

Федеральное агентство научных организаций

Федеральный исследовательский центр

Институт прикладной физики РАН

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2016

Нижний Новгород
ИПФ РАН
2017

УДК 534.222.2

ББК 22.312я4

Н49

Издано по решению редакционно-издательского совета
ФИЦ «Институт прикладной физики РАН»

Ответственные редакторы:

академик РАН *А. М. Сергеев*,

доктор физико-математических наук *А. В. Слюняев*

Организационно-финансовая поддержка:

- Федеральное агентство научных организаций;
- Министерство образования и науки Российской Федерации (мегагрант № 14.Z50.31.0007 «Лабораторные и численные исследования плазменных явлений в экстремальных астрофизических объектах», ведущий ученый Фукс Жулиан Соуля; № 14.Z50.31.0033 «Новые подходы к исследованиям климатических процессов и прогнозы экстремальных явлений», ведущий ученый Куртц Юрген Герберт Густав);
- Российский научный фонд (проект № 15-12-30021);
- Российский фонд фундаментальных исследований (проект № 16-02-20129);
- Научно-производственное предприятие «ГИКОМ».

Н49 **Нелинейные волны' 2016** / Федер. агентство науч. орг., Федер. исслед. центр Ин-т приклад. физики РАН ; отв. ред. А. М. Сергеев, А. В. Слюняев. — Нижний Новгород : ИПФ РАН, 2017. — 320 с.

ISBN 978-5-8048-0119-0

В сборник включены статьи, написанные по материалам лекций, прочитанных на XVII школе по нелинейным волнам (Нижний Новгород, февраль — март 2016 г.). В работах обсуждаются разнообразные аспекты нелинейной науки — физика экстремальных световых полей, процессы в геофизике, в астрофизике и космологии, динамика вихрей в квантовых системах и гидродинамике, исследования нейросистем.

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

УДК 534.222.2

ББК 22.312я4

ISBN 978-5-8048-0119-0

© ИПФ РАН, 2017

СОДЕРЖАНИЕ

<i>Предисловие</i>	5
Физика экстремальных световых полей и мощных лазеров	
<i>Хазанов Е. А.</i> Обнаружение гравитационных волн. Вклад ИПФ РАН.....	9
<i>Яшин В. Е.</i> Нелинейные процессы в мощных фемтосекундных лазерных системах.....	26
<i>Башинов А. В., Ким А. В.</i> О нелинейной электродинамике плазмы в экстремально сильном световом поле	47
<i>Коржиманов А. В.</i> Лазерное ускорение ионных пучков как уникальный инструмент для физики высоких плотностей энергии.....	63
<i>Емелина А. С., Емелин М. Ю., Рябикин М. Ю.</i> Генерация высоких гармоник в газах лазерными источниками среднего ИК-диапазона: влияние магнитного поля импульса на форму спектра гармоник	69
<i>Гречин С. Г.</i> Критичные и некритичные процессы при нелинейно- оптическом преобразовании частоты лазерного излучения	80
<i>Попруженко С. В.</i> Метод мнимого времени в физике сильных лазерных полей	99
Нелинейные процессы в геофизике	
<i>Мохов И. И.</i> Атмосферные блокинги и связанные с ними климатические аномалии	115
<i>Елисеев А. В.</i> Линейные и нелинейные модели бароклинных волн в атмосфере	129
<i>Яковлев Н. Г.</i> Северный Ледовитый океан: особенности гидродинамики, динамики морского льда и проблемы численного моделирования крупномасштабного состояния	143
<i>Иудин Д. И., Давыденко С. С.</i> Фрактальная модель компактного внутриоблачного разряда.....	154
<i>Корнев А. А.</i> Численное решение задач стабилизации	172

Астрофизика и космология

<i>Блинников С. И.</i> Радиативные ударные волны и их роль в объяснении загадки сверхмощных сверхновых.....	193
<i>Деряшев Е. В.</i> Релятивистские ударные волны: структура фронта и излучение	208
<i>Попель С. И.</i> Нелинейные волновые явления в плазменно-пылевых системах	220

Нелинейные процессы в нейросистемах

<i>Храмов А. Е., Короновский А. А., ван Лютааалар Ж., Лютжоханн А., Максименко В. А., Макаров В. А.</i> Предсказание непредсказуемого: новые возможности подавления абсанс-эпилепсии с использованием интерфейса «мозг — компьютер».....	237
<i>Трубецков Д. И.</i> Некоторые исторические аспекты математического моделирования живых систем (докомпьютерная эра)	248

Современные проблемы теории нелинейных колебаний и волн

<i>Ванаг В. К.</i> Связанные химические осцилляторы	263
<i>Пелиновский Е. Н., Шургалова Е. Г., Слюняев А. В.</i> Солитонный газ: взаимодействия, турбулентность и волны-убийцы	279
<i>Мельников А. С.</i> Физика вихрей в сверхпроводниках как пример задачи нелинейной физики	294
<i>Агафонцев Д. С., Кузнецов Е. А., Майлыбаев А. А.</i> Опрокидывание вихревых линий как предтеча колмогоровской турбулентности.....	304

Предисловие

В настоящий сборник вошли обзорные и оригинальные статьи, написанные по материалам лекций, прочитанных на очередной (семнадцатой) научной школе «Нелинейные волны — 2016». Школа проводилась Институтом прикладной физики РАН 27 февраля — 4 марта 2016 года. Ее главной целью было сориентировать молодежь в проводимых в научных центрах нашей страны исследованиях, которые соответствуют наиболее интересным и перспективным направлениям развития мировой науки в области фундаментальной нелинейной физики и ее приложений.

Тематика школы представила несколько отобранных программным комитетом школы срезов исследований нелинейных процессов в физике, геофизике, математике, астрономии, биологии, теории информации. Основу программы составили 6 главных тематических направлений:

- современные проблемы теории нелинейных колебаний и волн;
- нелинейные процессы в геофизике;
- астрофизика и космология;
- физика экстремальных световых полей и мощных лазеров;
- нелинейные процессы в нейросистемах;
- квантовые системы и конденсированные среды.

Хотя в ряде случаев принадлежность лекции одному из направлений можно было назвать лишь условно.

Особенностью XVII школы стал выбор удлиненного формата лекций — полуторачасовых выступлений. Кроме выступлений 31 приглашенного лектора, в программу школы вошло 64 семинара более узкой тематики, проведенные как уже состоявшимися учеными, так и молодыми специалистами, а также более 100 стендовых докладов молодых ученых. В работе школы принимали участие иностранные ученые из Болгарии, Германии, Франции, Швеции и Японии.

Кроме трудов, представленных в настоящем сборнике, на школе были прочитаны следующие лекции: К. В. Анохин «Нейроны-хабы и кротовые норы в когнитивных сетях мозга», Л. Е. Голуб «Фотогальванические эффекты в полупроводниках и топологических изоляторах», Г. Т. Гурия «Нелинейные явления в процессах гидродинамической активации внутрисосудистого тромбообразования», А. А. Короновский «Нелинейные феномены в когнитивной деятельности человека, связанной с визуальным восприятием», В. И. Некоркин «Сложные динамические сети», Р. И. Нигматулин «Фокусировка энергии при кавитации. Термоядерные эффекты

в углеводородных жидкостях», К. А. Постнов «Обнаружение гравитационных волн», А. М. Пухов «Взаимодействие лазерного излучения с плотной плазмой: от наноструктур до КЭД-режима», В. А. Семенов «Аномальная погода в условиях меняющегося климата», А. А. Старобинский «Исследование истории нашей Вселенной до Большого взрыва», А. М. Фейгин «Прогностические эмпирические модели сложных систем: метод построения и приложение к исследованию климата», С. В. Янчук «Свойства уравнений с запаздыванием и примеры их применения в моделировании оптоэлектронных и нейронных систем»; а также Ю. Куртц «Климатические сети» (J. Kurths “Climate Networks”) и Ж. Муру «Пути к физике зептосекунд» (G. Mourou “Routes to Zeptosecond Physics”).

В феврале 2016 года — в канун проведения школы — было официально объявлено о первой в истории регистрации гравитационных волн, существование которых оставалось лишь гипотезой на протяжении 100 лет. Сотрудники исследовательской группы из ИПФ РАН и ученые из МГУ являются российскими соавторами этого открытия. В объединенной лекции профессор К. А. Постнов и член-корреспондент РАН Е. А. Хазанов дали основные понятия теории гравитационных волн, описали устройство детектора LIGO, рассказали о проблемах и идеях на пути достижения уникальной точности измерений.

Специальный лекционный блок был посвящен памяти профессора Н. Б. Нарожного. Своими воспоминаниями поделились и прочитали лекции Ж. Муру, С. В. Попруженко, А. М. Пухов, А. А. Старобинский.

В подготовке сборника принимали участие сотрудники ИПФ РАН М. Р. Волков, А. Г. Вяткин, И. Ю. Костюков, Вл. В. Кочаровский, В. В. Курин, Е. А. Мареев, В. И. Некоркин, Е. А. Перевезенцев, А. В. Старобор, А. М. Фейгин, Е. А. Хазанов, А. И. Яковлев, Е. И. Якубович.

Выражаем признательность всем лекторам школы и авторам за предоставленные для сборника материалы.

А. М. Сергеев, А. В. Слюняев

**Физика
экстремальных
световых полей
и мощных лазеров**

ОБНАРУЖЕНИЕ ГРАВИТАЦИОННЫХ ВОЛН. ВКЛАД ИПФ РАН

Е. А. Хазанов

Институт прикладной физики РАН, г. Нижний Новгород

Введение

Проект LIGO (Laser Interferometer for Gravitation Wave Observatory) стартовал 25 лет назад [1]. В США было построено два детектора, которые введены в строй в 2002 году в городах Ливингстон, штат Луизиана, и Ханфорд, штат Вашингтон, и к 2007 году достигли проектных параметров. Эта так называемая начальная версия LIGO работала до 2010 года, однако обнаружить сигнал гравитационной волны так и не удалось. Следующие несколько лет ушли на обновление детекторов, что включало в себя замену системы сейсмической изоляции, системы подвеса зеркал, лазера, входной оптики, фотодетектора и т. д. В 2015 году на запланированный уровень чувствительности (в несколько раз лучше, чем у начальной версии) была выведена так называемая промежуточная версия детекторов. В результате 14 сентября 2015 года был зарегистрирован сигнал гравитационной волны, пришедший от слияния двух черных дыр, которое произошло 1,3 миллиарда лет назад [2].

В основе детектора LIGO лежит интерферометр Майкельсона с длиной плеча 4 км. Согласно общей теории относительности Эйнштейна, при прохождении гравитационной волны через детектор расстояние между зеркалами изменится на чрезвычайно малую величину. Достигнутая чувствительность позволяет обнаружить изменение длины плеч меньше, чем на одну десятитысячную диаметра протона (10^{-19} м). Физические проблемы, которые были решены для достижения столь фантастической чувствительности, обсуждаются в первой части этой статьи.

Институт прикладной физики Российской академии наук (ИПФ РАН) вступил в проект LIGO в 1997 году и активно работает в ней до настоящего времени. За этот период ученые ИПФ РАН выполнили целый ряд научных исследований: создание широкоапертурного интерферометра белого света для входного контроля зеркал LIGO, создание интерферометра белого света для дистанционного контроля зеркал, контроль искажений волнового фронта с помощью эффекта самофокусировки, дистанционный контроль качества поверхности зеркал LIGO с помощью обращения волнового фронта, обнаружение загрязнений поверхности зеркал с помощью генерации второй гармоники при отражении от поверхности, активное управление тепловой линзой зеркал LIGO с помощью нагрева излучением вспомогательного CO₂-лазера, поиск новых лазерных сред для лазера детектора гравитационных волн следующего поколения и др. Сотрудники

института участвовали как в создании детекторов гравитационных волн, так и в проведении измерений.

Наиболее существенным вкладом ИПФ РАН в LIGO является изобретение уникальных оптических изоляторов (изоляторов Фарадея), работающих при большой мощности лазерного излучения. Изолятор Фарадея — один из ключевых оптических элементов, обеспечивающий оптическую развязку между лазером и интерферометром для детектирования гравитационных волн. Поглощение излучения в магнитоактивной среде изолятора неизбежно приводит к ее нагреву и термонаведенным поляризационным и фазовым искажениям лазерного пучка. Такое самовоздействие ограничивает использование изоляторов Фарадея в лазерах с высокой средней мощностью.

Сотрудники ИПФ РАН впервые поняли причину, из-за которой параметры стандартных изоляторов резко ухудшались при увеличении мощности излучения; предложили способы уменьшить эти паразитные эффекты; экспериментально продемонстрировали уникальные параметры созданного прототипа; провели его испытания в LIGO и, наконец, изготовили и установили изоляторы на оба детектора LIGO. Изоляторы успешно работают и по сей день. Во второй части статьи приведен анализ всех искажений лазерного пучка с точки зрения ухудшения параметров изолятора, определены механизмы и ключевые физические величины, отвечающие за различные виды искажений, описаны методы компенсации и подавления паразитных тепловых эффектов.

1. Лазерный интерферометр как детектор гравитационных волн

Идея использовать лазерный интерферометр для обнаружения (детектирования) гравитационных волн впервые была предложена М. Е. Герценштейном и В. И. Пустовойтом еще в 1962 году [3]. В этой пионерской работе была показана принципиальная возможность интерферометра Майкельсона детектировать гравитационные волны, а также предложено использовать для «запитки» интерферометра лазер — всего через два года после создания Т. Мейманом первого лазера [4]. Однако принципиальная возможность долгие годы оставалась лишь принципиальной, поскольку существующие в то время технологии не позволяли даже близко подойти к необходимой чувствительности детектора — относительное изменение длины 10^{-23} . При длине плеча интерферометра 4 км абсолютное изменение длины, которое необходимо измерить, составляет порядка 10^{-19} м. Уровень шумов, прежде всего сейсмических и оптических, превышал эту цифру на много порядков.

Перевести разговор в практическую плоскость позволили две ключевые идеи.

Во-первых, был найден способ на несколько порядков повысить сигнал, увеличив эффективную длину плеч интерферометра Майкельсона, сохранив их физическую длину равной 4 км. Идея заключается в том, чтобы заставить свет многократно проходить путь от светоделителя до зеркала интерферометра и лишь после этого возвращаться на светоделитель и далее на фотоприемник. Использование многократных переотражений между дополнительными зеркалами, как это часто делается в оптических линиях задержки, наткнулось на ряд непреодолимых трудностей. Однако на помощь пришла идея использовать в каждом из плеч интерферометра Майкельсона высокодобротный интерферометр Фабри — Перо с длиной 4 км (рис. 1). Этот интерферометр, будучи настроенным в резонанс с длиной волны лазера, накапливает в себе оптическую мощность примерно на два порядка большую, чем мощность, попадающая в плечо интерферометра Майкельсона. В результате фотон, попавший внутрь интерферометра Фабри — Перо, прежде чем вернуться обратно к светоделителю, примерно сто раз пробегает между зеркалами интерферометра Фабри — Перо, т. е. проходит путь не 8 км (4 км туда и 4 км обратно), а 800 км. Таким образом, сигнал возрастает на два порядка. Этот же прием используется еще раз с помощью так называемого рециркуляционного зеркала (power recycling mirror) с коэффициентом отражения около 97 %. Дело в том, что фотоприемник располагается в темном порту интерферометра Майкельсона, т. е. вся лазерная мощность (за вычетом технических потерь) направляется обратно в сторону лазера. Другими словами, весь интерферометр Майкельсона работает как одно стопроцентное зеркало и в сочетании с рециркуляционным

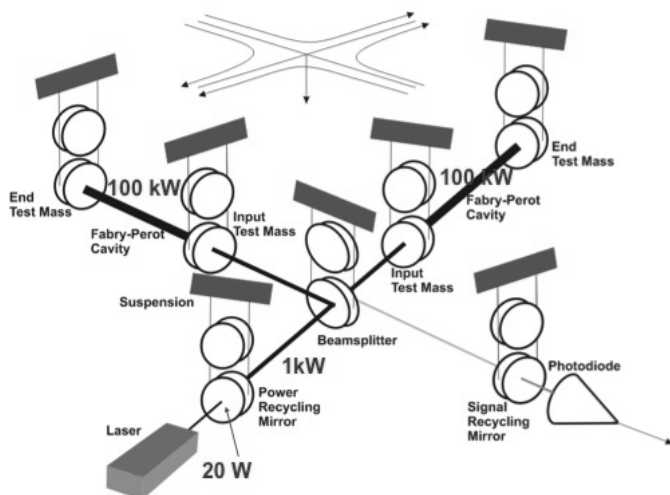


Рис. 1. Схема интерферометра детектора гравитационных волн LIGO

зеркалом образует еще один высокодобротный интерферометр Фабри — Перо, в котором запасается мощность примерно в 50 раз большая, чем мощность, падающая на интерферометр со стороны лазера. Таким образом, сигнал возрастает еще почти на два порядка. И наконец, аналогичную работу выполняет рециркуляционное зеркало (signal recycling mirror), расположенное перед фотоприемником. Заметим, что расплатой за это увеличение сигнала является не только многократно возросшая сложность юстировки и управления детектором, но и соответствующее увеличение мощности излучения и связанные с этим паразитные тепловые эффекты, рассказу о борьбе с которыми можно посвятить не одну статью.

Во-вторых, радикально понизить сейсмические шумы позволила идея использовать не жестко закрепленные, а свободно подвешенные на тонких нитях зеркала. Такая конструкция обеспечивает низкие резонансные частоты (около 1 Гц) механических колебаний зеркал и одновременно высокую добротность этих колебаний (время затухания составляет несколько лет). Столь узкий резонанс делает систему чрезвычайно мало восприимчивой к воздействию внешних сил на частотах в диапазоне предполагаемого сигнала гравитационных волн (100—1000 Гц), поскольку эти частоты очень далеки от резонанса. Таким образом, данная конструкция, допускающая гигантские (на десять порядков больше) сейсмические шумы на частотах несколько герц, обеспечивает в нужном диапазоне частот шумы ниже уровня 10^{-23} . Такая парадигма — не демпфирование колебаний на основных модах, а, наоборот, максимальное увеличение их добротности и плюс (по возможности) вынесение резонансов как можно дальше от рабочего диапазона детектора 100—1000 Гц — широко используется в LIGO для подавления и других видов шума.

Сейсмические шумы далеко не единственный вид шумов. Можно смело сказать, что вся история LIGO — это история борьбы с шумами. На рисунке 2 основные виды шумов показаны схематично. Чрезвычайно важны тепловые шумы самих зеркал и нитей подвеса. Эти шумы являются фундаментальными и не равны нулю при ненулевой температуре (LIGO работает при комнатной температуре). Фактически речь идет о тепловых флуктуациях толщины зеркал и длины нитей подвеса. Для борьбы с этими шумами толщина нитей подвеса выбрана минимальной — только чтобы нити выдержали вес зеркал. Кроме того, используемые изначально стальные нити были заменены на менее «шумные» кварцевые волокна. С целью увеличения добротности колебаний формы зеркал их боковая поверхность отполирована с оптическим качеством. Для работы интерферометра, разумеется, качество боковой поверхности вообще не важно.

Чрезвычайно важна вакуумная система LIGO, обеспечивающая остаточное давление 10^{-9} Торр 24 часа в сутки в огромном объеме (диаметр 1,24 м, длина 8 км). Такой глубокий вакуум необходим для подавления шумов, связанных с отличием от единицы показателя преломления газа.

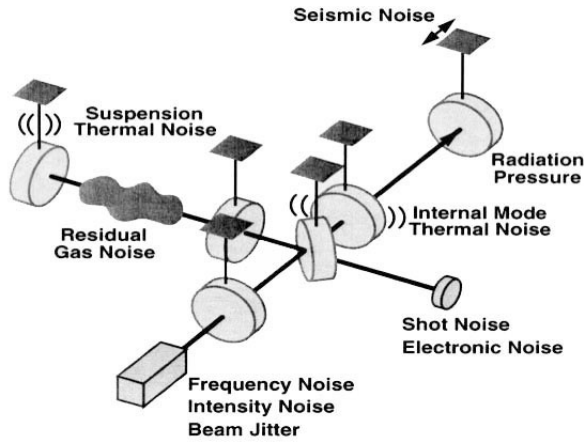


Рис. 2. Основные шумы детектора гравитационных волн

До начала строительства 4-километровых интерферометров LIGO в Калифорнийском технологическом институте был построен 40-метровый прототип, на котором отрабатывались все идеи. Для работы прототипа вполне достаточно было вакуумной трубы диаметром несколько сантиметров, однако была изготовлена вакуумная труба диаметром 1,24 м, чтобы доказать, что технически возможно создать и круглосуточно поддерживать столь глубокий вакуум в таких больших объемах. В интерферометрах LIGO через каждые 20 м расположены круглосуточно работающие вакуумные насосы.

Удивительно, что даже давление света на зеркала приводит к увеличению шума интерферометра, несмотря на то что масса зеркал 40 кг. Более традиционны для оптических измерений дробовой шум фотоприемника, шумы электроники и, разумеется, шумы, связанные с флуктуациями длины волны, мощности и направления волнового вектора лазера. Уникален чрезвычайно низкий уровень этих шумов, чего удалось добиться ценой кропотливой многолетней работы большого количества ученых из разных стран. Для примера укажем, что относительное изменение мощности лазера составляет порядка 10^{-9} .

В заключение этого раздела ответим на два важных концептуальных вопроса. Первый: представляет ли физический смысл измерение расстояния с точностью 10^{-19} м между двумя зеркалами, поверхности которых имеют шероховатость порядка 10^{-10} м? Ответ на вопрос, поставленный в таком провокационном виде, — нет, не имеет. Однако интерферометр измеряет не расстояние, а *изменение* расстояния во времени. Это в корне меняет дело. Для измерения изменения расстояния между двумя поверхностями характерные масштабы их формы не столь важны, лишь бы сама

форма не менялась. Второй вопрос касается непосредственно способа измерения. Реальная фаза волны, отраженной от зеркала, меняется в пределах поперечного сечения на 10^{-10} м. Интерферометр измеряет разность фаз двух оптических волн, пришедших на светоделитель. Эта разность фаз (как и фазы обеих волн) одинакова для всего поперечного сечения пучка. Как можно, измерив изменения этой средней фазы, говорить об измерении расстояния до отражающей поверхности с точностью 10^{-19} м? Ответ можно получить, воспользовавшись принципом суперпозиции. Отраженную от неидеальной поверхности волну, фаза которой действительно флуктуирует в поперечном сечении, можно разложить на идеальный гауссов пучок, отраженный от идеальной поверхности, плюс остальное поле в виде суммы высоких пространственных мод, каждая из которых имеет свою амплитуду и фазу. Однако при распространении от зеркала до фотоприемника высокие пространственные моды из-за их большой расходимости просто не попадают на фотодетектор: они фильтруются как самим интерферометром Майкельсона, так и специальной «чистящей» оптикой, расположенной перед фотодетектором (на рис. 1 не показана). Таким образом, фотодетектор «видит» только волну, отраженную от виртуальной идеально гладкой поверхности, а ее неидеальность приводит лишь к (незначительным) потерям мощности. В результате интерферометр измеряет изменение расстояния между двумя виртуальными идеальными плоскостями, однако эти плоскости жестко привязаны к реальным шероховатым поверхностям двух зеркал.

2. Изоляторы Фарадея для лазеров с большой средней мощностью

2.1. Тепловые эффекты в изоляторах Фарадея

С момента создания первого лазера и по настоящее время одним из важнейших направлений развития физики лазеров является увеличение средней мощности излучения. Уже в 1960-х годах встала проблема паразитных тепловых эффектов в активных элементах лазеров, особенно твердотельных. При ламповой накачке мощность тепловыделения может превышать (иногда существенно) мощность излучения лазера. Это приводит к четырем паразитным эффектам: увеличению средней по объему температуры, механическому разрушению из-за термонаведенных напряжений, тепловой линзе и двулучепреломлению. Как правило, разрушение происходит при мощности тепловыделения существенно больше той, при которой появляются сильные фазовые и поляризационные искажения: тепловая линза и термонаведенная деполяризация. Тепловая линза вызывается зависимостью длины от температуры (линейное расширение), зависимостью показателя преломления от температуры (dn/dT -эффект) и зависимо-

стью показателя преломления от деформаций (фотоупругий эффект). Деполяризация происходит из-за двулучепреломления, вызванного фотоупругим эффектом, в результате которого в исходно изотропной среде появляются термонаведенные собственные поляризации. Их направление и разность фаз (величина двулучепреломления) зависят от поперечных координат. В результате излучение после прохождения образца становится деполаризованным.

Под деполаризованным излучением здесь мы понимаем излучение, у которого поляризация постоянна во времени, но изменяется от точки к точке поперечного сечения. Деполаризованное излучение описывается комплексной амплитудой электрического поля, которая является комплексной векторной функцией поперечных координат. Квадрат модуля комплексной амплитуды определяет интенсивность излучения так же, как и для поляризованного излучения. Однако деполаризованное излучение нельзя характеризовать фазой (и, соответственно, волновым фронтом), так как для комплексного вектора нельзя ввести понятие фазы.

Первые работы по термонаведенной деполаризации начались еще в 1960-е годы [5]. Интенсивное исследование тепловых эффектов в активных элементах лазеров в 1970-х годах привело к появлению нового направления физики лазеров — термооптики твердотельных лазеров [6]. В эти же годы были предложены наиболее популярные до сих пор оптические схемы компенсации деполаризации: два идентичных активных элемента и 90-градусный вращатель поляризации, расположенный между ними [7], а также активный элемент, 45-градусный вращатель Фарадея и зеркало [8].

С увеличением средней мощности лазеров актуальными стали самонаведенные тепловые эффекты, вызванные поглощением не накачки, а собственно лазерного излучения. Коэффициент поглощения α в так называемых прозрачных диэлектриках составляет величину от 10^{-6} до 10^{-2} см $^{-1}$. Длина оптических элементов L меняется от миллиметров до сантиметров, т. е. мощность тепловыделения может составлять от ничтожных долей до единиц процентов от мощности лазерного излучения. Пик исследований самонаведенных тепловых эффектов пришелся на 1980-е и 1990-е годы. Были исследованы поляризационные и фазовые искажения, в том числе с учетом формы пучка и условий охлаждения. Изучалось влияние поглощения на выходные окна лазеров, зеркала, электрооптические модуляторы, а также удвоители частоты и другие нелинейно-оптические элементы.

Особняком в этом ряду стоят изоляторы Фарадея (ИФ) [9], которые являются одним из ключевых элементов мощного лазера, обеспечивая оптическую развязку либо между отдельными частями лазера (как правило, между задающим генератором и мощным усилителем), либо, как в LIGO, между собственно лазером и объектом, на который направляется лазерный пучок, — рециркуляционным зеркалом (power recycling mirror) (на рис. 1 ИФ не показан).

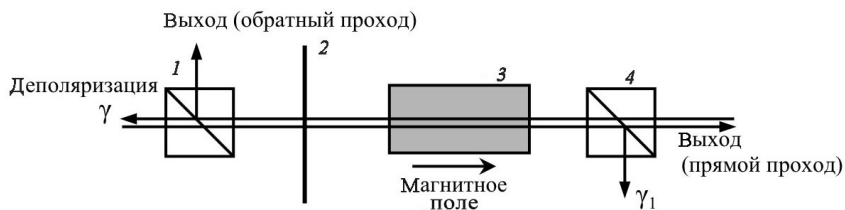


Рис. 3. Традиционный изолятор Фарадея: 1, 4 — поляризаторы; 2 — пластинка $\lambda/2$; 3 — фарадеевский элемент (угол поворота 45°)

На рисунке 3 представлена схема ИФ, в основе которого находящаяся в магнитном поле B магнитоактивная среда — фарадеевский элемент (ФЭ). Эффект Фарадея [10] приводит к возникновению в ФЭ циркулярного двулучепреломления, в результате которого плоскость поляризации лазерного излучения поворачивается на угол

$$\Phi = V \int_0^L B(z) dz, \quad (1)$$

где V — постоянная Верде, L — длина ФЭ. Из-за невязки эффекта Фарадея при $\Phi = 45^\circ$ после прямого прохода (слева направо) пучок сохраняет горизонтальную (в плоскости рисунка) поляризацию и проходит через поляризатор 4, а на обратном проходе меняет ее на вертикальную и в идеальном случае полностью отражается поляризатором 1. Деполяризованное излучение проходит через поляризатор 1, и его долю

$$\gamma = \frac{P_{dep}}{P_0} \quad (2)$$

будем называть невязкой ИФ. Здесь P_0 и P_{dep} — полная мощность и мощность излучения, прошедшего через поляризатор 1. Невязка, выраженная в децибелах, $I = -10 \log \gamma$, называется степенью изоляции и является основной характеристикой ИФ. Обычно достаточной считается $I = 30\text{—}40$ дБ. Вклад в невязку γ вносят поперечная неоднородность магнитного поля, неточность юстировки и контраст поляризаторов, не связанное с нагревом «холодное» двулучепреломление, а также тепловые эффекты.

Поглощение излучения в ФЭ, кроме увеличения средней по объему температуры, приводит к неоднородному по поперечному сечению распределению температуры, что дает три физических механизма воздействия на лазерное излучение:

1) искажение волнового фронта (тепловая линза), вызванное зависимостью изотропной части показателя преломления от температуры и механических напряжений, а также тепловым расширением;

2) неоднородное распределение угла поворота плоскости поляризации Φ , вызванное зависимостью постоянной Верде от температуры, а также тепловым расширением;

3) появление одновременно с циркулярным (эффект Фарадея) линейного двулучепреломления вследствие механических напряжений из-за градиента температуры (фотоупругий эффект).

Первый механизм не вызывает никаких поляризационных изменений в лазерном излучении и, следовательно, не влияет на неразвязку и степень изоляции. Последние два механизма, напротив, уменьшают степень изоляции. Температурная зависимость постоянной Верде и тепловое расширение приводят к изменению разности фаз между собственными поляризациями, которые при этом остаются циркулярными. Фотоупругий эффект влечет за собой не только изменение разности фаз между собственными поляризациями, но и изменение самих собственных поляризаций, которые становятся в этом случае эллиптическими.

С точки зрения деполяризации излучения ИФ имеют несколько важных отличий от всех ранее упомянутых оптических элементов. Во-первых, термонаведенное двулучепреломление появляется на фоне циркулярного двулучепреломления. Во-вторых, требования, предъявляемые к деполяризации в ИФ, значительно жестче: например, деполяризация $\gamma = 1\%$, вполне допустимая в большинстве других элементов, в ИФ приводит к степени изоляции 20 дБ, что, как правило, неудовлетворительно. В-третьих, относительно высокое поглощение в магнитоактивных средах $\alpha \approx (1 \div 3) \cdot 10^{-3} \text{ см}^{-1}$ и большая длина ФЭ $L \approx (1 \div 3) \text{ см}$ приводят к большой мощности тепловыделения — не менее 0,1 % от мощности лазерного излучения.

Тепловая линза в ИФ впервые наблюдалась в работах [11, 12]. Первые исследования деполяризации в ИФ были инициированы именно проектом LIGO в 1998 году [13], что связано не только с большой мощностью излучения, но и с высокими требованиями ко всем характеристикам ИФ. В течение нескольких лет были проведены первые систематические исследования как самих тепловых эффектов, так и методов их компенсации и подавления [14—24]. Эти результаты легли в основу термооптики магнитоактивных сред, которая бурно развивалась в 2000-е годы и развивается в настоящее время [25].

Термонаведенные искажения в ИФ характеризуются следующими параметрами. Во-первых и в главных, неразвязкой γ — величиной обратной степени изоляции ИФ, которая состоит из двух слагаемых: γ_p и γ_v , обусловленных фотоупругим эффектом и зависимостью $V(T)$ соответственно. Во-вторых, поляризационными потерями мощности на прямом проходе через ИФ γ_1 (слева направо на рис. 3). В-третьих, амплитудно-фазовыми искажениями пучка на прямом проходе, которые состоят из анизотропных потерь γ_a , обусловленных фотоупругим эффектом, и изотропных потерь

из-за зависимости показателя преломления от температуры и изотропной части фотоупругого эффекта. В зависимости от приложений изотропные искажения удобно характеризовать уменьшением интеграла перекрытия γ_i , уменьшением числа Штреля γ_s или увеличением параметра M^2 γ_M . Выражения для всех γ приведены в работе [25]. Заметим, что γ_1 , γ_a и γ_i очень важны именно для детекторов гравитационных волн.

Поглощение излучения в магнитоактивной среде имеет три механизма воздействия на лазерное излучение: деполяризацию из-за фотоупругого эффекта, деполяризацию вследствие зависимости постоянной Верде от температуры $V(T)$ и тепловую линзу. Эти три эффекта полностью определяются тремя параметрами: p , p_v и p_i , имеющими смысл безразмерной мощности лазерного излучения P_0 , а также параметром оптической анизотропии ξ :

$$p = \frac{L}{\lambda} \frac{\alpha Q}{\kappa} P_0, \quad p_v = \frac{\alpha}{8\kappa} \frac{1}{V} \frac{dV}{dT} P_0, \quad p_i = \frac{L}{\lambda} \frac{\alpha P}{\kappa} P_0, \quad (3)$$

$$\xi = \frac{\pi_{44}}{\pi_{11} - \pi_{12}}, \quad (4)$$

где Q , P — термооптические постоянные [6], κ — теплопроводность, π_{ij} — элементы пьезооптического тензора.

Параметры p , p_v и p_i определяются длиной волны λ , длиной фарадеевского элемента L , которая при заданной величине магнитного поля B обратно пропорциональна V (1), и характеристиками магнитоактивной среды. Таким образом, при заданных B и λ тепловые эффекты определяются только материальными константами среды. Следовательно, для сравнения различных сред, а также для сравнения различных тепловых эффектов между собой удобно ввести [25] три критические мощности: $P_{кр}$, $P_{кр,v}$ и $P_{кр,i}$, т. е. мощности лазерного излучения, при которых параметры традиционного ИФ достигают определенных критических значений, а именно: степень изоляции 30 дБ и изотропные потери 10 % после компенсации параболической фазы. Чем больше критическая мощность, тем лучше среда. Как правило, $P_{кр} \ll P_{кр,v}$. Это означает, что $\gamma_v \ll \gamma_p$ и температурной зависимостью постоянной Верде можно пренебречь. Исключение составляют криогенные ИФ [26] и кристалл TSAG [27] с ориентацией [001].

2.2. Компенсация деполяризации

Под компенсацией самонаведенных тепловых искажений мы понимаем их уменьшение с помощью ИФ с более сложными оптическими схемами, чем приведенная на рис. 3. Основная идея заключается в использовании двух оптических элементов, искажения пучка в которых компенсируют друг друга — накапливаются в одном и вычитаются в другом. Хорошо известно, что 90-градусный вращатель поляризации, расположенный меж-

ду двумя любыми одинаковыми фазовыми пластинками с линейными собственными поляризациями, приводит к полной компенсации деполаризации [7]. При наличии в пластинках также и циркулярного двулучепреломления это справедливо, только если оно имеет в фазовых пластинках разный знак. Для создания ИФ такое неприемлемо, так как при этом пропадают не взаимные свойства изолятора. В работе [14] предложено заменить один 45° -й ФЭ на два $22,5^\circ$ -х и пластинку $\lambda/2$ (рис. 4, *a*) или $67,5^\circ$ -й взаимный вращатель (ВВ) поляризации (рис. 4, *б*) между ними. Будем называть эти новые изоляторы ИФ с $\lambda/2$ и ИФ с ВВ. Последние используются в детекторах гравитационных волн LIGO и VIRGO [28—30].

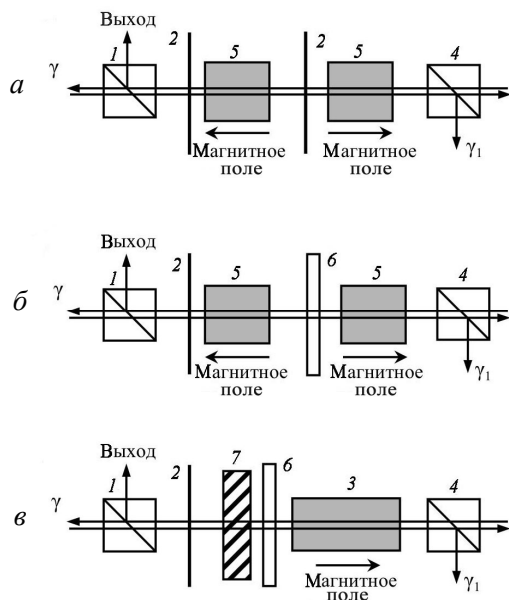


Рис. 4. ИФ с пластинкой $\lambda/2$ (*a*), ИФ с ВВ (*б*), ИФ с ПЭ (*в*): 1, 4 — поляризаторы; 2 — пластинки $\lambda/2$; 3 — ФЭ (45°); 5 — ФЭ ($22,5^\circ$); 6 — взаимный вращатель поляризации ($67,5^\circ$); 7 — поглощающий оптический элемент

Интерес представляет также возможность использовать для компенсации деполаризации оптический элемент, расположенный вне магнитной системы традиционного ИФ. Идея заключается в создании фазовой пластинки, в которой бы вычитались все фазовые набеги, приобретенные пучком в ФЭ. Для этого фазовая пластинка должна иметь такое же, как и ФЭ, поперечное распределение собственных поляризаций и такое же по амплитуде, но противоположное по знаку поперечное распределение разности фаз. В работе [31] предложен изображенный на рис. 4, *в* изолятор Фарадея,

где роль такой пластинки играет $67,5^\circ$ -й вращатель поляризации и поглощающий излучение оптический элемент (ПЭ), в котором возникает термонаведенное двулучепреломление. Такой изолятор будем называть ИФ с ПЭ.

Эффективность компенсации деполяризации в ИФ с $\lambda/2$ и в ИФ с ВВ впервые проверена экспериментально для ФЭ из стекла в работе [16], где было подтверждено предсказание теории о том, что именно фотоупругий эффект определяет степень изоляции при большой средней мощности излучения. Для лазеров с высокой средней мощностью наиболее актуальными являются не магнитоактивные стекла, а кристалл TGG, для которого также в 2000 году было получено экспериментальное подтверждение эффективности ИФ с ВВ [19]. ИФ с ПЭ был впервые реализован в работе [31], где при мощности 330 Вт продемонстрировано уменьшение γ в 36 раз по сравнению с традиционным ИФ.

ИФ с ВВ и ИФ с ПЭ были исследованы во многих работах при мощности лазерного излучения вплоть до 1,5 кВт. На рисунке 5 приведены результаты экспериментов с мощностью более 1 кВт [32]. При малой

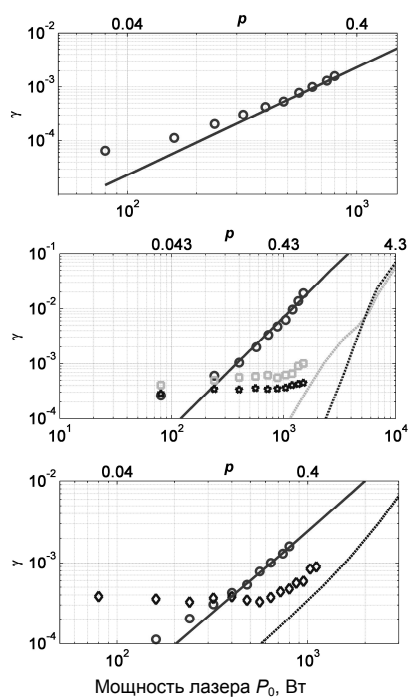


Рис. 5. Зависимость неразвязки от мощности излучения P_0 [32]: традиционного ИФ γ_p (кружки), ИФ с ВВ γ_R (квадраты — ВВ $67,5^\circ$, звездочки — ВВ 73°) и ИФ с ПЭ γ_A (ромбы)

мощности γ определяется «холодной» деполяризацией ФЭ, контрастом используемых поляризаторов и неоднородностью магнитного поля, а при увеличении мощности γ растет, стремясь к теоретическим зависимостям, изображенным сплошными линиями.

Выше мы пренебрегали увеличением средней температуры ФЭ, так как единственное негативное следствие этого — уменьшение угла поворота поляризации из-за зависимости $V(T)$. Чтобы уменьшить нагрев ФЭ, необходим хороший тепловой контакт, а при очень большой мощности требуется активное охлаждение водой или элементом Пельтье. Следует заметить, что в детекторах гравитационных волн ИФ расположены в вакууме и проблема увеличения средней температуры стоит весьма остро [28—30].

2.3. Компенсация тепловой линзы

Распределение температуры в оптическом элементе, а следовательно, и фазы аберрированного лазерного пучка близко к параболическому. Значительную часть фазовых искажений можно компенсировать при помощи обычной (параболической) линзы или телескопа, который вносит в волновой фронт дополнительную кривизну [18]. Этот способ компенсации называют телескопным. Альтернативный способ — адаптивный — заключается в следующем. Перед поляризатором I (см. рис. 3) располагается поглощающий элемент, параметры которого выбираются таким образом, чтобы тепловая линза имела ту же амплитуду и форму, что и в ИФ, но была при этом отрицательной (у большинства магнитоактивных материалов она положительна). Адаптивный способ имеет два несомненных и чрезвычайно важных для детекторов гравитационных волн преимущества: нет необходимости юстировки при изменении мощности лазера и более высокая точность компенсации. Благодаря этим преимуществам именно этот способ был применен в интерферометре LIGO [29, 30]. В качестве поглощающего элемента использовался кристалл DKDP (дейтерированный дигидрофосфат калия). Двумерные распределения фазы, полученные в работе [33] с использованием DKDP толщиной 5,5 мм, показаны на рис. 6.

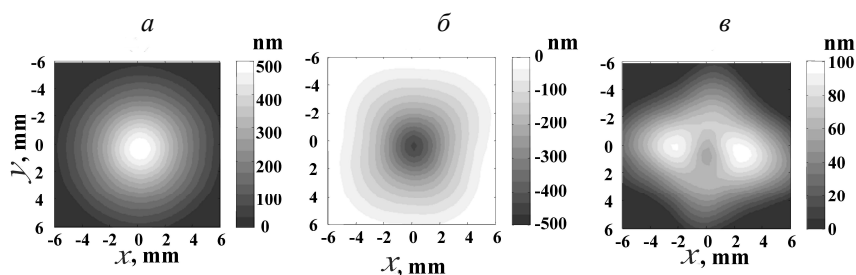


Рис. 6. Измеренные пространственные распределения набега фазы пучка после ИФ (а), DKDP (б), ИФ и DKDP (в) [33]

Экспериментально продемонстрировано, что при мощности излучения 45 Вт термонаведенные потери в гауссовом пучке γ_i уменьшились с 26 до 0,5 %.

2.4. Подавление тепловых эффектов в изоляторах Фарадея

Под подавлением искажений мы понимаем их уменьшение при сохранении схемы ИФ, изображенной на рис. 3. Наиболее радикальным методом представляется охлаждение ИФ до температуры кипения жидкого азота, что в разы улучшает многие ключевые характеристики кристалла тербий-галлиевого граната — до недавнего времени единственного кристалла, используемого в ИФ для лазеров с большой средней мощностью на длине волны $\lambda = 1$ мкм. В последние годы появилось много реальных альтернатив ему. Подавляют тепловые эффекты также увеличение магнитного поля и использование нестержневой геометрии отвода тепла от ФЭ. Представленные методы компенсации и подавления тепловых эффектов в ИФ могут применяться в различных сочетаниях друг с другом. Подробный обзор методов подавления тепловых искажений в ИФ можно найти в обзоре [25].

Заключение

Проект LIGO послужил толчком к развитию целого ряда технологий, используемых в настоящее время для решения научных и технических задач, никак не связанных с гравитационными волнами. Это касается лазерных, оптических, вакуумных, тепловых, механических и компьютерных технологий. Один из примеров — ИФ для лазеров с высокой средней мощностью (рис. 7), параметры которых в настоящее время обогнали требования к изоляторам Фарадея для детекторов гравитационных волн, изображенным на рис. 8.

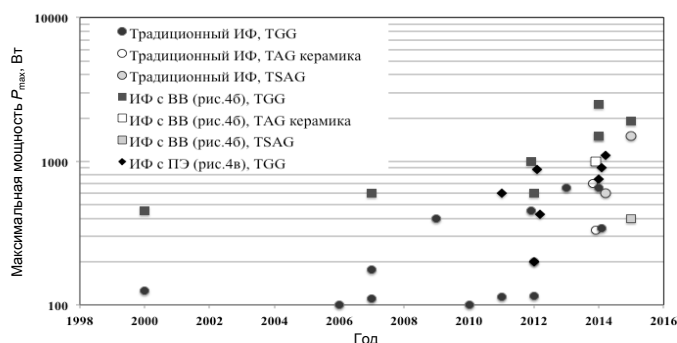


Рис. 7. Критическая мощность ИФ (мощность, при которой степень изоляции 30 дБ) [25]

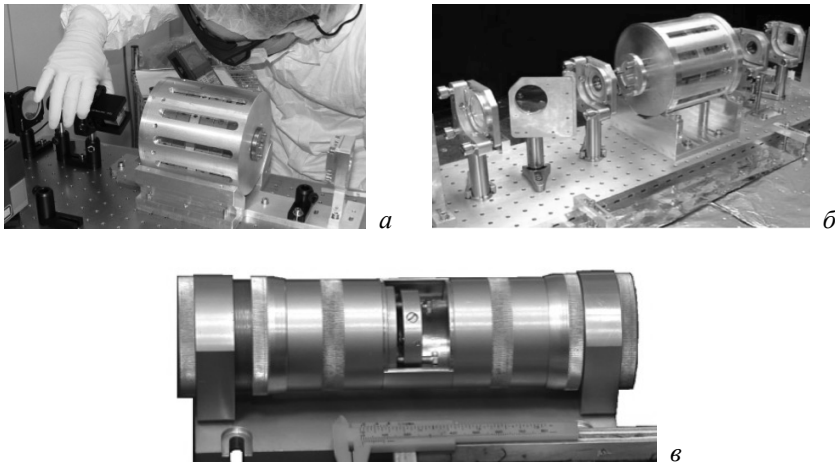


Рис. 8. Изоляторы Фарадея для детекторов гравитационных волн LIGO (а), VIRGO (б), GEO (в)

Литература

1. *Abramovici, A.* LIGO – the Laser-Interferometer-Gravitational-Wave Observatory / A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, M. E. Zucker // *Science*. — 1992. — V. 256. — P. 325—333.
2. *Abbott, B. P.* Observation of Gravitational Waves from a Binary Black Hole Merger / B. P. Abbott [et al.] // *Phys. Rev. Lett.* — 2016. — V. 116, № 6. — P. 061102.
3. *Герценштейн, М. Е.* К вопросу об обнаружении гравитационных волн малых частот / М. Е. Герценштейн, В. И. Пустовойт // *ЖЭТФ*. — 1962. — Т. 42, вып. 2. — С. 605—607.
4. *Maiman, T.* Stimulated optical radiation in ruby // *Nature*. — 1960. — V. 187. — P. 493.
5. *Ананьев, Ю. А.* Термическая деформация резонатора твердотельного ОКГ / Ю. А. Ананьев, Н. А. Козлов, А. А. Мак, А. И. Степанов // *Ж-л приклад. спектроскопии*. — 1966. — Т. 5, вып. 1. — С. 51—55.
6. *Мезенов, А. В.* Термооптика твердотельных лазеров / А. В. Мезенов, Л. Н. Сомс, А. И. Степанов. — Л.: Машиностроение, 1986. — 199 с.
7. *Scott, W. C.* Birefringence compensation and TEM₀₀ mode enhancement in a Nd:YAG laser / W. C. Scott, M. de Wit // *Appl. Phys. Lett.* — 1971. — V. 18, № 1. — P. 3—4.
8. *Giuliani, G.* Polarization flip cavities: a new approach to laser resonators / G. Giuliani, P. Ristori // *Opt. Commun.* — 1980. — V. 35, № 1. — P. 109—112.
9. *Rayleigh, L.* On the constant of magnetic rotation of light in bisulphide of carbon // *Philos. T. R. Soc. Lond.* — 1885. — V. 176. — P. 343—366.
10. *Faraday, M.* Experimental Researches in Electricity // *Philos. T. R. Soc. Lond.* — 1846. — V. 136. — P. 1—20. — (Nineteenth Series).

11. Андреев, Н. Ф. Четырехканальный импульсно-периодический YAG:Nd-лазер с дифракционным качеством выходного излучения / Н. Ф. Андреев, О. В. Палашов, Г. А. Пасманик, Е. А. Хазанов // Квантовая электроника. — 1997. — Т. 24, вып. 7. — С. 581—585.
12. Andreev, N. A two-channel repetitively pulsed Nd:YAG laser operating at 25 Hz with diffraction-limited beam quality / N. Andreev, E. Khazanov, O. Kulagin, B. Movshevich, O. Palashov, G. Pasmanik, V. Rodchenkov, A. Scott, P. Soan // IEEE J. Quantum Electron. — 1999. — V. 35, № 1. — P. 110—114.
13. Khazanov, E. A. Investigation of self-induced distortions of laser radiation in lithium niobate and terbium gallium garnet / E. A. Khazanov, O. V. Kulagin, S. Yoshida, D. Reitze // Proc. Conference on Lasers and Electro-Optics, San Francisco, CA, 3–8 May. — 1998. — P. 250—251.
14. Хазанов, Е. А. Компенсация термонаведенных поляризационных искажений в вентилях Фарадея // Квантовая электроника. — 1999. — Т. 26, вып. 1. — С. 59—64.
15. Khazanov, E. A. Investigation of self-induced depolarization of laser radiation in terbium gallium garnet / E. A. Khazanov, O. V. Kulagin, S. Yoshida, D. Tanner, D. Reitze // IEEE J. Quantum Electron. — 1999. — V. 35, № 8. — P. 1116—1122.
16. Khazanov, E. Suppression of self-induced depolarization of high-power laser radiation in glass-based Faraday isolators / E. Khazanov, N. Andreev, A. Babin, A. Kiselev, O. Palashov, D. Reitze // J. Opt. Soc. Am. B. — 2000. — V. 17, № 1. — P. 99—102.
17. Андреев, Н. Ф. Исследование термооптических постоянных магнитоактивных стекол / Н. Ф. Андреев, А. А. Бабин, Т. В. Зарубина, А. М. Киселев, О. В. Палашов, Е. А. Хазанов, О. С. Щавелев // Опт. ж-л. — 2000. — Т. 67, № 6. — P. 66—69.
18. Хазанов, Е. А. Особенности работы различных схем изолятора Фарадея при высокой средней мощности лазерного излучения // Квантовая электроника. — 2000. — Т. 30, вып. 2. — С. 147—151.
19. Андреев, Н. Ф. Изолятор Фарадея с развязкой 45 дБ при средней мощности излучения 100 Вт / Н. Ф. Андреев, О. В. Палашов, А. К. Потемкин, Д. Х. Райтци, А. М. Сергеев, Е. А. Хазанов // Квантовая электроника. — 2000. — Т. 30, вып. 12. — С. 1107—1108.
20. Хазанов, Е. А. Новый вращатель Фарадея для лазеров с большой средней мощностью // Квант. электроника. — 2001. — Т. 31, вып. 4. — С. 351—356.
21. Khazanov, E. Effect of terbium gallium garnet crystal orientation on the isolation ratio of a Faraday isolator at high average power / E. Khazanov, N. Andreev, O. Palashov, A. Poteomkin, A. Sergeev, O. Mehl, D. Reitze // Appl. Opt. — 2002. — V. 41, № 3. — P. 483—492.
22. Mueller, G. Method for compensation of thermally induced modal distortions in the input optical components of gravitational wave interferometers / G. Mueller, R. S. Amin, D. Guagliardo, D. McFeron, R. Lundock, D. H. Reitze, D. B. Tanner // Classical and Quantum Gravity. — 2002. — V. 19. — P. 1793—1801.
23. Khazanov, E. A. Compensation of birefringence in active elements with a novel Faraday mirror operating at high average power / E. A. Khazanov, A. A. Anastasiyev, N. F. Andreev, A. Voytovich, O. V. Palashov // Appl. Opt. — 2002. — V. 41, № 15. — P. 2947—2954.
24. Андреев, Н. Ф. Использование кристаллического кварца для компенсации термонаведенной деполяризации в изоляторах Фарадея / Н. Ф. Андреев, Е. В. Катин, О. В. Палашов, А. К. Потемкин, Д. Х. Райтци, А. М. Сергеев, Е. А. Хазанов // Квантовая электроника. — 2002. — Т. 32, вып. 1. — С. 91—94.

25. Хазанов, Е. А. Термооптика магнитоактивной среды: изоляторы Фарадея для лазеров с высокой средней мощностью // УФН. — 2016. — Т. 186, вып. 9. — С. 975—1000.
26. Железнов, Д. С. Криогенный изолятор Фарадея / Д. С. Железнов, В. В. Зеленогорский, Е. В. Катин, И. Б. Мухин, О. В. Палашов, Е. А. Хазанов // Квантовая электроника. — 2010. — Т. 40, вып. 3. — С. 276—281.
27. Snetkov, I. L. Thermo-optical and magneto-optical characteristics of terbium scandium aluminum garnet crystals / I. L. Snetkov, R. Yasuhara, A. V. Starobor, E. A. Mironov, O. V. Palashov // IEEE J. Quantum Electron. — 2015. — V. 51, № 7. — P. 1—7.
28. VIRGO-Collaboration. In-vacuum optical isolation changes by heating in a Faraday isolator // Appl. Opt. — 2008. — V. 47, № 31. — P. 5853—5861.
29. Palashov, O. V. High-vacuum-compatible high-power Faraday isolators for gravitational-wave interferometers / O. V. Palashov, D. S. Zheleznov, A. V. Voitovich, V. V. Zelenogorsky, E. E. Kamenetsky, E. A. Khazanov, R. M. Martin, K. L. Dooley, L. Williams, A. Lucianetti, V. Quetschke, G. Mueller, D. H. Reitze, D. B. Tanner, E. Genin, B. Canuel, J. Marque // J. Opt. Soc. Am. B. — 2012. — V. 29, № 7. — P. 1784—1792.
30. Katherin, L. D. Thermal Effects in the Input Optics of the Enhanced Laser Interferometer Gravitational-Wave Observatory Interferometers / L. D. Katherin, M. A. Arain, D. Feldbaum, V. V. Frolov, M. Heintze, D. Hoak, E. A. Khazanov, A. Lucianetti, R. M. Martin, G. Mueller, O. Palashov, V. Quetschke, D. H. Reitze, R. L. Savage, D. B. Tanner, L. F. Williams, W. Wu // Rev. Sci. Instrum. — 2012. — V. 83. — P. 033109.
31. Snetkov, I. L. Compensation of thermally induced depolarization in Faraday isolators for high average power lasers / I. L. Snetkov, I. B. Mukhin, O. V. Palashov, E. A. Khazanov // Opt. Express. — 2011. — V. 19, № 7. — P. 6366—6376.
32. Snetkov, I. L. Review of Faraday isolators for kilowatt average power lasers / I. L. Snetkov, A. V. Voitovich, O. V. Palashov, E. A. Khazanov // IEEE J. Quantum Electron. — 2014. — V. 50, № 6. — P. 434—443.
33. Zelenogorsky, V. Adaptive compensation of thermally induced phase aberrations in Faraday isolators by means of a DKDP crystal / V. Zelenogorsky, O. Palashov, E. Khazanov // Opt. Commun. — 2007. — V. 278, № 1. — P. 8—13.

НЕЛИНЕЙНЫЕ ПРОЦЕССЫ В МОЩНЫХ ФЕМТОСЕКУНДНЫХ ЛАЗЕРНЫХ СИСТЕМАХ

В. Е. Яшин

Государственный оптический институт им. С. И. Вавилова, г. Санкт-Петербург

Введение

Одна из фундаментальных задач лазерной физики — получать лазерные импульсы с предельными параметрами, позволяющими максимально концентрировать мультитераваттную мощность излучения в малых объемах вещества. При этом оказывается возможно изучить целый ряд принципиально новых физических явлений, протекающих при взаимодействии излучения с веществом (см., например, [1—3]). Такие лазеры широко применяются как в фундаментальных и прикладных научных исследованиях, так и для решения особо значимых конкретных проблем.

Одним из основных параметров лазеров с высокой пиковой мощностью излучения, определяющих возможность ее концентрации, является яркость излучения, которая прямо связана с интенсивностью (плотностью мощности) излучения, достижимой при его фокусировке или передаче на большие расстояния. Максимальная интенсивность излучения при его фокусировке идеальным объективом или зеркалом с фокусным расстоянием F , $I_\varphi = B_\varphi S/F^2$, пропорциональна площади апертуры пучка S и яркости излучения B_φ :

$$B_\varphi = \frac{4W}{\pi S t_p \varphi^2},$$

где W — общая энергия, φ — угловая расходимость лазерного излучения, t_p — длительность импульса.

Как видно из приведенного соотношения, для получения максимального уровня интенсивности необходимы два условия: генерация импульсов малой длительности и их жесткая фокусировка в малые объемы, т. е. достижение предельной яркости излучения. Этому препятствует целый ряд линейных и нелинейных оптических эффектов, возникающих при усилении, распространении и фокусировке сверхкоротких лазерных импульсов. Поэтому при создании лазеров с предельной яркостью излучения ключевое значение имели изучение этих эффектов и разработка методов их подавления, которые и будут рассмотрены ниже.

1. Активные лазерные среды фемтосекундных лазеров

Существует большое количество газовых, жидкостных и твердотельных лазерных сред, на которых создано также огромное количество лазеров различного применения. Вместе с тем число лазерных сред, исполь-

зуемых для создания лазеров с высокой пиковой мощностью излучения, весьма ограниченно (их можно пересчитать буквально по пальцам одной руки), что связано со специфическими требованиями, предъявляемыми к таким средам. Одним из главных требований является возможность запасть большое количество энергии в больших объемах оптически качественной активной среды на достаточно долгое время. Другое требование, более подробно рассмотренное ниже, связано в идеале с отсутствием нежелательных эффектов (или с возможностью их подавления), которые бы препятствовали энергообмену с активной среды и (или) существенно ограничивали пиковую мощность излучения. И наконец, поскольку пиковая мощность излучения напрямую связана с возможностью генерации и усиления коротких и сверхкоротких импульсов. Активные лазерные среды при подходе, основанном на прямом усилении таких импульсов (о другом подходе будет сказано ниже), должны обладать широкой полосой усиления (люминесценции).

Многочисленные исследования и разработки, ведущиеся во всем мире уже более 50 лет, показали, что по комплексу этих параметров наиболее оптимальными для лазеров со сверхкороткой длительностью импульса являются твердотельные активные среды на основе кристаллов и стекол. Так, для лазерных систем с относительно низкой частотой следования импульсов это неодимовое или иттербиевое стекло, а для лазерных систем с высокой средней мощностью — кристаллические активные среды [3, 4]. Особый класс мощных фемтосекундных лазеров составляют системы, основанные на параметрическом усилении световых пучков в кристаллах [3].

Рассмотрим далее кратко основные эффекты, ограничивающие пиковую мощность в лазерных системах, имея в виду в основном твердотельные лазеры.

2. Эффекты, ограничивающие мощность и яркость излучения

Как было выяснено в многочисленных исследованиях, среди всего разнообразия нелинейных явлений только три эффекта оказывают наибольшее ограничивающее влияние на мощность и энергию излучения в лазерных системах. Это оптический пробой, самофокусировка и вынужденное рассеяние Манделъштама — Бриллюэна [4].

2.1. Оптический пробой

Для твердотельных лазеров, работающих в видимом и ближнем инфракрасном диапазонах спектра, оптический пробой для импульсов длительностью более нескольких пикосекунд обусловлен тепловым механизмом, т. е. поглощением энергии излучения малыми включениями в объеме или на поверхности материала, их разогревом до высоких температур с последующим плавлением и испарением материала, приводящим к разру-

шению. Этот механизм характеризуется степенной зависимостью максимальной плотности энергии от длительности импульса:

$$W_{th} \approx A(t_p)^n,$$

где W измеряется в джоулях на сантиметр в квадрате, а t — в наносекундах. Для проведения оценок можно считать, что $n \approx 0,5$. Постоянная A зависит от типа материала и включений, а также от формы импульса [5, 6]. Для надежной работы лазера рабочая плотность энергии выбирается, естественно, существенно меньше порога разрушения и зависит от возможной величины модуляции интенсивности в пучке, т. е. в конечном счете от так называемого коэффициента заполнения апертуры:

$$F = \frac{\iint I(x, y) dx dy}{I_{\max} S},$$

где $I(x, y)$ — распределение интенсивности по сечению пучка, площадь апертуры которого равна S , I_{\max} — максимальная интенсивность в пучке. В современных лазерных системах, использующих методы аподизации и ретрансляции изображения пучка по усилительной системе, $F > 0,8$, и модуляции интенсивности соответственно малы. В этом случае плотность энергии в лазерной системе может очень близко подходить к плотности энергии разрушения. Для лазерных импульсов короче 10—20 пс механизм разрушения изменяется с теплового на лавинный [7], что приводит к изменению зависимости пороговой плотности энергии разрушения от длительности импульса.

2.2. Самофокусировка

Самофокусировка излучения является важным фактором, ограничивающим пиковую мощность импульсов в мощных лазерах. При этом различают крупномасштабную самофокусировку (КМС), когда пучок самофокусируется как целое, и мелкомасштабную самофокусировку (ММС).

Мелкомасштабная самофокусировка пучков — второй после оптического пробоя наиболее важный механизм ограничения мощности излучения в лазерных системах [4, 8, 9]. Она приводит к распаду пучка на нити — филаменты, в которых интенсивность излучения может нарастать вплоть до уровня, вызывающего разрушение оптических элементов. На промежуточной стадии, когда разрушение еще не происходит, ММС приводит к увеличению угловой расходимости излучения, а значит, уменьшению его яркости. Интегральной характеристикой, позволяющей оценивать ММС, является так называемый интеграл распада, определяемый в единицах СГСЕ следующей формулой:

$$B = \frac{8\pi^2 n_2}{\lambda c n} \int_0^L I(x, y, z) dz,$$

где λ — длина волны и c — скорость света в вакууме, n и n_2 — линейный и нелинейный показатели преломления нелинейной среды, L — длина нелинейной среды. Интеграл распада определяет нарастание интенсивности в мелкомасштабных пространственных возмущениях интенсивности: $I = I_0 \exp(2B)$, а также нелинейный набег фазы в мощном лазерном пучке: $\varphi = kLn + B$.

В простейшем случае сплошной нелинейной среды возмущения интенсивности нарастают в определенной полосе пространственных частот

$$0 < \kappa_{\perp} < \kappa_{\perp \text{кр}},$$

где

$$\kappa_{\perp \text{кр}} = K \left(\frac{2n_2}{n} \right)^{\frac{1}{2}} |E_0|$$

— критическое значение поперечного волнового числа, K — волновое число лазерной волны. Угол распространения возмущений по отношению к основному мощному пучку, который служит накачкой, определяется выражением

$$\Theta_{\perp} = \frac{\kappa_{\perp}}{K}.$$

Максимум усиления с инкрементом, определяемым выражением для B , достигается при $\kappa_{\perp \text{max}} = 2^{-1/2} \kappa_{\perp \text{кр}}$. Поперечный размер наиболее быстро развивающихся возмущений равен

$$\Lambda_{\perp} = \frac{2\pi}{\kappa_{\perp \text{max}}} = \frac{2\pi}{K \sqrt{\frac{n_2}{n} |E_0|^2}} \approx \sqrt{\frac{\pi P_{\text{кр}}}{I}},$$

причем одна неоднородность поля площадью $S = \pi \Lambda_{\perp}^2 / 4$ содержит мощность, примерно равную критической мощности самофокусировки $P_{\text{кр}}$, а продольная длина, на которой начальная амплитуда возмущений нарастает в $\exp \pi \approx 23$ раза, определяется следующим выражением:

$$l_{\text{ММС}} = \frac{\pi}{B} = \frac{2\pi}{K \sqrt{\frac{n_2}{n} |E_0|^2}} = \frac{\lambda n_0 c}{8\pi^2 n_2 I}.$$

При характерных для мощных лазеров параметрах излучения ($I \approx 5 \text{ ГВт/см}^2$, $n_2 \approx 1,3 \cdot 10^{-13}$ ед. СГСЕ) получаем $\Lambda_{\perp} \approx 0,5$ мм и $l_{\text{ММС}} \approx 7$ см. Эти цифры иллюстрируют важную роль мелкомасштабной самофокусировки в ограничении пиковой мощности излучения. Длины мелкомасштабной и крупномасштабной самофокусировки соотносятся как

$$\frac{l_{\text{ММС}}}{l_{\text{КМС}}} \approx \sqrt{\frac{P_{\text{кр}}}{P}}.$$

При $P \gg P_{кр}$ выполняется неравенство $I_{КМС} \gg I_{ММС}$, что и предопределяет доминирование мелкомасштабной самофокусировки.

Допустимая величина B в одном фрагменте нелинейной среды усилителей, при превышении которой наступает катастрофическое ухудшение угловой расходимости, а то и пробой нелинейной среды, зависит от уровня начальных возмущений интенсивности и колеблется в диапазоне $B = 1 \div 5$. Существуют хорошо разработанные методы подавления самофокусировки [4], которые будут рассмотрены далее. Они позволяют увеличить суммарную величину B , а значит, и интенсивности до 10—30 при сохранении, тем не менее, величины B в одном фрагменте усилительной системы на прежнем уровне.

2.3. Вынужденное рассеяние

Наконец, третий ограничивающий энергию лазерных систем эффект — это вынужденное рассеяние Мандельштама — Бриллюэна (ВРМБ), в результате развития которого излучение спонтанных шумов на стоксовой частоте экспоненциально нарастает в поле излучения накачки: $I = I_0 \exp G$. Этот эффект начинает проявляться, когда инкремент нарастания интенсивности G достигает некоторого порогового уровня $G_{th} = 20 \div 30$:

$$G = g \int_0^L I(x, y, z) dz > G_{th},$$

где g — локальный инкремент, являющийся параметром нелинейной среды. Например, для неодимовых стекол $g = (1 \div 2,5)$ см/ГВт [10].

Развитие рассмотренных нелинейных эффектов, а значит, и их значение зависит от длительности импульса. Так, для импульсов, короче 0,1—1 нс, определяющее влияние оказывает самофокусировка. Для $t_p > (1 \div 10)$ нс основным эффектом является оптический пробой, а для более длинных импульсов $t_p > (3 \div 50)$ нс в игру может вступать ВРМБ. В пограничных областях могут проявляться несколько эффектов, и относительное влияние того или другого будет зависеть от конкретных обстоятельств.

Возможность развития тех или иных нелинейных эффектов приводит к ограничению максимальной плотности энергии W в лазерных усилителях, а значит, к существенному возрастанию их апертуры и апертуры других оптических элементов в случае, если требуется получить большую выходную энергию. Для сред с большой плотностью энергии насыщения W_{sat} , к которым относится, например, неодимовое стекло ($W_{sat} \approx h\nu/\sigma = (3 \div 10)$ Дж/см² в зависимости от типа стекла), ограничение W на уровне значительно меньше W_{sat} ведет к низкому энергоэффекту.

Еще более заметные ограничения налагает самофокусировка — вследствие более сильной зависимости максимальной плотности энергии от длительности импульса. Это влечет значительные трудности при получе-

нии пикосекундных и фемтосекундных импульсов достаточно большой энергии в лазерных системах, использующих прямое усиление излучения, и привело в 1980-х годах к поиску другой архитектуры лазерных систем [11—13].

3. Классические методы подавления самофокусировки в лазерах прямого усиления

Прежде чем перейти к рассмотрению существующих методов борьбы с самофокусировкой, отметим наиболее прямой, но, к сожалению, не очень радикальный метод ее подавления — уменьшение n_2 лазерных материалов. Пределы варьирования n_2 в лазерных стеклах и кристаллах очень ограничены [4]. Не нашел практического применения и другой метод, заключающийся во введении в усилительный тракт дополнительных элементов с отрицательным значением нелинейного показателя преломления. К сожалению, не нашлось соответствующих сред с высокими значениями пропускания и высокой лучевой стойкостью.

