное волновое поле) значительное усиление возможно, когда длинные волны, имеющие большую скорость, обгоняют медленно движущиеся более короткие волны.

Дисперсионное сжатие волновых пакетов в линейной теории можно проиллюстрировать, используя известные кинематические уравнения для частоты и амплитуды [4, 5]:

$$\frac{\partial c_{gr}}{\partial t} + c_{gr} \frac{\partial c_{gr}}{\partial x} = 0, \quad \frac{\partial A^2}{\partial t} + \frac{\partial}{\partial x} (c_{gr} A^2) = 0, \quad (1)$$

где c_{gr} — групповая скорость поверхностных волн, легко находимая из дисперсионного соотношения, и A — амплитуда волны. Обычно первое уравнение в (1) приводят для частоты волны, однако после умножения на $dc_{gr}/d\omega$ его можно свести к уравнению только для групповой скорости. Решение этой системы легко находится:

$$c_{gr}(x, t) = c_0(\xi) = c_0(x - c_{gr}t), \quad \frac{dx}{dt} = c_{gr}(x, t) \quad (2)$$

$$A(x, t) = \frac{A_0(\xi)}{\sqrt{1 + t(dc_0/d\xi)}}, \quad (3)$$

где $c_0(x)$ и $A_0(x)$ — начальные распределения характеристик волны. Временная эволюция параметров волны показана на рис. 2. С течением времени наступает «градиентная» катастрофа — «обрушение волны» групповой скорости — приводящая к обращению амплитуды волны в бесконечность. Однако в окрестности каустики кинематические уравнения становятся несправедливыми (нарушается медленность изменения огибающей волновых пакетов), поэтому кинематическая модель, по существу, может быть использована только для выяснения условий дисперсионного сжатия. С физической точки зрения очевидно, что максимальное сжатие возможно, когда все спектральные компоненты сходятся в одну точку L_f в одно время T_f . Этот процесс описывается автомодельным решением уравнения (1):

$$c_{gr} = \frac{x - L_f}{t - T_f}. \quad (4)$$

С учетом нелинейности уравнения (1) становятся связанными, поскольку дисперсионное соотношение включает амплитуду волны [4]. Решение системы тяжело получить в наглядной форме, тем более что для волн на глубокой воде она оказывается эллиптической. Поэтому ниже мы рассмотрим другой подход, основанный на прямом решении волновых уравнений. Отметим здесь, что дисперсионное сжатие неоднократно

моделировалось в бассейне, при этом генерировались огромные волны [6]; условие оптимального сжатия выбиралось из (4), а затем эмпирически модифицировалось для учета нелинейности волны.

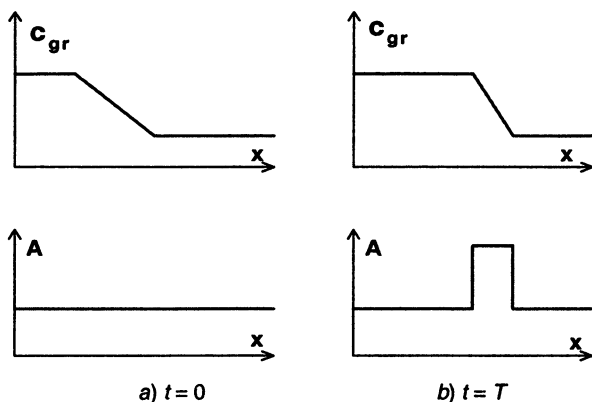


Рис. 2. Деформация характеристик волны в рамках кинематической модели

В рамках волновых уравнений процесс дисперсионного сжатия может быть рассмотрен как обратный к дисперсионному расплыванию в силу инвариантности уравнений к смене знака времени и координат (в пренебрежении диссипацией). Процесс дисперсионного расплывания общеизвестен как в линейной, так и в нелинейной дисперсионной теории. Если мы хотим получить аномальную волну очень большой амплитуды, то, по существу, должны решать волновые уравнения с сингулярными начальными условиями и находить функцию Грина (или ее нелинейный аналог). Поясним эту процедуру на примере уравнения Кортевега — де Вриза

$$\frac{\partial \eta}{\partial t} + c \left(1 + \frac{3\eta}{2h} \right) \frac{\partial \eta}{\partial x} + \frac{ch^2}{6} \frac{\partial^3 \eta}{\partial x^3} = 0, \quad (5)$$

написанного для смещения уровня воды $\eta(x,t)$ в жидкости малой глубины h . Пусть в начальный момент времени задано сингулярное (дельта-образное) начальное условие: $\eta(x_0) = Q\delta(x)$. Если пренебречь нелинейностью, то решение уравнения (5) выражается через функцию Эйри:

$$\eta(x,t) = Q \left(\frac{2}{cth^2} \right)^{1/3} Ai \left[\left(\frac{2}{cth^2} \right)^{1/3} (x - ct) \right] \quad (6)$$

и описывает преобразование импульса в диспергирующий пучок с уменьшающейся амплитудой, причем длинные волны, имеющие большую скорость, бегут впереди коротких. С учетом нелинейности решение сингулярной задачи Коши для (5) также общеизвестно, оно получается ме-

тодом обратной задачи рассеяния [7]. Вместе с осциллирующим пакетом типа функции Эйри (он модифицируется нелинейностью) впереди распространяется солитон с амплитудой $4Q^2/4h^3$ [8]. Очевидно, что если получаемое решение (различное в линейной и нелинейной теориях) обратить в пространстве, то оно за конечное время трансформируется в дельта-функцию, а затем снова расплывется. Этот процесс характеризуется рис. 3, полученным в ходе численного моделирования уравнения Кортевега — де Вриза; цифры у кривых соответствуют моментам времени, а координата есть $x-ct$ [8]. Ясно видно, как солитон и бегущий перед ним частотно-модулированный цуг сливаются в аномальную волну, а затем разбегаются вновь. Нарастание максимальной амплитуды волнового поля происходит очень быстро (рис. 4), объясняя внезапность появления волны-убийцы на море. Аномальная волна может появиться в результате сжатия волновой группы на фоне случайного (не когерентного) ветрового волнения, если нелинейность не слишком велика (рис. 5).

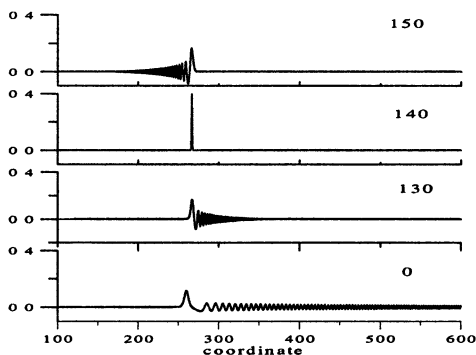


Рис. 3. Дисперсионное сжатие нелинейного волнового пакета на мелкой воде

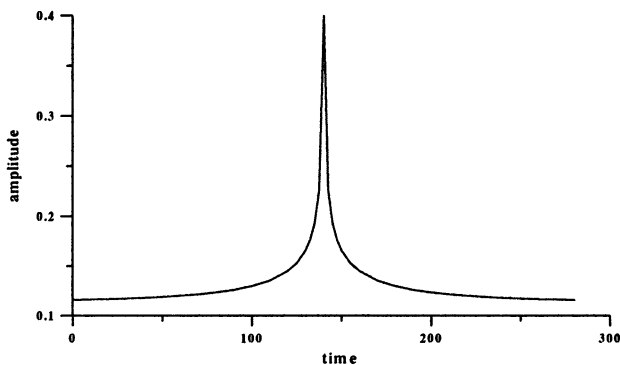


Рис. 4. Изменение максимальной амплитуды волнового поля в процессе формирования аномальной волны.

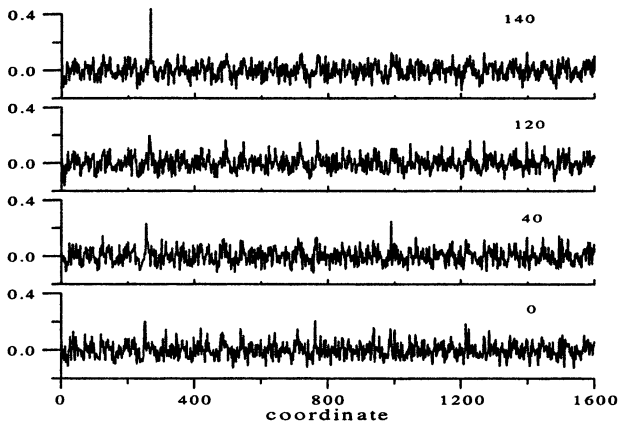


Рис. 5. Формирование аномальной волны в случайном поле на мелкой воде

Важно отметить, что в процессе дисперсионного сжатия солитоны не играют заметной роли. В рамках уравнения Кортевега — де Вриза, как известно, солитоны отталкиваются и не могут при своих столкновениях образовать волну большой амплитуды. Это легко увидеть из прямой задачи эволюции импульса, рождающего большое количество солитонов (для этого большим должен быть параметр AL^2 , где A и L — характерные амплитуда и ширина импульса). В этом случае максимальная амплитуда возникающих солитонов примерно в два раза выше высоты начального возмущения, и при обратном процессе солитоны сливаются в импульс меньшей высоты.

2. Модуляционная неустойчивость волновых пакетов

Как известно, волна Стокса на поверхности глубокой воды неустойчива по отношению к медленным возмущениям ее огибающей. Этот процесс часто демонстрируют в рамках нелинейного уравнения Шредингера

$$i \frac{\partial A}{\partial t} = \frac{\partial^2 A}{\partial x^2} + 2|A|^2 A, \quad (7)$$

где A — пропорционально крутизне волны ($\sqrt{2ka}$), x — безразмерная координата ($2kx - \omega t$) и t — безразмерное время ($\omega t/2$); ω и k — частота и волновое число несущей. Нелинейное уравнение Шредингера (НУШ) для волн на воде было выведено около 40 лет назад. Оно оказалось полностью интегрируемым, что было показано Захаровым и Шабатом [9] (см. также [4]). Здесь в динамике импульсных возмущений важную роль

играют солитоны огибающей. В случае же периодических возмущений в результате модуляционной неустойчивости с течением времени появляются и исчезают волновые группы большой амплитуды (рис. 6), которые можно назвать волнами-убийцами, поскольку они удовлетворяют амплитудному критерию аномальных волн.

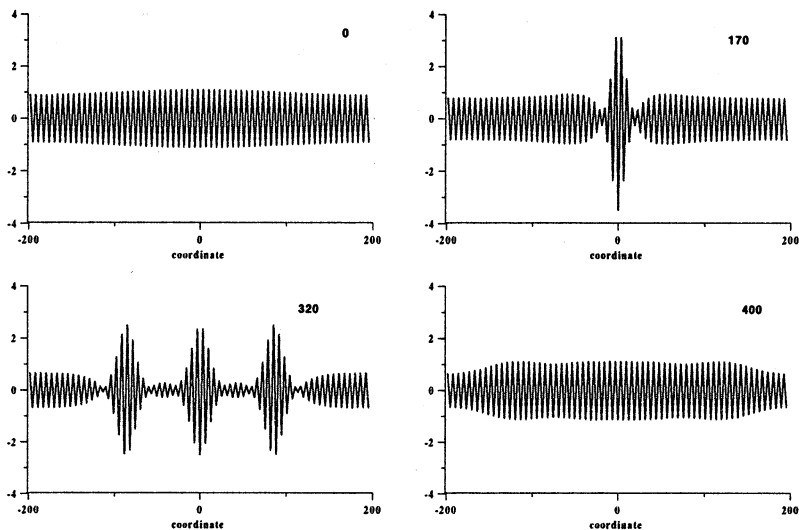


Рис. 6. Эволюция слабomodulированной волны с образованием группы большой амплитуды (числа — время, нормализованное на период несущей волны).

Интерпретация таких «вспыхивающих» групп связана с так называемыми бризерными решениями НУШ [10—13], имеющими вид

$$A(x, t) = e^{-2it} \times \frac{\cos \eta \cos(2\gamma(x - vt) - 2i\psi) - \cosh \psi \cosh(2\Gamma(x - Vt) - 2i\eta)}{\cos \eta \cos(2\gamma(x - vt)) - \cosh \psi \cosh(2\Gamma(x - Vt))},$$

$$\Gamma = -\sinh \psi \cos \eta, \quad \gamma = \cosh \psi \sin \eta, \quad \lambda = \cos(\eta + i\psi),$$

$$V = -\frac{\cosh 2\psi \sin 2\eta}{\sinh \psi \cos \eta}, \quad v = \frac{\sinh 2\psi \cos 2\eta}{\cosh \psi \sin \eta}, \quad (8)$$

где η и ψ — произвольные действительные числа (параметры решения), λ задает собственное значение задачи рассеяния. Бризер огибающей (8) является суперпозицией солитона огибающей и плоской волны [12, 13]. Как следует из (8), максимальное усиление волны (в 3 раза) получается в предельном случае алгебраического бризера (при $\eta = \psi \rightarrow 0$, он постро-

ен на рис. 7). Аномальная группа волн вырастает словно ниоткуда и так же исчезает.

При наличии случайных возмущений ситуация значительно усложняется. Уже в работе [14] было показано, что случайное волновое поле уменьшает границы и инкремент неустойчивости. Недавно было выполнено численное моделирование НУШ со случайными начальными возмущениями гауссовой формы [15]. Со временем спектр расширялся симметрично, достигая квазиравновесной ширины, соответствующей границе модуляционной неустойчивости. Таким образом, в среднем модуляционная неустойчивость является фактором релаксации спектра к устойчивому состоянию.

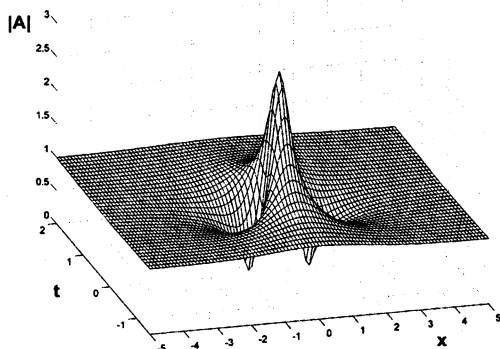


Рис. 7. Алгебраический бризер как модель волны-убийцы на фоне периодической волны

Главный вывод из сделанного анализа: волны-убийцы могут рождаться как результат модуляционной неустойчивости, так что нелинейность дает другой механизм возникновения волн-убийц дополнительно к дисперсионному сжатию волновых частотно-модулированных пакетов.

Реальное ветровое волнение, будучи однородным в среднем, может содержать волновые пакеты как с амплитудной, так и с частотной модуляцией, а также участки немодулированной волны. Следовательно, в реальных условиях аномальные волны могут возникнуть как в результате модуляционной неустойчивости, так и дисперсионного сжатия [16]. В работах [17, 18] выполнены численные эксперименты в рамках НУШ с реализациями поля ветровых волн, подчиняющимся спектру JONSWAP. В частности, было показано, что если спектр очень узкий, то вероятность появления волн-убийц возрастает по сравнению с рэлеевским значением, типичным для линейных волновых полей.

Учет трехмерности волнового поля приводит к дополнительному механизму геометрического фокусирования волн. Этот механизм обще-

известен в линейной теории и является предметом специальных исследований в нелинейной оптике. Мы же приведем здесь только один результат формирования волны-убийцы в случайном поле ветровых волн за счет геометрической и дисперсионной фокусировки трехмерных волн (рис. 8); этот процесс исследован в рамках уравнения Дэви — Стюартсона, обобщающего НУШ для «трехмерного» моря конечной глубины [13].

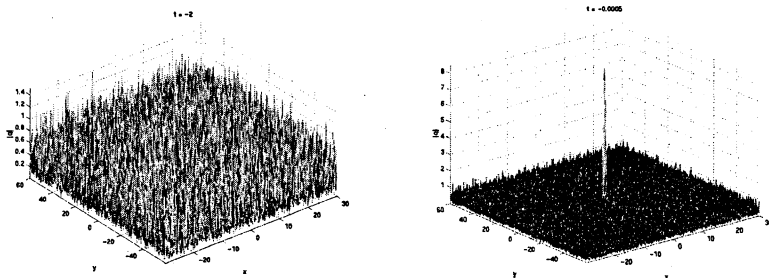


Рис. 8. Формирование волны-убийцы в случайном трехмерном поле

3. Взаимодействие волн с течением

Сильные течения в океане могут служить причиной блокировки волн, когда их групповая скорость сравнивается со скоростью течений (и противоположна по знаку). Этот процесс хорошо известен в линейной теории (см., например [19]). Работа [20] была первой, в которой рассматривалась нелинейная модель появления гигантских волн применительно к наблюдению волн-убийц на течении у мыса Игольного вблизи побережья Южной Африки. В ней, в частности, выведено модифицированное НУШ с добавочным членом $|dU/dx|xA$ в окрестности точки блокировки:

$$i \frac{\partial A}{\partial t} = \frac{\omega}{8k^2} \frac{\partial^2 A}{\partial x^2} + \frac{\omega k^2}{2} |A|^2 A - k \left| \frac{dU}{dx} \right| xA, \quad (9)$$

где dU/dx вычислено в точке блокировки (на глубокой воде)

$$\frac{1}{2} \sqrt{\frac{g}{k}} + U(x_0) = 0, \quad (10)$$

и уравнение (9) записано в размерных переменных. Последний член в (9) может быть легко получен из линейного дисперсионного соотношения, принимая во внимание доплеровский сдвиг частоты волны на течении. Если нелинейность незначительна, то уравнение (9) совпадает с уравнением Эйри (для монохроматических процессов) и описывает волновое поле на каустике. В нелинейном случае решение уравнения (9) выражается в терминах специальных функций через трансценденты

Пенлеве второго рода. Качественно оно похоже на функцию Эйри с переходом от затухающей синусоиды к экспоненциальному затуханию, однако его максимум смещен от линейной точки блокировки (см. детали в оригинальной работе [20]). Таким образом, нелинейность качественно не влияет на сценарий блокирования монохроматической волны встречным течением.

Ситуация, однако, становится более сложной, если волна не является монохроматической. В частности, уравнение (9) имеет точное солитонное решение [21]:

$$A(x, t) = B \operatorname{sech} \left[\sqrt{2k} B \left(x - B\omega t / \sqrt{8} + \omega \left| \frac{dU}{dx} \right| t^2 / 8k \right) \right] \times \\ \times \exp \left[i \left(\sqrt{2k^2} Bx + \omega k B t^2 \left| \frac{dU}{dx} \right| / \sqrt{32} - \omega t^3 \left| \frac{dU}{dx} \right|^2 / 24 \right) \right]. \quad (11)$$

Этот солитон движется с переменной скоростью, подходит к точке блокировки, проникает за нее и затем отражается. Неожиданностью здесь является постоянство амплитуды солитона огибающей в неоднородной среде в силу баланса дисперсионного сжатия (растяжения) и ослабления (усиления) на неоднородном течении. Более того, следующее преобразование [21, 22]:

$$A(x, t) = A'(x', t) \exp \left[i \left(kxt \left| \frac{dU}{dx} \right| + \frac{\omega t^2}{32} \left| \frac{dU}{dx} \right|^2 \right) \right], \quad x' = x + \frac{\omega t^2}{8k} \left| \frac{dU}{dx} \right|, \quad (12)$$

приводит (9) к НУШ (7) с постоянными коэффициентами для функции $A'(x', t)$, и следовательно, существование волн-убийц на неоднородных течениях становится очевидным.

4. Аномальные краевые волны

Подходя к суше, ветровые волны возбуждают краевые волны, распространяющиеся вдоль берега. Считается, что они ответственны за формирование искривленной береговой линии, так называемых параболических бухт [23]. Существенно, что краевые волны, будучи трехмерными по своей природе, могут быть описаны одномерными эволюционными уравнениями (вдоль берега). В частности, волны Стокса

$$\eta(x, y, t) = A \exp(-ky) L_n(2ky) \exp(i(\omega t - kx)), \quad (13)$$

(L_n — полином Лагерра, описывающий структуру волны в поперечном к берегу направлении y , n — номер моды) даже на мелкой воде обладают сильной дисперсией:

$$\omega = \sqrt{(2n+1)\alpha g k} \quad (14)$$

(α — угол откоса). Можно показать также, что эти волны являются модуляционно неустойчивыми [24]. Поэтому для них оба механизма — дисперсионное сжатие и модуляционная неустойчивость — будут работать и могут приводить к эффективной генерации аномальных волн [24—26]. Пример расчета аномальной локализованной краевой волны демонстрируется на рис. 9; максимум волнового поля приходится на прибрежную зону, так что такая волна может вызывать аномальное и кратковременное затопление берега.

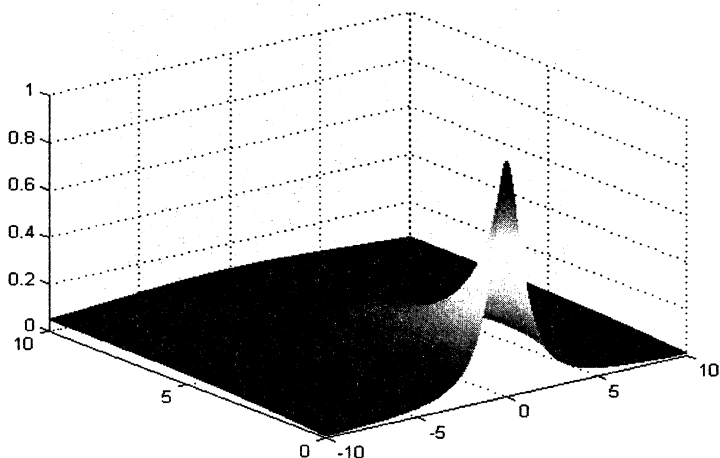


Рис. 9. Форма аномального импульса, полученного в результате дисперсионного сжатия краевых волн

5. Описание «новогодней волны»

Мы уже приводили запись аномальной «новогодней» волны в Северном море (см. рис. 1). В этом случае НУШ может быть выбрано в качестве первого приближения для анализа (оно не является полностью адекватной моделью, так как аномальная высота сильно локализована и нелинейна). Нужно также отметить, что для применения НУШ требуется определение частоты несущей, что может быть сделано неоднозначно.

На рис. 10 представлен текущий спектр «новогодней волны», построенный для достаточно длинной выборки (83 с, при периоде волны 12 с). Вертикальной белой линией отмечена определенная несущая частота, а горизонтальной — момент регистрации аномальной волны. Спектр демонстрирует существование нескольких пиков для фиксированного времени, что указывает на сильную модуляцию и возможную роль модуляционной неустойчивости в формировании аномальной вол-

ны. Видно, что в момент регистрации anomalно высокой волны имеет место раздвоение спектра. Оценки области модуляционной неустойчивости (неустойчивыми являются частоты $\omega - \omega_{BF} < \omega + \omega_{BF}$, где ω — несущая частота, и $\omega_{BF} \approx 0,074$ рад/с, см. подробности в [12]) довольно хорошо соответствуют видимой спектральной картине.

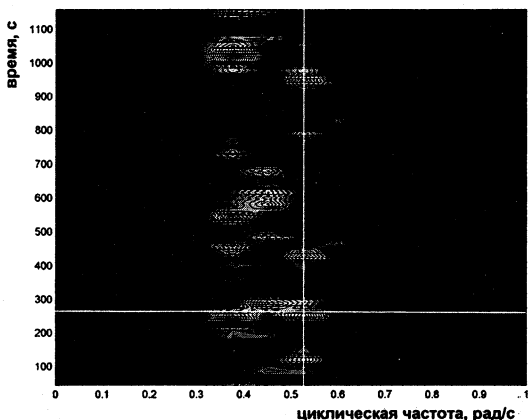


Рис. 10. Текущий спектр «черноморской волны»

Была выполнена оценка солитонных свойств наблюдаемой записи ветровой волны. Для этого полная запись разрезалась на участки и каждый кусок записи использовался в качестве потенциала для нахождения дискретного спектра ассоциированной задачи в методе обратной задачи рассеяния. На рис. 11 построены амплитуды наиболее высоких солитонных пакетов (т. е. если в выборке обнаруживалось несколько солитонов, то брался максимально высокий) и видно, что в момент регистрации anomalно высокой волны определяется солитон огибающей с амплитудой около 10,7 м (закрашенные кружки). В суперпозиции с фоновым квазилинейным волнением такой солитон может порождать высокие волны и вполне может объяснить «новогоднюю волну». Это другая интерпретация модуляционной неустойчивости групп волн: когда группа представляет собой суперпозицию квазилинейной плоской волны и интенсивного солитона огибающей, имеют место сильные биения в соответствии с формулой (8).

Проведенный анализ «новогодней волны» сделан в предположении справедливости модели НУШ. В работе [27] было выполнено численное моделирование пространственной эволюции «новогодней волны» в рамках модели Диста, более корректной в приложении к достаточно широким спектрам поверхностных волн. Модель Диста учитывает в эволюционном уравнении (7) слагаемые нелинейной дисперсии следующего

порядка и имеет дисперсионный закон линейных волн как в случае бесконечно глубокой воды; но она остается слабонелинейной. Нами было повторено это численное моделирование и сопоставлено с результатами моделирования в рамках модели НУШ. «Новогодняя волна» сперва была пересчитана на 500 м против движения волн (после преобразования поля, эквивалентного обращению времени), а затем по ходу волн до дистанции 500 м после точки записи.

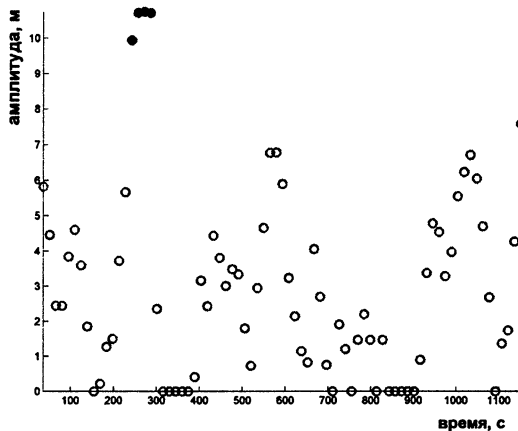


Рис. 11. Амплитуды солитонов огибающей, определенные для выборок записи «новогодней волны». Закрашены значения, соответствующие интенсивному солитону, образующему пиковую волну.

На рис. 12 построены максимумы и минимумы волнового поля, определенные в различных пространственных точках (дистанция 0 м соответствует месту инструментальной регистрации). Моделью Диста описывается область спектра вблизи несущей частоты, а для моделирования в рамках НУШ был использован весь спектральный диапазон записи, хотя это уравнение реально применимо лишь к узкой области вблизи несущей. В результате, хотя спектр вблизи несущей частоты в модели Диста точно описан, высокочастотная часть спектра отброшена, что повлекло уменьшение амплитуды аномальной волны, видимое на рис. 12. Из рисунка следует, что вблизи точки регистрации описание в рамках обеих моделей дает схожие расстояния проявления волны-убийцы: около 100 м против и по движению волн. На больших удалениях кривые сильно разнятся: в рамках интегрируемого НУШ солитон не может исчезнуть и он проявляет себя, время от времени генерируя интенсивные волны, сопоставимые с волной-убийцей. В рамках модели Диста солитон не сохраняется и, как следует из рис. 12, имеет меньшую энергию вдали от точки записи «новогодней волны».

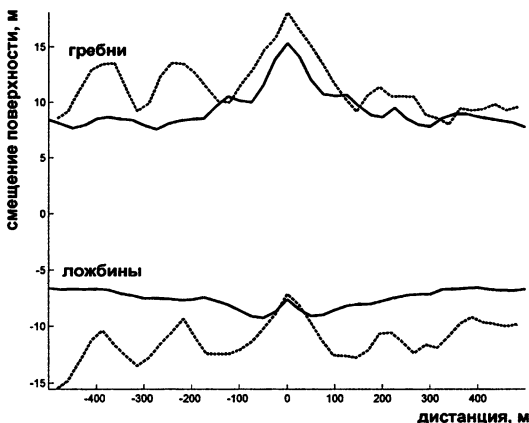


Рис. 12. Максимальные высоты гребней и глубины ложбин, определенные в разных пространственных точках регистрации по результатам моделирования всей 20-минутной записи «новогодней волны» в рамках уравнения НУШ (пунктирная линия) и уравнений Диста (сплошная линия).

Рассмотренные модели описывают однонаправленное движение ветровых волн. Очевидно, что учет поперечной волновой динамики может существенно изменить динамику аномальных волн (см. [13]). Упомянутой ранее фокусировке поверхностных волн в результате геометрического схождения посвящен раздел обзора [2]. Для таких исследований регистрации волн в одной точке недостаточно.

Работа выполнена при поддержке грантов ИНТАС 01-2156 и 03-51-4286, РФФИ № 02-05-65107 и 03-05-64975 и Научной школы академика В. И. Таланова № НШ-1637.2003.2.

Литература

1. Olagnon M. Vagues extrêmes – Vagues scélérates. <http://www.ifremer.fr/webcom/molagnon/jpo2000>.
2. Kharif C., Pelinovsky E. Physical mechanisms of the rogue wave phenomenon // Eur. J. Mech. B–Fluid. 2003. V. 22. P. 603–634.
3. Куркин А. А., Пелиновский Е. Н. Волны-убийцы: факты, теория и моделирование. Н. Новгород: ННГУ, 2004.
4. Уизем Дж. Линейные и нелинейные волны. М.: Мир, 1977. 622 с.
5. Островский Л. А., Потапов А. И. Введение в теорию модулированных волн. М.: Наука, 2003. 400 с.
6. Clauss G. Dramas of the sea: episodic waves and their impact on offshore structures // Appl. Ocean Res. 2002. V. 24. P. 147–161.
7. Лэмб Дж. Л. Введение в теорию солитонов. М.: Мир, 1983. 294 с.
8. Pelinovsky E., Talipova T., Kharif C. Nonlinear dispersive mechanism of the freak wave formation in shallow water // Physica D. 2000. V. 147, № 1–2. P. 83–94.
9. Захаров В. Е., Шабат А. Б. Точная теория двумерной самофокусировки и одномерной самомодуляции волн в нелинейных средах // ЖЭТФ. 1971. Т. 61. С. 118–134.

10. *Dysthe K. B., Trulsen K.* Note on breather type solutions of the NLS as a model for freak-waves // *Physica Scripta*. 1999. V. T82. P. 48—52.
11. *Osborne A. R., Onorato M., Serio M.* The nonlinear dynamics of rogue waves and holes in deep-water gravity wave train // *Phys. Lett. A*. 2000. V. 275. P. 386—393.
12. *Пелиновский Е. Н., Слюняев А. В., Талипова Т. Г., Хари́ф К.* Нелинейное параболическое уравнение и экстремальные волны на морской поверхности // *Изв. вузов. Радиофизика*. 2003. Т. 46, № 7. С. 499—512.
13. *Slunyaev A., Kharif C., Pelinovsky E., Talipova T.* Nonlinear wave focusing on water of finite depth // *Physica D*. 2002. V. 173, № 1—2. P. 77—96.
14. *Alber I. E.* The effects of randomness of the stability of two-dimensional surface wavetrains // *Proc. Roy. Soc. London*, 1978. V. A363. P. 525—546.
15. *Dysthe K. B., Trulsen K., Krogstad H. E., Socquet-Juglard H.* Evolution of a narrow-band spectrum of random surface gravity waves // *J. Fluid Mech.* 2003. V. 478. P. 1—10.
16. *Kharif C., Pelinovsky E., Talipova T., Slunyaev A.* Focusing of nonlinear wave groups in deep water // *Письма в ЖЭТФ*. 2001. Т. 73, № 4. С. 170—175.
17. *Onorato M., Osborne A., Serio M., Bertone S.* Freak wave in random oceanic sea states // *Phys. Rev. Lett.* 2001. V. 86, № 25. P. 5831—5834.
18. *Onorato M., Osborne A.R., Serio M.* Extreme wave events in directional, random oceanic sea states // *Phys. Fluids*, 2002. V. 14, № 4. P. L25—L28.
19. *Лавренов И. В.* Математическое моделирование ветровых волн в пространственно-неоднородном океане. СПб: Гидрометеоиздат, 1998. 500 с.
20. *Smith R.* Giant waves // *J. Fluid Mech.* 1976. V. 77. P. 417—431.
21. *Chen H. H., Liu C. S.* Solitons in nonuniform media // *Phys. Rev. Lett.* 1976. V. 37. P. 693—697.
22. *Peregrine D. H., Smith R.* Nonlinear effects upon waves near caustics // *Phil. Trans. Royal Soc. London*, 1979. V. A292. P. 341—370.
23. *Masselink G.* Alongshore variation in beach cusp morphology in a coastal embayment // *Earth Surf. Process. Landforms*. 1999. V. 24. P. 335—347.
24. *Дубинина В. А., Куркин А. А., Полухина О. Е.* Фокусировка краевых волн на шельфе моря // *Изв. РАН. ФАО*. 2003. Т. 39, № 6. С. 838—848.
25. *Kurkin A., Pelinovsky E.* Focusing of edge waves above sloping beach // *Eur. J. Mech. B-Fluid*. 2002. V. 21, № 5. P. 561—577.
26. *Дубинина В. А., Куркин А. А., Пелиновский Е. Н., Полухина О. Е.* Слабонелинейные периодические краевые волны Стокса: Препринт ИПФ РАН № 644. Н. Новгород, 2003.
27. *Trulsen K.* Simulating the spatial evolution of a measured time series of a freak wave // *Proc. Workshop «Rogue Waves 2000»*. 2000. P. 265.

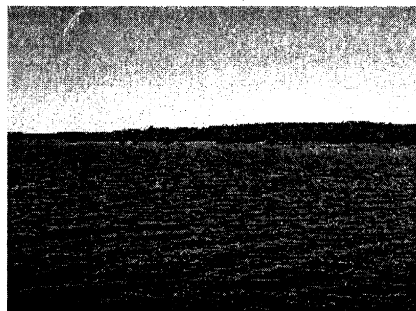
НЕЛИНЕЙНОЕ РЕЗОНАНСНОЕ ВЗАИМОДЕЙСТВИЕ ВОЛН С ПОТОКАМИ В ОКЕАНЕ

Ю. И. Троицкая

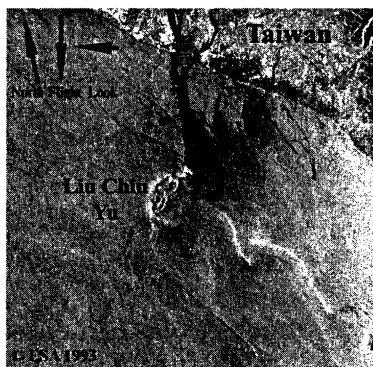
1. Введение.

Плазменно-гидродинамическая аналогия

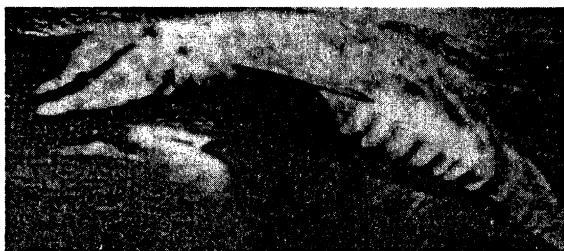
Рассмотрим ряд явлений, типичных для океана: ветровые волны (рис. 1, *а*), возникающие при обдуве поверхности воды турбулентным воздушным потоком; волнообразные структуры на коже плывущего дельфина (рис. 1, *б*); турбулентный след за препятствием в набегающем потоке (рис. 1, *в*). Эти и многие другие явления, различающиеся по пространственным и временным масштабам, по энергии, по физической природе и т. п., имеют сходный механизм возникновения — резонансное взаимодействие волн с потоками.



а



б



в

Рис. 1. Примеры резонансного взаимодействия волн с потоками в геофизической гидродинамике: ветровые волны (*а*); волнообразный след за островом (*б*); структуры на коже плывущего дельфина (*в*).

Чтобы была понятна природа такого резонансного взаимодействия, рассмотрим плазменно-гидродинамическую аналогию. Резонансное взаимодействие хорошо изучено в физике плазмы. Воспользуемся определением этого взаимодействия, данным в работе [1]. Хорошо известно, что из-за дальнего действующего характера кулоновских сил в бесстолкновительной плазме существуют коллективные степени свободы (например, волны пространственного заряда) и сохраняются индивидуальные (u отдельных частиц). Резонансное взаимодействие — это взаимодействие между индивидуальными и коллективными степенями свободы. Эффективный энергообмен между этими степенями свободы возможен, если фаза волны постоянна на траектории частицы (это так называемый фазовый резонанс). Такое взаимодействие и называется резонансным. Наличие резонансного взаимодействия между ленгмюровской волной (коллективная степень свободы) и частицами (индивидуальная степень свободы) приводит к «бесстолкновительному» затуханию Ландау [2].

Впервые о плазменно-гидродинамической аналогии стали говорить в связи с интерпретацией линейных задач теории гидродинамической устойчивости [1, 3—5]. При этом основная идея состояла в следующем: роль частиц (индивидуальная степень свободы) играют струйки жидкости в плоскопараллельном течении с профилем скорости $U_0(y)$. При этом распределение по скоростям связано с тем, что U_0 меняется в зависимости от y . Распределение частиц по скоростям осуществляется за счет «развертки» в пространстве. При этом резонансные «частицы», у которых скорость равна фазовой скорости волны, располагаются вблизи критического уровня, где $U(y_c) = c$, c — фазовая скорость волны. Дальше, путем подсчета «частиц», попадающих в интервал $[U, U + \Delta U]$, в [5] было показано, что аналогом функции распределения является величина $f_0(U) = |1/U_{0,y}|_{U_0(y)=U}$. Такой вариант плазменно-гидродинамической аналогии, однако, эффективен только в линейных задачах.

Другой вариант плазменно-гидродинамической аналогии был предложен в [6]. Такая аналогия основана на принципиальном сходстве уравнений для завихренности двумерного течения невязкой жидкости и бесстолкновительного кинетического уравнения для функции распределения частиц плазмы. Известно, что в двумерной гидродинамике идеальной жидкости завихренность сохраняется на траектории жидкой частицы, а бесстолкновительное кинетическое уравнение выражает сохранение функции распределения на траектории частиц в фазовом пространстве. Из сравнения уравнений (строка 1 в табл. 1) очевидно следующее соответствие переменных: $\Omega \leftrightarrow f$, $w \leftrightarrow eE/m$; $y \leftrightarrow v$; $x \leftrightarrow x$. Специфика среды входит лишь в материальные уравнения (строка 2 в табл. 1), которые, конечно, различны в плазме и гидродинамике, но и в том, и в другом «коллективное» поле (w или E) выражается через

величину, которая сохраняется на фазовой траектории (ω или f). Это не просто формальная аналогия. В плазме заряженные частицы с функцией распределения f создают «коллективное» электрическое поле E (см. материальное уравнение), которое управляет их движением. В гидродинамике вихри создают поле скорости (см. систему материальных уравнений), и частицы, на которых «сидят» эти вихри, в свою очередь, переносятся этими полями скорости.

Таблица 1

Плазменно-гидродинамическая аналогия

<p>1. Уравнение для завихренности в идеальной жидкости</p> $\frac{\partial \omega}{\partial t} + u \frac{\partial \omega}{\partial x} + w \frac{\partial \omega}{\partial y} = 0$	<p>1. Бесстолкновительное кинетическое уравнение</p> $\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \frac{e}{m} E \frac{\partial f}{\partial v} = 0$
<p>2. «Материальные» уравнения</p> $\omega = \frac{\partial u}{\partial y} - \frac{\partial w}{\partial x}; \quad \frac{\partial u}{\partial x} + \frac{\partial w}{\partial y} = 0$	<p>2. «Материальные» уравнения</p> $\frac{\partial E}{\partial x} = 4\pi e \int f dv$
<p>3. Критический уровень</p> $U(y_c) = \frac{\omega}{k}$	<p>3. Резонансные частицы</p> $v = \frac{\omega}{k}$
<p>4. Инкремент</p> $\lim_{\omega \rightarrow \frac{\omega}{k}} \omega \sim \frac{d\omega_0}{dy} \Big _{U_0 = \frac{\omega}{k}} = \frac{d^2 U_0}{dy^2} \Big _{U_0 = \frac{\omega}{k}}$	<p>4. Инкремент</p> $\text{Im } \omega \sim \frac{df_0}{dv} \Big _{v = \frac{\omega}{k}}$
<p>5. Радиационное напряжение</p> $\int \langle w_1 \Omega_1 \rangle dy = \int \frac{\partial}{\partial y} \langle w_1 u_1 \rangle dy$	<p>5. Работа «коллективного» поля над частицами плазмы</p> $\left\langle e E_1 \int_{-\infty}^{\infty} f_1 dv \right\rangle = e \int_{-\infty}^{\infty} \langle E_1 f_1 \rangle dv$
<p>6. Линии равной завихренности вблизи критического слоя</p> $\frac{dx}{dt} = U_y y$ $\frac{dy}{dt} = w_1 \cos(\omega t - kx)$	<p>6. Уравнения движения частиц вблизи области захвата</p> $\frac{dx}{dt} = v$ $\frac{dv}{dt} = \frac{e}{m} E_1 \cos(\omega t - kx)$
<p>7. Ширина нелинейного критического слоя $\delta_{kc} = \left(\frac{ w_1 }{kU_y} \right)^{1/2}$</p>	<p>7. Ширина области захвата</p> $\delta_v = \left(\frac{e E_1 }{m k} \right)^{1/2}$

Заметим, что в гидродинамике уравнения сложнее по форме, поскольку u и w зависят от координат x и y . А в плазме поле E принципиально зависит только от x (и не зависит от y). В рамках этого варианта плазменно-гидродинамической аналогии зависимость инкремента волн от второй производной профиля средней скорости очевидна. Поскольку ω_0 аналогично f_0 , то $d\omega_0/dy|_{U_0=\frac{\omega}{k}} = d^2U_0/dy^2|_{U_0=\frac{\omega}{k}}$ соответствует $df_0/dv|_{v=\frac{\omega}{k}}$.

Аналогом фазировки частиц плазмы, зависящей от производной функции распределения, в гидродинамике будет формирование возмущений завихренности в критическом слое (КС), знак которых, очевидно, зависит от знака производной завихренности или знака второй производной от профиля скорости.

В рамках «новой плазменно-гидродинамической аналогии» [6] существует и аналог нелинейного взаимодействия монохроматической волны с частицами плазмы. При этом в окрестности резонансного уровня линии равной завихренности имеют характерный вид, аналогичный фазовой плоскости захваченных частиц плазмы. Это ясно из характеристических уравнений (строка 6 табл. 1). При этом аналогом области захвата является ширина нелинейного КС (строка 7 табл. 1). Если ширина нелинейного КС мала по сравнению с характерным масштабом течения, то волновые возмущения поперечной скорости w_1 почти постоянны на масштабе КС, и уравнение для возмущений завихренности и функции распределения удовлетворяют одним и тем же уравнениям с точностью до обозначений. Такой случай был рассмотрен в [6] для волн на воле и в [7] для волн в пограничном слое.

2. Механизм «отрицательной вязкости»

В гидродинамике существует еще один механизм неустойчивости сдвигового потока, который в конечном счете сводится к резонансному энергообмену волны и потока. Это механизм «отрицательной вязкости». Поясним его на частном примере. Рассмотрим задачу о генерации волн на воде ветром. При этом для профиля скорости ветра примем простую модель: в узком слое вблизи поверхности коэффициент вязкости $\nu(y)$ много меньше, чем во всей остальной жидкости (рис. 2, а). Тогда градиент скорости велик вблизи поверхности, а выше существенно меньше. Для простоты можно считать, что скорость в верхнем слое постоянна и равна U_0 . Это модель реального логарифмического пограничного слоя с вязким подслоем, в котором существует очень резкий градиент скорости. Пусть на поверхности воды распространяется волна, фазовая скорость которой $c < U_0$, т. е. критический уровень лежит в вязком подслое.

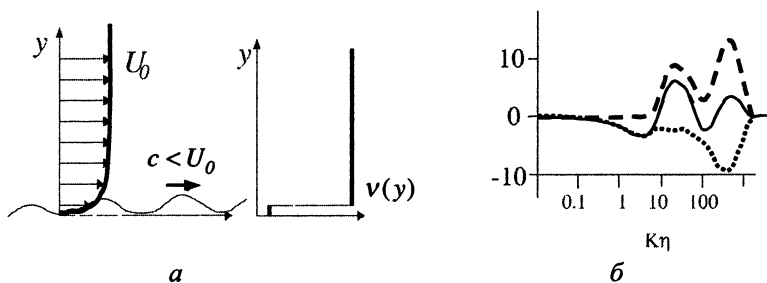


Рис. 2. Модельные профили скорости ветра и коэффициента турбулентной вязкости над взволнованной водной поверхностью (а) и характеристики энергообмена волны с потоком (б): - - - работа радиационных сил; работа вязких сил; — профиль плотности потока энергии.

Рассмотрим, как будет происходить энергообмен волны на воде с таким ветровым потоком. Для этого получим уравнение, которому удовлетворяет энергия волн на воде. Это можно сделать строго. Надо записать систему уравнений гидродинамики вязкой несжимаемой жидкости в воде и воздухе:

$$\begin{aligned} \frac{\partial u_1}{\partial t} + U(y) \frac{\partial u_1}{\partial x} + w_1 \frac{dU}{dy} + \frac{1}{\rho} \frac{\partial p_1}{\partial x} &= \frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y}, \\ \frac{\partial w_1}{\partial t} + U(y) \frac{\partial w_1}{\partial x} + \frac{1}{\rho} \frac{\partial p_1}{\partial y} &= \frac{\partial \sigma_{12}}{\partial x} + \frac{\partial \sigma_{22}}{\partial y}, \\ \frac{\partial u_1}{\partial x} + \frac{\partial w_1}{\partial y} &= 0. \end{aligned} \quad (1)$$

Здесь $\sigma_{ij} = v(y) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$ — выражение для вязких напряжений в рамках принятой модели, u_1 и w_1 — компоненты скорости. Системы уравнений в воде и воздухе выглядят одинаково, но в воде $U = 0$.

Из системы (1) можно получить уравнение для плотности энергии волновых возмущений (кинетической + потенциальной). Проинтегрировав систему (1) по y от $-\infty$ до η_1 (где η_1 — возвышение поверхности) и осреднив по периоду волны, можно получить, что с точностью до величин порядка квадрата амплитуды волны

$$\frac{\partial E_w}{\partial t} = \langle p_1 w_1 \rangle \Big|_{y=0} + \rho_w \int_{-\infty}^0 \text{Vis}_w(y) dy. \quad (2)$$

Здесь $E_w = \rho_w \left(\int_{-\infty}^0 \frac{\langle u_1^2 + w_1^2 \rangle}{2} dy + \frac{g \langle \eta_1^2 \rangle}{2} \right)$ — энергия волн на воде [8],

$\langle p_1 w_1 \rangle|_{y=0}$ — поток энергии в воздухе на поверхности воды, так назы-

ваемый поток энергии от ветра к волнам, $\text{Vis}_w(y) = - \left\langle v_w \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2 \right\rangle$

— работа вязких напряжений в воде.

Рассмотрим теперь возмущения воздушного потока. Умножив первое уравнение системы (1) на u_1 , второе — на w_1 и усреднив по x , получим

$$\begin{aligned} \frac{\partial}{\partial t} \rho_a \left(\frac{\langle u_1^2 \rangle + \langle w_1^2 \rangle}{2} \right) + \frac{\partial}{\partial y} \langle p_1 w_1 \rangle + \rho_a \left(\langle u_1 w_1 \rangle \frac{dU}{dy} \right) = \\ = - \rho_a \left\langle v_a \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2 \right\rangle. \end{aligned} \quad (3)$$

Следует заметить, что энергия волновых возмущений в системе вода — воздух определяется главным образом движениями в воде из-за малости плотности воздуха. Первое слагаемое в левой части уравнения (3) представляет собой скорость изменения плотности кинетической энергии возмущений воздушного потока — малую величину, которой мы в дальнейшем будем пренебрегать. Интегрируя это уравнение от некоторого $y = \delta$ до ∞ , получим

$$\langle p_1 w_1 \rangle|_{y=\delta} = \rho_a \int_{\delta}^{\infty} \left[\left((U(y) - U(\delta)) \frac{\partial}{\partial y} \langle u_1 w_1 \rangle - \text{Vis}(y) \right) \right] dy, \quad (4)$$

где 1-е слагаемое под знаком интеграла по смыслу равно работе радиационных сил, а 2-е представляет собой работу вязких напряжений в воздушном потоке. Смысл выражения (4) очень простой: поток энергии на произвольном уровне $y = \delta$ обусловлен работой радиационных и вязких сил при $y > \delta$, это естественно, если поток энергии направлен вниз.

При $\delta = 0$ выражение (4) определяет поток энергии на поверхности воды, который входит в уравнение волновой энергии (2). Принимая во внимание граничное условие $U(0) = 0$, получим

$$\frac{\partial E_w}{\partial t} = \rho_a \int_0^{\infty} \left[\left(U(y) \frac{\partial}{\partial y} \langle u_1 w_1 \rangle + \text{Vis}(y) \right) \right] dy + \rho_w \int_{-\infty}^0 \text{Vis}_w(y) dy. \quad (5)$$

Величины $\text{Vis}(y)$ и $\text{Vis}_*(y)$ всегда отрицательны, т. е. вязкость приводит к потерям волновой энергии. Нарастание волн возможно, только если положительна работа радиационных напряжений

$$T = U(y) \partial \langle u_1 w_1 \rangle / \partial y.$$

Покажем теперь, что только при наличии резонансного взаимодействия возможно нарастание волн. Рассмотрим уравнение (4). Если пограничный слой тонкий по сравнению с длиной волны на воде, то он смещается как целое. Но тогда

$$w_1(\delta) = \frac{\partial \eta_1}{\partial t} + U(\delta) \frac{\partial \eta_1}{\partial x} = ik \eta_1 (U(\delta) - c).$$

Кроме того, в пограничном слое давление меняется слабо: $p_1(\delta) \approx p_1(0)$. Значит, поток энергии на уровне $y = \delta$ можно выразить через поток энергии на поверхности:

$$\langle p_1 w_1 \rangle_{y=\delta} = \langle p_1 w_1 \rangle_{y=0} \frac{c}{c - U(\delta)}.$$

Отсюда следует, что поток энергии на поверхности, который входит в уравнение для инкремента волн на воде, имеет вид

$$\langle p_1 w_1 \rangle_{y=0} = \rho_a \int_{\delta}^{\infty} \left[\left((U(y) - U(\delta)) \frac{\partial}{\partial y} \langle u_1 w_1 \rangle - \text{Vis}(y) \right) \right] dy \frac{c}{c - U(\delta)}. \quad (6)$$

Если мы выберем $\delta = 0$, то получим выражение (5), если же возьмем δ больше толщины пограничного слоя, то $U(y) = U(\delta) = U_0$, тогда в подынтегральное выражение входит только работа вязких сил и $\langle p_1 w_1 \rangle_{y=0}$ оказывается положительной величиной при $c < U_0$. Это эквивалентно замене коэффициента вязкости ν , входящего в $\text{Vis}(y)$, на $\nu c / (c - U(\delta))$. В связи с этим такой механизм неустойчивости называется механизмом «отрицательной» вязкости. Его можно интерпретировать на языке волн отрицательной энергии. Если мы перейдем в систему отсчета, которая движется со скоростью U_0 , то работа сил вязкости будет совершаться в покоящейся среде, но энергия волн на воде будет отрицательная, такие волны нарастают за счет потерь энергии.

Из (6) очевидно, что $\langle p_1 w_1 \rangle > 0$ только при $c < U_0$, но при этом в потоке обязательно существуют частицы, находящиеся в резонансе с волной. Такое резонансное взаимодействие является причиной неравновесности потока. Это видно из рис. 2, б, где показаны зависимости работы вязких и резонансных напряжений: $\text{Vis}(y)$ всегда отрицательно, а $T > 0$,

в результате $\text{Vis}(y) + T > 0$ и $\langle p_1 w_1 \rangle > 0$. В этом случае описать резонансное взаимодействие волны с потоком количественно не очень просто, поскольку w_1 в уравнении для завихренности зависит от y и ширина критического слоя сравнима с характерным масштабом течения, но качественное описание такое же, как раньше.

3. Квазилинейное приближение

Перейдем теперь к рассмотрению квазилинейного приближения. Это приближение хорошо известно в физике плазмы [9]. Оно применяется при рассмотрении взаимодействия пакета волн, имеющих случайные фазы, и частиц плазмы. Из плазменно-гидродинамической аналогии [6] следует, что аналог квазилинейного приближения должен существовать и в гидродинамике. Соответствие уравнений квазилинейного приближения в гидродинамике и физике плазмы представлено в табл. 2.

Таблица 2

Квазилинейное приближение

<p>1. Пакет ленгмюровских волн</p> $\Phi = \sum_k \Phi_k(t) e^{ikx + i\varphi_k}$	<p>1. Возмущения функции тока</p> $\Psi = \sum_k \Psi_k(t) e^{ikx + i\varphi_k}$
<p>2. Медленная и быстрая компоненты функции распределения</p> $f = f_0(v, t) + \sum_k f_k(v, t) e^{ikx + i\varphi_k}$	<p>2. Медленная и быстрая компоненты завихренности</p> $\Omega = \Omega_0(y, t) + \sum_k \Omega_k(y, t) e^{ikx + i\varphi_k}$
<p>3. Квазилинейное уравнение диффузии</p> $\frac{\partial f_0}{\partial t} = -i \frac{e}{m} \sum_k k \Phi_{-k} \frac{\partial f_k}{\partial v}$	<p>3. Квазилинейное уравнение диффузии</p> $\frac{\partial \Omega_0}{\partial t} = -i \sum_k k \Psi_{-k} \frac{\partial \Omega_k}{\partial y}$
<p>4. Линейное уравнение для k-й гармоники</p> $\frac{\partial f_k}{\partial t} + ikv f_k = ik \frac{e}{m} \Phi_{-k} \frac{\partial f_0}{\partial v}$ $ik \Phi_k = 4\pi e \int f_k dv$	<p>4. Линейное уравнение для k-й гармоники</p> $\frac{\partial \Omega_k}{\partial t} + ik U_0(y) \Omega_k = ik \Psi_{-k} \frac{\partial \Omega_0}{\partial y}$ $\frac{\partial^2 \Psi_k}{\partial y^2} - k^2 \Psi_k = \Omega_k$

В левом столбце представлены уравнения плазменной задачи, а в правом — ее гидродинамического аналога. В гидродинамике имеется плоскопараллельное течение и возмущение на его фоне в виде пакета

волн со случайными фазами. Функция тока Ψ аналогична потенциалу Φ (строка 1 в табл. 2). Предположим, что пакет можно считать широким в том смысле, что разброс положений критических слоев у разных гармоник, составляющих спектр, много больше ширины нелинейного критического слоя для одной гармоник. Принимая во внимание рассмотренную выше аналогию между функцией распределения $f(v, t)$ в плазме и завихренностью Ω в гидродинамике, представим Ω в виде суммы медленной и быстрой части (строка 2 табл. 2). Для средней завихренности получается квазилинейное уравнение диффузии, аналог уравнения для средней функции распределения (строка 3 табл. 2), а для возмущений поля завихренности и функции тока, так же как для возмущения функции распределения и потенциала, надо решать линейную систему (строка 4 табл. 2). Заметим, что в гидродинамике надо учитывать адиабатическое изменение амплитуды волны за счет изменения параметров среды.

Для задач гидродинамики понятие квазилинейного приближения имеет и второй смысл, когда при описании взаимодействия даже монохроматической волны и сдвигового потока с хорошей точностью можно описывать волновое поле в линейном приближении [10—14]. Это часто реализуется, когда неустойчивость волнового возмущения обусловлена механизмом «отрицательной» вязкости. В этом случае в уравнении для завихренности нужно учитывать вязкую диссипацию. Тогда уравнения квазилинейного приближения модифицируются следующим образом. Уравнение для средней завихренности имеет вид

$$\frac{\partial \Omega_0}{\partial t} = -i \sum_k k \Psi_{-k} \frac{\partial \Omega_k}{\partial y} + \frac{1}{\text{Re}} \frac{\partial^2 \Omega_0}{\partial y^2},$$

а комплексные амплитуды гармоник удовлетворяют системе линейных уравнений

$$\frac{\partial \Omega_k}{\partial t} + ikU_0(y)\Omega_k - \frac{1}{\text{Re}} \left(\frac{\partial^2 \Omega_k}{\partial y^2} - k^2 \Omega_k \right) = ik \Psi_{-k} \frac{\partial \Omega_0}{\partial y},$$

$$\frac{\partial^2 \Psi_k}{\partial y^2} - k^2 \Psi_k = \Omega_k.$$

Видно, что в уравнениях появляется дополнительный параметр — число Рейнольдса Re . При этом нелинейность определяется значением числа Рейнольдса, которое находится по параметрам течения вблизи критического уровня, где максимальны амплитуды полей, $\text{Re}_{\text{КС}}$. Очень часто $\text{Re}_{\text{КС}}$ оказывается небольшим, и это означает, что для описания поля вблизи критического уровня можно применять линейное приближение. Практически оказывается, что при $\text{Re}_{\text{КС}}$ порядка нескольких единиц хорошо работает линейное приближение для гармоник волнового поля.

В то же время, среднее поле скорости будет меняться во времени (возможно также установление стационарного поля, когда радиационное напряжение уравнивается вязким). В этом случае по форме уравнения остаются уравнениями квазилинейного приближения, хотя «почти линейность» волнового поля обеспечивается не случайными фазами гармоник, а вязкостью. Следует заметить, что часто квазилинейное приближение работает за пределами его формальной применимости [11—14].

4. Примеры применения квазилинейного приближения при взаимодействии волн и течений в океане

4.1. Взаимодействие волн с ветром

В геофизической гидродинамике плазменно-гидродинамическая аналогия впервые обсуждалась применительно к майлсовскому механизму генерации волн ветром [15, 16]. В рамках этой модели пренебрегают турбулентными флуктуациями в воздушном потоке и для волновых возмущений используют уравнения гидродинамики идеальной жидкости. Но тогда, как мы видели, сразу можно воспользоваться плазменно-гидродинамической аналогией. Следует заметить, что в задаче о генерации волн ветром существует еще одно упрощающее обстоятельство: из-за большой разности плотностей воды и воздуха дисперсионные свойства волн определяются динамикой водного слоя, а инкремент, вызванный резонансным взаимодействием с ветром, мал и может быть вычислен по теории возмущений. Это полностью аналогично взаимодействию волн с частицами плазмы «на хвосте» функции распределения или со слабым пучком, когда дисперсионные свойства волны определяются основной частью функции распределения, т. е. нерезонансными частицами.

Для механизма Майлса рассмотрены:

- 1) плазменно-гидродинамическая интерпретация линейного инкремента [3, 5];
- 2) квазилинейная деформация среднего профиля скорости (и завихренности) [17];
- 3) нелинейный майлсовский инкремент [6].

Следует, однако, заметить, что квазиламинарная модель Майлса плохо применима к описанию взаимодействия ветровых волн с реальным ветровым потоком. Это, прежде всего, связано с присутствием турбулентных флуктуаций в воздушном потоке. Вопрос об описании взаимодействия волн с турбулентными флуктуациями не прост и не решен до конца к настоящему времени.

Вообще для описания турбулентных потоков обычно используется система уравнений Рейнольдса [17]:

$$\frac{\partial \langle u_i \rangle}{\partial t} + \langle u_j \rangle \frac{\partial \langle u_i \rangle}{\partial x_j} + \frac{1}{\rho_a} \frac{\partial \langle p \rangle}{\partial x_i} = \frac{\partial \langle u'_i u'_j \rangle}{\partial x_j}.$$

Она незамкнута, для определения входящих в нее неизвестных турбулентных напряжений $\langle u'_i u'_j \rangle$ мы воспользуемся простейшим способом параметризации турбулентности — гипотезой турбулентной вязкости или градиентной гипотезой замыкания [17]. В этом случае турбулентные напряжения имеют вид

$$\langle u'_i u'_j \rangle = \nu_t \left(\frac{\partial \langle u_i \rangle}{\partial x_j} + \frac{\partial \langle u_j \rangle}{\partial x_i} \right).$$

Здесь ν_t — коэффициент турбулентной вязкости. В рамках простейшей градиентной модели ν_t полагается функцией y . Вид функции $\nu_t(y)$ таков, что получается логарифмический профиль скорости с вязким подслоем [17], толщина которого в типичных океанских условиях не превышает нескольких миллиметров. В этом слое формируется основной перепад скорости. Это означает, что рассмотренная выше ступенчатая модель скорости воздушного потока хорошо моделирует реальный профиль скорости ветра и его взаимодействие с волнами, т. е. реализуется механизм «отрицательной» вязкости.

Как уже упоминалось выше, при этом можно рассматривать взаимодействие волн с потоком в квазилинейном приближении, которое справедливо даже для монохроматической волны. Это будет выполнено, если число Рейнольдса в критическом слое Re_{KC} невелико. По определению

$Re_{KC} = \frac{w_1 \delta}{\nu}$, где w_1 — вертикальная скорость, δ — ширина КС, ν —

кинематическая вязкость воздуха. Оценивая $w_1 \sim ckak\delta$, где c — фазовая скорость волны, k — ее волновое число, для типичных параметров волн на воде $c = 1$ м/с, $\delta = 10^{-3}$ м, $ka = 0,05$, $\nu = 0,15$ см²/с имеем $Re_{KC} \sim 10^{-2}$. При этом хорошо применимо линейное приближение для описания волновых возмущений.

Основным нелинейным эффектом при этом является деформация среднего профиля скорости. Поскольку ветер генерирует волны и, значит, передает им некоторое количество движения, то результатом нелинейного взаимодействия волн и ветра должно быть появление отрицательной добавки к средней скорости ветра. Следствием этого является уменьшение инкремента ветровых волн, т. е. отрицательная нелинейная добавка к инкременту (рис. 3). Необходимо принять во внимание, что в океанских условиях толщина пограничного слоя мала по сравнению с высотой волны. Для того чтобы исключить сильную

геометрическую нелинейность, которая не влияет на резонансное взаимодействие волн с потоком, нужно перейти в криволинейную систему координат, в которой координатная линия совпадает с взволнованной водной поверхностью [11]. В рамках этой системы можно найти линейный инкремент волны и нелинейную поправку к нему, которая вызывается деформацией среднего профиля скорости. Причем оказалось, что коэффициент при нелинейной поправке существенно превосходит линейную компоненту [11].

Учет нелинейной поправки к инкременту особенно важен при вычислении не постоянного инкремента, а при учете вариаций инкремента ветровых волн в присутствии неоднородных течений. В частности, такая задача возникает при построении теории радиоизображения длинных поверхностных волн, которые имеют характерные масштабы от десятков до сотен метров. При этом короткая волна, которая эффективно рассеивает радиоволны сантиметрового диапазона в соответствии с механизмом Брэгга, имеет длину порядка нескольких сантиметров.

За счет чего проявляется длинная волна в рассеянном радиосигнале? В разных фазах длинной волны наблюдается разная амплитуда коротких волн. В результате радиосигнал, рассеянный длинной волной, оказывается модулированным [19]. Можно выделить два основных механизма, приводящих к модуляции волн сантиметрового диапазона в поле длинных волн [20]:

1) длинная волна — это неоднородное поле течений, которое приводит к модуляции короткой волны;

2) модуляция инкремента.

Первый механизм определяется градиентом скорости течения на поверхности в длинной волне, а второй — возмущением скорости ветра, который пропорционален скорости на поверхности воды. С ростом длины и периода волны при ее фиксированной крутизне градиент скорости падает, в связи с этим в случае очень длинных волн основным механизмом является модуляция инкремента.

При этом возникает проблема фазы модуляции, которая по существу определяет, в каком месте длинной волны амплитуда коротких волн максимальна. Для случая очень длинных волн фаза модуляции инкремента линейных волн в противофазе с возвышением поверхности воды. Это объясняется тем, что во впадинах больше скорость ветра относи-

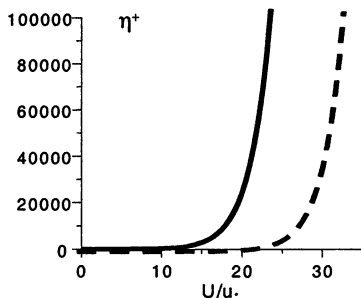


Рис. 3. Профили скорости ветра без учета (---) и с учетом (—) нелинейного взаимодействия ветра с волнами.

тельно воды (рис. 4, а). Это противоречит данным наблюдений, которые показывают, что фаза модуляции близка к $\pi/6 + \pi/3$. Разрешить это противоречие позволяет модель, учитывающая конечность амплитуды ветровых волн. Действительно, в областях, где больше амплитуда волн, меньше скорость ветра (рис. 4, б), это обусловлено нелинейным искажением профиля скорости ветра (см. рис. 3). В результате максимум амплитуды волн смещается в сторону гребня и фаза становится ближе к измеренной (рис. 4, в).



Рис. 4

Для определения величины модуляции радиосигнала, рассеянного радиоволной, используется модуляционная передаточная функция (МПФ), которая определяется следующим образом. Если поле смещения в длинной волне $\eta = \cos(kx - \omega t)$, то мощность рассеянного радиосигнала $P = P_0 \left(1 + ka \operatorname{Re} \left[|M| e^{i(kx - \omega t - \varphi_M)} \right] \right)$. МПФ называется комплексная величина $M = |M| e^{-i\varphi_M}$.

На рис. 5 представлены результаты различных вычислений МПФ: с учетом и без учета эффекта модуляции инкремента, в рамках линейной и квазилинейной моделей. Они сравнивались с измерениями [21], которые проводились на океанографической платформе в Северном море. Видно, что учет модуляции инкремента брэгговских волн принципиально важен для определения модуля МПФ, т. е. глубины модуляции (рис. 5, а), причем линейная и квазилинейная модели дают близкие значения. В то же время, фаза модуляции коротких

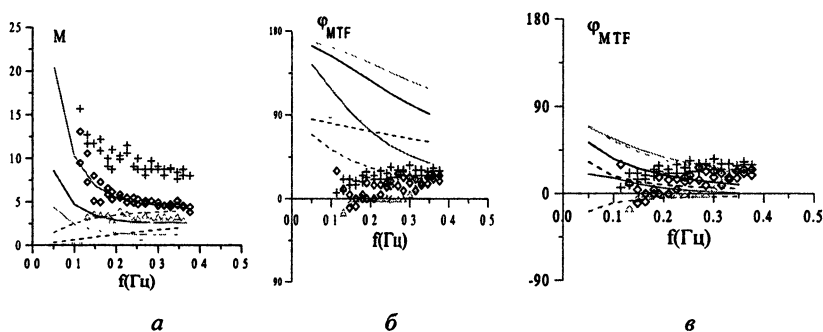


Рис. 5. Зависимость модуля и фазы МПФ от частоты длинной волны: модуль МПФ (а); фаза МПФ, рассчитанная в рамках линейной модели (б) и квазилинейной модели (в). Пунктир — расчеты МПФ без учета модуляции инкремента.

волн, полученная в рамках линейной модели, сильно расходится с результатами измерений (рис. 5, б). Учет нелинейности при взаимодействии волн с ветром в присутствии длинной волны дает значения фазы, близкие к измеренным (рис. 5, в).

Подобные возмущения скорости ветра могут создаваться не только длинными поверхностными волнами (зыбью), но и другими неоднородными течениями на поверхности, например интенсивными внутренними волнами. Они также могут вызывать неоднородное распределение ветровой ряби на морской поверхности [22].

4.2. Структуры на коже морских млекопитающих

Вообще, подобные задачи о резонансном взаимодействии поверхностных волн различного происхождения с обтекающим их сдвиговым потоком возникают достаточно часто. Приведем еще один «океанский» пример. Довольно давно изучаются особенности гидродинамики плавания морских млекопитающих, в частности дельфинов. И до сих пор не ясно, за счет какого механизма их сопротивление является аномально низким. В настоящее время считается, что оно обеспечивается за счет особенностей строения кожи дельфина. В последнее время скоростная фотосъемка обнаружила, что на коже плывущего дельфина наблюдаются волновые структуры (см. рис. 1, в). Формирование подобных структур наблюдалось на искусственных вязко-упругих покрытиях, обтекаемых потоком жидкости. Исследования [23] показали, что в системе сдвиговый поток — волна в покрытии существует 2 типа волн: одна ветка имеет положительную энергию, другая — отрицательную.

Волны положительной энергии нарастают за счет резонансного обмена энергией с потоком, а волны отрицательной энергии — за счет затухания в покрытии. В зависимости от величины этого затухания возбуждаются либо те, либо другие волны. Взаимодействие с турбулентным потоком, как и в случае поверхностных волн на воде, можно рассматривать в квазилинейном приближении. При этом, поскольку волна нарастает за счет взаимодействия с потоком, поток будет тормозиться, т. е. энергия у потока отбирается, это эквивалентно увеличению сопротивления. В искусственных покрытиях так и происходит, а у кожи дельфина сопротивление уменьшается неизвестно почему. Этот вопрос остается открытым.

4.3. Квазидвумерные следы за телами в потоках

Рассмотрим, наконец, последний в настоящей лекции пример, который удобно рассматривать с точки зрения энергообмена волн со сдвиговым потоком и в котором применимо квазилинейное приближение. Это задача о турбулентном следе за буксируемыми телами в стратифицированной жидкости. Она представляет интерес в различных приложениях,

в частности в связи с необходимостью дистанционной диагностики подводных препятствий. Турбулентные следы за препятствиями затухают очень медленно, и можно ожидать, что они будут вызывать значительные возмущения фоновых волн на больших расстояниях от препятствия. Об этом свидетельствуют результаты многочисленных экспериментов по протяжке тел в стратифицированной жидкости [24—26]. С точностью до системы отсчета эта постановка эксперимента эквивалентна задаче об обтекании подводного препятствия внешним течением.

За счет чего происходит «расплывание» следа? Была предложена теоретическая модель [27], в рамках которой основным механизмом эволюции следа является развитие гидродинамической неустойчивости этого струйного течения, которая обусловлена резонансным взаимодействием (энергообменом) между волнами и средним течением в следе. Нарастание возмущений сопровождается передачей им импульса от среднего течения, в результате плотность его среднего импульса падает, средняя скорость падает и след расплывается. В рамках модели [27] этот процесс передачи средней энергии к возмущениям описывается в квазилинейном приближении. При этом оказывается, что средняя скорость в следе удовлетворяет диффузионному уравнению:

$$\frac{\partial U}{\partial t} = D(t) \frac{\partial^2 U}{\partial y^2},$$

коэффициент диффузии которого $D(t)$ определяется нарастанием волновых возмущений и зависит только от времени. Решением такого уравнения является функция Гаусса. Это объясняет обычно наблюдаемую форму профиля скорости в следе.

К настоящему времени проведены детальные измерения скорости на оси следа в стратифицированной жидкости и его ширина в зависимости от времени [24—26]. На рис. 6 точками показаны результаты измерений, приведенные в [25], а кривые получены в результате расчетов в рамках квазилинейной модели. Видно хорошее согласие теории и эксперимента.

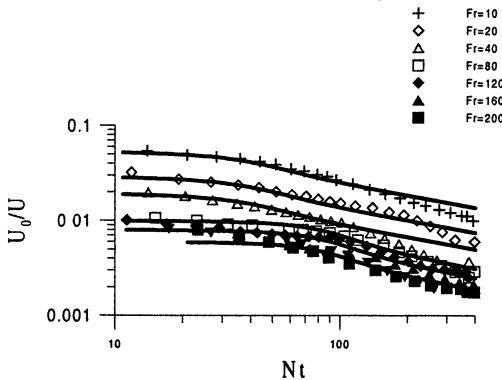


Рис. 6. Зависимость скорости на оси следа от времени: квазилинейная теория и лабораторный эксперимент.

Следует заметить, что результаты расчетов эволюции среднего течения в квазилинейном приближении зависят от начального спектра возмущений, который не измерялся в экспериментах. Для более детальной проверки квазилинейной модели было проведено ее сопоставление с результатами прямого численного моделирования эволюции двумерного следа. Прямое численное моделирование основано на решении полных уравнений гидродинамики без упрощений и поэтому может рассматриваться как разновидность эксперимента.

В численном эксперименте исследовалась эволюция двумерной струи с наложенными на нее двумерными возмущениями поля скорости с нулевым средним и значением кинетической энергии, составляющей 20 % от кинетической энергии среднего течения. Спектр возмущений имел максимум, соответствующий волновому числу наиболее неустойчивой моды. В численном эксперименте было получено 10 реализаций поля скорости, удовлетворяющих таким начальным условиям. На струе развивалась в основном изгибная мода, которая характеризуется наибольшим линейным инкрементом. Проводилось сравнение результатов расчетов в рамках квазилинейной модели и прямого численного моделирования. Для этого в рамках квазилинейной модели были проведены расчеты, соответствующие начальным условиям 10 реализаций прямого

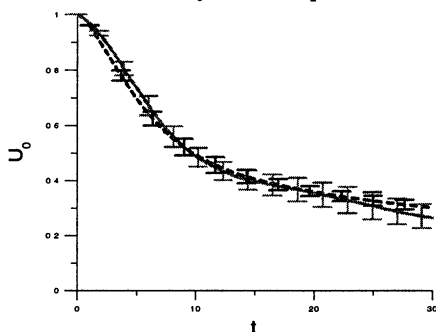


Рис. 7. Зависимость скорости на оси следа от времени: квазилинейная теория (—) и численный эксперимент (---).

численного моделирования. Полученные в модели и численном эксперименте величины осреднялись и сравнивались между собой. На рис. 7 показаны зависимости средней продольной скорости на оси следа, а на рис. 8 — средние спектры возмущений функции тока в разные моменты времени. Видно, что квазилинейная модель очень хорошо воспроизводит временную зависимость скорости на оси следа (рис. 7). При вычислении спектров отличия более существенны (рис. 8), но квазилинейная модель хорошо воспроизводит основные особенности эволюции спектра — уменьшение волнового числа, соответствующего спектральному максимуму, и зависимость величины спектрального максимума от времени.

Итак, квазилинейная модель находится в хорошем количественном согласии с результатами численного эксперимента, т. е. в двумерном следе основным нелинейным процессом является возникновение среднего течения. Взаимодействие гармоник, которое не учитывается в квазилинейной модели, по-видимому, оказывает существенное влияние

на формирование спектра возмущений, но слабо влияет на параметры среднего течения.

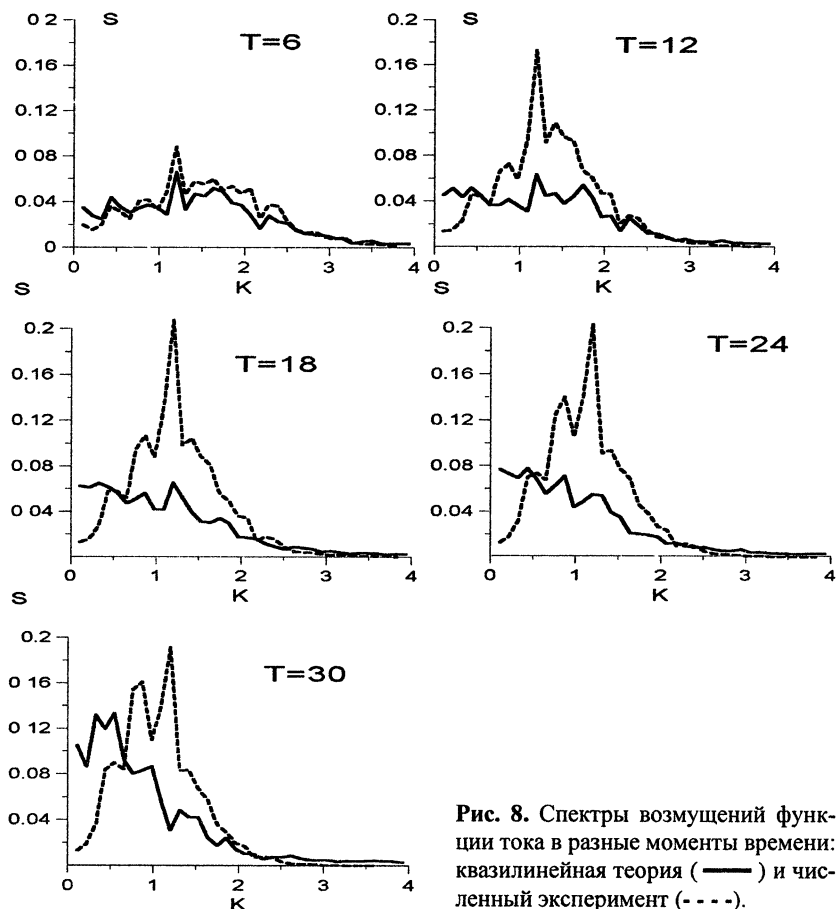


Рис. 8. Спектры возмущений функции тока в разные моменты времени: квазилинейная теория (—) и численный эксперимент (---).

Работа выполнена при частичной финансовой поддержке грантов РФФИ (проект 04-05-64264), Президентской программы по поддержке ведущих научных школ России (проект А-0046) и Программы Президиума РАН «Математические методы нелинейной динамики».

Литература

1. Тимофеев А. В. Резонансные явления в колебаниях плазмы. М.: Физматлит, 2000.
2. Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. М.: Наука, 1981.
3. Тимофеев А. В. Резонансные эффекты в колебаниях неоднородных течений сплошных сред // Вопр. теории плазмы. Вып. 17. М.: Энергоатомиздат, 1989. С.157—244.

4. Михайловский А. Б. Теория плазменных неустойчивостей. Т. 1, 2. М.: Атомиздат, 1975.
5. Андронов А. А., Фабрикант А. Л. Затухание Ландау, ветровые волны и свисток // *Нелинейные волны* / Ред. А. В. Гапонов-Грехов. М.: Наука, 1979. С. 68—104.
6. Реутов В. П. Плазменно-гидродинамическая аналогия и нелинейная стадия неустойчивости ветровых волн // *Изв. АН СССР. ФАО*. 1980. Т. 16, № 12. С. 1266—1275.
7. Реутов В. П. Нестационарный критический слой и нелинейная стадия неустойчивости в плоском течении Пуазейля // *ПМТФ*. 1982. № 4. С. 43—54.
8. Филлипс О. М. Динамика верхнего слоя океана. Л.: Гидрометеиздат, 1980. 320 с.
9. Вопросы теории плазмы / Ред. Б. Б. Кадомцева. Вып. 7. М.: Энергоатомиздат, 1973.
10. Jenkins A. D. A quasilinear eddy-viscosity model for the flux of energy and momentum to wind waves using conservation-law equations in a curvilinear coordinate system // *J. Phys. Oceanogr.* 1992. V. 22, № 8. P. 843—858.
11. Реутов В. П., Троицкая Ю. И. О нелинейных эффектах при взаимодействии волн на воде с турбулентным ветром // *Изв. РАН. ФАО*. 1995. Т. 31, № 5. С. 825—834.
12. Troitskaya Yu. I. Viscous diffusion nonlinear critical layer in a stratified shear flow // *J. Fluid Mech.* 1991. V. 233. P. 25—48.
13. Troitskaya Yu. I., Reznik S. N. Quasi-steady dissipative nonlinear critical layer in a stratified shear flow // *Phys. Fluids*. 1996. V. 8, № 12. P. 3313—3328.
14. Churilov S. M., Shukhman I. G. Nonlinear stability of a stratified shear flow: a viscous critical layer // *J. Fluid Mech.* 1987. V. 180. P. 1—20.
15. Miles J. W. On the generation of surface waves by shear flows // *J. Fluid Mech.* 1957. V. 3. P. 185—204.
16. Miles J. W. On the generation of surface waves by shear flows. Part 2 // *J. Fluid Mech.* 1959. V. 6. P. 568—582.
17. Фабрикант А. Л. Квазилинейная теория ветровых волн // *Изв. АН СССР. ФАО*. 1976. Т. 12, № 8. С. 858—862.
18. Монин А. С., Яглом А. М. Статистическая гидромеханика. Ч. 1 // СПб: Гидрометеиздат, 1992. 696 с.
19. Hasselman K, Raney R. K., Plant W. J., Alpers W., Shuchman R. A., Lyzenga D. R., Rufenach C. L., Tucker M. J. Theory of synthetic aperture radar imaging: a MARSEN view // *J. Geophys. Res.* 1985. V. 90. P. 4659—4686.
20. Smith J. A. Modulation of short wind waves by long waves // *Surface Waves and Fluxes*. V. 1. P. 247—284. Kluwer. Netherlands: Academic Publishers, 1990.
21. Plant W. J., Keller W. C., Cross A. Parametric dependence of ocean wave-radar modulation transfer functions // *J. Geophys. Res.* 1983. V. 88. P. 9747—9756.
22. Горшков К. А., Долина И. С., Соустова И. А., Троицкая Ю. И. Модуляция коротких ветровых волн в присутствии интенсивных внутренних волн. Эффект модуляции инкремента // *Изв. вузов. Радиофизика*. 2003. Т. 46, № 7. С. 513—536.
23. Реутов В. П., Рыбушкина Г. В. Квазилинейная теория генерации гидроупругих волн в турбулентном пограничном слое на упругом покрытии // *Изв. РАН. ФАО*. 2001. Т. 40, № 1. С. 118—133.
24. Spedding G. R., Browand F. K., Fincham A. M. Turbulence, similarity scaling and vortex geometry in the wake of a towed sphere in a stably stratified fluid // *J. Fluid Mech.* 1996. V. 314. P. 53—103.
25. Spedding G. R. The evolution of initially-turbulent bluff-body wakes at high internal Froude number // *J. Fluid Mech.* 1997. V. 337. P. 283—301.
26. Bonnier M., Eiff O. Experimental investigation of the collapse of a turbulent wake in a stably stratified fluid // *Phys. Fluids*. 2002. V. 14, № 2. P. 791—801.
27. Баландина Г. Н., Панко В. В., Сергеев Д. А., Троицкая Ю. И. Эволюция дальнего турбулентного следа за объектом, буксируемым в стратифицированной жидкости при больших числах Рейнольдса и Фруда // *Изв. РАН. ФАО*. 2004. Т. 40, № 1. С. 118—133.

ОСНОВЫ ФИЗИКИ ИНТЕНСИВНЫХ АКУСТИЧЕСКИХ ВОЛН

О. В. Руденко

1. Об акустической нелинейности

Первые вопросы, возникающие у физика при изучении интенсивных акустических волн, обычно таковы: откуда берутся нелинейные члены в математических моделях, описывающих процессы возбуждения и распространения волн? Какая среда является более нелинейной и чем определяется величина нелинейных модулей упругости?

Ниже используется следующая условная классификация. Мы будем говорить о нелинейностях границ и объема среды; каждая из них, в свою очередь, подразделяется на геометрическую, физическую и структурную нелинейности.

1.1. Граничная нелинейность

1.1.1. Рассмотрим плоский поршень, совершающий колебания в направлении нормали (совпадающей с осью x) к своей поверхности и возбуждающий в упругой среде, которая занимает полупространство $x > 0$, бегущую продольную волну (рис. 1).

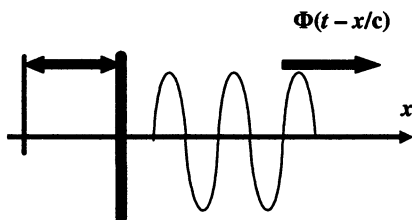


Рис. 1. Поршень, совершающий колебания большой амплитуды, и нелинейность, обусловленная конечностью смещения границы

Пусть смещение поршня из среднего положения $x = 0$ известно и описывается законом $x = X(t)$. От поршня убегает волна, скорость частиц среды в которой изменяется как

$$u = u_0 \Phi \left(t - \frac{x}{c} \right). \quad (1)$$

Поскольку скорость движения поверхности поршня должна совпадать со скоростью частиц среды на этой поверхности, для неизвестной функции Φ получаем функциональное уравнение

$$u_0 \Phi \left(t - \frac{X(t)}{c} \right) = \frac{dX}{dt}. \quad (2)$$

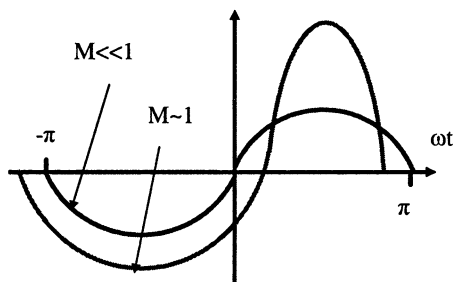
В частности, если поршень движется по гармоническому закону $X(t) = -X_0 \cos \omega t$, для нахождения Φ имеем

$$\frac{u}{u_0} = \Phi \left(\omega t + \frac{u_0}{c} \cos \omega t \right) = \sin \omega t, \quad u_0 = \omega X_0. \quad (3)$$

Очевидно, что при гармоническом законе движения поршня $X(t)$ форма бегущей волны $\Phi(t)$ содержит как постоянную составляющую, так и высшие гармоники. Различие между $X'(t)$ и $\Phi(t)$ усиливается с ростом числа Маха $M = u_0/c$ и становится особенно заметным при скоростях движения поршня, сравнимых со скоростью звука. Такая ситуация может возникнуть в жидкостях с пузырьками газа, в которых скорость звука может быть очень малой, а также в резонаторах, где эффекты граничной нелинейности могут накапливаться со временем.

Поскольку нелинейность связи (2) не зависит от свойств среды, ее имеет смысл называть «геометрической» или «кинематической» нелинейностью.

Рис. 2. Бегущая волна, искаженная граничной нелинейностью.



Форма одного периода бегущей волны (3) изображена на рис. 2. Для малых чисел Маха волна почти синусоидальна, а при конечных значениях M положительный и отрицательный «полупериоды» временного профиля искажены неодинаково. Для расчета спектрального состава сигналов, заданных неявными функциями, в акустике развит ряд простых приемов (см., например, [1]). Используя их, найдем коэффициенты C_n разложения волны (3) в ряд Фурье:

$$\frac{u}{u_0} = \Phi(\omega t) = \sum_{n=-\infty}^{\infty} C_n \exp(in\omega t) = \sum_{n=0}^{\infty} [A_n \cos n\omega t + B_n \sin n\omega t],$$

$$C_n = \frac{i^{-n}}{n} J_n'(nM). \quad (4)$$

В частности, для постоянной составляющей и первых двух гармоник имеем

$$A_0 = -\frac{M}{2}, \quad B_1 = J_0(M) - J_2(M), \quad A_2 = -\frac{1}{2}[J_1(2M) - J_3(2M)]. \quad (5)$$

Постоянная составляющая скорости направлена в сторону поршня. Ее появление связано с идеализированным представлением колеблющейся границы в виде бесконечной плоскости; из-за одинакового давления в любом ортогональном к x направлению (в отличие от поршня конечного размера) жидкость не может подтекать к оси (см., например, [2]).

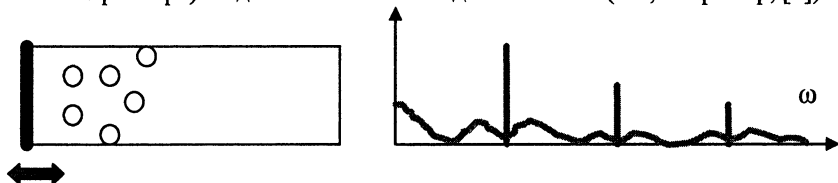


Рис. 3. Уширение спектра кавитационного шума при взаимодействии с колеблющейся границей.

На рис. 3 изображено облако кавитационных пузырьков, возникающее вблизи мощного ультразвукового излучателя. Скорость звука в этой области может уменьшиться до 20—40 м/с. Пузырьки колеблются, схлопываются, излучают широкополосный шум. Этот шум взаимодействует с колебаниями границы, и его спектр воспроизводится у «подножия» дискретных составляющих — высших гармоник основной частоты УЗ-источника. В результате широкополосный шум усиливается, а его спектр эффективно уширяется [3]. Описанное явление дополняет хорошо известные [4] механизмы уширения спектров под действием объемной нелинейности.

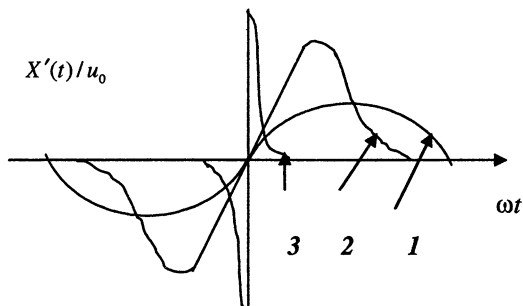


Рис. 4. Форма одного периода колебаний границы, обеспечивающая накопление энергии в полости резонатора в условиях проявления граничной нелинейности. Кривые 1, 2 и 3 отвечают растущим значениям u_0 .

Другие задачи, требующие учета конечности смещения границы, связаны с накоплением нелинейных искажений в добротных резонаторах. Поскольку форма интенсивных «стоячих» волн сильно искажена, происходит их насыщение при использовании гармонической накачки. Следовательно, для того чтобы продолжить процесс «закачивания» энергии в резонатор, нужно изменить форму колебаний границы [5] (подобно тому, как мы увеличиваем промежуток времени между толчками при раскачивании качелей). Необходимая для поддержания резонансно-го возбуждения форма колебаний источника изображена на рис. 4.

1.1.2. В реальной ситуации заданным следует считать не закон смещения поршня $x = X(t)$, а приложенную к поршню силу $F(t)$. Например, если на возбуждающий волну электромеханический преобразователь подается переменное электрическое напряжение V , известен именно закон $V(t)$ и, следовательно, $F(t)$. Уравнение колебаний для поршня массой m и площадью поверхности S при этом таково:

$$m \frac{d^2 X}{dt^2} = F(t) - p_a S. \quad (6)$$

Здесь p_a — давление на поршень, создаваемое убегаящей волной. Давление связано со скоростью нелинейным соотношением, известным из теории ринановых волн; если волна слабо нелинейна, эта связь содержит только линейный и квадратично-нелинейный члены:

$$\frac{p_a}{c^2 \rho} = \frac{u}{c} + \frac{\varepsilon}{2} \left(\frac{u}{c} \right)^2. \quad (7)$$

Поскольку на поверхности поршня $u = X'(t)$, уравнение (6) примет вид

$$\frac{d^2 X}{dt^2} + \frac{c \rho S}{m} \frac{dX}{dt} \left(1 + \frac{\varepsilon}{2c} \frac{dX}{dt} \right) = \frac{1}{m} F(t). \quad (8)$$

Ясно, что реакция убегаящей волны на поршень также нелинейна. В частности, при гармоническом изменении силы $F(t)$ закон движения поршня $X(t)$ будет содержать высшие гармоники. Нелинейность проявляется тем сильнее, чем больше произведение нелинейного параметра среды и числа Маха, т. е. εM . В отличие от предыдущего случая, реакция среды обусловлена нелинейной зависимостью давления от скорости (в твердых телах можно говорить о нелинейной зависимости напряжения от деформации, т. е. об отклонении от закона Гука), поэтому описанная нелинейность содержит, наряду с «геометрической», еще и «физическую» составляющую.

1.1.3. Рассмотрим теперь нелинейность границы, на которой происходит контакт двух тел. Классическим примером служит задача о двух сферах (контакт Герца [6], рис. 5). Сила, с которой отталкиваются сферы радиусов R_1 и R_2 , зависит нелинейно от разности смещений их центров $(\xi_2 - \xi_1)$, появившейся вследствие деформирования:

$$F = E \sqrt{\frac{R_1 R_2}{R_1 + R_2}} (\xi_2 - \xi_1)^{3/2} \theta(\xi_2 - \xi_1). \quad (9)$$

Здесь E — эффективный модуль, зависящий от модулей Юнга E_1 и E_2 материала сфер и коэффициентов Пуассона σ_1 и σ_2 :

$$\frac{1}{E} = \frac{3}{4} \left(\frac{1 - \sigma_1^2}{E_1} + \frac{1 - \sigma_2^2}{E_2} \right). \quad (10)$$

При отрицательных значениях разности смещений $(\xi_2 - \xi_1)$ сферы расходятся, не деформируясь, и сила равна нулю. При положительных значениях этой разности сила зависит от нее как $(\xi_2 - \xi_1)^{3/2}$. Очевидно, что при колебаниях системы из двух шариков под действием периодической внешней силы нелинейность существенна только в тех случаях, когда сильно изменяется площадь соприкосновения тел. Если контакт «поджат» большим статическим усилием, нелинейность уменьшится.

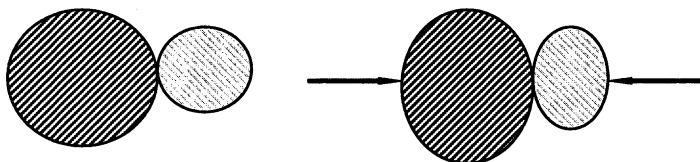


Рис. 5. Нелинейный контакт Герца

Свойства герцевских контактов используются при описании нелинейной динамики гранулированных сред (см. [7]), нелинейного контакта шероховатых поверхностей [8] и ряда других систем.

На рис. 6 изображена система из двух пластин, одна из которых (нижняя) является гладкой, а вторая — шероховатой.

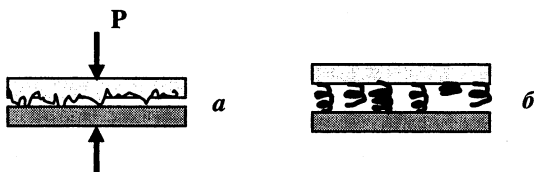


Рис. 6. Шероховатая поверхность (а) и модель (б) в виде ансамбля пружин разной длины, но одинаковой жесткости.

Очевидно, с ростом прижимающего усилия P все большее число «зубцов» (т. е. микровыступов на контакте) деформируется и жесткость контакта возрастает. Если $P \rightarrow 0$, падающая волна с частотой ω почти полностью отразится от границы раздела. Напротив, если $P \rightarrow \infty$, волна пройдет через границу, т. к. акустический контакт будет идеальным. В обоих случаях отраженной волны на частоте 2ω не возникнет. Вторая гармоника рождается при промежуточных значениях P , а зависимость ее амплитуды от P имеет максимум. Интересно, что эта кривая представляет статистическое распределение высот микровыступов (при использовании модели, показанной на рис. 6, б).

Другим примером «структурной» граничной нелинейности служит так называемая «хлопающая» нелинейность (рис. 7). Сжимающая сила вызывает упругую реакцию, а при растяжении контакт разрывается и тела расходятся. Осциллирующее по гармоническому закону напряжение вызывает сильные искажения деформационных осцилляций, в спектре которых появляются высшие гармоники.

В более общем случае нелинейность существенна для динамики любых контактов, по-разному реагирующих на сжатие и растяжение.

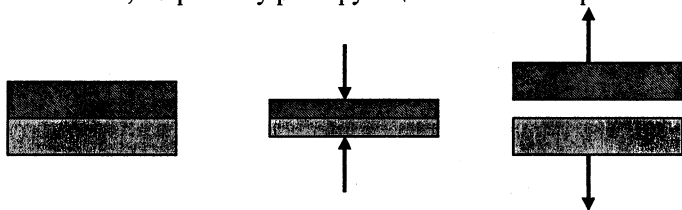


Рис. 7. К пояснению механизма «хлопающей» нелинейности

1.1.4. Следующий тип граничной нелинейности наблюдается при обтекании осциллирующим потоком тел, имеющих острые кромки. Градиенты потока имеют порядок $u/\max(r, \delta)$ где $\delta = \sqrt{\nu/\omega}$ — толщина акустического погранслоя, r — минимальный радиус кривизны препятствия, ν — кинематическая вязкость. Для острых кромок толщина δ является определяющей. Нелинейность существенна при числе Рейнольдса $Re \sim 1$, пропорциональном отношению членов в уравнении движения (12), которое можно оценить как

$$Re \sim \left| \frac{(\mathbf{u}\nabla)\mathbf{u}}{\partial\mathbf{u}/\partial t} \right| \sim \frac{u}{\sqrt{\omega\nu}} = \sqrt{\frac{2I}{c\omega\eta}}. \quad (11)$$

Согласно этой формуле, в воздухе ($\eta = 0,018 \cdot 10^{-3}$ Па·с) на частоте 500 Гц, заметная нелинейность должна наблюдаться при уровнях звука ~ 120 дБ. Эта оценка согласуется с данными экспериментов; если же на краях отверстий образуются вихри (рис. 8), нелинейные явления наблюдаются, начиная с 90 дБ [9].

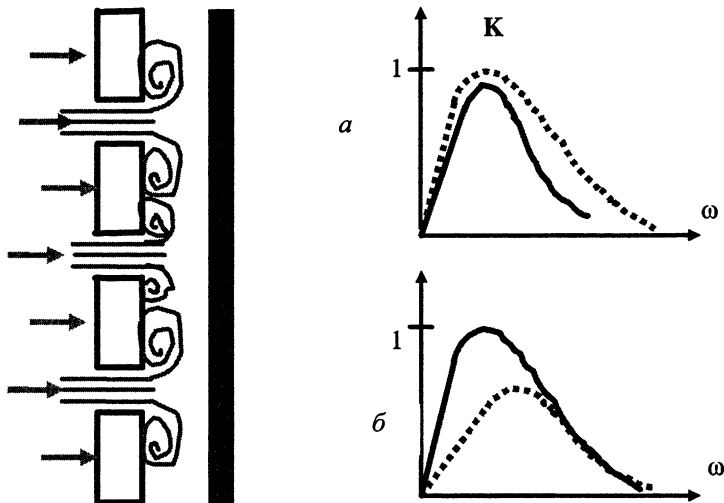


Рис. 8. Схема резонансного звукопоглотителя

На рис. 8 изображена схема резонансного звукопоглотителя. Колебания воздуха в отверстиях перфорированной панели затухают из-за трения о стенки, сеточки или специальный волокнистый материал. При правильном выборе параметров на определенной частоте можно добиться полного поглощения ($K = 1$). Нелинейность может уменьшить коэффициент K для линейно согласованного поглотителя (*a*) или увеличить его для несогласованного (*b*); линейные частотные характеристики даны штриховыми кривыми, нелинейные — сплошными [9]. Таким образом, при создании поглотителей для интенсивного звука нужно учитывать эффекты граничной нелинейности.

1.2. Объемная нелинейность

Объемная нелинейность, в отличие от граничной, может приводить к нелинейным эффектам, накапливающимся при распространении волны в среде. Если, например, на входе в среду волна была гармонической, при ее распространении будет происходить усиливающаяся с расстоянием генерация высших гармоник. Расстояние, на котором амплитуда 2-й гармоники достигает $1/2$ от амплитуды 1-й (без учета реакции), есть длина образования разрыва (характерная «нелинейная длина»). На этом расстоянии укладывается много длин волн, их число равно $(2\pi\epsilon M)^{-1}$ [1].

1.2.1. Напомним, как выделяют физическую и геометрическую нелинейности в уравнениях гидродинамики. Уравнения движения, непрерывности и состояния имеют вид [2]

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \nabla) \mathbf{u} \right] = -\nabla p + \eta \Delta \mathbf{u} + \left(\zeta + \frac{\eta}{3} \right) \text{grad} (\text{div} \mathbf{u}), \quad (12)$$

$$\frac{\partial p}{\partial t} + \text{div}(\rho \mathbf{u}) = 0, \quad p = p(\rho). \quad (13)$$

Пусть невозмущенное состояние среды есть $\rho = \rho_0$, $p = p_0$, $\mathbf{u} = 0$. Возмущения параметров, связанные с волной, обозначим штрихованными буквами и положим в (12), (13)

$$\rho = \rho_0 + \rho', \quad p = p_0 + p', \quad \frac{\rho'}{\rho} \sim \frac{p'}{p_0} \sim \frac{|\mathbf{u}|}{c_0} \sim \mu \ll 1.$$

Уравнения движения и непрерывности примут вид

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} + \nabla p' - \eta \Delta \mathbf{u} - \left(\zeta + \frac{\eta}{3} \right) \text{grad} (\text{div} \mathbf{u}) = -\rho' \frac{\partial \mathbf{u}}{\partial t} - (\rho_0 + \rho') (\mathbf{u} \nabla) \mathbf{u},$$

$$\frac{\partial \rho'}{\partial t} + \rho_0 \text{div} \mathbf{u} = -\text{div}(\rho' \mathbf{u}).$$

В левых частях этих уравнений собраны линейные, в правых — нелинейные члены. Появление последних обусловлено нелинейностью исходных уравнений, а не свойствами среды, поэтому указанную нелинейность называют «геометрической».

Напротив, нелинейность уравнения состояния

$$p' = c_0^2 \rho' + \frac{1}{2} \left(\frac{\partial^2 p}{\partial \rho^2} \right) \rho'^2 + \dots = c_0^2 \rho' + \frac{(\gamma - 1) c_0^2}{2 \rho_0} \rho'^2 + \dots$$

связана с нелинейной зависимостью между приращениями давления и плотности, т. е. это — физическая нелинейность. Заметим, что для адиабатического уравнения состояния переход к линейной среде соответствует $\gamma = c_p / c_v = 1$. При этом, однако, волна останется нелинейной, для нее коэффициент нелинейности ϵ обращается не в ноль, а в единицу.

1.2.2. Обсудим теперь объемные нелинейности для твердых тел. Уравнение движения и связь тензоров напряжений σ_{ik} и деформаций e_{ik} (которая в отсутствие нелинейных членов известна как закон Гука) имеют вид

$$\rho_0 \frac{\partial^2 U_i}{\partial t^2} = \frac{\partial}{\partial x_k} \sigma_{ik}, \quad \sigma_{ik} = K e_{ll} \delta_{ik} + 2\mu \left(e_{ik} - \frac{1}{3} \delta_{ik} e_{ll} \right) + O(e_{ik}^2). \quad (14)$$

Здесь U_i — вектор смещений, K , μ — модули всестороннего сжатия и сдвига [6],

$$e_{ik} = \frac{1}{2} \left(\frac{\partial U_i}{\partial x_k} + \frac{\partial U_k}{\partial x_i} + \frac{\partial U_l}{\partial x_k} \frac{\partial U_l}{\partial x_i} \right). \quad (15)$$

Посредством $O(e_{ik}^2)$ обозначена группа нелинейных членов:

$$O(e_{ik}^2) = \left(\mu + \frac{A}{4} \right) \left(\frac{\partial U_l}{\partial x_i} \frac{\partial U_l}{\partial x_k} + \frac{\partial U_i}{\partial x_l} \frac{\partial U_k}{\partial x_l} + \frac{\partial U_l}{\partial x_k} \frac{\partial U_l}{\partial x_i} \right) + \\ + \frac{1}{2} \left(K - \frac{2}{3} \mu + B \right) \left(\frac{\partial U_l}{\partial x_m} \frac{\partial U_l}{\partial x_m} \delta_{ik} + 2 \frac{\partial U_i}{\partial x_k} \frac{\partial U_l}{\partial x_l} \right) + \frac{A}{4} \frac{\partial U_k}{\partial x_l} \frac{\partial U_l}{\partial x_i} + \\ + B \left(\frac{\partial U_l}{\partial x_m} \frac{\partial U_m}{\partial x_l} \delta_{ik} + 2 \frac{\partial U_k}{\partial x_i} \frac{\partial U_l}{\partial x_l} \right) + C \frac{\partial U_l}{\partial x_i} \frac{\partial U_l}{\partial x_k} \delta_{ik},$$

где A, B, C — нелинейные модули упругости 3-го порядка (коэффициенты Ландау [6] в разложении внутренней энергии по степеням тензора деформаций):

$$E = \mu e_{ik}^2 + \left(\frac{K}{2} - \frac{\mu}{3} \right) e_{ii}^2 + \frac{A}{3} e_{ik} e_{ii} e_{ii} + B e_{ik}^2 e_{ii} + \frac{C}{3} e_{ii}^3.$$

Очевидно, геометрическая нелинейность появляется в связи (15) между тензором деформаций и компонентами вектора смещений, а физическая нелинейность — это все члены в выражении $O(e_{ik}^2)$.

1.2.3. Рассмотрим теперь твердые среды с наследственными, в том числе с гистерезисными, свойствами. Начало исследования таких систем связано с работами Л. Больцмана (1876) и В. Вольтерра (1913). Вместо «уравнений состояния», содержащих физическую нелинейность, здесь принято говорить об «определяющих уравнениях». Еще Вольтерра при построении нелинейной теории наследственной упругости использовал представление Фреше для нелинейного функционала в виде ряда кратных интегралов, аналогичное ряду Тейлора. Для одномерного случая (скалярный вариант) разложение Вольтерра — Фреше имеет вид

$$\sigma(t) = \int_{-\infty}^t G_1(t-\tau) d\epsilon(\tau) + \int_{-\infty}^t \int_{-\infty}^t G_2(t-\tau_1, t-\tau_1) d\epsilon(\tau_1) d\epsilon(\tau_2) + \dots, \\ \epsilon(t) = \int_{-\infty}^t J_1(t-\tau) d\sigma(\tau) + \int_{-\infty}^t \int_{-\infty}^t J_2(t-\tau_1, t-\tau_1) d\sigma(\tau_1) d\sigma(\tau_2) + \dots$$

Здесь ядра G_i называют функциями релаксации, а ядра J_i — функциями ползучести. Иногда вместо общего разложения (16) используют модельные нелинейные функционалы — уравнения Работнова и Линдермана — Розовского:

$$\varphi[\epsilon(t)] = \sigma(t) + \int_{-\infty}^t K(t-\tau) \sigma(\tau) d\tau, \\ \epsilon(t) = \psi[\sigma(t)] + \int_{-\infty}^t K(t-\tau) \chi[\sigma(\tau)] d\tau.$$

Выбор ядер и аппроксимирующих функций — сложная проблема; она основана на экспериментальных данных и проводится отдельно для каждого материала и заданной области внешних условий (различные диапазоны температур, скоростей деформирования и других параметров).

Теории типа Вольтерра относятся к процессам с запаздыванием, но обратимым в том смысле, что кривые мгновенного деформирования при нагрузке и разгрузке совпадают. В гистерезисных средах ситуация может быть иной. На рис. 9 изображены зависимости напряжения от деформации, учитывающие пластическое деформирование с разгрузкой для металла и для материала с деструкционной псевдопластичностью (пример: армированные пластики с концентрацией напряжений на изгибах волокон, трещинах) [10]. Нелинейность в металлах связана с накоплением существенно необратимой пластической деформации, и закон разгрузки близок к линейному (случай *a*). В случае *б* при разгрузке часть трещин «закрывается», график разгрузки приближается к началу координат.

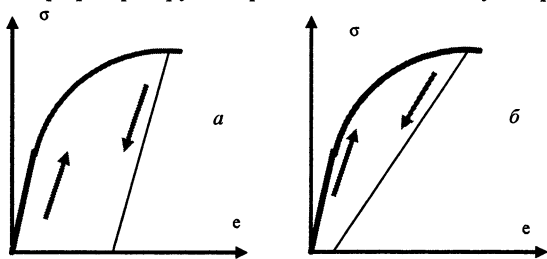


Рис. 9. Нагружение и разгрузка материалов с гистерезисом

Для циклических процессов удобна модель зависимости напряжения от деформации в форме Рэлея (1887).

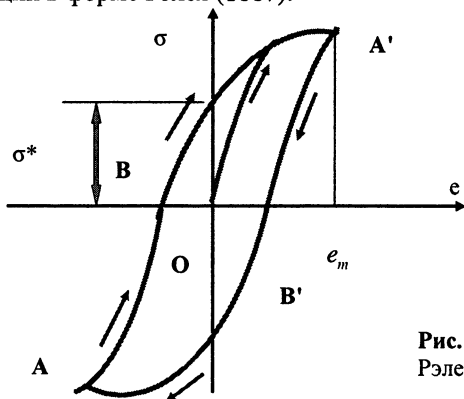


Рис. 10. Гистерезис Рэлея.

Петля на рис. 10 задана формулами

$$\sigma = (E + be_m)e + \frac{b}{2}(e^2 - e_m^2), \quad \sigma = (E + be_m)e - \frac{b}{2}(e^2 - e_m^2)$$

на участках $A'B'A$, ABA' и формулой $\sigma = Ee - be^2$ — на участке OA' . Остаточные напряжения и гистерезисные потери (за цикл) даются следующими выражениями:

$$\sigma^* = \frac{b}{2}e_m^2, \quad W = \oint e \cdot d\sigma = \frac{4}{3}e_m^3.$$

Гистерезисные свойства акустических материалов и сред отдельно рассмотрены в статье Л. А. Островского в данном издании (см. также [11]).

В последние годы в этой области акустики появилось значительное число публикаций, развивается международное сотрудничество. Однако большинство работ, как и в механике [10], носит скорее материаловедческий характер. Вместе с тем, имеет смысл постановка и решение прямых и обратных нелинейных волновых задач, в том числе для целей нелинейной диагностики (см. следующий раздел). Волны ведут себя в наследственных средах иначе, чем при обычной алгебраической зависимости $p'(p')$ или $\sigma(e)$. Для «релаксирующих» (в смысле Мандельштама — Леонтовича [2]) зависимостей между параметрами нелинейные волны исследованы достаточно подробно [1], но для гистерезисных сред таких результатов мало. В качестве примера на рис. 11 сравнивается искажение симметричного треугольного импульса, распространяющегося в обычной нелинейной среде и в среде с гистерезисом.

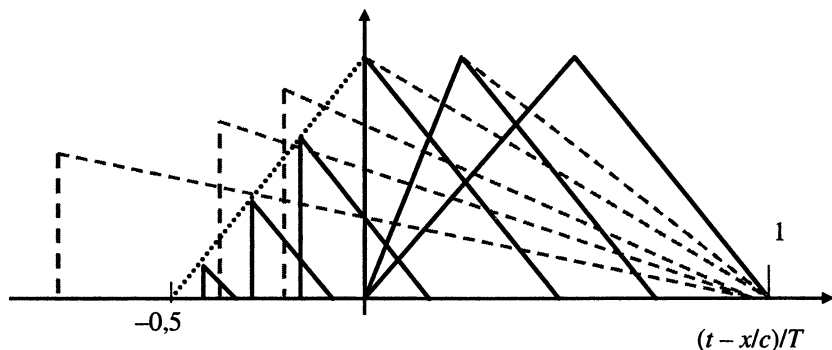


Рис. 11. Искажение профиля однополярного импульса при распространении в гистерезисной среде (сплошные кривые) и в обычной нелинейной среде (штриховые кривые) [12].

Если обычно площадь импульса (количество движения) сохраняется, а длительность возрастает, то в гистерезисной среде уменьшаются как площадь, так и длительность. Волна не может проникнуть в среду на расстояние больше 5—10 нелинейных длин как из-за сильной диссипа-

ции энергии, так и из-за потерь на формирование остаточных деформаций в среде. В примере на рис. 11 волна «исчезает» в точке $-0,5$ на оси абсцисс.

1.2.4. Рассмотрим теперь так называемую «структурную» объемную нелинейность. Этот тип физической нелинейности связан не с нелинейным характером сил межмолекулярного взаимодействия, а с дефектами среды на надмолекулярном (мезоскопическом) уровне. В жидкостях такой тип нелинейности возникает при появлении пузырьков газа; известно, что нелинейный параметр газированной воды может на 2—3 порядка превысить нелинейность каждого из составляющих смеси — как газа ($\epsilon = 1,2$), так и воды ($\epsilon = 3,5$). Гигантская нелинейность пузырьков жидкостей открыла возможность детектирования небольшого числа (и даже одиночных) пузырьков для целей диагностики кессонной болезни при декомпрессии или в различных промышленных технологиях. Эти работы в 80-х годах интенсивно велись в ИПФ РАН [13, 14]. Сейчас большое развитие получили медицинские диагностические технологии, основанные на инъекции в кровеносные сосуды специальных контрастных агентов на основе стабильных микропузырьковых взвесей (см. следующий раздел).

Сильный рост нелинейности в среде, содержащей хорошо сжимаемые включения, можно пояснить простым примером. Пусть тонкий слой легкой сжимаемой среды окружен более плотной средой с большим значением скорости звука (пример: тонкий слой воздуха в воде или водном геле). В задаче есть два малых параметра: отношение акустических импедансов и волновая толщина слоя: $\rho_0 c_0 / \rho_1 c_1 \sim \omega / c_0 h \ll 1$. Решение задачи о прохождении волны через этот слой [15] показывает, что 2-я гармоника на его выходе оказывается в K раз более сильной, когда легкий слой окружен плотной средой (по сравнению со случаем, когда плотной среды нет). Величина «усиления» нелинейности для генерации 2-й гармоники равна

$$K = \frac{P_{2\omega}^{inhg}}{P_{2\omega}^{hg}} = \frac{2}{3} \frac{\epsilon_0}{\epsilon_1} \left(\frac{c_1^2 \rho_1}{c_0^2 \rho_0} \right)^2.$$

Для воздушного слоя в воде $K \approx 5,5 \cdot 10^7$. Если слой состоит из воды с пузырьками, объемная концентрация которых порядка 10^{-4} , усиление будет порядка 5000.

Объяснение гигантскому усилению нелинейности очень простое. Пусть в слое воды распространяется волна с характерным давлением порядка нескольких атмосфер. Это слабая волна, потому что внутреннее давление в воде порядка 23000 атм. Зато когда такое давление начинает действовать на слой воздуха, оно изменяет его объем в несколько раз, то есть приводит к сильно нелинейным деформациям.

В твердых структурно-неоднородных средах (гранулированные, флюидонасыщенные, трещиноватые, пористые и др.) K также достигает $10^2 - 10^3$. Рост K — признак наличия дефектов. Нелинейность гранулированных сред обсуждалась в п. 1.1.3. Другим важным примером сред с большой структурной нелинейностью являются трещиноватые среды. Известно, что приложенное к образцу среды напряжение усиливается на кончике острой трещины в $K = \sqrt{l/r_0}$ раз, где l — длина трещины, r_0 — радиус кривизны ее конца (рис. 12). Когда $r_0 \rightarrow 0$, $K \rightarrow \infty$. При больших напряжениях вблизи конца среда пластически деформируется, и r_0 принимает некоторое конечное значение.

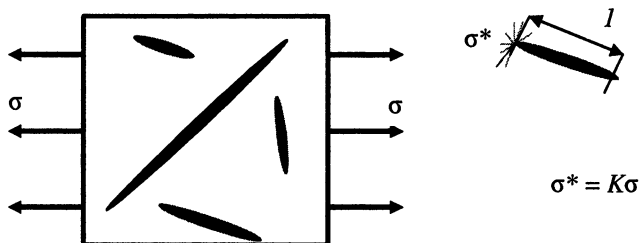


Рис. 12. Вблизи острого конца трещины приложенное извне напряжение усилено в $K \gg 1$ раз.

Прямое подтверждение усиления нелинейных свойств среды с ростом числа трещин дано В. А. Робсманом [16] (рис. 13). Бетонная балка нагружалась на стенде статическим напряжением, растущим вплоть до ее разрушения. С ростом напряжения увеличивалось число трещин и структурная нелинейность среды возрастала. При одних и тех же длине (6 м) и уровне исходного сигнала наблюдалось усиление нелинейных волновых взаимодействий (на рис. 13 как пример показано взаимодействие узкой линии сигнала с широким шумоподобным спектром).

Очевидно, с развитием системы трещин уменьшается прочность материала, поэтому усиление нелинейных эффектов в дефектной среде может служить критерием потери прочности [17].

2. О нелинейной диагностике

Идеи, лежащие в основе нелинейных методов диагностики и неразрушающего контроля, хорошо известны (см., например, [18]). С ростом амплитуд (интенсивностей) волн нарушается принцип суперпозиции: пересекаясь во времени и в пространстве, достаточно сильные волны начинают взаимодействовать, обмениваясь энергией. При этом каждая из волн «запоминает» как характеристики своих «партнеров», так и материальные характеристики среды (в той области, где это взаимодействие произошло). Эта информация может быть «доставлена» к приемни-

ку самими исходными волнами, испытывавшими кросс-модуляцию, а также излучаться непосредственно из области взаимодействия в виде новых спектральных компонент, отсутствовавших в исходном волновом спектре.

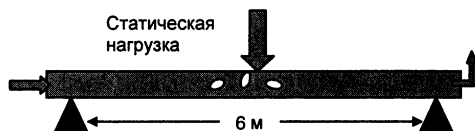
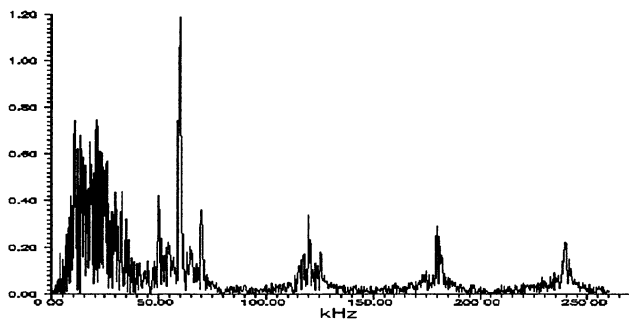
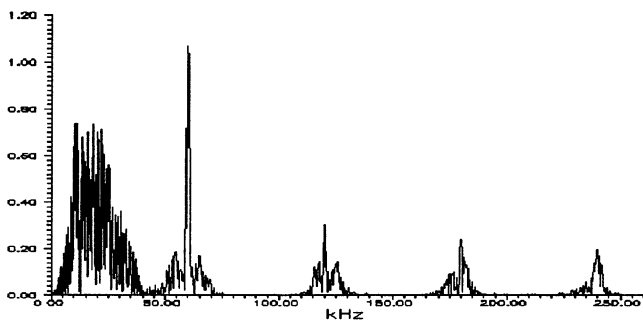
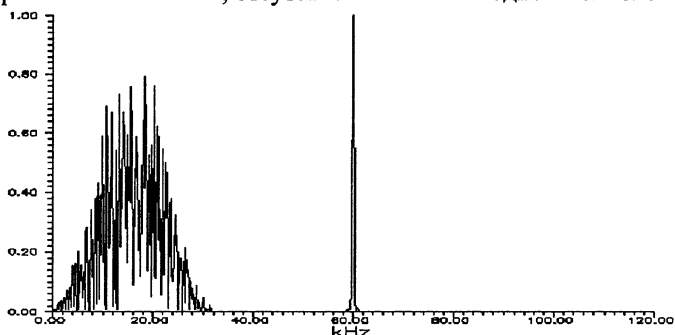


Рис. 13. Нелинейные взаимодействия широкополосного шума и монохромной волны на разных уровнях развития трещин в железобетонной балке: *а* — при отсутствии трещин; *б* — при появлении одиночных несвязанных микротрещин; *в* — при развитии системы связанных трещин и перед внезапным разрушением балки.

Общий подход к восстановлению границ и внутренней структуры нелинейных рассеивателей должен базироваться на методах решения обратных задач — нелинейной акустической томографии [19]. Однако даже простейший способ диагностики, основанный на генерации гармоник и комбинационных частот, весьма эффективен, особенно когда среда содержит сильно нелинейные включения типа пузырьков или трещин. Фактически любое аналитическое решение прямой нелинейной задачи содержит параметры (нелинейные модули, амплитуды волн, геометрические характеристики), которые можно определить экспериментально по измерениям нелинейно искаженного волнового поля или его спектра.

К числу конкретных достижений в этой области следует отнести создание контрастных веществ типа «Albunex» для целей визуализации кровотока и медицинской диагностики. О масштабах работ говорит такой факт: на создание этих веществ и проведение клинических испытаний 10 компаний из США, Германии, Англии, Швейцарии и других стран уже израсходовали более 1 миллиарда долларов. Препарат представляет собой стабильную взвесь микропузырьков микронных размеров, каждый из которых окружен биологически инертной оболочкой. Пузырьки инжицируются в сосудистое русло (рис. 14), после чего их перемещение потоком крови регистрируется как по обычному рассеянию звука, так и по второй гармонике. Большая локальная нелинейность и отсутствие помех от линейных рассеивателей на высших гармониках позволяют следить либо за одиночным пузырьком или небольшой их группой, либо за перемещением фронта облака пузырьков, либо за участком сосудистого русла в целом.

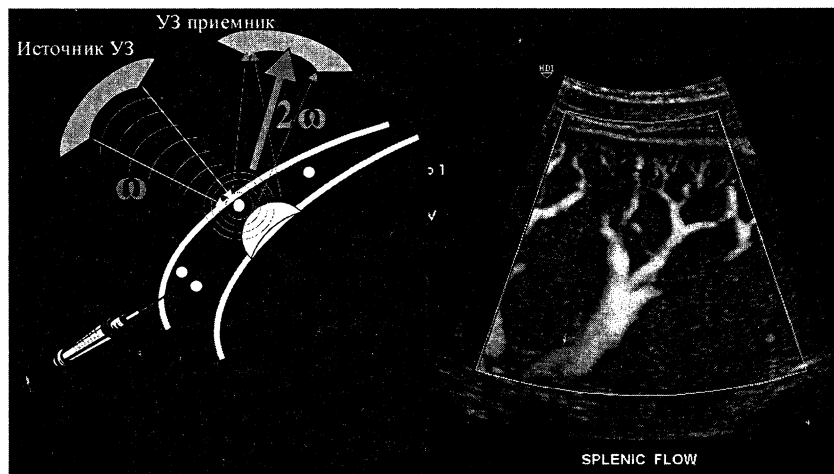


Рис. 14. Визуализация кровотока с помощью инъекции воздушных микропузырьков. Акустические контрастные агенты радикально улучшают ультразвуковую визуализацию кровотока и сосудистого русла.

Аналогичная задача решена при обнаружении дефектов в объеме твердотельных изделий (рис. 15, предоставлен И. Ю. Солодовым). Верхняя часть рис. 15 содержит изображение рельефа стоячих волн на поверхности плиты, колеблющейся на частоте 20 кГц. Внутренних дефектов не видно. Зато на нижней половине рис. 15 отчетливо видна объемная трещина. Ее изображение получено на частоте седьмой гармоники — 140 кГц, где помехи от линейных рассеивателей (в том числе отражения от поверхности тела) отсутствуют.

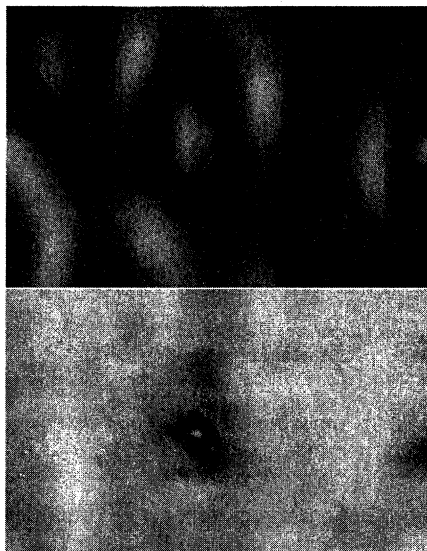
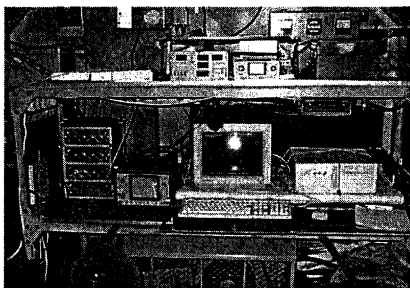


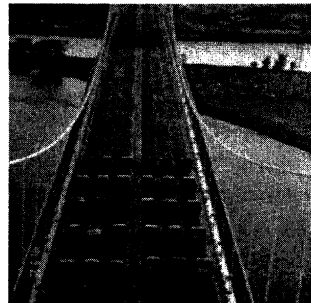
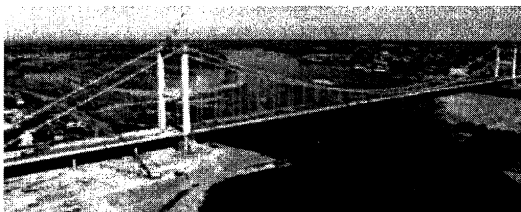
Рис. 15. Визуализация объемной трещины на частоте седьмой гармоники

Среди отечественных достижений в прикладной области необходимо выделить результаты использования методов нелинейной диагностики в строительной индустрии.

Начало этой деятельности относится к концу 80-х годов. Группа В. А. Робсмана (ЦНИИС) работала в Армении в связи с сооружением тоннеля Севан — Арпа. В это время произошло катастрофическое Спитакское землетрясение, повлекшее многочисленные разрушения и жертвы. Ученых попросили провести акустическую диагностику поврежденных зданий и дать заключение о том, какие здания можно восстановить, а какие следует разрушить в целях безопасности. В процессе ультразвукового «просвечивания» ключевых элементов (балок, несущих стен, перекрытий) выяснилось, что акустические спектры претерпевают искажения тем более сильные, чем сильнее повреждена конструкция. Позднее эмпирически установленным признакам было дано объяснение, и в настоящее время «нелинейные» прогнозы стали весьма достоверными.



Виброустановка и измерительная лаборатория мобильного комплекса



Динамические испытания моста через реку Иртыш в г. Семипалатинске автотранспортной нагрузкой

Рис. 16. Аппаратура и примеры испытаний, сопровождаемых использованием методов нелинейной диагностики.

ЦНИИС успешно использовал методы нелинейной диагностики при строительстве сооружений 3-го транспортного кольца в Москве, при реконструкции и реставрации архитектурно-исторических памятников, строительстве метрополитенов, обследовании энергетических установок с целью разработки проектов усиления конструкции и сейсмозащиты, для диагностики дефектов пролетных строений и опор больших и внеклассных мостов (более 30) для оценки фактической грузоподъемности и проектов реконструкции.

3. О нелинейных моделях и явлениях

Число математических моделей, используемых к настоящему времени при решении задач нелинейной акустики, очень велико. Имеет смысл указать только на простейшие «базовые» модельные уравнения с целью обратить внимание на их специфику, а также связь с другими моделями, используемыми в физике нелинейных волн.

Общее одномерное уравнение, описывающее распространение расходящихся и сходящихся волн (рупоры, концентраторы, лучевые трубки в приближении нелинейной геометрической акустики), имеет вид

$$\frac{\partial p}{\partial x} + \frac{p}{2} \frac{d}{dx} \ln S(x) - \frac{\varepsilon}{c_0^3 \rho_0} p \frac{\partial p}{\partial \tau} - \frac{b}{2c_0^3 \rho_0} \frac{\partial^2 p}{\partial \tau^2} = 0.$$

Здесь $S(x)$ — площадь сечения трубки, $\tau = t - x/c$. Важные частные случаи: при $S = \text{const}$ это уравнение переходит в уравнение Бюргера для плоских волн, $S \sim x$ соответствует цилиндрическому, $S \sim x^2$ — сферически-симметричным волнам.

Диссипация (вторая производная) связана с эффектами вязкости и теплопроводности. Для учета потерь за счет рассеяния на мелких неоднородностях используется аналогичное уравнение с 4-й производной [20].

Интегродифференциальное уравнение

$$\frac{\partial p}{\partial x} - \frac{\varepsilon}{c_0^3 \rho_0} p \frac{\partial p}{\partial \tau} - \frac{m}{2c_0} \frac{\partial}{\partial \tau} \int_{-\infty}^{\tau} K(\tau - \tau') \frac{\partial p}{\partial \tau'} d\tau' = 0$$

описывает нелинейные волны в наследственных средах. Важный случай $K(t) = \exp(-t/T)$ соответствует релаксирующей среде, для биотканей используются степенные ядра, для расплавов и сильно вязких жидкостей — модели с несколькими временами релаксации.

Базовым уравнением для интенсивных акустических пучков является уравнение ХЗК (Хохлова — Заболотской — Кузнецова):

$$\frac{\partial}{\partial \tau} \left[\frac{\partial p}{\partial x} - \frac{\varepsilon}{c_0^3 \rho_0} p \frac{\partial p}{\partial \tau} - \frac{b}{2c_0^3 \rho_0} \frac{\partial^2 p}{\partial \tau^2} \right] = \frac{c_0}{2} \Delta_{\perp} p.$$

Здесь Δ_{\perp} — оператор Лапласа, действующий на координаты в поперечном сечении пучка. Общая конструкция уравнения дифрагирующих пучков такова:

$$\frac{\partial}{\partial \tau} [\hat{\Pi}(p)] = \frac{c_0}{2} \Delta_{\perp} p, \quad \hat{\Pi}(p) = 0.$$

Здесь $\Pi(p) = 0$ — соответствующее уравнение для плоских волн (например, любое из первых двух уравнений этого раздела). Разумеется, во всех уравнениях квадратичная нелинейность может быть заменена на нелинейный член более общего вида

$$\gamma \partial f(p) / \partial \tau$$

и даже на интегральный нелинейный член типа (16).

Математическим свойствам и методам анализа модельных уравнений нелинейной акустики имеет смысл посвятить отдельную лекцию. Возможно, это будет сделано в будущем.

Литература

1. Руденко О. В., Солуян С. И. Теоретические основы нелинейной акустики. М.: Наука, 1975.
2. Ландау Л. Д., Лифшиц Е. М. Гидродинамика. М.: Наука, 1986.
3. Руденко О. В. Нелинейные взаимодействия регулярных и шумовых спектров при формировании интенсивного излучения поршнем в линейной среде // Акуст. журн. 1998. Т. 44, № 6.
4. Руденко О. В. Взаимодействия интенсивных шумовых волн // УФН. 1986. Т. 149, № 3.
5. Руденко О. В. Нелинейные колебания линейно деформируемой среды в закрытом резонаторе, возбуждаемые конечными смещениями его границы // Акуст. журн. 1999. Т. 45, № 3.
6. Ландау Л. Д., Лифшиц Е. М. Теория упругости. М.: Наука, 1986.
7. Rudenko O. V., Hedberg C. M. Nonlinear dynamics of grains in a liquid-saturated soil // Nonlinear Dynamics. 2004. V. 35. P. 187—200.
8. Руденко О. В., Чинь Ань Ву. Нелинейные акустические свойства контакта шероховатых поверхностей и возможности акустодиагностики статистических характеристик неровностей // Акуст. журн. 1994. Т. 40, № 4.
9. Руденко О. В., Хирных К. Л. Модель резонатора Гельмгольца для поглощения интенсивного звука // Акуст. журн. 1990. Т. 36, № 3.
10. Работнов Ю. Н. Элементы наследственной механики твердых тел. М.: Наука, 1977.
11. Ostrovsky L. A., Johnson P. A. Dynamic nonlinear elasticity in geomaterials // La Rivista del Nuovo Cimento. 2001. V. 24, № 7.
12. Руденко О. В., Робсман В. А. Нелинейные процессы в средах с акустическим гистерезисом и проблемы динамического взаимодействия свай и грунтового основания // Акуст. журн. (в печати).
13. Мансфельд А. Д., Рейман А. М. Особенности обнаружения газовых пузырьков в неоднородных нелинейных средах // Ультразвуковая диагностика. Горький: ИПФ РАН, 1983. С. 151—161.
14. Островский Л. А., Сутин А. М. Нелинейные акустические методы диагностики газовых пузырьков в жидкости // Ультразвуковая диагностика. Горький: ИПФ РАН, 1983. С. 139—150.
15. Hedberg C. M., Rudenko O. V. Pulse response of nonlinear layer // J. Acoust. Soc. Am. 2001. V. 110, № 5, Pt. 1. P. 2340—2350.
16. Робсман В. А. Трансформация акустических спектров в неоднородных твердых средах при нелинейной деформации // Акуст. журн. 1992. Т. 38, № 1. С. 129—143.
17. Shkolnik I., Zarebo L., Krasilnikov V. On structural nonlinear diagnostics of solids and estimation of their strength // Frontiers of Nonlinear acoustics / Eds. M. Hamilton and D. Blackstock. Elsevier, London: 1990. P. 589—594.
18. Rudenko O. V. Nonlinear methods in acoustic diagnostics // Rus. J. Nondestructive Testing. 1993. V. 29, № 8. P. 583—589.
19. Burov V. A., Gurinovich I. E., Rudenko O. V., Tagunov T. Ya. Reconstruction of the spatial distribution of the nonlinearity parameter and sound velocity in acoustic nonlinear tomography // Acoust. Phys. 1994. V. 40, № 6. P. 816—823.
20. Руденко О. В., Робсман В. А. Уравнение нелинейных волн в рассеивающей среде // ДАН. Физика. 2002. Т. 384, № 6. С. 434—439.

СТАТИСТИЧЕСКАЯ НЕЛИНЕЙНАЯ АКУСТИКА. ВОЛНЫ И СТРУКТУРЫ В СРЕДАХ БЕЗ ДИСПЕРСИИ

С. Н. Гурбатов, А. И. Саичев

Из статьи О. В. Руденко видно, что набор математических моделей, применяемых в статистической нелинейной акустике, очень велик. Но даже в рамках «базовых» модельных уравнений (Римана и Бюргерса) анализ статистики случайных волн весьма сложен. В этой лекции мы обсудим постановку статистических задач нелинейной акустики, а также рассмотрим некоторые ключевые решения уравнения Бюргерса, помогающие лучше понять основные законы эволюции нелинейных случайных волн [1].

1. Статистическая нелинейная акустика и турбулентность Бюргерса

Напомним основной эффект, присущий интенсивным шумовым акустическим волнам. На первых нелинейных школах часто вывешивали плакаты с условиями синхронизма $\omega_1 + \omega_2 = \omega_3$, $k_1 + k_2 = k_3$, а на лекциях демонстрировали — какая тройка волн эффективней взаимодействует в данной среде. В акустике отыскать подобные волны очень легко: дисперсионное уравнение звуковых волн $\omega = ck$ означает, что условия синхронизма выполняет любая тройка параллельно бегущих волн. В итоге для гармонического входного сигнала идет лавинообразный рост числа высших гармоник $\omega_n = n\omega_1$. Причем на начальной стадии амплитуда n -й гармоники $A_n \sim A_1^{n-1} z^{n-1}$, где A_1 — амплитуда входного сигнала, а z — расстояние от излучателя.

Особенности поведения случайно-модулированного квазигармонического сигнала на указанной стадии легко объяснить с помощью очевидного статистического неравенства $\langle A^2 \rangle > \langle A \rangle^2$. Согласно ему нелинейность сильнее реагирует на большие случайные выбросы, а нелинейные процессы у шумового сигнала протекают быстрее, чем у гармонической волны. Так, при гауссовой статистике квазигармонического шума интенсивность n -й гармоники в $n!$ раз больше, чем у гармонического сигнала той же начальной интенсивности. Этот факт наблюдался в экспериментах [2] по распространению узкополосного акустического шума интенсивностью 200 Па в заполненной воздухом 75-метровой трубе.

Лавинообразный рост числа высших гармоник ведет к появлению разрывов, на которых энергия волны диссипирует. При этом гар-

монический сигнал трансформируется в последовательность пилообразных импульсов с амплитудой $A(z) \sim A_1/(1+\beta z A_1 \omega_1)$. На большом расстоянии z от входа предельная амплитуда $A_*(z) \sim 1/\beta z \omega_1$ зависит лишь от частоты ω_1 и параметра нелинейности β . Таким образом, для амплитудно-модулированной волны $A_1 = A_1(t)$ идет подавление больших выбросов, а значит, вероятностное распределение шума со временем сужается, обращается в нуль вне интервала $[-A_*, A_*]$ и стремится к равномерному шириной $2A_*(z)$, не зависящей от начальной амплитуды и пропорциональной периоду волны.

При решении статистических волновых проблем обычно интегрируются эволюцией спектра волны $E_0(\omega) \rightarrow E_0(\omega, z)$. При конечной ширине начального спектра $\Delta\omega_1$ ширина спектра n -й гармоники растет с номером гармоники. Причем, в зависимости от вида исходной фазовой модуляции, $\Delta\omega_n \simeq n\Delta\omega_1$ или же $\Delta\omega_n \simeq n^{1/2}\Delta\omega_1$. При $\Delta\omega_n \simeq \omega_1$ спектры соседних гармоник перекрываются и формируют сплошной спектр $\sim \omega^{-2}$.

На разрывной стадии фазовая модуляция исходной волны порождает флуктуации положений разрывов, и как следствие случайную модуляцию амплитуды пилообразной волны. На рис. 1 показана эволюция спектра интенсивного квазигармонического шума в заполненной воздухом трубе [3] с центральной частотой 1,5 кГц, шириной спектра 316 Гц и интенсивностью 150 дБ. Здесь кривая 1 — спектр вблизи излучателя $z = 0,35$ м, кривая 2 — на расстоянии 10,35 м (изображена с усилением 20 дБ). Точками указаны амплитуды тонального сигнала той же интенсивности при 10,35 м.

Качественное отличие поведения квазигармонического сигнала от тонального состоит в генерации разностных комбинационных частот. Хотя эффективность генерации низкочастотных компонент мала, они затухают медленнее высокочастотных и формируют асимптотические свойства поля на больших временах. Амплитуда низкочастотной компоненты примерно равна $v_n \simeq (\Delta\omega_1/\omega_1)A_1$. На разрывной стадии ее появление вызвано движением пилообразных фронтов, ведущим к дополнительному уширению спектральных линий высокочастотных гармоник. Из-за быстрого затухания высокочастотных компонент энергия волны постепенно сосредоточивается в низкочастотной части спектра, а ее затухание начинается, когда разрывы пилообразной волны будут сливаться. Это происходит на больших расстояниях $z \sim z_n(\omega_1/\Delta\omega_1)^2 A_1$, так как скорости разрывов сильно коррелированы.

Для эволюции широкополосного шума характерны как генерация высокочастотных компонент, вызванная появлением разрывов,

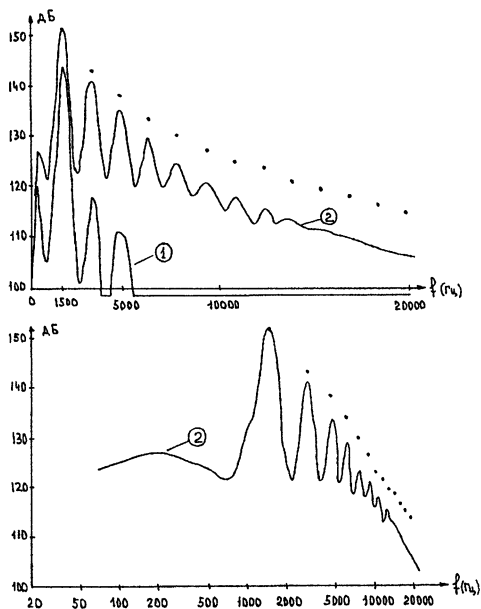


Рис. 1. Эволюция спектра квазимонохроматического шума

так и возникновение низкочастотной части спектра. Разрывы при этом слабо коррелированы, а постепенное слияние мелких разрывов ведет к росту характерного масштаба шума и смещению спектра в сторону низких частот. Из-за появления разрывов теряется информация о тонкой структуре входного сигнала, а из-за многократного слияния разрывов спектр приобретает универсальный вид, фактически не зависящий от формы начального спектра. Эти нелинейные явления объясняют эффект anomalно сильного высокочастотного «крыла», формирующегося в атмосфере при распространении шума мощных авиационных и ракетных двигателей.

При взаимодействии шумов и регулярных сигналов появление разрывов также кардинально меняет спектральный состав волны. В частности, шумовые компоненты возникают в очень широком диапазоне частот [4] (рис. 2). Кроме того, при взаимодействии интенсивной высокочастотной гармонической волны и низкочастотного шума последний дает лишь квазистатический сдвиг фазы ВЧ-сигнала, пропорциональный НЧ-сигналу, частоте гармоники и расстоянию от

входа. Иными словами, спектр шумовой компоненты n -й гармоники повторяет спектр низкочастотного шума.

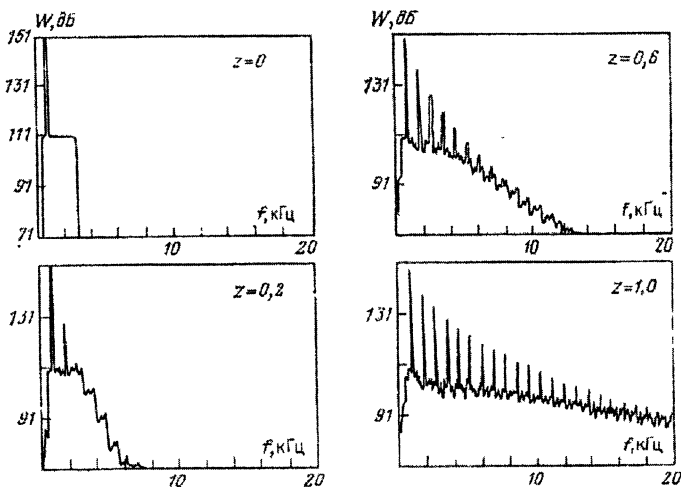


Рис. 2. Взаимодействие шума и тонального сигнала

На разрывной стадии, когда амплитуды гармоник обратно пропорциональны $zn\omega_1$, амплитуда спектра шумовой компоненты n -й гармоники не зависит от расстояния z и равна амплитуде НЧ-шума для большого числа гармоник. Этот эффект естественно назвать эффектом аномального усиления шума.

Подчеркнем еще раз: для акустических волн дисперсия пренебрежимо мала в широком диапазоне частот. В средах без дисперсии эффективное взаимодействие большого числа гармоник сопровождается сильной когерентностью их фаз. Поэтому здесь неприменим метод медленно меняющихся амплитуд, полезный в нелинейной оптике, где из-за сильной дисперсии достаточно ограничиться учетом взаимодействия малого числа линейных мод. Сильная когерентность спектральных компонент не позволяет также применить приближение хаотических фаз, используемое в теории слабой турбулентности, где описание взаимодействия большого числа линейных мод удается свести к кинетическим уравнениям для интенсивности мод.

Одним из базовых уравнений нелинейной акустики служит уравнение Бюргерса. В статье О. В. Руденко дано общее одномерное уравнение, описывающее эволюцию расходящихся и сходящихся волн (рупоры, концентраторы, лучевые трубки в приближении нелинейной

геометрической акустики). Распространение плоской нелинейной волны описывается при этом уравнением

$$\frac{\partial P}{\partial z} - \beta P \frac{\partial P}{\partial \tau} = \mu \frac{\partial^2 P}{\partial \tau^2}, \quad P(\tau, r = 0) = P_0(\tau). \quad (1)$$

В теории турбулентности часто используют также нелинейное уравнение диффузии:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = \nu \frac{\partial^2 v}{\partial x^2}, \quad v(x, t = 0) = v_0(x). \quad (2)$$

Оно носит имя Бюргерса, предложившего это уравнение в качестве модельного уравнения сильной гидродинамической турбулентности [5, 6], а решение этого уравнения со случайными начальными условиями называют турбулентностью Бюргерса (иногда даже [7]). В самом деле, уравнение (2) имеет много общего с известным уравнением Навье — Стокса: тот же тип нелинейности, общие инварианты, одинаковая частотная зависимость энергетических потерь и т. д. [7]. Заметим однако, что различия уравнений Бюргерса (2) и Навье — Стокса столь же поучительны, как и их подобие. Впоследствии выяснилось, что уравнение (2) описывает целый ряд нелинейных волновых явлений в акустике, физике плазмы и т. д. [8, 9, 10, 11]. Простой заменой переменных «акустическое» уравнение (1) сводится к уравнению (2). Поэтому решения уравнения Бюргерса со случайными начальными условиями называют одномерной акустической турбулентностью, так как большая часть результатов по турбулентности Бюргерса легко переносится в статистическую нелинейную акустику.

Заметим еще, что многомерное уравнение Бюргерса с внешними случайными силами широко используют для моделирования гидродинамической турбулентности Навье — Стокса без давления. Трехмерный вариант уравнения (2) $\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \mu \nabla^2 \mathbf{v}$ вместе с уравнением для плотности привлекают к анализу крупномасштабных структур Вселенной на нелинейной стадии гравитационной неустойчивости, когда силы давления пренебрежимо малы. Данная модель, известная астрофизикам как модель «слипания» (adhesion model) [12, 13, 14], описывает формирование сильно неоднородных структур в распределении вещества из случайного начального возмущения.

Другие проблемы, приводящие к многомерному уравнению Бюргерса или его вариантам, — это рост границы раздела двух сред из-за случайного осаждения вещества на поверхности, распространение фронтов пламени [11, 15] и оптических волн.

Продемонстрируем происхождение уравнения Бюргерса в теории роста поверхностей и движения волновых фронтов, взяв некоторую

линию $z = h(x, t)$. Будем полагать, что ее участки движутся перпендикулярно линии со скоростью c . На рис. 3 показаны фрагменты линии $h(x, t)$ в близкие моменты времени t и $t + \Delta$.

Из рисунка видно, что приращение высоты линии в произвольной выбранной точке $\Delta h = h(x, t + \Delta) - h(x, t) \approx c\Delta / \cos \theta$ обратно пропорционально $\cos \theta$. Отсюда сразу следует, что рост достаточно пологой поверхности подчиняется уравнению

$$\frac{\partial h}{\partial t} = c \sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2} \approx c + \frac{c}{2} \left(\frac{\partial h}{\partial x}\right)^2. \quad (3)$$

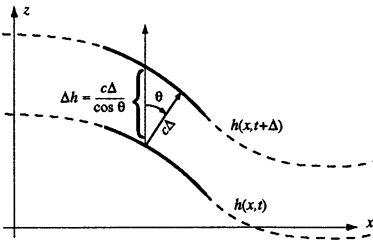


Рис. 3. Иллюстрация к росту поверхности.

ся также из-за эффектов поверхностного натяжения, учитываемых диффузионным слагаемым. Кроме того, берут во внимание неоднородность потока падающих на линию $z = h(x, t)$ частиц, добавив в (3) случайное слагаемое $F(x, t)$. В итоге уравнение растущей линии примет вид

$$\frac{\partial \psi}{\partial t} = \frac{1}{2} \left(\frac{\partial \psi}{\partial x}\right)^2 + \nu \frac{\partial^2 \psi}{\partial x^2} + F(x, t). \quad (4)$$

Мы пришли к двумерному варианту известного уравнения Kardar – Parisi – Zhang (KPZ-уравнение) [16, 15, 11]. Однородному (при $F \equiv 0$) KPZ-уравнению удовлетворяет потенциал поля скорости многомерного уравнения Бюргерса ($\mathbf{v} = -\nabla \psi$). Среднеквадратичный градиент $E(t) = \langle \mathbf{v}^2(\mathbf{x}, t) \rangle$ характеризует при этом изрезанность поверхности и может как убывать, так и возрастать с течением времени.

Исследованию динамики и статистики решений одномерного, а в последние годы и многомерного уравнения Бюргерса посвящено

много работ (см., например, библиографию книг [10, 11]). Хотя уравнение Бюргерса имеет точное решение — решение Хопфа — Коула [17, 18], анализ его статистики остается весьма сложной математической задачей. Так, первые серьезные результаты для броуновского начального потенциала были получены лишь 30 лет спустя [6] после появления самого уравнения [5], а точное статистическое описание этого специального случая сделано совсем недавно [19]. Мы обсудим ниже некоторые свойства решений уравнения Бюргерса, помогающие лучше понять законы эволюции шумовых сигналов.

2. Асимптотические решения уравнения Бюргерса

Будем использовать ниже классическую форму уравнения Бюргерса (2). При этом, переходя к акустическим задачам, надо помнить, что в (2) время t пропорционально расстоянию z до излучателя, координата x пропорциональна бегущему времени τ , а волновое число k эквивалентно частоте ω .

Уравнение Бюргерса имеет точное решение — решение Хопфа — Коула [17, 18]:

$$v(x, t) = \frac{\partial}{\partial x} \psi(x, t) = -2\mu \frac{\partial}{\partial x} \ln \varphi(x, t) = -\frac{2\nu}{\varphi} \frac{\partial \varphi}{\partial x},$$

где φ подчиняется линейному уравнению диффузии

$$\frac{\partial \varphi}{\partial t} = \nu \frac{\partial^2 \varphi}{\partial x^2}, \quad \varphi_0(x) = \exp\left(-\frac{\psi_0(x)}{2\nu}\right), \quad \psi_0(x) = \int^x v_0(x') dx'.$$

В пределе исчезающе малой вязкости $\nu \rightarrow 0$ применение метода перевала к решению Хопфа — Коула дает так называемый *принцип абсолютного максимума* [17, 10, 14]:

$$\psi(x, t) = \max_y \Phi(x, y, t), \quad (5)$$

$$\Phi(x, y, t) = \psi_0(y) - \frac{(x - y)^2}{2t}, \quad (6)$$

$$v(x, t) = \frac{x - y(x, t)}{t} = v_0(y(x, t)). \quad (7)$$

Здесь $\psi_0(y)$ — начальный потенциал: $v_0(x) = -\nabla \psi_0(x)$. Функция $y(x, t)$ в (7) есть лагранжева координата точки, где потенциал $\Phi(x, y, t)$ при данных эйлеровой координате x и времени t достигает абсолютного максимума. Заметим, что y совпадает с начальной координатой

частицы, попадающей в точку x в момент t [10]. Таким образом, решение уравнения (2) сводится к отысканию координат абсолютных максимумов, что позволяет в статистических задачах привлекать теорию выбросов случайных процессов [20]. На данном принципе основан эффективный численный алгоритм решения уравнения Бюргера, известный как быстрое преобразование Лежандра [21].

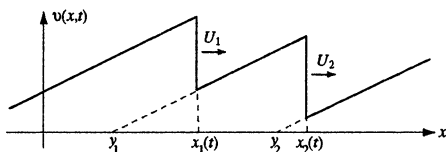


Рис. 4. Скелет пилообразной волны. Указаны «нули» линейных участков и координаты разрывов.

Из приведенного решения уравнения Бюргера видно, что на больших временах непрерывное поле трансформируется в последовательность ячеек с универсальным поведением поля внутри них: $v(x, t) = (x - y_k)/t$ (рис. 4). Известно еще, что решение уравнения Бюргера при $\nu \rightarrow 0$ эквивалентно потоку неупру-

го слипающихся частиц [10]. При этом из газа «легких» частиц постоянной начальной плотности образуются «тяжелые» макрочастицы, чьи координаты и скорости совпадают с положением и скоростью разрывов. Следовательно, скорость разрыва $v^*(t) = dx^*(t)/dt$ определяется законом сохранения импульса $v^*(t) = p^*(t)/m^*(t)$, где $m^*(t) = \int_{y^-(t)}^{y^+(t)} \rho_0 dy$ — масса, а $p^*(t) = \int_{y^-(t)}^{y^+(t)} \rho_0 v_0(y) dy$ — импульс макрочастицы. Законы слияния разрывов

$$V_{k,n} = \frac{\psi_{0,n} - \psi_{0,k}}{y_n - y_k} = \frac{M_{k,m} V_{k,m} + M_{m,n} V_{m,n}}{M_{k,m} + M_{m,n}}$$

тождественны законам сохранения масс $M_{k,m}$ и импульсов $M_{k,m} V_{k,m}$ при абсолютно неупругом столкновении (рис. 5).

Обсудим так называемый «поздний предел» решения уравнения Бюргера, когда при фиксированном коэффициенте вязкости $\nu \neq 0$ время t стремится к бесконечности. Рассмотрим класс возмущений с ограниченным начальным потенциалом $\langle \psi_0(x)^2 \rangle < \infty$, полагая $\psi_0(x)$ периодическим или статистически однородным случайным полем. Для подобных сигналов в скалярном поле $\varphi(x, t)$ удастся выделить постоянную компоненту $\bar{\varphi}$ и относительное возмущение $u(x, t)$:

$$\varphi(x, t) = \bar{\varphi} + \tilde{\varphi}(x, t) = \bar{\varphi} [1 + u(x, t)]. \quad (8)$$

С ростом времени диффузия ведет к сглаживанию неоднородностей решения $\varphi(x, t)$ линейного уравнения диффузии и, следовательно, к уменьшению амплитуды возмущения $u(x, t)$. На временах, когда

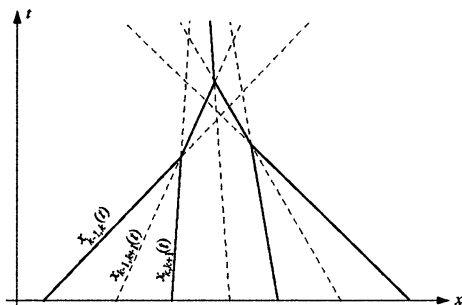


Рис. 5. Узор траекторий сливающихся разрывов. Сплошными линиями обозначены траектории действительно существующих разрывов.

($|u| \ll 1$), решение Хопфа — Коула можно линеаризовать, что дает

$$v(x, t) = -2\nu \frac{\partial u}{\partial x}. \quad (9)$$

Так как $u(x, t)$ подчиняется линейному уравнению диффузии, то и $v(x, t)$ при этом также описывается линейным уравнением. Последнее означает наступление линейной стадии эволюции. Суммарные нелинейные эффекты присутствуют в решении лишь из-за нелинейной интегральной связи между начальным полем $v_0(x)$ и полями $\tilde{\varphi}(x, 0)$, $\tilde{\varphi}$, а также начальным числом Рейнольдса $Re_0 \sim |\psi_0|/\nu$.

3. Эволюция основных типов возмущений

Для одномасштабных полей $v_0(x) = U u_0(x/\ell)$ удобно ввести число Рейнольдса $Re = U\ell/\nu = t_{dis}/t_{nl}$, где $t_{nl} = \ell/U$ — нелинейное время, а $t_{dis} = \ell^2/\nu$ — время диссипации. Изменение относительной роли нелинейных и диссипативных эффектов учитывается при этом текущим числом Рейнольдса $Re(t) = U(t)\ell(t)/\nu = \ell(t)/\delta(t)$, где $\ell(t)$ — внешний, а $\delta(t)$ — внутренний масштабы нелинейного поля $v(x, t)$. Обсудим эволюцию основных классов возмущений в рамках одномерного уравнения Бюргера.

Исходный гармонический сигнал $u_0(x) = U \sin(2\pi x/\ell_0)$ при $Re \gg 1$ трансформируется на больших временах $t \gg t_{nl}$ в пилообразную волну с градиентом $\partial v/\partial x = 1/t$. На этой стадии амплитуда $U(t) = \ell_0/t$ и энергия $E(t) = \ell_0^2/12t^2$ поля не зависят от начальной амплитуды. Из-за уширения ударного фронта число Рейнольдса

уменьшается $Re(t) \sim Re_0(t_{nl}/t)$, а при $t_{lin} = t_{dis} = t_{nl}Re$ начинается линейная стадия эволюции.

Произвольный однополярный импульс стремится к универсальному автомодельному решению (рис. 6):

$$v(x, t) = \sqrt{\frac{\nu}{\pi t}} \mathcal{V}_{Re} \left(\frac{x}{2\sqrt{\nu t}} \right).$$

Характерный масштаб и характерная амплитуда импульса при этом меняются как $l(t) \approx \sqrt{2\nu t}$, $U(t) \approx l(t)/t \sim t^{-1/2}$.

Для сложных сигналов имеются ситуации, когда число Рейнольдса может даже расти с течением времени. Обсудим, к примеру, эволюцию двух разнесенных однополярных импульсов одинаковой площади m . На начальной стадии оба они описываются упомянутым автомодельным решением с числом Рейнольдса $Re_0 = m/\nu$. После слияния решение снова становится автомодельным, но число Рейнольдса возрастает в 2 раза (рис. 7).

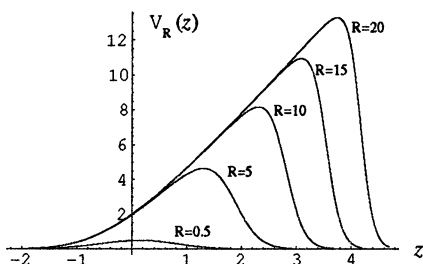


Рис. 6. График функции $\mathcal{V}_{Re}(z)$ в зависимости от z и при разных значениях $Re = m/\nu$, где m — площадь начального импульса.

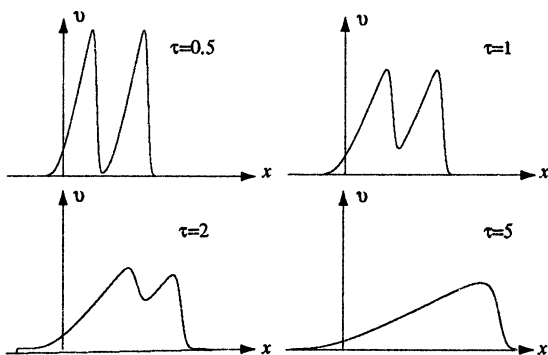


Рис. 7. Слияние треугольных волн, $Re(\tau) = 2Re_0!$

Принято считать, что диссипация ведет к уменьшению роли нелинейных эффектов, а нелинейная эволюция завершается линейной стадией. Однако возможна и обратная последовательность, когда линейная стадия эволюции сменяется нелинейной. Пусть начальное поле имеет вид последовательности N разнесенных импульсов с начальным числом $Re_0 \ll 1$. Из-за линейной диффузии импульсы расплываются, а затем сливаются в один импульс с числом $Re(t) = NRe_0$. Если при этом $NRe_0 \gg 1$, то последующая эволюция носит ярко выраженный нелинейный характер.

Пусть теперь начальное поле состоит из двух однополярных импульсов с площадями $-m$ у левого и m у правого. Тогда начальный потенциал $\psi_0(x)$ локализован и имеет единственный максимум $\psi_{\max} = m$ в некой центральной точке $x = y_0$. При $\nu \rightarrow 0$ поле трансформируется в так называемую N -волну [8] с градиентом $\partial v / \partial x = 1/t$ и положением разрыва $|x_s - y_0| = \ell(t) = (2mt)^{\frac{1}{2}}$. Благодаря росту характерного размера поле спадает по закону $x_s(t)/t \sim m^{\frac{1}{2}}t^{-\frac{1}{2}}$, а энергия уменьшается как $x_s^3/t^2 \sim m^{\frac{3}{2}}t^{-\frac{1}{2}}$ (рис. 8). Таким образом, асимптотическое поведение указанного поля на нелинейной стадии определяется лишь значением максимума $\psi_{\max} = m$ начального потенциала и не зависит от исходной пространственной структуры.

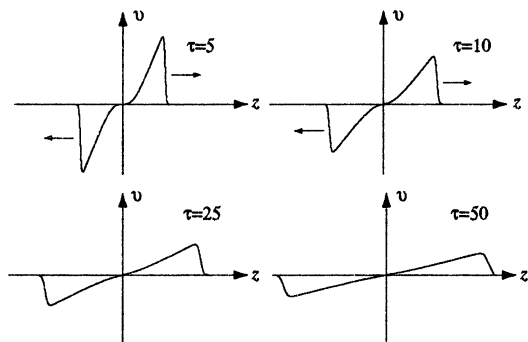


Рис. 8. Решение уравнения Бюргера типа N -волны, с $Re_0 = 25$. Здесь $Re(t) \sim 1/\ln(t)$, а время выхода на линейную стадию $t_l = t_{dis} 10^{10}$.

Для больших, но конечных чисел $Re_0 = m/2\nu \gg 1$ разрывы имеют конечную ширину $\delta = \nu t / \ell(t)$, и, казалось бы, число Рейнольдса каждого импульса должно сохраняться. Однако из точного решения видно, что координата разрыва равна $x_s(t) = [2t(m - \nu \ln(4\pi\nu t / L_{ef}^2))]^{\frac{1}{2}}$, где $L_{ef}^2 \sim L_0^2 / Re_0$, и растет медленнее чем при $\nu \rightarrow 0$. В итоге волна все же выходит на линейную стадию, но время

выхода чрезвычайно велико. Из (9) легко получить, что на линейной стадии ($t \gg L_0^2 \exp(2\text{Re}_0)/\text{Re}_0\nu$) поле $v(x, t)$ имеет универсальную структуру

$$v(x, t) = B \frac{x}{\sqrt{4\pi\nu t^3}} \exp\left(-\frac{x^2}{4\nu t}\right), \quad (10)$$

где постоянная $B \simeq L_0(2\pi/\text{Re}_0)^{\frac{1}{2}} \exp(\text{Re}_0)$.

Для локализованных волн со сложной случайной внутренней структурой форма волны также асимптотически универсальна [22]. При этом нелинейные эффекты генерируют среднее поле (крупномасштабную структуру) из начальной волны с нулевым средним.