3.1. Пространственная фильтрация

Один из методов борьбы с ММС заключается в уменьшении амплитуды затравочных возмущений интенсивности. Появление таких возмущений связано как с регулярными, так и со случайными причинами. Регулярными источниками модуляции интенсивности служат интерференция света и дифракция на формирующих и ограничивающих апертурах. Например, дифракционные возмущения от формирующих диафрагм могут попадать в область самофокусировочной неустойчивости и усиливаться до больших значений, вызывая даже разрушение лазерных активных элементов, что было уставлено еще в 1970-х годах [4, 14]. Приближение формирующей диафрагмы к нелинейной среде позволяет вывести дифракционные возмущения из полосы самофокусировочной неустойчивости, что в соединении с оптической ретрансляцией изображения диафрагмы по усилительной системе позволяет формировать световые пучки с высоким фактором заполнения [15]. Другой метод подавления дифракционных возмущений заключается в аподизации световых пучков. В современных сверхмощных лазерных усилительных системах обычно используется как аподизация, так и передача изображения формирующей апертуры по усилительному тракту с помощью телескопов-ретрансляторов [4].

Наконец, одним из главных и практически неустраняемых источников возмущений является рассеяние мощной волны на различного рода мелкомасштабных неоднородностях — пыли, включениях и воздушных пузырьках в стекле, покрытиях и других материалах. Нарастание возмущений до уровня амплитуды основной волны, когда начинается нелинейная стадия их развития, происходит при интегралах распада $B = 1 \div 5$ в зависимости от количества «шумов» в лазерной системе.

Основным методом подавления самофокусировки в настоящее время стала пространственная фильтрация — оптическая ретрансляция лазерного излучения. Пространственный фильтр представляет собой телескоп Кеплера, в общем фокусе двух линз которого располагается диафрагма. Эта диафрагма отбирает пространственные возмущения, нарастающие в результате ММС, не давая им развиваться в следующем усилителе. Оценку диаметра d_d диафрагмы, необходимого для блокировки наиболее опасных возмущений интенсивности, можно получить из выражения для угла распространения этих возмущений:

$$\Theta_{\max} = \frac{\kappa_{\perp \max}}{K} = \sqrt{\frac{8\pi n_2 I}{cn^2}}.$$

Возмущения задерживаются диафрагмой, если $\Theta_{\max} F_1 > d_d/2$ (F_1 — фокусное расстояние входной линзы). Отсюда следует выражение для полосы пропускания пространственного фильтра:

$$\Delta\Theta_{\text{пф}} < 2\Theta_{\max}.$$

Разделив это выражение на дифракционную расходимость $\Theta_d \approx 2\lambda/D$ (D — диаметр апертуры пучка), получим полосу пропускания в единицах дифракционного угла:

$$\frac{\Delta\Theta_{\text{пф}}}{\Theta_d} = \Theta_{\max} D/\lambda.$$

Например, для лазеров на неодимовом стекле с $n_2 = 1,3 \cdot 10^{-13}$ ед. СГСЕ, $n = 1,55$, $I = 5$ ГВт/см² получаем $2\Theta_{\max} \approx 3$ мрад, что в десятки раз превышает дифракционную расходимость пучков диаметром более 1 см с длиной волны излучения в районе 1 мкм. На практике диаметр диафрагмы пространственных фильтров в мощных лазерных системах выбирается так, чтобы полоса пропускания пространственного фильтра превышала дифракционную расходимость в 10—20 раз. Это связано с необходимостью предотвращать перекрытие диафрагмы плазмой, которая может возникнуть при воздействии крыльев распределения интенсивности на материал диафрагмы. Кроме того, в пространстве между двумя линзовыми объективами создается вакуум, чтобы избежать пробоя воздуха вблизи фокуса.

3.2. Использование фазовых эффектов для ограничения ММС

Наличие промежутков между участками нелинейной среды (например, между дисками в дисковых усилителях) влияет на развитие мелкомасштабных возмущений. Это связано с тем, что промежутки вносят сдвиг фаз между пространственными гармониками $\cos(\kappa_{\perp} r)$ (r — поперечная координата) и основной волной: $\Delta\varphi = \kappa_{\perp}^2 L/(2K)$ (K — волновое число в воздухе, L — длина промежутка), а развитие возмущений существенным образом зависит от $\Delta\varphi$. Так, например, если $\Delta\varphi = \pi/2$, то возмущение, на-

растающее в первом нелинейном элементе, затухает во втором. Казалось бы, это открывает возможность подавлять ММС путем разделения среды на отдельные фрагменты, как это делается, например, в дисковом усилителе. Однако, как показывает анализ развития ММС [16], такое разделение не позволяет снизить коэффициент усиления возмущений для всей области пространственных частот. Физически это связано с тем, что сдвиг фаз в промежутке приводит к затуханию возмущений на одних пространственных частотах, но к усилению возмущений на других.

Существует другой, более эффективный, способ подавления ММС с помощью фазовых эффектов. Он заключается в использовании оптических ретрансляторов — двух софокусных линз. При определенных ограничениях на интеграл распада в нелинейных элементах и геометрические размеры ретрансляторов фазовые возмущения, нарастающие в первой нелинейной среде, расположенной до ретранслятора, в результате введенной фазовой задержки преобразуются в убывающие во второй нелинейной среде, за ретранслятором [17].

Наиболее просто условия подавления ММС записываются для системы пассивных нелинейных фрагментов, разделенных ретрансляторами с параметрами $L = l/n$, $B < \pi/2$, где L — длина ретрансляторов, l — длина нелинейных фрагментов.

Для усиливающей нелинейной среды сохраняется такое же, как и для пассивной среды, ограничение на максимальный интеграл распада на одном элементе: $B < \pi/2$, а геометрические размеры ретрансляторов изменяются. Оптимальные геометрические размеры можно подбирать на основе численного расчета коэффициента передачи.

Учет фазовых эффектов при выборе длины ретрансляторов немного больше оптимального значения $L = l/n$ для пассивной среды позволяет отодвинуть пространственную частоту наиболее опасных возмущений в высокочастотную область, что, в свою очередь, дает возможность в 2—3 раза расширить полосу пропускания пространственных фильтров вплоть до 10 мрад. Это важно, например, для предотвращения перекрытия диафрагмы плазмой.

3.3. Выбор поляризации излучения

Скорость развития мелкомасштабных возмущений при ММС зависит от поляризации излучения [4, 18]. Как и в случае линейно поляризованных волн, можно найти инкремент нарастания возмущений для эллиптически поляризованной волны, которую можно представить в виде суммы право- и левополяризованных полей. Такие возмущения в лазерных системах распадаются на две группы волн с ортогональными поляризациями, некоторые из которых затухают, а некоторые нарастают в поле мощной волны [18]. Однако даже для нарастающих возмущений инкремент уменьшается по сравнению с линейно поляризованной волной, что связано с уменьшением эффективного коэффициента нелинейности n_2 .

При изменении коэффициента эллиптичности от линейной поляризации до круговой граничное значение пространственной частоты в сплошной нелинейной среде уменьшается в 1,5 раза и, что более важно, в 1,5 раза уменьшается значение интеграла распада.

Существуют и другие способы подавления самофокусировки, заключающиеся, например, в нарушении пространственной или временной когерентности излучения [4, 19]. Однако эти способы меньше применяются в современных мощных лазерных системах, и мы не будем их рассматривать.

4. Ограничения на параметры лазерных систем, накладываемые нелинейными эффектами

Как уже отмечалось, нелинейные эффекты приводят к ограничению максимальной плотности энергии W в лазерных усилителях и возрастанию их апертуры и апертуры других оптических элементов. Для сред с большой плотностью энергии насыщения W_{sat} , как, например, неодимовое стекло, ограничение W на уровне значительно меньше W_{sat} ведет к низкому энергосъему. В качестве примера на рис. 1 приведена зависимость эффективности съема энергии η для неодимовых стекол с различными значениями поперечного сечения вынужденного излучения σ от длительности импульса t_u при ограничении плотности энергии лучевой стойкостью. Видно, что даже для фосфатных стекол с большим значением σ эффективность съема для импульсов короче 2—5 нс невелика и еще больше снижается для стекол с малой величиной σ .

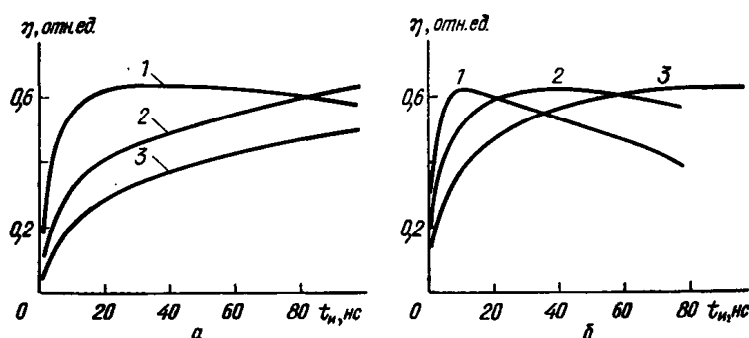


Рис. 1. Эффективность энергосъема в различных неодимовых стеклах в зависимости от длительности импульса при ограничении плотности выходной энергии лучевой стойкостью на уровне $W_{s \text{ вых}} = 3(t_u)^{1/2}$ (а) и $W_{s \text{ вых}} = 6(t_u)^{1/2}$ (б): 1 — ГЛС-21, 2 — ГЛС-1, 3 — ГЛС-6. Размерности: W_s — Дж/см², t_u — нс. Параметры усилителя: усиление на проход $G_0 = 10$, потери $k_{1,06} = 2 \cdot 10^{-3}$ см, плотность энергии насыщения $W_{sн} = 0,75h\nu/\sigma$

5. Лазеры с компрессией импульса

Основная идея лазеров с компрессией импульса [11—13] основана на том, что излучение усиливается при относительно большой длительности, а компрессия производится уже на выходе лазерной системы. Для того чтобы получить большой выигрыш в пиковой мощности и эффективности энергосъема, величина компрессии, определяемая как отношение усиленного и сжатого импульсов, должна быть значительной (несколько десятков или даже сотен раз), а сами схемы сжатия должны обладать высокой эффективностью.

В настоящее время наиболее эффективными и широко применяемыми схемами компрессии являются сжатие импульса с использованием вынужденного рассеяния и компрессии фазово-модулированных (чирпированных) импульсов дисперсионными оптическими элементами. Последний подход, имеющий англоязычную аббревиатуру CPA (chirped pulse amplification) [12], обеспечивает существенно более высокую степень компрессии, и поэтому мы рассмотрим его более подробно.

5.1. Лазеры с дисперсионной компрессией импульса

Наиболее высокие коэффициенты сжатия импульсов реализуются в CPA лазерных системах. Такой подход, впервые предложенный в работе [12], основан на сжатии импульсов с предварительно введенной фазовой самомодуляцией в диспергирующих элементах, в которых временная задержка какой-то части импульса зависит от мгновенной частоты. В качестве таких элементов выступают призмы, дифракционные решетки и в последнее время так называемые чирпированные зеркала.

В соответствии с этим подходом типичная блок-схема лазера с компрессией импульса выглядит, как это изображено на рис. 2. В качестве активных сред таких лазеров используются широкополосные среды. В настоящее время созданы лазеры практически на всех твердотельных активных средах. Выбор той или иной активной среды зависит от решаемых задач. Так, наиболее короткие импульсы, длительностью около 5 фс, получены в лазере на основе титан-сапфира $Ti:Al_2O_3$ [20], обладающего наибольшей шириной полосы усиления. Лазеры на основе этой и других кристаллических активных сред способны работать с большой частотой повторения импульсов, а значит, и с относительно высокой средней мощностью излучения. Наибольшая энергия и пиковая мощность получены в лазере, где в качестве оконечных усилительных каскадов использовалось неодимовое стекло, позволяющее, как известно, изготавливать активные элементы большой апертуры [4]. Среди сред, которые широко применяются в лазерах с компрессией в последнее время, следует отметить активированные Yb-кристаллические активные среды (Yb:YAG, Yb:KGW и др.), а также фосфатное стекло [21]. Перспективность этих сред заклю-

чается в возможности использовать прямую лазерную диодную накачку, малом стоксовом сдвиге между длинами волн накачки и генерации, а значит, и малом тепловыделении и широком спектре люминесценции, допускающем работу с фемтосекундными импульсами.

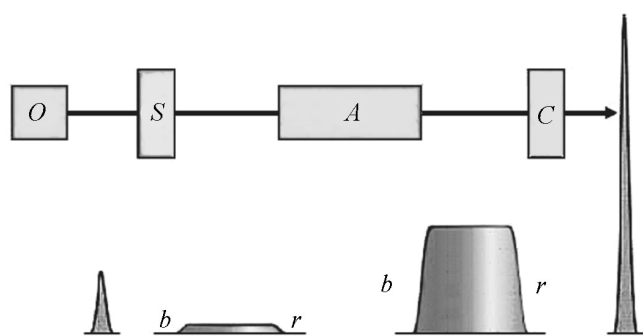


Рис. 2. Примерная блок-схема лазерной системы, использующей принцип усиления chirпированного (фазово-модулированного) импульса: *O* — задающий генератор, *S* — стретчер (система удлинения импульса), *A* — усилительная система, *C* — компрессор, *b* и *r* — коротковолновая и длинноволновая части спектра импульса

Рассмотрим основные элементы приведенной блок-схемы. На первом этапе необходимо получить chirпированный импульс. В настоящее время наиболее перспективным путем получения такого импульса является удлинение первоначально сверхкороткого импульса, генерируемого, как правило, в лазере с самосинхронизацией мод. Длительность первоначального короткого импульса, определяемого в конечном счете соотношением неопределенности $t_p = k/\Delta\nu$ (k — коэффициент порядка 1, зависящий от формы импульса, $\Delta\nu$ — ширина спектра люминесценции), лежит в пределах от 10 фс до 1 пс в зависимости от используемой активной среды. Затем такой импульс удлиняется в дисперсионной линии задержки с одновременным приобретением фазовой модуляции (чирпа). В качестве такой линии задержки можно использовать, например, пару дифракционных решеток с телескопом, размещенным между ними [22]. Телескоп служит для изменения знака дисперсии этой оптической системы. Задержка оптического пути, вносимая стретчером (и компрессором также [23]), описывается следующим выражением:

$$\Delta\tau = \frac{L\lambda N \Delta\lambda}{c[1 - (\lambda N \sin \theta)^2]},$$

где θ — угол падения света на решетку, N — количество штрихов на единицу размера у дифракционной решетки, $\Delta\lambda$ — ширина спектра в едини-

цах длин волн, L — эквивалентная оптическая длина между решетками, рассчитываемая с учетом того, что телескоп переносит изображение назад. Линзовый телескоп вносит в пучок хроматические aberrации, что может приводить к ухудшению качества компрессии. Поэтому в устройстве увеличения длительности предпочтительно использовать зеркальные оптические элементы, которые позволяют устранить aberrации [24].

Такая система, называемая обычно стретчером, обеспечивает линейный chirp (зависимость мгновенной частоты от времени), а энергия импульса получается достаточно большой (от нано- до микроджоулей — в зависимости от типа генератора). Поэтому практически во всех лазерах с компрессией, создаваемых по этому принципу, применяется такая система удлинения импульса. Используемая на ранней стадии система уширения спектра и удлинения импульса в одномодовом световолокне [12] практически вышла из употребления из-за вносимых такой системой фазовых искажений, что приводит к искажению формы сжатого импульса и снижению его контраста.

Как следует из предыдущего изложения, для подавления нелинейных эффектов в процессе усиления необходимо максимально удлинять импульс. Степень удлинения так же, как и последующей компрессии, в рассматриваемом подходе определяется шириной спектра и длиной стретчера и компрессора. При разумной длине этих устройств (метры) и использовании многопроходных схем максимальная длительность растянутого импульса составляет около 2—3 нс при коэффициенте удлинения примерно 10^4 — 10^6 . Это хотя и обеспечивает огромный выигрыш в подавлении нелинейных эффектов по сравнению со схемой прямого усиления, но не устраняет их полностью.

Генерация исходного короткого импульса происходит обычно в генераторах с самосинхронизацией аксиальных мод. Не останавливаясь на подробностях их работы и условиях реализации стабильной синхронизации мод (см., например, обзор [25]), отметим, что длительность импульса на выходе этих генераторов определяется в конечном счете двумя факторами: шириной линии усиления активной среды и временем релаксации нелинейного устройства, используемого для синхронизации. Так, например, наиболее короткие импульсы, длительностью около 5 фс, генерируются в $\text{Ti:A1}_2\text{O}_3$ -лазерах с самосинхронизацией мод керровской нелинейной линзой, наводимой самим лазерным импульсом в активном элементе. Эти лазеры работают в непрерывном режиме, так что энергия одиночного импульса весьма мала (нано-джоули). В импульсных лазерах на неодимовом стекле с самосинхронизацией мод органическими красителями и с отрицательной обратной связью, применяемой для стабилизации выходных параметров, типичная длительность импульса составляет около 0,5—1 пс. Использование другого типа нелинейного элемента на основе полупроводниковых структур с квантовыми ямами и непрерывной полу-

проводниковой накачки позволило уменьшить длительность импульса в лазере на неодимовом стекле до ~ 100 фс [26]. Нелинейные отражающие зеркала на основе полупроводниковых структур подобного типа с квантовыми ямами и точками используются в последнее время в большинстве фемтосекундных генераторов, так как позволяют реализовать самостартующий режим синхронизации мод.

Усиление растянутого во времени импульса происходит, как правило, в два этапа, что связано с необходимостью реализовать большое усиление (порядка 10^6 — 10^9). На первом из них энергия импульса увеличивается с нано- или микроджоульного уровня до миллиджоульного уровня. Делается это обычно в регенеративных или многопроходовых усилителях, когда излучение проходит через один и тот же объем активной среды несколько раз.

Для дальнейшего усиления фазово-модулированных импульсов должны применяться усилители большего диаметра, соответствующего заданному уровню выходной энергии. Выходной диаметр усилителя, как это уже обсуждалось ранее, определяется двумя эффектами, ограничивающими пиковую мощность излучения: оптическим пробоем и самофокусировкой.

Для подавления мелкомасштабной самофокусировки в лазерной усилительной системе может использоваться широкий набор средств, рассмотренных выше (см. также [4]): усилители с высоким коэффициентом усиления, а значит, малой длиной усиливающей среды, пространственная фильтрация и ретрансляция пучка, круговая поляризация излучения и т. д. Все эти меры позволяют поднять плотность мощности в усилительных каскадах до 5 — 6 ГВт/см², что уже близко к лучевой стойкости оптических элементов.

Для упрощения конструкции и снижения стоимости усилительной системы могут применяться те же методы, что и для оптимизации лазеров для термоядерного синтеза. В частности, можно использовать многопроходовые усилительные схемы, минимизирующие количество усилителей и оптических элементов [4].

При усилении импульса в регенеративных усилителях (РУ), а также в усилительной системе спектр импульса, а значит, и длительность, и форма сжатого импульса могут модифицироваться вследствие различных эффектов: ограниченности полосы усиления активной среды и полосы пропускания различных элементов усилительной системы, фазовой самомодуляции, обусловленной нелинейностью показателя преломления, насыщения усиления, временного перекрытия нескольких импульсов из задающего генератора в регенеративном усилителе. Влияние модуляции спектра, вызываемой, например, паразитными элементами Фабри — Перо в ячейках Поккельса, линзах, поляризаторах, зеркалах подчеркивается многопроходным характером усиления в РУ, что приводит к необходимости очень

тщательно отбирать и разрабатывать элементы РУ, просветляя их, изменяя наклон поверхностей элементов, исключая эффекты двулучепреломления. Мелкомасштабная модуляция спектра импульса, приводящая к низкому контрасту сжатого импульса, может возникать и при недостаточно высоком контрасте выделения одиночного импульса после задающего генератора. В этом случае основной выделенный импульс и предимпульсы или послепульсы могут перекрываться во времени в регенеративном усилителе, приводя к модуляции спектра [27]. Для устранения этого эффекта необходимо тщательно отделять одиночный импульс с помощью оптических развязок.

Обужение спектра сигнала при его усилении в активной среде с полосой усиления $\Delta\lambda_{lum}$ гауссовской формы описывается следующей формулой [28]:

$$\Delta\lambda_{out} = \Delta\lambda_{in} \left[1 + \beta \ln(G) (\Delta\lambda_{in} / \Delta\lambda_{lum})^2 \right]^{-1/2},$$

где $\Delta\lambda_{out}$ и $\Delta\lambda_{in}$ — ширина спектра выходного и входного импульсов, G — полное усиление в системе, $\beta = 2[(\lambda_{in} - \lambda_{lum}) / \Delta\lambda_{lum}]^2 - 1 \approx 1$. При общем усилении в лазерной системе на фосфатном неодимовом стекле в 10^9 раз ширина спектра исходного 100 фс импульса уменьшается в 4—5 раз, а длительность сжатого импульса соответственно возрастает пропорционально этой величине.

В том же регенеративном усилителе или в основной усилительной системе могут быть использованы неодимовые стекла различного состава, имеющие разные ширины и положения максимума линии люминесценции [29]. Спектр усиления можно еще более расширить, задействовав большее количество марок стекол. Применяя этот метод в лазерных системах на неодимовом стекле, можно, как показывают расчеты и эксперименты, получать импульсы длительностью около 150 фс с энергией более 100 Дж [3, 30]. В этом случае необходимо также дополнительно профилировать спектр импульсов с помощью спектральных фильтров [31, 32]. При более высоких выходных энергиях длительность импульса несколько увеличивается. Так, в первом петаваттном лазере она составляла около 440 фс при выходной энергии 600 Дж [33].

Насыщение усиления может также исказить длительность и форму импульса, поскольку при этом меняется форма растянутого импульса в процессе усиления: временной максимум импульса сдвигается к переднему фронту, что изменяет баланс энергии импульса для разных участков чирпа. Этот эффект наиболее заметен лишь при значительном превышении плотности энергии насыщения активной среды $W_{sat} = hv/\sigma$. В лазерах на неодимовом стекле с $W_{sat} = (4 \div 10)$ Дж/см² для импульсов длительностью 1—2 нс после стретчера такую плотность создать не удастся из-за ограничений со стороны самомодуляции и самофокусировки (см. выше),

поэтому влиянием этого эффекта по сравнению с другими можно пренебречь.

Для эффективного сжатия мощных чирпированных импульсов используются решеточные компрессоры [23] как в однопроходовой, так и в двухпроходовой конфигурациях. При этом геометрические параметры компрессора выбираются на основе вышеприведенного соотношения для оптической задержки. Однопроходовая схема, обладая большей энергетической эффективностью, имеет, тем не менее, и существенный недостаток — эллиптическое сечение пучка на выходе, что увеличивает размер пятна и требует оптики большой апертуры при его фокусировке. Поэтому большее распространение получили двухпроходовые компрессоры, в которых излучение два раза проходит через компрессор, компенсируя тем самым эллиптичность. Для вывода излучения из компрессора используется смещение пучка на втором проходе в плоскости, перпендикулярной дисперсии, с помощью системы двух зеркал или призм. Поскольку длительность сжатого импульса на выходе из компрессора 0,1—1 пс занимает в пространстве длину доли миллиметра, это накладывает на точность юстировки компрессора очень высокие требования, особенно для фемтосекундных импульсов.

Так как из-за сжатия импульса интенсивность излучения в компрессоре достигает величин более 100 ГВт/см^2 , уже кубичной нелинейности воздуха $n_2 \approx 10^{-15}$ ед. СГСЕ достаточно для развития самофокусировки ($B > 3$), что резко снижает фокусируемость излучения. Поэтому компрессор больших лазерных систем вакуумируют. По этой же причине стараются применять и зеркальные фокусирующие системы при фокусировке излучения на мишень.

Для компрессоров используются как нарезные, так и голографические дифракционные отражательные решетки с количеством штрихов на 1 мм $n = 1500 \div 1800$ при работе в первом порядке дифракции. С увеличением n увеличивается угол падения света на решетки и уменьшается расстояние между ними, что позволяет делать конструкции компрессоров более компактными. С другой стороны, при этом уменьшается световая апертура решеток и увеличиваются требования к точности их изготовления. Так, например, при $n = 1800$ и угле падения около 70° даже небольшой кривизны поверхности решеток достаточно, чтобы интенсивность излучения в фокусе линзы уменьшилась более чем в 100 раз из-за астигматизма, возникающего при падении веера лучей на слегка изогнутые наклонные поверхности. По этим причинам выбирают n в диапазоне 1200—1740 штрихов на 1 мм.

Из-за многопроходового характера компрессора предъявляются высокие требования к дифракционной эффективности решеток. Так, для получения общего коэффициента пропускания двухпроходового компрессора более 60 % дифракционная эффективность решетки должна превы-

шать 90 %. Такие эффективности достигаются как в нарезных, так и в голографических решетках для первого порядка дифракции [34—36]. Однако изготовить нарезные решетки большой апертуры с требуемым количеством штрихов на 1 мм затруднительно из-за износа резца в процессе нарезания.

Ключевым требованием к дифракционным решеткам является их высокая лучевая стойкость. Для импульса длительностью 1 нс порог разрушения металлических решеток составляет менее 1 Дж/см² [36—39], что более чем на порядок меньше порога разрушения оптических материалов и покрытий в усилительной системе. К примеру, для создания петаваттных лазеров необходимы решетки метрового размера. Технологию изготовления таких решеток успешно развивают в Ливерморской национальной лаборатории [37]. В петербургской компании «Хологрейт» могут изготавливать металлизированные голографические дифракционные решетки для лазерных применений размером до 40 см.

Таким образом, дифракционные решетки являются слабым элементом лазерной системы, что стимулировало поиск методов создания более стойких решеток. Один из таких методов состоит в нанесении дополнительных диэлектрических покрытий на металлическую дифракционную решетку. Эти дополнительные слои могут значительно (в несколько раз) уменьшать долю энергии, доходящую до металла, способствуя тем самым повышению лучевой стойкости. Проведенные эксперименты подтвердили тенденцию повышения лучевой стойкости до 2 раз при использовании четырехслойного покрытия на основе пары ZrO₂/SiO₂ [37—39].

Другой способ повышения лучевой стойкости — получить дифракционную структуру в многослойных диэлектрических покрытиях [40, 41], которые, как известно, имеют значительно более высокий порог разрушения, чем металлы. Проведенные исследования [41] показали, что возможно создавать подобные решетки с достаточно высокой дифракционной эффективностью (больше 90 %). При этом лучевая прочность таких решеток примерно в 5 раз превышает прочность металлизированных решеток [42]. Основная проблема связана с высокой стоимостью таких решеток.

5.2. Дополнительная компрессия сверхкоротких импульсов

Как следует из вышесказанного, предельные по длительности характеристики лазерного излучения часто нельзя реализовать на выходе лазерной системы. Для лазеров с компрессией на вынужденном рассеянии это связано с нестационарностью последнего, а для лазеров с усилением чирпированного импульса — с ограниченностью полосы усиления активной среды. В связи с этим представляет интерес дополнительное сжатие импульса. Это можно сделать, например, используя эффект фазовой самомодуляции. Однако применять этот эффект мешает мелкомасштабная самофокусировка, ограничивающая максимальный интеграл распада, а значит,

и степень сжатия на уровне $B = 2\div 5$. Поэтому необходимым условием уширения спектра и дальнейшей компрессии является подавление ММС.

Подавлять ММС можно в одномодовом стеклянном или кварцевом световоде. Однако энергия импульсов в этом случае невелика из-за малого размера волновода. Для увеличения выходной энергии были предложены два способа. В первом из них уширение спектра производится в полом стеклянном волноводе, заполненном благородным газом [43]. Благодаря тому что поперечный размер моды в полом волноводе (около 100 мкм для видимого и ближнего ИК-диапазона) превышает аналогичный размер в стекловолокне (~10 мкм), выходная энергия также существенно повышается — с микро- до субмиллиджоульного уровня.

Значительно большее увеличение энергии можно получить при уширении спектра в объеме нелинейной среды, подавляя ММС с помощью пространственных фильтров-ретрансляторов [44—47]. Общий интеграл распада B , характеризующий величину уширения спектра в системе, равен $B = NB_1$, где B_1 — интеграл распада на одном нелинейном элементе. В такой системе с помощью простых пассивных электрооптических или магнитооптических развязок можно легко реализовать многопроходовой или регенеративный режим работы, что увеличивает значение B в m раз, где m — число проходов. Так, например, даже при $N = 3$ и $m = 2$ можно получить $B = 6B_1 \approx 20\div 40$, т. е. такое же сжатие импульса.

При очень высокой интенсивности импульсов (несколько тераватт на сантиметр в квадрате) поперечный размер самофокусирующихся в результате ММС филаментов становится настолько малым, а их угловая расходимость настолько большой, что для их селекции могут использоваться уже не пространственные фильтры, а обычные диафрагмы, расположенные на некотором сравнительно небольшом расстоянии от нелинейной среды, или просто отрезок свободного пространства [48, 49]. Это сильно упрощает оптическую схему уширения спектра.

Для сжатия импульса после схемы уширения спектра могут использоваться chirпированные зеркала, обладающие высокой лучевой стойкостью и большой апертурой [50].

Заключение

Нелинейные эффекты, такие как оптический пробой и самофокусировка, являются основными факторами, ограничивающими пиковую мощность излучения мощных лазерных систем со сверхкороткой длительностью импульса. Разработанные за последние десятилетия методы подавления этих нелинейных эффектов наряду с разработкой новых лазерных сред и методов их накачки позволили поднять мощность лазерных систем в десятки раз, доведя ее до «рутинного» петаваттного уровня [3, 51]. На повестке дня стоит задача создания мультипетаваттных и даже эксаватт-

ных систем [52, 53], что потребует развития существующей СРА-архитектуры, например на основе когерентного фазирования многих лазерных каналов. Эти лазерные системы позволяют создавать сверхсильные световые поля с интенсивностью более 10^{23} Вт/см², что делает возможным переход к экспериментальному изучению новых явлений фундаментальной физики.

Литература

1. *Андреев, А. А.* Генерация и применение сверхсильных лазерных полей / А. А. Андреев, А. А. Мак, В. Е. Яшин // Квант. электроника. — 1997. — Т. 24. — С. 99—114.
2. *Tajima, T.* Zettawatt-exawatt lasers and their applications in ultrastrong-field physics / T. Tajima, G. Mourou // Phys. Rev. ST Accel. Beams. — 2002. — V. 5, № 3. — P. 031301-1—031301-9.
3. *Danson, C.* Petawatt class lasers worldwide / C. Danson, D. Hillier, N. Hopps, D. Neely // High Power Laser Science and Engineering. — 2015. — V. 3. — P. 1—14.
4. *Мак, А. А.* Лазеры на неодимовом стекле / А. А. Мак, Л. Н. Сомс, В. А. Фромзель, В. Е. Яшин. — М. : Наука, 1990. — 288 с.
5. *Шашкин, В. В.* Оптимизация формы импульса в твердотельных лазерных усилителях с учетом ограничений, накладываемых нелинейными эффектами / В. В. Шашкин, В. Е. Яшин // Известия АН СССР. Сер. Физическая. — 1991. — Т. 55. — С. 303—307.
6. *Carr, C. W.* Effect of temporal pulse shape on optical damage / C. W. Carr, J. B. Trenholme, M. L. Spaeth // Appl. Phys. Lett. — 2007. — V. 90. — P. 041110-1—3.
7. *Stuart, B. C.* Nanosecond-to-femtosecond laser-induced breakdown in dielectrics / B. C. Stuart, M. D. Feit, M. D. Perry, A. M. Rubenchik, B. W. Shore // Phys. Rev. B. — 1996. — V. 53. — P. 1749—1761.
8. *Беспалов, В. И.* О нитевидной структуре пучков света в нелинейных жидкостях / В. И. Беспалов, В. И. Таланов // Письма в ЖЭТФ. — 1966. — Т. 3. — С. 471—476.
9. *Баянов, В. И.* Исследование самофокусировки в лазерных усилителях на неодимовом стекле и ее подавления с помощью пространственной фильтрации // В. И. Баянов, А. А. Мак, В. А. Серебряков, В. Е. Яшин // Квант. электроника. — 1979. — Т. 6. — С. 902—910.
10. *Ешметьева, Е. В.* О предельных энергетических параметрах излучения в лазерных системах на неодимовом стекле / Е. В. Ешметьева, В. И. Королев, Е. П. Меснянкин, В. А. Серебряков, В. В. Шашкин, В. Е. Яшин // Квант. электроника. — 1992. — Т. 19. — С. 837—841.
11. *Мак, А. А.* Твердотельные лазеры с высокой яркостью излучения / А. А. Мак, В. В. Любимов, В. А. Серебряков, В. А. Фромзель, В. Е. Яшин // Известия АН СССР. Сер. Физическая. — 1982. — Т. 46. — С. 1858—1871.
12. *Strickland, D.* Compression of amplified chirped optical pulses // Optics communications. — 1985. — V. 56, № 3. — P. 219—221.
13. *Maine, P.* Generation of ultrahigh peak power pulses by chirped pulse amplification / P. Maine, D. Strickland, P. Bado, M. Pessot, G. Mourou // IEEE J. Quant. Electron. — 1988. — V. 24. — P. 398—403.

14. Баранова, Н. Б. Нелинейные процессы в оптической среде мощных неодимовых лазеров / Н. Б. Баранова, Н. Е. Быковский, Ю. В. Сенатский, С. В. Чекалин // Труды ФИАН. — 1978. — Т. 103. — С. 84—117.
15. Крыжановский, В. И. Формирование пространственной структуры излучения в твердотельных лазерных системах аподизирующими и жесткими апертурами / В. И. Крыжановский, Б. М. Седов, В. А. Серебряков, А. Д. Цветков, В. Е. Яшин // Квант. электроника. — 1983. — Т. 10, вып. 2. — С. 354—359.
16. Власов, С. Н. Стабилизация неустойчивости плоской волны в периодической системе // Письма в ЖТФ. — 1978. — Т. 13, вып. 4. — С. 795—797.
17. Власов, С. Н. Подавление самофокусировки в лазерных системах на неодимовом стекле с помощью оптических ретрансляторов / С. Н. Власов, В. Е. Яшин // Квант. электроника. — 1978. — Т. 8, вып. 3. — С. 510—518.
18. Власов, С. Н. Использование световых пучков с круговой поляризацией для подавления самофокусировочной неустойчивости в нелинейной кубичной среде с ретрансляторами / С. Н. Власов, В. И. Крыжановский, В. Е. Яшин // Квант. электроника. — 1982. — Т. 9, вып. 1. — С. 14—20.
19. Крыжановский, В. И. Применение явления обращения волнового фронта для подавления мелкомасштабной самофокусировки / В. И. Крыжановский, А. А. Мак, В. А. Серебряков, В. Е. Яшин // Письма в ЖТФ. — Т. 7, вып. 5. — С. 400—404.
20. Sartania, S. Generation of 0.1-TW 5-fs optical pulses at a 1-kHz repetition rate / S. Sartania, Z. Cheng, M. Lenzner, G. Tempea, Ch. Spielmann, F. Kraus, K. Ferencz // Opt. Lett. — 1997. — V. 22. — P. 1562—1565.
21. Krupke, W. Ytterbium Solid-state Lasers. The First Decade // IEEE J. Sel. Top. Quant. Electron. — 2000. — V. 6. — P. 1287—1296.
22. Martinez, O. E. 3000 times grating compressor with positive group velocity dispersion : Application to fiber compensation in 1.3–1.6 μm region // IEEE J. Quant. Electron. — 1987. — V. 23. — P. 59—64.
23. Treacy, E. B. Optical pulse compression with diffraction gratings // IEEE J. Quant. Electron. — 1969. — V. 5. — P. 454—458.
24. Cheriaux, G. Aberration-free stretcher design for ultrashort-pulse amplification / G. Cheriaux, P. Rousseau, F. Salin [et al.] // Opt. Lett. — 1996. — V. 21. — P. 414—416.
25. Spielmann, C. Femtosecond solid-state lasers / C. Spielmann, P. F. Curley, T. Brabec, F. Krausz // IEEE J. Quant. Electron. — 1992. — V. 28. — P. 2097—2122.
26. Korf, D. All-in-one dispersion-compensating saturable absorber mirror for compact femtosecond laser sources / D. Korf, F. X. Kartner, K. J. Weingarten, U. Keller // Opt. Lett. — 1995. — V. 21. — P. 468—488.
27. Van'kov, A. 1-TW Nd:glass laser system / A. Van'kov, A. Kozlov, S. Chizhov, V. Yashin // Proc. SPIE. — 1994. — V. 2095. — P. 87—90.
28. Rouyer, C. Generation of 50-TW femtosecond pulses in aTi:sapphire/Nd:glass chain / C. Rouyer, E. Mazataud, I. Allais, A. Pierre, S. Seznec, C. Sauteret, G. Mourou, A. Migus // Opt. Lett. — 1993. — V. 18. — P. 214—216.
29. Rouyer, C. Production and characterization of intensities above $2 \cdot 10^{19}$ W/cm^2 , obtained with 30-TW 300-fs pulses generated in a Ti:sapphire/Nd-doped mixed-glass chain / C. Rouyer, N. Blanchot, I. Allais, E. Mazataud, J. L. Miquel, M. Nail, A. Pierre, C. Sauteret, A. Migus // JOSA B. — 1996. — V. 1. — P. 55—58.

30. *Gogoleva, N. G.* Modeling of chirped-pulse-amplification laser / N. G. Gogoleva, V. A. Gorbunov // Proc. SPIE. — 1996. — V. 2770. — P. 23—30.
31. *Barty, C. P. J.* Regenerative pulse shaping and amplification of ultrabroadband optical pulses / C. P. J. Barty, G. Korn, F. Raksi, C. Rose-Petruck, J. Squier, A. C. Tien, K. R. Wilson, V. V. Yakovlev, K. Yamakawa // Opt. Lett. — 1996. — V. 21. — P. 219—221.
32. *Kim, G. H.* Генератор фемтосекундных импульсов на кристаллах Yb:KYW с подавлением сужения спектра в регенеративном усилителе и спектральным профилированием импульса / G. H. Kim, J. H. Yang, D. S. Lee; В. Е. Яшин, А. В. Кулик, Е. Г. Салль; С. А. Чижов, U. Kang // Оптич. ж-л. — 2013. — Т. 80, вып. 3. — С. 22—29.
33. *Perry, M. D.* Petawatt Laser Pulses / M. D. Perry, D. Pennington, B. C. Stuart, G. Tietzbohl, J. A. Britten, C. Brown, S. Hermann, B. Golick, M. Kartz, J. Miller, H. T. Powell, M. Vergino, V. Yanovsky // Opt. Lett. — 1999. — V. 24. — P. 160—162.
34. *Герке, Р. П.* Голограммная оптика в ГОИ им. С. И. Вавилова / Р. П. Герке, С. Н. Корешев, Г. В. Семенов, В. В. Смирнов // Оптич. ж-л. — 1994. — Т. 61, № 1. — С. 26—39.
35. *Loewen, E.* Diffraction efficiency of echelles working in extremely high orders / E. Loewen, D. Maystre, E. Popov, L. Tsonev // Appl. Optics. — 1995. — V. 34. — P. 1700—1704.
36. *Boyd, R. D.* High-efficiency metallic diffraction gratings for laser applications / R. D. Boyd, J. A. Britten, D. E. Decker, B. W. Shore, B. C. Stuart, M. D. Perry // Appl. Optics. — 1995. — V. 34. — P. 1697—1706.
37. *Yashin, V. E.* Diffraction gratings for lasers with pulse compression : Comparative research / V. E. Yashin, A. A. Mak, L. Bakh, E. Yakovlev, R. Gerke, I. Usupov // Proc. SPIE. — 1998. — V. 3291. — P. 199—201.
38. *Винокурова, В. Д.* Численное и экспериментальное исследование дифракционных решеток для компрессии мощных лазерных импульсов / В. Д. Винокурова, Е. Г. Салль, А. В. Чарухчев, В. Е. Яшин // Оптич. ж-л. — 2005. — Т. 72. — С. 27—31.
39. *Винокурова, В. Д.* Металлизированные голографические дифракционные решетки с повышенной лучевой стойкостью для систем компрессии лазерных импульсов / В. Д. Винокурова, Р. П. Герке, Т. Г. Дубровина, М. Д. Михайлов, Е. Г. Салль, А. В. Чарухчев, В. Е. Яшин // Квант. электроника. — 2005. — Т. 35. — С. 569—572.
40. *Свахин, А. С.* Стойкие к оптическому излучению дифракционные решетки для использования в лазерных резонаторах / А. С. Свахин, В. А. Сычугов, А. Е. Тихомиров // Квант. электроника. — 1994. — Т. 21. — С. 250—252.
41. *Perry, M. D.* High-efficiency multilayer dielectric diffraction gratings / M. D. Perry, R. D. Boyd, J. A. Britten, D. Decker, B. W. Shore, C. Shannon, E. Shults // Opt. Lett. — 1995. — V. 20. — P. 940—942.
42. *Jovanic, C.* Precision damage tests of multilayer dielectric grating for high energy PW lasers / C. Jovanic, G. Brown, B. C. Stuart, W. A. Molander, N. D. Nielsen, B. F. Wattellier, J. A. Britten, D. M. Pennington, C. P. Barty // Proc. SPIE. — 2004. — V. 5647. — P. 34—39.
43. *Nisoli, M.* Compression of high-energy laser pulses below 5 fs / M. Nisoli, S. De Silvestri, O. Svelto, R. Szipocs, K. Ferencz, Ch. Spielmann, S. Sartania, F. Krausz // Opt. Lett. — 1997. — V. 22. — P. 522—525.

44. *Мак, А. А.* О возможности сжатия лазерных импульсов большой энергии в компрессоре на основе квазипериодической системы нелинейных элементов и диспергирующей среды / А. А. Мак, В. Е. Яшин // *Оптика и спектроскопия*. — 1991. — Т. 70. — С. 3—5.
45. *Yashin, V. E.* Compression of high-energy laser pulses after self-phase modulation in a bulk nonlinear medium / V. E. Yashin, A. B. Vankov, A. A. Kozlov, S. A. Chizhov, A. A. Andreev, A. N. Sutyagin // *Proc. SPIE*. — 1997. — V. 3047. — P. 1027—1032.
46. *Высотина, Н. В.* Моделирование систем уширения спектра и нелинейной компрессии петаваттных лазерных импульсов / Н. В. Высотина, Н. Н. Розанов, В. Е. Яшин // *Оптика и спектроскопия*. — 2011. — Т. 110, вып. 6. — С. 1001—1010.
47. *Власов, С. Н.* Уширение спектра и сжатие мощных лазерных импульсов в квазипериодических системах с кубичной нелинейностью // С. Н. Власов, Е. В. Копосова, В. Е. Яшин // *Квант. электроника*. — 2012. — Т. 42, вып. 11. — С. 989—995.
48. *Миронов, С. Ю.* Компрессия фемтосекундных импульсов с гауссовыми временным и пространственным распределениями интенсивности / С. Ю. Миронов, В. В. Ложкарев, Е. А. Хазанов, Ж. Муру // *Квант. электроника*. — 2013. — Т. 43, вып. 8. — С. 711—714.
49. *Гинзбург, В. Н.* Влияние кубической фазы спектра мощных лазерных импульсов на их фазовую самомодуляцию / В. Н. Гинзбург, А. А. Кочетков, И. В. Яковлев, С. Ю. Миронов, А. А. Шайкин, Е. А. Хазанов // *Квант. электроника*. — 2016. — Т. 46, вып. 2. — С. 106—108.
50. *Fedulova, E.* Highly-dispersive mirrors reach new levels of dispersion / E. Fedulova, K. Fritsch, J. Brons, O. Pronin, T. Amotchkina, M. Trubetskov, F. Krausz, V. Pervak // *Opt. Express*. — 2015. — V. 23, № 11. — P. 13788—13793.
51. Список самых мощных лазерных установок [Электронный ресурс] : материал из Википедии — свободной энциклопедии. — Электрон. дан. — Режим доступа: https://ru.wikipedia.org/wiki/Список_самых_мощных_лазерных_установок. — Загл. с экрана. — (Дата обращения: 14.02.17).
52. Exawatt Center for Extreme Light Studies [Электронный ресурс]. — Электрон. дан. — Режим доступа: <http://www.xcels.iapras.ru>. — Загл. с экрана. — Яз. рус., англ. — (Дата обращения: 14.02.17).
53. *Kessler, T.* ICUIL Worldmap // *ICUIL News*. — 2016. — V. 7, № 7. — P. 3—5.

О НЕЛИНЕЙНОЙ ЭЛЕКТРОДИНАМИКЕ ПЛАЗМЫ В ЭКСТРЕМАЛЬНО СИЛЬНОМ СВЕТОВОМ ПОЛЕ

А. В. Башинов, А. В. Ким

Институт прикладной физики РАН, г. Нижний Новгород

Введение

К одной из важных тенденций в развитии современной лазерной физики следует отнести исследование взаимодействия сверхмощного лазерного излучения с веществом. Данная тенденция в первую очередь основывается на прогрессе лазерных технологий последних лет, уже приведших к созданию лазеров петаваттной мощности [1] и к построению в ближайшее время сверхмощных лазерных систем 10-петаваттного класса [2—4], а также к проработке проектов лазера эксаваттной мощности [4, 5]. Сверхмощные лазерные системы можно использовать как при решении прикладных проблем в медицине для онкотерапии, в промышленности для утилизации ядерных отходов, в оборонных технологиях для тестирования работы приборов в экстремальных условиях радиационного фона, так и в фундаментальных исследованиях физики высоких энергий для построения уникальных источников гамма-излучений и достижений экстремальных плотностей высокоэнергичной электрон-позитронной плазмы (см. [6, 7] и цитированную там литературу). Отличительной особенностью взаимодействия лазерного излучения такой экстремально высокой интенсивности с веществом, а по сути с плазмой, является эффективная трансформация лазерной энергии колеблющимися в лазерном поле электронами в энергию гамма-квантов. Данная особенность вводит два принципиально новых момента в описание классической электродинамики бесстолкновительной плазмы. Во-первых, сила отдачи излученного высокоэнергичного фотона модифицирует уравнение движения. Во-вторых, излученные фотоны должны быть самосогласованно учтены в потерях энергии (лазерной) электромагнитной волны. В данной работе будут построены некоторые простейшие нелинейные электродинамические модели лазерно-плазменного взаимодействия, позволяющие, с одной стороны, провести строгий анализ задачи и ввести макроскопические электродинамические характеристики такого взаимодействия, а с другой — представить картину взаимодействия экстремально сильного поля с плазмой в интуитивно наглядной форме, что может оказаться полезным в решении более сложных и практически важных задач, где основным методом остается суперкомпьютерное моделирование.

1. Исходная система уравнений. Классическое приближение

Естественно начать описание взаимодействия экстремально сильного лазерного поля с бесстолкновительной плазмой с наиболее простого слу-

чая классического приближения, когда ускоренный полем электрон излучает в соответствии с классическими законами. Очевидно, что применимость такого приближения нарушается, когда энергия излученного кванта становится соизмеримой с энергией самого электрона, и в этом случае необходим квантовый анализ процесса излучения. Способ обобщения полученных результатов на квантовый случай будет предложен в разделе 3. Для определения исходной системы уравнений электродинамики плазмы будем отталкиваться от общеизвестного факта, что в плазме действующее на одиночный (пробный) заряд поле равно макроскопическому (среднему) электромагнитному полю, которое, в свою очередь, формируется полями всех зарядов, кроме самого пробного заряда. В этом случае в уравнение движения одиночного заряда можно ввести силу реакции излучения [8], а под внешними полями подразумевать макроскопические электромагнитные поля в плазме.

$$\frac{d\vec{P}}{dt} = -e\vec{E} - \frac{e}{c}[\vec{V}, \vec{H}] + \vec{F}_{rr}, \quad (1)$$

где \vec{P} и \vec{V} — импульс и скорость частицы соответственно, \vec{E} и \vec{H} — макроскопические поля в плазме, $-e$ — заряд электрона, c — скорость света в вакууме, \vec{F}_{rr} — сила реакции излучения. Далее, используя естественное предположение, что в релятивистски сильных полях влияние соударений заряженных частиц пренебрежимо мало и, соответственно, температура плазмы много меньше энергии осцилляторного движения, переход к гидродинамическому описанию холодной плазмы можно совершить обычным способом:

$$\frac{\partial \vec{p}}{\partial t} + (\vec{v}\nabla)\vec{p} = \frac{e}{c}\frac{\partial \vec{A}}{\partial t} + e\nabla\varphi - \frac{e}{c}[\vec{v}, \text{rot}\vec{A}] + \vec{F}_{rr}, \quad (2)$$

$$\vec{F}_{rr} = \frac{2e^2\gamma^2}{3c^3}\left\{\frac{d^2\vec{v}}{dt^2} + \frac{3\gamma^2}{c^2}\left(\vec{v}\frac{d\vec{v}}{dt}\right)\frac{d\vec{v}}{dt} + \frac{\gamma^2}{c^2}\left[\vec{v}\frac{d^2\vec{v}}{dt^2} + \frac{3\gamma^2}{c^2}\left(\vec{v}\frac{d\vec{v}}{dt}\right)^2\right]\vec{v}\right\}, \quad (3)$$

$$\frac{\partial N_e}{\partial t} + \text{div}(N_e\vec{v}) = 0, \quad (4)$$

$$\Delta\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = \frac{1}{c}\frac{\partial\nabla\varphi}{\partial t} + \frac{4\pi e N_e\vec{v}}{c}, \quad (5)$$

$$\Delta\varphi = 4\pi e(N_e - N_i), \quad (6)$$

где импульс единицы объема $N_e\vec{p} = \gamma m N_e\vec{v}$, $\gamma = (1 - v^2/c^2)^{1/2}$ — релятивистский фактор, m — масса покоя электрона, сила реакции излучения может быть взята в форме Лоренца — Абрагама — Дирака (ЛАД) или Ландау — Лившица (ЛЛ), но для определенности здесь записана в форме ЛАД. Для простоты анализа не будем учитывать движения ионов как тяжелых частиц и будем считать плотность ионов плазмы неизменной и постоянной $N_i = N_0$. Здесь осуществлен переход к эйлеровым переменным ($\frac{d}{dt} = \frac{\partial}{\partial t} + (\vec{v}\nabla)$) и для удобства описания вместо электрического и магнит-

ного полей введены векторный и скалярный потенциалы, для которых принята кулоновская калибровка $\text{div}\vec{A} = 0$. Данная калибровка удобна для лазерно-плазменных задач потому, что векторный потенциал описывает только вихревые поля в плазме, в то время как скалярный потенциал ответственен за электростатические поля разделения зарядов. Уравнения (2)—(6) представляют собой самосогласованную систему уравнений, описывающую взаимодействие релятивистски сильных электромагнитных волн с плазмой с учетом силы реакции излучения электронов. Далее мы проанализируем данную систему уравнений для нескольких важных случаев, которые позволят, с одной стороны, значительно упростить анализ, а с другой — привести к интуитивно наглядной картине взаимодействия, и даже к практически важным рекомендациям.

2. Квазистационарные модели

Обратимся в первую очередь к волнам с циркулярной поляризацией поля, для которых световое давление на заряженную частицу не имеет осциллирующей (на удвоенной частоте волны) компоненты, что, например, использовалось в схемах ускорения ионов посредством пондеромоторной силы падающего лазерного импульса [9]. Такая постоянно действующая сила может быть в принципе скомпенсирована внешними полями или же силой разделения зарядов в плазме. Это, в частности, позволяет нам перейти к простой квазистационарной модели взаимодействия электромагнитной волны с плазмой, в которой продольные движения электронов подавлены. Полагая, что переменные величины гармонически зависят от времени, рассмотрим далее класс одномерных структур поля и плазмы, которые могут быть результатом взаимодействия экстремально сильного светового поля с плазмой произвольной плотности (не ограниченной изначально условием прозрачности).

Итак, полагая, что $\vec{A}(\vec{r}, t) = \text{Re}[\vec{e}A(z)e^{i\omega t}]$ и, соответственно, $\vec{v}(\vec{r}, t) = \text{Re}[\vec{e}v(z)e^{i\omega t}]$, где $A(z)$ и $v(z)$ — комплексные функции, ω — круговая частота волны, $\vec{e} = (\vec{e}_x + i\vec{e}_y)$, \vec{e}_x и \vec{e}_y — единичные базисные векторы вдоль осей x и y , из уравнения (3) нетрудно получить следующее выражение для величины силы радиационного трения $\vec{F}_{rr} = \text{Re}[F_{rr}(z)(\vec{x} + i\vec{y})e^{i\omega t}]$:

$$F_{rr} = -\frac{2e^2}{3c^3}\gamma^4\omega^2v, \quad (7)$$

которая в отсутствие продольных движений имеет только поперечные компоненты. Для удобства дальнейшего изложения введем следующие безразмерные величины: продольную координату вдоль направления распространения волны $\xi = z\omega/c$; векторный, $a(\xi) = eA/(mc^2)$, и скалярный, $\Phi = e\varphi/(mc^2)$, потенциалы; скорость электронов $v \rightarrow v/c$, нормирован-

ную на скорость света в вакууме; $n_0 = N_0/N_c$ — отношение невозмущенной плотности электронов к критическому значению плотности для данной частоты волны $N_c = m\omega^2/(4\pi e^2)$; плотность электронов $n_e = N_e/N_0$, нормированную на N_0 . Используя выражение для силы радиационного трения (7), уравнение (2) для поперечных компонент можно свести к алгебраическому соотношению

$$a = \gamma(1 - i\delta\gamma^3)v, \quad (8)$$

где введен новый параметр $\delta = 2e^2\omega/(3mc^3)$ ($= 4\pi e^2/(3mc^2\lambda)$), определяемый кроме фундаментальных констант только частотой лазерной волны (длиной волны). Например, для длины волны λ в один микрометр он равен примерно 10^{-8} . Используя выражение для релятивистского фактора $\gamma(|v|)$, точнее его обратное выражение $|v|^2 = 1 - \gamma^{-2}$, с помощью формулы (8) можно определить зависимость γ (энергии) вращающегося в поле волны электрона от ее амплитуды. Соотношение (8) сводится к решению алгебраического уравнения

$$\delta^2\gamma^6(\gamma^2 + 1) + \gamma^2 - (1 + |a|^2) = 0. \quad (9)$$

В области релятивистски сильных полей, но недостаточно сильных для учета сил реакции излучения, $1 \ll |a| \ll \delta^{-1/3} \approx 450$ ($\lambda = 1$ мкм), из (9) имеем общеизвестную зависимость релятивистского фактора от амплитуды волны: $\gamma = \sqrt{1 + |a|^2}$. Однако в области экстремально сильных полей данная зависимость существенно корректируется. Например, в предельном случае, когда $|a| \gg \delta^{-1/3}$,

$$\gamma \simeq (|a|/\delta)^{1/4}, \quad (10)$$

т. е. энергия электрона, равная γmc^2 , увеличивается существенно слабее с ростом интенсивности волны из-за возросшей роли радиационных потерь и, соответственно, радиационной силы трения.

Проектируя уравнение (2) на продольную ось, имеем

$$\frac{1}{2\gamma} \left[\frac{(|a|^2)'}{1+\delta^2\gamma^6} + i\delta\gamma^3 \frac{aa' - a^*a'}{1+\delta^2\gamma^6} \right] = \Phi', \quad (11)$$

где штрих означает дифференцирование по продольной координате ξ . Смысл полученного уравнения легко находится, если принять во внимание, что оно означает равенство сил, действующих в плазме. В левой части собраны силы, действующие на электроны со стороны электромагнитного поля волны, которые, естественно, при отсутствии внешних сил могут быть скомпенсированы лишь силами разделения зарядов в плазме. Последние при выборе кулоновской калибровки полностью описываются скалярным потенциалом, что отражено в правой части уравнения (11). Обратим внимание, что сила со стороны электромагнитной волны состоит из двух слагаемых, первое из которых есть не что иное, как пондеромо-

торная сила, модифицированная с учетом эффекта реакции излучения, снижающего ее дополнительно по сравнению со случаем чисто релятивистской коррекции массы электрона. Появление второго слагаемого, однако, обусловлено только учетом силы реакции излучения, и в этом смысле оно принципиально новое. По аналогии со столкновительной плазмой его появление можно связать с поглощением электромагнитной волны. Действительно, излучательные потери колеблющихся электронов, с которыми и связана сила радиационного трения, должны вести к диссипации поля, а поглощенный импульс электромагнитной волны должен быть передан излучаемым электронам. Таким образом, в неоднородном экстремально сильном поле на заряженные частицы действуют два типа усредненных сил: пондеромоторная сила и сила, связанная с поглощением волны из-за радиационных потерь.

Для интересующего нас квазистационарного случая исходные уравнения для потенциалов (5) и (6) могут быть приведены к виду

$$a'' + \left[1 - \frac{n_0 n_e}{\gamma(1 - i\delta\gamma^3)}\right] a = 0, \quad (12)$$

$$\Phi'' = n_0(n_e - 1). \quad (13)$$

Теперь уравнения (11)—(13) представляют собой самосогласованную систему уравнений, определяющую возможные структуры поля и плазмы в зависимости от граничных условий задачи. Обратим внимание, что для вывода исходной системы уравнений квазистационарной задачи нам не понадобилось уравнение непрерывности, которое необходимо при решении начальной, эволюционной задачи. Потому при решении квазистационарной задачи требуется наложить дополнительное условие, например глобальной квазинейтральности плазмы [10].

2.1. Однородная плазма

В качестве наглядного примера рассмотрим идеализированный случай однородной и безграничной плазмы, электроны которой в текущей системе координат не имеют продольных скоростей. Тогда можно ввести привычные для анализа электродинамические характеристики, например диэлектрическую проницаемость плазмы. Полагая, что $a \propto a_0 e^{ikz}$ (где k — волновое число плоской электромагнитной волны), из уравнения (12) стандартным способом можно получить нелинейное дисперсионное уравнение $k^2 = \omega^2 \epsilon / c^2$ с диэлектрической проницаемостью

$$\epsilon = 1 - \frac{n_0}{\gamma(1 + \delta^2 \gamma^6)} (1 + i\delta\gamma^3), \quad (14)$$

которое наряду с действительной частью содержит также и мнимую часть, ответственную за диссипацию электромагнитной волны, т. е. введение силы трения из-за радиационных потерь самосогласованно ведет к учету

поглощения волны. В отсутствие силы трения ($\delta = 0$) выражение для диэлектрической проницаемости принимает стандартный вид с учетом релятивистской коррекции массы $\varepsilon = 1 - n_0/\sqrt{1 + |a|^2}$. В области сильных полей мнимая часть диэлектрической проницаемости становится соизмеримой с действительной частью или больше ее при амплитудах волн $|a| \geq \delta^{-1/3} \approx 450$ ($\lambda = 1$ мкм). Диэлектрическая проницаемость также позволяет произвести оценку усредненных сил, действующих на единицу объема плазмы со стороны электромагнитного поля ($\vec{E} = \vec{e}E(\vec{r}, t)\exp(i\omega t)$) и связанных как с реальной частью диэлектрической проницаемости (ponderomotorная сила), $f_p = \frac{\text{Re } \varepsilon - 1}{8\pi} \nabla |E|^2$, так и с мнимой частью (поглощение импульса волны), $f_i = -\frac{\text{Im } \varepsilon}{4\pi} \text{Im}(E\nabla E^*)$, которые точно совпадают с соответствующими выражениями в (11) для рассмотренного нами строго продольного случая.

Отметим также, что знание диэлектрической проницаемости дает возможность весьма просто получить выражение для релятивистского томсоновского сечения рассеяния, которое определяется как отношение поглощенной на одну частицу электромагнитной энергии $\sigma |E|^2/N_e$ ($\sigma = \omega \text{Im } \varepsilon / (4\pi)$ — проводимость плазмы) к падающему потоку энергии $c|E|^2/(4\pi)$:

$$\sigma_R = \sigma_T \frac{\gamma^2}{1 + \delta^2 \gamma^6}, \quad (15)$$

где $\sigma_T = 8\pi r_e^2/3$ — томсоновское сечение рассеяния ($r_e = e^2/(mc^2)$ — радиус электрона). Когда сила реакции излучения не существенна, т. е. для амплитуд $|a| \ll \delta^{-1/3}$, мы имеем хорошо известное выражение для томсоновского рассеяния, обобщенное на релятивистский случай. В противоположном пределе экстремально сильных полей, $|a| \gg \delta^{-1/3}$, оно принимает вид $\sigma_R = \sigma_T/(\delta|a|)$.

2.2. Симметричное облучение плазменного слоя

В качестве второго примера рассмотрим симметричное облучение плазменного слоя, которое можно использовать для оценки эффективности трансформации лазерной энергии в энергию гамма-излучения. Для поиска плазменно-полевых структур представим комплексную амплитуду волны в виде $a = ue^{i\theta}$. Тогда уравнение для векторного потенциала (12) распадается на два уравнения для действительных функций $u(\xi)$ и $\theta(\xi)$:

$$u'' - \theta'^2 u + \left[1 - \frac{n_0 n_e}{\gamma(1 + \delta^2 \gamma^6)}\right] u = 0, \quad (16)$$

$$(u^2 \theta')' = \frac{\delta \gamma^2 n_0 n_e}{1 + \delta^2 \gamma^6} u^2. \quad (17)$$

Теперь можно вести поиск решений самосогласованной системы уравнений (11), (13), (16) и (17). Сначала отметим, что аналогичная задача, но без учета сил трения, в которой можно положить $\theta = 0$ ввиду отсутствия поглощений, рассматривалась в работе [10], где и был предложен алгоритм поиска всех возможных решений. Исходя из симметрии задачи, ограничимся поиском плазменно-полевых структур в классе симметричных функций, более того — рассмотрим только однослойные решения, которые всегда реализуются для тонких, меньше длины волны, плазменных слоев. Естественно, в силу симметрии задачи решение удобно начать конструировать с центральной точки $\xi = 0$, в которой $u' = \theta' = \Phi' = 0$, и далее сшивать ее вне слоя с соответствующим вакуумным решением, представляющим собой сумму падающей, отраженной и прошедшей волн.

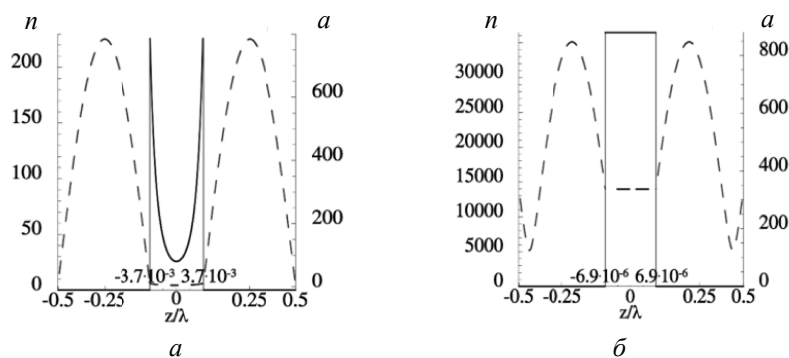


Рис. 1. Распределение вектор-потенциала (штриховая линия) и концентрации электронов (сплошная линия) при обжиге мишени толщиной $0,5\lambda$ и плотностью $n_0 = 500$ падающими циркулярно поляризованными импульсами с интенсивностью $6,6 \cdot 10^{23}$ Вт/см² (а) и 10^{24} Вт/см² (б). Эффективность преобразования лазерной энергии в энергию гамма-излучения 0,0034 % (а) и 44,8 % (б)

На рис. 1 показаны два типичных распределения плотности электронов и поля, которые реализуются при обжиге тонких плазменных слоев. Первое, на что следует обратить внимание, — что электронный слой эффективно сжимается. Достижимые плотности электронов могут значительно, на многие порядки, превосходить твердотельные плотности, что, в свою очередь, приводит к сильной экранировке электромагнитного поля в центральной области (рис. 1, а). Однако при выборе определенной интенсивности падающей волны и / или толщины плазменного слоя возможно достижение и высоких значений полей, и высоких плотностей электронов, при которых может происходить эффективная конверсия лазерного излучения в энергию гамма-квантов (рис. 1, б).

3. Коррекция в квантовом случае

Приведенный выше анализ основывался на классическом представлении процесса излучения электрона, что может быть некорректным для экстремально сильного поля, когда энергия излученных квантов может стать соизмерима с энергией самого электрона. Действительно, для вращающегося высокоэнергичного электрона классические формулы излучения дают завышенные энергии излученных квантов и, соответственно, излученную мощность [11]. Для коррекции в квантовом случае можно воспользоваться тем, что в сильно релятивистском пределе, когда, в общем-то, и необходимо рассмотрение квантового характера эмиссии электрона, сила реакции излучения направлена против мгновенного импульса электрона, т. е. $\vec{F}_{rr} = -\mu\vec{p}$, где μ — скалярный множитель. Тогда для циркулярно поляризованного поля соотношение (8) можно записать в виде

$$a = \gamma(1 - i\mu)v, \quad (18)$$

откуда, используя определение γ , можно получить

$$\gamma = \sqrt{1 + \frac{|a|^2}{1+\mu^2}}. \quad (19)$$

Выражение для диэлектрической проницаемости, а также для усредненных сил со стороны электромагнитного поля (левая часть уравнения (11)) можно легко переписать в виде

$$\varepsilon = 1 - \frac{n_0}{\gamma(1-i\mu)}, \quad (20)$$

$$f = \frac{1}{2\gamma} \left[\frac{(|a|^2)'}{1+\mu^2} + i\mu \frac{aa^* - a^*a'}{1+\mu^2} \right]. \quad (21)$$

3.1. Релятивистский фактор

Для вычисления силы реакции излучения при криволинейном движении электрона воспользуемся выражением для мощности излучения в квантовом случае [12] и тем фактом, что излучаемая мощность с точностью $1/\gamma^2$ совпадает с работой силы реакции излучения. Отсюда следует, что

$$\mu = \frac{\alpha}{3\sqrt{3}\pi\gamma\eta} \int_0^\infty \frac{4u^2+5u+4}{(1+u)^4} u K_{2/3} \left(\frac{2u}{3\chi} \right) du, \quad (22)$$

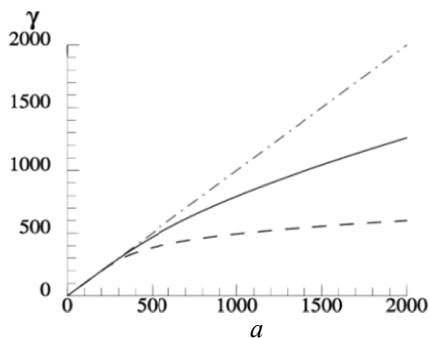
где $\eta = \hbar\omega/(mc^2)$, $K_\nu(x)$ — модифицированная функция Бесселя второго типа порядка ν , χ — квантовый параметр, который можно представить в неявной форме в виде

$$\chi = \eta|a| \sqrt{1 + \frac{|a|^2}{(1+\mu^2)^2}}. \quad (23)$$

Параметр χ определяет важность квантового характера процесса излучения, его физический смысл наиболее легко понять при $\chi \ll 1$, где он пред-

ставляет собой, по сути, отношение средней энергии излученного кванта к энергии самой частицы. В случае $\chi \ll 1$ процесс излучения почти классический, асимптотическое выражение для скалярного множителя имеет вид $\mu \approx \delta\gamma^3$, если амплитуда полей удовлетворяет неравенству $a \ll \eta^{-1}$, что для текущих лазерных частот достаточно хорошо выполняется. При этом анализ на основе классического подхода удовлетворителен.

Рис. 2. Зависимость гамма-фактора от амплитуды поля $|a|$ при учете классической силы реакции излучения (штриховая линия), при учете квантовой коррекции (сплошная линия) и без учета радиационных потерь (штрихпунктирная линия)



Для $\chi \gtrsim 1$ необходим уже квантовый анализ. Заметим, что с учетом (19) и (23) соотношение (22) для фиксированной лазерной частоты, т. е. параметра η , представляет собой интегральное уравнение для одной переменной, например μ , при заданной амплитуде поля $|a|$. Численное решение уравнения (22) позволяет согласно (19) определить зависимость релятивистского фактора γ от амплитуды поля $|a|$. Эта зависимость для длины волны лазерного излучения $\lambda = 0,8$ мкм представлена на рис. 2 (сплошная кривая). На этом же рисунке для сравнения представлены кривые, соответствующие учету классической силы реакции излучения (штриховая кривая) и пренебрежению радиационными потерями (штрихпунктирная кривая). Как видно, в диапазоне полей, когда сила радиационного трения существенно модифицирует движение электрона, необходим учет квантового характера процесса излучения, который может сильно изменить среднюю энергию электрона и, соответственно, спектр излученных гамма-квантов.

3.2. Плазменный слой

Имея представления о диэлектрической проницаемости и силах, действующих на плазму со стороны поля, нетрудно выполнить анализ плазменно-полевых структур в общем случае с учетом квантового характера радиационной силы трения. В частности, проанализируем структуры в плазменном слое, как и в разделе 2.2, при симметричном облучении. Для этого перепишем уравнение (12) с учетом диэлектрической проницаемости (20):

$$a'' + \left[1 - \frac{n_0 n_e}{\gamma(1-i\mu)}\right] a = 0. \quad (24)$$

Это уравнение необходимо рассмотреть уже совместно с уравнением Пуассона (13) и уравнением динамического равенства сил со стороны электромагнитного поля и поля разделения зарядов (11), в котором в левой части следует использовать силу в форме (21). Применяя теперь табулированную форму зависимости $\gamma(|a|)$ (см. рис. 2) и, соответственно, $\mu(|a|)$, можно, представляя поле в виде $a = ue^{i\theta}$, вести поиск самосогласованных распределений поля и электронной плотности так же, как и в классическом случае 2.2. На рисунке 3 представлен пример плазменно-полевой структуры. Из сравнения с рис. 1, б видно, что классическое описание силы реакции излучения ведет к большей степени сжатия электронного слоя и эффективности преобразования лазерной энергии в энергию гамма-излучения ввиду завышения радиационных потерь.

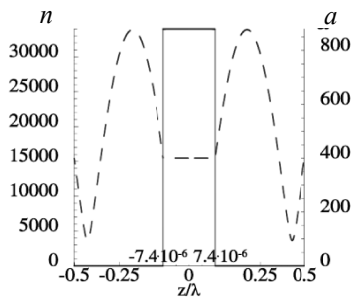


Рис. 3. Распределение вектор-потенциала (штриховая линия) и концентрации электронов (сплошная линия) при обжати мишени с теми же параметрами, что и на рис. 1, падающими импульсами с интенсивностью 10^{24} Вт/см². Эффективность преобразования лазерной энергии в энергию гамма-излучения 35 %

3.3. Самосогласованный слой электрон-позитронной плазмы

Интересным применением рассмотренного выше квазистационарного электродинамического подхода может стать проблема создания плотной электрон-позитронной (e^-e^+) плазмы в экстремально сильных лазерных полях. В полях встречных лазерных импульсов циркулярной поляризации, как известно, возможно рождение (e^-e^+)-плазмы уже при интенсивностях порядка 10^{24} Вт/см² [13], что связано с эффективным формированием так называемого квантового электромагнитного каскада, когда излученные вращающимися электронами (и позитронами) высокоэнергичные фотоны в лазерном поле распадаются, порождая электрон-позитронную пару. В достаточно сильном поле это приводит к быстрому лавинообразному росту (e^-e^+)-плазмы. Естественно, что рост (e^-e^+)-пар ограничивается обратным влиянием плазмы на порождающее их поле, приводя к его снижению в областях лавинообразного роста до некоторого порогового уровня, когда этот рост фактически прекращается (или же заметно замедляется в масштабе длительности лазерного импульса). Таким образом, в данном

случае мы имеем дело с нелинейным электродинамическим механизмом ограничения поля, в котором снижение поля самосогласованно определяется сгенерированной плазмой. Применим электродинамический подход, который использовался раньше для определения плазменно-полевых структур. Для простоты анализа учтем, что на электроны и позитроны при симметричном облучении действуют силы сжатия, а вследствие квазинейтральности (из-за отсутствия полей разделения зарядов) они могут быть сжаты в очень узкий слой (схематически это показано на рис. 4, *a*) — вплоть до ограничений, накладываемых лишь законами квантовой механики, а именно квантовым вырождением плазмы. Будем полагать, что сжатие происходит до толщины слоя много меньше длины лазерной волны. Это позволяет заменить реальное распределение (e^-e^+) -плазмы δ -слоем, что, с одной стороны, существенно упрощает решение электродинамической задачи, а с другой — позволяет определить такой важный параметр, как полное число рожденных пар. Для решения задачи представим $n_0 n_e = N_p \delta(z)/N_c$, где $N_p = \int_{-\infty}^{\infty} N_e dz$ — полное число пар на единицу поперечного сечения, $N_c = m\omega^2/(4\pi e^2)$ — критическое значение плотности для данной частоты волны, и тогда, интегрируя уравнение (24) вдоль слоя и используя условие непрерывности поля, можно связать его производные по обе границы слоя (a'_+ , a'_-) со значением самого поля в слое:

$$a'_+ - a'_- = \frac{2n_p}{\gamma(a_{th})[1-i\mu(a_{th})]} a_{th}, \quad (25)$$

где a_{th} — значение поля в слое, принятое равным пороговому значению, т. е. максимальному значению поля, при котором останавливается ионизация вакуума, а $n_p = N_p \omega/(cN_c)$. Представляя вакуумные поля по обе границы слоя в виде суммы исходных падающих лазерных волн с амплитудой a_i и волн, распространяющихся в противоположных направлениях, из условий (25) и непрерывности полей нетрудно получить соотношение, связывающее полное число рожденных пар на единицу поперечного сечения с амплитудой падающей волны:

$$a_i = \frac{a_{th}}{2} \sqrt{\left(1 + \frac{\mu n_p}{\gamma(1+\mu^2)}\right)^2 + \left(\frac{n_p}{\gamma(1+\mu^2)}\right)^2}. \quad (26)$$

На рисунке 4, *b* показаны результаты расчетов уравнения (26) для двух случаев: квантового и классического, в которых в качестве порогового значения принята интенсивность $5 \cdot 10^{23}$ Вт/см². Однако отметим, что для инициации каскада затравочными электронами (или гамма-квантами) требуются большие интенсивности ($\approx 10^{24}$ Вт/см² [13]), чем для его поддержания или же для его формирования с помощью плотной мишени. Пороговое значение для непрерывной волны определяется из условия равенства

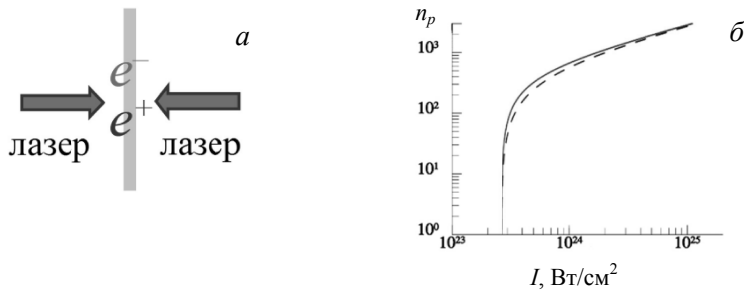


Рис. 4. Схематическое изображение предельно сжатого электрон-позитронного слоя под действием падающих циркулярно поляризованных лазерных импульсов (*а*) и число рожденных электрон-позитронных пар при классическом (штриховая линия) и квантовом (сплошная линия) подходе к описанию силы реакции излучения в сжатом циркулярно поляризованными лазерными импульсами δ -слое в зависимости от их интенсивности (*б*)

скоростей рождения пар и потерь частиц (их ухода из области сильного поля), и данная задача здесь не рассматривается. Принятое значение соответствует интенсивности, при которой скорость роста пар ($6 \cdot 10^{-3}$ обратных периодов поля [14]) порядка обратного времени действия лазерного импульса длительностью 17 периодов поля при длине волны 0,8 мкм. Для квантовых расчетов использовались выражения (19) и (22), описывающие γ и μ соответственно, в то время как для классических расчетов $\mu = \delta\gamma^3$, а γ — результат решения алгебраического уравнения (9). Обратим внимание: квантовый расчет дает большее количество рожденных пар, что может быть объяснено переоценкой поглощенной энергии в классическом случае. Действительно, для фиксированного поля и заданной плотности заряженных частиц величина мнимой части диэлектрической проницаемости при классическом описании больше, чем в квантовом случае, что при заданной амплитуде падающей волны дает меньшее значение поля в слое и тем самым порождает меньшее число пар.

4. Гамма-источник

Проведенный в разделах 2.2 и 3.2 анализ плазменно-полевых структур при симметричном облучении слоя плазмы дает возможность построить эффективный источник высокоэнергичных фотонов.

В силу того что под действием встречных лазерных импульсов может формироваться очень тонкий слой электронов, можно ожидать формирования узкой диаграммы направленности гамма-излучения. В плоскости этого слоя под действием вращающегося электрического поля электроны синхронно вращаются на частоте лазерных импульсов и генерируют синхротронное излучение. Как известно [12], в этом случае фотоны излучают в ко-

нус вдоль мгновенного импульса электрона с углом раскрытия порядка γ^{-1} , причем значение гамма-фактора определяется локальной амплитудой поля. Такой источник может быть получен, например, при облучении твердотельной мишени, выполненной в виде фольги. Подбирая соответствующим образом толщину фольги, можно добиться, с одной стороны, предельного сжатия электронного слоя, а с другой — достижения в нем достаточно сильного поля, позволяющего генерировать высокоэнергичные гамма-кванты. Схематическая реализация такого источника представлена в статье [15]. По методике, изложенной в разделах 2.2 и 3.2, были проведены серии расчетов с целью оптимизации параметров мишени и падающих лазерных волн для получения наибольшего выхода гамма-квантов.

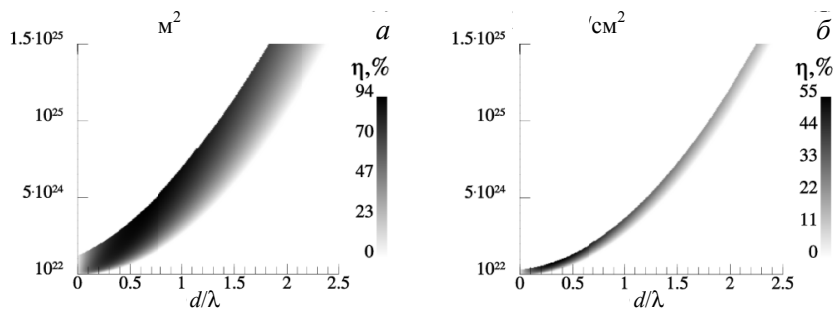


Рис. 5. Эффективность преобразования лазерной энергии в энергию гамма-излучения в зависимости от интенсивности встречных циркулярно поляризованных лазерных импульсов и толщины плазменной мишени при классическом (а) и квантовом (б) подходах к описанию силы реакции излучения. Графики соответствуют длине волны 0,8 мкм

На рисунке 5 представлены результаты этих расчетов для классического и квантового подходов к описанию силы реакции излучения. Как видно, при классическом подходе в оптимальном варианте возможна конверсия лазерной энергии с максимальной эффективностью 94 % в энергию гамма-квантов, в то время как наиболее полный квантовый анализ указывает на максимальную эффективность на уровне 50 %.

5. Численное моделирование

В качестве подтверждения реализуемости рассмотренного нами квазистационарного подхода приведем результаты одномерного численного моделирования методом частиц в ячейках (PIC-код) задачи симметричного облучения тонкого плазменного слоя падающими лазерными импульсами. Вопросы практического осуществления такой постановки задачи, несомненно, останутся открытыми; для их анализа необходимо провести трехмерное суперкомпьютерное моделирование. Тем не менее одномерное

моделирование дает возможность рассмотреть вопросы принципиального характера и проследить за качественными отличиями в динамике взаимодействия. Оно также позволяет выявить предельные характеристики, получаемые в результате такого взаимодействия, которые могут оказаться полезными при решении значительно более сложных реальных задач. Мы рассмотрим два случая с отличными друг от друга типами поляризаций лазерного излучения: циркулярной и линейной. Динамика взаимодействий для этих двух случаев имеет качественные отличия. Для циркулярно поляризованных волн непрерывно действующее давление со стороны падающих импульсов способно привести к сильному сжатию электронной компоненты плазмы, несмотря на то что ионы, как тяжелые частицы, могут распределиться достаточно широким фоном. Для линейно поляризованных волн из-за действия магнитного поля волны и соответствующих осцилляций силы Лоренца на удвоенной частоте электронная динамика выглядит намного сложнее, она нестационарная и в итоге не приводит к заметной компрессии электронов. На рисунке 6 показаны результаты численных расчетов взаимодействия плазменной мишени плотностью $8,7 \cdot 10^{23} \text{ см}^{-3}$ и толщиной $0,28 \text{ мкм}$ с падающими лазерными волнами (длина волны $0,8 \text{ мкм}$) циркулярной поляризации ($a, в, д$) с амплитудой $a_0 = 300$ и линейной поляризации ($б, з, е$) с амплитудой $a_0 = 420$.

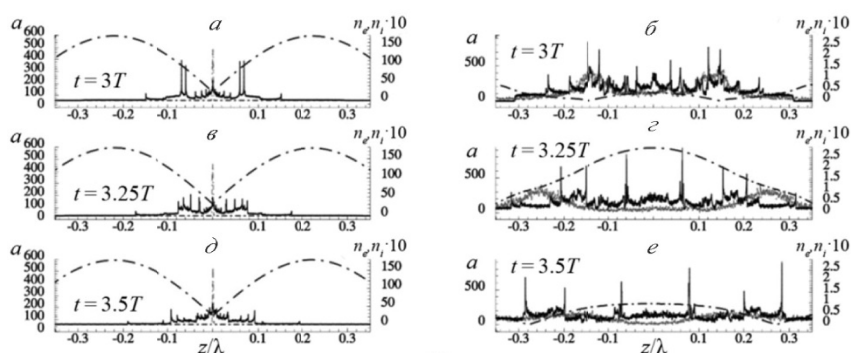


Рис. 6. Динамика симметричного облучения плазменной мишени падающими волнами с циркулярной (a), ($в$), ($д$) и линейной ($б$), ($з$), ($е$) поляризациями поля: штрихпунктирная линия соответствует распределению поля, штриховая — электронной плотности, сплошная — распределению ионов. T обозначает период поля. Плотность ионов нормирована на N_0

В случае циркулярно поляризованных волн происходит непрерывное, планомерное сжатие электронов мишени, в то время как ионы динамически распределяются в большей области. Как видно из графиков, распределение электронов практически стационарно, что не подвергает сомнению

правомерность использования квазистационарного подхода. В случае же линейно поляризованных волн взаимодействие качественно иное, динамика принципиально нестационарная, и, конечно, квазистационарный подход не может быть применен.

Заключение

В данной работе мы попытались взглянуть на весьма сложную проблему лазерно-плазменного взаимодействия в экстремально сильном поле, когда важную роль начинают играть процессы излучения высокоэнергичных фотонов, с позиций более простых квазистационарных моделей, но объединенных нелинейным электродинамическим подходом. В отличие от «обычной» классической электродинамики плазмы здесь принципиальным для динамики электронов в лазерном поле является учет силы радиационного трения. Более того, оказалось возможным учесть самосогласованным образом квантовые особенности процесса излучения фотонов. Рассмотренные электродинамические модели позволили описать в бесстолкновительной плазме новый механизм потерь лазерной энергии, связанный с излучением заряженными частицами высокоэнергичных фотонов.

Авторы выражают признательность Российскому научному фонду (РНФ) за поддержку настоящей работы (проект № 16-12-10486).

Литература

1. *Danson, C.* Petawatt class lasers worldwide / C. Danson, D. Hiller, N. Hopps, D. Neely // High Power Laser Sci. Eng. — 2015. — V. 3. — P. e3-1—e3-14.
2. The Vulcan 10 Petawatt Project [Электронный ресурс] // Central Laser Facility / Science and Technology Facilities Council. — Swindon, UK, 2010. — Режим доступа: <https://www.clf.stfc.ac.uk/Pages/The-Vulcan-10-Petawatt-Project.aspx/>. — Загл. с экрана. — (Дата обращения: 09.02.2017).
3. Apollon 10 PW facility [Электронный ресурс] // Laboratoire Charles Fabry / Institut d'Optique. — France, 2017. — Режим доступа: <https://www.lcf.institutoptique.fr/Groupes-de-recherche/Lasers/Thematiques/Projet-Apollon-10-PW/>. — Загл. с экрана. — (Дата обращения: 09.02.2017).
4. ELI Beamlines [Электронный ресурс] / ELI Beamlines research center. — Czech Republic, 2017. — Режим доступа: <http://www.eli-beams.eu/>. — Загл. с экрана. — (Дата обращения: 09.02.2017).
5. Exawatt Center for Extreme Light Studies (XCELS) [Электронный ресурс] / Институт прикладной физики РАН. — Нижний Новгород, 2016. — Режим доступа: <http://www.xcels.iapras.ru/>. — Загл. с экрана. — Яз. рус., англ. — (Дата обращения: 09.02.2017).
6. *Mourou, G.* Optics in the relativistic regime / G. Mourou, T. Tajima, S. Bulanov // Rev. Mod. Phys. — 2006. — V. 78. — P. 309–371.
7. *Mourou, G.* The future is fiber accelerators / G. Mourou, B. Brocklesby, T. Tajima, J. Limpert // Nat. Photon. — 2013. — V. 7. — P. 258–261.

8. *Ландау, Л. Д.* Теория поля / Л. Д. Ландау, Е. М. Лифшиц. — М. : Физматлит, 2006. — 536 с.
9. *Macchi, A.* Laser Acceleration of Ion Bunches at the Front Surface of Overdense Plasmas / A. Macchi, F. Cattani, T. V. Liseykina, F. Cornolti // Phys. Rev. Lett. — 2005. — V. 94, № 16. — P. 165003-1—165003-4.
10. *Kim, A. V.* New Regime of Anomalous Penetration of Relativistically Strong Laser Radiation into an Overdense Plasma / A. Kim, F. Cattani, D. Anderson, M. Lisak // Письма в ЖЭТФ. — 2000. — Т. 72, вып. 5. — С. 355—360.
11. *Берестецкий, В. Б.* Квантовая электродинамика / В. Б. Берестецкий, Е. М. Лифшиц, Л. П. Питаевский. — М. : Физматлит, 2006. — 720 с.
12. *Байер, В. Н.* Излучение релятивистских электронов / В. Н. Байер, В. М. Катков, В. С. Фадин. — М. : Атомиздат, 1973. — 376 с.
13. *Bell, A. R.* Possibility of prolific pair production with high-power lasers / A. R. Bell, J. G. Kirk // Phys. Rev. Lett. — 2008. — V. 101, № 20. — P. 200403-1—200403-4.
14. *Grismayer, T.* Laser absorption via quantum electrodynamics cascades in counter propagating laser pulses / T. Grismayer, M. Vranic, J. L. Martin, R. A. Fonseca, L. O. Silva // Phys. Plasmas. — 2016. — V. 23, № 5. — P. 056706-1—056706-10.
15. *Bashinov, A. V.* On the electrodynamic model of ultra-relativistic laser-plasma interactions caused by radiation reaction effects / A. V. Bashinov, A. V. Kim // Phys. Plasmas. — 2013. — V. 20, № 11. — P. 113111-1—113111-6.

ЛАЗЕРНОЕ УСКОРЕНИЕ ИОННЫХ ПУЧКОВ КАК УНИКАЛЬНЫЙ ИНСТРУМЕНТ ДЛЯ ФИЗИКИ ВЫСОКИХ ПЛОТНОСТЕЙ ЭНЕРГИИ

А. В. Коржиманов

Институт прикладной физики РАН, г. Нижний Новгород

Физика высоких плотностей энергии — относительно молодой и активно развивающийся раздел науки. Предметом его исследования является вещество при сверхвысоких давлениях. Формально системами с высокой плотностью энергии называют системы, находящиеся при давлении выше 1 Мбар, или, другими словами, 10^{11} Па, что эквивалентно плотности энергии 10^{11} Дж/м³. С физической точки зрения эта величина равна отношению потенциальной энергии электрона в основном состоянии атома водорода к условному объёму этого атома. То есть при давлении порядка 1 Мбар никакое вещество не может считаться несжимаемым, а его атомы ионизируются, образуя плазму при любой плотности вещества [1].

Состояние вещества с высокой плотностью энергии зачастую не может быть описано ни в рамках физики плазмы, ни в рамках физики конденсированных сред: в отличие от плазменных сред требуется учёт квантовой коррелированности частиц, а в отличие от конденсированных сред важны также процессы ионизации и учёт дальнедействующего кулоновского взаимодействия. Это серьёзно затрудняет теоретическое описание и требует построения полуэмпирических моделей, основанных на данных эксперимента. С другой стороны, создание подобных условий в экспериментах является нетривиальной задачей, поскольку достичь столь высоких давлений непросто.

В то же время понимание того, как вещество ведёт себя при сверхвысоких давлениях, обладает как фундаментальным, так и прикладным значением. С одной стороны, взаимодействие релятивистских, квантовых и плазменных эффектов приводит к возникновению необычной и сложной динамики, в том числе с возможностью наблюдать самоорганизацию в подобных системах. С другой стороны, высокие плотности энергии достигаются в недрах планет и звёзд, а также при осуществлении управляемого термоядерного синтеза. Таким образом, знание законов, которым подчиняется вещество при сверхвысоких давлениях, требуется в астрофизике и при проектировании термоядерных реакторов.

На данный момент существует довольно много лабораторных способов создания вещества с высокой плотностью энергии. Все они обладают теми или иными преимуществами и недостатками, а также диапазоном достижимых значений температуры вещества и его плотности. Для наиболее всестороннего изучения всех возможных свойств желательно иметь возможность создавать вещество с произвольными температурой и плотностью. Однако это невозможно сделать каким-то одним методом, поэто-

му исследования в физике высоких плотностей энергии носят комплексный характер и задействуют экспериментальные возможности многих научных групп.

Одним из первых и наиболее распространённых методов создания вещества с высокой плотностью энергии стали установки Z-пинча. При приложении к тонкой проволочке достаточно большого внешнего напряжения протекающий через неё ток вызывает радиальное сжатие вещества благодаря силам Ампера. Современные системы способны выдавать импульсы напряжения величиной более 1 МВ и длительностью порядка сотен наносекунд. Например, крупнейшая из них — Z Pulsed Power Facility, расположенная в одной из Сандийских национальных лабораторий США, — способна создавать импульс тока величиной более 25 МА и длительностью около 100 нс, выделяемая мощность при этом составляет сотни тераватт. Такие машины способны создавать вещество с плотностью чуть выше твердотельной и температурой до нескольких сотен тысяч кельвин. Недостатком Z-пинчей при их использовании для изучения свойств вещества является сильная неоднородность мишени как с точки зрения плотности, так и с точки зрения температуры. Размеры мишени не позволяют проводить локальные измерения, а интегральные характеристики оказываются сильно усреднёнными.

Альтернативой Z-пинчам, работающей практически в том же диапазоне давлений и энергий, являются плоские ударные волны. Такие волны могут инициироваться в веществе относительно короткими мощными лазерными импульсами. Типичный пример — импульсы с длительностью несколько наносекунд и с энергией от нескольких кило-джоулей до нескольких мегаджоулей. В результате облучения поверхности мишени начинается абляция вещества, которая вызывает резкое повышение давления в веществе (внутри мишени) и приводит к возникновению ударной волны. Волна, в свою очередь, распространяясь вглубь мишени, приводит к некоторому уплотнению вещества и к его заметному нагреву [2]. Очевидные недостатки метода — это, во-первых, малость области параметров, которую можно исследовать в данном веществе (эта область фактически определяется адиабатой Гюгонно), а во-вторых, малые пространственные масштабы ударной волны и её довольно быстрое движение, что осложняет диагностику вещества.

Усложнённым вариантом метода плоских волн является сферическое сжатие вещества. В этом случае запуск ударных волн в сферических мишенях производится с помощью как можно более равномерного облучения их поверхности мощным излучением. В качестве излучения может использоваться как непосредственно лазерное излучение, так и рентгеновское излучение, получаемое в результате конверсии лазерного излучения на внутренней поверхности так называемого хольраума, в который помещается мишень. В последнем случае увеличивается однородность нагрева

и, как следствие, устойчивость ударных волн относительно азимутальных возмущений. Сферическое сжатие позволяет достичь значительно более высокой степени сжатия вещества (фактически в 10—100 раз выше твердотельной плотности) и более высоких температур (вплоть до нескольких десятков миллионов кельвин) [3]. К сожалению, сложность экспериментальной установки не позволяет широко использовать данный метод.

Отметим также, что высокие плотности энергии достигаются и в плазме, возникающей в результате абляции вещества мощным лазерным излучением. Эта плазма, однако, имеет плотность ниже твердотельной, но зато её температура может составлять миллионы кельвин [4]. При использовании лазерных импульсов фемтосекундной длительности и петаваттной мощности возможно создание релятивистской лазерной плазмы, в которой температура электронов достигает релятивистских значений — порядка ста миллионов кельвин и выше [5]. Плотность такой плазмы, однако, обычно невелика, а её размеры микроскопичны.

Наконец, наиболее универсальным методом является использование для нагрева глубоко проникающих в вещество пучков энергичных ионов [6]. Характерная длина торможения тяжёлого иона с энергией порядка 1 ГэВ/нуклон в веществе составляет десятки и сотни сантиметров в зависимости от типа вещества. При этом энерговыделение, связанное в основном с процессом ионизации вещества, относительно равномерно распределяется вдоль всей длины за исключением последних миллиметров, где наблюдается так называемый брэгговский пик. Современные ускорители способны выдавать пучки необходимой мощности (около 1 ГВт) малой длительности (порядка сотен наносекунд) и диаметром около 1 мм. Таким образом, при использовании тонких мишеней — толщиной около 1 см — возможно создание макроскопических объёмов практически равномерно прогретого вещества, не успевающего за время нагрева испытать заметное гидродинамическое расширение. Такая ситуация идеальна для проведения измерений.

При необходимости ионный пучок можно профилировать в поперечном направлении для нагрева только цилиндрического слоя, который, расширяясь в дальнейшем из-за цилиндрического сжатия, создаст уплотнённое вещество на оси цилиндра. При этом, однако, теряются преимущества однородного прогрева.

Современные источники тяжёлых ионов не способны создать вещество с температурой выше нескольких сотен тысяч кельвин. Альтернативой им могут выступать источники, основанные на лазерном ускорении ионов [7]. В данной статье мы хотим показать, что они способны создавать вещество с параметрами, не достижимыми другими методами.

Лазерные источники ионных пучков обладают, однако, некоторыми недостатками. Во-первых, на данный момент не существует эффективных методов получения моноэнергетических пучков. Во-вторых, энергия ионов в пучке относительно мала — в лучшем случае порядка

100 МэВ/нуклон. Это приводит к невозможности однородного нагрева вещества. Ещё одна проблема связана с малыми размерами пучка и малой длиной торможения (для указанных энергий она не превышает миллиметра). В результате нагрев осуществляется только в малом объёме, исследование которого представляет определённые сложности.

Несмотря на это, уникальность достигаемых характеристик вызывает определённый интерес. Давайте оценим, какие плотности энергии могут быть получены этим методом. Лазерный импульс, необходимый для ускорения ионов до 100 МэВ/нуклон, должен иметь мощность порядка 1 ПВт. Лазерные системы подобной мощности делятся на пикосекундные (с длительностью импульса в сотни фемтосекунд) и фемтосекундные (с длительностью импульса в десятки фемтосекунд). Для наших целей важна не только большая мощность, но и как можно большая энергия импульса, поэтому рассмотрим пикосекундные лазерные системы. Примерами могут служить лазерные системы NOVA (Ливерморская национальная лаборатория, США, в настоящее время разобрана, длительность импульса 440 фс, энергия — до 660 Дж), Vulcan (Резерфордская лаборатория, Великобритания, длительность импульса 500 фс, энергия — до 500 Дж), RHILIX (Институт тяжёлых ионов, Германия, длительность импульса 400 фс, энергия — до 400 Дж), OMEGA EP (Рочестерский университет, США, длительность импульса 1—10 пс, энергия 1—2,6 кДж), LFEX (Осакий университет, Япония, длительность импульса 1000 фс, энергия — до 1000 Дж). Известны также планы по строительству более мощных систем с близкой длительностью импульса [8]. Упомянем также нижегородский проект XCELS, поддержанный правительством России в рамках программы проектов класса мегасайенс. Его расчётные характеристики — длительность импульса 20 фс, энергия 4 кДж.

Таким образом, для своей оценки мы можем принять, что лазерный импульс имеет энергию 1 кДж и длительность около 1000 фс. Диаметр пучка на мишени в подобных системах может составлять величину порядка 10 мкм. Эффективность конверсии энергии лазерного импульса в энергию ионов в известных схемах составляет величину порядка 1—10 %, а их теоретически рассчитанная энергия — порядка 100 МэВ/нуклон. Несложные вычисления показывают, что характерный пучок ионов будет иметь порядка 10^{10} частиц и интегральную по длине плотность потока энергии 10^7 Дж/см². Характерная длина торможения в веществе с высоким зарядом ядер, например в золоте, будет составлять величину порядка 1 мм, а плотность энерговыделения — 10^{14} Дж/м³. Для золота это эквивалентно 10 МДж/г, в то время как, например, на тяжёлоионном ускорителе FAIR, строящемся в Германии, рассчитывают получить рекордные для традиционных ускорителей 300 кДж/г [9]. Поскольку длительность ионного пучка, сгенерированного лазерным импульсом, составляет величину порядка длительности лазерного импульса, т. е. 1 пс, то нагрев происходит чрез-

вычайно быстро и может считаться изохорическим. Таким образом, плотность нагреваемого вещества не изменяется, оставаясь твердотельной, а температура может достигать миллионов кельвин. Такие условия, как мы указывали выше, другими методами не достижимы.

Отметим, что рассмотрение задачи о торможении столь плотных ионных пучков в веществе представляет собой новую фундаментальную задачу. Дело в том, что энерговыделение на уровне 10 МДж/г эквивалентно 10 кэВ/атом, что неизбежно ведёт к многократной ионизации атомов и существенному нагреву электронов уже в процессе взаимодействия ионного пучка с веществом. Обычно рассматривается задача о торможении ионов или в холодном веществе, или в ионизированном и горячем, но созданном неким внешним источником [10]. Энергия, приходящаяся в этом случае на один атом, оказывается малой, что позволяет считать, что взаимодействие происходит с неизменяемой средой. В случае сверхплотных ионных пучков изменением свойств тормозящей среды пренебречь уже невозможно, и требуется решать самосогласованную нелинейную задачу. Насколько нам известно, подобная задача на данный момент систематически никем не рассматривалась.

На это обстоятельство было обращено внимание в недавней работе [11], и было показано, что при превышении плотностью ионного тока определённой величины возникают нелинейные эффекты, заключающиеся, в частности, в более глубоком проникновении ионов в вещество, а также в развитии филаментационных неустойчивостей, разрушающих пучок и вызывающих сильно неоднородный нагрев вещества. Эти задачи, однако, всё ещё требуют более детального изучения.

Литература

1. *Drake, R. P.* High-Energy-Density Physics : Fundamentals, Inertial Fusion, and Experimental Astrophysics / R. P. Drake. — Berlin ; Heidelberg : Springer-Verlag, 2006. — 534 p.
2. *Millot, M.* Shock compression of stishovite and melting of silica at planetary interior conditions / M. Millot, N. Dubrovinskaia, A. Černok, S. Blaha, L. Dubrovinsky, D. G. Braun, P. M. Celliers, G. W. Collins, J. H. Eggert, R. Jeanloz // Science. — 2015. — V. 347, № 6220. — P. 418—420.
3. *Nora, R.* Gigabar Spherical Shock Generation on the OMEGA Laser / R. Nora, W. Theobald, R. Betti, F. J. Marshall, D. T. Michel, W. Seka, B. Yaakobi, M. Lafon, C. Stoeckl, J. Delettrez, A. A. Solodov, A. Casner, C. Reverdin, X. Ribeyre, A. Vallet, J. Peebles, F. N. Beg, M. S. Wei // Phys. Rev. Lett. — 2015. — V. 114. — Art. 045001.
4. *Гаранин, С. Г.* Мощные лазеры и их применение в исследованиях физики высоких плотностей энергии // УФН. — 2011. — Т. 181. — С. 434—441.
5. *Purvis, M. A.* Relativistic plasma nanophotonics for ultrahigh energy density physics / M. A. Purvis, V. N. Shlyaptsev, R. Hollinger, C. Bargsten, A. Pukhov, A. Prieto, Y. Wang, B. M. Luther, L. Yin, S. Wang, J. J. Rocca // Nature Photonics. — 2013. — V. 7. — P. 796—800.

6. *Фортвов, В. Е.* Интенсивные ионные пучки для генерации экстремальных состояний вещества / В. Е. Фортвов, Д. Хоффманн, Б. Ю. Шарков // УФН. — 2008. — Т. 178. — С. 113—138.
7. *Macchi, A.* Ion acceleration by superintense laser-plasma interaction / A. Macchi, M. Borghesi, M. Passoni // Rev. Mod. Phys. — 2013. — V. 85. — P. 751.
8. *Коржиманов, А. В.* Горизонты петаваттных лазерных комплексов / А. В. Коржиманов, А. А. Гоносков, Е. А. Хазанов, А. М. Сергеев // УФН. — 2011. — Т. 181. — С. 9—32.
9. *Tahir, N. A.* Studies of high-energy density states using isochoric heating of matter by intense heavy ion beams: the HEDgeHOB Collaboration / N. A. Tahir, P. Spiller, A. R. Piriz, A. Shutov, I. V. Lomonosov, M. Schollmeier, A. Pelka, D. H. H. Hoffmann, C. Deutsch // Physica Scripta. — 2008. — V. T132. — Art. 014023.
10. *Deutsch, C.* Ion Stopping in Dense Plasma Target for High Energy Density Physics / C. Deutsch, G. Maynard, M. Chabot, D. Gardes, S. Della-Negra, R. Bimbot, M. F. Rivet, C. Fleurier, C. Couillaud, D. H. H. Hoffmann, H. Wahl, K. Weyrich, O. N. Rosme J., N. A. Tahir, J. Jacoby, M. Ogawa, Y. Oguri, J. Hasegawa, B. Sharkov, A. Golubev, A. Fertman, V. E. Fortov, V. Mintsev // Open Plasma Phys. J. — 2010. — V. 3. — P. 88—115.
11. *Kim, J.* Self-Consistent Simulation of Transport and Energy Deposition of Intense Laser-Accelerated Proton Beams in Solid-Density Matter / J. Kim, B. Qiao, C. McGuffey, M. S. Wei, P. E. Grabowski, F. N. Beg // Phys. Rev. Lett. — 2015. — V. 115. — Art. 054801.

ГЕНЕРАЦИЯ ВЫСОКИХ ГАРМОНИК В ГАЗАХ ЛАЗЕРНЫМИ ИСТОЧНИКАМИ СРЕДНЕГО ИК-ДИАПАЗОНА: ВЛИЯНИЕ МАГНИТНОГО ПОЛЯ ИМПУЛЬСА НА ФОРМУ СПЕКТРА ГАРМОНИК

А. С. Емелина^{1,2}, М. Ю. Емелин^{1,2}, М. Ю. Рябкин^{1,2}

¹ Институт прикладной физики РАН, г. Нижний Новгород

² Институт общей физики РАН, г. Москва

Введение

Достигнутый в последнее время значительный прогресс в создании мощных фемтосекундных параметрических лазерных источников излучения ближнего и среднего ИК-диапазонов (с длинами волн λ в диапазоне от 1,3 до 4 мкм [1—3]) стимулировал активные исследования в области взаимодействия такого излучения с веществом. Одной из областей исследований, где эти новые возможности проявляются особенно ярко [4—9], является физика процессов, связанных с перерассеянием электронов [10—11] при туннельной ионизации. К числу таких процессов относится генерация гармоник высокого порядка (ГГВП) лазерного излучения [12—13].

Важной характеристикой различных процессов, связанных с ионизацией газов интенсивным лазерным излучением, является средняя энергия осциллирующей электрона в переменном электрическом поле с амплитудой E и частотой ω

$$U_p = e^2 E^2 / 4m\omega^2 \sim I\lambda^2, \quad (1)$$

где e и m — заряд и масса электрона, I и λ — соответственно интенсивность и длина волны лазерного излучения. Величина U_p , в частности, определяет высокоэнергетическую границу платообразного распределения в спектре ГГВП в газах; положение этой границы определяется универсальным выражением [11, 14]

$$\hbar\Omega_{\max} \approx I_p + 3,17 U_p, \quad (2)$$

где I_p — потенциал ионизации атома, \hbar — постоянная Планка.

Пропорциональность осцилляторной энергии электрона U_p квадрату длины волны лазерного излучения дает принципиальную возможность значительного расширения плато в спектре ГГВП в высокочастотную область [4, 15] при использовании источников среднего ИК-диапазона. При этом энергия эффективно генерируемых фотонов может достигать нескольких килоэлектрон-вольт и даже десятков килоэлектрон-вольт [16, 17], что практически недостижимо при использовании лазерных источников с длинами волн менее 1 мкм. Яркой демонстрацией такой возможности стал недавний эксперимент [5], где был получен когерентный широкополосный суперконтинуум с максимальной энергией фотонов, дости-

гающей 1,6 кэВ при ионизации атомов гелия излучением источника [2] с центральной длиной волны 3,9 мкм.

Особенностью ГГВП лазерного излучения среднего ИК-диапазона является также специфическая форма спектра генерируемого излучения, характеризующаяся резким подъемом интенсивности на самом краю плато. Эта особенность может быть использована для формирования аттосекундных биений при ионизации газа на заднем фронте предельно короткого лазерного импульса [9]. Характерный временной масштаб этих биений может составлять единицы аттосекунд, а при оптимальных параметрах лазерного излучения может быть менее 1 ас.

Однако с ростом длины волны лазерного излучения возрастает роль недипольных эффектов. Магнитная составляющая силы Лоренца приводит к отклонению («магнитному дрейфу») траектории электрона от прямолинейной при его движении в свободном пространстве. В результате снижаются вероятность рекомбинации электрона с родительским ионом и, соответственно, эффективность преобразования лазерного излучения в рентгеновский диапазон [18—20]. При взаимодействии излучения видимого или ближнего ИК-диапазонов с нейтральными атомами этот фактор несущественен, так как скорость свободного электрона при тех интенсивностях излучения, при которых происходит ионизация атомов, в этих случаях значительно меньше скорости света. Однако, в силу приведенного выше скейлинга (1) осцилляторной энергии электрона, с ростом λ влияние магнитной составляющей силы Лоренца возрастает. Поэтому при генерации высоких гармоник источниками среднего ИК-диапазона даже умеренной интенсивности роль недипольных эффектов обычно весьма существенна.

В данной работе показано, что магнитный дрейф электрона может приводить не только к общему снижению эффективности генерации высоких гармоник, но и к изменению формы спектра генерируемого излучения. Также продемонстрировано, что магнитный дрейф электрона не всегда играет отрицательную роль, а в ряде случаев дает положительный эффект.

1. Теоретическое описание

В основе используемого в настоящей статье теоретического подхода лежит предложенная нами в работе [17] модификация развитой М. Левенштейном с соавторами аналитической теории ГГВП [21].

Теория Левенштейна была модифицирована таким образом, чтобы в ее рамках стало возможным одновременное описание эффектов опустошения атомарных уровней и магнитного дрейфа электрона. В рамках настоящей статьи опустошение атомарных уровней не будет играть существенной роли в силу использования умеренных интенсивностей лазерных импульсов. Поэтому общее выражение для переменного дипольного момента атома во внешнем линейно поляризованном (вдоль оси x) электро-

магнитном поле, полученное в работе [17], может быть упрощено (здесь и далее используются атомные единицы):

$$x(t) = i \int_0^{\infty} d\tau \left(\frac{\pi}{\varepsilon + i\tau/2} \right)^{3/2} d_x(\mathbf{p}_{st}(t, \tau) - \mathbf{A}(t - \tau)) E(t - \tau) \times \\ \times d_x^*(\mathbf{p}_{st}(t, \tau) - \mathbf{A}(t)) \exp[-iS(p_{st,x}, t, \tau) - iS_m(\mathbf{p}_{st}, t, \tau)] + c. c. \quad (3)$$

Здесь $c. c.$ обозначает комплексное сопряжение; $E(t)$ и $\mathbf{A}(t)$ — электрическое поле и векторный потенциал лазерного импульса соответственно; τ — время движения электрона в свободном пространстве; ε — параметр регуляризации, который может быть выбран малым; $d_x(\mathbf{p})$ — x -компонента дипольного матричного элемента, соответствующего переходу из основного связанного состояния в непрерывный спектр. Рассматривая атом гелия в приближении одного активного электрона, мы используем для данного матричного элемента такое же выражение, как и для водородоподобных атомов:

$$d_x(\mathbf{p}) = i \frac{2^{7/2} (2I_p)^{5/4}}{\pi} \frac{p_x}{(\mathbf{p}^2 + 2I_p)^3}. \quad (4)$$

$S(p_x, t, \tau)$ — квазиклассическое действие, описывающее свободное движение электрона в поле лазерного импульса без учета влияния магнитного поля:

$$S(p_x, t, \tau) = \int_{t-\tau}^t \left[\frac{1}{2} \left(p_x - \frac{A(t')}{c} \right)^2 + I_p \right] dt', \quad (5)$$

где $A(t')$ — проекция вектора $\mathbf{A}(t)$ на ось x . $S_m(\mathbf{p}, t, \tau)$ — найденная аналогично работе [20] поправка к действию, возникающая в результате разложения векторного потенциала $\mathbf{A}(t - z/c)$ в ряд по координате распространения лазерного излучения. Мы ограничиваемся линейным членом этого разложения, дающим поправку первого порядка по v/c к электродипольному приближению $\mathbf{A}(t)$, что в случае длинноволнового лазерного излучения достаточно для адекватного учета влияния магнитного поля [18]. В случае линейно поляризованного вдоль оси x лазерного импульса, распространяющегося в направлении оси z , интересующая нас поправка к действию имеет вид

$$S_m(\mathbf{p}, t, \tau) = \int_{t-\tau}^t \frac{1}{2} \left\{ p_z - \left[\frac{p_x A(t')}{c^2} - \frac{A^2(t')}{2c^3} \right] \right\}^2 dt'. \quad (6)$$

$\mathbf{p}_{st}(t, \tau)$ — стационарные значения компонент канонического импульса электрона, соответствующие траектории, начинающейся на ядре в момент времени $t - \tau$ и заканчивающейся на нем же в момент времени t :

$$p_{st,x}(t, \tau) = \frac{1}{\tau} \int_{t-\tau}^t \frac{A(t')}{c} dt', \quad (7)$$

$$p_{st,z}(t, \tau) = \frac{p_{st,x}^2}{c} - \frac{1}{2c\tau} \int_{t-\tau}^t \frac{A^2(t')}{c^2} dt'. \quad (8)$$

Входящие в выражение (3) для переменного дипольного момента подынтегральные множители $d_x(\mathbf{p}_{st}(t, \tau) - \mathbf{A}(t - \tau))E(t - \tau)$ и $d_x^*(\mathbf{p}_{st}(t, \tau) - \mathbf{A}(t))$ определяют соответственно амплитуды переходов из основного состояния в непрерывный спектр в момент времени $t - \tau$ и из непрерывного спектра в основное состояние в момент времени t .

2. Влияние магнитного дрейфа электрона на форму спектра гармоник

При взаимодействии излучения видимого или ближнего ИК-диапазонов с нейтральными атомами влияние магнитного поля лазерного импульса на процесс ГВП сводится к равномерному по всему спектру генерируемого излучения небольшому снижению эффективности преобразования. Изучим влияние магнитного дрейфа электрона на процесс ГВП в случае взаимодействия лазерного излучения среднего ИК-диапазона с атомами и ионами.

На рисунках 1—3 приведены результаты расчетов, полученные с помощью развитой аналитической теории. В представленных расчетах лазерное поле задавалось в виде гауссова импульса с длительностью на полувысоте максимума интенсивности, равной шести периодам поля. Расчеты проведены для атома и иона гелия в приближении одного активного электрона.

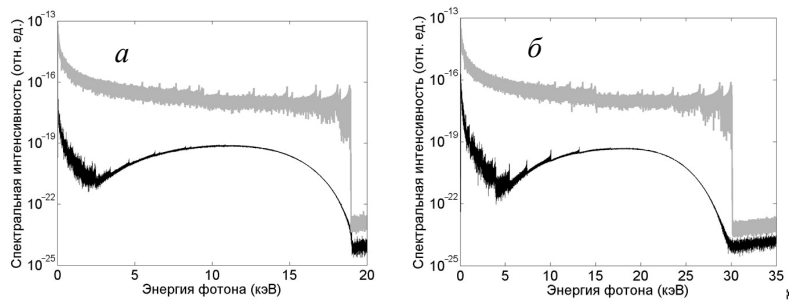


Рис. 1. Спектры высоких гармоник для атома и иона гелия, полученные с учетом (черные линии) и без учета (серые линии) магнитного дрейфа электрона: *a* — атом гелия, $\lambda = 8$ мкм, $I = 10^{15}$ Вт/см²; *б* — ион гелия, $\lambda = 3,9$ мкм, $I = 6,7 \cdot 10^{15}$ Вт/см²

Из рисунка 1 хорошо видно, что при учете магнитного дрейфа электрона характер спектров высоких гармоник может кардинально меняться:

вместо платообразных структур с резкими отсечками наблюдаются плавные дугообразные распределения. Это изменение характера спектра гармоник объясняется тем, что вклады различных типов электронных траекторий в поляризационный отклик системы оказываются в различной степени подавленными из-за магнитного дрейфа электронов [16, 17, 22]. В отсутствие магнитного дрейфа сигнал высоких гармоник содержит вклады коротких ($\tau < 0,65T$, где T — период лазерного поля) и длинных ($\tau \geq 0,65T$) электронных траекторий. Когерентное сложение вкладов этих траекторий приводит к формированию платообразного распределения с четко выраженной высокочастотной отсечкой и спектральной интерференцией (см. серые кривые на рис. 1). При субрелятивистском режиме взаимодействия с лазерным полем электроны испытывают магнитный дрейф, причем для траекторий с большим τ этот дрейф больше. Соответственно, вклады длинных траекторий оказываются полностью, а коротких — частично подавленными. Как результат, спектр приобретает наблюдаемую на рис. 1 дугообразную форму (черные кривые), а спектральная интерференция вкладов различных траекторий исчезает, за исключением низкочастотной области.

Все сказанное о вкладах различных электронных траекторий в сигнал высоких гармоник наглядно иллюстрируется приведенными на рис. 2 и 3 спектрограммами этого сигнала для атома и иона гелия соответственно. На рисунках представлены временные зависимости квадрата коэффициентов вейвлет-преобразования сигнала гармоник. Вейвлет-анализ осуществлялся с помощью анализирующей функции $T_{t_0, \Omega}(t) = \sqrt{\Omega} \Phi(\Omega(t - t_0))$, где в качестве материнского вейвлета использовался вейвлет Морле $\Phi(\eta) = (1/\sqrt{\tau'}) \exp(i\eta) \exp[-\eta^2/(2\tau'^2)]$ (см., например, [23]). Здесь t_0 и Ω — переменные, используемые соответственно для сдвига и масштабирования окна вейвлет-преобразования. При выборе параметра τ' равным $5\pi/\sqrt{2 \ln 2} \approx 13,34$ полная ширина на полувысоте анализирующей функции равна 5 периодам лазерного поля.

В отсутствие магнитного дрейфа электрона (электродипольное приближение, рис. 2, *а* и 3, *а*) сигнал высоких гармоник, как отмечалось выше, содержит вклады различных типов траекторий: коротким траекториям соответствуют высокочастотные компоненты с положительным чирпом (рост частоты генерируемых фотонов со временем), длинным траекториям с $0,65T \leq \tau \leq T$ — высокочастотные компоненты с отрицательным чирпом (спад частоты генерируемых фотонов со временем), а еще более длинным траекториям ($\tau > T$) — низкочастотные компоненты как с положительным, так и с отрицательным чирпом. Из-за влияния магнитного дрейфа (расчеты за рамками электродипольного приближения, рис. 2, *б* и 3, *б*)

вклады всех длинных траекторий оказываются полностью, а коротких — частично подавленными.

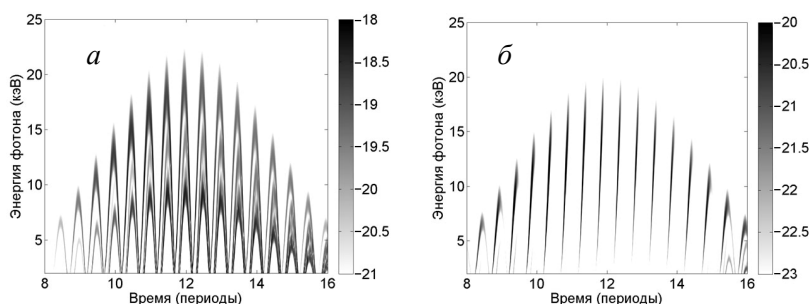


Рис. 2. Спектрограммы сигнала высоких гармоник для атома гелия в поле лазерного импульса с $\lambda = 8$ мкм и $I = 10^{15}$ Вт/см². Представлены результаты расчетов в электродипольном приближении (*a*) и за его рамками (*б*)

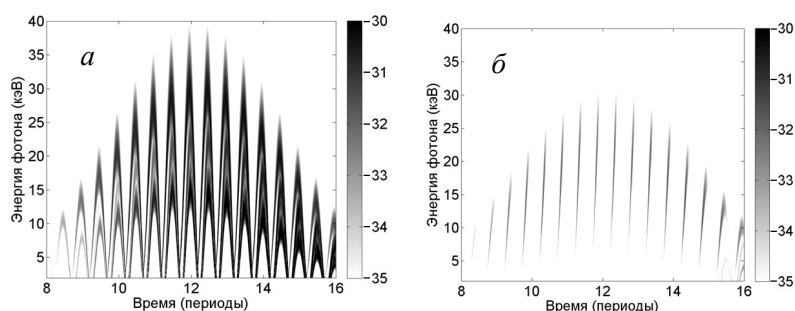


Рис. 3. То же, что на рис. 2, но для иона гелия в поле лазерного импульса с $\lambda = 3,9$ мкм и $I = 6,7 \cdot 10^{15}$ Вт/см²

Таким образом, рис. 1—3 наглядно демонстрируют, что в субрелятивистском режиме взаимодействия электрона с лазерным полем магнитный дрейф электрона приводит не только к уменьшению эффективности механизма ГВП, но и к качественному изменению формы спектра генерируемого излучения.

3. Магнитный дрейф электрона как положительный фактор

В предыдущем разделе было продемонстрировано подавление вкладов различных электронных траекторий в сигнал гармоник как следствие магнитного дрейфа электрона. Обычно такое подавление является негативным фактором, так как приводит к уменьшению эффективности генерации

высокоэнергичных фотонов. В данном разделе будет рассмотрен пример, когда магнитный дрейф электрона может играть положительную роль.

Особенностью ГТВП лазерного излучения среднего ИК-диапазона является наличие острых пиков в спектре гармоник. Эти пики являются следствием многократных возвратов групп электронов, оторванных на различных полупериодах лазерного поля, к родительскому иону с высокой энергией. Интерференция вкладов электронных траекторий, начавшихся на различных полупериодах лазерного поля и вернувшихся к родительскому иону в один и тот же момент времени, может приводить к образованию биений во временном представлении сигнала гармоник на заднем фронте предельно короткого лазерного импульса [9]. При этом характерный временной масштаб этих биений может составлять единицы аттосекунд и менее.

Авторы работы [9] рассматривали процесс генерации таких биений в электродипольном приближении. Однако, как следует из предыдущего раздела, в среднем ИК-диапазоне, необходимом для предлагаемого в статье [9] механизма формирования аттосекундных биений, влияние магнитной составляющей лазерного импульса может быть очень сильно. В результате магнитного дрейфа все многократные возвраты оторванного электрона к родительскому иону могут полностью исчезнуть, что приведет к полному исчезновению биений во временном представлении сигнала гармоник. Поэтому для анализа возможности формирования аттосекундных биений и выявления минимально достижимого временного масштаба биений необходимо рассмотрение за рамками электродипольного приближения.

Рассмотрим процесс генерации аттосекундных биений при воздействии на атом гелия лазерным импульсом следующего вида (такая же форма импульса использовалась в работе [9]):

$$E(t) = E_0 \sin^2\left(\frac{\pi ct}{4\lambda}\right) \sin\left(\frac{2\pi ct}{\lambda} - \phi\right). \quad (9)$$

Пиковая интенсивность лазерного импульса была зафиксирована на уровне $I = 3,4 \cdot 10^{14}$ Вт/см². На рисунке 4 показано электрическое поле лазерного импульса и временное окно, в котором анализировался сигнал гармоник, при фазе $\phi = \pi/8$ (в электродипольном приближении данная фаза является оптимальной [9]).

Сигнал гармоник, генерируемый в атоме гелия под действием лазерного импульса, представленного на рис. 4 черной линией, во временном представлении состоит из двух всплесков в окрестностях $t = 2,5T$ и $t = 3T$. Интересующие нас сверхвысокочастотные биения сигнала гармоник наблюдаются во втором из этих всплесков, который селективируется для анализа маской, показанной на рис. 4 серой линией. Биения обусловлены интерференцией преимущественно двух электронных траекторий, одна из

которых проводит в свободном пространстве менее периода лазерного поля ($\tau < T$), а для второй выполняется условие $T < \tau < 1,5T$.

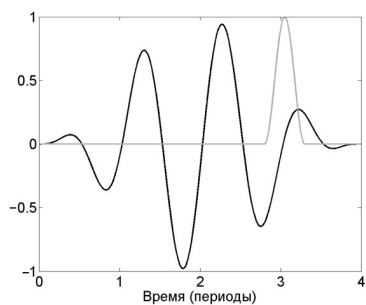


Рис. 4. Электрическое поле лазерного импульса, нормированное на величину E_0 (черная линия), и временное окно, в котором анализировался сигнал гармоник (серая линия); $\varphi = \pi/8$

На рисунке 5 показан спектральный состав рентгеновского всплеска, генерируемого на заднем фронте лазерного импульса (9) при $\lambda = 9$ мкм и $\varphi = \pi/8$. Серой линией представлены результаты, полученные в рамках электродипольного приближения. В спектре наблюдаются два ярко выраженных пика вблизи энергий фотонов 0,6 и 3,05 кэВ. Положения этих пиков соответствуют положениям отсечек в спектрах гармоник, генерируемых вследствие второго и первого возврата электронного волнового пакета к родительскому иону соответственно [9]. Высота каждого плато определяется весом соответствующей электронной траектории.

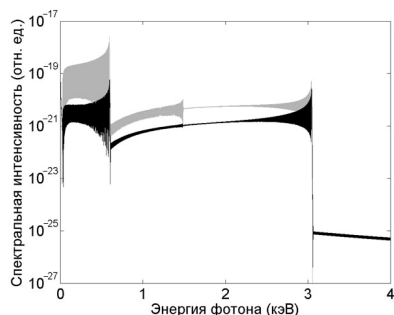


Рис. 5. Спектральный состав рентгеновского всплеска, генерируемого на заднем фронте лазерного импульса (9) при $\lambda = 9$ мкм и $\varphi = \pi/8$. Представлены результаты расчетов в электродипольном приближении (серая линия) и за его рамками (черная линия)

Наличие двух ярко выраженных пиков в спектре излучения, как на рис. 5, приводит к появлению биений во временном представлении сигнала. Очевидно, что контраст этих биений максимален, когда два спектральных пика имеют одинаковые амплитуды. В спектре на рис. 5, найденном

в рамках электродипольного приближения (серая линия), амплитуды высоко- и низкоэнергетического пиков сильно отличаются. Однако амплитуда низкоэнергетической компоненты спектра может быть подстроена путем пропускания сигнала гармоник, например, через алюминиевую фольгу [9]. В результате сигнал гармоник во временном представлении будет представлять собой биения с коэффициентом амплитудной модуляции 1 и характерной длительностью каждого всплеска около 1 ас.

Черной линией на рис. 5 представлены результаты расчетов за рамками электродипольного приближения. Как хорошо видно, магнитный дрейф электрона в данном случае не разрушает механизм генерации биений — в спектре генерируемого излучения по-прежнему наблюдаются два ярко выраженных максимума. Более того, магнитное поле лазерного импульса, которое обычно приводит к негативным последствиям в процессе ГГВП, в данном случае выступает в роли спектрального фильтра, выравнивающего интенсивности двух пиков.

Такой эффект легко объясняется на основе результатов, полученных в предыдущем разделе. Низкоэнергетический пик генерируется электронами, проводящими в свободном пространстве более периода лазерного поля. Высокоэнергетический пик обусловлен вкладом электронов с временем движения в свободном пространстве менее периода лазерного поля. Чем больше времени электрон проводит в свободном пространстве, тем в большей степени его траектория подвержена магнитному дрейфу, поэтому амплитуда низкоэнергетического пика в спектре анализируемого сигнала гармоник в результате учета магнитного поля лазерного импульса уменьшилась сильно, в то время как амплитуда высокоэнергетического пика практически не изменилась. В итоге для наблюдения аттосекундных биений требуются гораздо менее плотные спектральные фильтры, чем предсказывают расчеты в рамках электродипольного приближения, а если осуществить более тонкую подстройку параметров лазерного импульса (амплитуды или фазы заполнения), то можно обойтись совсем без дополнительной спектральной фильтрации. В работе [24] подобраны оптимальные параметры лазерного излучения для наблюдения аттосекундных биений, так что дополнительная спектральная фильтрация излучения гармоник не требуется, и показано, что минимально достижимый временной масштаб биений достигается при использовании лазерных источников с длинами волн 12—14 мкм и составляет приблизительно 0,8 ас.

Заключение

Таким образом, магнитный дрейф электрона может приводить не только к общему снижению эффективности ГГВП, но и к качественному изменению формы спектра гармоник. В силу различной степени влияния магнитного дрейфа электрона на короткие и длинные электронные траек-

тории спектр гармоник может приобретать дугообразный вид с максимумом в области средних энергий генерируемых фотонов и плавными спадами в областях высоких и низких энергий. Такой селективный характер подавления вкладов различных электронных траекторий может играть положительную роль в процессе формирования аттосекундных биений сигнала гармоник. Магнитный дрейф электрона не разрушает механизм генерации таких биений на заднем фронте интенсивного предельно короткого лазерного импульса, а выступает в качестве спектрального фильтра, позволяющего упростить наблюдение генерируемых биений в эксперименте.

Работа выполнена при поддержке гранта РФФИ № 16-12-10279.

Литература

1. *Schmidt, B. E.* Compression of 1.8 μm laser pulses to sub two optical cycles with bulk material / B. E. Schmidt, P. B  jot, M. Gigu  re, A. D. Shiner, C. Trallero-Herrero,  . Bisson, J. Kasparian, J.-P. Wolf, D. M. Villeneuve, J.-C. Kieffer, P. B. Corkum, F. L  gar   // *Appl. Phys. Lett.* — 2010. — V. 96, № 12. — P. 121109.
2. *Andriukaitis, G.* 90 GW peak power few-cycle mid-infrared pulses from an optical parametric amplifier / G. Andriukaitis, T. Balciunas, S. Ali  sauskas, A. Pug  zlys, A. Baltu  ska, T. Popmintchev, M.-C. Chen, M. M. Murnane, H. C. Kapteyn // *Opt. Lett.* — 2011. — V. 36, № 15. — P. 2755—2757.
3. *Deng, Y.* Carrier-envelope-phase-stable, 1.2 mJ, 1.5 cycle laser pulses at 2.1 μm / Y. Deng, A. Schwarz, H. Fattahi, M. Ueffing, X. Gu, M. Ossiander, T. Metzger, V. Pervak, H. Ishizuki, T. Taira, T. Kobayashi, G. Marcus, F. Krausz, R. Kienberger, N. Karpowicz // *Opt. Lett.* — 2012. — V. 37, № 23. — P. 4973—4975.
4. *Colosimo, P.* Scaling strong-field interactions towards the classical limit / P. Colosimo, G. Doumy, C. I. Blaga, J. Wheeler, C. Hauri, F. Catoire, J. Tate, R. Chirla, A. M. March, G. G. Paulus, H. G. Muller, P. Agostini, L. F. DiMauro // *Nature Phys.* — 2008. — V. 4, № 5. — P. 386—389.
5. *Popmintchev, T.* Bright coherent ultrahigh harmonics in the keV X-ray regime from mid-infrared femtosecond lasers / T. Popmintchev, M.-C. Chen, D. Popmintchev, P. Arpin, S. Brown, S. Ali  sauskas, G. Andriukaitis, T. Balciunas, O. D. M  cke, A. Pug  zlys, A. Baltu  ska, B. Shim, S. E. Schrauth, A. Gaeta, C. Hern  ndez-Gars  a, L. Plaja, A. Becker, A. Jaron-Becker, M. M. Murnane, H. C. Kapteyn // *Science.* — 2012. — V. 336, № 6086. — P. 1287—1291.
6. *Blaga, C. I.* Imaging ultrafast molecular dynamics with laser-induced electron diffraction / C. I. Blaga, J. Xu, A. D. DiChiara, E. Sistrunk, K. Zhang, P. Agostini, T. A. Miller, L. F. DiMauro, C. D. Lin // *Nature.* — 2012. — V. 483, № 7388. — P. 194—197.
7. *DiChiara, A. D.* Inelastic scattering of broadband electron wave packets driven by an intense midinfrared laser field / A. D. DiChiara, E. Sistrunk, C. I. Blaga, U. B. Szafruga, P. Agostini, L. F. DiMauro // *Phys. Rev. Lett.* — 2012. — V. 108, № 3. — P. 033002.
8. *Vozzi, C.* Strong-field phenomena driven by mid-infrared ultrafast sources / C. Vozzi, M. Negro, S. Stagira // *J. Mod. Opt.* — 2012. — V. 59, № 15. — P. 1283—1302.
9. *Hern  ndez-Gars  a, C.* Zeptosecond high harmonic keV X-ray waveforms driven by midinfrared laser pulses / C. Hern  ndez-Gars  a, J. A. P  rez-Hern  ndez, T. Popmint-

chev, M. M. Murnane, H. C. Kapteyn, A. Jaron-Becker, A. Becker, L. Plaja // Phys. Rev. Lett. — 2013. — V. 111, № 3. — P. 033002.

10. *Schafer, K. J.* Above threshold ionization beyond the high harmonic cutoff / K. J. Schafer, B. Yang, J. L. Krause, K. C. Kulander // Phys. Rev. Lett. — 1993. — V. 70, № 11. — P. 1599—1602.

11. *Corkum, P. B.* Plasma perspective on strong field multiphoton ionization // Phys. Rev. Lett. — 1993. — V. 71, № 13. — P. 1994—1997.

12. *McPherson, A.* Studies of multiphoton production of vacuum-ultraviolet radiation in the rare gases / A. McPherson, G. Gibson, H. Jara, U. Johann, T. S. Luk, I. A. McIntyre, K. Boyer, C. K. Rhodes // J. Opt. Soc. Am. B. — 1987. — V. 4, № 4. — P. 595—601.

13. *Ferray, M.* Multiple-harmonic conversion of 1064 nm radiation in rare gases / M. Ferray, A. L'Huillier, X. F. Li, L. A. Lompre, G. Mainfray, C. Manus // J. Phys. B: At. Mol. Opt. Phys. — 1988. — V. 21, № 3. — P. L31—L35.

14. *Krause, J. L.* High-order harmonic generation from atoms and ions in the high intensity regime / J. L. Krause, K. J. Schafer, K. C. Kulander // Phys. Rev. Lett. — 1992. — V. 68, № 24. — P. 3535—3538.

15. *Shan, B.* Dramatic extension of the high-order harmonic cutoff by using a long-wavelength driving field / B. Shan, Z. Chang // Phys. Rev. A. — 2001. — V. 65, № 1. — P. 011804(R).

16. *Емелина, А. С.* О возможности генерации высоких гармоник с энергиями квантов более 10 кэВ при взаимодействии интенсивного излучения среднего ИК-диапазона с нейтральными газами / А. С. Емелина, М. Ю. Емелин, М. Ю. Рябикин // Квант. электроника. — 2014. — Т. 44, вып. 5. — С. 470—477.

17. *Emelina, A. S.* Multi-keV ultrahigh-order harmonics produced in gases with subrelativistically intense mid-IR laser pulses / A. S. Emelina, M. Yu. Emelin, M. Yu. Ryabikin // J. Opt. Soc. Am. B. — 2015. — V. 32, № 12. — P. 2478—2487.

18. *Ким, А. В.* От фемтосекундных к аттосекундным импульсам / А. В. Ким, М. Ю. Рябикин, А. М. Сергеев // УФН. — 1999. — Т. 169, вып. 1. — С. 58—66.

19. *Taranukhin, V. D.* Relativistic high-order harmonic generation // Laser Physics. — 2000. — V. 10, № 1. — P. 330—336.

20. *Walser, M. W.* High harmonic generation beyond the electric dipole approximation / M. W. Walser, C. H. Keitel, A. Scrinzi, T. Brabec // Phys. Rev. Lett. — 2000. — V. 85, № 24. — P. 5082—5085.

21. *Lewenstein, M.* Theory of high-harmonic generation by low-frequency laser fields / M. Lewenstein, Ph. Balcou, M. Yu. Ivanov, A. L'Huillier, P. B. Corkum // Phys. Rev. A. — 1994. — V. 49, № 3. — P. 2117—2132.

22. *Chirilă, C. C.* Nondipole effects in photon emission by laser-driven ions / C. C. Chirilă, N. J. Kylstra, R. M. Potvliege, C. J. Joachain // Phys. Rev. A. — 2002. — V. 66, № 6. — P. 063411.

23. *Strelkov, V. V.* High-order harmonic generation by atoms in an elliptically polarized laser field: harmonic polarization properties and laser threshold ellipticity / V. V. Strelkov, M. A. Khokhlova, A. A. Gonoskov, I. A. Gonoskov, M. Yu. Ryabikin // Phys. Rev. A. — 2012. — V. 86, № 1. — P. 013404.

24. *Emelina, A. S.* Subattosecond keV beats of the high-harmonic X-ray field produced with few-cycle mid-IR laser pulses: magnetic-field effects / A. S. Emelina, M. Yu. Emelin, M. Yu. Ryabikin // Phys. Rev. A. — 2016. — V. 93, № 4. — P. 043802.

КРИТИЧНЫЕ И НЕКРИТИЧНЫЕ ПРОЦЕССЫ ПРИ НЕЛИНЕЙНО-ОПТИЧЕСКОМ ПРЕОБРАЗОВАНИИ ЧАСТОТЫ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ

С. Г. Гречин

Московский государственный технический университет им. Н. Э. Баумана

Введение

При всех достоинствах лазеров, формирующих высокоэнергетичное излучение с небольшой расходимостью и шириной спектра, у них имеется одно значительное ограничение — генерация происходит на фиксированных длинах волн или в ограниченном диапазоне. Расширить функциональные возможности лазеров при формировании излучения в диапазоне длин волн от вакуумного ультрафиолета до терагерцовых, а в последние годы и до миллиметровых волн позволяет нелинейно-оптическое преобразование частоты.

В процессе преобразования частоты для получения максимальной эффективности должно обеспечиваться условие когерентного накопления формируемой в каждой точке среды волны при ее распространении в нелинейном кристалле [1]. Для этого необходимо обеспечивать отличие от нуля коэффициента эффективной нелинейности, $d_{\text{эфф}} \neq 0$ (достаточное условие), и выполнение условия фазового синхронизма, $\Delta k = 0$ (необходимое условие). В нецентросимметричных кристаллах существуют направления, в которых $d_{\text{эфф}} \neq 0$. Вследствие этого формируется волна суммарной или разностной частоты. Фазовый синхронизм (ФС) обеспечивается определенным соотношением фазовых скоростей взаимодействующих волн, или, как это принято описывать на практике, разностью волновых векторов всех взаимодействующих волн ($\Delta k = k_3 - k_2 - k_1 = 0$). Оптимальная величина двулучепреломления в кристаллах определяет возможность реализации ФС.