4. Эволюция сложных сигналов

Затухание поля качественно иное для сигналов, имеющих фрактальную структуру [23]. Пусть, к примеру, есть N однополярных импульсов с периодом L_1 и данная структура воспроизводится N раз с периодом $L_2 = \beta L_1$. Тогда число импульсов в масштабе L_2 равно N^2 . Повторим процедуру M раз, взяв $L_i = \beta L_{i-1}$, $\beta = (N-1)\gamma$ и $\gamma > 1$. В масштабе L_M длительность импульса равна $T_M = L_M(N-1)$, а число импульсов последовательности N^M . При $M \gg 1$ можно ввести фрактальную размерность такого сигнала как $D = \lim(\ln R / \ln L)$, где $R = N^i$ — число импульсов в интервале $L = L_1 N \beta^{i-1}$ и

$$D = \ln N / \ln[(N-1)\gamma].$$

Если $(N-1)\gamma = N$, то $D = 1$ и мы имеем периодический сигнал с периодом L_1 . Если же внешний период $T_M = 1$, а $M \rightarrow \infty$, то легко видеть, что последовательность координат импульсов образует канторовское множество. Так, если $N = 2$, $\beta = 3$, то $D = \ln 2 / \ln 3$.

Обсудим эволюцию сигнала в наименьшем масштабе L_1 . Число Рейнольдса одного импульса равно $\text{Re}_0 = m/2\nu$. Если $\text{Re}_0 \ll 1$, то имеем линейную стадию, ширина каждого импульса растет как $l(t) = (2\nu t)^{1/2}$. На временах, когда $l(t) \gg L_1 N$, все N импульсов из-за *линейной диффузии* сольются в один с числом Рейнольдса $\text{Re}_1 = mN/2\nu = N\text{Re}_0 > \text{Re}_0$. Таким образом, вследствие линейной диффузии число Рейнольдса растет, линейная стадия сменяется сильно нелинейной, и справедлив анализ в пределе $\nu \rightarrow 0$. В наименьшем масштабе каждый отдельный импульс превращается в треугольный длиной $x_s(t) = (2mt)^{1/2}$, а его энергия затухает как $E(t) = x_s^3/3t^2 = (2m)^{3/2}t^{1/2}$. Взаимодействие импульсов начнется при

$t > t_0$, где $t_0 = L_1^2/2m$ находят из условия $x_s(t) = L_1$. После слияния всех разрывов в наименьшем масштабе N импульсов превращаются в один треугольный импульс длиной $x_s(t) = (2Nmt)^{1/2}$, а его энергия в $N^{1/2}$ раз больше энергии N исходных импульсов. В момент $t_1 = L_2 N^2/2m = t_0 \beta^2/N$ начинается слияние импульсов во втором масштабе. В итоге эволюция поля логарифмически периодически автомодельна, а именно, в интервале времени $t \in [t_0(\beta^2/N)^i, t_0(\beta^2/N)^{i+1}]$ поле повторяет эволюцию в интервале $t \in [t_0, t_0 \beta^2/N]$ (рис. 9). Усредненный закон затухания энергии (рис. 10) определяется при этом фрактальной размерностью сигнала D :

$$E(t) = C_E(t/t_0)^{-p}, \quad p = (1 - D)/(2 - D), \quad 0 < p < 1/2.$$

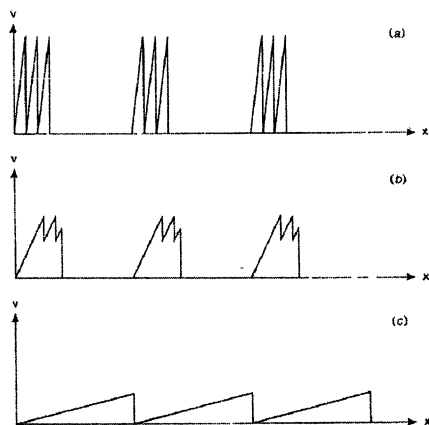


Рис. 9. Эволюция «канторовского» множества

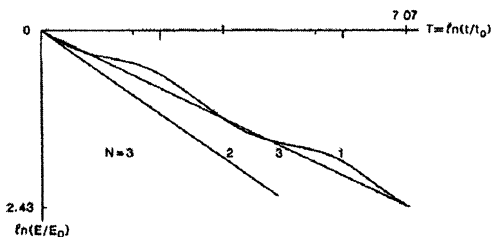


Рис. 10. Затухание энергии «канторовского» множества — кривая 1; кривая 2 — усредненный закон затухания, кривая 3 — затухание одиночного импульса.

Второй пример связан с затуханием *многомасштабных* сигналов [27], когда начальное поле имеет вид

$$v_0(x) = \sum_{p=0}^{\infty} a_p A_{\text{inv}}(k_p x),$$

где волновые числа мод k_p образуют спектр Вейерштрасса (аккумулируются в нуле), а амплитуды берутся равными $a_p^2 = E(k_p) \Delta_p$, $E(k) \sim k^n$. Здесь $E(k)$ — спектр энергии непрерывного сигнала

$$\begin{aligned} k_p &= k_0 \epsilon^p, & a_p &= a_0 \epsilon^{-hp}, \\ h &= -(n+1)/2, & a_0 &= \alpha k_0^{(n+1)/2}. \end{aligned}$$

В качестве мод выберем периодическую инверсную пилообразную функцию $A_{\text{inv}}(x)$, особенностью которой является образование разрыва конечной амплитуды, что эквивалентно слиянию всех легких частиц в тяжелые в один и тот же момент времени (рис. 11).

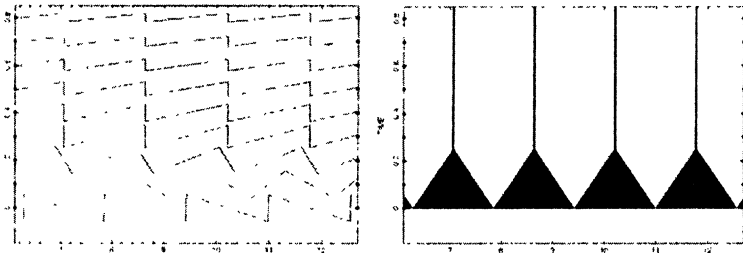


Рис. 11. Эволюция инверсной моды и слияние легких частиц в кластеры

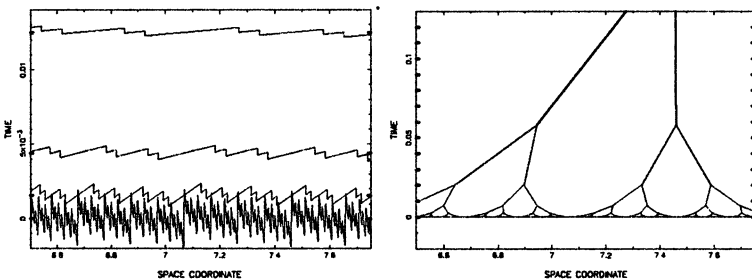


Рис. 12. Поле в моменты подобия и траектории кластеров (разрывов)

Для такой волны можно построить точное решение. Ему так же присуща логарифмически периодическая автомодельность, когда

эволюция поля в разных временных масштабах $[\tilde{t}_q, \tilde{t}_{q+1}]$ и $[\tilde{t}_p, \tilde{t}_{p+1}]$ подобна, с соответствующим растяжением пространственной $x_p/x_q = \epsilon^{p-q}$ и временной $t_p/t_q = (\epsilon^{-\frac{n+3}{2}})^{p-q}$ шкал. При этом усредненный масштаб поля (масса частиц) растет как $L(t) = (\alpha t)^{2/(3+n)}$ (рис. 12).

5. Акустическая турбулентность

Непрерывные случайные поля также трансформируются в набор ячеек с универсальным поведением поля внутри них: $v(x, t) = (x - y_k)/t$, но случайными положениями разделяющих ячейки разрывов. Слияние ячеек увеличивает внешний масштаб турбулентности $L(t)$ (средней массы макрочастицы), поэтому энергия поля $E(t) \sim L^2(t)/t^2$ спадает медленнее энергии периодического сигнала.

Темп слияния разрывов определяется свойствами случайных скоростей разрывов, которые, в свою очередь, зависят от начального потенциала $\psi_0(x)$. На больших временах после многократного слияния разрывов свойства турбулентности задаются статистикой приращений $\psi_0(x + L) - \psi_0(x)$ на больших расстояниях L , а при гауссовой статистике начального поля — асимптотикой структурной функции начального потенциала. Иными словами, сценарий развития турбулентности связан с поведением крупномасштабной части начального спектра $E_0(k)$ поля скорости

$$E_0(k) = \alpha^2 k^n b_0(k), \quad E_\psi(k) = E_0(k)/k^2 \sim \alpha^2 k^{n-2}. \quad (11)$$

Здесь $E_\psi(k)$ — спектр потенциала, а $b_0(k)$ — быстро спадающая при $k > k_0 \sim l_0^{-1}$ функция. Поскольку начальная энергия турбулентности определяется спектром $E_0(k)$, а дисперсия начального потенциала — спектром $E_\psi(k)$, то имеются критические индексы $n_* = -3, -1, 1, 2$, связанные с расходимостью спектра скорости или потенциала на малых k . Общим свойством здесь является установление автомодельности на больших временах, в частности, для спектра турбулентности (рис. 13):

$$E(k, t) = \frac{L^3(t)}{t^2} \tilde{E}(kL(t)).$$

Однако свойства автомодельности для n из разных интервалов между критическими индексами качественно различны. В [24] проводилось численное моделирование турбулентности Бюргера и было продемонстрировано, что даже при наличии обрезających множителей в области малых или крупных масштабов эволюция спектра

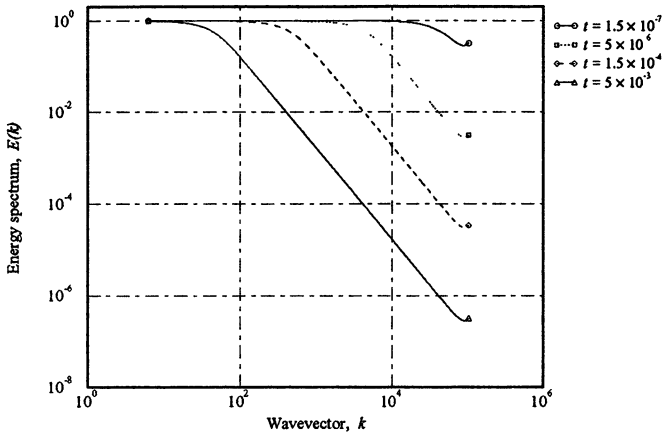


Рис. 13. Эволюция спектра $E(k, t)$, $E_0(k) \sim k^n$, $n = 0$

в ограниченной частотно-временной области по-прежнему автомодельна. Метод быстрого преобразования Лежандра позволил произвести усреднение по 3000 реализаций. Ниже даны полученные в [24] картины эволюции спектра для разных n .

Для $n < 1$ дисперсия потенциала бесконечна, а для $n < -1$ бесконечна и дисперсия начального поля. При чисто степенном начальном спектре начальный потенциал представляет собой броуновскую или фрактальную броуновскую реализацию, что позволяет провести ресэйлинг в глобальном решении уравнения Бюргера [14]:

$$\psi_0(\lambda x) = \lambda^{h+1} \psi_0(x), \quad h = -\frac{1}{2}(n+1),$$

$$\psi_0(x) \sim |x|^{h+1} = |x|^{\frac{1}{2}(n-1)},$$

$$v(x, t) = t^{\frac{h}{1-h}} v\left(x t^{-\frac{h}{1-h}}, 1\right).$$

Соответственно турбулентность оказывается автомодельной с самого начала эволюции. Ее спектр сохраняется при малых k , а характерный масштаб равен $L(t) = (\alpha t)^{2/(3+n)}$:

$$E(k, t) = \begin{cases} \alpha^2 k^n & \text{при } k < 1/L(t), \\ \frac{L(t)}{t^2} k^{-2} & \text{при } k > 1/L(t). \end{cases}$$

Асимптотика турбулентности со степенным начальным спектром, обрезанным на высоких частотах, также автомодельна и не зависит от масштаба l_0 начального спектра. Более того, на больших временах даже форма отдельных реализаций случайного поля слабо зависит от мелкомасштабных флуктуаций и определяется крупномасштабными компонентами начального поля [26]. При конечной вязкости ($\nu \neq 0$) число Рейнольдса

$$\text{Re}(t) = \frac{V_c(t)L_c(t)}{\nu} \sim t^{\frac{1-n}{n+3}}$$

растет с течением времени. Последнее связано с тем, что характерный масштаб турбулентности растет быстрее, чем уширение (из-за вязкости) ударного фронта.

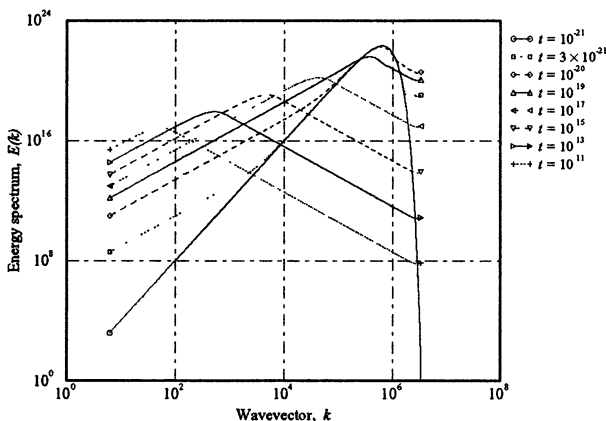


Рис. 14. Эволюция спектра $E(k, t)$, $E_0(k) \sim k^n$, $n = 4$

Если в спектре начального поля крупномасштабные компоненты отсутствуют, то реализуется иной сценарий вырождения турбулентности (рис. 14). При $n > 2$ закон затухания энергии сильно зависит от статистики начального поля [11, 25]. Для гауссова начального поля характерный масштаб $L(t)$ и энергия $E(t)$ турбулентности

$$L(t) = (t\sigma_\psi)^{1/2} l n^{-1/4} \left(\frac{t\sigma_\psi}{2\pi l_0^2} \right), \quad E(t) = t^{-1} \sigma_\psi l n^{-1/2} \left(\frac{t\sigma_\psi}{2\pi l_0^2} \right) \quad (12)$$

определяются лишь двумя параметрами начального спектра: дисперсиями начального потенциала $\sigma_\psi^2 = \langle \psi_0^2 \rangle$ и скорости $\sigma_v^2 = \langle v_0^2 \rangle$. Здесь $l_0 = \sigma_\psi / \sigma_v$ — характерный масштаб начального поля. При гауссовой статистике начального поля энергия затухает как t^{-1} с некой

логарифмической коррекцией. В этом случае удается дать исчерпывающее статистическое описание вырождающейся турбулентности. К примеру, показать, что спектры, корреляционные функции и вероятностные распределения становятся автомодельными [10, 28]. Так,

$$E(k, t) = \begin{cases} \frac{k^2 L^5}{t^2} & \text{при } k < 1/L, \\ \frac{2\tilde{L}(t)k^{-2}}{\pi^{3/2}t^2} & \text{при } k > 1/L. \end{cases} \quad (13)$$

При конечных числах Рейнольдса $Re_0 = \sigma_\psi/2\nu$ из-за многократных слияний разрывов число Рейнольдса спадает медленно $Re(t) = V(t)L(t)/\nu \sim Re_0 \ln^{-1/2} \left(\frac{t}{2\pi t_{nl}} \right)$, линейный режим возникает поздно, при $t \gg t_{nl} \exp(Re_0^2)/Re_0$, где $t_{nl} = \sigma_\psi/l_0^2$ – время проявления нелинейных эффектов. На линейной стадии энергия затухает как $Ct^{-3/2}$, где $C \simeq l_0 \exp(Re_0^2)/Re_0$.

Неполная автомодельность возникает при $1 < n < 2$ [29], когда в некотором интервале волновых чисел сохраняются крупные масштабы $E(k, t) \simeq \alpha^2 k^n$, $k < k_s(t)$, где $k_s \cdot L(t) = k_s(t)/k_L(t) \sim t^{-\frac{n-1}{2(2-n)}}$. Параметрическая генерация крупномасштабных структур порождает универсальный закон $E(k, t) \sim k^2$ в интервале $1/L(t) > k > k_s(t)$. При этом затухание энергии по-прежнему описывается уравнением (12) (рис. 15).

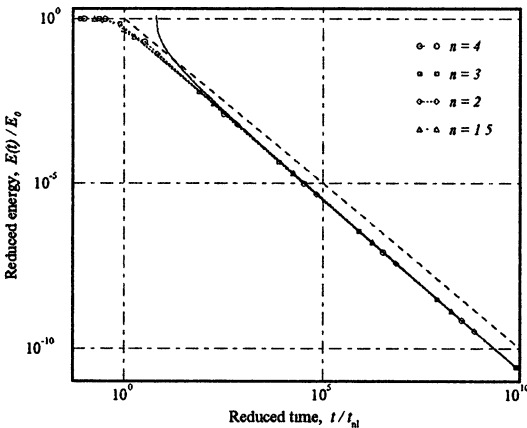


Рис. 15. Затухание энергии при разных n и теоретическая кривая, пунктиром изображено затухание энергии без логарифмической поправки.

Работа выполнена при поддержке грантов РФФИ 03-02-16680, 05-02-16517, «Ведущие научные школы России» НШ-838.2003.2 и программы «Университеты России».

Литература

1. Гурбатов С. Н., Саичев А. И. Введение в теорию нелинейных волн гидродинамического типа. Н. Новгород: Изд-во ННГУ, 2003. 156 с.
2. Pernet D. F., Payne R. C. Nonlinear propagation of signals in air // *J. Sound and Vib.* 1971. V. 17, № 3. P. 383–387.
3. Бьерно Л., Гурбатов С. Н. Об установлении высокочастотных асимптотик при распространении интенсивного акустического шума // *Акуст. журн.* 1985. Т. 31, № 3. С. 303–308.
4. Webster D. F., Blackstock D. T. Collinear interaction of noise with a finite-amplitude tone // *JASA.* 1978. V. 63, № 3. P. 687–693.
5. Burgers J. M. Mathematical examples illustrating relations occurring in the theory of turbulent fluid motion // *Kon. Ned. Akad. Wet. Verh.* 1939. V. 17. P. 1–53; *Selected Papers of J. M. Burgers* / Eds. F. T. M. Nieuwstadt & J. A. Steketee. Kluwer, 1995. P. 281–334.
6. Burgers J. M. *The nonlinear diffusion equation.* Dordrecht: Reidel, 1974.
7. Frisch U. *Turbulence: the legacy of A. N. Kolmogorov.* Cambridge: Cambridge University Press, 1995. 343 p.; Фриш У. Турбулентность. Наследие А. Н. Колмогорова. М.: ФАЗИС, 1998. 346 с.
8. Whitham G. B. *Linear and nonlinear waves.* N. Y.: Wiley, 1974.
9. Руденко О. В., Солуян С. И. *Теоретические основы нелинейной акустики.* М.: Наука, 1975.
10. Гурбатов С. Н., Малахов А. Н., Саичев А. И. *Нелинейные случайные волны в средах без дисперсии.* М.: Наука, 1990. 288 с.; Gurbatov S. N., Malakhov A. N., Saichev A. I. *Nonlinear random waves and turbulence in nondispersive media: waves, rays, particles.* Manchester: Manchester University Press, 1991. 308 p.
11. Woyczynski W. A. *Burgers–KPZ turbulence. Gottingen lectures.* Berlin: Springer-Verlag, 1998.
12. Gurbatov S. N., Saichev A. I., Shandarin S. F. The large-scale structure of the Universe in the frame of the model equation of non-linear diffusion // *Mon. Not. R. Astr. Soc.* 1989. V. 236. P. 385–402.
13. Shandarin S. F., Zeldovich Ya. B. The large-scale structure of the Universe: turbulence, intermittency, structures in a self-gravitating medium // *Rev. Mod. Phys.* 1989. V. 61. P. 185–220.
14. Vergassola M., Dubrulle B., Frisch U., Noullez A. Burgers' equation, Devil's staircases and the mass distribution for large-scale structures // *Astron. Astrophys.* 1994. V. 289. P. 325–356.
15. Barabási A.-L., Stanley H.E. *Fractal Concepts in Surface Growth.* Cambridge: Cambridge University Press, 1995.

16. Kardar M., Parisi G., Zhang Y. C. Dynamical scaling of growing interfaces // Phys. Rev. Lett. 1986. V. 56. P. 889–892.
17. Hopf E. The partial differential equation $u_t + uu_x = u_{xx}$ // Comm. Pure Appl. Mech. 1950. V. 3. P. 201–230.
18. Cole J. D. On a quasi-linear parabolic equation occurring in aerodynamics // Quart. Appl. Math. 1951. V. 9. P. 225–236.
19. Frachebourg L., Martin Ph. A. Exact statistical properties of the Burgers equation // J. Fluid Mech. 2000. V. 417. P. 323–349.
20. Leadbetter M. R., Lindgren G., Rootzen H. Extremes and related properties of random sequences and processes. N. Y.: Springer, 1983.
21. Noullez A., Vergassola M. A fast algorithm for discrete legendre transforms // J. Sci. Comp. 1994. V. 9. P. 259–281.
22. Gurbatov S. N., Enflo B. O., Pasmanik G. V. The decay of pulses with complex structure according to Burgers' equation // Acta Acustica. 1999. V. 85. P. 181–196.
23. Gurbatov S. N., Crighton D. B. The nonlinear decay of complex signals in dissipative media // Chaos. 1995. V. 5, №3. P. 524–530.
24. Noullez A., Gurbatov S. N., Aurell E., Simdyankin S. I. The global picture of self-similar and not self-similar decay in Burgers turbulence (in print).
25. Gurbatov S. N. Universality classes for self-similarity of noiseless multidimensional Burgers turbulence and interface growth // Phys. Rev. E. 2000. V. 61, № 3. P. 2595–2604.
26. Гурбатов С. Н., Пасманик Г. В. О самосохранении крупномасштабных структур в нелинейной вязкой среде, описываемой уравнением Бюргерса // ЖЭТФ. 1999. Т. 115, № 2. С. 564–583.
27. Gurbatov S. N., Troussov A. V. The decay of multiscale signals — a deterministic model of Burgers turbulence // Physica D. 2000. V. 145. P. 47–64.
28. Molchanov S. A., Surgailis D., Woyczynski W. A. Hyperbolic asymptotics in Burgers' turbulence and extremal processes // Commun. Math. Phys. 1995. V. 168. P. 209–226.
29. Gurbatov S. N., Simdyankin S. I., Aurell E., Frisch U., Toth G. On the decay of Burgers' turbulence // J. Fluid. Mech. 1997. V. 344. P. 339–374.

НЕКЛАССИЧЕСКАЯ НЕЛИНЕЙНАЯ АКУСТИКА

Л. А. Островский

Эта небольшая статья отражает содержание лекции, прочитанной на Школе по нелинейным волнам в марте 2004 года недалеко от Нижнего Новгорода. Она кратко описывает идеи «неклассической» нелинейной акустики (НА), относящейся к средам и материалам со структурной нелинейностью, связанной с макронеоднородностями, такими как зерна, трещины и т. д. Как лекция, так и эта короткая статья частично основаны на недавнем обзоре [1], а частично на еще более недавних публикациях как западных, так и российских исследователей. Поэтому я не буду давать детальных ссылок на работы, опубликованные до 2001 г., в надежде на то, что интересующийся читатель найдет их в [1]. Тем не менее, я процитирую несколько работ, на которые непосредственно опирается изложение, или те, которые появились позже и не упоминались в [1]. Эта заметка может рассматриваться как краткое введение в проблему для тех, кто не занимался ею раньше (или, возможно, занимался, но как у Мольера — не задумываясь, что «говорит прозой»).

О «классической» нелинейной акустике

Для того чтобы объяснить, что мы называем «неклассической» НА (как всякое определение, оно достаточно условно), вначале нам придется объяснить, что может быть определено как «классическая» НА. Исторически НА развивалась как часть механики сжимаемых сред, имеющей отношение к волнам (звуку) почти без дисперсии, с малой, но конечной нелинейностью, чаще всего квадратичной. «Классическими» проблемами, рассматриваемыми в различные периоды времени, можно считать простые (римановы) волны, слабые (тейлоровские) разрывы, параметрические антенны, нелинейные звуковые пучки, акустические взаимодействия в кристаллах и т. д.

Продемонстрируем вначале «классический» подход к одномерным продольным волнам в изотропном твердом теле. Известное уравнение движения, связывающее вектор смещения \mathbf{u} и тензор напряжений σ_{ik} , имеет вид

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ik}}{\partial x_j}. \quad (1)$$

Здесь, как обычно, по повторяющимся индексам производится суммирование. Тогда для продольных компонент тензора напряжений и деформаций $\sigma_{xx} = \sigma$ и $\epsilon_{xx} = \epsilon$ справедливо следующее уравнение:

$$\rho \frac{\partial^2 \epsilon}{\partial t^2} = \frac{\partial^2 \sigma(\epsilon)}{\partial x^2}. \quad (2)$$

Эти уравнения записаны в Лагранжевых переменных, как это принято в механике твердого тела. Эйлерова форма может содержать дополнительные нелинейные члены, однако в рассматриваемом ниже классе материалов преобладает «физическая» нелинейность, связанная с зависимостью $\sigma(\epsilon)$, и тогда разница между Лагранжевым и Эйлеровым представлениями несущественна.

Для слабой нелинейности, представляя $\sigma(\epsilon)$ в виде ряда

$$\sigma = M(\epsilon + \beta\epsilon^2 + \delta\epsilon^3 + \dots), \quad (3)$$

мы также получаем локальную скорость волны $c(\epsilon)$ в виде ряда:

$$c = \sqrt{\rho^{-1} d\sigma/d\epsilon} = c_0[1 + \beta\epsilon + (1/2)(3\delta - \beta^2)\epsilon^2 + \dots], \quad (4)$$

где $c_0 = \sqrt{M/\rho}$ — линейная скорость волны, а M характеризует жесткость материала (например, модуль Юнга в тонком стержне).

Одномерное уравнение, описывающее распространение волн в положительном направлении, может быть получено здесь в предположении, что ϵ зависит от «бегущих» переменных $\eta_+ = x - c_0 t$, $\eta_- = x + c_0 t$, таким образом, зависимость от η_- слабая (заметим, что в литературе чаще вводят медленную зависимость от t или x , что эквивалентно). Тогда из (2) следует:

$$\frac{\partial \epsilon}{\partial \eta_-} = -\frac{1}{4M} \frac{d\sigma_N}{d\epsilon} \frac{\partial \epsilon}{\partial \eta_+}, \quad (5)$$

где $\sigma_N = \sigma - M\epsilon$ — малая нелинейная часть напряжения. Ясно, что в отсутствие нелинейности решение этого уравнения есть произвольная функция от η_+ , т. е. линейная бегущая волна.

В общем же случае из этого уравнения следует (рассмотрите это сами), что бегущая волна конечной амплитуды (простая волна) испытывает искажение профиля вплоть до «опрокидывания», что означает формирование ударной волны. Заметим, что для квадратичной нелинейности, когда член с ϵ^3 в (3) пренебрежимо мал, и для кубичной нелинейности, когда квадратичный член исчезает (это возможно, например, для поперечно поляризованных волн, таких как сдвиговые), эволюция волны происходит совершенно различно (рис. 1).

Наконец, упомянем волны в резонаторе в виде стержня со свободными концами. В случае квадратичной нелинейности встречные волны в резонаторе в среднем не действуют друг на друга, а резонансные кривые остаются такими же, как и в линейном случае. Правда, если разрыв успевает образоваться на одном пробеге от одного конца стержня к другому, резонанс может стать шире вследствие дополнительных потерь на разрыве (рис. 2, а). С другой стороны, в кубичной среде встречные волны испытывают нелинейный сдвиг скорости и в среднем взаимодейст-

вуют друг с другом; резонансный отклик на каждой моде такого резонатора подобен поведению кубичного осциллятора с зависимостью резонансной частоты от амплитуды (рис. 2, б).

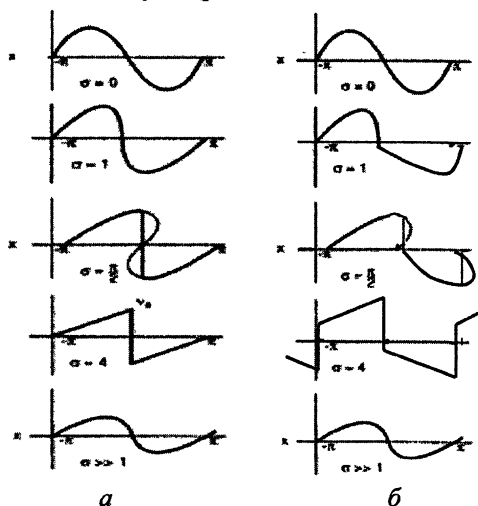


Рис. 1. Схема эволюции волнового профиля в среде с квадратичной нелинейностью (а) и в среде с кубичной нелинейностью (б); σ — время, нормированное на время формирования разрывов (ударных волн) в первоначально синусоидальной волне.

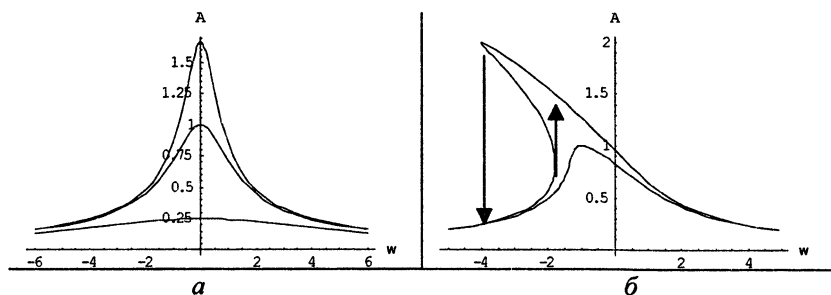


Рис. 2. Резонансные кривые при различных амплитудах возбуждения для квадратичной (а) и для кубичной нелинейности (б).

Начиная с 1960-х годов был достигнут существенный прогресс в классической НА. Он связан с изобретением «параметрических» антенн (которые было бы справедливее назвать просто нелинейными антеннами, потому что параметрический резонанс в них отсутствует) и с уравнением Хохлова — Заболотской — Кузнецова, широко используемого для исследования нелинейных звуковых пучков. Эти и другие результаты описаны более подробно в книгах по нелинейной акустике [2—3].

О «неклассической» НА

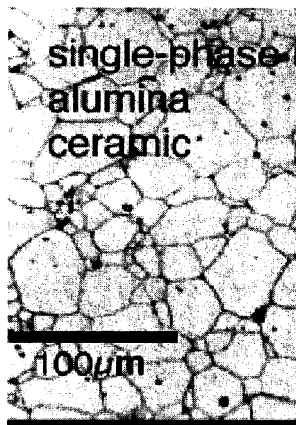
«Неклассические» материалы характеризуются аномально сильной нелинейностью, иногда сильной дисперсией и нередко неаналитическим (гистерезисным) откликом на воздействие. Жидкость с пузырьками газа может рассматриваться как типичная неклассическая нелинейная среда. При малом объемном содержании газа эта смесь характеризуется очень сильной квадратичной нелинейностью (на три порядка выше, чем в воде или газе по отдельности). Кроме того, по причине наличия у пузырьков резонансной частоты, газожидкостная смесь может быть сильно диспергирующей; это относится к скорости звука и/или затуханию: в случае широкого распределения по радиусам пузырьки порождают затухание, зависящее от их концентрации, но не зависящее от индивидуальных потерь в пузырьке (аналог затухания Ландау в плазме). Этот пример хорошо известен и описан в книгах, так что мы не будем рассматривать его далее. Отметим только, что подобное поведение обнаружено у упругой среды с порами, если среда «водоподобная» или «резиноподобная», т. е. если ее сдвиговой модуль мал [3].

Сильно нелинейные среды вовсе не редкое исключение. Для того чтобы в этом убедиться, давайте посмотрим на таблицу квадратичных параметров β (см. ниже) для различных материалов (в этом виде она взята из обзорной статьи [4]). Видно, что эти материалы могут быть поделены на две группы: «классические», имеющие β в интервале от 1 до 10, и «неклассические» с β порядка сотен и тысяч! При этом скорость звука в этих двух группах вполне сопоставима.

Скорость звука (продольной волны) и коэффициент квадратичной нелинейности, $-\beta$, для различных сред при комнатной температуре

Среда	Продольная скорость, c (м/с)	Коэффициент нелинейности, $-\beta$
Воздух	340	1.2
Дистиллированная вода	1497	3.5
Этанол	1180	5.5
Мягкие биологические ткани	1540 (в среднем) 1450	3.5—7.0
Ртуть	5968	4.9
Кварцевое стекло	5957	-3.9
Железо	4759	4.4
Медь	3704	2.4
Серебро	3240	2.0
Золото	~2000	3.1
Пенопласт	~3000	~100
Мрамор		~800
Морские глинистые осадки	~3000	100—1000
Земные породы		1000—10000

В чем же причина? Легко заметить, что первая группа представляет собой макроскопически однородные материалы: газы, жидкости, конденсированные твердые тела или монокристаллы. Материалы из второй группы явно неоднородны, они содержат поры, зерна и микротрещины. Они иллюстрируют общий принцип: сильная нелинейность ожидается в, так сказать, двухфазных материалах, в которых основной объем занят жесткой матрицей, но присутствуют и мягкие включения, занимающие сравнительно небольшой объем. Такими включениями могут быть пузырьки, контакты зерен, малые трещины и т. д. Одна иллюстрация приведена на рис. 3, где показан микроскопический срез одного из многих неклассических материалов. Здесь легко просматриваются зерна с близким контактом поверхностей.



Трудно проследить, когда начались исследования в области, которую мы сейчас называем неклассической НА. Изучение акустического воздействия на дислокации в металлах начато довольно давно. Отвлекаясь от статических экспериментов при сильных деформациях, отметим, что уже в сборнике «Методы ультразвуковых исследований», опубликованном в 1950-х, содержатся данные о гистерезисных свойствах металлов при ультразвуковом воздействии, так же как и модель Гранато-Люкке, о которой будет упомянуто ниже. Зарембо и Сердобольская отмечали аномально сильную нелинейность у ферроэлектриков вблизи фазовых переходов. Наши исследования начались в 1970-х [5] с экспериментов, в которых волна возбуждалась в кольцевом резонаторе, чтобы исключить сложности, связанные с взаимодействием встречных волн. Наблюдались следующие основные эффекты:

Рис. 3. Пример структуры (среза) «неклассического» материала [5].

1. Генерация очень широкого спектра гармоник: при возбуждении на частоте 10 кГц спектр расширился до нескольких десятков килогерц.

2. Генерация субгармоник: спектр расширялся как вниз относительно частоты накачки, так и вверх; были обнаружены дробные частоты на всем протяжении спектра.

3. Другой, тогда неожиданный, эффект заключался в том, что после акустического воздействия порог параметрической генерации изменялся, а потом возвращался к исходному значению за время порядка одного часа и более. Такая релаксация (или медленное время, как это называют сейчас), является весьма характерным эффектом структурной нелинейности.

Позднее, в 1980-х и 1990-х, подобные исследования интенсифицировались, особенно в двух местах: в ИПФ РАН и Лос-Аламосской на-

циональной лаборатории. Соответствующие эксперименты описаны в нашей недавней обзорной статье [1], поэтому здесь мы приведем лишь некоторые примеры для иллюстрации. Отметим, что большинство недавних экспериментов было проведено в стержневых резонаторах со свободными концами (Лос-Аламос) или с одним свободным и одним закрепленным концом (ИПФ РАН). В обоих случаях существуют технические нюансы, касающиеся контактов с излучателями, температуры и т. д., которые здесь обсуждаться не будут.

Естественно начать с генерации гармоник. На рис. 4 представлены зависимости амплитуд второй и третьей гармоник от амплитуды возбуждения в различных материалах, полученные как в Лос-Аламосе, так и в ИПФ РАН. Как и ожидалось, вторая гармоника растет пропорционально квадрату амплитуды волны основной частоты в резонаторе. Полностью неожиданный результат, однако, в том, что третья гармоника также

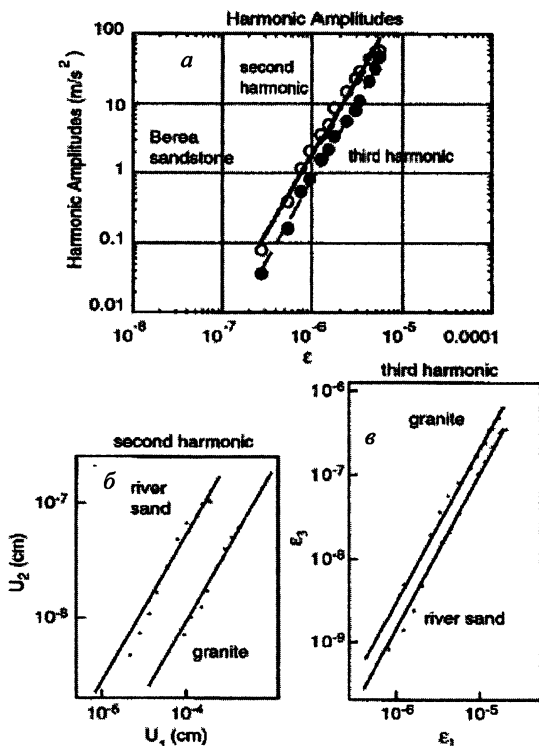


Рис. 4. Зависимости амплитуд второй и третьей гармоник от амплитуды основной гармоники (смещения и деформации) в резонаторах из различных материалов: обе гармоники в песчанике (Johnson et al., 1996) (a); вторая (б) и третья (в) гармоники в речном песке и граните (Johnson et al., 1996). См. обзор [1] и ссылки в нём.

пропорциональна квадрату, а не кубу амплитуды возбуждения! Легко показать, что такая зависимость просто невозможна, если связь «напряжение — деформация» представлена в виде аналитического ряда (3). Кроме того, амплитуда третьей гармоники может превышать вторую, что тоже необычно для акустики.

Другим важным признаком является сдвиг резонансного максимума для фиксированной моды резонатора. На рис. 5, *а* представлены резонансные кривые при разных амплитудах для бездефектного образца, а на рис. 5, *б* — для поврежденного образца (с трещиной). Очевидно, что во втором случае с увеличением амплитуды возбуждения пик резонанса намного сильнее сдвигается в область низких частот. Это похоже на рис. 2, *б* для кубического осциллятора, но с одним важным отличием: здесь резонансная кривая не опрокидывается, потому что потери тоже растут с амплитудой. Количественный анализ показывает, что нелинейный сдвиг пика резонансной кривой обычно пропорционален амплитуде вибраций в резонаторе. Это также невозможно для нелинейности, описываемой уравнением (3), для классических материалов, где эта зависимость квадратична.

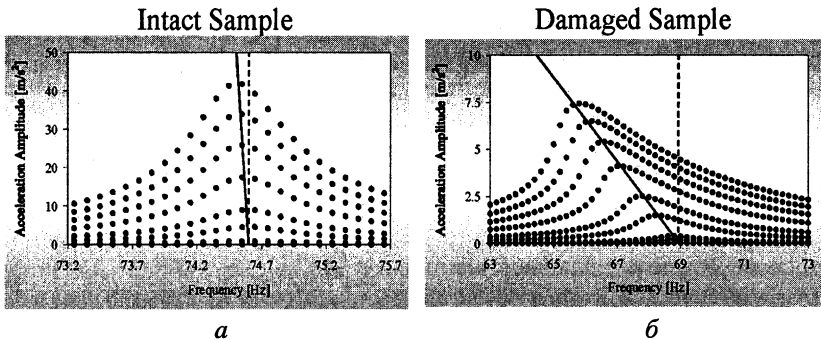


Рис. 5. Резонансные кривые для нормального (*а*) и дефектного (*б*) образцов композитного материала [1, 7].

Упомянем еще два примера «странного» поведения. Один из них — так называемое медленное время. Как уже отмечалось, релаксация к исходному состоянию может длиться более часа. Необычно для многих материалов и то, что возврат частоты линейного резонанса к его начальному значению после кратковременного воздействия мощного сигнала происходит в соответствии с законом $\log t$, тогда как обычно релаксационные процессы моделируются экспонентой от времени.

Другое интересное явление связано с водонасыщением материала. Многие материалы имеют пустоты, и заполнение водой меняет их акустические свойства. Один предел довольно очевиден: если заполнение

полное (100%-ное насыщение), то можно ожидать, что материал станет тверже, так что скорость звука в нем увеличится, а нелинейность уменьшится. Именно это и наблюдается в экспериментах. Но когда насыщение относительно малое, процесс становится более сложным. Действительно, вода может концентрироваться в небольшом зазоре между зернами, создавая таким образом силу поверхностного натяжения или, если зазор очень узкий (десятки или единицы молекулярных размеров) — также и силу Ван-дер-Ваальса. В результате нелинейность может увеличиться, а не уменьшиться, что также наблюдалось в экспериментах.

Все эти факты способны убедить любого, что классическая модель, основанная на аналитической зависимости $\sigma(\epsilon)$ и ее разложении в степенной ряд по ϵ , неадекватна для рассматриваемого класса материалов (хотя нелинейность все еще остается слабой в том смысле, что относительные нелинейные возмущения для, скажем, скорости звука, хотя и намного сильнее, чем в классической среде, но все еще малы по сравнению с самой этой скоростью).

О механизмах неклассической нелинейности

Мы до сих пор далеки от полного понимания специфических механизмов неклассической нелинейности. Однако очерченные выше основные принципы, а именно наличие мягких включений в жесткой матрице, в общем применимы.

Эта идея дает основу для более конкретных моделей. Одна из простейших — модель контактирующих сфер. Контакты описываются известной моделью Герца, связывающей силу F сжатия двух сфер и изменение расстояния Δ между их центрами. Для сфер с одинаковыми радиусами R это соотношение имеет вид

$$\Delta = \left(\frac{3(1-\nu^2)F}{4ER^{1/2}} \right)^{3/2}, \quad (6)$$

где E и ν — соответственно модуль Юнга и коэффициент Пуассона для материала сфер. Для среды, состоящей из таких сфер, после суммирования смещений по всем контактам получаем: $\sigma = K\epsilon^{3/2}$, где коэффициент K зависит от количества контактов, приходящихся на одну сферу. Такая среда, очевидно, сильно нелинейна; кроме того, растягивающая сила и вовсе не испытывает сопротивления. Соответственно скорость волны, пропорциональная $\sqrt{d\sigma/d\epsilon} \propto \sqrt{\epsilon}$, стремится к нулю для волн малой амплитуды. Однако в акустике волны обычно накладываются на постоянное (статическое) деформирование ϵ_0 , так что возмущения, связанные со звуком, много меньше ϵ_0 и нелинейность мала. В этом случае модуль M в уравнении (3) пропорционален $\epsilon_0^{1/2}$, а нелинейные параметры равны

$\beta = 1/2\epsilon_0$ и $\delta = 1/6\epsilon_0^2$. Интересно отметить, что эти параметры зависят только от статической деформации, но не зависят от радиуса сфер или других параметров материала. Рассматривались и модификации этой модели, включающие в себя различные размеры сфер и слабые контакты. Модель проверялась в ИПФ РАН для сферических зерен. Она легко дает сильную нелинейность, но не может объяснить гистерезис. Возможный шаг к учету последнего эффекта — рассмотрение поперечных (сдвиговых) смещений зерен. В соответствии с моделью Миндлина, нормальное давление создает еще тангенциальное смещение s и соответствующую поперечную силу

$$F_i = \frac{4\mu a(\Delta)s}{1-\nu}, \quad (7)$$

где μ — модуль сдвига. Параметр a может быть ненулевым, если контакт шероховатый (не скользящий). Эта сила может работать по-разному при сжатии и растяжении, создавая таким образом гистерезис. Однако её реальный эффект до сих пор не ясен.

Физическая модель гистерезисного поведения в металлах была предложена Гранато и Люкке в начале 1950-х. В ней дислокации в металле рассматриваются как «струны», закрепленные сильно к узлам решетки и слабо — к примесным атомам. Под действием механических деформаций эти слабые соединения разрываются, делая таким образом материал мягче, так что возврат в исходное состояние происходит другим путем на плоскости $\sigma - \epsilon$. Совсем недавно подобные идеи были применены к материалам, в которых конкретный механизм гистерезисного поведения не известен. В такой модели, («РМ-модель») материал рассматривается как конгломерат гистерезисных элементов с прямоугольной петлей на плоскости «длина — давление» со статистически распределенными толщинами петель, когда скачки давления меняются от нуля (без гистерезиса) до некоторого максимума. В результате зависимость напряжение — деформация также гистерезисная.

В теоретическом моделировании обычно используются два упрощенных типа петель $\sigma - \epsilon$. Одна из них, «неупругий гистерезис» (рис. 6, а), в принципе необратима и может, строго говоря, описывать только периодические движения. Она дается формулой

$$\sigma - M\epsilon = \sigma_N = -[a\epsilon_m \epsilon \pm \gamma(\epsilon^2 - \epsilon_m^2)], \quad \epsilon_i \geq 0; a, \gamma > 0. \quad (8)$$

Второй тип, «упругий гистерезис» (рис. 6, б), возвращает движение в начальную точку. Соответствующая зависимость $\sigma(\epsilon)$ имеет вид

$$\sigma - M\epsilon = \sigma_N = -M\gamma \begin{cases} \epsilon^2, & \epsilon > 0, \epsilon_i > 0; \\ -\epsilon^2 + 2\epsilon_m \epsilon, & \epsilon > 0, \epsilon_i < 0; \\ -\epsilon^2, & \epsilon < 0, \epsilon_i < 0; \\ \epsilon^2 + 2\epsilon_m \epsilon, & \epsilon < 0, \epsilon_i > 0. \end{cases} \quad (9)$$

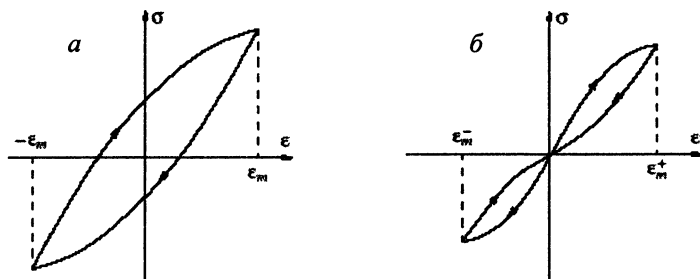


Рис. 6. Две упрощенные модели гистерезисной зависимости напряжение — деформация.

Пока не совсем ясно, какая из этих нелинейности более реалистична для малоамплитудных воздействий в виброакустических экспериментах. Вероятно, что-то промежуточное имеет место, когда некоторое остаточное напряжение сохраняется после каждого периода.

Акустические волны в гистерезисных средах

Волновые процессы в таких средах в общем случае весьма сложны для теории. Если дисперсия пренебрежимо мала (частотные зависимости гистерезисной нелинейности пока не изучены достаточно хорошо), то распространение соответствует простой (римановой) волне: каждая точка ее профиля распространяется с постоянной скоростью, как и в уравнении (5). Соответствующее уравнение для бегущей волны выглядит следующим образом (здесь $\eta_{\pm} = x \mp ct$ и зависимость от η_{-} медленнее):

$$2M \frac{\partial \epsilon}{\partial \eta_{-}} = \alpha \langle \epsilon_m \rangle \frac{\partial \epsilon_{+}}{\partial \eta_{+}} \pm \gamma \frac{\partial \epsilon_{+}^2}{\partial \eta_{+}}, \quad (10)$$

где знак последнего слагаемого соответствует знакам в (8) и (9).

Качественно это процесс показан на рис. 7, а для случая уравнения (8), а на рис. 7, б для (9) (см. [6]). В первом случае ударная волна не образуется вообще, вместо этого две половины периода волны распространяются (в «бегущей» системе отсчета) навстречу друг другу и в результате «съедаются» друг друга на максимумах и минимумах профиля, образуя в итоге пилообразный профиль, отличающийся от приведенного на рис. 1. Соответствующее затухание — это, по существу, потери, обусловленные гистерезисом. Отметим, что в отличие от «нормальной» среды с квадратичной нелинейностью, где сильное затухание начинается после образования разрыва (ударной волны), здесь процесс «съедания» (хотя и слабый вначале) существует с самого начала. В случае (9) такой эффект существует наряду с образованием ударных волн (рис. 7, б).

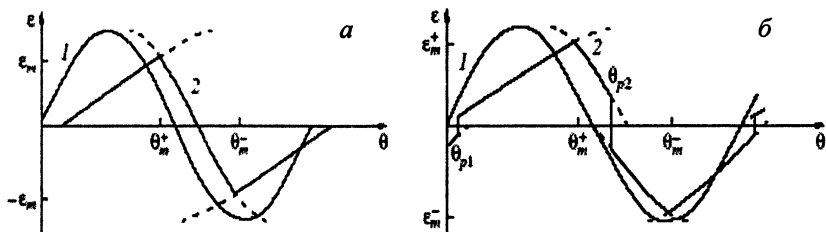


Рис. 7. Схема эволюции синусоидального возмущения для двух моделей, показанных на рис. 6.

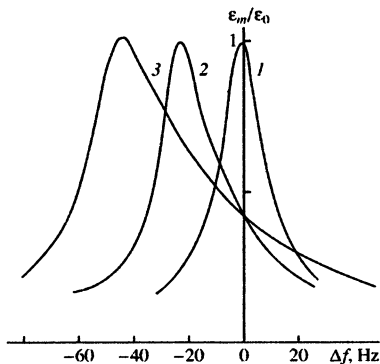
Более трудна теория резонаторов с гистерезисом. Относительно простая модель — кольцевой резонатор, в котором каждая мода представляет собой волну, распространяющуюся по кольцу длины L . Если источник с амплитудой ϵ_0 действует в некоторой точке $x = 0$ резонатора, установившаяся бегущая волна постоянной амплитуды ϵ_m должна удовлетворять граничному условию:

$$\epsilon_m(L) = \epsilon_m(0) \exp\{ikL[1 + \Delta_p + \Delta_h(\epsilon_m) - i\delta(\epsilon_m)]\} + \epsilon_0.$$

Здесь Δ_p и Δ_h — соответственно линейный и нелинейный сдвиги резонанса, δ характеризует потери, которые обычно связаны с гистерезисом. Результирующий нелинейный отклик имеет вид

$$\epsilon_m = \frac{\epsilon_0}{\sqrt{[\Delta_p + \Delta_h(\epsilon_m)]^2 + \delta^2(\epsilon_m)}}. \quad (11)$$

Рис. 8. Теоретические резонансные кривые для кольцевого резонатора из поликристаллической меди, возбуждаемого на частоте 4,35 кГц при различных амплитудах.



Пример семейства таких резонансных кривых (для различных амплитуд возбуждения), взятый из [6], приведен на рис. 8. На первый взгляд, они похожи на резонансные кривые для случая кубичной нели-

нейности (см. рис. 2). Однако есть важное отличие: как и в экспериментах, из-за нелинейных потерь в гистерезисной среде кривые не опрокидываются, поскольку они не только сдвигаются, но еще и расширяются, что предотвращает опрокидывание.

Однако большинство экспериментов проводится в стержневых резонаторах, в которых волны отражаются от концов и взаимодействуют друг с другом. Такое взаимодействие впервые рассматривалось нами в 1988 г. для случая «слабой» нелинейности, когда поле смещений близко к гармонической стоячей волне. Такое поле может быть представлено в виде суммы мод:

$$u \approx \sum_n A_n \cos[(\omega_n + \delta_n)t + \psi] \sin(k_n x + \phi), \quad (12)$$

где ω_n и k_n соответствуют n -му линейному резонансу, $\delta_n = \omega - \omega_n$, ω — частота возбуждения, ψ и ϕ — фазовые сдвиги, зависящие от граничных условий на концах стержня и его материала. Подставляя это выражение в уравнение (2), где зависимость $\sigma(\epsilon)$ теперь дается гистерезисной формулой (8) или (9), можно оставить только основную временную гармонику функции $\epsilon(x, t)$ и $\sigma(x, t)$ и таким образом замкнуть уравнение. В результате резонансный отклик получится в форме (11), но с другими нелинейным сдвигом частоты и потерями. Аналогично высшие гармоники могут быть найдены выбором соответствующих гармоник в искомым переменных.

Важно заметить, что многие качественные особенности гистерезисных резонаторов, наблюдаемые в экспериментах, такие как зависимости $\Delta_n \propto \epsilon_m$ и $A_3 \propto \epsilon_m^2$ (A_3 — амплитуда третьей гармоники), описываются уже в рамках рассмотренных простых моделей.

Еще труднее исследовать ситуации, когда волны в резонаторе не близки к синусоидальным. Для бегущих волн в кольцевом резонаторе эти искажения происходят тем же образом, как показано на рис. 7, а, б. Однако для стержневого резонатора необходимо учесть взаимодействие встречных волн. Такие процессы были изучены теоретически совсем недавно. Здесь мы будем следовать работе [7], которая дает простой способ для построения решения в форме

$$\epsilon \approx \epsilon_+(\eta_+, \eta_-) + \epsilon_-(\eta_-, \eta_+), \quad (13)$$

где $\eta_+ = x - ct$, $\eta_- = x + ct$, тогда как ϵ_+ медленно зависит от η_- (так что $\partial\epsilon/\partial\eta_- \ll \partial\epsilon/\partial\eta_+$), а ϵ_- — медленная функция η_+ (обычно в качестве медленных переменных берутся x или t по отдельности, но здесь, ввиду симметрии «лево-право», удобнее сохранить те же «бегущие» переменные). Подставляя это в уравнение (2) с данной (включая гисте-

резисную) зависимостью $\sigma(\epsilon)$, мы используем интегральную форму уравнения:

$$\epsilon = \epsilon_{0+}(\eta_+) + \epsilon_{0-}(\eta_-) + \epsilon_N, \quad (14)$$

$$2M\epsilon_N = \int \frac{\partial \sigma_N}{\partial \eta_-} d\eta_- + \int \frac{\partial \sigma_N}{\partial \eta_+} d\eta_+ + 2\sigma_N.$$

Далее учитываются только накапливающиеся члены в ϵ_N , которые не исчезают после интегрирования по периоду, иначе они добавляют только малые возмущения к линейному решению. Отсюда, в частности, без труда могут быть получены «классические» бегущие волны, например, изображенные на рис. 2, а также бегущие волны в средах с гистерезисом, описываемые уравнением (10). Для резонатора с гистерезисом типа (8) относительно простой результат следует в случае высокой добротности, когда амплитуды и профили встречных волн почти одинаковы, а именно

$$2M \frac{\partial \epsilon_+}{\partial \eta_-} = \alpha \langle \epsilon_m \rangle \frac{\partial \epsilon_+}{\partial \eta_+} \pm \gamma \frac{\partial \epsilon_+^2}{\partial \eta_+}, \quad (15)$$

где $\langle \epsilon_m \rangle > 0$ — усредненный по пространству максимум суммарного поля (сумма обеих волн, распространяющихся навстречу друг другу; в случае синусоид, $\epsilon_m(x)$ — это просто пик положительной фазы стоячей волны). Аналогичное уравнение справедливо для ϵ_- . Как и раньше, знаки \pm соответствуют случаям $\epsilon_i > 0$ и $\epsilon_i < 0$. На первый взгляд, это уравнение очень похоже на уравнение для одной бегущей волны с заменой $\langle \epsilon_m \rangle$ на постоянную или медленно меняющуюся амплитуду волны ϵ_m . Однако последний член значительно отличается: его знак определяется суммарным полем обеих волн, которое может иметь весьма сложный профиль. Анализ этого процесса основан на том, что переменная ϵ_+ «бежит» вместе с волной и, следовательно, фиксирована или только медленно меняется в каждой точке профиля (так же, как в простой волне). Действие встречной волны знакопеременно и в результате приводит к более медленным искажениям профиля, чем в бегущей волне.

Отсюда можно сделать следующие выводы. Средняя скорость волны меняется пропорционально величине $\langle \epsilon_m \rangle$, которая может отличаться от ϵ_m в одной бегущей волне той же амплитуды (они эквивалентны для пилообразной волны рис. 7, но для синусоиды $\langle \epsilon_m \rangle$ дает несколько большее нелинейное изменение скорости). Но в принципе изменение средней скорости и, соответственно, нелинейный сдвиг частоты происходят таким же образом, как и в кольцевом резонаторе с бегущей вол-

ной. В то же время искажение профиля каждой из встречных волн происходит медленнее, что приводит к меньшим амплитудам высших гармоник. Похоже, что этот вывод подтверждается экспериментальными данными.

Неразрушающий контроль материалов

Выше мы рассматривали основные физические эксперименты и модели. В заключение кратко остановимся на приложениях в неразрушающем контроле (или NDE – non-destructive evaluation). Нетрудно догадаться, что здесь мы имеем в виду «нелинейный неразрушающий контроль», или NNDE.

Известно много апробированных акустических методов NDE, таких как измерение скорости звука и потерь в материалах, отражение акустических импульсов и резонансная акустическая спектроскопия. Эти методы основаны на линейных акустических свойствах материала, в которых условия распространения и отражения не зависят от амплитуды пробной волны, так что последняя важна только для получения достаточного отношения сигнал/шум в принятом сигнале. Эти методы полезны в множестве приложений. Однако в других важных случаях дефекты в тестируемом материале не оказывают достаточного влияния на скорость звука и потери в материале, и их наличие не может быть достоверно обнаружено стандартными линейными методами. В таких случаях нелинейные методы зачастую более эффективны. Действительно, даже малые структурные дефекты материала, незначительно влияющие на линейное распространение звука, могут резко менять нелинейный отклик. Например, трещины в материале обычно обладают сильной неклассической нелинейностью и могут увеличивать нелинейный отклик материала на порядки, тогда как относительные изменения скорости звука и других линейных параметров остаются малыми. При этом могут быть успешно использованы практически любые из эффектов, рассмотренных выше (такие, как нелинейный сдвиг частоты, генерация гармоник и кросс-модуляция).

Использование нелинейных эффектов в NDE началось в нескольких местах, и Нижний Новгород был одним из первых. В настоящее время это широко востребованный метод (или скорее группа методов), который развивается в различных странах.

Один пример, когда дефект может быть обнаружен по сдвигу резонанса, по существу уже был показан на рис. 5. Рисунок 9 иллюстрирует другой способ. В этом случае использовалась кросс-модуляция: образец (колокол) возбуждался одной высокой и одной низкой частотами. Отклик бездефектного образца существенно линейный: его спектр содержит только те же две частоты. В поврежденном образце, однако, появились значительные комбинационные частотные компоненты, что свидетельствует о наличии повреждений. Рисунок 10 показывает часть набора

деталей, тестируемых таким образом в Лос-Аламосской национальной лаборатории. NNDE – перспективный и быстро развивающийся раздел NDE и в то же время наглядная демонстрация практической стороны неклассической НА.

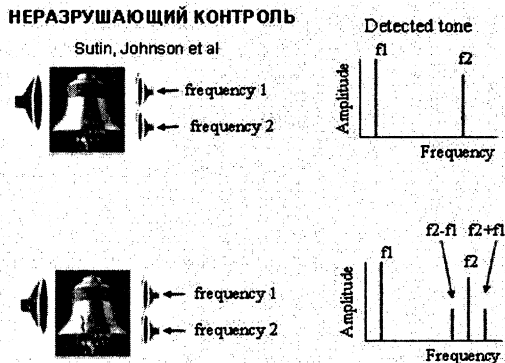


Рис. 9. Метод кросс-модуляции для диагностики дефектов [10]



Рис. 10. Некоторые образцы тестируемых деталей (предоставлено П. Джонсоном и А. Сутиним).

Заключение

Данная область исследований относительно нова и находится в развитии, много вопросов пока остается без ответа. Необходимо заметить, что физика, лежащая в основе описанных выше процессов, весьма богата: в принципе она включает в себя широкий спектр масштабов, от

атомных через «мезоскопические» (зерна, трещины) до «макроскопических» (длины акустических волн). К тому же модели нелинейных волн в таких средах достаточно специфичны (заметим, что их электродинамический аналог, включающий в себя ударные электромагнитные волны в ферритах с гистерезисом — популярная тема 1960-х и 1970-х). Наконец, практические приложения — многообещающие. Я бы хотел пригласить молодых ученых поработать в этой области.

Литература

1. *Ostrovsky L. A., Johnson P. A.* Dynamic nonlinear elasticity in geomaterials // *La Rivista del Nuovo Chimento*. 2001. V. 24, № 7. P. 1—46.
2. *Руденко О. В., Солуян С. И.* Теоретические основы нелинейной акустики. М.: Наука, 1975.
3. *Наугольных К. А., Островский Л. А.* Нелинейные волновые процессы в акустике. М.: Наука, 1990.
4. *Zheng Y., Maev R. G., Solodov I. Yu.* Nonlinear acoustic applications for material characterization: A review // *Can. J. Phys.* 1999. V. 77. P. 927—967.
5. *Ostrovsky L. A., Soustova I. A., Sutin A. M.* Nonlinear and parametric phenomena in dispersive acoustic systems // *Acustica*. 1978. V. 5. P. 298—306.
6. *Назаров В. Е., Радостин А. В., Островский Л. А., Соустова И. А.* Волновые процессы в средах с гистерезисной нелинейностью // *Акуст. журн.* 2004. Т. 48, № 3. Ч. 1 и 2.
7. *Ostrovsky L. A.* Wave interaction in acoustic resonators with and without hysteresis // *J. Acoust. Soc. Am.* 2004 (in print).

ВОЛНЫ ДЕФОРМАЦИИ В СРЕДЕ С ВНУТРЕННЕЙ СТРУКТУРОЙ

А. И. Потапов

В классической механике сплошной среды физически бесконечно малый (представительный) объем рассматривается как материальная точка, не имеющая размеров. В физике этим понятием обычно пользуются в тех случаях, когда размеры и форма частиц при изучении движения среды не играют роли, а важна только их масса. Эта модель весьма успешно описывает движение сплошной среды, а также распределения напряжений и деформаций в достаточно гладких областях при достаточно плавных нагрузках. В последние десятилетия были разработаны новые классы ультрадисперсных и нанокристаллических материалов. Для них математические модели, основанные на понятии материальной точки оказались неадекватными [1]. Кроме того, они не могут описывать высокочастотные вибрации [2, 3] в сложных кристаллических решетках органических веществ, в вершинах трещин и на фронте волны разрушения [4, 5], при структурных превращениях и возникновении дефектов [5, 6]. Подобные ситуации имеют место, например, при распространении коротких акустических волн в нанокристаллических средах, жидких кристаллах и полимерах [7—9]. В этих случаях необходимо учитывать дискретный характер структуры вещества, состоящего из неточечных частиц (гранул), связанных сложными силами взаимодействия. В роли частиц могут выступать крупные молекулы, домены, нанокристаллы и т. п. В подобных структурах появляются качественно новые волновые эффекты, которые не могут быть объяснены в рамках классической теории упругости.

Соответствие между теорией и экспериментом может быть достигнуто ценой отказа от классической модели сплошной среды и замены ее другими более отвечающими опыту моделями. На этом пути приходим к необходимости разработки и исследования новых математических моделей сред, способных отражать специфику внутренней структуры вещества [8, 10—13].

1. Обобщенные континуальные модели

По-видимому, наиболее универсальна модель среды как дискретной системы частиц, связанных определенными силами взаимодействия. Однако анализ этих сил в реальных системах сам по себе представляет сложную проблему [14, 15]. Более простой выход состоит в такой модификации классической модели, которая при сохранении гипотезы сплошности среды [8, 10—12] наделила бы ее главными специфическими свойствами среды с дискретной структурой [10, 16].

Одна из основных гипотез классической механики сплошных сред — принцип напряжений Коши — предполагает эквивалентность действия всех внутренних сил, приложенных к элементарной площадке, действию их равнодействующей, приложенной к ее центру. Из механики известно, что в общем случае действие произвольной системы сил эквивалентно действию главного вектора и главного момента. При этом в среде возникают не только силовые, но и моментные напряжения, образующие, вообще говоря, несимметричные тензоры [12]. Для учета этих факторов необходимо отказаться от гипотезы Коши и учитывать внутренние пары сил (собственные моменты) [17]. Это приводит к необходимости рассмотрения представительного объема среды не как материальной точки, а как более сложного объекта, обладающего дополнительными внутренними степенями свободы и способностью к «микродеформации» [8].

Фундаментальные законы сохранения массы, импульса и момента импульса приводят к универсальным (базовым) уравнениям движения, справедливым для любых типов сплошных сред. Однако они не замкнуты, и их необходимо дополнить материальными соотношениями (уравнениями), несущими информацию о структуре среды [17]. Например, связью между напряжением и деформацией в твердом теле или связью между напряженностью электрического поля и поляризацией среды в электродинамике и т. п. В общем случае это могут быть дифференциальные или интегро-дифференциальные уравнения. Если в материальные уравнения среды вводятся переменные, изначально отсутствующие в фундаментальных законах сохранения, то их называют *внутренними* переменными. Для них вводятся дополнительные уравнения, отражающие внутреннюю динамику среды. Примерами внутренних переменных могут служить, например, вектор магнитной или электрической поляризации среды, параметр порядка в теории фазовых переходов, параметр поврежденности в теории разрушения [4, 7, 17].

Исторически одной из первых моделей среды, которая не могла быть описана в рамках классической теории упругости, является континуум Коссера [10, 12], состоящий из твердых недеформируемых тел, обладающих тремя трансляционными и тремя ротационными степенями свободы. Уравнения динамики микрополярной среды могут быть получены из вариационного принципа Гамильтона — Остроградского:

$$\delta J = \delta \int_{t_0}^{t_1} \left(\iiint_{V_0} L dV \right) dt = 0.$$

Здесь L — лагранжиан, V_0 — начальный объем, занимаемый средой. В зависимости от того, какие полевые функции, определяющие динамическое состояние среды, выбираются в качестве искомого, вариационный

принцип может иметь несколько эквивалентных форм [13, 17]. Его исходная формулировка связана с варьированием закона движения частиц при неизменных лагранжевых координатах. Такая формулировка удобна в задачах механики твердого деформируемого тела, и уравнения движения получаются наиболее коротким путем. Лагранжиан записывается как разность между объемными плотностями кинетической и внутренней энергии среды:

$$L = \frac{1}{2} \rho_0 u_{\alpha,t}^2 + \frac{1}{2} \rho_0 J_{\alpha\beta} \phi_{\alpha,t} \phi_{\beta,t} - \rho_0 U(u_{\alpha,\beta}, \phi_{\alpha,\beta}). \quad (1)$$

Здесь поля смещений $u_\alpha(x, t)$ и микроповоротов $\phi_\alpha(x, t)$ рассматриваются как независимые функции лагранжевых координат и времени. После проведения варьирования получаются две системы уравнений:

$$\rho_0 \ddot{u}_\alpha = \frac{\partial}{\partial x_\beta} \left(\frac{\partial \rho_0 U}{\partial u_{\alpha,\beta}} \right), \quad \rho_0 J_{\alpha\beta} \ddot{\phi}_\beta = \frac{\partial}{\partial x_\beta} \left(\frac{\partial \rho_0 U}{\partial \phi_{\alpha,\beta}} \right). \quad (2)$$

Первая из них описывает макродвижения среды (движения центров масс частиц), а вторая — внутреннюю динамику среды, связанную с ориентационными движениями микроэлементов. Производная от внутренней энергии по градиенту перемещений $\sigma_{\alpha\beta} = \partial(\rho_0 U) / \partial u_{\alpha,\beta}$ представляет собой лагранжев тензор макронапряжений, а производные по градиенту внутренней переменной $m_{\alpha\beta} = \partial(\rho_0 U) / \partial \phi_{\alpha,\beta}$ — лагранжев тензор микронапряжений в среде. Для неполярных сред $J_{\alpha\beta} = 0$ и первая система уравнений в (2) представляет собой уравнения классической теории упругости, а из второй системы, выражающей условия отсутствия микронапряжений, следует симметричность тензора макронапряжений $\sigma_{\alpha\beta} = \sigma_{\beta\alpha}$ [8, 17].

Объемная плотность внутренней энергии однородного, изотропно-го, центрально-симметричного континуума Коссера может быть представлена в виде степенного ряда по инвариантам несимметричного тензора деформаций $\Psi_{\alpha\beta} = u_{\beta,\alpha} + \epsilon_{\alpha\beta\gamma} \phi_\gamma$ и тензора изгиба-кручения

$$\kappa_{ji} = \phi_{i,j} :$$

$$\begin{aligned} \rho_0 U = & \frac{\mu + \alpha}{2} \Psi_{ji} \Psi_{ji} + \frac{\mu - \alpha}{2} \Psi_{ji} \Psi_{ij} + \frac{\lambda}{2} (\Psi_{kk})^2 + \frac{\gamma + \epsilon}{2} \kappa_{ji} \kappa_{ji} + \\ & + \frac{\gamma - \epsilon}{2} \kappa_{ji} \kappa_{ij} + \frac{\beta}{2} (\kappa_{kk})^2 + \frac{\nu_1}{6} (\Psi_{kk})^3 + \frac{\nu_2 + \delta_1}{2} \Psi_{ji} \Psi_{ji} \Psi_{kk} + \\ & + \frac{\nu_2 - \delta_1}{2} \Psi_{ij} \Psi_{ji} \Psi_{kk} + \frac{4\nu_3 + 2\delta_2 + 2\delta_3}{9} \Psi_{ij} \Psi_{jk} \Psi_{ki} + \frac{4\nu_3 - 2\delta_2}{9} \Psi_{ij} \Psi_{jk} \Psi_{ik} + \end{aligned}$$

$$+ \frac{4\nu_3 - 2\delta_3}{9} \Psi_{ij} \Psi_{kj} \Psi_{ik} + \sigma_1 (\Psi_{kk})^2 \kappa_{nn} + \sigma_2 \Psi_{ij} \Psi_{ji} \kappa_{kk} + \sigma_3 \Psi_{ij} \Psi_{ij} \kappa_{kk},$$

где λ и μ — константы Ламе, α , γ , ε — константы линейной микроупругости, $\nu_{1,2,3}$ — константы нелинейной упругости, $\delta_{1,2,3}$ и $\sigma_{1,2,3}$ — константы нелинейной микроупругости.

В случае коллинеарного распространения плоских волн вдоль оси x уравнения движения записываются в виде

$$\rho u_{1,tt} - (\lambda + 2\mu) u_{1,xx} = \frac{\partial}{\partial x} N_{u1} (u_{j,x}, \phi_j, \phi_{j,x}), \quad (3a)$$

$$\rho u_{2,tt} - (\mu + \alpha) u_{2,xx} + 2\alpha \phi_{3,x} = \frac{\partial}{\partial x} N_{u1} (u_{j,x}, \phi_j, \phi_{j,x}), \quad (3б)$$

$$J \phi_{1,tt} - (\beta + 2\gamma) \phi_{1,xx} + 4\alpha \phi_1 = \frac{\partial}{\partial x} N_{\phi 1} (u_{j,x}, \phi_j, \phi_{j,x}), \quad (3в)$$

$$J \phi_{2,tt} - (\gamma + \varepsilon) \phi_{2,xx} + 4\alpha \phi_2 + 2\alpha u_{3,x} = N_{\phi 2} (u_{j,x}, \phi_j, \phi_{j,x}). \quad (3г)$$

Здесь u_1 — продольное, u_2 и u_3 — поперечные смещения частиц среды, $\phi_{1,2,3}$ — проекции вектора микроповорота частиц на оси координат, J — момент инерции частиц. Уравнение для переменной u_3 получается из (3б) заменой u_2 на u_3 , а для переменной ϕ_3 — из (3г) заменой ϕ_2 на $-\phi_3$. Конкретный вид нелинейных слагаемых, стоящих в правых частях уравнений, приведен в [13]. Первые два уравнения в (3) описывают упругие (продольные и поперечные) волны и являются следствиями закона сохранения импульса среды. Два других уравнения являются следствием закона сохранения полного кинетического момента среды, включающего в себя внутренний кинетический момент, связанный с микроповоротами частиц [8, 17]. Они описывают волны микровращений, которые отсутствуют в классической теории упругости.

В случае плоских движений ($u_{1,2} \neq 0$, $u_3 = 0$, $\phi_3 \neq 0$, $\phi_{1,2} = 0$) дисперсионные свойства волн, распространяющихся в континууме Коссера, описываются уравнением (рис. 1)

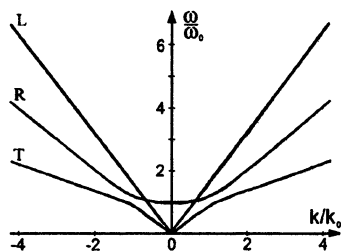
$$(\omega^2 - c_l^2 k^2) [(\omega^2 - c_\tau^2 k^2)(\omega^2 - c_\phi^2 k^2 - \omega_0^2) - \beta^2 k^2] = 0,$$

где $c_l = \sqrt{(\lambda + 2\mu)/\rho}$, $c_\tau = \sqrt{(\mu + \alpha)/\rho}$ — скорости продольной и сдвиговой волн соответственно, $c_\phi = \sqrt{(\gamma + \varepsilon)/J}$ — скорость волны микровращения, а $\omega_0 = \sqrt{4\alpha/J}$ — частота ее отсечки, β — параметр межмодовой связи.

Из анализа дисперсионных свойств системы следует, что продольная волна u_1 (мода L) не обладает дисперсией и в линейном приближении не взаимодействует ни с одной из волн. Учет микроповоротов частиц приводит к возникновению высокочастотной дисперсии у сдвиговых волн $u_{2,3}$ (мода T) и к изменению их скорости. Кроме того, появляется волна *микровращений* ϕ_3 (мода R), характеризующая внутренние вращения в среде.

Рис. 1. Дисперсионные ветви нормальных волн в континууме Коссера:

L — продольная мода,
 T — поперечная мода,
 R — спиновая мода.



Дисперсионные свойства волн микровращений аналогичны свойствам спиновых волн в магнитоупругой среде [18], поэтому их также называют *спиновыми*. Спиновые волны обладают предельной частотой ω_0 , ниже которой они являются не распространяющимися. Продольная спиновая волна ϕ_1 , также как и продольная упругая волна, является изолированной и в линейном приближении не взаимодействует с другими типами волн. В то же время поперечные спиновые волны $\phi_{2,3}$ и сдвиговые волны взаимодействуют между собой.

Большинство результатов механики деформируемого твердого тела получены в рамках континуального подхода [8, 10—12]. Дискретные же модели использовались в основном в физике твердого тела и теории кристаллической решетки [13—15]. В механике интерес к дискретным моделям возобновился к середине прошлого века. Такой интерес связан, в частности, с тем, что дискретные модели позволили обнаружить в задачах разрушения ряд эффектов, не улавливаемых в континуальном приближении [3, 4]. Кроме того, дискретные модели представляются привлекательными в силу моделирования реальной атомной структуры вещества.

2. Дискретные модели. Двухмассовая решетка

Решетка называется простой, если в элементарной ячейке находится один атом, а если несколько атомов разных сортов, то сложной.

Рассмотрим продольные колебания частиц (атомов) одномерной решетки (цепочки), когда на элементарную ячейку с размером $2a$ приходится два атома с массами m_1 и m_2 (рис. 2). Данную структуру можно

рассматривать как две подрешетки, которые совмещаются сдвигом на постоянный вектор a , являющийся параметром сложной решетки. Будем считать, что смещения частиц малы по сравнению с межатомным расстоянием, а силы взаимодействия между ними — упругие [15].

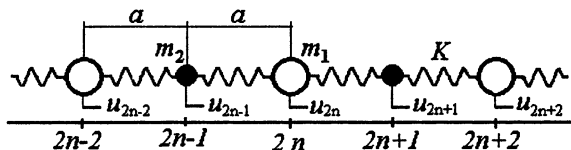


Рис. 2. Двухатомная цепочка без центра симметрии

Учитывая взаимодействие лишь ближайших соседей, из второго закона Ньютона получим уравнения движения частиц:

$$m_1 \frac{d^2 u_{2n}}{dt^2} = K(u_{2n+1} + u_{2n-1} - 2u_{2n}), \quad (4)$$

$$m_2 \frac{d^2 u_{2n+1}}{dt^2} = K(u_{2n+2} + u_{2n} - 2u_{2n+1}),$$

где K — постоянная упругого взаимодействия частиц. Заметим, что систему уравнений (4) можно представить в виде одного уравнения второго порядка с переменным коэффициентом:

$$m_c [1 + \mu \cos(\pi n)] \frac{d^2 u_n}{dt^2} - K(u_{n+1} - 2u_n + u_{n-1}) = 0, \quad (5)$$

где $m_c = (m_1 + m_2)/2$ — среднее значение, а $\mu = (m_1 - m_2)/(m_1 + m_2)$ — коэффициент модуляции массы (погонной плотности) в цепочке.

2.1. Анализ волновых процессов в периодической структуре

Так как колебания атомов разных масс могут происходить с различными амплитудами u_1 и u_2 , то решение уравнений (4) будем искать в виде двух полей, описывающих колебания легких и тяжелых частиц:

$$u_{2n} = u_1 \exp[i(2n\varphi - \omega t)], \quad u_{2n+1} = u_2 \exp[i(2n+1)\varphi - \omega t],$$

где ω — частота колебаний, φ — набег фазы на одной ячейке. Подставляя эти выражения в (4), получаем дисперсионное уравнение:

$$2\omega_{\pm}^2(\varphi) = (\omega_1^2 + \omega_2^2) \pm \sqrt{(\omega_1^2 + \omega_2^2)^2 - 4\omega_1^2 \omega_2^2 \sin^2 \varphi}. \quad (6)$$

Здесь $\omega_1 = \sqrt{2K/m_1}$, $\omega_2 = \sqrt{2K/m_2}$, $\omega_m = \sqrt{2K(m_1 + m_2)/m_1 m_2}$ — критические частоты. Дисперсионные кривые, описываемые уравнением (6), имеют две ветви (рис. 3). Нижнюю кривую $\omega_-(\varphi)$ называют акустиче-

ской ветвью колебаний (акустической модой), а верхнюю $\omega_+(\varphi)$ — оптической модой. При малых φ фазовая $v_\varphi = \omega_+/ \varphi$ и групповая $v_{гр} = d\omega_+/d\varphi$ скорости акустической моды совпадают. При $\varphi = \pm\pi/2$ (на границах первой зоны Бриллюэна) частота достигает своего максимального значения ω_1 , а групповая скорость обращается в нуль.

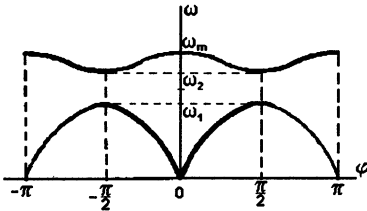


Рис. 3. Дисперсионные ветви колебаний двухатомной цепочки.

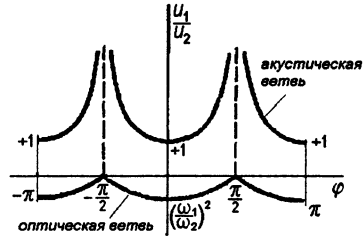


Рис. 4. Распределение амплитуд колебаний в нормальных модах.

Оптическая ветвь начинается от максимального значения ω_m . С ростом φ она опускается вниз и при $\varphi = \pm\pi/2$ достигает нижнего критического значения ω_2 . Акустическая и оптическая моды разделены полосой запрещенных частот, т. е. в области $\omega_1 < \omega < \omega_2$ уравнения (4) не имеют решения. Однако если в цепочке заменить одну или несколько частиц массы m_2 на частицы массы m_1 , т. е. ввести в структуру дефекты, то в запрещенной области частот появятся решения, которые называются *локальными модами* [14, 15].

Распределение амплитуд колебаний частиц в акустической и оптической модах описывается выражением (рис. 4)

$$\frac{u_1}{u_2} = (1 - \omega_\pm^2 / 2\omega_1^2)^{-1} \cos \varphi. \quad (7)$$

В акустической моде при низких частотах и больших длинах волн ($\varphi \ll 1$) колебания соседних частиц в цепочке происходят в фазе и имеют одинаковую амплитуду ($u_1/u_2 = 1$) (рис. 5, а). В оптической моде частицы в ячейке колеблются в противофазе $u_1/u_2 = -m_2/m_1$ (рис. 5, б), а центр масс каждой ячейки остается на месте, так как $u_1 m_1 + u_2 m_2 = 0$. Частицы m_1 образуют подрешетку, движущуюся как целое, и точно так же движется как целое подрешетка, образованная частицами m_2 .

Вблизи границы зоны Бриллюэна $\varphi \approx \pi/2$ отношение амплитуд для акустической волны стремится к бесконечности. Это означает уменьшение амплитуды колебаний легких частиц. При этом соседние частицы колеблются в фазе ($u_1/u_2 > 0$). При $\varphi = \pi/2$ амплитуда колебаний легких частиц обращается в нуль, а тяжелые частицы колеблются со сдвигом по фазе на 180° по отношению к соседним тяжелым частицам (рис. 6, а).

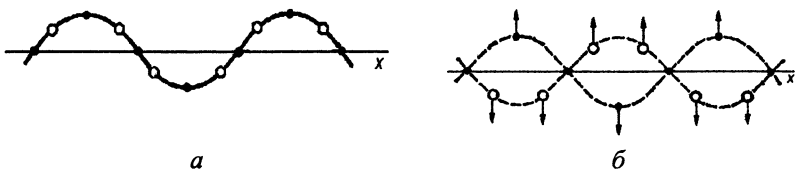


Рис. 5. Колебания частиц в длинноволновой области $\varphi \ll 1$:
акустическая ветвь (а); оптическая ветвь (б)

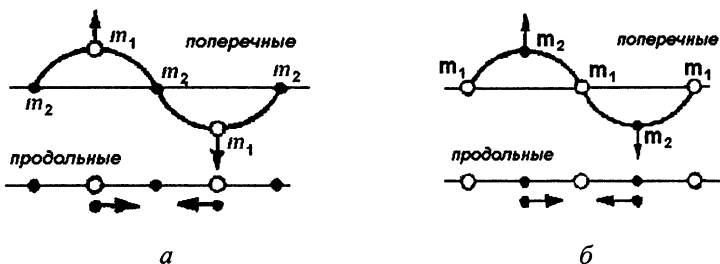


Рис. 6. Колебания частиц в коротковолновой области $\varphi \approx \pi/2$:
акустическая ветвь (а); оптическая ветвь (б)

Для оптической ветви отношение u_1/u_2 при $\varphi \approx \pi/2$ стремится к нулю, а соседние частицы колеблются в противофазе ($u_1/u_2 < 0$). Когда $\varphi = \pi/2$, имеем $u_1/u_2 = 0$. В этом случае $u_1 = 0$ и движутся только легкие частицы с амплитудой u_2 и сдвигом по фазе на 180° (рис. 6, б). Таким образом, во всем интервале значений $0 < \varphi < \pi/2$ в акустической моде соседние частицы всегда движутся в фазе, а в оптической моде соседние частицы всегда движутся в противофазе.

2.2. Континуальная аппроксимация

Проведем континуализацию дискретных уравнений движения, считая, что соседние частицы каждой подрешетки испытывают малые смещения без нарушения их ближнего порядка. В этом приближении каждая подрешетка движется как подконтинуум. Для введения такого подконтинуума принимается, что существуют гладкие, дифференцируемые нужное число раз функции¹ $u(x, t)$ и $v(x, t)$ пространственных положений частиц подрешеток в начальном и в конечном состоянии:

¹ Плавность, гладкость полей смещений относятся лишь к атомам каждой подрешетки. Соседние же атомы в каждой элементарной ячейке могут испытывать в общем случае большие относительные перемещения, что важно при изучении структурных превращений материала при больших деформациях [6]

$$u_{2n\pm 2}(t) = u(x \pm 2a, t) = u(x, t) \pm 2au_x + 2au_{xx} \pm \dots$$

$$u_{2n\pm 1}(t) = u(x \pm a, t) = v(x, t) \pm av_x + \frac{a^2}{2}v_{xx} \pm \dots \quad (8)$$

Подобное «натяжение» гладких функций на дискретную структуру приводит к тому, что ее фурье-спектр оказывается обрезанным со стороны высоких пространственных частот [10]. Удерживая в разложении (8) пространственные производные до второго порядка включительно и подставляя их в уравнения (4), приходим к системе дифференциальных уравнений в частных производных:

$$m_1 u_{tt} = Ka^2 v_{xx} - 2K(u - v),$$

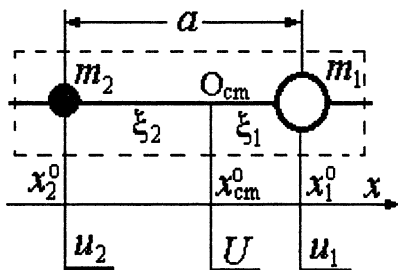
$$m_2 \left(v_{tt} + \frac{a^2}{2} v_{ttxx} \right) + Ka^2 v_{xx} = -2K(u - v) - 2Ka^2 u_{xx}. \quad (9)$$

На этом этапе фактически совершился переход от двух подрешеток к двум взаимно проникающим подконтинуумам.

2.3. Коллективные переменные

Уравнения (9) имеют несимметричный вид и трудны для физической интерпретации, поэтому в них целесообразно перейти от «индивидуальных» переменных (u и v) к «коллективным» переменным, характеризующим перемещение центра масс ячейки и относительные перемещения частиц.

Рис. 7. Введение коллективных переменных в ячейке



Введем в рассмотрение переменную U , описывающую смещение центра масс ячейки — *макросмещение* (рис. 7):

$$U = \frac{m_1 u_1 + m_2 u_2}{m_1 + m_2}. \quad (10)$$

Кроме этого, определим *микродоформацию* как средневзвешенную деформацию ячейки [10]:

$$\eta = \frac{m_1 \xi_1 u_1 + m_2 \xi_2 u_2}{m_1 \xi_1^2 + m_2 \xi_2^2} = \frac{u_1 + u_2}{a}. \quad (11)$$

Здесь $\xi_1 = am_2/M$, $\xi_2 = -am_1/M$ — начальные положения частиц относительно центра масс ячейки ($\xi_1 m_1 + \xi_2 m_2 = 0$), $M = m_1 + m_2$ — полная масса ячейки. Коллективные переменные связаны с индивидуальными переменными $u(x,t)$ и $v(x,t)$ соотношениями

$$U = \frac{m_1 u + m_2 v}{M}, \quad \eta = \frac{u - v}{a}. \quad (12)$$

В коллективных переменных длинноволновая аппроксимация уравнений (4) принимает симметричный вид:

$$\begin{aligned} \frac{\partial^2 U}{\partial t^2} - c_1^2 \frac{\partial^2 U}{\partial x^2} + \beta_1 \frac{\partial^4 U}{\partial x^2 \partial t^2} &= \chi_1 \eta + \chi_2 \frac{\partial^2 \eta}{\partial x^2} + \chi_3 \frac{\partial^4 \eta}{\partial x^2 \partial t^2}, \\ \frac{\partial^2 \eta}{\partial t^2} + c_2^2 \frac{\partial^2 \eta}{\partial x^2} + \omega_m^2 \eta + \beta_2 \frac{\partial^4 \eta}{\partial x^2 \partial t^2} &= \gamma_1 \frac{\partial^2 U}{\partial x^2} - \gamma_2 \frac{\partial^4 U}{\partial x^2 \partial t^2}, \end{aligned} \quad (13)$$

где $c_1 = \sqrt{2Ka/M}$, $c_2 = \sqrt{Ka^2(M + 2m_1)/M}$ — соответственно скорости распространения макро- и микровозмущений, $\omega_m = \sqrt{2KM/m_1 m_2}$ — максимальная частота микроколебаний, $\beta_{1,2}$ — параметры дисперсии, а $\gamma_{1,2}$ и $\chi_{1,2,3}$ — константы связи между макро- и микродвижениями.

3. Двухмассовая решетка с центром симметрии

Если в ячейку внести еще одну массу m_2 , как показано на рис. 8 ($a = 3b$), то решетка будет обладать центром симметрии.

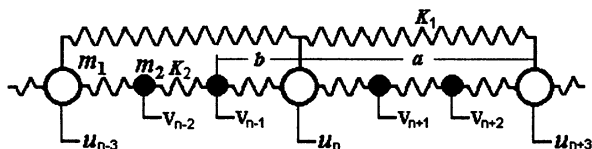


Рис. 8. Двухмассовая цепочка с центром симметрии

Колебания такой структуры будут описываться уравнениями

$$\begin{aligned} m_1 \frac{d^2 u_n}{dt^2} &= K_1 (u_{n+3} + u_{n-3} - 2u_n) + K_2 (v_{n+1} + v_{n-1} - 2u_n), \\ m_2 \frac{d^2 v_{n+1}}{dt^2} &= K_2 (u_n + v_{n+2} - 2v_{n+1}), \\ m_2 \frac{d^2 v_{n-1}}{dt^2} &= K_2 (u_n + v_{n-2} - 2v_{n-1}). \end{aligned} \quad (14)$$

Последние два уравнения получаются одно из другого заменой индекса n на $n - 2$ и эквивалентны одному уравнению:

$$m_2 \frac{d^2}{dt^2} (v_{n+1} + v_{n-1}) = K_2 (2u_n + v_{n+2} + v_{n-2} - 2v_{n+1} - 2v_{n-1}). \quad (15)$$

Как и в предыдущем случае, воспользуемся континуальной аппроксимацией перемещений:

$$u_{n \pm 3}(t) = u(x \pm a, t) = u(x, t) \pm au_x + \frac{1}{2} a^2 u_{xx} \pm \dots$$

$$v_{n \pm k}(t) = u(x \pm kb, t) = u(x, t) \pm kbu_x + \frac{1}{2} (kb)^2 u_{xx} \pm \dots$$

Тогда уравнения движения (14) в квазиконтинуальном приближении примут вид

$$m_1 u_{tt} - K_1 a^2 u_{xx} = 2K_2 (v - u),$$

$$m_2 (v_{tt} + b^2 v_{xxxx}) - K_2 b^2 v_{xx} = -K_2 (v - u), \quad (16)$$

где m_2 , b , K_2 — параметры, характеризующие внутреннюю структуру среды.

В данной цепочке координата центра масс ячейки совпадает с координатой тяжелой частицы m_1 , а внутренняя переменная, характеризующая микродеформацию среды, имеет вид

$$\eta = (v - u)/b. \quad (17)$$

В коллективных переменных уравнения движения выглядят следующим образом:

$$u_{tt} - c_1^2 u_{xx} = \gamma_1 \eta,$$

$$\eta_{tt} - c_2^2 \eta_{xx} + \beta^2 \eta_{xxxx} + \omega_m^2 \eta = \gamma_2 u_{xx}, \quad (18)$$

где $c_1 = \sqrt{K_1 a^2 / m_1}$, $c_2 = \sqrt{K_2 b^2 / m_2}$ — соответственно скорости волн макро- и микровозмущений, $\omega_m^2 = \sqrt{K_2 (m_1 + 2m_2) / m_1 m_2}$, $\beta = b/\sqrt{2}$ — пространственный масштаб микроструктуры, $\gamma_{1,2}$ — константы связи макро- и микродеформации. В области низких частот $\omega < \omega_m$ внутренние возмущения являются нераспространяющимися, а при $\omega > \omega_m$ представляют собой волны микродеформации.

4. Решетка с внутренними вращениями частиц

Решетка из «механических» диполей (решетка молекулярного типа) также относится к сложным (рис. 9). В ней существенную роль играют

нецентральные (моментные) взаимодействия частиц и ротационные степени свободы. В конце 30-х годов XX века Я. И. Френкель рассмотрел цепочку из диполей с закрепленными центрами тяжести и показал, что в ней могут распространяться «волны вращательных качаний». Первая микроскопическая модель, учитывающая взаимодействие трансляционных и вращательных (ориентационных) колебаний в молекулярной решетке, была предложена в работе [2]. Авторами было показано, что в таких решетках распространяются в основном смешанные ориентационно-трансляционные колебания, частоты которых зависят как от массы, так и от момента инерции частиц в решетке.

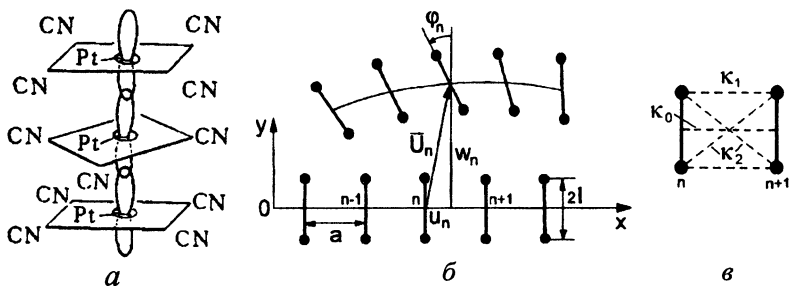


Рис. 9. Структура квазиодномерного слоистого кристалла (а).
 Механическая модель молекулярного кристалла (б).
 Схема силового взаимодействия частиц (в)

Многие динамические свойства «ориентированных» сред, состоящих из анизотропных молекул и имеющих кристаллическую решетку, относящуюся к типу молекулярных решеток (рис. 9, а), можно выяснить на примере плоских колебаний цепочки «гантелевидных» частиц [19, 20]. Примерами таких сред служат сероуглерод (CS), нитрат калия (KNO), нитрит натрия (NaNO_2), бензол (C_6H_6), нафталин (C_{10}H_8), жидкие кристаллы, композитные материалы, упругие ферромагнитные и ферроэлектрические кристаллы [1, 7, 9].

Рассмотрим плоские колебания цепочки «гантелевидных» частиц («диполей»), в исходном состоянии расположенных на расстоянии a друг от друга и ориентированных перпендикулярно линии, соединяющей их центры масс (рис. 9, б). Диполь моделируется двумя одинаковыми точечными массами $m/2$, соединенными жесткими невесомыми стержнями длины $2l$, и его центр масс совпадает с геометрическим центром.

Движение каждой частицы, имеющей три степени свободы, складывается из горизонтального u_n и вертикального v_n перемещений ее центра масс и поворота на угол φ_n (см. рис. 9, б). Считается, что каждая масса диполя взаимодействует со всеми массами двух ближайших сосе-

дей по цепочке. Это взаимодействие моделируется с помощью нелинейных упругих пружин (рис. 9, в) и описывается кубическим потенциалом, зависящим от относительных удлинений пружин, возникающих при отклонениях масс от положения равновесия. Предполагается, что частицы колеблются вблизи положений равновесия и углы их поворотов малы $\varphi \ll \pi/2$. В общем случае пружины в начальном состоянии могут быть деформированы. Введение предварительных деформаций пружин позволяет учесть начальные внутренние напряжения, присущие достаточно широкому кругу материалов и дает дополнительную возможность варьирования параметров микромоделей. Следует отметить, что начальные растяжения пружин не могут быть произвольными. Условие нахождения конфигурации в равновесии накладывает на них определенные ограничения [19].

В длинноволновом континуальном пределе (когда характерная длина волны много больше расстояния между частицами $\lambda \gg a$) система дифференциальных уравнений, описывающая распространение и взаимодействие различных типов волн в рассматриваемой среде, принимает вид [19]:

$$\begin{aligned} u_{tt} - c_u^2 u_{xx} &= \frac{1}{2} \frac{\partial}{\partial x} (3\alpha_1 u_x^2 + \alpha_2 w_x^2 + \alpha_3 \varphi^2 + 2\alpha_4 w_x \varphi), \\ w_{tt} - c_w^2 w_{xx} - \beta_1 \varphi_x &= \frac{\partial}{\partial x} \left(\alpha_2 u_x w_x + \alpha_4 u_x \varphi + \frac{1}{2} \alpha_5 w_x^3 + 3\alpha_6 w_x^2 \varphi \right), \\ \varphi_{tt} - c_\varphi^2 \varphi_{xx} + \omega_0^2 \varphi + \beta_2 w_x &= -\alpha_3 u_x \varphi - \alpha_4 u_x w_x - \alpha_6 w_x^3. \end{aligned} \quad (19)$$

Здесь $c_{u,w,\varphi}$ — скорости распространения продольных, поперечных и ротационных волн, ω_0 — критическая частота спиновой волны, $\beta_{1,2}$ — коэффициенты межмодовой связи, α_j ($j = 1+6$) — коэффициенты нелинейностей. Первые два уравнения описывают упругие (продольные и поперечные) волны и являются следствиями закона сохранения импульса среды. Третье уравнение описывает волну микроповоротов (спиновая волна), которая отсутствует в решетке из точечных частиц. Правые части уравнений (19) описывают нелинейные взаимодействия между различными типами волн.

Отметим, что структура уравнений (19) полностью совпадает со структурой уравнений континуума Коссера (3), в котором геометрические размеры частиц в явном виде не фигурируют. Форма частиц сказывается лишь на выражениях, входящих в уравнения коэффициентов:

$$\frac{c_u}{c_0} = \left[1 + \frac{2(a/R)^4 k_{20} + (\delta'_0 + 2\delta'_1 k_{10})(1 - (a/R)^2)}{1 + 2k_{10}} \right]^{1/2},$$

$$\frac{c_w}{c_0} = \left[\frac{(8a^2 l^2 / R)^4 k_{20} - 2\delta'_{,2} (1 - (a/R)^2) k_{20}}{1 + 2k_{10}} \right]^{1/2}, \quad (20)$$

$$\frac{c_\varphi}{c_0} = \left[\frac{2k_{10} + (a/R)^6 k_{20} + 2\delta'_{,1} k_{10}}{1 + 2k_{10}} \right]^{1/2}.$$

Здесь $c_0 = \sqrt{(k_0 + 2k_1)/m}$ — скорость распространения продольных волн в цепочке материальных точек, $k_{j0} = k_j / k_0$ — безразмерные параметры упругого взаимодействия, $\delta'_{,0,1,2}$ — начальные растяжения пружин.

Связь скоростей волн с параметрами внутренней структуры среды подробно проанализирована в работах [19, 20]. С помощью соотношений (20) по известным характеристикам упругих волн можно восстановить параметры микромоделей, и тем самым будет решена задача параметрической идентификации системы. Дисперсионные свойства нормальных волн в модели молекулярного кристалла идентичны дисперсионным свойствам волн в континууме Коссера (см. рис. 1).

5. Экспериментальные исследования

На важность учета микроструктуры среды, а именно вращательных степеней свободы частиц, в середине 30-х — начале 40-х годов обратили внимание физики-экспериментаторы. Ими наблюдался эффект изменения длины волны рассеянного света в жидкости, связанный с флуктуациями ориентации анизотропных молекул, и было отмечено, что оси молекулы могут поворачиваться на значительную величину. В дальнейшем было установлено, что и у кристаллов некоторых органических веществ спектр рассеяния малых частот связан с вращательными колебаниями молекул [2]. Наиболее интенсивен спектр рассеяния был у веществ, молекулы которых обладают большой оптической анизотропией (сероуглерод, нафталин, бензол). В конце 50-х годов стали проводиться опыты по наблюдению оптико-акустического эффекта в жидкостях и твердых телах, в частности опыты по изучению спектральной зависимости оптико-акустического эффекта в сегнетоэлектрических кристаллах. Эти исследования представляются интересными с точки зрения проблем, связанных с молекулярным механизмом пьезоэлектрического явления. Однако этот вопрос до конца не изучен [7].

Первые эксперименты с акустическими волнами в твердых телах с микроструктурой были проведены в начале 1970-х годов [22, 23]. Авторы установили корреляцию между размером зерна в различных металлах и алюминиевых сплавах и дисперсионным параметром акустической вол-

ны. Дисперсию ультразвуковых волн наблюдали также в искусственном зернистом композите — ферритовая дробь в эпоксидной смоле [24].

Среды с внутренней структурой содержат новые константы упругости. Для определения этих констант необходимо провести соответствующее число независимых экспериментальных измерений [25]. Для этого использовались как статические, так и динамические (волновые) эксперименты. Попытки измерения констант упругости на основе статических экспериментов не увенчались успехом из-за их недостаточной точности. Более точны динамические эксперименты, особенно ультразвуковые. В работах [22, 23] с помощью акустического метода определены параметры микроструктуры для некоторых металлов и сплавов. В [26] были измерены линейные модули зернистого композита – алюминиевая дробь в эпоксидной матрице. Следует заметить, что отсутствие экспериментальных данных о материальных константах структурированных сред и слабая разработанность методов их экспериментального исследования [27, 28] являются одними из основных факторов, сдерживающих изучение моделей неклассических сред и их практическое применение.

Заключение

Из рассмотренных выше примеров следует, что сложную решетку (дискретную модель среды с внутренней структурой) можно рассматривать либо как микронеоднородную среду, либо как однородную среду с внутренними степенями свободы. В последнем случае для ее описания удобно использовать коллективные переменные, в которых уравнения движения приобретают более симметричную форму. В качестве коллективных переменных удобно выбирать смещение центра масс ячейки (макросмещение) и относительные перемещения частиц внутри ячейки (микродеформации). Коллективные переменные описывают динамику сложных структур как движение взаимопроникающих подконтинуумов. Теорию движения подконтинуумов можно строить по аналогии с теорией многофазных (многокомпонентных) сред.

Работа выполнена по гранту Президента РФ для поддержки ведущих научных школ России (проект НШ-1638.2003.8) и при частичной финансовой поддержке грантами РФФИ (проект 04-02-17156), министерства образования РФ (проект Е02-3.5-224) и ФЦП «Интеграция» (проект Б-0106/1247).

Литература

1. Гусев А. И., Ремпель А. А. Нанокристаллические материалы. М.: Физматлит, 2001. 224 с.
2. Гросс Е., Коршунов А. Вращательные колебания молекул в кристаллической решетке органических веществ и спектры рассеяния // ЖЭТФ. 1946. Т. 16, № 1. С. 53—59.
3. Ансельм А. И., Порфирьева Н. Н. Ориентационно-трансляционные волны в молекулярных кристаллах // ЖЭТФ. 1949. Т. 19, № 5. С. 438—446.

4. Слепян Л. И. Механика трещин. Л.: Судостроение, 1990. 296 с.
5. Морозов Н. Ф., Паукино М. В. Дискретные и гибридные модели механики разрушения. СПб: Изд-во СПбГУ, 1995. 160 с.
6. Аэро Э. Л. Существенно нелинейная микромеханика среды с изменяемой периодической структурой // Успехи механики. 2002. Т. 1, № 3. С. 130—176.
7. Maugin G. A. Nonlinear waves in elastic crystals. Oxford: Oxford Univ. Press, 1999. 320 p.
8. Eringen A. C. Microcontinuum field theories. 1: Foundation and solids. Springer Verlag, New York Inc., 1999. 326 p.
9. Маневич Л. И., Рягусов С. В. Нелинейная плоская динамика молекулы полиэтилена // Физика твердого тела. 1992. Т. 34, № 5. С. 1554—1560.
10. Кунин И. А. Теория упругих сред с микроструктурой. М.: Наука, 1975. 416 с.
11. Аэро Э. Л., Кувшинский Е. В. Основные уравнения теории упругости сред с вращательным взаимодействием частиц // Физика твердого тела. 1960. Т. 2, № 7. С. 1399—1409.
12. Пальмов В. А. Основные уравнения теории несимметричной упругости // ПММ. 1964. Т. 28, № 3. С. 401—408.
13. Лисина С. А., Потапов А. И. Уравнения нелинейной динамики микрополярной среды. Вариационный подход // Изв. вузов. Северо-Кавказ. регион. Естественные науки. Нелинейные проблемы механики сплошных сред. Ростов-на-Дону, 2003. С. 249—255.
14. Борн М., Хуан К. Динамическая теория кристаллических решеток. М.: ИЛ, 1958.
15. Косевич А. М. Теория кристаллической решетки (физическая механика кристаллов). Харьков: Вища школа, 1988. 304 с.
16. Askar A. Lattice dynamics foundation of continuum theory. Singapore: World-Scientific, 1985.
17. Седов Л. И. Механика сплошной среды. Т. 1. М.: Наука, 1973. 536 с.
18. Ахизер А. И., Барьяхтар В. Г., Пелетминский С. В. Спиновые волны. М.: Наука, 1967. 368 с.
19. Дразунов Т. Н., Павлов И. С., Потапов А. И. Ангармонические взаимодействия упругих и ориентационных волн в одномерных кристаллах // ФТТ. 1997. Т. 39, № 1. С. 137—144.
20. Potapov A. I., Pavlov I. S., Maugin G. A. Nonlinear wave interactions in 1D crystals with complex lattice // Wave Motion. 1999. V. 29. P. 297—312.
21. Лисина С. А., Потапов А. И., Нестеренко С. Ф. Нелинейная гранулированная среда с вращением частиц. Одномерная модель // Акуст. журн. 2001. Т. 47, № 5. С. 666—674.
22. Савин Г. Н., Лукашев А. А., Лыско Е. М., Веремеенко С. В., Агафьев Г. Г. Распространение упругих волн в континууме Коссера со стесненным вращением частиц // Прикладная механика. 1970. Т. 6, № 6. С. 37—41.
23. Савин Г. Н., Лукашев А. А., Лыско Е. М. Распространение упругих волн в твердом теле с микроструктурой // Прикладная механика. 1970. Т. 6, № 7. С. 48—52.
24. Ерофеев В. И., Родюшкин В. М. Наблюдение дисперсии упругих волн в зернистом композите и математическая модель для ее описания // Акуст. журн. 1992. Т. 38, № 6. С. 1116—1117.
25. Askar A. Molecular crystals and the polar theories of continua: experimental values of material coefficients for KNO_3 // Int. J. Eng. Sci. 1972. V. 10. P. 293—300.
26. Gauthier R. D. Experimental investigations on micropolar media // Mechanics of micropolar media. Singapore: World Scientific, 1982. P. 395—463.
27. Yang J. F. S., Lakes R. S. Experimental study of micropolar and couple stress elasticity in compact bone bending // J. Biomechanics. 1982. V. 15. № 2. P. 91—94.
28. Потапов А. И., Родюшкин В. М. Экспериментальное исследование волн деформации в материалах с микроструктурой // Акуст. журн. 2001. Т. 47, № 1. С. 407—412.

ПРИМЕНЕНИЕ ЯВЛЕНИЯ САМОДИФРАКЦИИ СВЕТОВЫХ ВОЛН В ФОТОРЕФРАКТИВНЫХ КРИСТАЛЛАХ ДЛЯ ОБРАБОТКИ НЕСТАЦИОНАРНЫХ СПЕКЛОВЫХ ПОЛЕЙ

*Ю. Н. Кульчин, Р. В. Ромашко, С. М. Шандаров, Н. И. Буримов,
Ю. Ф. Каргин, В. В. Волков, О. Т. Каменев*

Спекл-эффект [1], возникающий в результате рассеяния когерентного излучения в оптически неоднородных средах и последующей интерференции рассеянных волн, широко используется при создании бесконтактных измерительных систем (ИС) для определения различных физических величин [2—4]. В таких системах, как правило, пространственные изменения спекл-картины, вызванные воздействием измеряемой физической величины на исследуемый объект (или чувствительный элемент ИС), преобразуются в модулированный по интенсивности оптический сигнал. Такое преобразование достигается за счет применения для обработки спекловых полей амплитудных [5] или голографических [6] корреляционных фильтров, записываемых в фоторегистрирующих средах. Однако случайные изменения параметров исследуемого объекта, вызванные флуктуациями внешних условий, могут привести к полной перестройке спекловой картины и, как следствие, к последующей неправильной работе корреляционного фильтра. Если же придать корреляционному фильтру свойства адаптивности, т. е. сделать его способным подстраиваться под неконтролируемые изменения внешних факторов, то проблема устойчивой длительной работы измерительной системы была бы решена. На сегодняшний день существует ряд схем, реализующих адаптивные корреляционные фильтры (АКФ) для обработки спекловых картин. В некоторых из них осуществляется непрерывная запись спекл-картины и ее последующая цифровая [7] или аналоговая [8] обработка. К недостаткам таких способов обработки стоит отнести необходимость использования высокоскоростных ЭВМ, что в результате ведет к ограничениям в использовании такой измерительной системы в реальном времени. Другой изъян — это недостаточно высокая пространственная разрешающая способность ПЗС-матриц, используемых для регистрации спекл-картины, что влечет за собой снижение чувствительности ИС.

Применение перезаписываемых голографических материалов для формирования динамических голограмм (ДГ), способных подстраиваться под случайные неконтролируемые воздействия, позволит решить задачу адаптивной обработки спекловых полей оптических ИС. В качестве перезаписываемых материалов могут быть использованы фоторефрак-

тивные кристаллы (ФРК), коэффициент преломления которых изменяется под действием света. Однако в большинстве схем формирования динамических голограмм используется двухволновое смешение, необходимым условием которого является наличие мощного стационарного опорного светового луча [9], что делает измерительную систему более сложной и менее надежной.

В настоящей статье представлено исследование двух возможных схем формирования динамических голограмм в фоторефрактивных кристаллах без использования дополнительной опорной волны. Обе эти схемы основаны на явлении самодифракции световой волны в фоторефрактивном кристалле. Первая схема строится в пропускающей геометрии, вторая — в отражательной.

1. Модель работы адаптивного корреляционного фильтра

1.1. Формирование АКФ в пропускающей геометрии на основе эффекта фаннинга. Появление фаннинга в ФРК связано с рассеянием когерентной световой волны на поверхностных и объемных неоднородностях и дефектах кристалла. Это рассеяние создает когерентный шум изначально очень малой интенсивности, компоненты которого интерферируют с введенной в ФРК световой волной, образуя хаотически ориентированные ДГ [10]. Последующая дифракция основной волны на этих голограммах приводит к усилению волн рассеяния и, как следствие, к увеличению дифракционной эффективности ДГ. Наблюдающаяся конкуренция волн рассеяния обуславливает создание стационарного набора ДГ, которые и определяют пространственно-угловой спектр волн фаннинга. Любое изменение пространственного распределения фазы или амплитуды введенной в ФРК волны приведет к нарушению стационарного режима фаннинга вследствие отстройки от брэгговского резонанса, что приведет к уменьшению мощности волны фаннинга. Для увеличения эффективности формируемых ДГ и, соответственно, для усиления эффекта фаннинга к кристаллу прикладывают внешнее знакопеременное электрическое поле [11].

Пусть в ФРК вводится волна, имеющая спекловое распределение интенсивности. Известно, что любая спекловая волна может быть представлена в виде суперпозиции большого числа плоских волн-фазоров [1]. Если объект, порождающий спекловую волну, изменяет свое положение относительно падающего на него когерентного светового луча, то фаза каждого фазора изменяется. Это приводит к изменению распределения амплитуды результирующей спекловой волны. Мощность волн фаннинга будет пропорциональна значению функции корреляции распределений амплитуды спеклового поля на входе в ФРК до и после воздействия на исследуемый объект. Можно показать, что если объект, по-

рождающий спекловую волну, подвержен синусоидальным вибрациям на частоте Ω , то изменение мощности дифрагированного на ДГ излучения во времени может быть представлено в виде

$$\Delta P_{fan}(t) \sim m \exp(\Gamma L) \sum_q c_q d_q \sum_{k=1}^{\infty} J_{2k}(\Phi_q) \sin 2k\Omega t, \quad (1)$$

где m — коэффициент начального рассеяния светового луча в фоторефрактивном кристалле; Γ — коэффициент двухволнового взаимодействия; L — длина кристалла; c_q и d_q — коэффициенты разложения светового спеклового поля, сформированного после рассеяния на стационарном и смещенном объекте соответственно; $J_k(\xi)$ — функция Бесселя k -го порядка; Φ_q — амплитуда изменения дополнительной фазы q -го фазора. Как видно из выражения (1), электрический сигнал фотодетектора, который будет регистрировать мощность волн фаннинга, должен содержать только четные гармоники частоты вибраций исследуемого объекта. Анализ выражения (1) показывает, что максимальный вклад в переменный сигнал дает первый член ряда, пропорциональный $\sin 2\Omega t$.

Необходимо отметить, что каждая ДГ формируется в кристалле в течение определенного времени τ_R . Таким образом, по прошествии этого времени старый набор ДГ исчезнет и запишется новый, соответствующий измененному распределению амплитуды спеклового поля. Соответственно, на фотодетекторе появится переменный сигнал лишь в том случае, если изменения параметров спеклового поля будут происходить быстрее, чем τ_R . Все же медленные изменения ($\tau > \tau_R$) не будут влиять на мощность сигнала фаннинга. Тем самым достигается адаптивность измерительной системы к медленным изменениям параметров обрабатываемого оптического сигнала, вызванным неконтролируемыми воздействиями внешних факторов.

1.2. Формирование АКФ в отражательной геометрии. Как было отмечено в п. 1.1, для усиления дифракционной эффективности динамических голограмм, формируемых в фоторефрактивных кристаллах в пропускающей геометрии необходимо прикладывать к кристаллу внешнее электрическое поле. Напряженность этого поля иногда достигает 10—20 кВ/см. Вместе с тем, эффективность голограммы может быть существенно повышена в случае ее записи в отражательной геометрии, когда лучи распространяются в кристалле навстречу друг другу [12]. При этом могут быть достигнуты высокие коэффициенты усиления без приложения к кристаллу внешнего электрического поля [13]. Любая световая волна, введенная в ФРК, будет частично отражаться от выходной грани кристалла. В результате в кристалле будут распространяться две когерентные волны в противоположных направлениях. В результате в толще фоторефрактивного кристалла сформируется ДГ, на которой введенная в кристалл световая волна будет отражаться. Если эта свето-

вая волна имеет спекловый характер распределения амплитуды, то дифракционная эффективность формируемой в кристалле ДГ будет пространственно неоднородной. В этом случае мощность отраженного на ДГ излучения будет определяться значением функции корреляции распределений амплитуды спеклового поля до и после его изменения. Можно показать, что в случае синусоидальных модуляций фазы спекловой волны на частоте Ω изменение мощности отраженного на ДГ излучения будет определяться выражением

$$\Delta P_{\text{refl}}(t) \sim [\exp(\Gamma L) - 1] \sum_q J_0(\Phi_q) J_2(\Phi_q) \sin 2\Omega t. \quad (2)$$

Как видно из выражения (2), мощность отраженного света модулируется на частоте второй гармоники 2Ω , так же как и в случае, описанном в п. 1.1.

2. Эксперименты

2.1. Пропускающая геометрия. Схема экспериментальной установки представлена на рис. 1, а. Спекловая волна, полученная в результате рассеивания излучения He-Ne лазера ($\lambda = 633$ нм) на шероховатой поверхности *Ш*, после прохождения поляризатора *П* фокусировалась линзой *Л* на входную грань фоторефрактивного кристалла $\text{Bi}_{12}\text{TiO}_{20}:\text{Ca}:\text{Ga}$ среза (110). Длина кристалла — 8 мм, размеры поперечного сечения — $1,5 \times 5$ мм². Внешнее знакопеременное электрическое поле с амплитудой 20 кВ/см и частотой от 5 до 30 Гц прикладывалось к кристаллу вдоль оси $[1\ 1\ 0]$ посредством серебряных электродов, напыленных на боковые грани кристалла. Мощность дифрагированного на ДГ излучения (фаннинга) регистрировалась фотоприемником *ФП*. Поверхность *Ш* приводилась в колебательное движение с частотой $\Omega/2\pi = 1000$ Гц электродинамическим вибровозбудителем *ЭВ*. Спектральный анализ сигнала фотоприемника показал, что максимальный переменный сигнал наблюдается на частоте второй гармоники 2Ω . На рис. 2 представлена зависимость амплитуды модуляции мощности сигнала фаннинга на второй гармонике от амплитуды дополнительного фазового набега спекловой волны, вызванного колебаниями рассеивающей поверхности (кривая *1*). Как видно, эти результаты находятся в хорошем соответствии с теоретическими расчетами, выполненными в соответствии с выражением (1).

Временная зависимость мощности сигнала фаннинга, которая отражает динамику формирования голограммы в ФРК, представлена на рис. 3. На протяжении всего временного интервала спекловая картина на входе ФРК была неизменной, и лишь в момент времени $t = 6$ с распределение амплитуды спекловой волны было резко изменено. Это привело к соответствующему уменьшению мощности сигнала фаннинга из-за несоответствия нового распределения параметров волны старой ДГ. Од-

нако в течение времени записи (~ 2 с) старая ДГ исчезла и сформировалась новая, соответствующая новым условиям. В результате этого сигнал фаннинга вернулся к первоначальному значению, что демонстрирует адаптивные свойства корреляционных фильтров на основе ДГ, формируемых в фоторефрактивных кристаллах.

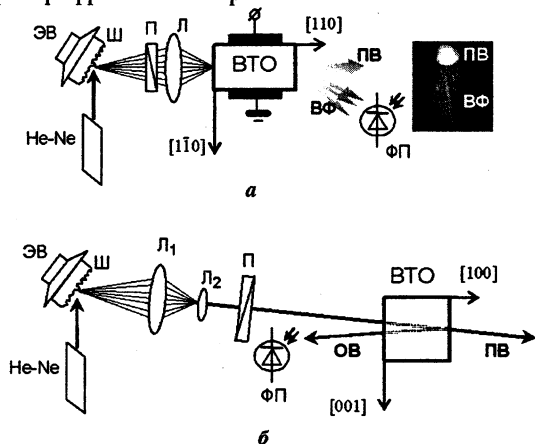


Рис. 1. Схема формирования адаптивного корреляционного голографического фильтра в фоторефрактивном кристалле без использования внешнего опорного луча: *а* — пропускающая геометрия; *б* — отражательная геометрия. He-Ne — лазер; ЭВ — электродинамический вибровозбудитель; Ш — шероховатая поверхность; Л, Л₁, Л₂ — линзы; П — поляризатор; ВТО — кристалл $\text{Bi}_{12}\text{TiO}_{20}$; ФП — фотоприемник; ПВ — прошедшая волна; ОВ — отраженная волна; ВФ — волны фаннинга.

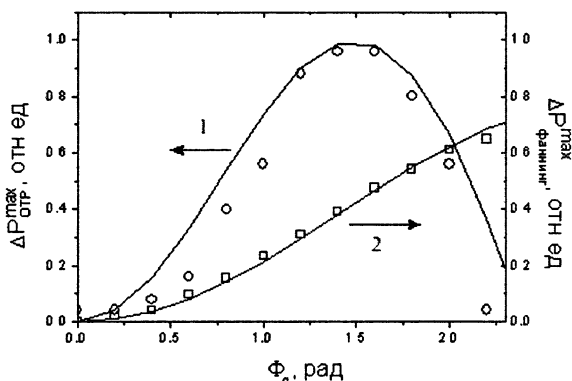


Рис. 2. Зависимость амплитуды модуляции мощности излучения, дифрагированного на ДГ, от амплитуды модуляции фазы спекловой волны для случаев пропускающей (1) и отражательной (2) геометрий. Сплошные линии — расчет, маркеры — экспериментальные результаты.

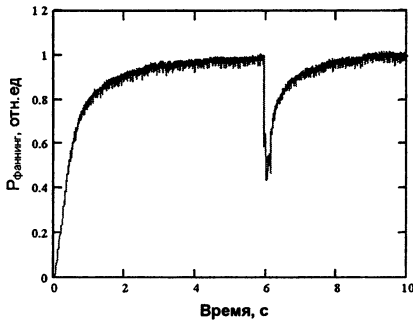


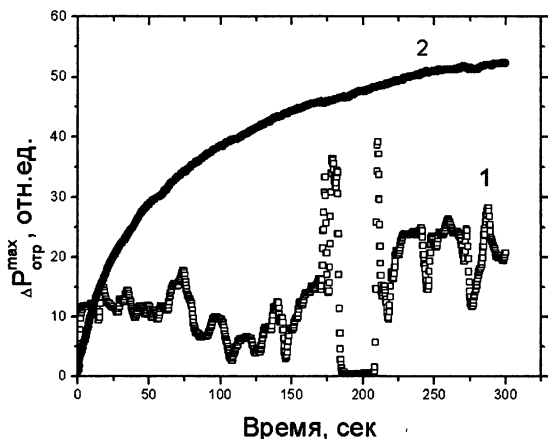
Рис. 3. Временная зависимость мощности сигнала фаннинга.

2.2. Отражательная геометрия. Схема экспериментальной установки показана на рис. 1, б. Расходящаяся спекловая волна, полученная в результате рассеяния на шероховатой поверхности, преобразуется в квазиплоскую волну посредством линз L_1 . Линза L_2 преобразует ее в слабосходящийся узкий световой пучок, который после прохождения поляризатора P попадает в фоторефрактивный кристалл $Bi_{12}TiO_{20}$ среза (100) длиной 8 мм. Посредством поляризатора угол между плоскостью поляризации входной волны и осью [001] кристалла задавался равным 30 град, что соответствовало максимальному значению коэффициента двухволнового взаимодействия. Волна, прошедшая сквозь кристалл (PB), и волна, отраженная от его выходной грани (OB), распространяются внутри кристалла в противоположных направлениях, интерферируют и формируют динамическую голограмму отражательного типа. Небольшая клиновидность кристалла позволила разделить в пространстве пучки, отраженные от его входной и выходной граней (пучок, отраженный от входной грани, не показан на рисунке). Благодаря большому углу между взаимодействующими лучами (~180 град) период следования интерференционных полос на порядок меньше, чем в случае пропускающей геометрии, когда угол между лучами много меньше 90 град. С уменьшением пространственного периода ДГ увеличивается ее дифракционная эффективность, что позволяет в свою очередь не использовать для этой цели внешнее электрическое поле.

С целью уточнения спектрального состава переменного сигнала была исследована динамика изменения амплитуды модуляции мощности отраженной на голограмме световой волны на первой и второй гармониках частоты модуляционного сигнала ($\Omega/2\pi = 2870$ Гц) в процессе формирования в кристалле отражательной решетки. На рис. 4 приведены временные зависимости этих сигналов. Видно, что амплитуда первой гармоники (кривая 1) достигает гораздо меньших значений по сравнению со второй гармоникой (кривая 2) и, кроме того, ведет себя нестабильно. Присутствие в переменном сигнале первой гармоники объясняется модуляцией спеклов на апертуре оптических элементов, а также

случайными флуктуациями спекловой картины. Основной же стабильный переменный сигнал наблюдается на второй гармонике, что находится в соответствии со сделанными ранее выводами. Экспериментально полученная зависимость амплитуды модуляции мощности отраженного пучка на частоте второй гармоники от амплитуды модуляции фазы входной волны представлена на рис. 2 (кривая 2). Как видно, полученные результаты находятся в хорошем соответствии с численными расчетами, выполненными на основе выражения (2).

Рис. 4. Временная зависимость глубины модуляции мощности отраженного от ДГ светового пучка на частоте первой (1) и второй (2) гармоник.



Заключение

В настоящей работе исследованы процессы самодифракции световых волн на ДГ, формируемых в фоторефрактивных кристаллах в двух геометриях — пропускающей и отражательной. В обоих случаях ДГ формируются в кристаллах без внешних опорных лучей, а с использованием либо рассеянного на неоднородностях излучения (пропускающая геометрия), либо частично отраженного от выходной грани кристалла (отражательная геометрия). Это существенно упрощает схему формирования адаптивных корреляционных фильтров на основе ДГ.

Для увеличения дифракционной эффективности ДГ в пропускающей геометрии используется внешнее электрическое поле, которое прикладывается к кристаллу. В случае же отражательной геометрии дифракционная эффективность голограммы повышается за счет уменьшения пространственного периода интерференционной картины. Это снимает необходимость в использовании внешнего электрического поля, что делает измерительные системы, построенные на ДГ отражательного типа, более простыми и менее энергозатратными. Кроме того, существует ряд приложений, где использование высокого напряжения вообще недопустимо.

Было показано теоретически и подтверждено экспериментально, что переменный сигнал на выходе адаптивного корреляционного фильтра, сформированного в ФРК, в обоих случаях наблюдается на второй гармонике частоты модуляции параметров обрабатываемой спекловой волны.

Полученные результаты могут быть использованы для создания высокоэффективных интерферометрических измерительных систем, способных в реальных условиях осуществлять длительный мониторинг параметров физических полей в реальном времени.

Эти исследования проводились при частичной финансовой поддержке грантов РФФИ и БРФФИ (проекты 02-01-00710-а и 02-02-81044), Министерства образования и науки РФ (проект Е02-3.2-486), а также Президентской программы по поддержке молодых российских ученых России (проект МК-3865.2004.2).

Литература

1. Goodman J. W. Statistical properties of laser speckle pattern // Laser speckle and related phenomena. N.-Y.: Springer-Verlag, 1975. P. 10—75.
2. Li E. B., Tieu A. K., Wang K. F. Dynamic laser speckle method for determining the relative velocity between two objects // Opt. Commun. 2003. V. 219. P. 1—8.
3. Rose B., Imam H., Hanson S. G. Non-contact laser speckle sensor for measuring one- and two-dimensional angular displacement // J. Optics. 1998. V. 29. P. 115—120.
4. Weiss B., Klein M., Sossna E., Volland B., Rangelow I. W. Noncontacting laser-based techniques for the determination of elastic constants of thin silicon membranes // Microelectronic Engineering. 2001. V. 57—58. P. 475—479.
5. Быковский Ю. А., Витрик О. Б., Кульчин Ю. Н. Амплитудная пространственная фильтрация в обработке сигналов одноволоконного многомодового интерферометра // Квантовая электроника. 1990. Т. 17, № 10. С. 1377—1378.
6. Быковский Ю. А., Кульчин Ю. Н., Витрик О. Б., Ларкин А. И. Голографическая согласованная фильтрация сигналов интерференционных датчиков на многомодовых волоконных световодах // Квант. электроника. 1990. Т. 17, № 1. С. 95—98.
7. Spillman Jr. W. B., Huston D. R., Wu J. Seismic event monitoring using very long guage integrating fiber optic sensors // Proc. SPIE. 2001. V. 4357. P. 54—65.
8. Kulchin Yu., Vitrik O., Kirichenko O., Kamenev O., Petrov Yu., Maksaev O. Method of single fiber multimode interferometer speckle signal processing // Opt. Engineering. 1997. V. 36, № 5. P. 1494—1499.
9. Delaye P., Blouin A., de Montmorillon L. A., Biaggio I., Drolet D., Monchalain J. P., Roosen G. Detection of ultrasonic vibrations on rough surface through the photorefractive effect // Proc. SPIE. 1996. V. 2782. P. 464—474.
10. Cronin-Golomb M., Yariv A. Optical limiters using photorefractive nonlinearities // J. Appl. Phys. 1985. V. 57. P. 4906—4910.
11. Stepanov S. I., Petrov M. P. Efficient unstationary holographic recording in photorefractive crystals under an external alternating electric field // Opt. Commun. 1985. V. 53. P. 292—295.
12. Kukhtarev N. V., Chen B. S., Venkateswarlu P., Salamo G. J., Klein M. B. Reflection holographic gratings in (111) cut $\text{Bi}_{12}\text{TiO}_{20}$ crystal for real-time interferometry // Opt. Commun. 1993. V. 104. P. 23—28.
13. Агеев Е. Ю., Шандаров С. М., Веретенников С. Ю., Мартыянов А. Г., Карташов В. А., Камшилин А. А., Прокофьев В. В., Шепелевич В. В. Двухволновое взаимодействие на отражательной решетке в кристалле $\text{Bi}_{12}\text{TiO}_{20}$ // Квант. электроника. 2001. Т. 31. С. 343—345.

ИНЖИНИРИНГ ЭЛЕКТРОННЫХ ВОЛНОВЫХ ПАКЕТОВ В АТОМАХ И МОЛЕКУЛАХ ДЛЯ УПРАВЛЕНИЯ ГЕНЕРАЦИЕЙ АТТОСЕКУНДНЫХ ИМПУЛЬСОВ

М. Ю. Рябикин, А. М. Сергеев

Введение

Первые экспериментальные данные, свидетельствующие о получении субфемтосекундных импульсов в процессе генерации высоких гармоник оптического излучения в газах [1—3], дали старт бурному развитию физики аттосекундных световых импульсов [4]. Преодоление аттосекундного барьера открывает возможности для осуществления прямых экспериментальных наблюдений [5] процессов, протекающих на временных масштабах, сравнимых с характерным временем движения электронов в атомах: $t_a = \hbar^3 / (m_e e^4) \approx 24$ ас ($1 \text{ ас} = 10^{-18}$ с). Длительности импульсов, полученных в недавних экспериментах [2, 6], составляют около 250 ас, а максимальный коэффициент преобразования оптического излучения в вакуумный ультрафиолетовый (ВУФ) диапазон достигает величины порядка 10^{-5} [7]. Актуальными задачами являются исследование путей получения еще более коротких импульсов и поиск условий для наиболее эффективной трансформации энергии фемтосекундных лазерных импульсов в ВУФ и мягкое рентгеновское излучение аттосекундного диапазона длительностей.

В основе эффекта генерации высоких гармоник интенсивного линейно поляризованного оптического излучения лежит синхронизованное лазерным полем тормозное излучение электронов, вырываемых из атомов, ускоряемых оптическим полем и соударяющихся вновь с родительскими ионами [8]. Такие соударения являются причиной возникновения высокочастотного плато в спектре нелинейного поляризационного отклика атома. Излучение в области края плато обусловлено электронами, имеющими при возвратных соударениях максимально возможную кинетическую энергию. Для случая синусоидального электрического поля $E(t) = E_0 \sin \omega_0 t$ эта закономерность выражается в формуле для частоты обрезания плато: $\hbar \omega_{\max} \approx 3,17 U_p + I_p$, где $U_p = e^2 E_0^2 / 4m\omega_0^2$ — средняя осцилляторная энергия электрона в поле $E(t)$, I_p — потенциал ионизации атома. Благодаря повторяемости с полупериодом оптического поля актов ионизации, происходящих с различными атомами в каждом физически малом объеме, излучение атомного ансамбля представляет собой набор высоких нечетных гармоник частоты лазерного поля. По-

сколько вклад в высокочастотное излучение дают лишь электроны, высвобождаемые из атомов в течение коротких промежутков времени вблизи моментов $\omega t_n \approx 0, 6\pi + \pi l$, генерация высокоэнергичных фотонов происходит в виде периодической последовательности аттосекундных всплесков [1—2, 9].

Важной задачей для исследований различных быстропротекающих процессов является получение одиночных ультракоротких импульсов. В настоящее время одним из основных способов получения одиночных аттосекундных импульсов является генерация высоких гармоник излучения оптического импульса, содержащего малое число периодов поля [10, 11]. Сильная временная локализация процесса генерации высоких гармоник при этом обусловлена резкой зависимостью скорости ионизации от величины электрического поля лазерного импульса. В силу этого же обстоятельства сверхкороткие всплески излучения гармоник могут быть получены в условиях быстрой ионизации атомов на переднем фронте мощного фемтосекундного лазерного импульса [12].

Фундаментальные ограничения на эффективность описанного выше процесса генерации аттосекундных импульсов связаны с распылением электронного волнового пакета при свободном движении. Поскольку скорость распыления пакета зависит от масштабов его начальной локализации, увеличение эффективности генерации аттосекундных импульсов может быть достигнуто, наряду с выбором подходящего профиля оптического поля, также и за счет оптимизации исходного электронного состояния [13—16]. С этой точки зрения большой интерес представляет генерация высоких гармоник в молекулярных газах, так как молекулы обладают большим по сравнению с атомами числом степеней свободы, воздействие на которые позволяет манипулировать состоянием частиц. Оптимальная подготовка среды и выбор профиля лазерного импульса могут также быть эффективными средствами управления спектральным составом [15] и длительностью [16] генерируемых аттосекундных импульсов. В данной статье обсуждаются возможности использования «инжиниринга» электронных волновых пакетов для управления временными, спектральными и энергетическими характеристиками аттосекундного излучения, генерируемого при ионизации атомов и молекул мощными фемтосекундными лазерными импульсами.

1. Генерация аттосекундных всплесков при быстрой ионизации возбужденных атомов фемтосекундным лазерным импульсом

Рассмотрим задачу о возбуждении одиночного аттосекундного всплеска при быстрой ионизации атомов мощным оптическим импуль-

сом. В случае, если пиковая интенсивность лазерного импульса превышает критическое значение, соответствующее возникновению надбарьерной ионизации [17], отрыв электрона может происходить за промежуток времени, существенно меньший по сравнению с периодом поля. Волновой пакет освободившихся электронов движется затем вне атома как единое целое и может, будучи ускоренным в электрическом поле лазерного импульса, возвратиться к родительскому иону и столкнуться с ним, вызвав всплеск тормозного излучения аттосекундной длительности [12, 18]. В случае надбарьерной ионизации атомов, находившихся первоначально в основном состоянии, генерация гармоник малоэффективна из-за быстрого расплывания электронного волнового пакета [19]. Однако, как будет показано ниже, эффективность этого процесса значительно повышается при использовании исходно возбужденных атомных состояний. Более того, при этом появляется возможность генерации импульсов длительностью порядка или даже меньше 10 ас.

При рассмотрении ограничимся простейшим примером атома водорода, для которого в предположении полной ионизации на одном периоде лазерного поля удастся получить аналитические выражения, описывающие эволюцию электронного волнового пакета и профиль аттосекундного всплеска при тормозном излучении электронов, возвращающихся к родительским ионам. Будут также приведены результаты трехмерных численных расчетов, которые, благодаря высокой степени локализации волнового пакета при надбарьерной ионизации, удастся произвести на сетке сравнительно небольших размеров за счет интегрирования по области, перемещающейся вслед за электронным пакетом. На примере s -состояний мы рассмотрим зависимость эффективности генерации аттосекундного всплеска и его длительности от главного квантового числа n начального состояния атома.

1.1. Генерация одиночного аттосекундного всплеска при надбарьерной ионизации атома водорода

Для иллюстрации особенностей рассматриваемого процесса приведем вначале результаты трехмерных численных расчетов для атома водорода, ионизируемого из исходного возбужденного $2s$ -состояния оптическим импульсом с линейно поляризованным электрическим полем:

$$E(t) = 2 \exp\left[-5(\omega_0 t / 2\pi - 1)^4\right] \sin \omega_0 t \quad (1)$$

(здесь и далее используются атомные единицы: $e = m = \hbar = 1$; частота поля $\omega_0 = 0,114$ соответствует второй гармонике титан-сапфирового лазера).

В результате ионизации атома в поле (1) на малом отрезке периода $T = 2\pi/\omega_0$ образуется волновой пакет свободных электронов, движущийся по траектории классической частицы в суммарном поле лазерного излучения и иона и испытывающий квантово-механическую диффузию. Повторное взаимодействие электронного пакета с кулоновским центром происходит при $t \approx 0,9 T$. На рис. 1, *a* и 1, *б* приведены временная зависимость и спектр нелинейного атомного отклика

$$R(t) = \iiint |\Psi(\vec{r}, t)|^2 \frac{\partial V}{\partial z} dx dy dz \quad (2)$$

(обусловленного атомным потенциалом вклада в параллельную электрическому полю компоненту вектора $\ddot{\mathbf{d}}(t) = d^2 \langle \mathbf{r} \rangle / dt^2$ [20]).

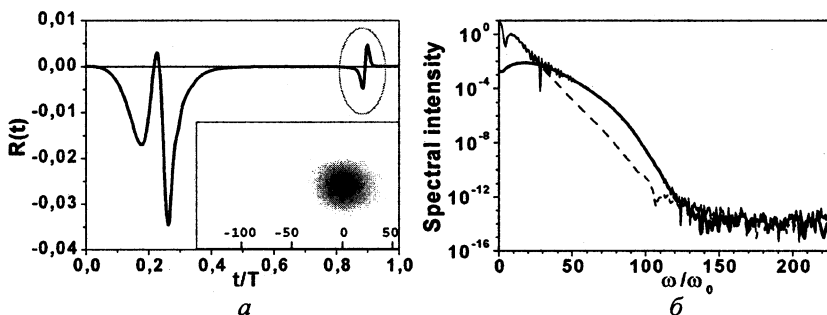


Рис. 1. Поляризационный отклик атома водорода (*a*), ионизируемого из $2s$ -состояния оптическим импульсом (1) (на врезке — «мгновенный снимок» электронного пакета при столкновении с родительским ионом) и спектр атомного поляризационного отклика (*б*); наряду с полным спектром (тонкая линия) представлены спектры фрагментов отклика, соответствующих движению волнового пакета до (пунктирная линия) и после (жирная линия) точки поворота ($t \approx 0,66 T$).

На рис. 1, *a* отчетливо видна генерация аттосекундного всплеска (выделенная область) при пролете ускоренного волнового пакета мимо родительского иона («мгновенный снимок» электронного пакета при столкновении с родительским ионом показан на врезке). На рис. 1, *б*, кроме полного спектра, представлены также отдельно вычисленные спектры фрагментов атомного отклика, соответствующих движению волнового пакета до и после точки поворота.

Приведенные зависимости показывают, что в спектре могут быть выделены три различные частотные области, обусловленные действием разных механизмов. Низкочастотная часть спектра возбуждается в основном при вырывании электронов из атома. Центральная часть спектра

обусловлена ускорением основной массы электронов при возвратном соударении. Поскольку кинетическая энергия этих электронов, пролетающих в основном на достаточно больших прицельных расстояниях от иона, превышает потенциальную, их движение носит безотражательный характер. Наконец, наиболее жесткое излучение обусловлено очень небольшой группой электронов, пролетающих на малых прицельных расстояниях по отношению к кулоновскому центру и поэтому наиболее сильно взаимодействующих с ним. Именно эта группа электронов формирует на профиле волнового пакета (см. врезку к рис. 1, *a*) мелкомасштабные сгустки, обусловленные интерференцией налетающей и отраженной компонент волновой функции. В обычно рассматриваемом случае генерации высоких гармоник в поле достаточно длинного лазерного импульса умеренной интенсивности такого рода интерференционная модуляция электронного волнового пакета становится ярко выраженной вследствие многократного взаимодействия пакета с кулоновским центром [21], что приводит к формированию высокочастотного плато в спектре излучения атома. Именно с областью плато обычно связывается возбуждение аттосекундного излучения. Однако в приведенном выше примере подавляющая часть энергии аттосекундного всплеска сосредоточена не в высокочастотной, а в промежуточной области спектра, определяющий вклад в которую вносит основная масса далеко пролетающих электронов. Это позволяет найти аналитическое решение для формы аттосекундного всплеска, пренебрегая вкладом наиболее сильно рассеивающихся электронов в его генерацию.

1.2. Аналитические расчеты

Как уже отмечалось, в рассматриваемом нами случае отрыв электрона происходит за малую долю периода оптического поля. В приведенном выше примере частота поля $\omega_0 = 0,114$ сравнима с собственной частотой $2s$ -состояния, равной $0,125$. Это фактически означает, что в поле (1) кулоновский потенциал открывается за время, существенно меньшее периода изменения Ψ -функции невозмущенного атомного состояния. При этом естественно полагать, что электронный пакет оказывается вне атома практически неизменным, и при аналитическом рассмотрении его эволюции под действием оптического поля в качестве начальных условий можно выбрать волновую функцию невозмущенного состояния электрона в атоме. Аналитическое рассмотрение дальнейшей эволюции волновой функции мы проводим в приближении сильного поля (см., например, [21]), пренебрегая кулоновским потенциалом иона, что оправдано для рассматриваемого нами случая высокоинтенсивного лазерного поля, оказывающего определяющее влияние на динамику

электрона в континууме. Задачу будем рассматривать в системе отсчета, связанной с центром масс пакета, движущегося, в соответствии с теоремой Эренфеста, по классической траектории (подобный подход в упрощенном варианте применялся в работе [19], где рассмотрение ограничивалось случаем надбарьерной ионизации из основного состояния, волновая функция которого для простоты вычислений аппроксимировалась гауссовой функцией). В этом случае в используемом нами приближении уравнение для радиальной функции, соответствующей произвольному s -состоянию атома водорода, имеет вид

$$i \frac{\partial X}{\partial t} = -\frac{1}{2} \frac{\partial^2 X}{\partial r^2} - \frac{1}{r} \frac{\partial X}{\partial r} \quad (3)$$

(r — расстояние от центра пакета), а начальными условиями являются невозмущенные волновые функции

$$X_n(r, 0) = \frac{2}{n\sqrt{n}} \exp\left(-\frac{r}{n}\right) F\left(-n+1, 2, \frac{2r}{n}\right), \quad (4)$$

где $F(a, b, c)$ — вырожденная гипергеометрическая функция.

После замены $X = \chi/r$ получим диффузионное уравнение с мнимым коэффициентом диффузии, решение которого с граничным условием $\chi(0, t) = 0$ может быть записано в интегральном виде:

$$\chi(r, t) = \frac{1}{\sqrt{2\pi it}} \int_0^\infty \chi(x, 0) \left[\exp\left(\frac{i}{2t}(x-r)^2\right) - \exp\left(\frac{i}{2t}(x+r)^2\right) \right] dx. \quad (5)$$

Рассмотрим поведение электронной плотности в центре волнового пакета в зависимости от времени, что является удобной характеристикой степени распыливания пакета. Из (4) и (5) можно получить [22]:

$$\begin{aligned} \Psi_n(0, t) = A^{(n)} \sum_{m=0}^{n-1} C_m^{(n)} \left[B_{1m}^{(n)} F\left(-n, m-n+\frac{3}{2}, z\right) + \right. \\ \left. + B_{2m}^{(n)} (-z)^{n-m-\frac{1}{2}} F\left(-m-\frac{1}{2}, n-m+\frac{1}{2}, z\right) \right], \quad (6) \end{aligned}$$

где

$$A^{(n)} = -\frac{\exp(-z)}{\sqrt{\pi^3 n^3}}, \quad B_{1m}^{(n)} = \frac{\Gamma(m+\frac{3}{2})\Gamma(n-m-\frac{1}{2})}{n!}, \quad B_{2m}^{(n)} = \Gamma(m-n+\frac{1}{2}),$$

$$C_m^{(n)} = (-1)^{n-m} \binom{2n}{2m+1}, \quad z = -\frac{it}{2n^2}.$$

Асимптотический анализ формулы (6) при малых и больших t дает:

$$|\Psi_n(0, t \rightarrow 0)|^2 \cong (1 - 4t^{1/2}/\pi n^2)/\pi n^3; \quad |\Psi_n(0, t \rightarrow \infty)|^2 \cong 8n^5/\pi^2 t^3. \quad (7)$$

Отсюда следует, что, хотя при малых t электронная плотность в центре пакета монотонно убывает с ростом главного квантового числа n , при больших t эта тенденция меняется на противоположную.

Зависимость $|\Psi_n(0, t)|^2$ в широком интервале времен для $n = 1-4$ представлена на рис. 2. Видно, что уже на временах, значительно меньших периода оптического поля, электронная плотность в центре волнового пакета для исходно невозбужденного состояния становится меньше, чем для возбужденных. Из полученных зависимостей следует, что для каждого заданного значения частоты лазерного импульса существует оптимальное возбужденное состояние атома n^* , ионизация из которого обеспечивает максимальное число электронов, участвующих в генерации аттосекундного всплеска. Как следует из рис. 2, при ионизации атома водорода излучением на основной частоте Ti:Sa-лазера данному условию удовлетворяет $n^* = 3$, а для второй гармоники — $n^* = 2$.

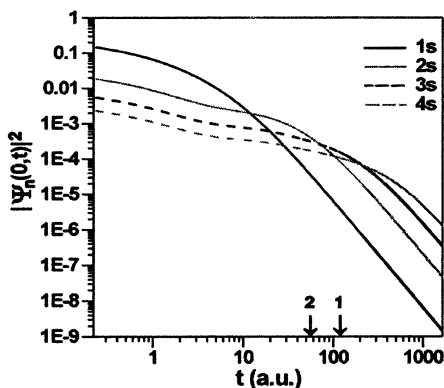


Рис. 2. Зависимость электронной плотности в центре волнового пакета от времени для начальных s -состояний с $n = 1-4$. Стрелками указаны значения t , соответствующие периоду поля (1) основной частоты ($\lambda = 800$ нм) и (2) второй гармоники ($\lambda = 400$ нм) титан-сапфирового лазера.

Более полное рассмотрение задачи может быть проведено на основе асимптотического представления для волновой функции при больших t , когда в результате расплывания размеры электронного пакета становятся существенно больше исходных. В этом случае в формуле (5) при $t \gg n^2$ можно положить $\exp(ix^2/2t) \approx 1$, что, с формальной точки зре-

ния, аналогично приближению Фраунгофера в задаче о дифракции волновых пучков в свободном пространстве. В результате вычислений получаем асимптотическое выражение для волновой функции, совпадающее, с точностью до замены r/t на волновое число, с исходной волновой функцией в импульсном пространстве, откуда получим

$$|\Psi_n(r, t \rightarrow \infty)|^2 = \frac{2n^3}{\pi^2 t^3 (1 + \rho^2)^{2n+2}} \left[\sum_{k=0}^{n-1} \binom{2n}{2k+1} (-\rho^2)^k \right]^2, \quad (8)$$

где $\rho = nr/t$.

Из уравнения (8) следует, что размер пакета $a_n(t)$ (полная ширина по полувысоте) при больших t возрастает пропорционально времени: $a_1(t) = 2t\sqrt{2^{1/4} - 1} \approx 0,87t$, $a_2(t) \approx 0,297t$, $a_3(t) \approx 0,144t$, $a_4(t) \approx 0,0833t$ и т. д. Таким образом, скорость расплывания волнового пакета падает с ростом n , что согласуется с рис. 2. Длительность всплеска в нашей задаче определяется длительностью пролета центрального сгустка волнового пакета вблизи кулоновского центра: $\tau_n \approx a_n(t^*)/\dot{z}(t^*)$, где t^* — интервал времени между событиями ионизации и соударения, рассчитываемый для классической траектории. Для используемого нами профиля лазерного поля (1) $\tau_1 \approx 75$ ас, $\tau_2 \approx 25$ ас, $\tau_3 \approx 12$ ас и $\tau_4 \approx 7,2$ ас.

В рамках используемых выше приближений, подставляя асимптотическое выражение (8) для волнового пакета свободных электронов в соотношение (2), удается аналитически вычислить дипольный отклик системы в процессе столкновения пакета с кулоновским центром:

$$R_n(t) = \frac{2}{\pi z^2(t)} \left[\arctg \lambda + \frac{\lambda K_n(\lambda)}{n(4n^2 - 1)(1 + \lambda^2)^{2n+1}} \right], \quad \lambda = \frac{nz(t)}{t}. \quad (9)$$

Здесь $z(t)$ — зависящее от времени расстояние между ядром и центром волнового пакета, которое определяется траекторией классического движения электрона в лазерном поле $E(t)$; $K_n(\lambda)$ — конечные полиномы степени $4n$, которые могут быть найдены аналитически при любом n [16, 22] и выражения для которых здесь для краткости опускаем.

Вычисления по формуле (9) показывают, что с увеличением главного квантового числа начального состояния энергия тормозного излучения существенно возрастает, а длительность генерируемого всплеска уменьшается в соответствии с приведенными выше оценками. Для более строгого анализа, в особенности для высоких s -состояний, необходим выход за рамки асимптотического приближения. Эту задачу удается ре-

шить аналитически для произвольного начального состояния $|nlm\rangle$ атома водорода и для любого момента времени t [22]. В случае s -состояний из (4) и (5) получается следующее решение для $\Psi_n(r, t)$:

$$\Psi_n(r, t) = \frac{1}{2\pi r} \exp\left(i \frac{r^2}{2t}\right) \times \sum_{k=0}^{n-1} \frac{(-1)^{k+1} \sqrt{n!(n-1)!}}{n^2 (n-1-k)! (k+1)! k!} \left(\frac{2}{n}\right)^k (2it)^{\frac{k+1}{2}} \times$$

$$\times \left\{ \Gamma\left(\frac{k+2}{2}\right) \cdot \left[F\left(\frac{k+2}{2}, \frac{1}{2}, z_2^2\right) - F\left(\frac{k+2}{2}, \frac{1}{2}, z_1^2\right) \right] - \right. \quad (10)$$

$$\left. - 2\Gamma\left(\frac{k+3}{2}\right) \cdot \left[z_2 F\left(\frac{k+3}{2}, \frac{3}{2}, z_2^2\right) - z_1 F\left(\frac{k+3}{2}, \frac{3}{2}, z_1^2\right) \right] \right\},$$

где $z_{1,2} = (\mp r + it/n) / \sqrt{2it}$, а поляризационный отклик на отрезке времени, соответствующем возбуждению аттосекундного импульса, записывается в виде следующего простого интегрального выражения:

$$R_n(t) = \frac{1}{z(t)|z(t)|} \int_0^{|z(t)|} |\Psi_n(r, t)|^2 4\pi r^2 dr, \quad (11)$$

где $z(t)$ — мгновенное расстояние между центром пакета и ядром.

Интеграл (11) с функцией $\Psi_n(r, t)$, задаваемой выражением (10), вычислить аналитически не удастся, поэтому интегрирование производилось численно. Результаты этих вычислений для четырех низших s -состояний представлены на рис. 3. Для сравнения приведены также отклики, полученные из трехмерных численных расчетов.

Сравнение результатов аналитических и численных расчетов показывает их хорошее согласие, особенно для $2s$ -состояния. Различия для $1s$ -состояния, обладающего наибольшим потенциалом ионизации, объясняются искажением волновой функции основного состояния в процессе ионизации атома. Расхождения между аналитическими и численными результатами при $n \geq 3$ связаны с возрастающим (с увеличением n) влиянием кулоновской фокусировки на эволюцию электронного волнового пакета после ионизации атома.

Важные выводы, которые следуют из всех проведенных расчетов, состоят в том, что генерация аттосекундного всплеска при использовании возбужденных состояний гораздо эффективнее, чем при ионизации из основного состояния, а длительность центральной наиболее высокочастотной части аттосекундного отклика существенно уменьшается с ростом главного квантового числа начального состояния атома.

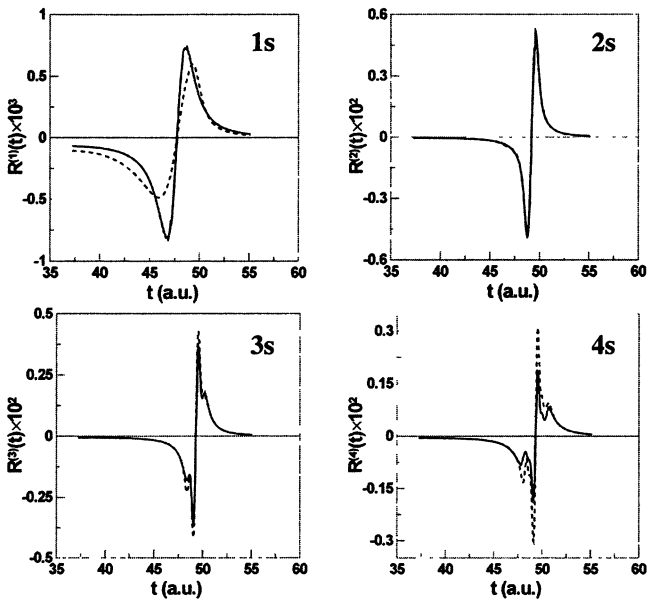


Рис. 3. Профили аттосекундного всплеска, полученные из аналитических (сплошная линия) и численных (пунктирная линия) расчетов для начальных s -состояний с $n = 1-4$. При аналитических расчетах траектория волнового пакета в поле (1) вычислялась классически с моментом ионизации, взятым из численных квантово-механических расчетов.

Обсудим кратко физические ограничения на эффективность генерации аттосекундных импульсов и их длительность при использовании исходно возбужденных атомных состояний. Одно из основных ограничений связано с тем, что с увеличением главного квантового числа n атомных водородных состояний вероятность локализации электронов вблизи ядра уменьшается и, соответственно, возрастает доля электронов, локализованных в периферийных структурах. В связи с этим в аттосекундном отклике системы доля высокочастотной составляющей, обусловленной центральной группой электронов, при больших n снижается, а возрастание роли периферийных электронов приводит к общему увеличению длительности всплеска и снижению эффективности генерации тормозного излучения при больших n .

Другим фактором, ограничивающим возможности высокоэффективной генерации аттосекундных импульсов при высоких интенсивностях оптического излучения, является влияние магнитного поля лазерной волны [18], искривляющего траекторию электронного волнового пакета. Степень влияния этого фактора зависит от поперечных размеров электронного пакета при его прохождении вблизи родительского иона и в

рассматриваемой задаче возрастает с ростом главного квантового числа начального атомного состояния из-за уменьшения пространственных размеров электронного сгустка в центральной части пакета.

2. Управление процессом генерации аттосекундных импульсов в молекулярных газах

Так как молекулы являются более сложными системами, чем атомы, использование молекулярных газов предоставляет дополнительные возможности для оптимизации и управления процессом преобразования частот по сравнению с атомарными газами. Такими дополнительными степенями свободы, обеспечивающими возможность контроля, являются размеры и конфигурация молекулы, ориентация молекулярной оси по отношению к вектору электрического поля лазерной волны и др. Ниже рассматриваются возможности управления процессом генерации аттосекундных импульсов за счет предварительного выстраивания и колебательного или электронного возбуждения молекул.

2.1. Факторы, влияющие на эффективность генерации гармоник в молекулярных газах

Выстраивание молекул. В сильном световом поле может происходить перестройка вращательного движения молекул, приводящая к выстраиванию молекулярных осей вдоль направления поляризации поля. К настоящему времени разработано несколько способов выстраивания молекул [23]. Результаты экспериментальных и теоретических исследований генерации гармоник в ансамбле выстроенных молекул [24—26] демонстрируют широкие возможности использования ориентации молекул для управления процессом нелинейного преобразования частот. Ориентационная зависимость нелинейного отклика молекулы, ионизируемой в сильном лазерном поле, может быть обусловлена зависимостью скорости ионизации молекулы от угла между ее осью и электрическим полем лазерного импульса, анизотропией распределения по импульсам в исходном электронном состоянии молекулы, взаимным расположением ядер, на которых рассеивается электрон, и др.

Размеры и конфигурация молекул. Геометрия молекул оказывается другим важным фактором, который может влиять на эффективность генерации гармоник и характеристики генерируемого излучения [13—15, 27]. Влияние этого фактора определяется зависимостью скорости ионизации молекулы от межъядерного расстояния, зависимостью скорости распыливания электронного волнового пакета от масштабов локализации исходной волновой функции, возможностью генерации излучения

при столкновениях электронов с соседними ядрами в молекуле и др. Большая протяженность молекул обуславливает наличие состояний, характеризующихся большой долей электронов с малым значением импульса. В результате появляется возможность увеличения числа частиц, эффективно участвующих в генерации тормозного излучения в процессе возвратных соударений ускоренных лазерным полем электронов с молекулярным остовом, благодаря чему эффективность генерации аттосекундных импульсов и высоких гармоник излучения при ионизации молекулярных структур фемтосекундным оптическим импульсом может значительно превосходить эффективность аналогичных процессов в атомах [13]. Сильная зависимость эффективности генерации гармоник в молекулярных газах от межъядерного расстояния может быть использована для управления этим процессом путем воздействия на молекулы последовательностью двух импульсов — накачки (для электронно-колебательного возбуждения молекулы) и зондирующего с варьируемой временной задержкой (для генерации гармоник) [14].

Тип молекулярной орбитали. Как показывают недавние теоретические и экспериментальные исследования, структура и симметрия электронной волновой функции оказывают существенное влияние на процессы взаимодействия сильного лазерного поля с молекулами, в том числе и на процесс генерации высоких гармоник [28, 29]. Этих факторы, в частности, влияют на скорость ионизации молекул и на динамику электронного волнового пакета после ионизации.

2.2. Генерация аттосекундных импульсов при ионизации молекул в возбужденных электронных состояниях

Как было показано выше, эффективность генерации аттосекундного импульса при ионизации атомов мощным оптическим импульсом с малым числом периодов поля может существенно увеличиться при использовании исходно возбужденных электронных состояний, поскольку ионизация из них приводит к формированию волновых пакетов, в меньшей степени подверженных расплыванию при свободном движении. Ниже мы показываем, что аналогичный эффект должен иметь место при ионизации электронно-возбужденных молекул.

При численных расчетах нами использовалась широко принятая модель (см., например, [25—27]) простейшей молекулярной системы — двумерный аналог иона H_2^+ со сглаженным кулоновским потенциалом $V(x, y)$, образованным двумя одинаковыми V однозарядными центрами, удерживающими единственный электрон:

$$V(x, y) = -\left[x^2 + (y - D/2)^2 + a^2\right]^{-1/2} - \left[x^2 + (y + D/2)^2 + a^2\right]^{-1/2}. \quad (12)$$

Здесь D — расстояние между ядрами, $a = (0,5)^{1/2}$ — параметр сглаживания кулоновской сингулярности, выбираемый из условия совпадения энергии ионизации модельного иона из нижнего электронного терма с ее величиной для реального молекулярного иона. Молекулы считаем выстроенными вдоль оси y , перпендикулярной линейно поляризованному электрическому полю $E(t)$. Ядра считаем неподвижными, что для интересующего нас здесь режима ионизации молекулы оптическим импульсом с малым числом периодов поля вполне оправдано даже для молекул с легкими ядрами.

На рис. 4 представлены результаты расчетов для случая ионизации молекулярного иона H_2^+ на переднем фронте лазерного импульса

$$E(t) = E_0 [\exp(2\omega_0 t / \pi) - 1] \sin \omega_0 t, \quad (13)$$

где $E_0 = 0,36$, $\omega_0 = 0,114$. Приведены исходные волновые функции основного и двух различных возбужденных состояний, «мгновенные снимки» электронного волнового пакета перед столкновением с молекулярным остовом и спектры нелинейного молекулярного отклика, рассчитанные для данных начальных электронных состояний иона H_2^+ при оптимальных межъядерных расстояниях ($D = 7,2, 15$ и 9 соответственно). Из приведенных зависимостей следует, что спектральная интенсивность излучения в области частот $\omega > 30 \omega_0$ при ионизации из возбужденных состояний иона H_2^+ увеличивается на несколько порядков по сравнению со случаем ионизации из основного электронного состояния.

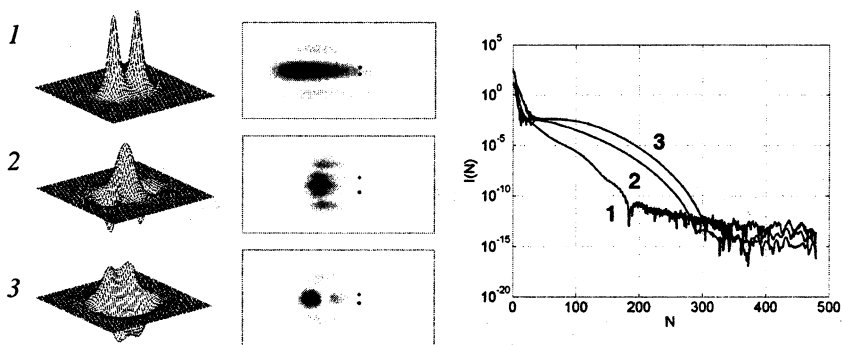


Рис. 4. Слева — волновые функции основного (1) и двух различных возбужденных состояний (2, 3); в центре — «мгновенные снимки» электронного волнового пакета перед столкновением с молекулярным остовом для начальных состояний 1—3; справа — спектры нелинейного молекулярного отклика, рассчитанные для начальных состояний 1—3.

2.3. Генерация перестраиваемого аттосекундного излучения с использованием квантовой интерференции

В молекуле из-за наличия нескольких ядер имеется несколько источников волн де Бройля, испускаемых при ионизации. При расстоянии между ядрами, превышающем несколько боровских радиусов, эти волны при дальнейшем распространении могут интерферировать, что в известном смысле аналогично интерференции волновых пучков при дифракции света на двух щелях. В случае ионизации молекулы в течение одного периода быстро нарастающего лазерного поля образующийся волновой пакет имеет регулярную структуру с характерными пространственными масштабами, зависящими от геометрии задачи. Квазипериодическая интерференционная модуляция электронного волнового пакета приводит к наличию в спектре тормозного излучения, генерируемого при его столкновении с молекулярным остовом, ярко выраженного максимума, который при увеличении межъядерного расстояния или при повороте оси молекулы от поперечной по отношению к полю ориентации к продольной смещается в высокочастотную область. Это смещение объясняется уменьшением продольного масштаба интерференционной модуляции электронного пакета с увеличением межъядерного расстояния и косинуса угла между осью молекулы и вектором электрического поля лазерного импульса.