Для заданных длин волн и температуры кристалла ФС имеет место в строго определенных направлениях. Изменение любого из параметров приводит к его нарушению. Необходимо узнать, возможно ли обеспечить процесс преобразования с высокой эффективностью в большом диапазоне изменения всех параметров. В этой статье дано общее описание возможности реализации критичных и некритичных фазовых синхронизмов при нелинейно-оптическом преобразовании частоты лазерного излучения. Они позволяют выделить «базовые» параметры кристаллов, которые в значительной степени определяют основные особенности развития нелинейных волновых процессов преобразования частоты в условиях сильного энергообмена. Работа базируется во многом на результатах, которые были получены коллективом сотрудников МГТУ им. Н. Э. Баумана и исследователями из других организаций.

1. Ширины синхронизма

Определение ширины синхронизма как допустимого изменения параметра, влияющего на процесс преобразования, следует из зависимости эффективности преобразования (η) от величины волновой расстройки (Δk) между тремя взаимодействующими волнами [1]. Не уточняя связь с параметрами, которые влияют на величину η , представим изменение расстройки в виде ряда Тейлора [1] по параметру p :

$$\Delta k(p) = \Delta k(p_0) + \frac{d\Delta k}{dp} \Delta p + \frac{1}{2} \frac{d^2 \Delta k}{dp^2} \Delta p^2 + \frac{1}{3!} \frac{d^3 \Delta k}{dp^3} \Delta p^3 + \dots, \quad (1)$$

где $\Delta k = k_3 - k_2 - k_1 = 2\pi(n_3/\lambda_3 - n_2/\lambda_2 - n_1/\lambda_1)$; φ, θ — углы среза кристалла; λ_i — длины волн взаимодействующих импульсов излучения; T — температура кристалла; n_i — показатели преломления среды на длинах волн λ_i .

При выполнении условия ФС $\Delta k(p_0) = 0$. В общем случае рассматривается аддитивный вклад в волновую расстройку отклонения оси лазерного излучения от направления ФС ($\Delta\varphi$ и $\Delta\theta$), изменения длин волн ($\Delta\lambda$) и температуры кристалла (ΔT). Допустимое изменение параметра Δp , при котором эффективность преобразования $\eta(\Delta p)$ уменьшается до заданной величины, определяет соответствующую ширину синхронизма. Ширины характеризуют свойства среды и устанавливаются по каждому параметру p независимо, в приближении плоских монохроматических волн и небольшой эффективности преобразования (приближение заданного поля). Из выражения для эффективности преобразования (см. [1]) можно выделить множители, обуславливающие ее зависимость от величины $\Delta k(\Delta p)$ при длине кристалла $L_{кр}$ и от коэффициента эффективной нелинейности $d_{эфф}$ среды:

$$\eta \propto d_{эфф}^2 \text{sinc}^2 \Delta k L_{кр} / 2. \quad (2)$$

При доминирующем вкладе одного из слагаемых в разложении (1) синхронизм является критичным при $d\Delta k/dp \gg 0,5 d^2 \Delta k/dp^2 \Delta p$ (кривая 1 на рис. 1) или некритичным при $d\Delta k/dp = 0$ (кривая 2 на рис. 1). Во втором случае ключевую роль играет $d^2 \Delta k/dp^2$ (некритичный синхронизм 2-го порядка). В промежуточном случае, когда вклад производных 1-го и 2-го порядков в (1) соизмерим, зависимость эффективности преобразования от волновой расстройки имеет вид кривой 3 на рис. 1.

Во всех случаях нарушение ФС может быть определено изменением показателей преломления n_i всех взаимодействующих волн (dn_i/dT). Поэтому ширины задаются изменением n_i по соответствующему параметру p . Так, при изменении температуры $d\Delta k/dT = 2\pi[(dn_3/dT)/\lambda_3 - (dn_2/dT)/\lambda_2 - (dn_1/dT)/\lambda_1]$.

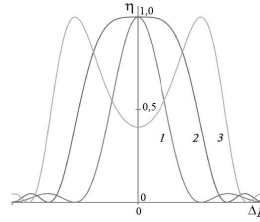


Рис. 1. Зависимость эффективности преобразования от волновой расстройки

В частном случае генерации второй гармоники (ГВГ, $\lambda_1 = \lambda_2 = \lambda$, $\lambda_3 = \lambda/2$) в кристалле длиной $L_{кр}$ полуширина на уровне 0,5 критичного и некритичного синхронизмов m -го порядка устанавливается из (2) с учетом (1) (при действии только одного параметра p) как

$$\Delta p_{ФС} = m \sqrt{\frac{\lambda m!}{4 L_{кр} d^m \Delta n / dp^m}}. \quad (3)$$

Принято, что $\Delta p_{ФС}$ определяется для кристалла длиной 1 см. Поэтому для критичного синхронизма ($m = 1$), например по температуре, размерность ширины — $^{\circ}\text{C}\cdot\text{см}$. Ее величина обратно пропорциональна $L_{кр}$. Для некритичного синхронизма с $m = 2$ величина ширины обратно пропорциональна корню квадратному из $L_{кр}$ — $^{\circ}\text{C}\cdot\text{см}^{1/2}$. Когда в волновую расстройку вклад делают одновременно производные разного порядка, ширина синхронизма задается корнями полиномиального выражения (1).

Некритичный синхронизм представляет практический интерес, так как в большом диапазоне изменения параметра Δp сохраняется условие ФС, т. е. возможно эффективное преобразование частоты для излучения с большой расходимостью и шириной спектра или в большом диапазоне изменения температуры. В следующих разделах мы рассмотрим реализацию углового, спектрального и температурного синхронизмов, а в последнем разделе — процесс преобразования импульсов фемтосекундной длительности при некритичном и критичном по частоте синхронизме.

2. Угловая ширина синхронизма

Величина угловой ширины синхронизма важна в двух случаях: при преобразовании частоты излучения с большой расходимостью и при фокусировке излучения в кристалл. Как в одноосных, так и в двухосных кристаллах в общем случае ФС по одной угловой координате всегда является

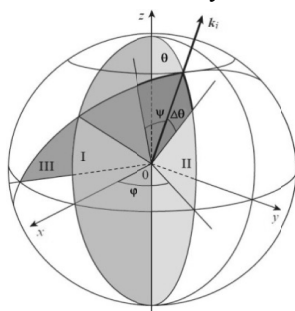


Рис. 2. Определение угловых расстройок оси излучения в зависимости от направления фазового синхронизма

некритичным. На рисунке 2 угловые расстройки оси лазерного излучения (задаваемого вектором k_i) определены в зависимости от направления ФС, задаваемого углами φ и θ . Расстройки находятся в секущей плоскости II ($\Delta\theta$) и ортогональной ей плоскости III ($\Delta\psi$). Для одноосных кристаллов направления синхронизма представляют собой поверхность конуса, ось которого совпадает с осью z . В плоскости II зависимость эффективности преобразования от $\Delta\theta$ имеет вид, соответствующий кривой 1 на рис. 1, а зависимость в плоскости III от $\Delta\psi$ — кривой 2 на рис. 1.

Встает вопрос о реализации не критичного по углу синхронизма в двух ортогональных плоскостях. В одноосных кристаллах не критичный синхронизм реализуется только на определенной длине волны (соответствующей границе области ФС) в плоскости xu ($\theta = 90^\circ$).

При этом величина угловой ширины синхронизма по $\Delta\psi$ равна 2π , а по угловой расстройке $\Delta\theta$ достигает единиц градусов. На всех других длинах волн области ФС $\theta < 90^\circ$ и синхронизм является критичным по углу θ .

Другая ситуация имеет место в двухосных кристаллах. На рисунке 3 приведены все возможные распределения направлений ФС для процесса ГВГ и переходы между ними при изменении длины волны в кристалле LBO (трибората лития, LiB_3O_5). Для типа взаимодействия *ssf* (slow-slow-fast) на некоторой длине волны вдоль оси y возникает ФС, не критичный по обоим углам [3]. При изменении длины волны направления ФС являются коническими поверхностями 4-го порядка (рис. 3). На некоторых длинах волн направление ФС совпадает с одной из главных осей кристалла: x или z .

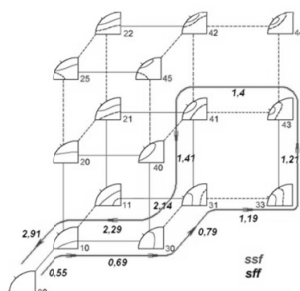


Рис. 3. Изменение направлений ФС в двухосном кристалле LBO в зависимости от длины волны (см. описание в работе [2])

На рисунке 4 показано распределение эффективности преобразования по $\Delta\phi$ и $\Delta\theta$ для случая, когда ФС совпадает по направлению с осью x . Это имеет место, например, при ГВГ *ssf*-типа взаимодействия на длине волны 1,0642 мкм в кристалле LBO при нагреве его до 148 °С [4]. В обеих плоскостях (xu и xz) синхронизм является не критичным и зависимость эффективности преобразования от угловой расстройки имеет вид, соответствующий кривой 2 на рис. 1. Угловые ширины синхронизма ($\Delta\phi_{\text{ФС}}$ и $\Delta\theta_{\text{ФС}}$) в этом случае составляют единицы градус Цельсия-сантиметров в степени 1/2. Небольшое различие $\Delta\phi_{\text{ФС}}$ и $\Delta\theta_{\text{ФС}}$ позволяет использовать фокусировку лазерного излучения в кристалл с помощью сферических линз.

На рисунке 4 видно, что под некоторыми углами к главным плоскостям кристалла xu и xz эффективность преобразования остается максимальной в большом диапазоне угловых расстроек [5]. В этих направлениях угловая ширина синхронизма может быть значительно больше, чем в главных плоскостях, и не критичность синхронизма будет более высокого порядка, чем параллельно главным плоскостям кристалла.

На рисунке 5 приведена зависимость угловой полуширины синхронизма от длины кристалла (сплошная кривая). Там же пунктирными линиями показаны корневые зависимости разных степеней от длины кристалла. Из этого графика следует, что при небольших длинах кристалла

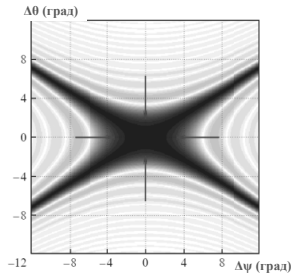


Рис. 4. Некритичный по двум углам синхронизм вдоль оси x в двухосном кристалле

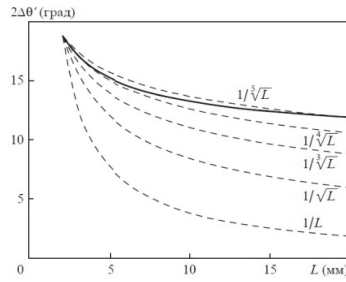


Рис. 5. Зависимость угловой ширины синхронизма от длины кристалла

возникает некритичный синхронизм 4-го порядка, а при увеличении длины — 5-го. Существуют срезы кристалла, при которых угловая ширина синхронизма может превышать 20° . Но этот результат получается только по одной координате, по другой — синхронизм является критичным.

Такая зависимость эффективности преобразования от угловых расстройек дает возможность получать высокую интенсивность лазерного излучения в нелинейном кристалле, используя фокусировку излучения короткофокусными цилиндрическими линзами, фокусные расстояния которых в ортогональных плоскостях согласованы с угловыми ширинами синхронизма.

3. Спектральная ширина синхронизма

Величина спектральной ширины синхронизма важна в двух случаях — при преобразовании частоты излучения с большой шириной спектра (в частности, импульсов фемтосекундной длительности) или перестраиваемого по длине волны. В большинстве случаев в одноосных кристаллах при изменении длины волны угол синхронизма $\theta_{\text{ФС}}$ изменяется, достигая некоторого минимального значения (рис. 6). Величина $\theta_{\text{ФС}}$ никогда не достигает нуля (соответствующего оси z кристалла), так как это запрещено дисперсией показателей преломления. При $d\theta_{\text{ФС}}/d\lambda = 0$ в достаточно большом диапазоне изменения длины волны величина угла синхронизма не изменяется, т. е. сохраняется условие ФС. В этом случае имеет место некритичный по частоте (длине волны) синхронизм

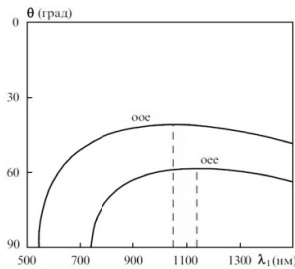


Рис. 6. Углы синхронизма $\theta_{\text{ФС}}$ и точки некритичного по длине волны синхронизма в одноосном кристалле

(НКЧС). Величина спектральной ширины синхронизма может достигать нескольких десятков нанометров. В более общем случае синхронизм является критичным по частоте ($d\theta_{\text{ФС}}/d\lambda \neq 0$) и спектральная ширина составляет доли — единицы нанометров.

Общий принцип изменения угла синхронизма в двухосных кристаллах определяет наличие режима НКЧС аналогично одноосным кристаллам (см. рис. 3) [3]. Например, для *ssf*-типа взаимодействия ФС появляется вдоль оси *y* (вырожденный случай). При изменении длины волны излучения формируется коническая поверхность направлений ФС, которая первоначально расширяется, а в дальнейшем стягивается к направлению оптической оси кристалла (расположенной в плоскости *xz*), однако никогда не достигает ее, потому что это запрещено дисперсией показателей преломления. На некоторой длине волны изменение угла синхронизма меняется на противоположное (конус расширяется, отходя от оптической оси). Это показывают переходы между распределениями 33—43—41 на рис. 3 (см. подробно в работе [3]). На рисунке 7 приведено распределение направлений ФС (белые линии) при ГВГ в кристалле LBO на 1307 нм для *ssf*-типа взаимодействия. Там же показано распределение коэффициента эффективной нелинейности (градации серого).

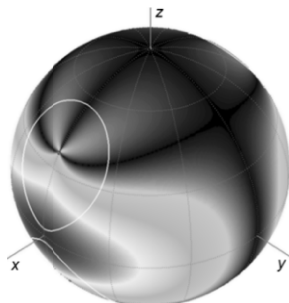


Рис. 7. НКЧС в плоскости *xz* в двухосном кристалле LBO

Для удобства качественного и количественного анализа не критичного по частоте синхронизма, по аналогии с ФС, удобно определить направления не критичных взаимодействий — НКЧВ ($d\Delta k(\varphi, \theta)/d\lambda = 0$). Как и для ФС ($\Delta k(\varphi, \theta) = 0$), НКЧВ имеет распределение по φ и θ и представляет собой коническую поверхность. Те направления, вдоль которых распределения ФС и НКЧВ пересекаются, т. е. одновременно выполняется условие ФС ($\Delta k(\varphi, \theta) = 0$) и равна нулю первая производная ($d\Delta k(\varphi, \theta)/d\lambda = 0$), являются направлениями НКЧС.

На рисунке 8 приведены распределения ФС и НКЧВ в области реализации НКЧС при ГВГ *ssf*-типа взаимодействия в кристалле LBO. На длине волны 1029,9 мкм кривые не пересекаются. Вдоль всех направлений ФС синхронизм является критичным по частоте. При увеличении длины волны до 1260 нм в плоскости *xz* имеется направление, вдоль которого поверхности ФС и НКЧВ касаются друг друга; оно соответствует НКЧС. Но в этом направлении $d_{\text{эфф}} = 0$ (см. рис. 7). Поэтому оно не имеет практической ценности. При дальнейшем увеличении длины волны конусы ФС и НКЧВ пересекаются при других значениях углов φ и θ , так как зависимости скорости их изменения от длины волны различны. В этом случае

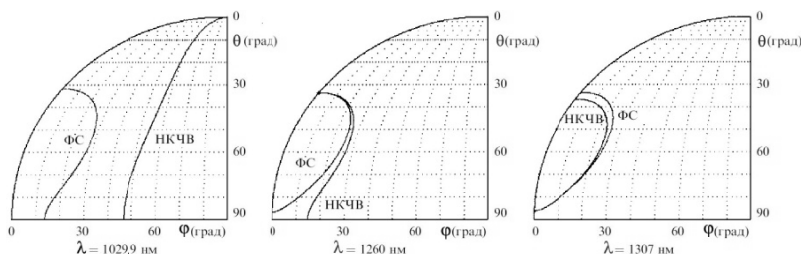


Рис. 8. Дисперсия направлений фазового синхронизма при ГВГ *ssf*-типа в двухосном кристалле LBO в области НКЧС

направление ФС с НКЧС лежит не в главных плоскостях кристалла. На длине волны 1307 нм конусы ФС и НКЧВ опять касаются в плоскости *xz*. В этом направлении величина $d_{эфф}$ имеет максимальное значение (см. рис. 7). Итак, для рассматриваемого случая ГВГ *ssf*-типа взаимодействия режим НКЧС реализуется в диапазоне длин волн 1260—1307 нм. Уменьшение величины $d_{эфф}$ на длине волны менее 1307 нм часто не является критичным при преобразовании частоты импульсов фемтосекундной длительности. Интенсивность излучения таких импульсов столь высока, что можно использовать кристаллы небольшой длины или с небольшим значением $d_{эфф}$.

Таким образом, в одноосных кристаллах режим НКЧС реализуется на фиксированных длинах волн для различных типов взаимодействия. В двухосных кристаллах режим НКЧС имеет место в некотором диапазоне длин волн. Необходимо отметить, что режим НКЧС соответствует равенству групповых скоростей взаимодействующих волн.

4. Температурная ширина синхронизма

С самого начала развития теории нелинейно-оптического преобразования частоты вопрос о возможности реализации не критичного по температуре синхронизма (НКТС) не ставился. Традиционно рассматривался только критичный по температуре синхронизм (КТС). Очевидно, это было связано с тем, что соотношения между температурными производными для главных значений показателей преломления могут быть самыми разными.

Типовые значения температурных ширин синхронизма для различных процессов в используемых кристаллах составляют в подавляющем большинстве величину от единиц до нескольких десятков градус Цельсия-сантиметров. Но экспериментально неоднократно получались аномально большие значения температурных ширин синхронизма, превышающие 100 °С. В таблице 1 из монографии [4] приведены значения температурных ширин синхронизма для процессов ГВГ *ssf*-типа (*slow-fast-fast*) при $T = 293$ К в разных главных плоскостях для широко используемого кристалла КТР (титанил-фосфата калия, KTiOPO_4).

Таблица 1

Длины волн, мкм	φ , град	θ , град	ΔT , °C
Плоскость xu			
1,062 \rightarrow 0,531	25	90	25
1,0642 \rightarrow 0,5321	23	90	20
1,0642 \rightarrow 0,5321	23,2	90	24
1,0642 \rightarrow 0,5321	25,2	90	25,7
1,0642 \rightarrow 0,5321	25,2	90	17,5
1,0642 \rightarrow 0,5321	23	90	23,3
1,3188+0,6594 \rightarrow \rightarrow 0,4396	0	90	8,5
1,0642+0,809 \rightarrow \rightarrow 0,45961	90	90	122
1,0796 \rightarrow 0,5398	0	90	30
Плоскость yz			
0,9943 \rightarrow 0,49715	90	90	175
1,0642 \rightarrow 0,5321	90	69	100

В работе [6] исследовались температурные зависимости для двух длин кристалла RTP при ГВГ *sff*-типа на длине волны излучения 1,0642 мкм с изменением угла φ от 55,45° до 56,335° в плоскости xu (рис. 9). Хорошо видно, что при изменении угла φ распределение изменяется от критичного к не критичному по температуре синхронизму, достаточно точно соответствуя характерным зависимостям на рис. 1, а НКТС имеет второй порядок. Наличие двух максимумов на графиках показывает, что при уменьшении угла φ от величины 56,4° с ростом температуры $d\varphi_{FC}/dT > 0$, а при температуре 100 °C $d\varphi_{FC}/dT = 0$. При дальнейшем увеличении температуры $d\varphi_{FC}/dT < 0$. Таким образом, режим НКТС имеет место при $T = 100$ °C и $\varphi = 56,4$ °. Очевидно, что при комнатной температуре режим НКТС также может быть получен для процесса преобразования частоты с другой длиной волны. Величина угла φ , соответствующая ему, будет отличаться от 56,4°. В работе [6] такой режим не исследовался.

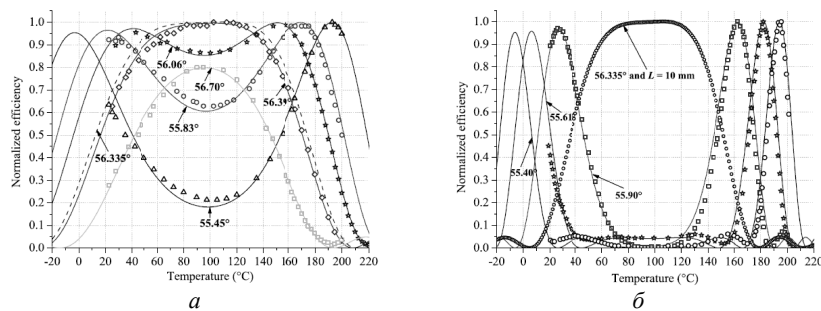


Рис. 9. Зависимость эффективности преобразования от температуры в кристалле RTP при ГВГ *sff*-типа на 1,0642 мкм [6]: *a* — $L = 5$ мм, *б* — $L = 10$ мм

В работе [7] в кристалле КТР длиной 3,5 мм при генерации суммарной частоты (ГСЧ) излучения с $\lambda_1 = 1,0642$ мкм и $\lambda_2 = 0,809$ мкм была получена температурная ширина 348 °C (рис. 10, *a*), а в работе [8] в кристалле КТР длиной 5 мм при ГВГ излучения $\lambda = 0,9943$ мкм — температурная ширина 350 °C (рис. 10, *б*). Характер обеих зависимостей показывает, что режим не является критичным по температуре.

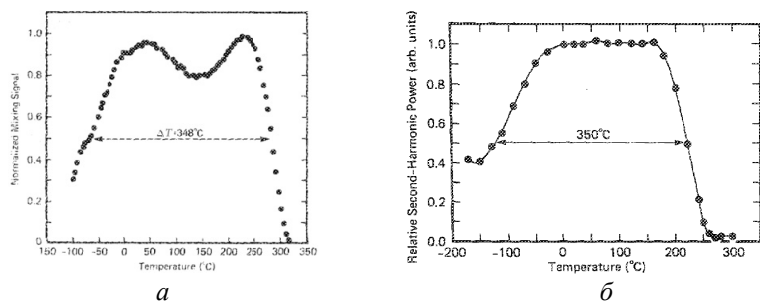


Рис. 10. Зависимость эффективности преобразования от температуры при ГСЧ (а) и ГВГ (б) в кристалле КТР (*sff*-тип взаимодействия)

Необходимо отметить одну систематическую ошибку, которая встречается в подавляющем числе публикаций. Измерения температурной ширины синхронизма производятся для кристаллов разной длины. В справочной литературе принято указывать этот параметр в пересчете на 1 см. Для критичного синхронизма справочная величина, определяемая по данным измерений, $\Delta T [^{\circ}\text{C}\cdot\text{см}] = \Delta T_{\text{изм}} \cdot L_{\text{изм}} [^{\circ}\text{C}] \cdot [\text{см}]$. Для некритичного синхронизма второго порядка температурная ширина определяется как $\Delta T [^{\circ}\text{C}\cdot\text{см}^{1/2}] = \Delta T_{\text{изм}} \cdot L_{\text{изм}}^{1/2} [^{\circ}\text{C}] \cdot [\text{см}^{1/2}]$ (см. (3)). Но в подавляющем большинстве публикаций этот режим определяется как нечувствительный по температуре и производится линейный пересчет на длину 1 см. Это хорошо видно из сравнения ширин кривых на рис. 10 с данными, приведенными в табл. 1. На самом деле зависимость от длины является корневой (как на рис. 9) или не описывается простой связью (как на рис. 10). Сравнение результатов, представленных на рис. 9, а и б, показывает, что при увеличении длины кристалла в 2 раза температурная ширина не уменьшается в 2 раза. В общем случае для некритичных по температуре процессов в разложении (1) соизмеримый вклад дают производные различного порядка (вплоть до 5—7-го) и строго определить зависимость ширины от длины кристалла не представляется возможным.

Аномально большие значения температурных ширин синхронизма получены и для других кристаллов (табл. 2) [4].

Таблица 2

Кристалл	Тип	Процесс	$\lambda_{\text{ПГС}}$, МКМ	$\lambda_{\text{ГСЧ}}$, МКМ	(θ ; φ)	ΔT , $^{\circ}\text{C}\cdot\text{см}$
ВІВО	1	ГСЧ	0,8186	0,4627	44,8°; 0	22
	2	ГСЧ	1,5203	0,6260		154
	1	ГВГ	0,8769	0,4385	47,1°; 0	12
	2	ГСЧ	1,3532	0,5957		54
УСОВ	1	ГСЧ	0,8077	0,4592	90°; 45,3°	Большая
	2	ГСЧ	1,5594	0,6325		151
	1	ГВГ	0,8664	0,4332	90°; 51,4°	75
	2	ГСЧ	1,379	0,6007		79

Точное определение режима НКТС приведено в работах [9, 10]. Было найдено, что при ГВГ на длине волны 1,0642 мкм в кристалле КТР для *sff*-типа взаимодействия существует срез ($\varphi = 67^\circ$, $\theta = 71^\circ$), для которого $d\Delta k/dT = 0$. В окрестности этого направления $d\Delta k/dT \neq 0$. В соответствии с общим определением для процессов по другим параметрам (длине волны и угловым отклонениям оси лазерного излучения от направления ФС) реализуется режим НКТС. Экспериментально при нагреве кристалла была получена зависимость, приведенная на рис. 11. Ширина кривой при нагреве составила 210 °С.

Очевидно, что это не единственный процесс преобразования частоты, для которого имеет место режим НКТС. Для анализа функциональных возможностей получения этого режима на различных длинах волн было предложено по аналогии с режимом НКЧС (см. рис. 8) определить совокупность направлений не критичных по температуре взаимодействий (НКТВ), для которых $d\Delta k(\varphi, \theta)/dT = 0$ [11]. Вдоль этих направлений в общем случае величина $\Delta k(\varphi, \theta)$ изменяется, но является термостабильной. Вдоль направлений ФС $\Delta k(\varphi, \theta) = 0$, но в общем случае величина $d\Delta k(\varphi, \theta)/dT \neq 0$. Направления, в которых пересекаются распределения ФС и НКТВ, являются направлениями НКТС: (НКТС) = (ФС) + (НКТВ). На рисунке 12 приведены распределения эффективности преобразования (2), максимальные значения которых (светлые области) соответствуют направлениям ФС, и направления НКТВ для случая ГВГ *sff*-типа взаимодействия (сплошная линия). Из этих результатов видно, что НКТВ существует в диапазоне 0,994÷1,205 мкм, а режим НКТС — в диапазоне 1,0265÷1,0967 мкм. Все они имеют различные направления (значения длин волн и углы даны в подписных полях).

Таким образом, в широком диапазоне длин волн может быть получен режим НКТС. Очевидно, что это возможно для различных процессов преобразования частоты (ГСЧ, генерации разностной частоты и ПГС).

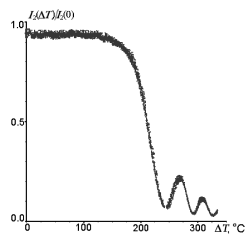


Рис. 11. Зависимость эффективности преобразования от температуры при ГВГ в кристалле КТР *sff*-типа

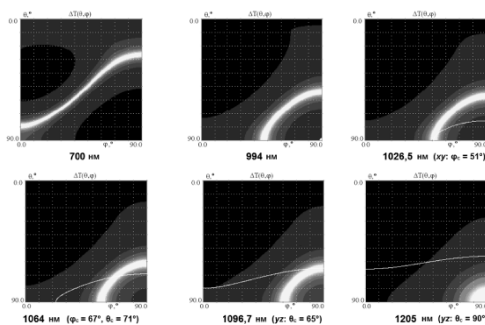


Рис. 12. Дисперсия направлений НКТВ и НКТС при ГВГ в кристалле КТР *sff*-типа

В работе [12] рассматривалась возможность реализации режима НКТС в кристалле LBO длиной 14,1 мм при генерации третьей гармоники (ГТГ) на длине волны 1,0642 мкм *ssf*-типа взаимодействия. Для среза кристалла $\varphi = 39,3^\circ$; $\theta = 74,8^\circ$ была получена температурная ширина 73 °С (рис. 13).

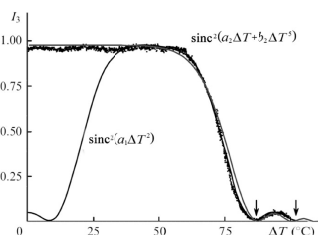


Рис. 13. Зависимость эффективности преобразования от температуры в кристалле LBO при ГТГ

на длине волны 1,0642 мкм *ssf*-типа взаимодействия. Для среза кристалла $\varphi = 39,3^\circ$; $\theta = 74,8^\circ$ была получена температурная ширина 73 °С (рис. 13).

С точной привязкой к нулям измеренной зависимости $\eta(T)$ был построен график в соответствии с выражением (2), аргументом которого является вторая производная волновой расстройки по температуре (сплошная кривая на рис. 13). Сравнение его с экспериментально измеренным позволяет

определить, что получен НКТС не ниже 5-го порядка. Очевидно, что в случае ГВГ в кристалле КТР (см. рис. 11) порядок НКТС также не ниже.

Все основные результаты по НКТС даны для двухосных кристаллов. Но этот режим возможен также в одноосных кристаллах. В работе [13] представлены зависимости угла синхронизма от длины волны для *eeo*-типа взаимодействия при ГВГ в кристалле ZGP ($ZnGeP_2$) при различных температурах (рис. 14). Из этих результатов следует, что при $\lambda = 4,5$ мкм в диапазоне температур от -180 °С до $+400$ °С угол синхронизма остается неизменным. Это соответствует режиму НКТС. Но прямые измерения (температурной зависимости угла синхронизма) не проводились.

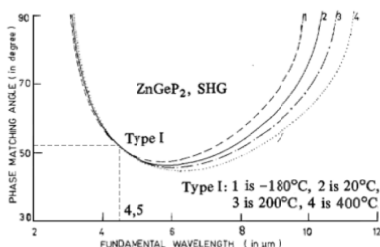


Рис. 14. Дисперсионная зависимость угла ФС при различных температурах кристалла ZGP [13]

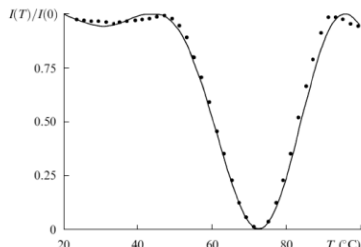


Рис. 15. Температурная зависимость двулучепреломления в кристалле LBO

До настоящего времени для новых синтезируемых кристаллов измеряются только производные первого порядка dn_i/dT . С использованием этих данных можно определить направление НКТС, но невозможно вычислить температурную ширину. Необходимы более точные измерения.

Выше рассмотрены не критичные по температуре процессы преобразования частоты. Но они не единственные процессы, для которых осуществим этот режим. В работах [14, 15] теоретически и экспериментально до-

казано, что возможна реализация не критичного по температуре двулучепреломления, и для кристалла LBO при увеличении температуры от комнатной получена температурная ширина $62\text{ }^\circ\text{C}$ (рис. 15). Показано, что если режим НКТС при преобразовании частоты существует в ограниченном диапазоне длин волн, то для двулучепреломления он существует во всем диапазоне прозрачности кристалла, так как вдоль оптической оси кристалла показатели преломления для двух волн всегда равны при любой температуре.

В работах [14, 15] сделан общий вывод: для всех процессов, эффективность которых определяется разностью волновых векторов, возможна реализация не критичного по температуре процесса. Пока исследовано небольшое количество процессов, остальное еще впереди. Обобщение результатов исследований влияния изменения температуры на процессы в оптических кристаллах приведено в работе [16].

Говоря о температурных процессах, нельзя не остановиться на действии еще одного механизма — влияния термодформационных изменений кристалла [17]. Проявляется оно в изменении формы кристалла при нагревании, его повороте и изменении угла преломления на входной грани. Практически все кристаллы имеют анизотропию коэффициента линейного теплового расширения (КЛТР) α_i . Кристалл LBO отличается большими значениями КЛТР и противоположными знаками главных значений коэффициентов: в диапазоне температур $25\div 70\text{ }^\circ\text{C}$ $\alpha_x = (58,7\div 71,1)\cdot 10^{-6}\text{ K}^{-1}$, $\alpha_y = -(48,9\div 59,0)\cdot 10^{-6}\text{ K}^{-1}$, $\alpha_z = (25,1\div 29,7)\cdot 10^{-6}\text{ K}^{-1}$. Если кристалл вырезан вдоль главных осей, то при нагревании размеры меняются пропорционально, без изменения его формы.

При нагревании кристалла, имеющего, например, срез $\varphi = 45^\circ$ в плоскости xu ($\theta = 90^\circ$), анизотропия КЛТР приводит к тому, что если первоначально он имел форму квадрата, то в дальнейшем приобретает форму ромба. На рисунке 16, а показано изменение формы в плоскости xu . Вдоль оси u кристалл сжимается, а вдоль оси x — расширяется. На практике требуется решить, как фиксировать кристалл — прижимом к основанию входной грани (рис. 16, б) или боковой грани (рис. 16, в). Фиксация производится легким прижимом, без механического напряжения.

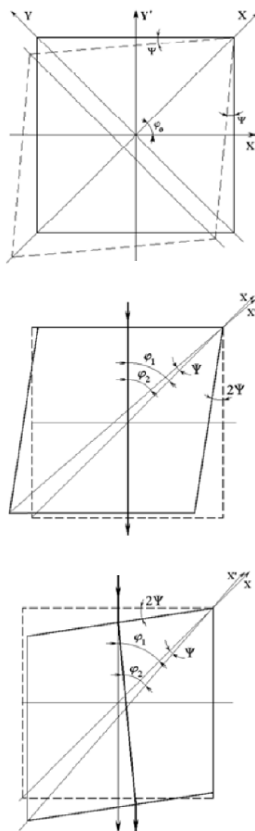


Рис. 16. Изменение формы кристалла LBO при нагреве

В первом случае (рис. 16, б) ось x кристалла поворачивается по часовой стрелке (изменение угла между осью излучения и осью x кристалла $d\Delta\varphi/dT > 0$), а во втором (рис. 16, в) — в обратном направлении ($d\Delta\varphi/dT < 0$). Также во втором случае происходит изменение угла преломления на входной грани кристалла. Из этих результатов следует, что направление распространения излучения изменяется относительно оси x при изменении температуры кристалла и знак этого изменения противоположный при разных способах фиксации кристалла. Одновременно при нагревании кристалла меняется направление ФС ($d\varphi_{\text{ФС}}/dT$) из-за температурных изменений показателей преломления кристалла. Знак и величина этого изменения не зависят от способа фиксации кристалла.

Величина температурной ширины синхронизма определяется величиной температурной производной изменения направления распространения

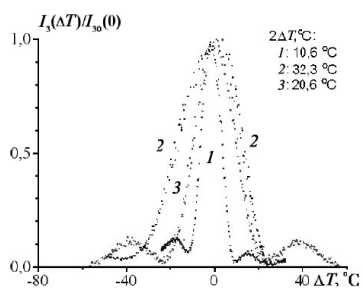


Рис. 17. Температурные зависимости ГТГ в кристалле LBO при различных фиксациях

излучения относительно направления ФС [17]. Величина $d\Delta\varphi/dT > 0$ может увеличивать или уменьшать суммарную расстройку по углу между осью излучения и направлением ФС. Это подтверждают результаты для различных фиксаций кристалла, представленные на рис. 17. Измерения проводились при генерации третьей гармоники на длине волны 1,0642 мкм в кристалле LBO длиной 5,33 мм со срезом в плоскости xz ($\varphi = 37,3^\circ$), тип взаимодействия *ssf*. Кривая 1 соответствует фиксации кристалла прижимом входной грани

к основанию, кривая 2 — прижиму боковой грани, а кривая 3 — прижиму по вертикали. Для исследованного случая максимальная величина температурной ширины синхронизма превышает минимальную более чем в 3 раза.

Очевидно, что при $\varphi = 45^\circ$ влияние термодформационных изменений кристалла на температурную ширину процесса будет максимальным и отношение ширин синхронизма при различных фиксациях кристалла будет наибольшим. Также нетрудно видеть, что и при действии этого эффекта может быть реализован режим не критичного по температуре синхронизма.

5. Преобразование частоты импульсов фемтосекундной длительности

Рассмотрим особенности нелинейно-оптического преобразования частоты при ГСЧ импульсов фемтосекундной длительности. У них большая ширина спектра, и для получения высокой эффективности преобразования необходима реализация не критичного по частоте синхронизма. Далеко не

все кристаллы имеют возможность обеспечить НКЧС на заданных длинах волн. Возможны варианты решений, когда этот режим выполняется для двух из трех взаимодействующих волн: 1-й (λ_1) и 2-й (λ_2), 1-й (λ_1) и 3-й (λ_3), 2-й (λ_2) и 3-й (λ_3). Как отмечалось выше, режим НКЧС соответствует равенству групповых скоростей ($v_{гр,i}$) для соответствующих комбинаций взаимодействующих волн. Рассмотрим основные варианты ГСЧ импульсов длительностью 100 фс.

1. НКЧС для всех трех волн ($v_{гр,1} = v_{гр,2} = v_{гр,3}$) (рис. 18).

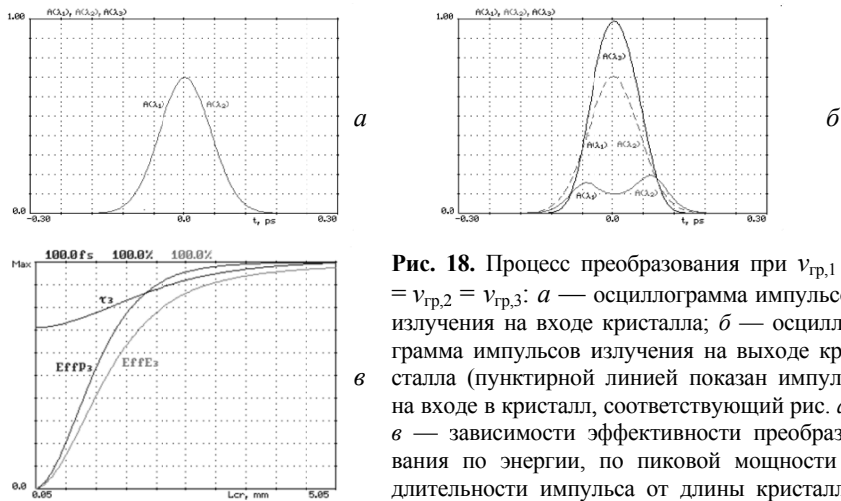


Рис. 18. Процесс преобразования при $v_{гр,1} = v_{гр,2} = v_{гр,3}$: *a* — осциллограмма импульсов излучения на входе кристалла; *б* — осциллограмма импульсов излучения на выходе кристалла (пунктирной линией показан импульс на входе в кристалл, соответствующий рис. *a*); *в* — зависимости эффективности преобразования по энергии, по пиковой мощности и длительности импульса от длины кристалла. Максимальные значения для оси ординат приведены над сеткой графика

На входе в кристалл импульсы с λ_1 и λ_2 согласованы во времени. И это сохраняется в процессе преобразования частоты на всей длине кристалла. Равенство групповых скоростей обеспечивает согласование во времени и для импульса формируемого излучения. Общий характер зависимостей показывает, что эффективности преобразования по энергии и пиковой мощности асимптотически стремятся к 100 %. Длительность импульса формируемой волны λ_3 в начале кристалла равна $\tau_1/\sqrt{2}$ и в процессе преобразования увеличивается до τ_1 . Некоторое симметричное изменение формы импульсов 1-й и 2-й волн на выходе кристалла обусловлено тем, что наибольшая эффективность преобразования будет при максимальной плотности мощности этих волн. Это имеет место на их вершине.

При $v_{гр,1} = v_{гр,2} = v_{гр,3}$ практически нет никакого ограничивающего действия волновой расстройкой для спектров всех импульсов взаимодействующих волн.

2. НКЧС для 1-й и 2-й волн ($v_{гр,1} = v_{гр,2} < v_{гр,3}$) (рис. 19).

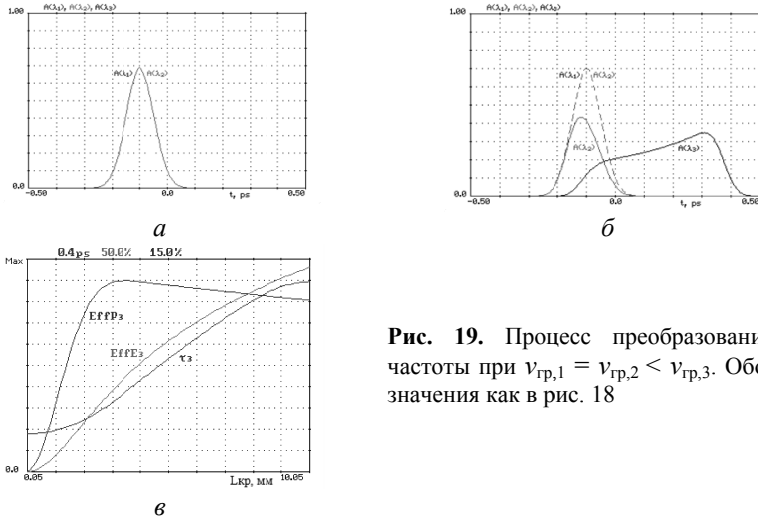


Рис. 19. Процесс преобразования частоты при $v_{гр,1} = v_{гр,2} < v_{гр,3}$. Обозначения как в рис. 18

Импульсы обеих волн входного излучения в процессе преобразования всегда согласованы во времени и формируют излучение 3-й, результирующей волны на всей длине взаимодействия. Но отличие групповой скорости формируемой волны с λ_3 от входных приводит к значительному увеличению ее длительности. На начальной длине кристалла длительность импульса формируемой волны такая же, как в предыдущем случае. Расстройка групповых скоростей еще не дает заметной задержки во времени. В процессе преобразования передняя часть импульса формируется в начале кристалла, а задняя — в его конце. Расстройка групповых скоростей приводит к значительному смещению во времени передней части импульса относительно 1-го и 2-го импульсов. В результате получается излучение с достаточно большой эффективностью преобразования по энергии, но небольшой эффективностью по пиковой мощности. Длительность импульса будет значительно большей по сравнению со случаем 1. Уменьшение эффективности преобразования по пиковой мощности (рис. 19, в) связано с наличием нерезонансных потерь в среде.

Видно, что при меньших значениях плотности мощности входного излучения импульс результирующей волны будет практически прямоугольным.

Аналогичная ситуация имеет место и при противоположном отношении групповых скоростей ($v_{гр,1} = v_{гр,2} > v_{гр,3}$). Но в этом случае импульс 3-й волны будет отставать от формирующих его волн.

3. НКЧС для 1-й и 3-й волн ($v_{гр,1} = v_{гр,3} \neq v_{гр,2}$).

Рассмотрим три частных случая при разных начальных условиях для двух волн на входе нелинейного кристалла.

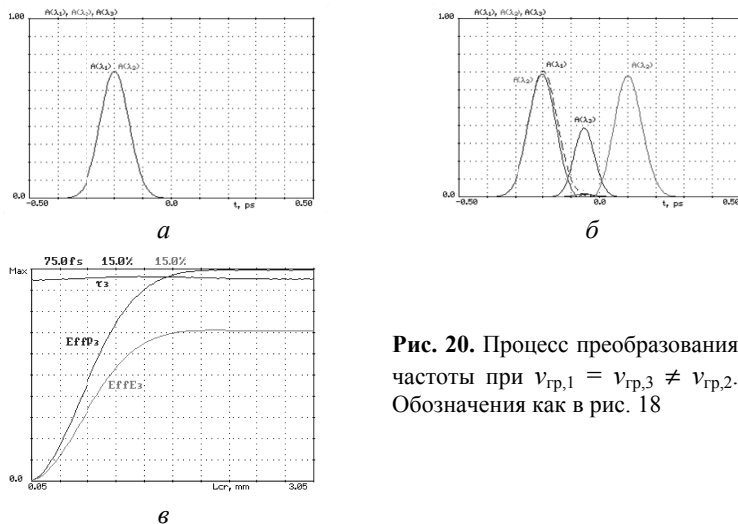


Рис. 20. Процесс преобразования частоты при $v_{гр,1} = v_{гр,3} \neq v_{гр,2}$. Обозначения как в рис. 18

3«а». На входе в кристалл импульсы излучения 1-й и 2-й волн согласованы во времени (рис. 20).

До тех пор пока импульсы излучения перекрываются во времени на конечной длине взаимодействия, происходит преобразование частоты и формируется волна суммарной частоты. Преобразование частоты возможно только в области перекрытия импульсов на конечной длине кристалла. При смещении одного импульса входного излучения относительно другой область их перекрытия смещается во времени. Это приводит к тому, что импульс результирующего излучения также смещается во времени. Как только время задержки одного входного импульса относительно другого превышает длительности импульсов, процесс ГСЧ прекращается и начинается процесс обратного преобразования (генерация разностной частоты, ГРЧ) — из 1-й и 3-й волны во 2-ю. Это слабо, но заметно на рис. 20, б по характеру зависимости длительности импульса 3-й волны от длины кристалла. Кроме того, длительность этого импульса практически не изменяется в процессе распространения.

Аналогичная ситуация имеет место и при $v_{гр,2} = v_{гр,3} > v_{гр,1}$. Импульс 3-й волны будет отставать от формирующих его волн.

3«б». Несколько улучшить ситуацию позволяет предварительное смещение импульса 2-й волны во времени на величину не менее длительности импульса 1-й волны (рис. 21).

Как и в случае 3«а», длительность импульса результирующей волны остается практически неизменной на всей длине взаимодействия. Но удастся получить значительный выигрыш в эффективности преобразования по энергии и пиковой мощности.

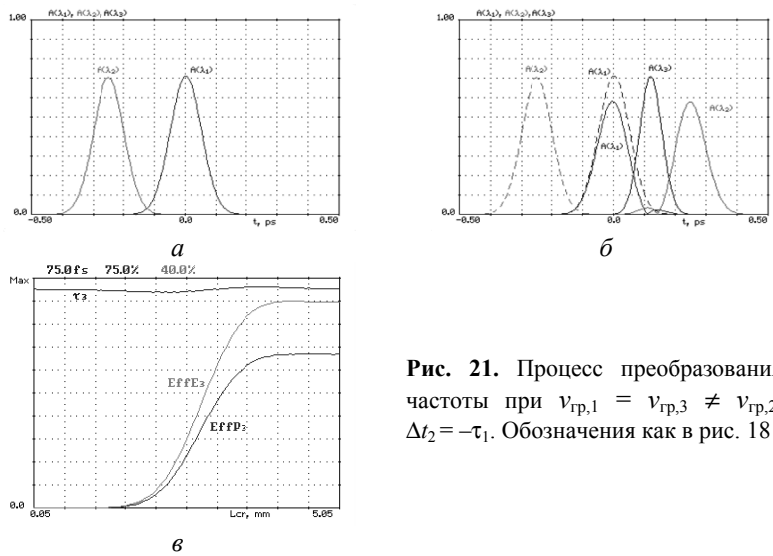


Рис. 21. Процесс преобразования частоты при $v_{гр,1} = v_{гр,3} \neq v_{гр,2}$, $\Delta t_2 = -\tau_1$. Обозначения как в рис. 18

3«в». Еще один вариант преобразования частоты при $v_{гр,1} = v_{гр,2} < v_{гр,3}$ — использование одного из импульсов входного излучения (например, 2-го) с длительностью большей, чем у другого, как это показано на рис. 22. Одновременно аналогично случаю 3«б» задается начальная задержка 1-го импульса на величину длительности 2-го импульса ($\Delta t_1 = -\tau_2$). Импульс 1-й волны «пробегает» через импульс 2-й, и в области их перекрытия происходит эффективное преобразование частоты. Импульс результирующей волны имеет небольшую длительность, величина которой слабо изменяется на всей длине взаимодействия (как и в случаях 3«а» и 3«б»). Величина ее определяется длительностью импульса 1-й волны. В этом случае удастся получить высокую эффективность преобразования, соизмеримую с той, которая получается при полном групповом синхронизме (случай 1). Эффективность преобразования по пиковой мощности может превышать 100%. В этом нет противоречия, так как мощность излучения 3-й волны формируется из всех частей 2-й.

Отличительной особенностью этого случая от всех предыдущих является то, что импульсы 1-й и 3-й волн всегда согласованы во времени. В области небольших значений плотности мощности 2-го импульса (на его переднем фронте при заданных параметрах начальных условий) начинается процесс обратного преобразования частоты — генерации разностной частоты. Происходит возрастание мощности 2-й волны. Это проявляется в наличии экстремума для зависимости эффективности преобразования по энергии и пиковой мощности от длины кристалла (рис. 22, в). Искажается форма импульса 1-й волны, что практически не сказывается на форме 3-й волны и длительности ее импульса.

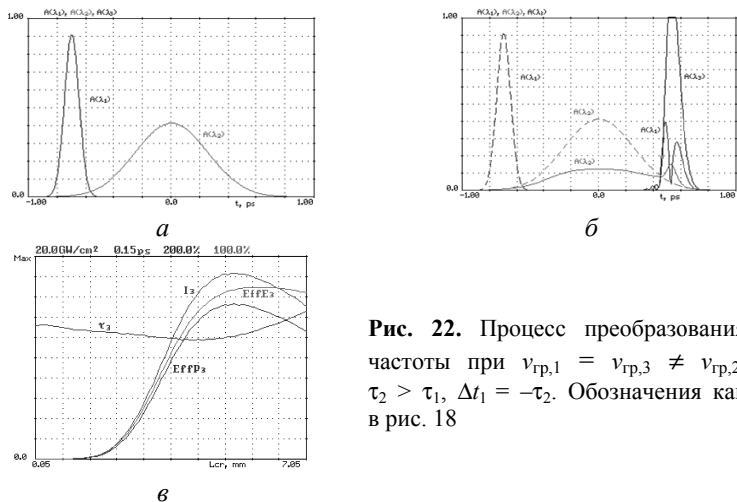


Рис. 22. Процесс преобразования частоты при $\nu_{гр,1} = \nu_{гр,3} \neq \nu_{гр,2}$, $\tau_2 > \tau_1$, $\Delta t_1 = -\tau_2$. Обозначения как в рис. 18

Подводя итог, необходимо отметить, что формирование импульсов с большой шириной спектра, которую имеют, например, импульсы фемтосекундной длительности, возможно для генерации суммарной частоты не только при полном групповом синхронизме, но и при выполнении его для одной из пар взаимодействующих волн.

Заключение

В этой статье представлены в наиболее общем виде результаты для критичных и некритичных процессов преобразования частоты в нелинейных кристаллах по основным параметрам, изменение которых влияет на эффективность, — длине волны и угловой расстройке оси лазерного излучения относительно направления фазового синхронизма, температуре кристалла; а также особенности преобразования частоты импульсов фемтосекундной длительности, когда групповой синхронизм выполняется для всех трех взаимодействующих волн или только для двух пар из них.

В недавно опубликованных работах [18, 19] приведены результаты анализа функциональных возможностей одноосных и двухосных кристаллов для различных задач преобразования частоты. Они являются хорошим дополнением к нашей статье.

Исследования температурных процессов, описанные в этой работе, были выполнены при поддержке гранта РФФИ 15-02-07760.

Литература

1. Дмитриев, В. Г. Прикладная нелинейная оптика / В. Г. Дмитриев, Л. В. Тарасов. — М.: Наука, 2004. — 512 с.

2. Гречин, С. Г. Полная классификация типов взаимодействия при генерации второй гармоники в двухосных нелинейных кристаллах / С. Г. Гречин, С. С. Гречин, В. Г. Дмитриев // Квантовая электроника. — 2000. — Т. 30, № 5. — С. 377—386.
3. Гречин, С. Г. Фазовый синхронизм и не критичные по частоте взаимодействия при преобразовании частоты импульсов фемтосекундной длительности // Квантовая электроника. — 2006. — Т. 36, № 1. — С. 45—50.
4. Nikogosyan, D. Nonlinear optical crystals : a complete survey / D. Nikogosyan. — N. Y. : Springer, 2005. — 437 p.
5. Гречин, С. Г. Угловые ширины синхронизма в двухосных нелинейных кристаллах преобразователей частоты // Квантовая электроника. — 2010. — Т. 40, № 9. — С. 822—827.
6. Mangin, J. Thermal expansion, normalized thermo-optic coefficients, and condition for second harmonic generation of a Nd:YAG laser with wide temperature bandwidth in RbTiOPO₄ / J. Mangin, G. Mennerat, P. Villeval // JOSA, B. — 2011. — V. 28, № 4. — P. 873—881.
7. Baumert, J. S. Generation of blue cw coherent radiation by sum frequency mixing in KTiOPO₄ / J. S. Baumert, F. M. Schellenberg, W. Lenth, W. P. Risk, G. Bjorklund // Appl. Phys. Lett. — 1987. — V. 51, № 26. — P. 2192—2194.
8. Risk, W. P. Noncritically phase-matched frequency doubling using 994 nm dye and diode laser radiation in KTiOPO₄ / W. P. Risk, R. N. Payne, W. Lenth, C. Harder, H. Meier // Appl. Phys. Lett. — 1989. — V. 55, № 12. — P. 1179—1181.
9. Гречин, С. Г. Аномально-не критичный по температуре фазовый синхронизм при преобразовании частоты в нелинейных кристаллах / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 1998. — Т. 25, № 11. — С. 963—964.
10. Гречин, С. Г. Некритичный по температуре синхронизм при ГВГ в кристалле КТР / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 1999. — Т. 26, № 1. — С. 77—81.
11. Гречин, С. Г. Дисперсия не критичных по температуре преобразования частоты и двулучепреломления в двухосных оптических кристаллах / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 2004. — Т. 34, № 5. — С. 461—466.
12. Гречин, С. Г. Некритичная по температуре генерация третьей гармоники в кристалле LBO / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 2004. — Т. 34, № 6. — С. 565—568.
13. Bhar, G. C. Temperature dependent phase-matching nonlinear optical device using CdSe and ZnGeP₂ / G. C. Bhar, G. C. Ghosh // IEEE J. Quantum Electron. — 1980. — V. 16, № 8. — P. 838—843.
14. Гречин, С. Г. Аномально-не критичное по температуре двулучепреломление в двухосных оптических кристаллах / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 2000. — Т. 30, № 1. — С. 1—2.
15. Гречин, С. Г. Аномально-не критичное по температуре двулучепреломление в двухосном оптическом кристалле LBO / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Квантовая электроника. — 2000. — Т. 30, № 4. — С. 285—286.
16. Гречин, С. Г. Некритичные по температуре процессы при распространении лазерного излучения в оптических кристаллах / С. Г. Гречин, В. Г. Дмитриев, В. А. Дьяков, В. И. Прялкин // Изв. АН РФ. Сер. физическая. — 2002. — Т. 66, № 8. — С. 1103—1107.
17. Grechin, S. G. The influence of thermal deformation processes on frequency conversion in an LBO crystal / S. G. Grechin, Yu. D. Arapov, V. A. Dyakov, I. V. Kasyanov // Laser Physics Letters. — 2014. — V. 11, № 12. — P. 125402.
18. Андреев, Ю. М. Функциональные возможности нелинейных кристаллов для преобразования частоты : одноосные кристаллы / Ю. М. Андреев, Ю. Д. Арапов, С. Г. Гречин, И. В. Касьянов, П. П. Николаев // Квантовая электроника. — 2016. — Т. 46, № 1. — С. 33—38.
19. Андреев, Ю. М. Функциональные возможности нелинейных кристаллов для преобразования частоты : двухосные кристаллы / Ю. М. Андреев, Ю. Д. Арапов, С. Г. Гречин, И. В. Касьянов, П. П. Николаев // Квантовая электроника. — 2016. — Т. 46, № 11. — С. 995—1001.

МЕТОД МНИМОГО ВРЕМЕНИ В ФИЗИКЕ СИЛЬНЫХ ЛАЗЕРНЫХ ПОЛЕЙ

С. В. Попруженко

Национальный исследовательский ядерный университет «МИФИ», г. Москва

Введение

В теоретической физике широко используется подход, основанный на формальном рассмотрении некоторых величин, являющихся по своему физическому смыслу вещественными, как комплексных функций комплексного аргумента. Такое расширение области определения позволяет применять мощные методы теории функции комплексного переменного и другие эффективные математические приемы для решения задач, рассмотреть которые иными способами оказывается гораздо сложнее или вообще не удается. В качестве общеизвестных примеров можно указать на исследование свойств амплитуды рассеяния как аналитической функции комплексных энергии или волнового вектора в квантовой механике [1] или на метод комплексного потенциала в гидродинамике несжимаемой идеальной жидкости [2].

В связи с быстрым прогрессом лазерной техники, позволяющим получать в лабораториях все более сильные и короткие лазерные импульсы, развитие аналитических методов описания взаимодействия сильных электромагнитных полей с атомами, молекулами и другими квантовыми системами остается актуальной задачей в течение нескольких последних десятилетий. В данной статье кратко излагаются принципы одного из таких методов, известного как «метод мнимого времени». Метод мнимого времени, впервые сформулированный в работах А. М. Переломова и В. С. Попова в середине 60-х годов прошлого века [3], имеет много общего с методом Ландау — Дыхне [4], предложенным несколько ранее для решения задачи о возбуждении квантового осциллятора с переменной во времени частотой.

В статье показано, каким образом метод мнимого времени может быть введен на примере вычисления амплитуд сильно нелинейных квантовых процессов в интенсивном электромагнитном поле и как он позволяет ввести в рассмотрение классические траектории заряженной частицы в комплексном пространстве и времени. Представление амплитуд на языке классических комплексных траекторий не только значительно упрощает численные расчеты, но и позволяет построить наглядную (хотя и формальную) картину квантовых процессов, обладающую большой предсказательной силой и позволяющую исследовать новые эффекты. В качестве примера в статье дается приложение метода мнимого времени к описанию кулоновских эффектов в спектрах нелинейной ионизации атомов — задача, для решения которой этот метод оказался наиболее эффективным,

а возможно, и единственным, позволившим получить количественно корректные результаты ценой относительно несложных вычислений.

Подробное изложение техники метода мнимого времени и его приложений в атомной физике сильных лазерных полей можно найти в обзорах [5—8].

1. Приближение сильного поля

Основная идея, позволяющая аналитически развить непertурбативное описание квантовых процессов в интенсивном лазерном поле, была предложена в известной работе Л. В. Келдыша [9]. Она заключается в том, чтобы использовать в качестве волновых функций конечного состояния системы «атом + сильное лазерное поле» плоские волковские волны — точные решения уравнения Шрёдингера (или — для релятивистских частиц — уравнений Кляйна — Гордона или Дирака) в поле плоской электромагнитной волны [10]. В случае, когда релятивистские эффекты несущественны, что, например, для лазера с длиной волны около 1 мкм отвечает интенсивности, не превышающей 10^{18} Вт/см², можно использовать простейший вариант волковской функции — решение нерелятивистского уравнения Шрёдингера в дипольном приближении:

$$\Psi_p(r, t) = \frac{1}{(2\pi\hbar)^{3/2}} \exp \left\{ i \frac{\mathbf{v}_p(t) \mathbf{r}}{\hbar} - \frac{im}{2\hbar} \int_0^t dt' \mathbf{v}_p^2(t') \right\}. \quad (1)$$

Здесь m — масса электрона, а $\mathbf{v}_p(t) = \frac{\mathbf{p} - e\mathbf{A}(t)}{m}$ — его скорость в поле лазерной волны, задаваемой векторным потенциалом $\mathbf{A}(t)$; e — заряд электрона, а \mathbf{p} — его канонический импульс — величина, сохраняющаяся в поле $\mathbf{A}(t)$ и совпадающая с кинематическим импульсом, измеряемым детектором, при условии, что электрическое поле волны достаточно быстро выключается во времени. В случае фемтосекундных лазерных импульсов это условие всегда выполняется. В более длинных импульсах нужно учитывать эффект пондеромоторного рассеяния (см., например, раздел 3 обзора [7] и ссылки в нем). Далее мы будем использовать атомную систему единиц: $|e| = m = \hbar = 1$.

В *приближении Келдыша*, которое также часто называется в литературе *приближением сильного поля*, амплитуда ионизации атомной системы записывается в виде

$$M(\mathbf{p}) = -i \int_{-\infty}^{+\infty} \langle \Psi_p | V(t) | \Psi_0 \rangle dt, \quad (2)$$

где $\Psi_0(\mathbf{r}, t) = \phi_0(\mathbf{r}) \exp\{iI_p t\}$ — волновая функция связанного состояния с потенциалом ионизации I_p , а энергия взаимодействия электрона с электрическим полем волны $\mathbf{E}(t) = -\dot{\mathbf{A}}(t)$ имеет вид $V(t) = \mathbf{E}(t) \mathbf{r}$.

2. Метод мнимого времени

После вычисления пространственного интеграла амплитуду (2) можно представить в виде

$$\begin{aligned} M(\mathbf{p}) &= \int_{-\infty}^{+\infty} P(\mathbf{p}, t) \exp\left\{iI_p t + \frac{i}{2} \int_0^t (\mathbf{p} + \mathbf{A}(t'))^2 dt'\right\} dt \equiv \\ &\equiv \int_{-\infty}^{+\infty} \exp\{iW_0(\mathbf{p}, t)\} dt. \end{aligned} \quad (3)$$

Здесь $P(\mathbf{p}, t)$ — предэкспоненциальный фактор, возникающий от интегрирования по пространству и определяемый видом волновой функции связанного состояния. Для дальнейшего изложения конкретный вид этой функции неважен. Как будет показано ниже, фаза $W_0(\mathbf{p}, t)$ представляет собой классическое действие для электрона в поле лазерной волны. В сильном низкочастотном поле, определяемом неравенствами

$$K_0 = \frac{I_p}{\omega} \gg 1, \quad z_F = \frac{E_0^2}{\omega^3} \gg 1, \quad (4)$$

где E_0 есть амплитуда электрического поля $\mathbf{E}(t)$, а ω — его частота, фаза в (3) велика в силу параметров K_0 и z_F (которые называются соответственно параметром многоквантовости [5] и параметром сильного поля [11]), что позволяет вычислить интеграл по времени в (3) методом перевала. Условие для перевальных точек t_s имеет вид

$$\frac{\partial W_0}{\partial t} = 0, \quad \text{т. е. } (\mathbf{p} + \mathbf{A}(t_s))^2 + 2I_p = 0. \quad (5)$$

Из уравнения (5) очевидно, что все его решения t_s комплексны. Несложный анализ функции $W_0(\mathbf{p}, t)$, который мы здесь не приводим, показывает, что контур интегрирования всегда должен проводиться через точки, лежащие в верхней полуплоскости комплексного времени. Нумеруя эти точки индексом α , можем записать амплитуду (3) в виде суммы вкладов всех таких точек:

$$M(\mathbf{p}) \approx \sum_{\alpha} P(\mathbf{p}, t_{s\alpha}) \exp\{iW_0(\mathbf{p}, t_{s\alpha})\}. \quad (6)$$

Формула (6) представляет собой стандартный результат применения метода перевала для вычисления интегралов от быстроосциллирующих функций. Важно, однако, что это выражение допускает простую классическую интерпретацию. Действительно, если рассматривать комплексную величину t_s как начальный момент времени движения электрона в поле $\mathbf{E}(t)$, то несложно убедиться в следующем: величина $W_0(\mathbf{p}, t_s)$ представляет собой действие, вычисленное вдоль траектории

$$\mathbf{r}_0(\mathbf{p}, t) = \mathbf{p}(t - t_s) + \int_{t_s}^t \mathbf{A}(t') dt', \quad (7)$$

являющейся решением уравнения Ньютона в комплексном времени

$$\ddot{\mathbf{r}}_0 = -\mathbf{E}(t) \quad (8)$$

с граничными условиями

$$\mathbf{r}(t_s) = 0, \quad \mathbf{v}_p(T) = \mathbf{p}. \quad (9)$$

В отличие от обычной задачи классической механики, в которой выбор начального момента времени произволен, здесь он определяется условием (5); кроме того, скорость задается не в начальный, а в конечный момент времени T (в аналитических расчетах удобно устремлять T к бесконечности), в начальный же момент времени она оказывается чисто мнимой величиной. Фаза $W_0(\mathbf{p}, t_s)$ может быть представлена в виде

$$W_0(\mathbf{p}, t) = \int_{t_s}^{+\infty} (L_0 - I_p) dt - \mathbf{v}_p \mathbf{r}_0 \Big|_{t_s}^T, \quad (10)$$

где $L_0(\mathbf{r}, \mathbf{v}) = \mathbf{v}^2/2 + \mathbf{E}(t)\mathbf{r}$ — функция Лагранжа электрона в поле $\mathbf{E}(t)$, и поэтому (10) представляет собой классическое укороченное действие (термин «укороченное» связан с наличием слагаемого $-I_p$, имеющего смысл начальной энергии системы [12]).

Таким образом, амплитуда нелинейной ионизации атома в сильном лазерном поле может быть представлена в виде суммы вкладов:

$$M(\mathbf{p}) \approx \sum_{\alpha} \exp\{iW_0[\mathbf{r}_0(\mathbf{p}, t_{s\alpha}, T)]\}, \quad (11)$$

вычисленных вдоль классических траекторий (7), которые представляют собой решение уравнения движения (8) в комплексном времени и пространстве. Поскольку действие $W_0(\mathbf{p}, t)$ — аналитическая функция времени (предполагается, что таковой является поле волны $\mathbf{E}(t)$), то интегрирование по времени в (10) можно осуществлять вдоль любого контура, соединяющего точки t_s и T (рис. 1). Изложенная выше процедура составляет суть метода мнимого времени (ММВ). Амплитуды более сложных процессов в сильных лазерных полях, таких как генерация высоких гармоник

лазерного излучения, многоэлектронная ионизация и др., также выражаются в виде интегралов, содержащих в показателе экспоненты большие по величине классические действия, поэтому и они могут быть представлены в виде суммы вкладов классических комплексных траекторий.

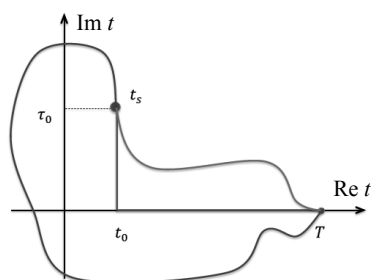


Рис. 1. Контуры интегрирования в комплексном времени, соединяющие начальный момент движения t_s , определяемый из (5), и момент T , в который электрон регистрируется детектором. Обычно в ММВ используется контур, состоящий из двух прямых отрезков, однако интегрирование вдоль любого другого контура приведет к тому же результату.

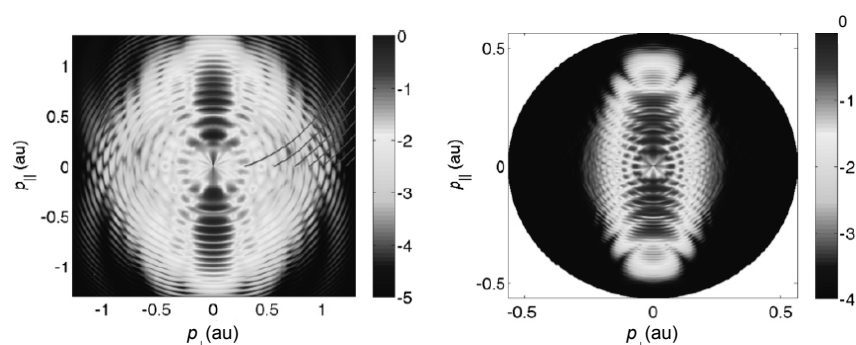


Рис. 2. Импульсные распределения фотоэлектронов при ионизации атома водорода линейно поляризованным (вдоль вертикальной оси) лазерным полем с длиной волны 800 нм и интенсивностью 10^{14} Вт/см². Слева — распределение, полученное из точного численного решения зависящего от времени уравнения Шредингера, справа — распределение, вычисленное в приближении Келдыша

Можно указать на два существенных преимущества применения ММВ по сравнению с вычислениями, не привлекающими понятия классических траекторий. Во-первых, язык траекторий позволяет наглядным образом (хотя и упрощенно) интерпретировать квантовую динамику атомной системы в поле лазерного излучения и поэтому обладает значительной предсказательной силой. Именно понятие об электроном, движущемся вдоль определяемой лазерным полем классической траектории, которая может

возвращаться к атому, позволило объяснить эффекты генерации высоких гармоник [13] и многократной некаскадной ионизации [14] и научиться управлять этими эффектами. Управление излучением высоких гармоник атомами в сильном лазерном поле лежит в основе *аттосекундной физики*, интенсивно развивающейся в последние годы (см., например, обзор [15] и ссылки в нем).

Во-вторых, ММВ позволяет развить специфическую *теорию возмущений по действию*, учитывающую влияние кулоновского взаимодействия между фотоэлектроном и родительским атомом на процесс ионизации. Из формул (1)—(3) очевидно, что в приближении Келдыша это взаимодействие полностью отброшено. Между тем оно существенно влияет как на вид импульсных распределений фотоэлектронов, так и на полную вероятность ионизации. Приведенные на рис. 2 импульсные распределения, рассчитанные для случая ионизации атома водорода, дают понятие о характерном масштабе этих отличий. В следующем разделе показано, каким образом влияние кулоновского взаимодействия на ионизацию может быть количественно учтено с помощью ММВ.

3. Влияние кулоновского взаимодействия

Предположим, что помимо лазерного поля $\mathbf{E}(t)$, где траектория электрона (7) легко вычисляется аналитически, присутствует еще дополнительная сила $\mathbf{F}(\mathbf{r})$, которая хотя и мала в большей части пространства по сравнению с лазерной, все же может оказывать существенное влияние на движение электрона. Наиболее важным примером такого рода является кулоновское взаимодействие электрона с имеющим заряд Z атомным остовом, определяемое выражениями

$$\mathbf{F}_C(\mathbf{r}) = -\frac{Z\mathbf{r}}{r^3}, \quad U_C(r) = -\frac{Z}{r}. \quad (12)$$

При выполнении неравенств (4) амплитуда колебаний электрона в лазерном поле, задающая пространственный масштаб траектории (7), обычно велика по сравнению с атомным размером $r_a \approx 1/\sqrt{2I_p}$, а также с расстоянием $r_0 \approx \sqrt{Z/E_0}$ от атома, на котором кулоновская сила равна по абсолютной величине лазерной. Поэтому влияние кулоновского взаимодействия на траекторию и его вклад в действие фотоэлектрона можно учесть по теории возмущений. Вместо уравнения (8) нужно решить уравнение

$$\ddot{\mathbf{r}} = -\mathbf{E}(t) - \frac{Z\mathbf{r}}{r^3}, \quad (13)$$

принимая величины (7), (10) за нулевое приближение и отыскивая кулоновские поправки к ним:

$$\mathbf{r}(t) = \mathbf{r}_0(t) + \mathbf{r}_1(t) + \dots, \quad \mathbf{v}(t) = \mathbf{v}_p(t) + \mathbf{v}_1(t) + \dots \quad (14)$$

Ограничиваясь членами нулевого и первого порядка по кулоновскому полю, получим, что к действию прибавляются поправки двух типов: из-за добавления к функции Лагранжа энергии кулоновского взаимодействия и из-за искажения кулоновским полем траектории фотоэлектрона, $W = W_0 + W_{C1} + W_{C2}$, где

$$W_{C1} = -\int_{t_s}^T U_C[\mathbf{r}_0(t)] dt = Z \int_{t_s}^T \frac{dt}{\sqrt{\mathbf{r}_0^2(t)}}, \quad (15)$$

$$W_{C2} = \int_{t_s}^T [L_0(\mathbf{r}_0 + \mathbf{r}_1) - L_0(\mathbf{r}_0)] dt. \quad (16)$$

Внеинтегральное слагаемое в (10) также получает соответствующее приращение. Таким образом, вклад в амплитуду нелинейной ионизации от каждой стационарной точки умножается на два комплексных фактора, зависящих от импульса, параметров лазерного поля и атома:

$$\tilde{M}_\alpha(\mathbf{p}) \propto \exp\{i[W_{0\alpha} + W_{C1\alpha} + W_{C2\alpha}]\} \equiv Q_{1\alpha} Q_{2\alpha} M_\alpha(\mathbf{p}). \quad (17)$$

Несмотря на то что действия (15) и (16) малы по абсолютной величине относительно $W_0(\mathbf{p}, t_s)$, они велики по сравнению с единицей, и поэтому их учет приводит к значительной модификации фотоэлектронных спектров.

Прежде чем обсудить некоторые примеры применения изложенной выше техники для расчета спектров фотоионизации, сформулируем кратко принципиальные трудности, которые могут возникать при вычислении кулоновских поправок ММВ.

1. Кулоновская сила не везде мала по сравнению с лазерной, а в начале координат, откуда стартует невозмущенная траектория (7), расходится.

2. Кулоновская поправка (15), вычисленная вдоль невозмущенной траектории в лазерном поле, логарифмически расходится на нижнем пределе.

3. Поправки к функции Лагранжа, входящие в действия (15) и (16), не являются аналитическими функциями комплексного времени, имеют на комплексной плоскости точки ветвления и разрезы, и поэтому выбор контура интегрирования требует специального рассмотрения.

Первые две проблемы решаются путем сшивания решения уравнения (13) и действий (15), (16), найденных вдали от атома, в области $r \gg r_0$, с соответствующими величинами, вычисленными вблизи атома с учетом только кулоновского поля. Такая сшивка оказывается возможной благодаря характерному для квазиклассического приближения перекрытию областей применимости противоположных приближений. Технически процедура сшивания достаточно трудоемка. Ее подробное описание можно найти в разделе 6.2 обзора [7].

Свойства функции $1/\sqrt{\mathbf{r}_0^2(\mathbf{p}, t)}$ в плоскости комплексного времени для траектории в линейно поляризованном лазерном поле также описаны

в обзоре [7] и более подробно — в недавней работе [16]. Оказывается, точки ветвления образуют пары, разделенные щелью, величина которой растет с ростом проекции импульса фотоэлектрона на направление поляризации поля [16]. От точек ветвления начинаются разрезы, уходящие в бесконечность вдоль мнимой оси. Топология расположения точек ветвления и разрезов такова, что всегда можно проложить контур интегрирования, соединяющий точки t_s и T без пересечения разрезов, но общепринятый контур с вертикальным опусканием на вещественную ось от стационарной точки возможен не при любых значениях импульса. Характерное расположение точек ветвления, разрезов и возможного контура интегрирования показано на рис. 3.

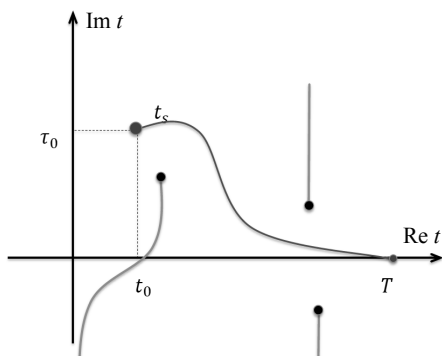


Рис. 3. Качественная картина расположения точек ветвления и разрезов функции $1/\sqrt{r_0^2(\mathbf{p}, t)}$ в плоскости комплексного времени при ионизации линейно поляризованным лазерным полем. В случае достаточно большого поперечного импульса точки ветвления расположены так, что интегрирование вдоль «стандартного» контура, показанного на рис. 1, невозможно.

4. Примеры применения ММВ

Приведем несколько примеров использования ММВ для расчета спектров фотоэлектронов и скорости ионизации атомов и ионов с учетом кулоновских эффектов.

4.1. Скорость ионизации атомов и ионов в сильном линейно поляризованном лазерном поле

Введенные выше амплитуды (11) и (17) определяют дифференциальную вероятность ионизации на элемент импульсного пространства $d^3 p$:

$$dw(\mathbf{p}) = |M(\mathbf{p})|^2 d^3 p. \quad (18)$$

Полная вероятность (скорость) ионизации получается интегрированием (18) по всем возможным значениям энергий электронов и углов их вылета и делением результата на длительность лазерного импульса. На практике, с учетом экспоненциального спада вероятности ионизации при росте p , оказывается достаточно разложить действие квадратично по импульсам вблизи точки $\mathbf{p} = 0$ и вычислить получившийся гауссов интеграл. По этой же причине для вычисления кулоновской поправки к полной вероятности ионизации нужно рассмотреть лишь траекторию, отвечающую $\mathbf{p} = 0$, которая в линейно поляризованном монохроматическом поле $\mathbf{E}(t) = E_0 \cos(\omega t)$ имеет вид $\mathbf{r}_0(\mathbf{p} = 0, t) = -(\mathbf{E}_0/\omega^2)(\cos(\omega t_s) - \cos(\omega t))$. Вдоль этой траектории кулоновские поправки могут быть вычислены в замкнутой аналитической форме. Подробности вычислений приведены в обзоре [7]. Их результат сводится к возникновению значительного по величине кулоновского фактора в скорости ионизации W :

$$W = Q_C W_0, \quad Q_C = \left(\frac{2}{F}\right)^{2\nu} \left(1 + \frac{2\gamma}{e}\right)^{-2\nu}, \quad e = 2,718\dots \quad (19)$$

Здесь W_0 — скорость ионизации, вычисленная без учета кулоновского взаимодействия (формально это отвечает случаю отрицательного иона, имеющего тот же потенциал ионизации, что и реальный атом), $F = E_0/(2I_p)^{3/2} \ll 1$ — приведенное лазерное поле (величина $E_{ch} = (2I_p)^{3/2}$ имеет смысл характерной напряженности электрического поля в атоме, на уровне с потенциалом ионизации I_p), $\nu = Z/\sqrt{2I_p}$ — эффективное главное квантовое число уровня и $\gamma = \sqrt{2I_p} \omega / E_0$ — параметр Келдыша [9].

Фактор (19) увеличивает скорость ионизации на много порядков. В туннельном режиме, $\gamma \ll 1$, второй множитель в (19) близок к единице — в таком виде кулоновский фактор был получен в работе [3] с помощью ММВ, впервые сформулированного там же. В многофотонном режиме, $\gamma \gg 1$, второй множитель становится малым и отчасти подавляет кулоновский фактор, который, тем не менее, остается большим. Масштаб кулоновских поправок иллюстрируется графиками зависимости скорости ионизации многозарядного иона ксенона от интенсивности лазерного излучения, показанными на рис. 4. Скорость ионизации, вычисленная в приближении Келдыша без учета кулоновского взаимодействия, показана кривой *st*. Домножение на кулоновский фактор приводит к увеличению вероятности ионизации на 9 (!) порядков (кривая *new*), причем оба множителя одинаково важны: их величина составляет порядка 10^{18} и 10^{-9} соответственно. Кривой *tunnel* показана скорость ионизации, вычисленная по статической туннельной формуле, т. е. для $\gamma = 0$.

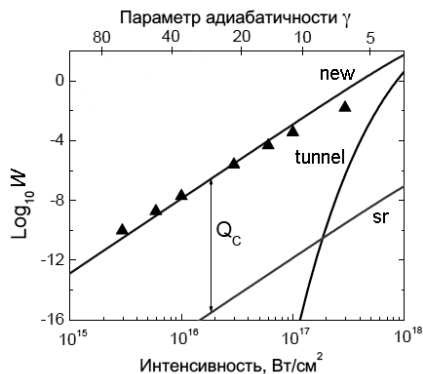


Рис. 4. Скорость ионизации Xe^{17+} полем рентгеновского лазера с энергией фотона 93 эВ (13 нм). Состояние $4p^0$ с потенциалом ионизации 434 эВ. Треугольники — скорость, найденная численным интегрированием зависящего от времени уравнения Шредингера; sr — результат расчета по теории Келдыша без учета кулоновского взаимодействия, т. е. для уровня в короткодействующем потенциале; new — с учетом кулоновского фактора Q_c ; tunnel — скорость ионизации, рассчитанная по туннельной формуле с учетом кулоновского взаимодействия, но для нулевого значения параметра адиабатичности

4.2. Новые типы траекторий

Кулоновское поле не только количественно деформирует траектории электронов, но и порождает новые, качественно меняя характер интерференционной картины в спектре фотоионизации. В поле с линейной поляризацией на каждом лазерном периоде имеется две траектории (7), отвечающие одному и тому же значению конечного импульса. На рисунке 5 они обозначены номерами I и II. Вдоль этих траекторий сохраняется поперечная по отношению к поляризации скорость электрона. Кулоновское поле разрушает этот интеграл движения, в результате чего появляются новые типы траекторий, вдоль которых поперечная проекция скорости меняет знак (обозначены номерами III и IV). Вклады новых траекторий интерферируют друг с другом и с вкладами траекторий I и II, создавая характерную картину, которая была обнаружена в эксперименте по ионизации метастабильных состояний ксенона полем инфракрасного лазера с длиной волны около 7 мкм [17]. Импульсное распределение фотоэлектронов, измеренное в эксперименте [17], показано справа на рис. 5. Расчеты, выполненные ММВ с учетом кулоновских траекторий [17, 18], продемонстрировали, что именно эти траектории приводят к появлению вертикальных интерференционных полос, хорошо различимых на экспериментальном графике. Интерференционные структуры, порождаемые кулоновским

взаимодействием, было предложено использовать для *фотоэлектронной голографии* атомных и молекулярных систем [17].

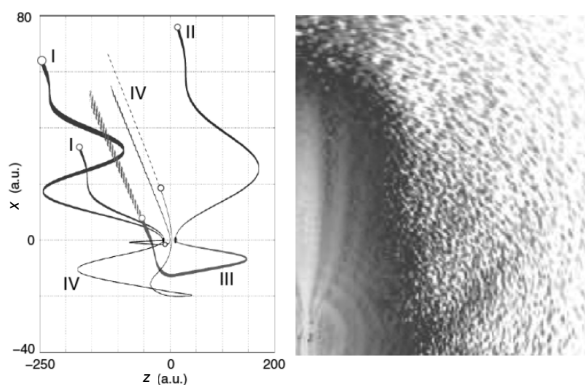


Рис. 5. Слева — возможные траектории фотоэлектрона в лазерном поле, поляризованном горизонтально (вдоль оси z), вычисленные с учетом кулоновского взаимодействия с атомным остатком. Все траектории отвечают одному и тому же значению импульса. Справа — импульсное распределение фотоэлектронов в плоскости (p_x, p_z) , измеренное в работе [17] для случая ионизации метастабильного состояния $6s$ атома ксенона лазерным полем с длиной волны 7 мкм и интенсивностью порядка 10^{12} Вт/см²

4.3. Асимметрия угловых распределений

В линейно поляризованном лазерном поле импульсные распределения фотоэлектронов существенно деформируются кулоновским взаимодействием, но остаются при этом симметричными относительно направления поляризации и перпендикулярной ему плоскости (речь идет о достаточно длинных импульсах, содержащих большое количество оптических периодов; в ультракоротких лазерных импульсах возникает дополнительная асимметрия, обусловленная наличием резкой временной огибающей). В лазерном поле с эллиптической или круговой поляризацией траектории электронов становятся двумерными, но в отсутствие кулоновского поля они остаются симметричными относительно большой и малой осей эллипса поляризации. На рисунке 6, *a* такие траектории, отвечающие углу вылета 30° относительно малой оси эллипса, показаны сплошными линиями. Угловое распределение фотоэлектронов обладает тем же уровнем симметрии. Характерное угловое распределение, вычисленное в рамках приближения сильного поля без учета кулоновского взаимодействия, показано на рис. 6, *б*; углы вылета, отвечающие показанным слева траекториям, обозначены точками.

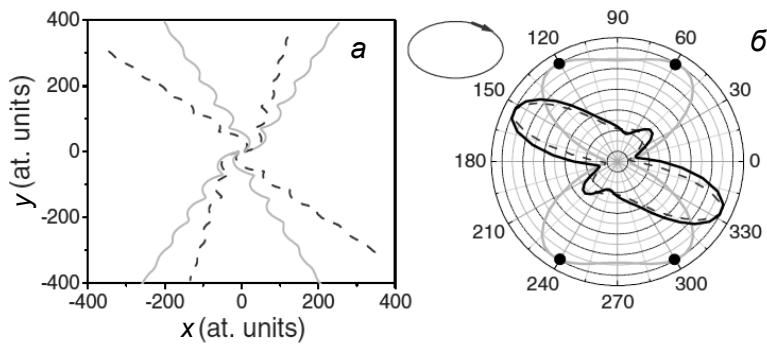


Рис. 6. Слева — траектории фотоэлектрона, отвечающие направлению вылета 60° по отношению к большой оси эллипса поляризации в случае ионизации атома водорода лазерным импульсом с длиной волны 800 нм, интенсивностью 10^{14} Вт/см² и эллиптичностью 0,5: сплошные линии — (симметричные) траектории, вычисленные без учета кулоновского взаимодействия; штриховые линии — траектории, стартующие с тем же начальным условием, но затем искаженные кулоновским полем. Соответствующие угловые распределения изображены справа. Там же сплошной черной линией показано распределение, полученное из точного численного решения зависящего от времени уравнения Шредингера. Из работы [20]

Кулоновское взаимодействие приводит к закручиванию траекторий (показаны на рис. 6, *a* пунктирными линиями) и деформации угловых распределений. В результате возникает качественный эффект разрушения четырехкратной симметрии угловых распределений в эллиптически поляризованном лазерном поле, наблюдающийся в эксперименте [19, 20]. Применение ММВ позволяет найти траектории с учетом обоих полей и вычислить угловые распределения, которые, как и следовало ожидать, обладают только центральной симметрией [20, 21].

Отметим, что так же, как и в случае кулоновских поправок к скорости ионизации, эффект оказывается относительно большим по величине: угол поворота распределения, показанного на рис. 6, составляет около 60° . На самом деле распределение не только поворачивается, но и деформируется из-за того, что вдоль разных траекторий электроны по-разному замедляются кулоновским полем. Суммарное действие этих двух факторов создает показанный на рис. 6 эффект. Рассчитанное методом мнимого времени угловое распределение хорошо воспроизводит результат численного интегрирования нестационарного уравнения Шредингера, подтверждая, таким образом, количественную точность ММВ.

Заключение

Помимо упомянутых в этой статье, метод мнимого времени имеет и другие эффективные приложения к теории нелинейных квантовых эффек-

тов в интенсивном электромагнитном поле. В частности, он применялся для рассмотрения ионизации в релятивистском режиме, вычисления спиновой поправки к скорости ионизации и расчета импульсных распределений фотоэлектронов в сверхкоротких лазерных импульсах. Эти приложения ММВ подробно обсуждаются в обзоре [8]. Поправки к действию и вероятности ионизации могут, конечно, быть найдены не только для кулоновского взаимодействия, но и в других случаях, в частности для систем, связанных дипольным взаимодействием (так называемые дипольные анионы) либо имеющих значительный постоянный или наведенный дипольный момент (некоторые большие молекулы), эффект от которого может конкурировать с кулоновским. Задача о вычислении кулоновского фактора, возникающего вследствие дипольного взаимодействия, остается пока нерассмотренной.

Литература

1. *Ландау, Л. Д.* Квантовая механика. Нерелятивистская теория / Л. Д. Ландау, Е. М. Лифшиц. — М. : Наука, 1989. — 768 с. — (Теоретическая физика : в 10 т. ; т. 3).
2. *Ландау, Л. Д.* Гидродинамика / Л. Д. Ландау, Е. М. Лифшиц. — М. : Наука, 1988. — 736 с. — (Теоретическая физика : в 10 т. ; т. 6).
3. *Переломов, А. М.* Ионизация атомов в переменном электрическом поле. III / А. М. Переломов, В. С. Попов // ЖЭТФ. — 1967. — Т. 52, вып. 2. — С. 514—526.
4. *Дыхне, А. М.* Адиабатическое возмущение состояний дискретного спектра // ЖЭТФ. — 1962. — Т. 41, вып. 4. — С. 1324—1327.
5. *Попов, В. С.* Туннельная и многофотонная ионизация атомов и ионов в сильном лазерном поле (теория Келдыша) // УФН. — 2004. — Т. 174, вып. 9. — С. 921—951.
6. *Попов, В. С.* Метод мнимого времени в атомной и ядерной физике // ЯФ. — 2005. — Т. 68, вып. 4. — С. 717—738.
7. *Popruzhenko, S. V.* Keldysh theory of strong field ionization: history, applications, difficulties and perspectives // J. Phys. B: At. Mol. Opt. Phys. — 2014. — V. 47, № 10. — P. 204001.
8. *Карнаков, Б. М.* Современное развитие теории нелинейной ионизации атомов и ионов / Б. М. Карнаков, В. Д. Мур, С. В. Попруженко, В. С. Попов // УФН. — 2015. — Т. 185, вып. 1. — С. 3—34.
9. *Келдыш, Л. В.* Ионизация в поле сильной электромагнитной волны // ЖЭТФ. — 1964. — Т. 47, вып. 5. — С. 1945—1956.
10. *Берестецкий, В. Б.* Квантовая электродинамика / В. Б. Берестецкий, Е. М. Лифшиц, Л. П. Питаевский. — М. : Наука, 1988. — 728 с. — (Теоретическая физика : в 10 т. ; т. 4).
11. *Reiss, H. R.* Effect of an intense electromagnetic field on a weakly bound system // Phys. Rev. A. — 1980. — V. 22, № 5. — P. 1786—1813.
12. *Ландау, Л. Д.* Механика / Л. Д. Ландау, Е. М. Лифшиц. — М. : Наука, 1988. — 216 с. — (Теоретическая физика : в 10 т. ; т. 1).

13. *Lewenstein, M.* Theory of high-harmonic generation by low-frequency laser fields / M. Lewenstein, Ph. Balcou, M. Yu. Ivanov, A. L'Huillier, P. B. Corkum // *Phys. Rev. A.* — 1994. — V. 49, № 3. — P. 2117—2132.
14. *Кучиев, М. Ю.* Атомная антенна // *Письма в ЖЭТФ.* — 1987. — Т. 45, вып. 7. — С. 319—321.
15. *Krausz, F.* Attosecond physics / F. Krausz, M. Ivanov // *Rev. Mod. Phys.* — 2009. — V. 81, № 1. — P. 163—234.
16. *Pisanty, E.* Slalom in complex time: Emergence of low-energy structures in tunnel ionization via complex-time contours / E. Pisanty, M. Ivanov // *Phys. Rev. A.* — 2016. — V. 93, № 4. — P. 043408.
17. *Huismans, Y.* Time-resolved holography with photoelectrons / Y. Huismans, A. Rousee, A. Gijsbertsen [et al.] // *Science.* — 2011. — V. 331. — P. 61—64.
18. *Yan, Tian-Min.* Low-energy structures in strong field ionization revealed by quantum orbits / Tian-Min Yan, S. V. Popruzhenko, M. J. J. Vrakking, D. Bauer // *Phys. Rev. Lett.* — 2010. — V. 105. — P. 253002.
19. *Bashkansky, M.* Asymmetries in above-threshold ionization / M. Bashkansky, P. H. Bucksbaum, D. W. Schumacher // *Phys. Rev. Lett.* — 1988. — V. 60. — P. 2458—2461.
20. *Popruzhenko, S. V.* Coulomb-corrected quantum trajectories in strong-field ionization / S. V. Popruzhenko, G. G. Paulus, D. Bauer // *Phys. Rev. A.* — 2008. — V. 77. — P. 053409.
21. *Goreslavski, S. P.* Coulomb asymmetry in above-threshold ionization / S. P. Goreslavski, G. G. Paulus, S. V. Popruzhenko, N. I. Shvetsov-Shilovski // *Phys. Rev. Lett.* — 2004. — V. 93. — P. 233002.

Нелинейные процессы в геофизике

АТМОСФЕРНЫЕ БЛОКИНГИ И СВЯЗАННЫЕ С НИМИ КЛИМАТИЧЕСКИЕ АНОМАЛИИ

И. И. Мохов

Институт физики атмосферы им. А. М. Обухова РАН, г. Москва
Московский государственный университет им. М. В. Ломоносова
Московский физико-технический институт, г. Долгопрудный

Введение

Сильнейшие региональные погодно-климатические аномалии, зависящие от глобальных изменений климата, связаны с атмосферными блокингами — блокированием западного зонального переноса в тропосфере средних широт с формированием длительных, до двух месяцев, аномалий циркуляционного и термического режимов, приводящим к засухам летом и сильным морозам зимой. Согласно имеющимся оценкам, есть существенный риск усиления региональной блокинговой активности со значительными последствиями при продолжении глобального потепления. При этом для процессов формирования атмосферных блокировок, их изменений и последствий характерны существенно нелинейные эффекты [1—44].

На фоне тенденции глобального потепления климата три последних десятилетия были по данным наблюдений самыми теплыми у поверхности Земли (с XIX века). И хотя с начала XXI века отмечается эффект замедления глобального потепления, последние годы были самыми теплыми на Земле. При этом скорость потепления в высоких широтах значительно больше глобальной и полушарной скорости (так называемое полярное усиление).

В России, как в северной стране, потепление идет существенно быстрее, чем для Земли в целом. В последние десятилетия скорость приповерхностного потепления в России в целом более чем вдвое превышала глобальную, а в отдельных регионах более чем в 4 раза. При этом количество опасных гидрометеорологических явлений со значительным социально-экономическим ущербом и общее число метеорологических опасных явлений в России заметно увеличивались в течение последних двух десятков лет. На рисунке 1 представлено изменение числа опасных атмосферных

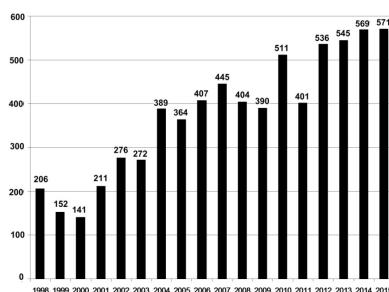


Рис. 1. Изменение числа опасных атмосферных (метеорологических) явлений в России (по данным Росгидромета)

(метеорологических) явлений в России по данным Росгидромета. В частности, в последние 5 лет количество гидрометеорологических явлений в России со значительным социально-экономическим ущербом было как минимум вдвое больше, чем в последние 5 лет XX века.

Особенности режимов блокирования и механизмы их формирования

Физические механизмы формирования атмосферных блокировок достаточно хорошо исследованы, но при этом остается много проблем. Формирование блокировок плохо предсказуемо, и результаты модельных оценок тенденций изменения блокировочной активности сильно различаются. В значительной степени неопределенности количественных характеристик атмосферных блокировок связаны с разными их проявлениями и различием критериев их выделения, в том числе используются разные критерии для минимальной продолжительности блокировок — 5, 7, 10 сут. [2—24].

Скорость зонального переноса в тропосфере средних широт определяется геострофическим балансом силы Кориолиса и градиента давления. При характерной для современной Земли трехъячейковой меридиональной циркуляции в тропосфере с нисходящими ветвями ячеек Хэдли и Ферреля и повышенным давлением в субтропических широтах (около 30°) и с восходящими ветвями ячейки Ферреля и полярной ячейки и пониженным давлением в субполярных широтах (около 60°) геострофический ветер в тропосфере средних широт направлен с запада на восток. Блокирование такого зонального потока должно быть связано с изменением градиента давления в средних широтах. Это достигается, в частности, когда в более высоких широтах находится блокирующий антициклон, а в более низких — циклон. Подобный вихревой диполь может характеризоваться с позиций классической гидродинамики как пара точечных вихрей в ведущем потоке, скорость которого равна противоположно направленной скорости движения вихрей [25] (см. также [26]). При этом относительно поверхности вихревая пара неподвижна.

Блокирование может быть связано с особенностями обтекания поверхности (блокирование меридионального типа) или с крупномасштабной неустойчивостью зонального потока, приводящей к образованию мощной системы взаимодействующих антициклона и циклона (блокирование расщепляющегося типа) либо мощного гребня и двух циклонов (омега-блокирование). Блокирование с расщеплением линий тока в основном наблюдается над Атлантикой и Евразией, омега-блокирование наблюдается над океанами, блокирование меридионального типа практически всегда отмечается в районе Уральских и Скалистых гор (см. [25]).

Относительно просто диагностировать блокировки по изменению знака градиента геопотенциала между субполярными и субтропическими широтами для достаточно протяженного (в несколько десятков градусов) долгот-

ного сектора, при этом в течение достаточно продолжительного времени (обычно не менее 5 сут.). На этом основаны критерии Лежёнаса — Окленда [5], Тибальди — Молтени [6] и их модификации (см., например, [10, 13]).

При обычном зональном переносе в тропосфере средних широт Северного полушария меридиональный градиент геопотенциала направлен в сторону более низких широт с большой величиной градиента в широтном поясе $40^\circ \text{ N} — 60^\circ \text{ N}$. Согласно простейшему критерию Лежёнаса — Окленда [5], если выполняются условия для LO-индекса I , определяемого на высоте H изобарической поверхности 500 гПа,

$$I(\lambda) = H(\lambda)_{40^\circ \text{N}} - H(\lambda)_{60^\circ \text{N}} < 0,$$

$$[I(\lambda - 10) + I(\lambda) + I(\lambda + 10)] < 0,$$

в течение не менее 5 суток, то в данном широтном поясе и диапазоне долгот (λ) не менее 30° (с центрами 10° -х ячеек от $\lambda - 10$ до $\lambda + 10$) осуществляется блокирование. В работе [13] дополнительно предложен индекс интенсивности блокирования BI , характеризующий перепад высот геопотенциала на уровне 500 гПа для блокинга.

Структура блокирующего образования расщепляющегося типа с двумя одинаковыми по интенсивности вихрями разного знака завихренности, циклоном и более полярным антициклоном, понятна с точки зрения гидродинамики [25]. Пара точечных вихрей на расстоянии d с равными по величине и противоположными по знаку вихревыми зарядами $\chi_1 = -\chi$, $\chi_2 = \chi$, $\chi > 0$ движется перпендикулярно соединяющей их оси со скоростью

$$U = \chi/2\pi d.$$

При этом вихрь с отрицательной (антициклонической) завихренностью находится справа (полярнее) по направлению движения. Такая пара вихрей (диполь), движущаяся с востока на запад в ведущем западном потоке со скоростью геострофического ветра, равной по величине, но противоположной по направлению скорости движения вихрей, будет неподвижна относительно Земли.

Омега-блокирование, как и блокирование расщепляющегося типа, связывается с неустойчивостью струйного течения в атмосфере. Этот тип блокирования можно представить комбинацией из трех вихрей, расположенных в вершинах правильного треугольника (триполь): более полярного антициклона с вихревым зарядом $\chi_1 = -\chi$ и двух циклонов с вихревыми зарядами $\chi_2 = \chi_3 = \chi/2$. Такой триполь должен двигаться вдоль оси, соединяющей вихри χ_2 и χ_3 , со скоростью

$$U = \frac{\chi\sqrt{3}}{4\pi r},$$

где r — сторона треугольника [25]. При этом антициклонический вихрь χ_1 находится справа по направлению движения. В ведущем потоке, как в тропосфере средних широт (с запада на восток), подобный вихревой три-

поль, движущийся с востока на запад со скоростью, равной по величине, но противоположной по направлению скорости зонального потока, стационарен относительно Земли.

Существуют значительные различия в характеристиках блокирований, обнаруженных разными методами. В исследованиях [12, 27] проведено сравнение характеристик блокирований, детектируемых различными методами, в том числе использовавшимися в работах [14] и [15], наряду с простейшим LO-индексом и его модификациями [10, 13]. Индекс блокирования в работе [14] основан на анализе меридионального градиента потенциальной температуры и потенциальной завихренности. Методом детектирования блокирующих антициклонов, использовавшимся в исследовании [15], определяется область повышенного давления или геопотенциальной высоты и внешний замкнутый контур вокруг локального максимума. К блокирующим относятся антициклоны, для которых выполняются условия квазистационарности (скорость смещения центра антициклона не превышает определенных пороговых значений).

Явления блокирования, характеризуемые сочетанием квазистационарных циклонических и антициклонических региональных режимов, обычно ассоциируются с блокирующими антициклонами. Это связано с тем, что при продолжительных квазистационарных антициклонах без облаков реализуются экстремальные температурные режимы — аномальная жара летом и морозы зимой. На рисунке 2 представлены кумулятивные распределения для циклонов и антициклонов в зависимости от параметра превышения, пропорционального их кинетической энергии, которая не меньше $E_k \sim (\Delta p)^2$, где Δp — перепад давлений между центром вихря и его периферией [28, 29]. Согласно рис. 2 для крупных (за исключением самых мощных) атмосферных вихрей характерно экспоненциальное распределение в зависимости от энергии вихрей, как в статистике Больцмана. Можно считать рассматриваемую систему каноническим ансамблем Гиббса, в котором океан играет роль термостата, а атмосфера является подсистемой, где вихри — это флуктуации, распределенные по Больцману.

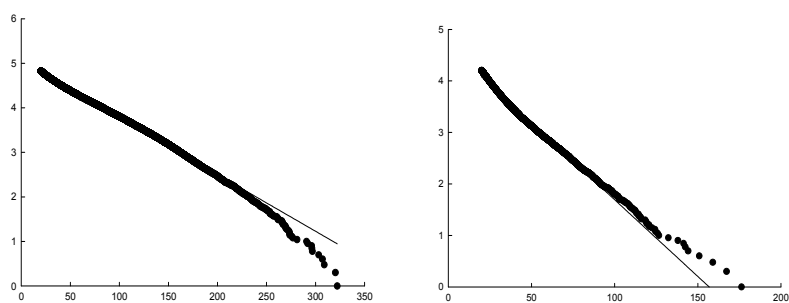


Рис. 2. Кумулятивное число ($\lg N$) циклонов (*слева*) и антициклонов (*справа*) с кинетической энергией, превышающей $E_k \sim (\Delta p)^2$ [(гПа)²] (ось абсцисс)

Существенные особенности отмечены на рис. 2 для хвостов распределений: с более быстрым, чем по экспоненте, уменьшением повторяемости наиболее мощных циклонов и более медленным — для повторяемости наиболее мощных антициклонов, к которым относятся блокирующие антициклоны. Это свидетельствует о повышенном риске региональных погодно-климатических аномалий, связанных с блокирующими антициклонами, по сравнению с экспоненциальным распределением.

Блокирование зонального переноса в тропосфере средних широт можно объяснить как на вихревом «языке», так и на волновом — стационарованием волн Россби [8, 30]. Волны Россби, связанные с зависимостью силы Кориолиса от широты (при наличии β -эффекта), распространяются с востока на запад со скоростью U_R , которую в простейшей баротропной модели можно представить в виде

$$U_R \sim \beta m^{-2},$$

где m — волновое число, $\beta = 2\Omega \cos\varphi/R$, R и Ω — радиус и угловая скорость вращения Земли, φ — широта. В тропосфере средних широт с зональным потоком с запада на восток со скоростью U при выполнении условия $U = U_R$ волна Россби стационарна относительно Земли. При этом вдоль широтного пояса чередуются m долготных секторов повышенного давления с антициклоническим режимом и пониженного — с циклоническим. Зимой с большими, чем летом, меридиональным температурным градиентом и геострофической скоростью зонального переноса в среднеширотной тропосфере условия стационарирования волн Россби реализуются при меньших значениях m . Характерный пример стационарирования волн Россби в тропосфере средних широт с $m = 4$ (или $m = 5$) — явления бабьего лета [30].

Тенденции изменений блокинговой активности

В работах [31, 32] по многолетним данным наблюдений оценено увеличение характерных длительностей блокингов при общем потеплении. Эту связь, подтверждаемую и по многолетним данным реанализа, можно качественно объяснить с помощью модельных оценок. Например, параметры чувствительности размера блокинга L и продолжительности τ к изменению температуры T могут быть представлены в виде

$$\frac{1}{L} \frac{dL}{dT} \sim \frac{1}{H_m} \frac{dH_m}{dT} - \frac{1}{U} \frac{dU}{dT},$$

$$\frac{1}{\tau} \frac{d\tau}{dT} \sim \frac{1}{L} \frac{dL}{dT} - \frac{1}{U} \frac{dU}{dT},$$

где U — характерная скорость ветра, H_m — интенсивность блокирующего антициклона, определяемая положительной аномалией геопотенциала [31, 32].

При простейшей оценке размера блокингов L масштабом Обухова для баротропной атмосферы $L_0 \sim T^{1/2}$ их протяженность должна расти с поте-

плением ($dL/dT = dL_0/dT > 0$). Согласно этим оценкам, в частности, блокинги формируются и живут дольше в более теплой тропосфере при более слабой зональной циркуляции в тропосфере. Это следует ожидать в связи с уменьшением меридионального температурного градиента и геострофической скорости ветра в тропосфере средних широт. При этом блокинги с уменьшающимися, неизменными или слабо увеличивающимися размерами при тропосферном потеплении должны становиться более слабыми, тогда как блокинги с достаточно сильно увеличивающимися при тропосферном потеплении размерами должны интенсифицироваться.

Отмеченные тенденции в целом согласуются с оценками исследования [10], полученными по расчетам с использованием климатической модели общей циркуляции с увеличением содержания CO_2 в атмосфере. В целом при ослаблении зональной циркуляции в тропосфере средних широт увеличивается риск прорыва холодных масс воздуха из Арктики и теплого воздуха из субтропиков. Это способствует увеличению изменчивости («нервозности») климата. Следует отметить, что при общем потеплении зональная циркуляция в тропосфере в отдельных широтных зонах может усиливаться в связи с выхолаживанием стратомезосферы, усилением струйного течения и квазициклическими процессами типа Северо-Атлантического колебания и Эль-Ниньо. Это может способствовать проявлению тенденции уменьшения характерного времени τ и характерного размера L для блокингов или увеличению их интенсивности. В работе [33] показано, что резкие изменения крупномасштабной структуры атмосферных течений могут приводить к быстрому распаду блокинга. Анализ фазовых траекторий и показателей Ляпунова позволил идентифицировать, когда происходят резкие изменения в общей атмосферной динамике. В частности, можно выделить четыре различных сценария распада блокингов: при недостатке синоптической подпитки, при активной роли синоптических процессов или при каждом из отмеченных режимов во взаимодействии с резким изменением характера атмосферных течений планетарного масштаба [33].

В работе [34] для оценки интегрального эффекта климатических структурных образований в атмосфере и океане, в частности атмосферных блокингов, было предложено использовать специальную величину — аналог действия S , которое имеет размерность [энергия]×[время] и определяется интегрированием функции Лагранжа по времени. Для отдельных блокирующих антициклонов с интенсивностью I (характеризуемой перепадом давления между центром антициклона и его периферией), размером L (площадью $\sim L^2$) и временем жизни τ действие S можно оценить согласно

$$S \sim \int I(t)L^2(t)dt.$$

Интегрирование производится по времени t от 0 до τ , где τ — время жизни блокинга. Интегральное действие для ансамбля N вихрей определяется суммой значений действия для отдельных вихрей:

$$S_{\Sigma} \sim \Sigma \int I(t)L^2(t)dt.$$

Оценить интегральное действие N блокирующих антициклонов можно согласно

$$S_{\Sigma} \sim N I_a L_a^2 \tau_a$$

с использованием средних значений для интенсивности (I_a), площади (L_a^2) и времени жизни τ_a блокирующих антициклонов. Согласно оценкам, полученным в работе [34] в результате модельных расчетов, при потеплении в связи с увеличением содержания CO_2 в атмосфере ожидается общий рост действия блокингов над континентами, в частности в холодный период года, и для Северного полушария в целом. Полученные оценки свидетельствуют о том, что зимние морозы не противоречат глобальному потеплению.

При климатических изменениях, в том числе температурного и циркуляционного режимов, изменяются и режимы волн Россби, включая условия их стационарирования. При $U = U_R$

$$dU/dT = -2U_R(0)(dm/dT)/m.$$

В случае

$$dU/dT < 0$$

следует

$$dm/dT > 0.$$

Это означает, что ослабление зонального потока при потеплении способствует стационарированию волн Россби с бóльшим волновым числом и следует ожидать проявления в тропосфере средних широт стационарной структуры с чередованием менее протяженных по долготе антициклонических и циклонических областей. При изменениях климата наряду с изменением условий стационарирования волн Россби можно ожидать проявления эффектов пространственного резонанса планетарных волн [30, 35, 36].

В исследовании [27] проведен анализ возможных изменений атмосферных блокирований по расчетам 10 климатических моделей в рамках международного проекта CMIP5 при различных RCP (Representative Concentration Pathways) сценариях антропогенных воздействий в XXI веке. Сопоставлялись характеристики блокирований, детектированных различными методами, в том числе использовавшимися в работах [5, 10, 13—15, 37, 38]. При этом использовались разные версии метода, представленного в работе [15], с различными условиями блокирования (в частности, со смещением блокирующего антициклона с радиусом R за сутки не более чем на pR , где $p = 0,3; 0,4; 0,5$). При анализе результатов расчетов с полным ансамблем из 10 климатических моделей проводился также отбор моделей, наиболее адекватно воспроизводящих современные сезонные распределения повторяемости среднеширотных атмосферных блокирований в зависимости от долготы, в сопоставлении с полученными по данным реанализа. Ансамбли «лучших» моделей определялись из условия максимальной корреляции долготных распределений среднесезонной частоты

блокирований по модельным расчетам и по данным NCEP/NCAR-реанализа для 30-летнего периода 1976—2005 годов.

На рисунке 3 представлены сделанные в работе [27] оценки вероятности превышения общей длительности летних и зимних блокирований в Евро-Атлантическом регионе ($60^\circ \text{ W} — 60^\circ \text{ E}$) в зависимости от длительности блокирования по ансамблевым модельным расчетам с ансамблем «лучших» моделей для последнего 30-летия XXI века (2071—2100 годов) при разных сценариях RCP в сопоставлении с 30-летним периодом 1976—2005 годов (кривая XX на рис. 3). При этом условия блокирования определялись аналогично [38]. Согласно полученным модельным оценкам, при продолжении глобального потепления в XXI веке увеличивается вероятность очень длительных летних и зимних атмосферных блокирований в средних широтах Северного полушария, в частности для сектора от Атлантики до Урала. Отмечено, что значимость возможных изменений блокирований различна для разных сезонов и сценариев, при этом оценки изменений экстремально длительных блокирований для полного ансамбля моделей (без выделения «лучших») получены менее значимыми.

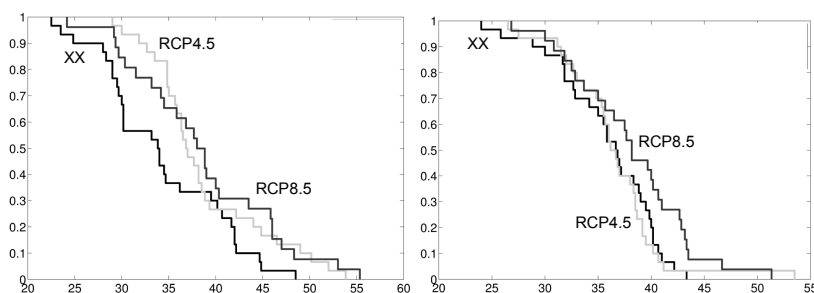


Рис. 3. Вероятность превышения общей длительности (сут.) летнего (*слева*) и зимнего (*справа*) блокирования в Евро-Атлантическом регионе в зависимости от длительности блокирования по ансамблевым модельным расчетам для последнего 30-летия XXI века при сценариях RCP4.5 и RCP8.5 в сопоставлении с 30-летним периодом 1976—2005 годов

Наряду с долгопериодными изменениями климата с учетом антропогенных воздействий заметное влияние на блокировочную активность в атмосфере оказывают более короткопериодные квазициклические глобальные и региональные климатические процессы, включая Эль-Ниньо / Южное колебание, Тихоокеанскую десятилетнюю осцилляцию, Северо-Атлантическое колебание, Атлантическую долгопериодную осцилляцию и др. [13, 34, 39—43].

Связанные с блокировками региональные погодно-климатические и экологические аномалии

Сильнейшие региональные погодно-климатические аномалии, в том числе летняя засуха 2010 года на европейской территории России и ре-

кордное наводнение на Амуре в 2013 году, связаны с атмосферными блокингами. На рисунке 4 приведено меридиональное распределение повторяемости атмосферных блокирований в Северном полушарии для разных лет с 1969 года по данным Global Climate Change Group (<http://solberg.snr.missouri.edu/gcc>). Лето 2013 года выделяется высокой продолжительностью блокирования в атмосфере над Тихим океаном, а лето 2010 года — над Европой (см. [39—44]).

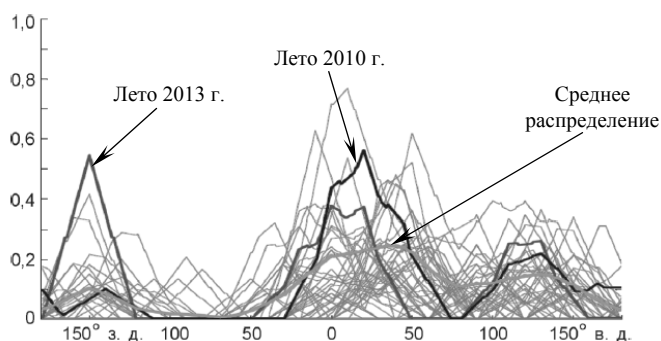


Рис. 4. Меридиональное распределение повторяемости атмосферных блокирований в Северном полушарии по данным для разных лет с 1969 года (тонкие серые кривые)

Насколько редки события с длительным атмосферным блокированием, в частности блокирования над Тихим океаном, инициировавшим рекордное наводнение на Амуре в 2013 году, можно оценить по рис. 5. Он характеризует период повторяемости двух последовательных лет с длительностью летнего блокирования в атмосфере не менее отмеченного на горизонтальной оси числа блоко-дней (τ_2) в западно-тихоокеанском секторе $130^\circ \text{ E} — 170^\circ \text{ W}$ средних широт Северного полушария по данным реанализа для периода 1969—2013 годов. Период повторяемости τ_r в западно-тихоокеанском секторе в двух последовательных годах общей длительности летнего блокирования 2013 года в соответствии с экспоненциальной аппроксимацией на рис. 5 — более 100 лет. На основании этих оценок в исследовании [41] возможность подобного длительного блокирования летом 2014 года была охарактеризована как очень маловероятная.

С атмосферными блокирующими антициклонами связано и формирование аномальных режимов атмосферных примесей, от содержания которых зависит экологическая обстановка. Особое значение имеет проблема образования озоновых «мини-дыр» (ОМД) в атмосфере — региональных особенностей полей озона с пониженным общим содержанием озона (ОСО) в атмосферном столбе [45—48]. От величины ОСО зависит интенсивность проходящей солнечной ультрафиолетовой радиации. В областях

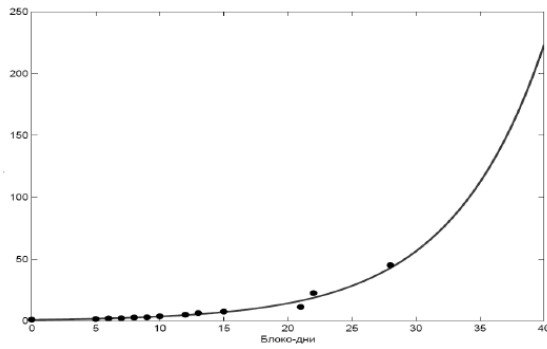


Рис. 5. Период (годы) повторяемости τ_r двух последовательных лет с длительностью летнего блокирования в атмосфере не менее отмеченного на горизонтальной оси числа блоки-дней τ_2 (сут.) в западно-тихоокеанском секторе Северного полушария по данным реанализа для периода 1969—2013 годов. Кривая — экспоненциальная аппроксимация

формирования ОМД могут отмечаться опасные уровни биологически активной солнечной радиации. В работах [47, 48] отмечено, что во время атмосферного блокирования летом 2010 года над европейской территорией России отрицательные аномалии ОСО достигали значений около 40 ед. Д. В работах [49—51] представлены результаты соответствующего анализа аномального содержания в атмосфере российских регионов озона, угарного газа, формальдегида и двуокиси азота в связи с блокированиями и пожарами в летние сезоны 2010, 2012 и 2016 годов.

Заключение

Согласно модельным оценкам, при глобальном потеплении можно ожидать увеличение вероятности очень продолжительных блокирующих режимов в атмосфере, инициирующих сильнейшие региональные погодно-климатические и экологические аномалии.

При невозможности точного долгосрочного (климатического) прогноза таких нелинейных процессов и явлений тем не менее потенциально возможно повысить предсказуемость режимов блокирования и их последствий при учете эффектов ключевых глобальных и региональных климатических процессов, проявляющихся в годовом цикле и межгодовой изменчивости. С точки зрения краткосрочного и среднесрочного прогноза есть перспектива более адекватного учета процессов, характеристики которых могут быть предикторами для формирования атмосферных блокингов. Полезными индикаторами при сохранении суммарного углового момента системы «Земля — атмосфера» являются угловая скорость вращения Земли и суммарный момент импульса атмосферы. В частности, про-

является связь блокирования в тропосфере с вариациями струйного течения выше по потоку относительно области формирования блокинга [39].

Чрезвычайные события последних лет (в частности, аномальная жара и пожары на европейской территории России летом 2010 года и рекордное наводнение на Амуре в 2013 году) с особой остротой поднимают вопросы, связанные с климатическими изменениями и их последствиями, в том числе экологическими. Согласно модельным оценкам, необычайно продолжительные режимы блокирования, с которыми связаны отмеченные аномалии, можно было ожидать [40, 41]. При оценке возможных изменений блокинговой активности в атмосфере на фоне вековых тенденций необходимо учитывать эффекты, связанные с внутривековыми квазициклическими глобальными и региональными процессами, в том числе явлениями Эль-Ниньо, Северо-Атлантического и Арктического колебаний, Тихоокеанской десятилетней осцилляции, Атлантического и Тихоокеанского долгопериодных колебаний.

Работа выполнена в рамках проектов РФФИ и программ РАН.

Литература

1. *IPCC, 2013 : Climate Change, 2013 : The Physical Science Basis : Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change / [T. F. Stocker, D. Qin, G.-K. Plattner [et al.] (eds)]. — Cambridge : Cambridge Univ. Press, 2013. — 1535 p.*
2. *Rex, D. F. Blocking action in the middle troposphere and its effect on regional climate. Part I : An aerological study of blocking action // Tellus. — 1950. — V. 2. — P. 196—211.*
3. *Rex, D. F. Blocking action in the middle troposphere and its effect on regional climate. Part II : The climatology of blocking action // Tellus. — 1950. — V. 2. — P. 275—301.*
4. *Charney, J. G. Multiple flow equilibria in the atmosphere and blocking / J. G. Charney, J. G. DeVore // J. Atmos. Sci. — 1979. — V. 36. — P. 1205—1236.*
5. *Lejenas, H. Characteristics of Northern Hemisphere blocking as determined from a long time series of observational data / H. Lejenas, H. Okland // Tellus A. — 1983. — V. 35. — P. 350—362.*
6. *Tibaldi, S. On the operational predictability of blocking / S. Tibaldi, F. Molteni // Tellus. — 1990. — V. 42A. — P. 343—365.*
7. *Груза, Г. В. Сезонные особенности пространственного распределения индексов блокирования в Северном полушарии / Г. В. Груза, Л. В. Коровкина // Метеорология и гидрология. — 1991. — № 3. — С. 108—110.*
8. *Hansen, A. R. A comparison between planetary-wave flow regimes and blocking / A. R. Hansen, A. Sutera // Tellus A. — 1993. — V. 45A. — P. 281—288.*
9. *Mokhov, I. I. Extratropical cyclones and anticyclones : Tendencies of change / I. I. Mokhov, T. N. Doronina, V. M. Gryanik, R. R. Khairullin, L. V. Korovkina, V. E. Lagun, O. I. Mokhov, E. P. Naumov, V. K. Petukhov, A. O. Senatorsky, M. V. Tevs // The Life of Extratropical Cyclones. — Vol. II / S. Gronas, M. A. Shapiro (eds). — Bergen : Geophysical Institute : University of Bergen, 1994. — P. 56—60.*

10. Lupo, A. R. Climatological features of blocking anticyclones : A study of Northern Hemisphere CCM1 model blocking events in present-day and double CO₂ concentration atmospheres / A. R. Lupo, R. J. Oglesby, I. I. Mokhov // *Climate Dyn.* — 1997. — V. 13. — P. 181—195.
11. Mokhov, I. I. Atmospheric blocking characteristics in the Northern Hemisphere : Diagnostics of changes / I. I. Mokhov, E. A. Tikhonova // *Research Activities in Atmospheric and Oceanic Modeling : Report № 987* / ed. by H. Ritchie. — Geneva, 2000. — P. 2.20—2.21. — (World Climate Research Programme / WMO).
12. Mokhov, I. I. Atmospheric blocking characteristics in the Northern Hemisphere : Comparison of two climatologies / I. I. Mokhov, E. A. Tikhonova, A. R. Lupo, J. M. Wiedenmann // *Research Activities in Atmospheric and Oceanic Modelling : Report № 1064* / ed. by H. Ritchie. — Geneva, 2001. — P. 2.20—2.21. — (World Climate Research Programme / WMO).
13. Wiedenmann, J. M. The climatology of blocking anticyclones for the Northern and Southern Hemispheres : Block intensity as a diagnostic / J. M. Wiedenmann, A. R. Lupo, I. I. Mokhov, E. A. Tikhonova // *J. Climate.* — 2002. — V. 15. — P. 3459—3473.
14. Pelly, J. L. A new perspective on blocking / J. L. Pelly, B. J. Hoskins // *J. Atmos. Sci.* — 2003. — V. 60. — P. 743—755.
15. Bardin, M. Quasi-stationary anticyclones in the Northern Hemisphere : An analysis of interannual and interdecadal variability and long-term trends at 1000 hPa and 500 hPa using geometric definition / M. Bardin, G. V. Gruza, A. R. Lupo, I. I. Mokhov, V. A. Tikhonov // *Proc. 16th Symp. on Global Change and Climate Variation, 85th Ann. Meet / AMS.* — 2005. — P. 9—13.
16. Barriopedro, D. A climatology of Northern Hemisphere blocking / D. Barriopedro, R. Garcia-Herrera, A. R. Lupo, E. Hernandez // *J. Climate.* — 2006. — V. 19. — P. 1042—1063.
17. Diao, Y. A new blocking index and its application : Blocking action in the Northern Hemisphere / Y. Diao, J. Li, D. Luo // *J. Climate.* — 2006. — V. 19. — P. 4819—4839.
18. Scherer, S. C. Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region / S. C. Scherer, M. Croci-Maspoli, C. Schwierz, C. Appenzeller // *Int. J. Climatol.* — 2006. — V. 26. — P. 233—249.
19. Croci-Maspoli, M. A multifaceted climatology of atmospheric blocking and its recent linear trend / M. Croci-Maspoli, C. Schwierz, H. C. Davies // *J. Climate.* — 2007. — V. 20. — P. 633—649.
20. Tyrlis, E. Aspects of a Northern Hemisphere atmospheric blocking climatology / E. Tyrlis, B. J. Hoskins // *J. Atmos. Sci.* — 2008. — V. 65. — P. 1638—1652.
21. Шакина, Н. П. Блокирование : условия лета 2010 года в контексте современных знаний / Н. П. Шакина, А. Р. Иванова, Б. А. Бирман, Е. Н. Скриптунова // *Анализ условий аномальной погоды на территории России летом 2010 года.* — М. : Трида, 2011. — С. 6—21.
22. Matsueda, M. Predictability of Euro-Russian blocking in summer of 2010 // *Geophys. Res. Lett.* — 2011. — V. 38. — Art. L06801. — doi:10.1029/2010GL046557.
23. Barnes, E. A. A methodology for the comparison of blocking climatologies across indices, models and climate scenarios / E. A. Barnes, J. Slingo, T. Woollings // *Clim. Dyn.* — 2012. — V. 38. — P. 2467—2481.

24. Barnes, E. A. Exploring recent trends in Northern Hemisphere blocking / E. A. Barnes, E. Dunn-Sigouin, G. Masato, T. Woollings // *Geophys. Res. Lett.* — 2014. — V. 41. — P. 638—644. — doi:10.1002/2013GL058745.
25. Обухов, А. М. Динамические условия возникновения засух и других крупномасштабных погодных аномалий / А. М. Обухов, М. В. Курганский, М. С. Татарская // *Метеорология и гидрология.* — 1984. — № 10. — С. 5—13.
26. Курганский, М. В. О движении пары вихрей на бета-плоскости // Исследования вихревой динамики и энергетики атмосферы и проблема климата. — Л. : Гидрометеоиздат, 1990. — С. 123—130.
27. Mokhov, I. I. Changes of atmospheric blockings in the 21st century from CMIP5 ensemble simulations with RCP scenarios / I. I. Mokhov, A. V. Timazhev // *Research Activities in Atmospheric and Oceanic Modelling : Report № 46* / E. Astakhova (ed.). — Geneva, 2016. — Sect. 2. — P. 6—7. — (World Climate Research Programme / WMO).
28. Акперов, М. Г. Функции распределения вероятностей циклонов и антициклонов по данным реанализа NCEP/NCAR и модели климата ИВМ РАН / М. Г. Акперов, М. Ю. Бардин, Е. М. Володин, Г. С. Голицын, И. И. Мохов // *Изв. РАН. Физика атмосферы и океана.* — 2007. — Т. 43, № 6. — С. 764—772.
29. Голицын, Г. С. Функции распределения вероятности циклонов и антициклонов в период 1952—2000 гг. : инструмент для определения изменений глобального климата / Г. С. Голицын, И. И. Мохов, М. Г. Акперов, М. Ю. Бардин // *Доклады РАН.* — 2007. — Т. 413, № 2. — С. 254—256.
30. Агаян, Г. М. Квазистационарные осенние режимы атмосферы Северного полушария в период ПГЭП / Г. М. Агаян, И. И. Мохов // *Изв. АН СССР. Физика атмосферы и океана.* — 1989. — Т. 25, № 11. — С. 1150—1156.
31. Mokhov, I. I. Sensitivity of storm track activity and blockings to global climatic changes : Diagnostics and modelling / I. I. Mokhov, V. K. Petukhov, A. O. Senatorsky // *Publ. Acad. Finland.* — *Painatuskaskus*, 1995. — № 6/95. — P. 438—441.
32. Мохов, И. И. Блоклинги и тенденции их изменения / И. И. Мохов, В. К. Петухов // *Доклады РАН.* — 1997. — Т. 357, № 5. — С. 687—689.
33. Лупо, А. Р. Оценка влияния на распад блокингов процессов планетарного масштаба с анализом фазовых траекторий и энтропии / А. Р. Лупо, И. И. Мохов, С. Достоглоу, А. Р. Кунц, Дж. П. Буркхардт // *Изв. РАН. Физика атмосферы и океана.* — 2007. — Т. 43, № 1. — С. 52—60.
34. Мохов, И. И. Действие как интегральная характеристика климатических структур : оценки для атмосферных блокингов // *Доклады РАН.* — 2006. — Т. 409, № 3. — С. 403—406.
35. Мохов, И. И. Внутрисезонные температурные режимы и их эволюция в трехмерной модели климата ИФА РАН / И. И. Мохов, В. К. Петухов, В. А. Семенов // *Изв. РАН. Физика атмосферы и океана.* — 1998. — Т. 34, № 2. — С. 145—152.
36. Petoukhov, V. Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes / V. Petoukhov, S. Rahmstorf, S. Petri, H. J. Schellnhuber // *PNAS.* — 2012. — V. 110. — P. 5336—5341.
37. Mokhov, I. I. Changes in atmospheric blocking characteristics within Euro-Atlantic region and Northern Hemisphere as a whole in the 21st century from model simulations using RCP anthropogenic scenarios / I. I. Mokhov, A. V. Timazhev, A. R. Lupo // *Glob. Planet. Change.* — 2014. — V. 122. — P. 265—270.
38. Мохов, И. И. Модельные оценки возможных изменений атмосферных блокингов в Северном полушарии при RCP-сценариях антропогенных воздействий / И. И. Мохов, А. В. Тимажев // *Доклады РАН.* — 2015. — Т. 460, № 2. — С. 210—214.

39. Мохов, И. И. Особенности формирования летней жары 2010 г. на европейской территории России в контексте общих изменений климата и его аномалий // Изв. РАН. Физика атмосферы и океана. — 2011. — Т. 47, № 6. — С. 709—716.
40. Мохов, И. И. Блокинги в Северном полушарии и Евро-Атлантическом регионе : оценки изменений по данным реанализа и модельным расчетам / И. И. Мохов, М. Г. Акперов, М. А. Прокофьева, А. В. Тимажев, А. Р. Лупо, Э. Ле Трет // Доклады РАН. — 2013. — Т. 449, № 5. — С. 582—586.
41. Мохов, И. И. Гидрологические аномалии и тенденции изменения в бассейне р. Амур в связи с климатическими изменениями / И. И. Мохов, В. Ч. Хон, А. В. Тимажев, А. В. Чернокульский, В. А. Семенов // Экстремальные паводки в бассейне р. Амур : причины, прогнозы, рекомендации : (сб. докладов). — М. : Росгидромет, 2014. — С. 81—120.
42. Lupo, A. R. Studying summer season drought in Western Russia / A. R. Lupo, I. I. Mokhov, Y. G. Chendev, M. G. Lebedeva, M. Akperov, J. A. Hobbart // Adv. Meteorol. — 2014. — V. 2014. — Art. ID 942027. — Режим доступа: <http://dx.doi.org/10.1155/2014/942027>.
43. Мохов, И. И. Оценки предсказуемости климатических аномалий в российских регионах в связи с явлениями Эль-Ниньо / И. И. Мохов, А. В. Тимажев // Доклады РАН. — 2015. — Т. 464, № 6. — С. 722—726.
44. Мохов, И. И. Погодно-климатические аномалии в российских регионах и их связь с глобальными изменениями климата / И. И. Мохов, В. А. Семенов // Метеорология и гидрология. — 2016. — № 2. — С. 16—28.
45. James, P. M. A climatology of ozone miniholes over the Northern Hemisphere // Int. J. Climatol. — 1998. — V. 18. — P. 1287—1303.
46. Barriopedro, D. Atmospheric blocking signatures in total ozone and ozone miniholes / D. Barriopedro, M. Antón, G. A. García // J. Climate. — 2010. — V. 23. — P. 3967—3983.
47. Ситнов, С. А. Формирование озоновой «мини-дыры» в условиях продолжительного блокирующего антициклона в атмосфере над европейской территорией России летом 2010 года / С. А. Ситнов, И. И. Мохов // Доклады РАН. — 2015. — Т. 460, № 1. — С. 74—78.
48. Ситнов, С. А. Особенности поля общего содержания озона при атмосферном блокировании над европейской территорией России летом 2010 г. (по спутниковым данным) / С. А. Ситнов, И. И. Мохов // Метеорология и гидрология. — 2016. — № 1. — С. 41—51.
49. Ситнов, С. А. Связь задымления атмосферы европейской территории России летом 2016 года с лесными пожарами в Сибири и аномалиями атмосферной циркуляции / С. А. Ситнов, И. И. Мохов, Г. И. Горчаков // Доклады РАН. — 2017. — Т. 472. — № 4. — С. 456—461.
50. Ситнов, С. А. Общее содержание оксида углерода в атмосфере над российскими регионами по данным спутникового прибора МОРПТ (2000—2014 гг.) / С. А. Ситнов, И. И. Мохов, А. В. Джола // Изв. РАН. Физика атмосферы и океана. — 2017. — Т. 53, № 1. — С. 38—55.
51. Sitnov, S. A. Atmospheric CH₂O and NO₂ contents during severe heat waves and wildfires in European Russia in 2010 and Siberia in 2012 / S. A. Sitnov, I. I. Mokhov // Proc. SPIE. — 2016. — V. 10035. — Art. 100356R-1.

ЛИНЕЙНЫЕ И НЕЛИНЕЙНЫЕ МОДЕЛИ БАРОКЛИННЫХ ВОЛН В АТМОСФЕРЕ

А. В. Елисеев

Институт физики атмосферы им. А. М. Обухова РАН, г. Москва
Казанский (Приволжский) федеральный университет

Введение

Бароклидная неустойчивость — основной механизм развития погодных возмущений в земной атмосфере. Она представляет собой наклонную конвекцию во вращающейся жидкости [1].

Физически бароклинный циклогенез можно описать следующим образом [2] (рис. 1). При приходе начальной циклонической (с завихренностью $\delta\zeta > 0$ — показана большой серой круговой стрелкой на рис. 1) аномалии в верхнюю тропосферу данного региона она передаётся в нижнюю тропосферу в случае выполнения условий бароклидной неустойчивости (малая серая круговая стрелка). Адвекция температуры (точнее, потенциальной температуры) приводит к возникновению индуцированной циклонической аномалии у земной поверхности (малая незаштрихованная круговая стрелка) к востоку от исходного циклона. Благодаря восходящим движениям эта аномалия передаётся в верхнюю тропосферу (большая незаштрихованная круговая стрелка), усиливая исходный циклон.

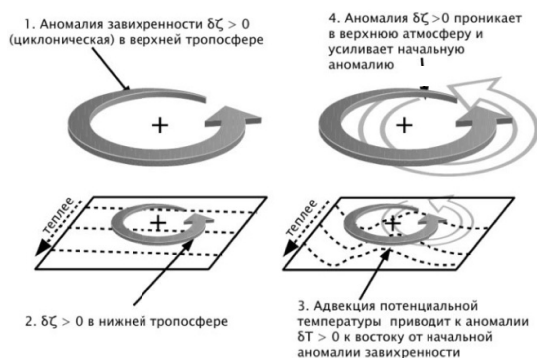


Рис. 1. Физическая схема развития бароклидной неустойчивости. Круговыми стрелками показано направление вращения в вихрях (серая стрелка — исходная аномалия, незаштрихованная стрелка — индуцированная аномалия). Пунктирными линиями показано распределение температуры на подстилающей поверхности, пунктирной стрелкой — направление её увеличения

Методы теоретического анализа бароклидной неустойчивости можно подразделить на две группы. К первой из них относится линейный метод анализа нормальных мод. Решения уравнения термогидромеханики для

любой векторной переменной $Y(\mathbf{r}, t)$, где $\mathbf{r} = (x, y, z)$ — радиус-вектор, а t — время, в этой группе методов ищутся в виде

$$Y = \text{Re} [\hat{Y} e^{i\mathbf{k}(\mathbf{r}-ct)}]. \quad (1)$$

Здесь i — мнимая единица, $\mathbf{k} = (k_x, k_y, k_z)$ — волновой вектор, $c = c_r + ic_i$ — комплексная фазовая скорость, $\hat{Y}(\mathbf{r}, t)$ — комплексная амплитуда, медленно меняющаяся в пространстве и времени:

$$\left| \frac{1}{\hat{Y}} \frac{\partial \hat{Y}}{\partial x} \right| \ll k_x, \quad \left| \frac{1}{\hat{Y}} \frac{\partial \hat{Y}}{\partial y} \right| \ll k_y, \quad \left| \frac{1}{\hat{Y}} \frac{\partial \hat{Y}}{\partial z} \right| \ll k_z, \quad \left| \frac{1}{\hat{Y}} \frac{\partial \hat{Y}}{\partial t} \right| \ll |c\mathbf{k}|. \quad (2)$$

При этом предполагается, что наиболее неустойчивая волна (т. е. волна с максимальным положительным $\text{Im } c_i |\mathbf{k}|$) определяет структуру развивающихся бароклиных возмущений. В приложениях метода нормальных мод уравнения, как правило, линеаризуются, что исключает исследование нелинейных процессов.

При использовании второй группы задача Коши решается при заданных начальных условиях для нелинейных уравнений термогидромеханики. Это позволяет анализировать не только линейные, но и нелинейные процессы бароклиной неустойчивости.