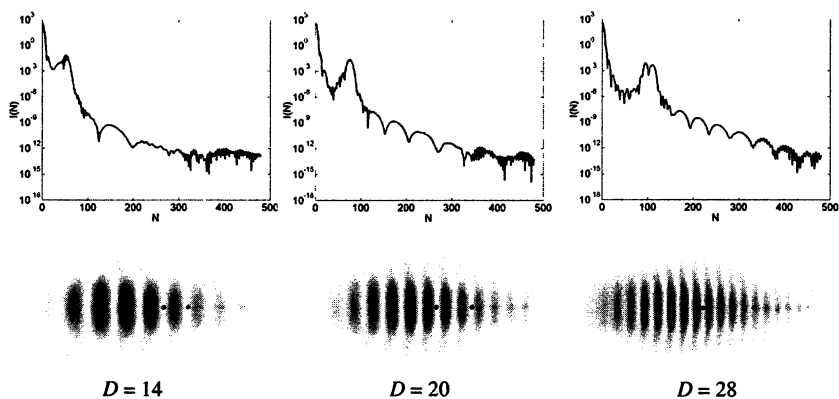


Рис. 5. Спектр поляризационного отклика (вверху) и электронный волновой пакет при соударении с молекулярным остовом (внизу) в случае ионизации молекулярного иона H_2^+ лазерным импульсом (14) при параллельной лазерному полю ориентации молекулярной оси и различных межъядерных расстояниях D .

В качестве примера на рис. 5 представлены спектры поляризационного отклика иона H_2^+ , ионизируемого из основного электронного состояния лазерным импульсом

$$E(t) = E_0 \exp\left[-5\left(\omega_0 t/2\pi - 1\right)^4\right] \sin \omega_0 t \quad (14)$$

(здесь $E_0 = 1,5$, $\omega_0 = 0,114$) при параллельной лазерному полю ориентации молекулярной оси и различных значениях межъядерного расстояния. Рисунок 6 демонстрирует возможности широкой перестройки спектра аттосекундного излучения с использованием квантовой интерференции.

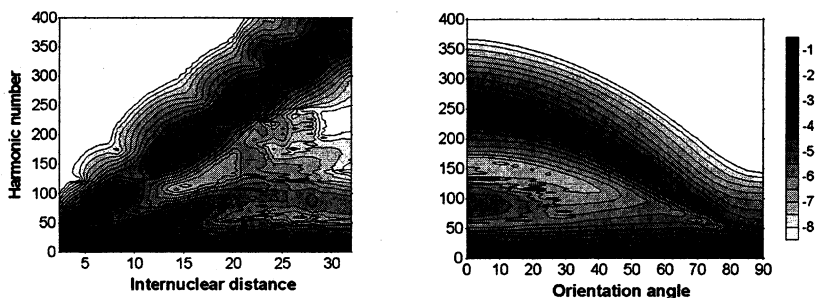


Рис. 6. Перестройка спектра аттосекундного излучения при изменении расстояния между ядрами (слева) и угла ориентации молекулярной оси иона H_2^+ по отношению к электрическому полю лазерного импульса (13) (справа).

Данная работа выполнена при частичной поддержке Президентской программы поддержки ведущих научных школ РФ (школы «Нелинейные преобразования лазерного излучения» и «Нелинейная динамика оптических систем и высокочувствительные оптические измерения»).

Литература

1. Papadogiannis N. A., Witzel B., Kalpouzos C., Charalambidis D. Observation of attosecond light localization in higher order harmonic generation // *Phys. Rev. Lett.* 1999. V. 83, № 21. P. 4289—4292.
2. Paul P. M., Toma E. S., Breger P. et al. Observation of a train of attosecond pulses from high harmonic generation // *Science*. 2001. V. 292, № 5522. P. 1689—1692.
3. Hentschel M., Kienberger R., Spielmann Ch. et al. Attosecond metrology // *Nature*. 2001. V. 414, № 6863. P. 509—513.
4. Agostini P., DiMauro L. F. The physics of attosecond light pulses // *Rep. Prog. Phys.* 2004. V. 67, № 6. P. 813—855.
5. Drescher M., Hentschel M., Kienberger R. et al. Time-resolved atomic inner-shell spectroscopy // *Nature*. 2002. V. 419, № 6909. P. 803—807.
6. Kienberger R., Goulielmakis E., Uiberacker M. et al. Atomic transient recorder // *Nature*. 2004. V. 427, № 6977. P. 817—821.
7. Hergott J.-F., Kovacev M., Merdji H. et al. Extreme-ultraviolet high-order harmonic pulses in the microjoule range // *Phys. Rev. A*. 2002. V. 66, № 2. Art. no. 021801.
8. Corkum P. B. Plasma perspective on strong-field multiphoton ionization // *Phys. Rev. Lett.* 1993. V. 71, № 13. P. 1994—1997.
9. Antoine P., L'Huillier A., Lewenstein M. Attosecond pulse trains using high-order harmonics // *Phys. Rev. Lett.* 1993. V. 77, № 7. P. 1234—1237.

10. *Christov I. P., Murnane M. M., Kapteyn H. C.* High-harmonic generation of attosecond pulses in the «single-cycle» regime // *Phys. Rev. Lett.* 1997. V. 78, № 7. P. 1251—1254.
11. *Spielmann Ch., Burnett N. H., Sartania S. et al.* Generation of coherent x-rays in the water window using 5-femtosecond laser pulses // *Science.* 1997. V. 278, № 5338. P. 661—664.
12. *Sergeev A. M., Kim A. V., Vanin E. V. et al.* Atoms in a superstrong laser field: towards subfemtosecond XUV sources // *Laser Optics '95 and ICONO '95: Superintense Laser Fields: Proc. SPIE / Eds. A. A. Andreev, V. M. Gordienko.* 1996. V. 2770. P. 36—45.
13. *Емелин М. Ю., Рябикин М. Ю., Сергеев А. М. и др.* О генерации аттосекундных всплесков и высоких гармоник излучения при ионизации молекул сверхкоротким лазерным импульсом // *Письма в ЖЭТФ.* 2003. Т. 77, вып. 5. С. 254—259.
14. *Pfeifer T., Walter D., Gerber G. et al.* Transient enhancement of high-harmonic generation in expanding molecules // *Phys. Rev. A.* 2004. V. 70, № 1. Art. no. 013805.
15. *Емелин М. Ю., Рябикин М. Ю., Сергеев А. М.* О возможностях управления процессом генерации аттосекундных рентгеновских импульсов при ионизации молекул фемтосекундным лазерным излучением // *Изв. вузов. Радиофизика.* 2004. Т. 47, вып. 10—11.
16. *Emelin M. Yu., Ryabikin M. Yu., Sergeev A. M. et al.* High-efficiency generation of attosecond pulses during atomic ionization from excited electronic states // *Europhys. Lett.* (in print).
17. *Augst S., Meyerhofer D. D., Strickland D., Chin S. L.* Laser ionization of noble gases by Coulomb-barrier suppression // *J. Opt. Soc. Am. B.* 1991. V. 8, № 4. P. 858—867.
18. *Ким А. В., Рябикин М. Ю., Сергеев А. М.* От фемтосекундных к аттосекундным импульсам // *УФН.* 1999. Т. 169, вып. 1. С. 58—66.
19. *Fedorov M. V., Peatross J.* Strong-field photoionization and emission of light in the wave-packet-spreading regime // *Phys. Rev. A.* 1995. V. 52, № 1. P. 504—513.
20. *Burnett K., Reed V. C., Cooper J., Knight P. L.* Calculation of the background emitted during high-harmonic generation // *Phys. Rev. A.* 1992. V. 45, № 5. P. 3347—3349.
21. *Protopapas M., Keitel C. H., Knight P. L.* Atomic physics with super-high intensity lasers // *Rep. Prog. Phys.* 1997. V. 60, № 4. P. 389—486.
22. *Емелин М. Ю., Рябикин М. Ю., Сергеев А. М.* Генерация одиночного аттосекундного всплеска при ионизации возбужденных атомов мощным сверхкоротким лазерным импульсом (направлено в печать).
23. *Stapelfeldt H., Seideman T.* Aligning molecules with strong laser pulses // *Rev. Mod. Phys.* 2003. V. 75, № 2. P. 543—557.
24. *Velotta R., Hay N., Mason M. B. et al.* High-order harmonic generation in aligned molecules // *Phys. Rev. Lett.* 2001. V. 87, № 18. Art. no. 183901.
25. *Lappas D. G., Marangos J. P.* Orientation dependence of high-order harmonic generation in hydrogen molecular ions // *J. Phys. B: At. Mol. Opt. Phys.* 2000. V. 33, № 21. P. 4679—4689.
26. *Lein M., Hay N., Velotta R. et al.* Role of the intramolecular phase in high-harmonic generation // *Phys. Rev. Lett.* 2002. V. 88, № 18. Art. no. 183903.
27. *Moreno P., Plaja L., Roso L.* Ultrahigh harmonic generation from diatomic molecular ions in highly excited vibrational states // *Phys. Rev. A.* 1995. V. 55, № 3. P. R1593—R1596.
28. *Shan B., Tong X.-M., Zhao Z. et al.* High-order harmonic cutoff extension of the O₂ molecule due to ionization suppression // *Phys. Rev. A.* 2002. V. 66, № 6. Art. no. 061401(R).
29. *Lein M.* Antibonding molecular orbitals under the influence of elliptically polarized intense light // *J. Phys. B: At. Mol. Opt. Phys.* 2003. V. 36, № 10. P. L155—L162.

РАМАНОВСКОЕ УСИЛЕНИЕ КОРОТКИХ ЛАЗЕРНЫХ ИМПУЛЬСОВ В ПЛАЗМЕ

А. Н. Степанов, Г. М. Фрайман

1. Введение

За последние двадцать лет достигнуты огромные успехи в генерации коротких (субпикосекундных) лазерных импульсов высокой интенсивности и исследовании их взаимодействия с веществом. Говоря о большой мощности получаемых импульсов, назовем цифру $P = 1$ ПВт (10^{15} Вт) — это, практически, максимальная мощность, достигнутая к настоящему времени в Ливерморской национальной лаборатории (США). При фокусировке такого излучения можно получить фантастические интенсивности электромагнитного излучения вплоть до $I \sim 10^{21} — 10^{22}$ Вт/см². Чтобы понять, много это или мало, расставим несколько характерных точек на оси интенсивностей.

При интенсивности $I \approx 3 \cdot 10^{16}$ Вт/см² поле в электромагнитной волне становится примерно равным электрическому полю, действующему на электрон, находящийся на первой бортовой орбите в атоме водорода. При таких полях ионизация атома происходит очень быстро — за время порядка периода изменения лазерного поля. Если интенсивность превосходит величину порядка $I \sim 5 \cdot 10^{18}$ Вт/см² (при длине волны излучения лазера около 1 мкм), то осцилляторная энергия электрона в таком поле становится больше энергии покоя электрона $E = mc^2$, и мы переходим в область релятивистских интенсивностей.

Таким образом, экспериментально полученные в настоящее время интенсивности лазерного излучения являются сильно релятивистскими, а величины электрических полей в фокусе мощных лазерных систем превосходят поля в атомах на 2—3 порядка.

Зачем нужны такие высокие интенсивности? В первую очередь — для исследования фундаментальных проблем взаимодействия полей, не достижимых никаким другим способом, с веществом. Можно привести несколько конкретных научных проблем, для решения которых нужны высокие интенсивности электромагнитного излучения. Более тридцати лет огромные усилия прилагались для реализации идеи управляемого термоядерного синтеза, осуществляемого с помощью лазеров. К настоящему времени получен большой объем научного материала, который, в частности, показывает, что для реализации лазерного термоядерного синтеза необходима энергия порядка мегаджоуля в импульсе лазерного излучения с характерной длительностью несколько наносекунд. Для снижения требований к энергии лазерного импульса была предложена идея быстрого поджига термоядерной мишени. Термоядерная мишень, сжатая наносекундным импульсом, дополнительно облучается сильно сфокусированным коротким субпикосекундным лазерным им-

пульсом высокой интенсивности. Этот короткий импульс, интенсивность которого согласно расчетам должна быть более 10^{20} Вт/см², локально нагревает мишень выше температуры, необходимой для начала термоядерной реакции, и далее горение термоядерного топлива распространяется по всей мишени.

При облучении высокоинтенсивными фемтосекундными импульсами поверхности твердых тел часть лазерной энергии поглощается, причем поглощение происходит очень локализованно: теплопроводностные процессы не успевают произойти. Поэтому часть вещества мишени оказывается сильно нагретой, в ней развиваются огромные давления. Таким образом, использование коротких лазерных импульсов высокой интенсивности позволяет в такого рода экспериментах исследовать вещество, находящееся в условиях экстремально высоких температур и давлений. В частности, можно определять уравнение состояния вещества в таких условиях.

Еще одной физической проблемой, решение которой возможно при применении коротких интенсивных лазерных импульсов, является проблема ускорения заряженных частиц (электронов) лазерным полем. Несмотря на то что в фокусе современных лазерных систем достигаются огромные электрические поля, их нельзя напрямую использовать для ускорения частиц. Однако если возбудить с помощью фемтосекундного лазерного импульса в ионизованном газе плазменную волну большой амплитуды, то тогда электрон, захваченный полем такой волны, может двигаться с ускорением в 10^2 — 10^3 раз большим, чем в лучших нынешних линейных ускорителях.

Основной схемой, используемой в настоящее время для получения фемтосекундных лазерных импульсов высокой интенсивности, является схема усиления частотно-модулированных импульсов [1]. Задающий генератор является источником последовательности фемтосекундных импульсов с энергией $W \approx 1$ нДж, следующих с высокой частотой повторения. Если усиливать такие короткие импульсы в усилителе, то при плотностях потока энергии, превышающих 1 ГВт/см², в оптических элементах схемы будут наблюдаться нелинейные искажения импульса или происходить пробой элементов оптики. Чтобы избежать этого, импульсы с задающего генератора растягиваются во времени в специальном устройстве — стретчере — в 10^3 — 10^4 раз, чтобы снизить интенсивность усиливаемого излучения. Такие растянутые частотно-модулированные импульсы затем усиливаются до необходимого уровня по энергии в нескольких каскадах усиления. В конце усиленные импульсы сжимаются во времени практически до длительности импульсов, определяемых задающим генератором. Сжатие происходит в устройстве, называемом компрессором. Временное растяжение и сжатие импульсов происходит при помощи дифракционных решеток, которые и являются самым слабым звеном, лимитирующим стойкость элементов лазерной системы. Для достижения большей энергии в лазерном импульсе

требуется увеличение размеров дифракционных решеток. Решетки больших размеров трудно технологически изготавливать и они чрезвычайно дороги. К настоящему времени максимальные размеры изготавливаемых решеток составляют примерно 300 – 400 мм, их стойкость не превышает 200—300 мДж/см². Использование описанной схемы позволило достичь приведенную выше максимальную мощность $P = 10^{15}$ Вт.

Что же можно предложить для получения экзаваттных уровней мощности (10^{18} Вт)? Разумеется, можно пытаться идти по экстенсивному пути, стараясь увеличить размеры кристаллов в каскадах усиления и размеры решеток в стрэтчере и компрессоре. Однако, как уже отмечалось, этот путь, фактически, уже исчерпал себя. Дело не только в прочностном пределе усилителя и других элементов схемы собственно лазера, но и в том, что излучение столь большой мощности нельзя фокусировать твердотельными оптическими элементами, поскольку и на них мы также будем иметь дело с пробоем. Но если уж нам приходится всюду опасаться образования плазмы, то нельзя ли собственно плазму и использовать как среду, в которой будет осуществляться усиление. Действительно, что нам было нужно от твердотельного усилителя? Широкая полоса усиления. Но в плазме есть достаточное число неустойчивостей, которые также позволяют осуществлять усиление в широкой полосе. При этом, чем больше интенсивность излучения, тем естественнее работать с полностью ионизованной плазмой. Поэтому схемы с использованием плазмы в качестве активной среды принципиально нацелены на высокие интенсивности. Одна из схем применения плазмы для получения сверхкоротких лазерных импульсов сверхвысокой интенсивности и обсуждается в этой лекции. Речь идет о возможности усиления слабой электромагнитной волны в плазме при ее взаимодействии с волной накачки на сдвинутой частоте, когда плазменные колебания играют роль третьей волны [2]. Это, так называемые, рамановские усилители. Первые эффекты резонансного трехволнового взаимодействия были открыты Раманом (Индия) в газах и Мандельштамом (СССР) с Бриллюэном (Франция) в жидкости. За открытие этого эффекта Раману в 1930 году была присуждена Нобелевская премия по физике.

В своей работе [2] авторы предложили реализовать в полностью ионизированной плазме процесс рассеяния длинного лазерного импульса накачки в короткий усиливаемый лазерный импульс, бегущий навстречу. Разность несущих частот обоих импульсов предполагалась равной электронной плазменной частоте. В области перекрытия волны накачки и усиливаемого импульса на электроны плазмы действует усредненная пондеромоторная сила («сила Миллера») [3], частота которой равна разности частот накачки и усиливаемого импульса, а волновое число — сумме волновых чисел взаимодействующих волн. Поскольку эта сила имеет частоту, совпадающую с плазменной частотой, она эффективно возбуждает плазменную волну, в которой плотность плазмы оказывается промодулированной в пространстве. В результате этого

и волна накачки и усиливаемый импульс в области перекрытия «видят» волну неоднородности плотности с волновым вектором, направленным в сторону распространения волны накачки и равным по величине сумме волновых чисел накачки и усиливаемого импульса. Поскольку, к тому же, частота этой волны равна разности частот взаимодействующих лазерных волн, то поляризационный ток, возникающий при поляризации этой волны плотности в поле волны накачки, будут резонансными с волной усиливаемого импульса, а поляризационный ток в поле усиливаемого импульса в резонансе с волной накачки.

Реализация этой идеи в плазме может быть осуществлена разными способами. Авторы идеи [2] предложили использовать слой однородной плазмы толщиной 1 см с концентрацией порядка 10^{18} см^{-3} . Посылая в эту плазму пикосекундный импульс накачки с интенсивностью порядка 10^{14} Вт/см^2 и встречный затравочный импульс фемтосекундной длительности, они показали, что на выходе можно получить короткий импульс с интенсивностью до 10^{17} Вт/см^2 . Более того, как было показано в [4], если на входе в слой плазмы затравочный импульс имеет параболический фазовый фронт, то при усилении в слоях плазмы с толщиной меньшей дифракционной длины кривизна фазового фронта усиленного импульса остается неизменной, т. е. в дальнейшем усиленный импульс будет эффективно фокусироваться, и тогда в фокальной области могут быть достигнуты интенсивности до 10^{25} Вт/см^2 . Экспериментальная проработка этой идеи успешно проводится в США [5] в настоящее время.

Другая идея [6] реализована в ИПФ РАН [7]. Суть ее состоит в следующем. Осуществим трехволновое взаимодействие в тонком диэлектрическом капилляре, заполненном газом. В этом случае длина взаимодействия будет определяться длиной капилляра и, разумеется, будет много больше дифракционной длины, определяющей длину взаимодействия в свободном пространстве. Усиление и компрессия импульса, как показано в [6], могут быть реализованы в этой схеме автоматически, что позволило бы отказаться и от усиливающего кристалла, и от решетки компрессора в схеме стандартного тераваттного лазера.

Именно вопросам, связанным с возможностью экспериментальной реализации этой схемы, и посвящена данная работа. В разд. 2 мы приведем и кратко обсудим основную модель, используемую для численного и аналитического анализа проблемы. Далее (разд. 3) обсуждается простейшая одномерная модель процесса. В разд. 4 приведены результаты экспериментов по реализации процесса обратного рамановского рассеяния в капилляре. В заключении (разд. 5) кратко обсуждаются основные проблемы на пути создания реального рамановского усилителя.

2. Основные уравнения

В результате мы имеем дело с пространственно-временным синхронизмом трех волн (волны накачки, усиливаемой волны и плазменной волны). В общем виде эти условия в терминах частот ω , и волновых

векторов \mathbf{k}_i взаимодействующих волн могут быть записаны следующим образом:

$$\omega_1 = \omega_2 + \omega_3, \quad \mathbf{k}_1 = \mathbf{k}_2 + \mathbf{k}_3. \quad (1)$$

В приближении параксиального распространения импульсов вдоль оси z уравнения для поля волны накачки усиливаемого импульса и плазменной волны могут быть представлены в следующем виде (см., например, [8]):

$$\begin{aligned} \partial_t a + c \partial_z a - \frac{ic^2}{2\omega_a} \Delta a &= -\omega_p b f, \\ \partial_t b - c \partial_z b - \frac{ic^2}{2\omega_b} \Delta b &= \omega_p a f, \\ \partial_t f + i\delta\omega f &= \frac{\omega}{2} b^* a + WB[f]. \end{aligned} \quad (2)$$

Здесь a и b — медленные амплитуды электрических полей

$$\begin{aligned} E_a &= \frac{m_e c \omega_a}{e} \left\{ ia \exp^{ik_a z - i\omega_a t} + \text{к.с.} \right\}, \\ E_b &= \frac{m_e c \omega_b}{e} \left\{ ib \exp^{ik_b z - i\omega_b t} + \text{к.с.} \right\} \end{aligned} \quad (3)$$

накачки и усиливаемого импульса соответственно, а f есть нормированный потенциал электрического поля плазменной волны

$$E_d = k_f^{(0)} \frac{m_e c \omega_p}{e} \left\{ f \exp^{ik_f z - i\omega_f t} + \text{к.с.} \right\}, \quad (4)$$

где $k_f^{(0)} = k_f / k_f$, $k_f = z^{(0)}(k_a - k_b)$, $\omega_f = \omega_a - \omega_b = \omega_p - \delta\omega$, $\omega_p = \sqrt{4\pi n_e e^2 / m_e}$ — плазменная частота, n_e — электронная концентрация, e и m_e — заряд и масса электрона соответственно, $\delta\omega$ — частотная расстройка трехволнового синхронизма. Плазма предполагается разреженной ($\omega_p \ll \omega_a \approx \omega_b \equiv \omega$) и холодной ($k_f \lambda_D \ll 1$, где λ_D — дебаевский радиус). Мы пренебрегаем дисперсией электромагнитных волн, считая $k_{a,b} \approx \omega_{a,b} / c$, а также дисперсией плазменной волны, полагая $\partial_{k_f} \omega_f \approx 0$. Последнее слагаемое в правой части уравнения для плазменной волны ответственно за эффект опрокидывания [8]. Отметим, что сам по себе вопрос об учете принципиально кинетического эффекта опрокидывания плазменных волн большой амплитуды в рамках квазигидродинамической модели является исключительно важным для решения задач взаимодействия лазерного излучения с плазмой. То, что сделано в [8] и использовано ниже при моделировании экспериментальной модели, фактически, состоит в ограничении амплитуды плазменной волны ее предельным уровнем. Напри-

мер, в холодной плазме этот уровень определяется из условия, что радиус осцилляции электрона в поле волны не должен превышать длину волны. Следует отметить, что эффект опрокидывания становится существенным только в редкой плазме, поэтому обычно при рассмотрении теоретических моделей его не учитывают, а ограничиваются лишь оценками. Именно так в теоретической части поступим и мы.

Для дальнейшего теоретического рассмотрения, чтобы исключить «лишние» параметры, удобно представить уравнения (2) в безразмерном виде:

$$\partial_{\tau} a + \partial_z a - (1 + \sigma) i \Delta a = -bf, \quad \partial_{\tau} b - \partial_z b - i \Delta b = af, \quad \partial_{\tau} f + i \delta \omega f = ab^*, \quad (5)$$

где время τ измеряется в единицах $t_0 = \sqrt{2/\omega \omega_p}$, продольная координата z — в единицах ct_0 , поперечная координата ρ — в единицах $c(2\omega_p \omega^3)^{-1/4}$, частотная расстройка $\delta \omega$ — в единицах t_0^{-1} ; $\sigma = (\omega_a - \omega_b) / \omega_b \sim \omega_p / \omega \ll 1$.

При нулевой частотной расстройке уравнения (5) инвариантны по отношению к замене переменных

$$a \rightarrow Ca, \quad b \rightarrow Cb, \quad f \rightarrow Cf, \quad \tau \rightarrow \tau/C, \quad z \rightarrow z/C, \quad \rho \rightarrow \rho/\sqrt{C}.$$

Таким образом, явное значение амплитуды накачки $a_0 = a(z \rightarrow -\infty)$ не является существенным: решения задач об эволюции волновых полей при двух различных значениях a_0 могут быть получены друг из друга путем перемасштабирования. (В нелинейном режиме зависимость от амплитуды затравочного импульса b_0 является слабой [2, 3]).

3. Одномерная задача

Простейшие оценки эффективности процесса обратного рамановского усиления могут быть получены в рамках одномерной модели. На линейной стадии, пока истощение накачки еще несущественно и амплитуда усиливаемого импульса мала по сравнению с амплитудой накачки, решение уравнений (5) может быть получено преобразованием Лапласа уравнений по координате ζ :

$$b(\zeta, z) = \frac{\partial}{\partial \zeta} \int G(\zeta - \zeta', z) b(\zeta', 0) d\zeta', \quad (6)$$

где мы пренебрегли расстройкой. Для $\eta \gg 1$ мы имеем в исходных переменных:

$$G = \exp\left(2\sqrt{\eta}\right) / 2\sqrt{\pi\sqrt{\eta}} \\ \eta = a_0^2 \omega_p (t+z)(-z) / 2c. \quad (7)$$

Следовательно, максимум G , т. е. максимум усиления на линейной стадии, реализуется при $z = ct/2$ и соответствует инкременту неустойчи-

ности $\gamma = a_0 \sqrt{\omega \omega_p} / 2$. В этой точке импульс растет по закону $\exp(\gamma t)$ с инкрементом, в точности совпадающим по величине с тем, что может быть получен в простейшей пространственно однородной задаче. Существенно, что на линейной стадии усиливаемый импульс расширяется в пространстве, т. е. процесс аналогичен тому, что мы имели с применением решетки стрэтчера.

Дальнейшее усиление с переходом на нелинейную стадию, когда становится существенным процесс истощения накачки, можно описать с помощью уравнения синус-Гордона:

$$\partial_{\zeta}^2 U = a_0^2 \sin U. \quad (8)$$

Для его получения достаточно сделать замены:

$$a = a_0 \cos(U/2), \quad f = \sqrt{2} a_0 \sin(U/2), \quad b = \partial_{\zeta} U / 2. \quad (9)$$

Уравнение (8) обладает важным классом автомодельных решений:

$$U(\tau, \xi) = U(\eta), \quad \eta U_{\eta\eta} + U_{\eta} = \sin U \quad (10)$$

или

$$U_{\xi\xi} + U_{\xi}/\xi = 0, \quad \xi = 2\sqrt{\eta}. \quad (11)$$

Соответствующее этой автомодельности решение приведено на рис. 1.

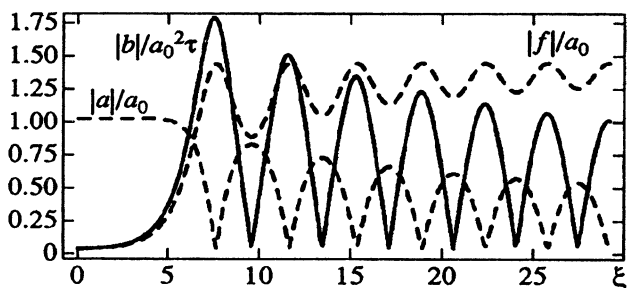


Рис. 1. Автомодельные решения для амплитуды накачки ($|a|/a_0$, пунктир); плазменной волны ($|f|/a_0$, штрихпунктир); усиливаемого импульса ($|b|/a_0^2 \tau$, сплошная линия) для начальных условий $b(z, 0) = 0, 01\delta(z)$.

Существенно, что это решение слабо зависит от начальных условий. Амплитуда первого импульса оказывается пропорциональной длине усиления, а его длительность — обратно пропорциональной его амплитуде. Таким образом, на нелинейной стадии короткий импульс с огибающей $b(z, t)$ не только усиливается, но и сжимается по длительности.

4. Экспериментальные результаты

Эксперимент по исследованию рамановского усиления фемтосекундных лазерных импульсов в плазме, создаваемой в диэлектрических

капиллярах, выполнялся на тераваттном фемтосекундном лазерном комплексе, функционирующем в ИПФ РАН. Схема экспериментальной установки приведена на рис. 2. Частотно-модулированный лазерный импульс с выхода последнего каскада усиления лазерной системы разделялся на два импульса с помощью поляризационного делителя D . Один из импульсов, содержащий примерно 75 % энергии и имеющий вертикальную поляризацию, использовался в качестве импульса накачки. После сжатия в оптическом компрессоре второй импульс, имеющий горизонтальную поляризацию, разделялся еще на две части с помощью зеркала с коэффициентом отражения 98 %. Отраженный от зеркала импульс использовался для создания плазмы, а прошедший за зеркало после поворота поляризации с помощью пластины $\lambda/2$ использовался в качестве затравочного импульса, взаимодействующего с накачкой. Энергия в импульсе накачки и затравочном импульсе менялась с помощью поляризационных ослабителей, образованных пластинами $\lambda/2$ и поляризаторами $P1$ и $P2$. Величины входных и выходных энергий всех трех импульсов контролировались калиброванными фотодиодами.

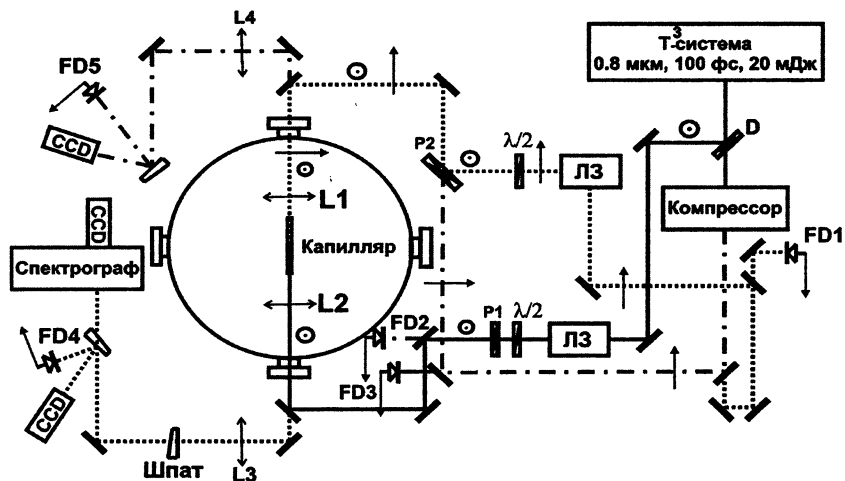


Рис. 2. Схема эксперимента по усилению фемтосекундных лазерных импульсов на обратном рамановском рассеянии в плазме. Стрелкой и кружком с точкой указана поляризация излучения, горизонтальная и вертикальная соответственно. Сплошная линия — тракт накачки, пунктирная — тракт усиливаемого импульса, штрихпунктирная — ионизирующего импульса. ЛЗ — линии задержки.

Оптический путь накачки и затравочного импульса регулировался с помощью линий задержки. Плазма в капилляре создавалась за 3 нс до прихода импульсов накачки и затравки путем опережающего прихода ионизирующего импульса. Излучение на вход капилляра фокусировалось линзами $L1$ и $L2$ с фокусным расстоянием 21 см, обеспечивающим оптимальное согласование лазерных пучков с фундаментальной модой

капилляра. Для диагностики спектра усиленного импульса и пространственной структуры выходных импульсов использовались спектрограф и CCD-камеры.

В эксперименте использовался стеклянный капилляр диаметром 30 мкм и длиной 1,5 см. Капилляр помещался в вакуумную камеру. Вакуумная камера после откачки наполнялась аргоном, давление которого варьировалось в диапазоне от 0,3 до 30 Торр. Плазма создавалась ионизирующим лазерным импульсом интенсивностью $\sim 5 \cdot 10^{15}$ Вт/см². Частотно-модулированный импульс накачки имел длительность 170 пс при спектральной ширине 14 нм. Затравочный импульс имел длительность 80 фс, спектр затравочного импульса идентичен спектру накачки и ионизирующего импульса. Энергия затравочного импульса менялась в диапазоне 0,1—1 мкДж, что соответствует максимальной интенсивности $\sim 3 \cdot 10^{13}$ Вт/см².

На рис. 3 представлена зависимость коэффициента усиления затравочного импульса от его входной энергии при максимальной энергии накачки. Под коэффициентом усиления мы понимаем отношение энергии усиливаемого импульса на выходе капилляра к энергии затравочного импульса. Как следует из рис. 3, эта зависимость имеет спадающий характер. Измерения показывают, что преобразование энергии импульса накачки в процессе усиления мало, т. е. мы находимся на линейной стадии усиления. Тогда коэффициент усиления не должен зависеть от энергии затравочного импульса (пунктирная линия на рис. 3). Чтобы объяснить падающую зависимость, мы предлагаем ввести в рассмотрение эффект опрокидывания плазменной волны. Учет опрокидывания плазменной волны в рамках «мягкой» модели дает хорошее согласие с экспериментом, «жесткая» модель дает заниженные результаты.



Рис. 3. Зависимость коэффициента усиления от энергии затравочного импульса. Точками нанесены экспериментальные данные, полученные при давлении нейтрального аргона 1 Торр, линиями представлены результаты численных расчетов для разных моделей опрокидывания.

Зависимость коэффициента усиления затравочного импульса от энергии в импульсе накачки для оптимальной задержки между усиленным импульсом и импульсом накачки приведена на рис. 4. Видно, что при максимальной энергии накачки максимальная величина коэффициента усиления составила ~ 130 . Это рекордная цифра из опубликованных в настоящее время экспериментальных работ по усилению мощных фемтосекундных лазерных импульсов при обратном рамановском рассеянии в плазме.

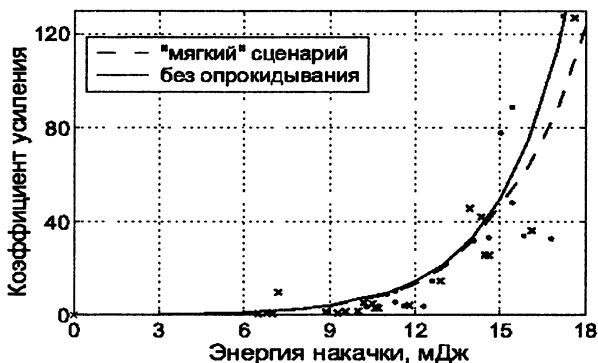
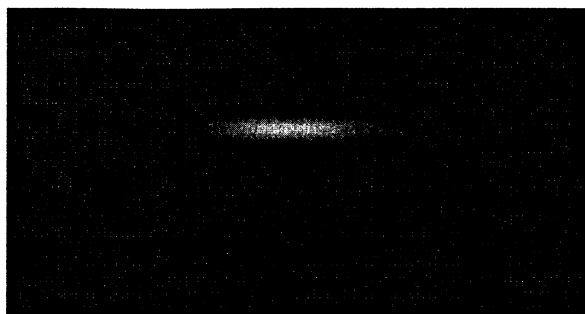


Рис. 4. Зависимость коэффициента усиления от энергии накачки. Точки и крестики — экспериментальные данные при давлении нейтрального аргона 0,3 Торр для двух значений задержки затравки относительно накачки, близких к оптимальной. Линии — численный расчет для разных моделей опрокидывания.

Результаты спектральных измерений для давления аргона $p = 1$ Торр представлены на рис. 5. Как следует из рис. 5, в спектре усиленного сигнала наблюдается появление узкого интенсивного пика (спайка), ширина которого существенно меньше ширины спектра затравочного импульса. При этом величина усиления по локальной спектральной интенсивности превышает 10^3 , что также является рекордной к настоящему времени величиной. В целом спектральная картина очень близка к картине, наблюдаемой при взаимодействии волн с узким спектром в обычной схеме обратного рамановского рассеяния с использованием разночастотных накачки и затравки. При уменьшении давления газа до 0,3 Торр структура спектра изменялась. В спектре усиленного импульса наблюдается появление нескольких узких локальных максимумов, спадающих по амплитуде. Подобная структура спектра, по нашим представлениям, может быть связана с модуляцией концентрации плазмы вдоль капилляра и соответствующим ей изменением инкремента усиления вдоль трассы для разных спектральных компонент.

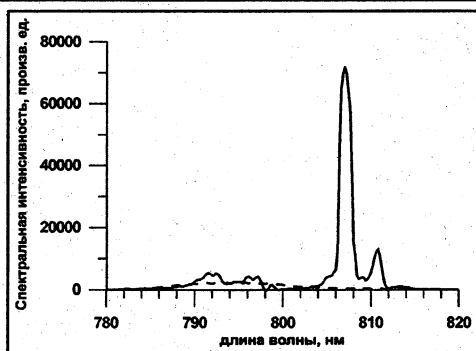
Пространственная структура усиленного сигнала, изображенная на рис. 6, свидетельствует о хорошей поперечной однородности плазмы в капилляре и квазиодномодовом режиме распространения излучения.



a



б



в

Рис. 5. Спектр усиленного сигнала при максимальной накачке и давлении нейтрального аргона 1 Торр: *a* и *б* — распределение на CCD-камере спектра затравочного импульса и усиленного сигнала, *в* — усредненный по поперечному сечению спектр, полученный оцифровкой изображений с CCD. Пунктир — спектр затравочного импульса.

Таким образом, резюмируя экспериментальные результаты, полученные при исследовании обратного рамановского усиления фемтосекундных лазерных импульсов с использованием плазмы, создаваемой в диэлектрических капиллярах, можно констатировать следующее. Эксперименты продемонстрировали, что применение длинных диэлектрических капилляров за счет увеличения длины взаимодействия позволило достичь рекордных на сегодняшний день значений усиления коротких

лазерных импульсов по энергии (~ 100). Капилляр, по которому распространяются импульсы, одновременно служит хорошим пространственным фильтром, приводящим к гладкому квазиодномодовому пространственному распределению усиливаемого импульса.

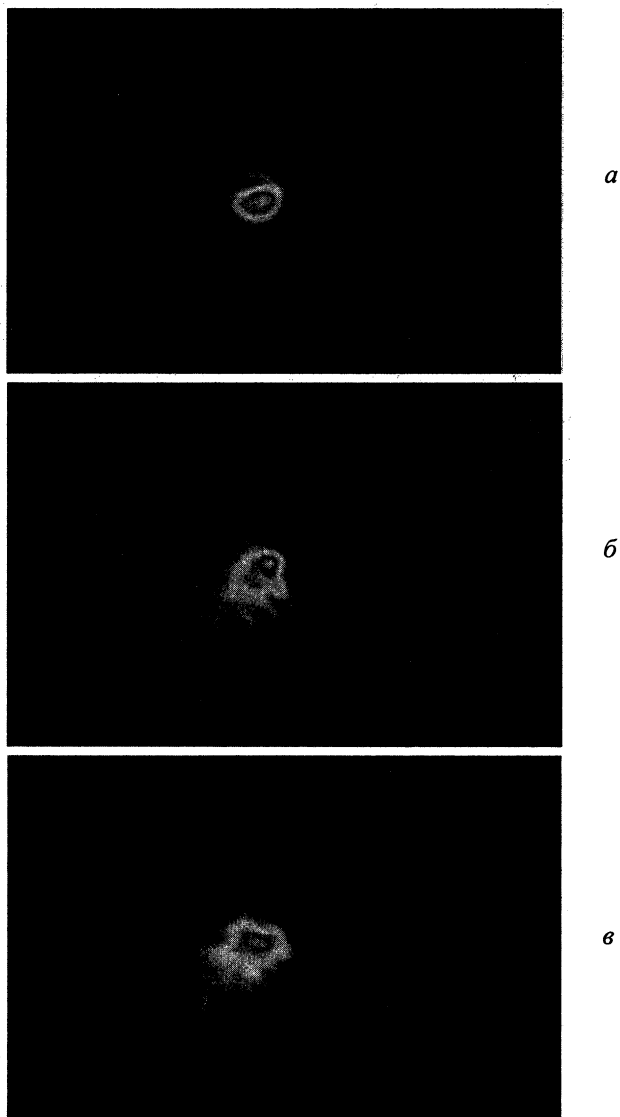


Рис. 6. Пространственное распределение интенсивности усиленного сигнала, полученное с CCD-камеры: *a* — в капилляре вакуум, *б* — давление нейтрального аргона 1 Торр, *в* — давление нейтрального аргона 0,3 Торр.

5. Заключение

В заключение обсудим основные проблемы, мешающие эффективной перекачке энергии из накачки в усиливаемый импульс в нашем эксперименте и переходе к наиболее привлекательной (как обсуждалось ранее) нелинейной стадии усиления. Прежде всего, поскольку затравочный импульс имеет достаточно узкую ширину спектра, а его несущая частота совпадает с частотой импульса накачки, чтобы выполнить условия синхронизма, мы вынуждены работать с очень редкой плазмой. В такой плазме легко достигается опрокидывание плазменных волн, и следовательно, предельная амплитуда плазменной волны оказывается весьма малой и наступает на малых длинах, т. е. весь капилляр не может быть использован достаточно эффективно.

Другая проблема обусловлена способом получения плазмы. В нашем случае она создавалась предварительным коротким пробойным импульсом. По мере распространения этого импульса по капилляру он излучался через стенки, так что поле в импульсе по длине капилляра было неоднородным. Поскольку интенсивность входного импульса была ограничена возможностями эксперимента, мы получали неоднородную плазму и по длине капилляра. Это приводило к дополнительному нарушению условий синхронизма и, опять-таки, к ограничению эффективности взаимодействия.

Серьезным фактором, усложняющим рамановское усиление лазерных импульсов в плазме в выполненных экспериментах, является сильная частотная модуляция импульса накачки. С одной стороны, она резко снижает паразитное спонтанное рассеяние импульса накачки на флуктуациях плазменной концентрации. С другой стороны, непрерывное изменение частоты ведет к нарушению условий синхронизма для спектральных компонент усиливаемого импульса и серьезному ограничению усиления. Решением проблемы было бы использование в качестве импульса накачки излучения от отдельного лазера со сдвинутым спектром и малым регулируемым частотным чирпом. Однако технические трудности, лежащие на пути получения такого излучения и прецизионной синхронизации отдельных лазерных систем, очевидны.

Отдельную проблему представляет твердотельная плазма, возникающая при ионизации материала, из которого изготовлен капилляр, при фокусировке как ионизирующего импульса, так и импульса накачки. Эта плазма приводит к уменьшению эффективности прохождения усиливаемого импульса, что будет важно, когда в экспериментах по рамановскому усилению в плазме речь будет идти о нелинейной стадии усиления и эффективности преобразования энергии импульса накачки в энергию усиливаемого импульса. Оптимизация геометрии капилляра (сужающиеся вход и выход) и увеличение энергии импульсов, используемых в эксперименте (это позволит работать с капиллярами большего диаметра), — вот пути минимизации влияния пристеночной твердотельной плазмы.

Таким образом, сама по себе идея использования рамановских усилителей в плазме для получения фемтосекундных лазерных импульсов сверхбольшой интенсивности является весьма привлекательной, и в настоящих экспериментах продемонстрирована возможность ее достижения. Но ее практическая реализация, как обычно, сопряжена с существенными трудностями, преодоление которых является типичным на пути от идеи к ее воплощению.

Работа выполнена при поддержке РФФИ (гранты 02-02-17275, 04-02-16684, 02-02-16065).

Литература

1. *Strickland D., Mourou G.* Compression of amplified chirped optical pulses // *Opt. Commun.* 1985. V. 56. P. 219—226; *Mourou G. A., J. Barty C. P., Perry M. D.* Ultra-high intensity lasers: physics of the extreme on the tabletop // *Phys. Today.* 1998. V. 51, № 1. P. 22—28.
2. *Malkin V. M., Shvets G., Fisch N. J.* Fast compression of laser beams to highly overcritical powers // *Phys. Rev. Lett.* 1999. V. 82. P. 4448—4451; *Phys. Plasmas.* 2000. V. 7. P. 2232.
3. *Gaponov A. V., Miller M. A.* Potential wells for charged particles in a high-frequency electromagnetic field // *Sov. Phys. JEPT.* 1958. V. 7. P. 168—174.
4. *Fraiman G. M., Yampol'sky N. A., Malkin V. M., Fisch N. J.* Robustness of laserphase fronts in backward Raman amplifiers // *Phys. Plasmas.* 2002. V. 9. P. 3617—3623.
5. *Ping Y., Geltner I., Suckewer S.* Raman backscattering and amplification in a gas jet plasma // *Phys. Rev. E.* 2003. V. 67, № 1. Part 2. Art. no. 016401—016406.
6. *Dodin I. Y., Fraiman G. M., Malkin V. M., Fisch N. J.* Amplification of short laser pulses by Raman backscattering in capillary plasma // *ЖЭТФ.* 2002. Т. 122. С. 723—730.
7. *Балакин А. А., Карташев Д. В., Киселев А. М., Скобелев С. А., Степанов А. Н., Фрайман Г. М.* Усиление лазерных импульсов при обратном рамановском рассеянии в плазме, создаваемой в диэлектрических капиллярах // *Письма в ЖЭТФ.* 2004. Т. 80. С. 15—20.
8. *Balakin A. A., Fraiman G. M., Malkin V. M., Fisch N. J.* Noise suppression and enhanced focusability in plasma Raman amplifier with multi-frequency pump // *Phys. Plasmas.* 2003. V. 10. P. 4856—4864.

З-ПИНЧ И ТЕРМОЯДЕРНЫЙ СИНТЕЗ

Л. С. Мхитарьян

После знаменитого доклада И. В. Курчатова в 1956 г. в Харуэлльском атомном центре (Англия) [1] о советских исследованиях по магнитному удержанию плазмы токамаки начали своё шествие по крупнейшим лабораториям мира. Довольно быстро выяснилось, что необходимое для развития термоядерной реакции время удержания плазмы не удастся достичь из-за возникающих неустойчивостей. Казалось, что при увеличении размеров плазмы (а значит, и масштабов установки) время удержания будет расти быстрее, чем время, за которое развиваются неустойчивости, выбрасывающие плазму на стенки. И хотя Л. А. Арцимович любил повторять, что природа на пути к управляемому термоядерному синтезу может воздвигнуть лишь конечное число препятствий, сегодня, по прошествии почти полувека, токамак, в котором бы началась самоподдерживающаяся термоядерная реакция, всё ещё ждёт своего часа.

Сразу после изобретения лазеров возникла идея об альтернативном подходе к управляемому синтезу: разгонять мощным световым потоком оболочку, окружающую термоядерное топливо, сжимать его, а время удержания обеспечивать за счёт инерции оболочки. Однако когерентность лазерного излучения оказалась фактором, препятствующим достижению сферически симметричного облучения оболочки: возникала мелкомасштабная неоднородность освещённости из-за интерференции световых пучков. Поэтому многие лаборатории стали искать выход, преобразуя лазерную энергию в рентгеновское излучение и уже это последнее используя для разгона оболочки. Наибольших успехов на этом пути добилась Ливерморская национальная лаборатория в США на установке НОВА — лазере, генерирующем 30 кДж световой энергии при длине волны 0,35 мкм за 3 нс.

Для зажигания термоядерной мишени, однако, требуется лазер с энергией около 2 МДж за ~5 нс, способный производить 1,5 — 2 МДж рентгеновского излучения. США и Франция начали строительство таких установок. Тем временем, эксперименты в Сандийской лаборатории США, в ходе которых на установке РВФА-Z был получен импульс рентгеновского излучения длительностью ~5 нс с энергией ~1,8 МДж [2, 3], показали, что электрофизические установки, способные создавать короткий (100—200 нс) импульс тока с амплитудой в десятков и более мегаампер, предоставляют альтернативную возможность для генерации мощного рентгеновского излучения.

Из-за большей, по сравнению с лазерами, эффективности использования электрической энергии при одинаковой энергоёмкости батареи электрофизические установки производят большее количество рентге-

новского излучения. Так, рентгеновский импульс, полученный на установке РВФА-Z, уже в полтора раза превосходит по энергии импульс, который планируется получить на строящейся в США лазерной установке NIF. Дело, однако, в том, что плотность энергии в лазерных установках более чем на порядок величины выше, чем та, что может быть получена в электрофизических установках. Поэтому для осуществления инерционного термоядерного синтеза требуется установка гораздо более мощная, чем РВФА-Z. Мне хотелось бы здесь показать, как простые и хорошо известные законы позволяют представить масштаб такой установки.

1. Поле рентгеновского излучения во вторичном боксе

Для генерации рентгеновского излучения в электрофизических установках используют цилиндрическую оболочку-лайнера (обычно он представляет собой сотню — другую тонких вольфрамовых проволочек, натянутых вдоль образующей цилиндра), через который пропускается мощный импульс тока. Возникающее при этом магнитное поле разгоняет лайнер до высокой скорости. При схлопывании на оси кинетическая энергия лайнера переходит в тепло. Разогретое вещество лайнера остывает, испуская рентгеновское излучение.

На рис. 1 приведено схематическое изображение конструкции, позволяющей использовать это излучение для сжатия мишени, содержащей термоядерное топливо (смесь трития с дейтерием).

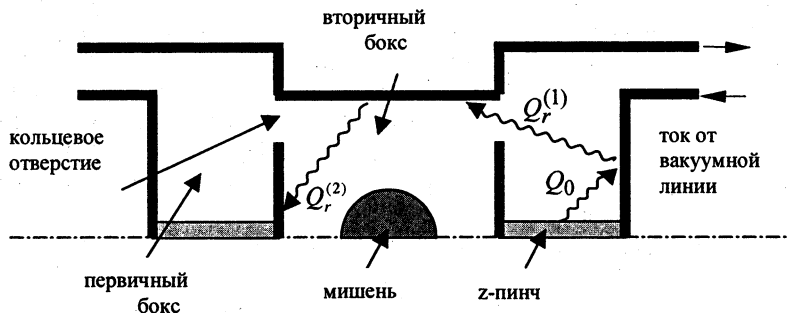


Рис. 1. Схема конструкции, передающей излучение Z-пинча к термоядерной мишени

Две первичные камеры (два бокса), в которых размещены лайнеры, соединены кольцевым отверстием со вторичным боксом. Излучение, генерируемое схлопнувшимся на оси лайнером, через это отверстие поступает во вторичный бокс и нагревает его стенки. Горячие стенки облучают мишень, обеспечивая её сжатие. Рисунок приведён на момент, когда лайнер, находившийся первоначально у верхней кромки кольцевого отверстия, схлопнулся на ось, сформировав так называемый Z-пинч.

Рассмотрим баланс рентгеновской энергии в этой системе.

Отвлекаясь от неоднородности потоков, падающих на стенки, будем оперировать с усреднёнными величинами. Полагая коэффициент отражения (альbedo) рентгеновского потока стенками α_i заданной величиной, запишем

$$Q_{1\text{ пад}} = Q_0 + Q_{r1} \frac{S_1}{S_1 + S_0} + Q_{r2} \frac{S_2}{S_2 + S_0} \cdot \frac{S_0}{S_1}, \quad (1)$$

$$Q_{2\text{ пад}} = Q_0 \frac{S_0}{S_2} + Q_{r1} \frac{S_1}{S_1 + S_0} \cdot \frac{S_0}{S_2} + Q_{r2} \frac{S_2}{S_2 + S_0}, \quad (2)$$

$$Q_{ri} = \alpha_i Q_{i\text{ пад}}. \quad (3)$$

Здесь Q_0 — поток рентгеновского излучения, создаваемый Z-пинчем на стенках первичного бокса, $Q_{r1,r2}$ — поток, излучаемый стенками первичного и вторичного боксов, $S_{1,2}$ — площадь стенок первичного и вторичного боксов, S_0 — площадь кольцевого отверстия между первичным и вторичным боксами.

Первый член в выражении (1) описывает рентгеновскую освещённость стенок первичного бокса Z-пинчем, второй учитывает поток, падающий на выбранный участок стенки первичного бокса от всех других участков стенки, третий описывает излучение, прошедшее через отверстие S_0 из вторичного бокса и освещающее стенки первичного.

Аналогичный смысл имеют члены в уравнении (2). Падающий и отражённый потоки связаны через альbedo выражением (3). Заменяя в (1) и (2) излучаемые потоки Q_{ri} на $Q_{i\text{ пад}}$ и подставляя в уравнение (2) $Q_{1\text{ пад}}$ из (1), мы найдём отношение $Q_{2\text{ пад}} / Q_0$ в зависимости от альbedo стенок, от относительной площади вторичного бокса S_2 / S_1 и от относительной площади отверстия, соединяющего боксы S_0 / S_2 .

2. Эффективность передачи энергии во вторичный бокс

Зная отношение потоков, легко определить долю рентгеновской энергии, поглощённой стенками вторичного бокса E_2 , от полной энергии рентгеновского излучения E_0 , сгенерированной Z-пинчем:

$$\frac{E_2}{E_0} = \frac{Q_{2\text{ пад}} (1 - \alpha_2) S_2}{Q_0 (S_1 + S_0)}.$$

На рисунке 2 представлена зависимость E_2 / E_0 от относительной площади отверстия S_0 / S_2 для фиксированных значений альbedo стенок и для нескольких значений S_2 / S_1 .

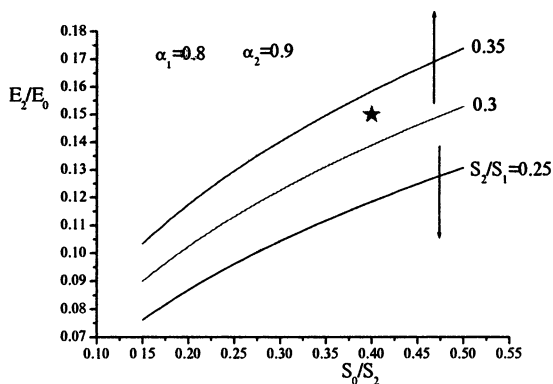


Рис. 2. Эффективность передачи энергии от Z-пинча во вторичный бокс.

Существеннейшее значение для установки имеет не сама по себе эффективность передачи энергии от Z-пинча во вторичную камеру E_2/E_0 , но абсолютное значение энергии рентгеновского излучения E_0 , сгенерированное Z-пинчем и обеспечивающее заданное значение потока во вторичном боксе Q_{r2} . Если надо достичь эффективной температуры излучения во вторичном боксе T_2 , то энергия, поглощённая стенками, будет

$$E_2 = \frac{1 - \alpha_2}{\alpha_2} \frac{\sigma c T_2^4}{4} \cdot S_1 \frac{S_2}{S_1} \left(1 - \frac{S_0}{S_2} \right) \tau, \quad (4)$$

где τ — длительность рентгеновского импульса. При заданных α_2 , T_2 , S_1 и τ энергия, сгенерированная Z-пинчем будет

$$E_0 = \frac{E_2}{E_2/E_0} = \text{const} \frac{S_2}{S_1} \left(1 - \frac{S_0}{S_2} \right) \frac{1}{E_2/E_0}.$$

Если на рис. 2 при заданном значении S_2/S_1 мы будем переходить от меньших значений S_2/S_1 к большим, то при $(S_2/S_1) \geq 0,35$ соотношение E_2/E_0 возрастает медленнее, чем S_2/S_1 , и E_0 увеличивается. При движении от больших значений S_2/S_1 к меньшим при $(S_2/S_1) \leq 0,3$ отношение E_2/E_0 падает быстрее, чем S_2/S_1 , и E_0 снова возрастает. Так что имеется область, внутри которой полная энергия рентгеновского излучения Z-пинча E_0 минимальна. На рис. 2 эта область обозначена звёздочкой

$$\left(\frac{S_0}{S_2} \approx 0,4; \frac{E_2}{E_0} \approx 0,15; \frac{S_2}{S_1} \approx 0,33 \right).$$

Из (4) видно, что выгодно уменьшать $S_1 = 2\pi RL \left(1 + \frac{R}{L} \right)$ — площадь стенок первичного бокса, где R — начальный радиус, а L — длина

лайнера. Нельзя, однако, уменьшать беспредельно. Использование лайнера с радиусом $\sim 1,5$ см представляется нелегко достижимой задачей, так как уменьшение радиуса ведёт к росту электрического напряжения, при котором должна работать установка. Длину лайнера меньше 1 см также нелегко реализовать, поскольку возникающая на боковых стенках плазма не должна шунтировать ток, протекающий по лайнеру. Представляется поэтому, что $S_1 = 2\pi \cdot 1,5 \cdot 1 \cdot 2,5 = 24 \text{ см}^2$ — оптимистическая оценка для площади первичного бокса. Не будем здесь специально анализировать условия зажигания мишени, а примем, что для вспышки достаточно, чтобы в течение ~ 3 нс мишень облучалась рентгеновским потоком с эффективной температурой $T_2 \sim 0,35$ кэВ [4]. Подставляя эти значения в (4) и полагая (см. рис. 2)

$\frac{S_0}{S_2} \approx 0,4$; $\frac{E_2}{E_0} \approx 0,15$; $\frac{S_2}{S_1} \approx 0,33$, найдём:

$$E_2 \approx 2,4 \text{ МДж при } \alpha_2 = 0,9. \text{ Следовательно } E_0 = \frac{E_2}{E_2/E_0} = \frac{2,4}{0,15} \approx 16 \text{ МДж.}$$

Нам теперь необходимо понять, какой энергией должен обладать лайнер, чтобы при схлопывании он смог высветить 16 МДж рентгеновской энергии. Рассмотрим поэтому процесс разгона лайнера.

3. Разгон лайнера, его кинетическая энергия

Пусть масса лайнера на единицу длины есть μ , его начальный радиус R_0 , и по нему протекает ток $I(t)$ (рис. 3).

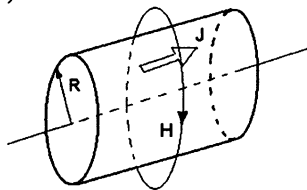


Рис. 3. Расчётная схема разгона лайнера

Движение лайнера следует закону: $\mu \frac{d^2 R}{dt^2} = -\frac{I^2}{c^2 R}$, а его кинетическая энергия

$$\left(\varepsilon = \frac{\mu}{2} \left(\frac{dR}{dt} \right)^2 \right) \text{ эволюционирует согласно } \frac{d\varepsilon}{dt} = -\frac{I^2}{c^2} \frac{d}{dt} \ln R.$$

Время нарастания тока выгодно выбрать так, чтобы лайнер схлопывался в момент времени около максимального значения тока. Но в этот момент $I = I_0 = \text{const}$, значит $\varepsilon = \frac{I^2}{c^2} \ln \frac{R}{R_0}$.

$$\varepsilon = \frac{I^2}{c^2} \ln \frac{R}{R_0}.$$

В действительности лайнер никогда не схлопывается в состояние с $R = 0$, так как он имеет конечную толщину, по которой распределён про-

текающий ток. Эта толщина, однако, не обусловлена нагревом лайнера за счёт джоулевых потерь, а определяется развитием рэлей-тейлоровских неустойчивостей при разгоне лайнера и составляет $\Delta \approx (0,1 + 0,2) R_0$. Учитывая, что \ln — медленно меняющаяся функция, для оценки кинетической энергии лайнера будем пользоваться выражением

$$\varepsilon \approx 2,5 \frac{I_0^2}{c^2} \approx 2,5 \cdot 10^{-3} I_{[MA]}^2 \left[\frac{\text{МДж}}{\text{см}} \right]. \quad (5)$$

На рис. 4 приведена зависимость кинетической энергии лайнера на единицу длины от максимального тока, протекающего через него.

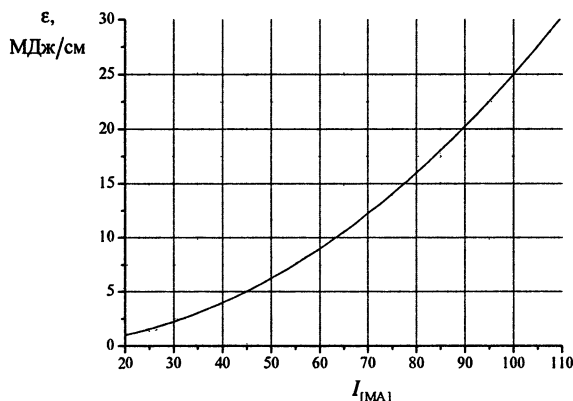


Рис. 4. Кинетическая энергия лайнера в зависимости от максимального значения тока.

4. Время нарастания тока

Вводя безразмерные переменные $r = \frac{R}{R_0}$, $\tau = \frac{t}{T}$, $T = \frac{cR_0}{I_0} \sqrt{\mu}$, $\frac{I}{I_0} = f(t)$,

из уравнения движения лайнера получим необходимое время нарастания тока

$$\frac{d^2 r}{d\tau^2} = -\frac{f^2(t)}{r}. \quad (6)$$

Для $f(t) = \sin\left(\frac{\pi}{2} \cdot \frac{T}{T_0}\right) = \sin\left(\frac{\pi}{2} \cdot \alpha \tau\right)$, $\alpha = \frac{T}{T_0}$, можно убедиться в том, что

решение (6), обеспечивающее схлопывание лайнера в момент, когда ток достигает максимального значения, существует при $\alpha \approx 0,45$.

На рисунках 5, 6 приведено решение (6) при $\alpha \approx 0,45$ и временная

эволюция кинетической энергии лайнера $\left(\varepsilon = \frac{1}{2} \left(\frac{dr}{d\tau} \right)^2 \right)$.

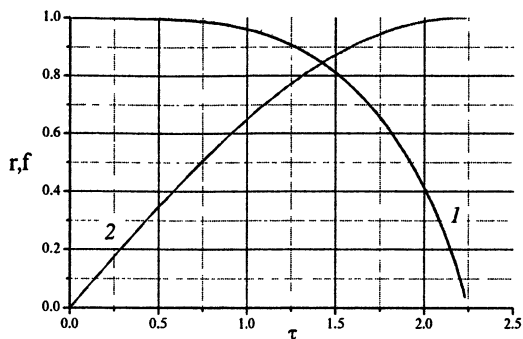


Рис. 5. Траектория лайнера (кривая 1) и ток через него (кривая 2)

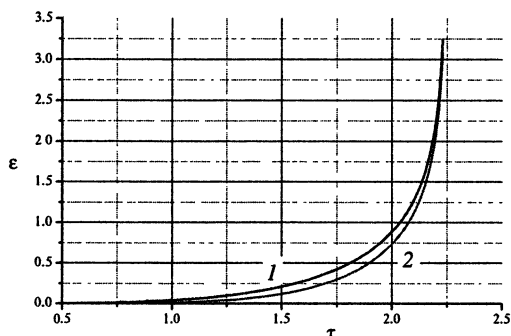


Рис. 6. Кинетическая энергия лайнера: кривая 1 — численное решение (6), 2 — оценка по (5)

Видно, что оценка (5) обладает вполне удовлетворительной точностью. Таким образом, время нарастания тока должно составлять

$$T_{0[c]} = \frac{1}{0,45} \frac{cR_0}{I_0} \sqrt{\mu} \approx 7 \cdot 10^{-7} \frac{R_0}{I_{[MA]}} \sqrt{\mu \left[\frac{MG}{CM} \right]}. \quad (7)$$

Зная кинетическую энергию лайнера, мы должны теперь оценить, какая часть этой энергии может перейти в рентгеновское излучение.

5. Генерация рентгеновского излучения Z-пинчем

Поток лучистой теплопроводности в схлопнувшемся и нагретом до температуры T лайнере есть $Q = \frac{lc\sigma}{3} \nabla T^4$, где l — усреднённый по Росселанду пробег квантов; σ — коэффициент Стефана—Больцмана. Пробег зависит от температуры и плотности лайнера: $l = a \frac{T^n}{\rho^k}$, $a = \text{const} > 0$.

Так что $Q = -a \frac{T^n}{\rho^k} \nabla T^4$, и высвечиваемая лайнером рентгеновская энергия есть $E \approx Q 2\pi \Delta \tau$, где Δ — радиус лайнера в момент схлопывания; $\tau \sim \Delta/u$ — время, по истечении которого схлопнувшийся лайнер начнёт расширяться (u — скорость лайнера перед схлопыванием). Длительность рентгеновского импульса как раз и определяется этим временем.

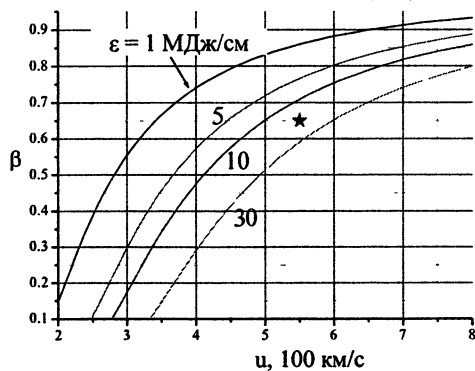
Коэффициент преобразования кинетической энергии лайнера в рентгеновское излучение есть, конечно,

$$\beta = \frac{E_0}{\varepsilon} \sim a \cdot \frac{T^{n+4}}{\rho^k \Delta} \cdot \frac{2\pi \Delta}{\frac{\mu u^2}{2}} \cdot \frac{\Delta}{u}, \quad (8)$$

где $a \cdot \frac{T^{n+4}}{\rho^k \Delta}$ — плотность рентгеновского потока, $2\pi \Delta$ — поверхность, через которую протекает поток, Δ/u — характерная длительность рентгеновского импульса и $\mu u^2/2$ — кинетическая энергия лайнера.

Выше уже отмечалось, что толщина лайнера $\Delta \sim (0,1-0,2) R_0$ определяется развивающимися при его сжатии неустойчивостями и для оценок её можно считать постоянной величиной. Так что плотность лайнера будет зависеть только от его погонной массы: $\rho \sim \mu \sim (\varepsilon/u^2)$. Температура же лайнера связана с его внутренней энергией: $c_v \mu T = \varepsilon(1-\beta)$, где c_v — удельная теплоёмкость вещества лайнера. Отсюда $T \sim u^2(1-\beta)$. Подставляя это в (8), найдём:

$$\beta \sim \frac{[u^2(1-\beta)]^{n+4}}{\left(\frac{\varepsilon}{u^2}\right)^{k+1} (u^2)^{3/2}}. \quad (9)$$



На рис. 7 приведена зависимость β от скорости лайнера, изготовленного из вольфрама, для различных значений скорости его полёта.

Рис. 7. Коэффициент преобразования кинетической энергии вольфрамового лайнера в рентгеновское излучение в зависимости от его скорости.

Теперь у нас есть всё, чтобы оценить масштаб установок.

6. Масштаб установки

Из рис. 7 видно, что стремиться к высоким коэффициентам преобразования β расточительно, так как $\partial\beta/\partial u$ при $\beta \geq 0,7$ стремительно падает. Выгодно работать в области, где $\beta(u)$ сильнее всего меняется

$\left(\frac{\partial^2 \beta}{\partial u^2} \right) \approx \max$. В этой области $\beta \approx 0,55 - 0,75$. Выбирая для оценок $\beta \approx \frac{2}{3}$,

видим, что скорость лайнера перед схлопыванием должна составить $u \sim 550$ км/с, а его кинетическая энергия $\epsilon \approx 24$ МДж/см. Максимальный ток через лайнер должен достигать (см. рис. 4) $I_0 \sim 100$ МА. Масса лайнера на единицу длины $\mu = 2\epsilon/u^2 \approx 160$ мг/см. Время нарастания тока, согласно (7), $T_0 \approx 130$ нс. Столь короткое время нарастания тока и составляет основную трудность при создании установки, ибо узлы, коммутирующие нагрузку с источником тока, должны работать при чрезвычайно высоких (мегавольты) напряжениях. И наконец, полная энергия, которую надо затратить на разгон при длине лайнера ~ 1 см, составит $E_{\text{полн}} = 2 \cdot 24 \text{ МДж} = 48 \text{ МДж} \approx 12 \text{ кг ВВ}$, поскольку для сжатия мишени мы использовали два первичных бокса. Видно, насколько экстремальны параметры, которыми должна обладать установка. И если они вообще достижимы, то определённо находятся на грани современных высоковольтных технологий. Нелишне заметить, что проведение на такой установке сколько-нибудь обширных экспериментальных программ при выделении в нагрузочном узле установки энергии порядка десятка килограммов ВВ в каждом опыте встретит специфические трудности. И всё же в Сандийской национальной лаборатории разработан проект стационарной установки X-I с параметрами, приближающимися к тем, что дают наши оценки: максимальный ток ~ 65 МА при времени нарастания около 150 нс. Во ВНИИЭФ также разработан проект установки ЭМИР с близкими характеристиками. Однако это должна быть разовая установка, поскольку источником энергии в ней служат мощные взрывомagnetные генераторы. Смогут ли электрофизические установки на основе Z-пинча стать реальной альтернативой лазерам при решении проблем управляемого термоядерного синтеза — покажет будущее.

Литература

1. Курчатов И. В. О возможности осуществления управляемой термоядерной реакции // Атомная энергия. 1956. Вып. 3. С. 65—75.
2. Spielman R. B., Deeney C., Chandler G. A et al. Dense Z-pinches // 4-th Int. Conf. Vancouver, Canada, 1997. P. 101—118.
3. Peterson D. L., Bowers R. L., Brownell J. H. et al. Dense Z-pinches // 4-th Int. Conf. Vancouver, Canada, 1997. P. 201—210.
4. Lindl J. Development of the indirect-drive approach to inertial confinement fusion and the target physics basis for ignition and gain // Phys. Plasmas. November, 1995.

КВАНТОВЫЙ ХАОС В ПРИЛОЖЕНИЯХ

В. Я. Демиховский

В последние годы активно развивалась квантовая механика тех динамических систем, чьи классические аналоги демонстрируют хаотическое поведение, и в настоящее время мы в основном понимаем, как классический хаос влияет на характеристики соответствующих квантовых объектов. Среди принципиально новых результатов, полученных в теории квантового хаоса, необходимо прежде всего отметить теорию случайных матриц, в которой рассматриваются статистические свойства энергетического спектра и волновых функций. Несомненным достижением теории является и развитие правил квазиклассического квантования хаотических систем (теория периодических орбит Гутцвиллера). Рассматривались и другие многочисленные проявления квантового хаоса, и в том числе теория квантовых бильярдных и различных мезоскопических рассеивающих систем. Одним из главных достижений теории квантового хаоса, несомненно, можно считать предсказание явления динамической локализации в системах, возбуждаемых внешним периодическим полем.

В это же время был достигнут значительный прогресс и в экспериментальных исследованиях квантового хаоса. В частности, были выполнены многочисленные эксперименты с ультрахолодными атомами, захваченными в магнитооптические ловушки, эксперименты по ионизации водородоподобных атомов сильным микроволновым полем, исследовался транспорт в низкоразмерных полупроводниковых мезоскопических структурах — квантовых бильярдах.

В настоящей лекции мы сосредоточимся в основном на обсуждении явлений, связанных с динамической локализацией. Это явление можно наблюдать в различных квантовых системах, на которые действует внешнее переменное поле. Динамическая локализация проявляется следующим образом. Если динамика классической системы имеет хаотический характер и в течение неограниченного времени происходит диффузия начальных состояний в ее фазовом пространстве, то в аналогичной квантовой системе, возбуждаемой переменным полем, диффузия через некоторое время останавливается. При этом интервал разбегания по квантовым состояниям можно характеризовать некоторой длиной локализации. Это явление впервые было исследовано в работе Чирикова, Израилева, Касати и Форда, которые численно изучали динамику одномерного ротатора, возбуждаемого периодическими δ -образными толчками [1]. Впоследствии теория этого явления получила развитие в работах [2—7]. Несмотря на то что динамическая локализация имеет много общего с локализацией Андерсона, она, как мы увидим, имеет другую природу.

Динамическая локализация, как было установлено в наших работах, может наблюдаться не только в условиях глобального хаоса, как это имеет место в случае ротатора, возбуждаемого периодическими толчками. Она ограничивает эволюцию квантовых систем и в условиях слабого хаоса, когда диффузия идет по сетке сепаратрис в вырожденной динамической системе или по паутине Арнольда в многомерной динамической системе [8—11]. Кроме теоретических моделей и методов, используемых в теории квантового хаоса, ниже будут рассматриваться и результаты последних экспериментов, в которых наблюдалась динамическая локализация, а также другие проявления хаоса в квантовых системах. К сожалению, у нас в стране последние достижения физики квантового хаоса пока не стали достоянием широкой научной общественности, что связано, прежде всего, с отсутствием соответствующих книг — учебников и монографий, посвященных данной проблеме. Поэтому, пользуясь случаем, я хотел бы обратить внимание читателей на то, что в 2004 году издательством «Физматлит» запланировано издание монографии профессора университета г. Марбург (Германия) Г. Штокмана, «Квантовый хаос. Введение» [12]. В этой книге весьма удачно сочетается достаточно полное изложение основ теории квантового хаоса с анализом последних экспериментальных результатов. Материалы этой монографии были широко использованы при подготовке настоящей лекции.

1. Квантовая динамическая локализация

Явление динамической локализации впервые было исследовано в работах Чирикова, Израйлева и Касати, которые изучали временную эволюцию одномерного квантового ротатора, возбуждаемого периодически повторяющимися δ -образными толчками [1]. Гамильтониан такой системы имеет сравнительно простой вид

$$H(t) = \frac{L^2}{2} + K \cos \theta \cdot \sum_n \delta(t - n). \quad (1)$$

В периоды между толчками ротатор свободно вращается с моментом импульса L . Коэффициент K определяет амплитуду толчков, момент инерции и период времени между толчками считаются равными единице. Уравнения движения, определяемые (1), сводятся к стандартному отображению Чирикова

$$\begin{cases} \theta_{n+1} = \theta_n + L_{n+1}, \\ L_{n+1} = L_n + K \sin \theta_n, \end{cases} \quad (2)$$

в котором L_n и θ_n — значения динамических переменных в момент сразу после толчка. Динамика, описываемая отображением (2), детально исследована. Если при малых значениях K она имеет вполне регулярный характер, то при $K > 0,9716\dots$, когда разрушается последний инвари-

антный тор, ограничивающий движение, становится диффузионной. Можно показать, что при этом среднее квадратичное отклонение L растёт по закону $\langle \Delta L^2 \rangle = K^2/2 \times N$, где N — число толчков. Следовательно, коэффициент диффузии можно определить как $D = K^2/2$.

Квантовую динамику плоского ротатора с δ -толчками можно описать, вводя оператор Флоке, который определяет эволюцию вектора состояния за период

$$\hat{F} = \exp\left(-\frac{i}{\hbar} k \cos \theta\right) \exp\left(-\frac{i}{\hbar} \frac{\tau}{2} \hat{L}^2\right), \quad (3)$$

где $\hat{L} = -i\hbar \frac{\partial}{\partial \theta}$ — оператор момента. Нетрудно убедиться, что в базе собственных функций этого оператора матрица оператора \hat{F} имеет вид

$$F_{nm} = \exp\left(-\frac{i}{\hbar} \frac{\tau}{2} m^2\right) \cdot i^{m-n} \cdot J_{m-n}\left(\frac{k}{\hbar}\right), \quad (4)$$

где J_{m-n} — функция Бесселя порядка $m-n$. Теперь, задавая вектор состояния ротатора A_0 в начальный момент, мы можем найти его в момент сразу после N -го толчка: $A_N = \hat{F}^N A_0$. Эволюцию начального состояния будем характеризовать величиной $\langle L^2 \rangle = A_N^+ \hat{L}^2 A_N$, для которой

$$\langle L^2 \rangle = \hbar^2 \sum_{l,m,n} l^2 (F^N)_{lm} (F^N)_{ln}^* a_m a_n^*, \quad (5)$$

где a_n — компоненты вектора начального состояния. Покажем, что при увеличении числа толчков, действующих на ротатор, величина $\langle L^2 \rangle$ стремится к постоянному пределу, т. е. квантовая диффузия останавливается. Для этого с помощью унитарного преобразования диагонализуем оператор Флоке, записав его как [12]

$$F_{ln} = \sum_k e^{-i\phi_k} U_{kl}^* U_{kn}.$$

Здесь ϕ_k — фазы Флоке, определенные на интервале 2π и связанные с квазиэнергиями соотношениями $E_k = \hbar/\tau \phi_k$. Пусть в начальный момент ротатор находился в состоянии с $l=0$, т. е. $a_0 = 1$, $a_{n \neq 0} = 0$. Тогда согласно (5) будем иметь

$$\langle L^2 \rangle = \hbar^2 \sum_{l,k,k'} l^2 e^{i(\phi_k - \phi_{k'})} U_{kl}^* U_{k0} U_{k'l} U_{k'0}^*. \quad (6)$$

Из выражения (4) следует, что матрица F_{nm} имеет ленточную структуру. То же самое можно сказать и о матрице оператора \hat{U} , осу-

существляющего диагональное преобразование \hat{F} . Следовательно, суммирование по k и k' в (6) эффективно идет лишь до некоторого конечного значения l_s . Фазы Флоке соответствующих состояний на интервале 2π распределены более или менее равномерно с плотностью $l_s/2\pi$, и поэтому наименьшая разность фаз имеет порядок $2\pi/l_s$. Если число произведенных толчков таково, что $N \ll l_s$, то система еще чувствует дискретность спектра. Это область классической диффузии. Если $N \gg l_s$, то фазовый множитель в (6) очень быстро осциллирует, и вклад в сумму дают лишь члены с $k \approx k'$. Тогда величина $\langle L^2 \rangle$ не зависит от N и равна

$$\langle L^2 \rangle = \hbar^2 \sum_{ik} l^2 |U_{ki}|^2 |U_{k0}|^2. \quad (7)$$

При этом квантовая диффузия останавливается, что свидетельствует о динамической локализации. Очевидно, что переход от режима диффузии к режиму локализации происходит при $N \sim l_s$ и в этот момент выражение (7) переходит в $\langle \Delta L \rangle^2 = l_s^2$, т. е. l_s играет роль длины локализации (напомним, что мы стартовали со значения $l = 0$).

На рис. 1, а приведена зависимость дисперсии момента импульса классического и квантового ротаторов от числа произведенных толчков. Расчет был выполнен при $K = 5$, когда классическая динамика является полностью хаотической [1]. Если для классического ротатора эволюция имеет характер диффузии при любых временах, то в случае квантового ротатора после примерно 50 толчков скорость диффузии резко уменьшается и после 1000 толчков она полностью останавливается.

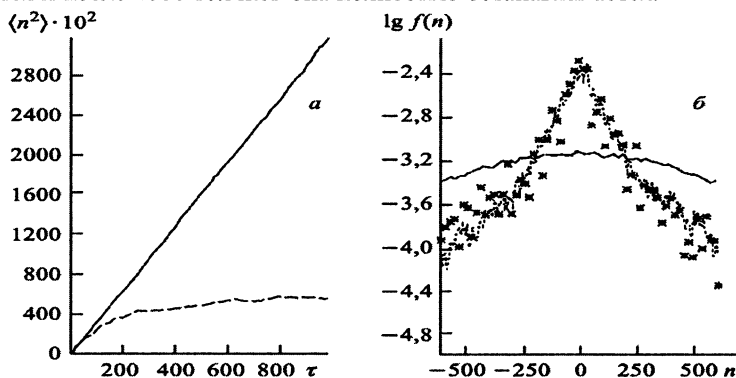


Рис. 1. Классическая и квантовая диффузия в фазовом пространстве ротатора с δ -толчками ($k = 5$). По вертикальной оси отложено среднеквадратичное отклонение момента импульса (здесь $\langle \Delta l^2 \rangle$) как функция числа толчков для классического (сплошная линия) и квантового (штриховая линия) ротатора (а). Классическое и квантовое распределение вероятности после 1000 толчков (б).

Рисунок 1, б иллюстрирует распределение вероятности различных состояний ротатора $f(n)$ после 1000 толчков. Если для классического ротатора это распределение имеет гауссовский характер, то в квантовой модели можно говорить об экспоненциальной динамической локализации с некоторой длиной локализации l_s .

Отметим еще одно важное соотношение, связывающее квантовую длину локализации с классическим коэффициентом диффузии [2, 6]:

$$l_s = aD. \quad (8)$$

Как показано в работах [7, 8], в модели Ллойда коэффициент пропорциональности a в этом соотношении равен 0,5.

Вскоре после открытия динамической локализации была установлена тесная связь этого явления с андерсоновской локализацией, которая имеет место в неупорядоченных одномерных системах [13]. Однако, несмотря на глубокую аналогию, существующую между этими явлениями, динамическая локализация по своей природе отличается от андерсоновской, поскольку последняя имеет место только в присутствии случайного потенциала, тогда как в рассматриваемых динамических моделях случайных функций нет.

Рассмотренная выше сравнительно простая модель одномерного ротатора с периодическими толчками позволяет описать динамику реальных атомных систем, на которые действует монохроматическое электромагнитное поле. Основные положения как классической, так и квантовой теории, включая эффект динамической локализации, нашли подтверждение в экспериментах с высоковозбужденными водородоподобными атомами, находящимися в сильном микроволновом поле, а также с атомами в магнитооптических ловушках. Обсудим основные идеи и результаты этих экспериментов.

1.1. Водородоподобные атомы в сильном микроволновом поле

Впервые динамическая локализация наблюдалась в экспериментах Байфилда и Коха с пучками высоковозбужденных атомов водорода, облучаемых сильным переменным микроволновым полем [14]. Описание экспериментальной установки, а также ссылки на последующие работы можно найти в монографии [12]. Здесь мы кратко рассмотрим только основные идеи эксперимента.

Предварительно с помощью CO_2 -лазера атомы водорода переводились в состояние с главным квантовым числом n , лежащим в интервале от 24 до 98. Далее пучок возбужденных таким образом атомов проходил через резонатор, в котором возбуждалось сильное микроволновое поле с частотой, изменяющейся в пределах от 7,6 ГГц до 36 ГГц. Под действием этого поля происходили переходы в состояния с различными главными квантовыми числами, причем часть атомов оказывалась ионизо-

ванной. Долю ионизованных атомов как функцию частоты и амплитуды поля можно было измерять.

Поскольку квантовая теория ионизации водородоподобного атома в переменном поле (теория Флоке) является достаточно громоздкой, в дальнейшем мы рассмотрим в основном выводы классической теории. При этом для простоты ограничимся моделью одномерного атома, считая, что электрон движется в атомном потенциале вида

$$V(z) = \begin{cases} -1/z, & z > 0, \\ \infty, & z < 0. \end{cases} \quad (9)$$

Уравнения Гамильтона рассматриваемой проблемы выглядят как

$$\dot{z} = \frac{\partial H}{\partial p} = p, \quad \dot{p} = \frac{\partial H}{\partial z} = -V'(z) + F \cos \omega t,$$

где $F = e\vec{E}$; e — элементарный заряд; \vec{E} — напряженность электрического поля; ω — частота. Интегрируя уравнения Гамильтона на интервале, равном одному периоду движения электрона T , при условии $\omega T \gg 1$ получим для фазы поля $\Phi = \omega t$ и энергии E дискретное отображение Кеплера. Обозначив значения этих величин перед n -м циклом как Φ_n и E_n , будем иметь

$$\Phi_{n+1} = \Phi_n + 2\pi\omega(2|E_{n+1}|)^{-3/2}, \quad E_{n+1} = E_n + \omega k \sin \Phi_n, \quad (10)$$

где $k = 2,58\omega^{-5/3}F$. Проведя линеаризацию по энергии вблизи некоторого начального значения главного квантового числа n_0 , можно свести (10) к стандартному отображению (2), где $L_n = 2\pi\hat{\omega}(1 - 3N_n\hat{h}\hat{\omega})$, $\theta_n = \Phi_n + \pi$, $\hat{\omega} = n_0^3\omega$. Здесь N_n — число поглощенных фотонов, n_0 — главное квантовое число исходного состояния атома, коэффициент K в стандартном отображении оказывается равным

$$K = 48,6\hat{\omega}^{1/3}F. \quad (11)$$

Тот факт, что отображение Кеплера приближенно сводится к стандартному отображению (2), очевидно, позволяет говорить об аналогии между динамикой электрона в атоме, облучаемом монохроматическим полем, и движением плоского ротатора, возбуждаемого δ -образными толчками. Как отмечалось выше, в случае стандартного отображения неограниченная диффузия в фазовом пространстве становится возможной при $K \geq 1$. В рассматриваемой задаче эта диффузия в итоге и приводит к ионизации. Следовательно, воспользовавшись приведенным выражением (11) для K , мы можем найти критическое значение амплитуды переменного поля при его фиксированной частоте. Согласно (11) критическая амплитуда оказывается пропорциональной $\omega^{-1/3}$. Именно

такая зависимость критического поля от частоты и была обнаружена в эксперименте [14].

При переходе в квантовую область амплитуда ионизирующего поля возрастает, что свидетельствует о проявлении динамической локализации. Здесь пороговую амплитуду можно определить из соотношения (8) между квантовой длиной локализации и классическим коэффициентом диффузии. Для того чтобы ионизация имела место, необходимо, чтобы длина локализации по энергии превышала расстояние от исходного ридберговского уровня до границы непрерывного спектра. Из этого условия, в частности, следует, что критическая амплитуда в квантовой области пропорциональна $\omega^{7/6}$. Оба упомянутых режима удалось наблюдать в эксперименте [14]. С деталями экспериментальной техники, а также с основными положениями квантовой теории динамической локализации в атомах, находящихся в сильном микроволновом поле, можно познакомиться, прочитав обзоры [15, 16].

2.1. Ультрахолодные атомы в магнитооптических ловушках

Динамическая локализация атомов, захваченных в магнитооптическую ловушку, впервые наблюдалась в [17]; основная идея эксперимента была сформулирована в [18]. Атомы захватывались и одновременно охлаждались в ловушке, где они двигались в неоднородном магнитном поле и в сильном поле лазера. Эксперимент проводился с атомами натрия, охлажденными до температуры порядка 10 мК, длина волны лазера соответствовала переходам между состояниями $2S_{1/2}$ и $3P_{3/2}$.

Эффективный гамильтониан, описывающий действие модулированной лазерной волны на электрон, находящийся в ловушке, и соответствующие уравнения движения можно представить в следующем безразмерном виде [10]:

$$H = \frac{p^2}{2} - k \cos(x - \lambda \sin t), \quad \dot{x} = \frac{\partial H}{\partial p} = p, \quad \dot{p} = \frac{\partial H}{\partial x} = -k \sin(x - \lambda t).$$

Здесь p — импульс атома, коэффициент k пропорционален квадрату амплитуды волны, λ определяет глубину модуляции волны [13]. При больших значениях λ аргумент синуса $\Phi(t) = x(t) - \lambda \sin t$ в последнем уравнении быстро осциллирует и средняя сила, действующая на атом, мала. Основное воздействие модулированная волна оказывает на атом в точках стационарной фазы t_s , где $\Phi'(t_s) = 0$ и $p(t_s) = \lambda \cos t_s$. Здесь сила действует на атом в фиксированном направлении, вызывая наибольшее изменение импульса. Это обстоятельство, как и в предыдущем случае, позволяет перейти к дискретному отображению, которое теперь можно записать как

$$\begin{aligned}
 x_{n+1} &= x_n + 2\pi p_{n+1}, \\
 p_{n+1} &= p_n - 2k\sqrt{\frac{2\pi}{\lambda}} \cos\left(\lambda - \frac{\pi}{4}\right) \sin x_n.
 \end{aligned}
 \tag{12}$$

Из сравнения с (2) следует, что данное отображение совпадает со стандартным. Чтобы убедиться в этом, достаточно в (12) произвести замену $2\pi p_n$ на L_n и ввести обозначение

$$K = 4\pi k \sqrt{\frac{2\pi}{\lambda} \cos\left(\lambda - \frac{\pi}{4}\right)}.$$

В результате можно ожидать, что динамика и в этой модели подобна динамике ротатора. Однако здесь при $K > 1$, когда движение имеет хаотический характер, классическая диффузия не продолжается сколь угодно долго, а останавливается в тот момент, когда будет выполнено условие $|p| \geq \lambda$ (см. условие существования точки стационарной фазы). При

этом в классическом режиме максимальная дисперсия импульса $\sqrt{\langle p^2 \rangle}$ будет иметь порядок λ . С другой стороны, как нам известно (см. (8)), квантовая длина локализации (в данном случае по импульсу) пропорциональна классическому коэффициенту диффузии, т. е. $D \sim \frac{1}{\lambda}$, и тогда должно иметь место соотношение вида $\sqrt{\langle p^2 \rangle} \sim \frac{1}{\lambda}$. Следовательно, в классическом режиме и в условиях квантовой динамической локализации конечное распределение по импульсам должно быть разным.

Основная цель эксперимента по наблюдению динамической локализации в системе атомов, захваченных и охлажденных в магнитооптической ловушке [17], состояла в сравнении их распределения по импульсам после лазерного облучения. Распределение атомов по импульсам при различных значениях глубины модуляции λ было проведено с использованием специальной экспериментальной техники (см., например, [13]). При слабой модуляции волны обнаружено линейное возрастание дисперсии импульсов с λ , а при больших значениях параметра модуляции наблюдался закон $\frac{1}{\lambda}$. Сравнение с результатами численного моделирования как в классической, так и в квантовой моделях говорит о хорошем согласии теории динамической локализации и эксперимента.

2. Слабый квантовый хаос в вырожденной динамической системе: диффузия по сетке сепаратрис

Как показано в работах [10, 11], динамическая локализация может наблюдаться не только в условиях сильного (глобального) квантового хаоса, как это имеет место в модели ротатора с периодическими толчка-

ми, но и в условиях слабого квантового хаоса. Нами изучалось движение заряженной частицы, на которую действует монохроматическая волна и одновременно постоянное магнитное поле, направленное перпендикулярно к волновому вектору. Такая классическая система относится к числу вырожденных, поскольку теорема Колмогорова—Арнольда—Мозера для нее не справедлива и уже при достаточно слабой амплитуде волны вблизи сепаратрис возникает стохастическая паутина, распространяющаяся на все фазовое пространство. Многочисленные примеры структур со стохастической паутиной приведены в монографии Заславского и Сагдеева [19]. Если стохастическая паутина занимает небольшую часть фазового пространства, можно говорить о слабом хаосе. В работах [10, 11] исследовалась квантовая диффузия по стохастической паутине. Гамильтониан изучавшейся вырожденной системы имел следующий вид:

$$\hat{H} = \frac{\left(\mathbf{p} + \frac{e}{c} \mathbf{A}\right)^2}{2m} + V_0 \cos(kx - \omega t), \quad (13)$$

где \mathbf{p} — импульс, e — заряд, m — масса частицы, \mathbf{A} — векторный потенциал, V_0 — амплитуда волны, k и ω — ее волновой вектор и частота. Квазиэнергетические функции, определяющие эволюцию начальных состояний, рассчитывались с помощью техники Флоке. Установлено, что диффузия вдоль квантовых сепаратрис может идти в двух различных режимах. При достаточно малой амплитуде волны она связана с так называемыми сепаратрисными квазиэнергетическими функциями, которые локализованы сразу в нескольких ячейках фазового пространства. Одна из таких функций показана на рис. 2.

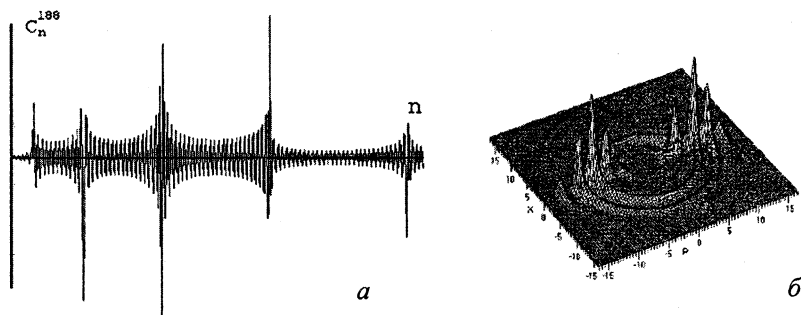


Рис. 2. Наиболее сильно делокализованная квазиэнергетическая функция (в представлении собственных функций Ландау) (а) и соответствующая ей функция Хушими (б).

Можно сказать, что в этом режиме квантовая диффузия обусловлена тунелированием между соседними ячейками фазового пространства. С ростом амплитуды возникает слабый квантовый хаос в области сепаратрис.

ратрис, что приводит к динамической локализации и подавлению диффузии. Однако при дальнейшем увеличении амплитуды, когда область стохастической паутины становится столь широкой, что в нее попадает много квантовых состояний, скорость диффузии вновь возрастает и стремится к классическому пределу.

Оригинальный способ экспериментальной реализации системы, рассмотренной выше, был предложен в [20]. Авторы этой работы обратили внимание на то, что динамика электрона в полупроводниковой сверхрешетке, помещенной в наклонное магнитное и перпендикулярное электрическое поле, подобна динамике в поле бегущей волны и в перпендикулярном магнитном поле (рис. 3).

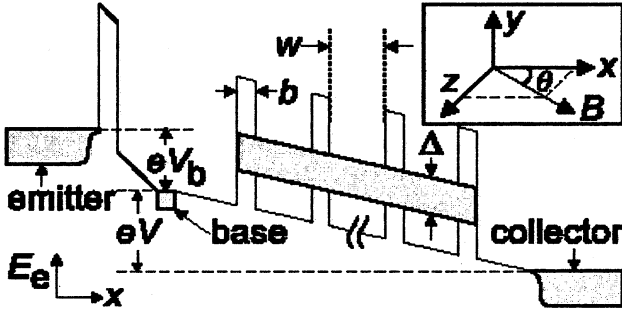


Рис. 3. Сверхрешетка в наклонном магнитном и перпендикулярном электрическом поле. Энергия отложена по вертикальной оси, b — ширина барьеров, W — ширина ям, Δ — ширина энергетической зоны.

Эффект бегущей волны здесь достигается за счет блоховских осцилляций с частотой $\omega_b = eEd/\hbar$ в наклонных электрическом поле энергетических зонах сверхрешетки. Соответствующий классический гамильтониан можно записать как

$$H = \frac{\Delta}{2} \left(1 - \cos \frac{p_x d}{\hbar} \right) + \frac{p_z^2}{2m} + \frac{m\omega_c^2}{2} \cdot [x \sin \theta - (z - z_0) \cos \theta]^2 - eEx. \quad (14)$$

«Стохастические» квантовые состояния, соответствующие классической паутине в фазовом пространстве этой системы, должны давать определенный вклад в транспорт (проводимость) электронного газа, находящегося в сверхрешетке.

3. Квантовая диффузия Арнольда

Проявления динамической локализации были обнаружены и в гамильтоновской динамической системе с числом степеней свободы $N = 2, 5$, где при произвольно малом возмущении существует большое число резонансов, по которым и идет диффузия Арнольда. Если при

$N = 2$ малые области хаотического движения в фазовом пространстве отделены друг от друга устойчивыми торами, то при $N \geq 2,5$ резонансные торы пересекаются и в результате независимо от величины возмущения все фазовое пространство оказывается пронизанным сеткой стохастических слоев, двигаясь по которым частица диффундирует, удаляясь сколь угодно далеко от невозмущенной траектории. Квантовая теория этого явления рассматривалась в наших работах [8, 9], соответствующая классическая модель ранее исследовалась Чириковым [21]. Нами изучалось движение двух связанных нелинейных осцилляторов, возбуждаемых переменным двухчастотным полем. Соответствующий оператор Гамильтона записывался как

$$\hat{H} = \frac{\hat{p}_x^2}{2} + \frac{x^4}{4} + \frac{\hat{p}_y^2}{2} + \frac{y^4}{4} - \mu xy - f_0 x (\cos \Omega_1 t + \cos \Omega_2 t), \quad (15)$$

где $[\hat{p}_x, x] = -i\hbar_0$; $[\hat{p}_y, y] = -i\hbar_0$; \hbar_0 — безразмерная постоянная Планка; f_0 — амплитуда переменного поля. Частоты Ω_1, Ω_2 выбирались кратными так, что общий период равен $T = 2\pi n/\Omega_1 = 2\pi m/\Omega_2$ ($n, m \in N$).

Можно убедиться в том, что при $f_0 = 0$ спектр соответствующей стационарной задачи вблизи резонанса связи ($\omega_x = \omega_y$) может быть представлен в виде $E_q = \hbar_0 \omega q + E_{q,s}^M$, где квантовое число q характеризует номер группы, а s — состояние внутри группы. Здесь $E_{q,s}^M$ — спектр типа Матье [13]. Квазиэнергетический спектр и функции рассчитывались с помощью теории Флоке. Установлено, что каждая квазиэнергетическая функция A_{qs} , записанная в представлении собственных функций двух взаимодействующих нелинейных осцилляторов, может характеризоваться «средним» номером группы \bar{q} и дисперсией, определяемой как

$\Delta_q = \sum_q (q - \bar{q})^2 \sum_s |A_{qs}|^2$. Проведенные в [8, 9] расчеты квантового коэффициента диффузии вдоль сепаратрисы резонанса связи двух осцилляторов показали, что при всех значениях параметров гамильтониана он меньше классического коэффициента диффузии. В том случае, когда число квантовых состояний, попадающих в стохастический слой, становится порядка единицы, квантовая диффузия останавливается.

Далее была рассчитана зависимость дисперсии от времени. Результат представлен на рис. 4. По характеру кривой можно сделать заключение о том, что диффузия продолжается лишь в течение ограниченного времени, а в дальнейшем при временах больших, чем $10^3 T$, наблюдаются осцилляции Δ_q вокруг некоторого среднего значения $\bar{\Delta}_q$, показанного на рис. 4, а пунктиром. Эта величина играет здесь роль длины локализации. Период осцилляций зависит от константы μ немонотонно.

Ясно, что такое поведение является следствием динамической локализации, которую можно объяснить локализацией квазиэнергетических сепаратрисных функций. Отметим однако, что в данном случае динамическая локализация имеет место в системе с числом степеней свободы $N = 2,5$, в то время как в случае ротатора с периодическими толчками $N = 1,5$. Необходимо подчеркнуть также, что в примерах, рассмотренных в п. 2 и п. 3, речь идет о динамической локализации в условиях слабого хаоса на сепаратрисах, а не глобального хаоса, как в случае ротатора с δ -толчками. На рис. 4, б показана зависимость длины локализации от параметра $1/\sqrt{\mu}$. Она имеет экспоненциальный характер, что не удивительно, поскольку рассчитанный коэффициент диффузии ведет себя таким же образом.

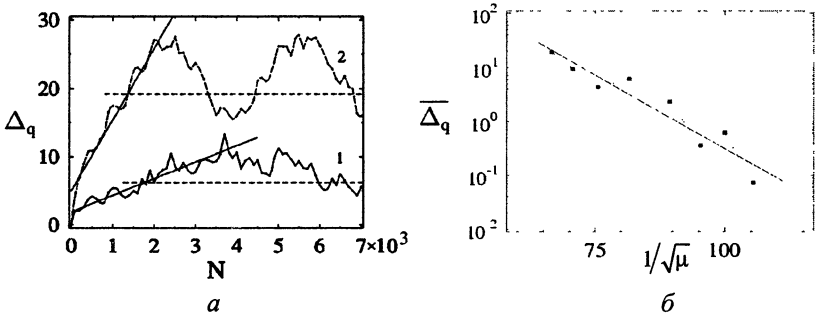


Рис. 4. Зависимость величины Δ_q от времени для начальных условий на сепаратрисе резонанса связи при $\mu = 1,5 \cdot 10^{-4}$ (кривая 1), $\mu = 2,25 \cdot 10^{-4}$ (кривая 2) (а). Распределение параметра локализации $\overline{\Delta_q}$ в зависимости от $1/\sqrt{\mu}$ (б).

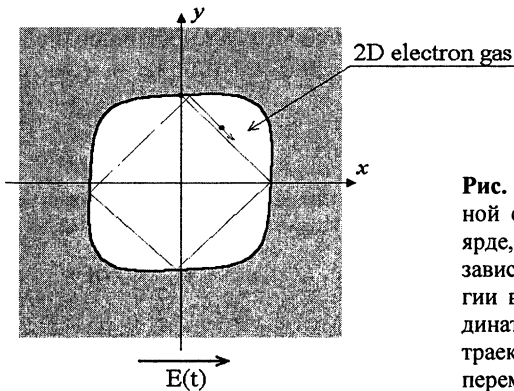


Рис. 5. Электронный газ в двумерной структуре — квантовом бильярде, форма которого определяется зависимостью потенциальной энергии в гамильтониане (15) от координат. Показана одна хаотическая траектория. $E(t)$ — напряженность переменного электрического поля.

В заключение настоящего раздела заметим, что различные проявления квантовой диффузии Арнольда, и в том числе динамическую локализацию в условиях слабого квантового хаоса, можно наблюдать в экспериментах с двумерным электронным газом, находящимся в полупроводниковой структуре — квантовом бильярде, где движение определяется гамильтонианом типа (15) (рис. 5).

Мы рассмотрели одно из наиболее ярких проявлений квантового хаоса — динамическую локализацию, ограничивающую диффузионные процессы в квантовых системах, находящихся во внешних переменных полях. Показано, что динамическая локализация может наблюдаться как в условиях сильного хаоса, когда движение практически во всем классическом фазовом пространстве имеет хаотический характер, так и в условиях слабого хаоса, а именно на сетке сепаратрис вырожденной динамической системы или на паутине Арнольда. В разделе 2.1 обсуждались эксперименты по ионизации предварительно возбужденных на высокие ридберговские уровни атомов водорода. В этих экспериментах впервые удалось наблюдать динамическую локализацию. Проявления динамической локализации были обнаружены и в экспериментах с атомами, захваченными в магнитооптические ловушки и охлажденными до температуры порядка нескольких микрокельвинов (раздел 2.2). Теория, описывающая два упомянутых эксперимента, основана на модели одномерного квантового ротатора с периодическими толчками.

Динамическая локализация обнаружена и при изучении квантовой диффузии по бесконечной сетке сепаратрис в вырожденной гамильтоновской системе (раздел 2), а также по паутине Арнольда в системе с числом степеней свободы больше двух (раздел 3). Обсуждались возможные пути экспериментальной реализации этих явлений.

Автор благодарен проф. Ф. М. Израйлеву за многочисленные полезные обсуждения вопросов, составивших предмет настоящей статьи, а также А. И. Малышеву за обсуждение и подготовку ее рукописи. Работа поддержана фондом РФФИ (проект №03-02-17054) и грантом программы «Университеты России».

Литература

1. *Casati G., Chirikov B. V., Izrailev F. M., Ford J.* Stochastic behavior of a quantum pendulum under a periodic perturbation // *Lecture Notes in Physics*. Berlin: Springer-Verlag, 1979. V. 93. P. 334.
2. *Chirikov B. V., Izrailev F. M., Shepelyansky D. L.* Quantum chaotic localization vs. ergodicity // *Physica D*. 1988. V. 33. P. 77.
3. *Casati G., Chirikov B.* The legacy of chaos in quantum mechanics // *Quantum Chaos*. Cambridge University Press, 1995. P. 3.
4. *Izrailev F. M.* Simple models of quantum chaos: Spectrum and eigenfunctions // *Phys. Rep.* 1990. V. 196. P. 299.
5. *Casati G., Guarneri I., Shepelyansky D. I.* Exponential photonic localization for the hydrogen atom in a monochromatic field // *Phys. Rev. A*. 1987. V. 66. P. 3501.

6. Chirikov B. V., Izrailev F. M., Shepelyansky D. L. Dynamical stochasticity in classical and quantum mechanics // *Sov. Sci. Rev.* 1981. V. 2. P. 209.
7. Shepelyansky D. L. Localization of quasienergy eigenfunction in action space // *Phys. Rev. Lett.* 1986. V. 56. P. 677.
8. Demikhovskii V. Ya., Izrailev F. M., Malyshev A. I. Manifestation of the Arnol'd diffusion in quantum system // *Phys. Rev. Lett.* 2001. V. 88. P. 154101.
9. Demikhovskii V. Ya., Izrailev F. M., Malyshev A. I. Quantum Arnol'd diffusion in a simple nonlinear system // *Phys. Rev. E.* 2002. V. 66. P. 036211.
10. Demikhovskii V. Ya., Kamenev D. I., Luna-Acosta G. A. Quantum weak chaos in a degenerate system // *Phys. Rev. E.* 1999. V. 59. P. 294.
11. Demikhovskii V. Ya., Kamenev D. I., Luna-Acosta G. A. Quantum resonance in an intrinsically degenerated system: Nonlinear cyclotron resonance // *Phys. Rev. E.* 1995. V. 52. P. 3351.
12. Штокман Х.-Ю. Квантовый Хаос. Введение. М.: Физматлит (план 2004 г.).
13. Fishman S. et al. Chaos, quantum recurrence and Anderson localization // *Phys. Rev. Lett.* 1989. V. 49. P. 509.
14. Bayfield J. E., Koch P. M. Multiphoton ionization of highly excited hydrogen atoms // *Phys. Rev. Lett.* 1974. V. 33. P. 258.
15. Koch P. M. Microwave ionization of excited hydrogen atoms // *Physica D.* 1995. V. 83. P. 178.
16. Casati G., Chirikov B. V., Shepelyansky D. L., Guarneri I. Relevance of quantum chaos in quantum mechanics. The hydrogen atom in a monochromatic field // *Phys. Rep.* 1987. V. 154. P. 77.
17. Moore et al. Atom optics realization of the quantum kicked rotor // *Phys. Rev. Lett.* 1995. V. 75. P. 4598.
18. Graham R. et al. Dynamical localization of atomic-beam deflection by a modulated standing light wave // *Physica A.* 1992. V. 45. P. 19.
19. Заславский Г. М., Сагдеев П. З. Введение в нелинейную физику. От маятника к турбулентности и хаосу. М: Наука, 1985. 368 с.
20. Fromhold T. M. et al. Effects of stochastic webs on chaotic electron transport in semiconductor superlattices // *Phys. Rev. Lett.* 2001. V. 87. P. 046803.
21. Chirikov B. V. A universal instability of many-dimensional oscillator systems // *Phys. Rep.* 1979. V. 52. P. 263.

НЕЛИНЕЙНОСТЬ ВАКУУМА И ПЕРСПЕКТИВЫ ЕЕ НАБЛЮДЕНИЯ

Вл. В. Кочаровский, В. Ю. Мартьянов

Введение

Согласно квантовой теории поля [1, 2] вакуум — это основное состояние квантованных полей, обладающее минимальной энергией и нулевыми значениями импульса, момента импульса, заряда и других квантовых чисел. В частности, в вакууме отсутствуют реальные частицы. Однако возможность виртуальных процессов обеспечивает нетривиальные физические свойства вакуума, в том числе его нелинейность. Нас будет интересовать одна из проблем нелинейности квантовоэлектродинамического вакуума, составляющая в настоящее время целую область исследований. В данной краткой лекции мы обсудим нелинейное смещение достаточно сильных электромагнитных волн, распространяющихся в вакууме. В присутствии электромагнитных волн нулевые колебания электрон-позитронного поля видоизменяются, т. е. виртуальные пары движутся иначе, чем в отсутствие поля, и следовательно, появляются нелинейные поляризация и намагниченность вакуума. Говоря квантовым языком, благодаря радиационным поправкам в вакууме возможно фотон-фотонное рассеяние [3–7]. Нелинейные члены в уравнениях Максвелла исчезают для случая параллельно распространяющихся плоских волн. Однако нелинейные вакуумные эффекты проявляются при пересечении электромагнитных пучков и для мод волноводов и резонаторов [8–11].

В данной статье мы рассмотрим как известные, так и новые схемы, позволяющие наблюдать указанную нелинейность, включая генерацию третьей гармоники в высокочастотных СВЧ и квазиоптических резонаторах и генерацию комбинационных частот при взаимодействии лазерных пучков с СВЧ-полями.

Легко понять происхождение нелинейности вакуума. Хорошо известно, что заряженная частица и ее античастица, например электрон и позитрон, могут аннигилировать с рождением пары фотонов. Соответственно существует и обратный процесс, в котором два фотона взаимодействуют с образованием электрон-позитронной пары. Эта пара может сразу исчезнуть или быть виртуальной, и опять породить два фотона. Тогда весь этот процесс в целом можно рассматривать как рассеяние фотонов друг на друге (рис. 1), а если

фотоны взаимодействуют, то электромагнитное поле должно описываться нелинейными уравнениями.

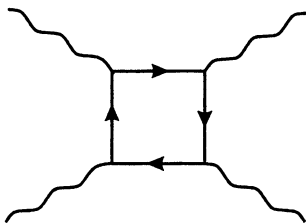


Рис. 1. Фотон-фотонное рассеяние на виртуальных парах

Учет рождения виртуальных электрон-позитронных пар приводит к появлению дополнительного слагаемого в лагранжиане, вычисленного Гейзенбергом и Эйлером в 1936 году [12]:

$$\mathcal{L}_{HE} = \frac{\mathbf{E}^2 - \mathbf{B}^2}{8\pi} + \frac{1}{8\pi^2} \int_0^\infty \left\{ 1 + \frac{e^2 \eta^2}{3} (\mathbf{B}^2 - \mathbf{E}^2) - e^2 \eta^2 \mathbf{B} \mathbf{E} \times \right. \\ \left. \times \frac{\operatorname{Re} \operatorname{ch} (e\eta \sqrt{\mathbf{B}^2 - \mathbf{E}^2 + 2i\mathbf{B}\mathbf{E}})}{\operatorname{Im} \operatorname{ch} (e\eta \sqrt{\mathbf{B}^2 - \mathbf{E}^2 + 2i\mathbf{B}\mathbf{E}})} \right\} e^{-m^2 \eta} \frac{d\eta}{\eta^3}.$$

Постоянная Планка \hbar и скорость света c приняты здесь за единицу, $\alpha \equiv e^2/\hbar c = e^2$ — постоянная тонкой структуры, e — заряд электрона, m — масса электрона. Другие частицы не принимаются во внимание, потому что электроны и позитроны имеют наименьшую массу и, соответственно, эффективнее других рождаются в электромагнитных полях \mathbf{E} , \mathbf{B} .

Нелинейность в слабых низкочастотных полях

Рождение электрон-позитронных пар экспоненциально мало, если работа электрического поля на комптоновской длине волны много меньше энергии покоя электрона, т. е. если электрическое поле меньше так называемого критического поля $E_c = mc^2/e\lambda$, где $\lambda = \hbar/mc$. В лабораторных условиях рождение реальных пар возможно только в экстремальных случаях, таких как столкновение атомных ядер, движущихся с релятивистскими скоростями, или фокусировка рентгеновских лучей [5, 11, 13]. Мы будем подразумевать, что все поля много меньше критических, а частоты достаточно малы, чтобы можно было пренебречь дисперсионными эффектами, связанными

с квантовой природой света [6, 7, 14]:

$$E, B \ll E_c = B_c = m^2 c^3 / e \hbar = 4,4 \cdot 10^{13} \text{ Гс} \approx 1,3 \cdot 10^{16} \text{ В/см},$$

$$\lambda \sim c/\omega \gg mc^2/e(E+B), \quad \text{т. е.} \quad \hbar\omega \ll \lambda e(E+B).$$

Оба эти предположения хорошо выполняются для современных мощных источников микроволнового и оптического излучения.

Используя теорию возмущений, можно получить хорошо известный вид лагранжиана электромагнитного поля [3]

$$8\pi\mathcal{L} = \mathbf{E}^2 - \mathbf{B}^2 + \frac{\alpha}{45\pi B_c^2} \left[(\mathbf{E}^2 - \mathbf{B}^2)^2 + 7(\mathbf{E}\mathbf{B})^2 \right]$$

и стандартный вид уравнений Максвелла, в которых поляризация и намагниченность кубически зависят от электрического и магнитного полей:

$$\begin{aligned} \operatorname{div} \mathbf{B} &= 0, & \operatorname{div} (\mathbf{E} + 4\pi\mathbf{P}) &= 0, \\ \operatorname{rot} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, & \operatorname{rot} (\mathbf{B} - 4\pi\mathbf{M}) &= \frac{1}{c} \frac{\partial (\mathbf{E} + 4\pi\mathbf{P})}{\partial t}; \\ \mathbf{P} &= \frac{\xi}{4\pi} [2(E^2 - B^2)\mathbf{E} + 7(\mathbf{E} \cdot \mathbf{B})\mathbf{B}], \\ \mathbf{M} &= -\frac{\xi}{4\pi} [2(E^2 - B^2)\mathbf{B} - 7(\mathbf{E} \cdot \mathbf{B})\mathbf{E}]. \end{aligned}$$

Характерная величина нелинейных полей определяется зарядом и массой электрона:

$$\xi = \frac{\hbar e^4}{45\pi m^4 c^7} \approx 0,26 \cdot 10^{-31} \text{ ед. СГС},$$

$$E_{NL} \sim \xi^{-1/2} \approx 0,6 \cdot 10^{16} \text{ ед. СГС} \approx 2 \cdot 10^{18} \text{ В/см}.$$

Во всех обсуждаемых в данной работе процессах реальные электрон-позитронные пары не рождаются, а рождение виртуальных пар проявляется только через появление нелинейности в уравнениях Максвелла. Другими словами, уравнения не содержат источников в виде реальных токов или зарядов. Важно также подчеркнуть, что все нелинейные слагаемые обращаются в ноль для случая одной плоской электромагнитной волны или для нескольких волн, распространяющихся в одном направлении.

В первых разделах мы кратко обсудим некоторые из предсказанных для вакуума нелинейных эффектов и возможные схемы их обнаружения. Далее остановимся на возможности обнаружения третьей гармоники СВЧ-излучения, порожденной вакуумом. Наконец,

в двух заключительных разделах рассмотрим новые комбинированные СВЧ и оптические схемы, которые являются естественным обобщением схем, ранее предлагавшихся либо в СВЧ, либо в оптическом диапазоне. (См., например, работы [9, 10, 13, 14], в которых, в частности, обсуждаются четырехволновое смешение и самовоздействие электромагнитных полей).

Намагниченный вакуум как анизотропная среда

Выражения для нелинейной поляризации и намагниченности вакуума содержат тройные произведения полей. Если считать часть этих полей заданными внешними полями, можно получить линейные, квадратичные и кубические эффекты.

Один из линейных эффектов — двулучепреломление в намагниченном вакууме. В присутствии сильного магнитного поля, две нормальные волны в уравнениях Максвелла имеют разные показатели преломления. Зависимость этих показателей преломления от значения внешнего магнитного поля легко вычисляется и приведена на рис. 2, где θ — угол между волновым вектором и магнитным полем \mathbf{B}_0 . Если внешнее магнитное поле много меньше критического, $B_0 \ll B_c$, и частота волны не слишком велика, $\hbar\omega \ll mc^2 |\sin \theta| B_c/B_0$, то поправка к фазовой скорости волны квадратично зависит от магнитного поля (см., например, [15, 16]):

$$\begin{cases} n_e \approx 1 + (7/2)a \sin^2 \theta \\ n_o \approx 1 + 2a \sin^2 \theta \end{cases}, \quad a = \frac{\alpha B_0^2}{45\pi B_c^2}.$$

Таким образом, вакуум в сильном магнитном поле может влиять на поляризацию электромагнитных волн, распространяющихся, например, через неоднородную магнитоактивную плазму. Наблюдение особенностей поляризации излучения астрофизических источников (например, нейтронных звезд) с сильными магнитными полями

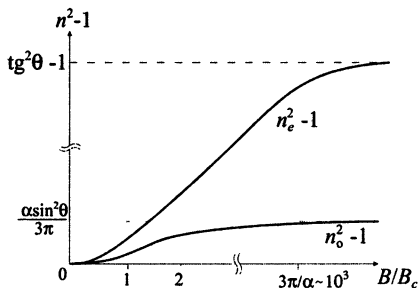


Рис. 2. Двулучепреломление в намагниченном вакууме; n_o и n_e — показатели преломления обыкновенной и необыкновенной волн соответственно.

может дать косвенные свидетельства нелинейности вакуума. Однако надежных экспериментальных подтверждений такого рода пока нет, несмотря на многочисленные обсуждения такой возможности и некоторые попытки наблюдений (см., например, [15—19]).

Нелинейный сдвиг фазы

Фазовый набег, приводящий к изменению поляризации излучения, может также появляться при распространении пучка через переменное электромагнитное поле. Рассмотрим ситуацию, в которой сфокусированный интенсивный лазерный пучок с амплитудой поля E и пробный пучок распространяются в противоположных направлениях на трассе длиной l (рис. 3). Из-за нелинейного взаимодействия пробный пучок приобретет фазовый набег, пропорциональный длине области взаимодействия и зависящий от поляризации пучка: $\delta\Phi \sim \alpha(|E^2|/E_c^2)(l/\lambda)$. Согласно оценкам [20] при длине волны $\lambda \sim 1$ мкм мощный субнаносекундный импульс с поперечным сечением $S \sim (300 \text{ мкм})^2$, длиной области взаимодействия $l \sim (S/\lambda) \sim 10$ см и энергией $\mathcal{E} \sim 10$ кДж дает очень небольшой фазовый набег $\delta\Phi \sim 10^{-10}$ рад, труднодоступный для детектирования.

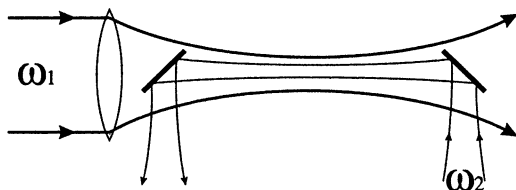


Рис. 3. Схема наблюдения нелинейного сдвига фазы пробного поля частоты ω_2 на мощном пучке излучения с частотой ω_1

В вакууме возможны и многие другие нелинейно-оптические эффекты, включая генерацию второй гармоники [8], формирование ударных волн [25, 26], параметрические неустойчивости, самофокусировку и каналирование во встречных пучках [9, 24]. В частности, критическая мощность самофокусировочной неустойчивости составляет [9]

$$P_{cr} \sim 10^3 E_c^2 \lambda^2 \sim 2,5 \cdot 10^{24} \text{ Вт.}$$

Еще одним интересным нелинейным эффектом является распад фотона в намагниченном вакууме [3, 21—23]. В присутствии магнитного поля обыкновенный фотон γ_0 с волновым вектором \mathbf{k}_0 может

распасться на два необыкновенных фотона γ_e с волновыми векторами \mathbf{k}_1 и \mathbf{k}_2 , образующими конечный угол (рис. 4).

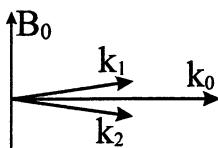


Рис. 4. Распад фотона $\gamma_0 \rightarrow \gamma_e + \gamma_e$ в намагниченном вакууме

Такое трехволновое взаимодействие также детально изучалось; в частности, известны автомодельные решения кинетических уравнений переноса фотонов при наличии рассеяния [21, 22].

Четырехволновое смещение в вакууме

Рождение и аннигиляция виртуальных электрон-позитронных пар приводит к фотон-фотонному рассеянию $\gamma_1 + \gamma_2 \rightarrow \gamma' + \gamma''$ (рис. 1). Его эффективность характеризуется хорошо известным сечением рассеяния σ , которое вычисляется методами квантовой электродинамики и имеет следующую асимптотику [3, 4]:

$$\sigma \approx 0,03\alpha^2 \left(\frac{e^2}{mc^2} \right)^2 \left(\frac{\hbar\omega}{mc^2} \right)^6, \quad \hbar\omega \ll mc^2;$$

$$\sigma \approx 4,7\alpha^4 (c/\omega)^2, \quad \hbar\omega \gg mc^2.$$

Максимальное значение $\sigma_{\max} \sim 10^{-30} \text{ см}^2$ достигается при $\hbar\omega \approx \approx 1,5mc^2$. Чрезвычайная малость значения σ_{\max} объясняет трудность наблюдения нелинейных эффектов в вакууме.

Процесс генерации суммарной частоты $\omega_s = \omega_1 + \omega_2 + \omega_3$ в вакууме запрещен. Для выполнения условий синхронизма и дисперсионного уравнения в этом случае потребовалось бы, чтобы волновые векторы всех волн, участвующих в процессе, были сонаправлены, $\mathbf{k}_s \uparrow\uparrow \mathbf{k}_1 \uparrow\uparrow \mathbf{k}_2 \uparrow\uparrow \mathbf{k}_3$, что обращает в нуль нелинейные поправки в уравнениях Максвелла.

Однако для генерации разностной частоты $\omega_s = \omega_1 + \omega_2 - \omega_3$ условие синхронизма $\mathbf{k}_s = \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3$ можно выполнить [14] (рис. 5).

Если зафиксировать три частоты $\omega_1, \omega_2, \omega_3$, а значит, и длины векторов k_1, k_2, k_3 , а также зафиксировать направления векторов \mathbf{k}_1 и \mathbf{k}_2 , то любая точка на пересечении сфер с радиусами k_s и k_3

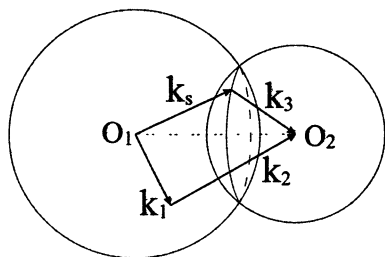


Рис. 5. Выполнение условия синхронизма для генерации разностной частоты

(рис. 5) будет удовлетворять условию синхронизма. Электрическое поле рассеянной волны будет пропорционально нелинейной поляризации и частоте ω_s и будет линейно расти с ростом координаты z вдоль волнового вектора k_s :

$$E_{sx}(z) \propto i\omega_s z (P_{sx}^{NL} + M_{sy}^{NL}).$$

Остановимся на конкретной реализации подобного процесса.

Схема обращения волнового фронта

А именно, частным случаем четырехволнового смешения является обращение волнового фронта, которое реализуется при выполнении условий

$$k_s = -k_3, \quad k_1 = -k_2, \quad \omega_s = \omega_1 = \omega_2 = \omega_3 = \omega.$$

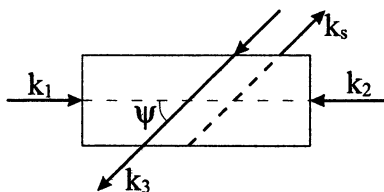


Рис. 6. Схема обращения волнового фронта.

В этом случае рассеянная волна распространяется строго противоположно третьей волне, чем и объясняется название процесса (рис. 6). Данная схема широко используется в нелинейной оптике, и амплитуду рассеянной волны легко найти [14]:

$$E_s(l) \sim i\xi(l\omega/c)(3 + \cos^2 \psi) E_1 E_2 E_3^*,$$

где $\xi \approx (\alpha/B_c)^2$, а l — путь, пройденный рассеянной волной в области взаимодействия.

Для случая нелинейности вакуума согласно [14] при равных интенсивностях трех лазерных пучков, $I_1 = I_2 = I_3 = \mathcal{E}/S\tau$,

энергия в рассеянном импульсе \mathcal{E}_s не зависит от длительности импульсов τ . Для случая малых углов, $\psi \ll 1$, рассеянная энергия $\mathcal{E}_s \sim 10^{-4} c^2 \mathcal{E}^3 / (E_c S \lambda)^2$. При $\lambda \sim 1$ мкм и $S \sim (100 \text{ мкм})^2$ для получения нескольких рассеянных фотонов требуется энергия $\mathcal{E} > 1$ кДж в каждом из лазерных импульсов.

Для СВЧ и радиочастотного диапазонов (для краткости будем говорить об СВЧ) предлагалось детектировать нелинейность вакуума с использованием волноводов [10]. При распространении моды вдоль волновода создаваемая ею нелинейная поляризация постепенно возбуждает третью гармонику. Теоретически последняя может быть измерена, однако простые оценки показывают, что для этого потребовался бы волновод более миллиона километров в длину. Поэтому более практичным является использование не волновода, а резонатора, на чем мы и остановимся ниже.

Генерация третьей гармоники в высокодобротном резонаторе

Если в резонаторе с высокой добротностью Q возбуждена основная мода с частотой Ω , то в нелинейной поляризации и намагниченности будут присутствовать компоненты, осциллирующие с частотой 3Ω . Если эта частота также является собственной для резонатора, то амплитуда третьей гармоники будет расти до некоторого уровня насыщения $E_{3\Omega}$. Чтобы третью гармонику было легче детектировать, удобно сконструировать резонатор таким образом, чтобы основная мода имела пренебрежимо малую амплитуду в некоторой части резонатора и не мешала измерениям третьей гармоники. Для этого можно сделать комбинированный резонатор, состоящий из двух частей, соединенных закритическим волноводом для частоты Ω (рис. 7, см. также [10]).

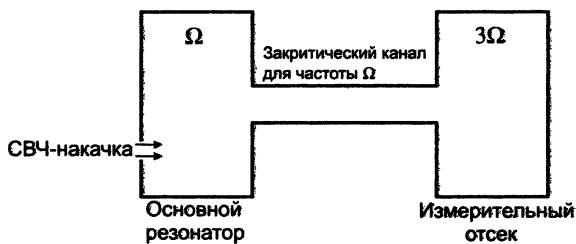


Рис. 7. Комбинированный резонатор для детектирования третьей гармоники

Для оценок рассмотрим модельную задачу с простейшей геометрией резонатора, поддающуюся аналитическому решению. Конкретно в качестве основной моды возьмем моду $m_1 n_1 0$ прямоугольного резонатора (рис. 8).

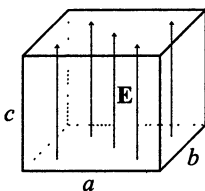


Рис. 8. Модельный прямоугольный резонатор

В этом случае легко находится условие того, что 3Ω — собственная частота резонатора:

$$\frac{m_3^2}{a^2} + \frac{n_3^2}{b^2} = 9 \left(\frac{m_1^2}{a^2} + \frac{n_1^2}{b^2} \right).$$

Стационарная амплитуда третьей гармоники, возбужденной нелинейными источниками P_3 и M_3 , может быть найдена стандартными методами теории возмущений:

$$E_{3\Omega} = \frac{-4\pi \cdot 3k_1}{(9k_1^2 - k_3^2)} \left[\frac{3k_1 \int P_3 E^{*(m)} dV}{\int E^{(m)} E^{*(m)} dV} + \frac{k_3 \int M_3 H^{*(m)} dV}{\int H^{(m)} H^{*(m)} dV} \right],$$

где $\mathbf{E}^{(m)}(\mathbf{r})$ и $\mathbf{H}^{(m)}(\mathbf{r})$ — нормированные на единицу (по амплитуде) поля возбуждаемой моды. С учетом конечной добротности сверхпроводящего резонатора Q на частоте 3Ω получаем оценку $E_{3\Omega} \sim \xi E_{\Omega}^3 Q$.

Здесь и в последующих оценках мы учитываем только слагаемые, обусловленные нелинейной электрической поляризацией вакуума. Слагаемые с нелинейной намагниченностью имеют тот же порядок, и эти два эффекта в принципе могут частично компенсировать друг друга, особенно в параксиальных схемах, подобных рассмотренным ниже.

Последняя оценка, очевидно, означает, что чем больше мощность накачки, тем легче детектировать третью гармонику. Однако металлические стенки не выдерживают слишком сильных электрических полей, и эмиссия электронов в сильных полях нарушает вакуум. Для предотвращения этих нежелательных эффектов электромагнитные поля не должны превышать значений порядка 40 МВ/м, или 1300 ед. СГС. Отсюда получаем оценку для

амплитуды третьей гармоники, зависящую только от добротности: $E_{3\Omega} < Q \cdot 0,6 \cdot 10^{-22}$ ед. СГС.

Какая добротность требуется для обнаружения эффекта? Очевидно, что детектируемое поле по крайней мере должно превышать уровень тепловых флуктуаций, который для типичного объема резонатора $V \sim 10^2$ см³ при температуре жидкого гелия составляет

$$E_{3\Omega\text{тепл}} \sim \sqrt{\frac{8\pi kT}{2V}} \sim 0,8 \cdot 10^{-8} \text{ ед. СГС.}$$

Для преодоления теплового порога потребовалась бы добротность $Q > 10^{14}$, вряд ли достижимая даже в сверхпроводящих резонаторах и накладывающая непомерно высокие требования на стабильность СВЧ-генератора.

Генерация третьей гармоники в квазиоптическом резонаторе

Чтобы обойти ограничения закрытого резонатора, можно воспользоваться открытым квазиоптическим резонатором, в котором фокусировка пучка позволяет достичь более сильных полей во внутренней части резонатора, сохраняя поля достаточно слабыми на его стенках. Третья гармоника будет возбуждаться в меньшем измерительном резонаторе, расположенном под углом к основному резонатору (рис. 9).

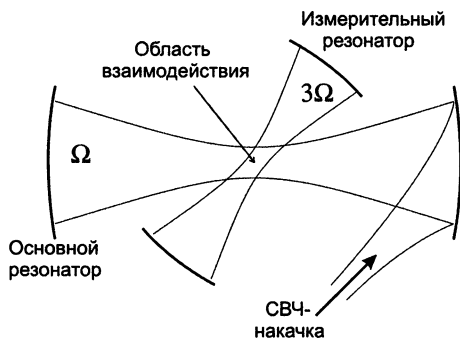


Рис. 9. Расположение квазиоптических резонаторов для генерации и детектирования третьей гармоники

Такое расположение резонаторов удобно для детектирования нелинейного эффекта, поскольку мощное СВЧ-излучение не попадает на охлаждаемые стенки высокодобротного измерительного резонатора, а детектор поля третьей гармоники можно расположить

в области, где поле основного пучка экспоненциально спадает. Однако на этот раз взаимодействие основной моды и третьей гармоники происходит в объеме, значительно меньшем объема резонатора.

Полагая пучки гауссовыми и пренебрегая их расходимостью, можно получить аналитическую оценку амплитуды насыщения третьей гармоники. Для перпендикулярно расположенных резонаторов имеем

$$E_{3\Omega} \sim \frac{\sqrt{\pi}}{8} \frac{Q a_3}{L_3} \frac{a_1^2}{a_3^2} \xi E_{\Omega}^3 \exp\left(\frac{-4a_1^2 k_1^2}{3}\right),$$

где Q и L_3 — добротность и длина измерительного резонатора; a_1 и a_3 — радиусы перетяжек квазиоптических пучков. Уменьшая угол между резонаторами, можно улучшить оценку не более чем в несколько раз, в основном за счет приближения экспоненциального фактора к единице. Добротность входит в последнее выражение в комбинации с длиной резонатора, то есть результат фактически зависит от свойств материала, из которого изготовлены квазиоптические зеркала. Экспоненциальный фактор отвечает за нарушение синхронизма между модами резонаторов и при желании он может быть сделан порядка единицы.

Для разумных (метровых) размеров зеркал квазиоптического резонатора на длине волны ~ 1 см поле в области взаимодействия ограничено величиной $E_{\Omega} \sim 4 \cdot 10^4$ ед. СГС. При $a_1^2/a_3^2 \sim 10$ и $L_3/a_3 \sim 10^2$ получаем $E_{3\Omega} \sim Q \cdot 1,6 \cdot 10^{-19} \exp(\dots)$ ед. СГС. Для наблюдения эффекта это поле должно превышать уровень тепловых флуктуаций амплитуды той же моды измерительного резонатора. При температуре $T \sim 4$ °К тепловое поле имеет порядок

$$E_{3\Omega\text{тепл}} \sim \sqrt{\frac{8\pi\kappa T}{2L_3 a_3^2}} \sim 3 \cdot 10^{-8} \text{ ед. СГС.}$$

Для достижения таких полей нужна добротность $Q > 10^{11}$, которая в принципе может быть реализована в сверхпроводящих резонаторах, но по-прежнему накладывает непомерно высокие требования на монохроматичность СВЧ-генератора.

Смещение двух СВЧ-волн с оптической волной

Другим способом обнаружения электродинамической нелинейности вакуума могло бы быть наблюдение когерентного рассеяния лазерного пучка на сильном СВЧ-поле. Пусть через область максимального поля квазиоптического СВЧ-резонатора под углом θ

к его оси проходит гауссов лазерный пучок частоты ω и радиуса a_p (рис. 10). В области взаимодействия индуцируются нелинейные поляризация и намагнитченность, которые, излучая, приводят к рассеянию лазерного пучка на СВЧ-поле с соответствующим изменением частоты ($\pm 2\Omega$) и направления (ϕ).

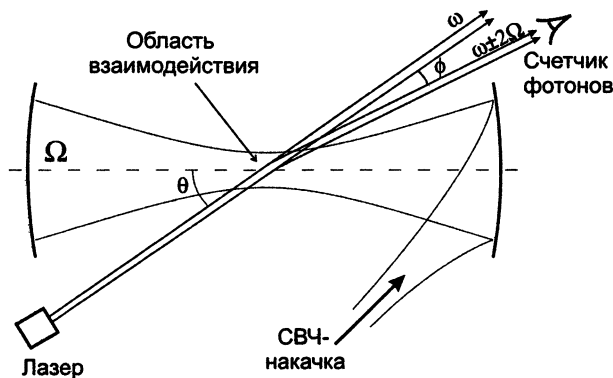


Рис. 10. Рассеяние лазерного пучка на СВЧ-волне

Для того чтобы этот процесс был эффективным, требуется соблюдение условия синхронизма $|\mathbf{k} + 2\mathbf{k}_{\text{СВЧ}}| = |\mathbf{k}| + 2|\mathbf{k}_{\text{СВЧ}}|$, которое, однако, может быть строго выполнено только при распространении лазерного пучка вдоль оси резонатора, когда нелинейная поляризация обращается в ноль. Баланс этих двух эффектов определяет оптимальные значения угла $\theta \approx \sqrt{2/(k_{\text{СВЧ}}a_{\text{СВЧ}})}$, где $k_{\text{СВЧ}}$ и $a_{\text{СВЧ}}$ — волновое число и радиус перетяжки СВЧ-пучка. Максимальная рассеянная мощность P_s оказывается пропорциональной мощности оптической накачки P :

$$P_s \approx \left[\frac{\sqrt{\pi}}{16} \frac{k}{k_{\text{СВЧ}}} \xi E_{\Omega}^2 k_{\text{СВЧ}} a_{\text{СВЧ}} \frac{\exp(-k_{\text{СВЧ}} a_{\text{СВЧ}}/2)}{\sqrt{2k_{\text{СВЧ}} a_{\text{СВЧ}}}} \right]^2 P.$$

Для $k/k_{\text{СВЧ}} \sim 10^4$, $k_{\text{СВЧ}} a_{\text{СВЧ}} \approx \pi$, $E_{\Omega} \sim 4 \cdot 10^4$ ед. СГС получаем

$$P_s \sim 1,5 \cdot 10^{-38} P.$$

Пучок рассеянных фотонов по своему оптическому качеству подобен лазерному пучку накачки и отклоняется от него на малый угол $\phi \sim k_{\text{СВЧ}}/k$. Энергия одного оптического фотона $\hbar\omega \sim 2 \cdot 10^{-19}$ Дж. Соответственно, для получения одного рассеянного фотона нужно затратить $\sim 10^{19}$ Дж энергии накачки. Практически эта энергия недостижима, так что в подобном лабораторном эксперименте не будет рассеян ни один фотон.

Смещение двух оптических волн с СВЧ-волной

Существенным недостатком рассмотренной в предыдущем разделе схемы являлась невозможность одновременно выполнить условия синхронизма и эффективного нелинейного взаимодействия волн. Преодолеть это ограничение можно при взаимодействии двух пересекающихся лазерных пучков с СВЧ-полем (рис. 11). Теперь условие синхронизма $c|\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_{\text{СВЧ}}| = 2\omega - \Omega \equiv c(|\mathbf{k}_1| + |\mathbf{k}_2| - |\mathbf{k}_{\text{СВЧ}}|)$ можно выполнить, если направить волновые вектора так, как изображено на диаграмме рис. 12.

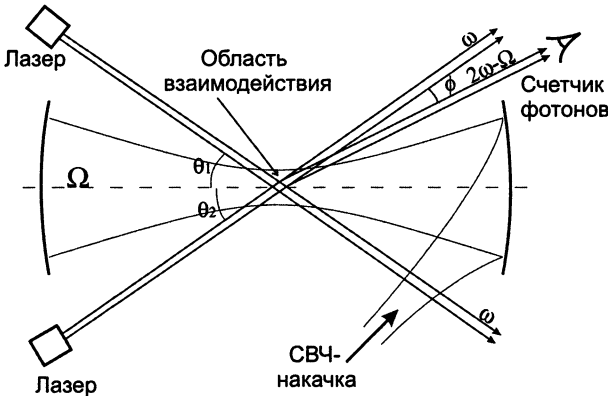


Рис. 11. Взаимодействие двух оптических волн с СВЧ-полем

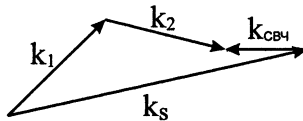


Рис. 12. Диаграмма волновых векторов для схемы рис. 11

В общем случае $\phi \sim \theta_1 \sim \theta_2$ и угол между векторами \mathbf{k}_s и $\mathbf{k}_{\text{СВЧ}}$ может быть произвольным. В симметричной модельной задаче с $\mathbf{k}_s \uparrow \downarrow \mathbf{k}_{\text{СВЧ}}$, когда $\theta_1 = \theta_2 = \theta$, условие синхронизма принимает вид $2k \cos \theta + k_{\text{СВЧ}} = 2k - k_{\text{СВЧ}}$ и угол θ однозначно определяется: $\theta \approx \sqrt{2k_{\text{СВЧ}}/k}$. В этом случае мощность P_s , рассеянная на частоте $2\omega - \Omega$, равна

$$P_s \approx \pi (\xi E_\Omega E_\omega)^2 (k_{\text{СВЧ}} a_{\text{опт}})^2 \left(\frac{k}{k_{\text{СВЧ}}} \right)^3 P \approx 4\pi (\xi E_\Omega)^2 \frac{k^3}{ck_{\text{СВЧ}}} P^2,$$

где $a_{\text{опт}}$ — радиус перетяжки лазерных пучков (в области их пересечения). Отметим, что особой монохроматичности СВЧ- и лазерных источников в этой схеме не требуется.

Чтобы за время лазерного импульса ~ 1 нс рассеялся хотя бы один фотон, необходимо выполнение условия $P_s > P_{s\text{min}} \sim 2 \cdot 10^{-10}$ Вт. Отсюда при тех же параметрах пучков, что и в предыдущей схеме ($k/k_{\text{СВЧ}} \sim 10^4$, $2\pi/k \sim 1$ мкм, $E_\Omega \sim 4 \cdot 10^4$ ед. СГС), получаем следующее требование к мощности накачки:

$$P > \left[\frac{P_{s\text{min}}}{4\pi(\xi E_\Omega)^2} \frac{ck_{\text{СВЧ}}}{k^3} \right]^{1/2} \sim 3 \cdot 10^{16} \text{ Вт.}$$

Выигрыш по сравнению с предыдущей нерезонансной схемой складывается из численного фактора $\sim 10^4$ и двух больших факторов ($k/k_{\text{СВЧ}}$) $\sim 10^4$ и $(E_\omega/E_\Omega)^2$. Однако требуемая энергия в импульсе наносекундного лазера $\sim 3 \cdot 10^7$ Дж, отвечающая полю $E_\omega \gtrsim 3 \cdot 10^7$ ед. СГС, все еще примерно на три порядка превосходит достижимый сейчас уровень.

В заключение, тем не менее, необходимо напомнить о стремительном прогрессе лазерной оптики и микроволновой электроники за последние два десятилетия, который позволяет надеяться, что нелинейность вакуума будет экспериментально обнаружена еще при нашей жизни.

Данные исследования проводились при частичной финансовой поддержке гранта 1744.2003.2 Совета по поддержке ведущих научных школ России, а также фонда «Династия» и МЦФФМ.

Литература

1. Боголюбов Н. И., Ширков Д. В. Квантовые поля. М.: Наука, 1993. 333 с.
2. Ициксон К., Зюбер Ж.-Б. Квантовая теория поля. Т. 1. М.: Мир, 1984. 448 с.
3. Берестецкий В. Б., Лифшиц Е. М., Питаевский Л. П. Квантовая электродинамика. М.: Наука, 1989. 724 с.
4. Ахиезер А. И., Берестецкий В. Б. Квантовая электродинамика. М.: Наука, 1981. 432 с.
5. Greiner W., Müller B., Rafelski J. Quantum electrodynamics of strong fields. Springer-Verlag, Berlin, 1985. 585 p.
6. Гриб А. А., Мамаев С. Г., Мостепененко С. Г. Вакуумные квантовые эффекты в сильных полях. М.: Энергоатомиздат, 1988. 287 с.
7. Никитшев А. И., Ритус С. И. Квантовая электродинамика явлений в интенсивном поле // Труды ФИАН, 1979, 111; Проблемы квантовой электродинамики интенсивного поля // Труды ФИАН, 1986, 168.

8. *Ding Y. J., Kaplan A. E.* // Phys. Rev. Lett. 1989. V. 63, N 25. P. 2725.
9. *Розанов Н. Н.* // ЖЭТФ. 1993. Т. 103, вып. 6. С. 1996; Proc. SPIE. 1993. V. 2097. P. 252.
10. *Brodin G., Marklund M., Stenflo L.* // Phys. Rev. Lett. 2001. V. 87, N 17. 171801.
11. *Alkofer R., Hecht M. B., Roberts C. D., Schmidt S. M., Vinnik D. V.* // Phys. Rev. Lett. 2001. V. 87, N 19. 193902.
12. *Heisenberg W., Euler H.* // Z. Phys. 1936. V. 98. P. 718.
13. *Mitter H.* Quantum electrodynamics in laser fields // Acta Phys. Austr. Suppl. 1975. V. 22. P. 1040.
14. *Розанов Н. Н.* // ЖЭТФ. 1998. Т. 113, вып. 2. С. 513.
15. *Павлов Г. Г., Гнедин Ю. Н.* Поляризация вакуума магнитным полем и ее астрофизические проявления // Пекулярные звезды, сильные магнитные поля и γ -астрономия (Астрономия, 22). М.: ВИНТИ, 1983. С. 172.
16. *Железняков В. В.* Излучение в астрофизической плазме. М.: Янус-К, 1997. 528 с.
17. *Шабад А. Е.* Поляризация вакуума и квантового релятивистского газа во внешнем поле // Труды ФИАН. 1988. Т. 192. С. 5.
18. *Shabad A. E.* // Proc. Workshop on strong magnetic fields and neutron stars. Havana, 2003; hep-th/0307214.
19. *Heyl J. S., Shaviv N. J.* // MNRAS. 2000. V. 311. P. 555.
20. *Александров Е. Б., Ансельм А. А., Москалев А. И.* // ЖЭТФ. 1985. Т. 89, вып. 4. С. 1181.
21. *Derishev E. V., Kocharovsky V. V., Kocharovsky Vl. V.* // Astron. and Astrophys. Trans. 2000. V. 19. P. 485.
22. *Derishev E. V., Kocharovsky V. V., Kocharovsky Vl. V.* // Radiophys. Quantum Electr. 1997. V. 40. P. 93.
23. *Baring M. G., Harding A. K.* // Ap. J. 2001. V. 547. P. 929.
24. *Soljasic M., Segev M.* // Phys. Rev. A. 2000. V. 62. 043817.
25. *Железняков В. В., Фабрикант А. Л.* // ЖЭТФ. 1982. Т. 82, вып. 5. С. 1366.
26. *Фабрикант А. Л.* // Письма в ЖЭТФ. 1982. Т. 36, вып. 3. С. 56.

ДИНАМИЧЕСКАЯ ПЕРКОЛЯЦИЯ

Д. И. Иудин, В. Ю. Трахтенгерц

В 1957 году математики Бломберген и Хаммерсли опубликовали статью [1], в которой поделились с читателями идеей вероятностной формализации процесса просачивания воды в перколяторе электрической кофеварки. Их построение было альтернативой классическому диффузионному описанию распространения одной физической субстанции, условно называемой *жидкостью*, в другой, условно называемой *средой*. Так возникла новая теория, названная впоследствии теорией перколяции или протекания (от английского глагола to percolate — просачиваться, протекать). Если первоначально перколяция была предложена как наглядная модель неупорядоченной среды, то сегодня мы являемся свидетелями проникновения перколяционных идей в сферу нелинейной динамики. Отсюда и название нашей статьи — динамическая перколяция. Но начнем с классических постановок статической перколяции.

1. Элементы статической перколяции

Простейшие решеточные задачи теории протекания формулируются следующим образом [2, 13]. Рассмотрим квадратную решетку, каждый узел которой занят с вероятностью p или пуст с вероятностью $1 - p$. Вероятность p можно интерпретировать как долю (концентрацию) занятых узлов при случайно-однородном заполнении решетки. Пустые и занятые узлы могут отвечать самым разнообразным физическим свойствам. Пусть для наглядности занятые (серые) квадратики на рис. 1 являются проводниками, а пустые (черные) — изоляторами. Предположим, что электрический ток может течь только между ближайшими проводящими соседями (на квадратной решетке каждый узел имеет четырех ближайших соседей: см. рис. 1). Очевидно, что при малых концентрациях $p \ll 1$ проводящие квадратики либо изолированы друг от друга, либо формируют небольшие группы — кластеры ближайших соседей. Два проводящих узла принадлежат одному кластеру только в том случае, когда они связаны проводящей цепочкой ближайших соседей. Максимальный кластер на рис. 1 изображен белыми пикселями. При $p \ll 1$ система в целом демонстрирует диэлектрические свойства, поскольку не существует длинных проводящих цепей, соединяющих противоположные стороны нашей решетки. Наоборот, при $p \sim 1$ проводящей

оказывается подавляющая доля пространства решетки, белые узлы формируют гигантский проводящий кластер, обеспечивающий прохождение электрического тока с одной стороны системы на другую (рис. 2). Серый цвет соответствует проводящим пикселям, изолированным от максимального кластера.

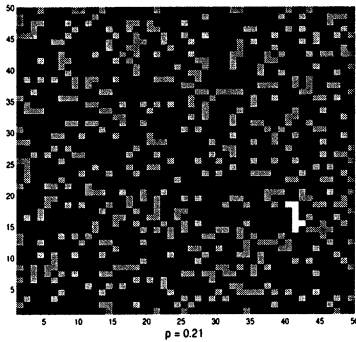


Рис. 1. Задача узлов на квадратной решетке. Серые ячейки квадратной решетки соответствуют проводящим пикселям, черные — непроводящим, белые — максимальному проводящему кластеру. Концентрация проводящих пикселей $p = 0,21$.



Рис. 2. Задача узлов на квадратной решетке. Концентрация проводящих пикселей $p = 0,73$. Серые ячейки квадратной решетки соответствуют проводящим пикселям, изолированным от максимального (белого) кластера.

Существует некоторое критическое значение $p = p_c$, при котором происходит переход от диэлектрического состояния (при $p < p_c$) к проводящему при ($p > p_c$). Критическое значение $p = p_c$ называется порогом протекания (перколяции). Для задачи узлов на квадратной решетке порог протекания приблизительно равен $p \approx 0,59$. Близкая к порогу ситуация отображена на рис. 3.

В противоположность обычным фазовым превращениям, где смена фаз происходит при некоторой критической температуре, описанный выше перколяционный переход является геометрическим фазовым превращением. Перколяционный переход характеризуется геометрическими свойствами проводящих кластеров вблизи $p = p_c$. При $p \ll 1$ существуют только кластеры небольших размеров. По мере роста концентрации p средний размер кластеров увеличивается. При концентрации близкой к критической появляется кластер, связывающий противоположные стороны решетки. Такой кластер

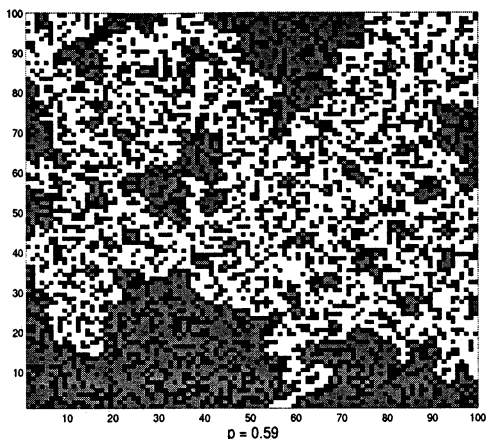


Рис. 3. Задача узлов на квадратной решетке. Концентрация проводящих пикселей $p = 0,59$.

называют перколяционным. В термодинамическом пределе бесконечно протяженной решетки перколяционный кластер называют бесконечным кластером (БК). С дальнейшим ростом концентрации доля принадлежащих перколяционному кластеру узлов (плотность перколяционного кластера) возрастает. Соответственно, средний размер конечных кластеров, не принадлежащих перколяционному кластеру, уменьшается, и при $p \sim 1$ все узлы принадлежат, очевидно, только одному кластеру.

Выше мы рассмотрели формулировку задачи узлов (site percolation), когда случайно занятыми или пустыми оказывались узлы квадратной решетки. Если случайно выбирать состояние ребер решетки (например, считать ребра проводящими или непроводящими), мы придем к формулировке задачи связей (bond percolation). Две занятых (проводящих) связи принадлежат одному кластеру только в том случае, когда они связаны цепочкой занятых (проводящих) связей. Критическая концентрация связей разделяет фазу конечных кластеров от фазы существования бесконечного кластера на плоской решетке.

1.1. Протекание как критическое явление. Аномалии коэффициентов переноса

Как было уже упомянуто, появление в системе БК связанных элементов можно рассматривать как геометрический фазовый пе-

реход: критическая концентрация $p = p_c$ отделяет фазу конечных кластеров $p < p_c$ от фазы $p > p_c$, где уже существует БК. Перколяционное превращение во многом напоминает обычные термодинамические фазовые переходы 2-го рода. В протекании концентрация p занятых узлов играет роль температуры в термических фазовых переходах. Как и в термических фазовых переходах длинные корреляционные связи контролируют перколяционное превращение, и поведение значимых величин вблизи перехода описывается степенными законами и критическими индексами.

Перколяционный переход характеризуется геометрическими свойствами кластеров вблизи $p = p_c$. Важной величиной является вероятность P_∞ того, что узел (или связь) принадлежит бесконечному кластеру. Когда $p < p_c$, существуют только конечные кластеры и $P_\infty = 0$. При $p > p_c$ P_∞ увеличивается с p по степенному закону:

$$P_\infty \sim (p - p_c)^\beta. \quad (1)$$

P_∞ описывает порядок в перколяционной системе и может быть определена как параметр порядка. Линейные размеры конечных кластеров ниже и выше p_c характеризуются корреляционной длиной ξ . Корреляционная длина определена как среднее расстояние между узлами одного и того же конечного кластера. С приближением к p_c ξ возрастает как

$$\xi \sim |p - p_c|^{-\nu} \quad (2)$$

с одинаковым показателем ν ниже и выше порога. Среднее число узлов (масса) конечного кластера тоже расходится:

$$S \sim |p - p_c|^{-\gamma}, \quad (3)$$

опять же с одинаковым показателем γ ниже и выше порога. Для получения S и ξ требуется вычисление средних по всем конечным кластерам на решетке.

Показатели β , ν и γ описывают критическое поведение типичных величин, связанных с перколяционным переходом, и называются критическими индексами. Индексы эти универсальны и не зависят ни от структурных особенностей решетки (квадратная или треугольная и т. д.), ни от типа протекания (по узлам, по связям или континуальное), но только от размерности пространства. Это свойство универсальности является характерной чертой фазовых переходов 2-го рода, где параметр порядка непрерывно исчезает в критической точке. Например, спонтанная намагниченность всех трехмерных магнитных материалов описывается одним и тем же индексом

β независимо от кристаллической структуры и типа взаимодействия между соседними спинами. Заметим, что все введенные выше величины определены в термодинамическом пределе бесконечно больших систем. В ограниченной системе P_∞ , например, не обязательно равно нулю ниже порога.

1.2. Структурные свойства перколяционного кластера

Начнем с рассмотрения БК на пороге протекания $p = p_c$. Типичный пример двумерного БК показан на рис. 3. На рисунке видно, что БК содержит дыры на всех масштабах, подобно салфетке Серпинского. Кластер самоподобен на всех линейных масштабах, больших периода решетки и меньших размера системы, и может рассматриваться как фрактал. Фрактальная размерность d_f описывает, как в среднем меняется масса кластера M (число его элементов) внутри евклидовой сферы с ростом ее радиуса r :

$$M(r) \sim r^{d_f}. \quad (4)$$

Мы знаем, что для случайных фракталов $M(r)$ представляет результат усреднения по многим различным конфигурациям БК или, что то же самое, по многим сферам с различными центрами на одном и том же БК. Выше и ниже порога средние размеры конечных кластеров описываются корреляционной длиной ξ . На пороге $p = p_c$ ξ расходится, и дыры в БК обнаруживаются на всех масштабах. Выше порога ξ представляет средний характерный размер дыр в БК. Поскольку ξ конечна выше порога, то самоподобие БК простирается только на масштабы меньшие ξ . Мы можем интерпретировать $\xi(p)$ как типичный масштаб самоподобия, в пределах которого БК демонстрирует фрактальные свойства. На масштабах больших ξ структура БК утрачивает самоподобие и становится однородной. На рис. 3 кроссовер от фрактального поведения на малых масштабах к однородному на больших r показан на примере регулярно паркета, составленного из ковриков Серпинского с размерами ξ . Итак, мы видим фрактальный кластер на линейных масштабах меньше чем ξ . На масштабах, превышающих ξ , мы обнаруживаем однородную структуру, которая набрана из многочисленных ячеек размера ξ . Математически это может быть записано в виде

$$M(r) \sim \begin{cases} r^{d_f}, & r \ll \xi, \\ r^d, & r \gg \xi. \end{cases} \quad (5)$$

Свяжем фрактальную размерность d_f перколяционного кластера с критическими индексами β и ν . Вероятность того, что некоторый выбранный наугад узел внутри евклидовой сферы радиуса r принадлежит БК, определяется отношением числа узлов, принадлежащих БК, к полному числу узлов в сфере:

$$P_\infty \sim \frac{r^{d_f}}{r^d}, r < \xi. \quad (6)$$

Последнее соотношение корректно для $r = \lambda\xi$, где λ — некоторая константа меньше 1. Подставляя $r = \lambda\xi$ в (6), получаем

$$P_\infty \sim \frac{\xi^{d_f}}{\xi^d}. \quad (7)$$

Здесь обе части соотношения являются степенными функциями от ξ . Подставляя (1) и (3) в (7), находим:

$$d_f = d - \frac{\beta}{\nu}. \quad (8)$$

Таким образом, фрактальная размерность d_f перколяционного кластера на пороге протекания не является новым независимым показателем, но выражается через термодинамические показатели β и ν . Поскольку β и ν являются универсальными индексами, d_f также является универсальной величиной, зависящей только от размерности пространства. Можно показать, что ниже порога и на пороге формула (8) определяет и фрактальную размерность конечных кластеров с линейными размерами, не превышающими ξ . Ниже порога встречаются, хотя и редко, кластеры с размерами, превышающими ξ . Такие кластеры называют решеточными зверями (lattice animals), их фрактальная размерность меньше чем d_f .

Благодаря фрактальности перколяционных структур, находящихся вблизи порога, их динамические характеристики демонстрируют аномальное поведение. Возьмем, например, транспортные свойства перколирующих узлов, предполагая, что занятые узлы являются проводниками, а пустые — изоляторами, и что ток электрических зарядов может проходить только между проводящими узлами, являющимися ближайшими соседями. В этом случае удельная проводимость уже не является константой, а зависит от размеров системы. Аналогично можно рассмотреть поведение упругих колебаний перколирующих систем в контексте задачи связей, предполагая, что занятые связи могут быть представлены пружинами с конечной жесткостью. Если все связи решетки заняты, то при возбуждении

мы обнаружим обычные фононы. Когда система «разбавлена», т. е. случайно выбранная часть пружин выброшена из решетки, мы обнаружим локализованные моды на больших частотах. Эти моды были введены Александром и Орбахом [3] и названы ими фрактонами. Плотность состояний фрактонов как функция частоты демонстрирует аномальное поведение.

Тот факт, что перколяционный переход является критическим явлением, подразумевает степенной характер поведения связанных с ним физических величин: они меняются как степени параметра $|p - p_c|$. Это утверждение сохраняет свою силу и для динамических свойств перколирующих объектов. Первое экспериментальное свидетельство степенного поведения экстенсивной динамической характеристики перколирующей системы было дано в 1971 году Ластом и Таулесом [4] при исследовании проводимости разбавленной двумерной системы. Анализ полученных ими экспериментальных данных показал, что выше порога $p_c \simeq 0,6$, в критической зоне проводимость σ ведет себя как

$$\sigma \sim (p - p_c)^\mu \quad (9)$$

с показателем μ больше, чем 1.

2. Направленная перколяция

В разделе 1 мы исследовали появление протекания при увеличении концентрации p структурных элементов и определили перколяционный переход как появление бесконечного кластера занятых узлов. Рассмотрим теперь динамическую модель протекания. Обратимся к задаче узлов. Предположим, что если в данный момент времени узел находится в возбужденном состоянии, то в следующий (дискретный) момент времени он возвращается в исходное состояние покоя, но может при этом с вероятностью p возбудить своих ближайших соседей по решетке. В результате разыгрывается стохастический процесс, напоминающий распространение инфекции.

2.1. Распространение инфекций

Ясно, что когда вероятность передачи возбуждения очень мала, число возбужденных элементов будет убывать со временем. В обратном пределе, когда вероятность передачи возбуждения близка к единице, число возбужденных элементов решетки неограниченно возрастает во времени. Критическое значение вероятности передачи возбуждения p_c , при котором появляется конечная вероятность

роста эпидемии при единственном инфицированном в начальный момент, соответствует перколяционному порогу. Траектория распространения инфекции представляет собой перколяционный кластер. Рассматривая наикратчайший путь между двумя его точками A и B , можно получить новую существенную информацию об устройстве фрактального кластера. Элементы кластера, принадлежащие этому пути, формируют самоподобную структуру с фрактальной размерностью d_{\min} . Таким образом, длина l этого наикратчайшего пути, называемая часто химическим расстоянием, меняется с увеличением евклидовой дистанции между A и B по закону

$$l \sim r^{d_{\min}}. \quad (10)$$

Характерное расстояние, на которое фронт развивающейся инфекции успевает уйти за время t , определяется именно химической размерностью перколяционного кластера.

С размерностями d_f и d_{\min} тесно связана так называемая химическая размерность d_l , которая описывает изменение числа элементов кластера M , удаленных от некоторого фиксированного элемента на химическое расстояние, не превышающее дистанцию l , с увеличением этой дистанции:

$$M(l) \sim l^{d_l}. \quad (11)$$

Если фрактальная размерность d_f характеризует рост массы кластера с увеличением его размера в евклидовом пространстве, то химическая размерность d_l описывает рост массы кластера в пространстве химических расстояний. Комбинируя (4), (10) и (11), находим связь:

$$d_l = \frac{d_f}{d_{\min}}. \quad (12)$$

Как и для других задач теории перколяции величина порога направленного протекания существенно зависит от числа ближайших соседей для заданной решетки, т. е. от координационного числа z . Чем большее число элементов находится в контакте с зараженным, тем выше вероятность размножения инфекции. Координационное число возрастает с увеличением размерности решетки. В пределе высоких координационных чисел для расчета порога протекания может быть использовано приближение среднего поля. Пусть в фиксированный момент времени t доля элементов среды, находящихся в возбужденном состоянии, составляет q_t . Если число ближайших соседей z у каждого элемента среды велико, можно считать, что среди его соседей в момент времени t имелось zq_t возбужденных элементов.

Выбранный нами элемент может быть инфицирован в следующий $(t + 1)$ -й момент времени с вероятностью

$$1 - (1 - p)^{zq_t}, \quad (13)$$

поскольку мы предполагаем независимость передачи инфекции через ближайших соседей. Величина (13) дает долю элементов среды, находящихся в возбужденном состоянии в следующий $(t + 1)$ -й момент времени, т. е.

$$q_{t+1} = 1 - (1 - p)^{zq_t}. \quad (14)$$

Выше порога перколяции величина q_t должна стремиться к отличному от нуля конечному пределу $\lim_{t \rightarrow \infty} q_t = \tilde{q}$, который дает долю возбужденных элементов в установившемся режиме, тогда как ниже порога перколяции $\lim_{t \rightarrow \infty} q_t = 0$.

Исследуя точечное отображение (14), нетрудно показать, что его аттракторами могут быть только стационарные точки, причем потеря устойчивости стационарной точки $q = 0$ и рождение новой устойчивой стационарной точки $q = \tilde{q}$ происходит при прохождении вероятности p через значение p_c , удовлетворяющее условию

$$1 - p_c = \lim_{\tilde{q} \rightarrow 0} (1 - \tilde{q})^{\frac{1}{z\tilde{q}}} = \exp\left(-\frac{1}{z}\right). \quad (15)$$

Таким образом, порог перколяции определяется соотношением

$$p_c = 1 - \exp\left(-\frac{1}{z}\right) = \frac{1}{z} - \frac{1}{2z^2} + \dots \quad (16)$$

При уменьшении координационного числа приближение среднего поля теряет применимость. В этом случае становятся существенными флуктуационные эффекты.