В нашей статье дан обзор основных результатов, полученных разными авторами к настоящему времени и касающихся обоих методов анализа бароклиной неустойчивости.

1. Линейная модель развития бароклиных возмущений

Теоретические модели развития бароклиных волн в атмосфере были впервые предложены в 1947 году Дж. Г. Чарни [3] и в 1949 году Ю. Т. Иди [4]. В обоих случаях рассматривались квазигеострофические возмущения (т. е. движения, в которых горизонтальный градиент давления уравновешивается силой Кориолиса [1, 2] и изменением скорости ветра) с зональным волновым числом $k_y = 0$. При этом также считалось, что скорость ветра \bar{u} , на фоне которой развиваются бароклиные волны, зависит от давления p (выбранного в качестве вертикальной координаты) линейно, с коэффициентом Λ . Существенное различие между моделями [3] и [4] состоит в том, что в первой работе движения рассматриваются на β -плоскости (т. е. в случае, когда параметр Кориолиса $f = 2\Omega \sin \varphi$, где Ω — угловая скорость вращения Земли, φ — географическая широта, представляется в линейном виде $f = f_0 + \beta y$ с постоянными f_0 и β ; $y = a\varphi$ — меридиональная координата, a — радиус Земли), тогда как во второй — на f -плоскости ($f \equiv f_0$).

В модели [4] для слоя атмосферы между изобарическими уровнями p_0 (поверхность) и $p_1 \ll p_0$ (тропопауза) уравнение для амплитуды Ψ синоптических возмущений функции тока имеет вид

$$ik_x[\bar{u}(p) - c] \left(-k_x^2 \Psi + n^2 \frac{d^2 \Psi}{dp^2} \right) = 0 \quad (3)$$

с граничными условиями

$$ik_x[\bar{u}(p_j) - c] \frac{d\Psi}{dp} + ik_x \Lambda \Psi = 0 \quad \text{при } j = 0, 1. \quad (4)$$

Здесь $n^2 = p^2 / L_R^2$, радиус деформации Россби $L_R = (g H_0)^{1/2} / f_0 \sim 10^3$ км, g — ускорение свободного падения, H_0 — высота однородной атмосферы (экспоненциальный масштаб уменьшения плотности воздуха с высотой) и для простоты считается, что $n^2 = \text{const}$. Из условия сохранения энергии следует, что если есть решение с фазовой скоростью распространения волн c_1 , то должно быть решение с фазовой скоростью $c_2 = c_1^*$. Поэтому для доказательства наличия бароклиной неустойчивости достаточно показать, что есть решения с чисто комплексными c . После деления уравнения (3) на $(\bar{u}(p) - c)$ получается уравнение, в котором множитель в круглых скобках из (3) приравнивается к нулю. Бароклиная неустойчивость возникает при условии $\alpha < \alpha_c \approx 2,4$, где $\alpha = (k_x / n)(p_0 - p_1)$. Скорость роста этих возмущений $k_x \text{Im } c_i$ максимальна (что соответствует наиболее неустойчивой моде) при $\alpha \approx 1,75$. Такая модель в целом воспроизводит основные особенности бароклиных возмущений в земной атмосфере.

Для простейших оценок в выражении для n можно заменить p на $\frac{1}{2}(p_0 + p_1)$, а затем положить $p_1 = p_0 / 3$ (это упрощает выкладки, хотя и несколько противоречит исходному предположению $p_1 \ll p_0$). В таком случае длина волны, соответствующая максимально неустойчивой моде, оказывается равной $3,6L_R$ (несколько тысяч километров), скорость её роста

$$(k_x \text{Im } c_i)_{\max} \approx 0,31 (\bar{u}_1 - \bar{u}_2) / L_R \quad (5)$$

(эта величина получила название показателя Иди), где $\bar{u}_1 = \bar{u}(p_1)$, $\bar{u}_2 = \bar{u}(p_2)$. Последнее при $u_1 - u_0 \sim 10$ м/с приводит к характерному времени развития возмущений $[(k_x \text{Im } c_i)_{\max}]^{-1} \sim 4$ сут. Эти оценки дают правильные по порядку значения для характерного горизонтального размера бароклиных волн в атмосфере и характерного временного интервала смены погоды.

Обобщение модели Иди [4] на более общий класс негеострофических движений проведено в статье [5].

2. Роль нелинейного взаимодействия между отдельными бароклиными волнами при циклогенезе

Пренебрежение корнями разности $\bar{u}(p) - c$ в уравнении (3) приводит к исчезновению у этого уравнения непрерывного спектра. При этом в его решениях остаётся только дискретный спектр счётной мощности [6]. Не-

смотря на то что о существовании этого спектра было известно ещё с начала XX века [7, 8], долгое время решениям из непрерывного спектра не придавалось значения, так как они растут пропорционально степенной функции от времени [9], т. е. значительно медленнее, чем экспоненциально растущие неустойчивые дискретные моды (см. раздел 1).

Для динамической системы с вектором состояния $Y \in \mathbf{H}_Y$ и эволюционным оператором A

$$\frac{dY(t)}{dt} = A(t)Y(t), \quad Y \in \mathbf{H}_Y, \quad (6)$$

справедлива спектральная теорема, согласно которой, если оператор $A(t)$ является нормальным (т. е. линейным, ограниченным в гильбертовом пространстве, а также перестановочным со своим сопряжённым), то он имеет дискретный счётный спектр $\lambda = \{\lambda_1, \lambda_2, \dots\}$ (при конечно-разностной аппроксимации этого оператора спектр всегда имеет конечную мощность), а в пространстве \mathbf{H}_Y можно построить ортонормированный базис $\Phi = \{\Phi_1, \Phi_2, \dots\}$ из собственных векторов этого оператора [10].

Из определения собственных векторов оператора

$$A(t) \Phi_j(t) = \lambda_j \Phi_j(t), \quad j = 1, 2, \dots, \quad (7)$$

с учётом (6) следует, что для всех j справедливо $\Phi_j(t) = \Phi_j(0)e^{\lambda_j t}$, т. е. собственные векторы дифференциального линейного оператора являются его нормальными модами.

Однако гидротермодинамический оператор (уравнения Навье — Стокса + термодинамическое уравнение + уравнение неразрывности) в трёхмерном евклидовом пространстве в общем случае не является нормальным. Строго говоря, он даже не является линейным. Как следствие, полного базиса из его собственных векторов построить невозможно. Собственные векторы его дискретного спектра могут быть подвергнуты процедуре ортогонализации, но полученный при этом набор векторов не будет полон в пространстве решений гидротермодинамического оператора. Система собственных векторов этого оператора может быть сделана полной в \mathbf{H}_Y добавлением собственных векторов из непрерывного спектра [6], однако мощность такого набора уже несчётна, и векторы из этого набора не могут быть сделаны взаимно ортогональными. Физически это выражается в том, что при нарастании бароклиных возмущений в земной атмосфере возникает взаимодействие между отдельными модами решения, влияющее на развитие этих возмущений.

Пионерские работы, в которых рассматривалось такое взаимодействие, были выполнены Дж. Фэррелом [11, 12]. В частности, в статье [11] было продемонстрировано, что межмодовое взаимодействие для классической задачи Иди (раздел 1) способно приводить к тому, что скорость нарастания со временем для мод, формально не являющихся наиболее неустойчивыми, может превышать скорость роста наиболее неустойчивой

моды. В целом наиболее неустойчивая бароклидная мода доминирует в первые часы развития бароклидной неустойчивости, а затем — начиная с временного горизонта, равного примерно 1 неделе (следует иметь в виду, что, как отмечено в разделе 1, одна неделя — характерная длительность существования бароклидных образований). Для наиболее интересных с точки зрения погодных явлений временных горизонтов (от нескольких часов до нескольких суток) её влияние оказывается не столь значимым по сравнению с модой, приток энергии к которой обеспечивается нелинейным межволновым взаимодействием. Более того, нелинейное взаимодействие между отдельными бароклидными волнами способно приводить к нарастанию даже волн, формально не являющихся неустойчивыми [12].

Диагностика важности такого взаимодействия может быть проведена с использованием подхода, предложенного в работе [13] и развитого в исследовании [14] (рис. 2). Он основан на анализе симметрии функций распределения вероятности (ФРВ) для погодных аномалий переменных. При этом для переменной Y выделяются среднее значение \bar{Y} и погодная аномалия $Y' = Y - \bar{Y}$. В связи с особенностями формирования погодных аномалий на разных временных масштабах ряд $Y'(t)$ можно затем подвергнуть фильтрации. В работе [14] таким образом выделены два интервала временных масштабов: синоптическая изменчивость (СИ) — 2—7 сут. (характерная для погодных процессов, см. выше) и низкочастотная изменчивость (НИ) — 9—30 сут. (которая часто связана с формированием блокирующих антициклонов в атмосфере — см., например, [15, 16]). В качестве характеристики близости ФРВ к симметричной удобно использовать асимметрию

$$S_Y = \frac{\overline{Y'^3}}{\sigma^3}, \quad (8)$$

где σ — стандартное отклонение ряда $Y'(t)$. В случае линейной динамики развития бароклидных возмущений можно ожидать симметричную ФРВ (рис. 2). Справедливо даже более сильное утверждение — в этом случае ФРВ должна быть нормальной. Поэтому любое статистически значимое отклонение S_Y от нуля можно рассматривать как проявление эффективной нелинейности динамики.



Рис. 2. Схема диагностики вклада нелинейных межволновых взаимодействий, основанная на оценке характеристик ФРВ для погодной переменной Y [18, 19]

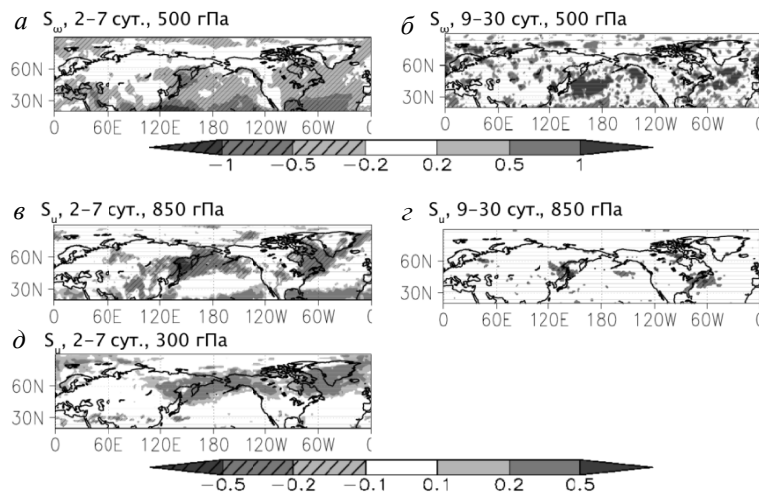


Рис. 3. Асимметрия выборочной функции распределения вероятности в октябре — марте для вертикальной скорости, определенной через производную от давления по времени, $\omega = dp/dt$, в изобарических координатах на изобарическом уровне 500 гПа (*a, б*) и для зональной скорости на изобарических уровнях 850 гПа (*в, г*) и 300 гПа (*д*). Рисунки представлены для интервалов временных масштабов 2—7 сут. (*a, в, д*) и 9—30 сут. (*б, г*). Поля сглажены 9-точечным квадратным пространственным фильтром. Отмечены лишь узлы расчётной сетки с $|s| > 2$ ($s = S_Y/\sigma_{S,G}$, оценка стандартного отклонения гауссова процесса $\sigma_{S,G} = (6/N_{eff})^{1/2}$, эффективная длина ряда $N_{eff} = N(1 - r_1)$, N — исходная длина ряда, r_1 — коэффициент автокорреляции этого ряда с единичным лагом)

В работе [14] с использованием данных реанализа JRA-55 (http://jra.kishou.go.jp/JRA-55/index_en.html) этот метод применён для анализа асимметрии функций распределения вероятности ряда атмосферных переменных, характеризующих состояние тропосферы Северного полушария в 1976—2014 годах. Статистически значимые отклонения от симметричного распределения вероятности наиболее сильно проявляются в регионах наиболее частого формирования атмосферных погодных возмущений — над западными частями океанов средних широт и ниже по потоку в атмосфере (рис. 3). В интервале СИ для всей толщи свободной тропосферы выявляется отрицательная асимметрия вертикальной скорости в изобарических координатах, что согласуется с общим доминированием циклонических аномалий в этом интервале временных масштабов. В интервале НИ асимметрия этой переменной во всей свободной тропосфере положительна, что указывает на доминирование антициклонических аномалий на этих временных масштабах. Для зональной скорости, температуры и геопотенциала знак асимметрии ФРВ при изменчивости с временными масштабами 2—7 сут. различен для верхней и нижней свободной тропосферы.

3. Взаимодействие бароклиных волн с основным потоком

Источником энергии бароклиных волн является фоновая (практически зональная) циркуляция [1, 2]. Частично энергия бароклиных вихрей передаётся фоновому потоку на стадии диссипации этих вихрей [17—19]. По имени автора работы [17] энергетический цикл бароклиных волн назван циклом Лоренца.

Оценка составляющих цикла Лоренца бароклиных вихрей по современным данным реанализа NCEP2 и ERA-40 была выполнена в исследовании [20] (рис. 4). Источником общей циркуляции атмосферы служит различие радиационного нагрева между тропиками и высокими широтами, которое порождает потенциальную энергию основного потока P_M . Этот поток, однако, неустойчив относительно бароклиных возмущений, что приводит к формированию вихревой потенциальной энергии P_E (небольшой вклад в неё также обусловлен генерацией потенциальной вихревой энергии $G(P_E)$). Так как развитие вихрей связано с движением в атмосфере, то при этом одновременно формируется и вихревая кинетическая энергия K_E , которая затрачивается в основном при диссипации вихрей (связанной с трением о поверхность и радиационной релаксацией). Однако около

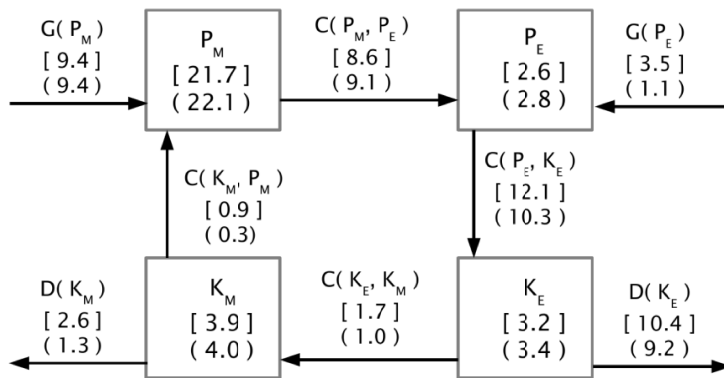


Рис. 4. Компоненты цикла Лоренца земной атмосферы по данным реанализа NCEP2 (<http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html>) — числа в квадратных скобках и ERA-40 (<http://apps.ecmwf.int/datasets/data/era40-daily/levtype=pl/>) — числа в круглых скобках соответственно [9]: P_M — потенциальная энергия основного потока, P_E — потенциальная энергия вихрей, K_M — кинетическая энергия основного потока, K_E — кинетическая энергия вихрей, $G(Y)$ — генерация Y , $C(Y, Z)$ — преобразование энергии из Y в Z (стрелками указано направление этого преобразования), $D(Y)$ — диссипация Y ($Y, Z = P_M, P_E, K_M, K_E$). Энергия выражена в 10^{20} Дж, преобразования энергии — в 10^{14} Вт

1/10 интенсивности преобразования P_E в K_E затем переходит в кинетическую энергию основного потока K_M . Большая часть кинетической энергии основного потока затрачивается на диссипацию, но некоторая часть этой энергии переходит в потенциальную энергию основного потока P_M .

Общая теория взаимодействия волн с зональным потоком изложена в статье [21], где в квазигеострофическом приближении на f -плоскости получены уравнения, связывающие зонально осреднённые значения атмосферных переменных (обозначенных ниже чертой сверху) с характеристиками вихревого воздействия (обозначенных штрихами):

$$\frac{\partial \bar{u}}{\partial t} - f_0 \bar{v} = -\frac{\partial(\overline{uv'})}{\partial y} + \bar{X}, \quad (9)$$

$$\frac{\partial \bar{T}}{\partial t} + \frac{N^2 H_0}{R} \bar{w} = -\frac{\partial(\overline{v'T'})}{\partial y} + \bar{J}, \quad (10)$$

где u , v и w — зональная, меридиональная и вертикальная компоненты скорости, T — температура, N — частота Брента — Вайсяля, вычисляемая по полям давления и температуры и служащая одной из характеристик устойчивости атмосферы [1, 2] (в [21] для простоты предположено, что $N = \text{const}$), X — интенсивность трения о поверхность, J — сумма всех неадиабатических источников изменения температуры. В связи с формулами (9) и (10) в работе [21] введено понятие остаточной среднезональной циркуляции

$$\bar{v}^* = \bar{v} - \frac{R}{\rho_0 H_0} \frac{\partial}{\partial z} \left(\frac{\rho_0}{N^2} \overline{v'T'} \right), \quad (11)$$

$$\bar{w}^* = \bar{w} - \frac{R}{H_0} \frac{\partial}{\partial y} \left(\frac{1}{N^2} \overline{v'T'} \right). \quad (12)$$

Здесь ρ_0 — плотность атмосферы в основном состоянии, подчиняющаяся гидростатическому закону (экспоненциальному уменьшению с ростом высоты с вертикальным масштабом H_0). Подставляя (11), (12) в (9), (10), можно получить элегантные уравнения для изменения характеристик основного (фонового) течения атмосферы:

$$\frac{\partial \bar{u}}{\partial t} - f_0 \bar{v}^* = \frac{1}{\rho_0} \nabla \cdot \mathbf{P} + \bar{X}, \quad (13)$$

$$\frac{\partial \bar{T}}{\partial t} + \frac{N^2 H_0}{R} \bar{w}^* = \bar{J}, \quad (14)$$

$$\frac{\partial \bar{v}^*}{\partial y} + \frac{1}{\rho_0} \frac{\partial}{\partial z} (\rho_0 \bar{w}^*) = 0. \quad (15)$$

В этих уравнениях вихревое воздействие на основное течение описывается так называемым потоком Элиассена — Пальма $\mathbf{P} = (P_y, P_z)$:

$$P_y = -\rho_0 \overline{u'v'}, \quad P_z = \rho_0 f_0 L_R \overline{v'T'}. \quad (16)$$

По приведённым в книге [2] данным, величина $\nabla \cdot \mathbf{P}$ максимальна в средних широтах.

4. Модель формирования блокирующих образований в атмосфере как следствия мультистабильности погодных режимов в атмосфере

Нелинейные взаимодействия бароклиных волн со средним потоком способны приводить к множественным положениям равновесия атмосферы. Такие примеры приведены в статье [15], где построена модель формирования блокирующего антициклона в результате быстрого перехода атмосферы из одного равновесного состояния в другое.

Эта модель является частным случаем модели взаимодействия баротропной волны, вызванной возмущением циркуляции из-за топографии, и среднего потока. Блокирующие антициклоны развиваются вследствие бароклиной неустойчивости зонального потока, однако становятся баротропными по мере своего развития. Как следствие, для анализа данной задачи можно использовать линеаризованное баротропное геострофическое уравнение возмущения завихренности ζ' , которое для такой топографической волны имеет вид [2]

$$\left(\frac{\partial}{\partial t} + \bar{u} \frac{\partial}{\partial x}\right) \zeta' + \beta v' + r \zeta' = -\frac{f_0}{H_0} \bar{u} \frac{\partial h_T}{\partial x}, \quad (17)$$

где v' — возмущение меридиональной скорости, соответствующее возмущению завихренности, r — коэффициент линейной диссипации, h_T — высота топографии. Зональная компонента скорости ветра среднего потока подчиняется уравнению

$$\frac{\partial \bar{u}}{\partial t} = -D(\bar{u}) - \kappa(\bar{u} - U_e), \quad (18)$$

где вихревое воздействие описывается первым слагаемым, а радиационная релаксация под влиянием «радиационно-равновесного» ветра U_e — вторым. Для баротропных возмущений, вызванных топографией, можно получить

$$D(\bar{u}) = -\overline{v' \zeta'} - \frac{f_0}{H_0} \overline{v' h_T}. \quad (19)$$

Если топография имеет вид линейного гребня с синусоидальной формой отдельных гор одинаковой высоты h_0 , ширины $2\pi/l$, с расстоянием между ними $2\pi/k$:

$$h_T(x, y) = \text{Re} [h_0 e^{ikx}] \cos(ly), \quad (20)$$

то

$$D(\bar{u}) = -\frac{f_0}{H_0} \overline{v' h_T} = \left(\frac{rK^2 f_0^2}{2\bar{u} H_0^2}\right) \frac{h_0^2 \cos(ly)}{[(K^2 - K_S^2)^2 + \varepsilon^2]}. \quad (21)$$

Здесь $\varepsilon = rK^2(k\bar{u})^{-1}$, $K^2 = k^2 + l^2$, $K_S^2 = \beta/\bar{u}$. Решение системы (18), (21) имеет три положения равновесия, одно из которых неустойчиво (и не имеет физического смысла), а два устойчивы. Одно из устойчивых положений

равновесия соответствует обычному западному переносу в тропосфере средних широт, а второе (с малым значением \bar{u}) — блокированному потоку.

Анализ блокирующих антициклонов в терминах «скрытых долгоживущих состояний» (в ряде случаев соответствующих отдельным состояниям равновесия динамической системы) был проведён в исследовании [22]. При этом выяснилось, что пространственная структура и временная динамика одного из таких «скрытых долгоживущих состояний» хорошо соответствует блокированию в атмосфере.

5. Режимы циркуляции

Результаты, полученные в работе [15], были эвристически обобщены Т. Н. Палмером в виде представлений о режимах циркуляции [23]. Согласно ему, вероятность ряда состояний атмосферы заметно превышает вероятность других её состояний [23]. При этом время пребывания атмосферы в состояниях первого типа также заметно превышает время её пребывания в состояниях второго типа. Первые состояния названы режимами циркуляции, а вторые интерпретируются как переходные между режимами циркуляции. Режимы циркуляции связываются с движениями на (возможно, несвязном) аттракторе. В свою очередь, любое изменение погодной статистики (в том числе климатические изменения) могут быть интерпретированы как изменение вероятности нахождения системы в режимах циркуляции, т. е. как изменение структуры погодного аттрактора.

Для системы (6) решение можно записать через пропагатор \mathfrak{L} :

$$Y(t) = \mathfrak{L}(t)Y(0). \quad (22)$$

Если на систему дополнительно действует внешнее (неавтономное) воздействие $f(t)$

$$\frac{dY(t)}{dt} = A(t)Y(t) + f(t), \quad (23)$$

то (22) превращается в

$$Y(t) = \mathfrak{L}(t)Y(0) + \int_0^t \mathfrak{L}(t)\mathfrak{L}^{-1}(s)f(s) ds. \quad (24)$$

Если f не зависит от времени, то (24) можно записать в виде

$$Y(t) = \mathfrak{L}(t)Y(0) + \mathbf{I}(t)f, \quad (25)$$

где

$$\mathbf{I}(t) = \int_0^t \mathfrak{L}(t)\mathfrak{L}^{-1}(s) ds. \quad (26)$$

Для линеаризованной системы (23) пропагатор (в этом случае он называется тангенциальным пропагатором)

$$\mathfrak{L} = e^{tF} \quad (F = dA/dt) \quad (27)$$

не зависит от времени и имеет вид квадратной матрицы. Если дополнительно к этому f также не зависит от времени (т. е. справедливо (25)), то сингулярное разложение имеет вид

$$\mathbf{I} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*, \quad (28)$$

где \mathbf{U} и \mathbf{V} — матрицы, состоящие из так называемых левых и правых сингулярных векторов \mathbf{u}_j и \mathbf{v}_j соответственно, $j = 1, 2, \dots, M$, $\mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_M)$, M — размерность \mathbf{F} . Без ограничения общности можно считать, что

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_M. \quad (29)$$

Согласно определению правых и левых сингулярных векторов, при всех $j = 1, 2, \dots, M$ справедливо

$$(\mathbf{I}\mathbf{I})\mathbf{u}_j = \sigma_j^2 \mathbf{u}_j, \quad (30)$$

$$(\mathbf{I}^*\mathbf{I})\mathbf{v}_j = \sigma_j^2 \mathbf{v}_j. \quad (31)$$

Так как \mathbf{F} в общем случае не является нормальной (раздел 2), то в общем случае $\mathbf{u}_j \neq \mathbf{v}_j$. Однако поскольку оператор $\mathbf{I}\mathbf{I}^* = \mathbf{I}^*\mathbf{I}$, очевидно, является нормальным, то каждый из наборов $\{\mathbf{u}_j\}$ и $\{\mathbf{v}_j\}$ ($j = 1, 2, \dots, M$) может быть выбран в качестве ортонормального базиса в пространстве решений (23). Тогда каждый вектор любого из этих наборов соответствует определённому режиму циркуляции. Более того, можно показать, что левые сингулярные векторы служат приближениями для ляпуновских векторов системы (23). В таком случае можно ожидать, что наиболее быстрый ляпуновский масштаб времени соответствует времени перехода между режимами циркуляции. Так как временной масштаб последнего перехода равен нескольким суткам, то такое соответствие может быть справедливо только в моделях, в которых более быстрые процессы (например, мезомасштабные синоптические) отфильтрованы. Примером таких фильтрованных уравнений являются квазигеострофические.

Согласно флуктуационно-диссипативной теореме [24], отклик системы на внешнее воздействие определяется автокорреляционной матрицей невозмущённой системы (6). Это согласуется с приведёнными выше соотношениями в случае, когда элементы автокорреляционной матрицы с наибольшими лагами соответствуют максимальным σ_j .

В работе [23] также показано, что аномалия с пространственной структурой \mathbf{Y}_0 наиболее эффективно возбуждается воздействием вида

$$\mathbf{f}_0 = \mathbf{V}\mathbf{\Sigma}(\mathbf{U}^*\mathbf{Y}_0). \quad (32)$$

Так как структура \mathbf{f}_0 при этом определяется в основном структурой сингулярных векторов с наибольшими σ_j , то отклик \mathbf{Y}_0 , в свою очередь, оказывается не очень чувствителен к географической структуре \mathbf{f}_0 . Последнее частично объясняет успех линейных моделей для анализа бароклинической неустойчивости.

Указанный подход в статье [25] был использован при анализе режимов циркуляции для данных геопотенциала на изобарическом уровне 500 мбар реанализа NCEP/NCAR (<http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html>) за 1949—1994 годы. Фильтрация быстрых возмущений

циркуляции была осуществлена разложением поля геопотенциала по эмпирическим ортогональным функциям с представлением состояния в виде двух основных векторов этого разложения. При этом ФРВ синоптических аномалий оказалась мультимодальной с 4 максимумами, причём в разные временные интервалы относительная роль этих максимумов (мод) оказалась различной. Работа [25] была подвергнута критике с точки зрения статистической значимости полученных максимумов ФРВ [26, 27]. Тем не менее такой подход может оказаться полезным при анализе климатических изменений.

Выводы

Методы теоретического анализа бароклиной неустойчивости — основного механизма развития погодных возмущений в земной атмосфере [1, 2] — можно разделить на две группы. К первой относится линейный метод анализа нормальных мод, в котором решения уравнений термогидромеханики ищутся в виде линейных бегущих волн без учёта взаимодействия между ними и при этом предполагается, что наиболее неустойчивая волна определяет структуру развивающихся бароклиных возмущений. Вторая группа методов получила название немодального анализа и представляет собой решение задачи Коши при заданных начальных условиях для нелинейных уравнений термогидромеханики. Это позволяет анализировать не только линейные, но и нелинейные процессы.

Классическими примерами приложения метода нормальных мод являются работы [3, 4]. При ряде дополнительных предположений в этих работах развиты теоретические модели бароклиных возмущений, которые в целом адекватно описывают физические особенности циклогенеза в земной атмосфере.

Однако тот факт, что гидродинамический оператор не является нормальным, указывает на важность межмодового взаимодействия при развитии бароклиной неустойчивости. В исследовании [11] было продемонстрировано, что межмодовое взаимодействие для классической задачи Иди [4] способно приводить к тому, что скорость нарастания со временем для мод, формально не являющихся наиболее неустойчивыми, может превышать скорость роста наиболее неустойчивой моды. Нелинейное взаимодействие между отдельными бароклиными волнами способно приводить к нарастанию даже волн, формально не являющихся неустойчивыми [12]. По данным реанализа получается, что взаимодействие между бароклиными модами наиболее значимо в регионах частого формирования атмосферных погодных возмущений — над западными частями океанов средних широт и ниже по потоку в атмосфере [13, 14].

Существенно нелинейным является взаимодействие бароклиных волн с основным (фоновым) потоком в атмосфере. Источником энергии бароклиных волн служит потенциальная энергия основного потока, ко-

торая переходит в потенциальную и кинетическую энергию бароклиных вихрей [1, 2, 17—19]. Вихревая кинетическая энергия в основном затрачивается на диссипацию вихрей, но часть её затем переходит в кинетическую и потенциальную энергию основного потока.

Нелинейное взаимодействие бароклиных волн с основным потоком может приводить к множественным положениям равновесия атмосферы. Например, в работе [15] построена модель формирования блокирующего антициклона благодаря быстрому переходу атмосферы из одного равновесного состояния в другое.

Результаты исследования [15] могут быть обобщены на понятие режимов циркуляции [23]. Такими режимами называются состояния атмосферы, вероятность реализации которых (и время пребывания в которых) заметно превышает вероятность реализации других состояний. Состояния второго типа интерпретируются как переходные между режимами циркуляции. Режимы циркуляции связываются с движениями на аттракторе. В свою очередь, любое изменение погодной статистики (в том числе климатические изменения) могут быть интерпретированы как взаимное изменение вероятности режимов циркуляции, т. е. как изменение структуры погодного аттрактора.

Автор выражает искреннюю благодарность М. В. Курганскому за полезные обсуждения при подготовке рукописи.

Литература

1. Курганский, М. В. Введение в крупномасштабную динамику атмосферы : (адиабатические инварианты и их применение) / М. В. Курганский. — СПб : Гидрометеоздат, 1993. — 168 с.
2. Holton, J. R. An Introduction to Dynamic Meteorology / J. R. Holton. — Cambridge : Academic Press, 2004. — 535 с.
3. Charney, J. G. The dynamics of long waves in a baroclinic westerly current // J. Meteorol. — 1947. — V. 4, № 5. — P. 135—162.
4. Eady, E. T. Long waves and cyclone waves // Tellus. — 1949. — V. 1, № 3. — P. 33—52.
5. Калашиник, М. В. О резонансном и квазирезонансном возбуждении бароклиных волн в модели Иди // Изв. РАН. Физика атмосферы и океана. — 2015. — Т. 51, № 6. — С. 648—657.
6. Педлоски, Дж. Геофизическая гидродинамика : в 2 т. / Дж. Педлоски. — Т. 2. — М. : Мир, 1984. — 416 с.
7. Orr, W. M. F. Stability or instability of the steady-motions of a perfect liquid and of a viscous liquid. Part I : A perfect liquid // Proc. Roy. Irish Acad. A. — 1907—1909. — V. 27. — P. 9—68.
8. Orr, W. M. F. Stability or instability of the steady-motions of a perfect liquid and of a viscous liquid. Part II : A viscous liquid // Proc. Roy. Irish Acad. A. — 1907—1909. — V. 27. — P. 69—138.

9. *Burger, A.* Instability associated with the continuous spectrum in a baroclinic flow // *J. Atmos. Sci.* — 1966. — V. 23, № 3. — P. 272—277.
10. *Рихтмайер, Р.* Принципы современной математической физики : в 2 т. / Р. Рихтмайер. — Т. 1. — М. : Мир, 1982. — 486 с.
11. *Farrell, B. F.* The initial growth of disturbances in a baroclinic flow // *J. Atmos. Sci.* — 1982. — V. 39, № 8. — P. 1663—1686.
12. *Farrell, B. F.* Transient growth of damped baroclinic waves flow // *J. Atmos. Sci.* — 1982. — V. 42, № 24. — P. 2718—2727.
13. *Petoukhov, V.* On statistics of the free-troposphere synoptic component: an evaluation of skewnesses and mixed third-order moments contribution to the synoptic dynamics and fluxes of heat and humidity / V. Petoukhov, A. V. Eliseev, R. Klein, H. Oesterle // *Tellus A.* — 2008. — V. 60, № 1. — P. 11—31.
14. *Логинов, С. В.* Негауссовость вариаций атмосферных переменных на временных масштабах от 2 до 30 суток / С. В. Логинов, А. В. Елисеев, И. И. Мохов // *Изв. РАН. Физика атмосферы и океана.* — 2017. — [в печати].
15. *Charney, J. G.* Multiple flow equilibria in the atmosphere and blocking / J. G. Charney, J. G. DeVore // *J. Atmos. Sci.* — 1979. — V. 36, № 7. — P. 1205—1216.
16. *Mokhov, I. I.* Changes in atmospheric blocking characteristics within Euro-Atlantic region and Northern Hemisphere as a whole in the 21st century from model simulations using RCP anthropogenic scenarios / I. I. Mokhov, A. V. Timazhev, A. R. Lupo // *Glob. Planet. Change.* — 2014. — V. 122. — P. 265—270.
17. *Lorenz, E. N.* Available potential energy and the maintenance of the general circulation // *Tellus.* — 1955. — V. 7, № 2. — P. 157—167.
18. *Williams, R. T.* Nonlinear, non-geostrophic effects in a baroclinic atmosphere // *J. Atmos. Sci.* — 1965. — V. 22, № 4. — P. 388—401.
19. *Simmons, A. J.* The life cycles of some nonlinear baroclinic waves / A. J. Simmons, B. J. Hoskins // *J. Atmos. Sci.* — 1978. — V. 35, № 3. — P. 414—432.
20. *Li, L.* Lorenz energy cycle of the global atmosphere based on reanalysis datasets / L. Li, A. P. Ingersoll, X. Jiang, D. Feldman, Y. L. Yung // *Geophys. Res. Lett.* — 2007. — V. 34, № 16. — P. L16813.
21. *Andrews, D. G.* Planetary waves in horizontal and vertical shear : The generalized Eliassen-palm relation and the mean zonal acceleration / D. G. Andrews, M. E. McIntyre // *J. Atmos. Sci.* — 1976. — V. 33, № 11. — P. 2031—2048.
22. *Horenko, I.* Metastable decomposition of high-dimensional meteorological data with gaps / I. Horenko, S. I. Dolaptchiev, A. V. Eliseev, I. I. Mokhov, R. Klein // *J. Atmos. Sci.* — 2008. — V. 65, № 11. — P. 3479—3496.
23. *Palmer, T. N.* A nonlinear dynamical perspective on climate prediction // *J. Climate.* — 1999. — V. 12, № 2. — P. 575—591.
24. *Leith, C. E.* Climate response and fluctuation dissipation // *J. Atmos. Sci.* — 1975. — V. 32, № 10. — P. 2022—2026.
25. *Corti, S.* Signature of recent estimate change in frequencies of natural atmospheric circulation regimes / S. Corti, F. Molteni, T. N. Palmer // *Nature.* — 1999. — V. 398, № 6730. — P. 799—802.
26. *Hsu, C. J.* Climate change in recurrent regimes and modes of northern hemisphere atmospheric variability / C. J. Hsu, F. Zwiers // *J. Geophys. Res. : Atmos.* — 2001. — V. 106, № D17. — P. 20145—20159.
27. *Stephenson, D. B.* On the existence of multiple climate regimes / D. B. Stephenson, A. Hannachi, A. O'Neill // *Q. J. Roy. Met. Soc.* — 2004. — V. 130, № 597. — P. 583—605.

СЕВЕРНЫЙ ЛЕДОВИТЫЙ ОКЕАН: ОСОБЕННОСТИ ГИДРОДИНАМИКИ, ДИНАМИКИ МОРСКОГО ЛЬДА И ПРОБЛЕМЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ КРУПНОМАСШТАБНОГО СОСТОЯНИЯ

Н. Г. Яковлев

Институт вычислительной математики им. Г. И. Марчука РАН, г. Москва

Введение

Цель настоящей статьи — дать представление об уровне физической постановки задачи в современном численном моделировании Северного Ледовитого океана и о фундаментальных проблемах, возникающих при переходе к моделям высокого пространственного разрешения.

Северный Ледовитый океан (СЛО) — один из наименее изученных уголков Мирового океана в силу своего географического положения. Интерес к Арктике вообще и к СЛО в частности связан как с чисто научными проблемами (наблюдаемые структурные изменения климатической системы, устойчивость полярной климатической системы; о природе «полярного усиления» см. работу [21]), так и с практическими вопросами судоходства и возможной добычи углеводородов на шельфе. Происходящие структурные изменения полярной климатической системы затрудняют прогнозирование в Арктике с помощью, например, статистических методов, которые основаны на исторических данных. При этом, несмотря на огромный прогресс в области развития методов измерений (прежде всего, спутниковые системы и система размещенных на льду буев и наблюдательных платформ), наблюдательная сеть в СЛО все еще недостаточно развита, особенно в глубоком океане под дрейфующим льдом. Поэтому в задаче диагноза и прогноза состояния СЛО велика роль численных моделей, построенных исходя из первых принципов.

Многолетний опыт международной активности по моделированию СЛО в рамках проекта АОМIP (Arctic Ocean Intercomparison Project, в настоящее время — FAMOS, The Forum for Arctic Ocean Modeling and Observational Synthesis, <http://www.who.edu/projects/famos/>) показывает, что многие задачи, поставленные еще 15 лет назад, до сих пор остаются на повестке дня. В то же время можно отметить и впечатляющие достижения в плане улучшения пространственного разрешения численных моделей, и развитие используемых физических параметризаций. Например, первый скоординированный эксперимент, проведенный в 1999 году, показал, что далеко не все модели воспроизводят наблюдаемую структуру распределения солености и только малая часть моделей была способна воспроизводить «классическую» схему циклонической циркуляции в слое атлантических вод (на глубине 300—500 м [19]). В настоящее время благодаря со-

трудничеству создателей численных моделей и экспедиционных океанологов и эта схема циркуляции скорректирована, и современные модели стали более реалистичными.

В то же время недавний эксперимент по воспроизведению циркуляции тихоокеанских вод [4] показал, что качество модели зависит от пространственного разрешения немонотонно и, вероятно, определяется не только разрешением. И этот результат можно трактовать как проявление сложной нелинейной физики полярного океана, морского льда и пограничного слоя подо льдом, которую имеет смысл изучить более внимательно.

Динамика океана

Для того чтобы разобраться в причинах неудач с моделированием СЛО, необходимо вспомнить принципы, на которых строятся современные модели крупномасштабной гидротермодинамики океана. Опуская вывод, полученный в книге [1], заметим, что традиционно используются приближения Буссинеска и бездивергентности (квазинесжимаемости) $\text{div } \vec{u} = 0$. Кроме того, применяется специальная запись уравнений в так называемой z системе координат, когда выделяются «горизонтальные» движения (соответствующие компоненты скорости движения обозначим нижним индексом h) в плоскости, нормальной к направлению суммы локальной силы тяжести и центробежного ускорения, и «вертикальные», направленные вдоль этой суммы. При этом в качестве вертикальной координаты выбирается расстояние от невозмущенной поверхности океана. Исходная система уравнений оказывается записанной на многообразии, которое можно назвать цилиндром над сферой. Вдобавок делается целый комплекс упрощений так, чтобы в новой системе координат существовали аналоги законов сохранения энергии и углового момента количества движения.

Использование приближения Буссинеска с постоянной референсной плотностью ρ_c требует переопределения давления в уравнении состояния $p \rightarrow p_0 = \rho_c g z$ — в противном случае будет нарушен закон сохранения энергии. Кроме того, вводятся отклонения плотности ρ' и давления p' от референсных значений ρ_c, p_0 , так что система уравнений движения, неразрывности, состояния, баланса тепла (в терминах потенциальной температуры θ) и солёности S в общепринятых обозначениях (вектор \vec{k} направлен вертикально вверх, f — параметр Кориолиса) имеет вид

$$\frac{D\vec{v}_h}{Dt} + f\vec{k} \times \vec{v}_h + \frac{1}{\rho_c} \vec{\nabla}_z p' = \vec{F},$$

$$\epsilon_{nh} \frac{Dw}{Dt} + \left[\frac{g\rho'}{\rho_c} + \frac{1}{\rho_c} \frac{\partial p'}{\partial z} \right] = \epsilon_{nh} F_w,$$

$$\begin{aligned}\bar{V}_z \cdot \bar{v}_h + \frac{\partial w}{\partial z} &= 0, \\ \rho' &= \rho(\theta, S, p_0) - \rho_c, \quad p_0(z) = \rho_c g z, \\ p' &= p - p_0, \\ \frac{D\theta}{Dt} &= Q_\theta, \\ \frac{DS}{Dt} &= Q_S.\end{aligned}$$

Уравнение для вертикальной компоненты скорости течений w в этом случае имеет непривычный вид, введен параметр ϵ_{nh} . Часто для расчетов крупномасштабных течений предполагается, что вертикальные ускорения частицы жидкости много меньше приведенного ускорения силы тяжести $\frac{g\rho'}{\rho_c}$ и можно пренебречь турбулентной вязкостью. В данном случае полагается $\epsilon_{nh} = 0$, и это называется приближением *гидростатики* (хотя правильнее было бы использовать термин *квазистатика*, так как скорость течений не равна нулю). В океане $\frac{\rho'}{\rho_c} = O(10^{-3})$. Получающееся при этом уравнение выделено квадратными скобками. Определенная путаница возникает из-за того, что часто при записи системы уравнений в приближении гидростатики штрихи опускаются, и тогда появляется желание сравнивать ускорение частицы с ускорением свободного падения, а не с *приведенным* ускорением свободного падения.

Таким образом, один из источников ошибки при моделировании — это неправильный учет процессов, зависящих от вертикальных движений, — например, из-за неустойчивой стратификации по плотности, которая является обычной ситуацией в СЛО, так как он охлаждается сверху. На практике эти процессы различным образом параметризуются в предположении, что вертикальная конвекция связана с так называемыми плюмами, горизонтальный масштаб которых порядка 100 м, и таких плюмов много на масштабе пространственного осреднения модели. Если горизонтальное разрешение модели достаточно мало (например, одного порядка с размером конвективного плюма), то использование приближения гидростатики может быть некорректным, и в этом случае нужно использовать негидростатическую модель динамики океана.

Второй источник ошибок — описание слагаемых, связанных с замыканием (или параметризацией) турбулентных диффузий и вязкости \bar{F}, Q_θ, Q_S . И здесь необходимо выделить несколько физических процессов.

1. Мезомасштабные вихри. Особенность СЛО — небольшие значения характерного размера квазигеострофических вихрей (радиуса деформации Россби). В глубоком Канадском бассейне бароклинный радиус деформации Россби ($R_{i,n} = \frac{NH}{fn}$, N — частота Брента — Вайсяля, H — глупина, n — номер вертикальной моды) составляет примерно 10 км для первой моды и 5 км — для второй [18], и еще меньше на Арктическом шельфе России. Это отличает СЛО от средних широт Мирового океана, где такие масштабы порядка 50—100 км и где для описания вихревого переноса температуры и солености (и других скаляров) в модели динамики океана разрешения порядка 15 км вполне достаточно [3].

Вероятно, именно с маленьким характерным горизонтальным масштабом вихрей и связан тот факт, что увеличение пространственного разрешения не дает ожидаемого эффекта в плане точности описания интенсивных струй и переноса тепла и солей [4]. При горизонтальном разрешении порядка 5 км спектр модели обрезается на масштабе отдельного вихря, что не позволяет правильно описать переходы энергии по спектру. Таким образом, переход от разрешения 20 км к разрешению 5 км может сопровождаться даже ухудшением результатов, так как организованные струи распадаются на цепочки вихрей, динамика которых описана неадекватно. В итоге струи разрушаются, а потоки тепла и солей резко падают. Это приводит к неверному описанию состояния океана в целом.

При низком пространственном разрешении, когда справедлив статистический подход в описании ансамбля вихрей, существуют параметризации так называемого вихревого переноса скаляра [8, 25]. Вихревой перенос скаляра τ можно связать со скоростью переноса \bar{u}_{GM} , которая определяется как диффузия с кососимметричным тензором коэффициентов диффузии $\kappa_{GM} \mathbf{K}_{GM}$:

$$\bar{u}_{GM} \tau = -\kappa_{GM} \mathbf{K}_{GM} \bar{\nabla} \tau,$$

$$\mathbf{K}_{GM} = \begin{pmatrix} 0 & 0 & -S_x \\ 0 & 0 & -S_y \\ S_x & S_y & 0 \end{pmatrix}.$$

Здесь S_x, S_y — тангенсы наклона изопикнической поверхности (поверхности равной потенциальной плотности) по отношению к горизонту. Коэффициент κ_{GM} может также быть функцией плотности морской воды. Пример использования такого рода параметризации в случае СЛО приведен в работе автора [2], где показано, что в результате улучшается воспроизведение переноса тепла и солей из Атлантики в Центральную Арктику в модели низкого пространственного разрешения.

Учитывая сказанное выше, можно ожидать, что такая параметризация вихревого переноса скаляров будет полезна при пространственном разрешении вплоть до 20 км. При более высоком пространственном разрешении мы попадаем в «серую зону», когда вихревые эффекты могут описываться моделью двояко — явно и через параметризацию. Достаточное разрешение могло бы быть порядка 1 км для большей части СЛО, однако при таком разрешении возникнут проблемы с описанием *субмезомасштабных* вихрей, которые обсудим ниже.

2. Субмезомасштабные вихри. Недавние измерения [22] показывают, что в Канадском бассейне в верхнем слое океана присутствуют маленькие вихри с горизонтальными размерами 2—4 км и максимальными орбитальными скоростями порядка 20 см/с, так что число Россби $Ro = \frac{U}{fL} = O(1)$,

где U , L — характерные масштабы скорости и длины. Эти параметры отличаются от случая мезомасштабных вихрей, в которых число Россби порядка 0,1—0,01. Подобные вихри классифицируются как субмезомасштабные и отвечают за восстановление стратификации в верхнем слое зимой. Важная особенность таких вихрей — относительно большая вертикальная скорость, $W \sim 10^{-2} \cdot U \sim 100$ метров в день, в отличие от мезомасштабных вихрей с вертикальными скоростями $W \sim (10^{-4} \div 10^{-3}) \cdot U \sim (1 \div 10)$ метров в день.

Интересно, что существование субмезомасштабных вихрей в океане было открыто благодаря численному моделированию и развитию спутниковых систем наблюдения за биотой океана (см. недавний обзор [15]).

В случае СЛО особенности субмезомасштабных вихрей, их роль в формировании наблюдаемых полей температуры и солености и возможность использования существующих параметризаций [7] в крупномасштабных моделях еще совершенно не изучены.

Для нас важно то, что в высоких широтах наблюдаемые пространственные масштабы субмезомасштабных вихрей оказываются близки к масштабам мезомасштабных вихрей (2 и 5 км соответственно), не давая возможности разделить процессы и построить параметризации для каждого из них отдельно, как сделано для случая средних широт. Это и представляет фундаментальную проблему замыкания турбулентности в СЛО на масштабе порядка нескольких километров.

Вероятно, оптимальным с точки зрения описания физики субмезомасштабных вихрей было бы аппроксимировать их явно, выбрав пространственное разрешение модели порядка первых сотен метров, — однако, как уже обсуждалось, в таком случае мы попадаем в зону перехода к негидростатическому режиму и должны использовать другие уравнения движения. Не говоря уже о резком увеличении вычислительных затрат.

3. Дифференциальная диффузия. Интересным вопросом, также требующим ответа, является возможность реализации в СЛО особого варианта двойной диффузии — «дифференциальной диффузии» [20, 23], приводящей к характерным ступенчатым структурам в вертикальных профилях температуры и солености. Такие структуры образуются в случае, когда пресная и холодная вода находится выше теплой и соленой, и обязаны своим существованием разной скорости молекулярной диффузии тепла и солей. Обычно в средних широтах или тропических областях океана, где, наоборот, теплая и соленая вода располагается над пресной и холодной, двойная диффузия приводит к увеличению эффективного вертикального перемешивания с образованием так называемых соляных пальцев (пример использования в глобальной модели климата океана — [16]). В случае же полярного океана можно ожидать обратного эффекта — уменьшения эффективного вертикального перемешивания (своего рода отрицательной диффузии). Вероятно, такой механизм может объяснить некоторые особенности, связанные с распространением теплых и соленых атлантических вод далеко от места их проникновения в проливе Фрама, что до сих пор не удается воспроизвести в численных моделях.

4. Внутренние приливы. Собственно приливная сила в Северном Ледовитом океане мала и приливы формируются за счет входящей волны. Полусуточный прилив M_2 при движении над неровным дном океана генерирует на той же частоте внутренний прилив. Из-за сингулярности дисперсионного соотношения для внутренних инерционно-гравитационных волн

$$\frac{d^2 w}{dz^2} + \frac{N^2(z)}{g} \frac{dw}{dz} + \frac{N^2(z) - \omega^2}{\omega^2 - f^2} k^2 w = 0,$$

где k — волновой вектор и ω — циклическая частота, распространение внутреннего прилива как свободной волны ограничено критической широтой 75° с. ш. Это означает, что на критической широте может происходить интенсивное вертикальное перемешивание, вызванное обрушением внутреннего прилива [17]. Роль приливов (и внутренних приливов в том числе, особенно в случае неплоского дна) в формировании крупномасштабного состояния СЛО также является предметом современных исследований.

Термодинамика морского льда

Термодинамика морского льда и снега на нем считается в первом приближении локально-одномерной (эффекты бокового таяния параметризуются особо), и, например, для определения температуры льда $T(^{\circ}\text{C})$ решается уравнение теплопроводности вида [11, 24]

$$c_{ice} \frac{\partial \rho_{ice} T}{\partial t} = \frac{\partial}{\partial z} \left(k_{ice} \frac{\partial T}{\partial z} \right) - \frac{\partial}{\partial z} [I],$$

$$c_{ice}(T, S_{ice}) = c_0 - \frac{L_{ice} \mu S_{ice}}{T^2}, \quad \mu = \text{const},$$

$$k_{ice}(T, S_{ice}) = k_0 + \frac{\beta S_{ice}}{T}, \quad \beta = \text{const}.$$

Здесь c_{ice} , ρ_{ice} , k_{ice} — удельная теплоемкость морского соленого льда, его плотность (вообще говоря, непостоянная) и коэффициент теплопроводности морского льда, c_0 , k_0 — теплоемкость и теплопроводность пресного льда, S_{ice} — соленость льда, I — поток энергии, связанный с проникающим в лед солнечным светом, L_{ice} — удельная теплота плавления пресного льда, а $\mu \approx 0,054$ — эмпирическая константа, связывающая соленость льда и температуру его плавления T_{melt} : $T_{melt} = -\mu S_{ice}$. Некоторое отличие от обычного уравнения теплопроводности состоит в том, что вид граничных условий зависит от свойств решения из-за фазовых переходов — это либо первое краевое условие (температура плавления льда и замерзания морской воды), либо обобщенное условие Неймана с потоками как функциями от решения.

Современные модели морского льда и снега на нем учитывают большое многообразие различных процессов, включая даже влияние цветения водорослей, живущих в толще и на нижней поверхности, на прозрачность толщи льда и поглощение солнечной радиации летом. Нам представляется, в этой области с точки зрения постановки задачи нет принципиальных проблем.

Динамика морского льда

Баланс импульса льда будет иметь вид (предполагается, что ледовый покров — двумерная среда и скорость дрейфа всех индивидуальных льдин на масштабе пространственного осреднения модели одинакова)

$$m \frac{\partial \vec{u}_i}{\partial t} + m f \vec{k} \times \vec{u}_i = -\vec{\nabla} P_0 + \vec{\tau}_a - \vec{\tau}_w + \vec{F}.$$

Здесь m — суммарная масса льда и снега (на единицу площади), \vec{u}_i — скорость дрейфа льда, $\vec{\tau}_a$ — касательное напряжение трения ветра, $\vec{\nabla} P_0$ — градиент давления на поверхности океана (в модели учитывается уровень океана и давление атмосферы), \vec{F} — средняя по толщине льда сила, вызванная реологией морского льда (т. е. напряжениями, возникающими в ледовом покрове вследствие его движения и различной толщины и сплоченности льда). Обычно делается предположение, что для морского

льда можно пренебречь инерционными слагаемыми $m(\bar{u}_i \cdot \bar{V})\bar{u}_i \approx 0$, так как масштабный анализ показывает, что в большей части океана, покрытого достаточно толстым льдом, эти слагаемые на порядок меньше локального ускорения. В сущности, это предположение не принципиально и может быть преодолено методом расщепления по физическим процессам.

Традиционно в крупномасштабных моделях предполагается, что число индивидуальных льдин на пространственном масштабе осреднения модели достаточно велико, так что ледовый покров можно рассматривать как непрерывную среду и ввести такую характеристику, как «сплоченность» (можно также назвать ее концентрацией) льда, — долю поверхности океана, занятую морским льдом. По определению $0 \leq A \leq 1$.

При построении реологии льда будем иметь в виду следующие соображения:

1. Разреженный лед (сплоченностью до 0,15) ведет себя как идеальный газ, $\bar{F} = 0$.

2. Если происходит разнообразного типа торошение (ломание и напозвание льдин друг на друга), то этот процесс можно аппроксимировать как пластичное течение.

3. Если нет торошения, то течение льда можно рассматривать как вязкое течение.

4. Если лед толстый и прочный и занимает всю область, то его можно рассматривать как твердое упругое тело.

5. На больших пространственных масштабах лед можно рассматривать как гранулированную среду.

Наиболее распространенные в настоящее время реологии — вязкопластичная [10] или упруговязкопластичная реология [11, 24]. При этом опять подразумевается существование большого числа льдин (заметим, что средний горизонтальный размер льдины в СЛО примерно 300 м, в окраинных морях меньше) на масштабе осреднения модели. Упруговязкопластичную реологию можно рассматривать не как отдельную реологию, а как вычислительный прием для реализации вязкопластичной.

Пусть компоненты тензора скоростей деформации определены обычным образом: $\dot{\epsilon}_{i,j} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$. Тогда общий вид тензора напряжений

будет иметь вид

$$\sigma_{i,j} = 2\eta(\dot{\epsilon}_{i,j}, P)\dot{\epsilon}_{i,j} + [\zeta(\dot{\epsilon}_{i,j}, P) - \eta(\dot{\epsilon}_{i,j}, P)]\dot{\epsilon}_{k,k}\delta_{i,j} - \frac{P}{2}\delta_{i,j}.$$

Здесь функции ζ, η — коэффициенты объемной и сдвиговой вязкости, а P — «прочность» льда. Наиболее простой способ задания прочности льда — по эмпирической формуле $P = P^* \cdot h \cdot e^{(-C^*(1-A))}$, где h, A — тол-

щина и сплоченность, а P^* и $C \approx 20$ — эмпирические константы (последняя выбирается так, чтобы обеспечить соотношение 1, приведенное выше).

В привычных декартовых координатах компоненты силы, связанной с реологией льда, будут иметь вид

$$F_x = \frac{\partial}{\partial x} \left\{ [\eta + \zeta] \frac{\partial u}{\partial x} + [\zeta - \eta] \frac{\partial v}{\partial y} - \frac{P}{2} \right\} + \frac{\partial}{\partial y} \left[\eta \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right],$$

$$F_y = \frac{\partial}{\partial y} \left\{ [\eta + \zeta] \frac{\partial v}{\partial y} + [\zeta - \eta] \frac{\partial u}{\partial x} - \frac{P}{2} \right\} + \frac{\partial}{\partial x} \left[\eta \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right].$$

При заданных параметрах льда напряжение в нем не может превышать некоторого критического значения, определяемого прочностью, что и приводит к разрушению льда (торошению), которое в двумерной геометрии аппроксимируется пластическим течением. Выбор вида коэффициентов вязкости неоднозначен и нетривиален и не выводится строго из первых принципов. Наиболее распространенная вязкопластичная реология — эллиптическая Хиблера, при этом

$$\zeta = \frac{P}{2\Delta}, \quad \eta = \frac{\zeta}{e^2}, \quad \text{где } \Delta = (\dot{\epsilon}_{11}^2 + \dot{\epsilon}_{22}^2)(1 + e^{-2}) + 4\dot{\epsilon}_{12}^2 e^{-2} + \dot{\epsilon}_{11}\dot{\epsilon}_{22}(1 - e^{-2}),$$

а параметр e равен примерно 2 — эксцентриситет эллиптической кривой пластичности. Возможны и другие варианты вязкопластичной реологии.

Сильная нелинейность в ледовом покрове связана с выходом на статический режим $\Delta \rightarrow 0$, когда формально коэффициенты сдвиговой и объемной вязкости становятся бесконечно большими, если не вводится специальная перенормировка давления. В этом случае может наблюдаться низкая скорость сходимости либо даже неустойчивость численного метода, приводящая к большим ошибкам в численном решении (это характерно для ранних вариантов упруговязкопластичной реологии). Для преодоления этих проблем предлагается либо использовать новые реологии [5, 9], либо применять более точные численные методы [6, 12—14].

Таким образом, фундаментальная проблема воспроизведения динамики морского льда — это формулировка математической постановки, реалистичной и вычислительно эффективной для масштаба осреднения порядка 1 км.

Заключение

Задача воспроизведения состояния Северного Ледовитого океана, как видится автору, представляет собой задачу моделирования широкого спектра разномасштабных и взаимосвязанных процессов, для описания

которого недостаточно использовать готовую, пусть даже и широко признанную, модель динамики океана и морского льда и формально увеличивать пространственное разрешение. В настоящее время мы находимся в такой точке, когда эволюционный период развития методом улучшения пространственного разрешения подходит к концу и переход к еще более высокому разрешению требует тщательной ревизии и изменений используемых постановок задачи, приближений, численных методов и программных реализаций. И к этому революционному этапу нужно быть готовыми, чтобы оставаться в клубе высоких технологий в области наук о Земле.

Литература

1. Каменкович, В. М. Океанология. Физика океана : в 2 т. / В. М. Каменкович, А. С. Монин. — Т. 1. — М. : Наука, 1978. — 455 с.
2. Яковлев, Н. Г. К вопросу о воспроизведении полей температуры и солености Северного Ледовитого океана // Изв. РАН. Физика атмосферы и океана. — 2012. — Т. 48, № 1. — С. 1—17.
3. Яковлев, Н. Г. Воспроизведение пространственно-временной изменчивости уровня Мирового океана моделью климата ИВМ / Н. Г. Яковлев, Е. М. Володин, А. С. Грицун // Изв. РАН. Физика атмосферы и океана. — 2016. — Т. 52, № 4. — С. 428—438. — doi: 10.7868/S000235151604012X.
4. Aksenov, Y. Arctic pathways of Pacific Water: Arctic Ocean Model Intercomparison experiments / Y. Aksenov, M. Karcher, A. Proshutinsky, R. Gerdes, B. de Cuevas1, E. Golubeva, F. Kauker, A. T. Nguyen, G. A. Platov, M. Wadley, E. Watanabe, A. C. Coward, A. J. G. Nurser // J. Geophys. Res. Oceans. — 2016. — V. 121. — P. 27—59.
5. Bouillon, S. Presentation of the dynamical core of neXtSIM, a new sea ice model / S. Bouillon, P. Rampal // Ocean Modelling. — 2015. — V. 91. — P. 23—37. — doi: <http://dx.doi.org/492.10.1016/j.ocemod.2015.04.005>.
6. Danilov, S. Finite-Element Sea Ice Model (FESIM) : version 2 / S. Danilov, Q. Wang, R. Timmermann, N. Iakovlev, D. Sidorenko, M. Kimmritz, T. Jung, J. Schröter // Geosci. Model Dev. — 2015. — V. 8. — P. 1747—1761. — doi:10.5194/gmd-8-1747-2015.
7. Fox-Kemper, B. Parameterization of mixed layer eddies. Part I : Theory and diagnosis / B. Fox-Kemper, R. Ferrari, R. Hallberg // J. Phys. Oceanogr. — 2008. — V. 38. — P. 1145—1165.
8. Gent, P. R. Isopycnal mixing in ocean circulation models / P. R. Gent, J. C. McWilliams // J. Phys. Oceanogr. — 1990. — V. 20, № 1. — P. 150—155.
9. Girard, L. A new modeling framework for sea-ice mechanics based on elastobrittle rheology / L. Girard, S. Bouillon, J. Weiss, D. Amitrano, T. Fichefet, V. Legat // Annals of Glaciology. — 2011. — V. 52, № 57. — P. 123—132.
10. Hibler, W. D., III. A dynamic-thermodynamic sea ice model // J. Phys. Oceanogr. — 1979. — V. 9, № 4. — P. 815—846.
11. Hunke, C. CICE : the Los Alamos Sea Ice Model : Documentation and Software : Version 5.0 : Rep. LA-CC-06-012 / C. Hunke, W. H. Lipscomb, A. K. Turner, N. Jeffery, S. M. Elliott ; Los Alamos National Laboratory Tech. — 2013.

12. *Kimmritz, M.* On the convergence of the modified elastic-viscous-plastic method for solving the sea ice momentum equation / M. Kimmritz, S. Danilov, M. Losch // *J. Comp. Phys.* — 2015. — V. 296. — P. 90—100.
13. *Kimmritz, M.* The adaptive EVP method for solving the sea ice momentum equation / M. Kimmritz, S. Danilov, M. Losch // *Ocean Modelling.* — 2016. — V. 101. — P. 59—67. — doi: <http://dx.doi.org/10.1016/j.ocemod.2016.03.004>.
14. *Losch, M.* A parallel Jacobian-free Newton-Krylov solver for a coupled sea ice-ocean model / M. Losch, A. Fuchs, J.-F. Lemieux, A. Vanselow // *J. Comp. Phys.* — 2014. — V. 257, part A(0). — P. 901—911.
15. *McWilliams, J. C.* Submesoscale currents in the ocean // *Proc. R. Soc. A.* — 2016. — V. 472, iss. 2189. — Art. 20160117. — <http://dx.doi.org/10.1098/rspa.2016.0117>.
16. *Merryfield, W. J.* A global ocean model with double-diffusive mixing / W. J. Merryfield, G. Holloway, A. E. Gargett // *J. Phys. Oceanogr.* — 1999. — V. 29. — P. 1124—1142.
17. *Morozov, E. G.* Internal Tidal Waves in the Barents Sea / E. G. Morozov, S. V. Pisarev, V. G. Neiman, S. Yu. Erofeeva // *Doklady Earth Sciences.* — 2003. — V. 393, № 8. — P. 1124—1126.
18. *Nurser, A. J. G.* The Rossby radius in the Arctic Ocean / A. J. G. Nurser, S. Bacon // *Ocean Sci.* — 2014. — № 10. — P. 967—975.
19. *Rudels, B.* On the intermediate depth waters of the Arctic Ocean / B. Rudels, E. P. Jones, L. G. Anderson, G. Kattner // *The Polar Oceans and their role in shaping the global environment : The Nansen Centennial Volume* / O. M. Johannessen, R. D. Muench, J. E. Overland (eds). — Washington : Am. Geophys. Un., 1994. — P. 33—46. — (Geophysical Monograph Series ; v. 85).
20. *Schmitt, R. W.* Double diffusion in oceanography // *Annu. Rev. Fluid Mechanics.* — 1994. — V. 26. — P. 255—285.
21. *Serreze, M. C.* Processes and impacts of Arctic amplification: A research synthesis / M. C. Serreze, R. G. Barry // *Global and Planetary Change.* — 2011. — V. 77. — P. 85—96.
22. *Timmermans, M.-L.* Eddies in the Canada Basin, Arctic Ocean, observed from Ice-Tethered Profilers / M.-L. Timmermans, J. Toole, A. Proshutinsky, R. Krishfield, A. Plueddemann // *J. Phys. Oceanogr.* — 2008. — V. 38. — P. 133—145.
23. *Timmermans, M.-L.* Ice-Tethered Profiler observations of the doublediffusive staircase in the Canada Basin thermocline / M.-L. Timmermans, J. Toole, R. Krishfield, P. Winsor // *J. Geophys. Res. : Oceans.* — 2008. — V. 113, № C1. — Art. C00A02. — doi:10.1029/2008JC004829.
24. *Vancoppenolle, M.* Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and validation / M. Vancoppenolle, T. Fichefet, H. Goosse, S. Bouillon, G. Madec, M. A. Morales Maqueda // *Ocean Modelling.* — 2009. — V. 27, № 1/2. — P. 33—53.
25. *Visbeck, M.* Specification of eddy transfer co-efficients in coarse resolution ocean circulation models / M. Visbeck, J. Marshall, T. Haine, M. Spall // *J. Phys. Oceanogr.* — 1997. — V. 27. — P. 381—402.

ФРАКТАЛЬНАЯ МОДЕЛЬ КОМПАКТНОГО ВНУТРИОБЛАЧНОГО РАЗРЯДА

Д. И. Иудин, С. С. Давыденко

Институт прикладной физики РАН, г. Нижний Новгород

Фрактальная динамика процессов в сильно неравновесных системах и её приложение к различным областям нелинейной физики, в частности к проблемам экстремальной геофизики, представляют собой стремительно развивающуюся область исследований [1—6]. Динамика сетей с изменяющейся архитектурой связей и резонансные области в хвосте земной магнитосферы, молнии и электрические разряды в верхней атмосфере, просачивание жидкости сквозь грунты со сложной реологией, сейсмическая активность и многое другое — всё это примеры фрактального поведения. При этом, несмотря на широкое разнообразие физических контекстов, в которых разворачивается фрактальная эволюция конкретных систем, последняя демонстрирует общие фундаментальные закономерности, определяющие независимость макроскопического поведения системы от особенностей взаимодействия её отдельных мелкомасштабных элементов. Поиск этих закономерностей наряду с построением базовых моделей фрактальной динамики представляется чрезвычайно актуальной задачей [7, 8].

Данная работа посвящена применению методов фрактального анализа для описания так называемых компактных внутриоблачных разрядов (КВР; в англоязычной литературе это явление носит название *compact intracloud discharge*, CID), история исследования которых начинается с 80-х годов прошлого века. Электромагнитный отклик КВР в дальней зоне представляет собой одиночный биполярный импульс электрического поля с длительностью от 10 до 30 мкс и сопровождается синхронным сверхмощным коротким всплеском высокочастотного излучения, интенсивность которого почти на порядок величины превосходит соответствующее излучение от типичных разрядов облако — земля. Как показано в работах [9, 10], указанные особенности электромагнитного излучения КВР не находят объяснения в рамках разработанных ранее моделей. Однако это удаётся сделать в рамках фрактального подхода, рассматривающего электрический разряд как динамическую иерархическую структуру проводящих каналов в среде с пространственно неоднородным распределением плотности заряда и напряжённости электрического поля. Остановимся на этом более подробно.

1. Молниевый разряд как система с самоорганизованной критичностью

Молниевые разряды в атмосфере представляют собой одно из самых ярких проявлений нелинейной динамики природных систем [11]. Сложность моделирования молниевых разрядов обусловлена, прежде всего,

широким диапазоном пространственно-временных масштабов этого явления. В частности, характерные пространственные масштабы процессов, связанных с разрядом, изменяются, по меньшей мере, на 10 порядков величины — от межатомных расстояний, актуальных при процессах электризации, до десятков километров, что соответствует размерам развитого грозового облака. Аналогично меняются и временные масштабы явления, простираясь от наносекунд при описании отдельных стримеров до десятых долей секунды при описании развитых молниевых разрядов. С другой стороны, широкий диапазон пространственно-временных масштабов молниевых разрядов даёт ключ к пониманию их основных закономерностей. Дело в том, что согласно данным натурных измерений динамика электрической структуры молниевых разрядов демонстрирует универсальные скейлинговые свойства, т. е. обладает самоподобием в достаточно широком диапазоне пространственно-временных масштабов, что проявляется в сильной, спадающей по степенному закону корреляции регистрируемых величин. Последнее свойство типично для критических явлений, поэтому качественно динамика электрического разряда проявляет свойства самоорганизующегося и самонастраивающегося критического режима. Другими словами, электрические разряды в атмосфере относятся к системам с самоорганизованной критичностью [12—15]. Такие системы оказываются сложными в том смысле, что для них не существует характерных временных или пространственных масштабов, которые бы управляли эволюцией этих систем. При этом, хотя динамический отклик системы сложен, его статистические свойства в некотором диапазоне параметров описываются простыми степенными законами. Разряд является результатом балансирования грозовой системы на пороге своеобразного геометрического фазового превращения, когда однородное поступление в облако свободной энергии уравновешивается потерями на фрактальных диссипативных структурах — фрактальных кластерах, образованных формирующимися при разряде проводящими каналами. Такое геометрическое превращение называют перколяционным фазовым переходом, который в отличие от обычных (термодинамических) фазовых превращений характеризуется геометрическими свойствами проводящих каналов. Последнее обстоятельство делает актуальным применение при моделировании молниевых разрядов современных методов фрактальной геометрии и теории перколяционных фазовых переходов. Покажем, как указанный подход можно применить для описания компактных внутриоблачных разрядов.