Важное значение геометрических флуктуаций, связанных с направленной перколяцией, очень выпукло проявляется в задачах нелинейной динамики. Рассмотрим флуктуационные эффекты на примере распределенной системы со случайным размножением и распадом.

2.2. Перколяция в средах с размножением, распадом и диффузией

Сосредоточенная система со случайным размножением и распадом описывается стохастическим дифференциальным уравнением

$$\dot{n} = -\alpha n + f(t)n, \quad (17)$$

где α — постоянная скорость распада, а $f(t)$ — случайно меняющаяся во времени скорость размножения с заданными статистическими характеристиками. Говорят, что для системы (17) превышен порог взрывной неустойчивости, если средняя по ансамблю реализаций плотность вещества $\langle n(t) \rangle$ неограниченно возрастает со временем.

Добавление в уравнение (17) диффузионного члена обобщает модель на случай распределенных систем, в которых возможны процессы распада, размножения и диффузии некоторого вещества. Предполагается, что скорость распада однородна в пространстве и во времени, а размножение происходит лишь внутри определенных центров размножения, которые случайно возникают во времени в случайных точках среды, но имеют одинаковую форму, интенсивность и продолжительность жизни. Соответствующей математической моделью является уравнение

$$\dot{n} = -\alpha n + f(\mathbf{r}, t)n + D\Delta n, \quad (18)$$

где n — плотность вещества, D — коэффициент его диффузии. Флуктуирующее поле $f(\mathbf{r}, t)$ задается суммой расположенных в случайных точках (\mathbf{r}_i, t_i) одинаковых импульсов $g(\mathbf{r}, t)$:

$$f(\mathbf{r}, t) = \sum_i g(\mathbf{r} - \mathbf{r}_i; t - t_i). \quad (19)$$

Среднее число импульсов, приходящихся в единицу времени на единицу объема, постоянно и равно m . Функция $g(\mathbf{r}, t)$ имеет вид

$$g(\mathbf{r}, t) = J\chi(\mathbf{r})\sigma(t), \quad (20)$$

где J характеризует интенсивность центра размножения, $\chi(\mathbf{r}) = 1$ при $r \leq r_0$ и $\chi(\mathbf{r}) = 0$ при $r > r_0$, $\sigma(t) = 1$ при $0 < t < \tau_0$ и $\sigma(t) = 0$ при $t < 0$ и $t > \tau_0$, так что r_0 и τ_0 дают, соответственно, характерный пространственный размер отдельного центра и время его жизни.

Существенным параметром задачи является безразмерная пространственно-временная концентрация центров размножения c , определяемая как

$$c = mr_0^d \tau_0, \quad (21)$$

где d — размерность среды. Когда параметр c мал, $c \ll 1$, различные центры размножения действуют независимо друг от друга

и расчет порога взрывной неустойчивости осуществляется в рамках стандартного анализа в приближении среднего поля. По мере роста с различные центры размножения начинают перекрываться в пространственно-временном континууме и их взаимное влияние становится существенным: последующие центры начинают действовать на фоне пятен населенности, оставленных предыдущими центрами. Появляются пространственно-временные цепочки центров размножения — кластеры размножающегося вещества, эволюционирующие на фоне практически нулевой населенности. Пространственно-временные масштабы этих кластеров расходятся при стремлении параметра c к некоторому критическому значению.

При численном исследовании уравнения (18) мы дополнили его слабой постоянной накачкой, поддерживающей населенность на уровне шума, и нелинейным диссипативным членом, обеспечивающим механизм запорогового ограничения роста населенности:

$$\dot{n} = \epsilon - \alpha n + f(\mathbf{r}, t)n + D\Delta n - \beta n^2. \quad (22)$$

Рисунок 4 воспроизводит эволюцию средней населенности для дискретного аналога уравнения (22) на квадратной решетке. Соответствующая динамика населенности в одной отдельно взятой ячейке решетки показана на рис. 5.

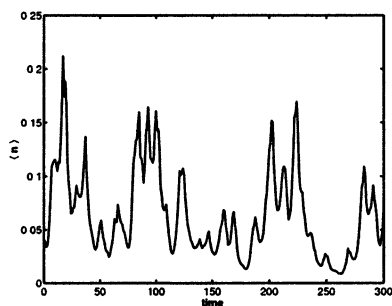


Рис. 4. Эволюция средней населенности на квадратной решетке 70×70 при $r_0 = 9$; $\tau_0 = 5$; $J = \alpha = 0, 3$; $\beta = 5 \cdot 10^{-2}$; $\epsilon = 3 \cdot 10^{-3}$ и $D = 10^{-2}$.

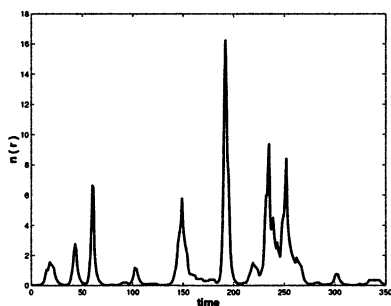


Рис. 5. Динамика населенности в одной отдельно взятой ячейке решетки.

Характерный вид пространственного распределения населенности в некоторый фиксированный момент времени представлен

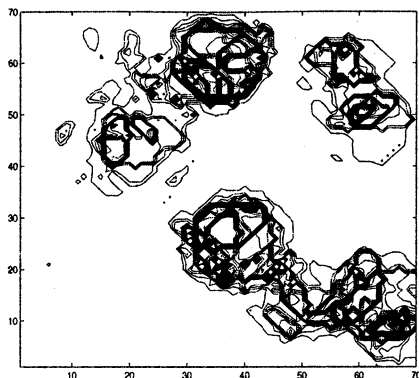


Рис. 6. Характерный вид пространственного распределения населенностей в фиксированный момент времени.

на рис. 6. Заметим, что среднее по пространству и времени значение параметра активности составляет величину $\langle f(\mathbf{r}, t) \rangle \approx 0,08$, которая значительно уступает коэффициенту линейной диссипации.

2.3. Эстафетная передача информации

Близкой дискретной иллюстрацией конструктивной роли геометрических флуктуаций может служить передача информации по одномерному информационному каналу. Рассмотрим одномерную цепочку клеточных автоматов, каждый из которых может либо блокировать, либо пропускать проходящую через него информацию. Цепочка характеризуется удельным числом p элементов, прозрачных для сигнала. Изображенная на рис. 7 цепочка автоматов длиной $L = 12$ соединяет расположенный слева источник информации А с приемником В. Черный цвет соответствует блокирующим ячейкам, серый — прозрачным, белый — прозрачным ячейкам, содержащим информацию. Пусть в каждый момент дискретного времени источник А подает на вход системы один бит информации. Этот бит может беспрепятственно перемещаться вправо только до ближайшего блокирующего узла. Очевидно, таким образом, что статическая цепь способна передавать информацию приемнику В только при значении $p = 1$.

Предположим, что элементы сети могут с течением времени изменять свое состояние, так что за один шаг дискретного времени t прозрачных элементов становятся непрозрачными, и наоборот,



Рис. 7. Одномерная цепочка клеточных автоматов

m непрозрачных элементов сети становятся прозрачными. При этом уровень прозрачности p остается постоянным, а доля обновляемых за один шаг автоматов составляет величину $q = 2m/L$. Исследуем возможность передачи информации по такой флуктуирующей цепи при условии, что информация способна сохраняться на прозрачных ячейках и пропадает только при изменении статуса ячейки. На каждом шаге модельного времени поступившая на вход системы информация передвигается вправо до ближайшего блокирующего узла. Предполагается, что информация передается через прозрачные ячейки за время, малое по сравнению с длиной шага модельного времени, так что бит информации всегда занимает крайнее правое из доступных ему положений. На каждом шаге модельного времени у бита информации, сохранившегося в прозрачной ячейке перед некоторым блокирующим узлом, существует три возможности:

1) он может продвинуться вперед с вероятностью $q(1-q)$, равной произведению вероятностей двух независимых событий: первое — изменение статуса соседней ячейки справа с непрозрачного на прозрачный (вероятность q) и второе — сохранение прозрачности самой ячейки, содержащей рассматриваемый бит информации (вероятность $1 - q$);

2) может остаться на месте с вероятностью $(1 - q)^2$;

3) может диссипировать с вероятностью q , где $q = 2m/L$ — вероятность смены статуса ячейки в единицу времени.

Для дальнейшего продвижения информации к потребителю необходимо выполнение следующих двух условий: во-первых, блокирующий узел, задерживающий движение информации, должен превратиться в прозрачный элемент и, во-вторых, прозрачный элемент, содержащий информацию, должен оставаться прозрачным вплоть до момента снятия блокировки. Очевидно, что при низком уровне прозрачности $p \ll 1$ вероятность одновременного выполнения этих условий чрезвычайно мала и информация не достигает потребителя. Напротив, при уровне прозрачности, близком к единице $p \simeq 1$, информация передается потребителю со скоростью, близкой к максимально возможной. Существует некоторое критическое значение уровня прозрачности $p = p_c$, при котором появляется отличная от

нуля вероятность прохождения информации через систему. При этом скорость передачи информации стремится к нулю вблизи порога.

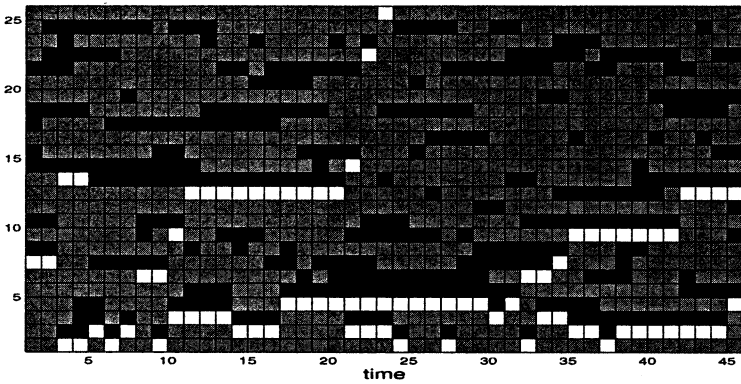


Рис. 8. Информация передается снизу вверх, время увеличивается слева направо.

Рисунок 8 демонстрирует эволюцию во времени цепочки автоматов длиной $L = 25$ при $q = 0,16$ для уровня прозрачности $p = 0,72$. Зависимость количества переданной информации от уровня прозрачности флуктуирующего канала при трех различных значениях параметра q представлена на рис. 9 и 10 в обычном и билогарифмическом масштабах.

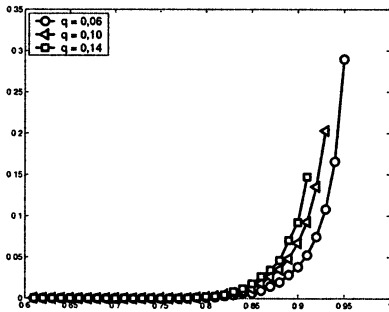


Рис. 9. Зависимость пропускной способности канала от уровня прозрачности p .

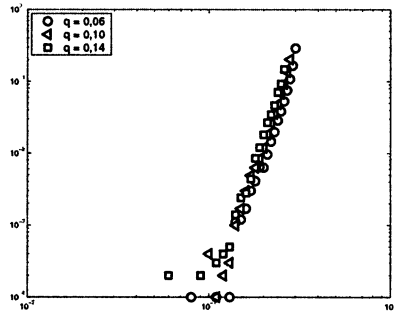


Рис. 10. То же, что слева, в билогарифмическом масштабе.

3. Динамическая перколяция

В предыдущем разделе мы рассмотрели несколько примеров, в которых геометрический, по сути, перколяционный переход приводил к нетривиальным эффектам нелинейной динамики. Вместе с тем, критический параметр, контролирующий близость системы к перколяционному переходу, оказывался там всего лишь инструментом внешней ручной настройки. В этом разделе мы обратимся к анализу ситуации, когда критический параметр перколяционного перехода сам становится внутренней динамической переменной системы.

3.1. Модель лесных пожаров

Мы начнем с простой, но содержательной модели, в которой роль перколирующих элементов играют деревья, растущие в узлах квадратной решетки. Критическим параметром геометрической части задачи здесь выступает, очевидно, плотность занятых деревьями узлов. Впервые модель лесных пожаров была предложена Баком, Ченом и Тангом в 1990 году [7] и позднее, в 1992 году, была усовершенствована Шваблом и Дросселем [8]. Мы рассмотрим именно этот последний вариант модели.

Итак, будем полагать, что динамика числа деревьев определяется следующими простыми правилами.

1. С вероятностью ϵ в единицу времени деревья появляются в узлах решетки. Мы, таким образом, не фиксируем рост деревьев в деталях, а предлагаем бинарную схему роста: с некоторого момента времени содержимое узла становится деревом, а до этого момента узел считается пустым.

2. С вероятностью θ в единицу времени молния ударяет в узел решетки. Если узел содержит дерево, то это дерево загорается. В противном случае состояние узла не меняется.

3. Если ближайшие соседи узла, содержащего загоревшееся дерево, содержат деревья, то с вероятностью единица эти деревья загораются на следующем шаге модельного времени. Ближайшими соседями узла мы считаем четырех его смежных соседей.

4. На следующем шаге модельного времени узел, содержащий загоревшееся дерево, становится пустым.

Перколяционные эффекты в этой задаче становятся существенными при выполнении двойного предела:

$$\epsilon \rightarrow 0, \quad \theta/\epsilon \rightarrow 0. \quad (23)$$

При этом система самостоятельно выходит на стационарный уровень заполнения деревьями $p \lesssim p_c$ с широким распределением кластеров по размерам. В отсутствие молний кластеры медленно растут и коагулируют. Бросая редкие молнии, мы выжигаем незначительную часть этих кластеров, практически не изменяя свойственного статической перколяции распределения кластеров по размерам. При выполнении соотношения (23) характерное время выгорания крупного кластера уступает характерной величине интервала между появлениями новых деревьев.

На верхней панели рис. 11 показана эволюция числа горящих деревьев, там же на нижней панели отражена зависимость от времени удельного числа растущих деревьев. Видно, что удельное число деревьев флуктуирует вблизи порогового уровня $p_c \simeq 0,58$, не превышая его. Этот критический уровень определяется равновесием между процессами роста деревьев и их уничтожения при возникновении пожаров.

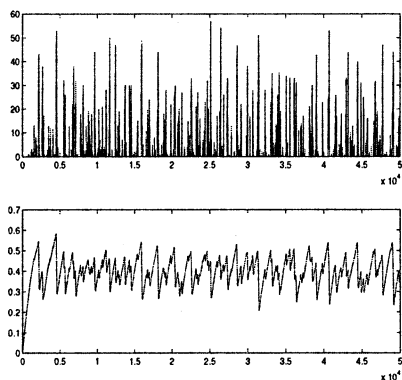


Рис. 11. Эволюция удельного числа деревьев.

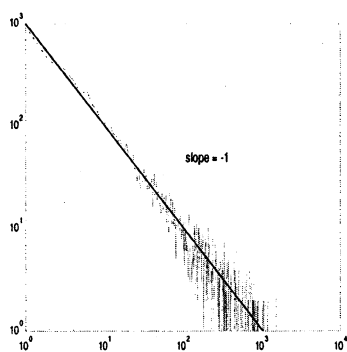


Рис. 12. Частота появления пожаров как функция их размеров в билогарифмическом масштабе.

На рис. 12 частота появления пожаров, охватывающих определенное число деревьев, представлена как функция этого числа в билогарифмическом масштабе. Очевиден степенной характер зависимости, указывающий на самоподобие пространственно-временной динамики пожаров – отсутствие в широком диапазоне значений характерных размеров кластеров и характерных частот их появления. Скейлинговый режим ограничен, с одной стороны, размерами систе-

мы и временем счета, с другой — размером кластера, состоящего из единственного дерева, и обратной частотой появления этих минимальных пожаров. Характерный интервал между частыми мелкими событиями составляет сотни шагов модельного времени.

Динамика рассмотренной модели примечательна в двух аспектах. Во-первых, она происходит в критическом режиме сильных флуктуаций: заранее не известно, в какой кластер, большой или маленький, угодит следующая молния. При этом сами кластеры имеют фрактальную перколяционную структуру и, следовательно, такой же структурой обладают паттерны происходящих пожаров. Другими словами, критический характер динамики — отсутствие характерных пространственно-временных масштабов — непосредственно связан здесь с фрактальностью реальных конфигураций системы. Во-вторых, мы имеем дело с самоорганизованной динамикой: при выполнении соотношений (23) система самостоятельно из исходного состояния с $p = 0$ выходит на уровень генерации фрактальных диссипативных структур $p \lesssim p_c$. Такое нетривиальное соединение критичности и самоорганизации нашло свое отражение в названии активно развивающейся сегодня парадигмы самоорганизованной критичности [9].

В следующем пункте мы рассмотрим физически более содержательный аналог модели лесных пожаров и увидим, что перколяция может обеспечивать нетривиальный флуктуационный механизм подавления роста неустойчивостей в активных средах.

3.2. Модели активного роста

Пусть теперь вместо деревьев в нашей модельной системе растет некоторое поле, например электрическое. Представим возбудимую среду в виде трехмерной решетки клеточных автоматов и поставим в соответствие каждому узлу решетки зависящую от времени скалярную величину u , характеризующую значение электрического потенциала [10]. В рассматриваемой модели клеточных автоматов основным управляющим параметром является отношение скорости роста ϵ перепадов потенциала в соседних ячейках, определяемой соотношением

$$\langle \Delta u^2 \rangle = \epsilon t, \quad (24)$$

где $\langle \Delta u^2 \rangle$ — среднеквадратичная разность потенциала между соседними ячейками, ϵ — скорости диссипации, принимаемой за единицу в нашей дискретной модели. Проведенные нами компьютерные эксперименты [11] показали, что при $\epsilon \gtrsim 1$ в рассматриваемой системе клеточных автоматов реализуются хорошо известные регулярные

режимы автоволн и релаксационные автоколебания. Интересующая нас критическая динамика обнаруживается в пределе $\epsilon \ll 1$. При этом происходит динамическая перколяция возбуждений на случайном потенциальном рельефе, аналогичная распространению пламени в модели лесных пожаров.

Эволюция потенциального рельефа на решетке осуществляется в рамках следующих моделей случайного роста.

А. Пространственный белый шум. В этом случае на каждом шаге модельного времени происходит распределение по узлам решетки нормально распределенных случайных добавок к величине скалярного поля (рис. 13, *а*). Период решетки отвечает оптимальному инкременту роста потенциального рельефа. Дисперсия распределения линейно увеличивается с течением времени. Каждый узел совершает независимое от соседей броуновское движение в пространстве значений скалярного потенциала. Подобное распределение потенциала можно наблюдать в плазмоподобной среде со случайной депозицией заряда при соблюдении локальной электронейтральности.

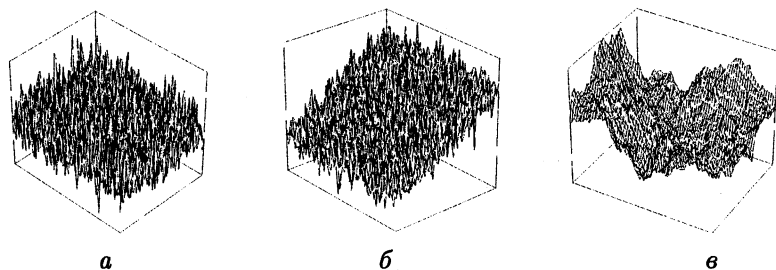


Рис. 13. Пример распределения потенциала на квадратной решетке 65×65 . В трехмерном моделировании это соответствует распределению потенциала в одном из ортогональных сечений простой кубической решетки $65 \times 65 \times 65$: *а* — пространственный белый шум; *б* — пространственный белый шум при наличии однородного поля; *в* — броуновский рельеф.

Б. Пространственный белый шум плюс регулярное линейное смещение. Здесь на каждом шаге модельного времени, помимо распределения по узлам решетки нормально распределенных случайных добавок, к величине скалярного потенциала прибавляется регулярное смещение, пропорциональное линейной комбинации координат ячеек. Такая ситуация отвечает наличию внешнего однородного электрического поля, линейно растущего во времени. Пример соответствующего распределения потенциала в одном из ортогональных сечений модельной решетки представлен на рис. 13, *б*.

В. Броуновский потенциальный рельеф. Рост потенциального рельефа может осуществляться в широком диапазоне инкрементов и иметь, следовательно, многомасштабный характер. При этом дисперсия распределения увеличивается не только во времени, но и в пространстве. Примером могла бы служить случайная депозиция заряда при соблюдении только глобальной электронейтральности. Потенциальный рельеф в этом случае имеет характер горного ландшафта и может быть промоделирован пространственным распределением, называемым обобщенным броуновским распределением. Для формирования обобщенного броуновского распределения потенциала на простой кубической решетке мы использовали трехмерный аналог алгоритма Фосса [12]. Генерируемый таким образом рельеф характеризуется пространственным ростом дисперсии распределения потенциала.

3.3. Перколяционный механизм диссипации

К модели роста потенциала мы должны добавить механизм, ограничивающий этот рост. В рамках развиваемого формализма подавление роста связывается с заселением среды диссипативными объектами и возникновением связей между ними. Естественно предположить, что дискриминация роста носит локальный характер. Пусть, например, развитие потенциального рельефа ограничено некоторым критическим значением разности потенциалов u_c между ближайшими узлами, по достижении которого происходит пробой — между соседними узлами возникает проводящая связь, выравнивающая соответствующую разность потенциалов. Предполагается далее, что некоторые участки потенциального рельефа метастабильны и на следующем шаге модельного времени возникший пробой может инициировать пробой соседних связей (инфицировать соседей), если приложенная к ним разность потенциалов превышает некоторый фиксированный уровень — уровень активации u_a , значение которого меньше критического. Важно подчеркнуть, что активация может иметь как скалярный, так и векторный характер. В первом случае принимается во внимание только абсолютное значение разности потенциалов, во втором — дополнительно учитывается знак. Мы начнем с рассмотрения скалярной активации и в заключение обсудим особенности векторного случая.

Итак, появление пары узлов с критической разностью потенциалов способно инициировать пробой между этими узлами и их ближайшими соседями по ортогональным направлениям (рис. 14). При этом разыгрывается стохастический процесс «металлизации» среды,

приводящий к выравниванию потенциального рельефа. Существенным, при этом, является вопрос о времени жизни возникающих проводящих связей. Если оно достаточно велико по сравнению с шагом модельного времени (с характерным временем передачи возбуждения), то динамика металлизации дополняется интереснейшими коллективными эффектами в ансамбле проводящих связей. В простейшем случае время жизни составляет некоторое фиксированное число шагов модельного времени и размер возникающих в процессе «металлизации» проводящих зон не превышает точно такого же числа периодов модельной решетки. Мы, однако, рассмотрим более интересную ситуацию, когда время жизни возникающих проводящих связей самосогласованно определяется их взаимной конфигурацией. Здесь возникают тонкие моменты, связанные с топологией образующихся проводящих зон, и мы остановимся на более подробном рассмотрении этого вопроса.

Заметим, прежде всего, что структура, возникающая в процессе «металлизации», представляет собой проводящий граф, вершины и ребра которого соответствуют узлам и связям проводящей структуры. Изолированные связные компоненты проводящих графов будем называть кластерами. Все вершины отдельного кластера можно разбить на три непересекающиеся группы: периферийные точки, т. е. вершины, образующие внешнюю границу кластера, промежуточные вершины, последовательно соединяющие смежные звенья отдельных ветвей, и точки соединения ветвей или точки ветвления — вершины общие для трех и более звеньев. Индексом ветвления вершины называется число приходящих в нее ребер, уменьшенное на единицу. Например, индекс ветвления периферийной вершины равен нулю, промежуточной вершины — единице и т. д. На рис. 15 вершины *A*, *B*, *E*, *G*, *K* являются периферийными, вершины *C*, *D*, *F* — точками ветвления с индексом ветвления 2.

Мы исследовали ситуацию, когда смена конфигурации кластера на каждом шаге моделирования целиком определяется состоянием его периферийных вершин. Такой подход дает уникальную возмож-

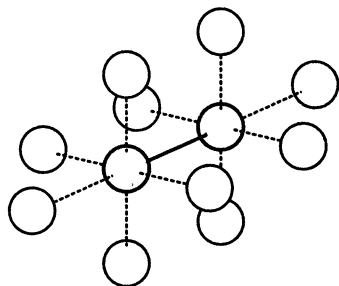


Рис. 14. Возможные направления инициации пробоев на простой кубической решетке. Серым цветом выделена пара с критической разностью потенциалов.

ность учета эффектов, связанных с наличием крупномасштабного поля, поскольку именно периферийные вершины являются точками приложения к кластеру внешней разности потенциалов, тем большей, чем сильнее поле и чем крупнее сам кластер.

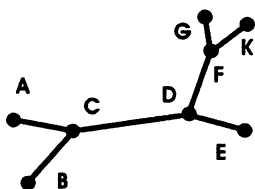


Рис. 15. Элементы кластера

Мы использовали следующий алгоритм анализа периферии. Если разность потенциалов между такой вершиной и хотя бы одним из ее ближайших соседей превышает уровень активации, то возникает одна или несколько новых проводящих связей и такое же количество новых периферийных вершин, опрос которых будет производиться уже на следующем шаге. Если же периферийная вершина не находит активированных соседей, то она отмирает вместе с соответствующим ей проводящим ребром, передавая роль периферии ближайшей своей предшественнице, если эта последняя не является точкой ветвления. В противном случае, когда вершина отмирающей связи является точкой ветвления, ее индекс ветвления просто уменьшается на единицу. Нетрудно убедиться в том, что непротиворечивая работа предложенного алгоритма возможна лишь при условии «самоизбегания» процесса металлизации, когда передача возбуждения некоторому узлу решетки возможна лишь при наличии у него единственного металлизированного соседа. Выполнение последнего условия запрещает появление петель в металлизированных кластерах. Возникающие при этом структуры без петель и самопересечений называются в теории графов деревьями, а их совокупность — лесом. В нашем лесу, возникающем в процессе эволюции потенциального рельефа, деревья не только растут и исчезают, но и объединяются друг с другом, причем алгоритм объединения сводится просто к увеличению индекса ветвления общей вершины соединившихся кластеров.

3.4. Результаты компьютерного моделирования

Пространственный белый шум

В этом простейшем случае при отсутствии внешнего поля и фиксированном времени жизни проводящих возбуждений максимальное число активированных связей не превышает величины 0,25, соответствующей обычной статической перколяции на кубической решетке. После некоторого переходного процесса система приходит в состоя-

ние динамического равновесия вблизи порога перколяции. Этот критический уровень определяется равновесием между процессами роста мелкомасштабного потенциального рельефа и процессами выравнивания потенциалов в соседних ячейках при возникновении пробоя между ними. Появление новых активированных связей компенсируется их исчезновением в процессе металлизации. Качественно эволюция системы ничем не отличается от динамики лесных пожаров.

Модификация перколяционной проблемы, связанная с дополнительным условием самоизбегания процесса активации, приводит к заметному увеличению порога протекания. Из рис. 16 видно, что в случае самоизбегающей кластеризации требуемый уровень активации достигает примерно 37 %.

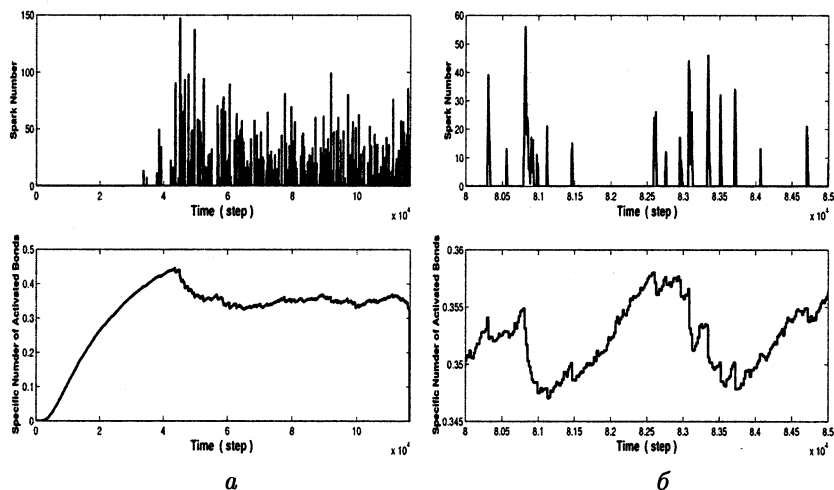


Рис. 16. Динамика трехмерной системы при условии самоизбегания и в отсутствие внешнего поля: *а* — период роста и выход на стационарный режим; *б* — фрагмент установившейся динамики.

Пространственный белый шум плюс внешнее поле

При наличии внешнего поля пороги динамической перколяции рассчитывались с учетом конечной проводимости возникающих кластеров. Если скорость роста внешнего поля чрезвычайно мала в сравнении со скоростью роста локальных флуктуаций, то динамика практически не меняется по сравнению со случаем самоизбегающей перколяции. Связано это с возникновением специфической средней про-

водимости системы, которая обусловлена эстафетным взаимодействием возникающих проводящих кластеров. Этот механизм, в основе которого лежит уже рассмотренный нами в разд. 2 эффект направленной перколяции, полностью нейтрализует сторонние источники, создающие внешнее поле. При этом характерные перепады потенциала, вызванные поляризацией проводящего кластера внешним полем, не превышают уровня мелкомасштабного шума.

Увеличение скорости роста макроскопического поля кардинально меняет картину эволюции системы. Теперь эстафетная проводимость уже не справляется с нейтрализацией сторонних источников, индуцированные внешним полем перепады потенциала на периферии проводящих кластеров начинают превышать уровень шума и процесс активации мелкомасштабных пробоев выходит за рамки отдельного перколяционного кластера, вовлекая в процесс металлизации значительные участки модельной системы (рис. 17, б). При этом амплитуда максимальных вспышек (число вспышек на одном шаге модельного времени) на порядок превосходит амплитуду вспышек, соответствующую самым крупным кластерам в ситуации с нулевым или слабым внешним полем (см. рис. 16). Эффект проиллюстрирован рис. 17 а, б. При еще более быстром восстановлении однородного внешнего поля практически любая критическая пара приводит к тотальной металлизации всей системы.

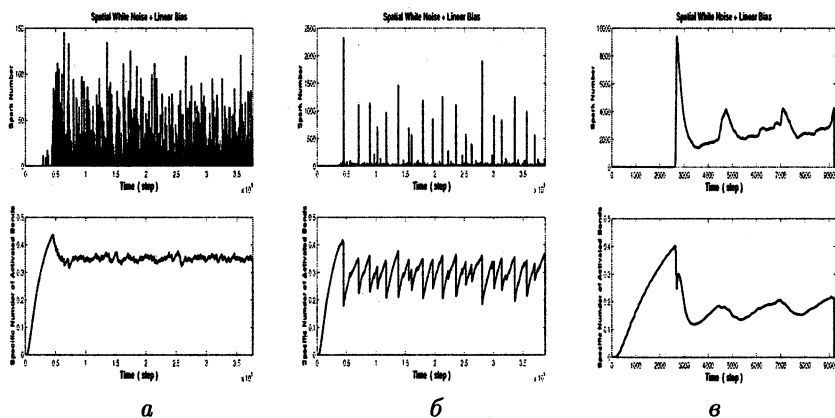


Рис. 17. Динамика в трехмерном случае в присутствии внешнего поля: а — скорость роста поля уступает скорости роста локальных неоднородностей; б — скорость роста поля увеличена в десять раз по сравнению с ситуацией на левом рисунке; в — вечный кластер.

Заметим, что наличие внешнего поля может привести к возникновению интересной ситуации, когда возникший кластер вообще не умирает, а лишь меняет свою конфигурацию (рис. 17, в). При этом сторонние источники, поддерживающие внешнее поле, полностью компенсируют токи выравнивания потенциала, которые непрерывно текут по кластеру, поддерживая его жизнеспособность. Возникает что-то вроде библейской неопалимой Купины или мифической Гидры. Плазменным прообразом такой модельной ситуации может служить шаровая молния, а информационным — система платежей в Интернете.

Самоаффинное потенциальное поле

Примененный нами пространственный аналог алгоритма Фосса, по существу, отражает многомасштабный характер развития неустойчивости, приводящей к росту потенциального рельефа. Действительно, результирующее поле в этом случае является суперпозицией флуктуаций с широким спектром характерных размеров. Их диа-

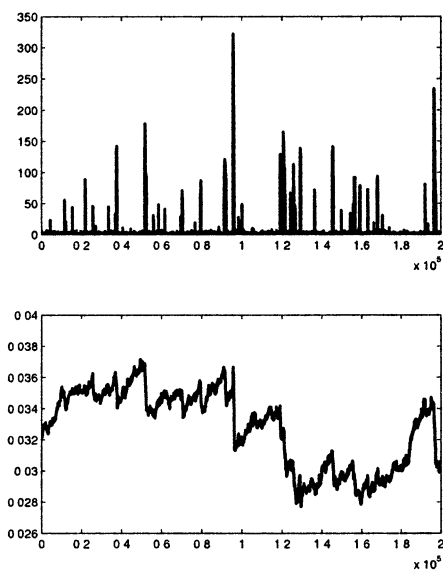


Рис. 18. Временная эволюция трехмерного броуновского потенциального рельефа. Верхний график соответствует числу ячеек, пробитых на одном шаге модельного времени, нижний — удельному числу активированных связей в системе.

пазон простирается от минимальной длины, равной шагу модельной решетки, до максимальной, сопоставимой с размерами системы. Механизм ограничения роста поля, связанный с возникновением и распространением мелкомасштабных пробоев, приобретает в многомасштабном потенциальном поле совершенно уникальный характер. Если в случае белого пространственного шума при распространении возбуждения последовательно использовались мелкомасштабные флуктуации, то здесь перенос возбуждения помимо мелкомасштабных флуктуаций последовательно включает флуктуации с непрерывно увеличивающимися характерными масштабами. Растущий кластер с размерами $r(t)$ замыкает собой постоянно увеличивающийся перепад потенциала, изменение которого на расстоянии r от фиксированной точки определяется асимптотикой $\delta u \sim r^H$, а соответствующие флуктуации поля убывают как $\delta E \sim \delta u / \delta r \sim r^{H-1}$, где H — показатель Херста обобщенного броуновского распределения, использованный нами в алгоритме Фосса. При этом происходит значительное понижение требуемого уровня активации. Для случая, представленного на рис. 18, этот уровень не превышает трех с половиной процентов!

* * *

Мы задумывали нашу статью как своеобразный пригласительный билет в интереснейшую область современных исследований, связанных с теорией перколяции и приложением перколяционных идей в области нелинейной динамики.

Во введении мы уже отмечали, что перколяция возникла как альтернатива диффузионному подходу к задачам проникновения одной субстанции через другую. Альтернатива эта оказалась не только содержательной, но и весьма привлекательной в прикладном аспекте, более того, она оказалась созвучной средствам дискретной математики, таким как теория графов, фрактальная геометрия, клеточные автоматы, актуализируя их как альтернативу средствам классического анализа.

Литература

1. Broadbent S. R., Hammersley J. M. Percolation processes. I. Crystals and mazes // Proc. Camb. Phil. Soc. 1957. V. 53. P. 629—641.
2. Bunde A., Havlin S. Fractals and Disordered Systems. Springer-Verlag, 1996.
3. Alexander S., Orbach R. Density of states on fractals // J. Phys. Lett. 1982. 43:L625.

4. *Last B. J., Thouless D. J.* Power-law behaviour in 2D diluted conducting material // *Phys. Rev. Lett.* 1971. V. 27. P. 1719.

5. *Gouyet J. F.* Dynamics of diffusion and invasion fronts: on the disconnection-reconnection exponents of percolation clusters // *Soft Order in Physical Systems* / Ed. Y. Rabin, R. Bruinsma. N. Y.: Plenum Press, 1994. P. 159–162.

6. *Mandelbrot B. B.* Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence // *Lecture Notes in Physics 12: Statistical Models and Turbulence* / Eds. M. Rosenblatt, C. Van Atta. N. Y.: Springer, 1972. P. 333–351.

7. *Bak P., Chen K., Tang C.* A forest fire model and some thoughts on turbulence // *Phys. Lett.* 1990. V. A 147. P. 297.

8. *Drossel B., Schwabl F.* Self-organized forest-fire model // *Phys. Rev. Lett.* 1992. V. 69. P. 1629.

9. *Bak P., Tang C., Wiesenfeld K.* Self-Organized Criticality // *Phys. Rev.* 1988. V. A 38. P. 364–376.

10. *Iudin D. I., Trakhtengerts V. Yu., Hayakawa M.* Fractal dynamics of electric discharges in a thundercloud // *Phys. Rev.* 2003. E 68. 016601.

11. *Трахтенгерц В. Ю., Иудин Д. И., Григорьев А. Н.* О фрактальной динамике активных сред // *Нелинейные волны'2002* / Отв. ред. А. В. Гапонов-Грехов, В. И. Некоркин. Н. Новгород: ИПФ РАН, 2003. С. 287–304.

12. *Voss R. F.* Random fractal forgeries // *Fundamental algorithms in computer graphics* / Ed. R. A. Earnshaw. Berlin: Springer-Verlag, 1985. P. 805–835.

ЭВОЛЮЦИЯ И КОЛЛАПС САМОГРАВИТИРУЮЩЕГО ТОРОИДАЛЬНОГО ВИХРЯ

Е. Ю. Банникова, К. Ю. Блюх, В. М. Конторович

В последнее время возобновился интерес к исследованию тороидальных конфигураций в контексте астрофизических задач (см., например, [1, 2]). Как известно, вихревые движения могут стабилизировать тороидальные конфигурации в силу сохранения углового момента, а нетривиальная топология приводит к существенным отличиям от поведения для вращающейся сферы или диска. Особую роль, на наш взгляд, должны играть тороидальные вихри с циркулирующей скоростью по малому контуру, самогравитирующие аналоги которых до последнего времени не рассматривались. В данной работе мы не учитываем орбитальное движение по большой окружности тора, используемое при моделировании толстых аккреционных дисков [3, 4], ограничиваясь только однонаправленным движением по малой окружности (его образом может служить «дымовое кольцо» курильщика), которое и будем называть в дальнейшем вихревым движением.

1. Эволюция тороидального вихря

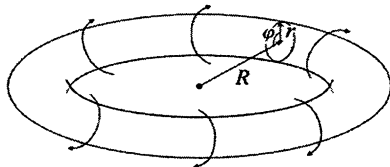
Тонкий вихрь

Предположим, следуя [5], что на начальной стадии форма вихря представляет собой тонкий тор (рис. 1), радиусы которого удовлетворяют неравенству

$$r \ll R. \quad (1)$$

(Примером такого вихря в гидродинамике может служить вихрь Максвелла [6, 7].)

Рис. 1. Тонкий тороидальный вихрь типа вихря Максвелла. Стрелки указывают направление движения частиц.



Будем следить за движением пробной частицы на поверхности тора¹. В данной системе помимо одной вращательной степени свободы

¹ Для определённости мы рассматриваем тор в качестве фигуры равновесия и компактного вихря, хотя истинная форма равновесия при данной топологии неизвестна.

имеются две поступательные: им отвечают изменения малого и большого радиусов r и R . (Мы не рассматриваем здесь движения вихря как целого.)

Запишем гамильтониан системы в виде

$$H = \frac{1}{2M} \left[p_r^2 + p_R^2 + \frac{p_\phi^2}{r^2} \right] + U(r, R). \quad (2)$$

Здесь M — полная масса вихря; p_s — импульсы, соответствующие координатам s ; ϕ — циклическая координата вращения; гравитационная потенциальная энергия системы U имеет вид

$$U(r, R) = G \frac{M^2}{\pi R} \ln \frac{\tilde{\alpha} r}{R}, \quad (3)$$

где $\tilde{\alpha} = \alpha/e$, α — численный коэффициент порядка единицы.

Выражение для потенциальной энергии получено из уравнений движения частицы [5], соответствующих гамильтониану (2), одно из которых описывает эволюцию вращающегося цилиндра, а другое — гравитационное сжатие тонкого кольца радиуса R . Логарифмический член возникает из-за расходимости силы, действующей на частицу на поверхности бесконечно тонкого кольца. С учётом выражения для потенциальной энергии (3) уравнения движения для двух поступательных степеней свободы имеют вид

$$\ddot{r} = \frac{p_\phi^2}{M^2 r^3} - G \frac{M}{\pi R r}, \quad (4a)$$

$$\ddot{R} = G \frac{M}{\pi R^2} \ln \frac{\alpha r}{R}, \quad (46)$$

и интеграл движения для вращательной:

$$p_\phi = M r^2 \dot{\phi} = \text{const}. \quad (5)$$

Последний представляет собой закон сохранения момента импульса (величина $2\pi p_\phi / M$ соответствует циркуляции скорости Γ). Вследствие условия (1) $(r/R) \ln(r/R) \ll 1$. Тогда сила гравитационного притяжения по r (4a) много больше силы гравитационного сжатия по R (46). Это дает основание разделить эволюцию системы по разным масштабам: быстрому по r и медленному по R (рис. 2). Быстрая эволюция (при практически неизменном R) может привести к установлению равновесия в уравнении (4a). Сила гравитационного притяжения при этом будет скомпенсирована центробежной силой вращения на расстоянии

$$r_{eq} = \sqrt{\frac{\pi p_\phi^2 R}{GM^3}} = \sqrt{\lambda R}, \quad \lambda = \Gamma^2 / (4\pi GM). \quad (6)$$

Это выражение определяет равновесный малый радиус как функцию большого, а неравенство $r > r_{eq}$ соответствует критерию гравитационной неустойчивости [8] с джинсовым масштабом λ .

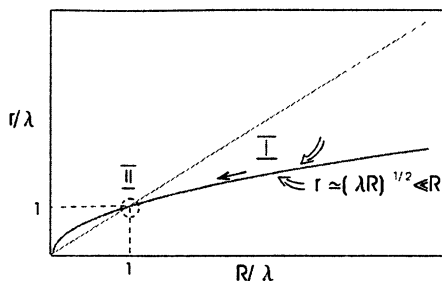


Рис. 2. Схема эволюции вихря

Компактный вихрь

Попробуем представить, как может происходить сжатие гравитирующего вихря, когда он является компактным объектом, топологически эквивалентным тору. Можно ожидать, что под действием гравитационных сил он будет приближаться к некоторой конфигурации, напоминающей вихрь Хилла [6, 7] (рис. 3).

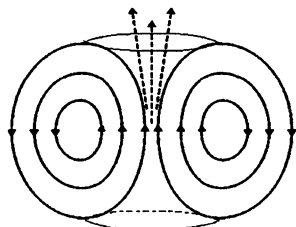


Рис. 3. Вихрь типа вихря Хилла

Рассмотрим условия равновесия компактного вихря в модели, в которой используются полученные выше соотношения для тонкого вихря. Подставляя (6) в (3), получим эффективную потенциальную энергию для эволюции по большому радиусу тора вдоль кривой равновесия. При $R \rightarrow 0$ $U(R) \propto (1/2R) \ln(\alpha^2 \lambda / e^2 R) \rightarrow +\infty$. Этот потенциальный барьер при малых R возникает за счёт центробежных сил, определяемых циркуляцией. При $R \rightarrow \infty$ $U(R)$ отрицательно и стремится к нулю. Вид $U(R)$ напоминает кривую для атомного потенциала. Условие равновесия $\partial U / \partial R = 0$ даёт нам значение $R_c = \alpha^2 \lambda e$. Условие $r = \sqrt{\lambda R} < R$, которое выполняется при $\lambda < R$, в точке минимума сводится к условию на параметр α : $\alpha^2 > e$. При этом неравенство $\lambda < r(R_c) < R_c$ выполняется автоматически. Значение $\ln(eR/\alpha r)$, которое велико для тонкого вихря, поскольку возникло из-за расходимости интеграла [5], в точке равновесия становится порядка единицы. В точке минимума $\alpha r/e < R$ в силу $1 < \sqrt{e}$. Внутренний радиус тора равен

$$R_c - r(R_c) = \lambda \left(\alpha / \sqrt{e} - 1 \right) \alpha / \sqrt{e} + \lambda_\theta, \quad \lambda_\theta = \frac{\pi p_\theta^2}{GM^3}, \quad (7)$$

где λ_θ — джинсова длина, связанная с введенным здесь в рассмотрение слабым ($\lambda_\theta \ll \lambda$) орбитальным движением, и в его отсутствие при $\alpha \rightarrow \sqrt{e}$ внутренний радиус стремится к нулю. Полученное состояние равновесия существенно связано с параметром α . Этот пример, конечно, не может служить доказательством устойчивости компактного вихря не только потому, что нам неизвестен этот параметр, но и вследствие того, что используемые соотношения для тонкого тора находятся здесь вне области их применимости и остаются справедливыми лишь по порядку величины. (Другой, менее формальный способ оценок R_c и других величин приведён в [5].) При сжатии тороидальный вихрь, несомненно, будет деформироваться, и нахождение его равновесной конфигурации представляет собой сложную и нерешённую задачу. Однако компенсация центробежных и гравитационных сил, тем более при наличии орбитального вращения, скорее всего, будет иметь место.

2. Притяжение частицы к тору

Покажем с помощью элементарных рассуждений, что на пробную частицу внутри тора действует сила, притягивающая ее к тору. Вначале рассмотрим тонкий тор, предельным случаем которого является цилиндр. Пробная частица, естественно, будет притягиваться к цилиндру. Уже это означает, что на пробную частицу будет действовать сила, направленная к тору, точнее — к ближайшей к частице части тора.

Представим тор в виде системы концентрических колец. Рассмотрим простейший случай: притяжение пробной частицы к диаметрально противоположным участкам кольца. Выделим два «сектора» с вершинами в пробной частице с малым угловым раствором, симметричных относительно диаметра кольца, проходящего через пробную частицу. Рассмотрим силы притяжения частицы к противоположащим дугам кольца внутри секторов. Если частица находится в центре кольца, они уравновешивают друг друга (рис. 4, а).

Смещая частицу из центра (вдоль выделенного диаметра), мы видим, что масса дуги растёт (или убывает) линейно с расстоянием от частицы, а сила меняется обратно квадрату расстояния (рис. 4, б). Поэтому сила притяжения от «дальней» дуги падает, хотя длина дуги растёт, а сила притяжения к ближайшей дуге растёт, хотя длина дуги падает с приближением к ней частицы. Возникает нескомпенсированная сила притяжения частицы к кольцу и, соответственно, к тору.

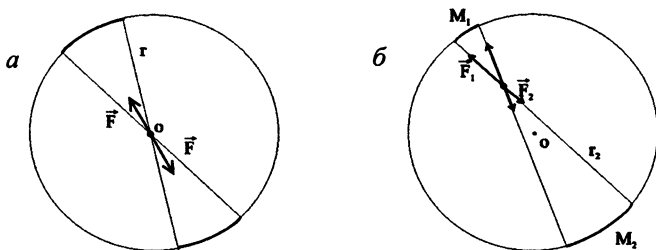


Рис. 4. Схема притяжения пробной частицы к кольцу

Расширяя угол раствора, мы вынуждены перейти от элементарных формул к интегралам, но суть дела и результат это никак не меняет. Заметим, что в случае сферы (если рассуждать аналогичным образом) притягивающая частицу масса пропорциональна площади, вырезаемой телесным углом на сфере. Поэтому при смещении частицы из центра точная компенсация сил сохраняется: масса изменяется квадратично с расстоянием и компенсируется обратной зависимостью силы от квадрата расстояния. Поэтому, как хорошо известно, внутри сферы (в отличие от тора) на пробную частицу сила гравитации не действует.

Предыдущие рассуждения подкрепляются вычислением траекторий пробных частиц. Для движений в «вертикальной» плоскости симметрии примеры траекторий приведены в [5]. Там же отмечено, что вблизи от оси тора возможно эффективное рассеяние частиц, которое способно приводить к возникновению однонаправленного джета (струи) вдоль оси (рис. 3). Примеры трехмерных движений приведены ниже.

3. Трёхмерные траектории частицы в гравитационном поле кольца

Рассмотрим движение пробной частицы в заданном гравитационном поле кольца фиксированного радиуса, дополнив проведенную ранее картину движений трехмерными траекториями (соответствующие уравнения очевидным образом обобщают приведенные в [5]).

Как видно из рис. 5, а и 5, б, при небольших энергиях частица вращается по малому радиусу, обвивая кольцо (рис. 5, а отвечает движению частицы в плоскости, 5, б — наличию орбитального движения). Это движение соответствует тонкому вихрю (возможная первая стадия эволюции). С увеличением энергии частицы появляются различные сложные траектории (рис. 5, в). Наконец, начиная с некоторого значения энергии частица переходит на почти замкнутые траектории типа «восьмёрки» (рис. 5, г). При этом радиус вращения частиц становится порядка

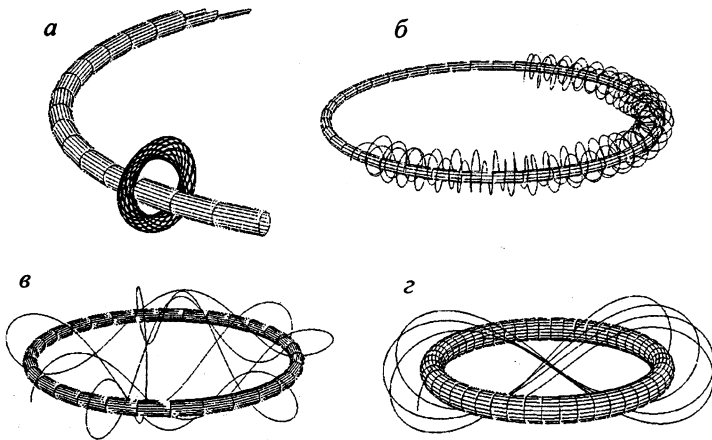


Рис. 5. Трёхмерные финитные траектории пробной частицы в гравитационном поле кольца

радиуса кольца, что характерно для компактной стадии сжатия вихря. Кинетическая энергия частицы на таких траекториях уже близка к энергии, необходимой для отрыва частиц. Движение частиц по «восьмёркам» приведет к эффективному их столкновению и рассеянию вблизи оси вихря. Существование потоков вещества типа «восьмёрок» следует и из гидродинамической модели вихря Максвелла (см., например, [7]).

4. Джинсова масса для тонкого и компактного вихрей

Чтобы связать давление с параметрами вихря, необходимо привлечь уравнение состояния вещества и процесса сжатия вихря.

Учёт теплового давления газа [9] приводит к замене (6) на

$$r_+ = \frac{\Gamma}{2\pi \sqrt{\frac{GM}{\pi R} - \frac{1}{2} c_s^2 j_1^2}}, \quad (8)$$

где c_s — скорость звука; j_1 — геометрический фактор порядка единицы (для тонкого вихря, который можно заменить цилиндром, это нуль функции Бесселя). Из соотношения (8) видно отклонение от корневой зависимости, связанное с влиянием давления газа на условие равновесия (равновесный радиус r_+ , играющий роль джинсова масштаба, увеличивается). Домножая r_+^2 на $2\pi^2 R\rho$, получим уравнение для линейной джинсовой массы $\chi(\rho, \Gamma) \equiv M/2\pi R$ тонкого вихря, решая которое приходим к следующему выражению:

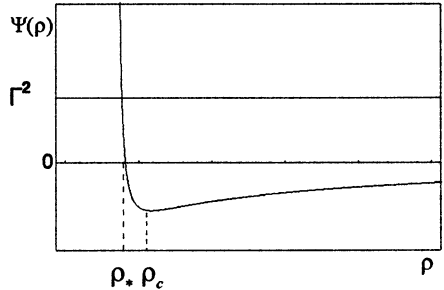
$$\chi_J = \frac{c_s^2 j_1^2}{8G} \left(1 + \sqrt{1 + \frac{8G\rho\Gamma^2}{\pi c_s^4 j_1^4}} \right). \quad (9)$$

Полагая χ_J равной заданной линейной массе вихря χ , находим неявное выражение для равновесной плотности тонкого вихря $\rho(\chi, \Gamma)$:

$$\Gamma^2 = \frac{8\pi\chi}{\rho} \left(G\chi - \frac{1}{4}c_s^2(\rho)j_1^2 \right), \quad (10)$$

графическое решение которого приведено на рис. 6.

Рис. 6. Графическое решение уравнения (10) для плотности тонкого вихря с циркуляцией Γ . Кривая $\Psi(\rho)$ соответствует правой части уравнения.



Для политропного процесса сжатия результаты приведены в [9] и существенно отличаются от известных результатов для сферы. Ниже мы будем интересоваться «поздними» (с точки зрения звездной эволюции) стадиями вихревого кольца, когда газ в нём можно считать вырожденным. Тогда c_s есть функция одной только плотности ρ , которую мы в данном рассмотрении считаем не зависящим от координат параметром, отслеживающим, как это принято в простейших моделях, процесс гравитационного сжатия (см., например, [8]).

При политропном уравнении состояния выражение (10), вообще говоря, представляет собой трансцендентное уравнение для ρ [9]. В интересующем же нас случае вырожденного газа выражение (10) является алгебраическим уравнением. Для вырожденного электронного и нейтронного газов мы получаем алгебраические уравнения соответственно 8-го и 5-го порядков. Но, как мы увидим, в наиболее важном для нас релятивистском случае¹ ситуация значительно упрощается и становится возможным явное решение задачи.

Заметим, что всегда существует равновесное решение (10). При этом плотность $\rho < \rho_*(M, \Gamma)$, где предельное значение ρ_* определяется

¹ Релятивистские обобщения рассмотренной проблемы могут оказаться весьма перспективными при анализе различных астрофизических приложений, в том числе к ядрам (активных) галактик, проблеме сверхновых и гиперновых, молодым звездным объектам — всюду, где актуальна проблема гравитационного коллапса в сочетании с вращением.

условием $4G\chi = c_s^2(\rho)j_1^2$. (При $\rho > \rho_*$ скорость звука слишком велика и среда ведёт себя как «практически» несжимаемая. Поэтому и нет соответствующих равновесных решений.) В случае давления вырожденного электронного газа, когда вклад в плотность, тем не менее, обеспечивается нерелятивистскими протонами (см. ниже), с ростом плотности скорость звука неограниченно растёт [10]. Поэтому всегда есть область плотностей $\rho > \rho_*$, для которых сжатие за счёт гравитации недостаточно, среда «несжимаема». В нейтронном случае, поскольку скорость звука, монотонно возрастаая, выходит на релятивистский предел $c^2/3$, при $12G\chi \rightarrow c^2 j_1^2$ значение ρ_* уходит на $+\infty$. Решение уравнения (10) возможно на всей оси.

Для компактного вихря, полагая $r = \varepsilon R$ ($\varepsilon < 1$), приходим из (8) к уравнению для джинсовой массы M_J , решение которого имеет вид

$$\mu M_J^{2/3} = \frac{c_s^2}{\rho^{1/3}} \left(1 + \sqrt{1 + \xi \Gamma^2 \frac{\rho}{c_s^4}} \right), \quad (11)$$

где $\mu = 4(2\pi^2 \varepsilon^2)^{1/3} G / (\pi \zeta) \sim G$; $\xi = 8G / \pi \zeta^2 \sim G$; $\zeta \sim 1$ — геометрический фактор.

5. Интервал устойчивых масс для тороидального вихря вырожденного нейтронного газа

Рассмотрим поведение тороидального вихря, состоящего из вырожденного электронного или нейтронного газа. Это поведение существенно регулируется зависимостью скорости звука от плотности и разительно отличается для случая вырожденного электронного газа, когда массовая плотность определяется нерелятивистскими частицами (нуклонами — протонами и нейтронами), и случая вырожденного нейтронного газа, когда нейтронами определяются и давление, и плотность.

Учитывая точные выражения [10] для плотности и давления вырожденного электронного и нейтронного газа, скорость звука для электронов можно привести к виду

$$c_e^2(\rho) = \beta \frac{\rho^{2/3}}{\sqrt{\rho_{0e}^{2/3} + \rho^{2/3}}}, \quad (12)$$

где

$$\rho_{0e} = \frac{\mu_e m_n}{3\pi^2 l_e^3}; \quad \beta = \frac{m_e c^2 l_e}{9\pi^2} \cdot \left(\frac{3\pi^2}{\mu_e m_n} \right)^{4/3};$$

l_e — электронная комптоновская длина; μ_e — число электронов на один нуклон; m_n — масса протона. Скорость звука монотонно растёт с ростом

плотности (рис. 7, а). Эффективный показатель политропы плавно изменяется от $\gamma = 5/3$ для нерелятивистского до $\gamma = 4/3$ для релятивистского газа. Поскольку остаётся выполненным условие $\gamma \geq 4/3$, срыва в коллапс не происходит. Иная ситуация в случае нейтронного газа. В этом случае

$$c_n^2(\rho) = \frac{c^2}{3} \frac{\rho^{2/3}}{(\rho_{0n}^{2/3} + \rho^{2/3})}, \quad (13)$$

где $\rho_{0n} = m_n^4 c^3 / (3\pi^2 \hbar^3)$.

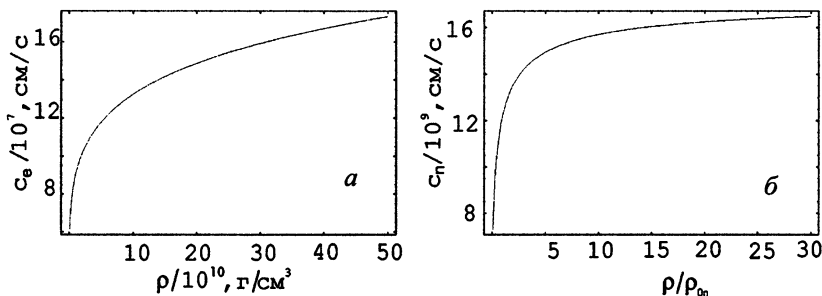


Рис. 7. Зависимость скорости звука от плотности для вырожденного электронного (а) и нейтронного (б) газа [10]

Видно, что с ростом плотности и переходом к ультрарелятивистскому пределу скорость звука стремится к постоянному значению $c_n^2 \rightarrow c^2/3$ (рис. 7, б). Постоянство скорости звука означает, что эффективный показатель политропы $\gamma \rightarrow 1 < 4/3$. Как известно, при односвязной топологии (например, для сферы) $\gamma < 4/3$ отвечает состоянию неустойчивости, при котором из-за уменьшения джинсовой массы с ростом плотности звезда срывается в коллапс. В случае тороидального вихря ситуация кардинально отличается. И действительно, решение уравнения для плотности компактного вихря (11) в этом (релятивистском) пределе имеет вид

$$\rho^{1/3}(\Gamma, M_J) = \frac{\mu^2 M_J^{4/3}}{2\xi\Gamma^2} \left(1 \pm \sqrt{1 - \frac{8\xi c^2}{3\mu^3 M_J^2} \Gamma^2} \right) \quad (14)$$

и описывает две ветви (рис. 8), одна из которых, соответствующая меньшим плотностям, становится неустойчивой (как и должно быть при $\gamma < 4/3$), а другая сохраняет устойчивость из-за вклада циркуляции.

Джинсова масса проходит через минимум при $\rho = \rho_{\min}(\Gamma) = 8c^4/(9\xi\Gamma^2) \sim c^4/(\Gamma^2)$, равный

$$M_{\min} = \sqrt{\frac{\xi}{6\pi}} \cdot \frac{c}{\varepsilon G} \Gamma \sim \frac{c\Gamma}{G}. \quad (15)$$

Происхождение минимальной массы M_{\min} здесь чисто динамическое и не имеет отношения к другим известным ограничениям на массу нейтронных звёзд со стороны малых масс [11]. Оценка M_{\min} в (15) следует также из соображений размерности при определяющих параметрах G , c и Γ . При больших, но еще физически достижимых значениях циркуляции $\Gamma \sim r_g c$, где r_g — гравитационный радиус эффективной сферической системы, это ограничение будет наиболее существенным. При этом минимальная масса может достаточно близко приближаться к максимальной массе нейтронной звезды M_{OV} [12, 13], создавая тем самым узкую щель для масс тороидальных нейтронных звезд (см. ниже).

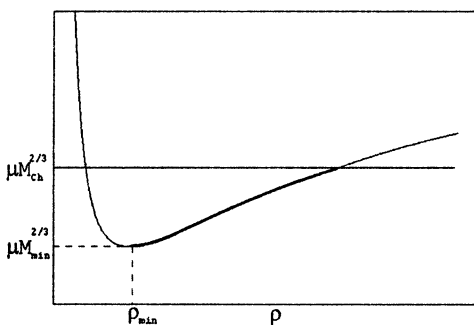


Рис. 8. Зависимость $\mu M^{2/3}$ от плотности ρ для компактного тороидального вихря вырожденного нейтронного газа.

Значение M_{OV} может быть получено из соображений размерности, если определяющими параметрами выбрать G , c и ρ . Известно, что из параметров G , c , \hbar из соображений размерности строится планковская масса $m_{pl} = \sqrt{\hbar c/G}$. В случае, когда \hbar входит только через концентрацию частиц, а концентрация однозначно связана с массовой плотностью, что имеет место для вырожденной нейтронной жидкости или газа, создающих квантовое давление, выбор плотности ρ в качестве определяющего параметра (вместо постоянной Планка) совершенно естествен. Это приводит из одних только соображений размерности к выражениям для предельной массы

$$M_{OV} \approx G^{-3/2} c^3 \rho^{-1/2}. \quad (16)$$

Оценивая для нейтронной звезды массовую плотность ρ как ядерную плотность согласно

$$\rho \approx \rho_n \approx \frac{m_n}{\lambda_n^3}, \quad (17)$$

где комптоновская длина волны нейтрона (протона) равна $\lambda_n \sim \hbar/m_n c$, приходим к оценке $\rho_n \sim m_n^4/(\hbar c)^3$. Отсюда, используя планковскую массу, можно переписать выражение для M_{OV} в виде

$$M_{OV} \approx \frac{m_{pl}^3}{m_n^2}. \quad (18)$$

Численно, как известно, это значение лежит вблизи от солнечной массы. Поскольку предельная масса не чувствительна к тому, какой именно вырожденный газ (электронный или нейтронный) создает квантовое давление, данная оценка правильна и для белых карликов (электронных вырожденных звезд). Для уравнения состояния идеального электронного вырожденного газа после введения поправок ОТО это приводит к известной оценке массы Ландау — Чандрасекара, которую можно с достаточной точностью записать в виде $M_{Ch} \approx \sqrt{2}M_\odot$. Для нейтронных звезд существенна неидеальность газа и вклад давления, что приводит для сферы к предельной массе Оппенгеймера — Волкова M_{OV} , отличающейся от M_{Ch} численным коэффициентом порядка единицы. Поскольку в случае тороидального вихря не известны решения уравнений Эйнштейна, учитываемые при вычислениях массы для сферических нейтронных звезд, мы вообще не будем на данном этапе различать M_{Ch} и M_{OV} .

Условие $v \leq c$, приводящее к предельной массе Ландау — Чандрасекара, переходит в случае тороидального вихря (с учётом циркуляции) в

$$M_{Ch}^{2/3}(\Gamma) = \frac{1}{2} M_{Ch}^{2/3}(0) \left(1 + \sqrt{1 + \kappa \left(\frac{m_n}{m_{pl}} \right)^2 \left(\frac{1}{l_n c} \right)^2 \Gamma^2} \right), \quad (19)$$

где κ — коэффициент порядка единицы; l_n — комптоновская длина нейтронов; m_{pl} — планковская масса. Действительно, исходя из

$$v = \frac{\hbar}{m_n^{4/3}} \rho_{eq}^{1/3} \leq c, \quad (20)$$

подставляя значение $\rho_{eq}^{1/3} = \rho^{1/3}(\Gamma, M)$ из (14) и разрешая полученное соотношение относительно массы, приходим к зависящему от циркуляции выражению для $M_{Ch}(\Gamma)$ (19). Существенно, что интервал между нижним и верхним пределами для допустимых масс с ростом циркуляции сужается, выходя на некоторое предельное значение. (За коэффициентами порядка единицы мы при этом следить не имеем возможности.) Таким образом, для тороидального вихря в случае вырожденного газа возникает конечный интервал устойчивых масс, определяемый циркуляцией скорости [14]

$$M_{\min}(\Gamma) \leq M \leq M_{Ch}(\Gamma). \quad (21)$$

Заметим, что существование подобного интервала является следствием ультрарелятивистского предела и не привязано именно к нейтронному газу. Удобно для оценки циркуляции характерный пространственный масштаб выбрать равным гравитационному радиусу, а скорость — равной скорости света. Тогда безразмерная циркуляция $\Gamma/(r_g c)$ определяется числом порядка единицы. Массу при этом удобно измерять в предельных массах Ландау — Чандрасекара при нулевой циркуляции $M_{Ch}(0)$. Отношение $M/M_{Ch}(0)$ также, как можно ожидать, будет порядка единицы. Для ширины полосы устойчивых масс

$$\Delta M = M_{Ch}(\Gamma) - M_{\min}(\Gamma) \quad (22)$$

получаем (коэффициенты порядка единицы при Γ опущены):

$$\Delta M = -M \frac{\Gamma}{r_g c} + \frac{1}{2\sqrt{2}} M_{Ch}(0) \left(1 + \sqrt{1 + \frac{M^2}{M_{Ch}^2(0)} \cdot \left(\frac{\Gamma}{r_g c} \right)^2} \right)^{3/2}, \quad (23)$$

где мы выразили M_{\min} через циркуляцию согласно $M_{\min} \sim M\Gamma/(r_g c)$. Используя полученные выше выражения для M_{Ch} и M_{\min} и считая ΔM известным из наблюдений, мы можем найти отсюда требуемое значение циркуляции $\Gamma(\Delta M, M)$, приводящее к заданному интервалу ΔM . Условия положительности Γ приводит к неравенству $M > \sqrt{2} (\Delta M)^2 / M_{Ch}(0)$.

В частности, при $\Delta M = 1M_{\odot}$ для этого требуются массы торов $M > 1M_{\odot}$. Как указано в работе А. М. Черепашука [15], массы нейтронных звезд¹ сосредоточены в полосе $1 \leq M/M_{\odot} \leq 2$. Если отождествить этот интервал с интервалом масс тороидальных вихрей, можно определить значение требуемой для этого циркуляции $\Gamma/r_g c \sim 1$.

* * *

Таким образом, эволюция самогравитирующего тороидального вихря обладает рядом качественных особенностей, отличающих ее от эволюции сферы и диска из-за сохраняющейся циркуляции скорости. Эти особенности могут быть существенны для астрофизики нейтронных звезд и ряда других астрофизических приложений.

Данная работа частично поддержана фондом Интас, грант 00-00292.

¹ Заметим, что эти данные относятся к тесным двойным системам, в которых требуемая циркуляция, по-видимому, может приобретаться при перетекании вещества через «сопло», вблизи внутренней точки Лагранжа.

Литература

1. *Antonucci R.* Unified models for active galactic nuclei and quasars // *Ann. Rev. Astron. Astrophys.* 1993. V. 31. P. 473—521.
2. *Spivey R. J.* Quasars: a supermassive rotating toroidal black hole interpretation // *Mon. Not. Roy. Astron. Soc.* 2000. V. 316. P. 856—874.
3. *Nishida S., Eriguchi Y.* Runaway instability of BH-neutron torus // *Astrophys. J.* 1996. V. 461. P. 320—326.
4. *Abramowicz M. A., Livio M., Soker N., Szuszkiewicz E.* Local stability of thick accretion disks II // *Astron. Astrophys.* 1990. V. 239. P. 399—403.
5. *Блиох К. Ю., Конторович В. М.* Эволюция и гравитационный коллапс тороидального вихря // *ЖЭТФ.* 2003. Т. 123, № 6. С. 1123—1130.
6. *Ламб Г.* Гидродинамика. М.; Л.: Огиз, Гостехиздат, 1947. 928 с.
7. *Лаврентьев М. А., Шабат Б. В.* Проблемы гидродинамики и их математические модели. М.: Наука, 1973. 416 с.
8. Происхождение и эволюция галактик и звезд / Под ред. С. Б. Пикельнера. М.: Наука, 1976. 408 с.
9. *Блиох К. Ю., Конторович В. М.* О гравитационном коллапсе и условиях равновесия тороидального вихря при учёте теплового давления // *ПАЖ.* 2003. Т. 29, № 11. С. 816—823.
10. *Зельдович Я. Б., Новиков И. Д.* Теория тяготения и эволюция звезд. М.: Наука, 1971. 484 с.
11. *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. Ч. 1. М.: Наука, 1951. 480 с.
12. *Липунов В. М.* Астрофизика нейтронных звёзд. М.: Наука, 1987. 296 с.
13. *Шapiro С., Тьюкольски С.* Чёрные дыры, белые карлики и нейтронные звёзды. М.: Мир, 1985. Т. 1. 655 с. Т. 2. 254 с.
14. *Банникова Е. Ю., Блиох К. Ю., Конторович В. М.* Об эволюции самогравитирующего тороидального вихря // *Вестник астрономической школы.* 2003. Т. 3, № 2. С. 100—112.
15. *Черепашук А. М.* Поиски чёрных дыр // *УФН.* 2003. Т. 173, № 4. С. 345—384.

СТАТИСТИКА И РЕАЛЬНОСТЬ В СТОХАСТИЧЕСКИХ ФИЗИЧЕСКИХ СИСТЕМАХ

В. И. Кляцкин

Введение

Часто физические процессы происходят в сложных средах, параметры которых можно рассматривать как некоторые реализации хаотических полей в пространстве и во времени. Решение различного рода динамических задач для этих конкретных реализаций параметров среды практически безнадежно из-за чрезвычайной математической сложности задач. Однако если бы даже удалось решить какую-то задачу в каком-либо конкретном случае, то пользы от этого было бы чрезвычайно мало, так как невозможно было бы сказать о том, что будет в другом случае. В то же время исследователей интересуют основные особенности протекающих явлений без отвлечения на частности. Поэтому очень привлекательной оказалась идея использовать хорошо развитый математический аппарат случайных процессов и полей, т. е. вместо отдельных реализаций исследуемых процессов рассматривать статистические средние по всему ансамблю возможных реализаций. Такая постановка существенно упрощает математический аппарат и позволяет исследовать чрезвычайно широкий круг задач. В настоящее время, например, практически все задачи физики атмосферы и океана в той или иной степени основываются на статистическом анализе.

Введение случайности в параметрах среды порождает стохастичность в самих физических полях. Индивидуальные реализации, например, скалярных двумерных полей $\rho(\mathbf{R}, t)$, где $\mathbf{R} = x, y$, напоминают сложный горный ландшафт со случайно распределенными пиками, провалами, хребтами и перевалами. Обычно используемые методы статистического усреднения (т. е. вычисления средних типа $\langle \rho(\mathbf{R}, t) \rangle$, $\langle \rho(\mathbf{R}_1, t) \rho(\mathbf{R}_2, t) \rangle$ и т. п., где через $\langle \dots \rangle$ обозначено усреднение по ансамблю реализаций случайных параметров) сглаживают качественные особенности типичных реализаций, и зачастую полученные статистические характеристики не только не имеют ничего общего с поведением отдельных реализаций, но и противоречат, на первый взгляд, друг другу. Часто возникают парадоксальные ситуации, когда статистические характеристики экспоненциально растут во времени, в то время как индивидуальные реализации экс-

пониженно убывают во времени. Полная статистика (например, n -точечные моменты) содержит, разумеется, всю информацию о динамической системе. На практике же удается исследовать лишь некоторые простейшие статистические характеристики, связанные главным образом с одновременными и односточечными распределениями вероятностей.

Существуют, однако, физические явления, происходящие с вероятностью единица, которые будем называть *когерентными явлениями*. Для таких явлений возможно установить связь между отдельными реализациями динамических систем со статистическими характеристиками, и их можно понять на основе единого подхода, основанного на анализе односточечных пространственно-временных плотностей вероятностей, соответствующих исходным динамическим стохастическим системам [1–6]. Это особенно актуально для геофизических проблем, связанных с атмосферой и океаном, где, вообще говоря, отсутствует соответствующий ансамбль усреднения и экспериментаторы, как правило, имеют дело с отдельными реализациями.

1. Примеры когерентных явлений в стохастических динамических системах

1.1. Явление переброса в стохастических динамических системах

Остановимся прежде всего на одном стохастическом аспекте, связанном с динамическими стохастическими уравнениями, а именно на *явлении переброса*, обусловленном случайными воздействиями.

Рассмотрим одномерное нелинейное уравнение

$$\frac{d}{dt}x(t) = x(1 - x^2) + f(t), \quad x(0) = x_0, \quad (1)$$

где $f(t)$ — случайная функция времени. При отсутствии случайности ($f(t) \equiv 0$) решение уравнения (1) имеет два устойчивых стационарных состояния $x = \pm 1$ и одно неустойчивое состояние $x = 0$. В зависимости от начального условия решение уравнения (1) придет в одно из устойчивых состояний. Однако в присутствии малых случайных возмущений $f(t)$ динамическая система (1) сперва придет в окрестность одного из устойчивых состояний, а затем, по прошествии некоторого времени, будет переброшена в окрестность другого устойчивого состояния. Очевидно, что подобная ситуация может осуществляться и в более сложных случаях.

Рассмотрим теперь простейший пример стохастической системы, возникающей в статистической теории волн и имеющей сингулярное поведение во времени:

$$\frac{d}{dt}x(t) = -\lambda x^2(t) + f(t), \quad x(0) = x_0, \quad \lambda > 0, \quad (2)$$

где $f(t)$ — случайная функция времени.

В отсутствие случайностей ($f(t) = 0$) решение уравнения (2) имеет вид

$$x(t) = \frac{1}{\lambda(t - t_0)}, \quad t_0 = -\frac{1}{\lambda x_0}.$$

Если $x_0 > 0$, то $t_0 < 0$ и решение задачи $x(t)$ монотонно стремится к нулю с ростом времени. Если же $x_0 < 0$, то решение $x(t)$ обращается в $\{-\infty\}$ в течение конечного времени $t_0 = -1/\lambda x_0$, т. е. становится *сингулярным* и имеет *взрывной характер*. В этом случае влияние случайной силы $f(t)$ на динамику системы не существенно. Оно становится важным только в случае положительного значения x_0 . В этом случае решение задачи, слегка флуктуируя, уменьшается с ростом времени, оставаясь положительным. При достижении достаточно малого значения $x(t)$ оно будет «перепрошено» под действием силы $f(t)$ в область отрицательных значений x и по прошествии некоторого конечного времени обратится в $\{-\infty\}$.

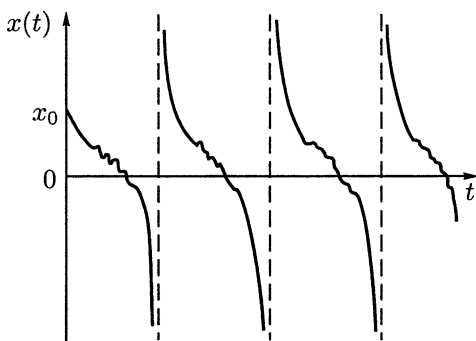


Рис. 1. Типичная реализация решения уравнения (2)

Таким образом, в стохастическом случае для любых значений величины x_0 решение задачи (2) имеет взрывной характер и обращается в $\{-\infty\}$ в течение конечного времени t_0 с вероятностью равной единице. Примерная реализация поведения решения задачи (2) $x(t)$ во времени для $t > t_0$ имеет как бы *квазипериодическую* структуру и схематически изображена на рис. 1.

Статистическое описание явлений переброса хорошо известно. Оно содержится практически во всех учебниках по теории случайных процессов и в простейшем случае описывается краевой задачей для обратного уравнения Фоккера — Планка.

1.2. Частицы в поле случайных сил и скоростей

Динамика малоинерционной частицы в случайном гидродинамическом потоке с гидродинамическим полем скоростей $\mathbf{u}(\mathbf{r}, t)$ описывается уравнениями Ньютона:

$$\begin{aligned} \frac{d}{dt}\mathbf{r}(t) &= \mathbf{V}(t), \quad \mathbf{r}(0) = \mathbf{r}_0, \\ \frac{d}{dt}\mathbf{V}(t) &= -\lambda[\mathbf{V}(t) - \mathbf{u}(\mathbf{r}(t), t)], \quad \mathbf{V}(0) = \mathbf{V}_0(\mathbf{r}_0). \end{aligned} \quad (3)$$

Линейный по полю скоростей $\mathbf{V}(\mathbf{r}, t)$ член в правой части (3) $\mathbf{F}(\mathbf{r}, t) = \lambda\mathbf{V}(\mathbf{r}, t)$ является известной *формулой Стокса* для силы сопротивления, действующей на медленно движущуюся частицу. При аппроксимации частицы шаром с радиусом a , параметр $\lambda = 6\pi a\eta/m_p$, где η — коэффициент динамической вязкости, а m_p — масса частицы.

Для безинерционных частиц параметр $\lambda \rightarrow \infty$ и, как следует из уравнения (3),

$$\mathbf{V}(t) = \mathbf{u}(\mathbf{r}(t), t), \quad (4)$$

и траектория частицы в случайном гидродинамическом потоке с полем скоростей $\mathbf{u}(\mathbf{r}, t)$ описываются простейшим стохастическим уравнением:

$$\frac{d}{dt}\mathbf{r}(t) = \mathbf{u}(\mathbf{r}(t), t), \quad \mathbf{r}(0) = \mathbf{r}_0, \quad (5)$$

т. е. для безинерционных частиц задача определения их траекторий в гидродинамическом потоке является чисто кинематической задачей.

В системе частиц, с формальной точки зрения, каждая частица движется независимым образом. Однако если случайное поле $\mathbf{u}(\mathbf{r}, t)$ имеет конечный пространственный радиус корреляции $l_{\text{кор}}$, то частицы, расстояние между которыми меньше $l_{\text{кор}}$, находятся в общей зоне влияния случайного поля $\mathbf{u}(\mathbf{r}, t)$ и могут появиться новые коллективные особенности в динамике такой системы частиц. В общем случае случайное поле $\mathbf{u}(\mathbf{r}, t)$ может иметь как соленоидальную ($\text{div } \mathbf{u}(\mathbf{r}, t) = 0$), так и потенциальную ($\text{div } \mathbf{u}(\mathbf{r}, t) \neq 0$) составляющие. Как показывают результаты численного моделирования, динамика системы частиц в этих случаях существенно различная.

На рис. 2, *a* для бездивергентного поля скоростей ($\text{div } \mathbf{u}(\mathbf{r}, t) = 0$) изображена эволюция во времени системы частиц (двумерный случай), равномерно расположенных в круге. В этом случае площадь, ограниченная контуром, сохраняется и частицы относительно равномерно заполняют область, ограниченную деформированным контуром. Возникает лишь сильная изрезанность фрактального характера этого контура. Для случая потенциального поля скоростей ($\text{div } \mathbf{u}(\mathbf{r}, t) \neq 0$) частицы, равномерно расположенные в квадрате в начальный момент времени, в процессе временной эволюции образуют кластерные области (компактные области повышенной концентрации частиц, расположенные в зонах в большей степени разреженных). Результаты численного моделирования для этого случая приведены на рис. 2, *b*. Подчеркнем, что это чисто кинематический эффект. Такая особенность динамики частиц исчезает при усреднении по ансамблю реализаций случайного поля скоростей.

Пусть гидродинамическое поле скоростей $\mathbf{u}(\mathbf{r}, t)$ — стационарное во времени и однородное и изотропное в пространстве гауссово случайное поле с дисперсией $\sigma_{\mathbf{u}}^2 = \langle \mathbf{u}^2(\mathbf{r}, t) \rangle$ и временным радиусом корреляции

$$\tau_0 = \frac{1}{\sigma_{\mathbf{u}}^2} \int_0^{\infty} d\tau \langle \mathbf{u}(\mathbf{r}, t + \tau) \mathbf{u}(\mathbf{r}, t) \rangle.$$

В приближении дельта-коррелированности поля $\mathbf{u}(\mathbf{r}, t)$ во времени легко получить выражения для дисперсии и временного радиуса корреляции скорости частицы $\mathbf{V}(t)$ в стационарном режиме ($\lambda t \rightarrow \infty$)

$$\sigma_{\mathbf{v}}^2 = \lambda \tau_0 \sigma_{\mathbf{u}}^2, \quad \tau_{\mathbf{v}} = 1/\lambda. \quad (6)$$

При этом пространственный коэффициент диффузии

$$D = \frac{1}{2} \frac{d}{dt} \langle \mathbf{r}^2(t) \rangle = \sigma_{\mathbf{u}}^2 \tau_0 \quad (7)$$

не зависит от параметра λ .

Одним из условий применимости приближения дельта-коррелированности случайного поля $\mathbf{u}(\mathbf{r}, t)$ во времени является условие $\lambda \tau_0 \ll 1$.

С другой стороны, для безынерционной частицы ($\lambda \rightarrow \infty$) имеет место равенство $\mathbf{V}(t) = \mathbf{u}(\mathbf{r}(t), t)$, т. е. дисперсия и временной радиус корреляции скорости частицы $\mathbf{V}(t)$ в этом предельном случае

$$\sigma_{\mathbf{v}}^2 = \sigma_{\mathbf{u}}^2, \quad \tau_{\mathbf{v}} = \tau_0. \quad (8)$$

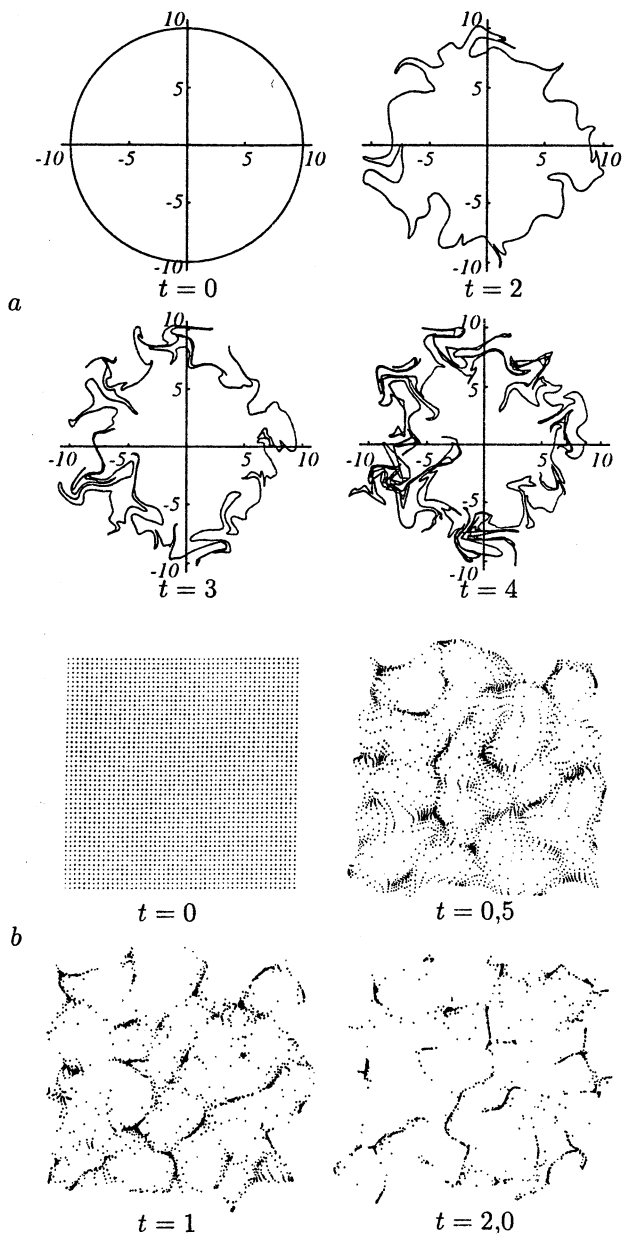


Рис. 2. Результаты численного моделирования диффузии системы частиц, описываемых уравнениями (5) в соленоидальном (а) и дивергентном (б) случайных полях скоростей $\mathbf{u}(\mathbf{r})$.

Сопоставляя равенства (6) с равенствами (8), видим, что они несовместимы, т. е. для выполнения равенства (8) требуется не только выполнение условий $\lambda t \gg 1$ и $t/\tau_0 \gg 1$, но и выполнение условия

$$\lambda\tau_0 \gg 1, \quad (9)$$

что противоречит условию применимости приближения дельта-коррелированности во времени случайного поля $\mathbf{u}(\mathbf{r}, t)$.

Таким образом, приближение дельта-коррелированности во времени случайного поля $\mathbf{u}(\mathbf{r}, t)$ для динамической системы (3) неправильно описывает статистику скорости частицы и ее корреляцию с положением частицы при переходе к приближению безынерционности частиц. В то же время для описания пространственной диффузии безынерционной частицы можно воспользоваться приближением дельта-коррелированности во времени случайного поля $\mathbf{u}(\mathbf{r}, t)$ в стохастическом уравнении (5), так как в нем уже не содержится параметр λ . При этом для пространственного коэффициента диффузии мы снова приходим к формуле (7).

Отметим, что выделение пространственного описания диффузии частицы из его пространственно-временного описания составляет так называемую *проблему Крамерса*.

Впервые, по-видимому, такого рода кластеризация для системы частиц была обнаружена в работах, в которых в рамках простейших уравнений динамики атмосферы проводилось численное моделирование так называемого Eole-эксперимента. В рамках этого глобального эксперимента в 1970 — 1971 годах было запущено в Аргентине 500 баллонов постоянной плотности, которые распространялись по всему южному полушарию на высоте примерно в 12 км.

На рис. 3 представлено распределение баллонов по южному полушарию на 105 сутки после начала численного моделирования, из которого ясно видно, что они концентрируются в группы, что соответствует кластеризации плавучей примеси.

Статистический анализ показывает, что *типичной реализацией* для расстояния между двумя частицами (при достаточно малом начальном расстоянии l_0 между ними), например, в двумерном случае будет экспоненциальная функция времени:

$$l^*(t) = l_0 \exp \left\{ \frac{1}{4} (D^s - D^p) t \right\},$$

растущая или затухающая в зависимости от знака разности $(D^s - D^p)$. Здесь величины D^s и D^p определяются спектральными

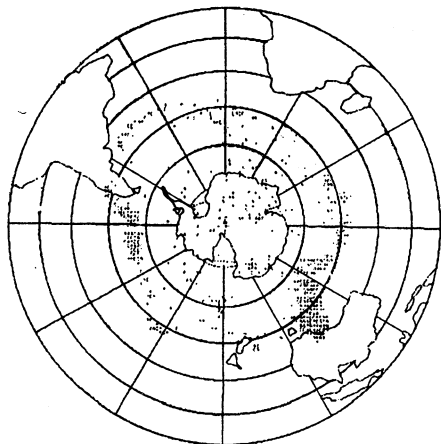


Рис. 3. Распределение баллонов в атмосфере через 105 суток после начала численного эксперимента.

частями соленоидальной и потенциальной составляющих поля скоростей. В частности, для соленоидального случая ($D^p = 0$) имеем экспоненциально растущую типичную реализацию, что соответствует экспоненциально быстрому разбеганию частиц. В другом предельном случае — потенциальном поле скоростей ($D^s = 0$) — типичной реализацией будет экспоненциально убывающая кривая, т. е. налицо стремление частиц «слиться». Это согласуется с результатами численного моделирования (рис. 2). В то же время все статистические моменты расстояния между двумя частицами экспоненциально растут со временем независимо от структуры поля скоростей.

1.3. Диффузия примеси в случайном поле скоростей

Рассмотрим теперь некоторые вопросы статистического описания переноса консервативного поля малоинерционной примеси $\rho(\mathbf{R}, t)$ (здесь для наглядности считаем вектор \mathbf{R} двумерным). Диффузия поля плотности в случайном гидродинамическом потоке описывается уравнением неразрывности:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \mathbf{r}} \mathbf{V}(\mathbf{r}, t) \right) \rho(\mathbf{r}, t) = 0, \quad \rho(\mathbf{r}, 0) = \rho_0(\mathbf{r}). \quad (10)$$

Здесь через $\mathbf{V}(\mathbf{r}, t)$ обозначено поле скоростей частиц в гидродинамическом потоке $\mathbf{u}(\mathbf{r}, t)$, которое для малоинерционных частиц можно

описывать квазилинейным уравнением в частных производных

$$\left(\frac{\partial}{\partial t} + \mathbf{V}(\mathbf{r}, t) \frac{\partial}{\partial \mathbf{r}} \right) \mathbf{V}(\mathbf{r}, t) = -\lambda [\mathbf{V}(\mathbf{r}, t) - \mathbf{u}(\mathbf{r}, t)]. \quad (11)$$

В общем случае возможны неединственность решения этого уравнения, существование в решениях разрывов и т. п. Однако в асимптотическом случае малой инерционности частиц (параметр $\lambda \rightarrow \infty$), который и представляет для нас интерес, решение задачи будет единственным на разумном интервале времени.

Общая масса примеси при этом сохраняется в процессе эволюции, т. е.

$$M = M(t) = \int d\mathbf{R} \rho(\mathbf{R}, t) = \int d\mathbf{R} \rho_0(\mathbf{R}) = \text{const.}$$

Случайное поле $\rho(\mathbf{R}, t)$, как указывалось во введении, в каждой реализации напоминает сложный горный ландшафт, который, как и в обычной топографии, удобно анализировать на основе линий уровня (контуров постоянной плотности), определяемых равенством $\rho(\mathbf{R}, t) = \rho - \text{const.}$ Перемещение таких контуров в пространстве подобно диффузии системы частиц, рассмотренной выше, с теми же статистическими параметрами и существенно зависит от того, является ли поле скоростей соленоидальным или потенциальным.

В предположении гауссовости, дельта-коррелированности во времени, однородности и изотропности в пространстве случайного поля $\mathbf{V}(\mathbf{r}, t)$ легко получить для одноточечной плотности вероятностей

$$P(\mathbf{r}, t; \rho) = \langle \delta(\rho(\mathbf{r}, t) - \rho) \rangle$$

уравнение вида

$$\begin{aligned} \left(\frac{\partial}{\partial t} - D_0 \frac{\partial^2}{\partial \mathbf{r}^2} \right) P(\mathbf{r}, t; \rho) &= D^{\text{P}} \frac{\partial^2}{\partial \rho^2} \rho^2 P(\mathbf{r}, t; \rho), \\ P(\mathbf{r}, 0; \rho) &= \delta(\rho_0(\mathbf{r}) - \rho), \end{aligned} \quad (12)$$

где коэффициенты диффузии

$$\begin{aligned} D_0 &= \frac{1}{2} \int_0^\infty d\tau \langle \mathbf{V}(\mathbf{r}, t + \tau) \mathbf{V}(\mathbf{r}, t) \rangle = \frac{1}{2} \tau_{\text{v}} \langle \mathbf{V}^2(\mathbf{r}, t) \rangle, \\ D^{\text{P}} &= \int_0^\infty d\tau \left\langle \frac{\partial \mathbf{V}(\mathbf{r}, t + \tau)}{\partial \mathbf{r}} \frac{\partial \mathbf{V}(\mathbf{r}, t)}{\partial \mathbf{r}} \right\rangle = \tau_{\text{divv}} \left\langle \left(\frac{\partial \mathbf{V}(\mathbf{r}, t)}{\partial \mathbf{r}} \right)^2 \right\rangle. \end{aligned} \quad (13)$$

Здесь τ_v и $\tau_{\text{div}v}$ — временные радиусы корреляции для случайных полей $\mathbf{V}(\mathbf{r}, t)$ и $\text{div} \mathbf{V}(\mathbf{r}, t)$. Отметим, что именно отличие от нуля коэффициента диффузии D^p и характеризует кластеризацию поля плотности примеси. Что касается коэффициента пространственной диффузии D_0 , то он совпадает с пространственным коэффициентом диффузии частиц и не играет никакой роли в кластеризации. В случае пространственной однородности поля плотности $\rho(\mathbf{r}, t)$ этот член вообще отсутствует в уравнении (12).

Очевидно, что даже в случае бездивергентного гидродинамического потока ($\text{div} \mathbf{u}(\mathbf{r}, t) = 0$) будет осуществляться кластеризация поля плотности (при учете, конечно, инерционности примеси), так как $\text{div} \mathbf{V}(\mathbf{r}, t) \neq 0$. И в этом случае невозможно использовать приближение дельта-коррелированности во времени случайного поля $\mathbf{u}(\mathbf{r}, t)$. В пределе же безынерционности примеси приближение дельта-коррелированности поля $\mathbf{u}(\mathbf{r}, t)$ может быть использовано, но кластеризация в этом случае осуществляться не будет.

Для соленоидального поля скоростей эволюция контуров совпадает с рис. 2. В этом случае сохраняются как площадь, ограниченная контурами, так и масса примеси, сосредоточенная в этих областях. Кроме того, сохраняется и общее число замкнутых контуров, которые не могут исчезать и порождаться в среде в процессе эволюции во времени. В этом случае, как видно из рис. 2, лишь линии уровня приобретают «фрактальный» характер, т. е. появляется все большая их «изрезанность», градиенты обостряются и этот процесс идет в сторону уменьшения пространственных масштабов вплоть до масштабов, на которых необходимо уже учитывать эффекты, связанные с молекулярной диффузией. Статистический анализ позволяет в этом случае вычислить среднее значение длины линии уровня $\langle l(t) \rangle = l_0 e^{D^p t}$, которое экспоненциально растет со временем.

При наличии у поля скоростей потенциальной составляющей все величины, перечисленные выше, эволюционируют во времени. Так, в частности, при $\tau \gg 1$, где $\tau = D^p t$, средняя площадь областей, где плотность выше заданного уровня ρ (т. е. $\rho(\mathbf{R}, t) > \rho$), убывает со временем по закону

$$\langle S(t, \rho) \rangle \approx \frac{1}{\sqrt{\pi \tau \rho}} e^{-\tau/4} \int d\mathbf{r} \sqrt{\rho_0(\mathbf{r})},$$

в то время как заключенная в этих областях средняя масса примеси

$$\langle M(t, \rho) \rangle \approx M - \sqrt{\frac{\rho}{\pi \tau}} e^{-\tau/4} \int d\mathbf{r} \sqrt{\rho_0(\mathbf{r})}$$

монотонно стремится к полной ее массе:

$$M = \int \rho_0(\mathbf{R})d\mathbf{R}.$$

Таким образом, подтверждается сделанный ранее вывод о том, что поле плотности со временем стремится собраться в кластеры (компактные области повышенной концентрации, окруженные разреженными областями) в случайном поле скоростей с потенциальной составляющей.

Динамику формирования кластеров можно проиллюстрировать на примере, когда первоначально примесь равномерно распределена по плоскости: $\rho_0(\mathbf{R}) = \rho_0 = \text{const}$. При этом удельная площадь областей (приходящаяся на единицу площади), внутри которых $\rho(\mathbf{R}, t) > \rho$, асимптотически убывает во времени, независимо от отношения ρ/ρ_0 , в то время как внутри них собирается практически вся масса примеси.

Характер же временной эволюции образования кластерной структуры существенно зависит от отношения ρ/ρ_0 . Так если $\rho/\rho_0 < 1$, то в начальный момент времени $s(0, \rho) = 1$ и $m(0, \rho) = 1$. Далее образуются небольшие области, где $\rho(\mathbf{R}, t) < \rho$, содержащие незначительную часть общей массы. С течением времени эти области быстро увеличиваются, а их масса перетекает в кластерные области.

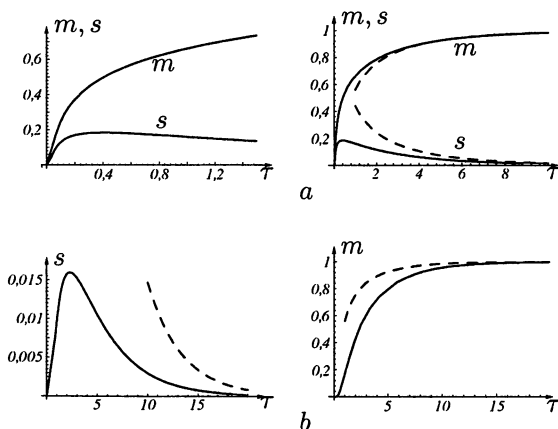


Рис. 4. Динамика образования кластеров для $\rho/\rho_0 = 1,5$ (а) и $\rho/\rho_0 = 10$ (б).

В другом, наиболее интересном случае $\rho/\rho_0 > 1$ в начальный момент времени как $s(0, \rho) = 0$, так и $m(0, \rho) = 0$. Сначала образуются небольшие кластерные области, где $\rho(\mathbf{R}, t) > \rho$, которые затем интенсивно втягивают в себя значительную часть общей массы. В дальнейшем площади этих областей с течением времени уменьшаются, а содержащаяся в них масса увеличивается до полной массы (рис. 4 для $\rho/\rho_0 = 1,5$ (а) и $\rho/\rho_0 = 10$ (б)).

1.4. Локализация плоских волн в слоистых случайных средах

Рассмотрим простейшую статистическую стационарную задачу о волнах в слоистых случайных средах. Пусть слой хаотически неоднородной среды занимает часть пространства $L_0 < x < L$ и из области $x > L$ на нее падает плоская волна $u_0(x) = e^{-ik(x-L)}$. Из-за наличия неоднородностей возникают отраженная от слоя волна с коэффициентом отражения $R = u(L) - 1$ и выходящая из слоя волна с коэффициентом прохождения $T = u(L_0)$ (рис. 5, а).

При отсутствии неоднородностей очевидно, что $R = 0$ и $T = 1$. Внутри слоя волновое поле описывается уравнением Гельмгольца

$$\frac{d^2}{dx^2} u(x) + k^2[1 + \varepsilon(x)]u(x) = 0, \quad (14)$$

где функция $\varepsilon(x)$, которую мы считаем случайной, описывает неоднородности среды. Данное уравнение следует дополнить краевыми условиями — условиями непрерывности волнового поля $u(x)$ и производной $du(x)/dx$ на границах слоя и условиями излучения на этих границах, которые можно записать в виде

$$u(L) + \frac{i}{k} \frac{du(x)}{dx} \Big|_{x=L} = 2, \quad u(L_0) - \frac{i}{k} \frac{du(x)}{dx} \Big|_{x=L_0} = 0. \quad (15)$$

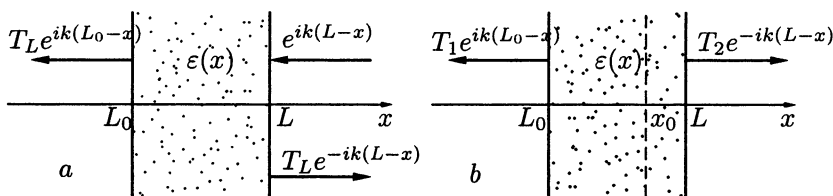


Рис. 5. Падение плоской волны на слой среды (а) и источник внутри среды (б).

Аналогичным образом поле точечного источника, расположенного в точке x_0 (рис. 5, b), описывается краевой задачей:

$$\frac{d^2}{dx^2}G(x; x_0) + k^2[1 + \varepsilon(x)]G(x; x_0) = 2ik\delta(x - x_0),$$

$$G(L; x_0) + \frac{i}{k} \frac{dG(x; x_0)}{dx} \Big|_{x=L} = 0, \quad G(L_0; x_0) - \frac{i}{k} \frac{dG(x; x_0)}{dx} \Big|_{x=L_0} = 0,$$

и решение задачи о падении волны на слой среды соответствует положению источника на границе $x_0 = L$, т. е. $u(x) = G(x; L)$.

Таким образом, мы имеем краевую задачу для комплексного волнового поля $u(x)$, состоящую в нахождении статистических характеристик как коэффициентов отражения и прохождения, так и интенсивности поля внутри слоя среды $I(x) = |u(x)|^2$, считая известными статистические характеристики функции $\varepsilon(x)$. Отметим, что коэффициенты отражения и прохождения связаны между собой равенством $|R|^2 + |T|^2 = 1$, являющимся следствием сохранения плотности потока энергии при отсутствии поглощения волны в среде.

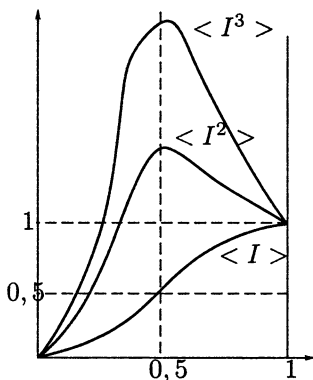


Рис. 6. Схематическое поведение моментов интенсивности волнового поля в задаче о падении волны на слой среды (стохастический волновой параметрический резонанс).

условий (которые «зжимают» значения волнового поля на границах слоя неоднородной среды) экспоненциальный рост поля может осуществляться только в глубине слоя, вдали от его границ.

На рис. 6 схематически изображено распределение моментов интенсивности волнового поля $I(x) = |u(x)|^2$, нормированных на их

Рассматриваемое волновое уравнение по своей форме напоминает уравнение для осциллятора с изменяющейся собственной частотой (если заменить пространственную переменную x на переменную t — время), для которого, как хорошо известно, имеет место явление *параметрического резонанса* на частотах $2k/n$ ($n = 1, 2, \dots$). Учитывая, что случайная функция $\varepsilon(x)$ имеет составляющие всех частот, включая и указанные, ясно, что и для рассматриваемой задачи должно иметь место аналогичное явление, которое может быть названо *явлением стохастического волнового параметрического резонанса*. Однако в силу краевых

значения на границе $x = L$, иллюстрирующее явление стохастического волнового параметрического резонанса. Из рисунка видно, что максимальное значение достигается примерно в середине слоя (за исключением средней интенсивности, которая монотонно убывает).

Для конкретных вычислений статистических характеристик решения волновой задачи обычно используется модель флуктуаций функции $\varepsilon(x)$ в виде гауссовского дельта-коррелированного случайного процесса (процесса *белого шума*), хотя основная принципиальная часть результатов и мало зависит от модели среды. Использование конкретной модели позволяет количественно вычислить основные параметры задачи.

Статистический анализ решения задачи показывает, что для достаточно толстого слоя среды, а именно $D(L - L_0) \gg 1$ (где величина D связана со статистическими характеристиками функции $\varepsilon(x)$), с вероятностью единица $|T| \rightarrow 0$ и, следовательно, $|R| \rightarrow 1$, т. е. полупространство ($L_0 \rightarrow -\infty$) случайно-неоднородной среды полностью отражает падающую волну за счет многократного преотражения в среде, т. е. имеет место *динамическая локализация волнового поля* в этом слое. При этом типичная реализация интенсивности волны в среде описывается экспоненциально спадающей кривой $I^*(x) = 2e^{-D(L-x)}$, где величина $l_{loc} = 1/D$ называется *длиной локализации* и определяет пространственный масштаб затухания интенсивности волнового поля в среде в отдельных его реализациях.

На рис. 7 приведены две реализации интенсивности волнового поля (белые и черные кружки), полученные численным моделированием рассматриваемой задачи. Из рисунка видно, что имеются большие выбросы интенсивности волнового поля на фоне общего быстрого спада, которое и можно отождествить с проявлением свойства динамической локализации данных реализаций.

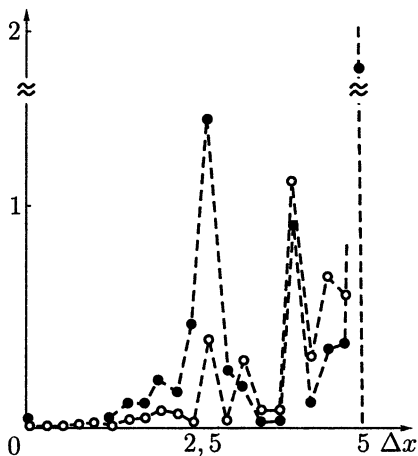


Рис. 7. Численное моделирование динамической локализации для двух реализаций неоднородностей среды.

Однако среднее значение интенсивности волнового поля для полупространства случайной среды $\langle I(L-x) \rangle = 2$, а более высокие моменты, нормированные на их значения на границе слоя, описываются выражением $\langle I^n(L-x) \rangle = e^{Dn(n-1)(L-x)}$, т. е. интенсивность волнового поля имеет логарифмически нормальное распределение вероятностей, моменты экспоненциально растут вглубь среды и формирование статистики осуществляется за счет больших выбросов относительно кривой типичной реализации, что означает отсутствие *статистической локализации волнового поля*.

1.5. Каустическая структура волнового поля в случайно-неоднородной среде

Распространение волн в случайно-неоднородной трехмерной среде обычно рассматривается на основе скалярного параболического уравнения, которое справедливо для описания распространения волн в среде с крупномасштабными трехмерными неоднородностями и описывает рассеяние волны на малые углы:

$$\frac{\partial}{\partial x} U(x, \mathbf{R}) = \frac{i}{2k} \Delta_{\mathbf{R}} U(x, \mathbf{R}) + \frac{ik}{2} \varepsilon(x, \mathbf{R}) U(x, \mathbf{R}), \quad U(0, \mathbf{R}) = U_0(\mathbf{R}).$$

Здесь x — координата в направлении распространения волны, \mathbf{R} — координаты в поперечной плоскости, а $\varepsilon(x, \mathbf{R})$ — отклонение диэлектрической проницаемости от единицы. Это уравнение является, естественно, приближенным.

Если ввести амплитуду и фазу волнового поля по формуле

$$U(x, \mathbf{R}) = A(x, \mathbf{R}) \exp \{iS(x, \mathbf{R})\},$$

то уравнение переноса для интенсивности волнового поля $I(x, \mathbf{R}) = |u(x, \mathbf{R})|^2$ можно записать в виде

$$\frac{\partial}{\partial x} I(x, \mathbf{R}) + \frac{1}{k} \nabla_{\mathbf{R}} \{ \nabla_{\mathbf{R}} S(x, \mathbf{R}) I(x, \mathbf{R}) \} = 0, \quad I(0, \mathbf{R}) = I_0(\mathbf{R}). \quad (16)$$

Уравнение (16) по форме совпадает с уравнением (10) и, следовательно, его можно трактовать как уравнение переноса консервативной примеси в потенциальном поле скоростей. Однако только в *приближении геометрической оптики* эту примесь можно считать пассивной. В общем же случае при учете дифракционных эффектов она является активной.

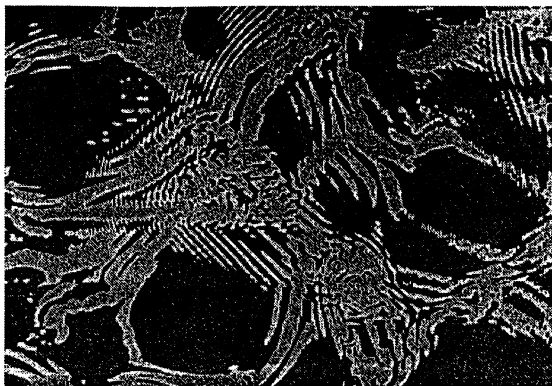


Рис. 8. Поперечное сечение лазерного пучка в турбулентной среде

Реализации поля интенсивности, согласно предыдущему разделу, должны иметь кластерный характер, который проявляется в виде каустических структур. Так, на рис. 8 приведена фотография поперечного сечения лазерного пучка, распространяющегося в турбулентной среде (лабораторный эксперимент); ясно видно возникновение каустической структуры волнового поля. На рис. 9 приведена фотография бассейна с четко выраженной каустической структурой волнового поля на его дне. Подобные структуры возникают

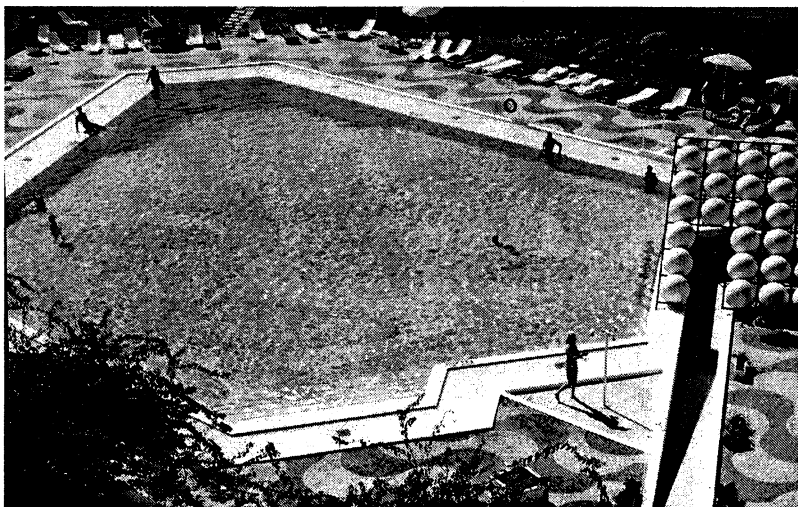


Рис. 9. Каустики в бассейне

при преломлении и отражении света взволнованной водной поверхностью, что соответствует рассеянию на так называемом *фазовом экране*. Отметим, что для плоской падающей волны, соответствующей начальному условию $U(0, \mathbf{R}) = U_0 - \text{const}$, как среднее значение интенсивности волнового поля для статистически однородной среды $\varepsilon(x, \mathbf{R})$ постоянно и не зависит от переменной \mathbf{R} ($\langle I(x, \mathbf{R}) \rangle = I_0$), так и корреляционная функция волнового поля имеет вид

$$\langle U(x, \mathbf{R}_1)U^*(x, \mathbf{R}_2) \rangle = I_0 e^{-k^2 D(\mathbf{R}_1 - \mathbf{R}_2)x},$$

где функция $D(\mathbf{R})$ связана со статистическими характеристиками поля $\varepsilon(x, \mathbf{R})$. Средняя интенсивность волнового поля и его корреляционная функция в этом случае не зависят от явления дифракции вообще, в отличие от примеров, приведенных на рис. 8 и 9.

2. Элементы математического аппарата для описания когерентных явлений

2.1. Статистическая топография случайных процессов и полей

Остановимся теперь кратко на математическом аппарате, позволяющем описать когерентные явления.

Прежде всего обсудим понятие типичной реализации случайного процесса $z(t)$, которое характеризует основные особенности поведения отдельной реализации процесса в целом на всем интервале времен.

2.1.1. Типичная реализация случайного процесса

Пусть $z(t)$ — случайный процесс по переменной t , которую для определенности будем считать временем. Статистические характеристики процесса $z(t)$ в фиксированный момент времени t полностью описываются ее плотностью вероятностей $p(z; t)$, параметрически зависящей от времени t , или интегральной функцией распределения, определяющей вероятность того, что в фиксированный момент времени t значение процесса $z(t) < z$ будет

$$F(z; t) = P(z(t) < z) = \int_{-\infty}^z dz' p(z'; t). \quad (17)$$

Назовем типичной реализацией случайного процесса $z(t)$ детерминированную кривую $z^*(t)$, которая является *медианой интегральной функции распределения* (17) и определяется как решение алгебраического уравнения $F(z^*(t); t) = 1/2$.

Основанием для этого является свойство медианы, заключающееся в том, что для любого интервала времени (t_1, t_2) случайный процесс $z(t)$ как бы «обвивает» кривую $z^*(t)$ таким образом, что среднее время, в течение которого выполняется неравенство $z(t) > z^*(t)$, совпадает со средним временем, в течение которого выполняется обратное неравенство $z(t) < z^*(t)$ (рис. 10), т. е.

$$\langle T_{z(t) > z^*(t)} \rangle = \langle T_{z(t) < z^*(t)} \rangle = \frac{1}{2} (t_2 - t_1).$$

Кривая $z^*(t)$ может, естественно, существенно отличаться от любой конкретной реализации процесса $z(t)$ и не описывает величину возможных выбросов. Таким образом, кривая типичной реализации $z^*(t)$ случайного процесса $z(t)$, полученная с помощью одновременной плотности вероятностей, определена, тем не менее, на всем интервале времени $t \in (0, \infty)$.

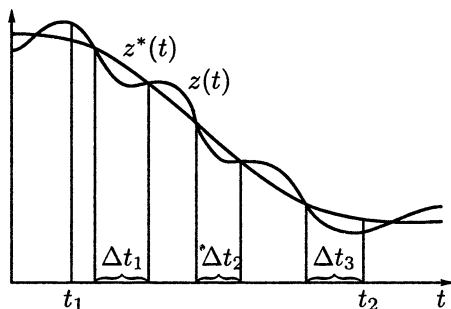


Рис. 10. К определению типичной реализации случайного процесса

Для конкретных типов случайных процессов можно получить дополнительную информацию, характеризующую уже выбросы относительно этой кривой.

2.1.2. Случайные поля

В статистической топографии случайных полей основным объектом изучения, как и в обычной топографии горных массивов, является система контуров — линий уровня (в двумерном случае) или

поверхностей (в трехмерном случае) постоянных значений, определяемых равенством $f(\mathbf{r}, t) = f = \text{const}$.

Для анализа системы контуров (для простоты в данном разделе будем говорить о двумерном случае ($\mathbf{r} = \mathbf{R}$)) удобно ввести сингулярную индикаторную функцию

$$\varphi(t, \mathbf{R}; f) = \delta(f(\mathbf{R}, t) - f), \quad (18)$$

сосредоточенную на них.

Через функцию (18) выражаются, например, такие величины, как общая площадь областей, ограниченных линиями уровня, где $f(\mathbf{R}, t) > f$:

$$S(t; f) = \int \theta(f(\mathbf{R}, t) - f) d\mathbf{R} = \int_f^\infty df' \int d\mathbf{R} \varphi(t, \mathbf{R}; f')$$

и общая «масса» поля, заключенная в этих областях:

$$M(t; f) = \int f(\mathbf{R}, t) \theta(f(\mathbf{R}, t) - f) d\mathbf{R} = \int_f^\infty f' df' \int d\mathbf{R} \varphi(t, \mathbf{R}; f').$$

Среднее значение индикаторной функции (18) по ансамблю реализаций определяет одновременную во времени и одноточечную в пространстве плотность вероятностей $P(t, \mathbf{R}; f) = \langle \varphi(t, \mathbf{R}; f) \rangle$, и, следовательно, средние по ансамблю реализации значения всех выражений непосредственно определяются этой плотностью вероятностей.

Дополнительную информацию о детальной структуре поля $f(\mathbf{R}, t)$ можно получить, включив в рассмотрение его пространственный градиент $\mathbf{p}(\mathbf{R}, t) = \nabla f(\mathbf{R}, t)$. Так, например, величина

$$l(t; f) = \int d\mathbf{R} |\mathbf{p}(\mathbf{R}, t)| \delta(f(\mathbf{R}, t) - f) = \oint dl$$

описывает общую длину контуров.

Включение в рассмотрение пространственных производных второго порядка позволяет оценить общее число контуров $f(\mathbf{R}, t) = f = \text{const}$ с помощью приближенной (с точностью до незамкнутых линий) формулы:

$$\begin{aligned} N(t; f) &= N_{\text{in}}(t; f) - N_{\text{out}}(t; f) = \\ &= \frac{1}{2\pi} \int d\mathbf{R} \kappa(t, \mathbf{R}; f) |\mathbf{p}(\mathbf{R}, t)| \delta(f(\mathbf{R}, t) - f), \end{aligned}$$

где $N_{\text{in}}(t; f)$, $N_{\text{out}}(t; f)$ — числа контуров, для которых вектор \mathbf{p} направлен по внутренней и внешней нормали соответственно, а $\kappa(t, \mathbf{R}; f)$ — кривизна линии уровня.

Отметим, что для пространственно однородного поля $f(\mathbf{R}, t)$, когда соответствующие одноточечные плотности вероятностей $P(t, \mathbf{R}; f)$ и $P(t, \mathbf{R}; f, \mathbf{p})$ не зависят от \mathbf{R} , статистические средние всех выражений (без интегрирования по \mathbf{R}) будут описывать соответствующие удельные (приходящиеся на единицу площади) значения этих величин, если они, конечно, существуют.

2.2. Уравнение Фоккера — Планка

Пусть векторная функция $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots, x_n(t)\}$ удовлетворяет динамическому уравнению

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{v}(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad (19)$$

где $v_i(\mathbf{x}, t)$ ($i = 1, \dots, n$) — детерминированные функции, а $f_i(\mathbf{x}, t)$ — случайные функции $(n + 1)$ переменной, обладающие следующими свойствами:

а) $f_i(\mathbf{x}, t)$ — гауссово случайное поле в $(n + 1)$ -мерном пространстве (\mathbf{x}, t) ;

б) $\langle f_i(\mathbf{x}, t) \rangle = 0$.

Будем для определенности считать, что t — временная координата, а \mathbf{x} — пространственная.

Статистические характеристики поля $f_i(\mathbf{x}, t)$ полностью описываются заданием его корреляционного тензора:

$$B_{ij}(\mathbf{x}, t; \mathbf{x}', t') = \langle f_i(\mathbf{x}, t) f_j(\mathbf{x}', t') \rangle.$$

Так как уравнение (19) является уравнением первого порядка с начальным условием, то для него выполняется условие динамической причинности

$$\frac{\delta}{\delta f_j(\mathbf{x}', t')} x_i(t) = 0 \quad \text{при } t' < t_0 \text{ и } t' > t, \quad (20)$$

т. е. его решение $\mathbf{x}(t)$ функционально зависит лишь от предшествующих по t значений $f_j(\mathbf{x}, t')$ из интервала $t_0 \leq t' \leq t$. При этом для вариационной производной имеем равенство

$$\frac{\delta}{\delta f_j(\mathbf{x}', t - 0)} x_i(t) = \delta_{ij} \delta(\mathbf{x}(t) - \mathbf{x}'). \quad (21)$$

Однако может существовать статистическая связь между $\mathbf{x}(t)$ и последующими значениями $f_j(\mathbf{x}, t'')$, где $t'' > t$, так как такие значения $f_j(\mathbf{x}, t'')$ коррелированы со значениями $f_j(\mathbf{x}, t')$ при $t' \leq t$. Ясно, что корреляция функции $\mathbf{x}(t)$ с последующими значениями $f_j(\mathbf{x}, t'')$ заметна лишь при $t'' - t \leq \tau_0$, где τ_0 — радиус корреляции поля $\mathbf{f}(\mathbf{x}, t)$ по переменной t .

Для достаточно большого класса реальных физических процессов характерный временной масштаб изменения функции $\mathbf{x}(t)$ имеет величину порядка $T \gg \tau_0$, и в этом случае существует малый параметр τ_0/T , который может быть использован для построения приближенного решения.

В первом приближении по этому параметру малости можно рассматривать асимптотику при $\tau_0 \rightarrow 0$. При этом значения $\mathbf{x}(t')$ при $t' < t$ будут не только функционально, но и статистически независимы от значений $\mathbf{f}(\mathbf{x}, t'')$ при $t'' > t$. Это приближение эквивалентно замене корреляционного тензора B_{ij} на некоторый эффективный тензор, определяемый равенством

$$B_{ij}^{\text{eff}}(\mathbf{x}, t; \mathbf{x}', t') = 2\delta(t - t')F_{ij}(\mathbf{x}, \mathbf{x}'; t). \quad (22)$$

Величина $F_{ij}(\mathbf{x}, \mathbf{x}'; t)$ при этом определяется из условия равенства интегралов от $B_{ij}(\mathbf{x}, t; \mathbf{x}', t')$ и $B_{ij}^{\text{eff}}(\mathbf{x}, t; \mathbf{x}', t')$ по t'

$$F_{ij}(\mathbf{x}, \mathbf{x}'; t) = \frac{1}{2} \int_{-\infty}^{\infty} dt' B_{ij}(\mathbf{x}, t; \mathbf{x}', t'),$$

что и соответствует переходу к гауссовому дельта-коррелированному по времени t случайному полю.

Введем индикаторную функцию

$$\varphi(\mathbf{x}, t) = \delta(\mathbf{x}(t) - \mathbf{x}),$$

где $\mathbf{x}(t)$ — решение уравнения (19), удовлетворяющую уравнению Лиувилля

$$\frac{\partial}{\partial t} \varphi(\mathbf{x}, t) = -\frac{\partial}{\partial \mathbf{x}} [\mathbf{v}(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t)] \varphi(\mathbf{x}, t). \quad (23)$$

При этом

$$\frac{\delta}{\delta f_j(\mathbf{x}', t-0)} \varphi(\mathbf{x}, t) = -\frac{\partial}{\partial x_j} \{ \delta(\mathbf{x} - \mathbf{x}') \varphi(\mathbf{x}, t) \}. \quad (24)$$

Уравнение для плотности вероятностей решения уравнения (19)

$$P(\mathbf{x}, t) = \langle \varphi(\mathbf{x}, t) \rangle = \langle \delta(\mathbf{x}(t) - \mathbf{x}) \rangle$$

получим, усредняя уравнение (23) по ансамблю реализаций поля $\mathbf{f}(\mathbf{x}, t)$:

$$\frac{\partial}{\partial t} P(\mathbf{x}, t) = -\frac{\partial}{\partial \mathbf{x}} \mathbf{v}(\mathbf{x}, t) P(\mathbf{x}, t) - \frac{\partial}{\partial \mathbf{x}} \langle \mathbf{f}(\mathbf{x}, t) \varphi(\mathbf{x}, t) \rangle. \quad (25)$$

Уравнение (25) можно переписать в виде

$$\begin{aligned} & \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \mathbf{v}(\mathbf{x}, t) \right) P(\mathbf{x}, t) = \\ & = -\frac{\partial}{\partial x_i} \int d\mathbf{x}' \int_{t_0}^t dt' B_{ij}(\mathbf{x}, t; \mathbf{x}', t') \left\langle \frac{\delta \varphi(\mathbf{x}, t)}{\delta f_j(\mathbf{x}', t')} \right\rangle. \end{aligned} \quad (26)$$

Здесь мы воспользовались формулой Фурутцу — Новикова

$$\begin{aligned} & \langle f_k(\mathbf{x}, t) R[t; \mathbf{f}(\mathbf{y}, \tau)] \rangle = \\ & = \int d\mathbf{x}' \int dt' B_{kl}(\mathbf{x}, t; \mathbf{x}', t') \left\langle \frac{\delta R[t; \mathbf{f}(\mathbf{y}, \tau)]}{\delta f_l(\mathbf{x}', t')} \right\rangle, \end{aligned} \quad (27)$$

справедливой для корреляции гауссова случайного поля $\mathbf{f}(\mathbf{x}, t)$ с произвольным функционалом $R[t; \mathbf{f}(\mathbf{y}, \tau)]$ от него, и условием динамической причинности (20).

Уравнение (26) показывает, что одновременная плотность вероятностей решения $\mathbf{x}(t)$ в момент времени t определяется функциональной зависимостью решения $\mathbf{x}(t)$ от поля $\mathbf{f}(\mathbf{x}', t)$ для всех времен в интервале (t_0, t) .

В общем случае плотность вероятностей $P(\mathbf{x}, t)$ не описывается замкнутым уравнением. Если же для корреляционной функции поля $\mathbf{f}(\mathbf{x}, t)$ воспользоваться приближением (22), то возникнут члены, связанные со значениями $\delta \varphi[\mathbf{x}, t; \mathbf{f}(\mathbf{y}, \tau)] / \delta f_j(\mathbf{x}', t')$ при совпадающих временных аргументах $t' = t - 0$, которые, согласно равенству (24), выражаются непосредственно через саму величину $\varphi[\mathbf{x}, t; \mathbf{f}(\mathbf{y}, \tau)]$. Таким образом мы приходим к замкнутому уравнению Фоккера — Планка

$$\begin{aligned} & \frac{\partial}{\partial t} P(\mathbf{x}, t) + \frac{\partial}{\partial x_k} [v_k(\mathbf{x}, t) + A_k(\mathbf{x}, t)] P(\mathbf{x}, t) = \\ & = \frac{\partial^2}{\partial x_k \partial x_l} [F_{kl}(\mathbf{x}, \mathbf{x}; t) P(\mathbf{x}, t)], \end{aligned} \quad (28)$$

где

$$A_k(\mathbf{x}, t) = \left. \frac{\partial}{\partial x'} F_{kl}(\mathbf{x}, \mathbf{x}'; t) \right|_{\mathbf{x}'=\mathbf{x}}.$$

Уравнение (28) следует решать с начальным условием $P(\mathbf{x}, t_0) = \delta(\mathbf{x} - \mathbf{x}_0)$ или же с начальным условием более общего вида $P(\mathbf{x}, t_0) = W_0(\mathbf{x})$, если начальные условия так же случайны, но статистически независимы от поля $f(\mathbf{x}, t)$.

Уравнение Фоккера — Планка (28) — уравнение в частных производных, и его дальнейший анализ существенно зависит от формулировки краевых условий по \mathbf{x} , которые формулируются для анализа конкретных задач.

2.2.1. Плотность вероятностей перехода

Вернемся теперь к динамической системе (19) и рассмотрим m -временную плотность вероятностей

$$P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m) = \langle \delta(\mathbf{x}(t_1) - \mathbf{x}_1) \dots \delta(\mathbf{x}(t_m) - \mathbf{x}_m) \rangle, \quad (29)$$

относящуюся к m различным моментам времени $t_1 < t_2 < \dots < t_m$. Дифференцируя (29) по времени t_m и используя затем динамическое уравнение (19), условие динамической причинности (20), определение функции $F_{kl}(\mathbf{x}, \mathbf{x}'; t)$, и формулу Фурутцу — Новикова (27), можно получить уравнение, аналогичное уравнению Фоккера — Планка (28):

$$\begin{aligned} & \frac{\partial}{\partial t_m} P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m) + \\ & + \sum_{k=1}^n \frac{\partial}{\partial x_{mk}} [v_k(\mathbf{x}_m, t_m) + A_k(\mathbf{x}_m, t_m)] P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m) = \\ & = \sum_{k=1}^n \sum_{l=1}^n \frac{\partial^2}{\partial x_{mk} \partial x_{ml}} [F_{kl}(\mathbf{x}_m, \mathbf{x}_m; t_m) P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m)]. \quad (30) \end{aligned}$$

Здесь суммирование по индексу m не производится. Начальное условие к (30) можно найти из формулы (29). Полагая $t_m = t_{m-1}$ в (29), получаем

$$\begin{aligned} & P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_{m-1}) = \\ & = \delta(\mathbf{x}_m - \mathbf{x}_{m-1}) P_{m-1}(\mathbf{x}_1, t_1; \dots; \mathbf{x}_{m-1}, t_{m-1}). \quad (31) \end{aligned}$$

Решение уравнения (30) можно искать в виде

$$P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m) = p(\mathbf{x}_m, t_m | \mathbf{x}_{m-1}, t_{m-1}) P_{m-1}(\mathbf{x}_1, t_1; \dots; \mathbf{x}_{m-1}, t_{m-1}). \quad (32)$$

Так как все дифференциальные операции в (30) относятся к t_m и \mathbf{x}_m , то подставляя (32) в (30) и (31), находим следующее уравнение для плотности вероятностей перехода $p(\mathbf{x}, t | \mathbf{x}_0, t_0) = \langle \delta(\mathbf{x}(t) - \mathbf{x}) | \mathbf{x}(t_0) = \mathbf{x}_0 \rangle$:

$$\begin{aligned} \frac{\partial}{\partial t} p(\mathbf{x}, t | \mathbf{x}_0, t_0) + \frac{\partial}{\partial x_k} [v_k(\mathbf{x}, t) + A_k(\mathbf{x}, t)] p(\mathbf{x}, t | \mathbf{x}_0, t_0) = \\ = \frac{\partial^2}{\partial x_k \partial x_l} [F_{kl}(\mathbf{x}, \mathbf{x}; t) p(\mathbf{x}, t | \mathbf{x}_0, t_0)], \\ p(\mathbf{x}, t | \mathbf{x}_0, t_0) |_{t \rightarrow t_0} = \delta(\mathbf{x} - \mathbf{x}_0). \end{aligned} \quad (33)$$

В уравнении (33) мы обозначили переменные \mathbf{x}_m, t_m через \mathbf{x}, t , а переменные $\mathbf{x}_{m-1}, t_{m-1}$ через \mathbf{x}_0, t_0 .

Применяя формулу (32) $(m-1)$ раз, получаем соотношение

$$P_m(\mathbf{x}_1, t_1; \dots; \mathbf{x}_m, t_m) = p(\mathbf{x}_m, t_m | \mathbf{x}_{m-1}, t_{m-1}) \dots p(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1) P(\mathbf{x}_1, t_1), \quad (34)$$

где $P(\mathbf{x}_1, t_1)$ — плотность вероятностей, определяемая уравнением (28) и относящаяся к одному моменту времени t_1 . Равенство (34) выражает многовременную плотность вероятностей через произведение плотностей вероятностей перехода и означает, что случайный процесс $\mathbf{x}(t)$ является марковским.

2.2.2. Об условиях применимости уравнения Фоккера — Планка

Для оценки границ применимости уравнения Фоккера — Планка необходимо учитывать конечность радиуса корреляции τ_0 поля $\mathbf{f}(\mathbf{x}, t)$ по временной координате. Для статистически однородного и изотропного в пространстве и стационарного во времени поля $\mathbf{f}(\mathbf{x}, t)$ его дисперсия и временной радиус корреляции τ_0 определяются равенствами

$$\sigma_{\mathbf{f}}^2 = \langle \mathbf{f}^2(\mathbf{x}, t) \rangle, \quad \tau_0 = \frac{1}{\sigma_{\mathbf{f}}^2} \int_0^{\infty} d\tau \langle \mathbf{f}(\mathbf{x}, t + \tau) \mathbf{f}(\mathbf{x}, t) \rangle.$$

При этом условие малости параметра τ_0/T является необходимым, но, вообще говоря, недостаточным для возможности описывать статистические характеристики решения уравнения (19) на основе приближения дельта-коррелированного случайного поля, которому соответствует уравнение Фоккера — Планка. Для каждой конкретной задачи необходимо проводить более детальные исследования. Далее мы приведем физически более наглядный метод, называемый *диффузионным приближением*, также приводящий к марковости решения уравнения (19), но учитывающий, в определенной мере, конечность временного радиуса корреляции.

Здесь же подчеркнем, что приближение дельта-коррелированного случайного поля не означает формальной замены случайного поля $f(x, t)$ в (19) на случайное поле с корреляционной функцией (22). Это приближение соответствует построению асимптотического разложения при стремлении временного радиуса корреляции τ_0 поля $f(x, t)$ к нулю. И при таком предельном переходе точные средние величины переходят в выражения, полученные с помощью формальной замены корреляционного тензора поля $f(x, t)$ на эффективный тензор (22).

Рассмотрим теперь несколько простейших аппроксимаций реальных случайных процессов, широко используемых в качестве моделей для параметров задач.

2.2.3. Гауссовский случайный процесс *белого шума*

Первая аппроксимация — это представление стационарного во времени гауссового случайного процесса $\xi(t)$ с нулевым средним значением и корреляционной функцией $B_\xi(t - t') = \langle \xi(t)\xi(t') \rangle$ в виде

$$B_\xi(t) \rightarrow B_\xi^{\text{eff}}(t) = 2\sigma_\xi^2\tau_0\delta(t), \quad (35)$$

где σ_ξ^2 — дисперсия реального процесса $\xi(t)$, а τ_0 — его временной радиус корреляции

$$\tau_0 = \frac{1}{\sigma_\xi^2} \int_0^\infty dt B_\xi(t).$$

Такая аппроксимация справедлива, если величина τ_0 много меньше любого временного масштаба в рассматриваемой задаче.

Рассмотрим теперь простейшие случайные процессы, построенные на основе процесса белого шума.

2.2.4. Винеровский случайный процесс

Интеграл от процесса $\xi(t)$ определяет *Винеровский случайный процесс (процесс броуновского движения)*:

$$w(t) = \int_0^t d\tau \xi(\tau).$$

Это гауссовский, нестационарный случайный процесс с характеристиками

$$\langle w(t) \rangle = 0, \quad \langle w(t)w(t') \rangle = 2\sigma_\xi^2 \tau_0 \min\{t, t'\}.$$

2.2.5. Логарифмически нормальный процесс

Определим логнормальный случайный процесс по формуле

$$y(t; \alpha) = e^{-\alpha t + w(t)} = \exp \left\{ -\alpha t + \int_0^t d\tau \xi(\tau) \right\}, \quad (36)$$

где $\xi(t)$ — гауссов процесс белого шума (35). Этот процесс описывается стохастическим уравнением

$$\frac{d}{dt} y(t; \alpha) = \{-\alpha + \xi(t)\} y(t; \alpha), \quad y(0; \alpha) = 1.$$

Логнормальный процесс $y(t; \alpha)$ является марковским процессом, и его одновременная плотность вероятностей описывается уравнением Фоккера — Планка:

$$\begin{aligned} \frac{\partial}{\partial t} P(y, t; \alpha) &= \alpha \frac{\partial}{\partial y} y P(y, t; \alpha) + D \frac{\partial}{\partial y} y \frac{\partial}{\partial y} y P(y, t; \alpha), \\ P(y, 0; \alpha) &= \delta(y - 1), \end{aligned}$$

решение которого имеет вид

$$P(y, t; \alpha) = \frac{1}{2y\sqrt{\pi Dt}} \exp \left\{ -\frac{\ln^2 (ye^{\alpha t})}{4Dt} \right\}, \quad (37)$$

где параметр $D = \sigma^2 \tau_0$. Характерной особенностью этого распределения является появление длинного пологого «хвоста» при

$\tau = 1$, означающего усиление роли больших выбросов процесса $y(t; \alpha)$ в формировании одновременной статистики. Соответственно, интегральная функция распределения определяется выражением

$$F(y, t; \alpha) = \Phi \left(\frac{1}{\sqrt{2Dt}} \ln (ye^{\alpha t}) \right), \quad (38)$$

где $\Phi(z) = \frac{1}{\sqrt{\pi}} \int_0^z dx e^{-x^2}$ — интеграл ошибок.

Зная только одновременные статистические характеристики процесса $y(t; \alpha)$, можно получить важную информацию о поведении реализаций процесса $y(t; \alpha)$ на всем интервале времен $(0, \infty)$. В частности:

1. Зная интегральную функцию распределения, можно вычислить типичную реализацию логнормального процесса $y(t; \alpha)$, которая оказывается экспоненциально спадающей кривой

$$y^*(t; \alpha) = e^{-\alpha t}.$$

2. Легко написать выражения для моментных функций процесса $y(t; \alpha)$:

$$\langle y^n(t; \alpha) \rangle = e^{n(n-\alpha/D)Dt}, \quad \langle y^{-n}(t; \alpha) \rangle = e^{n(n+\alpha/D)Dt}, \quad n = 1, 2, \dots,$$

которые экспоненциально растут во времени. Следовательно, экспоненциальный рост моментов обусловлен выбросами процесса $y(t; \alpha)$ относительно кривой типичной реализации $y^*(t; \alpha)$ как в сторону больших, так и малых значений y .

Для значения параметра $\alpha/D = 1$ среднее значение процесса $y(t; D)$ не зависит от времени и равно единице. При этом, однако, вероятность выполнения неравенства $y < 1$ при $Dt \gg 1$ быстро стремится к единице по закону

$$P(y(t; D) < 1) = \Phi \left(\sqrt{\frac{Dt}{2}} \right) = 1 - \frac{1}{\sqrt{\pi Dt}} e^{-Dt/4},$$

т. е. подавляющее время графики реализаций процесса лежат ниже уровня его среднего значения $\langle y(t; D) \rangle = 1$, хотя статистические моменты процесса $y(t; D)$ в основном и определяются его большими выбросами.

Таким образом, мы имеем явное противоречие между характером поведения статистических характеристик процесса $y(t; \alpha)$ и его реализаций.

3. Поведение реализаций процесса $y(t; \alpha)$ на всем интервале времени можно также оценить с помощью *p*-мажорантных кривых $M_p(t, \alpha)$, которые определим следующим образом. Назовем мажорантной кривой такую кривую $M_p(t, \alpha)$, для которой при любых временах t с вероятностью p выполняется неравенство $y(t; \alpha) < M_p(t, \alpha)$, т. е.

$$P \{y(t; \alpha) < M_p(t, \alpha) \text{ для всех } t \in (0, \infty)\} = p.$$

Для логарифмически нормальных процессов легко получить однопараметрический класс экспоненциально спадающих мажорантных кривых

$$M_p(t, \alpha, \beta) = \frac{1}{(1-p)^{D/\beta}} e^{(\beta-\alpha)t}. \quad (39)$$

Обратим внимание на тот факт, что несмотря на постоянство статистического среднего $\langle y(t; D) \rangle = 1$ и экспоненциальный рост высших моментов процесса $y(t; D)$, всегда можно указать экспоненциально спадающую мажорантную кривую (39), ниже которой будут лежать реализации процесса $y(t; D)$ с любой наперед заданной вероятностью $p < 1$. В частности, например, с вероятностью $p = 1/2$ выполняется неравенство

$$y(t; D) < 4 \exp \left\{ -\frac{Dt}{2} \right\} \quad (40)$$

для любого момента времени t из интервала $(0, \infty)$. Это еще раз подтверждает сделанный ранее вывод о том, что экспоненциальный рост моментов процесса $y(t; D)$ во времени — эффект чисто статистический, обусловленный усреднением по всему ансамблю реализаций.

Отметим, что площадь под экспоненциально убывающими мажорантными кривыми конечна. Следовательно, большие выбросы процесса $y(t; \alpha)$, вызывая экспоненциальный рост высших моментов, не вносят существенного вклада в площадь под реализациями, которая практически для всех реализаций также конечна, т. е. выбросы логнормального процесса $y(t; \alpha)$ достаточно узки.

Все эти свойства логнормального процесса проявляются в динамике конкретных физических систем в виде когерентных явлений — *локализации и кластеризации*.

2.3. Диффузионное приближение

Условием применимости приближения дельта-коррелированности случайного поля $f(\mathbf{x}, t)$ (т. е. уравнения Фоккера — Планка) является условие малости временного радиуса корреляции τ_0 случайного

поля $f(\mathbf{x}, t)$ по сравнению со всеми временными масштабами, имеющимися в рассматриваемой задаче. Учет конечности временного радиуса корреляции случайного поля $f(\mathbf{x}, t)$ можно провести в рамках диффузионного приближения. Это приближение более наглядно и физично, чем формальное математическое приближение дельта-коррелированного случайного поля. Это приближение справедливо также для достаточно малых флуктуаций параметров стохастической динамической системы и позволяет получить не только условия применимости дельта-коррелированного приближения, но и описать новые физические эффекты, порожденные конечностью временного радиуса корреляции случайных параметров. В рамках диффузионного приближения предполагается, что влияние случайных воздействий на временных масштабах порядка τ_0 не существенно, т. е. система на этих масштабах эволюционирует как свободная.

Пусть опять векторная функция $\mathbf{x}(t)$ удовлетворяет динамическому уравнению (19), где $\mathbf{v}(\mathbf{x}, t)$ — векторная детерминированная функция, а $f(\mathbf{x}, t)$ — случайное статистически однородное в пространстве и стационарное во времени гауссово векторное поле со статистическими характеристиками

$$\langle f(\mathbf{x}, t) \rangle = 0, \quad B_{ij}(\mathbf{x}, t; \mathbf{x}', t') = B_{ij}(\mathbf{x} - \mathbf{x}'; t - t') = \langle f_i(\mathbf{x}, t) f_j(\mathbf{x}', t') \rangle.$$

Для индикаторной функции $\varphi(\mathbf{x}, t) = \delta(\mathbf{x}(t) - \mathbf{x})$, где $\mathbf{x}(t)$ — решение уравнения (19), имеем уравнение Лиувилля (23) и, следовательно, уравнение для плотности вероятностей решения уравнения (19) $P(\mathbf{x}(t)) = \langle \varphi(\mathbf{x}, t) \rangle$ имеет вид (26).

В диффузионном приближении уравнение (26) является точным, а вариационная производная и индикаторная функция на временных масштабах порядка временного радиуса корреляции τ_0 случайного поля $f(\mathbf{x}, t)$ описываются системой динамических уравнений:

$$\begin{aligned} \frac{\partial}{\partial t} \frac{\delta \varphi(\mathbf{x}, t)}{\delta f_i(\mathbf{x}', t')} &= - \frac{\partial}{\partial \mathbf{x}} \left\{ \mathbf{v}(\mathbf{x}, t) \frac{\delta \varphi(\mathbf{x}, t)}{\delta f_i(\mathbf{x}', t')} \right\}, \\ \frac{\delta \varphi(\mathbf{x}, t)}{\delta f_i(\mathbf{x}', t')} \Big|_{t=t'} &= - \frac{\partial}{\partial x_i} \left\{ \delta(\mathbf{x} - \mathbf{x}') \varphi(\mathbf{x}, t') \right\}, \\ \frac{\partial}{\partial t} \varphi(\mathbf{x}, t) &= - \frac{\partial}{\partial \mathbf{x}} \{ \mathbf{v}(\mathbf{x}, t) \varphi(\mathbf{x}, t) \}, \quad \varphi(\mathbf{x}, t) \Big|_{t=t'} = \varphi(\mathbf{x}, t'). \end{aligned} \quad (41)$$

Решение задачи (26) и (41) справедливо для всех времен t . Однако в этом случае решение задачи (19) $\mathbf{x}(t)$ не является векторным марковским случайным процессом, так как ее многовременная плотность вероятностей не допускает факторизации с помощью

плотности вероятностей перехода. В асимптотическом случае $t \gg \tau_0$ решение исходной динамической системы (19) в диффузионном приближении уже будет марковским случайным процессом и условиями применимости его является малость всех статистических эффектов на временных масштабах порядка временного радиуса корреляции τ_0 .

Таким образом, диффузионное приближение позволяет снять основное ограничение, связанное с малостью временного радиуса корреляции τ_0 , оставаясь при этом в рамках марковского процесса. С помощью такого приближения можно, например, рассмотреть задачу о диффузии поля плотности малоинерционной примеси в случайном гидродинамическом бездивергентном потоке на основе уравнений (10) и (11), для которой неприменимо приближение дельта-коррелированности поля $\mathbf{u}(\mathbf{r}, t)$, и, следовательно, описать явление кластеризации в этой задаче.

* * *

Подход к анализу стохастических динамических задач на основе статистической топографии, позволяющий по одноточечным статистическим характеристикам случайных процессов и полей определить количественные и качественные характеристики отдельных реализаций этих величин на всем интервале времен (во всем пространстве), возник в процессе дискуссий с экспериментаторами, которые имеют дело в основном с отдельными реализациями. С их точки зрения, теоретики, занимающиеся вычислениями статистических средних, изучают неизвестно что, так как эти статистические средние зачастую не только не похожи на отдельные реализации, но и, как было показано выше, просто противоречат им. Предложенный выше подход позволил снять все эти противоречия и в какой-то мере «примирил» экспериментаторов и теоретиков.

Конечно, сама работа по вычислению статистических средних динамических систем — это сложная математическая задача, для решения которой используются различные простейшие статистические модели флуктуирующих параметров типа процессов и полей белого шума. Такие модели правильно описывают основные особенности динамических систем. Однако для более детального анализа «более тонких» особенностей необходимо в ряде случаев учитывать конечность временного (пространственного) радиуса корреляции.

Работа выполнена при поддержке РФФИ (проекты 04-05-64044 и 02-05-64375).

Литература

1. *Кляцкин В. И.* Стохастические уравнения глазами физика. Основные положения, точные результаты и асимптотические приближения. М.: Физматлит, 2001.
2. *Кляцкин В. И.* Динамика стохастических систем. Курс лекций. М.: Физматлит, 2002.
3. *Кляцкин В. И.* Кластеризация и диффузия частиц и плотности пассивной примеси в случайных гидродинамических потоках // УФН. 2003. Т. 173, № 7. С. 689—710.
4. *Кляцкин В. И.* Распространение электромагнитных волн в случайно-неоднородной среде как задача статистической математической физики // УФН. 2004. Т. 174, № 2. С. 177—195.
5. *Кляцкин В. И., Гуларий Д.* Когерентные явления в стохастических динамических системах // УФН. 1999. Т. 169, № 2. С. 171—207.
6. *Кляцкин В. И., Саичев А. И.* Стохастическая и динамическая локализация плоских волн в случайных слоистых средах // УФН. 1992. Т. 162, № 3. С. 161— 194.

СЛОЖНАЯ ДИНАМИКА РАСПРЕДЕЛЕННЫХ НЕЛИНЕЙНЫХ СИСТЕМ СВЕРХВЫСОКОЧАСТОТНОЙ ЭЛЕКТРОНИКИ

*Д. И. Трубецков, С. П. Кузнецов,
Н. М. Рыскин, А. Е. Храмов*

1. Сверхвысокочастотная электроника глазами нелинейной динамики (общий взгляд)

Когда-то С. М. Рытов определил радиофизику как «физику для радио и радио для физики». Право С. М. Рытова давать определение подтвердил тогда академик Минц, сказав, что «радиофизика — это то, чем занимается С. М. Рытов». Но термин родился в Нижнем Новгороде, тогда еще Горьком. В «Избранных очерках о зарождении и взрослении радиофизики в горьковско-нижегородских местах» М. А. Миллер отмечает, что В. И. Гапонов писал в «Горьковской правде»: «Радиофизика держится на трех китах: теории колебаний, электродинамике и электронике». Тема этой статьи — третий кит — электроника.

Правда, сегодня, согласно опять же М. А. Миллеру, «радиофизика ... вымахала за пределы, отпущенные ей при рождении: она проникла в акустику, в гидро-, гео- и астрофизику, в газо- и плазмодинамику, в твердотель, и уж конечно во все видимые и невидимые диапазоны электромагнетизма». Добавим сюда и науку о живом.

Радиофизика стала нелинейной, причем нелинейность стала понятием первичным. Появилось то, что называют нелинейной динамикой или синергетикой, то, что можно назвать современной наукой о колебаниях и волнах.

По Ю. Л. Климонтовичу «синергетика — не самостоятельная научная дисциплина, но новое междисциплинарное направление; цель синергетики — выявление общих идей, общих методов и общих закономерностей в самых разных областях естествознания, а также социологии и даже лингвистики; более того, в рамках синергетики происходит кооперирование различных специальных дисциплин».

Представляется, что на отдельные дисциплины разумно взглянуть глазами нелинейной динамики. Эта статья — подобный взгляд на вакуумную сверхвысокочастотную (СВЧ) электронику, точнее на ее отдельные объекты.

Что связывает нелинейную динамику и СВЧ-электронику, какие общие физические процессы и явления они изучают? Их много: самовозбуждение колебаний в неравновесной среде, возникновение и эволюция динамического хаоса, излучение в неоднородной среде — усиление и параметрические явления, нелинейные волны, образование паттернов, процессы синхронизации.

Нелинейная динамика изменила СВЧ-электронику, в первую очередь, благодаря решению нестационарных задач. Хокни и Банеман эмоционально писали применительно к магнетрону: «... вместо того, чтобы искать этот вечно ускользающий стационарный режим, можно следить за пространственно-временной эволюцией системы...»

И конечно, новый толчок исследованиям дало открытие сложной динамики в нелинейных системах, и особенно динамического хаоса, а также фрактальной геометрии.

Далее кратко рассмотрим с точки зрения нелинейной динамики «коллекцию» примеров в рамках пяти идей, создавших СВЧ-электронику. Более подробно остановимся на некоторых экземплярах коллекции, в частности на лампе обратной волны как распределенной автоколебательной системе, демонстрирующей сложную динамику и хаос. Кроме того, опишем образование паттернов и их взаимодействие в СВЧ-устройствах с виртуальным катодом.

Начнем с напоминаний об особенностях СВЧ-электроники по сравнению с электроникой низких частот. В любом электронном приборе взаимодействие переменного электромагнитного поля (электромагнитных колебаний или волн) с электронами (свободными или связанными) является основным в его работе. Поле воздействует на электронный поток, создает в нем переменный ток, который, в свою очередь, изменяет поле. Конечно, для конкретных расчетов нужно детализировать устройство электронного прибора, но всегда имеется связка поле — электроны (рис. 1).

Следуя [1], сформулируем, какими свойствами обладают низкочастотные лампы (например, вакуумный триод), а потом проследим, выполняются ли эти свойства на СВЧ (0,3—3000 ГГц).

Для низкочастотных ламп справедливо следующее.

1. Переменные поля в лампе не должны проявлять волновых свойств, в частности, вводы и выводы, выполненные в виде проводов, не должны существенно влиять на свойства лампы и ее цепи. Если D — характерный размер лампы, $k = 2\pi/\lambda$ — волновое число, λ — длина волны, то волновые свойства не проявляются при $kD \ll 1$, или $D \ll \lambda$.

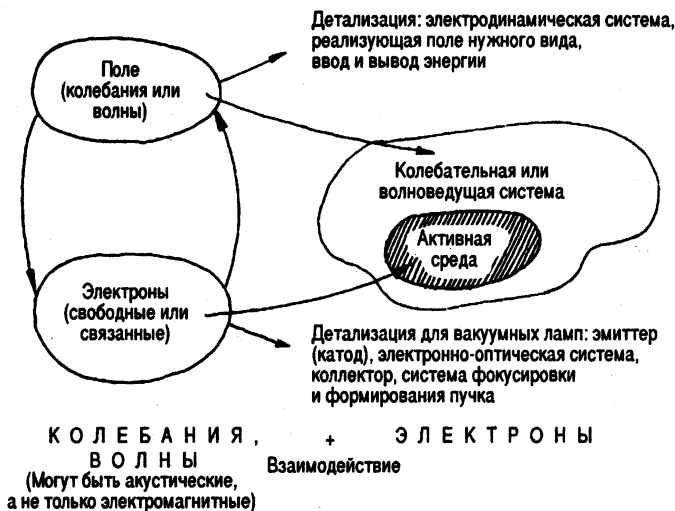


Рис. 1. Колебания, волны и электроны в электронных приборах

2. Электроны в лампе не должны проявлять своей инерционности (время пролета их между электронами $t_{\text{пр}}$ пренебрежимо мало по сравнению с периодом T колебаний поля). Это условие можно записать так: $\omega t_{\text{пр}} \ll 1$ или $t_{\text{пр}} \ll T$. Оба приведенных условия можно переписать единообразно. Если ввести $t_{\text{расп}} = D/c$ — время распространения волны через лампу и ее цепь (c — скорость света в пустоте), то оба приведенных выше неравенства можно представить следующим образом: $\omega t_{\text{расп}} \ll 1$. В СВЧ-электронике свойства, сформулированные выше, нарушаются, и нужно написать, что

$$D \sim \lambda, \quad t_{\text{пр}} \sim T \quad \text{или даже} \gg T,$$

кроме того, вводы и выводы энергии существенно влияют на работу лампы. Именно эти особенности привели к появлению пяти главных идей, создавших СВЧ-электронику, которые кратко изложим, выделяя основные колебательно-волновые явления и эффекты.

Идея первая. Модуляция электронов по скорости и группирование электронов в пространстве дрейфа.

Как известно, этой идее соответствует одна простая модель, в которой вдоль оси x движется поток невзаимодействующих частиц так, что в эйлеровых переменных (x, t) их скорость удовлетворяет уравнению

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = 0.$$

В электронике СВЧ это уравнение описывает в рамках так называемой кинематической теории поведение электронного потока в трубе дрейфа приборов клистронного типа (детали можно найти, например, в [2]). Различие в скоростях электронов приводит в трубе дрейфа к образованию электронных уплотнений — группированию электронов потока.

Пусть в момент времени $t = 0$ поток оказывается возмущенным по скорости по закону $a \sin kx$, где $k = \omega/v_0$ — волновое число, соответствующее волне скорости малой амплитуды с частотой ω ; $v_0 = \text{const}$. В системе координат, движущейся со скоростью v_0 ($v = v_0 + u$), уравнение $dv/dt = 0$ записывается как $\partial u/\partial t + u\partial u/\partial x = 0$, и его решение имеет вид простой волны $u = U(t - x/u)$, где выражение для U определяется начальным возмущением. Рисунок 2 иллюстрирует процесс эволюции потока на фазовой плоскости (u, x) .

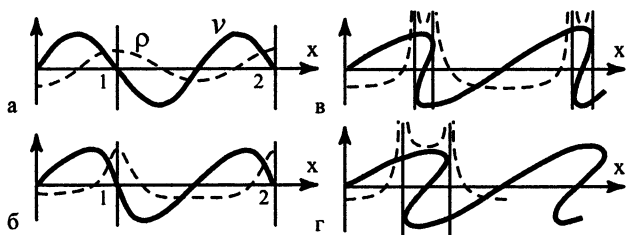


Рис. 2. Эволюция во времени синусоидального возмущения в пучке невзаимодействующих частиц: *а* — начальное состояние, соответствующее синусоидальному возмущению скорости; *б* — образование электронных уплотнений (группирование частиц около точек 1 и 2); *в, г* — опрокидывание волны скорости и образование удвоенного числа особенностей на кривой $\rho = \rho(x)$. Сплошными линиями обозначена скорость, штриховыми — плотность.

Начальное состояние потока — синусоида $a \sin kx$ на плоскости (u, x) ; здесь же штриховой линией показана зависимость плотности объемного заряда потока от x (рис. 2, *а*). С течением времени происходит искажение профиля волны: частицы с $u > 0$ уходят вперед, а с $u < 0$ отстают от волны. Одновременно образуются сгущения частиц вблизи точек 1 и 2, где $u = 0$, и происходит группирование потока (рис. 2, *б*). Волна постепенно становится все круче, и в конце концов производная $\partial u/\partial x$ на ее переднем фронте обращается в бесконечность (в бесконечность обращается в этой точке и плотность $\rho(x)$ объемного заряда потока). В следующий момент происходит

опрокидывание волны, и функция $u(x, t)$ перестает быть однозначной (рис. 2, в, з): у нее появляется точка поворота, т. е. образуются встречные пучки. После опрокидывания волны функция $\rho(x)$ имеет удвоенное число особенностей (рис. 2, в, з).

В СВЧ-электронике для описания процесса группирования электронов отказываются от переменных Эйлера (x, t) и переходят к переменным Лагранжа (t, t_0) или (x, t_0) , где t_0 — начальный момент влета электронов в трубу дрейфа. Наглядное представление о группировании в трубе дрейфа дает так называемая пространственно-временная диаграмма на плоскости $(x, \omega t_0)$ на рис. 3.

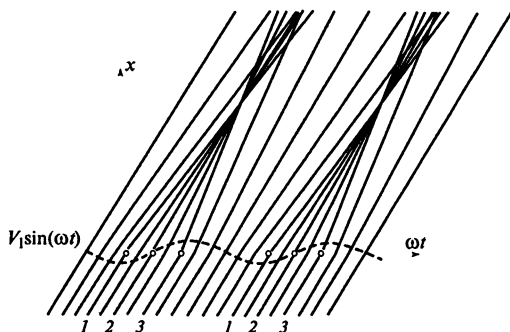


Рис. 3. Пространственно-временная диаграмма группирования электронов в пространстве дрейфа: 1 — электрон, который тормозится полем, 2 — электрон, не испытывающий воздействия со стороны поля, 3 — электрон, который ускоряется полем.

Из рис. 3 видно, что за период высокочастотного воздействия траектории сходятся (образуются уплотнения частиц) или расходятся (образуются разрежения частиц), что и иллюстрирует процесс группирования.

Сравнивая «волновой» и «корпускулярный» подходы к описанию группирования, можно сделать вывод, что процесс опрокидывания волны соответствует явлению обгона одних частиц другими.

Рассмотренная модель сильно идеализирована, поскольку не учитывает сил кулоновского взаимодействия между частицами. В частности, для волн малой амплитуды в системе дисперсии нет. Поскольку в модели, учитывающей влияние сил пространственного заряда, дисперсия появляется, естественным становится предположение о возможности существования в электронном потоке уединенных волн и даже солитонов. В большинстве работ, посвящен-

ных нелинейным волнам пространственного заряда (ВПЗ), исходные уравнения путем «насильственных» упрощений сводились к уравнению Кортевега — де Вриза (КдВ), на основе чего делался вывод о возбуждении в электронном потоке КдВ-солитонов. В работе [3] были найдены точные решения в виде уединенных волн, определены условия их устойчивости и проведено численное моделирование процессов в электронном потоке. Показано, что при эволюции широкого класса начальных возмущений образуются одна или несколько уединенных волн и «осциллирующий хвост» малой амплитуды. Пример такой эволюции иллюстрирует рис. 4, на котором приведены зависимости скорости (*a*) и плотности заряда (*b*) от координаты в различные моменты времени.

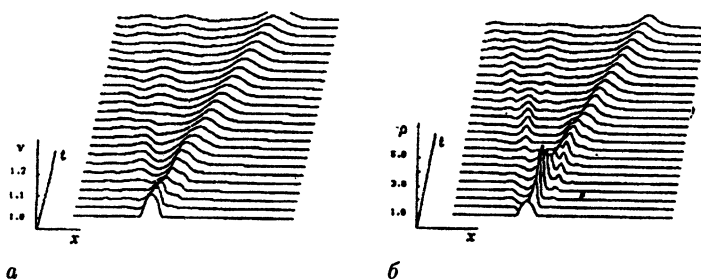


Рис. 4. Образование уединенной волны и осциллирующего хвоста при эволюции возмущения: *a* — скорость электронов, *b* — плотность заряда.

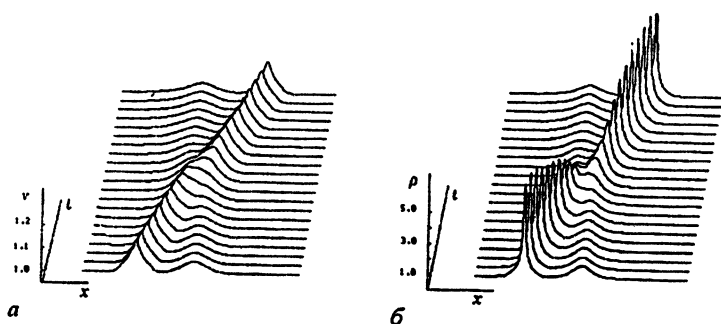


Рис. 5. Взаимодействие двух быстрых уединенных волн с сильно различающимися амплитудами: *a* — скорость электронов, *b* — плотность заряда.

Расчеты показали, что столкновение попутных волн является упругим с достаточно большой степенью точности (рис. 5).

Существует качественное отличие между взаимодействием волн с сильно различающимися и близкими амплитудами: в последнем случае столкновение происходит без образования одиночного пика. Столкновение встречных волн (т. е. быстрой и медленной) является существенно неупругим.

Особую группу приборов с динамическим способом управления электронным потоком (скоростная модуляция с последующим группированием) составляют генераторы со сверхпредельными токами (виркаторы), работа которых основана на формировании в пространстве взаимодействия виртуального катода (ВК) — области с потенциалом поля пространственного заряда, близким к потенциалу источника электронов. Одной из простейших моделей виркаторов на пролетном токе является модель диода Пирса, в которой поток электронов, статический пространственный заряд которого скомпенсирован фоном неподвижных ионов, поступает в эквипотенциальное дрейфовое пространство (подробное математическое описание модели и результатов численного эксперимента приведено ниже в самостоятельном разделе статьи). Применение современных методов нелинейной динамики (расчет размерностей восстановленных аттракторов колебаний ВК, вычисление максимальных ляпуновских характеристических показателей, определение взаимных корреляций между колебаниями плотности заряда в различных сечениях потока) позволило обнаружить области хаотических колебаний ВК в пространстве управляющих параметров системы.

Анализ физических процессов в электронном потоке проводится при помощи построения пространственно-временных диаграмм. Типичный результат для одного периода колебаний виртуального катода представлен на рис. 6. Каждая кривая показывает изменение местоположения x крупной частицы — электрона с течением времени t . На одном периоде колебаний можно выделить два сгущения кривых, которым соответствуют два сгустка, формирующихся в электронном потоке. Эти сгустки возникают при любых начальных условиях и представляют собой автоструктуры. Первый сгусток соответствует ВК. При его распаде поступающий в систему поток модулируется полем первого сгустка и в результате превращения модуляции по скорости в модуляцию по плотности формируется второй сгусток. Возникновению хаотических колебаний соответствует критическое значение плотности заряда во втором сгустке, при котором часть потока начинает отражаться от него, оказывая воздействие на формирование ВК на следующем периоде колебаний, что приводит к формированию распределенной обратной связи в системе.