2. Фрактальная модель внутриоблачного разряда

Будем рассматривать внутриоблачный электрический разряд как трёхмерную сеть клеточных автоматов, где каждому узлу i модельной решётки соответствует накопленный в соответствующей элементарной ячейке

(кубе с ребром a) электрический заряд q_i . Величина заряда в ячейках в начальный момент времени $t=0$ задаётся распределением внутриоблачного электрического поля $\mathbf{E}(\mathbf{r}, t=0)$. Совокупность полного числа N элементарных зарядов в начальный и последующие моменты времени определяет электрический потенциал φ_i в i -м узле решётки и электрическое поле E_{ij} между соседними i -м и j -м узлами:

$$\varphi_i = \sum_{\substack{j=1, \\ j \neq i}}^N \frac{q_j}{L_{ij}} + \frac{q_i}{a_{\text{эфф}}} - \mathbf{E}_0 \mathbf{r}_i, \quad E_{ij} = |\varphi_i - \varphi_j| / L_{ij}. \quad (1)$$

Здесь \mathbf{r}_i и \mathbf{r}_j — координаты i -го и j -го узлов соответственно, $L_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ — расстояние между ними, \mathbf{E}_0 — постоянная компонента внешнего электрического поля, $a_{\text{эфф}}$ — эффективный радиус ячейки, определяющий её электрическую ёмкость. Динамика заряда в каждой ячейке обуславливается наличием и свойствами электрических связей с соседними ячейками. Предполагается, что вероятность возникновения проводящего канала (пробоя) между соседними i -й и j -й ячейками зависит от напряжённости электрического поля между ними и описывается распределением Вейбула:

$$P_{ij} = \begin{cases} 1 - \exp\left(-\left|\frac{E_{ij} - E_n}{E_k - E_n}\right|^m\right), & E_{ij} \geq E_n; \\ 0, & E_{ij} < E_n. \end{cases} \quad (2)$$

Здесь m — индекс Вейбула, E_k и E_n — критическое поле и поле инициации разряда соответственно ($E_k \gg E_n$), зависящие от типа разряда. При наличии проводящей связи между соседними ячейками электрический ток между ними на каждом шаге модельного времени равен $I_{ij} = E_{ij} G_{ij}$, где G_{ij} — обратное погонное сопротивление связи, при этом полное изменение электрического заряда в i -й элементарной ячейке даётся соотношением

$$\delta q_i = \tau \sum_{\substack{j=1 \\ j \neq i}}^{N_i} I_{ij}, \quad (3)$$

где τ — длительность шага, N_i — число активных проводящих связей i -й ячейки на данном шаге модельного времени. Отметим, что использованное выше выражение для тока I_{ij} справедливо только в случае, когда шаг τ достаточно мал по сравнению с характерным временем релаксации

заряда: $\tau \ll CL_{ij} / G_{ij}$, где C — эффективная ёмкость ячейки. Важной особенностью модели является учёт изменения проводимости связи в зависимости от протекающего по ней тока. Обратное погонное сопротивление (проводимость) связи описывается уравнением

$$\frac{\partial G_{ij}}{\partial t} = \eta E_{ij}^2 G_{ij} - \beta G_{ij}, \quad (4)$$

представляющим собой дифференциальный аналог формулы Ромпе — Вейцеля для проводимости длинной искры в газе [16]. Первое и второе слагаемые в правой части уравнения (4) описывают влияние на проводимость омического нагрева и остывания канала соответственно. Отметим, что при снижении проводимости до некоторого порогового значения канал исчезает.

Система уравнений (1) — (4) описывает возникновение и эволюцию иерархической структуры проводящих каналов во внутриоблачной среде при наличии внешнего электрического поля и некоторого начального распределения плотности заряда. Возникающая при этом система токов представляет собой распределённый источник электромагнитного излучения. Элементами такой системы являются линейные токи между соседними ячейками, электромагнитное поле каждого из которых задаётся в рамках известного приближения передающей линии [17, 18].

Изложенная выше общая постановка задачи применима для описания электрических разрядов различного типа при соответствующем выборе характеристик канала, размера ячейки и шага модельного времени. В применении к компактным разрядам указанный подход обладает рядом особенностей. Прежде всего, это касается морфологии разряда. Как уже отмечалось, короткий и мощный всплеск излучения КВР не соответствует параметрам обычного внутриоблачного разряда, который длится гораздо дольше и излучает намного слабее как в низкочастотном, так и в высокочастотном диапазонах. При этом в рамках известных механизмов не удаётся объяснить необходимую для такого излучения силу тока источника, сопоставимую с током возвратного удара при разряде облако — земля (см., например, [19]), и сравнительно небольшой вертикальный размер, не превышающий нескольких сотен метров [20]. В связи с этим в данной модели предполагается, что КВР имеет две стадии, предварительную и основную. На предварительной стадии на сравнительно небольшом вертикальном расстоянии друг от друга развиваются две биполярные разрядные структуры, вертикальный размер и положение которых определяются масштабом неоднородности начального электрического поля. Время развития структур значительно превосходит длительность всплеска излучения КВР, и за время существования они накапливают на своих противоположных концах (в кронах разрядов) значительный электрический заряд разной полярности. Вторая, основная, стадия компактного разряда начи-

нается, когда противоположно заряженные кроны разрядов по мере их роста вступают в электрический контакт. При этом возникает токовый канал, соединяющий области с противоположным знаком заряда, в результате чего накопленный на предварительной стадии заряд стекает по возникшему проводящему каналу за короткий промежуток времени.

Рассмотренный выше сценарий развития компактного разряда предполагает, прежде всего, наличие начальной среднемасштабной неоднородности плотности заряда и, соответственно, электрического поля. Источником её может служить турбулентная составляющая конвективного потока, усиливающаяся вблизи верхней границы облака. Однако вследствие дисперсии масштабов возмущений плотности заряда следует ожидать значительного разброса параметров КВР, что слабо подтверждается наблюдениями. В связи с этим представляют интерес механизмы, способные сформировать квазирегулярное крупномасштабное возмущение электрического поля (плотности заряда). Вероятным механизмом формирования неоднородного по вертикали распределения плотности заряда с необходимым пространственным масштабом может служить потоковая неустойчивость, впервые описанная в работе [21]. Данная неустойчивость развивается в многокомпонентной внутриоблачной среде при наличии слабо проводящего воздушного потока относительно более тяжёлых внутриоблачных частиц. В результате в облаке формируется экспоненциально растущая волна пространственного заряда, движущаяся с конвективным потоком к вершине облака. Инкремент нарастания плотности заряда в системе отсчёта, связанной с потоком,

$$\gamma_{\max} = \begin{cases} 4\pi\sigma(\Omega/v-1)^2, & \Omega/v < 2; \\ 2\pi\sigma\Omega/v, & \Omega/v \geq 2, \end{cases} \quad (5)$$

определяется аналогом плазменной частоты для тяжёлых частиц $\Omega^2 = 4\pi Q^2 N / M$ (где Q , M и N — заряд, масса и концентрация тяжёлых частиц соответственно), эффективной частотой их соударений v и удельной проводимостью воздушного потока σ . При этом пространственный масштаб $2\pi/k_{\text{opt}}$ возникающей неоднородности плотности заряда зависит также от относительной скорости воздушного потока u :

$$k_{\text{opt}} = \begin{cases} \frac{v}{u} \left(\frac{\Omega}{v} - 1 \right)^{1/2}, & \Omega/v < 2; \\ \Omega/u, & \Omega/v > 2. \end{cases} \quad (6)$$

Согласно оценкам работы [21] характерное время развития потоковой неустойчивости составляет порядка 100 с, а пространственный масштаб неоднородности плотности заряда может варьироваться в широких пределах, от десятков до сотен метров. Важно, что указанный механизм форми-

рования неоднородности плотности заряда оказывается достаточно устойчивым по отношению к характеристикам тяжёлых частиц и величине внешнего электрического поля. Наличие таких частиц даже в верхних слоях развитых грозовых облаков недавно получило экспериментальное подтверждение [22]. В пользу существенного влияния локальных конвективных потоков на возникновение КВР говорит наблюдаемая пространственно-временная кластеризация компактных разрядов, которая проявляется в резком росте частоты их появления в ограниченной области вблизи границ заряженных слоёв в грозовом облаке [23].

Наличие среднемасштабной неоднородности начального распределения электрического поля позволяет естественно решить проблему пространственно-временной синхронизации разрядных структур на предварительной стадии КВР. Дело в том, что для возможности электрического контакта структур необходимо их приблизительно синхронное развитие на сравнительно небольшом расстоянии друг от друга вдоль электрического поля. Вообще говоря, электрическое поле вблизи краёв разрядной структуры по мере её эволюции усиливается и в некоторый момент времени может превысить порог инициации, что приведёт к возникновению ещё одной разрядной структуры. Однако при таком способе синхронизации разрядов расстояние между ними оказывается слишком малым (порядка наименьшего масштаба неоднородности поля) для накопления достаточного пространственного заряда до электрического контакта структур. В результате электрическое поле в промежутке между структурами быстро достигает пробойного значения и вторая структура становится частью более развитой первой структуры без формирования заметной основной стадии. При наличии вариаций плотности заряда с характерным масштабом порядка продольного (вдоль внешнего электрического поля) размера биполярной стримерной структуры поле может аналогично превысить пробойное значение и инициировать второй разряд на расстоянии до нескольких сотен метров от первой структуры. В итоге к моменту контакта обе структуры успевают накопить электрический заряд, достаточный для формирования заметной основной стадии КВР.

Вследствие небольшого относительного превышения электрического поля в области КВР над полем инициации предполагается, что формирующиеся на предварительной стадии проводящие структуры образованы разрядами стримерного типа (см. недавнюю работу [24]), скорость распространения которых составляет $v_{ст} = 5 \cdot 10^5$ м/с [25]. Выбирая размер элементарной ячейки a равным 10 м, определяем шаг модельного времени согласно соотношению $\tau = a / v_{ст} = 20$ мкс. Начальное сопротивление элементарной связи на предварительной стадии согласно модели [10] составляет 10^6 Ом. На основной стадии пробой происходит в среде с предварительной ионизацией и существенно неоднородным распределением

плотности электрического заряда, что позволяет в качестве механизма формирования проводящего канала рассматривать волну ионизации аналогично стадии возвратного удара разряда облако — земля. Важнейшей особенностью волны ионизации является резкое (на порядки величины) снижение погонного сопротивления канала разряда после прохождения фронта волны, что является следствием нагрева, ионизации и газодинамического расширения воздуха. Для описания этого процесса необходимо решать самосогласованную задачу о распространении скачка электрического поля вдоль оси канала с учётом радиального разлёта газа и его ионизации. Такая задача чрезвычайно сложна и в настоящее время далека от решения. Открытым вопросом является и зависимость скорости волны от уровня предварительной ионизации канала распространения. Типичная скорость фронта волны (импульса тока) при возвратном ударе лежит в интервале $(1/3 \div 2/3)c$ [26], где c — скорость света, что хорошо соответствует оценкам скорости распространения импульса тока на основной стадии КВР, полученным в рамках модели передающей линии [19, 20]. В случае КВР уровень предварительной ионизации каналов гораздо ниже, чем при обратном ударе, но по мере распространения фронта волны к центру стримерных структур уровень ионизации каналов растёт, что может привести к росту скорости импульса тока. Данный вопрос требует отдельного исследования, здесь же мы полагаем, что скорость распространения волны ионизации составляет $5 \cdot 10^7$ м/с, что определяет шаг модельного времени на основной стадии $\tau = 0,2$ мкс. Волна ионизации распространяется вверх и вниз от точки контакта, при этом возникающий канал считается хорошим проводником, так что выравнивание электрического потенциала возникающей проводящей структуры происходит за несколько шагов модельного времени. Сильный ток в канале обусловлен быстрым сбором зарядов элементарных ячеек, через которые проходит канал, при этом полный заряд канала стремится к нулю. Важной особенностью канала разряда на основной стадии КВР является его активное ветвление в областях с высокой плотностью заряда, сформированных на предварительной стадии КВР.

3. Результаты моделирования компактного разряда

Рассмотрим для примера развитие КВР в случае, когда в области формирования разряда исходное вертикальное электрическое поле дополнено гармоническим возмущением с пространственным масштабом 200 м (кривая 1 на рис. 1, б), обусловленным развитием потоковой неустойчивости. Высота формирования разряда (около 12 км) и величина электрического поля соответствуют области на границе основного положительного и верхнего экранирующего слоёв в грозном облаке. Видно, что по мере

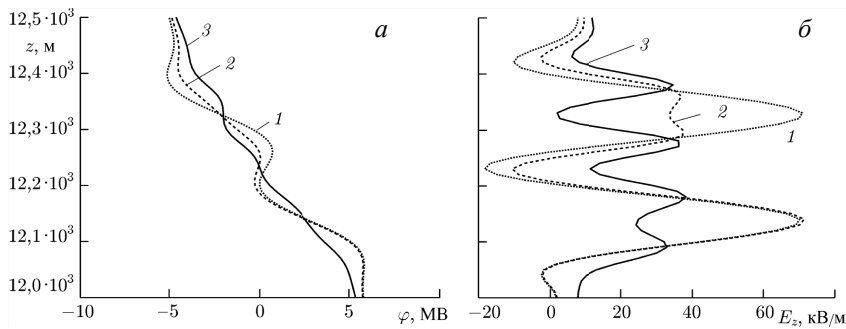


Рис. 1. Высотные профили потенциала (*a*) и вертикальной компоненты электрического поля (*б*) в различные моменты предварительной стадии КВР: кривые 1 соответствуют моменту возникновения первого разряда, 2 — второго разряда, 3 — моменту электрического контакта разрядов и возникновения токового канала основной стадии КВР. Зависимости построены вдоль линии, проходящей через точку инициации разряда

развития первого разряда (первой разрядной структуры на предварительной стадии), возникшего на высоте около 12,35 км, электрическое поле в его окрестности растёт. При этом выше и ниже первого разряда, на высотах полупериодов возмущения поля с отрицательной полярностью, суммарное поле остаётся значительно ниже поля инициации стримерного разряда. В результате в непосредственной близости от первого разряда формируется своего рода запрещённая зона для инициации второго разряда. Напротив, в соседних полупериодах положительного возмущения поля по мере развития первого разряда условия инициации становятся всё более благоприятными (виден, в частности, небольшой рост локального максимума поля на высоте около 12,15 км — см. рис. 1, *б*). В итоге в среде с рассмотренным квазипериодическим распределением плотности заряда развитие биполярной разрядной структуры с необходимостью приводит к инициации второй структуры, отделённой от первой областью слабого электрического поля. При этом задержка возникновения второй структуры мала по сравнению с длительностью предварительной стадии КВР и не приводит к существенному отличию структур в момент их электрического контакта. Отметим, что при формировании нескольких периодов пространственной неоднородности плотности заряда существует возможность последовательной инициации нескольких КВР.

На рис. 2 показаны мгновенные изображения различных стадий эволюции КВР в среде с пространственно неоднородным электрическим полем, распределение которого на оси области расчёта аналогично профилю, приведённому на рис. 1. Видно, что второй разряд возникает с задержкой около 3 мс после инициации первого разряда (рис. 2, *a, б*), после чего оба разряда длительное время существуют одновременно (рис. 2, *в*).

В результате перераспределения заряда провал электрического поля между разрядными структурами уменьшается, пока приблизительно через 11 мс после инициации второго разряда между стримерными структурами не возникает электрический контакт (рис. 2, *г*).

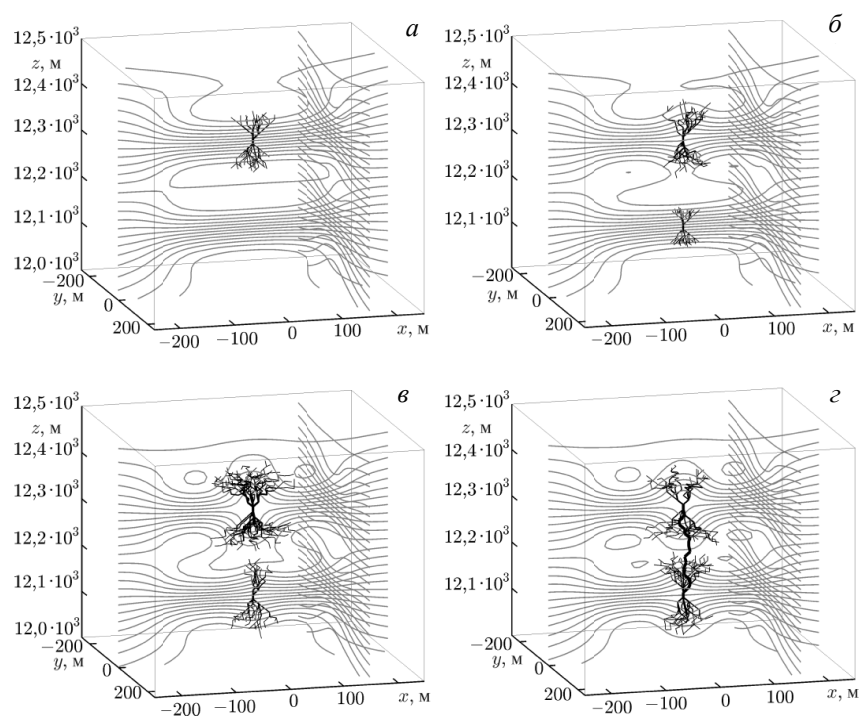


Рис. 2. Последовательные стадии развития КВР в среде с пространственно неоднородным внешним электрическим полем: *а* — развитие первого стримерного разряда за семь шагов модельного времени ($t = 140$ мкс); *б* — начало развития второго разряда ($t = 3,06$ мс); *в* — одновременное развитие пары разрядов ($t = 7,50$ мс); *г* — момент электрического контакта разрядов через 556 шагов модельного времени ($t = 11,12$ мс) с формированием канала с мощным током. Серые линии соответствуют эквипотенциалам в плоскостях $y = 0$ и $x = 150$ м

При расчёте электромагнитного поля на предварительной стадии КВР естественно предположить, что низкочастотная составляющая излучения токовой системы определяется её крупномасштабным медленно меняющимся сегментом. Излучение такого сегмента удобно рассчитать, рассматривая в качестве источника эффективный линейный ток, полученный путём сложения всех элементарных токов разряда в каждой горизонтальной плоскости. При этом благодаря большому числу элементов флуктуа-

ции такого суммарного тока эффективно усредняются. Напротив, высокочастотное излучение связано с первоначальным электрическим пробоем между соседними ячейками, т. е. с возникновением новых проводящих каналов. При этом характеристики высокочастотного излучения определяются длиной и ориентацией разрядного промежутка, скоростью распространения и формой импульса тока. Исходя из этого, в частности, следует ожидать, что высокочастотные компоненты излучения на предварительной и основной стадиях КВР существенно различны.

Используя описанный выше подход, можно найти электромагнитное поле на предварительной стадии рассматриваемого компактного разряда в рамках приближения передающей линии. Расчёты показывают, что в низкочастотной составляющей поля при наземных наблюдениях вблизи разряда, когда расстояние от его оси не превышает 10 км, доминирует электростатическая компонента, достигающая 0,45 В/м, тогда как индукционная и радиационная компоненты по порядку величины не превышают 3 мВ/м и 3 мкВ/м соответственно [10]. На достаточно большом расстоянии (около 100 км от оси разряда) все компоненты поля весьма малы и не могут быть зарегистрированы в современных экспериментах. Результат расчёта высокочастотного излучения на предварительной стадии КВР приведён на рис. 3.

Как было отмечено в предыдущем разделе, основная стадия КВР начинается в момент электрического контакта двух разрядных структур, сформировавшихся на предварительной стадии. На основной стадии КВР

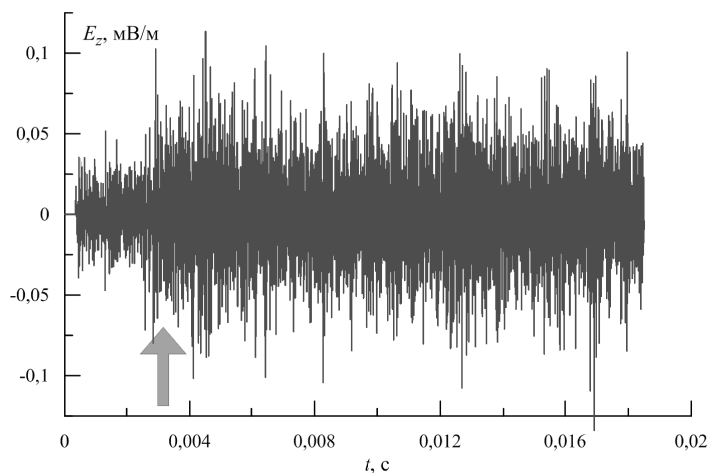


Рис. 3. Высокочастотное электромагнитное излучение биполярных стримерных структур на предварительной стадии КВР на расстоянии 100 км от оси разряда. Стрелкой показан момент появления второй структуры; правая граница реализации соответствует моменту начала основной стадии КВР

за сравнительно короткий промежуток времени (несколько микросекунд) развивающийся разряд нейтрализует электрические заряды, накопленные на смежных гранях разрядных структур. В результате на основной стадии КВР формируется короткий интенсивный импульс тока (сильный ток обусловлен быстрым сбором зарядов элементарных ячеек, через которые проходит канал разряда, при этом полный заряд канала близок к нулю), сопровождающийся интенсивным ветвлением канала в областях повышенной плотности заряда (см. [9]).

Для расчёта электромагнитного излучения на основной стадии КВР можно применить подход, аналогичный использованному выше для предварительной стадии. При этом удобно воспользоваться аналитической аппроксимацией пространственно-временного распределения средней линейной компоненты тока на основной стадии разряда [10]. Полагая, что основная стадия разряда начинается в момент времени $t = 0$, пространственно-временное распределение средней линейной компоненты тока основной стадии разряда можно аппроксимировать выражением

$$i_{oc}(z, t) = I_{oc}(t) \exp\left(-\frac{|z - z_{oc}|^\gamma}{(v_{oc}t)^\gamma} - \frac{|z - z_{oc}|^\gamma}{z_0^\gamma}\right), \quad (7)$$

где $z_{oc} = 12,25$ км — высота центра разряда, $v_{oc} = 5 \cdot 10^7$ м/с — скорость расширения области разряда вверх и вниз вдоль вертикальной оси, z_0 — вертикальный масштаб разряда, параметр γ определяет толщину переходной области на границе области тока. Временной фактор в выражении (7) имеет вид

$$I_{oc}(t) = \begin{cases} 0, & t < 0; \\ I_{oc0} \exp[-(t/\tau_{oc})]^{1,1} [1 - \exp[-(t/\tau_n)]^{4,2}], & t \geq 0, \end{cases} \quad (8)$$

где $\tau_n = 1,1 \cdot 10^{-6}$ с — время нарастания среднего тока, $\tau_{oc} = 1,4 \cdot 10^{-6}$ с, $I_{oc0} = 130$ кА. Используя аппроксимацию (7), (8), легко подсчитать, что полный заряд, нейтрализованный на основной стадии, составляет $Q_{oc} = \int_0^\infty i_{oc}(z_{oc}, t) dt \approx 0,08$ Кл и соответствует заряду меньшей (в данном случае — верхней) структуры, сформировавшейся на предварительной стадии разряда. Рассчитанный временной профиль среднего линейного тока на высоте центра разряда и его аппроксимация (8) показаны на рис. 4, а, пространственно-временное распределение модельного тока (7), (8) приведено на рис. 4, б.

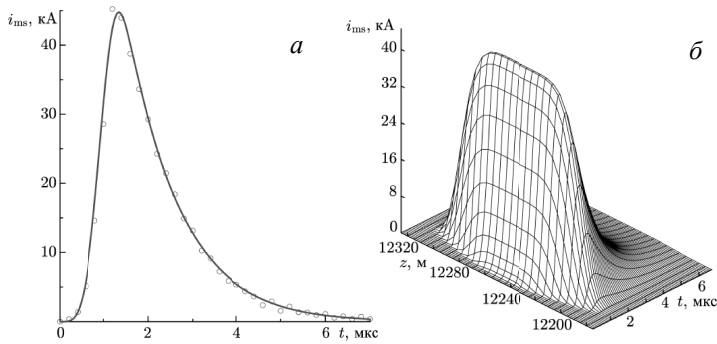


Рис. 4. Сопоставление временного профиля тока в центральном канале разряда (символы \circ) с аналитической аппроксимацией (8) (а) и пространственно-временное распределение среднего линейного тока (7), (8) на основной стадии КВР (б)

Электростатическую, индукционную и радиационную компоненты поля вертикального линейного тока $i_{oc}(z, t)$ можно рассчитать в приближении передающей линии аналогично тому, как это было сделано для средних токов на предварительной стадии КВР. Результаты расчёта указанных компонент поля на поверхности земли при разных расстояниях r от оси разряда приведены на рис. 5, а—в, на рис. 5, г показано суммарное поле на расстояниях 1 и 10 км от оси разряда. Из рисунка 5 видно, что, в отличие от предварительной стадии КВР, на небольших расстояниях от оси разряда (при $r = 1$ и 10 км) все компоненты электрического поля дают заметный вклад в полное поле. Практически под разрядом, при $r = 1$ км, доминирует индукционная компонента, превышающая 3 В/м, тогда как статическая и радиационная компоненты по порядку величины составляют около 0,15 и 0,7 В/м соответственно. Но уже на расстоянии $r = 10$ км наибольшей становится радиационная компонента, амплитуда которой достигает почти 35 В/м, тогда как индукционная компонента спадает до 0,8 В/м, а статическая компонента по порядку величины не превышает 30 мВ/м. То же соотношение между компонентами сохраняется и при больших расстояниях от разряда. При $r = 100$ км амплитуда радиационной компоненты поля, имеющей вид одиночного импульса, составляет приблизительно 14 В/м.

В качестве источника высокочастотного излучения на основной стадии КВР, как и на предварительной стадии, рассматривается совокупность пробойных токов разрядного древа, обеспечивающих его приращение на каждом шаге модельного времени. Пробойный ток, формирующий элементарный проводящий канал между соседними ячейками среды, характеризуется координатами и временем начала, пространственной ориентацией и временным профилем. Как и для предварительной стадии, считаем,

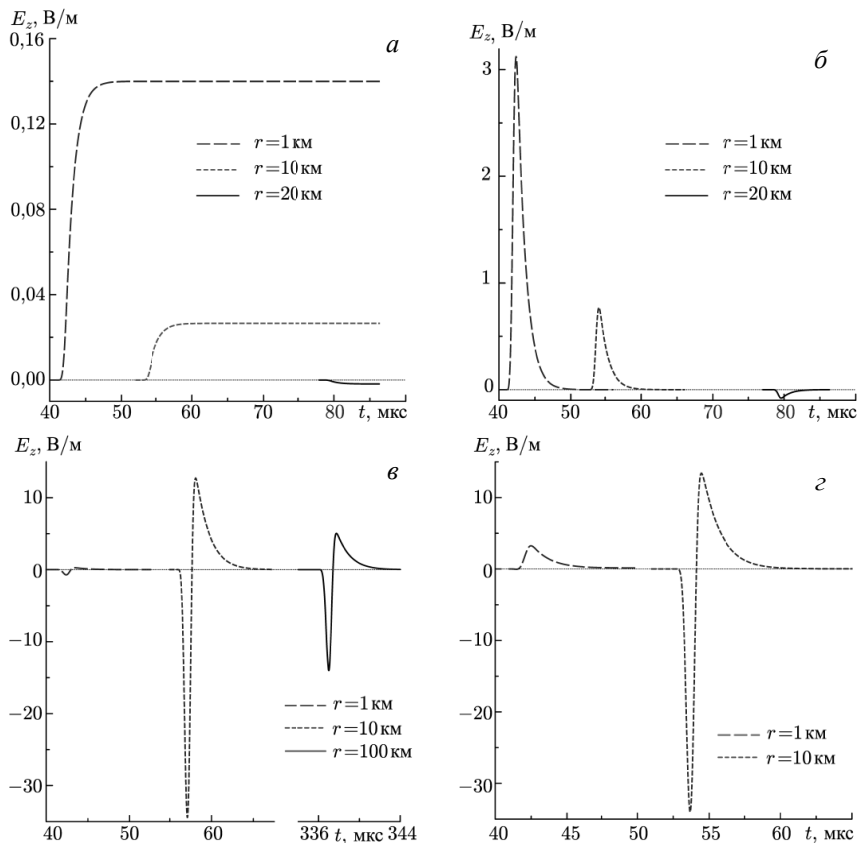


Рис. 5. Временные профили электростатической (*а*), индукционной (*б*) и радиационной (*в*) компонент приземного электрического поля, а также суммарное электрическое поле (*г*) на различных расстояниях r от оси основной стадии рассматриваемого компактного разряда

что при распространении вдоль канала профиль пробойного тока не меняет форму. Отличием пробойных токов на основной стадии КВР является, прежде всего, высокая скорость распространения импульса тока, равная скорости волны ионизации. Кроме того, форма импульса тока близка к функции единичного скачка, что качественно описывает распространение фронта ионизации. Определяя амплитуду, пространственную ориентацию и время начала каждого пробойного тока из модели (1) — (3) и складывая их поля на каждом шаге модельного времени, можно получить суммарное высокочастотное поле на основной стадии КВР. Временная реализация высокочастотной компоненты приземного электрического поля на расстоянии

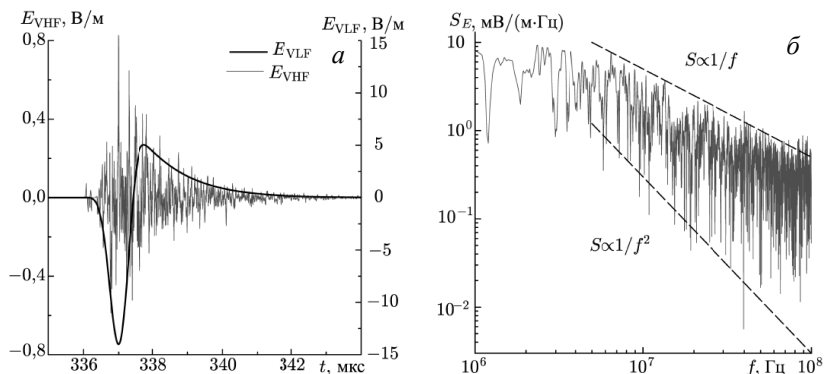


Рис. 6. Временная реализация высокочастотной компоненты электрического поля E_{VHF} на основной стадии КВР в сравнении с синхронным коротким импульсом низкочастотного электрического поля E_{VLF} (а) и спектр высокочастотной компоненты поля (б) на расстоянии $r = 100$ км от оси разряда

$r = 100$ км от рассматриваемого модельного разряда и её спектр показаны на рис. 6. Видно, что уровень высокочастотного излучения на основной стадии КВР на несколько порядков величины выше, чем на предварительной стадии (ср. рис. 3 и 6, а). Кроме того, рис. 6, а иллюстрирует хорошую синхронизацию всплеска высокочастотного излучения и короткого биполярного импульса низкочастотного электрического поля. Спектр высокочастотного излучения, как следует из рис. 6, б, лежит между двумя зависимостями вида f^{-1} и f^{-2} , при этом верхняя из указанных границ соответствует спектру процесса с критической динамикой, а нижняя связана с существованием в разряде нескольких независимых реализаций таких процессов, интенсивность которых распределена по нормальному закону.

4. Обсуждение результатов

После открытия в 80-е годы прошлого века [27] компактные разряды стали объектом интенсивных экспериментальных и теоретических исследований, достаточно подробный обзор результатов которых приведён в работе [9]. Накопленные к настоящему времени данные позволили статистически достоверно определить характерные параметры компонент излучения КВР в дальней зоне — узкого биполярного импульса электрического поля (narrow bipolar pulse, NBP) и синхронного с ним всплеска высокочастотного излучения [20, 28—30], а также оценить изменение зарядового момента при разряде [29, 31, 32].

Расчёты излучения КВР в рамках представленной выше фрактальной модели показывают хорошее соответствие данным наблюдений (см.,

в частности, [10]). Прежде всего, в модели на предварительной стадии КВР электромагнитное излучение оказывается сравнительно слабым: в большинстве случаев какая-либо электрическая активность в грозовом облаке до и после компактного разряда не регистрируется. Полученные в рамках численного эксперимента значения длительности и амплитуды импульса электрического поля в дальней зоне (до нескольких десятков микросекунд и до нескольких десятков вольт на метр на расстоянии 100 км от разряда), а также изменение зарядового момента (до нескольких десятых кулона на километр) соответствуют измерениям для типичных компактных разрядов. Отметим, что данные параметры излучения компактного разряда существенно зависят от его длины, т. е., фактически, от вертикального профиля возмущения электрического поля в области КВР. Что касается короткого мощного всплеска высокочастотного излучения, согласно предварительным оценкам его интенсивность и длительность хорошо соответствуют наземным измерениям, а спектральная плотность мощности — характеристикам регистрируемых спутниками высокочастотных сигналов атмосферного происхождения, так называемых пар трансферных импульсов (transionospheric pulse pair, TIPP), которые представляют собой прямой и отражённый от земли сигналы КВР [33]. Важное свидетельство в пользу представленной модели даёт сопоставление наблюдений компактных разрядов и данных метеорологических радаров, показывающее, что мощная конвекция является необходимым, но не достаточным условием возникновения КВР [23].

Таким образом, в рамках предложенной фрактальной модели получили объяснение важные особенности КВР, в частности слабое (с уровнем ниже установленных в экспериментах порогов обнаружения) излучение на предварительной стадии, формирование короткого биполярного импульса электрического поля и синхронного сверхмощного всплеска высокочастотного излучения. При этом вследствие особенности пространственно-временной структуры разрядного тока, область которого быстро расширяется вверх и вниз от точки электрического контакта разрядных структур, импульс электрического поля в дальней зоне остаётся узким в широком диапазоне параметров разряда [10]. Отметим, что известные модели компактного разряда не дают самосогласованного описания его широкополосного электромагнитного излучения (см., в частности, [27, 34—37]).

В заключение отметим, что представление разряда как сети клеточных автоматов в настоящее время, фактически, даёт единственную возможность описать его макроскопическую эволюцию, поскольку традиционные трёхмерные расчёты поведения разряда чрезвычайно трудоёмки и не отражают ряд ключевых особенностей рассматриваемого явления, в частности диссипацию и ветвление каналов разряда, а также не описывают его макроскопическую токовую систему. Более того, привлечение методов фрактальной геометрии для процессов в сильно неравновесных системах и

детальный учёт эффектов кластер-кластерной агрегации представляется весьма перспективным с точки зрения построения общей теории широкого класса электрических процессов, от инициации молнии до высотных разрядов (transient luminous events, TLE) в атмосфере.

Литература

1. Трахтенгерц, В. Ю. О фрактальной динамике активных сред / В. Ю. Трахтенгерц, Д. И. Иудин, А. Н. Григорьев // *Нелинейные волны* 2002 : сб. / под ред. А. В. Гапонова-Грехова и В. И. Некоркина. — Н. Новгород : ИПФ РАН, 2003. — С. 287—302.
2. Иудин, Д. И. Динамическая перколяция / Д. И. Иудин, В. Ю. Трахтенгерц // *Нелинейные волны* 2004 : сб. / под ред. А. В. Гапонова-Грехова, В. И. Некоркина. — Н. Новгород : ИПФ РАН, 2005. — С. 217—242.
3. Иудин, Д. И. Фрактальные лабиринты: структурная динамика / Д. И. Иудин, В. Ю. Трахтенгерц // *Нелинейные волны* 2006 : сб. / под ред. А. В. Гапонова-Грехова, В. И. Некоркина. — Н. Новгород : ИПФ РАН, 2007. — С. 360—377.
4. Иудин, Д. И. Перколяционные эффекты в биоэкологических системах / Д. И. Иудин, Д. Б. Гелашвили // *Нелинейные волны* 2010 : сб. / под ред. А. В. Гапонова-Грехова, В. И. Некоркина. — Н. Новгород : ИПФ РАН, 2011. — С. 412—434.
5. Иудин, Д. И. Фрактальные аспекты броуновского движения // *Нелинейные волны* 2012 : сб. / под ред. А. Г. Литвака, В. И. Некоркина. — Н. Новгород : ИПФ РАН, 2013. — С. 67—82.
6. Фракталы и мультифракталы в биоэкологии / Д. Б. Гелашвили [и др.] — Н. Новгород : ННГУ, 2013. — 370 с.
7. Iudin, D. I. Fractals: as simple as complex / D. I. Iudin, E. V. Kuposov. — Nizhny Novgorod : NiSOC, 2013. — 184 p.
8. Iudin, D. I. Percolation transition and related phenomena in terms of grossone infinity computations / D. I. Iudin, Ya. D. Sergeyev // *Advances in unconventional computing. V. 1 : Theory* / ed. A. Adamatzky. — Cham, Switzerland : Springer International Publisher, 2016. — P. 239—266.
9. Иудин, Д. И. Фрактальная модель компактного внутриоблачного разряда. I. Особенности строения и эволюции / Д. И. Иудин, С. С. Давыденко // *Изв. вузов. Радиофизика*. — 2015. — Т. 58, № 7. — С. 530—551.
10. Давыденко, С. С. Фрактальная модель компактного внутриоблачного разряда. II. Особенности электромагнитного излучения / С. С. Давыденко, Д. И. Иудин // *Изв. вузов. Радиофизика*. — 2016. — Т. 59, № 7. — С. 620—637.
11. Rakov, V. A. Lightning: Physics and Effects / V. A. Rakov, M. A. Uman. — N. Y. : Cambridge University Press, 2003. — 687 p.
12. Bak, P. How nature works : The science of self-organized criticality / P. Bak. — N. Y. : Springer-Verlag, 1996. — 212 p.
13. Iudin, D. I. Fractal dynamics of electric discharges in a thundercloud / D. I. Iudin, V. Yu. Trakhtengerts, M. Hayakawa // *Phys. Rev. E*. — 2003. — V. 68. — Art. 016601.

14. *Iudin, D. I.* Thundercloud cellular automaton model / D. I. Iudin, V. Yu. Trakhtengerts, A. N. Grigoriev // Nuclear instruments and methods in physics research. Sec. A : Accelerators, spectrometers, detectors and associated equipment. — 2003. — V. 502, № 2/3. — P. 526—528.
15. *Trakhtengerts, V. Yu.* Actual problems of thundercloud electrodynamics / V. Yu. Trakhtengerts, D. I. Iudin // Sprites, Elves and Intence Lightning Discharges / eds. M. Fullekrug, E. Mareev. — Dordrecht : Springer, 2006. — P. 341—376.
16. *Rompe, R.* Uber des Toeplersche Funkengesetz (On the Toepler spark theory) / R. Rompe, W. Weizel // Zeitschrift für Physik. — 1944. — V. 122. — S. 636—639.
17. *Uman, M. A.* The electromagnetic radiation from a finite antenna / M. A. Uman, D. K. McLain, E. P. Krider // Am. J. Phys. — 1975. — V. 43. — P. 33—38.
18. *Thottappillil, R.* Distribution of charge along the lightning channel: Relation to remote electric and magnetic fields and to return-stroke models / R. Thottappillil, V. Rakov, M. Uman // J. Geophys. Res. — 1977. — V. 102, № D6. — P. 6987—7006.
19. *Nag, A.* Compact intracloud lightning discharges: 2. Estimation of electrical parameters / A. Nag, V. A. Rakov // J. Geophys. Res. — 2010. — V. 115. — Art. D20103.
20. *Smith, D. A.* A distinct class of isolated intracloud discharges and their associated radio emissions / D. A. Smith [et al.] // J. Geophys. Res. — 1999. — V. 104, № D4. — P. 4189—4212.
21. *Трахтенгерц, В. Ю.* О природе электрических ячеек в грозовом облаке // Докл. АН СССР. — 1989. — Т. 308, № 3. — С. 584—586.
22. *Lazarus, S. M.* A Thermodynamic, kinematic and microphysical analysis of a jet and gigantic jet-producing Florida thunderstorm / S. M. Lazarus [et al.] // J. Geophys. Res. : Atmos. — 2015. — V. 120. — P. 8469—8490.
23. *Wiens, K. C.* Relationships among narrow bipolar events, “total” lightning, and radar-inferred convective strength in Summer 2005 Great Plains thunderstorms / K. C. Wiens [et al.] // J. Geophys. Res. — 2008. — V. 113. — Art. D05201.
24. *Kostinskiy, A. Yu.* Observation of a new class of electric discharges within artificial clouds of charged water droplets and its implication for lightning initiation within thunderclouds / A. Yu. Kostinskiy [et al.] // Geophys. Res. Lett. — 2015. — V. 42. — P. 8165—8171.
25. *Райзер, Ю. П.* Физика газового разряда / Ю. П. Райзер. — Долгопрудный : Интеллект, 2009. — 736 с.
26. *Rakov, V. A.* Lightning return stroke speed // J. Lightning Res. — 2007. — V. 1. — P. 80—89.
27. *Le Vine, D. M.* Sources of the strongest RF radiation from lightning // J. Geophys. Res. — 1980. — V. 85, № C7. — P. 4091—4095.
28. *Willett, J. C.* A class of unusual lightning electric field waveforms with very strong high-frequency radiation / J. C. Willett, J. C. Bailey, E. P. Krider // J. Geophys. Res. — 1989. — V. 94, № D13. — P. 16255—16267.
29. *Nag, A.* On phenomenology of compact intracloud lightning discharges / A. Nag [et al.] // J. Geophys. Res. — 2010. — V. 115. — Art. D14115.
30. *Karunarathne, S.* Observations of positive narrow bipolar pulses / S. Karunarathne [et al.] // J. Geophys. Res. : Atmos. — 2015. — V. 120. — P. 7128—7143.

31. *Eack, K. B.* Electrical characteristics of narrow bipolar events // *Geophys. Res. Lett.* — 2004. — V. 31, № 20. — Art. L20102.
32. *Rison, W.* Observations of narrow bipolar events reveal how lightning is initiated in thunderstorms / W. Rison [et al.] // *Nature Communications.* — 2016. — V. 7. — Art. 10721.
33. *Holden, D. N.* Satellite observations of transionospheric pulse pairs / D. N. Holden, C. P. Munson, J. C. Devenport // *Geophys. Res. Lett.* — 1995. — V. 22, № 8. — P. 889—892.
34. *Gurevich, A. V.* Runaway electron mechanism of air breakdown and preconditioning during a thunderstorm / A. V. Gurevich, G. M. Milikh, R. Roussel-Dupre // *Phys. Lett. A.* — 1992. — V. 165. — P. 463—468.
35. *Tierney, H. E.* Radio frequency emissions from a runaway electron avalanche model compared with intense, transient signals from thunderstorms / H. E. Tierney [et al.] // *J. Geophys. Res.* — 2005. — V. 110. — Art. D12109.
36. *Cooray, V.* Electromagnetic fields of a relativistic electron avalanche with special attention to the origin of lightning signatures known as narrow bipolar pulses / V. Cooray [et al.] // *Atmos. Res.* — 2014. — V. 149. — P. 346—358.
37. *da Silva, C. L.* Physical mechanism of initial breakdown pulses and narrow bipolar events in lightning discharges / C. L. da Silva, V. P. Pasko // *J. Geophys. Res. : Atmos.* — 2015. — V. 120, № 10. — P. 4989—5009.

ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧ СТАБИЛИЗАЦИИ

А. А. Корнев

Московский государственный университет им. М. В. Ломоносова
Институт вычислительной математики им. Г. И. Марчука РАН, г. Москва

Введение

Типичная задача вычислительной математики, возникающая при решении обыкновенных дифференциальных уравнений либо нестационарных уравнений в частных производных, заключается в разработке и обосновании алгоритмов, позволяющих по заданным начальным условиям восстановить с гарантированной точностью соответствующую траекторию. Если исследуемая система содержит управляющую функцию, то при наличии подобных алгоритмов появляется возможность не только исследовать динамику системы для конкретных данных, но и допустимым образом влиять на ее эволюцию, направляя к траектории с требуемыми свойствами. Для систем седлового типа решение подобных задач стабилизации удастся построить в терминах проектирования на локально устойчивые и неустойчивые многообразия в окрестности либо неподвижной точки, либо траектории седлового типа. Соответствующие численные алгоритмы основываются на классических результатах теории инвариантных многообразий для частично гиперболических динамических систем. В рамках данного подхода задача о построении управления, гарантирующего сближение заданных траекторий, сводится к задаче проектирования на многообразие, которое определяется решением некоторого функционального уравнения. Найденное таким образом управление, по сути, решает задачу стабилизации на неустойчивом подпространстве малой размерности, а сближение траекторий в остальной части пространства обеспечивается свойствами оператора эволюции. Соответствующие алгоритмы позволяют находить решения задач асимптотической стабилизации по начальным данным, краевым условиям, правой части для нестационарных конечно-разностных уравнений, обыкновенных дифференциальных уравнений, уравнений в частных производных. В том числе для задачи Лоренца, Бюргера, Навье — Стокса, мелкой воды, баротропного вихря на сфере.

Постановка задачи

Пусть на полном нормированном пространстве $\{H, \|\cdot\|\}$ задано отображение $S: H \rightarrow H$, имеющее неподвижную точку $z_0 \in H$, т. е. $S(z_0) = z_0$. Пусть имеется некоторая начальная точка $a_0 \in H$ и построена соответствующая траектория, порождаемая оператором S :

$$\{a_0, a_1 = S(a_0), a_2 = S(a_1) = S^2(a_0), \dots, a_i = S^i(a_0), \dots\}.$$

В данном случае $S^i(\cdot)$ означает i -кратное последовательное применение оператора $S(\cdot)$. Нас интересует задача нахождения множества $\mathcal{W}^- = \mathcal{W}^-(S, \mathcal{O}_{z_0})$ всех начальных данных a_0 из некоторой окрестности \mathcal{O}_{z_0} неподвижной точки z_0 , траектории которых стремятся к z_0 , т. е. $\|z_0 - S^i(a_0)\| \rightarrow 0$ при $i \rightarrow \infty$, а также методы проектирования a_0 на это множество \mathcal{W}^- .

Линейные отображения

Рассмотрим решение задачи в линейном случае, т. е. будем считать, что отображение имеет вид $S(z_0 + u) = S(z_0) + Lu$, где L — ограниченный линейный оператор, действующий из H в H . Наибольший (с точки зрения приложений) интерес представляют системы, обладающие следующими свойствами:

- a₁) $H = H_+ \oplus H_-$;
- a₂) $L(H_+) = H_+$, $L(H_-) \subset H_-$;
- a₃) $\|Lv\| \geq \mu_+ \|v\|$, $\forall v \in H_+$;
- $\|Lw\| \leq \mu_- \|w\|$, $\forall w \in H_-$, $0 \leq \mu_- < \mu < \mu_+$, $\mu \leq 1$.

Это означает, что исходное пространство H может быть разложено в прямую сумму двух подпространств, инвариантных относительно оператора L . Если указанные оценки выполняются для $\mu = 1$, то подпространство H_- сжимается, а H_+ растягивается под действием оператора L . В этом случае точка z_0 называется неподвижной точкой седлового типа. Для рассматриваемых задач стабилизации также допустимы условия $\mu < 1$ и $\mu_+ < 1$, означающие, что на части подпространства H_+ оператор L может сжимать, но не сильнее, чем на H_- . Отметим, что $\mu_+ > 0$, поэтому оператор L является невырожденным и обратимым на H_+ . Далее для простоты считаем, что $\mu = 1$.

При стабилизации линейных систем важную роль играет следующая теорема.

Теорема 1. Пусть линейное отображение S имеет указанный вид и выполнены условия «a₁₋₃». Тогда

$$\lim_{i \rightarrow \infty} \|z_0 - S^i(a_0)\| = 0 \Leftrightarrow (z_0 - a_0) \in H_-.$$

Условиям теоремы удовлетворяет, например, система, определенная на конечномерном действительном пространстве $H = \mathbf{R}^N$ с оператором L , имеющим в некотором базисе $\{e_i\}_1^N$ диагональный вид $\text{diag}(\mu_1, \dots, \mu_N)$, где $|\mu_i| \geq \mu_+$ при $1 \leq i \leq i_0$, $|\mu_i| \leq \mu_-$ при $i_0 + 1 \leq i \leq N$. В этом случае

$$H_+ = \text{span} \langle e_i, i = 1, \dots, i_0 \rangle, \quad H_- = \text{span} \langle e_i, i = i_0 + 1, \dots, N \rangle.$$

И если в качестве нормы векторов $u = \sum_{i=1}^N \alpha_i e_i$ взять $\|u\|_\infty = \max_{i=1, \dots, N} |\alpha_i|$, то условия теоремы будут выполнены.

Отметим, что для произвольной нормы можно гарантировать только оценки вида $\|L^i v\| \geq C_+ \mu_+^i \|v\|$ и $\|L^i w\| \leq C_- \mu_-^i \|w\|$, означающие асимптотическое растяжение и сжатие на H_+ и H_- соответственно. Для операторов L общего вида верна теорема 2.

Теорема 2. Пусть $H = \mathbf{R}^N$ и собственные числа оператора L удовлетворяют условию

$$|\lambda_1| \geq \dots \geq |\lambda_{i_0}| \geq \mu_+ > 1 > \mu_- \geq |\lambda_{i_0+1}| \geq \dots \geq |\lambda_N|.$$

Тогда собственные и присоединенные векторы L , отвечающие жордановым клеткам для $|\lambda_j| > 1$, образуют базис в инвариантном подпространстве H_+ , а оставшиеся собственные и присоединенные векторы, отвечающие жордановым клеткам для $|\lambda_j| < 1$, образуют базис в инвариантном подпространстве H_- . При этом для произвольного $\varepsilon > 0$ найдется такая константа C , что

$$\begin{aligned} \|L^i v\| &\geq C(\mu_+ - \varepsilon)^i \|v\|, \quad \forall v \in H_+, \\ \|L^i w\| &\leq C(\mu_- + \varepsilon)^i \|w\|, \quad \forall w \in H_-. \end{aligned}$$

Отметим важную специфику жордановых клеток. Пусть, например, в двумерном пространстве в естественном базисе матрица оператора L имеет вид $\begin{pmatrix} 0,5 & 1 \\ 0 & 0,5 \end{pmatrix}$. Тогда неустойчивое подпространство отсутствует и все векторы под действием L асимптотически стремятся к нулю. Однако евклидова норма вектора $a_0 = (0, 1)^T$ только после третьего шага начнет убывать. Отметим, что чем большего размера жордановы клетки присутствуют в матрице, тем хуже сжимающие свойства L . Для подпространства H_+ картина аналогична. Кроме того, в некоторых нормах оценки «а₃» ухудшаются при уменьшении угла между подпространствами H_+ и H_- .

В общем случае разложение $H = H_+ \oplus H_-$ позволяет определить операторы проектирования P_+ и P_- вдоль подпространств $H_- = P_-[H]$ и $H_+ = P_+[H]$ на подпространства H_+ и H_- соответственно. Подпростран-

ство $P_+[H]$ обычно конечномерно, т. е. $P_-[H]$ имеет конечную коразмерность, поэтому операторы P_{\pm} удобно задать следующими наборами базисных векторов: $P_+[H] = \text{span} \langle \xi_1, \dots, \xi_{i_0} \rangle$, $P_-[H] = \text{span} \langle \eta_1, \dots, \eta_{i_0} \rangle$ для $i = 1, \dots, i_0$, где $P_-[H] \perp P_+[H]$. Далее будем считать, что системы векторов $\{\xi_i\}$, $\{\eta_i\}$ известны и линейно независимы. Тогда построение $P_+[u]$ для произвольного $u \in H$ сводится, с учетом представления $u = \sum_{i=1}^{i_0} c_i \xi_i + P_-[u]$ и условия ортогональности $\{\eta_i\}$ и $P_-[u]$, к решению системы $\sum_{i=1}^{i_0} c_i (\xi_i, \eta_j) = (u, \eta_j)$, $j = 1, \dots, i_0$, и последующему нахождению $P_+[u] = \sum_{i=1}^{i_0} c_i \xi_i$. Проекция $P_-[u]$ вычисляется по формуле $P_-[u] = u - P_+[u]$.

Векторы $\{\xi_i, \eta_i\}$ можно найти с помощью стандартного пакета ARPACK с учетом следующей леммы и ее следствия.

Лемма 1. *Собственные числа матриц A и A^T совпадают. Если $A\xi = \lambda_1\xi$, $A^T\eta = \lambda_2\eta$ и $\lambda_1 \neq \lambda_2$, то $(\xi, \eta) = 0$, т. е. собственные векторы матриц A , A^T , отвечающие различным собственным числам, ортогональны.*

Отметим, что множества $\{\lambda_i(A)\}$ и $\{\lambda_i(A^T)\}$ совпадают согласно определению собственных чисел. Вторая часть леммы следует из равенства $\lambda_1(\xi, \eta) = (A\xi, \eta) = (\xi, A^T\eta) = \lambda_2(\xi, \eta)$.

Следствие 1. *Подпространства $P_+[H]$ и $P_-[H]$ состоят из векторов, отвечающих собственным числам по модулю больше μ для операторов L и L^T соответственно.*

Для рассмотренного класса линейных операторов, имеющих неподвижную точку седлового типа, рассмотрим так называемую задачу стабилизации по начальным данным: для начальной точки a_0 , фиксированного подпространства $\mathcal{L} \subset H$, заданного набором линейно независимых элементов $\mathcal{L} = \text{span} \langle e_1, \dots, e_{i_0} \rangle$, определить поправку $l = \sum_{i=1}^{i_0} c_i e_i$ из условия

$$\lim_{i \rightarrow \infty} \|S^i(a_0 + l) - z_0\| \leq C\mu^i, \quad i = 0, 1, 2, \dots,$$

т. е., согласно теореме 1, из условия $(a_0 + l - z_0) \in H_-$. Отметим, что искомые коэффициенты могут быть найдены из системы

$$\left(a_0 - z_0 + \sum_{i=1}^{i_0} c_i e_i, \eta_j \right) = 0, \quad j = 1, \dots, i_0.$$

Теорема 3. Пусть система векторов $\{e_i\}_1^{i_0}$ линейно независима и $\{e_i^+ = P_+[e_i]\}_1^{i_0}$ образуют базис в $P_+[H]$. Тогда для произвольного $a_0 \in H$ существует единственный вектор $l \in \mathcal{L}$, такой, что $(a_0 + l - z_0) \in H_-$, $l = \sum_{i=0}^{i_0} c_i e_i$.

Исследование линейных математических моделей, записанных в виде системы либо дифференциальных уравнений, либо уравнений в частных производных, обычно проводится на основе следующего утверждения.

Теорема 4. Пусть $H = \mathbf{R}^N$ и задана система линейных дифференциальных уравнений $x'(t) = Ax(t)$, $x(0) = a_0$, с постоянной квадратной матрицей A , порождающей оператор $S(t, \cdot)$ вида $x(t) = S(t, x(0))$. Пусть собственные числа матрицы A удовлетворяют условию $\operatorname{Re}(\lambda_i) \geq \lambda_+ > 0$ при $i \leq i_0$, $\operatorname{Re}(\lambda_i) \leq \lambda_- < 0$ при $i \geq i_0 + 1$. Тогда собственные и присоединенные векторы A , отвечающие жордановым клеткам для $\operatorname{Re}(\lambda_i) > 0$, образуют базис в инвариантном подпространстве H_+ , а оставшиеся собственные и присоединенные векторы, отвечающие жордановым клеткам для $\operatorname{Re}(\lambda_i) < 0$, образуют базис в инвариантном подпространстве H_- . При этом для произвольного $\varepsilon > 0$ найдется такая константа C , что

$$\|v(t)\| \geq Ce^{(\lambda_+ - \varepsilon)t} \|v(0)\|, \quad \forall v(0) \in H_+,$$

$$\|w(t)\| \leq Ce^{(\lambda_- + \varepsilon)t} \|w(0)\|, \quad \forall w(0) \in H_-, \quad t \geq 0.$$

Отметим, что при решении уравнений в частных производных пространство H бесконечномерно. Поэтому при выделении подпространств H_{\pm} и получении оценок типа «а₃» необходимо обращать внимание на конечномерность жордановых клеток и отделенность от нуля углов между корневыми подпространствами.

Пример 1. Пусть в пространстве $H = L_2[0, \pi]$ задано уравнение типа теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + bu, \quad u = u(t, x),$$

$$t \geq 0, \quad x \in \omega = [0, \pi], \quad (1)$$

$$u(t, 0) = u(t, \pi) = 0, \quad u(0, x) = a(x) = \sum_{k=1}^{\infty} c_k \sin(kx).$$

Его решение несложно получить методом разделения переменных (Фурье): $u(t, x) = \sum_{k=1}^{\infty} c_k e^{\lambda_k t} \sin(kx)$, где $\lambda_k = b - k^2$. Пусть существует такое i_0 , что $\lambda_{i_0} = b - i_0^2 > 0 > b - (i_0 + 1)^2 = \lambda_{i_0+1}$. Тогда условие $\|u(t)\|_{L_2(0,1)} \rightarrow 0$ при $t \rightarrow \infty$ эквивалентно условиям $c_k = 0$, $k = 1, 2, \dots, i_0$, т. е. вложению $u(0, x) \in P_- [L_2(0, 1)] = \text{span} \langle \sin(kx), k = i_0 + 1, \dots \rangle$. Отсюда, с учетом ортогональности функций $\{\sin(kx)\}$, получаем, что условие убывания с течением времени нормы решения эквивалентно условию $u(0, x) \perp P_- [L_2(0, 1)] = P_+ [L_2(0, 1)] = \text{span} \langle \sin(kx), k = 1, \dots, i_0 \rangle$.

Рассмотрим подробнее решение задачи стабилизации по начальным данным при $b = 7$, $i_0 = 2$, $u(0, x) = \sin x + \sin(2x)$. Отметим, что в этом случае $u(0, x) \in H_+$. Требуется, изменив $u(0, x)$ только на полуинтервале $[0, \pi/2[$, обеспечить выполнение $u(0, x) + l(x) \in H_-$. Таким образом, подпространство допустимых смещений \mathcal{L} состоит из всевозможных функций, равных нулю на $[\pi/2, \pi]$. Если выбрать базис в виде

$$\mathcal{L} = \left\langle l_i(x) = \begin{cases} \sin(ix), & x \in [0, \pi/2[\\ 0, & x \in [\pi/2, \pi] \end{cases}, i = 1, 2 \right\rangle,$$

то найденная поправка l будет иметь минимальную длину в подпространстве $L_2[0, \pi]$, однако при этом функция $u(0) + l$ будет разрывной. Для построения непрерывного и в некотором смысле оптимального продолжения можно выбрать базис в виде

$$\mathcal{L} = \left\langle l_i(x) = \begin{cases} \Delta_0^{-1} \sin(ix), & x \in [0, \pi/2[\\ 0, & x \in [\pi/2, \pi] \end{cases}, i = 1, 2 \right\rangle.$$

Здесь $\Delta_0^{-1} \sin(ix) = l_i$ означает $l_i'' = \sin(ix)$, $x \in]0, \pi/2[$, $l_i(0) = l_i(\pi/2) = 0$.

Если базис фиксирован, то для нахождения коэффициентов имеем систему

$$\begin{cases} (u(0) + c_1 l_1(x) + c_2 l_2(x), \sin x) = 0, \\ (u(0) + c_1 l_1(x) + c_2 l_2(x), \sin(2x)) = 0. \end{cases}$$

Отметим, что данный подход в том числе позволяет решить задачу стабилизации по краевым условиям для рассмотренного уравнения (1). Действительно, пусть требуется найти краевые условия, обеспечивающие стремление с течением времени решения задачи (1) к нулю. Рассмотрим вспомогательную задачу типа (1), но в области $\tilde{\omega} = [-\pi, 2\pi]$. При этом для расширенной задачи начальное условие $\tilde{a}(x)$ положим равным $a(x)$ в исходной области ω и нулевым в подобласти расширения $\hat{\omega} = \tilde{\omega} / \omega$. Решим

для полученной таким образом неустойчивой системы задачу стабилизации по начальным данным для $\tilde{a}(x)$, изменив функцию только в $\hat{\omega}$, а затем проинтегрируем по времени полученную задачу и найдем $\tilde{y}(t, x)$. Тогда след функции $\tilde{y}(t, x)$ при $x = 0, \pi$ задает искомые стабилизирующие граничные управления для исходной задачи.

Нелинейные отображения

Рассмотрим качественную картину динамики в окрестности неподвижной точки нелинейного отображения $S(\cdot)$. Далее будем считать, что $z_0 = 0$. Если же $z_0 \neq 0$, то можно рассмотреть оператор $\mathcal{S}(u) := S(z_0 + u) - z_0$ с нулевой неподвижной точкой, для него провести все последующие рассуждения, а затем переформулировать их для исходного оператора $S(u)$. Будем считать, что S — достаточно гладкий оператор и можно построить линеаризацию $S(u) = Lu + R(u)$ в малой окрестности \mathcal{O} нулевой точки. Пусть L ограничен, существует разложение H на устойчивое и неустойчивое подпространства и выполнены следующие условия частичной гиперболичности:

- a₁) $P_+ + P_- = I, \|P_+\| \leq C_+, \|P_-\| \leq C_-, H = P_+[H] \oplus P_-[H];$
- a₂) $L(P_+[H]) = P_+[H], L(P_-[H]) \subset P_-[H];$
- a₃) $\|Lv\| \geq \mu_+ \|v\|, \forall v \in P_+H, \|Lw\| \leq \mu_- \|w\|, \forall w \in P_-[H],$
 $\mu_- < \mu < \mu_+, \mu = 1;$
- a₄) $\|P_\pm[R(u_1) - R(u_2)]\| < \theta_\pm (\max\{\|u_1\|, \|u_2\|\}) \|u_1 - u_2\|, \forall u_i \in \mathcal{O}.$

Здесь C_\pm, μ, μ_\pm — константы, $\theta(\cdot), \theta(0) = 0$ — непрерывная положительная неубывающая функция, \mathcal{O} — некоторая окрестность нуля, для которой ведутся все последующие рассуждения, а операторы P_+ и P_- являются операторами проектирования вдоль подпространств $H_- = P_-[H]$ и $H_+ = P_+[H]$ на подпространства H_+ и H_- соответственно. В данных неравенствах v, w, u_i — произвольные элементы $\mathcal{O} \subset H$. В данном случае в условиях «a₃», «a₄» имеется спектральный зазор для нормы оператора L (т. е. $\mu_+ > \mu_-$) и точка 0 называется гиперболической [1].

Обозначим через $\mathcal{W}^-(S, \mathcal{O})$ устойчивое многообразие, так называемый входящий ус Адамара, подмножества \mathcal{O} :

$$\mathcal{W}^-(S, \mathcal{O}) = \{m^0 \in \mathcal{O} : \exists m^{i+1} \in \mathcal{O}, m^{i+1} = S(m^i), i = 0, 1, 2, \dots\},$$

а через $\mathcal{W}^+(S, \mathcal{O})$ неустойчивое многообразие, или «исходящий ус Адамара»:

$$\mathcal{W}^+(S, \mathcal{O}) = \{m^0 \in \mathcal{O} : \exists m^{i+1} \in \mathcal{O}, m^i = S(m^{i+1}), i = 0, 1, 2, \dots\}.$$

В окрестности \mathcal{O} неподвижной гиперболической точки многообразия \mathcal{W}^\pm существуют и определяют качественное поведение траекторий. В том числе известно [1, 6, 7], что

- а) многообразия \mathcal{W}^\pm локально инвариантны относительно S ;
- б) если $m^- \in \mathcal{W}^-$ и $i \geq 0$, тогда $S^i(m^-) \in \mathcal{O}$;
- в) под действием оператора S каждая траектория притягивается к \mathcal{W}^+ ;
- г) если $m^+ \in \mathcal{W}^+$, тогда для любого $i \geq 0$ существует $S^{-i}(m^+) \in \mathcal{O}$;
- д) если $u \notin \mathcal{W}^-$, тогда найдется $i = i(u) > 0$, такое, что $S^i(u) \notin \mathcal{O}$.

Соответствующий результат в гиперболическом случае принято называть теоремой Адамара — Перрона.

По сути, устойчивое многообразие представляет собой множество точек m^0 , всегда остающихся в \mathcal{O} при действии оператора S . Неустойчивое многообразие состоит из точек m^0 , для которых произвольная степень i оператора S обратима и $S^{-i}(m^0)$ принадлежит множеству \mathcal{O} . При этом если окрестность достаточно мала, то можно найти такое отображение $f: P_-[\mathcal{O}_0] \rightarrow P_+[\mathcal{O}_0]$, удовлетворяющее условиям $f(0) = 0$, $\|f(w_1) - f(w_2)\| \leq \gamma \|w_1 - w_2\|$, что

$$\mathcal{W}^-(\mathcal{O}) = \{w + f(w), w \in P_-[\mathcal{O}_0]\}.$$

Это означает, что для каждого элемента $w \in H_-$ однозначно вычисляется такой элемент $f(w) \in H_+$, что $(w + f(w)) \in \mathcal{W}^-$. Аналогично, существует отображение $g: P_+[\mathcal{O}_0] \rightarrow P_-[\mathcal{O}_0]$, удовлетворяющее условиям $g(0) = 0$, $\|g(v_1) - g(v_2)\| \leq \gamma \|v_1 - v_2\|$, такое, что

$$\mathcal{W}^+(\mathcal{O}) = \{v + g(v), v \in P_+[\mathcal{O}_0]\}.$$

Это означает, что для каждого элемента $v \in H_+$ однозначно вычисляется такой элемент $g(v) \in H_-$, что $(v + g(v)) \in \mathcal{W}^+$. Таким образом, подмножества \mathcal{W}^\pm для нелинейного оператора $S(u) = Lu + R(u)$ являются нелинейными, искривленными посредством функций f и g аналогами подпространств H_\pm для линейного отображения Lu .

Далее нас будет интересовать устойчивое многообразие \mathcal{W}^- (хотя некоторый класс задач стабилизации решается в терминах \mathcal{W}^+).

Известно, что отображение f может быть найдено из условия инвариантности: если $m = w + f(w)$, т. е. $m \in \mathcal{W}^-$, то для ее образа имеем $S(m) \in \mathcal{W}^-$, т. е. $P_+[S(m)] = f(P_-[S(m)])$. Отсюда получаем уравнение для функции f :

$$P_+[L(w + f(w)) + R(w + f(w))] = f(P_-[L(w + f(w)) + R(w + f(w))]),$$

то есть

$$Lf(w) + P_+R(w + f(w)) = f(Lw + P_-R(w + f(w))).$$

Второе тождество следует из условий инвариантности «а_{2,3}» подпространств H_{\pm} относительно L .