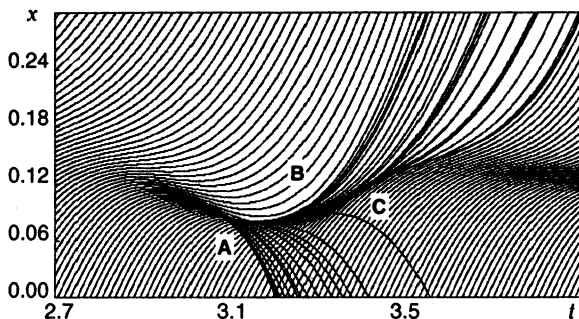


Рис. 6. Пространственно-временная диаграмма движения электронов в диоде Пирса, демонстрирующая образование двух взаимодействующих электронных структур.

Таким образом, возникновение хаотических колебаний в электронном потоке с ВК связано с усилением связи между структурами (возникновением распределенной внутренней обратной связи по потоку).

Идея вторая. *Взаимодействие электронного потока, движущегося в скрещенных статических электрическом и магнитном полях, с ВЧ электромагнитным полем (магнетронное взаимодействие).*

Нет сомнения в том, что спицы пространственного заряда в магнетроне представляют собой динамические автоструктуры, эволюция которых по-прежнему не понята до конца (см. «о трех загадках в магнетроне» [2, лекция 7]). И лишь в последние годы представления и методы нелинейной динамики позволяют продвинуться в понимании сложных процессов в облаке пространственного заряда и объяснении, например, больших шумов в магнетроне. Так, Г. Г. Моносовым [4] в компьютерном эксперименте обнаружен процесс хаотической модуляции в магнетроне. Качественно этот процесс можно описать следующим образом. Наложенное на однородный статический электронный поток высокочастотное поле резонаторной системы сначала приводит к образованию выпуклости, на которую постепенно накручиваются новые слои электронов, образуя локальную вращающуюся втулку. Когда верхние слои этой втулки поднимаются достаточно близко к аноду, они под действием высокочастотного поля резонаторной системы отрываются от локальной втулки и постепенно в виде сгустков импульсами оседают на аноде. После оседа-

ния сгустка локальная втулка постепенно восстанавливается и описанный процесс продолжается вновь. Нерегулярность этого процесса связана с тем, что время образования локальной втулки не связано со временем прохождения втулки от одной щели к другой. Поэтому процесс отрыва порций заряда от локальной втулки каждый раз начинается при новом исходном состоянии втулки. Вследствие этого от втулки каждый раз отрываются разные порции заряда. Такая хаотическая автомодуляция может являться одной из причин возникновения шумового пьедестала в частотном спектре сигнала, генерируемого магнетроном.

Сложная, в том числе хаотическая, динамика имеет место в усилителе обратной волны со скрещенными полями (амплитроне) и определяется возникающей обратной связью между входом и выходом прибора ([5], см. также [2]).

Идея третья. *Взаимодействие электронного потока с замедленной бегущей электромагнитной волной (лампа с бегущей волной — ЛБВ).*

Интересно, что нелинейная теория ЛБВ началась с поиска решений в виде нелинейных стационарных волн в работе Л. Бриллюэна (J. Appl. Phys. 1949. V. 20, № 12. P. 1196). Для СВЧ-электроники режимы возбуждения стационарных нелинейных волн, по крайней мере сегодня, интереса не представляют. Обзор теории нелинейных волн в электронных потоках и системах электронный поток — электромагнитное поле представлен в работе [6], где детально обсуждены вопросы, связанные с возможностью возбуждения и распространения в указанных системах уединенных волн, в частности солитонов.

Истории было угодно, чтобы ЛБВ-генератор с запаздывающей обратной связью (шумотрон), предложенный В. Я. Кисловым, стал первым СВЧ-генератором мощных шумовых (как впоследствии выяснилось, хаотических) сигналов (см., например, [7], а также [2]). Именно в натуральных экспериментах с шумотроном обнаруживались те или иные сценарии динамического хаоса и проводилось сопоставление с теорией.

Идея четвертая. *Взаимодействие электронного потока с обратной электромагнитной волной (лампа с обратной волной — ЛОВ или карсинотрон).*

Сложная динамика различных вариантов ЛОВ как распределенной автоколебательной системы исследована весьма полно как теоретически, так и экспериментально (см. [2], лекции 12, 13 и библиографию к ним). Тем не менее исследования продолжают, и об-

наруживается все больше и больше новых особенностей и тонкостей. Изложению основных результатов нестационарной теории ЛОВ посвящена отдельная часть статьи.

Идея пятая. *Взаимодействие криволинейных электронных потоков с электромагнитными полями. Мазеры на циклотронном резонансе (МЦР).*

Эта идея как бы с самого начала принадлежала не только СВЧ-электронике, но и нелинейной динамике: основополагающая статья называлась «Индукцированное излучение возбужденных классических электронов-осцилляторов и его использование в высокочастотной электронике» [8]. Наиболее удобной средой для будущих приборов оказался поток электронов-осцилляторов, вращающихся в магнитном поле. При взаимодействии таких осцилляторов с электромагнитным полем имеет место фазовая группировка, возникающая из-за неизохронности электронов-осцилляторов, т. е. вследствие релятивистской зависимости циклотронной частоты вращения от энергии электронов. Для МЦР изучены такие нелинейные явления, как конкуренция и синхронизация мод, автомодуляция и хаос.

Конечно, этот краткий обзор связи основных идей СВЧ-электроники с нелинейной динамикой далеко не полон: в частности, даже не упомянут параметрический способ управления электронным потоком, используемый в лазерах на свободных электронах; не обсуждена идея использования эффекта сверхизлучения ансамбля классических электронов-осцилляторов для генерации сверхмощных импульсов СВЧ-излучения.

«Коллекцию» примеров можно пополнять. Однако перейдем к более подробному обсуждению любимых «экземпляров коллекции».

2. Сложная динамика и хаос в ЛОВ¹

2.1. ЛОВ как динамическая система. Основные компоненты нестационарной нелинейной теории. В ЛОВ электронный пучок взаимодействует с волной в специальной электродинамической структуре (замедляющей системе), причем фазовая скорость волны близка к скорости электронов, а групповая скорость направлена навстречу пучку (рис. 7, а). Благодаря первому условию электроны подвергаются эффективному, накапливающемуся действию

¹См. [2], лекции 12 и 13 и библиографию к ним.

поля волны. По мере движения пучка вдоль системы в нем образуются сгустки, и возникает высокочастотная составляющая тока. Благодаря тому что распространение испущенного сгустками излучения (в смысле направления потока энергии) происходит навстречу пучку, обеспечиваются наличие в системе внутренней обратной связи и возможность автоколебаний. В этом отличие ЛОВ от ЛБВ, в которой перенос энергии волной осуществляется в направлении движения пучка и которая используется как усилитель.

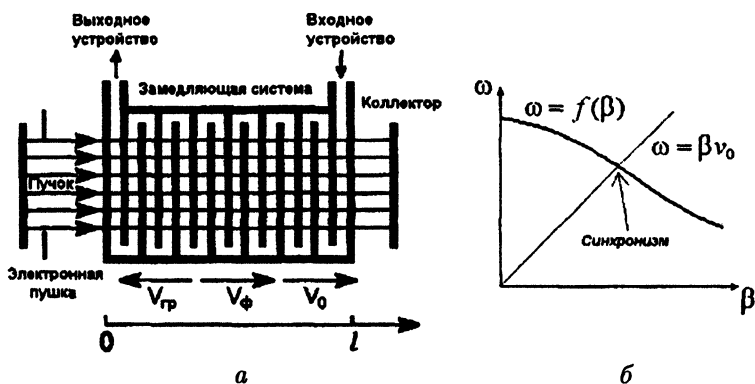


Рис. 7. Схематическое изображение ЛОВ типа О в «классическом» низковольтном варианте (а) и характерное расположение на диаграмме Бриллюэна дисперсионных характеристик, отвечающих синхронному взаимодействию электронного пучка с обратной волной — ЛОВ (б).

В этом разделе обсуждаются ключевые составляющие нестационарной нелинейной теории ЛОВ типа О, основанной на методе медленно меняющихся амплитуд. Естественно начать анализ с простейшего случая, когда динамика определяется единственным управляющим параметром. Таковым является параметр нормированной длины, появляющийся уже в простейшей модели ЛОВ, в которой не учитываются такие дополнительные факторы, как релятивистские эффекты, пространственный заряд, диссипация, отражения и др., $L = 2\pi CN$, где C — известный в электронике приборов с длительным взаимодействием параметр Пирса. Принцип длительного взаимодействия подразумевает, что длина системы l выбирается большой по сравнению с длиной замедленной волны, т. е. $N = l/\lambda_{\text{зам}} \gg 1$, а взаимодействие пучка и волны достаточно слабое. В этой ситуации существенно воздействовать на электронный пучок на протяжении всей длины системы и эффективно возбуждаться им могут только

спектральные составляющие поля, фазовые скорости которых близки к скорости движения электронов. Из-за характерной конфигурации взаимного расположения дисперсионных характеристик замедляющей системы и электронного пучка (рис. 7, б) это имеет место в узком интервале частот и волновых чисел вблизи точки пересечения (ω_0, β_0) на диаграмме Бриллюэна. На пространственно-временном языке узость спектра волновых чисел и частот означает медленность изменения амплитуды высокочастотного поля, соответственно, в пространстве и во времени. Вводя комплексную амплитуду $\tilde{E}(x, t)$ для взаимодействующей с электронным пучком составляющей поля волны, полагаем

$$E(x, t) = \text{Re}\tilde{E}(x, t)e^{i\omega_0 t - i\beta_0 x}. \quad (1)$$

Характерный пространственный масштаб изменения комплексной амплитуды $\tilde{E}(x, t)$ будет соответствовать по порядку величины длине системы L , а временной масштаб — характерному времени обхода сигнала с электронным пучком и волной по петле обратной связи. Оно дается соотношением $T_{\text{хар}} = L/v_0 + L/v_{\text{гр}}$, где v_0 — скорость пучка, а $v_{\text{гр}}$ — групповая скорость волны в точке пересечения дисперсионных характеристик: $v_{\text{гр}} = |d\omega(\beta_0)/d\beta|$. Высокочастотный ток, возникающий в электронном пучке в результате воздействия на него поля волны, представляется в виде

$$I(x, t) = \text{Re}[I_1(x, t)e^{i\omega_0(t-x/v_0)} + I_2(x, t)e^{2i\omega_0(t-x/v_0)} + \dots], \quad (2)$$

где $I_1(x, t), I_2(x, t), \dots$ — медленно меняющиеся амплитуды соответственно первой, второй и следующих гармоник. Только первая гармоника будет эффективно возбуждать электромагнитное поле, поскольку соответствующая ей волна тока находится в синхронизме с волной в замедляющей системе.

Уравнение Ньютона для продольного движения электрона под действием поля электромагнитной волны (1) имеет вид

$$\frac{d^2 x}{dt^2} = \frac{e}{m_0} \text{Re}\tilde{E}(x, t)e^{i\omega_0 t - i\beta_0 x}. \quad (3)$$

Перепишем его в эквивалентной форме относительно функции $t(x, t_0) = t_0 + x/v_0 + \tilde{t}(x, t_0)$, которая определяет время прибытия в точку x электрона, влетевшего в пространство взаимодействия в момент t_0 . Преобразование производной в левой части уравнения движения дает $d^2 x/dt^2 = -(\partial t/\partial x)_{t_0}^{-3}(\partial^2 t/\partial x^2)_{t_0}$. Из-за медленности изменения комплексной амплитуды во времени и малости изменения

скорости электронов в процессе взаимодействия мы вправе заменить $\tilde{E}(x, t(x, t_0))$ на $\tilde{E}(x, t_0 + x/v_0)$ и $(\partial t/\partial x)_{t_0}^3$ на v_0^{-3} . Тогда уравнение движения примет вид

$$(\partial \tilde{t}/\partial x^2)_{t_0} = -(e/mv_0^3) \operatorname{Re} \tilde{E}(x, t_0 + x/v_0) \exp i\omega_0 [t_0 + \tilde{t}(x, t_0)], \quad (4)$$

при этом на входе в пространство взаимодействия должно выполняться условие

$$\tilde{t}|_{x=0} = 0, \quad \partial \tilde{t}/\partial t|_{x=0} = 0. \quad (5)$$

Для вычисления комплексной амплитуды высокочастотного тока, которая будет фигурировать в уравнении возбуждения, необходимо отслеживать группу электронов, интервал времен влета которых в пространство взаимодействия отвечает одному высокочастотному периоду: $\omega_0 t_0 = 0 \div 2\pi \pmod{2\pi}$.

При взаимном расположении дисперсионных характеристик, показанном на рис. 7, б, основной нестационарный эффект, с точки зрения волн в замедляющей системе, — распространение огибающей с групповой скоростью. Оно учитывается при аппроксимации дисперсионной характеристики прямой линией, наклон которой определяется групповой скоростью в точке пересечения. Прочие частотные зависимости (непостоянство групповой скорости, зависимость сопротивления связи от частоты) можно не принимать во внимание из-за узости спектра частот и волновых чисел. В итоге требуемое уравнение можно «сконструировать» из принятого в СВЧ-электронике в стационарной теории, просто добавив член, ответственный за групповой перенос. Для комплексных амплитуд уравнение записывается так:

$$v_{\text{гп}}^{-1} \partial \tilde{E}/\partial t - \partial \tilde{E}/\partial x = -\beta_0^2 K_0 \tilde{I}_1/2, \quad (6)$$

где K_0 — сопротивление связи замедляющей системы для рабочей пространственной гармоники на частоте точного синхронизма ω_0 . Для дальнейшего удобно переписать уравнение в переменных $(x, t_0 = t - x/v_0)$. Тогда, используя правила вычисления производных, уравнение (6) можно представить в виде

$$(v_{\text{гп}}^{-1} + v_0^{-1})(\partial \tilde{E}/\partial t_0)_x - (\partial \tilde{E}/\partial x)_{t_0} = -\beta_0^2 K_0 \tilde{I}_1/2. \quad (7)$$

На правом конце пространства взаимодействия входной сигнал отсутствует, поэтому должно выполняться краевое условие

$$\tilde{E}(l, t) = 0. \quad (8)$$

Представим комплексную амплитуду первой гармоники тока пучка через функцию $t(x, t_0)$, фигурирующую в уравнении движения. Обозначая чертой усреднение за высокочастотный период, можно записать:

$$I_1(x, t) = \overline{2I(x, t)e^{-i\omega_0(t-x/v_0)}} = \frac{1}{\pi} \int I_0 e^{-i\omega_0(t_0 + \tilde{t}(x, t_0))} d(\omega_0 t), \quad (9)$$

где интегрирование ведется по периоду $2\pi/\omega_0$ и при преобразовании интеграла принято во внимание соотношение $|I(x, t)dt| = |I_0 dt_0|$, вытекающее из закона сохранения заряда. Введем следующие безразмерные переменные и параметры:

$$\xi = x/L, \quad \tau = \frac{t_0}{l/v_0 + l/v_{гp}} = \frac{t - x/v_0}{l/v_0 + l/v_{гp}},$$

$$\theta = \omega_0(t_0 + \tilde{t}(x, t_0)) \pmod{2\pi}, \quad F(\xi, \tau) = \frac{\tilde{E}}{2\beta_0 UC^2},$$

$$C = \sqrt[3]{\frac{I_0 K_0}{4U}} \text{ (параметр Пирса), } L = \beta_0 l C = 2\pi CN.$$

Тогда уравнения (4), (5), (7), (8) переписутся в виде

$$\frac{\partial^2 \xi^2}{\partial \theta} = -L^2 \operatorname{Re} F \exp[i\theta], \quad \partial F / \partial \theta - \partial F / \partial \xi = -\frac{L}{\pi} \int_0^{2\pi} e^{-i\theta} d\theta_0, \quad (10)$$

$$\theta|_{\xi=0} = \theta_0, \quad \partial \theta / \partial \xi|_{\xi=0} = 0, \quad F|_{\xi=1} = 0. \quad (11)$$

Заметим, что параметр Пирса обычно мал, $C \ll 1$. Можно показать, что это неравенство является основным условием применимости метода медленно меняющихся амплитуд. В дальнейшем мы будем ссылаться на уравнения (10)—(11) как на *основную теоретическую модель ЛОВ*. Для учета дополнительных факторов, таких как релятивистские эффекты, поле пространственного заряда, потери при распространении волн в замедляющей системе, отражения на концах и т. п., следует определенным образом модифицировать основную модель.

Имея уравнения (10)—(11), мы можем говорить о ЛОВ как о динамической системе. Представим себе какое-либо эволюционирующее во времени решение нестационарной краевой задачи (10)—(11). Траектуя τ как время, имеем при каждом τ некоторое распределение

поля по длине ξ , описываемое комплексной функцией $F(\xi)$. Если она задана в некоторый момент τ , то и при любом $\tau' > \tau$ в результате решения уравнений распределение поля определяется однозначно. По терминологии, принятой в теории динамических систем, множество комплексных функций $F(\zeta)$, удовлетворяющих естественно для комплексных амплитуд требованию гладкости и условию $F(1) = 0$, есть фазовое пространство данной модели ЛОВ. Каждой точке этого пространства, т. е. функции $F(\zeta)$, отвечает определенное состояние системы. Когда состояние меняется во времени в соответствии с уравнениями динамики системы (10) и (11), изображающая точка движется в пространстве функций по некоторой «фазовой траектории». Фазовое пространство модели бесконечномерное, тем не менее использование соответствующей терминологии оправдано, так как характерные для ЛОВ режимы и феномены оказываются сопоставимыми с общими понятиями теории динамических систем. Соответствующие геометрические образы, например фазовые портреты, которые можно интерпретировать как проекции аттрактора из бесконечномерного пространства на плоскость, полезны для наглядного представления и распознавания различных режимов регулярной и хаотической динамики. Более того, их несложно наблюдать также и в эксперименте.

2.2. Нелинейная динамика ЛОВ: от порога самовозбуждения до хаоса. Как уже говорилось, модель ЛОВ, заданная уравнениями (10)–(11), содержит единственный параметр, от которого может зависеть характер динамических режимов, — нормированную длину $L = 2\pi CN$. На практике ее обычно регулируют изменением тока электронного пучка I , варьируя напряжение на специальном управляющем электроде в электронной пушке, при этом $L \propto I_0^{1/3}$. Для электронной системы L играет роль основного управляющего параметра, подобную той, которая при анализе возникновения турбулентности отводится числу Рейнольдса в гидродинамических задачах. Обсудим динамические режимы и бифуркации в ЛОВ, реализующиеся при увеличении нормированной длины системы.

Начнем с того, что при малых значениях параметра L реализуется устойчивое тривиальное решение $F = 0$. Ему отвечает в фазовом пространстве неподвижная точка или состояние равновесия, ассоциирующееся с предгенерационным режимом ЛОВ. Из линейного анализа известно, что это состояние теряет устойчивость начиная с $L - L_{ст} = 1,97327$. Соответствующее значение тока пучка называют стартовым или пусковым током ($I_{ст}$). При превышении стартового

тока одна из мод линеаризованной задачи становится нарастающей во времени (две моды для $L > 3,69286$, три для $L > 4,79078$). Если начальный уровень колебаний очень мал, то процесс на некотором этапе описывается асимптотикой решения линеаризованной задачи $F = C_1 F_1(\xi) e^{\kappa \tau}$, где C_1 — постоянный коэффициент, $F_1(\xi)$ — собственная функция линейной задачи, κ — комплексное собственное число, отвечающие наиболее быстро растущей первой моде. Численное решение уравнений на компьютере показывает, что при не слишком большом превышении L над $L_{\text{ст}}$ в итоге переходного процесса устанавливается режим с постоянной амплитудой и частотой сигнала $F(\xi, \tau) = F^0(\xi) e^{i\Phi \tau}$, где $\Phi(L) = \Delta\omega l/v_0 < 0$ характеризует сдвиг частоты относительно точки синхронизма.

По мере увеличения параметра L характер переходного процесса меняется таким образом, что выход амплитуды на стационарный уровень сопровождается осцилляциями, затухающими все более медленно. Наконец, при $L > L_m \approx 2,937$ эти осцилляции становятся незатухающими, и при достаточно большом времени наблюдения устанавливается режим с периодической модуляцией сигнала по амплитуде и фазе. Численные расчеты свидетельствуют, что характеристики этого режима, в том числе форма модуляции, ее период и спектр, не зависят от начальных условий, так что его следует квалифицировать как автоколебательный. Будем использовать применительно к режимам такого рода термин автомодуляция. Согласно результатам расчетов, период автомодуляции вблизи порога ее возникновения составляет $\Delta\tau_m \approx 1,5$ и слабо меняется при вариации параметра L . В размерных величинах это соответствует $T_m = 1,5(l/v_0 + l/v_{\text{гр}})$.

Главную роль в возникновении автомодуляции играют два момента. Во-первых, распределенный механизм обратной связи: перенос возмущений осуществляется пучком со скоростью v_0 , и навстречу — с волной, со скоростью $v_{\text{гр}}$. Во-вторых, возникновение перегруппировки электронов. Амплитуда высокочастотного тока в пучке, движущемся в достаточно интенсивном поле волны, нарастает начиная от входа в пространство взаимодействия, достигает максимума, затем убывает, потом снова начинает нарастать, вновь достигает максимума, уже меньшей величины, и так далее. Такой характер зависимости связан с формированием электронных сгустков и их рассыпанием из-за различия скоростей электронов, попавших в разные фазы поля, что и имеют в виду, говоря о перегруппировке. Чем больше амплитуда воздействующего поля, тем ближе к входу в пространство взаимодействия она достигается.

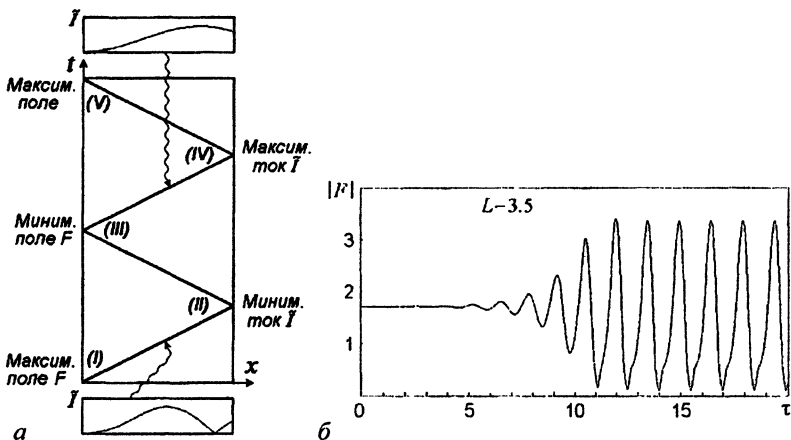


Рис. 8. Пространственно-временная диаграмма, поясняющая механизм автомодуляции (а); развитие неустойчивости стационарной одночастотной генерации в ЛОВ по отношению к малым возмущениям (б), приводящее в итоге к автомодуляционному режиму.

На рис. 8, а показана пространственно-временная диаграмма, поясняющая механизм автомодуляции. Пусть в некоторый момент I амплитуда высокочастотного поля на левом конце системы, где встреливается электронный пучок, относительно велика. Это приводит к перегруппировке электронов в пучке вдоль линии (характеристики) $x - v_0 t = \text{const}$: распределение амплитуды высокочастотного тока по длине выглядит так, как показано на нижней вставке. В результате амплитуда тока у правого конца в момент II оказывается относительно малой. Тогда на линии, отвечающей распространению волнового пакета справа налево с групповой скоростью — характеристике $x - v_0 t = \text{const}$, амплитуда поля будет меньше, поэтому в момент $t \cong L/v_0 + L/v_{гр}$ амплитуда сигнала на левом конце минимальна III . Меньшее поле более эффективно группирует пучок (верхняя вставка), поэтому на соответствующей характеристике достигается максимум тока у правого конца системы IV . В результате новый максимум поля V на левом конце возникнет через время $T \cong 2(L/v_0 + L/v_{гр})$. Это дает вполне удовлетворительную оценку периода автомодуляции: согласно численным расчетам фактор в этом соотношении оказывается 1,5 вместо 2.

При дальнейшем увеличении параметра L можно наблюдать усложнение характера автомодуляции, и в области $L \geq 4$ в опре-

деленных интервалах параметра она оказывается непериодической. Наконец, при достаточно больших величинах $L \geq 6$, соответствующих превышению тока над стартовым значением в 30 и более раз, изменение амплитуды и фазы сигнала во времени становится явно нерегулярным, хаотическим. Даже при наблюдении в течение очень большого времени не обнаруживается никакой тенденции к установлению режима с постоянной или периодически меняющейся амплитудой. Однако усредненные за время τ порядка нескольких единиц квадрат амплитуды $|F|^2$ и скорость изменения фазы сигнала становятся примерно постоянными, что позволяет говорить о достижении установившегося режима в статистическом смысле. Режим с теми же самыми статистическими характеристиками возникает и при других начальных условиях, так что его следует классифицировать как *хаотические автоколебания*. Как уже говорилось, при малых L реализуется устойчивое состояние равновесия, $F \equiv 0$. Возникновение режима генерации при $L = L_{\text{ст}} = 1,97327$ соответствует *бифуркации Андронова — Хопфа*. Неподвижная точка теряет устойчивость, и рождается предельный цикл. Имеет место режим генерации сигнала с постоянной амплитудой, величина которой зависит от превышения над порогом генерации по стандартному для бифуркации Андронова — Хопфа закону квадратного корня: $|F^0(0)| \propto \sqrt{L - L_{\text{ст}}}$. На плоскости $(\text{Re}F(0, \tau), \text{Im}F(0, \tau))$ портрет аттрактора представлен окружностью — проекцией предельного цикла системы на данную плоскость из бесконечномерного фазового пространства.

На рис. 8, б показано, как развитие неустойчивости стационарного состояния приводит в итоге к автомодуляционному режиму.

О дальнейших бифуркациях, наблюдаемых при увеличении параметра L , даст представление диаграмма, показанная на рис. 9, а. Для ее построения на компьютере при каждом значении L в установившемся режиме откладывались на графике значения максимумов амплитуды выходного сигнала.

Обсуждавшиеся до этого момента «простые» автомодуляционные режимы представлены одиночной линией вблизи левого края бифуркационной диаграммы. При $L = L_1 \approx 4$ эта линия раздваивается, что соответствует *бифуркации удвоения периода автомодуляции*: форма «пиков» повторяется через раз. Далее в узкой области изменения параметра $4 < L < 4,1$ имеют место дальнейшие бифуркации удвоения периода автомодуляции L_2, L_3, \dots и переход к хаосу. Соответствующий фрагмент бифуркационной диаграммы показан в увеличенном виде на рис. 9, б.

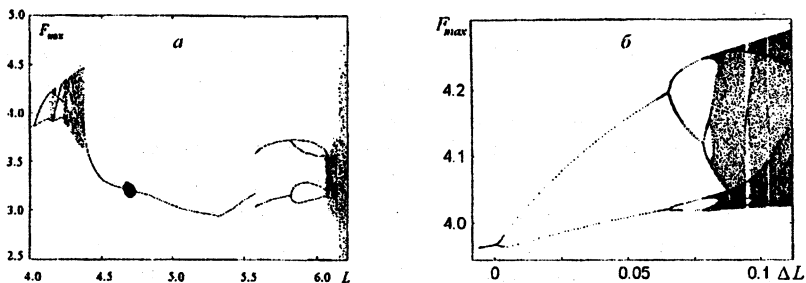


Рис. 9. Бифуркационная диаграмма модели ЛОВ (а), где по оси абсцисс отложен параметр нормированной длины L , по оси ординат — значения максимумов амплитуды выходного сигнала, которые достигаются в установившемся режиме и фрагмент бифуркационной диаграммы, отвечающий области перехода к хаосу через каскад бифуркаций удвоения периода автомодуляции (б) (по оси абсцисс — величина $\Delta L = L - L_1$, где $L_1 \approx 4$ отвечает первой бифуркации удвоения).

Вид бифуркационного дерева на рисунке 9, б с точностью до многочисленных деталей воспроизводит картину, известную для широкого класса нелинейных диссипативных систем как *сценарий перехода к хаосу по Фейгенбауму через бесконечный каскад бифуркаций удвоения периода*. Предельная точка этой последовательности — критическая точка Фейгенбаума — соответствует порогу возникновения хаоса. В закритической области можно видеть области хаоса и окна периодичности (например, хорошо заметны окна периода 5 и 7). Вблизи критической точки выполняются универсальные количественные закономерности, присущие сценарию Фейгенбаума, например, оценка по первым бифуркационным значениям L константы сходимости $\delta = (L_k - L_{k-1}) / (L_{k+1} - L_k)$ дает примерно 4,9, что близко к универсальному числу Фейгенбаума 4,6692. Отношения амплитуд субгармоник, возникающих в результате последовательных бифуркаций, близки к теоретическому значению 13,4 дБ.

Хаос, возникающий в результате последовательности бифуркаций удвоения периода автомодуляции, оказывается в известном смысле «не развитым», «слабым». Во-первых, он занимает весьма узкую область по параметру L . Во-вторых, зависимость амплитуды от времени выглядит как приблизительное повторение однотипных, близких по размеру «пиков» через примерно равные интервалы времени, а интенсивность шумовой составляющей в спектре невелика. Пространственное распределение поля не обнаруживает богатого разнообразия форм, какое должно было бы быть характерно для развитого хаоса — «турбулентности».

Вслед за областью удвоений периода автомодуляции в интервале $4,15 < L < 4,33$ присутствует большое количество окон периодичности в хаосе. В первом из этих окон имеет место автомодуляционный режим, в котором высота и форма максимумов повторяется через каждые три пика. В следующих окнах период повторения шаг за шагом увеличивается на единицу. (Такая последовательность «прибавлений периода» характерна для многих нелинейных систем, например, вблизи ситуации, ассоциирующей с наличием петли седлофокуса и применимостью теоремы Шильникова.) На рис. 10, а показан фрагмент бифуркационной диаграммы, соответствующий «области прибавлений периода». Выход из этой области с увеличением L сопровождается переходом от хаотической автомодуляции к периодической через перемежаемость. Зависимость амплитуды выходного сигнала от времени имеет вид прерываемых нерегулярными всплесками цугов почти периодической модуляции, длительность которых неограниченно растет по мере увеличения параметра L с приближением к критической величине $L_c \approx 4,33$. Зависимость средней длительности ламинарной фазы (в относительных единицах) от параметра L хорошо описывается функцией $T_{\text{лам}} \propto (L_c - L)^{-1/2}$, что характерно для перемежаемости 1-го рода по Помо и Манневиллю. Заметим, что переход через перемежаемость сопровождается значительным смещением частот в спектре. Безразмерная частота автомодуляции (разность между основной частотой и частотой сателлита) до перехода остается почти постоянной $\Omega_a \approx 1,33\pi$, затем скачком изменяется до $\Omega_a \approx 1,6\pi$ и далее монотонно увеличивается с ростом L .

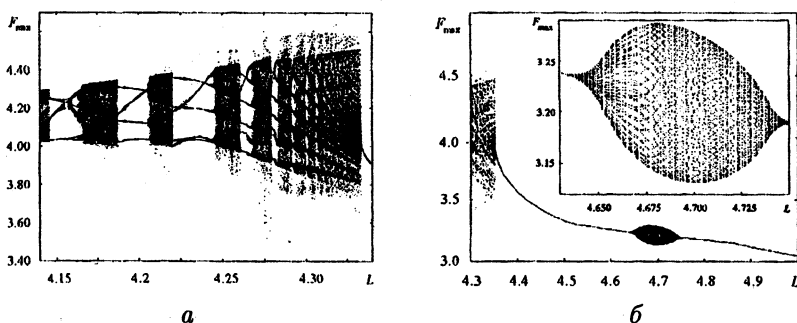


Рис. 10. Фрагменты бифуркационной диаграммы, отвечающие «области прибавления периода» (а) и области квазипериодической динамики (б).

В результате бифуркации при $L_c \approx 4,33$ хаотическая автомодуляция сменяется периодической. Затем она становится квазипериодической (в интервале примерно от 4,6 до 4,75), потом снова периодической. По-видимому, странный аттрактор в фазовом пространстве превращается здесь в метастабильное хаотическое множество. Об этом свидетельствует тот факт, что в ходе переходного процесса фазовая траектория на начальном участке ведет себя достаточно сложным, нерегулярным образом. Переходный процесс, предшествующий установлению регулярного автомодуляционного режима, может занимать длительное время, во много раз превышающее характерный период автомодуляции.

На рис. 10, б представлен фрагмент бифуркационной диаграммы для области квазипериодической автомодуляции. В этом режиме в спектре выходного сигнала присутствует дополнительная независимая частотная составляющая, не находящаяся в простом рациональном отношении с другими частотами. Однако существуют области квазипериодической автомодуляции и области, где наблюдается внутренняя синхронизация с образованием резонансных циклов на торе, когда частоты в спектре автомодуляции оказываются в рациональном соотношении. Наконец, при $L > 6$ наблюдаются режимы развитого хаоса, «турбулентная» динамика. При этом на проекции фазового портрета уже не различается какая-либо внутренняя структура. Спектр выходного сигнала становится сплошным, степень изрезанности формы спектра быстро убывает с ростом L .

В современной нелинейной динамике хорошо известно, что главным атрибутом динамического хаоса служит наличие чувствительности фазовых траекторий к малым вариациям начальных условий. На рис. 11 представлены результаты численного эксперимента, иллюстрирующего ее присутствие в динамике модели ЛОВ в области развитого хаоса. На графике наложены два десятка зависимостей амплитуды выходного сигнала от времени, полученных при задании в качестве начальных условий распределений амплитуды поля с малой случайной добавкой. Видно, что на начальном участке они совпадают, а с течением времени расходятся друг относительно друга все сильнее и сильнее: картина «замазывается».

Для режима развитого хаоса характерно, что зависимость амплитуды выходного сигнала от времени содержит нерегулярную последовательность «пиков» — максимумов разной высоты и формы. Из динамики распределения поля и тока по длине, после «ухода» (вследствие распространения с групповой скоростью) каждого очередного максимума возникает распределение поля по длине,

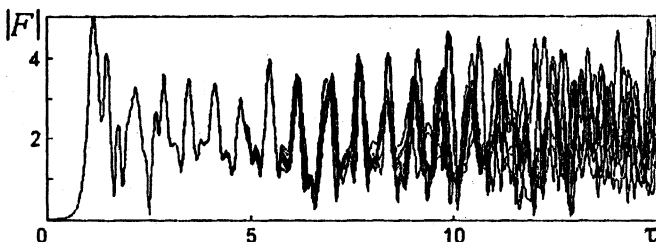


Рис. 11. Иллюстрация свойства чувствительности динамики к малому возмущению начальных условий: на графике наложены два десятка зависимостей амплитуды выходного сигнала от времени при $L = 6, 1$, причем в качестве начальных условий задавалось распределение амплитуды поля на уровне 0,01 со случайной добавкой порядка 0,001.

характеризующееся сравнительно небольшой и быстро осциллирующей амплитудой. Это поле наведено электронами, поступившими во время прохождения по системе максимума. Они оказываются сильно перегруппированными, и соответственно, излученное ими поле имеет амплитуду, относительно быстро меняющуюся вдоль системы. Мелкие детали этого распределения существенно влияют на высоту и время образования следующего максимума, формирующегося из одного из оставшихся «горбов». Таким образом, динамика оказывается чувствительной к малым вариациям начальных условий, что обуславливает хаотическую природу наблюдаемого режима.

2.3. Релятивистские эффекты. Если ограничиться случаем $C \ll 1$ ($C = \sqrt{\frac{I_0 K_0}{2U\gamma_0(1+\gamma_0)}}$, $\gamma_0 = 1/\sqrt{1-v_0^2/c^2}$ — релятивистский масс-фактор электронов на входе в пространство взаимодействия), то полная система уравнений принимает вид

$$\frac{\partial \epsilon}{\partial \xi} = \frac{\nu}{2} \operatorname{Re} F e^{j\theta}, \quad \frac{\partial \theta}{\partial \xi} = \frac{\epsilon^{-2} - 1}{\nu}, \quad \frac{\partial F}{\partial \tau} - \frac{\partial F}{\partial \xi} = -\frac{1}{\pi} \int_0^{2\pi} e^{-i\theta} d\theta_0,$$

$$\theta|_{\xi=0} = \theta_0, \quad \partial\theta/\partial\zeta|_{\xi=0} = 0, \quad F|_{\xi=L} = 0,$$

где введены следующие безразмерные переменные и параметры:

$$\xi = \beta_0 C x, \quad \tau = \omega_0 C (1 - x/v_0) (1 + v_0/v_{гр})^{-1},$$

$$F(\xi, \tau) = \frac{\tilde{E}}{2\beta_0 U C^2 \gamma_0 (1 + \gamma_0)},$$

$$L = \beta_0 l C = 2\pi C N, \quad \nu = 2C\gamma_0^2 = \frac{\gamma_0^2 L}{\pi N}, \quad \epsilon = 1 / \sqrt{1 + \nu \frac{\partial \theta}{\partial \xi}}.$$

В уравнениях присутствуют два независимых параметра — нормированная длина L и коэффициент ν , отвечающий за относительную величину релятивистских эффектов. При анализе динамики системы в качестве второго параметра, однако, предпочтительно использовать комбинацию ν/L , не зависящую от тока пучка, или приведенный релятивистский масс-фактор $\Gamma_0 = \gamma_0/\sqrt{N} = \sqrt{\nu\pi/L}$, где N — число длин волн на длине пространства взаимодействия.

При $\nu \ll 1$ безразмерная форма уравнения движения такая же, как в нерелятивистской теории. Однако выполнение этого условия само по себе еще не означает перехода к нерелятивистскому пределу, оно может быть справедливым и для релятивистской ЛОВ. В самом деле, параметр $\nu = 2C\gamma_0^2$ будет малым даже при больших значениях масс-фактора γ_0^2 , если достаточно мал модифицированный параметр Пирса ($C \ll \gamma_0^{-2}$). Данное замечание позволяет отнести теоретические результаты, полученные в рамках основной модели, к релятивистским ЛОВ, работающим в режимах с малым КПД, и создает основу для физического моделирования процессов в таких устройствах на нерелятивистских макетах (см. [2] и библиографию к лекции 13).

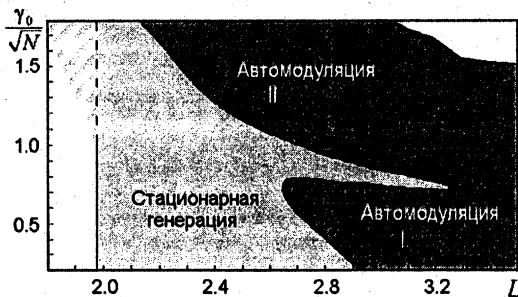


Рис. 12. Карта динамических режимов релятивистской ЛОВ на плоскости параметров L и γ_0/N , являющихся критериями подобия системы в асимптотике больших N . Белая область в левой части рисунка отвечает отсутствию генерации, чередующимися полосами белого и серого цвета показана область жесткого возбуждения. В белой области в правой верхней части диаграммы имеет место нарушение условий применимости уравнений («остановка электронов»).

На рис. 12 показана карта динамических режимов системы. По горизонтальной оси отложена нормированная длина L , по верти-

кальной — приведенный релятивистский масс-фактор γ_0/\sqrt{N} . Эти величины служат критериями подобия релятивистских ЛОВ в асимптотике больших N .

Граница области отсутствия колебаний на карте динамических режимов изображается вертикальной прямой, поскольку в принятой нормировке стартовое значение нормированной длины не зависит от релятивистского фактора, оставаясь таким же, как и в нерелятивистской теории, $L = L_{ст} = 1,97327$.

При этом, однако, в левом верхнем углу карты имеется область жесткого возбуждения, показанная чередующимися полосами белого и серого цвета. При задании начальных условий малой амплитуды колебания затухают, однако при достаточно большой амплитуде устанавливается стационарный режим генерации. Если представить себе систему, параметры которой можно медленно изменять, то будет наблюдаться гистерезис. При входе в «полосатую» область слева направо колебания будут отсутствовать — они возбуждятся только при достижении стартового значения параметра $L_{ст}$, а при входе справа налево будет сохраняться режим генерации.

Причина этого эффекта связана с нелинейностью, обусловленной релятивистской зависимостью массы от скорости. На начальном этапе процесса группирования половина частиц в пучке ускоряется и «тяжелеет», а вторая тормозится и уменьшает свою массу. «Легкие» электроны становятся более подвижными, что способствует росту эффективности группирования и относительному увеличению амплитуды высокочастотного тока. Иными словами, в системе присутствует нелинейный эффект, способствующий увеличению эффективности взаимодействия с ростом амплитуды.

Как можно заметить, в области умеренных L на карте представлены две зоны автомодуляции, что определенно указывает на наличие двух разных механизмов ее возникновения.

В нижней части диаграммы последовательность бифуркаций и наблюдаемые формы зависимостей амплитуды от времени аналогичны тем, которые имеют место в нерелятивистской ЛОВ. В частности, в области, обозначенной римской цифрой I , реализуется автомодуляция, механизм которой обусловлен группировкой — разгруппировкой электронов и эффектами запаздывания (см. выше). Понижение порога автомодуляции при увеличении релятивистского фактора связано, очевидно, с упоминавшимся возрастанием эффективности взаимодействия и, соответственно, перегруппировки из-за появления в пучке фракции более «легких» частиц.

При больших значениях параметра γ_0/\sqrt{N} возникает автомодуляция другого типа, обозначенная на карте цифрой II. Одной из ее особенностей является существенно меньший период.

Чтобы прояснить механизм возникновения автомодуляции типа II, начнем с замечания, что в процессе группирования электронный сгусток формируется как частицами, уменьшившими скорость, так и теми, которые ускорились полем волны. В слаборелятивистском случае движение обеих фракций приблизительно симметрично относительно центра сгустка. Поэтому его разрушение в области перегруппировки хорошо выражено: амплитуда тока падает почти до нуля. Это существенный момент, лежащий в основе возникновения автомодуляции I. В сильно релятивистском случае ускоренные электроны заметно увеличивают, а замедленные уменьшают массу. Далее процесс группирования протекает так, что «тяжелые» частицы мало изменяют скорость в поле волны, а у «легких», напротив, скорость меняется значительно. Выраженная асимметрия между ускоренными и замедленными частицами приводит к тому, что полной разгруппировки не происходит: амплитуда тока в минимуме падает лишь где-то до половины максимальной величины. В этой ситуации автомодуляция I подавляется, с чем связано наличие на карте «отростка» области стационарной генерации, простирающегося до значений нормированной длины около 3,2. В области выше него реализуется другой механизм автомодуляции, которой поясняется на рис. 13. Когда амплитуда высокочастотного поля на левом (выходном) конце системы относительно велика, вдоль линии движения пучка $x - v_0 t = \text{const}$ образуется «двугорбое» распределение тока, показанное на нижней вставке. Поле, излученное электронами в области второго максимума, распространяясь с групповой скоростью, порождает следующий максимум амплитуды на левом конце. Когда же амплитуда на выходном конце минимальна, возникает распределение тока с минимумом амплитуды у правого конца системы (верхняя вставка), и это приводит к появлению следующего минимума на левом конце. Оценка периода автомодуляции такого типа дает $T \cong 4(L/v_0 + L/v_{гр})$ в разумном соответствии с численными расчетами.

Характерным свойством переходного процесса в релятивистской ЛОВ, особенно при достаточно больших значениях фактора γ_0/\sqrt{N} , является присутствие первого интенсивного всплеска амплитуды излучения на выходе. Уровень мощности в момент всплеска в несколько раз превышает стационарный уровень. Объяснение этого феномена состоит в том, что максимум амплитуды, формирующийся на

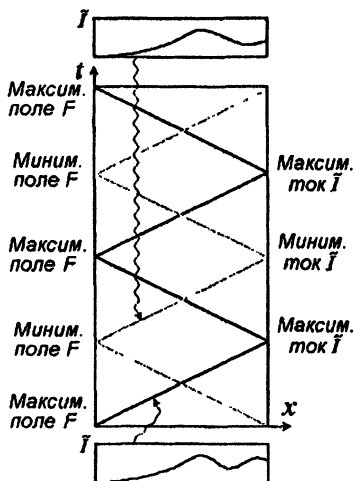


Рис. 13. Пространственно-временная диаграмма, поясняющая механизм автомодуляции в ЛОВ и ультрарелятивистской области.

начальной стадии процесса у коллекторного конца системы в области развитой группировки пучка, распространяется затем навстречу пучку с групповой скоростью. Встречая на своем пути слабо возмущенные участки пучка, он вызывает интенсивное формирование сгустков, излучение от которых добавляется и еще более увеличивает интенсивность всплеска. Фактически локализованный всплеск аккумулирует энергию электронов, поступивших в пространство взаимодействия, на протяжении всего времени его прохождения по длине пространства. Эффект усиливается благодаря специфике процесса группирования в релятивистском пучке, в котором, как отмечалось, в области слабой нелинейности имеет место увеличение интенсивности взаимодействия. Феномен формирования мощного импульса излучения в ходе переходного процесса находит применение для реализации эффективных импульсных режимов создаваемых релятивистских ЛОВ.

3. Диод Пирса: от регулярных автоколебаний к хаосу. Виркатор и другие

3.1. Диод Пирса и неустойчивость Пирса. Наиболее простой электронной моделью со сложной динамикой является диод

Пирса [9—13]. В простейшей модели диода одномерный моноскоростной электронный поток с постоянной плотностью пространственного заряда, нейтрализованный неподвижным ионным фоном, движется между заземленными сетками. Единственным бифуркационным параметром задачи является параметр Пирса $\alpha = \omega_p L / v_0$ — невозмущенный угол пролета электронов по плазменной частоте ω_p (v_0 — скорость электронов на входе в межсеточное пространство, L — расстояние между сетками). Неустойчивость в диоде Пирса возникает при $\alpha > \pi$ и определяется существованием внешней обратной связи через цепь, соединяющую сетки, поскольку они заземлены [2, 9].

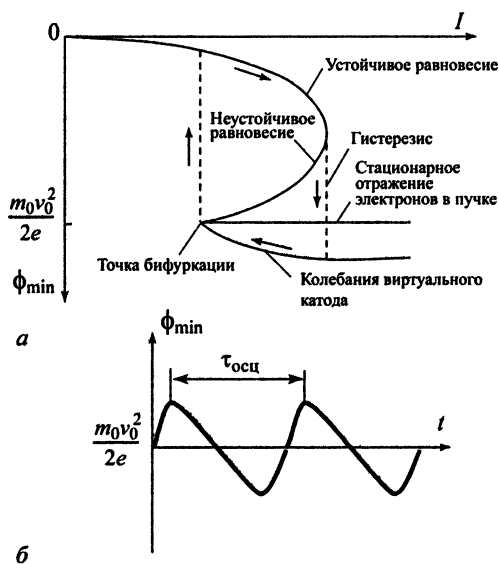


Рис. 14. Величина минимума потенциала ϕ_{\min} в пространстве взаимодействия как функция тока I инжектируемого пучка (а) и временная зависимость величины минимума потенциала поля пространственного заряда (б).

На рис. 14, а представлена зависимость величины минимума потенциала в диодном промежутке от тока I инжектируемого в него пучка со скоростью v_0 . Из рисунка видно, что с увеличением тока I глубина «провисания» потенциала ϕ_{\min} в пучке становится все больше, и при некотором критическом значении тока I_{SCL} величина ϕ_{\min} переходит скачком с устойчивой ветви, соответствующей полному стационарному прохождению потока через диодный промежуток, на

ветвь, которая соответствует колеблющемуся виртуальному катоду (ВК).

При токах пучка больше критической величины I_{SCL} высота потенциального барьера в диодном промежутке ВК становится больше кинетической энергии влетающих в пространство взаимодействия электронов. В этом случае электроны останавливаются перед ВК и разворачиваются, что приводит к смещению ВК и максимума плотности пространственного заряда в сторону плоскости инжекции. Кроме того, величина плотности пространственного заряда быстро увеличивается, так как практически все электроны оказываются захваченными движущимся к входной плоскости ВК, образующим сгусток электронов. По мере приближения к плоскости инжекции величина потенциального барьера уменьшается и в некоторый момент времени становится меньше кинетической энергии влетающих электронов. Снижение величины потенциального барьера продолжается и после того, как отраженные электроны покидают пространство взаимодействия. Инжектируемые электроны теперь легко преодолевают уменьшившийся потенциальный барьер и двигаются к выходной плоскости промежутка. ВК начинает смещаться к центру диода до тех пор, пока не восстановится потенциальный барьер достаточной высоты для отражения электронов. Далее процесс повторяется. Ток отраженных от ВК и пролетных электронов оказывается промодулированным на частоте $\omega_{ВК}$ колебаний ВК. Сказанное можно иллюстрировать и временной зависимостью величины минимума потенциала поля пространственного заряда в области дрейфа (рис. 14, б). Колебания потенциала напоминают релаксационные: нарастание происходит почти в 3 раза быстрее, чем спад. Положение минимума потенциала также осциллирует относительно своего стационарного положения. Частота осцилляций потенциала также определяется частотой колебаний ВК: $\omega_{осц} = 1/\tau_{осц} = \omega_{ВК}/2\pi$ и зависит от плазменной частоты ω_p нелинейно, имея тенденцию к увеличению с ростом тока пучка.

Вернемся к рис. 14, а. Если теперь в состоянии, когда в системе имеется колеблющейся ВК, начать уменьшать ток пучка I , то амплитуда колебаний $\Delta\phi_{min}(t)$ уменьшается, а координата ϕ_{min} в пространстве стремится к центру диодного промежутка. Состояние системы возвращается к прежнему устойчивому равновесию при достижении точки бифуркации. Таким образом, в окрестности $I \sim I_{SCL}$ существует гистерезисная петля, которая обозначена на рис. 14, а штриховой линией. За точкой бифуркации существует четыре возможных состояния, два из которых устойчивы и реализуются в зависимости

от предыстории, а остальные два — неустойчивы и физически нереализуемы.

Наиболее простой и фундаментальной моделью, которая позволяет изучить особенности физических процессов в потоке с ВК, является модель плоского диодного промежутка, в который инжектируется сверхкритический ток $I > I_{SCL}$ (модель диода Пирса и ее различные модификации). Впервые численно такая система была исследована в 60-х годах Бриджесом и Бёдселлом [14, 15]. Это был один из важных, наряду с работой по моделированию динамики электронного пучка в магнетронном диоде [16], вычислительных экспериментов в электронике. В нем сразу же была обнаружена нестационарная динамика электронного потока со сверхкритическим током в диодном промежутке. К достоинствам модели диода Пирса следует отнести простоту, легкость численного моделирования, возможность получения аналитических результатов, общность получаемых результатов². В последнее время активно исследуются более реалистичные модели генераторов на ВК (виркаторов [20, 21]). Обычно такие исследования проводятся численно в рамках электромагнитного моделирования в двух или даже трех измерениях (см., например, [2, 22, 23]).

3.2. Хаотические колебания виртуального катода в диоде Пирса. Для анализа нестационарной нелинейной динамики ВК в плоской геометрии необходимо использовать численное моделирование методом «частиц в ячейке» [2, 22]. В этом случае электронный поток представляется в виде совокупности крупных частиц (заряженных листов), динамика которых описывается нерелятивистскими уравнениями движения. Поле пространственного заряда находится с помощью численного решения уравнения Пуассона.

Рассмотрим динамику электронного потока с ВК в диоде Пирса при изменении управляющего параметра α [24–26]. На рис. 15 показаны характерные динамические состояния колебаний ВК при раз-

² Отметим, что при значении параметра Пирса $\alpha \sim 3\pi$ в диоде Пирса наблюдается возникновение колебаний в пучке со сверхкритическим током без образования ВК. В этом случае возможно описание процессов в системе в рамках гидродинамического приближения [2]. В работах [10–12] было показано, что при уменьшении параметра Пирса в диапазоне $\alpha \in (2, 8575\pi, 3\pi)$ наблюдаются бифуркации, которые приводят к возникновению хаотического аттрактора по сценарию Фейгенбаума. В работах [13, 17–19] было проведено исследование возможности управления сложными колебаниями в гидродинамической модели диода Пирса с помощью подключения внешней запаздывающей обратной связи и на основе идей стабилизации хаотических колебаний с помощью непрерывной обратной связи.

личных значениях параметра Пирса. Для каждого состояния представлены спектр мощности, проекция восстановленного аттрактора и часть временной реализации колебаний напряженности электрического поля на выходе системы.

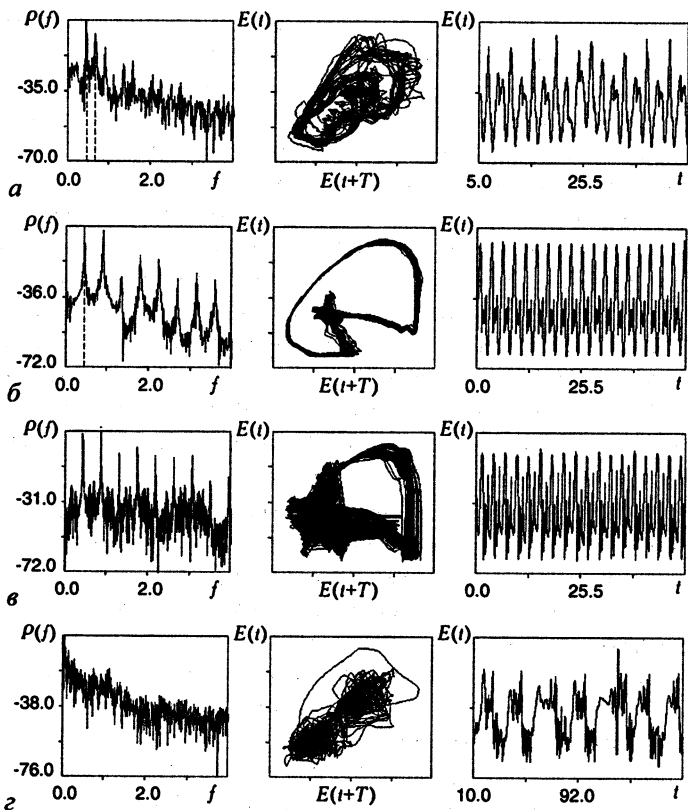


Рис. 15. Характеристики колебаний для следующих значений параметра Пирса: $\alpha/\pi = 1,20$ (а); $1,50$ (б); $1,76$ (в); $2,60$ (г). Штриховыми линиями в спектрах отмечены базовые частоты.

Сразу после превышения параметром Пирса предельного значения в потоке возникают нерегулярные колебания ВК (рис. 15, а). В спектре можно выделить две частоты (отмечены на рис. 15, а штриховыми линиями), находящиеся в иррациональном соотношении, при этом шумовой пьедестал в спектре имеет значительную высоту. Фазовый портрет соответствует сильно размытому предель-

ному циклу. При значении $\alpha > 1,3\pi$ поведение системы становится близким к регулярному, аттрактор превращается в узкую ленту, в спектре мощности выделяются только основная частота и ее гармоники (рис. 15, б). На фазовом портрете появляется неоднородность: в области, соответствующей отсутствию ВК в системе, фазовая траектория проводит относительно больше времени. Внутри каждого периода колебаний в реализации можно выделить промежуток почти неизменных значений поля и следующий за ним острый пик.

Дальнейший рост α приводит к расширению ленты аттрактора, что соответствует нерегулярности колебаний в электронном пучке. В спектре мощности при этом растет уровень шумового пьедестала и расплываются отдельные спектральные пики (рис. 15, в). При $\alpha > 1,86\pi$ нерегулярность колебаний резко уменьшается. Аттрактор представляет узкую ленту сложной формы. В спектре мощности можно выделить большое число высших гармоник, имеющих почти одинаковую с гармоникой основной частоты амплитуду, что свидетельствует о сильной нелинейности процессов в системе. При $\alpha > 2,48\pi$ в системе возникают шумоподобные колебания, которые на фазовом портрете представляют собой блуждание изображающей точки между двумя притягивающими центрами (рис. 15, г).

Для изучения физических процессов, приводящих к хаотизации колебаний, рассмотрим пространственно-временную диаграмму электронного потока для одного периода колебаний ВК (см. рис. 6). Буквами на ней отмечены различные характерные траектории заряженных частиц в пучке. Когда плотность заряда в области ВК достаточна, ВК перестает существовать, поскольку часть потока отражается от него (на рис. 6 эта часть потока обозначена буквой А), а другая часть, затормозившаяся в области ВК, уходит к выходной плоскости, создавая пролетный ток (часть потока, обозначенная символом В). В процессе распада ВК плотность заряда в нем падает, а следовательно, уменьшается напряженность создаваемого им тормозящего поля. Поэтому на электроны, поступающие в систему позже, действует меньшее тормозящее поле, и их скорость изменяется на меньшую величину, чем у влетевших ранее. В результате образуется сгусток электронов за ВК, проходящий по системе к выходной сетке. На пространственно-временной диаграмме формирования вторичного сгустка соответствует сближение кривых (область С). Образование вторичного сгустка приводит ко второму увеличению плотности заряда в потоке.

Рассмотрим зависимость величины $\rho_{\max 2}$ второго максимума плотности заряда от параметра Пирса, представленную на рис. 16

(сплошная линия). Величина второго максимума определяет характер динамики электронного потока: хаотическим режимам соответствует значение плотности заряда во втором сгустке большее 20, а его уменьшение приводит к слаберегулярным колебаниям. Анализ пространственно-временных диаграмм электронного потока показывает, что при значениях $\rho_{\max 2} > 20$ в пучке с ВК возникают отражения части потока от формирующегося вторичного сгустка.

Характеристикой сложности колебаний в электронном пучке в диоде Пирса является значение λ максимальной ляпуновской экспоненты. Рост амплитуды плотности заряда во втором максимуме приводит к увеличению величины λ (рис. 16, штриховая линия). Из рис. 16 видно, что рост плотности вторичного сгустка $\rho_{\max 2}$ сопровождается ростом ляпуновской экспоненты λ . Качественное различие между двумя хаотическими режимами связано с тем, насколько плотность пространственного заряда во втором сгустке превышает критическое значение. В первом режиме нерегулярных колебаний величина $\rho_{\max 2}$ настолько велика, что отраженный от него поток существует постоянно, а во втором режиме превышение небольшое и отраженный поток возникает не на каждом периоде колебаний.

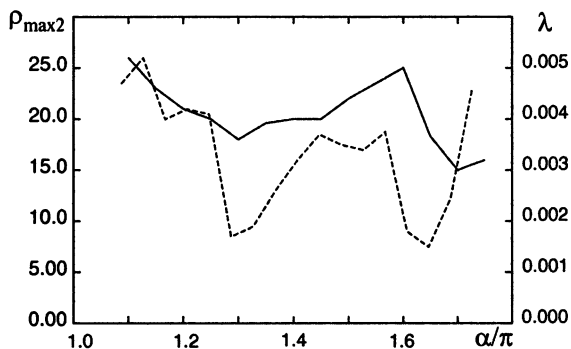


Рис. 16. Зависимость плотности заряда во вторичном сгустке (сплошная линия) и максимальной ляпуновской экспоненты (штриховая линия) от параметра Пирса α .

Значение $\rho_{\max 2}$ определяется амплитудой потенциала поля пространственного заряда в области ВК и расстоянием между входной сеткой и местом формирования ВК. При увеличении параметра Пирса величина плотности пространственного заряда в области ВК уменьшается, и для того чтобы остановить электрон, достаточно

меньшего уплотнения электронного потока. Это уменьшение наиболее резко происходит при увеличении α от π до $1,3\pi$, что объясняет первое уменьшение плотности заряда во втором максимуме и возникновение слабого нерегулярного режима при $\alpha \approx 1,3\pi$. Расстояние между входной сеткой и ВК также уменьшается с ростом α , но неравномерно с ростом надкритичности. Это расстояние сильно уменьшается при $\alpha \approx 1,7\pi$, что приводит к снижению плотности заряда во втором сгустке, так как он не успевает сформироваться.

Таким образом, возникновение хаоса в диоде Пирса определяется взаимодействием двух электронных структур (ВК и вторичного сгустка, формирование которого определяется кинематическим эффектом — превращением модуляции по скорости в области ВК в модуляцию по плотности). Усложнение хаотических колебаний связано с увеличением связи между этими электронными структурами.

3.3. Хаос и структуры в электронном пучке с виртуальным катодом: электромагнитное моделирование. Опишем теперь механизм развития хаотических колебаний в электронном пучке со сверхкритическим током в виркаторе. В этом случае оказывается важным рассмотрение неоднородных моделей динамики пучка в области ВК. В качестве модели виркатора рассмотрим отрезок цилиндрического волновода, помещённый в продольное магнитное поле (МП). Через сечение $z = 0$ волновода (плоскость инжекции) внутрь системы поступает моноскоростной трубчатый электронный поток с релятивистским фактором γ_0 . Рассмотрение физических процессов в такой системе будем проводить в рамках электромагнитного моделирования [2, 22, 23].

Изучим усложнение колебаний ВК в виркаторе при увеличении параметра тока пучка $\alpha = I/I_{SCL}$. Будем предполагать, что величина ведущего МП достаточно велика, так что движение электронов можно рассматривать близким к одномерному. Хаотическая динамика и процессы образования и взаимодействия когерентных структур в электронном пучке со сверхкритическим током в такой системе были детально исследованы в работах [27, 28].

Численное исследование показывает, что с увеличением тока пучка α наблюдается усложнение колебаний в электронном потоке. На рис. 17 приведены временные реализации, спектры мощности и фазовые портреты колебаний тока пучка в области ВК для различных режимов генерации.

При малых значениях α ($\alpha < 1,7$) в системе устанавливаются регулярные колебания релаксационного типа (рис. 17, а). Спектр мощности содержит узкие пики, являющиеся кратными гармоника-

ми основной частоты. С увеличением α происходит разрушение периодических колебаний, и последовательно с ростом надкритичности появляются два типа хаотического поведения. В первом случае ($1,7 < \alpha < 3$), как видно из рис. 17, б, хаотическая динамика появляется на базе одного неустойчивого предельного цикла, соответствующего динамике системы в регулярном режиме при меньшем токе пучка. Во втором случае ($\alpha > 3$; рис. 17, в) фазовый портрет колебаний более однороден, структура аттрактора сложна, спектр мощности сильно зашумлен.

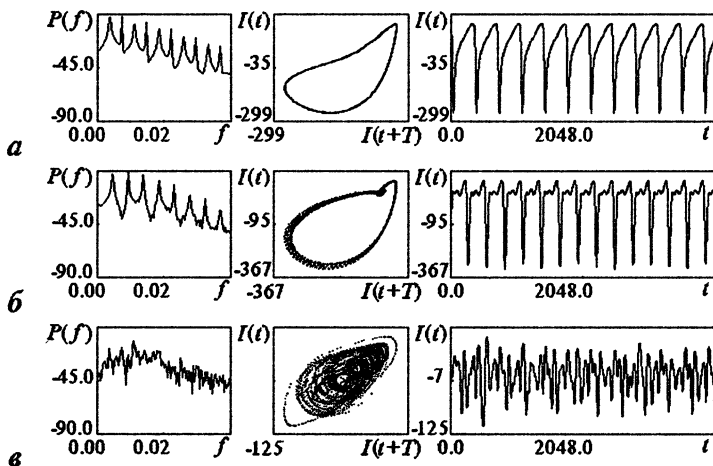


Рис. 17. Характеристики колебаний тока $I(t)$ пучка в области ВК для различных динамических режимов: $\alpha = 1,4$ (а); 2 (б); 4 (в).

Сложная динамика электронного потока определяется формированием в пространстве дрейфа нескольких областей отражения заряженных частиц (нескольких «виртуальных катодов»), которые связаны между собой через отраженные от них электроны [27, 28].

Это иллюстрирует функция распределения $\Phi(\tau)$ электронов по временам жизни τ в пространстве взаимодействия (рис. 18). В регулярном режиме (рис. 18, а, сплошная линия) $\Phi(\tau)$ имеет двугорбый вид. Площадь под кривой пропорциональна числу существующих в потоке пролётных и отражённых к плоскости инжекции частиц. Максимум отраженных частиц соответствует единственной структуре в электронном пучке — ВК. С увеличением α отражённые частицы начинают доминировать в общем числе инжектируемых частиц (рис. 18, а, пунктирная линия), область возможных времён жизни

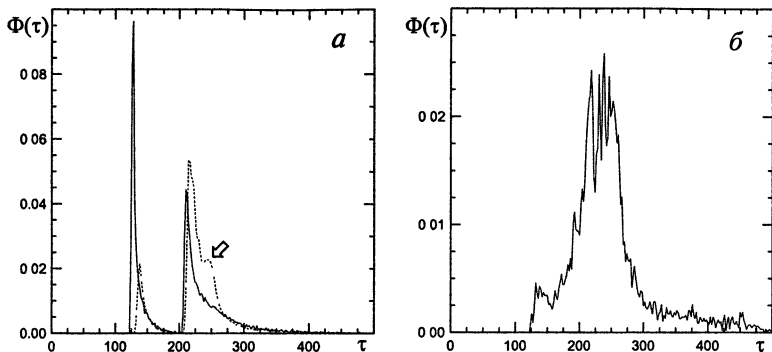


Рис. 18. Распределение электронов по временам жизни в пространстве взаимодействия для следующих параметров: сплошная линия — $\alpha = 1, 4$, пунктирная линия — $\alpha = 2$ (а); $\alpha = 4$ (б).

отражённых частиц увеличивается. За счёт появления долгоживущих частиц в потоке возникает внутренняя распределённая обратная связь, обеспечивающая связь между основной структурой (ВК) и возникающей вторичной структурой, которой соответствует третий максимум на кривой $\Phi(\tau)$ (отмечен стрелкой на рис. 18, а). Однако при небольшой надкритичности эффективность этой связи мала (общее количество частиц, отражённых от вторичного ВК, мало) и движение слабонерегулярно (в фазовом пространстве наблюдается размытый предельный цикл). Для развитого хаоса (рис. 18, б) характерна сильно изрезанная форма $\Phi(\tau)$, которая позволяет выделить несколько примерно равных по количеству групп заряженных частиц с различными временами жизни. Это означает, что в системе формируется несколько ВК (несколько пространственно-временных структур) на различном расстоянии от плоскости инжекции. Отражение части потока от каждого ВК оказывает влияние на условия формирования других структур, что обеспечивает несколько петель обратной связи с различными временами запаздывания. Такая распределённая связь между всеми структурами в потоке приводит к сильно нерегулярной динамике системы при большой надкритичности α . Заметим, что аналогичные результаты демонстрирует и анализ внутренней структуры электронного потока с помощью декомпозиции пространственно-временных данных по алгоритму Каруна — Лоэва (КЛ) [28].

Рассмотрим теперь динамику электронного пучка в виркаторе при изменении величины B ведущего МП. Будем исследовать вир-

каторную систему, представляющую собой замкнутый отрезок цилиндрического волновода длиной $L = 0,16$ м и радиусом $R = 0,3L$. Аксиально-симметричный трубчатый релятивистский пучок, инжектируемый в волноводную камеру, характеризуется нулевой поперечной скоростью и однородным распределением плотности тока в поперечном направлении. Энергия ускорения электронов 560 кэВ, внутренний радиус и толщина инжектируемого релятивистского пучка соответственно $r_b = 0,2L$ и $\Delta_b = 0,03L$. В качестве источника внешнего МП предполагается соленоид цилиндрического сечения, создающий однородное МП в пределах рабочей камеры виркатора.

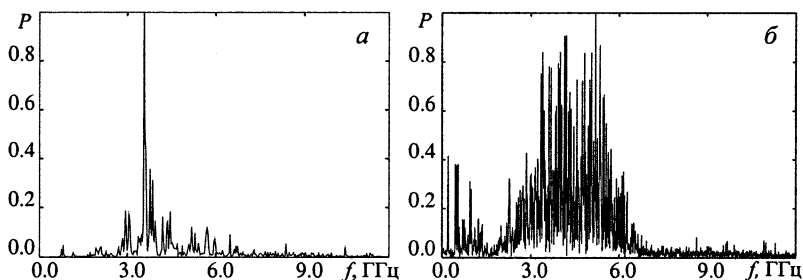


Рис. 19. Спектры колебаний продольного электрического поля в области ВК при одинаковом токе пучка $\alpha = 1,5$ для различных значений величины ведущего МП: 40 кГц (а) и 0 (б).

На рис. 19 показаны спектры мощности колебаний продольного электрического поля в области ВК при токе пучка $\alpha = 1,5$ и при различных значениях величины B ведущего магнитного поля. При значительных величинах B колебания в системе оказываются близкими к одночастотным (рис. 19, а; $B = 40$ кГц). В спектре мощности наблюдается один ярко выраженный пик, соответствующий частоте $f \approx 2,5(\omega_p/2\pi)$. С уменьшением величины ведущего МП наблюдается усложнение спектра генерации, заключающееся в появлении дополнительных частотных компонент и увеличении мощности шумового пьедестала до уровня $-30 \div -20$ дБ. При малых величинах B в спектре мощности наряду с ВЧ-компонентами в области 3–6 ГГц появляются НЧ-компоненты с частотами $f < 500$ МГц. При отсутствии ведущего МП спектр генерации становится сплошным в диапазоне 3,5–6 ГГц, причем шумовой пьедестал поднимается до уровня -10 дБ (рис. 19, б). Изрезанность хаотического спектра не превышает 3–10 дБ.

Таким образом, при фиксированном токе пучка с уменьшением величины фокусирующего МП в системе с ВК наблюдается последовательное усложнение спектра генерации начиная от одночастотных колебаний при близкой к одномерной динамике электронов до многочастотных широкополосных хаотических колебаний. При этом при малых $B < 5$ кГс динамика пучка оказывается принципиально двумерной, в частности в системе наблюдается интенсивное токооседание электронов на стенки рабочей камеры.

Выше было показано, что хаотизация выходного излучения с ростом тока пучка определяется формированием нескольких электронных структур (ВК) в пространстве дрейфа. Каждый из ВК воздействует на другие структуры через отраженный от него электронный поток. Как следствие, в системе формируется несколько цепей обратной связи, которые и приводят к усложнению динамики пространственного заряда электронного пучка со сверхкритическим током. При уменьшении величины МП наблюдается аналогичная ситуация, но теперь электронные структуры формируются не только в различных продольных сечениях пространства взаимодействия, но и имеется сложная картина динамики пучка в поперечном направлении.

Таким образом, усложнение динамики электронного пучка с ВК с увеличением тока пучка или уменьшением величины ведущего МП определяется одним и тем же физическим механизмом, а именно формированием и взаимодействием пространственных структур в пучке с ВК. Однако реализация данного механизма в каждом из случаев различна. Так, при уменьшении величины ведущего МП основную роль играет динамика пучка в поперечном направлении (т. е. двумерные эффекты), и, в первую очередь, начинается рост «поперечных» электронных структур в пучке со сверхкритическим током.

Заключение

В статью вошли не все даже «любимые» результаты. Не изложены результаты экспериментального исследования ЛОВ, которые хорошо согласуются с представленной в статье теорией. Опущены данные о влиянии пространственного заряда, потерь и отражений в замедляющей системе на динамику ЛОВ. В частности, не обсужден изящный результат перехода абсолютной неустойчивости в конвективную при определенном значении параметра диссипации. Не рассмотрены различные модификации ЛОВ, такие как ЛОВ со связан-

ными замедляющими системами, гирогенератор со встречной волной, ЛОВ на аномальном доплеровском эффекте, ЛОВ с поперечным полем, пениотрон на встречной волне, а также методы подавления автомодуляции, различные модификации генераторов на виртуальном катоде.

Что еще привнесла нелинейная динамика в СВЧ-электронику?

Предложена и подробно изучена феноменологическая модель электронной турбулентности (заметим, что в нелинейной динамике такие модели занимают важное место). Предполагается, что существует последовательность электронных структур, оказывающих друг на друга взаимное влияние. Под электронной структурой понимается в данном случае малый объем электронной среды, состоящий из электронов-осцилляторов (нелинейных, неизохронных осцилляторов). Для каждого электронного ансамбля имеет место сверхизлучение, возникающее, как известно, в результате взаимодействия осцилляторов через поле собственного излучения, являющегося для них фазирующим. Проведен анализ усложнения пространственно-временной динамики процесса в потоке взаимно связанных электронных структур; изучены также возможности управления динамикой такого потока с помощью расположенных на его пути объемных резонаторов [25].

В работе [29] предложена радиотехническая модель — аналог генератора на вакуумном микротриоде под внешним гармоническим воздействием и его математическая модель. Показано существование сложных режимов, включая хаотические, и продемонстрировано качественное соответствие результатов вычислительного и натурального эксперимента. Подобный подход типичен для нелинейной динамики.

Примеры использования методов и идей нелинейной динамики в СВЧ-электронике можно продолжить.

Авторы благодарят О. И. Москаленко и П. В. Попова за помощь в подготовке рукописи.

Работа выполнена при поддержке научно-образовательного центра «Нелинейная динамика и биофизика» при Саратовском госуниверситете им. Н. Г. Чернышевского (грант REC-006 of U. S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF)) и Программы поддержки ведущих научных школ РФ. А. Е. Х. благодарит также за финансовую поддержку Фонд «Династия» и Международный центр фундаментальной физики (г. Москва).

Литература

1. *Вайнштейн Л. А., Солнцев В. А.* Лекции по сверхвысокочастотной электронике. М.: Сов. радио, 1973.
2. *Трубецков Д. И., Храмов А. Е.* Лекции по сверхвысокочастотной электронике для физиков. Т. 1. М.: Наука: Физматлит, 2003.
3. *Рыскин Н. М.* Уединенные волны пространственного заряда // Изв. вузов. Прикладная нелинейная динамика. 1994. Т. 2, № 5. С. 93—100.
4. *Моносов Г. Г.* Исследование электромагнитных шумов в магнетроне с помощью двумерного компьютерного моделирования // Современные проблемы электроники и радиофизики СВЧ: Мат. межд. межвузовской конф. Саратов: Изд-во ГосУНЦ «Колледж», 2001. С. 114.
5. *Rizopoulos S. A.* Feedback-induced noise in crossed field devices // IEEE Trans. Plasma Sci. 1992. V. 20, № 3. P. 60.
6. *Рыскин Н. М., Трубецков Д. И.* Нелинейные электронные волны. Методы и результаты для приборов О-типа // Радиотехника и электроника. 1993. Т. 38, № 2. С. 193—206.
7. *Кислов В. Я.* Теоретический анализ шумоподобных колебаний в электронно-волновых системах и автогенераторах с запаздыванием // Лекции по электронике СВЧ и радиофизике. 5-я зимняя школа-семинар. Кн. 5. Саратов: Изд-во Сарат. ун-та, 1981. С. 78.
8. *Гапонов А. В., Петелин М. И., Юлпатов В. К.* Индуцированное излучение возбужденных классических электронов-осцилляторов и его использование в высокочастотной электронике // Изв. вузов. Радиофизика. 1967. Т. 10, № 9—10. С. 1414—1473.
9. *Pierce J.* Limiting currents in electron beam in presence ions // J. Appl. Phys. 1944. V. 15. P. 721.
10. *Godfrey B. B.* Oscillatory nonlinear electron flow in Pierce diode // Phys. Fluids. 1987. V. 30. P. 1553.
11. *Kuhn S., Ender A.* Oscillatory nonlinear flow and coherent structures in Pierce-type diodes // J. Appl. Phys. 1990. V. 68. P. 732.
12. *Анфиногентов В. Г., Трубецков Д. И.* Хаотические колебания в гидродинамической модели диода Пирса // Радиотехника и электроника. 1992. Т. 37. С. 2251.
13. *Hramov A. E., Rempen I. S.* Investigation of the complex dynamics and regime control in Pierce diode with the delay feedback // Int. J. Electronics. 2004. V. 91, № 1. P. 1—12.
14. *Bridges W. B., Birdsall C. K.* Space-charge instabilities in electron diodes. I // J. Appl. Phys. 1961. V. 32. P. 2611.
15. *Bridges W. B., Birdsall C. K.* Space-charge instabilities in electron diodes. II // J. Appl. Phys. 1963. V. 34, № 10. P. 2946.
16. *Vuneman O., Yu S. P., Koyers G. P.* Time-dependent computer analysis of electron-wave interaction in crossed fields // J. Appl. Phys. 1965. V. 35, № 8. P. 2550—2559.

17. *Короновский А. А., Ремпен И. С., Храмов А. Е.* Исследование неустойчивых периодических пространственно-временных состояний в распределённой автоколебательной системе со сверхкритическим током // Изв. РАН. Сер. физическая. 2003. Т. 67, № 12. С. 1705—1708.
18. *Ремпен И. С., Короновский А. А., Храмов А. Е.* Управление хаосом в электронном пучке со сверхкритическим током в гидродинамической модели диода Пирса // Письма в ЖТФ. 2003. Т. 29, № 23. С. 67—74.
19. *Храмов А. Е., Ремпен И. С.* Влияние обратной связи на сложную динамику в гидродинамической модели диода Пирса // Радиотехника и электроника. 2002. Т. 47, № 6. С. 732—738.
20. *Рухадзе А. А., Столбцов С. Д., Тараканов В. П.* Виркаторы (обзор) // Радиотехника и электроника. 1992. Т. 37, № 3. С. 385.
21. *Дубинов А. Е., Селемир В. Д.* Электронные приборы с виртуальным катодом // Радиотехника и электроника. 2002. Т. 47, № 6. С. 575.
22. *Birdsall C. K., Langdon A. B.* Plasma physics, via computer simulation. NY: McGraw-Hill, 1985.
23. *Anderson T. M., Mondelli A. A., Levush B., Verboncoeur J. P., Birdsall C. K.* Advances in modelling and simulation of vacuum electron devices // Proc. IEEE. 1999. V. 87, № 5. P. 804—839.
24. *Анфиногентов В. Г.* Хаотические колебания в электронном потоке с виртуальным катодом // Изв. вузов. Прикладная нелинейная динамика. 1994. Т. 2, № 5. С. 69.
25. *Trubetskov D. I., Mchedlova E. S., Anfingentov V. G., Ponomorenko V. I., Ryskin N. M.* Nonlinear waves, chaos and patterns in microwave devices // Chaos. 1996. V. 6, № 3. P. 358.
26. *Трубецков Д. И., Храмов А. Е.* Лекции по сверхвысокочастотной электронике для физиков. Т. 2. М.: Наука: Физматлит, 2004.
27. *Анфиногентов В. Г., Храмов А. Е.* К вопросу о механизме возникновения хаотической динамики в вакуумном СВЧ-генераторе на виртуальном катоде // Изв. вузов. Радиофизика. 1998. Т. 41, № 9. С. 1137—1146.
28. *Храмов А. Е.* Хаос и образование структур в электронном потоке с виртуальным катодом в ограниченном пространстве дрейфа // Радиотехника и электроника. 1999. Т. 44, № 5. С. 551—556.
29. *Пономаренко В. И., Трубецков Д. И.* Сложная динамика автогенератора на вакуумном микротриоде: вычислительный и аналоговый эксперименты на радиотехнической модели // Изв. вузов. Прикладная нелинейная динамика. 1994. Т. 2, № 6. С. 56—65.

НАБЛЮДАЕМЫЕ ПРОЯВЛЕНИЯ ХАОСА В СПИРАЛЬНЫХ ГАЛАКТИКАХ

А. М. Фридман, Р. З. Сагдеев, Е. В. Поляченко

Введение

Как следует из наблюдений, газовые диски спиральных галактик (набор характеризующих их основных параметров) лежат вблизи границы устойчивости [1—6]. Этот факт довольно естествен: в процессе развития неустойчивости дисперсия скоростей возрастает, и диск достигает границы устойчивости [7]. Поскольку неустойчивость, генерирующая спиральные рукава и вихри, насыщена, трехмерное движение газа должно быть стационарно в системе отсчета, вращающейся с волной плотности. Как показано Сагдеевым с соавторами [8], уравнения движения трехмерного стационарного поля скоростей можно свести к динамическим уравнениям в двумерном фазовом пространстве. Последние могут быть использованы как основа для исследования стохастической динамики стационарных полей скоростей газовых галактических дисков. Таким образом, основная задача данной лекции состоит в демонстрации того факта, что результаты анализа наблюдаемого поля скоростей галактического диска могут быть источником нашего понимания стохастической динамики галактик.

Эта задача имеет отношение скорее к газовому, чем к звездному диску. Несмотря на очевидный прогресс в измерениях полей лучевых скоростей звездных дисков, наше представление о стохастической звездной динамике внешних галактик складывается почти целиком из анализа теоретических моделей. Причина того, что информация, извлекаемая из наблюдательных данных по полям лучевых скоростей звезд, более чем скромна, заключается в следующем.

Как правило, галактики не разрешаются на отдельные звезды. Используя анализ абсорбционных линий, можно измерить поле лучевых скоростей звезд, усредненное по некоторой небольшой пространственной области. Траектория этой малой области звездного диска может существенно отличаться от траекторий самих звезд в данной области. Так, например, звездный бар, концы которого соединены со спиральными рукавами, вращается вместе с рукавами как твердое тело, в то время как траектории звезд в баре могут, многократно пересекаясь друг с другом, быть чрезвычайно сложными и сильно отличаться от траекторий простого вращения. Такое поведение типично для бесстолкновительных самогравитирующих систем.

Используя наблюдательные данные, можно построить гравитационный потенциал как функцию координат. Однако изменение потенциа-

ла в пределах ошибок измерения часто приводит к превращению регулярных звездных орбит в хаотические, и наоборот.

Поле лучевых скоростей газового диска, на первый взгляд, также не может быть напрямую использовано для изучения стохастической динамики диска. Однако использование метода восстановления трехмерного поля скоростей из наблюдаемого поля лучевых скоростей [7, 9—12] позволяет (путем вычисления характеристических показателей Ляпунова) определить регулярные и хаотические траектории [13].

Как хорошо известно (см., например, [14]), показатели Ляпунова вычисляются для траекторий в фазовом пространстве, тогда как метод восстановления дает трехмерное поле скоростей в координатном пространстве [7, 9—13]. Это затруднение, однако, можно преодолеть. Как показано в [8], гидродинамические уравнения, описывающие трехмерный стационарный несжимаемый поток, сводятся к нестационарным динамическим уравнениям в двумерном фазовом пространстве. То же справедливо и для сжимаемого трехмерного стационарного потока [15]. Идея преобразования состоит в следующем.

Стационарный трехмерный поток описывается набором уравнений

$$\frac{dx}{dt} = v_x(x, y, z), \quad \frac{dy}{dt} = v_y(x, y, z), \quad \frac{dz}{dt} = v_z(x, y, z), \quad (1)$$

которые можно переписать в виде

$$\frac{dx}{v_x} = \frac{dy}{v_y} = \frac{dz}{v_z} = dt, \quad (2)$$

и далее

$$\frac{dx}{dz} = \frac{v_x}{v_z} \equiv f_1(x, y, z), \quad \frac{dy}{dz} = \frac{v_y}{v_z} \equiv f_2(x, y, z). \quad (3)$$

Последние уравнения показывают, что мы имеем дело с «нестационарной» задачей для динамической системы в двумерном фазовом пространстве (x, y) , где z играет роль времени τ .

$$\frac{dx}{d\tau} = f_1(x, y, \tau), \quad \frac{dy}{d\tau} = f_2(x, y, \tau). \quad (4)$$

Эти уравнения описывают динамические системы, в которых могут сосуществовать хаотические и регулярные траектории.

1. Восстановленное поле скоростей спиральной галактики NGC 3631

На рис. 1 показано восстановленное поле скоростей галактики NGC 3631 в ее плоскости с наложенными линиями постоянной фазы

вертикальных (перпендикулярных к плоскости диска) скоростей. Решение некорректной задачи восстановления трехкомпонентного поля скоростей из наблюдаемого поля лучевых скоростей газовых дисков галактик с хорошо определенной спиральной структурой описано в работах [9—13] и обзоре [7].

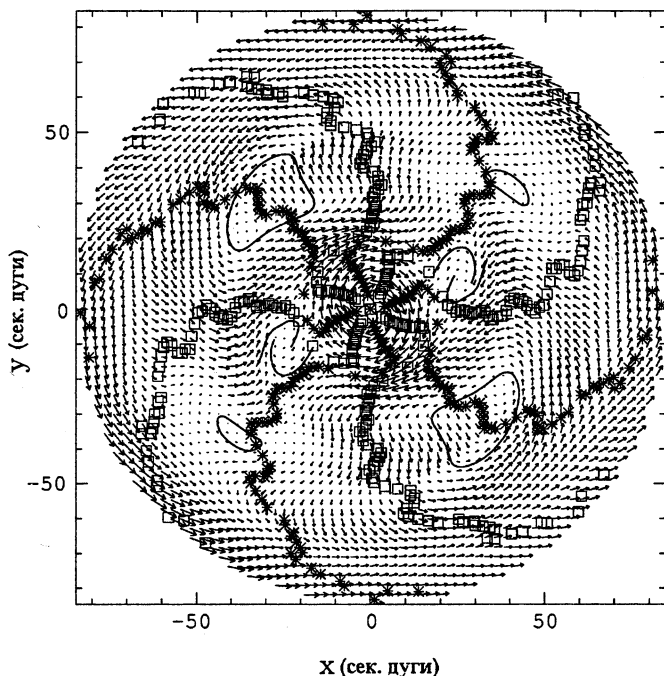


Рис. 1. Восстановленное поле скоростей NGC 3631 в плоскости галактики и наложенные линии постоянных фаз вертикальных движений. Квадратами отмечены азимутальные максимумы абсолютных значений вертикальной скорости газа для каждого радиуса. Звездочками показаны положения нулей вертикальной скорости; тонкими линиями обозначено положение нулей — антициклонов (верхний левый и правый нижний) и циклонов.

Как было отмечено во введении, поле скоростей в газовом диске спиральной галактики является стационарным. Это означает, что траектории совпадают с линиями тока в трехмерном координатном пространстве. Вблизи областей нулевых вертикальных движений каждая жидкая «частица» принимает участие лишь в двумерном движении и, следовательно, траектории должны совпадать с их двумерными линиями тока. Возможно, именно поэтому сепаратрисы двух антициклонов и двух циклонов, центры которых лежат вблизи нулей вертикальных скоростей, замкнуты. В то же время сепаратрисы двух других циклонов, центры

которых лежат между линиями максимума и нуля вертикальных скоростей, являются незамкнутыми.

«Нестационарный» характер движения жидких частиц в плоскости диска внутри этих циклонов может вызываться вертикальным движением. Его, в отличие от предыдущих случаев, уже нельзя считать пренебрежимо малым.

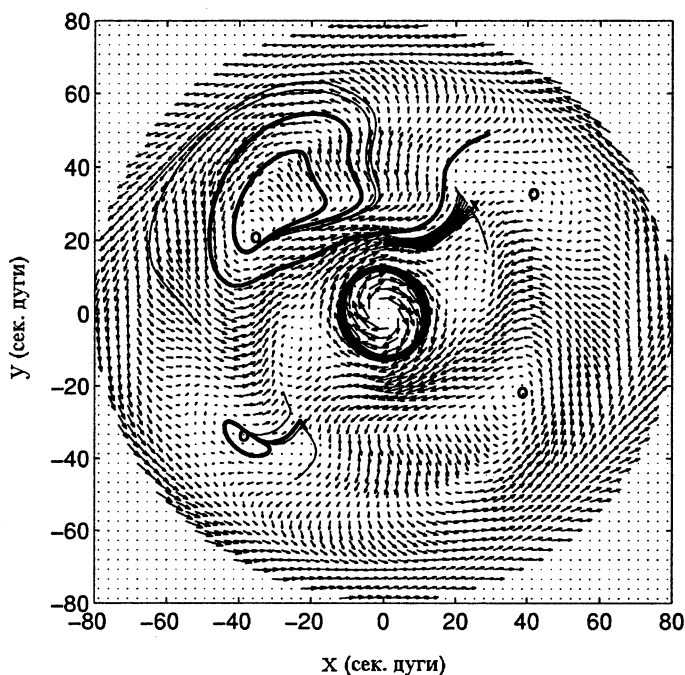


Рис. 2. Расходящиеся траектории вблизи сепаратрис вихрей и пучок траекторий вблизи седловой точки, наложенные на восстановленное двумерное поле скоростей. Жирные сплошные линии показывают траектории, которые используются для вычисления стохастических характеристик вблизи вихрей. Символами обозначены: о — положение центров вихрей, х — положение седловых точек. В центре диска приведено семейство регулярных траекторий.

Помимо особых линий — сепаратрис, окружающих вихри, — двумерное поле скоростей в плоскости галактики содержит особые точки — седловые (отмеченные крестами на рис. 2). Выбирая начала линий тока вблизи сепаратрис или вблизи седловой точки, можно видеть, что эти линии тока расходятся. Интересно поставить вопрос о скорости расхождения траекторий: является ли эта скорость экспоненциальной или нет, т. е. линии тока хаотические или регулярные. Для этого необходимо вычислить показатели Ляпунова.

2. Вычисление характеристических показателей Ляпунова

Под хаосом подразумевают быстрое перемешивание, которое возникает вследствие быстрой (экспоненциальной) расходимости близких траекторий. Рассмотрим две первоначально близкие точки на расстоянии d_0 друг от друга. Спустя некоторое время t они окажутся на расстоянии $d(t)$. Экспоненциальная расходимость траекторий означает, что

$$d(t) \sim d_0 e^{\lambda t}, \quad (5)$$

где усредненный за большой промежуток времени показатель степени $\lambda > 0$ есть (максимальный) показатель Ляпунова; он характеризует степень экспоненциальной расходимости в пределе $t \rightarrow \infty$ и $d_0 \rightarrow 0$. Иными словами, показатель Ляпунова определяется двойным пределом [14]:

$$\lambda = \lim_{\substack{t \rightarrow \infty \\ d_0 \rightarrow 0}} \frac{1}{t} \ln \frac{d(t)}{d_0}. \quad (6)$$

Можно видеть, что этот предел существует и одинаков почти для всех взаимных ориентаций первоначально выбранных точек. Лишь для некоторых взаимных ориентаций предел (6) будет меньше максимального значения [14]. Ниже мы будем рассматривать ситуацию общего положения и опускать термин «максимальный» для показателя Ляпунова (подробнее см. [14]). Математически удобная формулировка показателя Ляпунова оказывается неудобной в компьютерных расчетах даже при изучении идеализированных (с точки зрения реальных систем) теоретических моделей. Существенным допущением, принимаемым в таких расчетах, является предположение о том, что расстояние $d(t)$ остается достаточно малым, так что эволюция вектора $\mathbf{w}(t)$, соединяющего точки на двух близких линиях тока (рис. 3), описывается некоторым линеаризованным уравнением [14]. Действительно, согласно определению скорости $\mathbf{V}(\zeta)$ имеем следующие уравнения:

$$\frac{d\zeta}{dt} = \mathbf{V}(\zeta), \quad \zeta = \zeta_0 + \mathbf{w},$$

$$\mathbf{V}(\zeta) = \mathbf{V}_0(\zeta_0) + \left. \frac{d\mathbf{V}}{d\zeta} \right|_{\zeta=\zeta_0} \mathbf{w} + \frac{1}{2} \left. \frac{d^2\mathbf{V}}{d\zeta^2} \right|_{\zeta=\zeta_0} \mathbf{w}^2 + \frac{1}{6} \left. \frac{d^3\mathbf{V}}{d\zeta^3} \right|_{\zeta=\zeta_0} \mathbf{w}^3 + \dots$$

Используя эти соотношения, а также выражение $\mathbf{V}_0(\zeta_0) = d\zeta_0/dt$, найдем временную производную от малого вектора \mathbf{w} в виде суммы:

$$\frac{d\mathbf{w}}{dt} = \frac{d}{dt}(\zeta - \zeta_0) = \mathbf{V}(\zeta) - \mathbf{V}_0(\zeta_0) = \left. \frac{d\mathbf{V}}{d\zeta} \right|_{\zeta=\zeta_0} \mathbf{w} + \frac{1}{2} \left. \frac{d^2\mathbf{V}}{d\zeta^2} \right|_{\zeta=\zeta_0} \mathbf{w}^2 + \frac{1}{6} \left. \frac{d^3\mathbf{V}}{d\zeta^3} \right|_{\zeta=\zeta_0} \mathbf{w}^3 + \dots$$

Линейный член разложения по w содержит линейный оператор $dV(\zeta)/d\zeta|_{\zeta=\zeta_0} \equiv M(\zeta)$, с помощью которого определяются показатели Ляпунова. Следовательно, их вычисление может быть легко осуществлено в случае пренебрежения всеми нелинейными членами разложения по сравнению с линейным членом. Это возможно при условии, когда величина $d \equiv |w|$ оказывается много меньше характерной величины R_{ch} изменения поля скоростей. Таким образом, при условии $d/R_{ch} \ll 1$ мы приходим к линейному уравнению

$$\frac{dw}{dt} = M(\zeta_0)w. \quad (7)$$

Это уравнение может быть записано в эквивалентной форме в виде

$$\frac{dw_i}{dt} = M_{ik} w_k; \quad i, k = 1, 2, 3. \quad (7a)$$

Матрица

$$M_{ik} = \frac{dV_i}{d\zeta_k} \quad (7b)$$

является тензором второго порядка.

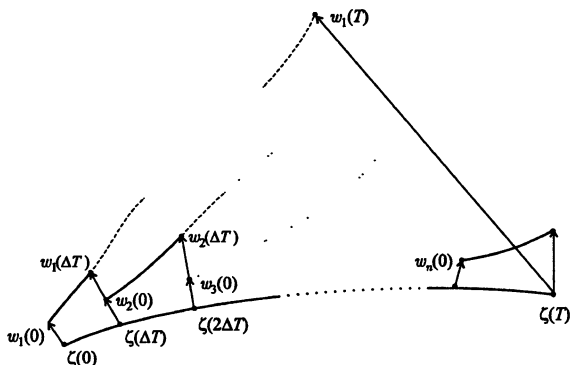


Рис. 3. Метод вычисления показателя Ляпунова по двум траекториям. Основная траектория — нижняя сплошная кривая, несколько вспомогательных траекторий показаны короткими сплошными линиями. Дальнейший ход вспомогательных кривых показан пунктирной линией.

Кроме теоретического ограничения — использования линейного приближения при нахождении показателя Ляпунова, имеется ряд ограничений, следующих из анализа наблюдательных данных:

1. Продолжительность наблюдений намного короче характерного времени ($-\lambda^{-1}$) экспоненциального расхождения двух точек, движущихся вдоль соседних траекторий.

2. Минимальное расстояние между двумя соседними траекториями d_{\min} не может быть меньше пространственного разрешения телескопа δ , $d_{\min} \geq \delta$. Для 6-метрового телескопа САО среднее δ оказывается равным примерно 1,5 угл. сек., что даже для рассматриваемой близкой галактики NGC 3631 составляет около 0,02 радиуса «рабочей» части ее газового диска.

3. Отношение пространственного разрешения δ к характерному масштабу изменения поля скоростей R_{ch} , δ/R_{ch} не всегда является малым параметром, как того требует условие линейного приближения. Для некоторых областей величина δ/R_{ch} оказывается порядка единицы.

Выше уже отмечалось, что, как следует из анализа наблюдений [1—6], поле скоростей в газовых галактических дисках должно быть стационарным в системе координат, вращающейся вместе со спиральными рукавами. Как известно (см., например, [16]), стационарные траектории можно рассматривать как линии тока. Две другие особенности реально наблюдаемых систем ограничивают допустимую максимальную длину траектории (и, следовательно, максимальное время вычисления показателя Ляпунова) в случае, когда рассматриваются только две соседние траектории. Действительно, поскольку первоначальное расстояние между соседними траекториями не может быть выбрано сколь угодно малым, при экспоненциальном расхождении траекторий расстояние между ними оказывается сравнительно большим (так что перестает выполняться линейное приближение) уже за относительно небольшой отрезок времени.

Для преодоления этой трудности воспользуемся методом, впервые описанным в [17, 18] и состоящим в следующем. Разобьем так называемую «основную» траекторию (нижнюю сплошную кривую на рис. 3) на равные временные промежутки ΔT . В конце каждого отрезка $\zeta(\Delta T)$, $\zeta(2\Delta T)$ и т. д. строятся векторы $w(\Delta T)$, $w(2\Delta T)$ и т. д., которые затем укорачиваются до величины $|w_1(0)| = d_0$. Из конца каждого укороченного вектора проводится «вспомогательная» траектория (короткий отрезок сплошной линии, соответствующий временному промежутку ΔT и затем — пунктирная кривая). Таким образом, при каждом временном шаге мы используем новую вспомогательную траекторию, расположенную ближе, нежели предыдущая, к основной траектории. В конце каждого временного шага n мы оцениваем соответствующий показатель Ляпунова согласно следующему определению:

$$\lambda^{(n)} = \frac{1}{n\Delta T} \sum_{i=1}^n \ln \frac{|w_i(\Delta T)|}{d_0}. \quad (8)$$

Покажем, что выражение (8) при достаточно большом T совпадает с физическим определением показателя Ляпунова (см. рис. 3):

$$\lambda = \frac{1}{T} \ln \frac{|\mathbf{w}_1(T)|}{d_0}. \quad (9)$$

Введем временной оператор \hat{T} в соответствии с определением (см. рис. 3):

$$\begin{aligned} \hat{T} \mathbf{w}_1(0) &= \mathbf{w}_1(\Delta T), |\mathbf{w}_1(0)| \equiv d_0; \\ \hat{T} \mathbf{w}_1(\Delta T) &= \mathbf{w}_1(2\Delta T) = \hat{T}^2 \mathbf{w}_1(0); \\ &\dots \\ \hat{T} \mathbf{w}_1((n-1)\Delta T) &= \mathbf{w}_1(n\Delta T) = \mathbf{w}_1(T) = \hat{T}^n \mathbf{w}_1(0). \end{aligned} \quad (10)$$

Подставляя (10) в (9), получим физическое определение показателя Ляпунова в форме

$$\lambda = \frac{1}{n\Delta T} \ln \frac{|\hat{T}^n \mathbf{w}_1(0)|}{d_0}. \quad (11)$$

Как следует из рис. 3,

$$\mathbf{w}_{i+1}(0) = \frac{\hat{T} \mathbf{w}_i(0)}{|\hat{T} \mathbf{w}_i(0)|} d_0, |\mathbf{w}_i(0)| = d_0, i = 1, 2, \dots, n, |\mathbf{w}_{i+1}(0)| = d_0. \quad (12)$$

Используя формулы (10) и (12), получаем

$$\begin{aligned} \frac{|\hat{T}^n \mathbf{w}_1(0)|}{d_0} &= \left| \frac{1}{d_0} \hat{T}^{n-1} \hat{T} \mathbf{w}_1(0) \right| = \left| \frac{1}{d_0} \hat{T}^{n-1} \mathbf{w}_2(0) \frac{|\hat{T} \mathbf{w}_1(0)|}{d_0} \right| = \\ &= \left| \frac{1}{d_0} \mathbf{w}_{n+1}(0) \frac{|\hat{T} \mathbf{w}_1(0)|}{d_0} \frac{|\hat{T} \mathbf{w}_2(0)|}{d_0} \dots \frac{|\hat{T} \mathbf{w}_n(0)|}{d_0} \right| = \left| \prod_{i=1}^n \frac{|\hat{T} \mathbf{w}_i(0)|}{d_0} \right|. \end{aligned}$$

Подставляя последнее выражение в (11), получаем

$$\lambda = \frac{1}{n\Delta T} \ln \left| \prod_{i=1}^n \frac{|\hat{T} \mathbf{w}_i(0)|}{d_0} \right| = \frac{1}{n\Delta T} \sum_{i=1}^n \ln \frac{|\hat{T} \mathbf{w}_i(0)|}{d_0} = \frac{1}{n\Delta T} \sum_{i=1}^n \ln \frac{|\mathbf{w}_i(\Delta T)|}{d_0}.$$

Последнее выражение тождественно совпадает с (8).

Стоящий в определении показателя Ляпунова (6) предел $t \rightarrow \infty$ (в реальной ситуации $T \rightarrow \infty$) означает, что нас интересует результат динамической эволюции, наблюдаемый за время, много большее характерного времени процесса. Характерное время динамического процесса есть время локальной неустойчивости $\tau \sim \lambda^{-1}(T)$. Следовательно, предел $T \rightarrow \infty$ для реальных систем означает

$$\xi \equiv T\lambda(T) \gg 1. \quad (13)$$

Очевидно, что за такое время (много большее времени локальной неустойчивости) хаос вполне может развиться.

В некоторых ситуациях траектории оказываются очень короткими, так что (13) не выполняется. В таком случае рассматривают не одну траекторию, а целое семейство, состоящее из N близких траекторий [19, 20]. Показатель Ляпунова для семейства траекторий определяется по формуле

$$\lambda^{(n)} = \frac{1}{N} \sum_{k=1}^N \lambda_k^{(n)}, \quad (14)$$

где $\lambda_k^{(n)}$ вычисляется стандартно с использованием (8), при этом k -я траектория рассматривается как основная, N — полное число основных траекторий. В этом случае условие (13) приобретает лишний большой множитель $N \gg 1$ и получает, таким образом, большую возможность быть выполненным:

$$\eta \equiv N\xi = NT\lambda(T) \gg 1. \quad (15)$$

Описанный выше способ определения показателя Ляпунова с помощью линеаризованного уравнения требует вычисления матрицы Якоби M_{ik} (76), элементами которой являются градиенты компонент скорости. Мы использовали метод расчета показателя Ляпунова по двум траекториям, проиллюстрированный на рис. 3. Отличие данного метода от описанного выше состоит лишь в способе вычисления эволюции вектора $w(t)$. Задав количество шагов по времени n , мы вычисляем траектории, начинающиеся в точках $\zeta(0)$ и $\zeta(0) + w_1(0)$ на протяжении времени $\Delta T = T/n$. Далее, перенормировав вектор $w_1(\Delta T)$ согласно рис. 3, находим новую начальную точку для вспомогательной траектории. После выполнения n шагов оценка для показателя Ляпунова вычисляется по прежней формуле (8).

Для вычисления показателя Ляпунова необходимо определить риманову метрику d . Используемый нами в данной работе общий вид метрики

$$d(w) \equiv \sqrt{g_1 x^2 + g_2 y^2} \quad (16)$$

зависит от двух постоянных положительных коэффициентов g_1 и g_2 (x и y — компоненты вектора w). Для краткости будем обозначать различные метрики указанием этих коэффициентов: (g_1, g_2) . Например, часто используемая метрика $d \equiv \sqrt{x^2 + y^2}$ записывается в виде (1,1). На рис. 4 показано поведение $\lambda^{(i)}$ в зависимости от $i \Delta T/T_0$ ($T_0 = 7,5 \cdot 10^7$ лет) для различных n , вычисленное для основной траектории, изображенной жирной линией вблизи сепаратрисы антициклона на рис. 2. Для этой траектории $\lambda^{(n)} = 0,87T_0^{-1}$ и $\xi = 9,57$ (см. формулу (13)). Как следует из рис. 4, поведение $\lambda^{(i)}$ слабо зависит от числа n шагов по времени вдоль

траектории. Начальное расстояние d_0 принято равным 1,4 сек. дуги, т. е. порядка углового разрешения поля скоростей галактик NGC 3631.

Теорема Оселедца [21] утверждает, что результат вычисления показателя Ляпунова не зависит от римановой метрики. Интересно отметить, что для рассмотренной реальной системы показатели Ляпунова, вычисленные с помощью различных метрик, тоже слабо отличаются (рис. 5). Вблизи сепаратрисы антициклона для метрики (1,1) $\lambda^{(n)} = 0,8768 T_0^{-1}$, для метрики (1,0) $\lambda^{(n)} = 0,8931 T_0^{-1}$, для метрики (0,1) $\lambda^{(n)} = 0,8513 T_0^{-1*}$.

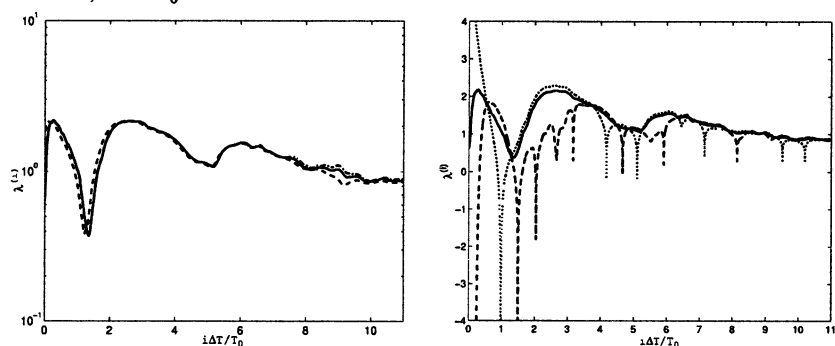


Рис. 4. Зависимость $\lambda^{(i)}$ от $i \Delta T/T_0$ для различного числа шагов n , вычисленная для основной траектории вблизи сепаратрисы антициклона: $n = 10000$ (жирная линия), $n = 1000$ (штрихпунктирная линия), $n = 100$ (пунктирная линия).

Рис. 5. Зависимость $\lambda^{(i)}$ от $i \Delta T/T_0$ для траектории вблизи сепаратрисы антициклона, вычисленная для различных метрик: (1,1) — сплошная линия, (1,0) — штриховая линия, (0,1) — пунктирная линия. Число шагов по времени $n = 1000$.

Для траектории вблизи сепаратрисы циклона, используя метрику (1,1), мы получили $\lambda^{(n)} = 1,04 T_0^{-1}$ (рис. 6) и $\xi = 6,24$ (для метрики (1,0) $\lambda^{(n)} = 1,0884 T_0^{-1}$, для метрики (0,1) $\lambda^{(n)} = 0,8892 T_0^{-1}$).

Из-за довольно коротких траекторий в окрестности седловой точки необходимо рассматривать не одну траекторию, а семейство близких траекторий. Мы взяли $N = 10$ начальных точек, расположенных равноудаленно на оси ординат между $y = 18,5''$ и $y = 20,5''$ (см. рис. 2). Зависимость $\lambda^{(i)}$, вычисленная по формуле (14) представлена на рис. 7. Оценка показателя Ляпунова $\lambda = 0,79 T_0^{-1}$. Принимая во внимание, что $T = 2T_0$ и $N = 10$, можно получить $\eta = 15,8$ (см. формулу (15)), т. е. результат вполне достоверный.

* Согласно определению метрики — неотрицательной функции — $d(x,y) - d = 0$ только в случае $x = y = 0$, поэтому метрические коэффициенты должны быть строго положительны. В случае метрик (1,0) и (0,1) подразумеваем, что соответствующие коэффициенты близки к нулю, оставаясь строго больше нуля.

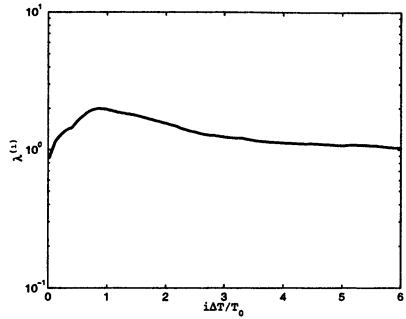
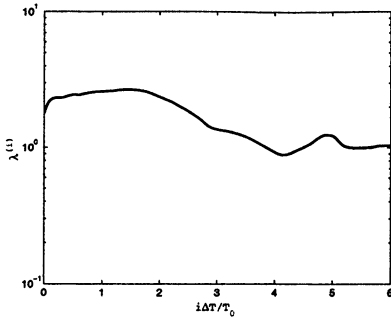


Рис. 6. Поведение $\lambda^{(i)}$, вычисленной для основной траектории вблизи сепаратрисы цикла.

Рис. 7. Зависимость $\lambda^{(i)}$ для семейства близких траекторий вблизи седловой точки.

Как следует из рис. 4—7, показатели Ляпунова для траекторий вблизи антициклона, циклона и седловой точки положительны, т. е. траектории хаотические. На рис. 8 можно видеть пример $\lambda \approx 0$ для одной траектории из семейства однотипных линий тока около центра галактики. Эта траектория регулярна.

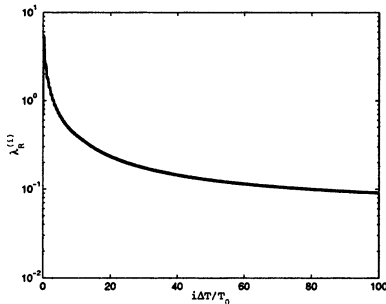


Рис. 8. Зависимость $\lambda^{(i)}$ для регулярной траектории около галактического центра.

Помимо использованного нами метода вычисления показателя Ляпунова по двум траекториям, мы пересчитали показатели Ляпунова методом, использующим линеаризованное уравнение (7). Оказалось, что для достаточно гладкой интерполяции векторного поля оба метода дают идентичные результаты.

3. Спектр коэффициентов растяжения

Используя определение

$$a_i = \frac{1}{\Delta T} \ln \frac{|\mathbf{w}_i(\Delta T)|}{d_0} \quad (17)$$

коэффициентов растяжения*, найдем спектр этой величины вдоль траектории. По определению, спектр коэффициента растяжения [19, 20]

$$S(a, x_0, y_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \frac{dn(a)}{da}, \quad (18)$$

где $dn(a)$ — число коэффициентов растяжения a_i , попадающих в интервал $(a, a + da)$. Спектры хаотических и регулярных траекторий, вычисленные с использованием разных метрик, показаны на рис. 9—12. Как видно из данных этих рисунков, спектры не инвариантны по отношению к изменению метрики. Аналитическое доказательство этого факта, по-видимому, впервые было дано в работе [23].

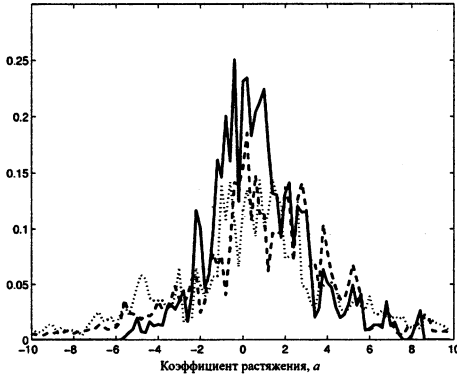
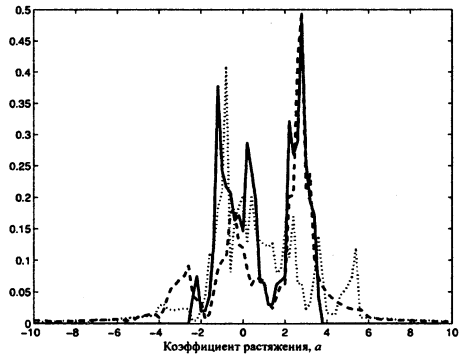


Рис. 9. Спектр для хаотической траектории в окрестности антициклона (см. рис. 2), вычисленный для различных метрик: (1,1) — сплошная линия, (1,0) — штриховая линия, (0,1) — пунктирная линия. Число шагов по времени $n = 1000$.

Рис. 10. Спектр для хаотической траектории в окрестности цикла (см. рис. 2), вычисленный для различных метрик: (1,1) — сплошная линия, (1,0) — штриховая линия, (0,1) — пунктирная линия. Число шагов по времени $n = 1000$.



Спектр коэффициентов растяжения играет в стохастической динамике важную роль, поскольку показатель Ляпунова является его первым моментом:

$$\lambda = \int aS(a)da. \quad (19)$$

* В данной работе мы следуем терминологии, впервые введенной Оселедцем [21]. Та же величина фигурирует в более поздних работах других авторов: в [22] — локальный показатель Ляпунова, в [19, 20] — коэффициент растяжения.

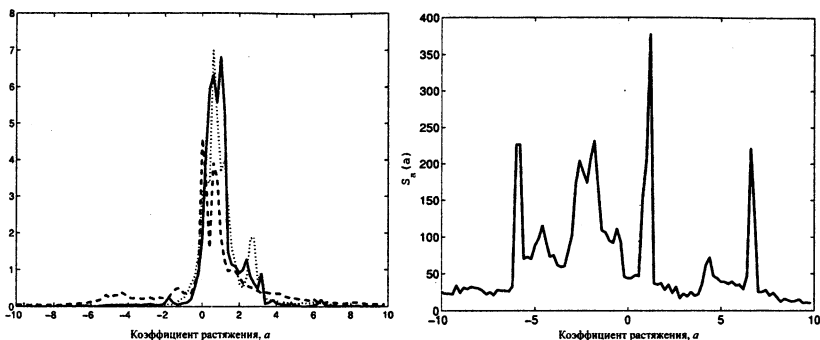


Рис. 11. Спектр для семейства хаотических траекторий в окрестности седловой точки (см. рис. 2), вычисленный для различных метрик: (1,1) — сплошная линия, (1,0) — штриховая линия, (0,1) — пунктирная линия. Число шагов по времени $n = 1000$.

Рис. 12. Спектр для регулярной траектории вокруг центра диска (см. рис. 2), вычисленный для метрики (1,1).

Инвариантность показателя Ляпунова относительно изменений метрики была доказана аналитически в работе [21] и много раз численно для различных моделей динамических систем. В данной работе, видимо, впервые приведены численные доказательства инвариантности λ и неинвариантности $S(a)$ относительно изменения метрики для реально существующей системы. Попробуем пояснить на качественном уровне различные поведения спектра и его первого момента по отношению к изменению метрики.

Неодинаковость спектров при различной метрике понять несложно. Предположим, что на шаге i вектор $w_i(\Delta T)$ оказался почти параллельным оси ординат. Тогда коэффициент растяжения a_i , вычисленный с использованием метрики (1,1), может быть больше нуля, тогда как с помощью метрики (1,0) мы получим $d_i \approx 0$ и соответствующий ей большой отрицательный коэффициент a_i . Именно выстраиванием векторов w_i параллельно координатным осям объясняется наличие широких крыльев при $a \rightarrow \pm\infty$ в форме спектров, найденных с помощью (1,0) и (0,1) метрик.

Для того чтобы понять причину независимости показателя Ляпунова от изменения метрики, рассмотрим две произвольные метрики:

$d(g_1, g_2)$ и $\bar{d}(\bar{g}_1, \bar{g}_2)$. Прежде всего заметим, что существуют такие положительные постоянные C_1 и C_2 , что для произвольного вектора w

$$C_1 \bar{d}(w) \leq d(w) \leq C_2 \bar{d}(w). \quad (20)$$

Действительно, это двойное неравенство будет выполнено, если положить

$$C_1 \leq \sqrt{\frac{\min(g_1, g_2)}{\max(\bar{g}_1, \bar{g}_2)}}, \quad C_2 \geq \sqrt{\frac{\max(g_1, g_2)}{\min(\bar{g}_1, \bar{g}_2)}}.$$

Для примера рассмотрим случай $g_1 \leq g_2$, $\bar{g}_1 \leq \bar{g}_2$, так что

$$C_1 \leq \sqrt{\frac{g_1}{\bar{g}_2}}, \quad C_2 \geq \sqrt{\frac{g_2}{\bar{g}_1}}.$$

Из написанных выше неравенств, выполненных для любого вектора, следует:

$$C_1 \bar{d}(\mathbf{w}) \leq \sqrt{\frac{g_1}{\bar{g}_2}} \sqrt{\bar{g}_1 x^2 + \bar{g}_2 y^2} = \sqrt{\frac{\bar{g}_1}{\bar{g}_2} g_1 x^2 + g_1 y^2} \leq d(\mathbf{w}),$$

$$C_1 \bar{d}(\mathbf{w}) \geq \sqrt{\frac{g_2}{\bar{g}_1}} \sqrt{\bar{g}_1 x^2 + \bar{g}_2 y^2} = \sqrt{g_2 x^2 + \frac{\bar{g}_2}{\bar{g}_1} g_2 y^2} \geq d(\mathbf{w}).$$

Из этих неравенств получим ограничения на отношение двух метрик

$$C_1 \leq \frac{d(\mathbf{w})}{\bar{d}(\mathbf{w})} \leq C_2, \quad (21)$$

совпадающие с (20). Используя физическое определение λ и $\bar{\lambda}$ (9) соответственно для метрик d и \bar{d} и систему неравенств (21), найдем

$$|\lambda - \bar{\lambda}| = \frac{1}{T} \left| \ln \frac{d(\mathbf{w}(T))}{d(\mathbf{w}_0)} - \ln \frac{\bar{d}(\mathbf{w}(T))}{\bar{d}(\mathbf{w}_0)} \right| = \frac{1}{T} \left| \ln \frac{d(\mathbf{w}(T)) \bar{d}(\mathbf{w}_0)}{\bar{d}(\mathbf{w}(T)) d(\mathbf{w}_0)} \right| \leq \frac{1}{T} \left| \ln \frac{C_2}{C_1} \right|.$$

Последняя формула показывает, что независимость показателя Ляпунова от метрики достигается в пределе $T \rightarrow \infty$. Таким образом, независимость от метрики для ограниченных траекторий в реальных системах является тестом на достоверность полученных оценок λ . На рис. 13, 14 представлены показатели Ляпунова и спектры, вычисленные с использованием различных метрик для хорошо известных динамических моделей: стандартной модели Чирикова и странного аттрактора Лоренца.

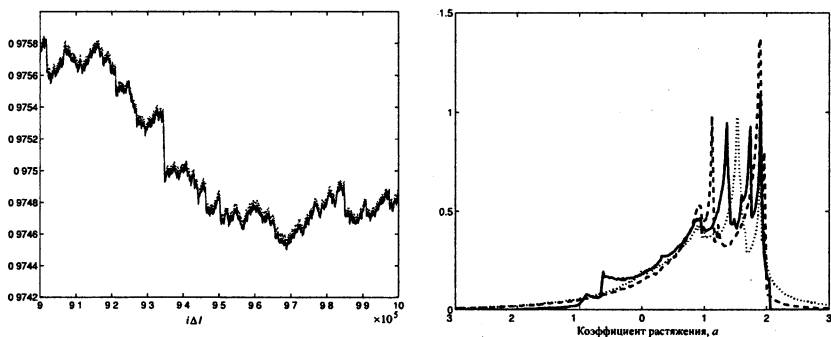


Рис. 13. Показатели Ляпунова (слева) и спектры стандартной модели Чирикова, вычисленные для различных метрик: (1,1) — сплошная линия, (0,1) — штриховая линия, (1,0) — пунктирная линия.

Можно видеть, что в обоих случаях показатель Ляпунова, вычисленный с использованием различных метрик, одинаков, в то время как формы спектров различны.

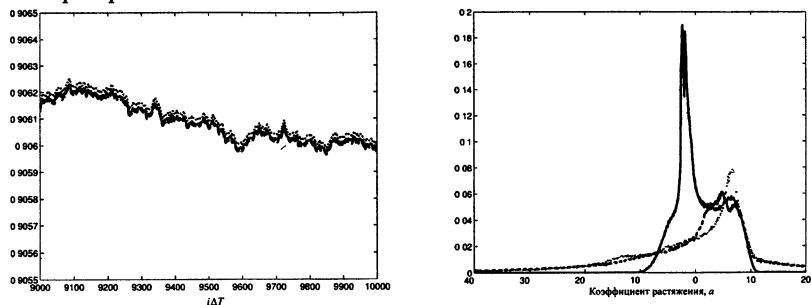


Рис. 14. Показатели Ляпунова (слева) и спектры странного аттрактора Лоренца, вычисленные для различных метрик: (1,1) — сплошная линия, (0,1) — штриховая линия, (1,0) — пунктирная линия.

В заключение мы резюмируем полученные динамические свойства газового диска галактики NGC 3631.

- Трехкомпонентное поле скоростей, восстановленное из наблюдаемого поля лучевых скоростей газового диска галактики NGC 3631, стационарно, и в нем сосуществуют как регулярные, так и хаотические траектории газа.

- Регулярные траектории наблюдаются вблизи центра диска, стохастические — вблизи седловых точек и сепаратрис гигантских вихрей, имеющих в поле скоростей.

- Тип траекторий был определен двумя методами: вычислением показателя Ляпунова по двум траекториям и с помощью линеаризованного уравнения для вектора отклонения. Оба метода дали одинаковые результаты.

- Показатели Ляпунова, полученные для реальной галактики, оказались инвариантными по отношению к изменению метрики, в полном соответствии с математической теоремой Оселедеца [21].

- Впервые было численно продемонстрировано, что форма спектра коэффициентов растяжения зависит от использованной метрики.

Последние два вывода справедливы также и для классических динамических систем: стандартной модели Чирикова и странного аттрактора Лоренца.

Авторы благодарны В. И. Арнольду, Дж. Контопулосу, В. И. Оселедцу, М. И. Рабиновичу, Я. Г. Синаю, Ю. А. Фридман и Б. В. Чирикову за многочисленные и плодотворные обсуждения. Благодарим Г. М. Заславского за указание на ссылку [15] и Х. Эфсимипулосу за указание на ссылку [23].

Эта работа была выполнена при поддержке РФФИ (гранты № 02-02-16878, № 04-02-16706), Минпромнауки (грант «Ведущие научные школы» НШ № 925.2003.2 и Государственные контракты № 40.022.1.1.1101 и 40.020.1.1.1167.ИА.1), Фонда содействия отечественной науке и грантов «Фундаментальные космические исследования. Астрономия»: № 1.2.3.1, № 1.7.4.3.

Литература

1. Засов А. В., Симаков С. // *Астрофизика*. 2002. Т. 29. С. 190.
2. Засов А. В., Рыбцова Т. В. // *Письма в Астрон. журн.* 1989. Т. 15. С. 118.
3. Fridman A. M., Polyachenko V. L., Zasov A. V. // *Dynamics of galaxies and molecular cloud distribution* / Eds. F. Combes and F. Casoli. Dordrecht; Boston; London: Kluwer Acad. Publishers, 1991. P. 109.
4. Martin C., Kennicutt R. C. // *Astrophys. J.* 2001. V. 555. P. 301.
5. Wong T., Blitz L. // *Astrophys. J.* 2002. V. 569. P. 157.
6. Boissier S., Prantzo N., Boselli A., Gavazzi G., in press (astro-ph 0309372).
7. Fridman A. M., Khoruzhii O. V. // *Space Sci. Rev.* 2003. V. 105. P. 1.
8. Заславский Г. М., Сагеев Р. З., Усиков Д. А., Черников А. А. Слабый хаос и квази-регулярные структуры. М.: Наука, 1991; Zaslavsky G. M., Sageev R. Z., Usikov D. A., Chernikov A. A. Weak chaos and quasi-regular patterns. Cambridge: Cambridge Univ. Press, 1991.
9. Ляхович В. В., Фридман А. М., Хоружий О. В., Павлов А. И. Метод восстановления полного поля скоростей в газовых дисках спиральных галактик // *Астрон. журн.* 1997. Т. 74. С. 509.
10. Fridman A. M., Khoruzhii O. V., Lyakhovich V. V., Avedisova V. S., Sil'chenko O. K., Zasov A. V., Rastorguev A. S., Afanasiev V. L., Dodonov S. N., Boulesteix J. Spiral-vortex structure in the gaseous disks of galaxies // *Astrophis. Space Sci.* 2001. V. 252. P. 115.
11. Fridman A. M., Khoruzhii O. V., Lyakhovich V. V., Sil'chenko O. K., Zasov A. V., Afanasiev V. L., Dodonov S. N., Boulesteix J. Restoring the full velocity field in the gaseous disk of the spiral galaxy NGC 157 // *Astron. Astrophis.* 2001. V. 371. P. 538.
12. Fridman A. M., Khoruzhii O. V., Polyachenko E. V., Zasov A. V., Sil'chenko O. K., Moiseev A. V., Burlak A. N., Afanasiev V. L., Dodonov S. N., Knapen J. H. Gas motions in the plane of the spiral galaxy NGC 3631 // *Mon. Not. R. Astr. Soc.* 2001. V. 323. P. 651.
13. Fridman A. M., Khoruzhii O. V., Polyachenko E. V. Observational manifestation of chaos in the gaseous disk of the grand-design spiral galaxy NGC 3631 // *Observational Manifestation of Chaos in Astrophysical Objects. Int. Workshop, Moscow, August 28—29, 2001* / Ed. by A. M. Fridman, M. Ya. Marov, R. H. Miller; *Space Sci. Rev.* 2002. V. 102. P. 51.
14. Lichtenberg A. J., Leiberman M. A. *Regular and Stochastic Motion*. N. Y.; Heidelberg: Berlin: Springer-Verlag, 1983.
15. Govorukhin V. N., Morgulis A., Yudovich V. I., Zaslavsky G. M. // *Phys. Rev. E.* 1999. V. 60, Issue 3. P. 2788.
16. Лойцянский Л. Г. *Механика жидкости и газа*. М.: Наука, 1973.
17. Casartelli M., Diana E., Galgani L., Scotti A. // *Phys. Rev.* 1976. V. A13. P. 1921.
18. Voglis N., Contopoulos G. // *J. Phys.* 1994. V. A27. P. 4899.
19. Contopoulos G., Voglis N. A fast method for distinguishing between ordered and chaotic orbits // *Astron. Astrophis.* 1997. V. 317. P. 73.
20. Bennetin G., Galgani L., Strelcyn J.-M. // *Phys. Rev.* 1976. V. A14, P. 2338.
21. Оседеу В. И. // *Тр. Моск. мат. общ.* 1968. Т. 19. С. 179.
22. Froeschlé C., Froeschlé Ch., Lohinger E. Generalized Lyapunov characteristic indicators and corresponding Kolmogorov like entropy of the standard mapping // *Cel. Mech. Dyn. Astron.* 1993. V. 56. P. 307.
23. Elskens Y. // *Physica D.* 1997. V. 100. P. 142.

**Общие вопросы
нелинейной динамики**

ФАЗОВО-УПРАВЛЯЕМЫЕ КОЛЕБАНИЯ В НЕЙРОДИНАМИКЕ

В. Б. Казанцев, В. И. Некоркин

1. Введение

Исследование динамики нейронных систем и разработка адекватных моделей их поведения становятся в последние годы одним из наиболее интенсивно развивающихся направлений современной нелинейной физики. Возрастающий интерес к этому направлению связан, с одной стороны, с современными возможностями нейрофизиологического эксперимента, позволяющего получить все более обширную информацию о функционировании различных нейронных систем [1—4]. С другой стороны, многообещающими выглядят перспективы разработки на основе нейронных моделей новейших систем медицинской диагностики, обработки и передачи информации, автоматического управления и навигации, робототехники, нейрокомпьютинга и др. [5—11]. С фундаментальной точки зрения интенсивно изучаются динамические механизмы функционирования нейронных систем, «алгоритмы» представления и преобразования сенсорной информации различными системами головного мозга.

Последние исследования нейрофизиологов свидетельствуют о том, что ключевую роль в нейропроцессах играет собственная колебательная активность нейронных систем [4]. Колебания в частотном диапазоне 1—100 Гц регистрируются в различных частях головного мозга. Так, например, сравнительно быстрые процессы (~ 40 Гц) формируются в таламо-кортикальной системе [12, 13]. Они связаны с такими явлениями, как пространственно-временное связывание сенсорной информации, приходящей по различным сенсорным каналам, что составляет основу ассоциативного восприятия. Колебания в диапазоне 8—12 Гц играют ключевую роль в динамике оливо-мозжечковых взаимодействий, ответственных за формирование моторных паттернов в системе моторного контроля [14—17]. Колебательные процессы в гиппокампе (5—10 Гц) связаны с процессами кратковременной памяти и фазовой прецессии, играющей ключевую роль при навигации объекта в пространстве [18—19]. Различные нейронные системы способны за счет собственной динамики формировать пространственно-временные структуры колебательной активности, изменяющиеся за счет внешнего воздействия (сенсорных сигналов). Вопрос о том, как именно нейронная система оперирует с такими структурами («алгоритм» функционирования), остается до сих пор открытым. Однако уже сейчас ясно, что не только сложная, нелокальная архитектура межнейронных связей важна при моделировании нейропроцессов (как, например, принято считать при разработке нейро-

компьютеров), но и наличие у нейронов собственной колебательной динамики [4].

При изучении колебательных свойств нейронных систем важнейшей информационной характеристикой является фаза колебаний. Именно она определяет моменты возникновения импульсов действия — спайков, которые, как считается, являются основой «нейронного кода», осуществляя также различные функции передачи и преобразования информации. Так, например, нейроны нижних олив, составляющие основу оливо-мозжечковой системы, демонстрируют квазипериодические колебания ниже порога возбуждения с фиксированными амплитудой (5—10 мВ) и частотой (8—12 Гц) [14]. Импульсы возбуждения формируются на пиках подпорогового сигнала, а моменты их возникновения однозначно определяются фазой колебаний. Таким образом, структуры коллективной активности нейронов нижних олив можно классифицировать как фазовые кластеры [18]. Согласно представлениям нейрофизиологов, эти фазовые кластеры и определяют моторные паттерны, задающие шаблоны мышечных сокращений [16]. Например, синфазные колебания нейронов определяют одновременное появление импульсов действия — спайков и, как следствие, синхронное сокращение определенных групп мышц. Изменение фазы колебаний нейрона нижних олив и, как следствие, конфигурации фазовых кластеров происходит при получении системой сенсорного или командного стимула. В таламо-кортикальной системе явление временного связывания выражается в совпадении моментов возникновения спайков или, другими словами, в фазовой синхронизации колебаний [12]. Явление фазовой пресессии в нейронах гиппокампа состоит в изменении фазы спайков на фоне опорного осцилляторного сигнала [18]. Это изменение происходит за счет сенсорного стимула, изменяющегося при перемещении объекта в пространстве. Таким образом, нейронная система способна как отслеживать текущее положение, так и предсказывать последующее движение.

Ключевой задачей при моделировании колебательных процессов в нейронных системах является исследование отклика нейрона на внешний импульсный сигнал. Рассматривая нейрон как нелинейную динамическую систему, мы приходим к классической задаче теории колебаний — изучению динамики неавтономной нелинейной системы. Здесь широко известны, например, явления вынужденной синхронизации, захвата фазы и хаотизации колебаний [20, 21]. Воздействие одиночного импульса на классический генератор Ван-дер-Поля приводит к фазовой переустановке. После окончания входного импульса система восстанавливает частоту и амплитуду автоколебаний, однако фаза колебаний сдвигается по отношению к исходной. Очевидно, что установившаяся фаза зависит от исходной фазы и параметров стимула (амплитуды и длительности). Однако совершенно по-другому реагируют на стимул нейроны нижних олив (рис. 1) [17].

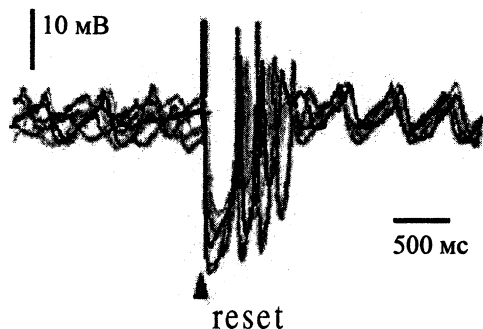


Рис. 1. Фазовая переустановка в нейронах нижних олив. Экспериментальные данные исследования срезов ствола головного мозга крысы. При наложении стимула колебания переустанавливаются к одинаковой фазе.

При получении входного импульса переустановка происходит таким образом, что колебания всегда возвращаются в одинаковую фазу, значение которой не зависит от момента прихода стимула (исходной фазы) и определяется только параметрами самого стимула (амплитудой и длительностью). Рассмотрим отличительные динамические механизмы фазовой переустановки в нейронных системах, фазового кодирования и обсудим возможность формирования фазовых кластеров заданной конфигурации в колебательных нейронных ансамблях. Для моделирования собственной динамики нейронов будем использовать известную модель ФитцХью — Нагумо [3, 22, 23]. Эта модель, с одной стороны, воспроизводит все ключевые черты динамики нейронов, с другой — представляется наиболее выгодной для теоретического рассмотрения в неавтономном случае и расчета динамики больших ансамблей нейронов.

2. Фазово-управляемые колебания в модели ФитцХью — Нагумо

2.1. Фазовая автоперустановка

Рассмотрим одну из простейших модификаций модели ФитцХью — Нагумо следующего вида:

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

Здесь переменная u определяет (качественно) мембранный потенциал клетки, v — «восстанавливающая» переменная (определяет динамику трансмембранных ионных токов, отвечающих за восстановление потенциала покоя нейрона) [3]. Нелинейность $f(u)$ — кубического типа (для определенности $f(u) = u(u - a)(1 - u)$, $0 < a < 1$), ε — параметр, оп-

ределяющий характерные временные масштабы импульса возбуждения, I — параметр, характеризующий уровень деполяризации нейрона (близость системы к порогу возбуждения). Сигнал $I(t)$ — внешний стимул, представляющий собой прямоугольные импульсы или последовательности импульсов различной амплитуды и длительности.

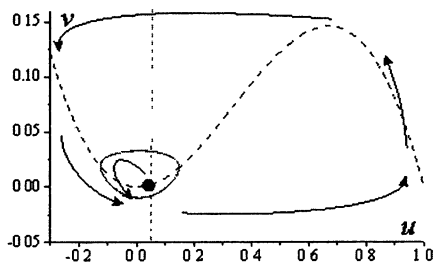


Рис. 2. Фазовая плоскость модели (1) в режиме подпороговых колебаний.

Система (1) имеет единственное состояние равновесия в точке $O(I, f(I))$. При $I > u_{\min}$, где u_{\min} — координата минимума функции $f(u)$, это состояние теряет свою устойчивость в результате бифуркации Андронова — Хопфа и в системе мягко рождается устойчивый предельный цикл, отвечающий квазисинусоидальным колебаниям модели (рис. 2). Зафиксируем для определенности следующие значения параметров: $a = 0,01$, $\epsilon = 0,02$, $I = 0,01$. Заметим, что в этом случае система (1) может рассматриваться как Ca^{2+} -подсистема модели нейрона нижних олив [6]. При этом квазисинусоидальные колебания можно интерпретировать как подпороговые колебания нейрона. Отметим также, что импульсы действия (спайки) возникают на пиках (по u -переменной) этих колебаний, следовательно однозначно определяются фазой последних.

Введем в рассмотрение некоторый периодический сигнал $x(t)$, период которого T совпадает с периодом подпороговых колебаний. Такой сигнал можно интерпретировать как опорный, генерируемый, например, соседним нейроном при отсутствии стимула ($I(t) = 0$ в системе (1)). Введем относительную фазу подпороговых колебаний следующим образом:

$$\varphi = 2\pi \frac{P_u - P_x}{T} \bmod 2\pi, \quad (2)$$

где p_u и p_x — времена максимумов подпорогового и опорного сигналов соответственно (рис. 3).

При внешнем импульсном воздействии амплитуда и частота подпороговых колебаний «восстанавливаются» к значениям, определяемым предельным циклом автономной (невозмущенной) системы. Однако фаза колебаний изменяет свое значение от некоторого φ_1 к φ_2 (рис. 3).

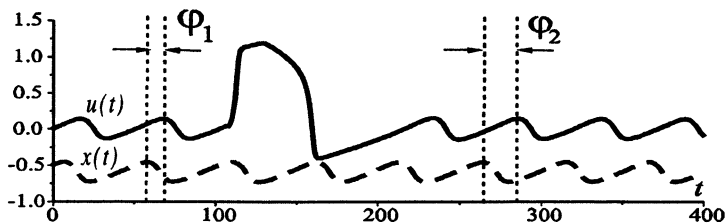


Рис. 3. Фаза подпороговых колебаний, определяемая как временной сдвиг относительно опорного сигнала.

Предположим теперь, что на систему действует одиночный прямоугольный импульс с амплитудой A_{st} и длительностью τ . Момент поступления импульса совместим с максимумом опорного сигнала. Будем изменять начальную фазу φ_1 в интервале $[0, 2\pi]$. На рис. 4 представлены временные реализации колебаний модели при различных начальных фазах.

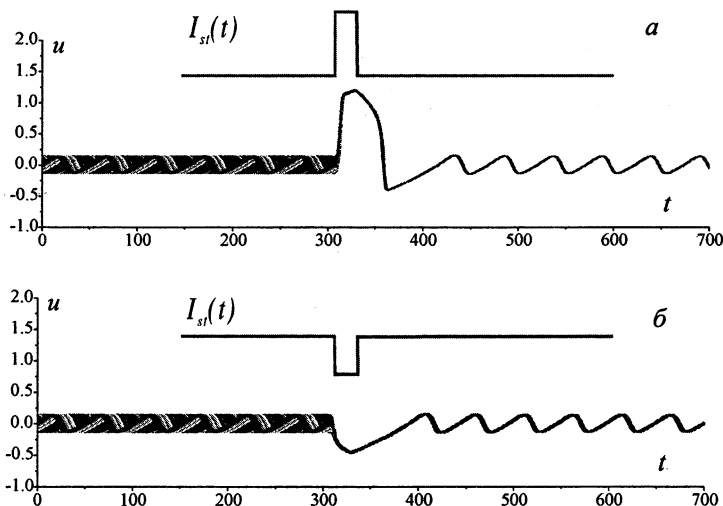


Рис. 4. Фазовая переустановка в модели ФитцХью — Нагумо. Временные реализации для различных начальных условий, синхронизованные относительно импульса-стимула. Значения параметров: $A_{st} = 1,15$, $\tau_{st} = 0,4 T$ (а); $A_{st} = -1$, $\tau_{st} = 0,4 T$ (б).

Очевидно, что фаза установившихся колебаний оказывается одинаковой ($\varphi_2 \approx \text{const}$), для любых значений начальной фазы. Отметим, что эффект фазовой переустановки в модели ФитцХью — Нагумо, фактически, повторяет экспериментальные данные с нейронов нижних олив (см. рис. 1). На рис. 5 представлена зависимость стандартного отклонения (разброса)

переустановленных фаз (нормированного на единицу) от амплитуды импульса-стимула, рассчитанная для $N = 100$ временных реализаций с различной начальной фазой. Эта зависимость описывается некоторой колоколообразной функцией с быстро спадающим «хвостом». При достаточно малой амплитуде импульса фазовой переустановки не происходит, фазовая переменная остается распределенной в интервале $[0, 2\pi]$. При увеличении амплитуды происходит сильное сжатие фазовой переменной к некоторому среднему значению φ^* , которое будет определено ниже в пункте 2.3. Заметим, что при увеличении амплитуды импульса (интенсивности воздействия) девиация фазовой переменной асимптотически уменьшается. Отметим также, что фазовая переустановка имеет место как для положительной амплитуды импульса ($A_{st} > 0$, рис. 4, а), так и для отрицательной ($A_{st} < 0$, рис. 4, б). Другими словами, переустановка происходит как в случае возбуждающего стимула, так и в случае подавляющего (тормозящего) воздействия на нейрон. При этом «точность» переустановки выше для возбуждающего импульса (рис. 4, 5).

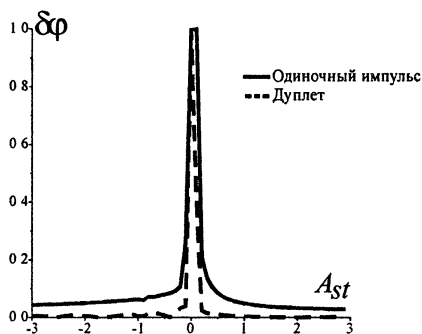
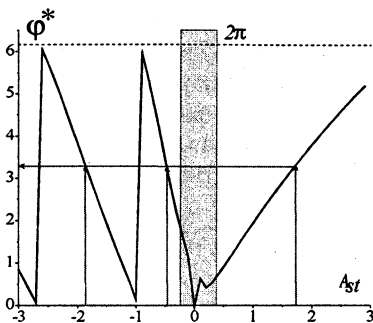


Рис. 5. Зависимость стандарта отклонения фаз колебаний от амплитуды импульса-стимула при фиксированной длительности импульса $\tau = 0,4 T$. Расчет проводился по $N = 100$ временных реализаций с начальными фазами, равномерно распределенными в интервале $[0, 2\pi]$. Пунктирная кривая соответствует отклику на дуплет импульсов с достаточно большим ($\sim 10 T$) межимпульсным интервалом.

Исследуем теперь зависимость установившейся фазы φ^* от амплитуды импульса (рис. 6).

При изменении амплитуды значение установившейся фазы может принимать любые значения в интервале $[0, 2\pi]$. Таким образом, любое требуемое значение фазы колебаний в модели (1) может быть установлено выбором соответствующей амплитуды импульсного воздействия (см. рис. 6) независимо от начальной фазы. Система, фактически, «забывает начальные условия». По аналогии с автоколебательными процессами в нелинейных диссипативных системах эффект можно трактовать как фазовую автопереустановку.

Рис. 6. Зависимость установившейся фазы от амплитуды импульса при фиксированной длительности $\tau_{st} = 0,4 T$.



2.2. Динамический механизм фазовой автопереустановки

Исследуем динамический механизм фазовой автопереустановки. При отсутствии стимула ($I_{st} = 0$) наличие в системе (1) малого параметра $\epsilon \ll 1$ позволяет разделить движения на быстрые и медленные. Медленные движения происходят в окрестности кривой медленных движений $v = f(u)$ в слое толщиной $\sim \epsilon$. Два участка этой кривой с $f'(u) < 0$ являются устойчивыми по отношению к быстрым движениям $v = v_0 = \text{const}$, происходящем вне слоя (см. рис. 2). Устойчивый предельный цикл сравнительно малой амплитуды, отвечающий подпороговым колебаниям модели, расположен в окрестности минимума функции $v = f(u)$ и определяется медленными движениями системы. Мгновенная фаза колебаний характеризуется угловым положением изображающей точки на предельном цикле при $I_{st} = 0$. Рассмотрим преобразование этой кривой при воздействии на систему одиночного прямоугольного импульса. Во время действия импульса $0 < t < \tau_{st}$ динамика модели описывается автономной системой ФитцХью — Нагумо с эффективным параметром деполяризации $I_{3\phi} = I + A_{st}$:

$$\begin{cases} \dot{u} = f(u) - v, \\ \dot{v} = \epsilon(u - I_{3\phi}). \end{cases} \quad (3)$$

Потребуем, чтобы амплитуда импульса была достаточно велика, $A_{st} > u_{\max} - I$. В этом случае система (3) имеет устойчивое состояние равновесия, расположенное на правой ветви кривой $f(u)$ (см. рис. 2). Потребуем, чтобы длительность импульса τ_{st} была также достаточно велика, чтобы за это время все траектории с начальными условиями на кривой оказались в области быстрых движений, расположенной между средней и правой ветвями функции $f(u)$ (рис. 7, а). Тогда образ исходной кривой, отвечающей предельному циклу, будет определяться некоторой замкнутой кривой, вытянутой вдоль быстрых движений. По окончании действия импульса, $t > \tau_{st}$, динамика системы определяется исходным параметром деполяризации I (см. рис. 2) и образ кривой продолжает эволю-

ционировать по траекториям невозмущенной системы (1). При движении в слое медленных движений вдоль устойчивой (по отношению к быстрым движениям) правой ветки кривой $f(u)$ происходит сжатие образа в «продольном» направлении. Далее, преодолев участок быстрых движений, образ вновь сжимается в окрестности левой ветки и возвращается в окрестность устойчивого предельного цикла (см. рис. 7, а).

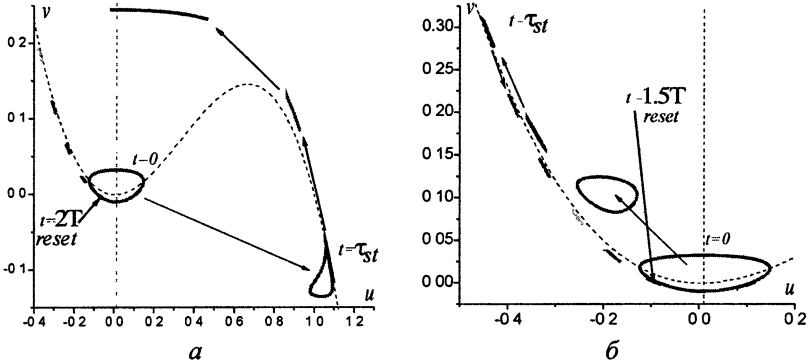


Рис. 7. Механизм фазовой переустановки на фазовой плоскости (u, v) . Преобразование замкнутой кривой, отвечающей предельному циклу под действием стимула. Последовательность мгновенных снимков $N = 100$ начальных условий, равномерно распределенных по предельному циклу при $t = 0$. Возбуждающий стимул: а — $A_{st} = 1,15$, $\tau_{st} = 0,4$ T; б — $A_{st} = -1$, $\tau_{st} = 0,4$ T.

В результате, любая начальная фаза под действием стимула оказывается переустановленной, приближаясь к единственному значению φ^* , которое определяется параметрами системы и импульса-стимула. Пренебрегая временем движения по быстрым траекториям «точность» переустановки можно оценить сжатием образа в окрестности кривой медленных движений. Полагая $v = v_0 = \text{const}$, получаем из первого уравнения системы (1), что правая и левая ветви медленных движений устойчивы по отношению к быстрым с ляпуновским показателем

$$\lambda \sim f'(u_0) < 0, u_0 = f(v_0). \quad (4)$$

Из второго уравнения системы (1) легко видеть, что время движения по медленным ветвям $\Delta t \sim 1/\epsilon$. Тогда показатель экспоненциального сжатия образа по переменной $\delta \lambda_u = \ln \delta u \sim f(v_0)/\epsilon$. Другими словами, при достаточно малом ϵ в системе происходит сильное сжатие фазового объема в окрестности кривой медленных движений. Таким образом, фазовая автоперустановка в модели ФитцХью — Нагумо определяется релаксационным характером движений на фазовой плоскости.

В случае тормозящего (ингибиторного) стимула $A_{st} < 0$ переустановка происходит подобным образом. При воздействии стимула образ

инвариантной кривой «перебрасывается» в окрестность левой ветви кривой медленных движений (рис. 7, б). Ясно, что в этом случае «точность» переустановки снижается по сравнению с возбуждающим стимулом (рис. 6), так как время эволюции по медленным движениям уменьшается.

2.3. Стимул-индуцированная синхронизация нейронных систем

Эффект фазовой автоперустановки можно использовать для фазовой синхронизации колебаний больших нейронных ансамблей. Так как устанавливаемая фаза не зависит от исходной фазы, то при воздействии одинаковым импульсом на элементы ансамбля фаза каждого из них будет переустановлена к одному и тому же значению φ^* (рис. 8). Заметим, что эффект не требует абсолютной идентичности элементов ансамбля, они могут иметь некоторый разброс по частотам. Этот разброс, как показано на рис. 8, приводит к расплыванию фазы и выходу элементов из режима синхронизации.

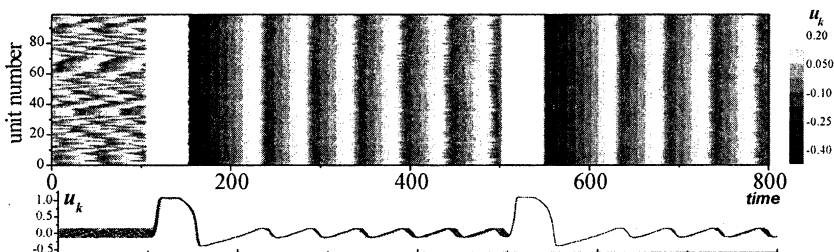


Рис. 8. Пространственно-временная диаграмма стимул-индуцированной синхронизации ансамбля из $N = 100$ невзаимодействующих элементов ФитцХью — Нагумо.

Повторное стимулирование системы импульсом вновь приводит к установлению синхронного режима. Заметим также, что в отличие от классической синхронизации ансамблей автоколебательных систем, осуществляемой за счет непосредственного взаимодействия между элементами (например, диссипативной связи), стимул-индуцированная синхронизация не требует непосредственного контакта между элементами. В пространстве они могут находиться на значительном удалении друг от друга и не взаимодействовать вообще. Ясно, что такой эффект требует некоторого «центрального» элемента, формирующего импульс-стимул. Более того, этот импульс должен поступать на вход каждого из элементов одновременно. Эти два условия определяют архитектуру построения системы стимул-индуцированной синхронизации (рис. 9). Такой архитектурой обладают, например, таламо-кортикальная и оливо-

мозжечковая нейронные системы. В первом случае роль «центрального» осциллятора играет таламус, который «опрашивает» различные области коры, в оливо-мозжечковой системе — нейроны нижних олив, импульсы возбуждения которых происходят на фоне подпороговых колебаний (опорного сигнала). Эти импульсы распространяются по аксонам — «лазающим» волокнам, возбуждая клетки Пуркинье коры мозжечка. Как было установлено в недавних экспериментах нейрофизиологов [15], это возбуждение происходит в высшей степени синхронизованно. Время распространения импульса оказывается одинаковым для клеток, расположенных в разных областях коры. Это обеспечивается различной скоростью распространения импульса по аксонам различной толщины.

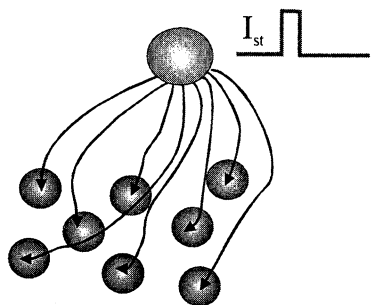


Рис. 9. Архитектура построения системы фазовой автоперестройки с «центральным» осциллятором.

2.4. Описание фазовой автоперестройки. Отображение фазы

Для описания эффекта фазовой автоперестройки построим отображение фазы. Будем предполагать, что входной сигнал $I_{st}(t)$ представляет собой последовательность прямоугольных импульсов с амплитудой A_{st} , длительностью τ_{st} и межимпульсным интервалом T_s , удовлетворяющую следующим предположениям. (А) Момент прихода i -го импульса в последовательности синхронизован с максимумом опорного сигнала, $t_i = p_x + mT$, $m \in \mathbb{Z}$. (Б) Интервал следования импульсов достаточно велик по сравнению с периодом подпороговых колебаний $T_s = nT$, $n \gg 1$.

Фазовое пространство неавтономной системы (1) является трехмерным. При выполнении условия (Б) на сечениях Пуанкаре $t = t_i = p_x + mT$, $m \in \mathbb{Z}$, фазовые траектории возвращаются в окрестность замкнутой кривой, отвечающей предельному циклу автономной системы (см. рис. 2). Используя определение относительной фазы (2), введем отображение фазы

$$\Phi: \varphi_i \rightarrow \varphi_{i+1}, \quad (5)$$

где φ_i — фаза колебаний на момент прихода i -го импульса стимула. Отображение (5) является отображением окружности и определено на интервале $[0, 2\pi]$. Оно описывается некоторой функцией $\varphi_{i+1} = \Phi(\varphi_i)$, которую называют функцией фазового отклика (*phase response curve*).

В случае, когда интенсивность стимула достаточно велика (п. 2.2), функция фазового отклика близка к константе (рис. 10). Отображение (5) имеет единственную неподвижную точку φ^* , координаты которой определяются параметрами стимула. Тогда любая начальная фаза отображается в малую окрестность этой неподвижной точки (рис. 10). Заштрихованная область, отвечающая образу интервала $[0, 2\pi]$, под действием отображения (5) задает «точность» фазовой автоперестройки. Ясно, что повторное итерирование отображения (5), соответствующее дуплету импульсов стимула, еще более сжимает этот образ (рис. 5, 10).

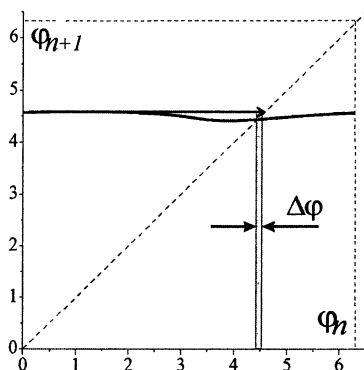


Рис. 10. Отображение фазы при фазовой автоперестройке

Так как импульсы стимула, синхронизованы с опорным сигналом (условие (A)), фазовая автоперестройка приводит к захвату фазы колебаний со сдвигом φ^* относительно опорного сигнала. Выбирая соответствующим образом амплитуду импульса, можно получить любое наперед заданное значение φ^* в интервале $[0, 2\pi]$.

3. Формирование фазовых кластеров заданной конфигурации

В ансамблях нейроноподобных элементов эффект фазовой перестройки позволяет формировать фазовые кластеры любой наперед заданной пространственной конфигурации. Путем выбора соответствующим образом интенсивности стимула фаза каждого элемента может быть установлена к любому значению независимо от текущего состояния. Процесс формирования фазовых кластеров иллюстрирует рис. 11. При воздействии стимула фазы элементов группируются согласно стимулу «клеточной» конфигурации. Заметим, что для формирования кластеров не требуется непосредственного взаимодействия между элементами. Такое взаимодействие может быть введено для поддержания заданной конфигурации на время, когда стимул отсутствует. Подобным образом, например, функционирует оливо-мозжечковая система. При

отсутствии стимула нейроны нижних олив взаимодействуют через гэп-контакт, что приводит к их взаимной синхронизации. При поступлении стимула за счет тормозящей петли обратной связи происходит подавление гэп-контакта и фазы элементов устанавливаются к требуемому значению под воздействием сенсорного стимула (см. рис. 1). Формируемый таким образом моторный паттерн любой сколь угодно сложной пространственной конфигурации задает требуемый синхронизм мышечных сокращений. Группы мышц, отвечающие фазовому кластеру, сокращаются одновременно [6].

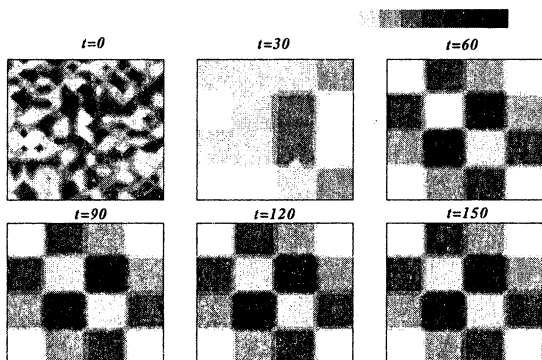


Рис. 11. Формирование фазовых кластеров в ансамбле из $N = 20 \times 20$ невзаимодействующих элементов (1). При $t = 0$ на элементы со случайно разбросанными фазами одновременно действует одиночный импульс-стимул длительностью $\tau_{st} = 0,4 T$. Амплитуда импульса для каждого элемента выбирается в соответствии с рис. 6, устанавливая требуемую фазу согласно цветовой шкале.

Таким образом, эффект фазовой автоперезагрузки, фактически, является сенсоромоторным преобразованием, при котором интенсивность сенсорного стимула конвертируется в шаблон мышечных сокращений, управляющий движением.

4. Фазовые последовательности. Нейронный код

Фазовые кластеры, формируемые с помощью фазовой автоперезагрузки подпороговых колебаний нейронов, определяют пространственные структуры спайковой активности нейронной системы, так как времена возникновения импульсов возбуждения (спайков) однозначно определяются фазой подпороговых колебаний. При этом временная динамика системы, ее «предыстория», фактически «стирается», поскольку устанавливаемая фаза никак не зависит от предыдущей. Однако очевидно, что различные функции запоминания, предсказания, осуществления моторных программ и др. требуют временного представления сенсорной информации. На вопрос о том, как нейронные системы формируют вре-

менной код, преобразуют и обрабатывают сенсорную информацию, до сих пор нет однозначного ответа. Как показывают исследования нейрофизиологов, временное кодирование информации осуществляется с помощью изменяющихся интервалов следования импульсов (межспайковых интервалов) [24, 25].

Рассмотрим один из возможных вариантов временного кодирования информации с помощью фазовой переустановки в модели (1). Будем предполагать, что длительность импульсов стимула, удовлетворяющего условиям (А) и (В), сравнительно мала. Положим для определенности, $\tau_{st} \sim 0,01 T$. Такое соотношение, в частности, выполняется для импульсов возбуждения нейронов нижних олив. Тогда модель (1) можно трактовать как нейрон, получающий импульсный сигнал от другого нейрона через синаптическую связь. Такая трактовка представляется обоснованной еще и потому, что синаптическая связь, обладающая порогом, фактически «фильтрует» подпороговые колебания. В результате остается импульсная последовательность с дискретно изменяющимся межимпульсным интервалом (условие (В)).

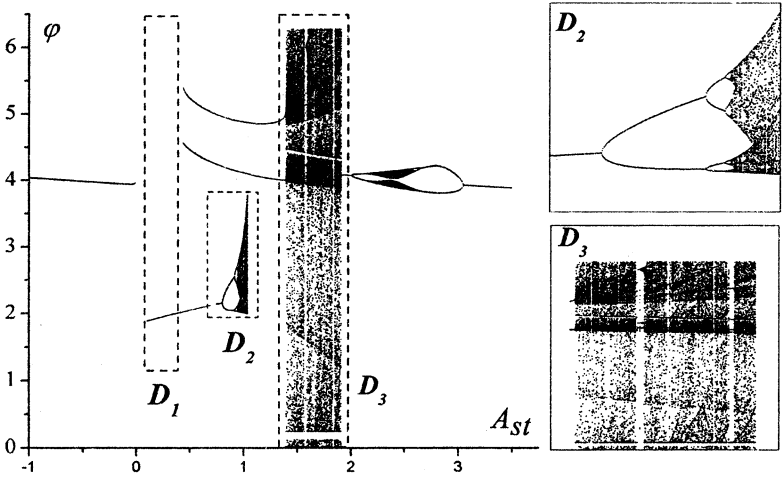


Рис. 12. Бифуркационная диаграмма отображения фазы в зависимости от амплитуды импульса при фиксированной длительности $\tau_{st} = 0,01 T$. Различные начальные точки (начальная фаза) итерировались с помощью отображения (5). Для каждого значения A_{st} на диаграмме отмечалось 100 точек после 1000 итераций отображения. Эти точки характеризуют распределение фаз колебаний на аттракторах модели (1).

Рассмотрим теперь фазовый отклик такой модели. Отображение фазы (5) определяет фазовые последовательности, изменяющие фазу подпороговых колебаний после прихода очередного импульса. На рис. 12 представлена бифуркационная диаграмма отображения фазы при изме-

нении амплитуды импульса. Рисунки иллюстрируют установившиеся режимы фазовой динамики отображения (5).

При $A_{sr} \ll 1$ отображение имеет единственную устойчивую неподвижную точку P с мультипликатором $\mu(P)$, близким к 1, и фазовая последовательность сходится к ней при большом числе итераций (рис. 13, *а*, область D_1 на бифуркационной диаграмме рис. 12). При увеличении A_{sr} точка P теряет устойчивость и в отображении происходит каскад бифуркаций удвоения периода (область D_2), ведущий к возникновению хаотического аттрактора, сосуществующего с устойчивым периодическим движением периода 2 (рис. 13, *б*). В области D_3 динамика отображения характеризуется перемежаемостью. Траектории отображения остаются длительное время в ограниченной области. Эти движения («ламинарная фаза»), характеризующиеся осцилляциями с плавно нарастающей амплитудой, прерываются выбросами («турбулентная фаза») на различные участки кривой отображения (рис. 13, *в*). Области хаотической динамики в интервале D_3 чередуются с устойчивыми периодическими окнами сложного профиля (рис. 13, *д*).

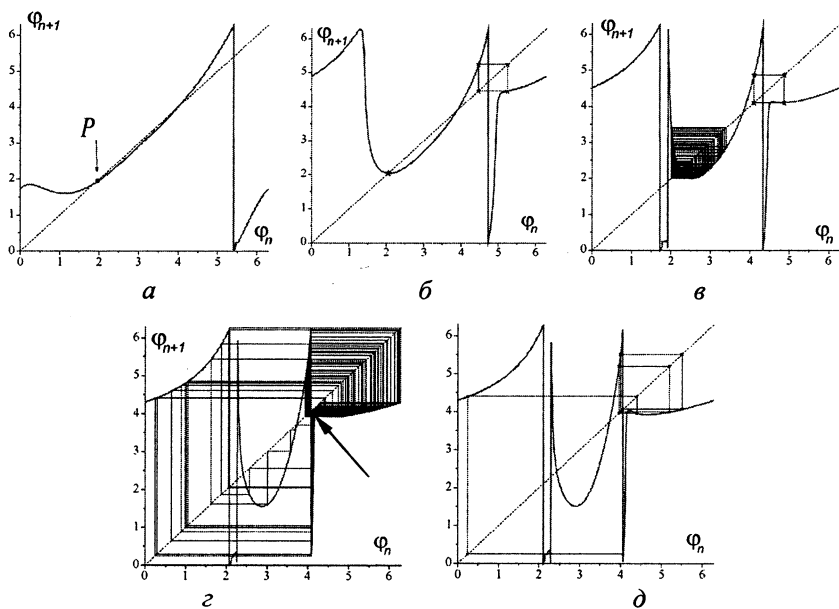


Рис. 13. Отображение фазы (5) при фиксированном $\tau_{sr} = 0,01 T$: *а* — $A_{sr} = 0,15$ (P — устойчивая неподвижная точка); *б* — $A_{sr} = 0,5$ (сосуществование устойчивой неподвижной точки и устойчивого периодического движения периода 2); *в* — $A_{sr} = 1,02$ (сосуществование хаотического аттрактора и устойчивого периодического движения периода 2); *з* — $A_{sr} = 1,62$ (хаотический аттрактор с перемежаемостью); *д* — $A_{sr} = 1,66$ (устойчивое периодическое движение в области D_3).

На рис. 14, *а* представлена временная динамика фазовой переменной для случая хаотического аттрактора в области D_2 (рис. 13, *в*). Фаза подпороговых колебаний хаотически осциллирует в ограниченной (инвариантной) области, отвечающей хаотическому аттрактору. В случае сценария перемежаемости (рис. 13, *з*) фазовая переменная оказывается распределенной во всем интервале $[0, 2\pi]$. Временная реализация (рис. 14, *б*) характеризуется «выбросами» на фоне «ламинарной» динамики колебаний с нарастающей амплитудой. Рисунок 14, *в* иллюстрирует сложное периодическое движение периода 7 в окне периодичности области D_3 . При увеличении амплитуды импульса $A_{st} > 3$ отображение (5) имеет единственную устойчивую неподвижную точку. Она соответствует захвату фазы подпороговых колебаний последовательностью импульсов стимула.

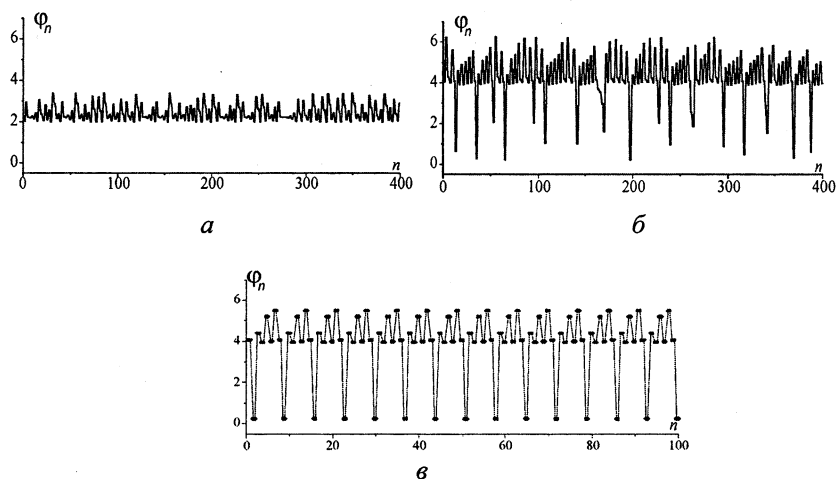


Рис. 14. Временные реализации хаотических аттракторов (*а*, *б*) отображения фазы рис. 13, *в* и *з* соответственно; *в* — устойчивое периодическое движение, отвечающее отображению рис. 13, *д*.

Таким образом, под воздействием импульсного сигнала модель нейрона (1) способна формировать различные временные структуры спайковой активности. Эти структуры представляют собой фазовые последовательности, временная динамика которых определяется отображением фазы (5). В свою очередь, фаза подпороговых колебаний задает моменты возникновения спайков нейронов. Другими словами, фазовые последовательности можно трактовать как элементы динамической кодировки, где определенной амплитуде (длительности) стимула соответствует определенный временной отклик. Так, например, в случае периодических движений (рис. 13, *д*, 14, *в*) фазовая последовательность периода 7 задается амплитудой стимула $A_{st} = 1,66$. Более того, динамика

отображения (5) может быть мультистабильной (рис. 13, б, в). В этом случае временной отклик зависит не только от параметров стимула, но и от того, в области притяжения какого из аттракторов находилась в начальный момента система. Заметим, что «специфика» того или иного фазового отклика (мультистабильность, периодические и хаотические траектории) не содержится в стимулирующем импульсном сигнале, а определяется динамическими механизмами самого нейрона, формирующего отклик.

5. Заключение

Отметим несколько ключевых моментов колебательной динамики модели (1) применительно к исследованию информационных процессов в нейронных системах.

1. При отсутствии внешнего стимула модель генерирует спонтанные подпороговые колебания квазисинусоидальной формы. Эти колебания могут служить опорным сигналом (при моделировании нейронов нижних олив) для формирования импульсов возбуждения спайков. Вследствие этого основной информационной характеристикой является фаза колебаний, однозначно определяющая моменты возникновения спайков. При воздействии на такую систему стимулом достаточно большой интенсивности (амплитуды и длительности) в модели происходит фазовая автопереустановка. Переустановленная таким образом фаза определяется только параметрами стимула и не зависит от начальных условий (начальной фазы). В частности, выбирая соответствующим образом амплитуду стимула, можно установить любую наперед заданную фазу подпороговых колебаний.

2. На основе эффекта фазовой автопереустановки можно синхронизовать колебания нейронных ансамблей. Для такой стимул-индуцированной синхронизации не требуется непосредственного контакта, связи между элементами. Таким образом, нейроны, находящиеся на значительном удалении друг от друга, могут быть эффективно синхронизованы единым импульсным (сенсорным) стимулом.

3. С помощью стимула различной интенсивности можно формировать кластерные структуры спайковой активности. Такие структуры ассоциируются, например, с моторными паттернами, формируемыми оливо-мозжечковой нейронной системой. Элементы с одинаковой фазой колебаний определяют синхронные сокращения различных групп мышц моторной системы.

4. Импульсный стимул способен формировать временные структуры спайковой активности. Такие структуры реализуются как фазовые последовательности, где фаза подпороговых колебаний последовательно переустанавливается по приходу очередного импульса-стимула. Фазовые последовательности можно трактовать как своего рода нейронный код, реализующий определенную последовательность «действий», например моторную программу.

Эти исследования проводились при финансовой поддержке гранта РФФИ 03-02-17135, гранта президента РФ для поддержки молодых ученых МК-4586.2004.2, гранта CRDF (REC-006), Фонда содействия отечественной науке.

Литература

1. Principles of neural science. Third Edition / Eds. E. R. Kandel, J. H. Schwartz, T. M. Jessell. Prentice-Hall Intern. Inc., 1991.
2. Николлс Дж., Мартин Р., Валлас Б, Фукс П. От нейрона к мозгу. Изд-во УРСС, 2003. 672 с.
3. Scott A. Neuroscience: a mathematical premier. Berlin: Springer-Verlag, 2002.
4. Llinas R. I of vortex. From neurons to self. Massachusetts: The MIT Press, 2002.
5. Казанцев В. Б., Некоркин В. И. Динамика колебательных нейронов. Информационные аспекты // Нелинейные волны – 2002. Н. Новгород: ИПФ РАН, 2003. С. 9—33.
6. Kazantsev V. B., Nekorkin V. I., Makarenko V. I., Llinas R. Olivo-cerebellar cluster-based universal control system // Proc. Natl. Acad. Sci. USA. 2003. V. 100. P. 13064—13068.
7. Koch C. Biophysics of computation: information processing in single neurons. Oxford University Press, 1998.
8. Grillner S. Neural Networks for Vertebrate Locomotion // Sci. Am. 1996. V. 64.
9. Ekeberg O. A combined neuronal and mechanical model of fish swimming // Biol. Cybern. 1993. V. 69. P. 363.
10. Потанов А. Б., Али М. К. Нелинейная динамика обработки информации в нейронных сетях // Новое в синергетике: Взгляд в третье тысячелетие / Ред. Г. Г. Малинецкий, С. П. Курдюмов. М.: Наука, 2002.
11. Казанцев В. Б., Некоркин В. И. Принципы контроля и координации движений на основе динамики нейронов головного мозга // Изв. вузов. Прикладная нелинейная динамика. 2001. Т. 9, № 1. С. 38—48.
12. Llinas R. Consciousness and the thalamocortical loop // Int. Congress Ser. 2003. V. 1250. P. 409—416.
13. Behrendt R. P. Hallucinations: Synchronisation of thalamocortical c oscillations under-constrained by sensory input // Consciousness and Cognition. 2003. V. 12. P. 413—451.
14. Llinas R., Yarom Y. Oscillatory properties of guinea-pig inferior olivary neurones and their pharmacological modulation: An *in vitro* study // J. Physiol. 1986. V. 376. P. 163.
15. Lang E. J., Sugihara I., Welsh J. P., Llinas R. Patterns of spontaneous purkinje cell complex spike activity in the awake rat // J. Neurosci. 1999. V.19, № 7. P. 2728—2739.
16. Welsh J. P., Lang E. J., Sugihara I., Llinas R. Dynamic organization of motor control within the olivocerebellar system // Nature. 1995. V. 374. P. 453—457.
17. Leznik E., Makarenko V. I., Llinas R. Electrotonically mediated oscillatory patterns in neuronal ensembles: An *in vitro* voltage-dependent dye-imaging study in the inferior olive // J. Neurosci. 2002. V. 22. P. 2804—2815.
18. Henze D. A., Buzsak G. Single cell contributions to network activity in the hippocampus // Int. Congress Ser. 2003. V. 1250. P. 161.
19. Magee J. C. A prominent role for intrinsic neuronal properties in temporal coding // Trends Neurosci. 2003. V. 26, № 1. P. 14.
20. Пиковский А., Розенблюм М., Куртс Ю. Синхронизация. Фундаментальное нелинейное явление. М.: Техносфера, 2003. 496 с.
21. Гласс Л, Мэски М. От часов к хаосу. Ритмы жизни. М.: Мир, 1991. 248 с.
22. FitzHugh R. Mathematical models of excitation and propagation in nerve // Biological Engineering / Ed. H. P. Schwan. 1969. P. 1—85.
23. Izhikevich E. M. Neural excitability, spiking and bursting // Int. J. Bifurc. Chaos. 2000. V. 10, № 6. P. 1171—1266.
24. Eguia M. C., Rabinovich M. I., Abarbanel H. D. I. Information transmission and recovery in neuron communication channels // Phys. Rev. E. 2000. V. 62. P. 7111—7122.
25. Brenner N., Strong S. P., Koberle R., Bialek W., de Ruyter van Steveninck R. R. Synergy in a neural code // Neural Computation. 2000. V. 12. P. 1531—1552.

МОДЕЛЬНЫЕ НЕЙРОНОПОДОБНЫЕ СИСТЕМЫ. ПРИМЕРЫ ДИНАМИЧЕСКИХ ПРОЦЕССОВ

В. Г. Яхно, И. В. Нуйдель, А. Е. Иванов

1. Введение

Примеры нейроноподобных систем широко известны (см., например, [1—7]) и чаще всего они представлены объектами биологической или химической природы, но к ним также относятся и многие искусственные неравновесные распределенные системы преобразования информационных сигналов. Наиболее распространены модельные описания для систем из биомембран, нейронов, нейронных ансамблей, нервных тканей, искусственные аналоги нейронных систем.

Результаты исследований режимов функционирования нейроноподобных систем сильно зависят от вида выбранных модельных описаний.

В предыдущих статьях [7—16] были приведены варианты математических моделей для описания динамики преобразования сигналов (изображений) в однородных нейроноподобных системах [7—12] и настраивающихся распознающих системах [12—16].

В данной работе рассматриваются примеры динамических процессов в модели из трех взаимосвязанных распределенных нейроноподобных подсистем. Каждая распределенная подсистема состоит из однотипных нейроноподобных элементов, описывающих работу нейронов коры, таламуса и нейронов ретикулярно-таламического ядра. Влияния этих подсистем друг на друга соответствуют известной архитектуре прохождения сенсорных сигналов в нейронных ансамблях у животных [17—20]. Определение функциональной адекватности выбранной модели «таламус — кора — ретикулярные ядра таламуса» опирается на сопоставление качественного вида динамики получаемых решений модельных уравнений с процессами, зарегистрированными в экспериментах. Рассмотрим некоторые примеры динамических режимов и возможных реакций на внешние стимулы для выбранной базовой нейроноподобной модели.

2. Режимы стробирования входного сенсорного сигнала на взаимодействующих моделях однородных нейроноподобных систем

Рассмотрим варианты преобразования входной информации (представленные в виде изображений) в однородных нейроноподобных системах на примере таламо-кортикальных цепочек, по которым проходят все сенсорные сигналы у животных [17—22] (рис. 1).

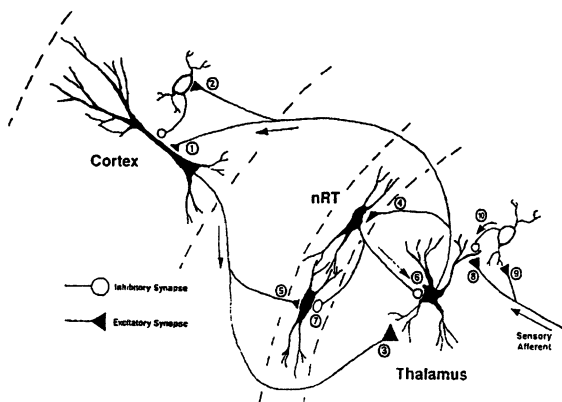


Рис. 1. Основная таламико-кортикальная цепь преобразования входных сенсорных сигналов [19]

На кору поступают потоки сигналов как от специфических, так и от неспецифических систем. Специфическими называют такие системы, в которых осуществляется передача и обработка информации от разных сенсоров в соответствующие зоны коры, а также системы, с помощью которых передаются команды от коры к нейронам спинного мозга, а затем и к мышцам. Неспецифические системы — организующие, контролирующие, регулирующие системы самоорганизации мозга как целого. Они получают информацию от всех органов чувств и регулируют все виды деятельности, имеют связи между собой и со всеми звеньями специфических систем [18—20].

Процесс прерывания потока сенсорной информации (или стробирования) содержит следующие этапы [20]. Когда в кору приходят сенсорные сигналы, каждая локальная область коры возбуждает паттерн активности глубоких пирамид, который остается неизменным в течение 10—100 мс. При этом, за то время, пока существует активность глубоких пирамид, выполняется одна из предварительных операций преобразования сигнала для его распознавания. Например, за счет изменения связей между нейронными ансамблями происходит настройка такой нейронной подсистемы на информативные признаки обрабатываемого сигнала — формируются ассоциативные связи.

На этапе включения активности в коре происходит конкуренция между нейронами: те глубокие пирамиды, которые первыми перешли к активному состоянию, тормозят через интернейроны всех соседей; образуются петли реверберационной активности с соответствующими клетками таламуса — в результате развивается динамический процесс формирования паттерна статической активности. Порядок, в котором переключаются различные кортикальные области, может задавать последо-

вательность операций обработки сигнала для выделения из него информационных признаков.

В конце каждого активного периода таламические ядра, поддерживающие активность глубоких пирамид в кортикальной области, сильно тормозятся из-за активности неспецифических ретикулярных ядер медиального таламуса (далее *nRT*). На следующем этапе таламические ядра перестают активировать глубокие пирамиды в коре. Это означает переход к этапу, когда кортикальная область неактивна. Переход к новому включению происходит из-за уменьшения активности нейронов в *nRT* и соответственного снижения торможения таламических нейронов, которые, в свою очередь, снова активируются под влиянием входного сенсорного сигнала и в результате вновь включают активность глубоких пирамид в коре. В течение всего времени активности глубоких пирамид, они поддерживаются сигналами от специфических и неспецифических клеток ствола мозга и таламуса, а также сигналами от других активных глубоких пирамид в коре.

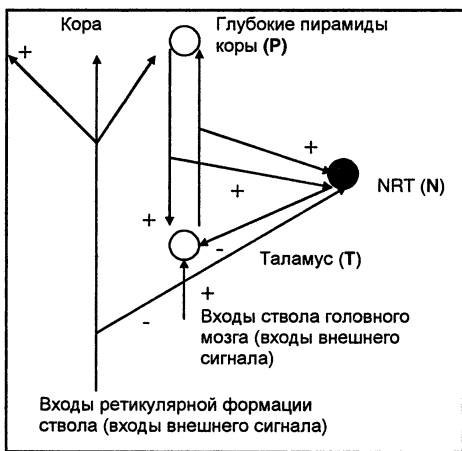


Рис. 2. Схема прохождения внешнего сигнала в кору.

следующая последовательность: активируется таламус, активируются пирамиды, таламус и пирамиды активируют *nRT*, *nRT* подавляет сигнал в таламусе, что приводит к подавлению сигнала в пирамидах, в результате сигнал с *nRT* перестает подавлять клетки таламуса, и они снова активируются. Затем процесс повторяется.

Описываемая функциональная схема прерывания сенсорного сигнала во многих деталях повторяет известную и давно используемую схему свержегенеративных приемников [23, 24]. Такая схема позволяет обеспечить весьма большой коэффициент усиления поступающего сиг-

В течение нерегулярных эпизодов длительностью от 10 до 1000 мс каждая локальная кортикальная область проходит цикл кортикальной активности: включение, активность, выключение, неактивность, новое включение и т. д.

На рис. 2 показаны взаимосвязи и пути поступления внешнего сенсорного сигнала на кору, таламус и *nRT*. Один только внешний сигнал не может активировать глубокие пирамиды в коре. Для этого необходимо также, чтобы в кору поступал еще и сигнал с таламуса. Происходит сле-

нала. Отметим, что в распределенной системе из взаимодействующих друг с другом сверхрегенеративных приемников режимы трансформации сигналов становятся весьма разнообразными и непростыми.

Классификация возможных режимов с помощью таких моделей может использоваться при интерпретации экспериментально регистрируемых сигналов. Понятно также, что в подобной системе, работающей вблизи порога самовозбуждения, возможны такие изменения параметров подсистемы стабилизации, при которых либо снижается чувствительность системы к внешнему сигналу, либо происходит самовозбуждение автоколебаний. Последний процесс вполне может трактоваться как патологический режим, потому что система уже не может обрабатывать входные сигналы и ее пространственно-временной режим определяется, в первую очередь, внутренними параметрами системы взаимосвязанных нейронных ансамблей.

Анализ известных экспериментальных данных подтверждает возможность аналогичных процессов при функционировании системы «таламус — кора — ретикулярные ядра таламуса» в мозге [17—20].

Для качественного теоретического описания рассмотренных динамических процессов использовались соответствующие «базовые» математические модели нейроноподобных систем [7—13, 21, 22].

2.1. Модельное описание

Анализ возможных динамических процессов в соответствии с приведенной схемой взаимосвязей между нейронными ансамблями (рис. 2) стал возможным после разработки специальных исследовательских программных систем. Расчеты проводились на двух вариантах таких систем. В первой системе рассмотрение проводилось на основе трех балансных уравнений, которые описывали величины активности глубоких пирамид, таламуса и nRT . Переменные во всех уравнениях характеризовали усредненные значения активности по ансамблю клеток в слое: $P(r, t)$ — активность глубоких пирамид, $T(r, t)$ — активность таламуса, $N(r, t)$ — активность клеток nRT , $U_{ex}(r, t)$ — внешний сенсорный сигнал, $F_{P,T,N}[Z]$ — соответствующая пороговая функция срабатывания нейроноподобных элементов.

$$\tau_P \frac{\partial P}{\partial t} + P = F_P \left[\alpha_{P,P} \int_D \Phi_{P,P}(\xi) P(r - \xi) d\xi + \alpha_{P,T} \int_D \Phi_{P,T}(\xi) T(r - \xi) d\xi + \right. \\ \left. + \alpha_{P,N} \int_D \Phi_{P,N}(\xi) N(r - \xi) d\xi + \beta_P U_{ex} \right], \quad (1)$$

$$\tau_T \frac{\partial T}{\partial t} + T = F_T \left[\alpha_{T,P} \int_D \Phi_{T,P}(\xi) P(r-\xi) d\xi + \alpha_{T,T} \int_D \Phi_{T,T}(\xi) T(r-\xi) d\xi + \right. \\ \left. + \alpha_{T,N} \int_D \Phi_{T,N}(\xi) N(r-\xi) d\xi + \beta_T U_{ex} \right], \quad (2)$$

$$\tau_N \frac{\partial N}{\partial t} + N = F_N \left[\alpha_{N,P} \int_D \Phi_{N,P}(\xi) P(r-\xi) d\xi + \alpha_{N,T} \int_D \Phi_{N,T}(\xi) T(r-\xi) d\xi + \right. \\ \left. + \alpha_{N,N} \int_D \Phi_{N,N}(\xi) N(r-\xi) d\xi - \beta_N U_{ex} \right]. \quad (3)$$

Для исследования вариантов перехода к самовозбуждению использовались примеры режимов нормальной работы на этапе выделения одного из возможных признаков входного сигнала. Для этого использовались функции пространственной связи типа «латеральное торможение» $\Phi(r) = (1 - br^2) \exp(-ar^2)$ — с положительным центром и отрицательными флангами (рис. 3).

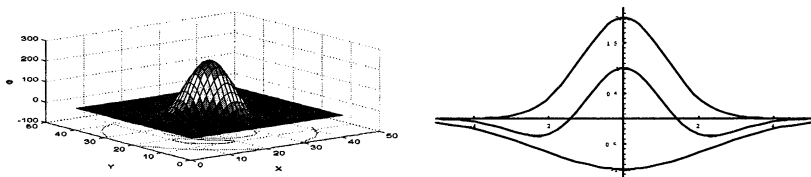


Рис. 3. Пример возможного вида функций связи типа «латеральное торможение» (положительный центр, отрицательные фланги) в системах (1)—(3) или (4), полученных в результате суммирования функций связей для возбуждающих нейронов (положительные связи) и тормозных нейронов (отрицательные связи).

В предыдущих работах [8—11] было показано, что параметр в виде отношения a/b определяет характер решений в однокомпонентной модели распределенной среды и соответственно управляет возможными режимами преобразования начального возмущения. Например, при значениях параметра $a/b < 0,51$ происходит затухание возмущения, а также выделение объектов заданных размеров; при $0,51 < a/b < 0,625$ — выделение объектов заданных размеров и концов линий; при $0,625 < a/b < 1$ — выделение контура заданной толщины; при $a/b > 1$ — зарастивание разрывов, сглаживание, скелетизация (получение центральных осей заданного изображения).

Анализ структур активности в модели (1)—(3) проводился с помощью специальной исследовательской системы, разработанной в интер-

фейсе операционной системы Windows. Такая система представляет собой инструментальное средство, разработанное для изучения разнообразия возможных решений в базовых моделях нейроноподобных распределенных систем. Программная система позволяет проводить анализ как сосредоточенной (без учета пространственных связей), так и распределенной модели (с учетом пространственных связей). Состояния переменных в распределенной модели визуализируются в виде изображений. В сосредоточенной модели отображаются три графика переменных по времени и фазовое пространство. Исследователь на основе полученных ранее данных [8—13, 21, 22, 26, 27] подбирает параметры характеристик сосредоточенной и распределенной моделей систем. Затем включается режим вычислений. Система позволяет наблюдать и регистрировать результаты расчетов изучаемых режимов. Система содержит средство для подбора параметров функций свертки и автоматического создания матриц свертки. Система реализована как 32-разрядное приложение MS Windows 95 (MS Windows NT) и обладает дружественным интерфейсом [21, 22].

Во второй программной системе была введена дополнительная возможность моделирования внутренних ритмов в каждом из трех взаимодействующих нейронных ансамблей. Для этого при описании каждого выделенного нейронного ансамбля вводилась дополнительная переменная G_i , которая могла регулировать порог перехода в активное состояние для основной популяции нейронов этого ансамбля. В итоге для описания динамики процессов использовалось уже шесть переменных.

$$\left\{ \begin{array}{l} \frac{\partial U_i(x, y)}{\partial t} = -\frac{U_i(x, y)}{\tau_{U_i}} + k_{U_i}^A \cdot A_{U_i}(x, y) \cdot F_{U_i} \times \\ \times \left[-T_{U_i} + k_{U_i}^S \cdot S_{U_i}(x, y) + k_{U_i}^G \cdot G_i(x, y) + \sum_{j=0}^2 k_{U_i}^j \cdot h_{ij} \otimes U_j(x, y) \right], \\ \frac{\partial G_i(x, y)}{\partial t} = -\frac{G_i(x, y)}{\tau_{G_i}} + k_{G_i}^A \cdot A_{G_i}(x, y) \cdot F_{G_i} \times \\ \times \left[-T_{G_i} + k_{G_i}^S \cdot S_{G_i}(x, y) + k_{G_i}^G \cdot G_i(x, y) + k_{G_i}^U \cdot U_i(x, y) \right], \\ F_{U_i}(a) = 0,5 + 0,5 \frac{a \cdot k_{U_i}^F}{\sqrt{1 + (a \cdot k_{U_i}^F)^2}}, \\ F_{G_i}(a) = 0,5 - 0,5 \frac{a \cdot k_{G_i}^F}{\sqrt{1 + (a \cdot k_{G_i}^F)^2}}, \end{array} \right. \quad (4)$$

где $i = 3$. Данное модельное описание содержит уравнение активирующего типа (U_i) и уравнение подавляющего типа (G_i), а также функции

активации (F_{U_i}, F_{G_i}) для каждого типа нейроноподобных активных элементов. Сочетание большого количества коэффициентов и высокая гибкость архитектуры позволяет получать большое количество различных типов связей. Схема связей между тремя нейронными ансамблями в такой модельной системе приведена на рис. 4.

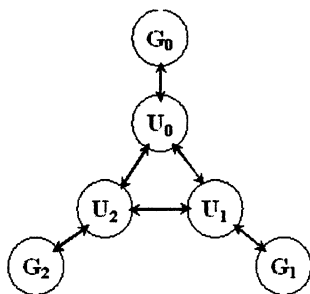


Рис. 4. Схема взаимодействия переменных в модели

Компьютерная реализация данной шестикомпонентной системы требует перехода к дискретным уравнениям и приведения процесса к итеративному виду. Дискретная система уравнений имеет вид

$$\begin{cases} U_i^{n+1} = U_i^n \left(1 - \frac{\Delta t}{\tau_{U_i}}\right) + k_{U_i}^A \cdot A_{U_i} \cdot F_{U_i} \left[-T_{U_i} + k_{U_i}^S \cdot S_{U_i} + k_{U_i}^G \cdot G_i^n + \sum_{j=0}^2 k_{U_i}^j \cdot h_{ij} \otimes U_j^n \right], \\ G_i^{n+1} = G_i^n \left(1 - \frac{\Delta t}{\tau_{G_i}}\right) + k_{G_i}^A \cdot A_{G_i} \cdot F_{G_i} \left[-T_{G_i} + k_{G_i}^S \cdot S_{G_i} + k_{G_i}^G \cdot G_i^n + k_{G_i}^U \cdot U_i^n \right]. \end{cases}$$

Дискретный вариант системы уравнений представлен в виде итеративного процесса с параметром итерации n .

Новая дискретная (интервал дискретизации времени Δt) система уравнений содержит следующие обозначения: U_i^n, G_i^n — распределение активности нейронов активирующего и подавляющего типа; $k_{U_i}^A, k_{G_i}^A$ — величина уровня концентрации нейронов; A_{U_i}, A_{G_i} — распределение концентрации нейронов; T_{U_i}, T_{G_i} — начальный порог величины активации нейронов; $k_{U_i}^S, k_{G_i}^S$ — величина уровня постоянной активности нейрона; S_{U_i}, S_{G_i} — распределение постоянной активности нейронов (или внешний сигнал); $k_{U_i}^G, k_{G_i}^G$ — коэффициенты вклада компонент G_i ; $k_{U_i}^j, k_{G_i}^U$ — коэффициенты вклада компонент U_i ; h_{ij} — профиль функции пространственной связи (такие же, как на рис. 3).

2.2. Примеры динамических режимов обработки входного сенсорного сигнала

Рассмотрим некоторые из полученных нами режимов преобразования сигналов в системе (4), моделирующей функционирование системы «таламус — кора — ретикулярные ядра таламуса» в мозге. Один из примеров нормального режима «включения — выключения» входного сенсорного сигнала приведен на рис. 5. Показаны последовательные во времени картины активности корковой переменной $U_0(x, y)$ в процессе прерывания входного изображения (сенсорного сигнала) и выделения на нем фрагментов с одним из искомым информационных признаков. Параметры модели подобраны так, что циклы кортикальной активности (включение — активность — выключение — отсутствие активности) повторяются и выделение искомого признака (контур) поступающего сигнала периодически прерывается. Повторяющиеся циклы активности коры (переменная $U_0(x, y)$) демонстрируют пример преобразований, которые осуществляет система с постоянным сенсорным сигналом на входе (на рис. 5, а приведено изображение, на котором можно увидеть и индейца, и эскимоса).

Обращаем внимание на интересную динамику выделения отдельных фрагментов изображения внутри каждого из периодов активации корковых элементов. Эти процессы явно не синхронные: в начальных вспышечных активациях выделение фрагментов идет сверху вниз (первым выделяется профиль индейца, а затем — эскимоса); в последующих вспышечных активациях последовательность выделения фрагментов становится уже слева направо (первым выделяется профиль эскимоса, а затем — профиль индейца). Фазовая рассинхронизация при выделении разных фрагментов образа на изображении в подсистеме, соответствующей корковым нейронам, может быть связана с волновыми процессами в подсистемах, описывающих активности нейронов таламуса или в nRT . Кроме того, к процессам такого же вида рассинхронизации, как это было показано в работе [27], может приводить немонотонная зависимость периода прерываний от амплитуды входного сигнала. Для уравнений типа (1)—(3), записанных для случая одного сосредоточенного элемента, было показано, что прерывание входного сигнала происходит только в определенном диапазоне амплитуд этого сигнала. Максимальный период режимов прерывания наблюдается для амплитуд внутри этого диапазона, близких к границам (для минимальной и максимальной амплитуд сигнала, при которых начинаются и срываются режимы прерывания). Минимальный период режимов прерывания соответствует середине получаемого диапазона амплитуд входного сигнала. Изучение особенностей таких динамических процессов, связанных с расфазировкой выделения различных фрагментов входного изображения, несомненно, должно быть продолжено.

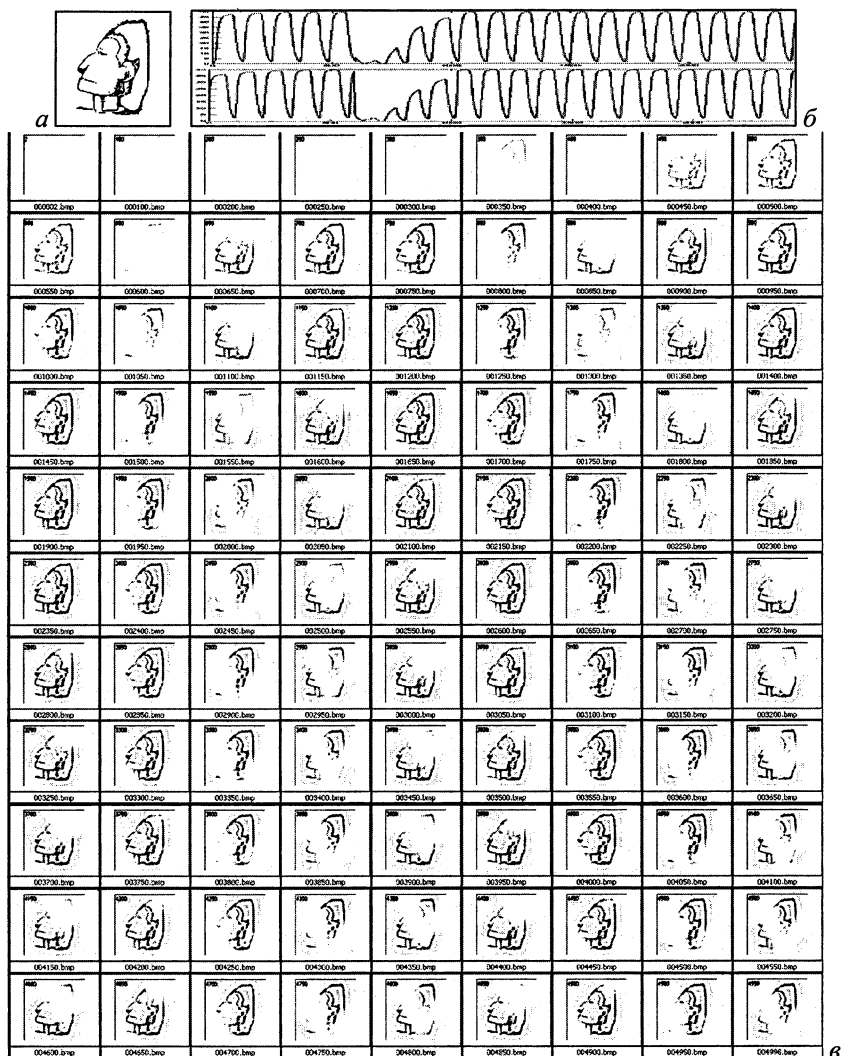


Рис. 5. Пример нормального режима «активации — подавления» входного сенсорного сигнала: *a* — исходное изображение, подаваемое в модельную систему для обработки; *б* — «пульсирующая» активация в двух выделенных сосредоточенных элементах в подсистеме, соответствующей корковым нейронам. Верхний график — ритмическая активность в точке с координатами (103;50) — волося индейца, по оси *X* — время, по оси *Y* — уровень активности клеток коры. Внизу — ритмическая активность в точке с координатами (33;50) — нос индейца — локоть эскимоса, по оси *X* — время, по оси *Y* — уровень активности клеток коры. *в* — «пульсирующая» активация подаваемого для обработки изображения в подсистеме, соответствующей корковым нейронам.

2.3. Примеры режимов самовозбуждения

Чтобы продемонстрировать возможные режимы на границе между нормой и патологией, были проведены расчеты при частично измененных параметрах системы (4). При переходе к режимам самовозбуждения в структурах пространственной активности пропадает зависимость от вида начального возбуждения (входного сигнала) и форма этих структур определяется уже внутренними параметрами системы.

Из параметров системы (4) был изменен только порог возбуждения для переменной U_0 , он уменьшился. Это соответствует увеличению возбудимости слоя нейронов, при большей возбудимости система переходит в режим вспышечной активности. При параметрах, показанных на рис. 6, система от прерывистого выделения признака (контур на портрете) переходит к вспышечной активности, в которой теряется зависимость от вида входного сигнала (рис. 7).

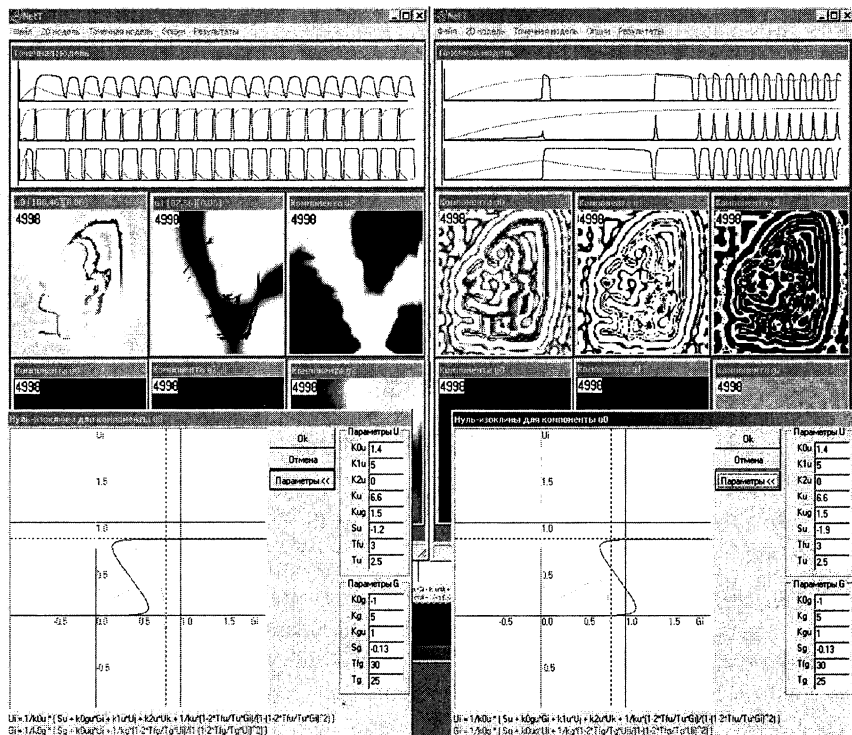


Рис. 6. Параметры исследовательской программной системы в нормальном режиме стробирования входного изображения и в режиме вспышечной активности.

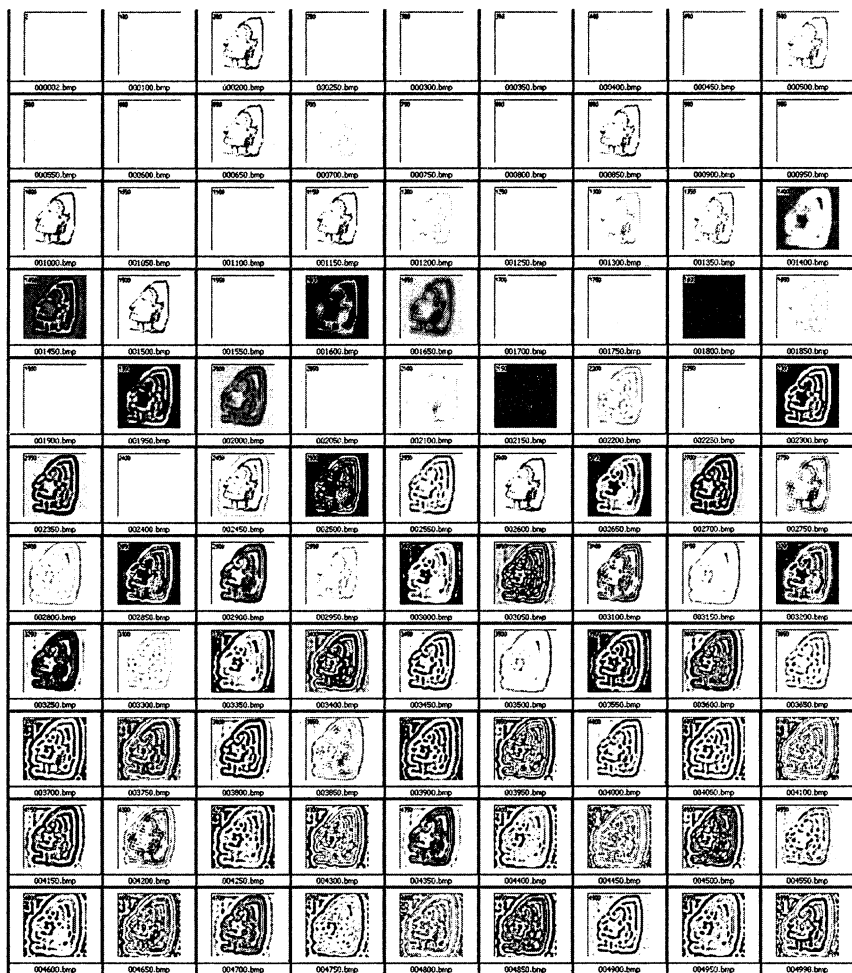


Рис. 7. Режим вспышечной активности и генерации собственных структур активности с потерей похожести на входное изображение.

Понятно, что изменением ряда других параметров системы (например, варьированием вида различных функций пространственной связи) также можно существенно модифицировать наблюдаемые при самовозбуждении режимы структурообразования. Эти режимы имеют большое разнообразие, и для их полной классификации требуются дальнейшие исследования.

Выполненные расчеты в первую очередь демонстрируют возможность изучения на одной и той же модели из распределенных нейроноподобных систем (1)—(3) или (4) разных режимов функционирования:

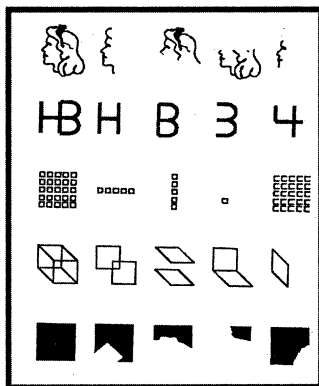
одни из этих режимов можно интерпретировать как нормальную обработку сенсорных сигналов, а другие — как патологические (например, эпилептиформная активность).

2.3. Фрагментация стабилизированного изображения

В психофизиологических экспериментах проводилась регистрация результатов восприятия стабилизированных изображений, неподвижно зафиксированных на сетчатке глаза человека [28]. Оказалось, что при предъявлении испытуемому стабилизированного изображения он в течение нескольких секунд воспринимает его как при нормальном зрении, но затем начинаются нарушения. Предъявляемое изображение может начать исчезать или появляться целиком, заменяясь черным или серым фоном. В ряде случаев изображение появляется частями (рис. 8). Интересен характер распада изображения на части (фрагментация). Фрагментация происходит не хаотично, а в соответствии с «истинной» структурой изображения, которая формировалась у человека в процессе обучения.

Например, контурный человеческий профиль всегда исчезает и восстанавливается отдельными дискретно организованными элементами. Передняя часть лица, задняя сторона головы, комплекс линий вокруг глаз или вокруг уха исчезают и появляются вновь как целые элементы в отдельности друг от друга или в различных комбинациях [28].

Рис. 8. Варианты восприятия испытуемыми людьми зрительных сигналов во времени (последовательности изменений образов в четырех правых колонках) при предъявлении им входных стабилизированных изображений (пять примеров предъявляемых изображений в левой колонке) [28].



Пульсирующая фрагментация воспринимаемых изображений, представленная на рис. 8, формируется как бы волновыми «затенениями» распознаваемого образа вдоль горизонтального, вертикального или наклонных направлений. Такие «затенения», связанные с волнами в подсистеме, моделирующей активность таламуса, видны также и в результатах численного расчета (см. рис. 5, б).

Таким образом, имеется хорошее качественное соответствие между модельным описанием и данными экспериментов.

3. Выводы

Проведенное рассмотрение пространственно-временных режимов преобразования изображений в системе из трех взаимосвязанных распределенных однородных нейроноподобных подсистем было ориентировано на выявление характерных особенностей динамических режимов преобразования сенсорных сигналов у животных.

Показано, что с помощью такого модельного описания удается воспроизводить динамические процессы, связанные с прохождением сенсорных сигналов по цепям «таламус — кора — ретикулярные ядра таламуса».

Таким образом, демонстрируется и, в определенной мере, обосновывается, что рассмотренные модельные системы (1)—(3), (4) позволяют изучать широкий набор нелинейных динамических процессов, которые аналогичны наблюдаемым и регистрируемым динамическим процессам в экспериментах с животными. В частности, воспроизведены варианты известных характерных режимов нормального преобразования сенсорных сигналов, а также некоторые варианты перехода в режимы патологического самовозбуждения. Результаты показывают хорошее качественное соответствие с известными данными экспериментов.

Рассматриваемая модель перспективна для дальнейшего изучения особенностей возможной динамики восприятия у животных. Для этого требуется более подробное изучение и классификация различных динамических режимов преобразования изображений в предложенных ранее базовых моделях [7—16].

Работа выполнялась при частичной поддержке грантов INTAS-01-0690 и АФГИР №РМО-10214-BNL №36943.

Литература

1. Hubel D., Wiesel T. Receptive fields, binocular integration and functional architecture in cat's visual cortex // *J. Physiol.* 1962. V. 160. P. 106—154.
2. Позин Н. В. Моделирование нейронных структур. М.: Наука, 1970
3. Соколов Е. Н., Шмелев Л. А. Нейробионика. М.: Наука, 1983.
4. Wilson H. R., Cowan J. D. A Mathematical theory of the functional dynamics of cortical and thalamic neuron tissue // *Kibernetik.* 1973. V. 13. P. 55—80.
5. Прибрам К. Языки мозга. М.: Прогресс, 1975.
6. Сбитнев В. И. Перенос спайков в статистических нейронных ансамблях. II. Нейрон — нелинейный источник спайков // *Биофизика.* 1976. Т. 21. С. 1072—1076.
7. Яхно В. Г. Модели нейроноподобных систем. Динамические режимы преобразования информации // *Нелинейные волны' 2002'* Сб. тр. Нижний Новгород: ИПФ РАН, 2003. С. 90—114.
8. Яхно В. Г., Нуйдель И. В., Иванов А. Е., Беллюстин Н. С., Будников Д. Н., Тельных А. А., Еремин Е. В., Коган А. Н., Костин М. А., Перминов А. О., Раджабова Ю. Х., Сорокин М. А., Тихомиров Д. А., Шилин С. Г., Чайкин А. В. Исследование динамических режимов нейроноподобных систем. Примеры приложений // *Информационные технологии и вычислительные системы.* 2004. № 1. С. 126—148.

9. Яхно В. Г., Беллюстин Н. С., Красильникова И. Г., Кузнецов С. О., Нуйдель И. В., Панфилов А. И., Перминов А. О., Шадрин А. В., Шевырев А. А. Исследовательская система принятия решений по фрагментам сложного изображения, использующая нейроноподобные алгоритмы // Изв. вузов. Радиофизика. 1994. Т. 37, № 8. С. 961—986.
10. Нуйдель И. В., Кузнецов С. О. Использование однородных нейроноподобных сред для обработки изображений // Изв. вузов. Радиофизика. 1994. Т. 37, № 8. С. 1053—1061.
11. Nuidel I. V., Yakhno V. G. Modelling of extraction of image fragments in the forms of crosses and rhombuses in the simulated receptive fields // Cellular Neural Networks and their Applications: Proc. 6th IEEE Int. Workshop. Catania, Italy, 2000. P. 45—49.
12. Yakhno V. G. Dynamics of autowave processes in Neuron-like systems and CNN technology // Cellular Neural Networks and their Applications: Proc. 6th IEEE Int. Workshop. Catania, Italy, 2000. P. 33—38.
13. Yakhno V. G. Basic models of hierarchy neuron-like systems and ways to analyse some of their complex reactions // J. Opt. Memory & Neural Network. 1995. V. 4. № 2. P. 141—155.
14. Telnykh A. A., Yakhno V. G. Neuron-like models of the second and third levels — adaptive recognition system // Proc. XII Int. Conf. Neurocybernetics. 1999. P. 164—168.
15. Яхно В. Г., Нуйдель И. В., Тельных А. А., Бондаренко Б. Н., Сборщиков И. Ф., Хилько А. И. Метод адаптивного распознавания информационных образов и система для его осуществления. Российский Патент № 2160467, 1999.
16. Яхно В. Г. Процессы самоорганизации в распределенных нейроноподобных системах. Примеры возможных применений // Нейроинформатика 2001. Лекции по нейроинформатике. М.: МИФИ, 2001. С. 103—141.
17. The WAG/Rij rat model of absence epilepsy: ten years of research / Eds. Gilles van Luitelar, Anton Coenen. Nijmegen: Institute for Cognition and Information, 1997. P. 1—433.
18. Maki Inoue. Spike-wave discharges in rats: mechanisms for generation and modulation. NICI, 1992.
19. Douglas A. Coulter thalamocortical anatomy and physiology, epilepsy: A comprehensive textbook / Eds. J. Engel, Jr., T. A. Pedley. Philadelphia: Liippincott Raven Publisher, 1997. 341 p.
20. Hecht-Nielsen R. A theory of the cerebral cortex // Proc. 6th Int. Conf. Molecular Electronics and Biocomputing. Future Electronic Devices Association of Japan, Okinawa, 28—30 November, 1995.
21. Kuznetsova G. D., Nuidel I. V., Khurlapov P. G., Yakhno V. G. Modeling of normal and pathological sensor activity image transformation versions in an animal's cortex // Optical Memory and Neural Networks: Proc. SPIE / Ed. Andrei L. Micaelian. 1998. V. 3402. P. 486—499.
22. Coenen A. M. L., van Luitelaar Egidius L. J. M., Kuznetsova G. D., Ivanov A. E., Nuidel I. V., Khurlapov P. G., Yakhno V. G. Modeling of transition regimes between normal and pathological transformation of sensor signals in Brain // Proc. Nijmegen Institute for Cognition and Information, 2004. P. 331—351.
23. Жеребцов И. П. Техника метровых волн. М.: 1955. С. 68—69.
24. Григоров И. Сверхрегенеративный приемник. Миф и реальность // Радиолобитель. 1994. № 1. С. 42—45.
26. Мастеров А. В., Рабинович М. И., Толков В. Н., Яхно В. Г. Исследование режимов взаимодействия автоволн и автоструктур в нейроноподобных средах // Коллективная динамика возбуждений и структурообразование в биологических тканях. Горький: ИПФ АН СССР, 1988. С. 89—104.; Мастеров А. В., Яхно В. Г. Анализ стационарных волн в однокомпонентной возбудимой среде с нелокальными связями // В том же сборнике. С. 198—207.
27. Стицын И. Г., Нуйдель И. В., Яхно В. Г. Моделирование таламо-кортикальных связей в сенсорных системах // Нейроинформатика-2004. Ч. 1. М.: МИФИ, 2004. С. 145—149.
28. Турчин В. Ф. Феномен Науки. М.: ЭТС, 2000; Prichard R. Stabilized image on retina // Sci. Am. 1961. V. 204, № 6. P. 72—78.

ВВЕДЕНИЕ В НЕЛИНЕЙНУЮ ФИЗИКУ ДНК

Л. В. Якушевич

Дезоксирибонуклеиновая кислота, или сокращенно ДНК, — одна из наиболее интересных и загадочных молекул биологии. Она принадлежит к классу биополимеров и обладает важнейшей биологической функцией, заключающейся в возможности сохранять и передавать генетическую информацию. Молекула ДНК — интереснейший объект и для физики. В настоящей статье я попробую показать, что молекула ДНК представляет собой сложную динамическую систему, состоящую из множества атомов, собранных в своеобразную квазидномерную структуру. Эта структура имеет вид двойной спирали, обладает множеством степеней свободы, большой внутренней подвижностью и специфическим распределением внутренних взаимодействий.

Я постараюсь показать также, что молекулу ДНК можно рассматривать как среду, в которой могут возникать и распространяться нелинейные волны. Исследование условий возникновения и распространения нелинейных волн, а также изучение свойств этих волн и их роли в биологической активности ДНК и составляет предмет нового, активно развивающегося сейчас направления — нелинейной физики ДНК.

1. Структура ДНК и внутренние взаимодействия

Молекула ДНК состоит из двух полинуклеотидных цепочек, слабо взаимодействующих между собой и свернутых в двойную спираль. Каждый отдельный нуклеотид состоит из трех атомных групп: сахара (S), основания (B) и фосфатной группы (Ph) (рис. 1).

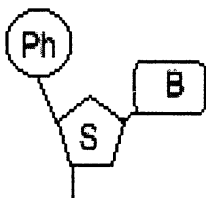


Рис. 1. Отдельный нуклеотид

Объединившись в полинуклеотидную цепочку, сахара и фосфаты образуют регулярную часть полинуклеотидной цепи, а основания — нерегулярную часть, так как в ДНК возможны четыре типа оснований: аденин (А), гуанин (G), цитозин (С) и тимин (Т). Чередуюсь между собой вдоль цепи, основания образуют последовательность, уникальную для каждого живого организма (рис. 2).

Две полинуклеотидные цепи, входящие в состав ДНК, слабо взаимодействуют между собой. Такое взаимодействие осуществляется в основном через водородные связи, образуемые между основаниями, причем между основаниями А и Т образуется две водородные связи, а между основаниями G и С — три (рис. 3).

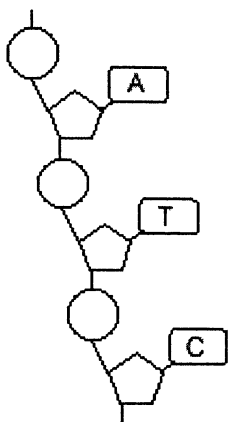


Рис. 2. Полинуклеотидная цепочка

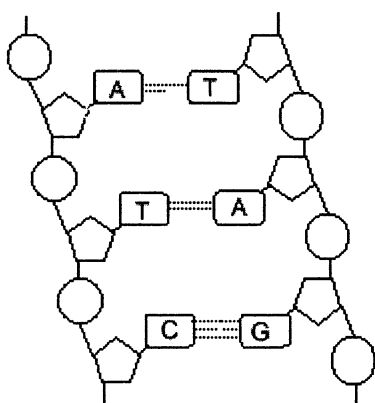
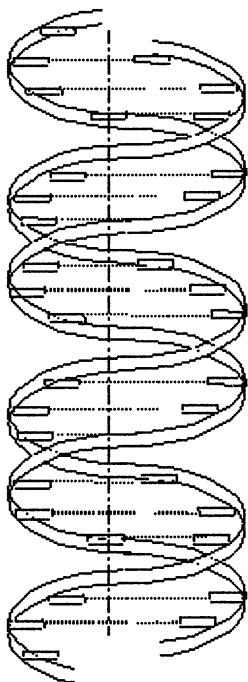


Рис. 3. Двойная полинуклеотидная цепочка. Водородные связи показаны штриховыми линиями

Чтобы получить структуру ДНК, остается только свернуть двойную полинуклеотидную цепочку в спираль (рис. 4).

Рис. 4. Двойная спираль. Сахаро-фосфатные цепочки изображены в виде лент, основания — в виде небольших поперечных стержней. Водородные связи показаны поперечными штриховыми линиями.



Стабильность двойной спирали обеспечивается продольными (стэкинг) и поперечными (водородные связи) взаимодействиями. И те, и другие взаимодействия значительно слабее ковалентных взаимодействий.

2. Внутренняя подвижность ДНК и динамические модели

Двойная спираль представляет собой довольно подвижную, динамическую систему. Наличие термостата, столкновений с «горячими» молекулами раствора, взаимодействий с белками, внешними полями и средой приводит к активации в ДНК самых разнообразных видов движений. Среди них и простые смещения атомов и атомных групп вблизи положений равновесия, и вращательные движения отдельных элементов структуры, и локальное расплетание двойной спирали, и переходы между различными конформационными формами, различающимися параметрами двойной спирали.

Сведения о разнообразных видах движений в ДНК собраны и систематизированы в работах [1—2]. А математические модели, имитирующие эти движения с различной точностью, подробно описаны в книге [3].

2.1. Модель синус-Гордона

Рассмотрим, для примера, модель синус-Гордона. Она считается первой и наиболее простой нелинейной динамической моделью ДНК. Эта модель описывает вращательные колебания оснований вокруг сахаро-фосфатной цепочки (рис. 5). Такие движения вносят существенный вклад в расплетание двойной спирали и поэтому очень важны для понимания динамической составляющей начального этапа процесса транскрипции.

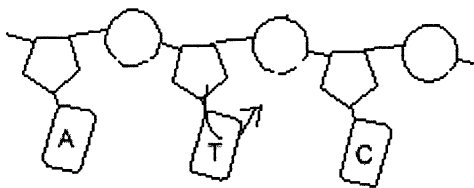


Рис. 5. Вращательные колебания оснований вокруг сахаро-фосфатной цепочки

Как заметила группа американских исследователей [4], вращательные колебания оснований вокруг сахаро-фосфатной цепочки во многом аналогичны вращательным колебаниям связанных маятников, подвешенных на горизонтальную нить. Последние можно описать с помощью уравнения синус-Гордона

$$I\partial^2\varphi/\partial t^2 = Ka^2\partial^2\varphi/\partial z^2 - mgl\sin\varphi, \quad (1)$$

где φ — угловое отклонение маятника; I , m и l — момент инерции, масса и длина маятника; K — коэффициент жесткости горизонтальной нити, на которой подвешены маятники; a — расстояние между ближайшими соседними маятниками.

В случае ДНК в роли маятников выступают основания, функция φ трактуется как угол отклонения оснований от положения равновесия, коэффициент I — как момент инерции оснований. Роль горизонтальной нити играет в ДНК сахаро-фосфатная цепочка, и следовательно, коэффициент K трактуется как жесткость этой цепочки. Наконец роль гравитационного поля в случае ДНК играет поле, наводимое второй нитью. В первом приближении его можно представить как сумму водородных взаимодействий между основаниями внутри пар. Поэтому множитель перед $\sin\varphi$ можно заменить константой, пропорциональной жесткости этих взаимодействий.

Нелинейная компонента в уравнении (1) чрезвычайно важна, поскольку в процессе расплетания двойной спирали основания значительно отклоняются от положений равновесия, и в таких случаях использование линейного (или гармонического) приближения становится некорректным.

Решения уравнения (1) хорошо изучены, и известно, что среди них имеются солитонные решения в виде кинков и антикинков, которые как раз и были интерпретированы американскими исследователями [4] как нелинейные конформационные волны, возникающие и распространяющиеся вдоль молекулы ДНК.

2.2. Усложненный вариант модели синус-Гордона

Модель синус-Гордона ограничивается описанием вращательных колебаний только в одной из полинуклеотидных цепей ДНК. Если учесть колебания оснований во второй цепи, а также взаимодействие между ними, то вместо уравнения (1) мы получим систему, состоящую из двух связанных уравнений

$$\begin{aligned} I_1(\partial^2\varphi_1/\partial t^2) &= K_1\alpha^2 l_1^2(\partial\varphi_1/\partial z^2) - k_{1,2}[l_1(l_1 + l_2)\sin\varphi_1 - l_2 l_1 \sin(\varphi_1 - \varphi_2)]; \\ I_2(d^2\varphi_2/dt^2) &= K_2\alpha^2 l_2^2(\partial\varphi_2/\partial z^2) - k_{1,2}[l_2(l_1 + l_2)\sin\varphi_2 - l_2 l_1 \sin(\varphi_2 - \varphi_1)]. \end{aligned} \quad (2)$$

В этих уравнениях φ_1 и φ_2 — угловые отклонения оснований в первой и второй цепях ДНК соответственно; I_i и l_i — момент инерции и расстояние от центра тяжести маятника i -ой цепи до сахаро-фосфатной цепочки; K_i — коэффициент жесткости i -ой сахаро-фосфатной цепочки ($i = 1, 2$); $k_{1,2}$ — коэффициент, характеризующий жесткость водородных связей между основаниями внутри пар.

Уравнения (2) были исследованы недавно в работе [6]. Оказалось, что они также обладают достаточно устойчивыми солитонными решениями.

2.3. Другие модели

Кроме вращательных движений оснований заметный вклад в процесс расплетания двойной спирали ДНК, а также в процесс денатурации молекулы вносят поперечные смещения полинуклеотидных цепей. Эти движения детально изучались в работе Пейярда и Бишопа [5]. Был построен соответствующий гамильтониан, получены динамические уравнения и найдены солитонные решения этих уравнений. Если к уравнениям (2) добавить уравнения Пейярда и Бишопа, а также дополнить все уравнения слагаемыми, учитывающими взаимодействия между вращательными движениями и поперечными смещениями оснований, то модель, имитирующая внутреннюю динамику молекулы ДНК, существенно улучшится.

Дальнейшее улучшение модели потребовало бы учета слагаемых, описывающих влияние внешней среды. И прежде всего необходимо было бы учесть воздействие термостата, ионного окружения и внешних полей. Необычайно важно было бы учесть также тот факт, что основания в ДНК неодинаковые. Для этого потребовалось бы ввести в коэффициенты уравнения (2), которые до сих пор считались константами, зависимость от переменной z . Такая зависимость отражала бы уникальную для каждого живого организма последовательность оснований, и введение ее в уравнения значительно приблизило бы модель к реальной ситуации.

Вся совокупность перечисленных выше проблем как раз и составляет математическую основу нелинейного подхода в исследованиях ДНК. Привлекательность этого направления исследований заключается в том, что у него имеется достаточно серьезный шанс объяснить многие черты и механизмы функциональной активности ДНК исходя из нелинейных динамических свойств этой молекулы.

Литература

1. *McCommon J. A., Harvey S. C.* Dynamics of Proteins and Nucleic Acids. Cambridge: Cambridge University Press, 1987.
2. *Yakushevich L. V.* Nonlinear dynamics of biopolymers: theoretical models, experimental data // *Quart. Rev. Biophys.* 1993. V. 26, № 2. P. 201—223.
3. *Englander S. W., Kallenbach N. R., Heeger A. J., Krumhansl J. A., Litwin A.* Nature of the open state in long polynucleotide double helices: possibility of soliton excitations // *Proc. Natl. Acad. Sci. USA.* 1980. V. 77, № 12. P. 7222—7226.
4. *Yakushevich L. V.* Nonlinear Physics of DNA. Weinheim: Wiley-VCH, 2004.
5. *Peyrard M., Bishop A. R.* Statistical mechanics of a nonlinear model for DNA denaturation // *Phys. Rev. Lett.* 1989. V. 62, № 23. P. 2755—2758.
6. *Yakushevich L. V., Savin A. V., Manevitch L. I.* On the internal dynamics of topological solitons in DNA // *Phys. Rev. E.* 2002. V. 66, 016614-29.

НЕКОТОРЫЕ ПРОБЛЕМЫ РЕКОНСТРУКЦИИ МОДЕЛЬНЫХ УРАВНЕНИЙ ПО ВРЕМЕННЫМ РЯДАМ И ПУТИ ИХ РЕШЕНИЯ

*Б. П. Безручко, М. Б. Бодров, Т. В. Диканев,
А. С. Караваев, В. И. Пономаренко, М. Д. Прохоров,
Е. П. Селезнев, И. В. Сысоев, Д. А. Смирнов*

Формирование моделей на основе результатов наблюдений и исследование их свойств — вот, по существу, основное содержание науки Л. Льюис [1], с. 15.

В исследовательской практике типичны ситуации, когда процессы, протекающие в объекте, обусловлены нечетко очерченной совокупностью явлений различной природы или общие законы (аналогичные законам Ньютона в механике) для рассматриваемой области не установлены. В этом случае основным источником информации о поведении объекта часто становятся данные измерений скалярной или векторной величины η (наблюдаемой), сделанные в отдельные моменты времени t_i — временные ряды $\{\eta_i\}$. Здесь $\eta_i = \eta(t_i)$, $t_i = i\Delta t$, Δt — интервал выборки, $i = 1, 2, 3, \dots, N$, N — длина ряда¹. Используя экспериментальные ряды, создают *эмпирические* модели поведения объекта во времени². Модели эти менее универсальны, чем конструкции, создаваемые на основе «первых принципов», и в некотором смысле подобны уникальной одежде, изготовленной методом завертывания клиента в куски ткани и сшивания по краям. Но развитие вычислительной техники, формирование концепции динамического хаоса и труд поколений исследователей привели к тому, что такие модели прошли путь от элементарных функций $x = f(t)$, аппроксимирующих множества экспериментальных точек на плоскости (x, t) [2], к многомерным системам нелинейных разностных и дифференциальных уравнений. Это направление моделирования актуально, например, для физики, метеорологии, сейсмографии, финансов, медицины,

¹ Так выглядит сигнал на выходе любого аналого-цифрового преобразователя, но задача принципиально не меняется, если говорить о дискретных отсчетах в различных точках пространства, а не в различные моменты времени.

² Это направление деятельности в математической статистике и теории автоматического управления получило название *идентификация систем* [1]. В нелинейной динамике оно именуется *реконструкцией динамических систем*, хотя, на наш взгляд, использование слова реконструкция (в значении «восстановление») оправдано лишь при восстановлении уравнений по их решениям, в случае же моделирования реальных систем больше подошел бы термин *конструирование*.

физиологии и используется для решения сложных фундаментальных и прикладных задач [3, 4].

Речь пойдет о глобальной³ реконструкции по зашумленным хаотическим рядам, которые типичны для природных и рукотворных нелинейных систем. Ранее задача описания столь сложных движений, как правило, решалась с помощью вероятностных моделей [5]. В настоящее время, когда известно, что хаотическое поведение свойственно и простым (маломерным) нелинейным динамическим системам, продуктом реконструкции стали и разностные и дифференциальные уравнения (обыкновенные, с запаздыванием, в частных производных). Список всех публикаций занял бы много страниц, поэтому мы упомянем лишь некоторые пионерские работы [6—9] и обзорные материалы [10—16].

Целью статьи является представление некоторых проблем, возникающих при эмпирическом моделировании, и демонстрация возможностей их преодоления. Речь идет о моделях в виде разностных или дифференциальных уравнений вида

$$\mathbf{x}(t_i) = \mathbf{F}(\mathbf{x}(t_{i-1}), \mathbf{c}). \quad (1)$$

Здесь $\mathbf{x} \in R^D$ — вектор состояния; \mathbf{F} — вектор-функция; $\mathbf{c} \in R^P$ — вектор параметров; t_i — дискретное время;

$$d\mathbf{x}(t)/dt = \mathbf{F}(\mathbf{x}(t), \mathbf{c}), \quad (2)$$

где время t непрерывно.

Материал излагается по принципу «от простого к сложному» — по мере увеличения неопределенности в постановке задачи: от случая, когда о структуре модели все известно и остается только вычислить коэффициенты в уравнениях (параметры модели), до ситуации, когда об объекте, а следовательно, и о виде его адекватной модели практически ничего не известно априори. При этом мы будем опираться на типовую «технология» реконструкции, а для иллюстрации по возможности использовать примеры из собственного опыта.

1. Типовая схема процесса моделирования по рядам данных

Несмотря на безграничное число ситуаций, объектов и целей, вносящих в процесс свое специфическое, можно выделить типичную схему моделирования (рис. 1). Работа над моделью начинается рассмотрением известного об объекте с позиций поставленных целей, с анализа имею-

³ Термин «глобальная» означает, что модельные уравнения описывают поведение объекта во всем фазовом пространстве (глобально), в отличие от «локальных» моделей, когда используется отдельная конструкция со своим набором параметров для малой окрестности каждой точки в фазовом пространстве.

щихся или получения доступных экспериментальных данных — этап 1, а заканчивается использованием полученной модели в приложении к конкретной задаче. Но этот процесс обычно сопровождается неоднократными повторениями, возвратами в исходную и промежуточные точки схемы, являясь вариантом метода последовательных приближений.

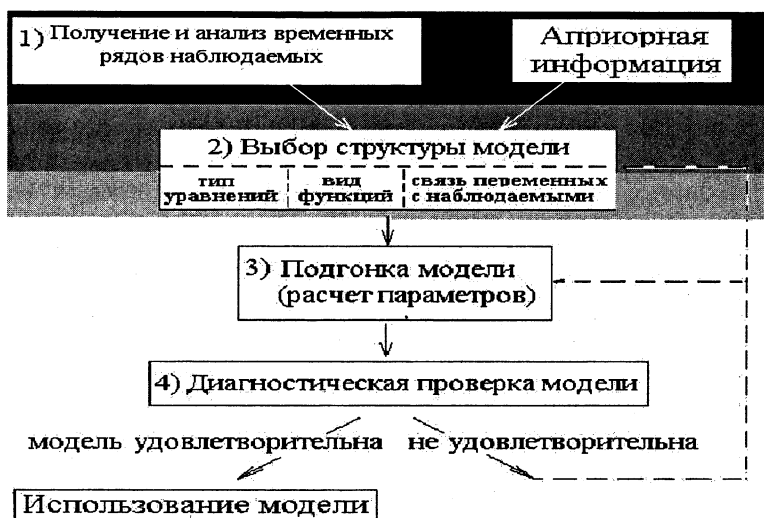


Рис. 1. Типовая схема моделирования по временному ряду

На этапе 2 формируется структура модели: выбираются тип и число уравнений, задается вид входящих в них функций F и динамических переменных x . В качестве переменных могут использоваться и непосредственно наблюдаемые величины, но в общем случае связь переменных с наблюдаемыми приходится специально выбирать: $\eta(t_n) = h(x(t_n)) + \xi(t_n)$. Функцию h , связывающую наблюдаемые и переменные модели (состояние), называют *измерительной функцией*, а ξ — измерительным шумом. На этапе 3 подбираются оптимальные значения параметров уравнений. При этом используется некий критерий качества, определяемый целью моделирования, например минимальное отклонение решений полученных уравнений от данных эксперимента или обеспечение максимальной длительности прогноза. Наконец, модель проверяется с использованием прибереженной для этой цели тестовой части ряда — этап 4. В зависимости от результата проверки полученная конструкция берется в дело или возвращается на доработку на любой из этажей технологической схемы. Фон, на котором изображена схема рис. 1, меняется от черного («тьма незнания») до белого, отражая степень неопределенности, с которой приходится сталкиваться при моделировании. Наименее благо-

приятна для моделирования ситуация, получившая название «черного ящика», когда информация о структуре возможной адекватной модели объекта отсутствует и начинать приходится с самого верха описанной схемы. Чем больше известно о том, как должна выглядеть модель, тем вероятнее успех — «ящик» становится «серым» и даже «прозрачным».

От решения проблем, встречающихся на нижних ярусах технологической схемы, уклониться невозможно, с ними неизбежно сталкивается исследователь, преодолевший специфические трудности менее определенной задачи. Поэтому есть смысл начать обещанное в заголовке рассмотрение возникающих при реконструкции проблем с наиболее простой для моделирования ситуации, когда о модели известно все, кроме конкретных значений ее параметров, на рисунке ей соответствует белый фон.

2. Структура уравнений адекватной модели полностью известна из «первых принципов» или других соображений — «прозрачный» ящик

Задача сводится к оценке P неизвестных параметров $\mathbf{c} = (c_1, \dots, c_p)$ по наблюдаемым данным. Такая постановка весьма востребована и встречается в различных приложениях. В отсутствие шумов в рядах наблюдаемых и при возможности сформировать все ряды переменных задача выливается в решение системы алгебраических уравнений. Но наличие шума качественно меняет ситуацию — вместо точного решения приходится ограничиться статистическими оценками.

Не вдаваясь в подробности, мы только перечислим известные подходы к оцениванию в порядке повышения статистической эффективности получаемых с их помощью результатов: *простого усреднения, статистических моментов, наименьших квадратов, максимального правдоподобия* (обсуждение методов см., например, в [17]) и, наконец, *метод Байеса* — целая область статистики, которая занимается получением распределения вероятностей для параметров по выборке [18]. Причем чаще всего используется метод наименьших квадратов, согласно которому минимизируется сумма квадратов отклонений наблюдаемых величин или величин, полученных по наблюдаемым $\boldsymbol{\eta}$, от соответствующих модельных величин \mathbf{y} :

$$S(\mathbf{c}, \mathbf{x}) = \sum_{i=1}^N (\boldsymbol{\eta}(t_i) - \mathbf{y}(t_i, \mathbf{x}, \mathbf{c}))^2 \rightarrow \min. \quad (3)$$

И далее сосредоточим внимание на проблеме «скрытых переменных», которая возникает, если отсутствует возможность сформировать по экспериментальным наблюдаемым ряды всех D динамических переменных D -мерной модели, необходимые для последующего подбора и оптими-

зации P параметров (коэффициентов модельных функций). Она возникает, например, из-за невозможности измерить нужные величины, сложности измерительной функции, чрезмерной зашумленности рядов наблюдаемых и приводит к нехватке данных для подбора параметров. Дело осложняется тем, что при заданной структуре модели, включая выбор переменных (т. е. все до этапа 3 на схеме рис. 1 строго определено), для формирования рядов переменных по наблюдаемым не удастся воспользоваться стандартными методами реконструкции вектора состояния по скалярному ряду. Полученные, например, последовательным дифференцированием или методом задержек переменные в общем случае не будут соответствовать *заложенным в модель переменным* ни по своему физическому смыслу, ни количественно. Поэтому для реконструкции значений параметров модели при наличии d скрытых переменных требуются специальные подходы. Их основная идея — расширить круг неизвестных величин, включив в него не только параметры, но и значения переменных состояния модели⁴, перенеся центр тяжести проблемы с реконструкции рядов переменных на многопараметрическую оптимизацию (метод «начального состояния» [19]). На этом этапе решаются модельные уравнения со стартовыми значениями набора ($P + d$) искомым величин. Последние затем подгоняются так, чтобы временные ряды известных ($D-d$) переменных модели (или ряды других величин, связанных с переменными модели и соответствующих наблюдаемым) были как можно ближе к наблюдаемым рядам.

Реализация описанной идеи обещает получение оценок не только искомым параметров, но и рядов скрытых переменных и сулит ряд соблазнительных приложений. Это — проверка адекватности заложенных в модель представлений, «измерение» величин, недоступных прибору экспериментатора, восстановление утерянных или искаженных участков временной зависимости измеряемой величины, входящей в априори известную модель, уход от необходимости численного дифференцирования зашумленной наблюдаемой⁵ и т. п.

Однако опыт реализации описанной процедуры при наличии характерной для хаотических систем неустойчивости к малым возмущениям показал невозможность использовать для подгонки всю длину экспери-

⁴ На практике приходится вводить в круг искомым параметров начальные значения всех компонент вектора состояния. Хотя, казалось бы, достаточно искать начальные условия только для скрытых переменных. Это диктуется технологией подгонки параметров при сшивании отдельных фрагментов (см. далее) и, кроме того, позволяет «исправить» начальные условия, если наблюдаемые существенно зашумлены.

⁵ Речь идет о случаях, когда динамические переменные в модели являются последовательными производными от наблюдаемых величин. При больших уровнях шума и порядках дифференцирования бывает выгоднее отказаться от плохих экспериментальных данных и считать эту переменную скрытой.

ментальных рядов в методе начального состояния. Использование же коротких рядов приводит к большой дисперсии оценок. Поэтому было предложено [19] рассматривать задачу расчета реализации модели в формуле (3) не как задачу Коши, а как краевую задачу, — разделить ряд длиной N точек на L сегментов по n точек каждый, $N = Ln$, и рассматривать начальный вектор \mathbf{x}_i на каждом сегменте ($i = 1, \dots, L$) тоже как неизвестную искомую величину. Но это не свободные параметры задачи, на них накладывается условие непрерывности реализации модели

$$\mathbf{x}_{i+1} = \mathbf{x}(t_{i+1} - t_i, \mathbf{x}_i, \mathbf{c}), \quad i = 1, \dots, L-1. \quad (4)$$

При таком подходе (который называют методом «множественной стрельбы») функция S (см. сноску 4) теперь зависит от $LD+p$ переменных, независимыми из которых являются $d+P$. Задача поиска оценок становится задачей условной минимизации S при условиях (4), что технически более сложно, чем безусловная минимизация. Она решается чаще всего с помощью обобщенного метода Гаусса — Ньютона. Результат коренным образом зависит от числа искомых параметров P и числа начальных условий искомых состояний DL , а также «удачности» стартовых догадок для них. Для хаотических временных рядов функция $S(\mathbf{c}, \mathbf{x})$ имеет множество минимумов, из которых только один — глобальный — с точностью до шумов и погрешностей расчёта соответствует истинным значениям параметров⁶. Попадание в локальный минимум приводит к большим ошибкам описания ряда и даже качественно отличному от наблюдаемого типу поведения полученной модели.

Опыт моделирования показывает, что в некоторых случаях результативным оказывается ослабление требований подгонки. Речь идет о выделении в длинном ряде v сегментов длиной n , на каждом из которых начальные состояния считаются независимыми (свободные параметры задачи), а параметры \mathbf{c} удерживаются *одинаковыми* на всех сегментах. Внутри каждого сегмента реализуется метод множественной стрельбы (деление сегмента на L «подсегментов» и т. д.). Такой метод назовем гибридным. В итоге имеем задачу условной минимизации функции S , но независимых переменных имеется уже $vD + p$. Но даже по длинному ряду при надлежащем подборе v, L, n находится глобальный минимум в тех случаях, где все предыдущие методики находили локальный, либо расходились.

Проиллюстрируем сказанное, используя в качестве критерия качества алгоритма реконструкции размеры области параметров, задание стартовых догадок в которой обеспечивает сходимость процесса оптимизации в глобальный минимум. При этом воспользуемся нормирован-

⁶ Если только такие значения существуют, т. е. если имеющаяся модель действительно способна хорошо описывать наблюдаемую динамику.

ными параметрами $b_i = (\gamma_i - c_i^0) / c_i^0$, характеризующими отклонение стартовых догадок γ_i для параметров от их истинных значений c_i^0 . Точка $\mathbf{b} = 0$ в пространстве этих параметров соответствует истинным значениям (или глобальному минимуму, если пренебречь шумами и погрешностями расчётов). С центром в этой точке задавалась область стартовых догадок в виде шара радиуса r . За меру μ принимался относительный объём концентрической шарообразной области, из которой обеспечивалась сходимость к глобальному минимуму. Значение μ можно понимать как вероятность попадания в глобальный минимум, если стартовые догадки для параметров заданы с ошибкой, не большей r . Другим критерием эффективности алгоритма может служить максимальное значение r , при котором имеет место стопроцентная сходимость ($\mu = 1$). Обозначим его индексом r_{100} . В качестве объекта использовалась система Лоренца:

$$dx/dt = c_1(y - x), \quad dy/dt = -y + x(c_3 - z), \quad dz/dt = -c_2z + xy. \quad (5)$$

Стартовые догадки для скрытых переменных задавались равными скалярной наблюдаемой. Использовались временные ряды, записанные при значениях параметров $\mathbf{c}^0 = \left(10, \frac{8}{3}, 46\right)$ и начальных условиях $\mathbf{x}^0 = (-7, 60, -12, 37, 38, 66)$, как в [20]. На рис. 2, *а*, *б* приведено сечение пространства параметров, где белым цветом обозначены точки, откуда достигается сходимость в глобальный минимум целевой функции S , оттенками серого — точки, откуда методы сходятся в какой-либо локальный минимум, чёрным — расходятся. Видно, что лучшие шансы получить хорошую модель имеются при использовании модифицированного метода: для метода множественной стрельбы $r_{100} = 1,062$ (рис. 2, *а*), для гибридного — $r_{100} = 1,304$ (рис. 2, *б*). Однако результаты моделирования зависят не только от стартовых догадок для параметров и скрытых переменных, но и от N, L, v , истинных значений параметров, от того, какая переменная является наблюдаемой, от уровня шума и т. д. Поэтому вывод о превосходстве одного метода над другим можно делать только на основе анализа зависимости меры r_{100} от всех перечисленных факторов. На рис. 2, *в* приведена зависимость $r_{100}(N)$ при оптимальном подборе количества фрагментов L . Видно, что как в среднем, так и по максимальным показателям (наибольший из всех r_{100} и максимальная длина ряда, при которой достигается сходимость с ненулевым радиусом) гибридный метод эффективнее.

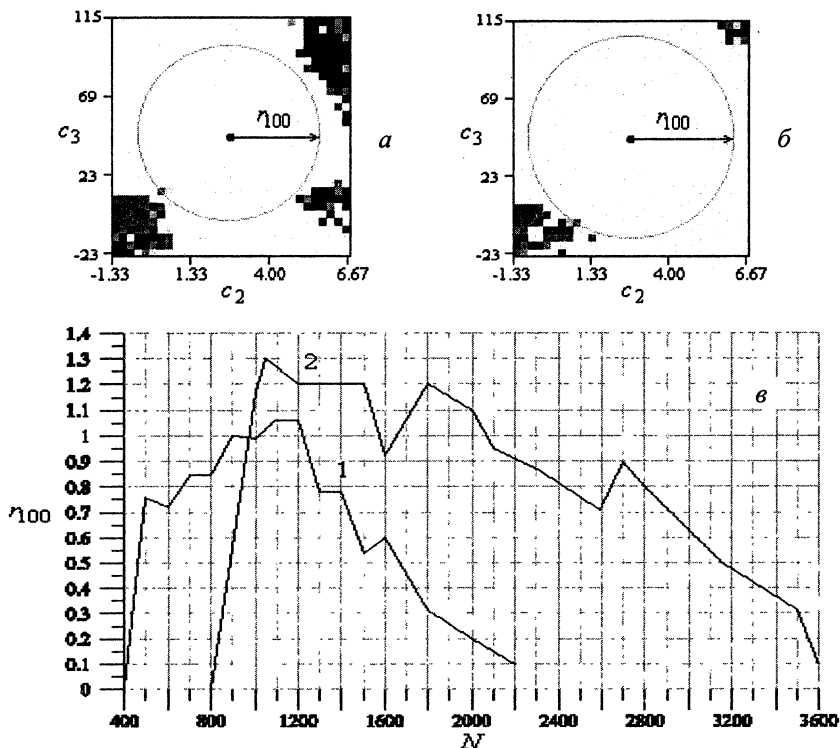


Рис. 2. Сечение пространства параметров системы Лоренца плоскостью $c_1 = 10$, в центре истинные значения параметров: *a* — метод множественной стрельбы при $L = 35$, $n = 30$; *б* — гибридный при $L = 15$, $n = 30$, $\nu = 2$; *в* — зависимость r_{100} от N для двух методов.

Мы также исследовали зависимость меры r_{100} от L , n , выбора наблюдаемой из трех имеющихся переменных, влияния шумов, показали, что $r_{100}(\nu)$ в общем случае не убывает с ростом ν . Аналогичное исследование было проведено для системы Рёсслера. Полученные результаты также свидетельствуют о преимуществе гибридного метода над методом множественной стрельбы.

3. Априорная информация позволяет частично определиться со структурой модели («серый ящик»)

Можно предположить множество вариантов не полностью определенной структуры модели. Часто неизвестны, например, размерность системы и вид аппроксимирующих функций. В таких случаях прибегают

к использованию стандартных структур (см. пример в п. 3.1) и универсальных функциональных представлений, в частности степенных полиномов. Это влечет за собой наличие слишком большого числа неизвестных параметров, оценка которых представленными в разделе 2 методами (особенно при наличии скрытых переменных) практически невозможна. Для решения задачи в такой ситуации приходится предпринимать не всегда успешные усилия по снижению размерности модели, сокращению числа параметров, подбору функционального базиса и т. п.). Другим путем к успеху может оказаться отказ от универсальности методики моделирования и специализация подхода применительно к классам объектов, о специфических чертах которых имеется априорная информация. Направление работы может подсказать анализ рядов наблюдаемых. Рассмотрим сказанное на двух примерах реконструкции обыкновенных дифференциальных уравнений (ОДУ).

3.1. Стандартная структура модели

Одна из теорем Такенса свидетельствует о том, что для почти любой динамической системы (с достаточно гладкими функциями в правых частях) ее движение может быть описано в терминах произвольной (типичной) наблюдаемой величины и ее последовательных производных. Поэтому есть основания (при предположении о «детерминированности, конечности и гладкости» исследуемого процесса)⁷ искать модель в виде

$$d^D x/dt^D = F(x, dx/dt, \dots, d^{D-1}x/dt^{D-1}), \quad (6)$$

где $\eta(t) = x(t) + \xi(t)$. Но размерность модели должна быть достаточно большой: $D \geq M + 1$, где M — размерность наблюдаемого движения (фрактальная размерность аттрактора, если движение происходит на аттракторе).

При глобальном моделировании искомая функция F представляется с помощью единой формулы, описывающей поведение модели во всем фазовом пространстве. Ее чаще всего ищут в виде суммы известных *базисных* функций с неизвестными коэффициентами:

$$F(\mathbf{x}) = \sum_{j=1}^P c_j f_j(\mathbf{x}). \quad (7)$$

Широко распространено использование стандартного полиномиального базиса — функций $1, x_1, \dots, x_D, x_1^2, x_1 x_2, \dots, x_D^2, \dots$, обоснованное тео-

⁷ Жаргон — речь идет о модели процесса или модели объекта, в котором он наблюдается. Кроме того, заметим, при моделировании *реальных* объектов, реверансы в сторону теорем Такенса скорее традиционны и наукообразны, чем обязательны.

ремой Вейерштрасса. Значения коэффициентов оцениваются, как правило, путем минимизации

$$\varepsilon^2 = \frac{1}{N} \sum_{i=1}^N \left(d^D \eta(t_i) / dt^D - F \left(\eta(t_i), d\eta(t_i) / dt, \dots, d^{D-1} \eta(t_i) / dt^{D-1} \right) \right)^2.$$

Производные наблюдаемой ищутся численно, что вызывает при наличии шумов большие проблемы. Иногда их можно решить с помощью фильтрации, например, фильтра Савицки — Голзя, но при достаточно больших шумах трудности непреодолимы.

Однако все известные нам успехи моделирования реальных объектов в представленном универсальном виде носят единичный характер. Как правило, структура (6) с алгебраическим многочленом в правой части приводит к очень громоздким уравнениям с расходящимися решениями и особенно неэффективна в случае функций многих переменных.

3.2. Специализация применительно к системам, находящимся под внешним воздействием [21]

При наличии априорной информации или диагностических предположений о наличии внешнего воздействия на систему успеха можно добиться, используя при моделировании *неавтономные* ОДУ. В этом случае в уравнения модели включаются описывающие воздействие функции, *явно* зависящие от времени⁸. Например, если воздействие гармоническое и аддитивное, целесообразна модель в виде

$$d^D x / dt^D = F(x, dx/dt, \dots, d^{D-1} x / dt^{D-1}) + a \cos(\omega t) + b \sin(\omega t), \quad (8)$$

размерность D которой на единицу меньше, чем у стандартной формы. Предпосылкой к выбору структуры (8) может быть дискретный пик в спектре мощности. По нему можно грубо оценить частоту воздействия ω , уточнив ее значение затем перебором в окрестности. При этом удается избавиться от необходимости для определения параметров решать нелинейные уравнения. Значения a , b и коэффициенты F находятся решением линейной системы уравнений. Оптимальная размерность модели и порядок многочлена определяются перебором значений в соответствии с выбранным критерием качества модели.

В случае произвольного регулярного воздействия (сложного периодического и квазипериодического) удобно представление

$$F(\mathbf{x}, t) = f(\mathbf{x}) + g(t), \quad (9)$$

⁸ Если повезло и воздействие задано временным рядом, оно может непосредственно (без промежуточных преобразований) использоваться в реконструкции при любой его сложности.

где f — алгебраический многочлен, а функция $g(t)$ описывает воздействие и также содержит неизвестные параметры. Для этого можно использовать тригонометрический многочлен:

$$g(t) = a_0 + \sum_{j=1}^{k_1} a_{1,j} \cos\left(\frac{2\pi jt}{T_1} + \varphi_{1,j}\right) + \dots + \sum_{j=1}^{k_m} a_{m,j} \cos\left(\frac{2\pi jt}{T_m} + \varphi_{m,j}\right). \quad (10)$$

Эффективность подхода проверялась нами на численных примерах реконструкции уравнений по зашумленной хаотической реализации при различных видах воздействия: периодическом импульсном, периодическом с субгармониками, квазипериодическом (см., например, рис. 3). Важно отметить, что адекватные модели с тригонометрическими многочленами могут быть получены и *при очень большом числе (сотни) используемых гармоник*. Примечательно, что при стандартной аппроксимации степенными полиномами увеличение членов ряда чревато неустойчивостью моделей.

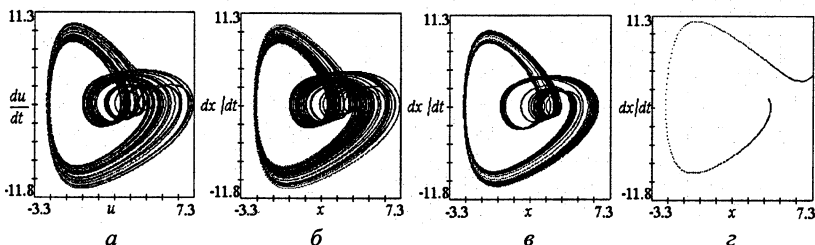


Рис. 3. Реконструкции нелинейного осциллятора Тода: $d^2u/dt^2 = -0,45 du/dt + (5 + 4 \cos t)(e^{-u} - 1) + 7 \sin t$: аттрактор объекта (а); фазовые траектории модели (б), (в) ($D = 2$, $K = 9$) (б); модели (8) ($D = 2$, $K = 10$) (в) и стандартной модели ($D = 4$, $K = 6$) (г). K — порядок полинома.

3.3. Специализация применительно к системам с запаздыванием [22]

Методика восстановления систем с запаздыванием опирается на особенности расположения экстремумов во временных реализациях наблюдаемых колебаний (рис. 5, б), которые выявляются при анализе временного ряда наблюдаемой и используются как диагностический признак принадлежности объекта к системам этого класса. Статистический анализ временных интервалов, разделяющих экстремумы во временных реализациях различных модельных и реальных систем с запаздыванием, позволяет установить, что зависимость числа N пар экстремумов временной реализации, удаленных друг от друга на время τ , от величины τ имеет четкий минимум при времени, соответствующем времени запаз-

дывания системы (рис. 5, б). Закономерность этой особенности можно пояснить с помощью уравнения с запаздыванием вида

$$\varepsilon_1 \dot{x}(t) = -x(t) + f(x(t - \tau_1)), \quad (11)$$

где τ_1 — время запаздывания, а ε_1 — параметр, характеризующий инерционность системы. При наличии инерционности ($\varepsilon_1 > 0$), что соответствует реальным ситуациям, экстремумы во временной реализации $x(t)$ близки к квадратичным, а следовательно, в экстремальных точках $\dot{x}(t) = 0$, $\ddot{x}(t) \neq 0$. Продифференцировав (11) по t , получим

$$\varepsilon_1 \ddot{x}(t) = -\dot{x}(t) + \frac{df(x(t - \tau_1))}{dx(t - \tau_1)} \dot{x}(t - \tau_1). \quad (12)$$

Если при $\dot{x}(t) = 0$ в типичном случае $\ddot{x}(t) \neq 0$, то, как видно из уравнения (2), при $\varepsilon_1 \neq 0$ должно выполняться условие $\dot{x}(t - \tau_1) \neq 0$. Это означает, что на удалении τ_1 от квадратичного экстремума не должно быть другого экстремума, а следовательно, значение $N(\tau_1) \rightarrow 0$.

Аналогичными свойствами обладают многие системы с запаздыванием при наличии инерционности (ограниченности полосы пропускания), делающей экстремумы во временной реализации колебаний «округлыми». В чем можно убедиться, продифференцировав уравнение с запаздыванием общего вида, например, $\dot{x}(t) = F(x(t), x(t - \tau_1))$, и повторив только что проведенные рассуждения. Далее с помощью найденного значения τ_1 по хаотической временной реализации находятся параметры инерционности ε_1 и нелинейной функции f . Для этого проводится проектирование бесконечномерного фазового пространства системы с запаздыванием в специальном образом выбранные подпространства малой размерности. Например, если записать уравнение (11) в виде $\varepsilon_1 \dot{x}(t) + x(t) = f(x(t - \tau_1))$ и построить на плоскости множество точек с координатами $(x(t - \tau_1), \varepsilon_1 \dot{x}(t) + x(t))$, то оно воспроизведет функцию f . Поскольку заранее величина ε_1 не известна, зависимости $\varepsilon_1 \dot{x}(t) + x(t)$ от $x(t - \tau_1)$ строят для различных значений ε , добываясь однозначной зависимости на плоскости $(x(t - \tau_1), \varepsilon \dot{x}(t) + x(t))$, которая возможна лишь при $\varepsilon = \varepsilon_1$. В качестве количественного критерия однозначности при таком поиске ε_1 используется минимальная длина линии $L(\varepsilon)$, соединяющей точки на плоскости $(x(t - \tau_1), \varepsilon \dot{x}(t) + x(t))$, упорядоченные по величине координаты $x(t - \tau_1)$. Минимум зависимости $L(\varepsilon)$ наблюдается при $\varepsilon = \varepsilon_1$, а множество точек на плоскости $(x(t - \tau_1), \varepsilon_1 \dot{x}(t) + x(t))$ при этом воспроизводит нелинейную функцию, которую при необходимости можно аппроксимировать.

Предложенный статистический подход к определению времени задержки может быть распространен на системы с запаздыванием более высокого порядка, чем (11), а также на системы с несколькими различными временами задержки. Подход экономичен по отношению к машинному времени и использует для восстановления нелинейной функции все точки временного ряда. Последнее позволяет по коротким временным рядам восстанавливать нелинейную функцию даже в случаях слаборазвитого хаоса. Он не критичен к присутствию шума в наблюдаемой.

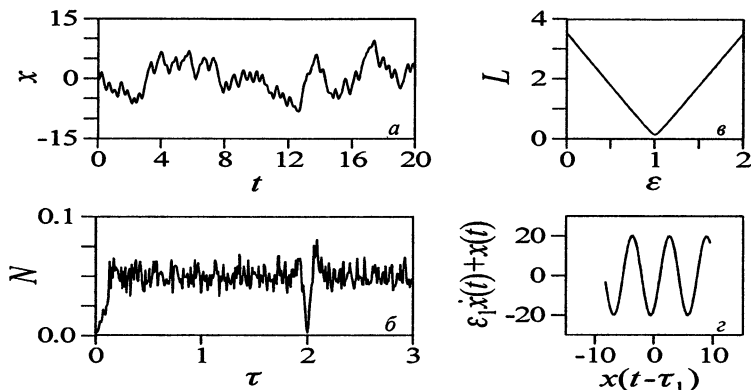


Рис. 4. Временная реализация уравнения Икеды (б) (а). Число пар экстремумов $N(\tau)$ нормировано на общее число экстремумов во временном ряду, $N_{\min}(\tau) = N(2,00)$ (б). — Длина L линии в зависимости от ε , $L(\varepsilon)$ нормирована на число точек, $L_{\min}(\varepsilon) = L(1,00)$ (в). Восстановленная нелинейная функция (д). Численные эксперименты с добавлением шума показывают возможность реконструкции при уровнях шума до 20 % от уровня сигнала.

На рис. 4 в качестве примера приведены результаты реконструкции по хаотической временной реализации (рис. 4, а) уравнения Икеды, описывающего динамику пассивного оптического резонатора:

$$\dot{x}(t) = -x(t) + \mu \sin(x(t - \tau_1) - x_0). \quad (13)$$

4. Дополнительные приемы реконструкции в ситуации «черного ящика»

Перейдем к обсуждению ситуации, когда о структуре уравнений заранее ничего не известно. Кроме гадания, остается надеяться на универсальные технологии, типа представленной в п. 3.1, и поиск возможностей хоть как-то сузить множество вариантов возможной структуры модели, опираясь на анализ экспериментальных рядов. Рассмотрим несколько примеров действий такого рода.

4.1. Оптимизация выбора динамических переменных [23]

Координаты векторов x могут быть получены, например, методом последовательных производных, временных задержек и т. д. Ряд «левых частей» модельных уравнений $\{y(t_i)\}$ обычно получают из ряда переменных $\{x(t_i)\}$: для ОДУ — численным дифференцированием ряда $\{x(t_i)\}$, а для отображений $x(t_{i+1}) = F(x(t_i))$ — сдвигом ряда $\{x(t_i)\}$ по времени. Затем задают вид функций F и находят значения их коэффициентов. Неудачный выбор динамических переменных может затруднить аппроксимацию зависимости $y(x)$ гладкой функцией или сделать ее вообще невозможной, если связь y с x окажется неоднозначной. Попыткой уберечься от этой неприятности является следующая процедура выбора более перспективного для глобального моделирования вида переменных (рис. 5). Область V , внутри которой содержится множество векторов $\{x(t_i)\}$, разбивается на одинаковые «гиперкубические» ячейки со стороной δ . Из них выбираются все ячейки, содержащие больше одного вектора каждая: s_1, s_2, \dots, s_M . Разность между максимальным и минимальным значениями y в пределах ячейки s_k назовем *локальным разбросом*: ϵ_k . Величина максимального локального разброса $\epsilon_{\max} = \max_{1 \leq k \leq M} \epsilon_k$ и график $\epsilon_{\max}(\delta)$ используются в качестве *основных* характеристик исследуемой зависимости $y(x)$. Пригодность рассматриваемых величин x и y для глобального моделирования оценивается из следующих соображений. Для построения глобальной модели переменные нужно выбирать так, чтобы график $\epsilon_{\max}(\delta)$ стремился к началу координат плавно, без изломов, причем с небольшим наклоном (рис. 5, б, нижняя кривая).

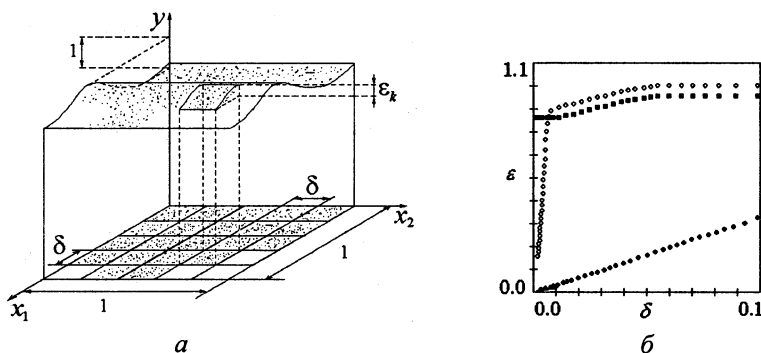


Рис. 5. Иллюстрация методики проверки зависимости $y(x)$ на однозначность и непрерывность тестирования в случае $D = 2$ (а) и графики $\epsilon_{\max}(\delta)$ для различных вариантов выбора переменных (б).

4.2. Оптимизация структуры модельных уравнений (выделение лишних слагаемых) [24]

Упомянутые в п. 3.1 недостатки стандартных структур модели обычно возникают в связи с очень большим числом членов в функциональных рядах, часть которых лишние. Их появление связано с неидеальностью рядов наблюдаемых, например, из-за шумов. Удаление лишних коэффициентов улучшает качество модели, оцениваемое критерием, диктуемым целью моделирования. Но как их выявить, скажем, из десятков членов аппроксимирующего степенного полинома? Одним из эффективных принципов идентификации лишних слагаемых является наблюдение за их поведением при реконструкции модели по различным участкам временного ряда — по точкам, лежащим в отличающихся областях фазового пространства. Наиболее естественно это делается с использованием временных рядов переходных процессов. Коэффициенты адекватной глобальной модели стационарной системы не должны зависеть от того, каких участков ее временного ряда ведется реконструкция. Коэффициенты же, появившиеся из-за искажений на отдельных тренировочных участках ряда, могут претерпевать в зависимости от выбора заметные изменения (рис. 6).

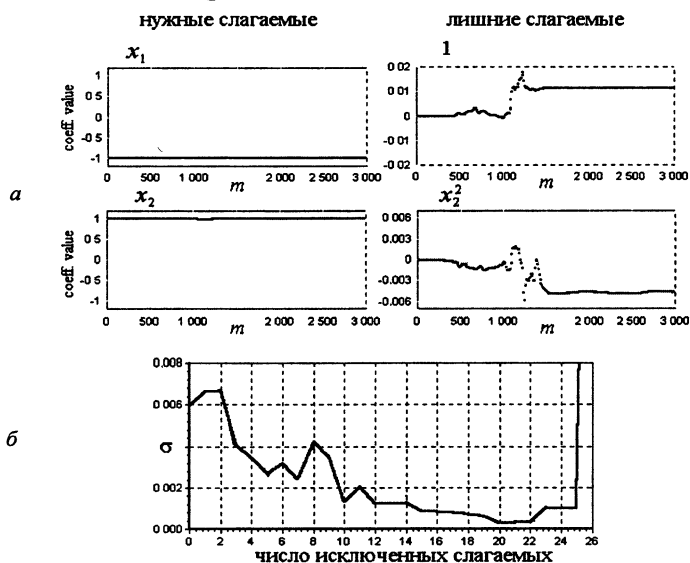


Рис. 6. Результаты реконструкции уравнения осциллятора Ван-дер-Поля — Тода ($d^2x/dt^2 = (1-x^2)dx/dt - 1 + e^{-x}$) по скалярному ряду: а — зависимость коэффициентов (при указанных рядом слагаемых полинома) модели от положения окна реконструкции; б — зависимость ошибки аппроксимации δ от числа исключенных лишних слагаемых.

Можно предложить процедуру улучшения структуры модели, основанную на последовательном удалении слагаемых, коэффициенты в которых наиболее нестабильны при изменении m . О степени стабильности (надежности) какого-либо коэффициента a можно судить по величине отношения его среднего значения к его стандартному отклонению, где усреднение проводится по ансамблю значений, полученных при разных m .

Заключение

Математическое моделирование по временным рядам – активно развивающаяся область математической статистики и нелинейной динамики. Построение и верификация эмпирических моделей имеют фундаментальное значение для лучшего понимания процессов самой различной природы. Их практические приложения очень разнообразны: от прогнозов будущего до технической и медицинской диагностики, но процедуры получения эффективных моделей не формализуются. Математическое моделирование, по-видимому, всегда останется в значительной степени искусством, но могут быть выделены некоторые общие принципы и частные рецепты (технологические приемы). Существенная часть из них описана в данной лекции. Мы попытались систематизировать их, опираясь на схему рис. 1, разные элементы которой были проиллюстрированы на примерах рассмотренных нами задач. Представлены в основном оригинальные методики, относящиеся к различным постановкам задач (от «прозрачного» до «черного ящика») и целям моделирования. Все они апробированы на численных примерах реконструкции по зашумленным решениям эталонных уравнений. Многие применялись нами для исследования реальных систем (нелинейных электрических цепей, функциональных систем живых организмов), а также для обработки записей биопотенциалов. Естественно, что изложенное является лишь «экскурсом в...», а не исчерпывающим изложением, наверное, вечного вопроса о возможностях математического моделирования объектов и явлений природы.

Исследования проводились при финансовой поддержке грантов РФФИ (проект № 03-02-17593, CRDF (REC-006), Минобразования РФ, Минпромнауки РФ, Президента РФ (МК-1067.2004.2), Фонда содействия отечественной науке, INTAS (No. 03-55-920).

Литература

1. *Льюнг Л.* Идентификация систем. Теория для пользователя. М.: Наука, 1991. 432 с.
2. *Демиденко Е. З.* Линейная и нелинейная регрессия. М.: Финансы и статистика, 1981. 302 с.
3. *Horbelt W., Timmer J., Bünner M. J., Meucci R., Ciofini M.* Identifying physical properties of a CO₂ laser by dynamical modeling of measured time series // Phys. Rev. E. 2001. V. 64. 016222.

4. Фейгин А. М., Мольков Я. И., Мухин Д. Н., Лоскутов Е. М. Прогноз качественного поведения динамической системы по хаотическому временному ряду // Изв. вузов. Радиофизика. 2001. Т. 44, № 5—6. С. 376—399.
5. Бокс Дж., Дженкинс Т. Анализ временных рядов. Прогноз и управление. Ч. 1. М.: Мир, 1974. 242 с.
6. Crutchfield J. P., McNamara B. S. Equations of motion from a data series // Complex Systems. 1987. V. 1. P. 417—452.
7. Cremers J., Hubler A. Construction of differential equations from experimental data // Z. Naturforschung A. 1987. V. 42. P. 797—802.
8. Casdagli M. Nonlinear prediction of chaotic time series // Physica D. 1989. V. 35. P. 335—356.
9. Smith L. A. Identification and prediction of low-dimensional dynamics // Physica D. 1992. V. 58. P. 50—76.
10. Kantz H., Schreiber T. Nonlinear time series analysis / Cambridge: Cambridge University Press, 1997. Chapter 12. P. 202—222.
11. Малинецкий Г. Г., Пономарев А. Б. Современные проблемы нелинейной динамики. М.: Эдиториал УРСС, 2000. 336 с. (гл. 11—14, с. 239—311).
12. Chaos and Its Reconstructions / Eds. G. Gouesbet, S. Meunier-Guttin-Cluzel and O. Menard. N. Y.: Nova Science Publishers, 2003. 320 p.
13. Rapp P. E., Schmah T. I., Mees A. I. Models of knowing and the investigation of dynamical systems // Physica D. 1999. V. 132. P. 133—149.
14. Павлов А. Н., Янсон Н. Б., Анищенко В. С. Реконструкция динамических систем // Радиотехника и электроника. 1999. Т. 44, вып. 9. С. 1075—1092.
15. Аносов О. Л., Бутковский О. Я., Кравцов Ю. А. Восстановление динамических систем по хаотическим временным рядам (краткий обзор) // Изв. вузов. Прикладная нелинейная динамика. 2000. Т. 8, № 1. С. 29—51.
16. Bezruchko B. P., Seleznev Ye. P., Ponomarenko V. I., Prokhorov M. D., Smirnov D. A., Dikanev T. V., Sysoev I. V., Karavaev A. S. Special approaches to global reconstruction of equations from time series // Изв. вузов. Прикладная нелинейная динамика. 2002. Т. 10, № 3. С. 137—158.
17. Pisarenko V.F., Sornette D. On statistical methods of parameter estimation for deterministically chaotic time series // Phys. Rev. E. 2004. V. 69. 036122.
18. Bremer C. L., Kaplan D. T. Markov chain Monte Carlo estimation of nonlinear dynamics from time series // Physica D. 2001. V. 160. P. 116—126.
19. Bock H. G. Numerical treatment of inverse problems in chemical reaction kinetics // Modelling of Chemical Reaction Systems / Eds. K. H. Ebert, P. Deuflhard, W. Jaeger. N. Y.: Springer, 1981. V. 18. P. 102—125.
20. Baake E., Baake M., Bock H. J., Briggs K. M. Fitting ordinary differential equations to chaotic data // Phys. Rev. A. 1992. V. 45. P. 5524—5529.
21. Bezruchko B. P., Smirnov D. A. Constructing nonautonomous differential equations from a time series // Phys. Rev. E. 2001. V. 63. 016207.
22. Bezruchko B. P., Karavaev A. S., Ponomarenko V. I., Prokhorov M. D. Reconstruction of time-delay systems from chaotic time series // Phys. Rev. E. 2001. V. 64. 056216.
23. Smirnov D. A., Bezruchko B. P., Seleznev Ye. P. Choice of dynamical variables for global reconstruction of model equations from time series // Phys. Rev. E. 2002. V. 65. 026205.
24. Bezruchko B. P., Dikanev T. V., Smirnov D. A. Role of transient processes for reconstruction of model equations from time series // Phys. Rev. E. 2001. V. 64. 036210.

РЕКОНСТРУКЦИЯ ДИНАМИЧЕСКИХ СИСТЕМ ПО ХАОТИЧЕСКИМ ВРЕМЕННЫМ РЯДАМ

Н. Г. Макаренко

Введение

В 1980 году Н. Паккард с коллегами [1] предложили следующий способ реконструкции фазовой геометрии 3D-динамической системы из скалярной проекции одной из ее координат: *эвристическая идея метода реконструкции состоит в том, что для описания состояния трехмерной системы в любой момент времени достаточно измерения любых трех независимых величин, где термин «независимый» пока не определен формально, но будет определен операционно*. Аналогами обычных координат $\{x_1, x_2, x_3\}$ могут быть значения единственной (наблюдаемой) переменной, например, $x_1(t)$ и две ее производные. В качестве альтернативы можно использовать также сдвиг $x_1(t)$ на временной лаг τ , так что новые координаты¹ имеют вид $\{x_1(i), x_1(i + \tau), x_1(i + 2\tau)\}$. Численные эксперименты показали, что фазовый портрет системы, построенный в новых координатах, похож на оригинал с точностью до непрерывных, но нелинейных деформаций.

Идея восстановления многомерной динамики из ее единственной проекции выглядела почти фантастической. Однако, в 1981 году Ф. Такенс [3] формализовал эвристику в форме строгой теоремы. Оказалось, что любую координату диссипативной системы или даже произвольную липшиц-непрерывную функцию этой координаты можно использовать для реконструкции универсальной модели — копии аттрактора в евклидовом пространстве R^n подходящей размерности. Необходимой для этого процедурой является топологическое вложение, известное в дифференциальной топологии. Для того чтобы копия наследовала свойства оригинала, который в общем случае может быть и неизвестен в явном виде, наблюдаемая проекция должна *допускать гладкое детерминированное объяснение* [4]. Это накладывает ряд математических требований на динамическую систему, точнее на ее аттрактор. Проверить эти требования практически невозможно: они являются *«символами веры»* в теореме Такенса. С точки зрения экспериментатора, они не слишком понятны, и следовательно, *ничтоже сумняшеся*, их просто можно принять как *«кредо идеального экспериментатора»* [5]. Возникающие в результате

¹ Такие координаты называют *запаздывающими*; они были предложены в 1980 году для описания предельных циклов в рамках пороговых AR-моделей [2].

возможности превышают даже то, что мы не смели предположить. Действительно, реконструкция, полученная вложением временного ряда в R^n , топологически эквивалентна оригиналу и, следовательно, имеет ту же самую качественную и количественную динамику. Таким образом, ее не только можно предъявить в качестве детерминированного объяснения наших наблюдений, но и использовать для вычислений размерности, ляпуновских показателей и энтропии неизвестного оригинала. Более того, можно построить многомерный нелинейный прогноз хаотических временных рядов [6] и даже восстановить явный вид уравнений [7]. Способ извлечения «геометрии из временных рядов» инициировал поток публикаций, посвященных численным алгоритмам для оценок параметров вложения и динамических инвариантов аттракторов, редукции шума и нелинейного прогноза. В результате возник новый раздел теории детерминированного хаоса. Тим Зауер и др. [8] предложили назвать ее *эмбедологией* — от английского слова *embedding* (вложение)².

Существует большое число обзоров, посвященных эмбедологии (см., например, [8—20]). Они написаны на разном уровне математической строгости. Большая их часть содержит технические аспекты эмбедологии: «алгоритма Такенса» и его приложения к нелинейному анализу временных рядов [16—20]. Меньшая часть, которая включает описание «топологического фона» теоремы, рассчитана на весьма подготовленного читателя [8—15]. Формализм, лежащий в основе «алгоритма Такенса», не просто красив, но и важен для понимания результатов реконструкции. Его основой являются понятие типичного вложения, догма о трансверсальности и принцип линеаризации, важный для доказательства теоремы. Данная статья содержит популярное введение в первые два принципа эмбедологии и следует лекции автора, прочитанной на Школе. Изложение рассчитано на физиков-экспериментаторов, так что все технические детали опущены. Более конструктивную часть эмбедологии, в которой рассматриваются проблемы моделирования, можно найти в статье А. М. Фейгина и др. [21] в этом же сборнике.

Отображения, типичность и многообразия

В этом разделе для удобства читателя собраны некоторые необходимые определения и понятия, которые используются в статье.

- Запись $x \in X$ означает, что x принадлежит X .
- Запись $\{x | d(a, x) < r\}$ читается: множество таких x , для которых расстояние от точки a меньше r .

² Синонимом является термин *алгоритмическое моделирование*, предложенный Rapp P. E. et al. [9].

- Запись $A \subset B$ ($A \subseteq B$) означает, что A содержится в B (содержится или совпадает с B).
- Объединением множеств A и B называют множество $C = A \cup B$, состоящее из элементов $c \in C$, каждый из которых принадлежит A или (*и*) B .
- Пересечением A и B называют множество $C = A \cap B$, каждый элемент c которого входит и в A , и в B .

Соответствия между различными множествами описываются функциями или отображениями $f(\dots)$ или $f: X \rightarrow Y$. Отображение $f: X \rightarrow Y$ называют *сюръекцией* или *отображением на*, если образ всего A совпадает с B , т. е. каждый элемент из B является образом по крайней мере одного элемента из A . Отображение $\varphi: A \rightarrow B$ называют *инъективным* или *взаимно однозначным*, если для каждого $y \in B$ существует не более одного элемента $x \in A$ такого, что $y = \varphi(x)$. Тогда $\varphi(x) = \varphi(y)$ влечет $x = y$. Обратную к $f(x)$ функцию будем обозначать $f^{-1}(x)$. Отображение, являющееся инъективным и сюръективным, называется *биекцией*. Отображение $f: X \rightarrow Y$ *непрерывно* в $x \in X$, если для любой окрестности V , открытой в Y и содержащей точку $f(x)$, найдется такая открытая окрестность $U(x) \subset X$, что $f(U) \subset V$. Напомним, что *открытым* шаром в точке $a \in M$ в метрическом пространстве (M, d) называют подмножество $B(a, r) = \{x \in M \mid d(a, x) < r\}$; для топологических пространств *открытые* подмножества (или топология) выделяются как подмножества, замкнутые относительно объединения и пересечения. Отображение φ открытого множества $U \in R^n$ называется *гладким*, если оно имеет непрерывные частные производные всех порядков: говорят, что φ имеет класс C^r , если это справедливо для производных до порядка r включительно. Известно важное свойство гладких функций: если дифференциал такой функции в точке x имеет некоторое свойство, то сама функция будет иметь такое же свойство по меньшей мере в окрестности точки x . Например, если производная $f'(x_0) > 0$ положительна, то в малой окрестности точки x_0 функция f возрастает. Это свойство называют обычно *принципом общего положения* или *принципом линеаризации*, поскольку дифференциал функции является линейной непостоянной частью отображения. Принципу общего положения удовлетворяет не каждая функция, поэтому важно сделать математически точными утверждения о выполнимости некоторого свойства

для «почти каждой» функции. Вообще говоря, любые свойства математических объектов можно классифицировать как обычные (*usual*) и необычные (*unusual*). Стандартный прием выразить распространенность некоторого свойства, выполняющегося на подмножестве $A \subset B$, сводится к тому, чтобы показать, что обычные свойства соответствуют подмножеству с большей мерой. Однако пространство функций является бесконечномерным, и поэтому здесь невозможно прибегнуть к замене выражения «почти каждая» на утверждение «с вероятностью единица». Другой способ сделать это основан на топологии. Рассмотрим два множества A и B таких, что $A \subset B$. Говорят [22], что A *плотно* в B , если каждое открытое подмножество B содержит точку A . Кажется разумным связать «обычные» свойства с множеством A , которое плотно в B , но в некоторых случаях дополнение A в B также плотно в B . Например, подмножество рациональных чисел плотно в R , но таким же является и его дополнение — подмножество иррациональных чисел. Однако, интуитивно ясно, что иррациональных чисел в R «больше». Чтобы описать такие ситуации, было введено понятие *остаточного* множества (*residual set*). Подмножество A *остаточное* в B , если оно является пересечением конечного или счетного числа открытых и плотных подмножеств в B . Любое счетное пересечение остаточных множеств плотно. Ради простоты мы будем понимать выражение «с точностью до предположения о типичности» в следующем не очень строгом смысле. Свойство *типично* для функции f , если оно выполняется в каждой точке соответствующего пространства и открытой окрестности этой точки. Если оно не выполняется в точке p , всегда можно найти достаточно близкую к ней точку, в которой это свойство справедливо. Аналогом является трактовка: свойство типично, если оно справедливо для f , либо, если это не так, становится справедливым для $f + \delta f$, где порядок возмущения δf заранее оговаривается.

Рассмотрим группу преобразований $G: R^2 \rightarrow R^2$, заданную выражением $\tilde{x} = Ax$. Если матрица A ортогональна, то обычную геометрию можно определить как совокупность тех свойств, которые сохраняются (инвариантны) относительно действия этой группы. Аффинная геометрия получается, если условие ортогональности заменить на требование $\det A \neq 0$. Следующим обобщением будет группа, заданная в виде

$$x' = f(x, y), y' = g(x, y),$$

где f и g — нелинейные непрерывные функции. Если такими свойствами обладают и обратные преобразования (f^{-1}, g^{-1}) , мы получим геометрию *многообразий* [23]. Это совокупность свойств, инвариантных относительно взаимно однозначных и непрерывных преобразований, т. е. *гомеоморфизмов*. Например, сфера S^2 и бесконечное число ее копий, не

различимых с точностью до гомеоморфизмов, являются многообразием. Примером многообразия, не гомеоморфного S^2 , является тор T^2 . Гомеоморфизмы позволяют определить многообразие M как топологическое пространство, локально гомеоморфное R^m . Примером может служить сфера, но не два конуса, склеенных вершинами (рис. 1).

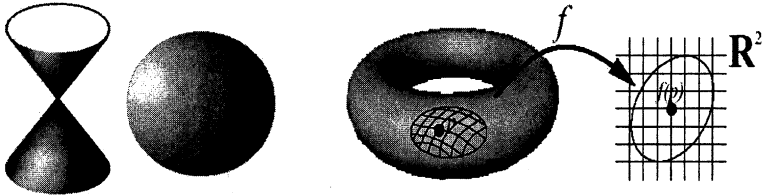


Рис. 1. Слева: окрестность общей вершины двух конусов локально не гомеоморфна плоскости в отличие от сферы. Справа: локальная карта тора.

Каждый гомеоморфизм $\varphi: M \rightarrow R^m$ определяет карту локальной окрестности V , содержащей точку p . Локальные координаты на карте используются для анализа структуры многообразия (рис. 1); размерностью последнего, $\dim M$, является размерность карты. Пусть $p \in U \cap V$ и $\varphi_1(U) \rightarrow R^2$ и $\varphi_2(V) \rightarrow R^2$ — две карты. Если функции $\varphi_2 \circ \varphi_1^{-1}$ и $\varphi_1 \circ \varphi_2^{-1}$ являются C^r -гладкими, то говорят, что M — C^r -гладкое многообразие. отображение $\varphi: X \rightarrow Y$ двух гладких многообразий или двух произвольных множеств в R^n называют диффеоморфизмом, если φ и φ^{-1} — гладкие гомеоморфизмы. Окружность в R^2 диффеоморфна эллипсу, но не треугольнику, которому она лишь гомеоморфна. Диффеоморфизмы определяют класс эквивалентных диффеоморфных пространств.

Кривая на M , проходящая через точку p , — это отображение $I \equiv [0, 1] \rightarrow M$, сопоставляющее каждой $\lambda \in I$ точку $c(\lambda) \in M$, так что $c(0) = p$. Две кривые различны, даже если их образы совпадают, но одинаковым точкам p в M соответствуют разные λ . Функция на M — отображение $M \rightarrow R^1$, сопоставляющее каждой точке $p \in M$ вещественное число. Определим касательный вектор к $p \in M$. Идея стрелки, «приклеенной» к $p \in M$, дает подходящее представление, недостатком которого является необходимость иметь некоторое объемлющее пространство, в которое выходит конец стрелки.

Пусть через точку $p \in M$ проходит кривая $I \rightarrow M: \lambda \rightarrow c(\lambda)$ и $c(0) = p$. Будем изучать смещение точки p с помощью произвольной

скалярной функции f : мы просто используем в качестве аргумента этой функции точки, принадлежащие кривой. В инфинитезимальной окрестности p изменение функции $f(c(\lambda)) - f(c(0))$ не зависит от самой функции — это просто сдвиг p и больше ничего! Но сдвиг — это оператор, который выражает общее свойство всех кривых, проходящих через p , иметь одно и то же изменение для произвольных функций, моделирующих смещение p . Этот оператор $d/d\lambda$ называют *касательным вектором* к p на M . Множество касательных векторов к p образуют *касательное пространство* T_p ; $\bigcup_{p \in M} T_p$ называют *касательным расслоением* к M .

Погружения и вложения

Пусть $f: X \rightarrow Y$ — гладкое отображение двух многообразий и $\dim X < \dim Y$. Этому отображению можно поставить в соответствие отображение касательных пространств $df: T_x(X) \rightarrow T_{f(x)}(Y)$, «приклеенных» к точкам x и $f(x)$ соответственно. Понятно, что в нашем примере df не может быть сюръекцией, потому что оно не в состоянии «накрыть» лишние координаты в Y . Самое лучшее, на что мы можем претендовать, — это считать df инъекцией. В этом случае f называют *иммерсией* или *погружением* $X \rightarrow Y$ [23—24]. Например, окружность S^1 можно погрузить в R^2 так, что ее образом будет «восьмерка». При этом взаимнооднозначное соответствие точек образа и прообраза отсутствует, но каждому вектору скорости на S^1 соответствует единственный касательный вектор «восьмерки» (рис. 2, справа).

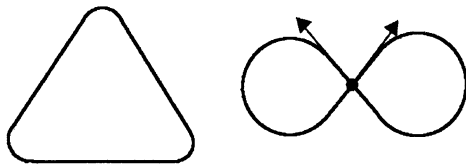


Рис. 2. Вложение окружности $S^1 \rightarrow R^2$ (слева) и ее иммерсия (справа)

Каноническая иммерсия получается, если евклидово пространство R^k погрузить в R^m ($k < m$). При этом точка $x \in R^k$ линейно отображается в точку $f(x) \in R^m$:

$$(x_1, x_2, \dots, x_k) \rightarrow (x_1, x_2, \dots, x_k, 0, \dots, 0).$$

В этом случае образ иммерсии — подпространство в R^m ; для многообразий, которые являются только локально евклидовыми, это не так. Например, в случае «восьмерки» окрестность точки самопересечения не является открытой в R^2 . В общем случае гладкое отображение F компактного гладкого многообразия A называют *иммерсией*, если дифференциал отображения $dF(x)$ является взаимно однозначным³ в каждой точке A . При этом безразлично, имеет ли само F такое свойство. Гомеоморфизм $A \rightarrow F(A)$ называют *топологическим вложением* (см. рис. 1, слева). *Дифференцируемое вложение* A — это гладкий диффеоморфизм из A на собственный образ $F(A)$. Если A — компактное (т. е. ограниченное и замкнутое) многообразие, то отображение F является вложением тогда и только тогда, когда F — взаимно однозначная иммерсия. Знаменитая теорема Уитни [23] показывает, что всякое компактное многообразие можно считать подмногообразием евклидова пространства достаточно большой размерности. Теорема гласит: *Пусть M — гладкое компактное k -мерное многообразие. Тогда существует вложение $F : M \rightarrow R^n$ для $n = 2k + 1$. Важно, что вложение может быть реализовано почти любой функцией, удовлетворяющей упомянутым условиям, или, что то же самое, множество отображений в R^{2k+1} , которые являются вложениями, всюду плотны в классе всех непрерывных отображений. Иными словами, если данное гладкое отображение есть вложение, его малое возмущение также является вложением. С другой стороны, любое гладкое отображение независимо от того, вложение это или нет, «лежит» произвольно близко от вложения. Условие теоремы Уитни для размерности R^n можно трактовать следующим образом: для взаимной однозначности образа и прообраза отображения следует задать k координат; для такого же соответствия между касательными векторами необходимо еще k чисел. Наконец, единица гарантирует дополнительную степень свободы для того, чтобы избежать ситуации с самопересечением. Тор T^2 требует, согласно теореме Уитни, R^5 для вложения. Этот избыток размерностей гарантирует, что любые диффеоморфные копии тора будут вложениями! Существует важное обобщение теоремы Уитни на случай, когда A является фракталом. Здесь не существует разумного определения топологической размерности. Вместо нее используют так называемую *box-размерность*, которая обычно не является целым числом [25]. Для компактного подмножества $A \in R^k$ с *box-размерностью* d поч-*

³ В многомерном случае $dF(x)$ задается матрицей Якоби. Тогда иммерсия эквивалентна условию, что якобиан имеет полный ранг T_x .

ти каждое гладкое отображение $F: R^k \rightarrow R^n$, где $n \geq 2d$, является взаимно однозначной иммерсией каждого компактного гладкого многообразия $C \supseteq A$ [10]. Геометрические истоки условия Уитни лежат в догме трансверсальности.

Трансверсальность

Трансверсальность является геометрическим аналогом типичности: во взаимном расположении геометрических тел в пространстве. Две линии в R^3 в общем случае не пересекаются: в самом деле, достаточно слегка «пошевелить» пересекающиеся прямые, чтобы достичь этого общего случая. Напротив, в R^2 пара прямых, как правило, пересекается. Две плоскости в R^3 либо параллельны, либо пересекаются вдоль линии; второй случай, очевидно, устойчив относительно «малых шевелений». Линия и плоскость имеют, как правило, общую точку — линия, параллельная плоскости, не устойчива. Эти наблюдения легко обобщаются на R^n . Пусть R^k и R^l — два подпространства в R^n . В общем случае они не пересекаются, если $(k+l) \leq n$; в противном случае они пересекаются так, что размерность их «общей части» равна $(k+l) - n$. Ситуации взаимного расположения геометрических объектов, устойчивые относительно малых шевелений, называют *трансверсальными*. На рис. 3 (вверху) приведены примеры трансверсальных пересечений в R^3 . В нижнем ряду все пересечения в R^2 не являются трансверсальными. Может случиться, что пересечение вообще отсутствует. Такая ситуация также считается трансверсальной. Поскольку локальной версией многообразий являются касательные пространства, приведенные рассуждения легко обобщаются для взаимных конфигураций двух подмногообразий K и L в M . Говорят, что K и L пересекаются в M трансверсально, если для каждой точки $p \in K \cap L$

$$T_p K + T_p L = T_p M.$$

Пусть $L \equiv R$, а $K \equiv R^2$, $M \equiv R^3$. Тогда в точке их трансверсального пересечения мы имеем тройку касательных векторов: два из них принадлежат плоскости, а третий — линии. Эта тройка — локальная модель R^3 . Таким образом, трансверсальное пересечение локально моделирует пространство, в котором оно имеет место. Адаптируем это определение для отображений. Пусть $Z \subset M$, $f: N \rightarrow M$ — гладкое отображение и $p \in N$ — точка. Говорят, что f — трансверсально Z в p , если либо $f(p) \notin Z$, либо $T_p f(T_p N) + T_{f(p)} Z = T_{f(p)} M$. Итак, условие трансвер-

сальности необходимо для описания типичной ситуации. В локальном варианте его можно редуцировать в принцип общего положения.

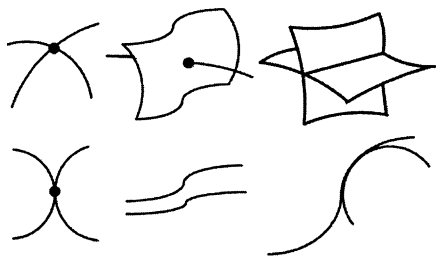


Рис. 3. Примеры взаимного расположения геометрических тел в R^3 (вверху) и R^2 (внизу): трансверсальными являются 2-й и 3-й примеры в верхнем ряду.

Рассмотрим теперь связь между трансверсальностью и условием теоремы Уитни. Пусть по-прежнему R^k и R^l — два подпространства в R^n , которые пересекаются трансверсально. Тогда верхняя часть рис. 3 учит нас, что в случае трансверсальности размерности k, l, n удовлетворяют неравенству $(k+l-n) \geq 0$. Следовательно, при отсутствии трансверсального пересечения должно выполняться неравенство $k+l-n < 0$. Если $k=l=d$, условие $n > 2d$ гарантирует нам отсутствие самопересечения. Самый простой способ удовлетворить этому неравенству — положить $n = 2d + 1$. Это и есть условие теоремы Уитни. Так, любую кривую ($d = 1$) можно вложить в R^3 : возможное самопересечение легко ликвидируется малым шевелением.

Эмбедология и теорема Такенса

Напомним (см., например, [10, 12]), что динамической системой в фазовом пространстве M называется однопараметрическая группа диффеоморфизмов $f^t(x): M \rightarrow M$. Параметр группы t — это время, которое может принимать непрерывные или дискретные значения. В качестве M выступают обычные евклидовы пространства R^n или гладкие многообразия. Диффеоморфизмы f^t могут быть заданы в виде дискретных или непрерывных преобразований; групповое свойство означает:

$$f^{t_2}(f^{t_1}(x)) = f^{t_1+t_2}(x).$$

При этом $f^0(x) = x$ — единица группы; если $x_1 = f^{t_1}(x_2)$, то $x_2 = f^{-t_1}(x_1)$. Если динамическая система задана дифференциальным уравнением $\dot{x} = F(x)$, то в первом приближении

$$x(t + dt) = x + dx = x + Fdt = f^{dt}(x) = f^0(x) + \frac{df^0(x)}{dt}dt + \dots = x + \frac{df^0(x)}{dt}dt.$$

Функцию $F(x) = \left. \frac{df^t(x)}{dt} \right|_{t=0}$ называют векторным полем, которое зада-

ет касательную к траектории в каждой точке фазового пространства. *Траекторией или орбитой* динамической системы называют последовательность точек

$$\dots, f^{-3}(x), f^{-2}(x), f^{-1}(x), x, f(x), f^2(x), \dots$$

Неподвижной точкой x_0 называется траектория, которая для всех t удовлетворяет условию: $f^t(x_0) = x_0$. *Периодической* называется траектория, не являющаяся неподвижной точкой, и такая, что для некоторого P и произвольного t выполнено равенство $f^{t+P}(x) = f^t(x)$. Точка p называется ω -предельной точкой для x , если $f^t(x) \rightarrow p$ при $t \rightarrow \infty$. Пусть $A \in M$ и $f^t(A)$ — множество образов всех точек из A ; тогда A инвариантно относительно f , если $f^t(A) = A$ для всех t . Замкнутое инвариантное множество $A \subset M$ называется *притягивающим*, если для него существует окрестность U такая, что для всех $x \in U$, $f^t(x) \rightarrow A$, когда $t \rightarrow \infty$. Инвариантное притягивающее множество, которое содержит все ω -предельные точки, называют *аттрактором*. Если аттрактор достаточно велик и содержит только конечное число неподвижных точек и периодических орбит, то поведение фазовых кривых на аттракторе и рядом с ним является хаотическим. Наиболее важным свойством хаотического аттрактора является существенная зависимость от начальных условий. В этом случае для f^t , точки $x \in M$ и некоторого числа $\delta > 0$ существует время $T > 0$ и точка y , принадлежащая окрестности x , такие, что расстояние $d(f^T(x), f^T(y)) > \delta$. Иными словами, погрешности начальных данных экспоненциально растут в фазовом потоке, так что начиная с некоторого момента времени будущее состояние системы становится непредсказуемым. Теперь мы располагаем всем необходимым, чтобы привести формулировку теоремы Такенса [11].

Теорема Такенса. Пусть D^r — множество C^r -диффеоморфизмов $\{f^t\}$, $f^t: M \rightarrow M$, компактного d -мерного многообразия M и пусть $C^r(M, R)$ — множество наблюдаемых функций $h(t)$, $h: M \rightarrow R$. Определим для $m \geq 2d + 1$ отображение запаздывающих координат $F_{f,h}: M \rightarrow R^m$:

$$F_{f,h}(\bar{x}) = [h(\bar{x}), h(f(\bar{x})), \dots, h(f^{m-1}(\bar{x}))].$$

Тогда множество (f^t, h) , для которого $F_{f,h}$ является вложением, открыто и всюду плотно в $D^r(M) \times C^r(M, R)$ для $r \geq 1$.

Первая часть теоремы содержит наши символы веры, сводящиеся фактически к тому, что достаточно гладкие (т. е. класса $C^r(M, R)$) наблюдения — временной ряд $h(t)$ — продуцируются $C^r(M)$ -гладкой динамической системой, которая имеет компактный d -мерный аттрактор.

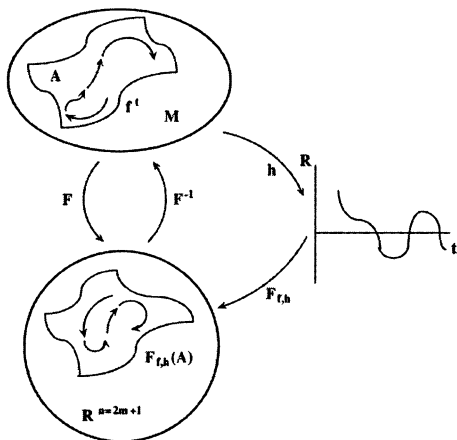


Рис. 4. Иллюстрация теоремы Такенса (из [15])

Вторая часть теоремы учит нас, как, используя временной ряд $h(t)$, построить копию аттрактора в евклидовом пространстве R^m размерности $m = 2d + 1$ (см. рис. 4). Для этого следует сперва взять m отсчетов временного ряда, начиная с произвольного номера n : $h_n, h_{n+1}, \dots, h_{n+m-1}$. Они рассматриваются как компоненты m -мерного вектора и образуют точку образа $z_n \in R^m$:

$$\begin{aligned} z_n &= [h(f^n(x_0)), h(f^{n+1}(x_0)), \dots, h(f^{n+m-1}(x_0))] = \\ &= [h(x_n), h(f(x_n)), \dots, h(f^{m-1}(x_n))] = F(x_n). \end{aligned}$$

Следующая точка получается сдвигом m -мерного набора на один отсчет, а итеративное продолжение этой процедуры даст упорядоченную последовательность точек — копию истинной орбиты. Действительно, поскольку $F_{f,h}$ — гладкое и обратимое преобразование, можно определить

отображение $\sigma = F \circ f \circ F^{-1}$. Так как точка z_n является образом F , применение σ в R^m дает

$$\begin{aligned}\sigma(z_n) &= F \circ f \circ F^{-1}(z_n) = F \circ f \circ F^{-1}(F(\mathbf{x}_n)) = \\ &= F \circ f(\mathbf{x}_n) = F(\mathbf{x}_{n+1}) = z_{n+1}.\end{aligned}$$

Заметим, что теорема справедлива для диффеоморфизмов f' , имеющих лишь конечное число неподвижных точек и периодических орбит с периодом меньше m [10]. В противном случае для такой периодической точки нарушается даже условие иммерсии (погружения) $F_{f,h}$. Наконец, теорема гарантирует, что копия будет вложением, т. е. сохранит свойство оригинала с точностью до диффеоморфизмов. Отображение F , которое реализует вложение, называют отображением запаздывающих координат, а пространство R^m называют пространством вложения. Хотя справедливость теоремы не зависит от величины сдвига (лага) между отсчетами (в приведенной выше формулировке он был выбран равным единице), на практике требуется выполнение условия «независимости» новых запаздывающих координат. В противном случае реконструкция «коллапсирует» к диагональной гиперплоскости в пространстве вложения. Образ вложения z_{n+1} является вектором, однако легко убедиться, что уравнение, моделирующее динамику в пространстве вложения, имеет единственную нетривиальную компоненту вида

$$h_{n+n} = F_m(h_n, h_{n-1}, \dots, h_{n-m+1}).$$

Это выражение и является основой для многомерного нелинейного прогноза. Фактически, это нелинейное обобщение традиционного авторегрессионного прогноза, в котором порядок авторегрессии заменяется размерностью вложения. Относительно функции F_m известно лишь, что она непрерывна и определена на конечном множестве прошлой истории ряда. Поэтому ее численная аппроксимация является некорректной задачей. Наиболее удачным аппроксиматором является искусственная нейронная сеть, которая позволяет представить функцию m переменных F_m как суперпозицию некоторых стандартных функций одной переменной [25].

В заключение уместно заметить, что практическое применение теоремы Такенса требует большой осторожности. Проблема шумов, конечность временных рядов и возможное присутствие скрытых параметров в модели вложения — далеко не полный перечень подводных камней эмбедологии. В этой области всегда полезно следовать мудрому совету Козьмы Пруткова: «Имея в виду какое-либо предприятие, помысли сперва, точно ли оно тебе удастся».

Литература

1. *Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S.* Geometry from a time series // *Phys. Rev. Lett.* 1980. V. 45. P. 712—716.
2. *Tong H., Lim K. S.* Threshold autoregression, limit cycles and cyclical data // *J. R. Statist. Soc. B.* 1980. V. 42, № 3. P. 245—292.
3. *Takens F.* Detecting strange attractors in turbulence // *Lecture Notes in Math.* 1981. V. 898. P. 366—381.
4. *Takens F.* Distinguishing deterministic and random systems // *Nonlinear dynamics and turbulence* / Eds G. J. Barenblatt, G. Jooss, D. D. Joseph. N.Y.: Pitman, 1983. P. 314—333.
5. *Афраймович В. С., Рейман А. М.* Размерности и энтропии в многомерных системах // *Нелинейные волны. Динамика и эволюция.* М.: Наука, 1989. С. 238—262.
6. *Farmer J. D., Sidorovich J. J.* Predicting chaotic time series // *Phys. Rev. Lett.* 1987. V. 59. P. 845—848.
7. *Gouesbet G., Meunier-Guttin-Cluzel S., Menard O.* Global reconstructions of equations of motion from data series, and validation techniques, a review // *Chaos and its reconstruction.* N.Y.: Nova Science Inc., 2003. P. 1—160.
8. *Sauer T., Yorke J. A., Casdagli M.* Embedology // *J. Stat. Phys.* 1991. V. 65. P. 579—616; URL:<http://math.gmu.edu/~tsauer>.
9. *Rapp P. E., Schah T. I., Mees A. I.* Models of knowing and the investigation of dynamical systems // *Physica D.* 1999. V. 132. P. 133—149.
10. *Ott E., Sauer T., Yorke J. A.* Coping with chaos: Analysis of chaotic data and the exploitation of chaotic systems. John Wiley and Sons, 1994.
11. *Noakes L.* The Takens embedding theorem // *Int. J. Bifurcation and Chaos.* 1991. V. 1. P. 867—872.
12. *Ruelle D.* Chaotic evolution and strange attractors. The statistical analysis of time series for deterministic nonlinear systems. Cambridge University Press, 1989.
13. *Gilmore R.* Topological analysis of chaotic dynamical systems // *Rev. Mod. Phys.* 1998. V. 70. P. 1456—1529.
14. *Eckmann J. P., Ruelle D.* Ergodic theory of chaos and strange attractors // *Rev. Mod. Phys.* 1985. V. 57. P. 617—656.
15. *Engel M.* Time series analysis. A part III. Essay // URL: <http://www.m-engel.de>.
16. *Schreiber T.* Interdisciplinary application of nonlinear time series methods // *Phys. Rep.* 1999. V. 308, № 2. URL:<http://xyz.lanl.gov/chaos-dyn/980700>.
17. *Lay Ying-Cheng, Ye N.* Recent developments in chaotic time series analysis // *Int. J. Bifurcation and Chaos.* 2003. V. 13, № 6. P. 1383—1422.
18. *Huke J.* The dynamical systems approach to nonlinear signal processing // URL: <http://personalpages.umist.ac.uk/staff/j.p.huke>.
19. *Bradley E.* Time-series analysis // *Intelligent Data Analysis. An Introduction* / Eds. M. Berthold, D. Hand. Springer Verlag, 2003. P. 199—227.
20. *Малинецкий Г. Г., Потанов А. Б.* Современные проблемы нелинейной динамики. М.: Эдиториал УРСС, 2000.
21. *Лоскутов Е. М., Мольков Я. И., Мухин Д. Н., Фейгин А. М.* Статистический подход к реконструкции динамических систем // *Настоящий сборник.* С. 411—425.
22. *Келли Дж. Л.* Общая топология. М.: Наука, 1968.
23. *Хирш М.* Дифференциальная топология. М.: Мир, 1979.
24. *Guillemin V., Pollak A.* Differential topology. Prentice-Hall. Inc., Englewood Cliffs, New Jersey, 1974.
25. *Макаренко Н. Г.* Фракталы, мультифрактальные меры и аттракторы // *Нелинейные волны – 2002.* Н. Новгород: ИПФ РАН, 2003. С. 381—394; *Макаренко Н. Г.* Эмбедология и нейропрогноз // *Нейроинформатика-2003.* Ч. 1. Москва, 2003. С. 86—148.

СТАТИСТИЧЕСКИЙ ПОДХОД К РЕКОНСТРУКЦИИ ДИНАМИЧЕСКИХ СИСТЕМ

*Е. М. Лоскутов, Я. И. Мольков,
Д. Н. Мухин, А. М. Фейгин*

В опубликованных недавно работах обсуждается применимость байесова (*статистического*) подхода к реконструкции динамических систем (ДС) по экспериментальным данным. Существенным достоинством подхода является универсальность: он «по происхождению» пригоден для анализа временных рядов (ВР), зашумленных произвольным образом (как в процессе генерации, так и при измерениях). С другой стороны, корректный (в смысле соблюдения условий лежащей в основе теоремы) байесов подход к реконструкции ДС практически не реализуем в наиболее интересном случае зашумленных хаотических ВР. В данной статье мы обсуждаем модификацию байесова подхода, позволяющую применить его для реконструкции ДС по зашумленным ВР. Мы демонстрируем эффективность модифицированного подхода при решении двух типов задач: 1) при отыскании значений параметров известной ДС по зашумленному ВР; 2) при классификации режимов поведения такой ДС по короткому сильно зашумленному ВР.

1. Статистический подход к реконструкции динамических систем

Под реконструкцией ДС по порожденному ей ВР обычно понимают отыскание ее оператора эволюции. В случае, когда ДС *известна*, речь идет о нахождении значений параметров, которые определяли эволюцию системы во время генерации ВР; такая постановка возникает, например, при использовании хаотических режимов поведения ДС для решения задачи скрытой передачи информации [1]. В ситуациях, характерных для большинства природных систем (атмосферно-океанических, тектонических, биологических), ДС, породившая наблюдаемый ВР, *неизвестна*. Реконструкция ДС в этом случае означает построение (на основе информации, содержащейся в ВР) параметризованной модели неизвестного оператора эволюции. Очевидно, что такая модель не может быть идеальной: для любого класса моделей не существует, вообще говоря, набора значений параметров, при котором модель полностью адекватна моделируемой ДС; неизбежное расхождение между ними называют «дефектом модели» [2]. В этом смысле две указанные постановки задачи о реконструкции ДС иногда классифицируют как «сценарий идеальной модели» и «сценарий неидеальной модели» [3].

Будем полагать, что по имеющемуся в нашем распоряжении ВР *одной* наблюдаемой (скалярной) величины может быть установлен факт динамичности системы, сгенерировавшей ВР, и определена минимальная размерность вложения d аттрактора, ответственного за ее наблюдаемую эволюцию. Методы получения этой информации и имеющиеся (весьма существенные) ограничения обсуждались многими авторами и суммированы в различных обзорах и монографиях (см., например, [4, 5]), основные из этих методов описаны в статье Н. Г. Макаренко в данном сборнике [6]. Сделанные предположения позволяют нам далее рассматривать в качестве наблюдаемой векторную величину $\mathbf{x} \in R^d$, построенную по исходному скалярному ВР методом координат с задержками ([4—6]).

Удобная для дальнейшего формулировка теоремы Байеса выглядит следующим образом [7]. Допустим, что система, над которой производится эксперимент, обладает набором свойств (параметров) μ , прямое измерение которых невозможно, и пусть в процессе эксперимента регистрируются значения величины \mathbf{x} . Тогда апостериорная условная плотность вероятности *ненаблюдаемых* параметров (часто называемая правдоподобием) $p(\mu | \mathbf{x})$ пропорциональна произведению их априорной плотности вероятности $p(\mu)$ и априорной условной плотности вероятности полученных экспериментальных результатов $p(\mathbf{x} | \mu)$:

$$p(\mu | \mathbf{x}) = C \times p(\mu) \times p(\mathbf{x} | \mu). \quad (1)$$

Как будет видно из дальнейшего, условная плотность вероятности $p(\mathbf{x} | \mu)$ целиком определяется «способом» зашумления ВР и плотностями вероятности всех присутствующих в ВР шумов. Фактор $p(\mu)$ позволяет учесть имеющуюся *a priori* информацию о системе. Если такая отсутствует, плотность вероятности $p(\mu)$ должна быть выбрана постоянной с шириной, заведомо допускающей все мыслимые значения параметров μ . Постоянная C в (1) определяется условием нормировки:

$$C = \left\{ \int p(\mu) p(\mathbf{x} | \mu) d\mu \right\}^{-1}.$$

Естественным основанием применения *вероятностного* байесова подхода к построению моделей ДС является наличие в экспериментальных данных шумовой компоненты [8]. Рассмотрим в качестве примера ДС с дискретным временем и предположим для определенности, что как погрешность («шум») измерений ξ , так и случайное воздействие η на систему в процессе генерации измеряемой величины \mathbf{x} («динамический шум») аддитивны:

$$\xi_t = \mathbf{x}_t - \mathbf{u}_t, \quad \eta_t = \mathbf{u}_t - f(\mathbf{u}_{t-1}, \mu). \quad (2)$$

Здесь индекс t нумерует отсчеты дискретного времени, вектор $\mathbf{u} = \{u_k\}_{k=1}^d$ задает «истинное» (латентное) состояние ДС в d -мерном фазовом пространстве (пространстве вложения), вектор-функция $\mathbf{f}(\mathbf{u}, \boldsymbol{\mu})$ описывает оператор эволюции ДС (определяемой вектором параметров $\boldsymbol{\mu} = \{\mu_m\}_{m=1}^M$), который мы представили в виде функции последования:

$$\mathbf{u}_{t+1} = \mathbf{f}(\mathbf{u}_t, \boldsymbol{\mu}). \quad (3)$$

Всюду далее мы будем считать, что случайные величины ξ и η взаимно независимы и не зависят от состояния системы.

Поскольку «истинные» состояния ДС неизвестны, фигурирующие в (1) плотности вероятности зависят не только от параметров $\boldsymbol{\mu}$, но и от латентной переменной \mathbf{u} : $p(\boldsymbol{\mu} | \mathbf{x}) \Rightarrow p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$; $p(\boldsymbol{\mu}) \Rightarrow p(\mathbf{u}, \boldsymbol{\mu})$, $p(\mathbf{x} | \boldsymbol{\mu}) \Rightarrow p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu})$; при этом априорная условная плотность вероятности $p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu})$ целиком определяется свойствами случайных величин ξ и η . Если шумы взаимно независимы, а их плотности вероятности описываются функциями w_ξ и w_η , выражение для $p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu})$ имеет вид:

$$p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu}) = \prod_{t=1}^T w_\xi(\mathbf{x}_t - \mathbf{u}_t) w_\eta(\mathbf{u}_t - \mathbf{f}(\mathbf{u}_{t-1}, \boldsymbol{\mu})). \quad (4)$$

В случае, когда ДС известна, соотношения (1), (4) решают (в вероятностной постановке) задачу отыскания значений параметров ДС, при которых сгенерирован экспериментальный ВР. Кроме того, они позволяют произвести фильтрацию шума: найти наиболее вероятные незамешанные значения измеряемой динамической переменной.

Если ДС не известна, фигурирующая в (1), (4) функция $\mathbf{f}(\mathbf{u}, \boldsymbol{\mu})$ является, как уже говорилось, моделью оператора эволюции ДС, которая, вообще говоря, при любых значениях параметров $\boldsymbol{\mu}$ не совпадает с «истинным» оператором эволюции. Отметим, в связи с этим, два существенных достоинства байесова подхода. Во-первых, неизбежный в обсуждаемом случае дефект модели может быть учтен в виде дополнительного слагаемого в правой части второго из соотношений (2); это дает возможность оценить его величину и, тем самым, оценить качество модели. Во-вторых, знание апостериорной условной плотности вероятности обеспечивает правильный выбор ценовой функции (учитывающей, в том числе, априорную информацию о ДС, если таковая имеется) как для оценки качества модели в случае, когда ДС не известна, так и для отыскания по измеренному ВР наиболее вероятных значений параметров и латентных переменных известной ДС.

К сожалению, применение байесова подхода к реконструкции ДС по зашумленным экспериментальным данным встречает, как обсуждалось в [3, 7], принципиальные трудности.

В случае, когда динамический шум существен, проблема связана с большим числом аргументов у плотности вероятности $p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$, определяемых, прежде всего, протяженностью (числом отсчетов T) измеренного ВР. Увеличение длины ВР, желательное с точки зрения снижения эффективного уровня шума измерений, приводит к практической невозможности использования $p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$ для расчета требуемых вероятностных характеристик. Отметим, что уже для простых (низкоразмерных) ДС оказываются чрезвычайно ресурсоемкими даже наиболее быстрые численные алгоритмы, основанные на МСМС (*Markov Chain Monte Carlo*) методе [9]. В случае, когда динамический шум мал, так что распределение случайной величины $\boldsymbol{\eta}$ можно считать δ -функцией, возникает трудность другого рода. В этом случае выражение (4) преобразуется к виду

$$p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu}) = \prod_{t=0}^{T-1} w_{\xi}(\mathbf{x}_t - \mathbf{f}'(\mathbf{u}_0, \boldsymbol{\mu})). \quad (5)$$

Здесь \mathbf{u}_0 — значение латентной переменной в начальный момент времени, $\mathbf{f}'(\cdot)$ — t -кратная (последовательно примененная t раз) функция последования, $\mathbf{f}^0(\mathbf{u}_0, \boldsymbol{\mu}) := \mathbf{u}_0$. Видно, что «полная динамичность» системы уменьшает количество латентных переменных до одной. Казалось бы, данное обстоятельство кардинально облегчает (по крайней мере, технически) построение апостериорной плотности вероятности параметров модели

$$p(\boldsymbol{\mu} | \mathbf{x}) = \int p(\mathbf{u}_0, \boldsymbol{\mu} | \mathbf{x}) d\mathbf{u}_0, p(\mathbf{u}_0, \boldsymbol{\mu} | \mathbf{x}) \propto p(\mathbf{u}_0, \boldsymbol{\mu}) \times p(\mathbf{x} | \mathbf{u}_0, \boldsymbol{\mu}) \quad (6)$$

и ее моментов, а также существенно упрощает отыскание (с помощью следующей из (5), (6) ценовой функции) наиболее вероятных значений параметров и *единственной* латентной переменной. Можно, однако, понять, что для хаотического ВР выражение (6) для значений T , больших некоторых, станет слишком сложным как для применения метода Монте-Карло, так и для отыскания наиболее вероятных значений параметров и начальных условий. Это связано с тем, что из-за экспоненциально быстрого «разбегания» первоначально близких фазовых траекторий рост T приводит к уменьшению характерного размера тех областей значений латентной переменной и параметров модели, которые обеспечивают пребывание модельной фазовой траектории в задаваемой уровнем шума окрестности траектории, реконструированной по исходному зашумленному ВР. (Заметим, что сама реконструированная траектория является, вследствие зашумленности исходного ВР, «следающей» — «shadowing», см. [10] — по отношению к «истинной» фазовой траектории, т. е. расположена в определяемой дисперсией шума окрестности истинной траектории ДС). В соответствии с этим, правдоподобие (6), как функция своих аргументов, приобретает все более многомодальный («изрезанный») вид, становясь непригодной для практического использования.

В данной статье излагается модификация байесова подхода, позволяющая преодолеть вторую из указанных проблем: несовместимость статистического подхода с динамичностью изучаемой системы. Способ решения первой проблемы, насколько нам известно, пока не предложен. В следующем разделе поясняется идея модификации и выводится выражение для модифицированной апостериорной плотности вероятности $p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$. Далее модифицированный подход применяется для отыскания значений параметров известной ДС по сгенерированному ею зашумленному ВР. В четвертом разделе описывается применение модифицированного подхода для классификации режимов поведения ДС по очень короткому и сильно зашумленному ВР. В заключении резюмируются рассмотренные в статье вопросы и обсуждаются проблемы, которые необходимо решить для эффективного применения байесова подхода в более сложных ситуациях, в частности для построения прогноза качественного поведения высокоразмерных ДС по зашумленным хаотическим ВР.

2. Модификация байесова подхода для динамических систем: кусочно-динамическая реконструкция

Итак, рассмотрим ситуацию, когда динамическим шумом можно пренебречь. Идея модификации «идеологически верного» в таком случае выражения (5) для априорной плотности условной вероятности состоит в следующем. Представим исходный ВР «длины» T в виде последовательности не перекрывающихся сегментов одинаковой протяженности w и потребуем, чтобы одна и та же модель $\mathbf{f}(\mathbf{u}, \boldsymbol{\mu})$ правильно воспроизводила эволюцию ДС на каждом из $[T/w]$ сегментов. Выражение (5) для априорной плотности вероятности должно быть теперь переписано следующим образом:

$$p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu}) = \prod_{l=0}^{[T/w]-1} \prod_{j=0}^{w-1} w_{\xi}(\mathbf{x}_{l \times w + j} - \mathbf{f}^j(\mathbf{u}_{l \times w}, \boldsymbol{\mu})). \quad (7)$$

Здесь индексы l и j нумеруют последовательные сегменты и отсчеты внутри каждого сегмента соответственно. Очевидно, что выражение (7) имеет смысл при $w \in [1, T/2]$: для $w \in (T/2, T]$ разделение ВР на сегменты одинаковой протяженности невозможно¹.

Переход от (5) к (7) увеличивает число латентных переменных (от одной до $[T/w]$), однако разумный выбор длины сегмента w позволяет избежать чрезмерной изрезанности распределения $p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu})$ при $T \rightarrow \infty$. Ясно, что в случае хаотического ВР «предельно-разумную»

¹ Формально при $w = T$ выражение (7) переходит в «классическое» выражение (5).

длину сегмента можно оценить обратной величиной старшего показателя Ляпунова, рассчитанной по восстановленной фазовой траектории. Нетрудно также понять, что чрезмерное уменьшение w приведет к снижению точности определения значений латентных переменных, а следовательно, и параметров модели.

Отметим, что следующая из (7) ценовая функция в случае гауссова распределения шума измерений $w_\xi = N(0, \epsilon^2)$ точно соответствует ценовой функции, отвечающей алгоритму «многократной стрельбы» («multiple shooting») [11]. Укажем также, что в недавней работе [12] предложено отыскивать оптимальный набор параметров μ независимо для каждого из сегментов; соответствующая ценовая функция (для $w_\xi = N(0, \epsilon^2)$) получается из (7) после замены $\mu \rightarrow \mu_t$. Нетрудно понять, однако, что такая коррекция выражения (7) соответствует предположению о наличии весьма специфического внешнего воздействия на ДС, изменяющего (случайным образом) параметры системы в процессе генерации ВР, причем распределение такого «параметрического» динамического шума определяется как самой ДС (функцией $f(\mathbf{u}, \mu)$), так и распределением шума измерений w_ξ .

Достаточно очевидный недостаток выражения (7) — его несимметрия по отношению к латентным переменным, зашумленные значения которых формируют ВР: величина плотности вероятности (7) зависит лишь от $\lfloor T/w \rfloor$ (из T) латентных переменных. Для проведения симметризации заметим, во-первых, что используемая в (7) выборка латентных переменных $\{\mathbf{u}_{t \times w}\}_{t=0}^{\lfloor T/w \rfloor - 1}$ однозначно задается выбором (зашумленного) состояния ДС \mathbf{x}_0 в качестве начального. Во-вторых, ясно, что в случае стационарного и достаточно протяженного ВР плотность вероятности (7) не должна зависеть от того, какое из (зашумленных) состояний \mathbf{x}_t , ($t = 0, \dots, (w - 1)$) будет начальным. Данные соображения позволяют записать априорную плотность вероятности как среднее геометрическое w выражений (7), отличающихся выбором начального состояния. Для сегментов протяженностью $w \in [1, T/2]$ с точностью до нормировки получаем:

$$p(\mathbf{x} | \mathbf{u}, \mu) \propto \left[\prod_{i=1}^{T-w+1} \prod_{j=0}^{w-1} w_\xi(\mathbf{x}_{t+j-1} - \mathbf{f}^j(\mathbf{u}_{t-1}, \mu)) \right]^{1/w}. \quad (8)$$

Заметим, что, в отличие от (7), симметризованное выражение для модифицированной плотности вероятности может быть записано и для сегментов протяженностью $w \in (T/2, T]$. Необходимо заменить в (8) показатель степени $w^{-1} \rightarrow (T - w)^{-1}$ и данная замена превращает (8) в классическое байесово выражение для априорной условной плотности

вероятности, но записанное для ВР протяженностью $w \leq T$ (см. выражение (5)). Необходимость использовать укороченный ВР обусловлена, как уже пояснялось, экспоненциально быстрым разбеганием фазовых траекторий, формирующих хаотический аттрактор.

Выражение (8) для априорной плотности вероятности является ключевым для статистического подхода к реконструкции ДС по зашумленным ВР. Апостериорные плотности вероятности $p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$ и $p(\boldsymbol{\mu} | \mathbf{x})$ выражаются через (8) очевидным (аналогичным (1), (6)) образом:

$$p(\boldsymbol{\mu} | \mathbf{x}) = \int p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x}) d\mathbf{u}; \quad p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x}) \propto p(\mathbf{u}, \boldsymbol{\mu}) \times p(\mathbf{x} | \mathbf{u}, \boldsymbol{\mu}). \quad (6^*)$$

Априорная плотность вероятности (8) максимально возможным в рамках байесова подхода образом учитывает динамичность реконструируемой системы. Мерой реконструкции динамичности (и фильтрации шума измерений) является длина сегмента w : чем больше w , тем менее (8) отличается от плотности вероятности (5), которая, как отмечалось, формально включает всю информацию о ДС, содержащуюся в исходном ВР. Максимально возможное для реконструкции ДС по конкретному ВР значение w определяется уровнем шума, дефектом выбранной модели и используемым *численным алгоритмом* построения условной плотности вероятности параметров модели. Примеры зависимости w от перечисленных факторов приведены далее в разд. 3 и 4.

Обратим внимание, что ценой адаптации байесова подхода к реконструкции ДС является увеличение числа латентных переменных: в выражении (8) их $(T - w) \leq T/2$. Таким образом, проводя реконструкцию по достаточно протяженным ВР, содержащим *шум измерений*, мы можем столкнуться с проблемой (чрезмерно большого числа аргументов у апостериорной плотности вероятности $p(\mathbf{u}, \boldsymbol{\mu} | \mathbf{x})$), «обязательной» для случая с существенным *динамическим шумом* (см. соотношение (4)). Способ решения данной проблемы, как уже говорилось, требует специального исследования. В двух следующих разделах мы продемонстрируем эффективность модифицированного байесова подхода на примере решения задач, в которых данная проблема не возникает.

3. Восстановление значений параметров известной ДС по зашумленному ВР: сравнение модифицированного и «классического» подходов

Неприменимость «классического» байесова подхода к реконструкции ДС по хаотическим ВР (см. разд. 1) была весьма убедительно продемонстрирована в [3] на примере логистического отображения — одной из простейших ДС, демонстрирующих хаотическое поведение. Опе-

ратор эволюции данной ДС представляет собой функцию последования 1-го порядка:

$$u_{n+1} = 1 - au_n^2. \quad (9)$$

В зависимости от значения единственного параметра a система может демонстрировать как регулярные (периодические), так и хаотический режимы; переход к динамическому хаосу при изменении a происходит через каскад удвоений периода. В [3] решалась задача восстановления значения параметра a по хаотическому ВР, в котором присутствует аддитивный шум измерений ξ :

$$\begin{cases} x_n = u_n + \xi_n; \\ u_n = 1 - au_{n-1}^2. \end{cases} \quad (10)$$

Предполагалось, что шум δ -коррелирован и имеет гауссово распределение с известной дисперсией σ_ξ :

$$w_\xi(\xi) \propto \exp\left(-\sum_i \xi_i^2 / 2\sigma_\xi^2\right).$$

Было показано, что уже при не слишком высоком уровне шума² $\sigma_\xi = 0,1$ и весьма коротком ВР ($T = 70$) при решении данной задачи в классической байесовой постановке (6) характерный масштаб изрезанности условной плотности вероятности $p(u_0, a | \mathbf{x})$ по латентной переменной u_0 становился меньше машинной точности: для отыскания правильного значения параметра среднеквадратичный разброс, характеризующий ширину априорного распределения $p(u_0, \alpha)$, должен быть $\sigma_{u_0} \leq 10^{-17}$. Другими словами, классический байесов подход требует недостижимо точной информации о начальном состоянии ДС. Покажем, что использование модифицированного подхода (выражения (8)) позволяет производить оценку параметра без какой-либо априорной информации о значении латентной переменной.

На рис. 1, a приведена апостериорная плотность вероятности $p(a | \mathbf{x})$ (см. (8), (6*)) для различных длин сегмента w . Как и в [3], исходный ВР был сгенерирован системой (10) при $a = 1,85$; $u_0 = 0,3$, случайная величина ξ имела дисперсию $\sigma_\xi = 0,1$, протяженность ВР равнялась $T = 70$. Видно (см. также рис. 1, б), что в соответствии с качественными соображениями, приведенными в разделе 1, с ростом w уменьшается как систематическое смещение распределения относительно пра-

² Для ВР, рассмотренного в [3] (сгенерированного системой (10) при $a = 1,85$; $u_0 = 0,3$), $\sigma_\xi = 0,1$ примерно соответствует отношению сигнал/шум, равному 6:1.

вильного значения параметра, так и ширина распределения, определяющая ошибки. В результате уже при длине сегмента $w = 3$ правильное значение параметра попадает в 65%-ный доверительный интервал, а при $w = 7$ края доверительного интервала отличаются от правильного значения менее, чем на 1%. Предельно допустимая для данного конкретного ВР длина сегмента оказалась равной $w = 9$. Она определяется, как уже отмечалось, возможностями численного алгоритма. В данном конкретном случае при $w > 9$ шаг интегрирования функции $p(u_0, a | x)$ по латентной переменной (см. (6*)) становился больше масштаба изрезанности этой функции. Можно, таким образом, заключить, что модифицированный подход успешно решает задачу отыскания правильных значений параметров известной ДС по зашумленному хаотическому ВР в ситуации, когда классический байесов алгоритм оказывается совершенно неэффективным.

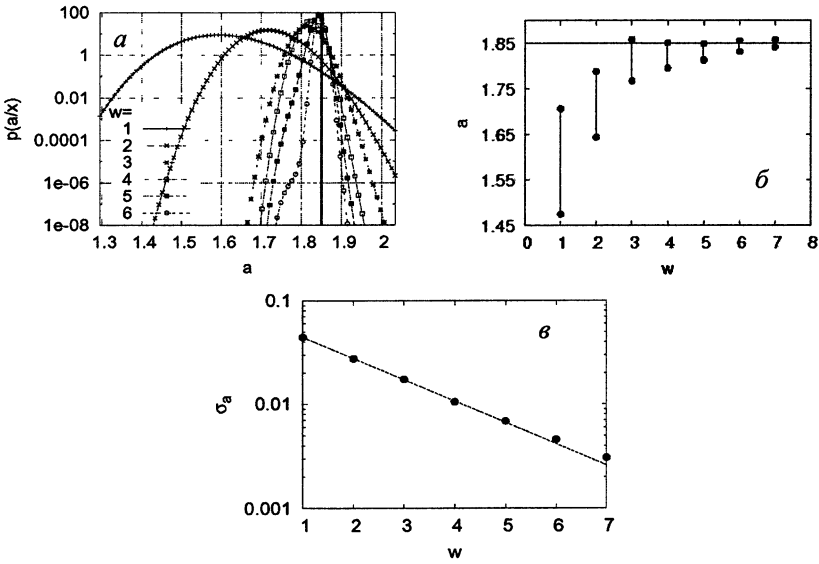


Рис. 1. Восстановление параметра a логистического отображения: *a* — зависимость апостериорной плотности вероятности от параметра a для различных w («правильное» значение a показано вертикальной линией); *б* — доверительные интервалы восстановления в зависимости от w (горизонтальной линией показано «правильное» значение a); *в* — зависимость σ_a от w : точками показан результат восстановления, линией — теоретическая зависимость.

В заключение данного раздела заметим, что для не слишком сильно зашумленного хаотического ВР можно показать, что ширина σ_a рас-

предела (8) связана со старшим показателем Ляпунова λ и длиной сегмента w соотношением:

$$\sigma_a \propto e^{-\lambda w}. \quad (11)$$

Рисунок 1, в демонстрирует, что теоретическая зависимость (11) σ_a от w (прямая линия на рис. 1, в; значение λ вычислялось по временному ряду при помощи программного пакета TISEAN [13]) находится в хорошем соответствии с результатами восстановления (точки на рис. 1, в).

4. Классификация режимов поведения известной ДС по коротким зашумленным ВР

Полученные в предыдущем разделе результаты позволяют поставить важную с практической точки зрения задачу классификации режимов поведения ДС. В качестве примера рассмотрим очень короткие ($T = 20$) ВР, сгенерированные более сложной ДС — отображением Эно (Enon). Оператор эволюции данной ДС представляет собой функцию последования 2-го порядка:

$$\begin{cases} u_{n+1} = v_n; \\ v_{n+1} = 1 - av_n^2 - bu_n. \end{cases} \quad (12)$$

В зависимости от значений параметров a и b система (12) демонстрирует различные регулярные и хаотический режимы. Области параметров, отвечающие различным режимам поведения, выделены оттенками серого на бифуркационной диаграмме (рис. 2).

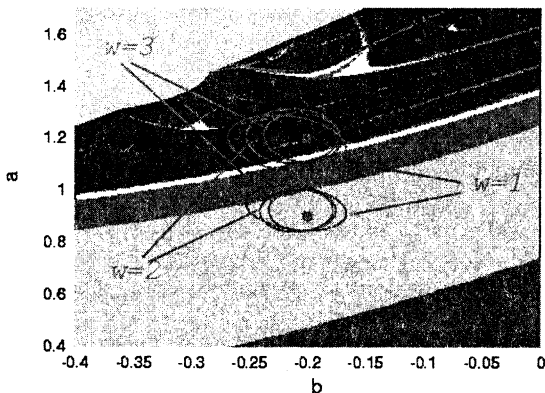


Рис. 2. Бифуркационная диаграмма отображения Эно. Оттенками серого цвета обозначены режимы, соответствующие (снизу вверх): устойчивому состоянию равновесия, двухпериодическому, четырехпериодическому, хаотическому режимам, области глобальной неустойчивости модели. Серыми точками показаны значения параметров, при которых генерировались ВР. Эллипсы являются границами доверительных интервалов восстановления параметров для $w = 1, 2$ и 3 .

В качестве «измеряемой» рассматривалась величина x , представляющая собой аддитивно зашумленную (латентную) переменную u : $x_n = u_n + \xi_n$. Шум измерений ξ был δ -коррелирован и имел гауссово распределение с дисперсией σ_ξ . Анализируются ВР, сгенерированные при различных значениях параметров a и b . В отсутствие шума ВР, отвечающие различным периодическим и хаотическому режимам, легко различимы. На рисунке 3, $a, б$ показаны для примера ВР, порожденные отображением (12), соответствующие двухпериодическому ($a = 0,85, b = -0,2$) и хаотическому ($a = 1,15, b = -0,2$) режимам. При добавлении даже сравнительно небольшого шума (см. рисунок 3, $в, г, \sigma_\xi = 0,3$) различить динамические режимы становится невозможно. Покажем, как модифицированный байесов подход решает эту проблему.

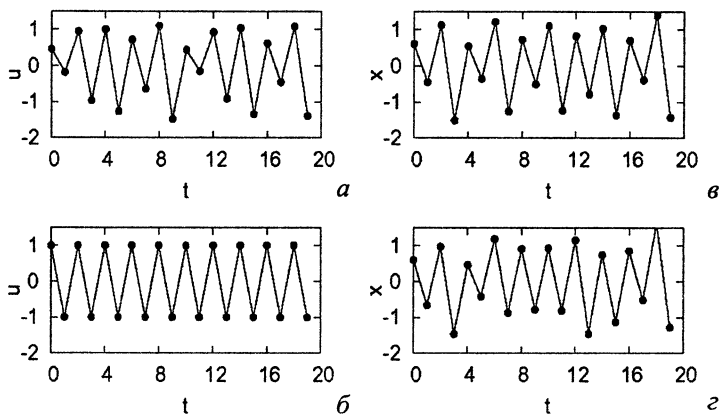


Рис. 3. Хаотический и периодический ВР: $a, б$ — отвечающие значениям параметров, отмеченным на рис. 2 точками; $в, г$ — с шумом $\sigma_\xi = 0,25$.

Задача классификации режима поведения ДС, породившей исходный зашумленный ВР, основана на исследовании статистического ансамбля незашумленных ВР, отвечающего статистическому ансамблю параметров, распределенных в соответствии с апостериорной плотностью вероятности (8). Подсчитывается, сколько элементов ансамбля незашумленных ВР соответствуют тому или иному динамическому режиму, затем вычисляется вероятность каждого из зарегистрированных режимов. На рисунке 4 показана зависимость вероятности «правильного» режима от уровня шума для хаотического и одного из периодических зашумленных ВР, сгенерированных отображением (12). Видно, что вплоть до уровня шума 25% в случае хаотического ВР и 50% в случае периодического модифицированный алгоритм выделяет правильный режим с вероятностью не менее 0,5.

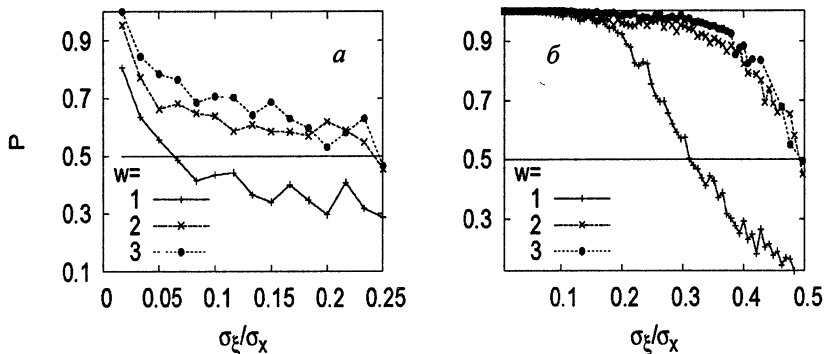


Рис. 4. Вероятность распознавания «правильного» динамического режима в зависимости от уровня шума для различных w : *a* — хаотический ВР; *б* — периодический ВР.

Отметим, что близость к единице вероятности наиболее вероятного режима является удобной количественной характеристикой качества классификации. Обсудим зависимость данной характеристики от длины сегмента w . Из обсуждавшегося выше ясно, что в случае хаотического режима с увеличением w эта вероятность будет, вообще говоря, приближаться к единице: из-за уменьшения с ростом w ширины апостериорной плотности распределения параметров вероятности ошибочных режимов будут почти всегда уменьшаться. Нетрудно понять, что этот вывод справедлив и в отношении зависимости от w вероятности «правильного» регулярного режима, хотя зависимость ширины распределения параметров от w становится несколько иной. Именно для регулярных режимов данная зависимость не является экспоненциально убывающей, как в случае хаотического режима (см. выражение (11)), а насыщается с ростом w . Этот факт является следствием независимости установившегося поведения системы, функционирующей в устойчивом периодическом режиме, от начальных значений динамических переменных³. Тем не менее, уменьшение ширины $p(\mu | x)$ с ростом w обеспечивает и для регулярных режимов поведения увеличение вероятности правильного режима при увеличении длины сегмента. Результаты восстановления системы (10), приведенные на рис. 2 и 4, подтверждают сказанное. На рисунке 2 выделены доверительные области восстановления параметров системы (10) по соответствующим (зашумленным) ВР. Видно, что как для хаотического, так и для регулярных режимов поведения качество классификации улучшается при увеличении длины сегмента.

³ Другими словами, слагаемые в (8), содержащие f^j достаточно высоких порядков, слабо зависят от конкретных значений латентных переменных.

Рисунок 5 демонстрирует насыщающуюся с ростом w зависимость величины доверительной области восстановления регулярного (двухпериодического) режима от длины сегмента.

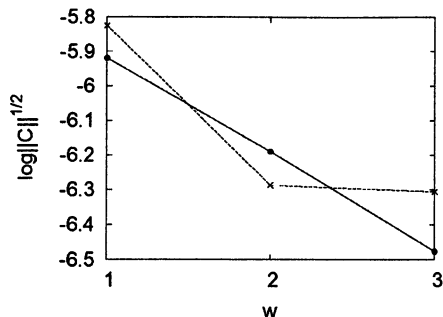


Рис. 5. Зависимость характерного размера доверительного интервала от длины сегмента в случае хаотического (линия с точками) и периодического (линия с крестиками) ВР.

4. Заключительные замечания

В статье обсуждался статистический байесов подход к реконструкции ДС по сгенерированным ими ВР. Именно такой подход наиболее, на наш взгляд, адекватен задаче реконструкции «реальных» ДС, потому что, во-первых, они подвержены нерегулярным внешним воздействиям в процессе генерации сигнала и, во-вторых, сам сигнал испытывает случайные возмущения как при распространении, так и во время регистрации. Статистический подход имеет целью не только реконструкцию оператора эволюции ДС, но и построение плотности вероятности параметров этого оператора, рассматриваемых как случайные величины. Одна из двух трудностей в реализации такого подхода возникает при реконструкции по зашумленным хаотическим рядам: экспоненциально быстрое разбегание фазовых траекторий на хаотическом аттракторе имеет следствием нарастающую с увеличением продолжительности ВР изрезанность плотности вероятности параметров. В результате уже для сравнительно коротких ВР построение плотности вероятности параметров становится невозможным. Нами предложена модификация классического байесова подхода, позволяющая преодолеть указанную трудность. В статье эффективность модифицированного подхода продемонстрирована на примере решения задач отыскания неизвестных значений параметров известной ДС и классификации режимов поведения ДС (как регулярных, так и хаотических) по короткому зашумленному ВР.

Вторая трудность, возникающая при использовании статистического подхода к реконструкции ДС, состоит в отсутствии в настоящее время «быстрого» алгоритма построения произвольной функции распределения (плотности вероятности), зависящей от очень большого числа аргументов.

В случае зашумленных ВР это число определяется, в первую очередь, протяженностью ВР (числом латентных переменных). В разделе 2 было продемонстрировано, что в рамках модифицированного байесова подхода сохранение информации о «динамичности» системы достигается путем увеличения числа латентных переменных до почти предельного для используемого ВР значения. Таким образом, предложенное нами решение первой проблемы делает еще более насыщенным создание эффективного алгоритма расчета многомерных функций распределения. Создание такого алгоритма мы надеемся завершить в ближайшее время.

В заключение отметим еще раз универсальность предлагаемого подхода к реконструкции ДС. Он, очевидно, применим к любым задачам извлечения информации из ВР, порожденных хотя бы «частично динамическими» системами. Чрезвычайно привлекательной задачей такого рода является прогноз качественного поведения *неизвестной* ДС по хаотическому ВР. Общая постановка этой задачи («прогноз бифуркаций») подробно обсуждалась в лекции, прочитанной на предыдущей школе по нелинейным волнам [7]. Напомним, что, во-первых, успех в решении данной задачи был достигнут за счет «превращения» ДС в стохастическую: неизбежный (в силу неизвестности ДС) дефект модели описывался в [7] как аддитивный динамический «шум». Во-вторых, использованный в [7] алгоритм прогноза бифуркаций представляет собой частный случай предложенного в данной лекции модифицированного байесова подхода, соответствующий выбору предельно короткого сегмента $w = 1$. Обсуждавшиеся в разд. 3 и 4 данной статьи результаты позволяют рассчитывать на повышение качества прогноза бифуркаций при увеличении w до максимально возможной величины. Наконец, в-третьих, пример построения прогноза бифуркаций, предьявленный в [7], и все другие опубликованные к настоящему моменты примеры (см. [14—17]), относятся к случаю «идеальных» (свободных от измерительного шума) ВР, порожденных ДС низкой размерности. Отсутствие продвижения в данной задаче связано опять-таки с необходимостью создания эффективного алгоритма построения многомерной плотности вероятности, произвольным образом зависящей от своих аргументов.

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

Литература

1. Anishchenko V. S., Paylov A. P. Global reconstruction in application to multichannel communication // Phys. Rev. E. 1998. V. 57, № 2. P. 2455—2457.
2. Бутковский О. Я., Кравцов Ю. А., Логунов М. Ю. Анализ погрешностей восстановления параметров нелинейного отображения по зашумленным хаотическим временным рядам // Изв. вузов. Радиофизика. 2002. Т. 45, № 1. С. 55—66.

3. Judd K. Chaotic-time-series reconstruction by the Bayesian paradigm: Right results by wrong methods // *Phys. Rev. E*. 2003. V. 67. 026212.
4. Abarbanel H. D. I. Analysis of observed chaotic data // N. Y.: Springer-Verlag, 1997. 272 p.
5. Kantz H., Schreiber T. Nonlinear time series analysis // *Nonlinear Science Series*. № 7. Cambridge: Cambridge University Press, 1997.
6. Макаренко Н. Г. Реконструкция динамических систем по хаотическим временным рядам // Настоящий сборник. С. 398—410.
7. Мольков Я. И., Фейгин А. М. Прогноз качественного поведения динамических систем по зашумленным временным рядам // *Нелинейные волны'2002: Сб. лекций* / Ред. А. В. Гапонов-Грехов, В. И. Некокоркин. Н. Новгород: ИПФ РАН, 2003. С. 34—52.
8. Meyer R., Christesen N. Bayesian reconstruction of chaotic dynamical systems // *Phys. Rev. E*. 2000. V. 62. P. 3535—3542.
9. Gilks W. R., Richardson S., Spiegelhalter D. J. Markov Chain Monte Carlo in practice. London: Chapman and Hall, 1996.
10. Farmer J. D., Sidorowich J. J. Optimal shadowing and noise reduction // *Physica D*. 1991. V. 47. P. 373—392.
11. Bock H. G. Recent advances in parameter identification techniques for O.D.E. // *Numerical Treatment of Inverse Problems in Differential and Integral Equations* / Eds. P. Deuffhard, E. Hairer. Birkhauser, Basel, 1983. P. 95—121.
12. Pisarenko V. S., Sornette D. Statistical methods of parameter estimation for deterministically chaotic time series // *Phys. Rev. E*. 2004. V. 69. 036122.
13. Hegger R., Kantz H., Schreiber T. Practical implementation of nonlinear time series methods: The TISEAN package // *Chaos*. 1999. V. 9. P. 413.
14. Фейгин А. М., Мольков Я. И., Мухин Д. Н., Лоскутов Е. М. Прогноз бифуркаций слабонеавтономных динамических систем на основе наблюдаемых временных рядов: Препринт ИПФ РАН №508. Н. Новгород, 1999. 53 с.
15. Фейгин А. М., Мольков Я. И., Мухин Д. Н., Лоскутов Е. М. Прогноз качественного поведения динамической системы по хаотическому временному ряду // *Изв. вузов. Радиофизика*. 2001. Т. 44, № 5—6. С. 376—399.
16. Feigin A. M., Molkov Y. I., Mukhin D. N., Loskutov E. M. Investigation of nonlinear dynamical properties by the observed complex behaviour as a basis for construction of the dynamical models of atmospheric photochemical systems // *Faraday Discussions*. 2002. V. 120. P. 105—123.
17. Feigin A. M. Prognosis of qualitative behaviour of a dynamic system by the noisy chaotic time series // *Topical Problems of Nonlinear Wave Physics: Proc. Int. Symp. Nizhny Novgorod*, 2003. P. 91—92.

НЕКОТОРЫЕ МАТЕМАТИЧЕСКИЕ ПРОБЛЕМЫ КЛАССИЧЕСКОЙ СИНХРОНИЗАЦИИ

Д. В. Тураев, А. Л. Шильников, Л. П. Шильников

1. Введение. Гомоклиническая петля с периодической внешней силой

Основными объектами классической теории синхронизации являются задача о воздействии внешнего периодического сигнала на автоколебательную систему и задача о взаимодействии двух автоколебательных систем. Цель, которая обычно здесь ставится, состоит в выделении в пространстве параметров областей существования грубых устойчивых периодических движений — областей синхронизации — и описании динамических явлений, возникающих при переходе через границы этих областей. Сразу отметим, что возникающие здесь динамические системы многомерны (размерность фазового пространства три и выше). Поэтому понятно, что изучение конкретных моделей требует использования широкого арсенала методов современной качественной теории.

Относительно просто рассматриваемые задачи решаются в случае малой внешней силы или малого взаимодействия, которые мы будем считать имеющими порядок μ . Здесь легко строится отображение Пуанкаре T_μ , которое в подходящих (так называемых нормальных) координатах представляется в виде

$$\begin{aligned}\bar{x} &= f(x, \theta, \mu), \\ \bar{\theta} &= \theta + \omega + g_0(x, \theta, \mu) \equiv g(x, \theta, \mu) \pmod{2\pi},\end{aligned}$$

где $x \in R^n$, $\theta \in S^1$, f и g_0 — периодические по θ функции с периодом 2π . При этом $f(0, \theta, 0) \equiv 0$, $g_0(0, \theta, 0) \equiv 0$ и $\|f'_x\| < 1$ всюду в некотором кольце $K_\delta = \{(x, \theta), \theta \in S^1, \|x\| \leq \delta\}$ при всех малых μ .

Согласно принципу кольца [3, 27], у отображения T_μ при всех достаточно малых μ будет существовать устойчивая гладкая инвариантная замкнутая кривая вида $x = h(\theta, \mu)$, к которой стремятся все траектории из K_δ . В фазовом пространстве инвариантной кривой отображения Пуанкаре отвечает двумерный устойчивый инвариантный тор. Поведение траекторий на торе определяется отображением инвариантной кривой, которое является диффеоморфизмом окружности вида

$$\bar{\theta} = \theta + \omega + \mu\alpha(\theta, \mu) \pmod{2\pi},$$

где $\alpha(\theta)$ — периодическая функция с периодом 2π . Здесь, как правило, к каждому ω , соизмеримому с 2π , при $\mu = 0$ будут примыкать области синхронизации на плоскости параметров (ω, μ) , также называемые языками Арнольда. Они схематически показаны на рис. 1.

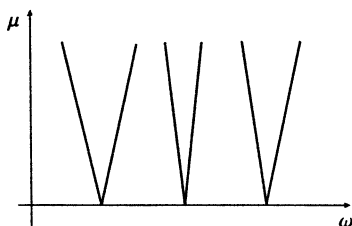


Рис. 1. Языки Арнольда на плоскости параметров (ω, μ)

Внутри каждого языка Арнольда число вращения Пуанкаре

$$r = \frac{1}{2\pi} \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^N (\theta_{n+1} - \theta_n)$$

рационально и постоянно, устойчивые периодические движения с данным числом вращения на торе отвечают режиму синхронизации.

При переходе через границы областей синхронизации структура множества траекторий на торе должна изменяться. При этом в силу непрерывной зависимости числа вращения от параметров должны реализовываться и иррациональные числа вращения, которым отвечает уже не режим синхронизации, а квазипериодический режим (биения). Заметим, что в прикладных исследованиях периодические режимы с большими периодами в узких языках Арнольда обычно не воспринимаются как режимы синхронизации, поскольку они практически неотличимы от квазипериодических.

Данная картина имеет место для классического уравнения Ван-дер-Поля

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + \omega_0^2 x = \mu A \sin \omega t \quad (1)$$

в квазилинейном случае, т. е. при $\mu \ll 1$ и $A^2 < \frac{4}{27}$. Это следует из известных исследований Ван-дер-Поля [35], Андронова и Витта [7], Крылова и Боголюбова [13], а также работы Афраймовича и Шильникова [1], объясняющей сохранение тора при исчезновении седлоузлов периодического движения в случаях главных резонансов.

Основная задача, которая будет обсуждаться в настоящей работе, состоит в изучении динамических эффектов, возникающих при

увеличении μ при выходе из области существования инвариантного тора. Поскольку исследования этого вопроса оказали значительное влияние на современное точное естествознание, мы более подробно остановимся на истории проблемы.

В 1927 году Ван-дер-Поль и Ван-дер-Марк опубликовали свои результаты по экспериментальному исследованию генератора с неоновой лампой, находящегося под воздействием периодического сигнала. Основные свои выводы они резюмировали следующим текстом [36]: «Часто в телефонных трубках перед скачком частоты к соседнему, меньшему значению слышится переходный шум. Однако это второстепенное явление, главный эффект заключается в регулярном делении частоты». А это значило, что они наблюдали сосуществование устойчивых периодических режимов с различными периодами (при этом очень большими, в эксперименте они были в 100—200 раз больше, чем период внешней силы), а также, говоря современным языком, сложное поведение траекторий (хоть и оцененное как второстепенный эффект). Последнее указывает если не на существование в фазовом пространстве странного аттрактора, то на наличие нетривиального множества, состоящего из седловых траекторий, ответственного за сложный переходный процесс.

Эти эксперименты не оставили без внимания Картрайт и Литлвуд. В 1945 году они в краткой форме излагают [8] удивительные результаты своего анализа уравнения Ван-дер-Поля (1) при $\mu \gg 1$. Так, они указывают на наличие двух сортов интервалов значений A в отрезке $(0, 1/3)$: в одних интервалах динамика простая, а в других, наряду с существованием двух устойчивых периодических движений с разными периодами, есть также нетривиальное неблуждающее множество, состоящее из неустойчивых траекторий и допускающее описание на языке символической динамики из двух символов. Тем самым впервые было установлено, что в трехмерных моделях неконсервативного происхождения неблуждающее множество может содержать счетное множество периодических движений и континуум непериодических траекторий. Развернутое изложение этих результатов было, однако, дано Литлвудом только позднее, в 1957 году [16, 17].

В 1949 году Левинсон [15] дал объяснение этим результатам на примере уравнения

$$\varepsilon \ddot{x} + \dot{x} \operatorname{sign}(x^2 - 1) + \varepsilon x = A \sin t, \quad (2)$$

где $\varepsilon \ll 1$. Поскольку уравнение (2) кусочно-линейно, изучение поведения траекторий значительно упростилось и могло быть изложено достаточно прозрачно.

Идея типичности сложного поведения траекторий для широко класса нелинейных уравнений была высказана Литлвудом в [16] практически в явном виде; здесь любопытно также отметить, что уже в первой работе [8] содержалось пророческое высказывание о том, что при значениях A из интервалов со сложным поведением траекторий, описываемых уравнениями вандерполевского типа, системы могут быть топологически эквивалентны, т. е. обнаруженный хаос виделся Картрайт и Литлвуду как грубое явление. После того как Левинсон указал на эти результаты Смейлу, тот обнаружил, что они допускают простую геометрическую интерпретацию, по крайней мере на качественном уровне, и на основе этого анализа построил в 1961 году диффеоморфизм подковы с нетривиальным гиперболическим множеством, топологически сопряженным схеме Бернулли из двух символов, с чего, собственно, и началась современная теория динамического хаоса.

В 70—80-х годах изучение уравнений вандерполевского типа продолжили Осипов [23] и Леви [14]. Для нелинейностей близких к кусочно-линейным ими было дано полное описание неблуждающих множеств в отмеченных интервалах и доказана их гиперболичность.

В указанных работах рассмотрение велось в конечном числе сегментов значений A . Оставшиеся интервалы в сумме составляли множество, мера которого стремилась к нулю при $\mu \rightarrow \infty$ (т. е. $\varepsilon \rightarrow 0$). Таким образом, в силу сингулярной природы данной задачи динамическими явлениями, происходящими в интервалах негиперболичности, можно, в принципе, и пренебречь в первом порядке. Заметим, что именно в этих интервалах происходит переход от простого поведения к сложному, сопровождающийся появлением и исчезновением устойчивых периодических движений, возникновением негрубых гомоклинических орбит, влекущих, в свою очередь, существование интервалов негрубости Ньюхауса, и т. д. В моделях, не содержащих малого параметра при старшей производной, все это немедленно требует настоящего анализа.

Естественным примером здесь служит автономная система

$$\dot{x} = X(x, \mu),$$

находящаяся под действием периодического возмущения малой амплитуды порядка μ , у которой при $\mu > 0$ существует устойчивый предельный цикл L_μ , влипающий при $\mu \rightarrow 0$ в гомоклиническую петлю седлового состояния равновесия. Исследованию таких систем был посвящен цикл работ Афраймовича и Шильникова [2, 3, 4]. Ниже

излагается ряд полученных результатов, имеющих принципиальное значение для теории синхронизации.

Для простоты ограничимся случаем двух уравнений. Предположим, что автономная система

$$\begin{aligned}\dot{x} &= \lambda x + F(x, y, \mu), \\ \dot{y} &= \gamma y + G(x, y, \mu)\end{aligned}$$

имеет при всех малых μ состояние равновесия $O(0, 0)$ типа седло, и в нем седловая величина отрицательна:

$$\sigma = \lambda + \gamma < 0. \quad (3)$$

Пусть также седло имеет петлю сепаратрисы при $\mu = 0$ (рис. 2, а).

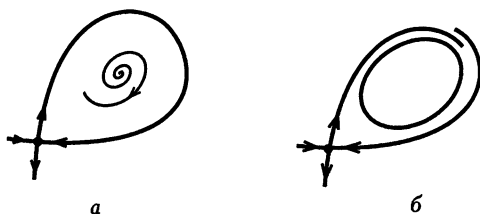


Рис. 2. Бифуркация рождения устойчивого предельного цикла из петли сепаратрисы на плоскости.

Как известно [5, 6], при малых μ из петли может родиться только одно, и притом устойчивое, периодическое движение L_μ . Пусть это имеет место при $\mu > 0$ (рис. 2, б). В общем случае период родившегося цикла будет порядка $|\ln \mu|$. Это роднит рассматриваемую задачу с уравнениями вандерполевского типа, у которых при $A = 0$ период релаксационного цикла имеет порядок $1/\varepsilon$, т. е. также стремится к бесконечности при стремлении малого параметра к нулю.

Относительно возмущенной системы

$$\begin{aligned}\dot{x} &= \lambda x + F(x, y, \mu) + \mu p(x, y, t, \mu), \\ \dot{y} &= \gamma y + G(x, y, \mu) + \mu q(x, y, t, \mu),\end{aligned} \quad (4)$$

где p и q — 2π -периодические функции по t , будем предполагать, что функция Мельникова положительна, т. е. при всех достаточно малых $\mu > 0$ у седлового периодического движения, близкого к $(x = 0, y = 0)$, устойчивое W_μ^s и неустойчивое W_μ^u многообразия не пересекаются (рис. 3).

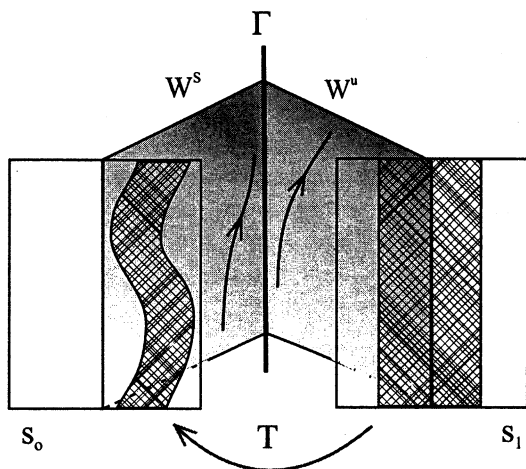


Рис. 3. Отображение Пуанкаре переводит секущую S_1 в плоскости $x = \delta$ к неустойчивому многообразию W^u седлового периодического движения Γ на секущую S_0 (в той же плоскости) к устойчивому многообразию.

Легко видеть, что плоскость $x = \delta$ в пространстве $\{x, y, t\}$ является секущей для системы (4) при малом δ . Соответствующее отображение Пуанкаре T_μ может быть записано в виде, близком к следующему модельному отображению:

$$\begin{aligned} \bar{y} &= [y + \mu(1 + f(\theta))]^\nu, \\ \bar{\theta} &= \theta + \omega - \frac{1}{\gamma} \ln[y + \mu(1 + f(\theta))], \end{aligned} \quad (5)$$

где $\nu = -\frac{\lambda}{\gamma} > 1$, ω — некоторая константа, и $\mu(1 + f(\theta))$ — функция Мельникова с $\langle f(\theta) \rangle = 0$. Правую часть второго уравнения следует вычислять по модулю 2π , т. е. θ — угловая переменная. Ее можно интерпретировать как разность фаз между внешним воздействием и откликом системы; таким образом, устойчивые неподвижные точки этого отображения (т. е. точки, для которых $\bar{\theta} = \theta \bmod 2\pi$) отвечают режиму синхронизации.

Предельное множество отображения T_μ при достаточно малых μ лежит в кольце $K_\mu = \{0 < x < C\mu^\nu, 0 \leq \theta < 2\pi\}$, для некоторого $C > 0$. После замены $y \rightarrow \mu^\nu y$ отображение Пуанкаре примет вид

$$\begin{aligned} \bar{y} &= [1 + f(\theta)]^\nu + \dots, \\ \bar{\theta} &= \theta + \tilde{\omega} - \frac{1}{\gamma} \ln[1 + f(\theta)] + \dots, \end{aligned} \quad (6)$$

где многоточиями обозначены члены, стремящиеся к нулю вместе с производными, а $\tilde{\omega} = \omega - \frac{1}{\gamma} \ln \mu$ стремится к бесконечности при $\mu \rightarrow +0$, т. е. $\tilde{\omega} \bmod 2\pi$ принимает произвольное значение из интервала $[0, 2\pi)$ счетное число раз. Таким образом, динамика отображения Пуанкаре во многом определяется свойствами семейства отображений окружности

$$\bar{\theta} = \theta + \tilde{\omega} + \mathcal{F}(\theta) \bmod 2\pi, \quad (7)$$

где $\mathcal{F}(\theta) = -\frac{1}{\gamma} \ln[1 + f(\theta)]$.

Утверждение [2, 3].

1. В случае, если

$$\frac{1}{\gamma} \frac{f'(\theta)}{1 + f(\theta)} < 1, \quad (8)$$

отображение T_μ имеет устойчивую гладкую инвариантную замкнутую кривую вида $y = h(\theta, \mu)$, которая содержит ω -предельное множество любой траектории из K_μ .

2. Пусть существует отрезок $I = [\theta_1, \theta_2]$, такой, что либо

$$f'(\theta) < 0 \quad (9)$$

всюду на I и

$$\frac{1}{\gamma} \ln \frac{1 + f(\theta_1)}{1 + f(\theta_2)} > 2\pi(m + 1), \quad m \geq 2, \quad (10)$$

либо

$$\frac{1}{\gamma} \frac{f'(\theta)}{1 + f(\theta)} > 2 \quad (11)$$

всюду на I и

$$\frac{1}{\gamma} \ln \frac{1 + f(\theta_2)}{1 + f(\theta_1)} > 2(\theta_2 - \theta_1) + 2\pi(m + 1), \quad m \geq 2. \quad (12)$$

Тогда при достаточно малых $\mu > 0$ отображение T_μ будет иметь гиперболическое множество Σ_μ , сопряженное со схемой Бернулли из m символов.

Так, в случае $f(\theta) = A \sin \theta$, если $A < \frac{\gamma}{\sqrt{1 + \gamma^2}}$, при всех малых μ аттрактором является инвариантная замкнутая гладкая кривая, а при $A > \tanh 3\pi\gamma$ имеет место сложная динамика.

Условия (11) или (9) обеспечивают растяжение по θ в области $\Pi : \theta \in I$ и, следовательно, гиперболичность отображения (5) в этой области (сжатие по переменной y всегда достигается при достаточно малых y за счет того, что $\nu > 1$). Когда выполнены условия (12) или (10), соответственно, образ области Π пересекает ее как минимум m раз (рис. 4, а). Таким образом, мы получаем конструкцию, вполне аналогичную подкове Смейла, и утверждение пункта 2 достигается, скажем, ссылкой на лемму о седловой неподвижной точке в счетном произведении банаховых пространств [26].

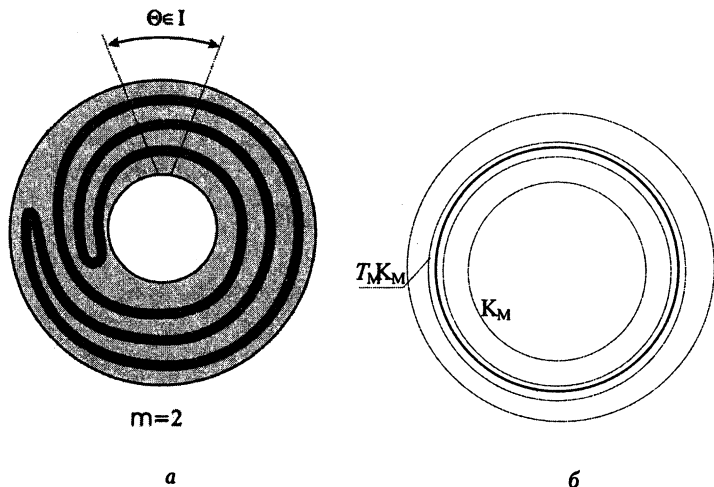


Рис. 4. Образ области Π пересекает ее как минимум m раз (а). Образ кольца K_μ под действием отображения T_μ не имеет складок (б).

В случае выполнения условия (8) образ кольца K_μ под действием отображения T_μ не имеет складок (рис. 4, б) при достаточно малых μ , т. е. также представляет собой кольцо, ограниченное кривыми вида $y = h_\pm(\theta)$. Образ этого кольца также имеет подобный вид и т. д. В результате мы получаем последовательность вложенных колец, и сжатие по переменной y обеспечивает их пересечение по единственной и гладкой замкнутой кривой. То, что эта кривая инвариантна и притягивает все траектории, следует из конструкции; формальное доказательство утверждения (пункта 1) достигается ссылкой на принцип кольца [3] (см. также [27]).

Для детального исследования областей синхронизации мы ограничим себя далее случаем $f(\theta) = A \sin \theta$ (в принципе, точно так же можно рассмотреть случай $f(\theta) = Ag(\theta)$, где $g(\theta)$ — функция

с заданными свойствами). Это позволяет построить весьма разумную бифуркационную диаграмму (рис. 5, 6) на плоскости параметров $(A, -\ln \mu)$ в области $0 \leq A < 1$, $0 < \mu < \mu_0$, где μ_0 достаточно мало.

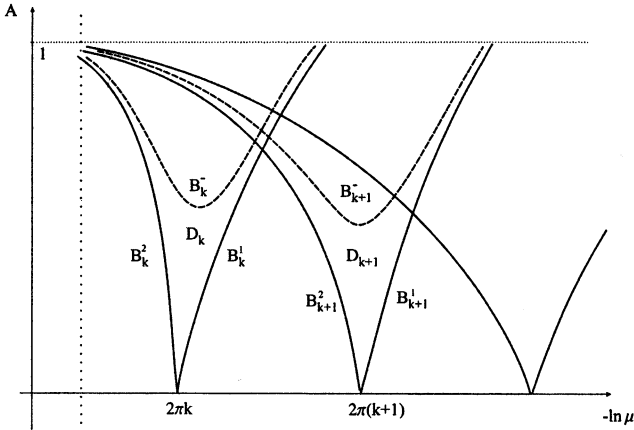


Рис. 5. Пересекающиеся резонансные зоны

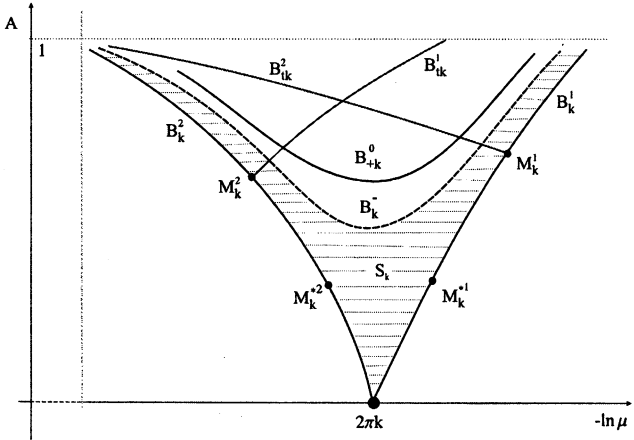


Рис. 6. Бифуркационные кривые в резонансной зоне

Можно показать, что к каждой точке на оси $-\ln \mu_0$ с координатами $(2\pi k, 0)$ (где k — достаточно большое целое число) примыкает область D_k , отвечающая существованию пары неподвижных точек отображения Пуанкаре таких, что $\bar{\theta} = \theta + 2\pi k$; в системе (4) им отве-

чают периодические движения периода $2\pi k$. Эта область ограничена двумя бифуркационными кривыми B_k^1 и B_k^2 , для значений параметров которых неподвижные точки сливаются в одну типа седло-узел. Кривые B_k^1 продолжаются до прямой $A = 1$, а кривые B_k^2 уходят влево (что соответствует возрастанию μ), оставаясь в области $A < 1$. При этом B_k^2 будет пересекать кривые B_m^1 и B_m^2 при $m < k$.

Внутри области D_k одна из неподвижных точек, Q_k , всегда седловая. Другая точка, P_k , устойчива в области S_k между кривыми B_k^1 , B_k^2 и B_k^- ; при переходе через B_k^- точка P_k теряет устойчивость и из нее рождается цикл периода два. Область S_k является областью синхронизации, так как отвечает существованию устойчивого периодического движения периода $2\pi k$. Заметим, что для любых достаточно больших целых k и m пересечение областей S_k и S_m не пусто. Таким образом, для всех достаточно больших k и m имеются области сосуществования устойчивых периодических точек периодов $2\pi k$ и $2\pi m$.

В области D_k замкнутая инвариантная кривая, которая существует, согласно утверждению (пункт 1), по крайней мере при $A < \frac{\gamma}{\sqrt{1+\gamma^2}}$, представляет собой объединение неустойчивого многообразия W^u седловой неподвижной точки Q_k и устойчивой точки P_k (рис. 7). При переходе через B_k^- инвариантная кривая, очевидно, разрушается (рис. 8).

Другой механизм разрушения замкнутой инвариантной кривой связан с появлением гомоклинических касаний, т. е. касаний неустойчивого и устойчивого многообразий точки Q_k . Такие касания имеют место на бифуркационных кривых B_{tk}^1 и B_{tk}^2 , каждая отвечает образованию гомоклинического касания на своей компоненте множества $W^u \setminus Q_k$ (рис. 9).

Кривые B_{tk}^1 и B_{tk}^2 замечательны тем, что они разграничивают область синхронизации S_k на области с простой и сложной динамикой. Ниже кривых B_{tk}^1 и B_{tk}^2 в области S_k устойчивая точка P_k является единственным аттрактором, к ней стремятся все траектории, кроме седловой неподвижной точки Q_k . В области выше кривых B_{tk}^1 и B_{tk}^2 точка Q_k имеет грубую гомоклиническую траекторию,

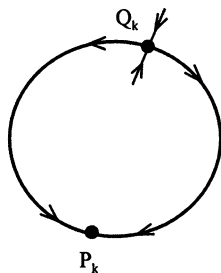


Рис. 7. Замыкание неустойчивого многообразия седловой неподвижной точки является замкнутой инвариантной кривой.

и соответственно отображение имеет нетривиальные гиперболические множества [26]. Отметим, что S_m ($m < k$) пересекает S_k всегда выше B_{tk}^2 . Поэтому в данной модели в области $S_k \cap S_m$, отвечающей существованию пары устойчивых периодических траекторий различных периодов $2\pi k$ и $2\pi m$, всегда есть и сложная динамика, аналогичная случаю уравнения Ван-дер-Поля.

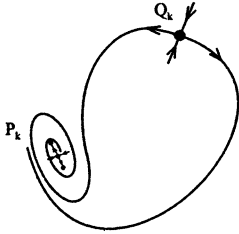


Рис. 8. После удвоения периода замыкание неустойчивого многообразия седловой неподвижной точки уже не гомеоморфно окружности.

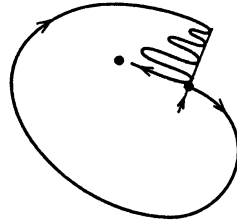


Рис. 9. Разрушение замкнутой инвариантной кривой: одна из компонент неустойчивого многообразия седловой неподвижной точки касается ее устойчивого многообразия.

В действительности в области выше кривых B_{tk}^1 и B_{tk}^2 возможно сосуществование устойчивой точки P_k и с устойчивыми периодическими точками других периодов: так как гомоклинические касания, отвечающие кривым B_{tk}^1 и B_{tk}^2 , невырождены (квадратичны), то согласно [20] выше этих кривых на плоскости параметров есть области Ньюхауса, в которых для плотного множества значений параметров система имеет одновременно счетное множество устойчивых периодических движений. Как следует из [21] (см. также [29] для многомерного случая), здесь также будут существовать области гиперболичности, причем можно показать, что для значений параметров из этих областей устойчивая точка P_k будет единственным аттрактором.

Заметим, что в области синхронизации S_k выше кривых B_{tk}^1 и B_{tk}^2 синхронизация всегда *неполная*. Во-первых, здесь возможно существование, наряду с устойчивым периодическим движением L_k , отвечающим точке P_k , устойчивых периодических траекторий других периодов. Но даже если таких траекторий нет и L_k — единственный аттрактор, то для траекторий из гиперболического множества вблизи грубой гомоклинической траектории к седловой точке Q_k разность фаз с устойчивым периодическим движением L_k растет асимптотически линейно, т. е. десинхронизация наблюдается по крайней мере для переходного процесса.

Надо отметить, что для десинхронизации важен не хаос сам по себе, а наличие грубой гомоклинической траектории у седла Q_k . Так, в области $D_k \setminus S_k$ ниже кривых B_{tk}^1 и B_{tk}^2 , где у Q_k нет гомоклинических траекторий, разность фаз между любыми двумя траекториями всегда остается ограниченной, т. е. можно говорить об *относительной синхронизации*. При этом динамика может быть хаотической: так, в области выше кривой B_{tk}^0 точка P_k , которая уже потеряла устойчивость и стала седловой, имеет грубую гомоклиническую траекторию. На кривой B_{tk}^0 точка P_k имеет траекторию гомоклинического касания (рис. 10), причем это касание третьего класса, поэтому сложная динамика имеет место и ниже кривой B_{tk}^0 [10].

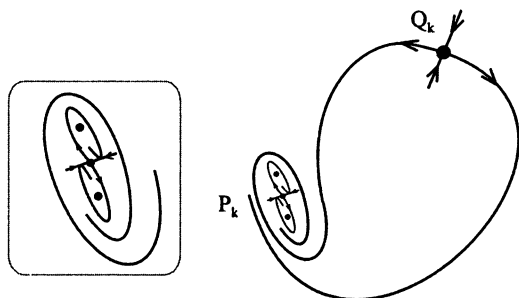


Рис. 10. Гомоклинические касания между неустойчивым и устойчивым многообразиями точки P_k .

Таким образом, в области D_k , отвечающей существованию периодического движения периода $2\pi k$, мы выделили области полной, неполной и относительной синхронизации. Режим десинхронизации, а также неполной синхронизации, т. е. сосуществование периодических траекторий различных чисел вращения, для значений параметров из области D_k всегда влечет сложную динамику. Кроме того, в области относительной синхронизации имеет место также и другой, «невращательный» тип хаотического поведения. Можно показать, что такая же картина поведения в резонансной области D_k имеет место не только для $f(\theta) = A \sin \theta$, но и в общем случае, для произвольной функции f .

Возникает вопрос: что происходит при выходе из зоны синхронизации S_k через границу B_k^2 или B_k^1 , т. е. при исчезновении седло-узла? Ответ существенно зависит от поведения неустойчивого многообразия седло-узла W^u . Выше точек M_k^1 и M_k^2 , в которых заканчиваются кривые B_{tk}^1 и B_{tk}^2 , отвечающие появлению гомоклинических

траекторий, неустойчивое многообразие седло-узла W^u имеет точки трансверсального пересечения с его сильно устойчивым многообразием W^{ss} (рис. 11).

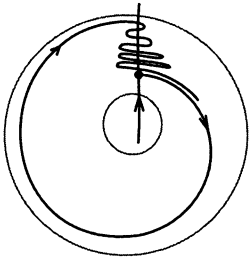


Рис. 11. Гомоклиническое пересечение неустойчивого и сильно устойчивого многообразий седло-узла.

сохраняется [18], т. е. при переходе через границы B_k^2 или B_k^1 выше точек M_k^1 и M_k^2 мы попадаем в область десинхронизации («вращательного» хаоса).

Как показано в [18], вблизи гомоклинической траектории, отвечающей этому пересечению, лежит нетривиальное гиперболическое множество, такое же, как вблизи трансверсальной гомоклинической траектории к грубому седлу. При переходе внутрь области D_k седло-узел разваливается на устойчивый узел и седло, при этом седло наследует гомоклиническую траекторию и гиперболическое множество сохраняется. При выходе из области D_k седло-узел исчезает, но большая часть гиперболического множества сохраняется.

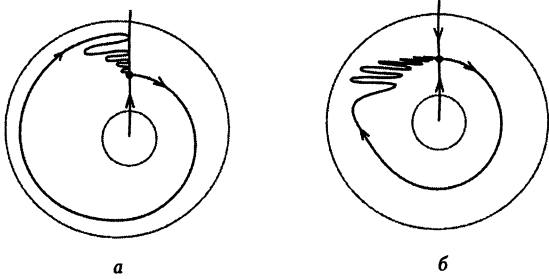


Рис. 12. Гомоклиническое касание между неустойчивым и сильно устойчивым многообразиями седло-узла (а). Складки на неустойчивом многообразии седло-узла сохраняются и после исчезновения гомоклинического касания (б).

Точки M_k^1 и M_k^2 отвечают касанию многообразий W^u и W^{ss} по негрубой гомоклинической траектории (рис. 12, а). При этом единственной предельной точкой траекторий из W^u служит сам седло-узел и W^u гомеоморфно окружности. Ниже этих точек многообразие W^u всегда возвращается к седло-узлу и лежит при этом

в узловой области. Оно по-прежнему гомеоморфно окружности. При этом при значениях параметров вблизи точек M_k^1 и M_k^2 складки на W^u сохраняются (рис. 12, б), т. е. оно не будет гладким многообразием (касательный вектор к W^u при подходе к седло-узлу со стороны узловой области осциллирует и не имеет предела). Это всегда так, пока W^u касается сильно устойчивого инвариантного слоения F^{ss} в какой-нибудь точке узловой области [22, 28]. Если же W^u пересекает слоение F^{ss} трансверсально, то оно примыкает к седло-узлу уже гладким образом. Такая ситуация имеет место при достаточно малых A : здесь единственным аттрактором служит гладкая инвариантная замкнутая кривая, и на линиях $B_k^{1,2}$ она совпадает с W^u . Обозначим через M_k^{*1} и M_k^{*2} такие точки на кривых B_k^1 и B_k^2 соответственно, что ниже этих точек W^u примыкает к седло-узлу гладко, а выше — негладко.

Покажем, что при выходе из зоны полной синхронизации через кривые B_k^1 или B_k^2 ниже точек M_k^{*1} и M_k^{*2} соответственно устойчивая гладкая инвариантная замкнутая кривая сохраняется. На ней будет существовать либо плотная квазипериодическая траектория при иррациональных числах вращения, либо периодические траектории очень больших периодов при рациональных числах вращения. С практической точки зрения это означает, что мы будем иметь здесь двухчастотный режим, то есть биения. При выходе из зоны синхронизации по любому пути, пересекающему границы B_k^1 или B_k^2 выше точек M_k^{*1} и M_k^{*2} , мы либо сразу попадаем в зону хаотической динамики, либо интервалы значений параметра, отвечающие как хаосу, так и простой динамике, будут чередоваться. При этом первый случай всегда имеет место вблизи точек $M_k^{1,2}$ (в том числе и ниже этих точек), а второй вблизи точек M_k^{*1} и M_k^{*2} (выше их).

2. Исчезновение седло-узла

Задачу о глобальных бифуркациях при исчезновении седло-узла будем рассматривать здесь не в рамках отображений, а для систем в непрерывном времени. Дело в том, что в случае систем, не допускающих глобальной секущей, здесь возникают новые необычные явления, представляющие несомненный интерес для многих проблем нелинейной динамики.

Рассмотрим однопараметрическое семейство C^2 -гладких $(n + 2)$ -мерных динамических систем, гладко зависящих от параметра μ , где $\mu \in (-\mu_0, \mu_0)$. Предположим, что выполнены следующие условия:

1. При $\mu = 0$ система имеет негрубое периодическое движение L_0 типа седло-узел. Это означает, что все мультипликаторы, кроме одного, равного 1, лежат внутри единичного круга и первая ляпуновская величина не равна нулю.

2. Все траектории неустойчивого многообразия W^u траектории L_0 стремятся к L_0 при $t \rightarrow \infty$, причём $W^u \cap W^{ss} = \emptyset$, т. е. W^u , когда возвращается к L_0 , целиком лежит в узловой области.

3. Рассматриваемое семейство трансверсально бифуркационно-му множеству систем с негрубым периодическим движением типа седло-узел. Это означает, в частности, что при изменении μ седло-узел бифурцирует: разваливается на седло и узел при, скажем, $\mu < 0$ и исчезает при $\mu > 0$.

Согласно [34], можно ввести координаты в малой окрестности траектории L_0 так, что система будет записываться в виде

$$\begin{aligned} \dot{x} &= \mu + x^2[1 + p(x, \theta, \mu)], \\ \dot{y} &= [A(\mu) + q(x, \theta, y, \mu)]y, \\ \dot{\theta} &= 1, \end{aligned} \quad (13)$$

где собственные числа матрицы A все лежат строго слева от мнимой оси. Здесь θ — циклическая переменная, определенная по модулю 1, при этом точка $(x, y, \theta = 0)$ подклеивается к точке $(x, \sigma y, \theta = 1)$, где σ — некоторая инволюция в R^n (см. [27]). Таким образом, p — функция периода 1 по θ и q — функция периода 2. При этом $p(0, \theta, 0) = 0$ и $q(0, \theta, 0) = 0$. Кроме того, данные координаты вводятся так, что p не зависит от θ при $\mu = 0$ (т. е. отображение Пуанкаре на центральном многообразии вложено в автономный поток [30]).

Седло-узловая периодическая траектория L_0 задаётся уравнением $x = 0, y = 0$ при $\mu = 0$. Ее сильно-устойчивое многообразие W^{ss} локально задается уравнением $x = 0$. Оно разделяет окрестность траектории L_0 на узловую и седловую области, $x < 0$ и $x > 0$ соответственно. Многообразие $y = 0$ инвариантно, это центральное многообразие. При $\mu < 0$ в нем лежат две периодические траектории L_1 и L_2 , одна устойчивая, а другая седловая, которые сливаются в одну траекторию L_0 при $\mu = 0$. При $\mu > 0$ периодических траекторий нет и все траектории покидают окрестность.

При $\mu = 0$ координата x монотонно возрастает. При этом $x(t) \rightarrow 0$ в области $x < 0$ при $t \rightarrow +\infty$, причем медленно, как $\sim 1/t$. Поскольку координата y экспоненциально убывает, то получаем, что все траектории в узловой области стремятся к L_0 при $t \rightarrow +\infty$, касаясь цилиндра $y = 0$. В седловой области $x(t) \rightarrow 0$ уже при $t \rightarrow -\infty$, и так как y экспоненциально возрастает с убыванием t , то множество тра-

екторий, которые стремятся к седло-узлу L_0 при $t \rightarrow -\infty$, т. е. его неустойчивое многообразие W^u имеет вид $\{y = 0, x \geq 0\}$.

С ростом t траектории из $W^u \setminus L_0$ покидают окрестность седло-узла, но в силу предположения 2 все они возвращаются в узловую область при $t \rightarrow +\infty$, т. е. стремятся к L_0 , касаясь цилиндра $y = 0$. Поэтому можно указать такое малое $d > 0$, что W^u будет пересекать секущую $S_0 : \{x = -d\}$. Очевидно, что $\bar{l} = W^u \cap S_0$ есть замкнутая кривая. Однако она может быть вложена в S_0 различным образом. Будем считать, что средняя линия $l_0 : \{y = 0\}$ секущей S_0 ориентирована в сторону возрастания θ , так же как и средняя линия $l_1 : \{y = 0\}$ секущей $S_1 : \{x = +d\}$. Так как $l_1 = W^u \cap S_1$, то кривая \bar{l} является образом кривой l_1 под действием отображения по траекториям системы, поэтому ориентация на l_1 определяет ориентацию на \bar{l} . Таким образом, с учетом ориентации кривая \bar{l} гомотопна ml_0 , где $m \in \mathbb{Z}$. В случае $n = 1$, т. е. когда система задана в R^3 и S_0 — двумерное кольцо, возможны только два случая: $m = 0$ и $m = +1$. В случае же $n \geq 2$ допустимы любые целые m . Поведение W^u в случае $m = 0$ изображено на рис. 13, а. В случае $m = 1$ многообразие W^u гомеоморфно тору (рис. 13, б), а в случае $m = -1$ — бутылке Клейна (рис. 13, в). В случае $|m| \geq 2$ множество W^u является $|m|$ -ветвленным многообразием (к каждой точке траектории L_0 со стороны узловой области подклеиваются $|m|$ кусков множества W^u). Заметим также, что в задаче о воздействии периодической внешней силы на автоколебательную систему, как, например, в системе (4), может реализовываться только случай $m = 1$, независимо от размерности.

Изучение поведения траекторий в окрестности множества W^u представляет интерес только при $\mu > 0$ (при $\mu \leq 0$ оно тривиально). При $\mu > 0$ эта задача сводится к изучению отображения Пуанкаре $T : S_1 \rightarrow S_1$, которое будем строить в виде суперпозиции двух отображений: $T_1 : S_1 \rightarrow S_0$ и $T_0 : S_0 \rightarrow S_1$.

Как показано в [34], когда система приведена к виду (13) и функция p не зависит от θ при $\mu = 0$, отображение $T_0 : (y_0, \theta_0) \in S_0 \mapsto (y_1, \theta_1) \in S_1$ имеет вид

$$\begin{aligned} y_1 &= \alpha(y_0, \theta_0, \nu), \\ \theta_1 &= \theta_0 + \nu + \beta(\theta_0, \nu), \end{aligned} \tag{14}$$

где $\nu(\mu)$ — время перехода с S_0 на S_1 . При $\mu \rightarrow +0$ время перехода ν монотонно стремится к бесконечности: $\nu \sim 1/\sqrt{\mu}$; при этом функции α и β равномерно стремятся к нулю вместе со всеми своими производными. Таким образом, образ секущей S_0 под действием отображения T_0 сжимается к средней линии l_1 при $\mu \rightarrow +0$.

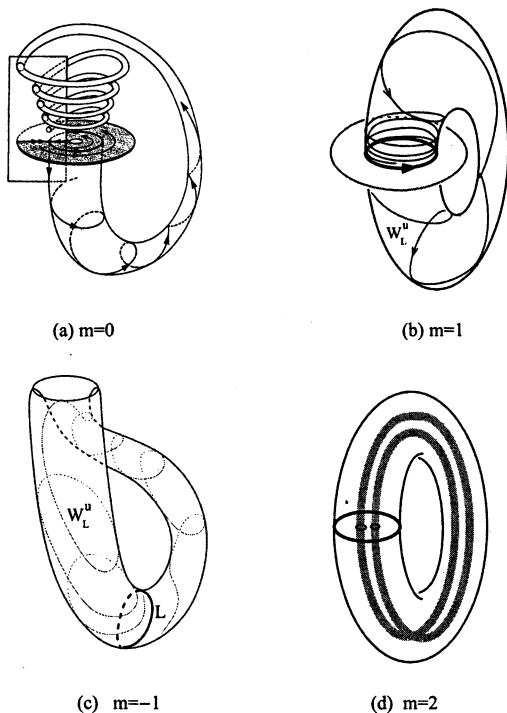


Рис. 13. Случай $m = 0$ — «катастрофа голубого неба» (a). Случаи $m = 1$ и $m = -1$ (замыкание неустойчивого многообразия седло-узловой периодической траектории) являются двумерным тором (b) или бутылкой Клейна (c) соответственно. Случай $m = 2$ — полноторий S_1 растягивается в два раза в направлении угловой переменной, сжимается в поперечном направлении и вкладывается сам в себя, образуя в пределе соленоид Виеториса — Ван Данцига (d).

Отображение $T_1 : S_1 \rightarrow S_0$ — это отображение по траекториям за конечное время, следовательно, это гладкое отображение, определенное при всех малых μ . Оно записывается в виде

$$\begin{aligned} y_0 &= G(y_1, \theta_1, \mu), \\ \theta_0 &= F(y_1, \theta_1, \mu). \end{aligned} \quad (15)$$

Образ линии l_1 будет, следовательно, иметь вид

$$y_0 = G(0, \theta_1, 0), \quad \theta_0 = F(0, \theta_1, 0). \quad (16)$$

На второе уравнение можно смотреть как на отображение одной ок-

ружности в другую. Его можно записать в виде

$$\theta_0 = m\theta_1 + f(\theta_1), \quad (17)$$

где $f(\theta)$ — периодическая функция периода 1, а степень отображения m есть как раз то самое целое число, о котором говорилось выше.

В силу (14), (15), (17) отображение Пуанкаре $T = T_0T_1 : S_1 \rightarrow S_1$ записывается в виде

$$\begin{aligned} \bar{y} &= g_1(y, \theta, \nu), \\ \bar{\theta} &= m\theta + \nu + f(\theta) + f_1(y, \theta, \nu), \end{aligned} \quad (18)$$

где функции f_1 и g_1 стремятся к нулю при $\nu \rightarrow +\infty$ вместе со всеми производными. Таким образом, мы видим, что если мы зафиксируем дробную часть ν , то при стремлении целой части ν к бесконечности отображение T вырождается в отображение окружности \bar{T} :

$$\bar{\theta} = m\theta + f(\theta) + \nu \pmod{1}. \quad (19)$$

Понятно, что во многом динамика отображения (18) определяется свойствами отображения (19). Значениям ν с одинаковой дробной частью отвечает одно и то же отображение \bar{T} . Тем самым область изменения малого параметра $\mu > 0$ представляется в виде объединения счетной последовательности отрезков $J_k = [\mu_{k+1}, \mu_k]$ (где $\nu(\mu_k) = k$) таких, что поведение отображения T на каждом из отрезков J_k в целом одинаково.

Прежде всего остановимся на двух замечательных случаях $m = 0$ и $|m| \geq 2$, рассмотренных в [32, 33, 34, 28].

Теорема [32, 34]. При $m = 0$, если $|f'(\theta)| < 1$ при всех θ , то отображение T при всех достаточно малых μ будет иметь единственную, притом устойчивую, неподвижную точку.

После того как отображение T приведено к виду (18), утверждение теоремы непосредственно следует из принципа сжимающих отображений. Из теоремы следует, что при исчезновении L_0 в этом случае появляется единственная и устойчивая периодическая траектория, длина и период которой стремятся к бесконечности при $\mu \rightarrow +0$ — как говорят, имеет место «катастрофа голубого неба». Вопрос о возможности неограниченного увеличения длины периодической траектории при приближении к моменту бифуркации был поставлен в [24]; первый пример (бесконечной коразмерности) был построен в [19]. Наша конструкция дает «катастрофу голубого неба» при бифуркации коразмерности 1. Отметим также, что, как показано в [28, 25], наш вариант «катастрофы голубого неба» описывает поведение, типичное

для сингулярно-возмущенных систем с как минимум двумя быстрыми переменными. Пример системы с явно выписанными правыми частями, в которой реализуется ситуация данной теоремы, был впервые построен в [11]. Приложение нашей конструкции к конкретным моделям нейронной динамики приведено в [38].

Теорема [32, 33]. Пусть $|m| \geq 2$ и $|m + f'(\theta)| > 1$ при всех θ . Тогда при всех достаточно малых $\mu > 0$ отображение T будет иметь гиперболический аттрактор Смейла — Вильямса.

В этих условиях отображение T действует таким же образом, как и в конструкции Смейла — Вильямса. Именно полноторий S_0 отображается сам в себя так (см. рис. 13, *d*), что предел $\Sigma = \bigcap_{k \geq 0} T^k S_0$ является соленоидом Виеториса — Ван Данцига, локально гомеоморфным прямому произведению канторовского множества на отрезок. Условия теоремы гарантируют равномерное растяжение по θ и сжатие по y , т. е. аттрактор Σ — равномерно-гиперболическое множество, и более того, так как отображение $T|_{\Sigma}$ топологически сопряжено пределу обратного спектра для растягивающего отображения окружности (степени $|m|$), то все точки Σ — неблуждающие, т. е. мы действительно имеем здесь гиперболический аттрактор [32, 33].

Как мы уже отмечали, в случае $m = \pm 1$ поверхность W^u при $\mu = 0$ может примыкать к седло-узлу L_0 как гладким, так и негладким образом, в зависимости от того, как W^u пересекает сильно устойчивое инвариантное слоение F^{ss} в узловой области. Когда система приведена к виду (13), слои этого слоения имеют вид $\{x = \text{const}, \theta = \text{const}\}$, т. е. на секущей S_0 слои F^{ss} имеют вид $\{\theta_0 = \text{const}\}$. Пересечение $W^u \cap S_0$ — это кривая (16). Следовательно (см. выражение (17)), W^u примыкает к L_0 гладко тогда и только тогда, когда

$$m + f'(\theta) \neq 0 \quad (20)$$

при всех θ . Это условие эквивалентно тому, что предельное отображение \tilde{T} (19) — диффеоморфизм окружности при всех ν .

Теорема [1]. Если \tilde{T} — диффеоморфизм, то при всех достаточно малых $\mu > 0$ отображение (18) будет иметь замкнутую устойчивую гладкую инвариантную кривую, к которой притягиваются все траектории.

Доказательство теоремы достигается ссылкой на принцип кольца [3, 27]. Заметим, что гладкой замкнутой инвариантной кривой отображения Пуанкаре T отвечает в случае $m = 1$ двумерный гладкий инвариантный тор исходной системы, а при $m = -1$ — инвариантная гладкая бутылка Клейна.

В случае, когда отображение \bar{T} не является диффеоморфизмом, ситуация намного сложнее. Мы оставим случай $m = -1$ в стороне, а основное внимание обратим на $m = 1$. Именно этот случай, как мы отмечали, характерен для задачи синхронизации.

Таким образом, мы предполагаем, что $m = 1$ и что предельное отображение \bar{T} имеет критические точки. Основные результаты для этого случая содержатся в трех теоремах ниже.

Введем величину δ , определяемую как

$$\delta = \sup_{\theta_1 < \theta_2} (\theta_1 + f(\theta_1) - \theta_2 - f(\theta_2)).$$

Очевидно, что $\delta = 0$ тогда и только тогда, когда отображение \bar{T} — гомеоморфизм при всех ν , т. е. когда его график — неубывающая функция. Если $\delta > 0$, то это отображение обязательно имеет по крайней мере одну точку максимума и одну точку минимума; по существу, δ измеряет размах колебания между минимальным значением правой части отображения (17) и предшествующим ему максимальным значением (рис. 14). Несложно подсчитать, что в случае $f = A \sin 2\pi\theta$, например, $\delta(A) = \frac{1}{\pi} (\sqrt{4\pi^2 A^2 - 1} - \arctan \sqrt{4\pi^2 A^2 - 1})$.

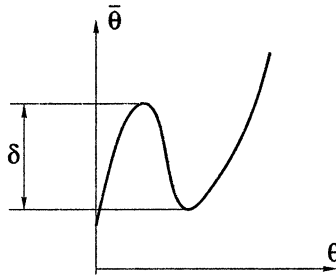


Рис. 14. Величина δ равна модулю разности между некоторым минимальным значением правой части отображения и предшествующим ему максимальным значением.

В случае $\delta \geq 1$ у каждого значения θ есть по крайней мере три прообраза относительно отображения (17). В терминах исходной системы легко видеть, что мы имеем $\delta > 0$ тогда и только тогда, когда некоторый слой слоения F^{ss} имеет более одного (на самом деле, по крайней мере, три) пересечения с неустойчивым многообразием седло-узла W^u при $\mu = 0$, и $\delta \geq 1$ тогда и только тогда, когда W^u пересекает каждый слой слоения F^{ss} по крайней мере трижды.

Заемствуя терминологию из [1], будем называть случай $\delta > 1$ случаем большой петли.

Теорема [31]. В случае большой петли отображение T имеет сложную динамику при всех достаточно малых $\mu > 0$.

Под сложной динамикой следует понимать наличие грубой гомоклинической траектории к грубому периодическому движению. Это влечет существование гиперболического множества, т. е. в малой окрестности гомоклинической структуры все траектории (их континуум), включая счетное множество периодических движений, являются седловыми и допускают полное символическое описание на языке топологической марковской цепи.

Хотя такая структура и грубая, т. е. гиперболическое множество сохраняется при малых изменениях μ , мы не можем, вообще говоря, ожидать, что одно и то же гиперболическое множество будет сохраняться при всех μ , т. е. следует иметь в виду, что различным интервалам значений μ в теореме могут отвечать различные гиперболические множества.

Теорема [31]. Если $\delta > 0$ в отображении (19) и все критические точки отображения имеют конечный порядок, то сколь угодно близко к $\mu = 0$ имеются интервалы значений μ , при которых отображение T имеет сложную динамику.

Теорема [31]. Если $2\delta \max_{\theta} f''(\theta) < 1$, то сколь угодно близко к $\mu = 0$ имеются интервалы значений μ , при которых отображение T имеет простую динамику: все траектории стремятся к непрерывной инвариантной кривой, гомеоморфной окружности, на которой имеется конечное число неподвижных точек (половина из них устойчивы, половина — седловые).

Таким образом, в противоположность случаю большой петли при малых $\delta > 0$ интервалы простой и сложной динамики чередуются при $\mu \rightarrow 0$.

3. Заключение

Как мы видим, классические задачи синхронизации сводятся к изучению семейств отображений кольца вида

$$\begin{aligned} \bar{x} &= G(x, \theta, \mu), \\ \bar{\theta} &= \theta + \omega + Af(\theta) + F(x, \theta, \mu) \pmod{1}, \end{aligned} \quad (21)$$

где f , F и G — периодические по θ функции с периодом 1, причем функции F и G равномерно стремятся к нулю вместе со всеми своими

производными при $\mu \rightarrow 0$. Естественно, что динамика отображения (21) связана с поведением отображения окружности

$$\bar{\theta} = \theta + \omega + Af(\theta) \pmod{1}. \quad (22)$$

При малых A это отображение является диффеоморфизмом, но с ростом A у него появляются критические точки, определяемые из уравнения

$$1 + Af'(\theta) = 0.$$

Пусть $A = A^*$ отвечает появлению первой критической точки (точки перегиба) и $A > A^{**} > A^*$ отвечает выполнению условий большой петли. Как мы видели, при $A < A^*$ отображение (21) будет при всех ω иметь гладкую инвариантную кривую, а при $A > A^{**}$ оно при всех ω будет иметь сложную динамику.

При $A^* < A < A^{**}$ мы можем гарантировать сложную динамику только в определенных интервалах значений ω , и эти интервалы могут чередоваться с интервалами простой динамики. Во всяком случае, такое чередование имеет место, когда $A - A^*$ достаточно мало.

Заметим, что при $A = A^*$ отображение (21), хотя и имеет точки перегиба, является гомеоморфизмом окружности, т. е. для него определено число вращения Пуанкаре $r(\omega)$. Как правило, для таких отображений (в противоположность диффеоморфизмам окружности) множество значений ω , отвечающих иррациональным числам вращения, образует множество меры нуль. Таким образом, типичным образом переход из области $A < A^*$ в область $A > A^*$ осуществляется внутри какой-нибудь резонансной зоны, отвечающей рациональному числу вращения. Внутри такой зоны при $A = A^*$ мы имеем на окружности конечное число периодических точек (половина устойчивых, половина неустойчивых). Ясно, что эта структура сохраняется при малых μ и при всех близких значениях A и ω , т. е. внутри резонансной зоны отображение (21) действительно будет иметь простую динамику при A , близких к A^* . Основные бифуркации, которые происходят в резонансной зоне с увеличением A , такие же, как описано в разделе 1: бифуркационные кривые, отвечающие гомоклиническим касаниям и удвоению периода, разбивают резонансную зону на области синхронизации (полной, частичной, относительной) и вращательного хаоса (см. рис. 6).

Границы резонансной зоны отвечают бифуркации седло-узла. Как следует из результатов предыдущего раздела, выходя из резонансной зоны в области $A > A^*$, всегда можно попасть в интервалы сложной динамики. Отсюда, в частности, вытекает, что на плоскости параметров (A, ω) при $A > A^*$ в любой окрестности любой точки

(A^*, ω) с иррациональным $r(\omega)$ имеются как области простой, так и области сложной динамики, в отличие от случая рационального $r(\omega)$.

Надо отметить, что в случае, когда функции F и G в отображении (21) не малы, а особенно когда отображения (21) уже не сжимает площади, т. е. задача становится более многомерной по существу, ситуация значительно усложняется. В частности, кроме бифуркаций седло-узла и удвоения периода здесь может встречаться бифуркация периодической траектории с парой комплексно сопряженных мультипликаторов на единичной окружности. Более того, когда соответствующая бифуркационная кривая выходит на границу области синхронизации, мы получаем периодическую траекторию с мультипликаторами $(1, e^{\pm i\varphi})$. Известно, что локальные бифуркации такой траектории (аналогичные бифуркациям состояния равновесия с характеристическими показателями $(0, \pm i\omega)$ [9, 12]) очень нетривиальны. Более того, на границе области синхронизации данная периодическая траектория должна иметь гомоклинику, таким образом, необходимо изучать также и глобальные бифуркации такой траектории. Заметим, что эта задача рассматривалась в связи с примером «катастрофы голубого неба» в [11]; в связи с задачей синхронизации подобное явление было замечено в [37]. Очевидно, что другие случаи коразмерности два — $(\pm 1, \pm 1)$, $(-1, e^{\pm i\varphi})$ — с гомоклиническими траекториями также нуждаются в дополнительном анализе.

Эта работа частично поддержана грантом CRDF RU-M1-2583-MO-04, грантом РФФИ № 02-01-00273, научной программой «Российские университеты», проект UR 03.01.015. Л. П. Шильников также благодарит фонд А. Гумбольта.

Литература

1. *Афраймович В. С., Шильников Л. П.* О глобальных бифуркациях, связанных с исчезновением неподвижной точки типа седло-узел // ДАН СССР. 1974. Т. 219. С. 1281—1284.
2. *Афраймович В. С., Шильников Л. П.* О малых периодических возмущениях автономных систем // ДАН СССР. 1974. Т. 214. С. 739—742.
3. *Афраймович В. С., Шильников Л. П.* Принцип кольца и задача о взаимодействии двух автоколебательных систем // ПММ. 1977. Т. 41. С. 618—627.
4. *Афраймович В. С., Шильников Л. П.* Инвариантные двумерные торы, их разрушение и стохастичность // Методы качественной теории дифференциальных уравнений. Горький: Изд-во ГГУ, 1983. С. 3—26.
5. *Андронов А. А., Леонтович Е. А.* Рождение предельных циклов из петли сепаратрисы и из сепаратрисы состояния равновесия типа седло-узел // Математический сборник. 1959. Т. 48. С. 335—376.

6. Андронов А. А., Леонтович Е. А., Гордон И. И., Ма́йер А. Г. Теория бифуркаций динамических систем на плоскости. М.: Наука, 1968.
7. *Andronov A. A., Vitt A. A.* Zur Theorie des Mithnehmens von van der Pol // Archiv für Elektrotechnik. 1930. Bd. 24. S. 99.
8. *Cartwright M. L., Littlewood J. E.* On nonlinear differential equations of the second order, I: The equation $\ddot{y} + k(1 - y^2)\dot{y} + y = b \lambda k \cos(\lambda t + a)$, k large // J. Lond. Math. Soc. 1945. V. 20. P. 180–189.
9. *Гаврилов Н. К.* О бифуркациях состояния равновесия с одним нулевым и парой чисто мнимых корней // Методы качественной теории дифференциальных уравнений. Горький: Изд-во ГГУ, 1978. С. 33–40.
10. *Гаврилов Н. К., Шильников Л. П.* О трехмерных динамических системах близких к системе с негрубой гомоклинической кривой. I // Математический сборник. 1972. Т. 88. С. 475–492.
11. *Gavrilov N. K., Shilnikov A. L.* An example of blue sky catastrophe // Methods of qualitative theory of differential equations and related topics. Amer. Math. Soc. Transl. Ser. 2. 2000. V. 200. Adv. Math. Sci. Providence, RI. P. 165–188.
12. *Guckenheimer J.* On a codimension two bifurcation // Lect. Notes Math. 1981. V. 898. P. 99–142.
13. *Крылов Н. М., Боголюбов Н. Н.* Фундаментальные проблемы нелинейной механики // ДАН СССР. 1933. Т. 4. С. 475–498.
14. *Levi M.* Qualitative analysis of the periodically forced relaxation oscillations // Mem. Amer. Math. Soc. 1981. V. 32. P. 1–146.
15. *Lewinson N.* A second order differential equation with singular solutions // Ann. Math. 1949. V. 2. P. 127–153.
16. *Littlewood J. E.* On non-linear differential equations of the second order. III. The equation $\ddot{y} - k(1 - y^2)\dot{y} + y = b?k \cos(?t + \alpha)$ for large k , and its generalizations // Acta Math. 1957. V. 97. P. 267–308.
17. *Littlewood J. E.* On non-linear differential equations of the second order. IV. The general equation $\ddot{y} + kf(y)\dot{y} + g(y) = bkp(\phi)$, $\phi = t + \alpha$ // Acta Math. 1957. V. 98. P. 1–110.
18. *Лукьянов В. И., Шильников Л. П.* О некоторых бифуркациях динамических систем с гомоклиническими структурами // ДАН СССР. 1978. Т. 243. С. 26–29.
19. *Медведев В. С.* О новом типе бифуркаций на многообразиях // Математический сборник. 1980. Т. 113. С. 487–492.
20. *Newhouse S. E.* The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms // Publ. Math. IHES. 1979. V. 50. P. 101–151.
21. *Newhouse S., Palis J.* Cycles and bifurcation theory // Asterisque. 1976. V. 31. P. 44–140.
22. *Newhouse S., Palis J., Takens F.* Bifurcations and stability of families of diffeomorphisms // Publ. Math. IHES. 1983. V. 57. P. 5–71.
23. *Осипов А. В.* Гиперболичность основного семейства для уравнения Левинсона // Дифференц. уравн. 1975. Т. 11. С. 1795–1800.

24. *Palis J., Pugh C.* Fifty problems in dynamical systems // Lect. Notes Math. 1975. V. 486. P. 334–353.
25. *Shilnikov A. L., Shilnikov L. P., Turaev D. V.* Blue sky catastrophe in singularly perturbed systems // Moscow Math. J. 2004 (in print).
26. *Шильников Л. П.* Об одной задаче Пуанкаре — Биркгофа // Математический сборник. 1967. Т. 74. С. 378–397.
27. *Shilnikov L., Shilnikov A., Turaev D., Chua L.* Methods of qualitative theory in nonlinear dynamics. Part I. Singapore: World Scientific, 1998.
28. *Shilnikov L., Shilnikov A., Turaev D., Chua L.* Methods of qualitative theory in nonlinear dynamics. Part II. Singapore: World Scientific, 2001.
29. *Стенькин О. В., Шильников Л. П.* Гомоклинический Ω -взрыв и области гиперболичности // Математический сб. 1998. Т. 189. С. 125–144.
30. *Takeya F.* Singularities of vector fields // Publ. Math. IHES. 1974. V. 43. P. 47–100.
31. *Тураев Д. В., Шильников Л. П.* Бифуркации квазиаттракторов тор-хаос // Математические механизмы турбулентности. Киев: Наукова думка, 1986. С. 113–121.
32. *Тураев Д. В., Шильников Л. П.* О катастрофах голубого неба // Доклады РАН. 1995. Т. 342. С. 596–599.
33. *Shilnikov L. P., Turaev D. V.* On simple bifurcations leading to hyperbolic attractors // Comput. Math. Appl. 1997. V. 34. P. 441–457.
34. *Shilnikov L. P., Turaev D. V.* A new simple bifurcation of a periodic orbit of blue sky catastrophe type // Methods of qualitative theory of differential equations and related topics: Amer. Math. Soc. Transl. 2000. II Ser. 200, AMS, Providence, RI. P. 165–188.
35. *B. van der Pol.* Forced oscillations in a circuit with nonlinear resistance (receptance with reactive triode) // London, Edinburgh and Dublin Phil. Mag. 1927. V. 3. P. 65–80.
36. *B. van der Pol, van der Mark.* Frequency demultiplication // Nature. 1927. V. 120. P. 363–364.
37. *Krauskopf B., Wicczorek S.* Accumulating regions of winding periodic orbits in optically driven lasers // Physica D. 2002. V. 173. P. 114–129.
38. *Shilnikov A. L., Cymbalyuk G.* Homoclinic saddle-node orbit bifurcations en route between tonic spiking and bursting in neuron models // Regular & Chaotic Dynamics. 2004. V. 9, № 3. P. 281–297.

НОРМАЛЬНЫЕ ФОРМЫ ДЛЯ СИНГУЛЯРНО ВОЗМУЩЕННЫХ СИСТЕМ ПАРАБОЛИЧЕСКОГО ТИПА И СИСТЕМ С ЗАПАЗДЫВАНИЕМ

С. А. Кащенко

Методы нормализации являются одними из основных методов анализа поведения решений нелинейных уравнений в окрестности установившегося режима. Подход, связанный с использованием известной теории инвариантных интегральных многообразий, позволяет для изучения широкого класса эволюционных уравнений воспользоваться хорошо разработанным для обыкновенных дифференциальных уравнений (см., например, [1, 2]) методом нормальных форм. Такими методами изучаются и параболические краевые задачи. Малость коэффициентов диффузии принципиально усложняет проблему и делает невозможным непосредственное применение метода нормальных форм. Дело в том, что в этом случае вырождения, определяющие порядок нормальных форм, оказываются бесконечными. При некоторых условиях, которые сформулируем ниже, будут построены в определенном смысле аналоги нормальных форм — квазинормальные формы. Последние будут записываться с помощью эволюционных уравнений. Некоторые из таких квазинормальных форм уже изучались в работах [3, 4].

Отметим, что «малость» коэффициента диффузии в некотором смысле эквивалентна «большой» области определения. Например, краевая задача с малым коэффициентом диффузии

$$\frac{\partial u}{\partial t} = \varepsilon D \frac{\partial^2 u}{\partial x^2} + F(u),$$

где $0 \leq x \leq 1$, $0 < \varepsilon \ll 1$, и граничными условиями

$$\left. \frac{\partial u}{\partial x} \right|_{x=0} = \left. \frac{\partial u}{\partial x} \right|_{x=1} = 0$$

и краевая задача с «большой» областью определения

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial y^2} + F(u),$$

где $0 \leq y \leq L = \varepsilon^{-1}$, и граничными условиями

$$\left. \frac{\partial u}{\partial y} \right|_{y=0} = \left. \frac{\partial u}{\partial y} \right|_{y=L} = 0$$

преобразуются одна в другую с помощью нормирующей замены

$$y = \varepsilon^{-1/2} x.$$

Подобная ситуация возникает и в уравнениях с запаздыванием. Так уравнение с «большим» запаздыванием

$$\frac{\partial u}{\partial t} = F(u(t - T)),$$

где $T = \varepsilon^{-1}$ ($0 < \varepsilon \ll 1$), в результате замены времени

$$\tau = Tt$$

преобразуется в сингулярно возмущенное уравнение

$$\varepsilon \frac{\partial u}{\partial \tau} = F(u(t - 1)).$$

Для локальной динамики сингулярно возмущенных уравнений с запаздыванием тоже характерны ситуации, когда размерность критического случая «бесконечномерна», т. е. неограниченно растет при $\varepsilon \rightarrow 0$.

Работа состоит из двух разделов, в первом из которых методы бесконечномерной нормализации применяются для нелинейных параболических систем с малой диффузией. Во втором разделе методами локального анализа исследуется динамика сингулярно возмущенных систем с запаздыванием.

1. Построение квазинормальных форм для нелинейных параболических уравнений с малой диффузией

В п. 1.1 приведены известные результаты о применимости методов нормальных форм для параболических уравнений, коэффициенты диффузии которых не являются малыми. В п.п. 1.2 и 1.3, напротив, предполагается, что эти коэффициенты достаточно малы. Отметим, что это предположение осмысленно, поскольку, как известно, во многих прикладных задачах коэффициенты диффузии являются относительно малыми (либо область изменения пространственной переменной достаточно «велика»). В п. 1.2 предполагается, что в сосредоточенной модели (при нулевой диффузии) реализуется случай близкий к критическому в задаче об устойчивости, а в п. 1.3 критические случаи определяются коэффициентами диффузии. Заметим, что квазинормальные формы, приведенные в последнем параграфе, оказываются самыми сложными.

1.1. О применимости метода нормальных форм для невырождающихся параболических уравнений

1. В качестве модельного класса уравнений рассмотрим на отрезке $0 \leq x \leq 1$ параболическую краевую задачу.

$$\frac{\partial u}{\partial t} = (D_0 + \varepsilon D_1) \frac{\partial^2 u}{\partial x^2} + (A_0 + \varepsilon A_1)u + F(u), \quad (1.1)$$

$$\left. \frac{\partial u}{\partial x} \right|_{x=0} = \left. \frac{\partial u}{\partial x} \right|_{x=1} = 0. \quad (1.2)$$

Здесь $u \in R^m$; D_0 — положительно определенная матрица; $0 < \varepsilon \ll 1$; $F(u)$ — достаточно гладкая вектор-функция, имеющая в нуле порядков малости выше первого. Ниже примем обозначения

$$F(u) = F_2(u, u) + F_3(u, u, u) + \dots,$$

где каждое из слагаемых, стоящее в правой части, линейно по каждому аргументу.

Предположим, что при $\varepsilon = 0$ реализуется критический случай в задаче об устойчивости нулевого состояния равновесия, т. е. ровно $m_0 > 0$ собственных значений совокупности матриц $A(n^2 \pi^2)$ ($n = 0, 1, \dots$), где

$$A(z) = A_0 - zD_0, \quad (1.3)$$

имеют нулевые вещественные части, а все отрицательные — отрицательные. В качестве фазового пространства выберем $C_{[0,1]}(R^m)$.

Поставим задачу об изучении поведения при $t \rightarrow \infty$ всех решений (1.1), (1.2) с начальными условиями из некоторого шара $S(r_0)$ с центром в точке $u \equiv 0$ достаточно малого (но независящего от ε) радиуса $r_0 > 0$.

Известно, что при $t \rightarrow \infty$ структура решений из $S(r_0)$ определяется лишь поведением тех решений, которые принадлежат некоторому инвариантному интегральному многообразию $V(\varepsilon)$ размерности m_0 (каждое решение $u(t, x)$ с начальным условием при $t = t_0$ из $S(r_0)$ определено при тех $t \geq t_0$, для которых $u(t, x) \in S(r_0)$ и расстояние от $u(t, x)$ до $V(\varepsilon)$ экспоненциально (равномерно по ε) убывает при увеличении t). Приведем здесь для примера построение уравнения на инвариантном многообразии для двух наиболее простых и важных критических случаев.

2. Пусть $m_0 = 1$, т. е. при некотором $n = n_0$ матрица $A(n_0^2 \pi^2)$ имеет нулевое собственное значение, а все остальные ее собственные значения и всех матриц $A(n^2 \pi^2)$ при $n \neq n_0$ имеют отрицательные

вещественные части. Через a и b обозначим такие собственные векторы матрицы $A(n_0^2\pi^2)$ и транспонированной к ней $A^*(n_0^2\pi^2)$ соответственно, что

$$A(n_0^2\pi^2)a = 0, \quad A^*(n_0^2\pi^2)b = 0, \quad (a, b) = 1.$$

Уравнение на многообразии $V(\varepsilon)$ в данном случае такое:

$$\dot{\xi} = \varepsilon k_1 \xi + k_2 \xi^2 + k_3 \xi^3 + O(\xi^4 + \varepsilon \xi^2 + \varepsilon^2 |\xi|). \quad (1.4)$$

При $n_0 = 0$ имеем

$$k_1 = (A_1 a, b), \quad k_2 = (F_2(a, a), b), \quad (1.5)$$

$$k_3 = (F_3(a, a, a) + F_2(A_0^{-1} F_2(a, a), a) + F_2(a, A_0^{-1} F_2(a, a)), b), \quad (1.6)$$

а при $n_0 \neq 0$

$$k_1 = ((A_1 - n_0^2 \pi^2 D_1) a, b), \quad k_2 = 0, \quad (1.7)$$

$$k_3 = \left(\frac{3}{4} F_3(a, a, a) + \frac{1}{2} F_2(A_0^{-1} F_2(a, a), a) + \frac{1}{2} F_2(a, A_0^{-1} F_2(a, a)) + \right. \\ \left. + \frac{1}{4} F_2(A^{-1}(4n_0^2 \pi^2) F_2(a, a), a) + \frac{1}{4} F_2(a, A^{-1}(4n_0^2 \pi^2) F_2(a, a)), b \right). \quad (1.8)$$

Состоянию равновесия $\xi_0(\varepsilon)$ уравнения (1.4) отвечает состояние равновесия $u_0(x, \varepsilon)$ краевой задачи (1.1), (1.2)

$$u_0(x, \varepsilon) = \xi_0(\varepsilon) a \cos n_0 \pi^2 x + o(\xi_0(\varepsilon))$$

той же устойчивости. Очевидно, иных, кроме состояний равновесия, установившихся режимов уравнения (1.4) (значит, и (1.1), (1.2)) нет.

3. Пусть теперь $m_0 = 2$ и собственные значения λ_{\pm} ($\text{Re } \lambda_{\pm} = 0$) матрицы $A(n^2\pi^2)$ при $n = n_0$ отличны от нуля: $\lambda_{\pm} = i\omega_0$, $\omega_0 > 0$. Вектора a и b введем так, что

$$A(n_0^2\pi^2)a = i\omega_0 a, \quad A^*(n_0^2\pi^2)b = -i\omega_0 b,$$

$$(a, b) = 1, \quad (\bar{a}, b) = 0$$

(скалярное произведение берется в смысле унитарного пространства). Уравнение на $V(\varepsilon)$ (нормальная форма) в данном случае можно представить в виде скалярного уравнения с комплексными коэффициентами:

$$\dot{\xi} = \varepsilon k_1 + i\varepsilon \alpha \omega_0 \xi + d|\xi|^2 \xi + O(|\xi|^5 + \varepsilon|\xi|^2 + \varepsilon^2|\xi|). \quad (1.9)$$

Здесь α — параметр, $k_1 = ((A_1 - n_0^2 \pi^2 D_1)a, b)$, а для первой ляпуновской величины d имеют место равенства: при $n = 0$ имеем $d = \beta$, где

$$\begin{aligned} \beta = & (F_3(a, a, \bar{a}) + F_3(a, \bar{a}, a) + F_3(\bar{a}, a, a) + \\ & + F_2(a, A_0^{-1}(F_2(a, \bar{a}) + F_2(\bar{a}, a))) + F_2(\bar{a}, (A_0 - 2i\omega_0 I)^{-1}F_2(a, a)) + \\ & + F_2((A_0 - 2i\omega_0 I)^{-1}F_2(a, a), \bar{a}) + F_2(A_0^{-1}(F_2(a, \bar{a}) + \\ & + F_2(\bar{a}, a)), a), b), \end{aligned} \quad (1.10)$$

а при $n_0 \neq 0$ имеем

$$\begin{aligned} d = & \frac{1}{2}\beta + \frac{1}{4}(F_3(a, a, \bar{a}) + F_3(a, \bar{a}, a) + F_3(\bar{a}, a, a) + \\ & + F_2(a, (A_0 - 4n_0^2 \pi^2 D_0)^{-1}(F_2(a, \bar{a}) + F_2(\bar{a}, a))) + F_2(\bar{a}, (A_0 - 4n_0^2 \pi^2 D_0 - \\ & - 2i\omega_0 I)^{-1}F_2(a, a)) + F_2((A_0 - 4n_0^2 \pi^2 D_0 - 2i\omega_0 I)^{-1}F_2(a, a), \bar{a}) + \\ & + F_2((A_0 - 4n_0^2 \pi^2 D_0)^{-1}(F_2(a, \bar{a}) + F_2(\bar{a}, a)), a), b). \end{aligned}$$

Поведение решений (1.9) определяется главным укороченным уравнением

$$\dot{\xi} = (\varepsilon k_1 + i\varepsilon \alpha \omega) \xi + d|\xi|^2 \xi. \quad (1.11)$$

Пусть $d_0 = \text{Re } d < 0$ и $\text{Re } k_1 > 0$. Тогда при некотором $\alpha = \alpha_0$ уравнение (1.11) имеет состояние равновесия $\xi_0(\varepsilon) = (\varepsilon(\text{Re } k) d_0^{-1})^{1/2}$, а краевая задача (1.1), (1.2) имеет периодическое решение

$$\begin{aligned} u(t, x, \varepsilon) = & [\xi_0(\varepsilon)a \exp i\omega_0 \tau + \bar{\xi}_0(\varepsilon)\bar{a} \exp(-i\omega_0 \tau)] \cos n_0 \pi x + O(\varepsilon), \\ \tau = & (1 + \varepsilon \alpha_0 + O(\varepsilon^2))t, \end{aligned}$$

той же устойчивости, что и $\xi_0(\varepsilon)$.

1.2. О квазинормальных формах уравнений с малой диффузией в случае, близком к критическому в точечной модели

1. Рассмотрим параболическое уравнение

$$\frac{\partial u}{\partial t} = \varepsilon D \frac{\partial^2 u}{\partial x^2} + (A_0 + \varepsilon A_1)u + F(u), \quad (1.12)$$

$$(0 < \varepsilon \ll 1, \quad F(u) = F_2(u, u) + F_3(u, u, u) + \dots)$$

с краевыми условиями двух типов:

$$1) \quad \left. \frac{\partial u}{\partial x} \right|_{x=0} = \left. \frac{\partial u}{\partial x} \right|_{x=2\pi} = 0 \quad (1.13)$$

и

$$2) \quad u|_{x=0} = u|_{x=2\pi}, \quad \left. \frac{\partial u}{\partial x} \right|_{x=0} = \left. \frac{\partial u}{\partial x} \right|_{x=2\pi} = 0. \quad (1.14)$$

Здесь $u \in R^m$; D — положительно определенная матрица. Основное предположение состоит в том, что матрица A_0 имеет собственные значения с нулевыми вещественными частями и не имеет — с положительными.

Другое ограничение состоит в том, что при каждом $z > 0$ все собственные значения матрицы $A(z) = A_0 - zD$ имеют отрицательные вещественные части. Отметим, что если при некотором $z_0 > 0$ матрица $A(z_0)$ имеет собственное значение с положительной вещественной частью, то каждое периодическое решение рассматриваемых краевых задач, лежащее в достаточно малой окрестности нуля, неустойчиво при малых ε .

Рассмотрим вопрос о существовании, структуре и асимптотике установившихся режимов краевой задачи (1.12), (1.13) ((1.14)), лежащих в достаточно малой окрестности (пространства $C_{[0,2\pi]}(R^m)$) нулевого состояния равновесия.

Поставленная задача существенно сложнее той, о которой говорилось в п. 1.1. Дело в том, что при $\varepsilon \rightarrow +0$ бесконечно много характеристических показателей линейной (при $F(u) \equiv 0$) краевой задачи (1.12), (1.13) ((1.14)) стремятся к мнимой оси. Метод нормальных форм привел бы к необходимости анализировать счетную систему уравнений довольно сложного вида (так как реализуется к тому же счетное число резонансов). Основу построения квазинормальных форм составляет применение векторного формализма построения обычных (п. 1.1) нормальных форм в новой ситуации. В следующих пунктах разберем основные (и наиболее интересные) случаи.

2. Предположим сначала, что $m_0 = 1$, т. е. матрица A_0 имеет простое нулевое собственное значение. Пусть

$$A_0 a = 0, \quad A_0^* b = 0, \quad (a, b) = 1.$$

В (1.1), (1.2) положим

$$u = \varepsilon \xi(\varepsilon t, x) a + \varepsilon^2 u_2(\varepsilon t, x) + \dots \quad (1.15)$$

и будем собирать коэффициенты при одинаковых степенях ε . На втором шаге тогда получим краевую задачу для $\xi(\tau, x)$:

$$\frac{\partial \xi}{\partial \tau} = (Da, b) \frac{\partial^2 \xi}{\partial x^2} + (A_1 a, b) \xi + (F_2(a, a), b) \xi^2, \quad (1.16)$$

$$\left. \frac{\partial \xi}{\partial x} \right|_{x=0} = 0 \quad (\xi|_{x=0} = \xi|_{x=2\pi}, \quad \left. \frac{\partial \xi}{\partial x} \right|_{x=0} = \left. \frac{\partial \xi}{\partial x} \right|_{x=2\pi}). \quad (1.17)$$

Эту краевую задачу назовем квазинормальной формой. Она отличается от «обычной» нормальной формы уравнения при $D = 0$ диффузионным слагаемым $(Da, b) \xi''$. В том случае, когда $(F_2(a, a), b) = 0$ и $k_3 \neq 0$ (k_3 определяется формулой (1.6)) разложение, подобное (1.15), будем вести по степеням $\sqrt{\varepsilon}$. В результате с краевыми условиями (1.17) получим уравнение

$$\frac{\partial \xi}{\partial \tau} = (Da, b) \frac{\partial^2 \xi}{\partial x^2} + (A_1 a, b) \xi + k_3 \xi^3. \quad (1.18)$$

Из условия отрицательности вещественных частей всех собственных значений матриц $A(z)$ ($z > 0$) вытекает, что $(Da, b) \geq 0$. Будем дополнительно предполагать, что выполнено условие невырожденности

$$(Da, b) > 0. \quad (1.19)$$

Отметим, что при $m = 2$ и $D = \text{diag}(d_1, d_2)$ ($d_{1,2} > 0$), $A_0 = \{a_{ij}\}_1^2$ сформулированные ограничения задаются неравенством

$$d_1 a_{22} + d_2 a_{11} < 0.$$

Установившимися режимами (1.16), (1.17) и (1.18), (1.17) могут быть только состояния равновесия. Пусть $\xi_0(x)$ — такое состояние равновесия.

Теорема 2.1. Пусть спектру линеаризованной на $\xi_0(x)$ краевой задачи не принадлежит нулевое собственное значение. Тогда найдется такое $\varepsilon_0 > 0$, что при $\varepsilon \in (0, \varepsilon_0)$ краевая задача (1.16), (1.17) ((1.18), (1.17)) имеет состояние равновесия

$$u_0(x, \varepsilon) = \varepsilon^k \xi_0(x) a + O(\varepsilon^{2k})$$

($k = 1$, если $\xi_0(x)$ — решение уравнения (1.16) и $k = \frac{1}{2}$, если — (1.18)). Это состояние равновесия экспоненциально устойчиво, если весь указанный спектр отрицателен, и неустойчиво, если найдется собственное значение с положительной вещественной частью.

Из общей теории краевых задач типа Штурма — Лиувилля получаем следующее утверждение.

Теорема 2.2. *В условии теоремы 2.1 для устойчивости решения $u_0(x, \varepsilon)$ необходимо и достаточно, чтобы $\xi_0(x) \equiv \text{const}$.*

3. Здесь рассмотрим случай, когда матрица A_0 имеет пару чисто мнимых собственных значений $\pm i\omega_0$ ($\omega_0 > 0$) (а все остальные ее собственные значения и все собственные значения $A(z)$, $z > 0$, имеют отрицательные вещественные части). Пусть

$$A_0 a = i\omega_0 a, \quad A_0^* b = -i\omega_0 b, \quad (a, b) = 1, \quad (a, \bar{b}) = 0.$$

Из свойств матриц $A(z)$ вытекает, что $\text{Re}(Da, b) \geq 0$. Будем дополнительно предполагать, что

$$\text{Re}(Da, b) > 0. \quad (1.20)$$

При $m = 2$ и $D = \text{diag}(d_1, d_2)$ сформулированные условия эквивалентны неравенству

$$\kappa(d_1 - d_2) > -2\sqrt{d_1 d_2}, \quad \text{где } \kappa = \frac{\text{Re } a_1 \bar{a}_2}{\text{Im } a_1 a_2}, \quad a = (a_1, a_2). \quad (1.21)$$

В этом случае $(Da, b) = \frac{1}{2}(d_1 + d_2) - i\kappa \frac{1}{2}(d_1 - d_2)$.

Полагаем в (1.12), (1.13) ((1.14))

$$t = (1 + \varepsilon\alpha + \dots)\tau,$$

$$u = \sqrt{\varepsilon}(\xi(\varepsilon t, x)a \exp(i\omega_0 \tau) + \bar{\xi}(\varepsilon t, x)\bar{a} \exp(-i\omega_0 \tau)) + \varepsilon u_2(\varepsilon t, \tau, x) + \varepsilon^{3/2} u_3(\varepsilon t, \tau, x) + \dots, \quad (1.22)$$

где $u_j(s, \tau, x)$ периодичны по τ . Стандартные действия приводят к уравнению

$$\frac{\partial \xi}{\partial s} = (Da, b) \frac{\partial^2 \xi}{\partial x^2} + [(A_1 a, b) + i\omega_0 \alpha + d|\xi|^2]\xi \quad (1.23)$$

соответственно с краевыми условиями

$$\left. \frac{\partial \xi}{\partial x} \right|_{x=0} = \left. \frac{\partial \xi}{\partial x} \right|_{x=2\pi} = 0 \quad (1.24)$$

или

$$\xi|_{x=0} = \xi|_{x=2\pi}, \quad \left. \frac{\partial \xi}{\partial x} \right|_{x=0} = \left. \frac{\partial \xi}{\partial x} \right|_{x=2\pi} = 0. \quad (1.25)$$

В (1.23) первая ляпуновская величина d определяется формулой (1.10). Полученную краевую задачу естественно назвать квазинормальной формой.

Остается заметить, что грубым состояниям равновесия (при некотором $\alpha = \alpha_0$) квазинормальной формы отвечают (согласно формуле (1.22)) периодические решения той же устойчивости. Грубому периодическому решению отвечает двумерный тор исходной краевой задачи.

4. Остановимся на наиболее интересном с точки зрения приложений случае, когда

$$d_0 = \operatorname{Re} d < 0, \quad \tau'_0 = \operatorname{Re} (A_1 a, b) > 0.$$

В этом случае уравнение (1.12) при $D = 0$ и краевая задача (1.12), (1.13) ((1.14)) имеют периодическое решение

$$u_0(t, \varepsilon) = \sqrt{\varepsilon}(\xi_0 a \exp(i\omega_0 t) + \bar{\xi}_0 \bar{a} \exp(-i\omega_0 t)) + \varepsilon u_2(t) + \dots,$$

$$t = (1 - \varepsilon[|\operatorname{Im} d| |\xi_0|^2 + \frac{1}{\omega_0} \operatorname{Im} (A_1 a, b)] + O(\varepsilon^2))\tau,$$

где

$$\xi_0 = (\tau'_0 |d_0|^{-1})^{1/2}$$

— состояние равновесия (1.23), (1.24) ((1.25)). Кроме этого, при условии

$$\tau'_0 > n^2 \operatorname{Re} (Da, b) \quad (n > 0 -)$$

краевая задача (1.23), (1.25) имеет решения

$$\xi_n \exp inx, \quad \text{где} \quad |\xi_n|^2 = -d_0^{-1} [\tau'_0 - n^2 \operatorname{Re} (Da, b)],$$

которым отвечают бегущие волны

$$u_{\pm n}(\tau \pm nx, \varepsilon) = \sqrt{\varepsilon}(\xi_n a \exp i(\omega_0 \tau \pm nx) + \bar{\xi}_n \bar{a} \exp[-i(\omega_0 \tau \pm nx)]) + O(\varepsilon),$$

$$t = (1 + O(\varepsilon))\tau.$$

Приведем результаты из работы [5] об устойчивости бегущих волн. Для этого введем несколько обозначений. Положим

$$x_0 = -\operatorname{Im} (Da, b), \quad y_0 = -\frac{\omega_0 \operatorname{Im} d}{\operatorname{Re} d},$$

$$Q(k, n) = \frac{1}{4} \left(\frac{k^2}{n^2} + z_n \right)^2 \left(\frac{k^2}{n^2} - Az_n \right) - \frac{k^4}{n^4} - \frac{k^2}{n^2} Az_n + Bz_n^2,$$

где $k = 1, 2, \dots$,

$$z_n = \tau_0' [n^2 \operatorname{Re}(Da, b)]^{-1} - 1,$$

$$A = 2(x_0 y_0 - 1)(1 + x_0^2)^{-1}, \quad B = (1 + y_0^2)(1 + x_0^2)^{-1}.$$

Теорема 2.3. Пусть выполнены сформулированные выше условия (на $A(z)$) и неравенство $x_0 y_0 < 1$. Тогда при всех достаточно малых ε периодическое решение $u_0(\tau, \varepsilon)$ экспоненциально орбитально устойчиво. Если же $x_0 y_0 > 1$, то $u_0(\tau, \varepsilon)$ устойчиво (неустойчиво) при условии

$$\operatorname{Re}(Da, b) - \frac{16\tau_0'(x_0 y_0 - 1)}{1 + x_0^2} > 0 \quad (< 0).$$

Таким образом устойчивое в точечной модели решение может потерять устойчивость при учете диффузии.

Отметим, что факт существования бегущих волн можно установить путем применения теории Андронова — Хопфа и методов из [6] к системе обыкновенных уравнений

$$\varepsilon Du'' - \omega_n(\varepsilon)u' + (A_0 + \varepsilon A_1)u + F(u) = 0.$$

Теорема 2.4. Пусть при всех $k = 1, 2, \dots$ выполнены неравенства $Q(k, n) > 0$. Тогда найдется такое $\varepsilon_0 > 0$, что при всех $\varepsilon \in (0, \varepsilon_0)$ периодическое решение $u_{\pm n}(\tau \pm nx, \varepsilon)$ экспоненциально орбитально устойчиво. Если для некоторого номера k_0 имеем $Q(k_0, n) < 0$, то решение $u_{\pm n}(\tau \pm nx, \varepsilon)$ неустойчиво.

Приведем один известный результат. Пусть $\xi_0(x)$ — состояние равновесия для некоторого $\alpha = \alpha_0$ задачи (1.23), (1.24) и пусть только одна точка спектра линеаризованной на $\xi_0(x)$ краевой задачи лежит на мнимой оси. Тогда при достаточно малых ε исходная краевая задача (1.12), (1.13) имеет периодическое решение

$$u_0(t, x, \varepsilon) = \sqrt{\varepsilon}(\xi_0(x)a \exp i\omega_0\tau + \bar{\xi}_0(x)\bar{a} \exp(-i\omega_0\tau)) + O(\varepsilon),$$

$$t = (1 + O(\varepsilon))\tau.$$

В указанных условиях для устойчивости $u_0(t, x, \varepsilon)$ необходимо и достаточно, чтобы не было собственных значений упомянутой линеаризованной задачи с положительными вещественными частями.

5. О построении квазинормальных форм в вырожденных случаях. Ситуацию, когда коэффициент при $\frac{\partial^2 \xi}{\partial x^2}$ обращается в ноль, назовем вырожденным случаем. Квазинормальные формы здесь сложнее. Приведем для примера такую форму при условиях $m_0 = 1$ и $(Da, b) = 0$ (в уравнении (1.16)).

Квазинормальной формой назовем краевую задачу

$$\begin{aligned} \frac{\partial \xi}{\partial \tau} = & -\varepsilon\{(Da_1, b) \frac{\partial^4 \xi}{\partial x^4} + [(Da_2, b) + (A_1 a, b)] \frac{\partial^2 \xi}{\partial x^2} + \\ & + (A_1 a_2, b) \xi\} + (A_1 a, b) \xi + (F_2(a, a), b) \xi^2, \\ \frac{\partial \xi}{\partial x} \Big|_{x=0} = & \frac{\partial \xi}{\partial x} \Big|_{x=1} = \frac{\partial^3 \xi}{\partial x^3} \Big|_{x=0} = \frac{\partial^3 \xi}{\partial x^3} \Big|_{x=1} = 0. \end{aligned}$$

Здесь a_1 и a_2 определяются из равенств $A_0 a_1 = Da$, $A_0 a_2 = A_1 a - (A_1 a, b)a$. Отметим, что имеет место неравенство $(Da_1, b) \geq 0$.

1.3. Квазинормальные формы для уравнений, критические случаи в которых определяются коэффициентами диффузии

1. Рассмотрим параболическую краевую задачу

$$\dot{u} = \varepsilon Du'' + (A_0 + \mu A_1)u + F(u), \quad (1.26)$$

$$\frac{\partial u}{\partial x} \Big|_{x=0} = \frac{\partial u}{\partial x} \Big|_{x=1} = 0. \quad (1.27)$$

Как и выше, здесь $u \in R^m$, $0 < \varepsilon \ll 1$, $0 < \mu \ll 1$, D — положительно определенная матрица, $F(u) = F_2(u, u) + F_3(u, u, u) + \dots$. В отличие от ограничений п. 1.2, ниже считаем, что все собственные значения матрицы A_0 имеют отрицательные вещественные части. Основное предположение состоит в том, что при всех $z > 0$ и $z \neq z_0$ ($z_0 > 0$) собственные значения матриц $A(z) = A_0 - zD$ тоже имеют отрицательные вещественные части, а при $z = z_0$ у матрицы $A(z_0)$ есть собственные значения с нулевыми вещественными частями и нет — с положительными. И в этом случае при $\varepsilon \rightarrow 0$ бесконечно много характеристических показателей (1.26), (1.27) (при $F(u) \equiv 0$) стремятся к нулю. Квазинормальные формы, с помощью которых исследуется краевая задача (1.26), (1.27), в этом случае существенно отличаются от тех, которые использовались в п. 1.2. Ниже приведем соответствующие квазинормальные формы для двух основных критических случаев.

2. В этом пункте будем предполагать, что мнимой оси принадлежит только одно (простое) нулевое собственное значение $A(z_0)$.

Пусть $A(z_0)a = 0$, $A^*(z_0)b = 0$, $(a, b) = 1$, $\mu = c\varepsilon$, $n(\varepsilon) = \frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}} + \theta$, где $\theta = \theta(\varepsilon) \in (-1, 0]$ и определяется равенством

$$\theta = \left\{ \frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}} \right\} - \frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}}.$$

Из наложенных на матрицу $A(z)$ условий вытекает, что

$$(Da, b) = 0, \quad (Da_1, b) \geq 0,$$

где a_1 — решение уравнения $A(z_0)a_1 = Da$. Дополнительно будем предполагать, что

$$\lambda = (Da_1, b) > 0.$$

Положим в (1.26), (1.27)

$$\begin{aligned} u &= \sqrt{\varepsilon} [\xi(\tau, x) \cos \pi n(\varepsilon)x + \eta(\tau, x) \sin \pi n(\varepsilon)x] + \varepsilon [u_{20}(\tau, x) + \\ &+ u_{21}(\tau, x) \cos \pi n(\varepsilon)x + u_{22}(\tau, x) \sin \pi n(\varepsilon)x + v_{21}(\tau, x) \cos 2\pi n(\varepsilon)x + \\ &+ v_{22}(\tau, x) \sin 2\pi n(\varepsilon)x] + \varepsilon^{3/2} u_3(t, x, \varepsilon), \end{aligned}$$

где $\tau = \varepsilon t$,

$$\frac{\partial \xi}{\partial x} \Big|_{x=0} = \frac{\partial \xi}{\partial x} \Big|_{x=1}, \quad \eta|_{x=0} = \eta|_{x=1} = 0. \quad (1.28)$$

Применяя стандартные процедуры, на втором шаге получаем выражения

$$\begin{aligned} u_{20}(\tau, x) &= -\frac{1}{2}(\xi^2 + \eta^2)c_0, \quad c_0 = A_0^{-1}F_2(a, a), \\ u_{21}(\tau, x) &= 2\sqrt{z_0} [\theta\pi\xi - \eta']a_1, \\ u_{22}(\tau, x) &= 2\sqrt{z_0} [\xi' + \theta\pi\eta]a_1, \\ v_{21}(\tau, x) &= -\frac{1}{2}(\xi^2 - \eta^2)c_1, \quad c_1 = A^{-1}(4z_0)F_2(a, a), \\ v_{22}(\tau, x) &= -\frac{1}{2}\xi\eta c_1. \end{aligned}$$

На следующем шаге, приравнявая коэффициенты при $\varepsilon^{3/2}$, из условия разрешимости краевой задачи относительно $u_3(t, x, \varepsilon)$ получим выражения

$$\dot{\xi} = 4z_0\lambda[\xi'' + 2\pi\theta\eta' - \pi^2\theta^2\xi] + c(A_1a, b)\xi + \xi(P_1\xi^2 + P_2\eta^2), \quad (1.29)$$

$$\dot{\eta} = 4z_0\lambda[\eta'' - 2\pi\theta\xi' - \pi^2\theta^2\eta] + c(A_1a, b)\eta + \eta(P_3\xi^2 + P_4\eta^2), \quad (1.30)$$

в которых

$$P_1 = \left(\frac{3}{4}F_3(a, a, a) - \frac{1}{2}(F_2(a, c_0) + F_2(c_0, a)) + \frac{1}{2}(F_2(a, c_1) + F_2(c_1, a)), b\right),$$

$$P_2 = \left(-\frac{1}{4}F_3(a, a, a) - \frac{1}{2}(F_2(a, c_0) + F_2(c_0, a)), b\right),$$

$$P_3 = \left(\frac{3}{4}F_3(a, a, a) - \frac{1}{2}(F_2(a, c_0) + F_2(c_0, a)), b\right),$$

$$P_4 = P_3 + \frac{1}{4}(F_2(a, c_1) + F_2(c_1, a)), b).$$

Краевую задачу (1.29), (1.30), (1.28) назовем квазинормальной формой. Переходя к комплексной переменной, этой краевой задаче можно придать более удобный вид. Особенно полезен такой переход при рассмотрении периодической краевой задачи (1.26) и

$$u|_{x=0} = u|_{x=2\pi}, \quad \frac{\partial u}{\partial x}\Big|_{x=0} = \frac{\partial u}{\partial x}\Big|_{x=2\pi}. \quad (1.31)$$

(Правда, в этом случае существуют две линейно независимые собственные функции $a \exp(\pm in(\varepsilon)x)$ соответствующего оператора. Тем не менее, специфика этой задачи такова, что дополнительных трудностей не появляется). Итак, здесь $n(\varepsilon) = \left\{ \sqrt{\frac{z_0}{\varepsilon}} \right\} + \theta$, $u = \sqrt{\varepsilon} (\xi(\tau, x) \exp in(\varepsilon)x + \bar{\xi}(\tau, x) \exp(-in(\varepsilon)x) + \dots$,

$$\dot{\xi} = 4z_0\lambda[\xi'' + i\theta\xi' - \theta^2\xi] + c(A_1a, b)\xi + d|\xi|^2\xi,$$

$$\xi|_{x=0} = \xi|_{x=2\pi}, \quad \frac{\partial \xi}{\partial x}\Big|_{x=0} = \frac{\partial \xi}{\partial x}\Big|_{x=2\pi}, \quad (1.32)$$

а для постоянной d получаем равенство

$$d = (3F_3(a, a, a) - F_2(a, c_0) - F_2(c_0, a) - F_2(a, c_1) - F_2(c_1, a), b).$$

Грубым состояниям равновесия (периодическим решениям) краевых задач (1.28) — (1.30) и (1.32) соответствуют диссипативные структуры (периодические решения с тем же периодом) той же устойчивости исходной краевой задачи (1.26), (1.27) или (1.26), (1.31). Отметим еще, что просто находятся состояния равновесия задачи (1.32) вида $\xi = \xi_k \exp ikx$. Их устойчивость (неустойчивость) определяется неравенствами $d < 0$, $4z_0\lambda(2k + \theta)^2 < -2d$ ($d > 0$ или $4z_0\lambda(2k + \theta)^2 > -2d$).

3. Всюду ниже предполагаем, что матрица $A(z_0)$ имеет пару чисто мнимых значений $\pm i\omega_0$ ($\omega_0 > 0$), а все остальные ее собственные значения имеют отрицательные вещественные части. Пусть

$$A(z_0)a = i\omega_0 a, \quad A^*(z_0)b = -i\omega_0 b, \quad (a, b) = 1, \quad (\bar{a}, b) = 0,$$

$$n(\varepsilon) = \frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}} + \theta, \quad \theta = \theta(\varepsilon) = -\frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}} + \left\{ \frac{1}{\pi} \sqrt{\frac{z_0}{\varepsilon}} \right\} \in (-1, 0].$$

Из свойств $A(z)$ вытекают соотношения

$$\operatorname{Re} (Da, b) = 0, \tag{1.33}$$

$$\lambda_1 = \operatorname{Re} [(Da_1, b) - (Da, b)(a_1, b)] \geq 0, \tag{1.34}$$

а вектор a_1 является решением уравнения

$$[A(z_0) - i\omega_0 I]a_1 = Da - (Da, b)a.$$

В следующих двух пунктах рассмотрим отдельно условия $\lambda_0 \neq 0$ и $\lambda_0 = 0$, где

$$i\lambda_0 = (Da, b).$$

Как оказывается, квазинормальные формы для этих случаев существенно отличаются друг от друга.

4. Предположим здесь, что

$$\lambda_0 \neq 0 \quad \text{и} \quad \mu = c\sqrt{\varepsilon}.$$

В краевую задачу (1.26), (1.27) подставим выражение

$$\begin{aligned} u = \varepsilon^{1/4} [\xi(\tau, x) \cos \pi n(\varepsilon)x + \eta(\tau, x) \sin \pi n(\varepsilon)x] a \exp i\omega_0 t_1 + \\ + \varepsilon^{1/2} [v(\tau, x, \varepsilon) + w(\tau, x, \varepsilon) \exp i\omega_0 t_1] + \\ + K + \varepsilon^{3/4} u_3(\tau, x, \varepsilon) + \dots, \end{aligned} \tag{1.35}$$

где $\tau = \varepsilon^{1/2} t$, $t_1 = (1 + \varepsilon^{1/2} \alpha_1 + \dots)t$, а через K обозначено выражение, комплексно сопряженное к выписанным в (1.35) коэффициентам при $\varepsilon^{1/4}$ и $\varepsilon^{1/2}$. Конкретизируем функции $v(\tau, x, \varepsilon)$ и $w(\tau, x, \varepsilon)$:

$$v(\tau, x, \varepsilon) = v_1(\tau, x) + v_2(\tau, x) \cos 2\pi n(\varepsilon)x + v_3(\tau, x) \sin 2\pi n(\varepsilon)x, \tag{1.36}$$

$$\begin{aligned} w(\tau, x, \varepsilon) = w_1(\tau, x) + w_2(\tau, x) \cos 2\pi n(\varepsilon)x + \\ + w_3(\tau, x) \sin 2\pi n(\varepsilon)x. \end{aligned} \tag{1.37}$$

В результате стандартных действий на втором шаге находим, что

$$v_1 = (|\xi|^2 + |\eta|^2)v_{10}, \quad v_{10} = -\frac{1}{2}A_0^{-1}(F_2(a, \bar{a}) + F_2(\bar{a}, a)), \quad (1.38)$$

$$v_2 = (|\xi|^2 - |\eta|^2)v_{20}, \quad v_{20} = -\frac{1}{2}A^{-1}(4\pi^2 z_0)(F_2(a, \bar{a}) + F_2(\bar{a}, a)), \quad (1.39)$$

$$v_3 = (\xi\bar{\eta} + \bar{\xi}\eta)v_{20}, \quad (1.40)$$

$$w_1 = (\xi^2 + \eta^2)w_{10}, \quad w_{10} = -\frac{1}{2}(A_0 - 2i\omega_0 I)^{-1}F_2(a, a), \quad (1.41)$$

$$w_2 = (\xi^2 - \eta^2)w_{20}, \quad w_{20} = -\frac{1}{2}(A(4\pi^2 z_0) - 2i\omega_0 I)^{-1}F_2(a, a), \quad (1.42)$$

$$w_3 = \xi\eta w_{20}. \quad (1.43)$$

На третьем шаге из условия разрешимости получающейся краевой задачи относительно u_3 получаем краевую задачу гиперболического типа для $\xi(\tau, x)$ и $\eta(\tau, x)$:

$$\begin{aligned} \frac{\partial \xi}{\partial \tau} = i2\sqrt{z_0}\lambda_0 \left[\frac{\partial \eta}{\partial x} - 2\pi\theta\xi \right] + [c(A_1 a, b) - i\omega_0 \alpha_1] \xi + R_1 |\xi|^2 \xi + \\ + R_2 |\eta|^2 \xi + R_3 \bar{\xi} \eta^2, \end{aligned} \quad (1.44)$$

$$\begin{aligned} \frac{\partial \eta}{\partial \tau} = -i2\sqrt{z_0}\lambda_0 \left[\frac{\partial \xi}{\partial x} + 2\pi\theta\eta \right] + [c(A_1 a, b) - i\omega_0 \alpha_1] \eta + \\ + R_1 |\eta|^2 \eta + R_2 |\xi|^2 \eta + R_4 \bar{\eta} \xi^2, \end{aligned} \quad (1.45)$$

$$\left. \frac{\partial \xi}{\partial x} \right|_{x=0} = \left. \frac{\partial \xi}{\partial x} \right|_{x=1} = 0, \quad \eta|_{x=0} = \eta|_{x=1} = 0. \quad (1.46)$$

Приведем выражение для R_j ($j = 1, 2, 3, 4$):

$$R_1 = \left(\frac{3}{4}f_0 + F_2(a, v_{10}) + F_2(v_{10}, a) + F_2(\bar{a}, w_{10}) + F_2(w_{10}, \bar{a}) + \right.$$

$$\left. + \frac{1}{2}(F_2(\bar{a}, w_{20}) + F_2(w_{20}, \bar{a}) + F_2(a, v_{20}) + F_2(v_{20}, a)), b \right),$$

$$f_0 = F_3(a, a, \bar{a}) + F_3(a, \bar{a}, a) + F_3(\bar{a}, a, a),$$

$$R_2 = \left(\frac{1}{2}f_0 + F_2(a, v_{10}) + F_2(v_{10}, a) + \frac{1}{2}(F_2(\bar{a}, w_{20}) + F_2(w_{20}, \bar{a})), b \right),$$

$$R_3 = R_1 - R_2, \quad R_4 = R_3 - \frac{1}{2}(f_0, b).$$

Отметим еще, что устойчивость состояния равновесия $(\xi_0, 0)$ ($\xi_0^2 = -\frac{c \operatorname{Re}(A_1 a, b)}{\operatorname{Re} R_1}$) краевой задачи (1.44) — (1.46) и соответствующего периодического решения краевой задачи (1.26), (1.27) определяется корнями уравнений (относительно g)

$$\det \begin{pmatrix} \xi_0^2 R_1 - g & \xi_0^2 R_1 & i\lambda \tilde{m} & 0 \\ \xi_0^2 \bar{R}_1 & \xi_0^2 \bar{R}_1 - \bar{g} & 0 & -i\lambda \tilde{m} \\ i\lambda \tilde{m} & 0 & \delta + \xi_0^2 R_2 - g & \xi_0^2 R_4 \\ 0 & -i\lambda \tilde{m} & \xi_0^2 \bar{R}_4 & \bar{\delta} + \xi_0^2 \bar{R}_2 - \bar{g} \end{pmatrix} = 0, \quad (1.47)$$

где $\tilde{m} = 0, 1, \dots$; $\lambda = 2\sqrt{z_0}\pi\lambda_0$; $\delta = -2i\theta\lambda + c(A_1 a, b) - i\omega_0\alpha_1$; $i\omega_0\alpha_1 = -2\lambda\theta + c \operatorname{Im}(A_1 a, b) - \xi_0^2 \operatorname{Im} R_1$.

Для случая периодической краевой задачи

$$u|_{x=0} = u|_{x=2\pi}, \quad \left. \frac{\partial u}{\partial x} \right|_{x=0} = \left. \frac{\partial u}{\partial x} \right|_{x=2\pi}$$

просто определяются состояния равновесия вида $\xi_k \exp ikx$, $\eta_k \exp ikx$ системы уравнений (1.36), (1.37) с краевыми условиями

$$(\xi, \eta)|_{x=0} = (\xi, \eta)|_{x=2\pi}, \quad \left(\frac{\partial \xi}{\partial x}, \frac{\partial \eta}{\partial x} \right) \Big|_{x=0} = \left(\frac{\partial \xi}{\partial x}, \frac{\partial \eta}{\partial x} \right) \Big|_{x=2\pi}.$$

5. В этом пункте считаем, что

$$i\lambda_0 = (Da, b) = 0, \quad \mu = c\varepsilon$$

и выполнено условие невырожденности $\lambda_1 \neq 0$ (тем самым, согласно (1.34), имеем неравенство $\lambda_1 > 0$).

Положим в (1.26), (1.27)

$$\begin{aligned} u &= \varepsilon^{1/2} [\xi(\tau, x) \cos \pi n(\varepsilon)x + \eta(\tau, x) \sin \pi n(\varepsilon)x] a \exp i\omega_0 t_1 + \\ &+ \varepsilon [v(\tau, x, \varepsilon) + g(\tau, x, \varepsilon) \exp i\omega_0 t_1 + w(\tau, x, \varepsilon) \exp 2i\omega_0 t_1] + K + \\ &+ \varepsilon^{3/2} u_3(t, x, \varepsilon), \end{aligned}$$

где $\tau = \varepsilon t$, $t_1 = (1 + \varepsilon\alpha_1 + \dots)t$, K имеет тот же смысл, что и в (1.35), для $v(\tau, x, \varepsilon)$ и $w(\tau, x, \varepsilon)$ сохраняются выражения (1.36) — (1.43), а вектор-функция $g(\tau, x, \varepsilon)$ такова:

$$g(\tau, x, \varepsilon) = g_1(\tau, x) \cos \pi n(\varepsilon)x + g_2(\tau, x) \sin \pi n(\varepsilon)x,$$

$$g_1(\tau, x) = -2\sqrt{z_0} [\eta' - 2\pi\theta\xi]a_1,$$

$$g_2(\tau, x) = 2\sqrt{z_0} [\xi' + 2\pi\theta\eta]a_1.$$

Собирая коэффициенты при $\varepsilon^{3/2}$ и применяя стандартные условия разрешимости, получаем параболическую краевую задачу для $\xi(\tau, x)$, $\eta(\tau, x)$ — квазинормальную форму:

$$\begin{aligned} \frac{\partial \xi}{\partial \tau} = 4z_0\lambda_1 \left[\frac{\partial^2 \xi}{\partial x^2} + 4\pi\theta \frac{\partial \eta}{\partial x} - 4\pi^2\theta^2\xi \right] + [c_1(A_1a, b) - i\omega_0\alpha_1]\xi + \\ + R_1|\xi|^2\xi + R_2|\eta|^2\eta + R_3\bar{\xi}\eta^2, \end{aligned} \quad (1.48)$$

$$\begin{aligned} \frac{\partial \eta}{\partial \tau} = 4z_0\lambda_1 \left[\frac{\partial^2 \eta}{\partial x^2} + 4\pi\theta \frac{\partial \xi}{\partial x} - 4\pi^2\theta^2\eta \right] + [c(A_1a, b) - i\omega_0\alpha_1]\eta + \\ + R_1|\eta|^2\eta + R_2|\xi|^2\xi + R_4\bar{\eta}\xi^2, \end{aligned} \quad (1.49)$$

$$\left. \frac{\partial \xi}{\partial x} \right|_{x=0} = \left. \frac{\partial \xi}{\partial x} \right|_{x=1} = 0, \quad \eta|_{x=0} = \eta|_{x=1} = 0. \quad (1.50)$$

Для состояния равновесия $\xi = \xi_0 > 0$, $\eta_0 = 0$ (когда оно существует) имеет место равенство

$$\xi_0 = [(16\pi^2\theta^2z_0\text{Re } \lambda_1 - c\text{Re}(A_1a, b)(\text{Re } R_1)^{-1})^{-1/2}].$$

Отсюда заключаем, что при $\varepsilon \rightarrow 0$ возможна ситуация, когда это состояние равновесия то возникает (из точки $\xi = \eta = 0$), то опять «влияет» в ноль. Здесь учтено, что при $\varepsilon \rightarrow 0$ выражение $\theta = \theta(\varepsilon)$ многократно пробегает значения от -1 до 0 . Устойчивость рассматриваемого состояния равновесия (и соответствующего периодического решения краевой задачи (1.26), (1.27)) определяется корнями уравнений (относительно g) при $k = 0, 1, \dots$

$$\det \begin{pmatrix} R_1\xi_0^2 - g & \delta k & R_1\xi_0^2 & 0 \\ \delta(k - \theta) & \sigma + R_2\xi_0^2 - g & 0 & R_4\xi_0^2 \\ \bar{R}_1\xi_0^2 & 0 & \bar{R}_2\xi_0^2 - \bar{g} & \bar{\delta}k \\ 0 & \bar{R}_4\xi_0^2 & \bar{\delta}(k - \theta) & \sigma + \bar{R}_2\xi_0^2 - \bar{g} \end{pmatrix} = 0,$$

где $\delta = 16\pi^2z_0\theta\lambda_1$, $\sigma = \text{Re}(c(A_1a, b)) - \text{Im}(R_1\xi_0^2) - \text{Im}(\delta\theta)$.

6. Заключение. С помощью приведенных в этом параграфе квазинормальных форм обнаруживаются быстро осциллирующие по пространственной переменной установившиеся режимы, параметры которых относительно медленно меняются в пространстве и во времени. Так же, как это было сделано в п. 1.2, можно рассмотреть вопросы, связанные с бифуркационными явлениями в квазинормальных формах при изменении устойчивости состояний равновесия. Особый интерес, по-видимому, представляют бифуркации за счет изменения параметров θ и коэффициентов диффузии. Некоторые результаты на эту тему приведены в работах [7, 8, 9, 10, 11, 12]. Здесь на этом останавливаться не будем. Укажем еще, что приведенный выше формализм получения квазинормальных форм допускает обобщения и на случаи большего числа собственных значений $A(z_0)$ с нулевыми вещественными частями, и на случай пространственной переменной большей размерности (с заменой оператора $\frac{\partial^2}{\partial x^2}$ на соответствующий эллиптический оператор, например оператор Лапласа Δ) [7]. Новые интересные явления возникают при построении нормализованных уравнений в окрестности цикла для сингулярно возмущенных параболических систем [13, 14] и при анализе локальной динамики контрастных структур [13, 15].

2. Уравнение Гинзбурга — Ландау — нормальная форма для дифференциально-разностного уравнения второго порядка с большим запаздыванием

Здесь исследуется локальная динамика нелинейного дифференциально-разностного уравнения второго порядка с «большим» запаздыванием. Показано, что в окрестности «критических» случаев структура решений определяется динамикой уравнения типа Гинзбурга — Ландау, которое играет роль нормальной формы [16].

Уравнения второго порядка с нелинейными запаздывающими обратными связями являются одними из наиболее популярных математических моделей, описывающих процессы и явления во многих областях науки. В настоящей работе рассматривается вопрос о динамике дифференциально-разностного уравнения

$$\ddot{x} + a\dot{x} + x = F(x(t - T)). \quad (2.1)$$

Уравнения вида (2.1) возникают при изучении многих прикладных задач (см., например, [17—22]). Предполагается, что параметр

a положителен: $a > 0$, т. е. при $F \equiv 0$ решение (2.1) затухает при $t \rightarrow \infty$. Функцию $F(x)$ считаем достаточно гладкой и $F(0) = 0$. Основное предположение состоит в том, что

$$T \gg 1. \quad (2.2)$$

При условии (2.2) удобно произвести замену времени $t \rightarrow Tt$ и замену $x(Tt) \rightarrow x(t)$. В итоге приходим к уравнению

$$\varepsilon^2 \ddot{x} + \varepsilon a \dot{x} + x = F(x(t-1)), \quad (2.3)$$

где $\varepsilon = T^{-1} \ll 1$. Это уравнение является сингулярно возмущенным. Отметим, что решение задачи об асимптотике решений (2.3) на конечном отрезке изменения времени следует из [6].

Положим в (2.3) формально $\varepsilon = 0$. Для получающегося в результате этого разностного уравнения

$$x(t) = F(x(t-1)) \quad (2.4)$$

в настоящее время построена достаточно полная теория [23]. Из многих численных экспериментов известно, однако, что динамические свойства решений уравнений (2.3) (при малых ε) и (2.4) могут быть принципиально различны. В работе автора [24] причина этого явления объяснена для случая квазилинейной функции $F(x)$ для дифференциально-разностных уравнений первого порядка и для уравнений типа (2.3) при дополнительном условии

$$a > \sqrt{2}. \quad (2.5)$$

В настоящей работе основное внимание уделено наиболее интересному случаю, когда условие (2.5) места не имеет, т. е.

$$0 < a < \sqrt{2}. \quad (2.6)$$

Как и в [24], предполагается, что функция $F(x)$ квазилинейна:

$$F(x) = bx + \mu f(x) \quad \text{и} \quad 0 < \mu \ll 1. \quad (2.7)$$

Будет показано, что при определенных условиях вопрос о динамике уравнения (2.3) сводится к изучению глобальной структуры решений уравнения Гинзбурга — Ландау (коэффициенты которого не содержат малых параметров).

Методика исследования основана на применении бифуркационных методов. Коротко остановимся на ней. Известно, что поведение

решений (2.3) в ограниченной (при $\varepsilon, \mu \rightarrow 0$) области во многом определяется расположением корней характеристического квазиполинома линейного при $\mu = 0$ уравнения

$$\varepsilon^2 \lambda^2 + \varepsilon a \lambda + 1 = b \exp(-\lambda). \quad (2.8)$$

В том случае, когда все корни (2.8) лежат слева от мнимой оси и отделены от нее при $\varepsilon \rightarrow +0$, решения (2.3) (из произвольно фиксированной ограниченной при $\varepsilon \rightarrow 0$ области Ω) стремятся к нулю при $t \rightarrow \infty$. Если при каждом достаточно малом значении ε найдется корень с положительной вещественной частью, отделенной от нуля при $\varepsilon \rightarrow 0$, то решения (2.3) с увеличением t покидают область Ω .

Ниже будет рассмотрен пограничный случай, когда нет корней (2.8) с отделенной от нуля при $\varepsilon \rightarrow 0$ положительной вещественной частью и есть корень $\lambda(\varepsilon)$, вещественная часть которого стремится к нулю при $\varepsilon \rightarrow 0$. Важной особенностью этого случая является то обстоятельство, что из условия $\operatorname{Re} \lambda(\varepsilon) \rightarrow 0$ ($\varepsilon \rightarrow 0$) следует, что вещественные части бесконечного множества корней уравнения (2.8) тоже стремятся к нулю при $\varepsilon \rightarrow 0$. Тем самым в задаче об устойчивости реализуется критический случай бесконечной размерности. Напомним, что исследование динамики в случаях близких к критическим конечной размерности базируется на методах интегральных многообразий и нормальных форм. В данной ситуации эти методы оказываются неприменимыми непосредственно, однако их формализм существенно используется.

Сделаем одно замечание. Рассмотрим разностное уравнение (2.4) с дискретным временем

$$x_{n+1} = F(x_{n-1}), \quad n = 0, 1, \dots, \quad (2.9)$$

где $F(x)$ задана формулой (2.7). При $|b| < 1$ и $0 < \mu \ll 1$ все решения (2.9) из ограниченной (при $\mu \rightarrow 0$) области стремятся к нулю при $n \rightarrow \infty$, а при $|b| > 1$ и $0 < \mu \ll 1$ решения покидают фиксированную (ограниченную при $\mu \rightarrow 0$) область с ростом аргумента n . Рассмотрим критические случаи $b = 1$ и $b = -1$. Установившимися режимами в первом из них могут быть только состояния равновесия ξ_i , которые являются корнями уравнения $F(\xi) = 0$. Их устойчивость зависит от знака величины $F'(\xi)$. При $b = -1$ установившимися режимами могут быть только состояние равновесия $x = 0$ и циклы периода 2, количество которых совпадает с числом положительных корней уравнения $\Phi(\xi) = 0$, где $\Phi(\xi) = \frac{1}{2}[F(-\xi) - F(\xi)]$. Устойчивость цикла, отвечающего корню ξ_0 этого уравнения, определяется знаком величины $\Phi'(\xi_0)$.

Следующее замечание касается ограничений на функцию $F(x)$. В ряде задач математической физики вместо условия квазилинейности (2.7) более естественно рассмотреть нелинейную функцию

$$F(x) = (b + \varepsilon^2 b_1)x + cx^2 + dx^3 + O(x^4) \quad (2.10)$$

без малого параметра μ , но с дополнительным требованием локальности задачи, т. е. динамика уравнения (2.3) в случае (2.10) рассматривается лишь в достаточно малой (но независимой от ε) окрестности нулевого состояния равновесия фазового пространства $C_{[-1,0]} \times R^1$.

Ниже будут рассмотрены оба случая (2.7) и (2.10). В п. 2.1 исследуются медленно осциллирующие решения уравнения (2.3), а в п. 2.2 — быстро осциллирующие. Отметим, что запас динамических свойств быстро осциллирующих решений существенно богаче. В частности, могут наблюдаться стохастические колебания. В п. 2.3 изучаются близкие вопросы о локальной структуре решений уравнения

$$\ddot{x} + ax\dot{x} + x = \frac{d}{dt}F(x(t-T)), \quad T \gg 1,$$

где гладкая функция $F(x)$ определяется формулами (2.7) или (2.10). Полученные результаты существенно дополняют представления о динамике уравнений второго порядка с нелинейными обратными запаздывающими связями.

2.1. Динамика медленно осциллирующих решений

В этом разделе будем предполагать, что имеет место неравенство (2.5). Из него следует, что критический случай в задаче об устойчивости решений (при $\mu = 0$) уравнения (2.3) реализуется либо при $b = 1$, либо при $b = -1$. Рассмотрим отдельно каждый из этих случаев.

Пусть сначала $b = 1$ и для $F(x)$ имеет место формула (2.7). Тогда имеется бесконечно много корней $\lambda_k(\varepsilon)$ ($k = 0, \pm 1, \dots$) уравнения (2.8), таких, что $\lambda_k(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$, а все остальные корни этого уравнения лежат слева и отделены при $\varepsilon \rightarrow 0$ от мнимой оси комплексной плоскости. Отметим, что

$$\lambda_k(\varepsilon) = 2k\pi i + \varepsilon\lambda_{k_1} + \varepsilon^2\lambda_{k_2} + \dots,$$

где $\lambda_{k_1} = -2ak\pi i$; $\lambda_{k_2} = 4k^2\pi^2[1 - \frac{1}{2}a^2] + 2a^2k\pi i$. Согласно общей идеологии бифуркационных методов установившиеся режимы урав-

нения (2.3) могут формироваться в «окрестности» критических частот $2k\pi i$ ($k = 0, \pm 1, \dots$). В связи с этим введем в рассмотрение формальный ряд

$$x(t, \varepsilon, \mu) = \sum_{k=-\infty}^{\infty} [\xi_k(t, \varepsilon, \mu) + \dots] \exp 2k\pi i t + \\ + \mu \sum_{k=-\infty}^{\infty} z_k(t, \varepsilon, \mu) \exp 2k\pi i t + \dots \quad (2.11)$$

Здесь функции $\xi_k(t, \varepsilon, \mu)$ и $z_k(t, \varepsilon, \mu)$ таковы, что их производные по t малы при $\varepsilon, \mu \rightarrow 0$, а многочлен означает слагаемые более высокого порядка малости при $\varepsilon, \mu \rightarrow 0$. Подставляя (2.11) в (2.3) и используя стандартные процедуры, для медленно меняющихся амплитуд $\xi_k(t, \varepsilon, \mu)$ получим систему уравнений

$$\dot{\xi}_k = [\varepsilon \lambda_{k_1} + \varepsilon^2 \lambda_{k_2} + \dots] \xi_k + \mu \varphi_k(\xi), \quad (2.12)$$

в которой через $\varphi_k(\xi)$ обозначен коэффициент при $\exp 2ik\pi t$ в разложении функции $f(\sum_{k=-\infty}^{\infty} \xi_k \exp 2k\pi i t)$ в ряд Фурье. Отбросив слагаемые более высокого порядка малости, систему (2.12) можно записать в виде одного параболического уравнения ($z = \sum_{k=-\infty}^{\infty} \xi_k \exp 2k\pi i y$)

$$\frac{\partial z}{\partial t} = \varepsilon^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 z}{\partial y^2} - \varepsilon a (1 - \varepsilon a) \frac{\partial z}{\partial y} + \mu f(z)$$

с периодическими краевыми условиями $z(t, y) \equiv z(t, y + 1)$. Имеется в виду, что формальное разложение функции $z(t, y)$ в ряд Фурье по собственным функциям периодической краевой задачи совпадает с выражением (2.12). Отсюда в результате замен

$$z(t, y) = u(t, y - \varepsilon a (1 - \varepsilon a) t), \quad y - \varepsilon a (1 - \varepsilon a) t = \tau, \quad \mu t = \tau \quad (2.13)$$

приходим к краевой задаче

$$\frac{\partial u}{\partial \tau} = \delta^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 u}{\partial r^2} + f(u), \quad u(\tau, r) \equiv u(\tau, r + 1), \quad (2.14)$$

где $\delta = \varepsilon \mu^{-1/2}$. Из неравенства (2.5) следует параболичность этой краевой задачи. Устойчивыми решениями (2.14) могут быть только однородные состояния равновесия. Таким образом, в случае $b = 1$

и при фиксированном δ динамические свойства решений уравнения (2.9) и (2.3) совпадают. Отметим, что краевая задача (2.14) может иметь неоднородные состояния равновесия $u_0(r)$, а значит, уравнение (2.3) имеет периодическое решение $u_0((1 + a\varepsilon + o(\varepsilon))t) + o(1)$.

В том случае, когда функция $F(x)$ определена формулой (2.10), аналог краевой задачи (2.14) имеет вид

$$\frac{\partial u}{\partial \tau} = \delta^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 u}{\partial r^2} + b_1 u + c u^2, \quad u(\tau, r) \equiv u(\tau, r + 1). \quad (2.15)$$

Случай $b = -1$ более интересен. В отличие от предыдущего случая все стремящиеся к мнимой оси при $\varepsilon, \mu \rightarrow 0$ корни $\lambda_k(\varepsilon)$ уравнения (2.8) имеют вид

$$\lambda_k(\varepsilon) = (2k + 1)\pi i + \varepsilon \lambda_{k_1} + \varepsilon^2 \lambda_{k_2} + \dots,$$

где $\lambda_{k_1} = -a(2k + 1)\pi i$; $\lambda_{k_2} = (2k + 1)^2 \pi^2 (1 - \frac{1}{2} a^2) + a^2 (2k + 1)\pi i$.

В формуле, аналогичной (2.11), вместо множителя $2k$ в показателе экспоненты следует писать $(2k + 1)$, а в системе (2.12) тогда через $\varphi_k(\xi)$ обозначен коэффициент при $\exp(2k + 1)i\pi t$ в разложении функции $\Phi(\sum_{k=-\infty}^{\infty} \xi_k \exp(2k + 1)i\pi t)$ в ряд Фурье. Полученную систему уравнений можно записать в виде одного параболического уравнения

$$\frac{\partial z}{\partial t} = \varepsilon^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 z}{\partial y^2} - \varepsilon a(1 - \varepsilon a) \frac{\partial z}{\partial y} + \mu \Phi(z)$$

с антипериодическими краевыми условиями $z(t, y + 1) \equiv -z(t, y)$. После замен (2.13) в итоге приходим к краевой задаче

$$\frac{\partial u}{\partial \tau} = \delta^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 u}{\partial r^2} + \Phi(u), \quad u(\tau, r + 1) \equiv -u(\tau, r). \quad (2.16)$$

Теорема. Пусть при некотором фиксированном $\delta = \delta_0 > 0$ краевая задача (2.16) имеет состояние равновесия $\varphi_0(r)$ и только один характеристический показатель линеаризованной на $\varphi_0(r)$ задачи равен по модулю 1. Тогда существует такое $\mu_0 > 0$, что при $\mu \in (0, \mu_0)$ и $\varepsilon = \delta_0 \sqrt{\mu}$ уравнение (2.3) имеет периодическое с периодом $1 + o(1)$ решение $x_0(t, \varepsilon, \mu)$ той же, что и $\varphi_0(r)$, устойчивости, причем $x_0(t, \varepsilon, \mu) = \varphi((1 + \varepsilon(1 + o(1))))t) + o(1)$.

Отметим, что в том случае, когда функция $F(x)$ определена формулой (2.10), аналог краевой задачи (2.16) имеет вид

$$\frac{\partial u}{\partial \tau} = \delta^2 \left(\frac{1}{2} a^2 - 1 \right) \frac{\partial^2 u}{\partial r^2} + b_1 u + (d + c^2) u^3, \quad u(\tau, r + 1) \equiv -u(\tau, r). \quad (2.17)$$

Краевые задачи (2.14) — (2.17) играют роль нормальных форм для уравнения (2.3). Они представляют собой упрощенный вариант — из-за вещественности коэффициентов и переменной $u(\tau, r)$ — уравнения Гинзбурга — Ландау.

Коротко остановимся на случае, когда $a = \sqrt{2}$. Тогда уравнения (2.14) — (2.17) теряют смысл. Для построения соответствующей нормальной формы необходимо учесть в линейных членах слагаемые до порядка ε^4 . На этом пути получаем опять параболическое уравнение вида

$$\frac{\partial u}{\partial \tau} = -\frac{\sqrt{2}}{3} \frac{\partial^3 u}{\partial r^3} + 4\varepsilon \left[\frac{\partial^4 u}{\partial r^4} + \frac{1}{3} \frac{\partial^3 u}{\partial r^3} - \frac{\partial^2 u}{\partial r^2} \right] + \Psi(u)$$

с периодическими при $b = 1$ и антипериодическими при $b = -1$ краевыми условиями. Соответственно, $\Psi(u)$ совпадают с нелинейностями из (2.14), (2.15), (2.16) или (2.17), а $\tau = \varepsilon^3 t$.

Схема обоснования этого утверждения стандартна (см., например, [25]). Основным моментом ее является исследование уравнения, линеаризованного на асимптотическом по невязке периодическом решении $\tilde{x}_0(t, \varepsilon, \mu) = \varphi_0((1 + \varepsilon)t) + \varepsilon x_1(t, \varepsilon, \mu)$ уравнения (2.3). В задаче об устойчивости решений этого линейного уравнения реализуется критический (при $\varepsilon, \mu \rightarrow 0$) случай бесконечной размерности. Соответствующий алгоритм исследования основывается на методике работы [26]. В результате не только удается дать ответ об устойчивости, но и получить оценки (в норме $C_{[0,1]}$) нормы обратного оператора, определяемого исходным линейным уравнением и действующего из банахова пространства $C_{(-\infty, \infty)}$ ограниченных на всей оси функций в $C_{(-\infty, \infty)}$. Этой оценки достаточно для завершения обоснования теоремы. Детальное изложение описанной здесь схемы имеется в работе [5].

Отметим, что функция $\Phi(\xi)$ нечетна. Отсюда и из свойств консервативности уравнения $\delta\varphi'' + \Phi(\varphi) = 0$ легко получаем условия разрешимости краевой задачи

$$\delta\varphi'' + \Phi(\varphi) = 0, \quad \varphi(r) = -\varphi(r + 1). \quad (2.18)$$

Особый интерес представляет исследование случая, когда фигурирующий в (2.18) параметр δ является достаточно малым. При этом условии те решения (2.18), которые остаются ограниченными при $\delta \rightarrow 0$, имеют релаксационный вид, близкий к ступенчатой функции, а сами «ступени», т. е. участки медленного движения, близки к корням уравнения $\Phi(\xi) = 0$. Используя результаты работы [6], можно выписать детальную асимптотику соответствующих решений при $\delta \rightarrow 0$.

Обозначим через ξ_1, ξ_2, \dots занумерованные в порядке возрастания все положительные корни уравнения $\Phi(\xi) = 0$ и предположим, что все они простые. Отображение (2.4) при малых μ имеет циклы Ξ_1, Ξ_2, \dots периода 2 вида $X_j = \{\xi_j + O(\mu), -\xi_j + O(\mu)\}$. Из приведенных выше результатов вытекает, что при $\varepsilon \rightarrow 0$ уравнение (2.3) будет иметь периодическое решение $x(t, \varepsilon, \mu)$, близкое на участке медленного движения только к циклу Ξ_1 и только если $\Phi'(0) > 0$ (при неустойчивом нулевом состоянии равновесия). Отсюда, например, следует важность предположения о глобальной устойчивости цикла в приведенном в работе [23] достаточном условии существования периодического решения (2.3), «близкого» к соответствующему циклу (2.4).

Отметим еще, что при $\delta \rightarrow 0$ (и при $\Phi'(0) > 0$, $\Phi'(\xi_1) < 0$) количество состояний равновесия краевой задачи (2.16) неограниченно возрастает, но все они, кроме, возможно, состояний равновесия с двумя нулями на полуинтервале длины 2, неустойчивы. Равномерного предельного при $\varepsilon, \mu \rightarrow 0$ перехода в рассматриваемой задаче, конечно, нет.

Интересно было бы исследовать нестационарные установившиеся режимы (2.16), (2.17). Это позволило бы выявить существование структур нового типа для уравнения (2.3). В связи с этим отметим результаты работы [21], полученные для некоторого довольно близкого к (2.17) параболического уравнения.

Из формул, связывающих решения нормализованных уравнений (2.14) — (2.17) с решениями исходного уравнения (2.3), следует, что основной вклад в формирование структур (2.3) вносят коэффициенты ξ_k при модах с «небольшими» номерами. Поэтому в рассматриваемом случае решения названы медленно осциллирующими. Ниже речь пойдет о решениях совсем другого типа.

2.2. Быстро осциллирующие структуры

Ниже предполагаем, что выполнено неравенство (2.6). Вернемся к рассмотрению характеристического квазиполинома (2.8). Пусть при некоторых значениях a, b, ε он имеет чисто мнимый корень $\lambda = i\omega$. Тогда приходим к равенствам

$$1 - \delta^2 = b \cos \omega, \quad a\delta = -b \sin \omega, \quad \text{где } \varepsilon\omega = \delta.$$

Отсюда получаем, что $b^2 = (1 - \delta^2)^2 + a^2\delta^2$. Обозначим через $b^2(a)$ наименьшее по всем $\delta \geq 0$ значение правой части этого равенства: $b^2(a) = \min_{\delta \geq 0} [(1 - \delta^2)^2 + a^2\delta^2]$. Очевидно, что $b^2(a) = a^2(1 - a^2/4)$.

Отметим, что при условии (2.6) верно неравенство $0 < b^2(a) < 1$.

Имеет место простое утверждение.

Лемма. Пусть выполнены неравенства (2.6) и

$$b^2 < b^2(a), \quad (2.19)$$

тогда найдется такое $c > 0$, что при всех достаточно малых $\varepsilon > 0$ все корни $\lambda_j(\varepsilon)$ уравнения (2.8) удовлетворяют условию

$$\operatorname{Re} \lambda_j(\varepsilon) \leq -c. \quad (2.20)$$

Если же $b^2 > b^2(a)$, то при каждом достаточно малом ε уравнение (2.8) имеет корень $\lambda_j(\varepsilon)$, удовлетворяющий неравенству

$$\operatorname{Re} \lambda_j(\varepsilon) \geq c.$$

В том случае, когда

$$b^2 = b^2(a), \quad (2.21)$$

уравнение (2.8) имеет бесконечно много корней $\lambda_j(\varepsilon)$, для которых $\lim_{\varepsilon \rightarrow 0} \operatorname{Re} \lambda_j(\varepsilon) = 0$, и для каждого $\sigma > 0$ найдется такое $\varepsilon(\sigma) > 0$, что при $\varepsilon \in (0, \varepsilon(\sigma))$ все корни (2.8) лежат в полуплоскости

$$\operatorname{Re} \lambda \leq \sigma.$$

Исследуем динамику уравнения при условии (2.21). Отметим, что при этом условии тоже реализуется критический случай бесконечной размерности. Итак, пусть $b^2 = b^2(a)$. Введем несколько обозначений. Через Ω обозначим корень уравнения $\operatorname{tg} \Omega = -\sqrt{4a^{-2} - 2}$, причем в случае $b = b(a)$ считаем, что $\Omega \in (-\pi/2, 0)$, а при $b = -b(a)$ $\Omega \in (\pi/2, \pi)$. Пусть $\Theta = \Theta(\varepsilon)$ — такое значение из полуинтервала $[0, 2\pi)$, для которого выражение $\varepsilon^{-1}\sqrt{1 - a^2/2} + \Theta$ является целым кратным 2π ($\Theta = \varepsilon^{-1}\sqrt{1 - a^2/2} - 2\pi\{(2\varepsilon\pi)^{-1}\sqrt{1 - a^2/2}\}$). Наконец, положим

$$\omega_k(\varepsilon) = \varepsilon^{-1}\sqrt{1 - a^2/2} + \Theta + \Omega + 2\pi k \quad (k = 0, \pm 1, \pm 2, \dots).$$

Уравнение (2.8) имеет корни $\lambda_k(\varepsilon)$ ($k = 0, \pm 1, \pm 2, \dots$), для которых $\operatorname{Re} \lambda_k(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$ (а все остальные корни удовлетворяют оценке (2.20)). Для $\lambda_k(\varepsilon)$ верна формула

$$\lambda_k(\varepsilon) = i\omega_k(\varepsilon) + \varepsilon\lambda_{k_1} + \varepsilon^2\lambda_{k_2} + O(\varepsilon^3),$$

где

$$\begin{aligned}
 \lambda_{k_1} &= -2ia^{-1}[\Theta + \Omega + 2\pi k], \\
 \lambda_{k_2} &= -4\pi^2 k^2(d_1 + id_2) + 2i\pi k(c_1 + ic_2) + g_1 + ig_2, \\
 d_1 &= 4a^{-2}(4 - a^2)^{-1}(2 - a^2) > 0, \\
 d_2 &= -8a^{-1}(4 - a^2)^{-1}\sqrt{1 - a^2/2}, \\
 c_1 &= 4a^{-2} + 16a^{-1}(4 - a^2)^{-1}\sqrt{1 - a^2/2}(\Theta + \Omega), \\
 c_2 &= 4(2 - a^2)a^{-2}(4 - a^2)^{-1}(\Theta + \Omega), \\
 g_1 &= -4(2 - a^2)a^{-2}(4 - a^2)^{-1}(\Theta + \Omega)^2, \\
 g_2 &= 4a^{-2}(\Theta + \Omega) + 8a^{-1}(4 - a^2)^{-1}\sqrt{1 - a^2/2}(\Theta + \Omega).
 \end{aligned}$$

Положим

$$f_0(z) = (2\pi)^{-1} \int_0^{2\pi} f(z \exp i\tau + \bar{z} \exp(-i\tau)) \exp(-i\tau) d\tau$$

и введем в рассмотрение формальный ряд

$$\begin{aligned}
 x(t, \varepsilon) &= \exp(i\omega_0(\varepsilon)t) \sum_{k=-\infty}^{\infty} \xi_k(t, \varepsilon) \exp 2\pi i k + \\
 &+ \exp(-i\omega_0(\varepsilon)t) \sum_{k=-\infty}^{\infty} \bar{\xi}_k(t, \varepsilon) \exp(-2\pi i k) + \\
 &+ \varepsilon x_1(t, \varepsilon) + \mu x_2(t, \varepsilon) + \dots,
 \end{aligned} \tag{2.22}$$

где $\xi_k(t, \varepsilon)$ — медленно меняющиеся по t функции, для которых

$$\dot{\xi}_k = [\varepsilon \alpha_{k_1} + \varepsilon^2 \alpha_{k_2}] \xi_k + \mu \varphi_k. \tag{2.23}$$

Подставляя ряд (2.22) в (2.3) и учитывая (2.23), получаем, что

$$\alpha_{k_1} = \lambda_{k_1}, \quad \alpha_{k_2} = \lambda_{k_2},$$

а φ_k является коэффициентом при $\exp(2k\pi i)$ в разложении функции $f_0(\exp(-\Omega) \sum_{k=-\infty}^{\infty} \xi_k(t, \varepsilon) \exp(2\pi i k))$ в ряд Фурье.

Рассмотрим краевую задачу параболического типа

$$\begin{aligned}
 \frac{\partial z}{\partial t} &= \varepsilon[\varepsilon(c_1 + ic_2) - 2a^{-1}] \frac{\partial z}{\partial s} + \varepsilon^2[d_1 + id_2] \frac{\partial^2 z}{\partial s^2} + \\
 &+ \varepsilon^2 g_1 z + \mu f_0(\exp(-i\Omega)z),
 \end{aligned} \tag{2.24}$$

$$z(t, s + 2\pi) \equiv z(t, s).$$

Для коэффициентов разложения функции $z(t, s) = v(t, s) \times \exp[i\epsilon t(-2a^{-1}(\Theta + \Omega) + ig_2)]$ в ряд Фурье по переменной s получаем равенства (2.23), поэтому уравнение (2.24) считаем эквивалентной записью системы (2.23). Произведем затем в (2.24) замены

$$s \rightarrow s + \epsilon(\epsilon c_1 - 2a^{-1})t, \quad t \rightarrow \epsilon^2 t, \quad \mu = \sigma \epsilon^2, \\ z(\epsilon^2 t, s + \epsilon(\epsilon c_1 - 2a^{-1})t) = \exp(i\Omega)u(\tau, r).$$

В итоге получим краевую задачу

$$\frac{\partial u}{\partial \tau} = (d_1 + id_2) \frac{\partial^2 u}{\partial r^2} + ic_2 \frac{\partial u}{\partial r} + g_1 u + \alpha \sigma f_0(u), \quad (2.25)$$

$$u(\tau, r + 2\pi) \equiv u(\tau, r). \quad (2.26)$$

Здесь $\alpha = 2[a(a + 2i(1 - 1/2a^2)^{1/2})]^{-1}$.

Основной результат состоит в том, что краевая задача (2.25), (2.26) играет роль нормальной формы для уравнения (2.3). Ее нелокальные установившиеся режимы определяют локальную структуру решений (2.3) при малых ϵ . Так, например, периодическому (по τ) решению $u_0(\tau, r)$ отвечает с точностью до $O(\epsilon)$ тор $x_0(t, \epsilon)$ уравнения (2.3) той же, что и $u_0(\tau, r)$, устойчивости, причем $x(t, \epsilon) = y + \bar{y} + O(\epsilon)$, где

$$y = y(t, \epsilon) = \exp[it(\omega_0(\epsilon) - 2a^{-1}\epsilon(\Theta + \Omega) + \epsilon^2 g_2) + i\Omega]u_0(\epsilon^2(1 + O(\epsilon))t, [1 + \epsilon(\epsilon c_1 - 2a^{-1}) + O(\epsilon^2)]t).$$

В том случае, когда функция $F(x)$ определена формулой (2.10), аналогом краевой задачи (2.25), (2.26) является классическое уравнение Гинзбурга — Ландау

$$\frac{\partial u}{\partial \tau} = (d_1 + id_2) \frac{\partial^2 u}{\partial r^2} + ic_2 \frac{\partial u}{\partial r} + g_1 u + \\ + \alpha[3d + 2c((1 - b)^{-1} + \Delta^{-1})]|u|^2 u, \quad (2.27)$$

$$u(\tau, r + 2\pi) \equiv u(\tau, r), \quad (2.28)$$

где $\Delta = [2a^2 - 3 + 2ia\sqrt{1 - 1/2a^2} - \exp(-2i\Omega)] \exp(-2i\Omega)$. Связь между решениями $u(\tau, r)$ этой краевой задачи и $x(t, \epsilon)$ уравнения (2.3) устанавливает формула $x(t, \epsilon) = \epsilon v + \epsilon \bar{v} + O(\epsilon^2)$, где

$$v = v(t, \epsilon) = \epsilon \exp[it(\omega_0(\epsilon) - 2a^{-1}\epsilon(\Theta + \Omega) + \epsilon^2 g_2) + i\Omega]u_0(\epsilon^2(1 + O(\epsilon))t, [1 + \epsilon(\epsilon c_1 - 2a^{-1}) + O(\epsilon^2)]t).$$

Таким образом, приходим к выводу, что даже локальная динамика уравнения (2.3) является «достаточно сложной» [21]. В частности, могут наблюдаться нерегулярные колебания, содержащие быстро осциллирующие составляющие с медленно и хаотически меняющимися амплитудами. Отметим, что по структуре такие колебания близки к решениям параболических краевых задач с малым коэффициентом диффузии [7].

Остановимся еще на одном интересном моменте. Коэффициенты краевых задач (2.24) и (2.27) зависят от «внутреннего» параметра $\Theta = \Theta(\varepsilon)$. При $\varepsilon \rightarrow 0$ величина Θ бесконечно много раз пробегает полуинтервал $[0, 2\pi)$. Существование установившихся режимов и их устойчивость может зависеть от Θ , а следовательно, при $\varepsilon \rightarrow 0$ может наблюдаться неограниченный процесс «рождения» и «гибели» установившихся режимов уравнения (2.3).

2.3. Динамика уравнения (2.11)

При рассмотрении уравнения (2.11) с нелинейной функцией (2.7) сначала произведем замену времени $t \rightarrow Tt$, в результате которой получим уравнение

$$\varepsilon^2 \ddot{x} + \varepsilon a \dot{x} + x = \varepsilon \frac{dx}{dt} F(x(t-1)) \quad (\varepsilon = T^{-1} \ll 1). \quad (2.29)$$

При условии $|b| < a$ (и при достаточно малых ε) все корни характеристического квазиполинома

$$\varepsilon^2 \lambda^2 + \varepsilon a \lambda + 1 = b \varepsilon \lambda \exp(-\lambda) \quad (2.30)$$

имеют отрицательные (и отделенные от мнимой оси при $\varepsilon \rightarrow 0$) вещественные части. Тем самым, все решения из произвольной (но фиксированной при $\varepsilon \rightarrow 0$) окрестности нулевого состояния равновесия стремятся к нулю при $t \rightarrow \infty$. Если же $|b| > a$, то существует такое $\gamma > 0$, что при каждом $0 < \varepsilon \ll 1$ найдется корень, удовлетворяющий условию $\operatorname{Re} \lambda \geq \gamma$. Отсюда заключаем, что в рассматриваемой окрестности устойчивых решений уравнения (2.29) быть не может и найдется решение, покидающее эту окрестность при $t \rightarrow \infty$.

Ниже рассмотрим критический случай, когда

$$|b| = a. \quad (2.31)$$

Пусть $\Theta = \Theta(\varepsilon) \in [0, 2\pi)$ — такое, что величина $\varepsilon^{-1} + \Theta$ является целым кратным 2π . При условии (2.31) уравнение (2.30) имеет корни

$$\lambda_k = i[\varepsilon^{-1} + P_k] + \varepsilon \lambda_{k_1} + \varepsilon^2 \lambda_{k_2} + \dots \quad (k = 0, \pm 1, \pm 2, \dots),$$

где

$$\begin{aligned} P_k &= \begin{cases} \Theta + 2\pi k, & \text{если } b = a, \\ \Theta + 2\pi(k + 1), & \text{если } b = -a, \end{cases} \\ \lambda_{k_1} &= -2ia^{-1}P_k, \\ \lambda_{k_2} &= a^{-1}[-2a^{-1}P_k^2 - 3iP_k^2 + ia^{-1}P_k], \end{aligned}$$

а все остальные корни лежат слева и отделены от мнимой оси при $0 < \varepsilon \ll 1$. Положим

$$\varphi(z) = (2\pi)^{-1} \int_0^{2\pi} f'(z \exp i\tau + \bar{z} \exp(-i\tau)) [iz - i\bar{z} \exp(-2i\tau)] d\tau$$

и

$$f_0(\varepsilon) = \begin{cases} \varphi(z), & \text{если } b = a, \\ -\varphi(z), & \text{если } b = -a. \end{cases}$$

Далее рассмотрим формальный ряд (2.22), в котором полагаем

$$\omega_0(\varepsilon) = \begin{cases} \varepsilon^{-1} + \Theta, & \text{если } b = a, \\ \varepsilon^{-1} + \Theta + \pi, & \text{если } b = -a. \end{cases}$$

Подставляя его в (2.29) и учитывая (2.23), получаем систему уравнений для медленно меняющихся «амплитуд» ξ_k :

$$\dot{\xi}_k = [\varepsilon \lambda_{k_1} + \varepsilon^2 \lambda_{k_2} + \dots] \xi_k + \mu \varphi_k(\xi), \quad (2.32)$$

где $\varphi_k(\xi)$ — коэффициент при $\exp(2ik\pi t)$, если $b = a$ и коэффициент при $\exp((2k + 1)\pi t)$, если $b = -a$ в разложении функции $\varphi(z)$ в ряд Фурье.

Введем наконец в рассмотрение параболическое уравнение

$$\begin{aligned} \frac{\partial z}{\partial t} &= -2\varepsilon a^{-1} \left[\frac{\partial z}{\partial s} + i\Theta z \right] + \mu a^{-1} f_0(z) + \\ &+ \varepsilon^2 a^{-1} [(2a^{-1} + 3i) \frac{\partial^2 z}{\partial s^2} + (4ia^{-1}\Theta - 6\Theta) \frac{\partial z}{\partial s} - \Theta^2(a^{-1} - 3i)z] \end{aligned}$$

с краевыми условиями

$$\text{и } \begin{cases} z(t, s + 1) \equiv z(t, s) & \text{при } b = a, \\ z(t, s + 1) \equiv -z(t, s) & \text{при } b = -a. \end{cases}$$

Формальное разложение решения $z(t, s)$ этого уравнения в ряд по собственным функциям приводит к системе (2.32), поэтому приведенное уравнение считаем эквивалентной формой записи системы (2.32). После замен

$$s = r - 2\varepsilon a^{-1}t, \quad z = u \exp[ia^{-1}\Theta \varepsilon t(3\varepsilon\Theta - 2)], \quad t = \varepsilon^2\tau, \quad \mu = \varepsilon^2\sigma$$

получаем итоговое уравнение — нормальную форму:

$$\frac{\partial u}{\partial \tau} = a^{-1}[2a^{-1} + 3i] \frac{\partial^2 u}{\partial r^2} + 2\Theta[2ia^{-1} - 3] \frac{\partial u}{\partial r} - \Theta^2 a^{-1} u + \sigma a^{-1} f_0(u), \quad u(\tau, r+1) \equiv u(\tau, r) \operatorname{sign} b. \quad (2.33)$$

Полезно привести аналогичную краевую задачу при условии (2.10):

$$\frac{\partial u}{\partial \tau} = a^{-1}[2a^{-1} + 3i] \frac{\partial^2 u}{\partial r^2} + 2\Theta[2ia^{-1} - 3] \frac{\partial u}{\partial r} - \Theta^2 a^{-1} u + (a\Delta)^{-1}[3d + 4ic^2] \operatorname{sign} b \cdot |u|^2 u, \quad (2.34)$$

$$u(\tau, r+1) \equiv u(\tau, r) \operatorname{sign} b. \quad (2.35)$$

Как и в разобранных выше случаях, решение $u(\tau, r)$ краевых задач (2.33), (2.34), (2.35) определяет в главном структуру и асимптотику решений исходного уравнения (2.3). Обратим внимание, что динамические свойства решений уравнений Гинзбурга — Ландау (2.33) и (2.34), (2.35) могут быть достаточно сложными [21].

Некоторые более общие по сравнению с уравнением (2.1) задачи исследованы в [14, 27, 28].

Отметим, что разработанный в этой части подход применялся при изучении локальной динамики для ряда задач лазерной физики [29, 30, 31, 32].

Литература

1. Брюно А. Д. Локальный метод нелинейного анализа. М.: Наука, 1979.
2. Старжинский В. М. Прикладные методы нелинейных колебаний. М.: Наука, 1977.
3. Kuramoto Y., Tsuzuki T. On the formation of dissipative structures in reaction — diffusion systems // Prog. Theor. Phys. 1975. V. 54, № 3. P. 687—699.
4. Азрмеева Т. С., Малинецкий Г. Г. Простейшие типы упорядоченности в двумерных диссипативных системах: Препринт ИПМ им. Н. В. Келдыша АН СССР № 112. 1984. 28 с.
5. Васильева А. Б., Кащенко С. А., Колесов Ю. С., Розов Н. Х. Бифуркация автоколебаний нелинейных параболических уравнений с малой диффузией // Математический сборник. 1986. Т. 130(172), № 4(8). С. 488—499.
6. Васильева А. Б., Бутузов В. Ф. Асимптотические разложения решений сингулярно возмущенных уравнений. М.: Наука, 1973.

7. *Кащенко С. А.* Пространственные особенности высокомодовых бифуркаций двухкомпонентных систем с малой диффузией // Дифференц. уравн. 1989. Т. 25, № 2. С. 9.

8. *Кащенко С. А.* О квазинормальных формах для параболических уравнений с малой диффузией // ДАН СССР. 1988. Т. 299, № 5. С. 1049—1053.

9. *Кащенко С. А.* О коротковолновых бифуркациях в системах с малой диффузией // ДАН СССР. 1989. Т. 307, № 2. С. 269—273.

10. *Кащенко С. А.* Быстро осциллирующие бегущие волны в системах с малой диффузией // Дифференц. уравн. 1992. Т. 28, № 2. С. 254—262.

11. *Кащенко С. А., Воколишвили И. В., Малинецкий Г. Г., Потанов А. В.* Complex ordering and stochastic oscillations in a class of reaction-diffusion systems with small diffusion // J. Nonlinear Science. 1994. V. 4. P. 545—562.

12. *Kaschenko S. A.* Normalization in the systems with small diffusion // Int. J. Bifurcations and chaos. 1996. V. 6, № 7. P. 1093—1109.

13. *Кащенко С. А.* О нормализации в окрестности цикла систем параболических уравнений с малой диффузией // Украинский матем. журн. 1991. Т. 43, № 9. С. 1155—1161.

14. *Кащенко С. А.* Бифуркационные особенности сингулярно возмущенного уравнения с запаздыванием // Сибирский матем. журн. 1999. Т. 40, № 3. С. 567—572.

15. *Кащенко С. А.* Локальная динамика двухкомпонентных контрастных структур в окрестности точки бифуркации // ДАН СССР. 1990. Т. 312, № 2. С. 345—350.

16. *Кащенко С. А.* Уравнения Гинзбурга — Ландау — нормальная форма для дифференциально-разностного уравнения второго порядка с большим запаздыванием // Вычисл. матем. и матем. физ. 1998. Т. 38, № 3. С. 457—465.

17. *Хэйл Дж.* Теория функционально-дифференциальных уравнений. М.: Наука, 1984.

18. *Ланда П. С.* Автоколебания в распределенных системах. М.: Наука, 1983.

19. *Дмитриев А. С., Кислов В. Я.* Стохастические колебания в радиофизике и электронике. М.: Наука, 1989.

20. *Кузнецов С. П.* Сложная динамика генераторов с запаздывающей обратной связью (обзор) // Изв. вузов. Радиофизика. 1982. Т. 25, № 12. С. 1410—1428.

21. *Ахромеева Т. С., Курдюмов С. П., Малинецкий Г. Г., Самарский А. А.* Нестационарные структуры и диффузионный хаос. М.: Наука, 1992.

22. *Колмановский В. Б., Носов В. Р.* Устойчивость и периодические режимы регулируемых систем с последействием. М.: Наука, 1981. 448 с.

23. *Шарковский А. Н., Майстренко Ю. Л., Романенко Е. Ю.* Разностные уравнения и их приложения. Киев: Наукова думка, 1981.

24. *Кащенко С. А.* Применение метода нормализации к изучению динамики дифференциально-разностных уравнений с малым множителем при производной // Дифференц. уравн. 1989. Т. 25, № 8. С. 1448–1451.
25. *Stokes A.* On the approximation of nonlinear oscillation // Тр. 5-й Международной конф. по нелинейным колебаниям. Киев: Наукова думка, 1970. С. 480–491.
26. *Кащенко С. А.* Исследование устойчивости решений линейных параболических уравнений с близкими к постоянным коэффициентами и малой диффузией // Тр. семинара им. И.Г. Петровского. 1991. Вып. 15. С. 128–155.
27. *Кащенко С. А.* Исследование локальной динамики нелинейных сингулярно возмущенных систем с запаздыванием // Известия РАЕН. Сер. МММИУ. 1998. Т. 2, № 2. С. 5–68.
28. *Кащенко С. А.* Локальная динамика нелинейных сингулярно возмущенных систем с запаздыванием // Дифференц. уравн. 1999. Т. 35, № 10. С. 1343–1355.
29. *Григорьева Е. В., Кащенко С. А.* Нормальные и квазинормальные формы для полуклассических лазерных уравнений с дифракцией // Изв. вузов. ПНД. 1995. Т. 3, № 1. С. 59–72.
30. *Bestehorn M., Grigorieva E. V., Haken H., Kaschenko S. A.* Order parameters for class-B lasers with a long time delayed feedback // Phys. D. 2000. V. 145. P. 111–129.
31. *Григорьева Е. В., Кащенко С. А.* Параметры порядка в моделях лазеров с запаздывающей обратной связью // Сб. статей ИПМ им. М. В. Келдыша. М.: РАН, 2001. 26 с.
32. Новое в синергетике: взгляд в третье тысячелетие / Отв. ред. сер. акад. И. М. Макаров. Отв. ред. издания проф. Г.Г. Малинецкий и чл.-корр. РАН С.П. Курдюмов. М.: РАН; Наука, 2002. 478 с.

СТРАННЫЙ НЕХАОТИЧЕСКИЙ АТТРАКТОР

С. П. Кузнецов, А. С. Пиковский, У. Фойдель

Концепция странного нехаотического аттрактора (СНА) введена в 1984 г. в работе Гребужи, Отта, Пеликана и Йорке [1] и относится к специальному классу диссипативных систем — системам с квазипериодическим внешним воздействием. В них мгновенное состояние задается набором некоторого числа обобщенных координат, относящихся собственно к системе, и некоторого числа фазовых переменных по количеству несоизмеримых частот внешнего воздействия. Эпитет «*странный*» противопоставляет СНА аттрактору в виде тора, имеющему гладкую зависимость координатных переменных от фазовых. Эпитет «*нехаотический*» противопоставляет его странному хаотическому аттрактору: для СНА отсутствует экспоненциальная неустойчивость траекторий и нет положительного показателя Ляпунова. Оказалось, что в области между порядком и хаосом СНА типичны в системах с квазипериодическим воздействием [2—5]. Можно сказать, что при переходе от простой динамики к сложной в этих системах сначала возникает «странность», а уже потом хаос.

1. Примеры систем со странным нехаотическим аттрактором

В современной нелинейной динамике на равноправной основе рассматривают системы с непрерывным и с дискретным временем. Первые задаются дифференциальными уравнениями, вторые — итерировуемыми отображениями. Чтобы СНА мог реализоваться в системе с непрерывным временем, в дифференциальном уравнении должна присутствовать квазипериодическая зависимость коэффициентов от времени, что требует наличия как минимум двух несоизмеримых частот. В системах, заданных отображениями, есть собственный внутренний «ритм» — шаги дискретного времени, так что для добавляемого внешнего воздействия достаточно одной частоты, несоизмеримой с частотой шагов. Простейший класс систем, демонстрирующих СНА, — отображения с одной координатной и одной фазовой переменной. При этом внешнее воздействие характеризуется одним частотным параметром, который при проведении конкретных исследований чаще всего полагают равным «золотому среднему» — числу $(\sqrt{5}-1)/2$ ¹.

¹ Этот выбор обусловлен не только упрощением теоретического анализа, но что обычно ссылаются, но также и тем, что тонкие структуры, выявление которых составляет предмет экспериментов и численных расчетов, оказываются лучше различимыми, чем это имеет место при других соотношениях частот.

1.1. Бифуркация вилки в системе с квазипериодическим воздействием

Возьмем отображение $x_{n+1} = \lambda \operatorname{th} x_n$, которое демонстрирует бифуркацию вилки: при $\lambda < 1$ аттрактором служит неподвижная точка в нуле, при $\lambda > 1$ она становится неустойчивой, но появляются две симметрично расположенных устойчивых точки $x = \pm x_0(\lambda)$. Следуя работе [1], модифицируем его, введя квазипериодическую зависимость коэффициента от дискретного времени:

$$x_{n+1} = 2\lambda \operatorname{th} x_n \cos 2\pi\theta_n, \quad \theta_{n+1} = \theta_n + w \pmod{1}, \quad (1)$$

где θ — фазовая переменная. Параметр $w = (\sqrt{5} - 1)/2$ определяет частоту внешнего воздействия, которое носит в данном случае мультипликативный характер.

Вычислим показатель Ляпунова для решения $x \equiv 0$. В линейном приближении

$$x_{n+1} = 2\lambda x_n \cos 2\pi\theta_n$$

и

$$\ln |x_N/x_0| = \sum_{n=0}^{N-1} \ln |2\lambda x_n \cos 2\pi\theta_n|.$$

В силу иррациональности w , значения фазы θ_n при $N \rightarrow \infty$ равномерно и плотно распределены на единичном интервале. Поэтому сумму заменяем интегралом и для показателя Ляпунова имеем

$$\begin{aligned} \Lambda &\equiv N^{-1} \ln |x_N/x_0| = N^{-1} \sum_{n=0}^{N-1} \ln |2\lambda \cos 2\pi\theta_n| \equiv \\ &\equiv \int_0^1 \ln |2\lambda \cos 2\pi\theta| \, d\theta = \ln \lambda. \end{aligned} \quad (2)$$

При $\lambda < 1$ получаем $\Lambda < 0$, т. е. решение устойчиво, и линия $x = 0$ на плоскости (θ, x) служит аттрактором. При $\lambda > 1$ оно характеризуется положительным показателем Ляпунова и аттрактором не является.

Что можно сказать об аттракторе при $\lambda > 1$?

Во-первых, он непременно содержит точки с $|x| > 0$.

Во-вторых, поскольку $|\operatorname{th} x| < 1$, аттрактор расположен в полосе $|x| < 2\lambda$.

В-третьих, на оси $x = 0$ присутствует множество точек, принадлежащих аттрактору.

В самом деле, при $\theta = 1/4$ косинус обращается в ноль, так что для всех $\theta_n = 1/4 + n w \pmod{1}$, где $n = 1, 2, \dots, \infty$, имеем $x_n = 0$. Это подмножество точек аттрактора бесконечное и всюду плотное, так что аттрактор в целом представляет собой нетривиальный геометрический объект, «странное» множество.

На рис. 1 представлены изображения СНА в системе (1), фазовый портрет на плоскости (θ_n, x_n) и итерационная диаграмма в координатах (x_n, x_{n+1}) .

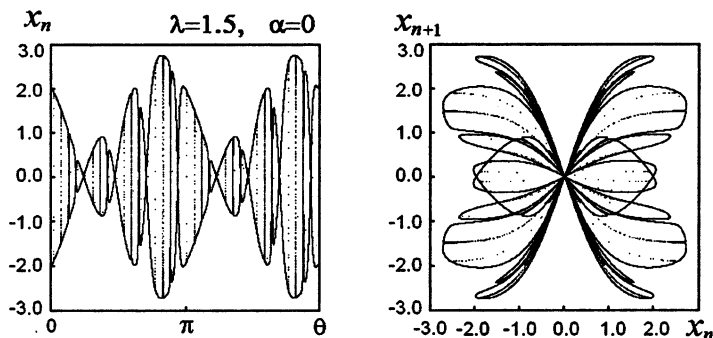


Рис. 1. Портреты странного нехаотического аттрактора модели (1) на фазовой плоскости и итерационной диаграмме при $\lambda = 1,5$.

Убедимся, что показатель Ляпунова для данного аттрактора отрицательный. Заметим, что для функции $f(x) = \text{th } x$ справедливо соотношение $|f'(x)| \leq |f(x)/x|$, причем равенство достигается лишь при $x = 0$. Для ляпуновского показателя получаем оценку

$$\Lambda = \lim_{N \rightarrow \infty} N^{-1} \sum_{n=0}^{N-1} \ln |2\lambda f'(x_n) \cos 2\pi\theta_n| < \lim_{N \rightarrow \infty} N^{-1} \sum_{n=0}^{N-1} \ln |x_{n+1}/x_n| = 0. \quad (3)$$

Таким образом, «странность» аттрактора *не связана* с неустойчивостью по отношению к начальным условиям, как это имеет место для хаоса.

2.2. Квадратичное отображение с квазипериодическим воздействием

Одна из популярных и содержательных моделей в нелинейной динамике — квадратичное отображение $x_{n+1} = \lambda - x_n^2$, в котором при увеличении параметра λ имеет место переход к хаосу через последовательность бифуркаций удвоения периода [6—9].

Вводя переменную добавку к λ , осциллирующую с иррациональной частотой w и амплитудой ϵ , приходим к модели [10—12]

$$x_{n+1} = \lambda - x_n^2 + \epsilon \cos 2\pi\theta_n, \quad \theta_{n+1} = \theta_n + w \pmod{1}, \quad (4)$$

в которой при некотором выборе параметров возникает СНА. На рис. 2 показана карта режимов на плоскости параметров (λ, ϵ) , а по периферии — итерационные диаграммы с портретами аттракторов в нескольких

характерных точках. Светло-серыми тонами обозначены области, где реализуются тор-аттракторы типа T1, T2, T4, изображающиеся на итерационной диаграмме соответствующим количеством замкнутых кривых, более темными — области ШНА и хаоса.

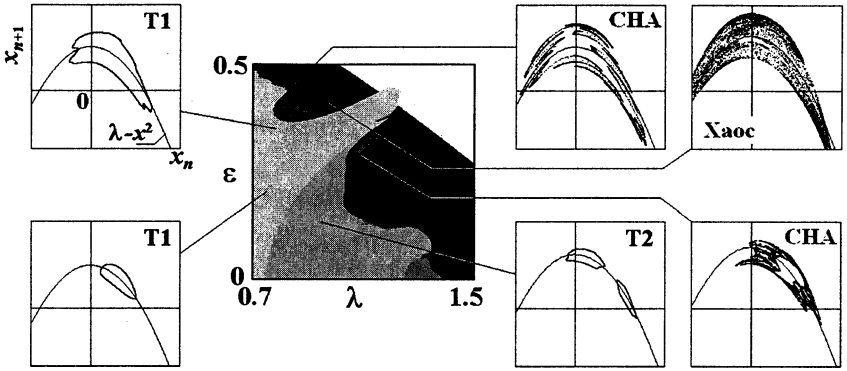


Рис. 2. В центре — карта режимов на плоскости параметров модели (4), по периферии — итерационные диаграммы для аттракторов в некоторых характерных точках. Частотный параметр $w = (\sqrt{5} - 1)/2$. Белым показана часть плоскости параметров, где наблюдается уход траекторий на бесконечность.

2.3. Модель с касательной бифуркацией

Рассмотрим отображение $x_{n+1} = f(x_n) + b$, где функция f выбрана так, что имеет место касательная бифуркация: при изменении b две неподвижные точки, устойчивая и неустойчивая, сближаются, сливаются и исчезают, оставляя узкий «коридор», прохождение которого требует большого числа итераций. Если динамика после прохождения «коридора» такова, что траектория обязательно вернется к месту входа в него, то возникает тип динамического поведения, известный как перемежаемость [13]. Имея целью локальный анализ бифуркации, функцию можно задать в виде $f(x) = x/(1 - x)$. Предполагая, что параметр b содержит квазипериодическую добавку с амплитудой ϵ , приходим к модели [14, 15]:

$$x_{n+1} = x_n / (1 - x_n) + b + \epsilon \cos 2\pi\theta_n, \quad \theta_{n+1} = \theta_n + w \pmod{1}. \quad (5)$$

Она пригодна для изучения перехода, являющегося аналогом касательной бифуркации и ведущего к рождению ШНА, но для описания ШНА как такового нуждается в модификации, вводящей механизм возврата. Интересно, что заменой $x_n = 1 - \psi_n / \psi_{n-1}$ уравнение (5) приводится к линейному разностному уравнению Харпера

$$\psi_{n+1} + \psi_{n-1} - 2\psi_n + (b + \varepsilon \cos 2\pi n w) \psi_n = 0, \quad (6)$$

которое известно в физике твердого тела в контексте анализа локализации и делокализации квантовых состояний в одномерной модели квазипериодической среды [14, 16].

2.4. Отображение окружности с квазипериодическим воздействием

Отображением окружности принято называть модель с дискретным временем

$$x_{n+1} = x_n + r + (K/2\pi) \sin 2\pi x_n,$$

где переменная x_n определена с точностью до целой части («по модулю 1»). Это можно интерпретировать как задачу об автогенераторе, на который действует периодическая последовательность импульсных толчков: x_n определяет фазу колебаний перед n -м толчком, K характеризует интенсивность толчков, а r — отклонение частоты следования импульсов от частоты автоколебаний. Добавляя член, отвечающий дополнительно воздействию с несоизмеримой частотой, приходим к модели [2, 17]:

$$\begin{aligned} x_{n+1} &= x_n + r + (K/2\pi) \sin 2\pi x_n + (\varepsilon/2\pi) \sin 2\pi \theta_n \pmod{1}, \\ \theta_{n+1} &= \theta_n + w \pmod{1}. \end{aligned} \quad (7)$$

Даже в отсутствие дополнительного члена отображение окружности демонстрирует разнообразную и богатую динамику. На плоскости параметров (r, K) можно видеть языки Арнольда, где динамика периодическая, и области квазипериодического поведения и хаоса (рис. 3, а). Нижняя часть плоскости параметров, где реализуется регулярная динамика (периодическая и квазипериодическая), и верхняя, где возможен хаос, разграничены критической линией $K = 1$ (пунктир).

При включении квазипериодического воздействия вместо критической линии возникает критическая зона (рис. 3, б). Ее ширина тем больше, чем больше амплитуда воздействия. Языки Арнольда трансформируются в зоны квазипериодической динамики с двумя несоизмеримыми частотами (частота шагов и частота воздействия), которые отвечают аттракторам в виде двумерных торов, а области между языками соответствуют теперь трехчастотной квазипериодичности или трехмерным торам (добавляется частота «прокручивания» циклической переменной x). В части плоскости параметров ниже критической зоны присутствуют указанные два типа регулярной динамики. Выше критической зоны имеются области хаоса. Наконец, в самой критической зоне в зависимости от параметров могут реализоваться двухчастотная квазипериодичность или СНА (см. пример на рис. 3, в).

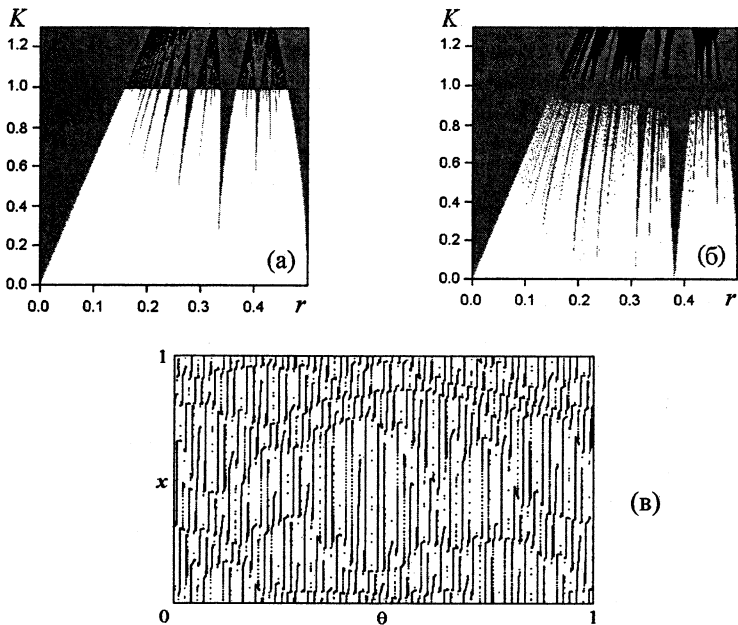


Рис. 3. Карты режимов отображения окружности в отсутствие и в присутствии квазипериодического воздействия: *а* — $C = 0$, светло-серым показаны области периодического поведения (языки Арнольда), белым — квазипериодические режимы, темно-серым — хаос; *б* — $C = 0,6$, светло-серые и белые области отвечают, соответственно, двумерному и трехмерному тор-аттрактору; *в* — СНА на фазовой плоскости (θ, x) при $C = 0,6, K = 1, r = 0,3$.

3. Свойства и количественные характеристики странных нехаотических аттракторов

3.1. Показатели Ляпунова

Для моделей в виде одномерных отображений с квазипериодическим воздействием нетривиальный показатель Ляпунова Λ один. Для СНА, как и для тора, он отрицательный. Поскольку аттрактор состоит из множества траекторий, не исключено присутствие на нем индивидуальных неустойчивых траекторий, у которых показатель Ляпунова положительный (их должно быть в определенном смысле мало, множество меры ноль, чтобы усредненная по аттрактору величина показателя Ляпунова была все же отрицательной). Например, для модели (1) при $\lambda > 1$ множество точек на оси $x = 0$ с фазами $\theta = 1/4 + 2\pi n\omega$ образует как раз такую исключительную траекторию. Именно это свойство является от-

личительной чертой СНА по сравнению с тором-аттрактором. В численных расчетах о наличии неустойчивых траекторий на аттракторе можно судить, рассматривая статистику *локальных показателей Ляпунова, определенных на конечном времени* [18, 19].

Для одномерного отображения с квазипериодическим воздействием $x_{n+1} = f(x_n, \theta_n)$, $\theta_{n+1} = \theta_n + w \pmod{1}$ определим величину

$$\Lambda_N = T^{-1} \ln |\partial x_N / \partial x_0| = T^{-1} \sum_{n=0}^{N-1} \ln |f'(x_n, \theta_n)|,$$

где штрихом обозначена производная функции по первому аргументу, а начальная точка (x_0, θ_0) предполагается расположенной на аттракторе. Проводя вычисления многократно, получим набор величин Λ_N , для которого можно ввести функцию распределения $F(\Lambda)$.

Рассмотрим в качестве примера квадратичное отображение (4).

На рис. 4, а показан портрет СНА при $\lambda = 0,8$, $\varepsilon = 0,45$ с показателем Ляпунова $\Lambda = -0,010$.

На рис. 4, б показаны полученные численно графики функций распределения локального показателя Λ_N при $N = 250, 500$ и 1000 . Они имеют вид колоколообразных кривых, причем точка расположения максимума приблизительно соответствует величине Λ . Ширина распределения убывает с ростом N , однако оно имеет «хвост», простирающийся в область положительных¹ значений Λ_N . Это указывает на существование на аттракторе локально неустойчивых участков траекторий.

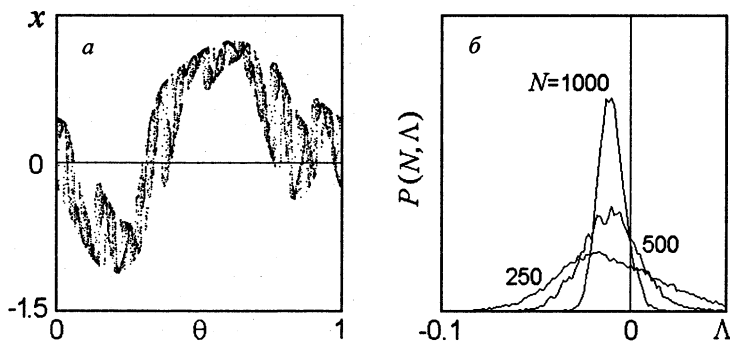


Рис. 4. Портрет СНА в отображении (4) при $\lambda = 0,8$, $\varepsilon = 0,45$ (а) и графики функций распределения локального показателя Ляпунова Λ_N (б).

¹ Заметим, что доля испытаний, приводящих к положительным значениям локального показателя Ляпунова Λ_N , уменьшается и, очевидно, стремится к нулю с ростом N .

3.2. Показатель фазовой чувствительности

Рассмотрим возможность выявить на количественном уровне наличие характерной для СНА негладкой зависимости координатной переменной от фазы воздействия. Для этого естественно попытаться отследить временную эволюцию величины $D_n = |\partial x_n / \partial \theta_0|$ [18]. Схема вычислений строится как совместные итерации исходного отображения и соотношения, полученного дифференцированием по фазе¹:

$$x_{n+1} = f(x_n, \theta_n), \quad \theta_{n+1} = \theta_n + w \pmod{1},$$

$$\partial x_{n+1} / \partial \theta_0 = f_x(x_n, \theta_n) \partial x_n / \partial \theta_0 + f_\theta(x_n, \theta_n).$$

На рис. 5, а показан полученный для случая СНА график зависимости D_N от числа итераций. Он выглядит как набор пиков, причем высота максимального пика, зарегистрированного в течение N итераций, растет с увеличением N .

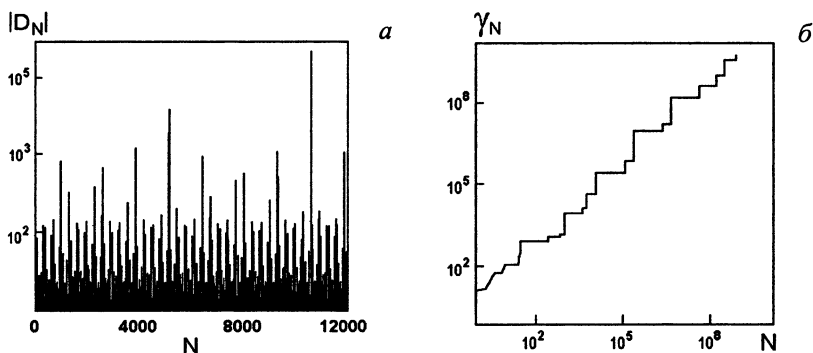


Рис. 5. Зависимости абсолютной величины производной $D_N = |\partial x_N / \partial \theta_0|$ от числа итераций для модели (1) при $\lambda = 1,5$, $w = (\sqrt{5} - 1)/2$ (а) и максимума высоты зарегистрированного пика от времени наблюдения (б) [18].

На рис. 5, б представлен график функции $\gamma_N(x_0, \theta_0, N) = \max_{0 \leq n \leq N} D_n$ в двойном логарифмическом масштабе, из которого видно, что рост происходит в среднем по степенному закону, $\gamma_N(x_0, N) \propto N^{\mu(x_0, \theta_0)}$, причем показатель степени, вообще говоря, зависит от выбранной траектории. Чтобы получить характеристику, относящуюся к аттрактору в целом, в работе [18] предложено определить показатель фазовой чувствительности

¹ Можно показать, что в случае СНА результат для D становится не зависящим от начального условия D_0 при большом числе шагов итераций в силу того, что нетривиальный показатель Ляпунова отрицательный.

сти μ как показатель роста для траектории, реализующей минимум этого показателя, $\Gamma_N = \min_{x_0, \theta_0} \gamma_N(x_0, \theta_0) \propto N^\mu$. Для рассмотренного примера $\mu \approx 0,97$. С другой стороны, для тора-аттрактора, очевидно, $\mu = 0$. Таким образом, в принципе показатель фазовой чувствительности дает средство отличить СНА от тора.

3.3. Фрактальная структура и размерность СНА

Хотя квалификация СНА как «геометрически странного» объекта — один из главных моментов в его определении, вопрос о фрактальных свойствах СНА изучен довольно слабо. Согласно работе [20] СНА в модели с бифуркацией вилки (1) и в отображении окружности (7) имеет фрактальную размерность (емкость)

$$D_0 = \lim_{\epsilon \rightarrow 0} [-\log N(\epsilon) / \log \epsilon] = 2,$$

а информационную размерность

$$D_1 = \lim_{\epsilon \rightarrow 0} \left[\sum_{i=1}^N p_i \log p_i / \log \epsilon \right] = 1.$$

Здесь N — число элементов покрытия ячейками (боксами) размера ϵ ; p_i — вероятность пребывания в i -й ячейке покрытия. На рис. 6 показан портрет аттрактора и график, использованный для оценки размерностей в отображении окружности.

Сведения об оценке размерностей СНА для некоторых моделей в экспериментах и численных расчетах имеются в работах [21—23].

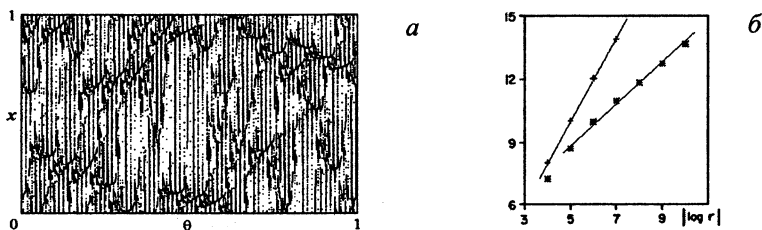


Рис. 6. Фазовый портрет СНА в (7) при $K = 0,95$, $\epsilon = 1,2$, $r = 0,2841$ (а) и график в двойном логарифмическом масштабе, использованный для оценки размерностей из работы [20] (б). Крестиками обозначены точки, относящиеся к оценке емкости D_0 , а звездочками — к оценке информационной размерности D_1 .

3.4. Спектральные свойства СНА

Фурье-анализ — один из общепринятых способов обработки сигналов при изучении динамических процессов, в том числе в эксперименте.

Говоря о типах спектров, будем рассуждать следующим образом. Имея некоторую последовательность x_n , подлежащую анализу, построим «накапливающуюся» сумму $S(\Omega, N) = \sum x_n e^{i\Omega n}$, где Ω — параметр, и отследим зависимость комплексной величины $S(\Omega, N)$ от числа членов суммы N . Для периодических и квазипериодических последовательностей на некотором дискретном множестве значений Ω получаем $S(\Omega, N) \propto N$, т. е. $|S(\Omega, N)|^2 \propto N^2$, а для остальных Ω имеем $S(\Omega, N) \rightarrow 0$. Это *дискретный спектр*. Для случайных или хаотических последовательностей эволюция величины S имеет характер случайного блуждания на комплексной плоскости, при этом $|S(\Omega, N)| \propto \sqrt{N}$ и $|S(\Omega, N)|^2 \propto N$. Это *непрерывный спектр*, такие спектры рассматриваются в теории стационарных случайных процессов Винера — Хинчина. Для последовательностей, порождаемых СНА, блуждание изображающей точки на плоскости при построении сумм $S(\Omega, N)$ носит фрактальный характер, образуя структуры с нетривиальными свойствами масштабного подобия (скейлинга). Асимптотическое поведение при увеличении N дается соотношением $|S(\Omega, N)|^2 \propto N^\beta$, где показатель β располагается между 1 и 2 и зависит от Ω . В этом случае говорят о *сингулярно-непрерывном спектре* [25, 26].

3.5. Метод рациональных аппроксимаций и природа СНА

Любое иррациональное число из интервала (0,1) представляется в виде цепной дроби $\rho = 1/(a_1 + 1/(a_2 + 1/(a_3 + \dots))) \equiv \langle a_1, a_2, a_3, \dots \rangle$, где a_i — целые числа. Подходящая дробь порядка k получается, если оборвать последовательность элементов на k -й позиции: $\rho_k = \langle a_1, \dots, a_k \rangle = p_k/q_k$. При этом $\lim_{k \rightarrow \infty} \rho_k = \rho$, и члены последовательности рациональных чисел ρ_k доставляют оптимальные приближения иррационального числа. Для «золотого среднего» имеем $(\sqrt{5} - 1)/2 = \langle 1, 1, \dots \rangle$, а подходящие дроби даются отношениями чисел Фибоначчи: F_k/F_{k+1} , $F_0 = 0$, $F_1 = 1$, $F_{k+1} = F_k + F_{k-1}$.

Если вместо иррационального частотного параметра взять k -ю рациональную аппроксимацию: $x_{n+1} = f(x_n, \theta_n)$, $\theta_{n+1} = \theta_n + w_k \pmod{1}$, $w_k = p_k/q_k$, то внешнее воздействие будет периодическим, с периодом q_k . В отличие от квазипериодического поведения, когда фазовая переменная эргодическим образом посещает плотное множество точек единичного интервала, теперь она обходит конечное множество $\{\theta_0, \theta_1, \dots, \theta_{q_k-1}\}$. При этом в зависимости от θ_0 будем получать, вообще говоря, разные режимы динамики и разные аттракторы. Итак, начальная

фаза θ_0 играет роль дополнительного параметра. Намереваясь на основе используемой рациональной аппроксимации составить представление, как будет вести себя система с иррациональной частотой, мы должны, очевидно, принимать во внимание все ситуации, которые получаются при произвольном задании θ_0 в интервале от 0 до $1/q_k$.

На рис. 7 показаны карты динамических режимов на плоскости параметров квадратичного отображения с внешним воздействием при рациональной аппроксимации ($a - \epsilon$) и при иррациональном значении частоты $w = (\sqrt{5}-1)/2$ (ϵ). Сравнивая карты, можно видеть, что области регулярной квазипериодической динамики на диаграмме (ϵ) (торы T1, T2, T4) ассоциируются с областями периодических режимов (циклов) на картах ($a - \epsilon$). В областях, где в системе с иррациональной частотой реализуется СНА, на картах для рациональных аппроксимаций можно видеть «мельтешение» режимов разного типа в зависимости от порядка аппроксимации и от начальной фазы.

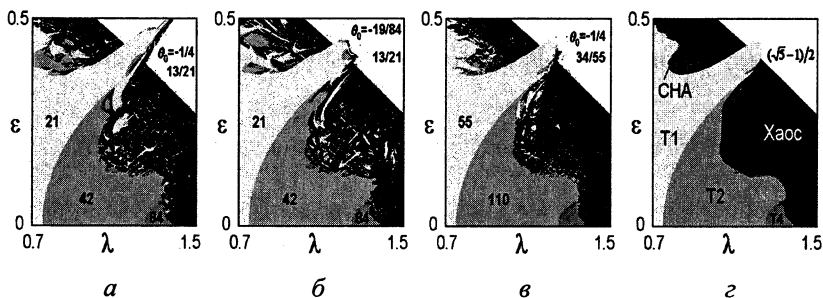


Рис. 7. Карты динамических режимов для модели (4) на плоскости параметров для рациональных аппроксимаций частоты воздействия $13/21$ с двумя разными начальными фазами (a) и (b), $34/55$ (v) и для иррациональной частоты воздействия $w = (\sqrt{5}-1)/2$ (ϵ). Цифрами обозначены периоды наблюдаемых периодических режимов в соответствующих областях.

В работе Пиковского и Фойдель [18] была высказана гипотеза, что необходимое условие существования СНА состоит в том, что при рациональной аппроксимации частотного параметра система демонстрирует бифуркации в зависимости от параметра начальной фазы, причем это свойство сохраняется при увеличении порядка аппроксимации. На качественном уровне, рассуждая в терминах систем, полученных на основе рациональной аппроксимации, можно представить себе, что при иррациональном значении частоты имеем как бы медленный дрейф параметра начальной фазы, и система по ходу динамики все время претерпевает бифуркации. В этом и состоит подоплека возникновения СНА.

4. Как рождается странный нехаотический аттрактор?

4.1. Бифуркации в системах с квазипериодическим воздействием и сценарии рождения СНА

Если в системе с квазипериодическим воздействием при одном значении управляющего параметра реализуется гладкий тор, а при другом — СНА, то какие переходы (бифуркации) могут наблюдаться в процессе движения по параметру от одного режима к другому? Один из сценариев, получивший название бифуркации расширения («blowout») [27], реализуется в модели с бифуркацией вилки (1) и уже обсуждался в общих чертах в п. 2.1. Интересный подход к изучению деталей возникновения СНА для нескольких ситуаций был развит на основе метода ренормгруппы [15, 28—30].

Говоря о бифуркационных сценариях в квадратичном отображении с квазипериодическим воздействием (4), естественно начать с *бифуркаций удвоения торов* [31—33]. Известно, что при $\varepsilon = 0$ имеем последовательность бифуркаций удвоения периода с переходом к хаосу по Фейгенбауму. Первая бифуркация состоит в том, что неподвижная точка теряет устойчивость и рождается аттрактор из двух точек, посещаемых попеременно, цикл периода 2. В присутствии квазипериодического воздействия вместо устойчивой неподвижной точки имеем притягивающую замкнутую инвариантную кривую (тор-аттрактор) $T1$, а бифуркация состоит в том, что от нее отделяются две замкнутые кривые, посещаемые попеременно (рис. 8). Это новый аттрактор — удвоенный тор $T2$. При малых ε , увеличивая λ , что соответствует движению на плоскости параметров по линии TD (рис. 9, а), можно наблюдать несколько бифуркаций удвоения торов ($T1 \rightarrow T2 \rightarrow T4 \rightarrow \dots$). В отличие от фейгенбаумовского каскада эта последовательность содержит конечное число бифуркаций (тем большее, чем меньше ε) [11, 31—33].

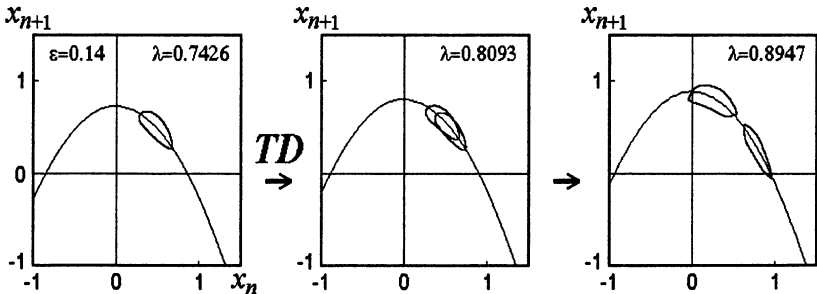


Рис. 8. Бифуркация удвоения тора (Torus Doubling, TD), представленная на итерационной диаграмме; парабола изображает график функции $\lambda - x^2$.

Хеги и Хэмел [34] рассмотрели механизм, ответственный за ограничение числа удвоения торов, и связали с ним сценарий рождения СНА. Он отвечает движению по пути **НН** на рис. 9, *a* и иллюстрируется фазовыми портретами на рис. 9, *б*. После рождения удвоенного тора, представленного на фазовом портрете двумя изогнутыми линиями, продолжает существовать также потерявший устойчивость родительский тор, обозначенный серой линией. С увеличением λ компоненты удвоенного тора приобретают все более извилистый вид, разбухая по ширине, и, наконец, в момент бифуркации касаются своими выступами располагающейся между ними родительской инвариантной кривой. После бифуркации притягивающее множество уже не имеет двух попеременно посещаемых компонент, а представляет собой единый СНА.

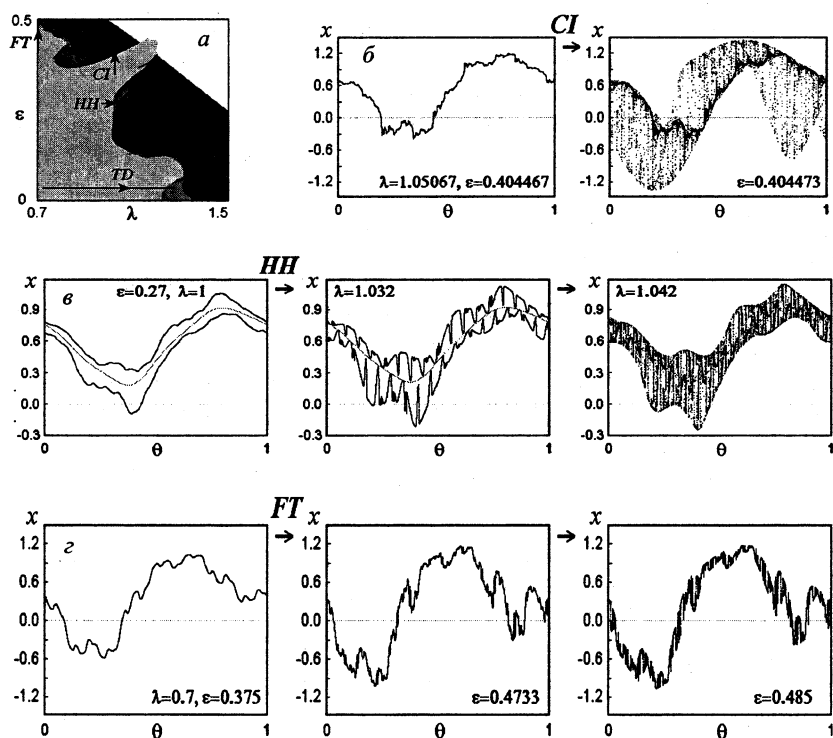


Рис. 9. Карта режимов квадратичного отображения с квазипериодическим воздействием (*a*), перемежаемость, связанная с кризисом **CI** (*б*) и фазовые портреты, иллюстрирующие сценарии рождения СНА при движении по плоскости параметров, как показано стрелками: сценарий Хеги — Хэмела **НН** (*в*), фрактализация тора по Нишикава — Канеко **FT** (*г*).

Другой сценарий описан Нишикавой и Канеко [35] и состоит в том, что притягивающая инвариантная кривая постепенно делается все более извилистой, теряет гладкость и становится фрактальной — возникает СНА. Это имеет место при движении на плоскости параметров по пути TF (рис. 9, *a*) и иллюстрируется фазовыми портретами на рис. 9, *в*.

Переход к СНА на пути CI состоит в том, что область, занятая аттрактором, резко расширяется, хотя система по-прежнему проводит большую часть времени в окрестности ранее существовавшей притягивающей инвариантной кривой, т. е. имеет место перемежаемость. В работе [36] этот переход гипотетически ассоциирован с бифуркацией торов, являющейся аналогом столкновения устойчивой и неустойчивой точек при касательной (седло-узловой) бифуркации. На самом деле происходит столкновение тора-аттрактора не с партнером в виде неустойчивого тора, а со специфическим неустойчивым инвариантным множеством, которое идентифицировано в работе [37] с использованием рациональных аппроксимаций («ring-shape invariant set»). По-видимому, феномен следует интерпретировать как внутренний кризис определенного типа, приводящий к рождению СНА [38].

Настоящим аналогом касательной бифуркации и перемежаемости по Помо и Манневиллю служит переход в отображении окружности [3, 17], а также в модели (5), модифицированной так, чтобы обеспечить присутствие механизма возврата [15]:

$$x_{n+1} = f(x_n) + b + \epsilon \cos 2\pi(nw + \theta_0), \quad f(x) = \begin{cases} x/(1-x), & x \leq 0.75 \\ 9/2x - 3, & x > 0.75 \end{cases}. \quad (8)$$

Оказывается, что существует критическое значение амплитуды воздействия ϵ_c такое, что при $\epsilon < \epsilon_c$ и $\epsilon > \epsilon_c$ переход с увеличением параметра b носит существенно разный характер¹.

В первом случае имеет место сближение двух гладких инвариантных кривых (устойчивой и неустойчивой) с последующим их слиянием и исчезновением, а результатом является рождение хаоса (рис. 10, *a*). Во втором случае приближение к точке бифуркации сопровождается образованием выступов («шипов») на инвариантных кривых, и в момент бифуркации они касаются друг друга остриями этих шипов (рис. 10, *б*). В силу эргодичности динамики фазовой переменной точки касания образуют плотное множество, так что возникающий в момент бифуркации объект фрактальный. О данной бифуркации говорят как о *фрактальном столкновении торов*, ее результатом является рождение СНА.

¹ В модели (8) ϵ_c точно равно 2, что находит объяснение при формулировке задачи в терминах уравнения Харпера (6), для которого критическое значение $\epsilon_c = 2$ отвечает переходу локализация — делокализация.

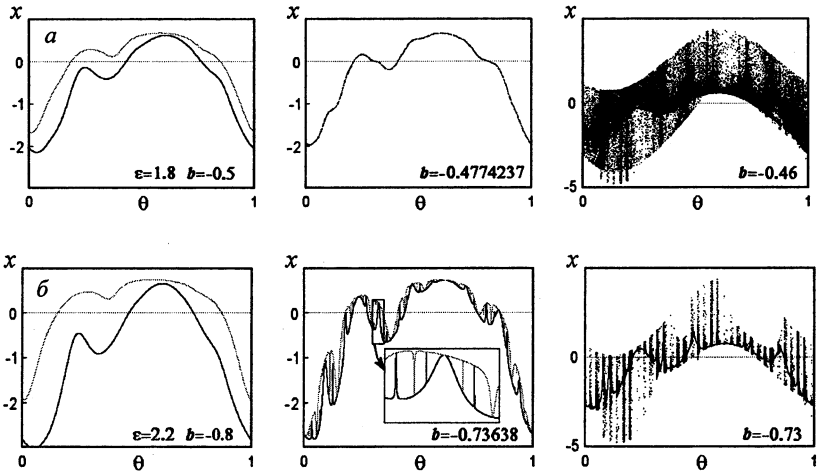


Рис. 10. Переход через перемежаемость в модели (8) через бифуркацию слияния гладких торов (а) и фрактального столкновения торов (б). Устойчивая инвариантная кривая показана черным цветом, неустойчивая — серым.

4.2. Ренормгрупповой подход к анализу рождения СНА

Метод ренормгруппы (РГ) изначально возник в квантовой теории поля и в теории фазовых переходов. В нелинейную динамику он был привнесен благодаря работам Фейгенбаума [6, 7, 9] и оказался полезным и мощным теоретическим инструментом для анализа переходов к хаосу.

Пусть имеем оператор эволюции некоторой системы на каком-то временном интервале. Применяя его несколько раз, можно построить оператор на большем интервале. Предметом анализа служит ситуация (*критическая точка*), когда посредством масштабной замены переменных новый оператор приводится точно или почти точно к старому. Эту процедуру, называемую *РГ-преобразованием*, можно повторять многократно и построить последовательность операторов для больших временных интервалов. В критической точке они определяются структурой РГ-преобразования. Поэтому поведение на больших временах для широкого класса систем оказывается одинаковым и не зависящим от деталей локальной динамики (*универсальность*). Поскольку получаемые операторы совпадают с точностью до масштабной замены, динамика на разных характерных временах оказывается подобной (*скейлинг*).

Более конкретно, для системы с дискретным временем при квазипериодическом воздействии на частоте $w = (\sqrt{5}-1)/2$ рассмотрим операторы эволюции за интервалы времени, отвечающие числам Фибоначчи

F_m . Пусть мгновенное состояние задается двумя переменными (x, u) , а эволюция за F_m шагов описывается уравнениями

$$x_{n+F_m} = f_m(x_n, u_n), \quad u_{n+F_m} = u_n + F_m w \pmod{1}, \quad (9)$$

причем по второму аргументу функция f имеет период 1. В соответствии с определением чисел Фибоначчи $F_{m+2} = F_{m+1} + F_m$, так что

$$x_{n+F_{m+2}} = f_m(f_{m+1}(x_n, u_n), u_n + wF_{m+1}). \quad (10)$$

В силу свойств золотого среднего $u + wF_{m+1} = u - (-w)^{m+1} \pmod{1}$. Для оператора эволюции за F_m шагов введем нормировку u на фактор $(-w)^m$ и x на a^{-m} , где a — пока неизвестная константа. Тогда оператор эволюции будет представлен функцией $g_m(x, u) = a^m f_m(xa^{-m}, u(-w)^m)$, а уравнение (10) примет вид [29, 30]

$$g_{m+2} = a^2 g_m(a^{-1} g_{m+1}(xa^{-1}, -wu), w^2 u + w). \quad (11)$$

В пространстве функций это уравнение может иметь неподвижные точки или циклы, отвечающие за те или иные типы критического поведения в системе с квазипериодическим воздействием. Рассмотрение малого возмущения неподвижной точки или цикла приводит к задаче на собственные функции. Существенные собственные числа δ , превышающие по модулю 1, ассоциируются со сдвигами по параметрам, выводящими систему из критической точки, и играют роль факторов скейлинга в пространстве параметров.

Первый тип критического поведения, проанализированный в рамках этого подхода, относится к рождению СНА в отображении с бифуркацией вилки (1) [28]. Ему отвечает решение уравнения РГ вида

$g_m(x, u) = Q_m(u)x / \sqrt{1 + H_m(u)x^2}$ в виде цикла периода 6, хотя с точностью до знака эти функции повторяются с периодом 3. Найденный численно фактор перенормировки $\alpha = a^3 = 1,323891$, а собственное число $\delta = w^{-3} = 4,236068$. Отсюда вытекает заключение о подобии режимов динамики вблизи критической точки $\lambda = 1$ при пересчете параметра $\lambda \rightarrow (\lambda - 1)/\delta$ и динамических переменных $x \rightarrow x/\alpha$, $u \rightarrow u/\beta$, где $\beta = -w^{-3}$ (рис. 11, а, б). Можно также получить нетривиальный степенной закон, которому подчиняется размер аттрактора вблизи критической точки: $|\Delta x| \propto \Delta \lambda^\nu$, где $\nu = \log_\delta \alpha$ (рис. 11, в).

Обратимся теперь к квадратичному отображению с квазипериодическим воздействием (4). На плоскости (λ, ϵ) имеется линия бифуркации удвоения тора $T1 \rightarrow T2$. Если двигаться вдоль бифуркационной линии, увеличивая амплитуду воздействия, то замкнутая инвариантная кривая $T1$, расположенная первоначально в области $x > 0$, увеличивается в размерах и в какой-то момент касается точки $x = 0$, отвечающей экстремуму

параболы. Это соответствует окончанию бифуркационной кривой удвоения тора — критической точке TDT (Torus Doubling Terminal) при $\lambda_c = 1,158096856726$ и $\epsilon_c = 0,360248020507$ [29]. Оказывается, что за поведение в точке TDT отвечает цикл периода 3 в уравнении РГ. Из решения функционального уравнения численно найдены соответствующие функции, а также фактор перенормировки $\alpha = a^3 = 3,96376647$ и собственные числа $\delta_1 = 10,5029$ и $\delta_2 = 5,1881$.

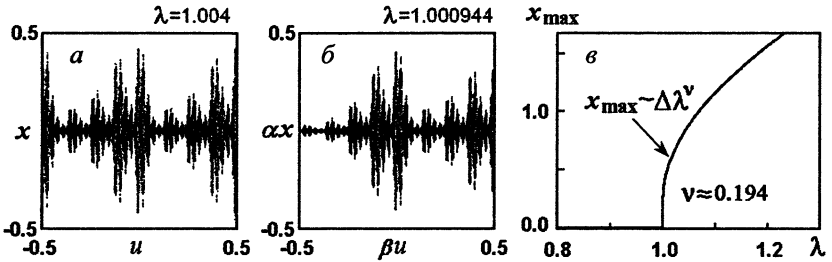


Рис. 11. Демонстрация скейлинга вблизи точки рождения СЧА на фазовых портретах отображения (1) (а) и (б). Во втором случае отклонение по параметру от критической точки уменьшено в δ раз, а масштаб по осям координат изменен на факторы $\beta = -w^{-3}$ и $\alpha \approx 1,32391$. Иллюстрация закона поведения размера аттрактора в зависимости от надкритичности (в).

Со стороны малых λ область T1 на плоскости параметров (λ, ϵ) ограничена линией бифуркации слияния устойчивого и неустойчивого торов. Будем двигаться вдоль нее в направлении увеличения ϵ . При этом замкнутая инвариантная кривая, располагавшаяся в области $x < 0$, в какой-то момент достигает краем точки $x = 0$. Это имеет место в критической точке TCT (Torus Collision Terminal): $\lambda_c = 0,09977122895$, $\epsilon_c = 1,01105609099$. За критическое поведение TCT отвечает неподвижная точка уравнения РГ [30]. Численное решение уравнения РГ позволило найти соответствующие функции, фактор перенормировки $\alpha = a = 1,7109605$ и собственные числа $\delta_1 = 3,600810$, $\delta_2 = 1,828329$.

На рис. 12 иллюстрируются свойства скейлинга критических аттракторов в точках TDT и TCT. Портрет аттрактора на плоскости (x, u) локально подобен себе и переходит в себя при пересчете масштаба по вертикали и горизонтали на соответствующие константы. Особенность, которая визуалью похожа на излом, имеет вид $\Delta x \propto |\Delta u|^\gamma$, где показатель $\gamma = \log_{\beta} \alpha$ для TDT равен 0,954, а для TCT 1,117. Благодаря эргодичности она «тиражируется» на плотном множестве точек, и аттрактор фактически представляет собой фрактальную кривую.

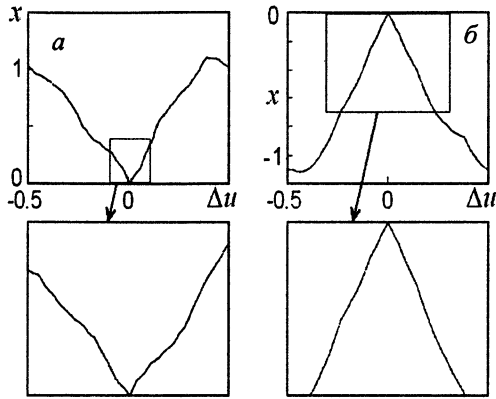


Рис. 12. Фазовые портреты аттрактора квадратичного отображения в точках TDT (а) и TCT (б) и иллюстрации скейлинга в окрестности особенности.

Еще один тип критического поведения — точка TF («Torus Fractalization») отвечает моменту изменения характера перехода через перемежаемость, который при малой амплитуде воздействия связан с бифуркацией слияния гладких торов, а при больших — с фрактальным столкновением [15]. Для отображения (5) точка TF располагается при $\epsilon_c = 2$, $b_c = -0,597515185376121$. За поведение этого типа отвечает неподвижная точка уравнения РГ в виде дробно-линейной функции $g(x, u) = (a(u)x + b(u)) / (c(u)x + d(u))$. Согласно численным результатам [15] масштабный фактор равен $\alpha = a = 2,890053525$, а существенные собственные числа $\delta_1 = 3,134272989$, $\delta_2 = w^{-1} = 1,6180340$. На рис. 13 иллюстрируется скейлинг аттрактора в точке TF: в специальной локальной системе координат картинка переходит в себя при пересчете масштаба по двум осям координат на факторы α и β .

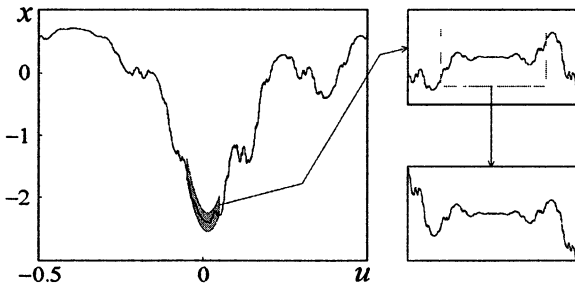


Рис. 13. Фазовый портрет аттрактора отображения (5) в критической точке TF и иллюстрация свойства скейлинга при пересчете масштаба по вертикали и горизонтали, соответственно, на факторы $\alpha \approx 2,89005$ и $\beta = -w^{-1}$ в системе координат $c_1 = (x + 2,34719526) + 5,9267u - 210,63u^2$, $c_2 = u$.

5. Системы с непрерывным временем и эксперименты по наблюдению феноменов, связанных с присутствием СНА

До сих пор мы обсуждали модели с дискретным временем, представленные итерируемыми отображениями. Рассмотрим вопрос о возможности реализации СНА в системах, заданных привычными физикам дифференциальными уравнениями. Ключевой момент – установление связи между дифференциальными уравнениями и описанием динамики в терминах отображений на основе идеи сечения Пуанкаре для интересующего нас класса систем с квазипериодическим воздействием.

Предположим для определенности, что система без внешнего воздействия имела двумерное фазовое пространство (x, y) и описывалась двумя дифференциальными уравнениями первого порядка. В случае двухчастотного воздействия функции в правых частях уравнений надо считать зависящими периодическим образом от фазовых переменных Θ и Φ , изменяющихся во времени по линейному закону:

$$\dot{x} = f_1(x, y, \Theta, \Phi), \quad \dot{y} = f_2(x, y, \Theta, \Phi), \quad \dot{\Theta} = \omega, \quad \dot{\Phi} = \Omega, \quad (12)$$

причем частоты ω и Ω полагаем находящимися в иррациональном отношении. Пусть в момент, отвечающий определенной фазе второй составляющей Φ , имеем значения переменных (x, y, Θ) ; какими они будут через один период этой составляющей $T = 2\pi/\Omega$? Из теоремы единственности решения дифференциального уравнения следует, что эти величины должны выражаться однозначно, как функции от исходных значений. Еще через период они получатся повторным применением тех же функций и так далее. Итак, возникает итерируемое отображение

$$\begin{aligned} x_{n+1} &= F_1(x_n, y_n, \Theta_n), & y_{n+1} &= F_2(x_n, y_n, \Theta_n), \\ \Theta_{n+1} &= \Theta_n + 2\pi\omega/\Omega \pmod{2\pi}. \end{aligned}$$

В данном контексте это и есть *отображение Пуанкаре* или *стробоскопическое отображение*. Теперь можно отвлечься от исходных уравнений и сосредоточиться на анализе динамики, порождаемой отображением. Чаще всего функции F_1 и F_2 можно получить только на основе численного решений уравнений (12). Однако, с принципиальной точки зрения, мы имеем задачу того же типа, что и обсуждавшиеся модели с дискретным временем. Добавим, что в системах с сильной диссипацией наблюдаемые по ходу временной эволюции состояния на секущей гиперплоскости имеют тенденцию располагаться при каждом фиксированном Θ вдоль одномерной кривой на плоскости (x, y) . В этом случае, введя вместо (x, y) одну координату вдоль этой кривой, приходим к описанию динамики в терминах одномерного отображения с квазипериодическим воздействием.

Представительный класс нелинейных систем, демонстрирующих СНА, составляют нелинейные диссипативные осцилляторы, возбуждаемые квазипериодической силой [21—23, 39—41].

В качестве примера на рис. 14 приводятся численные результаты, относящиеся к нелинейному осциллятору Дуффинга с квазипериодическим параметрическим возбуждением [23]:

$$\dot{x} = y, \quad \dot{y} = -\gamma y - x^3 + x(1 + A(R \cos \Theta + \cos \Phi)), \quad \dot{\Theta} = \omega, \quad \dot{\Phi} = \Omega, \quad (13)$$

где $\omega = 2\pi$, $\Omega = \pi(\sqrt{5} + 1)$. Показаны фазовые портреты аттракторов в проекции на плоскость $(x, y = \dot{x})$, фазовые портреты на плоскости $(\theta = \Theta/2\pi, x)$ в сечении Пуанкаре $\Phi = 0$ и спектры Фурье для режимов, соответствующих тору-аттрактору (а), СНА (б) и хаосу (в). Более подробные численные расчеты показывают, что в данной системе реализуются многие из феноменов, упоминавшихся в предыдущих разделах, в том числе последовательности удвоения торов, переходы к СНА по Хегги — Хэмелу, Нишикава — Канеко, перемежаемость.

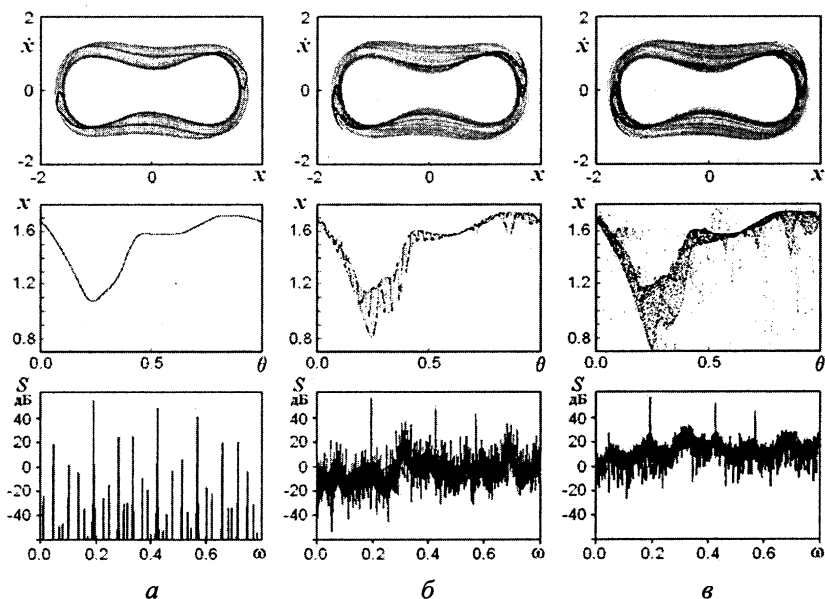


Рис. 14. Тор-аттрактор при $A = 0,38$ (а), СНА при $A = 0,396$ (б) и хаос при $A = 0,402$ (в) в осцилляторе Дуффинга с параметрическим возбуждением [23] — фазовый портрет в проекции на плоскость (x, \dot{x}) , где черные точки отвечают стробоскопическому сечению $\theta = 0$, фазовые портреты на плоскости (x, θ) и фурье-спектры. Остальные параметры: $\gamma = 0,1$, $R = 0,3$.

В ряде работ сообщалось об обнаружении СНА при численном решении задачи о квазипериодическом воздействии на систему автоколебательного типа – осциллятор Ван-дер-Поля [42, 43].

Первый описанный в литературе эксперимент по наблюдению СНА в реальной системе относится к механическому осциллятору в виде закрепленной на конце магнитно-упругой ленты, колебания которой возбуждаются переменным магнитным полем [21, 23]. Система описывается неавтономным уравнением Дуффинга (13). Режим СНА идентифицировался по виду фазовых портретов в стробоскопическом сечении, скейлинговому поведению составляющих в спектре Фурье, характерному для сингулярно-непрерывного спектра, и по размерностным характеристикам: оценки фрактальной и информационной размерности аттрактора в сечении Пуанкаре дали результаты, близкие, соответственно, к 2 и 1.

В работах [22, 44] сообщалось о наблюдении СНА в экспериментах с нелинейными осцилляторами в виде электронных схем, которые представляли собой аналоговые устройства, реализующие динамику, описываемую определенными дифференциальными уравнениями. Помимо фазовых портретов, спектров и оценок размерностей представлены результаты, относящиеся к показателям Ляпунова для дифференциального уравнения, аналоговой моделью которого служит исследуемая схема. В одной из этих работ реализован и изучен в эксперименте переход к СНА через бифуркацию «blowout» [44].

В работе [45] исследован экспериментально электронный осциллятор в виде колебательного контура, в котором роль нелинейного элемента выполнял полупроводниковый диод. Известно, что при одночастотном возбуждении в этой системе реализуется переход к хаосу через последовательность удвоений периода, как в квадратичном отображении. Добавление второй составляющей воздействия на частоте, находящейся с первой в иррациональном отношении (золотое среднее), приводит к появлению феноменов, имеющих место в логистическом отображении с квазипериодическим воздействием. В этом можно убедиться, сравнивая карты динамических режимов в эксперименте (по осям координат — амплитуды составляющих воздействующего переменного напряжения в вольтах) и в численных расчетах для квадратичного отображения (рис. 15). На рисунке показаны также полученные в эксперименте фазовые портреты типичных аттракторов (тор, удвоенный тор, СНА, хаос). Для их построения снимаемые с последовательно включенного в контур резистора значения потенциала вводились в компьютер с помощью аналого-цифрового преобразователя и на графике по осям координат откладывались величины, относящиеся к определенному моменту в пределах n -го и $n+1$ -го периодов основного воздействия. Для идентификации СНА использовалась идея Пиковского и Фойдель: соотношение частот зада-

валось как рациональная аппроксимация золотого среднего ($13/21$) и при фиксированных остальных параметрах отслеживалось присутствие или отсутствие бифуркаций в системе в зависимости от соотношения фаз частотных составляющих.

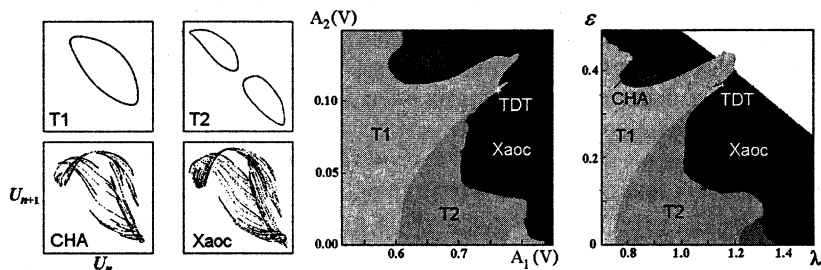


Рис. 15. Фазовые портреты аттракторов, полученные в эксперименте с нелинейным колебательным контуром при двухчастотном возбуждении с отношением частот, даваемым золотым средним, и области режимов на плоскости параметров, где по осям координат отложены амплитуды составляющих в вольтах. Справа для сравнения аналогичная карта для отображения (4).

В литературе сообщалось также о наблюдении СНА в электронных схемах, содержащих элемент с кусочно-линейной характеристикой с квазипериодическим возбуждением [46, 47], в химическом эксперименте с реакцией Белоусова — Жаботинского при воздействии на реактор двух других реакторов, где реализовались осцилляции на несоизмеримых частотах [48]. Отдельно следует упомянуть работу [49], в которой характерные для СНА особенности динамики обнаружены в распределенной системе — тлеющем разряде в неоновой трубке. В качестве зависящей от времени наблюдаемой величины регистрировался испускаемый разрядом световой поток. В этом случае внешнего квазипериодического воздействия не было; по-видимому, его роль играли какие-то внутренние квазинезависимые периодические процессы в разряде.

Заключение

Как любопытный экземпляр в «зоопарке» феноменов нелинейной динамики, СНА, несомненно, заслуживает, чтобы о нем знали специалисты, имеющие дело с колебательными системами, функционирующими в условиях внешнего воздействия со сложным спектром (механические конструкции, радиотехнические и электронные устройства и т. п.). С прикладной точки зрения вызывает интерес возможность применения СНА в схемах скрытой коммуникации [50, 51], поскольку осуществить синхронизацию систем с СНА в приемнике и передатчике проще, чем для систем с хаосом.

Здесь мы хотим подчеркнуть более фундаментальный аспект. Как известно, на раннем этапе обсуждения проблемы возникновения турбулентности был предложен сценарий Ландау, согласно которому с увеличением числа Рейнольдса в гидродинамических течениях в результате ряда бифуркаций последовательно возникают колебательные движения с несоизмеримыми частотами, а при наличии достаточно большого количества таких составляющих динамика воспринимается как турбулентная. Эта картина была подвергнута критике Рюэлем и Такенсом, которые выдвинули концепцию странного хаотического аттрактора, возникающего после рождения трех (в другой версии теоремы четырех) независимых частотных составляющих. Аргументация, впрочем, носит общий характер и с равным основанием может быть отнесена к системам самой разной физической природы. Со времени этих пионерских работ достигнут большой прогресс в исследовании перехода к хаосу через квазипериодичность, но в целом картина весьма сложна и все еще далека от полного понимания. Один из возможных путей продвижения — обратиться к упрощенной задаче, предположив, что реализующиеся в какой-то сложной системе с многомерным фазовым пространством колебания с несоизмеримыми частотами квазинеzависимы и действуют однонаправленным образом на одну дополнительную, вновь возникшую составляющую движения (моду). Это приводит к моделям с квазипериодическим воздействием, обсуждавшимся выше. В них значения частот выступают как независимые параметры, задаваемые исследователем, что создает возможность для более свободного обращения с моделью и анализа возникающих режимов динамики. Замечательно, что уже на этом пути история делает неожиданный поворот: наряду с регулярной (периодической и квазипериодической) динамикой и хаосом обнаруживается возможность нового типа поведения — СНА.

В системах с квазипериодическим воздействием СНА реализуется на множестве положительной меры в пространстве параметров и поэтому должен рассматриваться как типичный феномен, характерный для области между порядком и хаосом.

В автономных системах частоты колебательных составляющих порождаются динамикой самой системы и не допускают простого и непосредственного регулирования. В этой ситуации возможность реализации СНА как типичного феномена проблематична. Вообще говоря, в каких-то точках пространства параметров автономной системы, совершающей движение с несколькими независимыми колебательными составляющими, по-видимому, будут реализоваться СНА, но насколько заметную часть пространства параметров может занимать множество этих точек и имеет ли оно отличную от нуля меру, сказать трудно. Отдельные появившиеся в литературе сообщения относительно наблюдения СНА в автономных системах не кажутся убедительными.

Подводя итог, можно констатировать, что исследования, связанные с феноменом ШНА, обогатили представления о явлениях в нелинейных системах в области между порядком и хаосом, послужили стимулом для привлечения и использования в нелинейной динамике новых концепций. Были вскрыты интересные аналогии с фазовыми переходами, задачами теории колебаний и волн в неупорядоченных средах, теорией квантового хаоса, что способствует формированию междисциплинарных связей и взаимному обогащению теоретического инструментария.

Работа выполнена при поддержке РФФИ (грант 03-02-16074).

Литература

1. *Grebogi C., Ott E., Pelikan S., Yorke J.A.* Strange attractors that are not chaotic // *Physica D*. 1984. V. 13. № 1, 2. P. 261—268.
2. *Ding M., Grebogi C., Ott E.* Evolution of attractors in quasiperiodically forced systems: From quasiperiodic to strange nonchaotic to chaotic // *Phys. Rev. A*. 1989. V. 39. P. 2593—2598.
3. *Feudel U., Pikovsky A. S., Kurths J.* Strange nonchaotic attractor in a quasiperiodically forced circle map // *Physica D*. 1995. V. 88. P. 176—186.
4. *Anishchenko V. S., Vadivasova T. E., Sosnovtseva O.* Mechanisms of ergodic torus destruction and appearance of strange nonchaotic attractors // *Phys. Rev. E*. 1996. V. 53. P. 4451—4456.
5. *Prasad A., Negi S. S., Ramaswamy R.* Strange nonchaotic attractors // *ИВС*. 2001. V. 11. P. 291—311.
6. *Фейгенбаум М.* Универсальность в поведении нелинейных систем // *УФН*. 1983. Т. 141, № 2. С. 343—374.
7. *Лихтенберг А., Либерман М.* Регулярная и стохастическая динамика. М.: Мир, 1984.
8. *Рабинович М. И., Трубецков Д. И.* Введение в теорию колебаний и волн. М.: Наука, 1984.
9. *Кузнецов С. П.* Динамический хаос. М.: Физматлит, 2001.
10. *Kaneko K.* Oscillation and Doubling of Torus // *Progr. Theor. Phys*. 1984. V. 72, № 2. P. 202—215.
11. *Кузнецов С. П.* О воздействии периодического возмущения на систему, демонстрирующую переход порядок-хаос через бифуркации удвоения периода // *Письма ЖЭТФ*. 1984. Т. 39, № 3. С. 113—116.
12. *Kuznetsov S. P., Pikovsky A. S.* Renormalization group for the response function and spectrum of the period-doubling system // *Phys. Lett. A*. 1989. V. 140, № 4. С. 166—172.
13. *Берже П., Помо И., Видаль К.* Порядок в хаосе. О детерминистическом подходе к турбулентности. М.: Мир, 1991.
14. *Ketoja J. A., Satiya I. I.* Harper equation, the dissipative standard map and strange nonchaotic attractors // *Physica D*. 1997. V. 109. P. 70—80.
15. *Kuznetsov S. P.* Torus fractalization and intermittency // *Phys. Rev. E*. 2002. V. 65. 066209.
16. *Суслов И. М.* Локализация в одномерных несоизмеримых системах // *ЖЭТФ*. 1982. Т. 83. № 3. С. 1079—1088.
17. *Osinga H., Wiersig J., Glendinning P., Feudel U.* Multistability and nonsmooth bifurcations in the quasiperiodically forced circle map // *ИВС*. 2001. V. 11, № 12. P. 3085—3107.
18. *Pikovsky A. S., Feudel U.* Characterizing strange nonchaotic attractors // *Chaos*. 1995. V. 5, № 1. P. 253—260.

19. Prasad A., Ramaswamy R. Characteristic distributions of finite-time Lyapunov exponents // *Phys. Rev. E*. 1999. V. 60. P. 2761—2766.
20. Ding M., Grebogi C., Ott E. Dimensions of Strange Nonchaotic Attractors // *Phys. Lett. A*. 1989. V. 137, № 4, 5. P. 167—172.
21. Ditto W. L., Spano M. L., Savage H. T., Raueo S. N., Heagy J., Ott E. Experimental observation of a strange nonchaotic attractor // *Phys. Rev. Lett.* 1990. V. 65. P. 533—536.
22. Zhou T., Moss F., Bulsara A. Observation of a strange nonchaotic attractor in a multistable potential // *Phys. Rev. A*. 1992. V. 45. P. 5394—5400.
23. Heagy J., Ditto W. L. Dynamics of a two-frequency parametrically driven Duffing oscillator // *J. Nonlinear Sci.* 1991. V. 1. P. 423—455.
24. Datta S., Sharma A., Ramaswamy R. Thermodynamics of critical strange nonchaotic attractors // *Phys. Rev. E*. 2003. V. 68. 036104.
25. Pikovsky A. S., Feudel U. Correlations and spectra of strange nonchaotic attractors // *J. Phys. A: Math. Gen.* 1994. V. 27. P. 5209—5219.
26. Pikovsky A. S., Zaks M. A., Feudel U., Kurths J. Singular continuous spectra in dissipative dynamics // *Phys. Rev. E*. 1995. V. 52. P. 285—296.
27. Yalcynkaya T., Lai Y.-C. Blowout bifurcation route to strange nonchaotic attractors // *Phys. Rev. Lett.* 1996. V. 77. P. 5039—5042.
28. Kuznetsov S. P., Pikovsky A. S., Feudel U. Birth of a strange nonchaotic attractor: A renormalization group analysis // *Phys. Rev. E*. 1995. V. 51. P. 1629—1632.
29. Kuznetsov S., Feudel U., Pikovsky A. Renormalization group for scaling at the torus-doubling terminal point // *Phys. Rev. E*. 1998. V. 57. P. 1585—1590.
30. Kuznetsov S. P., Neumann E., Pikovsky A., Sataev I. R. Critical point of tori collision in quasiperiodically forced systems // *Phys. Rev. E*. 2000. V. 62. P. 1995—2007.
31. Анищенко В. С. Индуцированные внешним воздействием фазовые переходы в радиофизической системе со странным аттрактором // Флуктуационные явления в физических системах: Тез. докл. 3-й Всесоюз. конф. Вильнюс: Изд-во АН ЛитССР, 1982, 17 с.
32. Kaneko K. Oscillation and Doubling of Torus // *Progr. Theor. Phys.* 1984. V. 72, № 2. P. 202—215.
33. Arneodo A., Collet P. H., Spiegel E. A. Cascade of period doublings of tori // *Phys. Lett. A*. 1983. V. 94, № 1. P. 1—4.
34. Heagy J. F., Hammel S. M. The birth of strange nonchaotic attractors // *Physica D*. 1994. V. 70, № 1, 2. P. 140—153.
35. Nishikawa T., Kaneko K. Fractalization of a torus as a strange nonchaotic attractor // *Phys. Rev. E*. 1996. V. 54. P. 6114—6124.
36. Prasad A., Mehra V., Ramaswamy R. Intermittency route to strange nonchaotic attractors // *Phys. Rev. Lett.* 1997. V. 79. P. 4127—4130.
37. Kim S.-Y., Lim W., Ott E. Mechanism for the intermittent route to strange nonchaotic attractors // *Phys. Rev. E*. 2003. V. 67. 056203.
38. Witt A., Feudel U., Pikovsky A. S. Birth of strange nonchaotic attractors due to interior crisis // *Physica D*. 1997. V. 109. P. 180—190.
39. Bondeson A., Ott E., Antonsen T. M. Quasiperiodically forced damped pendula and Schrodinger equations with quasiperiodic potentials: Implications of their equivalence // *Phys. Rev. Lett.* 1985. V. 55. P. 2103—2106.
40. Venkatesan A., Lakshmanan M., Prasad A., Ramaswamy R. Intermittency transitions to strange nonchaotic attractors in a quasiperiodically driven Duffing oscillator // *Phys. Rev. E*. 2000. V. 61. P. 3641—3651.
41. Kuznetsov S. P., Neumann E. Torus fractalization and singularities in the current-voltage characteristics for the quasiperiodically forced Josephson junction // *Europhys. Lett.* 2003. V. 61, № 1. P. 20—26.
42. Kapitaniak T., Ponce E., Wojewoda J. Route to chaos via strange non-chaotic attractors // *J. Phys. A: Math. Gen.* 1990. V. 23. P. 383—387.

43. Pokorny P., Schreiber I., Marek M. On the route to strangeness without chaos in the quasiperiodically forced van der Pol oscillator // *Chaos, Solitons and Fractals*. 1996. V. 7, № 3. P. 409—424.
44. Yu Y. H., Kim D. C., Ryu J. Y., Hong S. R. Experimental Study on the Blowout Bifurcation Route to Strange Nonchaotic Attractor // *J. Korean Phys. Soc.* 1999. V. 34, № 2. P. 130—134.
45. Bezruchko B. P., Kuznetsov S. P., Seleznev Y. P. Experimental observation of dynamics near the torus-doubling terminal critical point // *Phys. Rev. E*. 2000. V. 62. P. 7287—7830.
46. Yang T., Bilingut K. Experimental results of strange nonchaotic phenomenon in a second-order quasi-periodically forced electronic circuit // *Phys. Lett. A*. 1997. V. 236. P. 494—504.
47. Zhu Z., Liu Z. Strange nonchaotic attractors of Chua's circuit with quasiperiodic excitation // *IJBC*. 1997. V. 7. P. 227—238.
48. Zeyer K.-P., Münster A. F., Schneider F. W. Quasiperiodic Forcing of a Chemical Reaction: Experiments and Calculations // *J. Phys. Chem.* 1995. V. 99. P. 13173—13180.
49. Ding W. X., Deutsch H., Dinklage A., Wilke C. Observation of a strange nonchaotic attractor in a neon glow discharge // *Phys. Rev. E*. 1997. V. 55. P. 3769—3772.
50. Ramaswamy R. Synchronization of strange nonchaotic attractors // *Phys. Rev. E*. 1997. V. 56. P. 7294—7296.
51. Zhou C., Chen T. Robust communication via synchronization between nonchaotic strange attractors // *Europhys. Lett.* 1997. V. 38. P. 261—265.

БИЛЬЯРДЫ

А. Ю. Лоскутов, А. Б. Рябов

1. Введение

Долгое время считалось, что перемешивание и хаотическое движение возможны только в системах с чрезвычайно большим числом степеней свободы, например в системах статистической механики. Однако в 1962 г. Я. Г. Синай доказал [1] (см. также [2]), что даже очень простая модель, состоящая из двух плоских упругих дисков в ящике, один из которых закреплен, может обладать свойством перемешивания (рис. 1, а). Если движущийся диск заменить материальной точкой, то мы приходим к так называемой бильярдной задаче (рис. 1, б) — динамической системе, отвечающей движению по инерции материальной точки («шара») внутри ограниченной области с условием упругого отражения от ее границы по закону «угол падения равен углу отражения» [3].

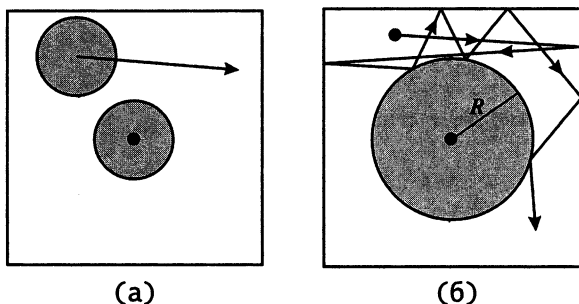


Рис. 1. Простейший вид рассеивающего бильярда (бильярда Синай)

Впервые понятие бильярда в теоретической и математической физике возникло после того, как Д. Биркгоф [4] рассмотрел задачу о движении шара на поверхности, ограниченной некоторой замкнутой кривой. Позже глубокие работы Н. С. Крылова [5], посвященные проблеме перемешивания в ансамбле упругих шаров в ящике, привели исследователей к необходимости рассмотрения задач бильярдного типа (см. [6 — 8] и цитируемую там литературу), поскольку они являются достаточно удобными моделями целого ряда физических систем. Например, многим динамическим задачам могут быть

поставлены в соответствие уравнения, описывающие траектории частиц в бильярдах заданной формы. Бильярдные модели возникают также в физике твердого тела и в некоторых геометрических приближениях отражения волн, особенно в акустике [9].

Математически плоский бильярд представляет собой обычный бильярд, только с произвольной конфигурацией стола и без луз. И хотя, строго говоря, бильярды относятся к сильно упрощенным моделям классической статистической механики, имеется естественная аналогия между некоторыми физическими задачами и системами бильярдного типа [8, 10]. Одной из таких моделей является так называемый газ Лоренца. На ее основе удастся получить ряд строгих и очень глубоких результатов.

Рассмотрим неограниченную область D в евклидовом пространстве. Допустим, что в этой области находятся определенным образом расположенные непересекающиеся шары B_i (рассеиватели) радиуса R (рис. 2). При условии, что шары неподвижны, такая система, очевидно, будет бильярдной. Она называется *газом Лоренца*: частица движется между рассеивателями, упруго отражаясь от них. Впервые газ Лоренца был введен в связи с проблемой описания движения электронов в металлах [3].

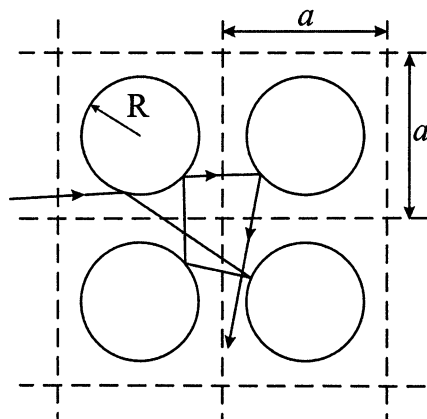


Рис. 2. Двумерный газ Лоренца с периодическим расположением рассеивателей на квадратной решетке

При периодическом расположении рассеивателей с периодом a одним из основных параметров газа Лоренца является отношение $(a/R)^2$. В зависимости от его значения различают газ Лоренца с ограниченным горизонтом ($(a/R)^2 < 4$) (рис. 3), открытым гори-

зонтом ($4 < (a/R)^2 < 8$) и бесконечным горизонтом ($(a/R)^2 > 8$). В первом случае, когда центры образуют квадратную решетку, движение частицы ограничено одной ячейкой. Поэтому для ограниченного горизонта динамика газа Лоренца аналогична динамике частицы в рассеивающем бильярде (рис. 1, б). Во втором и третьем случаях бильярдной частице будет доступна вся область D .

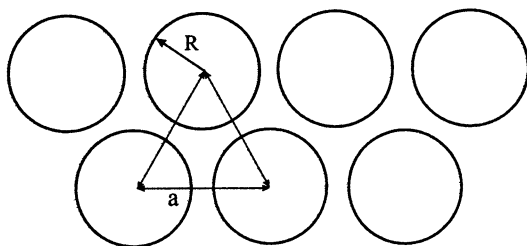


Рис. 3. Пример газа Лоренца с ограниченным горизонтом

Для бесконечного горизонта из-за увеличения вероятности длинных перелетов изменяются статистические свойства бильярда. В частности, замедляется скорость убывания корреляций и не сходится среднее значение длины свободного пробега [6, 11 — 14]. В то же время, для газа Лоренца с ограниченным и открытым горизонтом наблюдается экспоненциальный спад корреляций. Средняя длина свободного пробега определяется как $l = \pi A/P$, где A — доступная частице область бильярда, а P — периметр рассеивателей. Для системы с открытым горизонтом $l = (a^2 - \pi R^2) / 2R$, а для бильярда с бесконечным горизонтом значение l не ограничено сверху.

Для газа Лоренца с ограниченным и открытым горизонтами было доказано, что в двумерном (плоском) случае движение частицы является эргодическим, обладает свойством перемешивания и, более того, сводится к броуновскому. Также в такой системе существует и положителен коэффициент диффузии [11]. Это явилось первым строгим подтверждением рождения хаоса в динамических системах.

Формально описание бильярдных систем можно ввести следующим образом. Бильярдным столом Q является область M с кусочно-гладкой границей ∂Q . Бильярдная динамическая система в M порождается свободным движением материальной точки (бильярдного шара) с условием упругого отражения от ∂Q . Это означает, что точка всегда движется по прямым в M с постоянной скоростью до соударения с границей. Если граница в точке соударения гладкая,

то шар отражается от нее так, что тангенциальная компонента скорости не изменяется, в то время как нормальная составляющая изменяет знак. Если шар попадает в угол, то его дальнейшее движение определено неоднозначно или же не определено совсем.

Как правило, граница каждого бильярда состоит из m компонент ∂Q_i , $i = 1, 2, \dots, m$. Точки шивки компонент ∂Q_i (т. е. «углы» стола Q) называются особыми. Остальные точки q — регулярными. Если для каждой компоненты ∂Q_i рассмотреть единичные нормали $\mathbf{n}(q)$ в каждой точке $q \in \partial Q_i$, направленные внутрь области Q , то они будут определять кривизну $k(q)$ кривой ∂Q_i во всех регулярных точках q . Компонента ∂Q_i является *рассеивающей* и обозначается Γ^d , если $k(q) > 0$. Для $k(q) = 0$ и $k(q) < 0$ получим соответственно *нейтральную* Γ^0 и *фокусирующую* Γ^f компоненты границы бильярда. Таким образом, каждая из компонент границы может быть рассеивающей, фокусирующей или нейтральной. Объединение всех рассеивающих, нейтральных и фокусирующих компонент обозначим соответственно как ∂Q^+ , ∂Q^0 и ∂Q^- .

Когда граница состоит только из рассеивающих и нейтральных компонент, такой бильярд называется *рассеивающим* (или бильярдом Синая, рис. 1, б).

Бильярд, граница которого состоит только из фокусирующих или нейтральных и фокусирующих компонент называется *фокусирующим*. Самым популярным примером такого бильярда является «стадион», граница которого состоит из двух дуг окружностей одинакового радиуса и двух параллельных отрезков, соединяющих концы этих дуг (рис. 4). Несмотря на наличие фокусирующих компонент, для такого бильярда тоже характерно свойство перемешивания. Это связано с тем, что параллельные лучи после прохождения точки фокусировки начинают расходиться.

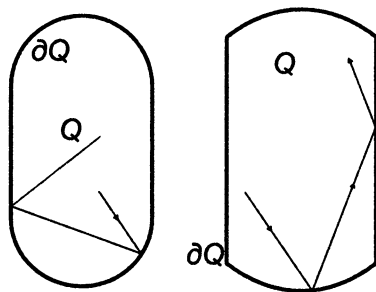


Рис. 4. Бильярд с границей в форме стадиона

При произвольных соотношениях между параметрами бильярда, если его граница имеет как рассеивающие, так и фокусирующие компоненты, динамика частицы будет определяться соотношением времен, которые она проводит в окрестности ∂Q^+ и ∂Q^- . Точный смысл этого утверждения описан в работах [15, 16], где доказано, что если на фокусирующую компоненту $\Gamma^f \subset \partial Q$, имеющую постоянную кривизну, падает параллельный пучок траекторий и испытывает серию из j отражений от Γ^f , то компонента Γ^f эффективно не влияет на уменьшение ширины фронта, а его кривизна остается ограниченной сверху.

Основные результаты, касающиеся бильярдov с фокусирующими и нейтральными компонентами, сводятся к следующему [15]. Если: а) кривизна каждой компоненты ∂Q^- постоянна, $k(q|_{\partial Q^-}) = = \text{const}$; б) никакие две фокусирующие компоненты не являются дугами одной и той же окружности; в) дуга, дополняющая любую фокусирующую компоненту до полной окружности, лежит внутри области Q , то бильярд в области, граница которого имеет рассеивающие и фокусирующие компоненты, обладает свойством перемешивания. При этом не накладывается никаких условий на соотношение длин рассеивающих и фокусирующих компонент.

Условие а можно немного ослабить, так чтобы граница фокусирующей компоненты слабо менялась. Кроме того, если Γ^f пересекает только нейтральные компоненты, то вместо условия в достаточно, чтобы центр окружности лежал внутри Q . Примеры таких бильярдov показаны на рис. 4 и 5.

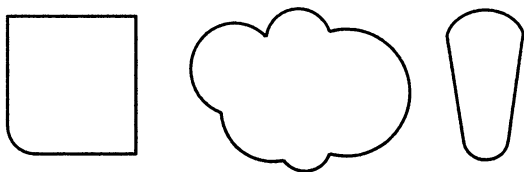


Рис. 5. Примеры перемешивающих бильярдov с фокусирующими компонентами

Более общие условия существования хаотичности в двумерных бильярдах на плоскости описаны в работе [17] (см. также приведенную там литературу).

Естественным обобщением бильярдных систем являются бильярды, границы которых могут изменяться с течением времени. К моделям такого типа сводятся некоторые важные проблемы ма-

тематической физики. Например, анализ движения частиц внутри возмущенных ядер привел к необходимости изучения квадрупольных или «динамических» резонансов. В таких задачах рассматривалось столкновение частиц с движущимися границами (полями) и, как следствие, обмен энергиями. В физике плазмы ускорение в магнитных ловушках моделируется бильярдами с изменяющейся во времени границей. Наконец, если рассматривать газ Лоренца как систему, описывающую динамику электронов в металлах, то рассеиватели должны слегка «дрожать» с небольшой амплитудой вблизи своего равновесного состояния. Более того, как сейчас ясно, бильярды с зависящими от времени границами — это область физики, открывающая большие перспективы в исследовании многих давно известных, но малоизученных проблем (см., например, [18 — 20] и приводимые там ссылки). Например, возмущение границы может приводить к замедлению частиц (см. ниже), что может оказаться очень важным в приложениях.

2. Бильярды с возмущаемыми границами

В бильярдах с возмущаемыми границами величина скорости шара меняется от столкновения к столкновению. При единичных актах столкновения шар приобретает или отдает энергию в зависимости от того, движется ли шар и граница навстречу друг другу (*встречное* столкновение), или же шар нагоняет границу (*сопутствующее* столкновение). Возникает вопрос, как будет меняться модуль скорости частицы со временем.

Оказывается, что для бильярда с зависящими от времени границами существенными оказываются его динамические свойства: если бильярд проявляет хаотическую динамику, то возмущение границы может привести к неограниченному росту скорости частицы. С другой стороны, бильярд с достаточно гладкой границей (например, в форме эллипса или окружности) не хаотичен, и колебания границы не приводят к разгону бильярдной частицы.

Первоначально идея стохастического ускорения частиц была предложена Энрико Ферми [21] для объяснения происхождения космических лучей высоких энергий. Она состояла в том, что при столкновениях заряженных частиц с беспорядочно движущимися магнитными облаками в межзвездном пространстве частицы должны в среднем ускоряться. Рассматривая облако как тело большой массы, нетрудно понять причину ускорения. Если скорости облаков,

с которыми сталкивается частица, распределены случайно, то можно сказать, что число облаков, движущихся в одном направлении, равно числу облаков, движущихся в обратном направлении. Поэтому частица будет чаще сталкиваться с теми из них, которые движутся ей навстречу. Отсюда следует, что частица чаще будет приобретать энергию, чем отдавать ее. Так возникает эффективное ускорение, называемое *ускорением Ферми*.

Для бильярдов ситуация оказалась значительно сложнее: не всякому хаотическому бильярду при возмущении границ присуще ускорение Ферми. Более того, в некоторых бильярдах, например в форме стадиона, когда кривизна фокусирующих компонент мала (см. правый рис. 4), наблюдается *замедление* частиц [20]. Это позволяет разделять частицы по скоростям и посредством простого периодического внешнего воздействия *охлаждать* частицы разреженного газа.

3. Динамика частиц в бильярдах с возмущаемыми границами

В данном разделе рассмотрено возникновение ускорения Ферми на примере обобщенного бильярда — газа Лоренца с открытым горизонтом с возмущаемой границей [18]. При этом изучаются два разных случая, соответствующие стохастическим и регулярным (гармоническим) колебаниям границы.

3.1. Газ Лоренца с неподвижной границей

Как известно, в качестве канонических переменных для такого бильярда можно выбрать азимутальный угол ϕ и угол падения α между внутренней нормалью к поверхности и вектором скорости частицы перед соударением. Введем угол отражения α^* между внешней нормалью и вектором скорости после соударения (рис. 6). Очевидно, что $\phi \in [0, 2\pi]$, а углы α и α^* изменяются в интервале $[-\pi/2, \pi/2]$.

Для описания динамики невозмущенного бильярда необходимо получить отображение $(\alpha_n, \phi_n) \rightarrow (\alpha_{n+1}, \phi_{n+1})$, переводящее значения переменных (α, ϕ) в момент перед n -м столкновением с ∂Q в их значения в момент перед $(n+1)$ -м столкновением. Из геометрических соображений понятно, что

$$\phi_n + \alpha_n^* + \pi = \phi_{n+1} + \alpha_{n+1} . \quad (1)$$

Кроме того, $\alpha_n^* = -\alpha_n$, так как углы отсчитываются в противоположные стороны.

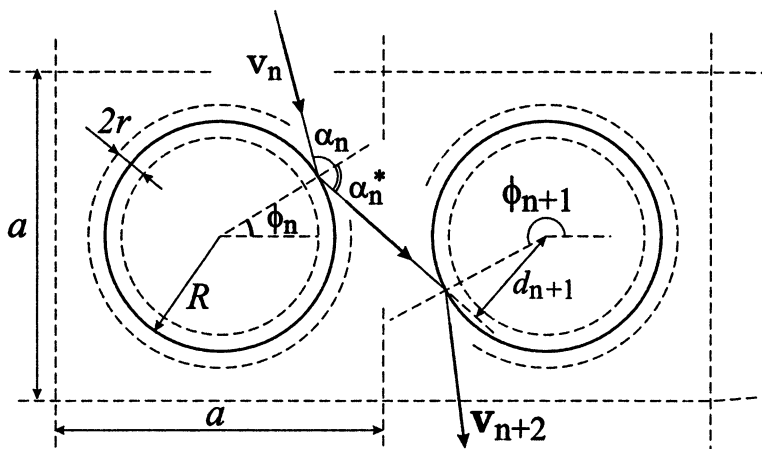


Рис. 6. Координаты в модели газа Лоренца

Введем систему координат с центром в круге, на котором произошло очередное рассеяние и найдем расстояние, на котором частица пройдет мимо другого центра, отстоящего на p ячеек по горизонтали и q ячеек по вертикали:

$$d_{n+1} = a [p \sin(\phi_n + \alpha_n^*) - q \cos(\phi_n + \alpha_n^*)] - R \sin \alpha_n^* . \quad (2)$$

Значение p считается положительным, если частица движется вправо, и отрицательным в противном случае; соответственно q положительно, если частица движется вверх, и отрицательно, если вниз. Величины p и q определяются из условия рассеяния, т. е. как минимальные по модулю целые числа, при которых выполняется $|d_{n+1}| \leq R$. Вычислив прицельный параметр d_{n+1} , легко найти угол, под которым произойдет столкновение со следующим рассеивателем:

$$\alpha_{n+1} = \arcsin \left(\frac{d_{n+1}}{R} \right) . \quad (3)$$

Якобиан отображения (1)–(3) равен $\partial(\phi_{n+1}, \alpha_{n+1})/\partial(\phi_n, \alpha_n) = \cos \alpha_n / \cos \alpha_{n+1}$. Таким образом, система сохраняет фазовый объем $\cos \alpha \, da \, d\phi$. Отсюда, в частности, следует, что если рассматриваемый бильярд является эргодическим, то величина α_n будет распределена по закону

$$\rho_\alpha(\alpha) = \frac{1}{2} \cos \alpha , \quad (4)$$

где $1/2$ — нормировочный множитель.

3.2. Газ Лоренца с осциллирующими границами рассеивателей

Допустим, что рассеивающая компонента ∂Q^+ границы ∂Q сжимается и расширяется (см. рис. 6) так, что ее радиус меняется по закону $R = R(t) = R + r(t)$, где $\max |r(t)| \ll R$. Тогда скорость границы будет зависеть от времени по закону $u(t) = \dot{r}(t)$. Далее для определенности будем считать, что $u(t) = u_0 \cos(\omega t)$, где $u_0 = \omega r_0$. В этом случае, помимо величин α и ϕ , для описания поведения системы необходимо ввести еще две динамические переменные: скорость частицы v и время столкновения t . Учитывая, что при отражении от рассеивателя меняется только нормальная (радиальная) компонента скорости, а тангенциальная остается неизменной, получим отображение для модуля скорости частицы после столкновения:

$$v_{n+1} = \sqrt{v_n^2 - 4u_n v_n \cos \alpha_n + 4u_n^2}, \quad (5)$$

где $u_n \equiv u_0 \cos \omega t_n$ — скорость границы рассеивателя в момент n -го столкновения. В свою очередь, связь между углом падения и углом отражения можно записать в виде $\alpha_n^* = -\arcsin(v_n \sin \alpha_n / v_{n+1})$. Теперь, вычисляя расстояние между последовательными соударениями, нетрудно получить отображение для момента столкновения t_n :

$$t_{n+1} = t_n + \frac{l_{n+1}}{v_{n+1}}, \quad (6)$$

$$l_{n+1} = \sqrt{\left(R(\cos \phi_{n+1} - \cos \phi_n) - pa\right)^2 + \left(R(\sin \phi_{n+1} - \sin \phi_n) - qa\right)^2},$$

где l_n — длина свободного пробега. В предположении $r \ll R$ отображения для переменной ϕ и прицельного параметра d останутся теми же, что и для невозмущенного бильярда (см. формулы (1) и (2)).

Во время соударений с возмущаемой границей величина скорости бильярдного шара постоянно меняется. Как показывают исследования [18, 19], эти изменения скорости носят случайный характер. Поэтому далее рассмотрим ансамбль частиц и найдем распределение частиц по скоростям и среднюю скорость в зависимости от времени t и числа столкновений n (число столкновений и время не пропорциональны друг другу, так как за одно и то же время «быстрая» частица испытывает больше соударений, чем «медленная»).

3.2.1. Стохастически возмущаемая граница рассеивателей

Пусть скорость границы рассеивателя, с которым происходит n -е столкновение, равна

$$u_n(t) = u_0 \cos \varphi_n, \quad (7)$$

где u_0 — это амплитудное значение скорости колебания границы, а совокупность $\{\varphi_n\}$ — множество некоррелированных между собой случайных величин, равномерно распределенных в интервале $[0, 2\pi]$. Найдем функцию распределения частиц по скоростям и зависимость средней скорости ансамбля частиц от числа соударений n и времени t . В случае малых скоростей частицы ($v \ll u_0$) в выражении для скорости (5) основную роль играет последний член и, следовательно, $v_{n+1} \approx 2|u(t_n)|$.

Если колебания границы заданы в соответствии с выражением (7), то $\langle v_{n+1} \rangle \approx 2 \langle |u(t_n)| \rangle_t = 4u_0/\pi$. Таким образом, уже после первого столкновения средняя скорость становится больше величины u_0 .

Теперь найдем изменение скорости в случае $v \gg u_0$. Раскладывая (5) в ряд Тейлора по параметру u/v , получим выражение для изменения скорости:

$$\Delta v_n = v_{n+1} - v_n = -2u_n \cos \alpha_n + 2 \frac{u_n^2}{v_n} \sin^2 \alpha_n + v_n O \left(\left(\frac{u_n}{v_n} \right)^3 \right), \quad (8)$$

где u_n — скорость границы рассеивателя при n -ом столкновении.

Используя соотношение (4) и условие равномерного распределения по фазе колебаний в момент соударения, найдем $\langle \Delta v_n \rangle$ и $\langle (\Delta v_n)^2 \rangle$:

$$\begin{aligned} \mu_s &\equiv \langle \Delta v_n \rangle = \frac{M_s}{v}, \\ \sigma_s^2 &\equiv \langle (\Delta v_n)^2 \rangle = \frac{4}{3} u_0^2. \end{aligned} \quad (9)$$

Здесь для простоты дальнейших выкладок введена величина $M_s \equiv u_0^2/3$; индекс s обозначает стохастический случай. При подсчете среднего ненулевой вклад дает только второе слагаемое, а при вычислении дисперсии достаточно учесть вклад, связанный с первым слагаемым.

При достаточно больших n от первого уравнения системы (9) можно перейти к дифференциальному уравнению

$$\mu_s = \frac{\langle \Delta v_n \rangle}{1} = \frac{\partial v(n)}{\partial n} = \frac{M_s}{v(n)}. \quad (10)$$

Его решение с начальным условием $v(0) = v_0$ дает зависимость наиболее вероятной скорости от числа соударений:

$$v(n) = \sqrt{2M_s n + v_0^2}. \quad (11)$$

Поскольку скорость частиц представляется в виде суммы случайных независимых величин Δv_n с известным средним и дисперсией, то на основании центральной предельной теоремы Ляпунова можно утверждать, что плотность распределения случайной величины $v_n = v_0 + \sum_{i=1}^n \Delta v_i$ будет стремиться к нормальной со средним $v(n)$ и дисперсией $n\sigma_s^2$. Таким образом, распределение частиц по скоростям имеет вид расплывающегося гауссова пакета. Максимум распределения дается наиболее вероятной скоростью $v(n)$, которая растет как корень из n .

Приведенные рассуждения верны только в случае достаточно *больших* скоростей частицы ($v \gg u_0$). Чтобы описать распределение при *малых* скоростях, введем еще одно условие — отсутствие потока частиц в область отрицательных скоростей: $(v\partial\rho/\partial v)|_{v=0} = 0$. Хорошо известно, что удовлетворяющее этому условию гауссово распределение имеет вид

$$\rho(v, n) = \frac{1}{\sigma_s \sqrt{2\pi n}} \left[\exp\left(-\frac{(v - v(n))^2}{2\sigma_s^2 n}\right) + \exp\left(-\frac{(v + v(n))^2}{2\sigma_s^2 n}\right) \right]. \quad (12)$$

Это позволяет найти среднюю скорость ансамбля частиц в зависимости от числа столкновений:

$$V(n) = \sigma_s \sqrt{\frac{2n}{\pi}} \exp\left(-\frac{v(n)^2}{2\sigma_s^2 n}\right) + v(n) \Phi\left(\frac{v(n)}{\sigma_s \sqrt{2n}}\right), \quad (13)$$

где $\Phi(x) = 2/\sqrt{\pi} \int_0^x \exp(-x^2) dx$ — известный интеграл ошибок. Здесь и далее большой буквой V обозначена средняя скорость ансамбля частиц. Подставляя все значения и раскладывая в ряд выражение для скорости, можно найти, что

$$V(n) = C\sqrt{n} + O\left(\frac{1}{\sqrt{n}}\right), \quad (14)$$

где постоянная $C = \sqrt{2} \left(\sigma_s e^{-M_s/\sigma_s^2} / \sqrt{\pi} + \Phi(\sqrt{M_s}/\sigma_s) \sqrt{M_s} \right) \approx \approx 1,143u_0$.

Таким образом, соотношения (12) и (14) дают *распределение по скоростям и среднюю скорость ансамбля частиц в зависимости от числа столкновений*.

Для нахождения *средней скорости частиц в зависимости от времени* воспользуемся уравнением Фоккера — Планка:

$$\frac{\partial \rho(v, t)}{\partial t} = -\frac{\partial}{\partial v} (A\rho(v, t)) + \frac{1}{2} \frac{\partial^2}{\partial v^2} (B\rho(v, t)),$$

где коэффициенты A и B имеют следующий вид: $A \equiv \langle \Delta v / \tau \rangle = M_s / l$, $B \equiv \langle \Delta v^2 / \tau \rangle = \sigma_s^2 v / l$. Здесь среднее время между столкновениями $\tau = l / v$, l — средняя длина свободного пробега, а Δv и Δv^2 взяты из (9). Теперь, подставляя полученные коэффициенты в уравнение, найдем:

$$\frac{\partial \rho(v, t)}{\partial t} = -\frac{M_s}{l} \frac{\partial}{\partial v} \rho(v, t) + \frac{1}{2} \frac{\sigma_s^2}{l} \frac{\partial^2}{\partial v^2} (v\rho(v, t)). \quad (15)$$

Если параметры M_s и σ_s заданы в соответствии с выражениями (9), то решение этого уравнения при скоростях частиц много больших начальной скорости, т. е. через достаточно большой интервал времени, будет стремиться к $\rho(v, t) = \exp\left(-\frac{v}{2tA}\right) / \sqrt{2tA\pi v}$, где $A = M_s / l$. Используя его, получаем среднюю скорость частиц:

$$V(t) = \frac{M_s}{l} t + v_0 = \frac{1}{3} \frac{u_0^2}{l} t + v_0. \quad (16)$$

Таким образом, в исследуемой системе наблюдается ускорение Ферми, причем скорость растет линейно со временем.

3.2.2. Периодически возмущаемые границы рассеивателей

Пусть все границы рассеивателей сжимаются и расширяются по некоторому периодическому закону и в одной и той же фазе. Тогда в течение одной половины периода скорость частицы при столкновениях будет увеличиваться, а в течение другой — уменьшаться. При достаточно больших скоростях время свободного пробега τ_s становится меньше периода колебаний радиусов рассеивателей. Это приводит к корреляциям изменений скорости частицы, так что ее последовательные приращения (8) уже нельзя считать независимыми.

Для проведения анализа изменений скорости в хаотических билльярдах с периодически осциллирующими границами можно предложить следующий подход. Рассмотрим приближенное отображение для скорости (8). Поскольку корреляции величин α_n по-прежнему

экспоненциально спадают (так как это определяется геометрией бильярда), то соотношения можно усреднить по α , пользуясь распределением (4). Тогда

$$\langle \Delta v \rangle_\alpha = -\frac{\pi}{2} u_0 \cos \omega t_n + \frac{u_0^2 \cos \omega t_n}{v_n}. \quad (17)$$

В течение периода колебаний наибольший вклад в изменение скорости дает первое слагаемое. Следовательно, для учета корреляций в первом приближении достаточно учитывать изменения скорости, связанные только с первым слагаемым, а вторым слагаемым можно пренебречь. С другой стороны, учет корреляционных поправок во втором слагаемом приводит к членам более высокого порядка малости, чем среднее от него. Таким образом, можно не учитывать корреляционные эффекты во втором слагаемом. Поэтому допустимо независимо вычислить две величины, связанные с первым и вторым слагаемым, т. е. $\langle \Delta v \rangle = \langle \Delta v_I \rangle + \langle \Delta v_{II} \rangle$, причем $\langle \Delta v_{II} \rangle = u_0^2 / (3v)$, что совпадает с величиной μ_s для стохастического случая (см. выражения (9)), а $\langle \Delta v_I \rangle$ — поправка, обусловленная наличием корреляций. Отбрасывая в выражении (17) второе слагаемое, получим следующее отображение для вычисления величины $\langle \Delta v_I \rangle$:

$$\begin{aligned} v_{n+1} &= v_n + \gamma \cos \theta_n, \\ \theta_{n+1} &= \theta_n + \frac{l_{n+1} \omega}{v_{n+1}}. \end{aligned} \quad (18)$$

Здесь $\gamma = -\pi u_0 / 2$ и вместо времени была введена фаза столкновения $\theta_n \equiv \omega t_n$. Приведенное отображение в точности соответствует известному отображению Улама [22 — 29] с той лишь разницей, что в данном случае длина свободного пробега l_n является случайной величиной, которая может принимать значения из определенного интервала.

Как показывает анализ, для данного отображения (при тех же значениях u_0 и ω , какие были выбраны при исследовании газа Лоренца) первый момент приращений скорости положителен, но при этом он все равно меньше наблюдаемого приращения скорости в газе Лоренца.

Оценим дисперсию и скорость спада корреляций изменений скорости. Предположим, что скорость частицы столь велика, что ее изменение за n соударений намного меньше ее значения. Понятно, что для выполнения этого условия при любом n достаточно соответствующим образом выбрать v и u_0 . Найдем корреляции приращений

скорости Δv_m и Δv_{m+n} (см. формулу (8)) при $n \rightarrow \infty$. Учитывая в первом приближении только первые слагаемые, из (8) получим

$$R(n) \equiv \langle \Delta v_m \Delta v_{m+n} \rangle = u_0^2 \frac{\pi^2}{4} \langle \cos \omega t_m \cos \omega t_{m+n} \rangle ,$$

где в соответствии с (4) учтено, что $\langle \cos \alpha_n \rangle = \pi/4$. Полагая частоту колебаний равной единице и вводя обозначения $S_n \equiv \sum_{i=1}^n \tau_{m+i}$, где $\tau_i = t_i - t_{i-1}$, найдем, что $\langle \cos t_m \cos t_{m+n} \rangle = \langle \cos t_m \cos (t_m + S_n) \rangle$. Величину S_n можно представить в виде $S_n = \sum_{i=1}^n (l + \Delta l_i)/v$, где Δl_i — отклонение от средней длины свободного пробега при i -м перелете. Поскольку S_n является суммой независимых случайных величин, то ее распределение при больших n будет стремиться к нормальному $N(nl, n\sigma_l^2)$, где σ_l^2 — дисперсия длины свободного пробега. Раскладывая косинус суммы и производя усреднение по S_n , получим следующее выражение для корреляционной функции приращений скорости:

$$R(n) \simeq \frac{\pi^2}{8} u_0^2 \cos(\omega n \tau) \exp\left(-\frac{n}{N}\right), \quad (19)$$

где ω — частота колебаний рассеивателя и $N = v^2 / (\omega^2 \sigma_l^2)$. Таким образом, последовательные изменения скорости скоррелированы тем сильнее, чем выше скорость. При этом число соударений N , после которых корреляции спадают в e раз, увеличивается как v^2 . Отметим, что число столкновений за период растет пропорционально v . Таким образом, для правильного вычисления дисперсии необходимо учитывать тем большее число периодов колебаний, чем выше скорость частицы. Однако вопрос о том, каким образом это можно сделать, пока остается открытым.

Для того чтобы в первом приближении оценить дисперсию, рассмотрим изменение скорости только после двух последовательных соударений с границей. При этом будем считать, что корреляции между тремя и более изменениями пренебрежимо малы. В пределе больших скоростей бильярдной частицы коррелятор последовательных приращений скорости можно оценить как

$$\langle \Delta v_n \Delta v_{n+1} \rangle = u_0^2 \frac{\pi^2}{4} \langle \cos^2 \omega t_n (1 - O(\tau^2)) \rangle = u_0^2 \frac{\pi^2}{8} + O\left(\frac{u_0^2}{v^2}\right) .$$

Учитывая это выражение и соотношение (9), получим:

$$\sigma_r^2 = \frac{\langle (\Delta v_n + \Delta v_{n+1})^2 \rangle}{2} \approx \left(\frac{4}{3} + \frac{\pi^2}{8} \right) u_0^2 . \quad (20)$$

Таким образом, оценки, приведенные в этом пункте, свидетельствуют о том, что ускорение частиц должно наблюдаться в хаотических бильярдах с периодически осциллирующими границами. По-видимому, можно утверждать, что *детерминированная случайность* — *достаточное условие возникновения ускорения Ферми*.

3.2.3. Численные результаты

В этом разделе приводятся численно полученные зависимости скорости частиц от числа столкновений и времени, которые сопоставляются с аналитическими, полученными выше. Вычисления производились на основе модели газа Лоренца со следующими параметрами: амплитудное значение скорости колебаний границы рассеивателя $u_0 = 0,01$; радиус рассеивателей $R = 0,4$; расстояние между центрами $a = 1$; частота осцилляций границы $\omega = 1$; начальная скорость $v_0 = 1$. Таким образом, соответствующее аналитически полученное значение длины свободного пробега составляет $l = 0,6216815$. Численно найденные длина свободного пробега (см. соотношения (6)) и ее дисперсия при данной геометрии бильярда составляют соответственно $l = 0,62163 \pm 0,00003$ и $\sigma_l^2 = 0,657 \pm 0,001$.

В каждом случае было построено 100 реализаций динамики бильярдной частицы, которые отличались друг от друга начальными значениями α и ϕ , которые выбирались случайным образом. Исследовались два различных случая: стохастические колебания границ рассеивателей при равномерном распределении фазы колебаний и регулярные колебания границ. Динамика бильярдного шара (частицы) определялась из отображения, полученного в п. 4.2. При этом скорость колебаний границ рассеивателей в момент столкновения в первом случае определялась как $u_n = u_0 \cos \varphi_n$ (где φ_n — равномерно распределенная на интервале $[0, 2\pi)$ случайная величина), а во втором — как $u_n = u_0 \cos \omega t_n$ (где t_n — момент соударения частицы с границей).

Усредненные зависимости скорости от числа столкновений и времени представлены на рис. 7, а, б. На обоих графиках сплошные линии отвечают регулярному случаю возмущения границы, а пунктир — стохастическому. При этом в стохастическом случае величины M_s и σ_s были найдены из уравнений (9), а в регулярном — взяты предельные значения M_r и σ_r , полученные численно.

Как следует из рис. 7, рост средней скорости близок к линейному, и аналитические аппроксимации средней скорости на основании соотношений (14), (16), дают неплохое согласие с численным экспериментом.

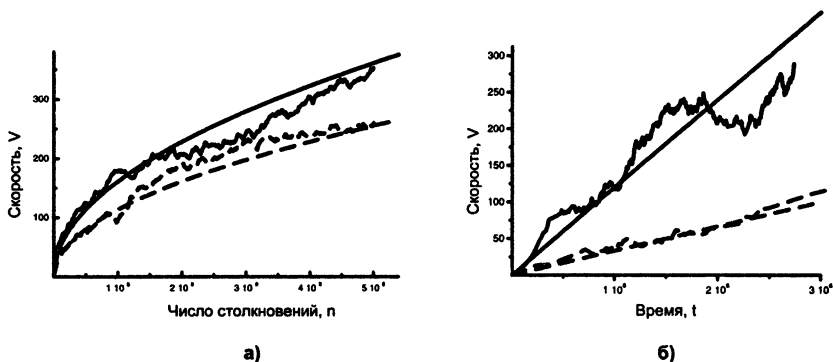


Рис. 7. Зависимости средней скорости частиц в газе Лоренца от числа столкновений (а); от времени (б). Их аппроксимации на основе уравнений (14) и (16) соответственно.

3.3. Бильярд в форме «стадион» с неподвижной границей

Механизмы возникновения хаоса в рассеивающих и фокусирующих бильярдах различны. В первом случае два близких пучка частиц, попав на рассеивающую компоненту, начинают сразу расходиться, тогда как во втором случае после отражения от фокусирующей компоненты два близких пучка сходятся до точки пересечения. Хаос в таком бильярде возникает, если время, в течение которого пучки сходятся, меньше времени расхождения.

Сначала рассмотрим *бильярд с невозмущенными границами* и сравним результаты точного и приближенного анализа в предположении, что кривизна фокусирующих компонент мала (рис. 8):

$$l \gg a \gg b. \quad (21)$$

Это позволит более полно понять динамику бильярдной частицы.

Пусть фокусирующие компоненты представлены дугами окружности радиуса R и угловой меры 2Φ , симметричными относительно вертикальной оси бильярда (рис. 8). Из геометрических соображений вытекают следующие соотношения:

$$R = \frac{a^2 + 4b^2}{8b}; \quad \Phi = \arcsin \frac{a}{2R}. \quad (22)$$

Условие стохастичности для такого бильярда выполнено, если дуга, дополняющая фокусирующую компоненту до полной окруж-

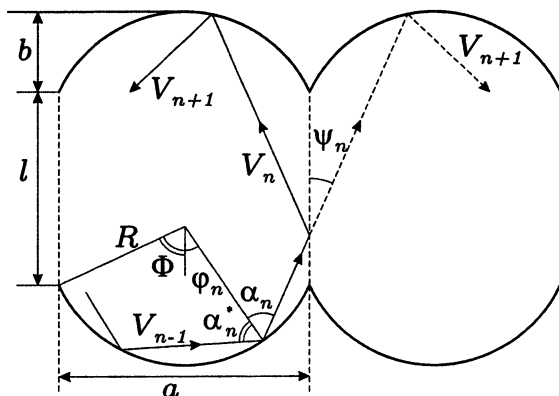


Рис. 8. Бильярд типа стадион с границей в виде дуги окружности

ности, принадлежит области Q [15]. При $b \ll a$ это дает:

$$\frac{L}{2R} \approx \frac{4bL}{a^2} > 1. \quad (23)$$

Введем динамические переменные, как показано на рис. 8. Положительными для углов φ_n и α_n^* примем направления против часовой стрелки, а для угла α_n — по часовой. Если граница бильярда неподвижна, то угол падения α_n^* равен углу отражения α_n . Пусть V_n — скорость частицы, а t_n — время n -го столкновения с границей.

Построим *точное отображение*, описывающее динамику частицы в таком бильярде. Для этого необходимо рассмотреть два случая: после очередного столкновения с фокусирующей компонентой частица сталкивается с ней же (парные столкновения), либо следующее столкновение происходит с другой фокусирующей компонентой.

В случае парных столкновений простые геометрические рассуждения [20] приводят к отображению вида (рис. 9):

$$\begin{aligned} \alpha_{n+1}^* &= \alpha_n, \\ \alpha_{n+1} &= \alpha_{n+1}^*, \\ \varphi_{n+1} &= \varphi_n + \pi - 2\alpha_n \pmod{2\pi}, \\ t_{n+1} &= t_n + \frac{2R \cos \alpha_n}{V_n}. \end{aligned} \quad (24)$$

Если $|\varphi_{n+1}| < \Phi$, то частица продолжает каскад столкновений с одной границей. В противном случае $(n+1)$ -е столкновение произойдет с другой фокусирующей компонентой.

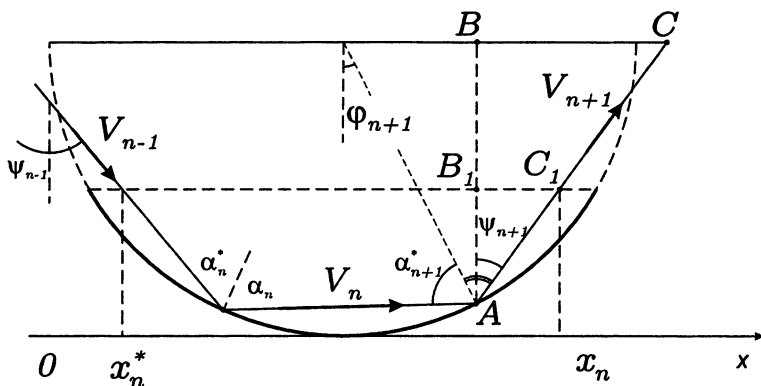


Рис. 9. Парные столкновения с границей

В случае перехода от одной фокусирующей компоненты к другой изображение запишется как

$$\begin{aligned}
 \alpha_{n+1}^* &= \arcsin \left[\sin(\psi_n + \Phi) - \frac{x_{n+1}^*}{R} \cos \psi_n \right], \\
 \alpha_{n+1} &= \alpha_{n+1}^*, \\
 \varphi_{n+1} &= \psi_n - \alpha_{n+1}^*, \\
 t_{n+1} &= t_n + \frac{R(\cos \varphi_n + \cos \varphi_{n+1} - 2 \cos \Phi) + l}{V_n \cos \psi_n},
 \end{aligned} \tag{25}$$

где введены дополнительные величины $\psi_n = \alpha_n - \varphi_n$, $x_n = R/(\cos \psi_n) [\sin \alpha_n + \sin(\Phi - \psi_n)]$, $x_{n+1}^* = x_n + l \tan \psi_n \pmod{a}$. При выводе данных отображений не использовались никакие упрощающие соотношения. Выражения (24) соответствуют серии последовательных столкновений частицы с одной фокусирующей компонентой, а (24) — переходу от одной границы к другой.

На рис. 10, а представлен фазовый портрет бильярда типа «стадион», задаваемый отображениями (24) — (25). В качестве координаты ξ_n выбиралась величина $\xi_n = 1/2 + (R \sin \varphi_n)/a$, равная нормированной на ширину бильярда проекции точки столкновения частицы с фокусирующей компонентой на ось Ox . Поскольку бильярд обладает осевой симметрией, то и исследуемые отображения должны обладать симметрией относительно замены $\xi \rightarrow 1 - \xi$, $\psi \rightarrow -\psi$. Следовательно, достаточно получить результаты в области неотрицательных ψ . При численном анализе интервал изменения величин ξ и ψ разбивался на 300×300 ячеек. Интенсивность черного цвета на рисунке пропорциональна числу точек, попавших в данную ячейку.

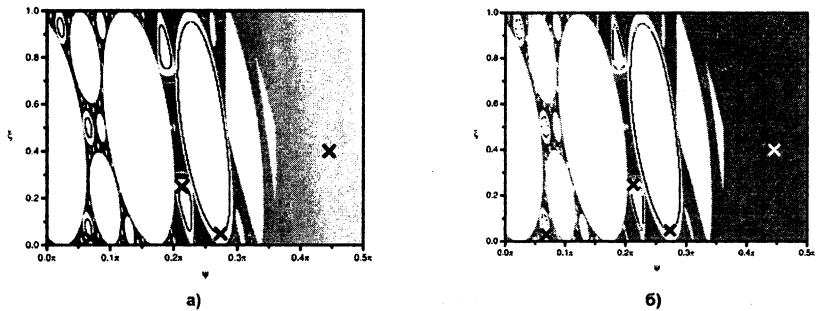


Рис. 10. Фазовый портрет бильярда типа «стадион» с границей в виде дуги окружности (а) и параболической границей (б) при $a = 0,5$, $b = 0,01$ и $l = 1$

Для построения графика использовалось четыре траектории. Одна из них начиналась в стохастической области и включала $5 \cdot 10^8$ итераций. Три другие начинались в регулярной области и содержали по 10^7 итераций. Начальные условия отмечены на графике крестиком.

Из рисунка следует, что в системе существуют устойчивые неподвижные точки, окруженные инвариантными кривыми и соответствующие различным резонансам. Динамика частиц в окрестностях этих точек является регулярной, отвечающей движению вдоль таких инвариантных кривых. Области, соответствующие различным резонансам, разделяются сепаратрисами, окруженными стохастическим слоем. Ширина этого слоя определяется величиной нелинейности в системе. Частица, начав свое движение в таком слое, посещает все доступные ей области случайным образом. С ростом нелинейности неподвижные точки могут потерять свою устойчивость, в результате чего образуется область глобальной стохастичности, в которой частице доступно уже все фазовое пространство и она может диффундировать в любом направлении. Таким образом, можно выделить три слоя, в которых динамика системы будет качественно различна.

Для нахождения стационарных точек и анализа их устойчивости используем определенные упрощения в отображении (24) — (25). Аппроксимируем фокусирующую компоненту бильярда частью параболы $\chi(x) = 4bx(x - a)/a^2$, а качестве координат выберем угол ψ между вектором скорости и вертикалью, $\psi \in [-\pi/2, \pi/2]$, а также величину x — проекцию на ось Ox точки столкновения частицы с границей, $x \in [0, a)$ (рис. 11). Тогда, принимая во внимание прибли-

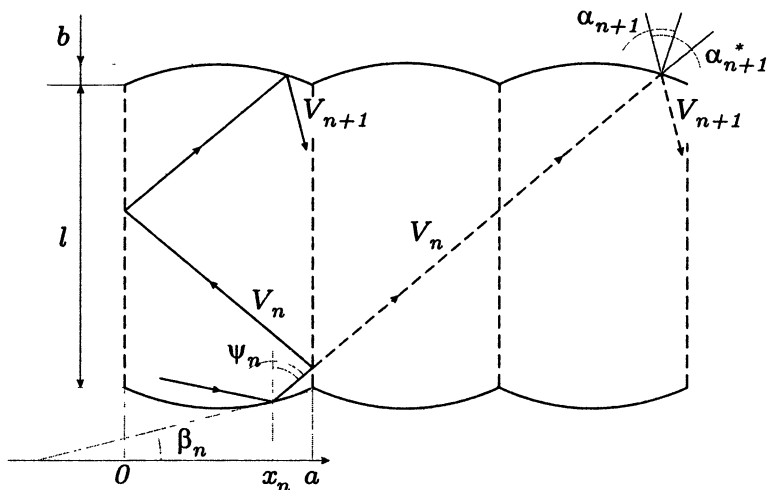


Рис. 11. Развертка бильярда типа «стадион»

жение (21), найдем:

$$\begin{aligned} x_{n+1} &= x_n + l \tan \psi_{n+1} \pmod{a}, \\ \psi_{n+1} &= \psi_n - 2\beta(x_{n+1}), \end{aligned} \quad (26)$$

где $\beta(x) = \arctan(\chi'(x))$ — угол наклона касательной в точке столкновения. В этом случае, очевидно, $\beta \approx 4b(2x - a)/a^2$.

Для дальнейших упрощений сделаем замену: $\xi = x/a$, $\xi \in [0, 1)$. В этих переменных отображение примет вид

$$\begin{aligned} \xi_{n+1} &= \xi_n + \frac{l}{a} \tan \psi_n \pmod{1}, \\ \psi_{n+1} &= \psi_n - \frac{8b}{a}(2\xi_{n+1} - 1). \end{aligned} \quad (27)$$

Одно из семейств неподвижных точек бильярда, очевидно, можно записать в виде $\{\xi = 1/2, \psi_s = \arctan(ma/l)\}$, $m \in Z$. Эти точки соответствуют столкновениям частицы с центром дуги. Причем, если $m = 0$, то частица движется строго вертикально, если $m = 1$, то при каждом соударении частица смещается на одну ячейку и т.д. Исследуем устойчивость таких точек. Для этого линеаризуем отображение, сделав замену $\xi_n = \Delta\xi_n + 1/2$, $\psi_n = \Delta\psi_n + \arctan(ma/l)$.

Тогда, раскладывая в ряд по $\Delta\psi$, найдем:

$$\begin{aligned}\Delta\xi_{n+1} &= \Delta\xi_n + \frac{l}{a \cos^2 \psi_s} \Delta\psi_n + O(\Delta\psi_n^2), \\ \Delta\psi_{n+1} &= \Delta\psi_n - \frac{16b}{a} \Delta\xi_{n+1},\end{aligned}$$

где $\psi_s = \arctan(ma/l)$. При этом матрица преобразования имеет вид

$$A = \begin{pmatrix} 1 & \frac{l}{a \cos^2 \psi_s} \\ -\frac{16b}{a} & 1 - \frac{16bl}{a^2 \cos^2 \psi_s} \end{pmatrix}.$$

Нетрудно убедиться, что $\det A = 1$. Следовательно, полученное отображение сохраняет площадь. Критерий устойчивости стационарных точек можно записать как $|Tr A| \leq 2$. Отсюда $\cos^2 \psi_s \geq \geq 4bl/a^2$ или $m^2 \leq l/(4b) - l^2/a^2$.

Разрушение всех резонансов и переход к хаосу произойдет, если $4bl/a^2 > 1$. Нетрудно понять [20], что движение частицы вокруг устойчивой точки описывается отображением поворота с числом вращения

$$\sigma = \arccos \left(1 - \frac{8bl}{a^2 \cos^2 \psi_s} \right).$$

Движение частицы внутри бильярда с определенной скоростью V порождает поток, в котором можно ввести время t . Время между последовательными столкновениями приблизительно равно $\tau \approx \frac{l}{\cos \psi_s} \frac{1}{V}$. Поэтому период вращения

$$T_1 = \frac{2\pi}{\sigma} \tau = \frac{2\pi l \cos \psi_s}{(\cos^2 \psi_s - 8bl/a^2) V}.$$

На рис. 10, б представлен фазовый портрет бильярда с фокусирующими компонентами в форме параболы, задаваемого *приближенным отображением* (27). Из его сравнения с рис. 10, а, построенным на основе точного отображения (24) — (25), видно, что во втором случае траектория равномерно распределена по всей стохастической области. Однако в первом случае большая плотность соответствует области $\psi \geq 0$, а область $\psi \leq \pi/2$ практически пуста. Разницу результатов объясняет рис. 12. Приближение малой глубины фокусирующей компоненты, сделанное при выводе отображения (27), буквально означает следующее: при пересечении отрезка, соединяющего концы параболы в точке B мы считаем, что соударение

происходит в точке A , являющейся проекцией точки B на дугу. На самом деле соударение происходит в точке C . В результате для точного отображения при большом ψ соударение будет происходить с большей вероятностью с правой (на рисунке) частью дуги, так что угол ψ будет уменьшаться. Таким образом, частицы как бы выталкиваются в область малых ψ и большая часть их движения проходит в окрестности регулярных областей.

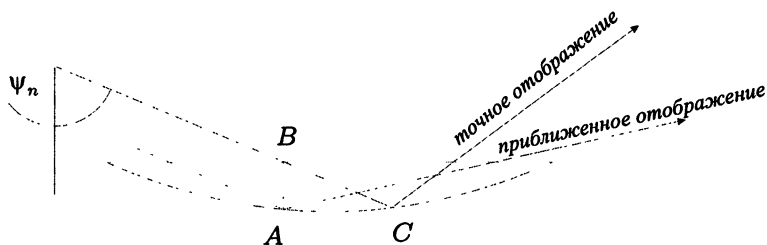


Рис. 12. Разница траекторий частицы в точном и приближенном отображении

3.4. Бильярд типа «стадион» с возмущаемой границей

В случае невозмущенного бильярда движение частицы вокруг устойчивой точки происходит по эллипсу и описывается в переменных действие — угол. Если граница бильярда возмущается, то частица может переходить из стохастической области в регулярную и обратно. Когда скорость частиц много больше скорости движения границ, можно показать [20], что частицы, попадая в окрестность устойчивых точек, будут двигаться по спирали, сначала приближаясь к центру, а потом удаляясь от него. При этом частота вращения вокруг устойчивой точки остается такой же, как и в невозмущенной системе. Кроме того, можно показать, что при определенной скорости частиц наступает резонанс между вращением частицы вокруг устойчивой точки и возмущением границ. В результате вся область в ее окрестности становится доступной. Резонансную скорость можно выразить следующим образом:

$$V_r = \frac{l}{2U_0 \cos \psi_s \arccos(1 - 8bl/(a \cos \psi_s)^2)} \quad (28)$$

Пусть фокусирующие компоненты возмущаются таким образом, что скорость их движения в каждой точке одинакова по модулю и

направлена по нормали к поверхности. Предположим, что модуль скорости границы зависит от времени по некоторому периодическому закону $U(t) = U_0 f(\omega(t + t_0))$, где ω — частота колебаний. Будем считать, что $U_0/\omega \ll l$, т. е. смещение границы мало. В этом случае отображение, описывающее движение бильярдного шара, можно записать как

$$\begin{aligned} V_n &= \sqrt{V_{n-1}^2 + 4V_{n-1} \cos \alpha_n^* U_n + 4U_n^2}, \\ \alpha_n &= \arcsin \left(\frac{V_{n-1}}{V_n} \sin \alpha_n^* \right), \end{aligned} \quad (29)$$

$$\left. \begin{aligned} \alpha_{n+1}^* &= \alpha_n \\ \varphi_{n+1} &= \varphi_n + \pi - 2\alpha_n \pmod{2\pi} \\ t_{n+1} &= t_n + \frac{2R \cos \alpha_n}{V_n} \end{aligned} \right\}, \quad \text{если } |\varphi_{n+1}| \leq \Phi; \quad (30)$$

$$\left. \begin{aligned} \psi_n &= \alpha_n - \varphi_n \\ x_n &= \frac{R}{\cos \psi_n} [\sin \alpha_n + \sin (\Phi - \psi_n)] \\ x_{n+1}^* &= x_n + l \tan \psi_n \pmod{a} \\ \alpha_{n+1}^* &= \arcsin \left[\sin (\psi_n + \Phi) - \frac{x_{n+1}^*}{R} \cos \psi_n \right] \\ \varphi_{n+1} &= \psi_n - \alpha_{n+1}^* \\ t_{n+1} &= t_n + \frac{R(\cos \varphi_n + \cos \varphi_{n+1} - 2 \cos \Phi) + l}{V_n \cos \psi_n} \end{aligned} \right\}, \quad (31)$$

если $|\varphi_n + \pi - 2\alpha_n| > \Phi$.

Отметим, что при его выводе не использовалось соотношение (21). Единственное приближение, которое было сделано, — учет малости смещения границ бильярда. Выражения (30) соответствуют серии последовательных столкновений частицы с одной фокусирующей компонентой, а (31) — переходу от одной границы к другой.

3.4.1. Численный анализ

Численное исследование бильярда типа «стадион» с возмущаемой границей, описываемого отображениями (29) — (31), проводилось при тех же значениях геометрических параметров, что и в предыдущем пункте. При этом амплитуда осцилляций границы $U_0 = 0,01$, а частота $\omega = 1$.

Как указывалось ранее, для разных скоростей частиц фазовые портреты такой системы должны отличаться. На рис. 13 приведена зависимость резонансной скорости от угла ψ_s , заданная выражением (28). Из него следует, что в области от 0 до $\psi_{s \max}$ ее значение меняется в узких пределах.

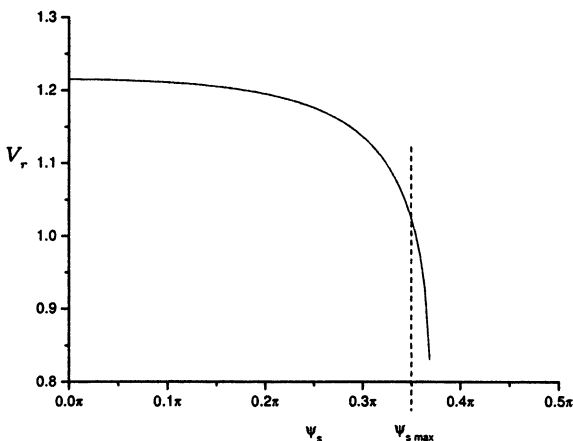


Рис. 13. Резонансная скорость в зависимости от ψ_s

Фазовые портреты для возмущенного бильярда представлены на рис. 14. Для исследования рассматривались ансамбли с различными начальными скоростями V_0 и начальными условиями, выбираемыми случайным образом в стохастической области невозмущенного бильярда. После того как скорость частицы отклонялась от начальной на 10%, инициализация частицы проводилась вновь. Для построения фазового портрета плоскость была поделена на 300×300 ячеек. При попадании координат частицы (ξ_n, ψ_n) в определенную ячейку, значение в ячейке увеличивалось на величину $\Delta V_n = V_n - V_{n-1}$. Таким образом, фазовый портрет дает представление о том, как в среднем происходит изменение скорости. Вертикальная штриховка соответствует увеличению скорости, горизонтальная — уменьшению. Черным цветом показаны области, недоступные частице.

Нетрудно видеть, что если $V_0 \neq V_r$, то вокруг устойчивых точек существуют области, окруженные инвариантными кривыми. В окрестности устойчивых точек, ставшей доступной частице в результате возмущения системы, хорошо выделяются области преимущественного увеличения и уменьшения скорости. Они меняются местами в зависимости от того, больше или меньше резонансной начальная скорость частиц. Если $V_0 = V_r$, то все окрестности устойчивых точек, кроме центральной ($\psi_0 = 0, \xi_0 = 1/2$), становятся доступными частице. При этом нет ярко выраженных областей ускорения частиц.

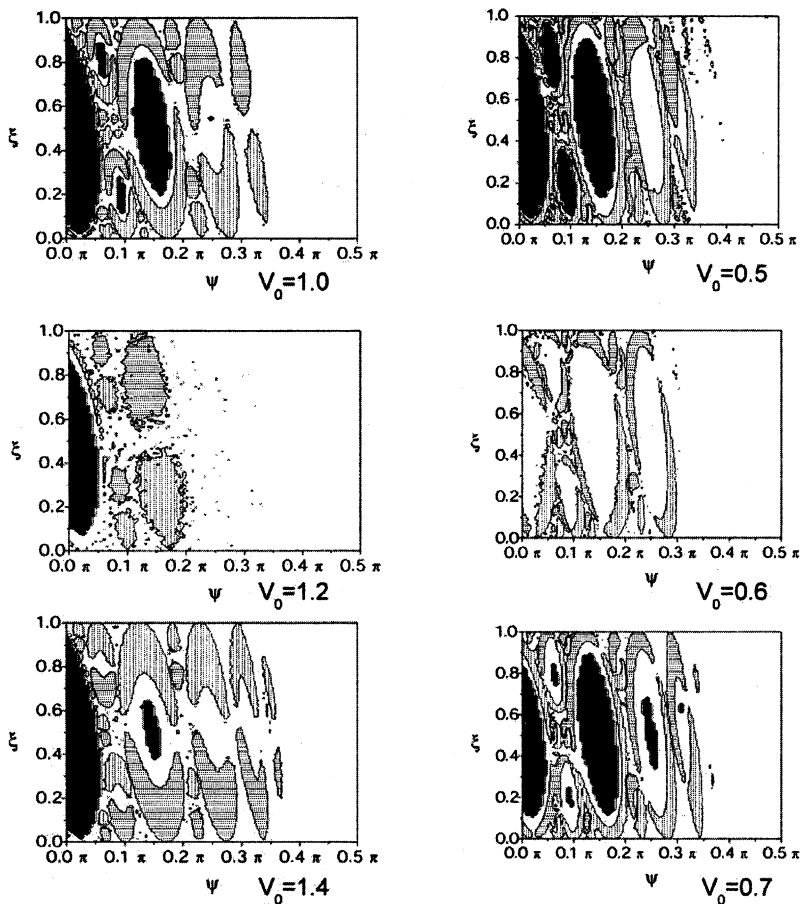


Рис. 14. Фазовые портреты изменения скорости в бильярде с возмущаемыми границами (29) — (31) при $b = 0,01$, $a = 0,5$, $l = 1$, $U_0 = 0,01$ и $\omega = 1$

Можно показать [20], что резонансная скорость в окрестности центральной устойчивой точки в два раза меньше, чем получается из уравнения (28), и будет равна $v_r^0 \approx 0,6$. Из рис. 14 следует, что при $V_0 = v_r^0$ все области фазового пространства становятся доступными частице. Кроме того, области ускорения и замедления частицы в окрестности центральной устойчивой точки меняются местами при переходе через v_r^0 .

Численные исследования отображения (30) — (31) были проведены в двух случаях: для полностью хаотического и близкого к интегрируемому бильярдов. В первом случае бильярд — классический стадион, т. е. $\Phi = \pi/2$, и граница бильярда состоит из двух полуокружностей и двух параллельных прямых, касательных к ним (см. левый рис. 4). Близкий к интегрируемому случай означает, что фокусирующие компоненты — почти прямые (см. правый рис. 4), так что выполняется соотношение (21). Для хаотического бильярда были выбраны следующие параметры: $a = 0,5$, $b = 0,25$, $l = 1$, $U_0 = 0,01$, $\omega = 1$ и $V_0 = 0,1$. Скорость частицы была вычислена как среднее значение ансамбля из 5000 траекторий с различными начальными условиями (сплошная кривая 1 на рис. 15). Начальные условия выбирались случайным образом на фокусирующей компоненте, так чтобы вектор скорости частицы был направлен внутрь бильярдной области. Как следует из численного анализа, полученная зависимость имеет приблизительно корневой характер ($V(n) \sim \sqrt{n}$).

В близком к интегрируемому бильярде параметр b достаточно мал и кривизна фокусирующей компоненты дает лишь слабую нелинейность в системе. В такой конфигурации фазовое пространство имеет области с регулярной и хаотической динамикой. Этот случай представляется значительно более интересным.

Невозмущенный бильярд имеет устойчивые точки, окруженные инвариантными кривыми. В их окрестности движение частицы квазипериодично и хорошо аппроксимируется отображением поворота. В то же время вне этих регионов динамика хаотична. В этом случае траектории заполняют хаотическую область целиком. Таким образом, если граница бильярда не возмущена, то в зависимости от начальных условий движение бильярдной частицы может быть либо хаотическим, либо регулярным.

Теперь обратимся к возмущаемому бильярду. В такой системе частица может двигаться из стохастической области в регулярную и обратно. В достаточно малой окрестности точки покоя поведение частицы имеет определенный период вращения $T_1 = 2\pi l / (\rho V)$, где ρ — число вращения, l — средняя длина свободного пробега и V — скорость частицы. В то же время период осцилляций границы $T_2 = 2\pi / \omega$. Таким образом, в системе при некоторой скорости V_r будет наблюдаться резонанс. Как следует из численного анализа, зависимости скорости частицы от числа столкновений по разные стороны резонанса отличаются. Если начальная скорость $V_0 < V_r$, тогда скорость частицы уменьшается до конечной величины $V_{fin} < V_r$ и распределение частиц по скоростям стремится к стационарному в

интервале $(0, V_{fin})$. Если же $V_0 > V_r$, то частицы могут достигать высоких скоростей. В этом случае распределение частиц нестационарно и неограниченно возрастает.

На рис. 15 показана зависимость скорости от числа столкновений (кривые 2 — 5, параметры бильярда те же, что и для бильярда «стадион» (кривая 1), за исключением $b = 0,01$).

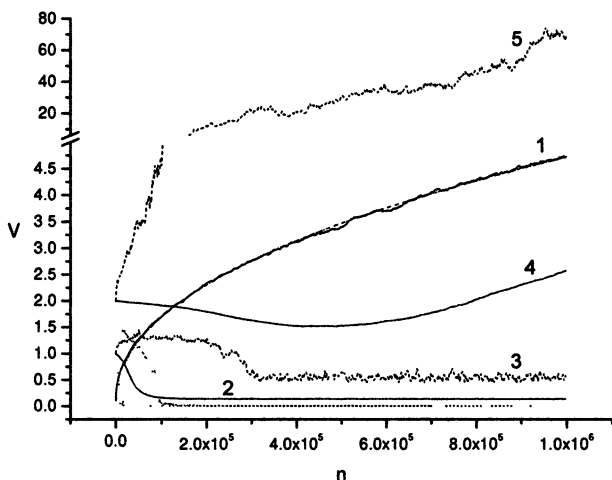


Рис. 15. Зависимость средней скорости частицы от числа итераций

На основании 5000 реализаций для каждой начальной скорости было построено три кривых: средняя, минимальная и максимальная скорости, достигнутые на n -ой итерации. Таким образом, можно определить интервал изменения скорости. Как следует из рисунка, если $V_0 < V_r$, тогда средняя скорость частиц (сплошная кривая 2) постепенно уменьшается и стремится к константе. Максимальная скорость частиц (пунктирная кривая 3) также уменьшается до величины V_{fin} и флуктуирует около этого значения. В конечном счете скорости частиц оказываются в интервале $0 < V < V_{fin}$. В случае $V > V_r$ минимальная скорость частиц также уменьшается. Это означает, что в ансамбле существуют частицы, попавшие в область малых скоростей. В численном анализе доля таких частиц составляла примерно 3/4. В то же время, существуют частицы с большой скоростью (штриховая кривая 5, которая соответствует максимальной скорости ансамбля). В результате средняя скорость (сплошная кривая 4) частиц будет расти.

4. Заключение

Бильярды являются достаточно удобными моделями целого ряда физических систем. Например, многим динамическим задачам могут быть поставлены в соответствие уравнения траектории частицы в бильярдах заданной формы. Более того, большинство подходов к проблеме перемешивания в системах из многих частиц сводятся к задачам бильярдного типа. Естественным обобщением бильярдных систем являются бильярды, границы которых не являются неподвижными, а изменяются по какому-либо закону. Это достаточно новая область, открывающая новые перспективы в исследовании многих давно известных, но малоизученных проблем.

В настоящей работе исследовалась проблема ускорения Ферми в бильярдах с возмущаемыми границами, относящихся к двум основным типам: рассеивающий бильярд на примере газа Лоренца и фокусирующий бильярд в форме стадиона.

В части, касающейся динамики рассеивающих бильярдов, можно выделить два основных механизма ускорения. Во-первых, это «сносовый» механизм, возникающий вследствие выполнения соотношения $\langle \Delta V \rangle > 0$ (см. (9)) и приводящий к сносу всех частиц в направлении положительных скоростей. Во-вторых, это дисперсионный (или флуктуационный) механизм, проявляющийся по двум причинам: а) $\langle \Delta V^2 \rangle > 0$, и поэтому вероятностный пакет «расплывается» со временем; б) модуль скорости не может быть отрицательным и, таким образом, расплывание пакета не симметрично, а направлено в сторону больших скоростей, что выражается в замене простого нормального распределения распределением (12). При этом, как показывают аналитические и численные исследования, флуктуации и среднее увеличение скорости частицы выше в случае *регулярных* колебаний границ рассеивателей, что приводит к большему росту скорости, таким образом, можно еще говорить о механизме, обусловленном корреляциями между последовательными изменениями скорости.

Нетрудно понять, что рассуждения, используемые при выводе зависимости скорости частицы от числа соударений и времени, легко перенести на любой другой бильярд, для которого известно вероятностное распределение угла α (угол между нормалью к поверхности в точке соударения и скоростью частицы). Поэтому развитая в данной работе техника может оказаться полезной в при изучении проблемы ускорения Ферми в общем случае.

В отличие от газа Лоренца в бильярде типа «стадион» динамика роста скорости зависит от его формы. Если форма такова, что

поведение частицы в нем полностью хаотично, то, как и в газе Лоренца, средняя скорость ансамбля частиц растет как корень из числа столкновений. В то же время, в близком к интегрируемому случаю наблюдается новое явление. В зависимости от начальной скорости ансамбль частиц может как ускоряться, так и замедляться, т. е. происходит разделение частиц по скоростям.

Существует критическое значение V_c такое, что если начальная скорость $V_0 < V_c$, то скорость ансамбля частиц убывает вплоть до некоторой близкой к нулю (но ненулевой) величины, а распределение частиц стремится к стационарному. При этом скорость всех частиц падает до достаточно малого значения (существенно меньшего начальной скорости) и устанавливается равновесное распределение частиц по скоростям. Данный эффект обусловлен тем, что в системе предпочтительными становятся траектории, соответствующие убыванию скорости. Это означает, что бильярдная частица испытывает в основном только сопутствующие столкновения с границей. Таким образом, изменения в динамике системы асимметричны по отношению к разгону или замедлению частиц, и такая система может быть использована для их охлаждения. Если $V_0 > V_c$, то скорость частиц может неограниченно возрастать.

С прикладной точки зрения такая модель представляется перспективной и реализуемой в экспериментах, на основе которых, по-видимому, можно создать устройства охлаждения частиц принципиально нового типа. Кроме того, данным способом возможно охлаждение частиц для ускорителей, где однородность частиц по скоростям очень важна.

Литература

1. *Синай Я. Г.* // Докл. АН СССР. 1963. Т. 153. С. 1261—1264.
2. *Синай Я. Г.* // Успехи математических наук. 1970. Т. 25. С. 141—192.
3. *Корнфельд И. П., Синай Я. Г., Фомин С. В.* Эргодическая теория. М.: Наука, 1980.
4. *Биркгоф Д.* Динамические системы. Ижевск: Изд. дом «Удмуртский университет», 1999.
5. *Крылов Н. С.* Работы по обоснованию статистической физики. М.; Л.: АН СССР, 1950.
6. *Бунимович Л. А.* // Динамические системы — 2. Сер. «Современные проблемы математики. Фундаментальные направления». ВИНТИ, 1985. С. 173—204.
7. *Tabachnikov A.* Billiards. Franse Mathematical Soc. Press, 1995.
8. *Заславский Г. М.* Стохастичность динамических систем. М.: Наука, 1984.
9. *Glans J.* // Science. 1994. V. 265. P. 474—475.
10. Fundamental Problems in Statistical Mechanics. V. 3 / Ed. R. H. Cohen. Amsterdam: Elsevier, 1975.

11. *Bunimovich L. A., Sinai Ya. G.* // Commun. Math. Phys. 1981. V. 78. P. 479—497.
12. *Baldwin P. R.* // J. Phys. A. 1991. V. 24. P. L941—L947.
13. *Chernov N.* // J. Stat. Phys. 1997. V. 88. P. 1—29.
14. *Garrido P. L.* // J. Stat. Phys. 1997. V. 88. P. 807—824.
15. *Бунимович Л. А.* // Математический сб. 1974. Т. 94. С. 49—73.
16. *Bunimovich L. A.* // Commun. Math. Phys. 1979. V. 65. P. 295—312.
17. *Bunimovich L. A.* // Chaos. 1991. V. 1. P. 187—193.
18. *Лоскутов А. Ю., Рябов А. Б., Акиншин Л. Г.* // ЖЭТФ. 1999. Т. 116. С. 1781—1797.
19. *Loskutov A., Ryabov A. B., Akinshin L. G.* // J. Phys. A. 2000. V. 33. P. 7973—7986.
20. *Loskutov A., Ryabov A.* // J. Stat. Phys. 2002. V. 108. P. 995—1014.
21. *Fermi E.* // Phys. Rev. 1949. V. 75. P. 1169—1174.
22. *Brahic A.* // Astron. Astrophys. 1971. V. 12. P. 98—110.
23. *Ulam S. M.* // Proc. 4th Berkeley Symp. Math. Stat. Probability. California Univ. Press, 1961. V. 3. P. 315.
24. *Лизттенберг А., Либерман М.* Регулярная и стохастическая динамика. М.: Мир, 1984.
25. *Заславский Г. М.* Стохастическая необратимость в нелинейных системах. М.: Наука, 1970.
26. *Lichtenberg A. J., Lieberman M. A., Cohen R. H.* // Physica D. 1980. V. 1. P. 291—305.
27. *Пустыльников Л. Д.* // ДАН СССР. 1987. Т. 292. С. 549—553.
28. *Пустыльников Л. Д.* // Математический сб. 1994. Т. 85. С. 113—124.
29. *Krüger T., Pustyl'nikov L. D., Troubetzkoy S. E.* // Nonlinearity. 1995. V. 8. P. 397—410.

CONTENTS

<i>Preface</i>	5
Nonlinear processes in physical systems	
<i>A. A. Abrashkin, E. I. Yakubovich.</i> Exact solutions in hydrodynamics.....	9
<i>A. B. Ezersky.</i> Spatio-temporal dynamics of solitons in resonators	23
<i>A. A. Kurkin, E. N. Pelinovsky, A. V. Slunyaev.</i> The physics of waves-killers in the ocean	37
<i>Yu. I. Troitskaya.</i> Nonlinear resonance wave–flow interaction in the ocean	52
<i>O. V. Rudenko.</i> Fundamentals of the physics of intense acoustic waves	70
<i>S. N. Gurbatov, A. I. Saichev.</i> Statistical nonlinear acoustics. Waves and structures in dispersion-free media.....	89
<i>L. A. Ostrovsky.</i> Non-classical non-linear acoustics	109
<i>A. I. Potapov.</i> Strain waves in a medium with internal structure.....	125
<i>Yu. N. Kulchin, R. V. Romashko, S. M. Shandarov, N. I. Burimov, Yu. F. Kargin, V. V. Volkov, O. T. Kamenev.</i> Application of the phenomenon of light wave self-diffraction in photorefractive crystals for processing of nonstationary speckle fields	141
<i>M. Yu. Ryabikin, A. M. Sergeev.</i> Engineering of electron wave packets in atoms and molecules for controlling attosecond pulse generation	149
<i>A. N. Stepanov, G. M. Fraiman.</i> Raman amplification of short laser pulses in plasma	165
<i>L. S. Mkhitarian.</i> Z-pinch and thermonuclear fusion	179
<i>V. Ya. Demikhovskiy.</i> Quantum chaos in applications	188
<i>Vi. V. Kocharovskiy, V. Yu. Martyanov.</i> Nonlinearity of vacuum and prospects for its observation.....	202
<i>D. I. Iudin, V. Yu. Trakhtenherts.</i> Dynamic percolation	217
<i>E. Yu. Bannikova, K. Yu. Bliokh, V. M. Kontorovich.</i> The evolution and collapse of a self-gravitating toroidal vortex.....	243

<i>V. I. Klyatskin.</i> Statistics and reality in stochastic physical systems	256
<i>D. I. Trubetskov, S. P. Kuznetsov, N. M. Ryskin, A. E. Hramov.</i> Complex dynamics of distributed nonlinear systems in super-high-frequency electronics	287
<i>A. M. Fridman, R. Z. Sagdeev, E. V. Polyachenko.</i> Observable manifestations of chaos in spiral galaxies	327

General problems of nonlinear dynamics

<i>V. B. Kazantsev, V. I. Nekorkin.</i> Phase-controlled oscillations in neurodynamics	345
<i>V. G. Yakhno, I. V. Nuidel, A. E. Ivanov.</i> Model neuron-like systems. Examples of dynamic processes	362
<i>L. V. Yakushevich.</i> Introduction to the nonlinear physics of DNA	376
<i>B. P. Bezruchko, M. B. Bodrov, T. V. Dikanev, A. S. Karavaev, V. I. Ponomarenko, M. D. Prokhorov, E. P. Seleznev, I. V. Sysoev, D. A. Smirnov.</i> Some problems of reconstruction of model equations from time series and methods of their solution	381
<i>N. G. Makarenko.</i> Reconstruction of dynamic systems by chaotic time series	398
<i>E. M. Loskutov, Ya. I. Molkov, D. N. Mukhin, A. M. Feigin.</i> Statistical approach to reconstruction of dynamic systems	411
<i>D. V. Turaev, A. L. Shilnikov, L. P. Shilnikov.</i> Some mathematical problems of classical synchronization	426
<i>S. A. Kashchenko.</i> Normal forms for singularly perturbed parabolic systems and systems with delay	451
<i>S. P. Kuznetsov, A. S. Pikovsky, U. Foidel.</i> Strange non-chaotic attractor	484
<i>A. Yu. Loskutov, A. B. Ryabov.</i> Billiards	510

NONLINEAR WAVES' 2004

Edited by A. V. Gaponov-Grekhov and V. I. Nekorkin

This book includes review lectures and papers based on the lectures given at the XII-th School in Nonlinear Waves (Nizhny Novgorod, 29 February–7 March, 2004). Spatio-temporal chaos, pattern formation, synchronization, dynamics of waves, nonlinear phenomena occurring in plasma and optics, as well as during the interaction of superstrong fields with substance, mathematical problems of nonlinear dynamics, and other problems concerned with nonlinear waves are considered.

The book is intended for specialists working in the field of nonlinear phenomena, as well as for graduate and postgraduate students of the corresponding specialities.

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2004

Сборник научных трудов

Утверждено к печати Институтом прикладной физики РАН,
603950 Нижний Новгород, ул. Ульянова, 46

Редактор *Л. О. Кузнецова*
Верстка *М. В. Башевой*
Корректор *Д. П. Семенова*

Подписано к печати 18.03.2005 г. Формат 60 × 90 ¹/₁₆. Бумага офсетная № 1.
Усл. печ. л. 34. Уч.-изд. л. 28,3. Темплан 2004 г. Поз. 6. Тираж 400 экз.

Отпечатано в полном соответствии с качеством
предоставленных диапозитивов в ГУП ППП «Типография "Наука"» АИЦ РАН
121099 Москва, Шубинский пер., 6

Заказ № 577