Пример 2. Рассмотрим отображение
$$S(u) = \begin{cases} \frac{1}{4} w_2, \\ \frac{1}{2} w_1, \\ 2v - \frac{7}{4} w_1^2 - \frac{127}{64} w_2^3, \end{cases}$$

в пространстве $u = (w_1, w_2, v) \in R^3$. Не сложно проверить, что точки вида $(w_1, w_2, v = w_1^2 + w_2^3)$ переходят в точки $(\frac{1}{2} w_1, \frac{1}{4} w_2, v = (\frac{1}{2} w_1)^2 + (\frac{1}{4} w_2)^3)$, т. е. удовлетворяют условию инвариантности. Таким образом, для указанного отображения множество точек $\mathcal{W}^- = (w_1, w_2, w_1^2 + w_2^3)$ является инвариантным и стремится к нулю с экспоненциальной скоростью: $u_n = S^n(u) = ((\frac{1}{2})^n w_1, (\frac{1}{4})^n w_2, (\frac{1}{2})^{2n} w_1^2 + (\frac{1}{4})^{3n} w_2^3)$.

В общем случае при решении нелинейных уравнений в частных производных отображение f удается построить только приближенно с помощью итерационных алгоритмов (см. [5, 7]), основанных на следующем утверждении: в малой окрестности \mathcal{O} имеем

$$a + l \in \mathcal{W}^- \Leftrightarrow \lim_{n \rightarrow \infty} P_+[S^n(a + l)] = 0.$$

Отсюда следует, что, выбрав достаточно большое n и решив уравнение

$$P_+[S^n(a + \sum_{i=1}^{i_0} c_i e_i)] = 0$$

относительно неизвестных коэффициентов c_i , мы получим приближенную проекцию a вдоль $\mathcal{L} = \text{span} \langle e_i, i = 1, \dots, i_0 \rangle$ на \mathcal{W}^- . При $n = 0$ имеем систему $P_+[a + \sum_{i=1}^{i_0} c_i e_i] = 0$, т. е. строим проекцию на H_- . Поскольку \mathcal{W}^- касается H_- , то часто такая проекция имеет достаточно высокую точность.

Пример 3. В качестве примера можно рассмотреть задачу построения устойчивого многообразия для уравнения Чафе — Инфанта:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 7u - u^3, \quad u = u(t, x),$$

$$t \geq 0, \quad x \in \omega = [0, \pi],$$

$$u(t, 0) = u(t, \pi) = 0, \quad u(0, x) = \sum_{k=1}^{\infty} c_k \sin(kx).$$

В данном случае нелинейная добавка $(-u^3)$ разрушает инвариантность подпространства $w \in H_- = \text{span} \langle \sin(\pi kx), k = 3, 4, \dots \rangle$ и элемент устойчивого многообразия \mathcal{W}^- имеет вид $w + v$, где $v = f(w) \in \text{span} \langle \sin(\pi kx), k = 1, 2 \rangle$. Таким образом, для произвольного $w \in H_-$, достаточно малого по норме пространства $L_2[0, \pi]$, можно найти такие коэффициенты c_1, c_2 , что решение $u(t, x)$ задачи Чафе — Инфанта с начальным условием $w(x) + c_1 \sin x + c_2 \sin(2x)$ стремится к нулю в норме $L_2[0, \pi]$ при $t \rightarrow \infty$. Так как соответствующие коэффициенты могут быть найдены только приближенно, то и убывание нормы $u(t)$ мы получим только на некотором конечном временном отрезке $[0, T]$. Поэтому, если мы хотим обеспечить дальнейшее убывание решения, для полученной функции $u(T)$ потребуется заново решить задачу перепроектирования на \mathcal{W}^- . Напомним, что в рамках данного подхода удается решить задачу стабилизации нелинейной системы по краевым условиям.

Как было сказано выше, в некоторых случаях для решения нелинейной задачи стабилизации достаточно осуществлять проектирование не на устойчивое многообразие \mathcal{W}^- , а только на H_- , так как даже такая проекция имеет достаточно высокую точность. В качестве примера рассмотрим задачу о стабилизации течения электропроводящей жидкости в кольцевой области на основе двумерных уравнений Навье — Стокса.

Пример 4. Экспериментальная установка, приближенное моделирование процессов которой мы рассматриваем, представляет собой горизонтально расположенную круглую ванну из органического стекла с двумя коаксиальными цилиндрическими электродами, между которыми находится раствор медного купороса. Дно ванны соприкасается с электромагнитом, создающим вертикальное рабочее электромагнитное поле с почти синусоидальным по радиусу и не зависящим от угла профилем. В результате аналитического исследования предсказано (см. [8]) и экспериментально получено (см. [4]), что при малых значениях протекаемого тока J в системе имеется локально устойчивое ламинарное двухпотокное течение $u_z^{(0)}$, такое, что ближняя к внешней границе часть жидкости вращается против часовой, а ближняя к внутренней границе — по часовой стрелке

(рис. 1, *a*). При увеличении модуля J такой квазистационарный (в смысле функции тока) режим становится неустойчивым с неустойчивостью седлового типа. Это приводит к появлению трех вращающихся против часовой стрелки вихрей $u_z^{(3)}$ с центрами в вершинах также вращающегося правильного треугольника (рис. 1, *б*). В работе [8] аналитически было показано, что помимо трехвихревой структуры для кольца с отношением радиусов порядка $1/3$ при больших числах Рейнольдса теоретически могут существовать неустойчивые структуры с волновыми числами до шести. Подобные режимы также наблюдались в некоторых экспериментах (см. [2, 3]).

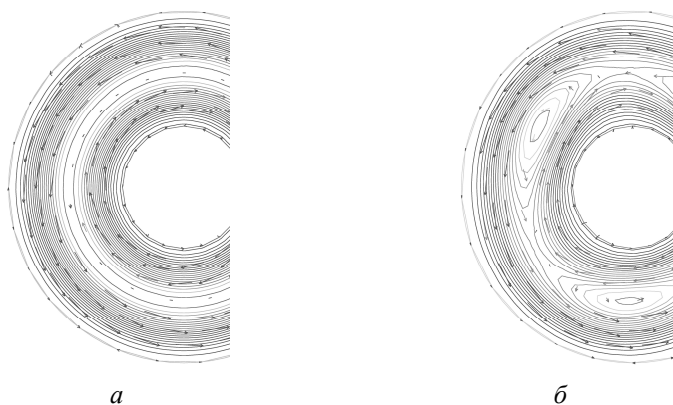


Рис. 1. Двухпоточковый режим $u_z^{(0)}$ (*a*) и трехвихревой режим $u_z^{(3)}$ (*б*)

Для численного моделирования данного эффекта и решения задачи стабилизации ламинарного двухпоточкового течения рассмотрим систему уравнений Навье — Стокса в переменных «скорость — давление»:

$$\begin{cases} u_t - \nu \Delta u + u \cdot \nabla u + \nabla p = f, & (x, y) \in \Omega, \\ \operatorname{div} u = 0, & u = (v, w), \quad f = (f_1, f_2), \end{cases} \quad (2)$$

с начально-краевыми условиями

$$\begin{aligned} u|_{\partial\Omega} &= 0, \quad u|_{t=0} = u(0, x, y), \\ \Omega &= \{(x, y) : r_1 \leq x^2 + y^2 \leq r_2\}, \quad r_1 = 0,3, \quad r_2 = 1, \end{aligned} \quad (3)$$

правой частью вида $f(r, \phi) = -0,345 \sin\left(2\pi \frac{r-r_1}{r_2-r_1}\right)$ и вязкостью $\nu = 170^{-1}$.

Функция $f(r, \phi)$ выбрана с учетом рекомендаций работ [4, 8], коэффициент ν найден в результате численных экспериментов. В какой степени по-

лученные далее результаты согласуются с течением реальной жидкости, можно, по-видимому, проверить только экспериментально.

Численное решение рассматриваемого уравнения (2), (3) в данной работе строилось в рамках пакета FreeFEM++ методом конечных элементов. Численно проверено, что для указанных значений параметров уравнения качественная картина динамики выглядит следующим образом: решение монотонно приближается к изображенному на рис. 1 течению $u_z^{(0)}$ и некоторое время находится в его окрестности, а далее за время 30—70 единиц постепенно формируется устойчивое трехвихревое течение $u_z^{(3)}$. Таким образом, система (2), (3) качественно правильно воспроизводит структуру течений реальной жидкости. Рассмотрим задачу стабилизации в окрестности неустойчивого течения $u_z^{(0)}$ по краевым условиям: для указанных параметров уравнения и фиксированной начальной функции $u_a|_{r=0}$, достаточно близкой к $u_z^{(0)}$, требуется найти такую граничную функцию $u_a(t)|_{r=r_1}$, определенную на внутренней границе области Ω , что полученное решение $u_a(t)$ уравнения (2) с найденным краевым условием на внутренней границе и нулевым на внешней остается в малой окрестности течения $u_z^{(0)}$ при всех $t \geq 0$, т. е. $u_a(t)$ имеет стабильную двухпотоковую структуру.

Решение сформулированной задачи стабилизации по краевым условиям будем строить в терминах проектирования на устойчивое инвариантное подпространство H . Это означает, что краевое условие строится для линеаризованной задачи, а найденное в результате граничное управление применяется для исходной нелинейной системы. Основу данного подхода составляют конструктивные теоретические результаты А. В. Фурсикова, говорящие о возможности стабилизации по краевым условиям в окрестности стационарного течения для уравнений Навье — Стокса (см. [9, 10]), конструктивные теоремы о существовании инвариантных многообразий О. А. Ладыженской и В. А. Солонникова для уравнений типа Навье — Стокса (см. [6]), Д. В. Аносова и Я. М. Песина для конечномерных динамических систем (см. [1, 7] и цитируемую в этих работах литературу).

Приведем формальные шаги алгоритма.

Шаг 1. Приближенно найдем основное квазистационарное решение $u_z^{(0)}$ для уравнения (2), (3).

Шаг 2. Сформулируем вспомогательную расширенную математическую задачу

$$\begin{cases} \tilde{u}_t - \tilde{\nu} \Delta \tilde{u} + \tilde{u} \cdot \nabla \tilde{u} + \nabla \tilde{p} = \tilde{f}, & (x, y) \in \tilde{\Omega}, \\ \operatorname{div} \tilde{u} = 0, \quad \tilde{u} = (\tilde{v}, \tilde{w}), \quad \tilde{f} = (\tilde{f}_1, \tilde{f}_2), \end{cases} \quad (4)$$

с начально-краевыми условиями

$$\begin{aligned} \tilde{u}|_{\partial\tilde{\Omega}} &= 0, \quad \tilde{u}|_{t=0} = \tilde{u}^0(x, y), \\ \tilde{\Omega} &= \{r_0 \leq x^2 + y^2 \leq r_2\} = \Omega \cup \bar{\Omega}, \quad 0 < r_0 < r_1. \end{aligned} \quad (5)$$

При этом выберем параметры расширения так, чтобы исходная задача (2), (3) и построенная расширенная задача (4), (5) совпадали в области Ω , т. е. положим $\tilde{v} = v$, $\tilde{f}_{1,2} \equiv f_{1,2}$ для $(x, y) \in \Omega$. В подобласти расширения $\bar{\Omega} = \tilde{\Omega} \setminus \Omega$ значения указанных параметров определим так, чтобы основное квазистационарное течение $\tilde{u}_z^{(0)}$ для расширенной задачи было близко к решению u_z при $(x, y) \in \Omega$. По сути, это соответствует погружению исходной кольцеобразной области в среду с другими характеристиками, но с приближенным сохранением интересующего нас квазистационарного состояния. Отметим, что взаимодействие между средами чисто математическое, формально описываемое задачей (4), (5).

Найдем основное течение $\tilde{u}_z^{(0)}$ для задачи (4), (5) и убедимся, что $\|\tilde{u}_z^{(0)} - u_z^{(0)}\|_{\Omega} \ll 1$. Это позволит заменить исходную задачу стабилизации в окрестности $u_z^{(0)}$ в области Ω на задачу стабилизации в окрестности $\tilde{u}_z^{(0)}$ в области $\tilde{\Omega}$.

Шаг 3. Пусть оператор L соответствует линеаризации стационарного уравнения Навье — Стокса (4), (5) в области $\tilde{\Omega}$ в окрестности решения $\tilde{u}_z^{(0)}$, а L^* — сопряженный к нему оператор. Вычислим для операторов L и L^* базисы в подпространствах $H_+ = \text{span} \langle \xi_i, i = 1, \dots, i_0 \rangle$ и $H_-^\perp = \text{span} \langle \eta_i, i = 1, \dots, i_0 \rangle$, отвечающие собственным числам с неотрицательной действительной частью. Если $\tilde{u}_z^{(0)}|_{\Omega} \equiv u_z^{(0)}$, то в силу неустойчивости исходного течения $u_z^{(0)}$ такие числа заведомо существуют и всегда $0 < i_0 < \infty$. Отметим, что все пространство H можно представить в виде прямой суммы $H = H_+ \oplus H_-$, где H_- соответствует корневым подпространствам оставшейся части спектра с отрицательной действительной частью. При этом $L(H_+) = H_+$, $L(H_-) \subset H_-$, т. е. указанные подпространства инвариантны относительно оператора L , и $H_-^\perp \perp H_-$. Далее будем считать, что соответствующие вектор-функции $\{\xi_i, \eta_i, i = 1, \dots, i_0\}$ найдены и линейно независимы.

Построим так называемое подпространство допустимых смещений $\mathcal{L} = \text{span} \langle e_i, i = 1, \dots, i_0 \rangle$ вида

$$e_i(x, y) = \begin{cases} 0, & (x, y) \in \Omega, \\ \bar{e}_i, & (x, y) \in \bar{\Omega}, \end{cases}$$

где функция $\bar{e}_i(x, y)$ является решением задачи Стокса в области $\bar{\Omega}$ с нулевыми краевыми условиями:

$$\begin{cases} -\Delta \bar{e}_i + \nabla q = \xi_i, \\ \operatorname{div} \bar{e}_i = 0, \\ \bar{e}_i|_{\partial\bar{\Omega}} = 0, \quad \bar{\Omega} = \{r_0 \leq x^2 + y^2 \leq r_1\}. \end{cases} \quad (6)$$

Такой выбор обеспечивает корректность и в некотором смысле оптимальность изменения поля скоростей при проектировании на многообразие $\mathcal{W}^-(\mathcal{O})$.

Шаг 4. Для стабилизации по границе решения задачи (2), (3) с начальным условием u_a продолжим функцию u_a из области Ω в область $\tilde{\Omega}$ до функции \tilde{u}_a так, что $\operatorname{dist}(\tilde{u}_a, \mathcal{W}^-(\mathcal{O}(\tilde{u}_z^{(0)}))) \ll 1$. Процесс продолжения состоит из двух подшагов.

4.1. Приблизительно продолжим u_a в область $\tilde{\Omega}$ до бездивергентной функции \bar{u}_a .

4.2. Для найденной функции \bar{u}_a методом нулевого приближения (см. [11]) определим приближенную проекцию на устойчивое подпространство H_- в виде $\tilde{u}_a = \bar{u}_a + \sum_{i=1}^{i_0} c_i e_i$, где коэффициенты c_1, \dots, c_{i_0} вычислим из условий $(\tilde{u}_a - \tilde{u}_z, \eta_j) = 0, j = 1, \dots, i_0$. В этом случае в разложении функции $\tilde{u}_a - \tilde{u}_z$ по базису корневых подпространств H_+ и H_- не содержатся компоненты подпространства H_+ , поэтому $(\tilde{u}_a^{(0)} - \tilde{u}_z^{(0)}) \in H_-$.

Найденное таким образом продолжение \tilde{u}_a является начальным условием для нелинейной задачи (4), (5). И если \tilde{u}_a принадлежит малой окрестности $\mathcal{W}^-(\mathcal{O})$, то норма разности $\tilde{u}_a(t) - \tilde{u}_z^{(0)}$ убывает для $0 \leq t \leq T_0$. Поэтому след $\tilde{u}_a(t)|_{y=0, y=l_2}$ дает искомое стабилизирующее управление $\theta^{(1,2)}(t)$ для исходной задачи в области Ω .

Относительно линейризованной задачи подпространство H_- инвариантно, поэтому если $\|\tilde{u}_z^{(0)} - u_z\|_{\Omega} = 0$ и все промежуточные задачи решаются абсолютно точно, то в нестационарном линейном случае (т. е. для задач Озеена в расширенной и исходной областях) время T_0 можно выбрать сколь угодно большим. Для нелинейных задач в разложении разности ре-

шений $\tilde{u}_a(t) - \tilde{u}_z^{(0)}$ в некоторый момент времени $t = T_1$ появятся значительные компоненты из подпространства H_+ , поэтому процесс стабилизации будет нарушен. Следовательно, шаг 4 необходимо циклически повторять для начальных функций $u_a|_{t=T_1}$, $u_a|_{t=2T_1}$ и т. д.

Расчеты проводились на квазиравномерной триангуляции области при $\tau, h \sim 10^{-2}$. Приближение к квазистационарному решению $u_z^{(0)}$ было найдено в результате решения стационарной линеаризованной системы (2). Отметим, что в некоторых случаях для приближенного вычисления $u_z^{(0)}$ лучше решать исходную нелинейную стационарную систему (2), например методом Ньютона. Численно получено, что в окрестности $u_z^{(0)}$ для исходной задачи имеются вихревые собственные функции с волновыми числами $n_v = 1, 2, 3, 4, 5, \dots$ и каждой моде соответствует пара комплексно-сопряженных собственных чисел, а также ламинарное двухпотокное течение. Наиболее неустойчивыми являются моды $n_v = 4, 3, 5$. Отметим, что структура собственной функции, например для $n_v = 3$, близка к структуре функции ξ_1 (рис. 2), а четырехвихревая мода близка к ξ_3 (рис. 3). При этом все численно найденные собственные числа имеют близкую к нулю, но отрицательную действительную часть, хотя решение нелинейной задачи с $u(0) = u_z^{(0)}$ также выходит за время порядка 30—40 единиц на трехвихревой режим. Данная несогласованность объясняется, на наш взгляд, тем, что линеаризация строится для функции, не являющейся стационаром.

В подобласти расширения $\bar{\Omega} = \tilde{\Omega} \setminus \Omega$ положим $\tilde{v} = v \cdot \left(50 \frac{r-r_1}{r_0-r_1} + \frac{r-r_0}{r_1-r_0} \right)$, $\tilde{f}_{1,2} \equiv 0$. Приближение к квазистационарному решению $\tilde{u}_z^{(0)}$ было найдено в результате решения расширенной линеаризованной стационарной задачи (4). Так как функция $\tilde{u}_z^{(0)}$ принадлежит малой окрестности $u_z^{(0)}$ в области Ω , следовательно процесс стабилизации в области $\tilde{\Omega}$ можно направлять к $\tilde{u}_z^{(0)}$, что обеспечит в случае успеха стабилизацию в окрестности $u_z^{(0)}$ в области Ω .

Выделим линейную часть L для оператора правой части расширенной системы Навье — Стокса в окрестности функции $\tilde{u}_z^{(0)}$ и вычислим для оператора L базис в H_+ и H_-^\perp . Численно получено, что в окрестности $\tilde{u}_z^{(0)}$ имеются вихревые собственные функции с волновыми числами $n_v = 1, 2, 3, 4, 5, \dots$, каждой моде соответствует пара комплексно-сопряженных собственных чисел, а также ламинарное двухпотокное течение (так как $\tilde{u}_z^{(0)}$ не

является стационаром). При этом $\lambda(n_v = 4) \approx 0,06 \pm 0,4i$, $\lambda(n_v = 5) \approx 0,04 \pm 0,4i$, $\lambda(n_v = 3) \approx -0,03 \pm 0,5i$. Отметим, что решение нелинейной расширенной задачи с $\tilde{u}(0) = \tilde{u}_z^{(0)}$ также выходит на трехвихревой режим. Соответствующие функции $\{\xi_1, \eta_1\}$ для $n_v = 3$ изображены на рис. 3. При этом функции ξ_2, η_2 имеют схожий вид с точностью до поворота на угол порядка $\pi/6$.

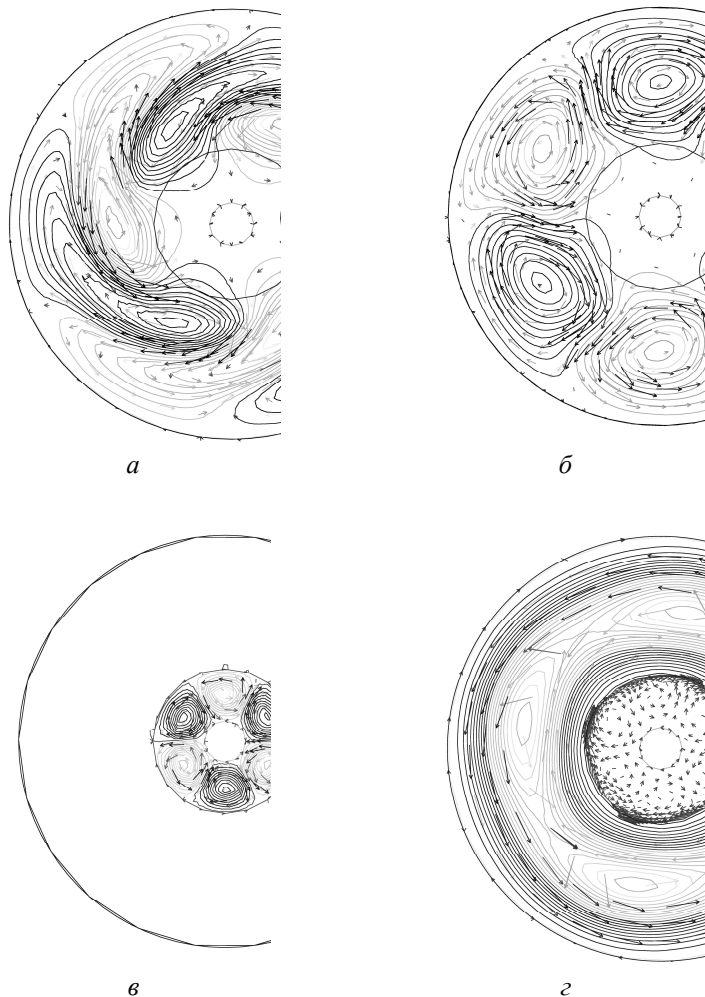


Рис. 2. Базисные функции ξ_1 (а), η_1 (б), e_1 (в) и типичная функция $\tilde{a}^{(4)}$ (г)

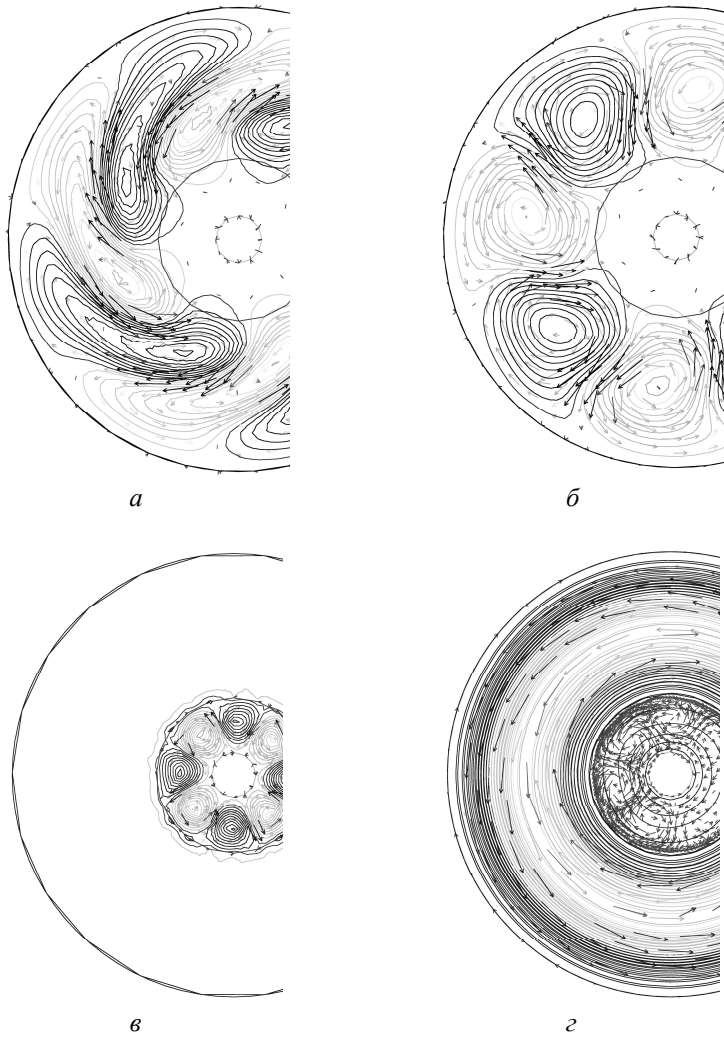


Рис. 3. Базисные функции ξ_3 (а), η_3 (б), e_3 (в) и типичная функция $\tilde{a}^{(0)}$ (г)

Процесс проектирования (т. е. четвертый пункт алгоритма) будем повторять регулярно (с некоторой частотой по времени). Если краевые условия равны нулю, т. е. стабилизация отсутствует, то решение стремится к вращающейся системе из трех вихрей $u_z^{(3)}$ (см. рис. 1). Если процесс стабилизации включить с момента времени $T = 10$ для решения задачи (2) при

нулевой начальной скорости и нулевых граничных условиях, а в процессе стабилизации учитывать только пару собственных функций для $n_v = 3$, то решение стремится к течению из четырех вращающихся вихрей $\tilde{a}^{(4)}$ (см. рис. 2). Таким образом, частичная стабилизация в окрестности $u_z^{(0)}$ в смысле блокирования только трехвихревой моды $n_v = 3$ обеспечивает стабилизацию в окрестности вторичного четырехвихревого течения.

Если в процессе стабилизации учитывать четыре собственные функции, отвечающие $n_v = 3, 4$, то решение стремится к двухпотокковому течению, т. е. обеспечивается процесс стабилизации в окрестности $u_z^{(0)}$. Типичные структуры $\tilde{a}^{(0)}$, образующиеся в этом случае для расширенной задачи, изображены на рис. 3. Модули найденных коэффициентов в разложении поправки по векторам $\{e_i(x)\}$ имеют порядок 0,03—0,003.

Еще раз отметим, что на всем отрезке стабилизации $t \geq 10$ реализуется управление с обратной связью: через равные промежутки времени $T_1 = 0,01$ осуществляется перепроектирование текущего течения в расширенной области на устойчивое подпространство H_- . Поэтому формально стабилизирующие краевые условия могут существенно зависеть от времени. Однако для данной задачи управление выходит на асимптотически постоянный режим типа мелких пульсаций.

Заключение

Многочисленные расчеты по решению задач стабилизации, проведенные для различных уравнений, показали высокую робастность данного подхода. Это связано с тем, что во многих случаях изложенный алгоритм устойчив относительно ошибок, возникающих при решении вспомогательных задач типа нахождения неподвижной точки u_z и вычисления базисов в подпространствах H_+ и H_-^\perp . Более того, для широкого класса уравнений математической физики \mathcal{W}^- существует в окрестности, имеющей в интегральной норме вид бесконечномерного эллипсоида, длины полуосей которого стремятся к бесконечности. Отметим также, что согласно структуре функций ξ_i, η_i , изображенных на рис. 2 и 3, основная неустойчивость сосредоточена в исходной области, так как вязкость в подобласти расширения почти на два порядка меньше. Однако в данном случае, на наш взгляд, процесс стабилизации оказался успешным вдоль базиса $\{e_i\}$ (рис. 2 и 3) благодаря общей глобальной структуре течений — в полученной расширенной задаче существуют либо двухпотокковое, либо трех-, либо четырехвихревое течения. Блокируя в процессе стабилизации воз-

никновение вихревых структур, мы обеспечиваем двухпотокное решение. Косвенно это подтверждается следующим расчетом: если указанный алгоритм стабилизации формально реализовать для $u(x)$, считая, что $\tilde{z}_0(x) \equiv 0$, т. е. подавлять возникающие трехвихревые структуры на фоне нулевого течения, то мы также выйдем на четырехвихревой режим. Если же подавлять в $u(x)$ возникающие на фоне нулевого течения трехвихревые и четырехвихревые структуры, то мы получим двухпотокное течение. Таким образом, нелинейность в данном случае играет ключевую, глобально стабилизирующую роль.

Работа выполнена при поддержке гранта РФФИ № 15-01-08023.

Литература

1. *Аносов, Д. В.* Многомерный аналог одной теоремы Адамара // Науч. докл. высш. шк. Физ.-мат. науки. — 1959. — № 1. — С. 3—12.
2. *Бондаренко, Н. Ф.* Лабораторная и теоретическая модели плоского периодического течения / Н. Ф. Бондаренко, М. З. Гак, Ф. В. Должанский // ФАО. — 1979. — Т. 15, № 10. — С. 1017-1026.
3. *Гак, Е. З.* Магнитные поля и водные электролиты / Е. З. Гак. — СПб. : Элмоп, 2013. — 535 с.
4. *Довженко, В. А.* Моделирование процесса генерации вихрей в аксиально-симметричном азимутальном поле магнитогидродинамическим методом / В. А. Довженко, Ю. В. Новиков, А. М. Обухов // ФАО. — 1979. — Т. 15, № 11. — С. 1199—1202.
5. *Корнев, А. А.* Классификация методов приближенного проектирования на устойчивое многообразие // Докл. РАН. — 2005. — Т. 400, № 6. — С. 736—738.
6. *Ладыженская, О. А.* О принципе линеаризации и инвариантных многообразиях для задач магнитной гидродинамики / О. А. Ладыженская, В. А. Солонников // Зап. научн. сем. ЛОМИ. — 1973. — Т. 38. — С. 46—93.
7. *Песин, Я. Б.* Характеристические показатели Ляпунова и гладкая эргодическая теория // Успехи матем. наук. — 1977. — Т. 32, № 4. — С. 55—112.
8. *Пономарев, В. М.* Об устойчивости одного класса осесимметричных течений несжимаемой жидкости // Механика жидкости и газа. — 1980. — № 1. — С. 3—9.
9. *Fursikov, A. V.* Feedback Stabilization for Navier–Stokes Equations : Theory and Calculations / A. V. Fursikov, A. A. Kornev // Mathematical Aspects of Fluid Mechanics. — Cambridge [etc.] : Cambridge University Press, 2012. — P. 130—172.
10. *Fursikov, A. V.* Stabilizability of Two-Dimensional Navier–Stokes Equations with Help of a Boundary Feedback Control // J. Math. Fluid Mech. — 2001. — V. 3. — P. 259—301.

Астрофизика и космология

РАДИАТИВНЫЕ УДАРНЫЕ ВОЛНЫ И ИХ РОЛЬ В ОБЪЯСНЕНИИ ЗАГАДКИ СВЕРХМОЩНЫХ СВЕРХНОВЫХ

С. И. Блинников

Институт теоретической и экспериментальной физики, г. Москва
Институт Кавли физики и математики Вселенной Токийского университета, г. Касива, Япония

1. Введение: три пути, то есть три сценария, предложенные для SLSNe

Типичные взрывы сверхновых (сокращенно SN) производят выбросы (ejecta) с кинетическими энергиями 10^{51} эрг \equiv 1 foe. Эта единица энергии была введена Гансом Бете (H. Bethe) и является сокращением от ten to fifty one ergs [1]. Свет, излучаемый в течение первого года сверхновой, составляет лишь небольшую часть этой энергии, около 0,01 foe. В последнее десятилетие обнаружены многие сверхновые с пиковой светимостью на один-два порядка выше, чем для нормальных сверхновых известных типов. Они излучают мощный свет с интегральной энергией, приближающейся к 1 foe, а иногда даже выше.

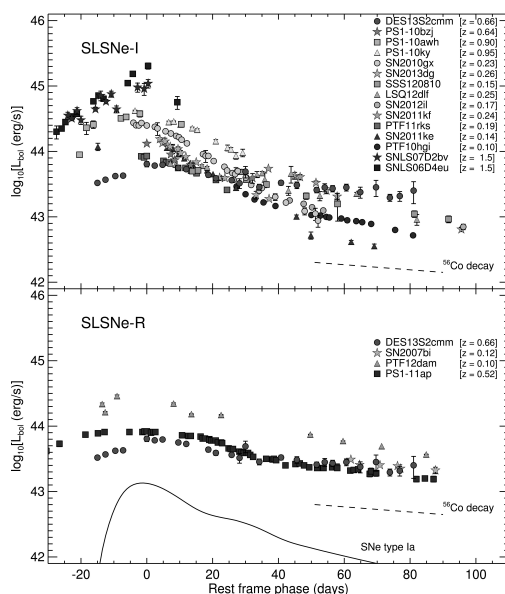


Рис. 1. БолOMETрические кривые блеска
некоторых сверхмощных сверхновых [2]

Пример приведен на рис. 1 вместе с типичной сверхновой SN Ia для сравнения (такие сверхновые используются в космологии из-за их высо-

кой светимости). Мы видим, что новые объекты излучают гораздо больше света. Их называют сверхмощными сверхновыми: SLSNe (Super-luminous Supernovae). Это вызов для теории, так как даже обычные сверхновые еще не полностью поняты из первых принципов.

В литературе по сверхмощным сверхновым обсуждается много моделей [3]. Среди них наиболее серьезными представляются следующие три сценария:

1. Неустойчивость при рождении пар (англ. Pair Instability Supernovae — PISN).

2. «Магнитарная» накачка. (Пишем в кавычках, так как наблюдаемые магнитары медленно вращаются. Их находят в повторяющихся вспышках мягких гамма-лучей, источниками которых являются сильно замагниченные нейтронные звезды с периодами вращения в несколько секунд, а тут нужны миллисекундные периоды.)

3. Радиативные ударные волны в окружающей плотной оболочке, которые могут рождаться в сверхновых, порожденных пульсационной неустойчивостью из-за рождения пар (англ. Pulsational Pair Instability Supernovae — PPISN).

2. Неустойчивость при рождении пар и сверхновые

Для того чтобы понять механизм взрыва сверхновых из-за неустойчивости при рождении пар, нужно немного изучить теорию звездной эволюции.

Рисунок 2 показывает эволюционные треки нормальных звезд с разными массами.

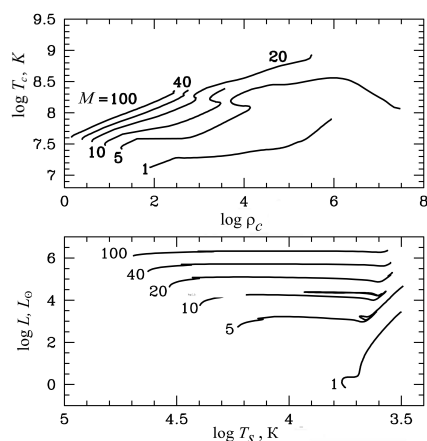


Рис. 2. Эволюция на диаграмме «центральная температура — центральная плотность» (вверху) и диаграмма Герцшпрунга — Рассела (внизу) для нескольких моделей нормальных звезд. Числа возле треков — это массы звезд в единицах массы Солнца M_{\odot}

Можно легко заметить тенденцию: центральная температура T_c растет с центральной плотностью ρ_c примерно как $T_c \propto \rho_c^{1/3}$. Этот закон более выражен для более массивных звезд. Такое соотношение легко понять из условий механического равновесия звезды.

Очень грубая оценка по порядку величины для силы притяжения двух половинок звезды с полной массой M дает

$$F \sim \frac{G_N M^2}{4R^2},$$

где R — радиус звезды, а G_N — постоянная Ньютона. Эта сила должна быть уравновешена градиентом давления P в звезде, которая находится в механическом равновесии.

Давление P практически равно нулю на поверхности, а в центре

$$P_c = \frac{F}{S} = \frac{F}{\pi R^2}.$$

Опуская все коэффициенты порядка единицы, получаем для давления и плотности в центре

$$P_c \simeq \frac{G_N M^2}{R^4},$$

$$\rho_c \simeq \frac{M}{R^3}$$

и находим, что равновесие требует (в ньютоновской гравитации)

$$P_c \simeq G_N M^{2/3} \rho_c^{4/3}.$$

Так что, если у нас есть классическая идеальная плазма с

$$P_c = \mathcal{R} \rho T / \mu,$$

где \mathcal{R} — это универсальная газовая постоянная, а μ — средняя молекулярная масса, то получаем для центральной температуры

$$T_c \simeq \frac{G_N M^{2/3} \rho_c^{1/3} \mu}{\mathcal{R}}.$$

Таким образом, $T_c \propto M^{2/3} \rho_c^{1/3}$ в невырожденных звездах (т. е. в звездах, где вклад эффектов вырождения в давление фермионов мал), и при фиксированной массе

$$T_c \propto \rho_c^{1/3}.$$

Та же степень $1/3$ получается для массивных звезд (массой около $100 M_\odot$ и выше) с преобладанием излучения в давлении (но тогда зависимость от массы слабей: $P_c \propto \rho_c^{1/3} M^{1/6}$).

Условие механического равновесия,

$$P_c \simeq G_N M^{2/3} \rho_c^{4/3},$$

говорит нам нечто важное о гидродинамической устойчивости звезды. Оно означает, что снижение показателя адиабаты до $\gamma < 4/3$ может привести к гидродинамической нестабильности. Действительно, если $\gamma < 4/3$, то глобальное сжатие звезды приводит к слишком слабому росту физического давления, ниже того, которое необходимо для равновесия. Тогда сила тяжести одерживает победу над давлением, и звезда начинает рушиться, т. е. коллапсировать.

Релятивистские частицы, т. е. фотоны, приводят к $\gamma \rightarrow 4/3$. В массивных звездах с $M \sim 100M_{\odot}$ мы имеем $\gamma \sim 4/3$ из-за высокой энтропии S (т. е. на каждый барион приходится много фотонов).

Массивная звезда, в давлении которой доминирует излучение, уже была на грани потери устойчивости, так как давление приближалось к соотношению $P \propto \rho^{4/3}$, а теперь звезда теряет устойчивость: γ становится меньше, чем $4/3$, поскольку создание (e^+e^-)-пар требует энергии. Импульсы частиц при сжатии не растут достаточно быстро: в парах создаются частицы, имеющие низкую скорость. Потеря механической устойчивости при $\rho_c \sim 10^4$ г/см³ приводит к началу гравитационного коллапса, который может быть остановлен при $\rho_c < 10^7$ г/см³ мощным взрывом, если звезда имеет достаточно ядерного горючего (в основном кислорода) в центре. Взрыв дает кинетическую энергию до 70 foe и достаточно света, чтобы объяснить некоторые из SLSNe, производящие огромные количества радиоактивного ⁵⁶Ni — до $\sim 20M_{\odot}$ (в то время как в стандартной термоядерной сверхновой масса радиоактивного материала в 20—50 раз меньше).

Эта модель, которая называется PISN, может объяснить некоторые из событий SLSN, но только медленные, из-за длительного времени диффузии фотонов в выбросах, перегруженных элементами железного пика (продуктами самого взрыва, а также распада ⁵⁶Ni).

Подобные медленные кривые блеска производятся и при накачке «магнитаром». В модели магнитара, в которой также нужно выделить огромную энергию в десятки foe, извлекая ее из энергии вращения нейтронной звезды, есть много факторов неопределенности. Я не обсуждаю этот сценарий здесь, а перейду прямо к самой экономичной модели, которая особенно хороша для быстрых SLSNe.

3. Модели с радиативными ударными волнами

Модели, объясняющие события SLSN с минимальным бюджетом энергии, основаны на сценарии, в котором перед основным взрывом в предсверхновой звезде произошло несколько мощных выбросов массы. Радиативные ударные волны, полученные при столкновении выброшенных слоев, могут обеспечить требуемую мощность света. Этот класс мо-

делей называют «взаимодействующими сверхновыми». Нелинейные эффекты играют важную роль на всех этапах этого явления: от начального выброса массы, потом при росте и вблизи максимума светимости и, наконец, на стадии остатка сверхновой с фрагментацией плотных оболочек.

Получим некоторые оценки для важных величин.

Светимость выражается через эффективную температуру T_{eff} и радиус фотосферы R_{ph} :

$$L = 4\pi\sigma T_{eff}^4 R_{ph}^2. \quad (1)$$

Для сверхновой с возрастом $t = 10$ сут и характерной скоростью на уровне фотосферы $u = 10^9$ см/с (т. е. 10000 км/с) получаем $R_{ph} = ut \approx 10^{15}$ см, и при типичной $T_{eff} \sim 10^4$ К светимость будет $L \sim 10^{43}$ эрг/с.

Светимость L спадает за несколько недель, таким образом, обычные, невзаимодействующие сверхновые производят порядка 10^{49} эрг = 0,01 foe в фотонах в течение первого года после взрыва, а порядка 10^{51} эрг = 1 foe остается в кинетической энергии выброса в «стандартных» взрывах SN.

Эта энергия излучается намного позже — в течение тысячелетий после взрыва (в основном в рентгеновских лучах) газовым остатком сверхновой в ударных волнах, образованных в обычной межзвездной среде с плотностью числа частиц порядка 1 см^{-3} . Если плотность околозвездного вещества (Circum-Stellar Matter — CSM) в миллиард раз выше, то большая часть кинетической энергии излучается гораздо быстрее — на временном масштабе порядка года, и фотоны будут гораздо мягче, чем рентген, они будут излучаться в основном в видимом или ультрафиолетовом диапазоне.

В таких условиях эффективная температура может быть такой же: $T_{eff} \sim 10^4$ К, в то время как $R_{ph} \sim 10^{16}$ см гораздо больше, а светимость нарастает, приближаясь к $L \sim 10^{45}$ эрг/с на некоторый период времени. Таким образом, сверхмощная сверхновая может получиться при энергии взрыва обычного масштаба $1 \text{ foe} \sim 10^{51}$ эрг, но теперь основная доля этой энергии теряется в течение первого года в виде фотонного излучения.

Если имеется сгусток вещества с массой m_1 и импульсом p_1 , его кинетическая энергия

$$E_1 = \frac{p_1^2}{2m_1}. \quad (2)$$

Если он сталкивается с другим сгустком — массой m_0 и с нулевым импульсом, мы получаем для конечной кинетической энергии двух слившихся сгустков в полностью неупругом столкновении

$$E_2 = \frac{p_1^2}{2(m_1 + m_0)}. \quad (3)$$

Импульс сохраняется, но энергия $E_1 - E_2$ высвечивается, поскольку $E_2 < E_1$. Если $m_0 \ll m_1$, то только совсем малая доля E_1 высвечивается, а при $m_0 \gg m_1$ имеем $E_2 \ll E_1$, и почти вся начальная E_1 излучается.

Это означает, что столкновения маломассивных и быстро движущихся выбросов с тяжелыми (плотными) медленно движущимися сгустками окружающей среды эффективно производят большой поток фотонов. Конечно, следует помнить, что в этом случае импульс двух слившихся сгустков может отличаться от начального p_1 , если порождается направленный поток фотонов, которые уносят некоторую долю полного импульса.

Нет особого смысла в детальном изучении этого эффекта с помощью порядковых оценок, так как детали производства фотонов могут быть сложными. Степень «неупругости» столкновения зависит от картины гидродинамического потока и от свойств излучения / поглощения плазмы, например от ее состава. Так или иначе, эти детали и сохранение импульсов и энергии должны быть приняты во внимание в полном радиационно-гидродинамическом моделировании.

Найдем теперь температуру за фронтом ударной волны. Опять же на уровне простых оценок для давления за фронтом ударной волны P_s у нас есть

$$P_s \sim \rho_0 D^2 = n_0 m_i D^2, \quad (4)$$

если плотность перед фронтом составляет ρ_0 , а D — скорость фронта. Плотность $\rho = nm_i$, если n — плотность числа частиц, а m_i — усредненная масса ионов. Оценка (4) вытекает из оценки для импульса: поток импульса равен $P + \rho u^2$ для течения со скоростью u , а когда давление P перед фронтом, где вещество холодное пренебрежимо мало, за фронтом получим (4). Более точное выражение для P_s легко выводится из законов сохранения.

Оценка (4) для нерелятивистской плазмы с уравнением состояния $P = nk_B T$ дает

$$k_B T_s \sim m_i D^2, \quad (5)$$

что указывает на очень высокие температуры в диапазоне килоэлектронвольт и выше для скоростей ударных волн больше, чем тысяча километров в секунду.

Теперь выведем «точные» коэффициенты в формуле (5).

Используя стандартные обозначения для плотности ρ , скорости u , давления P , термодинамической энергии E , определим компоненты вектора \vec{U} через плотность массы:

$$U_1 = \rho,$$

плотность импульса:

$$U_2 = \rho u,$$

плотность полной энергии:

$$U_3 = E + \frac{\rho u^2}{2}.$$

Определим соответствующие потоки массы:

$$F_1 = \rho u,$$

импульса:

$$F_2 = \rho u^2 + P,$$

и энергии:

$$F_3 = (E + \rho u^2/2 + P)u.$$

Выполняется общий закон сохранения:

$$\partial \bar{U} / \partial t = -\partial \bar{F} / \partial x.$$

В стационарном случае, т. е. $\partial \bar{U} / \partial t = 0$, получаем $\bar{F} = \text{const}$.

Введем $j \equiv \rho u = \text{const}$, $V \equiv 1/\rho$.

Из $F_2 = \rho u^2 + P = j^2 V = \text{const}$ получим

$$j^2 V_0 + P_0 = j^2 V_s + P_s \rightarrow P_s = P_0 + j^2 (V_0 - V_s).$$

Индекс «0» у ρ , V , u , P , E обозначает начальные значения вверх по течению (перед фронтом ударной волны), в то время как «s» соответствует значениям ниже по течению, в ударно-сжатом веществе.

Наиболее удобно работать в системе отсчета, где фронт находится в состоянии покоя, тогда скорость ударной волны D оказывается равна u_0 , потому что она по определению измеряется относительно вещества, еще не прошедшего фронт.

Далее, $F_3 = \text{const}$ дает

$$\left(E_0 + \frac{1}{2} j^2 V_0 + P_0 \right) u_0 = \left(E_s + \frac{1}{2} j^2 V_s + P_s \right) u_s.$$

Заменим $u_i \rightarrow j V_i$, тогда получим

$$\left(E_0 + \frac{1}{2} j^2 V_0 + P_0 \right) j V_0 = \left(E_s + \frac{1}{2} j^2 V_s + P_s \right) j V_s.$$

Отсюда

$$E_0 V_0 + \frac{1}{2} j^2 V_0^2 + P_0 V_0 = E_s V_s + \frac{1}{2} j^2 V_s^2 + P_s V_s$$

и

$$(E_0 + P_0) V_0 + \frac{1}{2} j^2 (V_0^2 - V_s^2) = (E_s + P_s) V_s.$$

Но $(V_0^2 - V_s^2) = (V_0 - V_s)(V_0 + V_s)$ и соотношение $P_s = P_0 + j^2 (V_0 - V_s)$, полученное выше, дают $V_0 - V_s = (P_s - P_0) / j^2$, так что j^2 сокращается в числителе и знаменателе:

$$(E_0 + P_0) V_0 + \frac{1}{2} \chi^2 \frac{(P_s - P_0)}{\chi^2} (V_0 + V_s) = (E_s + P_s) V_s.$$

Итак,

$$\left(E_0 + \frac{P_0 + P_s}{2}\right)V_0 = \left(E_s + \frac{P_0 + P_s}{2}\right)V_s,$$

и мы получаем общую формулу для сжатия вещества в течении (например, на ударном фронте):

$$\frac{V_s}{V_0} = \frac{2E_0 + P_0 + P_s}{2E_s + P_0 + P_s}.$$

Уравнение состояния $E = E(P, V)$ или $P = P(E, V)$ дает ударную адиабату. Для общего уравнения состояния в сильной ударной волне ($P_s \gg P_0, E_s \gg E_0$), что особенно важно в оболочках сверхновых,

$$\frac{V_s}{V_0} = \frac{2E_0 / (P_0 + P_s) + 1}{2E_s / (P_0 + P_s) + 1} \approx \frac{1}{2E_s / P_s + 1},$$

или

$$\frac{\rho_s}{\rho_0} = \frac{V_0}{V_s} \approx 1 + \frac{2E_s}{P_s}$$

в общем случае и

$$\frac{\rho_s}{\rho_0} = \frac{V_0}{V_s} \approx 1 + \frac{2}{\gamma - 1} = \frac{\gamma + 1}{\gamma - 1}$$

для случая уравнения состояния с $\gamma = \text{const}$.

Пусть $P = (\gamma - 1)E_{tr}$, где E_{tr} — трансляционная внутренняя энергия, т. е. кинетическая энергия частиц плазмы, и пусть $E = E_{tr} + Q$, где Q — внутренняя потенциальная энергия, например энергия ионизации. Тогда в сильной ударной волне

$$\frac{\rho_s}{\rho_0} = \frac{V_0}{V_s} \approx 1 + \frac{2E_{2tr} + 2Q}{P_s} = 1 + \frac{2}{\gamma - 1} + \frac{2Q}{(\gamma - 1)E_{2tr}},$$

то есть

$$\frac{\rho_s}{\rho_0} = \frac{V_0}{V_s} \approx \frac{\gamma + 1}{\gamma - 1} + \frac{2Q}{(\gamma - 1)E_{2tr}}.$$

При $\gamma = 5/3$ это дает

$$\frac{\rho_s}{\rho_0} = \frac{V_0}{V_s} \approx 4 + \frac{3Q}{E_{2tr}}$$

— формула (3.71) в знаменитой книге Я. Б. Зельдовича и Ю. П. Райзера по физике ударных волн [4].

Мы нашли из сохранения импульса ($F_2 = \text{const}$), что $P_s = P_0 + j^2(V_0 - V_s)$, то есть

$$j^2 = \frac{P_s - P_0}{V_0 - V_s} \approx \frac{P_s}{V_0 - V_s} = \frac{P_s}{V_0[1 - (\gamma - 1)/(\gamma + 1)]} = \frac{P_s(\gamma + 1)}{2V_0}.$$

Это верно для сильной ударной волны, постоянного γ и малого Q . Отсюда

$$\rho_0^2 u_0^2 = \frac{P_s(\gamma + 1)\rho_0}{2},$$

то есть

$$P_s = \frac{2}{\gamma + 1} \rho_0 u_0^2. \quad (6)$$

Заметим, что γ здесь надо брать для газа позади сильной ударной волны, поскольку давлением перед фронтом ударной волны P_0 можно пренебречь и уравнение состояния для него не играет роли.

Для нерелятивистской плазмы с давлением $P = \mathcal{R}\rho T/\mu$ получаем из (6)

$$\rho_0 u_0^2 = \frac{(\gamma + 1)\mathcal{R}\rho_s T_s}{2\mu},$$

откуда

$$u_0^2 = \frac{(\gamma + 1)\mathcal{R}\rho_s T_s}{2\rho_0\mu} = \frac{(\gamma + 1)^2 \mathcal{R} T_s}{2(\gamma - 1)\mu}.$$

Температура за фронтом сильной ударной волны T_s при постоянном γ и малом Q из последнего уравнения оказывается равной

$$T_s = \frac{2(\gamma - 1)u_0^2\mu}{(\gamma + 1)^2 \mathcal{R}}.$$

Для $\gamma = 5/3$ имеем

$$T_s = \frac{3u_0^2\mu}{16\mathcal{R}}. \quad (7)$$

Если подставим сюда $D_8 = u_0/10^8$ см/с, т. е. D_8 — скорость фронта в тысячах километров в секунду, то получим

$$T_s = 2,25 \cdot 10^7 \mu D_8^2 \quad (8)$$

в кельвинах, или

$$T_s = 1,94\mu D_8^2 \quad (9)$$

в килоэлектронвольтах. Здесь $\mu = A/(1 + Z)$ для плазмы (так как $n = n_{\text{baryon}}/\mu = n_{\text{ion}}A/\mu = n_{\text{ion}} + n_e = n_{\text{ion}} + Zn_{\text{ion}}$). Заметим, что типичное значение D в сверхновых около 10000 км/с, так что T получается из этой формулы порядка 10^9 К, т. е. сотни килоэлектронвольт.

Поскольку $\mathcal{R} \approx k_B / m_p$, где m_p — масса протона, имеем

$$k_B T_s \sim m_p D_s^2. \quad (10)$$

Эта оценка такая же, как в выражении (5), если там положить $m_i = m_p$.

Эти цифры, часто приводимые в статьях астрономов, совсем неверны в оболочках сверхновых, они только сбивают с толку и вводят в заблуждение!

В действительности плазма в условиях сверхновых, по меньшей мере частично, релятивистская: у нас есть огромное количество фотонов с $P = aT^4/3$, так что T_s заметно ниже из-за высокой теплоемкости фотонного газа. Приведенные ниже выражения (12) и (13) показывают, что с учетом излучения при D порядка тысячи километров в секунду и $\rho \sim 10^{-12}$ г/см³ мы имеем $T_s = 4,3 \cdot 10^4$ К, что значительно ниже рентгеновского диапазона температур, но достаточно высоко для поддержки высокой светимости L в течение длительного времени при больших радиусах R .

Во многих работах и моделированиях в астрофизике используется приближение $\gamma = \text{const}$, но в сверхновых это очень грубая идеализация, и часто она совсем не применима. На значение γ влияет ионизация / возбуждение атомов, и оно очень изменяется на фронте ударной волны, когда фронт проходит через холодные слои и нагревает плазму так сильно, что радиационное давление доминирует ниже по течению за фронтом. В этом случае, который является достаточно общим для периода выбегания ударной волны (supernova shock breakout), формулы (8) и (9) не применимы и вводят в заблуждение.

Уравнения для законов сохранения массы, импульса и энергии в радиационно-доминированных ударных волнах более сложны, потому что приходится учитывать передачу импульса и перенос энергии фотонов. Тем не менее есть два важных предельных случая для сильных ударных волн с излучением, когда могут быть получены простые выражения.

В первом случае мы можем иметь относительно холодный газ вверх по течению с $P_0 \ll P_s$ в сильной ударной волне, а ниже по течению газ может быть непрозрачен, причем в давлении доминирует излучение.

Благодаря высокой теплоемкости фотонного газа температура за фронтом оказывается на несколько порядков ниже, чем в формулах (8) и (9).

Подставим давление излучения как P_s в формулу (6), тогда получим

$$\frac{aT_s^4}{3} = \frac{2}{\gamma+1} \rho_0 u_0^2. \quad (11)$$

При преобладании вклада в давление фотонов $\gamma = 4/3$, тогда, подставив $u_0 = D$, получаем

$$T_s = \left(\frac{18}{7a} \rho_0 D^2 \right)^{1/4}. \quad (12)$$

То есть

$$T_s[K] = 4,3 \cdot 10^4 \rho_{-12}^{1/4} D_8^{1/2}, \quad (13)$$

если нормализовать плотность на $\rho = 10^{-12}$ г/см³ и взять D в тысячах километров в секунду. Итак, в реальности температура намного меньше, чем по формуле (8).

Второй важный случай имеет место, когда излучение не заперто за фронтом, его давлением и импульсом можно пренебречь, но оно очень эффективно при переносе тепла. Тогда энергия не сохраняется, а поток энергии F_3 больше не является постоянной величиной. Вместо этого может быть постоянной температура впереди и позади фронта. Сохранение массы и импульса дают, как и ранее,

$$P_s = P_0 + j^2(V_0 - V_s). \quad (14)$$

Теперь и перед, и за фронтом давление $P = \mathcal{R}\rho T / \mu$ при одинаковой температуре T , так что условие сильной ударной волны $P_s \gg P_0$ определяет не высокое значение T за фронтом, а то, что $\rho_s \gg \rho_0$; тогда $P_s \approx \rho_0 u_0^2$, которое мы извлекаем из (14), дает

$$\frac{\rho_s}{\rho_0} = \frac{\mu D^2}{\mathcal{R}T}. \quad (15)$$

Изотермическая температура T здесь намного меньше, чем температура, найденная в (8) и (9) для адиабатических ударных волн, следовательно степень сжатия в изотермических сильных ударных волнах может быть на несколько порядков больше, чем каноническое значение $(\gamma + 1)/(\gamma - 1)$ в адиабатических.

Это типичная ситуация для формирования холодных плотных оболочек взаимодействующих сверхновых. Точное значение T и степени сжатия зависит от деталей теплопроводных свойств плазмы, но следует помнить, что эти плотные оболочки могут стать неустойчивыми, поэтому точные цифры, найденные в идеализированных точных расчетах для плоскопараллельных или сферически-симметричных задач могут быть не очень полезны.

4. Численное моделирование кривых блеска

Кратко опишем некоторые результаты численного моделирования, которые учитывают эффекты запирания излучения в сверхновых, взаимодействующих с плотной окружающей средой. Для иллюстрации используются результаты из статьи [5].

Моделирование использует структуры предсверхновых, либо полученные из эволюционных расчетов, либо искусственно созданные. В каждой начальной модели есть быстро движущаяся часть, которую можно

назвать выбросом. Эта часть имеет массу M_{ej} и радиус R_{ej} . Значение M_{ej} может быть гораздо меньше, чем общая масса коллапсирующего ядра; это просто удобная форма параметризации моделей.

Для того чтобы создать взаимодействующую модель, выброс окружен довольно плотной оболочкой, «ветром», с массой M_w , которая распределена внутри радиуса R_w . Внешний радиус этой оболочки должен быть порядка 10^{16} см или даже больше в экстремальных случаях. Оболочка может иметь распределение плотности по степенному закону $\rho \propto r^{-p}$, который имитирует ветер, окружающий взорвавшуюся звезду. Для стационарного ветра $p = 2$, но на самых последних этапах эволюции предсверхновой звезды ветер не может быть стационарным, и параметр p можно варьировать в диапазоне от 1,5 примерно до 3,5. Другой вид оболочки, отделенной от выброса областью с более низкой плотностью, также рассматривается в нашем моделировании.

Кривые блеска были рассчитаны для сверхновых, взрывающихся в этих оболочках. Ударная волна образуется на границе между выбросом и оболочкой. Она очень эффективно преобразует энергию упорядоченного движения расширяющегося газа в энергию хаотического теплового движения частиц, которая может легко излучаться. В результате можно ожидать кривых блеска достаточно мощных для объяснения по крайней мере части сверхмощных сверхновых без предположений о необычно высокой энергии взрыва. Детальные вычисления подтверждают эти ожидания.

Для SLSNe типа II_n (т. е. событий с узкими линиями водорода в спектрах, отсюда индекс «n») используются богатые водородом оболочки. Для SLSN I (без водорода в спектрах) строятся модели из смеси углерода и кислорода с различными соотношениями C и O или же модели с гелиевыми оболочками. Модели могут содержать некоторое количество радиоактивных элементов, таких как ^{56}Ni , но это не обязательно в этом классе моделирования, так как эффект чистого взаимодействия выброса сверхновой и околозвездного вещества является достаточным для объяснения большинства SLSNe с нулевым количеством ^{56}Ni .

Синтетические кривые блеска в работе [5] вычисляются с помощью нашего радиационно-гидродинамического алгоритма с переносом излучения «STELLA» в стандартной установке. Код имитирует сферически симметричные гидродинамические потоки в сочетании с многогрупповой трактовкой переноса фотонов. Процедура расчета непрозрачности учитывает рассеяние на электронах, свободно-свободные и связанно-свободные процессы. Вклад спектральных линий (т. е. связанно-связанных процессов) учитывается в приближении непрозрачности расширяющейся среды и перекрытия спектральных линий из-за эффекта Доплера.

Взрывы были смоделированы в виде «тепловой бомбы» с переменной энергией E_{expl} порядка 2—4 фое, которая несколько больше, чем при стан-

дартном взрыве сверхновой с 1 foe, но значительно ниже, чем используется в гиперновых или в сверхновых, порожденных неустойчивостью из-за рождения пар (англ. pair-instability supernovae — PISNe).

Рисунок 3 показывает, как профили плотности, скорости, температуры и оптическая глубина (по росселандову среднему для непрозрачности) развиваются по времени для одной из моделей. Левые графики соответствуют эволюции до максимума кривой блеска (что происходит на 27-й день после взрыва для этой модели), правые графики показывают эволюцию после максимума блеска.

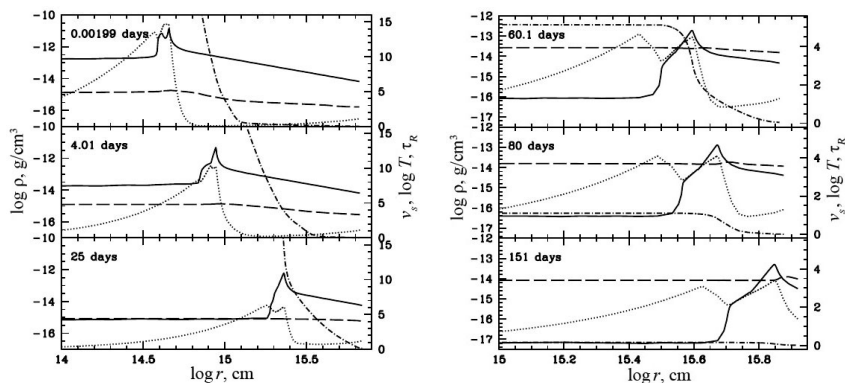


Рис. 3. Эволюция радиальных профилей плотности (сплошные линии), скорости в единицах 10^8 см/с (пунктир), температуры вещества (штриховые линии) и росселандова оптическая глубина (штрихпунктир) для одной из моделей [5]. Шкала плотности находится на левой оси Y , для всех других величин — на правой оси Y . Слева — эволюция гидродинамической структуры до максимума: сразу после взрыва и после 4 и 25 сут. Справа — те же параметры, но после максимума: в 60, 80 и 151 сут. Обратите внимание, что на левой и правой панелях используются различные масштабы для осей.

В самом начале формируется структура ударной волны в результате столкновения между выбросом и околозвездным веществом. Тогда излучение от фронта ударной волны нагревает газ в оболочке, что делает ее непрозрачной, и фотосфера довольно быстро перемещается в самые наружные слои. Когда радиус фотосферы достигает своего максимума, наблюдается предельное излучение от сверхновой.

Скорость роста радиуса фотосферы зависит от массы оболочки, поскольку для нагрева большей массы оболочки ударная волна должна излучить большее количество фотонов.

Другой параметр, который влияет на начальный рост радиуса фотосферы, — это химический состав оболочки. Например, кривая блеска поднимается быстрее для CO-оболочки, чем для гелиевой, так как в CO-смеси

высокая непрозрачность получается при более низкой температуре. Это поведение кривой блеска может помочь находить химический состав для некоторых наблюдаемых SLSNe.

Графики на правой панели рис. 3 показывают этапы, когда фотосфера медленно движется назад к центру, а оболочка и выброс наконец стали полностью прозрачными. В начале этой послемаксимальной стадии весь газ в оболочке уже нагрет фотонами, которые пришли из области ударных волн и продиффундировали через оболочку к внешнему краю, и вся система (выброс и оболочка) становится почти изотермической. Ударная волна слабеет со временем и излучает меньше фотонов, которые могут нагревать оболочку, поэтому температура внешней оболочки, куда еще не убежала ударная волна, падает.

Вещество между ударными волнами собирается в тонкий, плотный слой, который, в конце концов, будет содержать почти всю массу в системе. Формирование этого слоя приводит к численным трудностям, что существенно ограничивает временной шаг расчета. Еще одна проблема может появиться из-за образования этого слоя: тонкая, плотная оболочка с очень большим радиусом, скорее всего, будет неустойчивой и может фрагментироваться на более мелкие сгустки. Тогда задача станет существенно многомерной.

Если посмотреть на профили скорости, то можно увидеть, как с самого начала формируется структура с многократными отражениями. Картина течения очень быстро превращается в стандартную структуру из двух ударных волн (передней и возвратной). Этот переход мало зависит от начального профиля скорости в оболочке. Взаимодействие выброса с оболочкой приводит к аналогичным распределениям конечных скоростей.

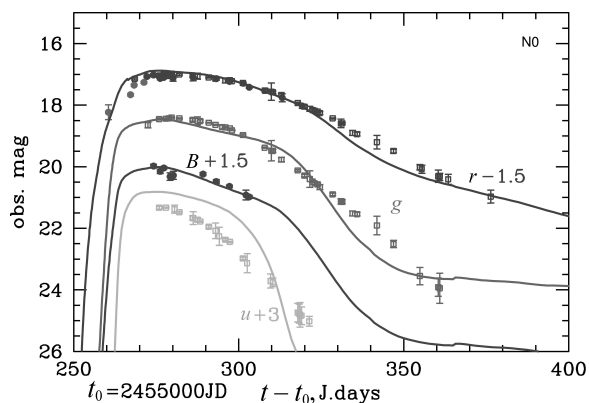


Рис. 4. Теоретические кривые блеска для модели с рисунка 3, одной из лучших для SN 2010gx, в фильтрах r , g , B и u в сравнении с наблюдениями Pan-STARRS (полосы u и g) и PTF (B и r)

Это течение похоже на автомодельное поведение в решении, найденном Д. К. Надёжиным и Р. Шевалье, но с излучением [6—8].

Рисунок 4 показывает, как модель с гидродинамикой течения, как на рис. 3, воспроизводит наблюдения в нескольких фильтрах для хорошо изученной сверхмощной сверхновой SN 2010gx.

Работа по расчетам кривых блеска SLSN поддержана грантом Российского научного фонда 14-12-00203.

Литература

1. *Bethe, H. A.* My Life in Astrophysics // *Annu. Rev. Astron. Astrophys.* — 2003. — V. 41. — P. 1—14.
2. *Papadopoulos, A.* DES13S2cmm: the first superluminous supernova from the Dark Energy Survey / A. Papadopoulos [et al.] // *Monthly Notices of the Royal Astronomical Society.* — 2015. — V. 449. — P. 1215—1227.
3. *Avishay, G.-Y.* Luminous Supernovae // *Science.* — 2012. — V. 337. — P. 927—932.
4. *Зельдович, Я. Б.* Физика ударных волн и высокотемпературных гидродинамических явлений / Я. Б. Зельдович, Ю. П. Райзер. — М.: Наука, 1966. — 688 с.
5. *Sorokina, E. I.* Type I Superluminous Supernovae as Explosions inside Non-Hydrogen Circumstellar Envelopes / E. I. Sorokina, S. I. Blinnikov, K. Nomoto, R. Quimby, A. Tolstov // *Astrophys. J.* — 2016. — V. 829. — P. 17.
6. *Надёжин, Д. К.* О начальной фазе взаимодействия разлетающейся оболочки звезды с окружающей средой: препринт № 1 / Д. К. Надёжин; ИТЭФ. — М., 1981. — 44 с.
7. *Nadyozhin, D. K.* On the initial phase of interaction between expanding stellar envelopes and surrounding medium // *Astrophysics and Space Science.* — 1985. — V. 112. — P. 225—249.
8. *Chevalier, R. A.* Self-similar solutions for the interaction of stellar ejecta with an external medium // *Astrophys. J.* — 1982. — V. 258. — P. 790—797.

РЕЛЯТИВИСТСКИЕ УДАРНЫЕ ВОЛНЫ: СТРУКТУРА ФРОНТА И ИЗЛУЧЕНИЕ

Е. В. Дершиев

Институт прикладной физики РАН, г. Нижний Новгород

Введение

Релятивистские ударные волны невозможно исследовать в лабораториях: они существуют в различных космических объектах. К числу объектов, порождающих релятивистские выбросы плазмы, можно отнести пульсары, микроквезары, активные ядра галактик и источники гамма-всплесков. Все эти объекты находятся очень далеко от Земли и единственный способ изучать их — это наблюдать излучения от образующихся там ударных волн.

Несмотря на различную природу источников, у релятивистских ударных волн в космосе есть общая особенность — все они бесстолкновительные. Отдельного внимания заслуживают источники гамма-всплесков. Существующие там ударные волны имеют лоренц-фактор Γ порядка нескольких сотен. Их можно считать эталонным примером релятивистских ударных волн. Кроме того, наблюдаемые спектры гамма-всплесков весьма разнообразны, и их объяснение представляет собой трудную задачу для теории. В силу этих причин, когда речь пойдет о характерных параметрах и спектрах, мы будем иметь в виду прежде всего источники гамма-всплесков.

Как преобразовать кинетическую энергию в излучение

Энергия источников гамма-всплесков и активных ядер галактик высвобождается в виде кинетической энергии релятивистской струи (джета), которая затем преобразуется в наблюдаемое излучение. Каким бы ни был механизм преобразования, он должен генерировать нетепловое излучение с достаточно высокой эффективностью.

Исторически релятивистские бесстолкновительные ударные волны были первыми кандидатами для объяснения нетеплового излучения гамма-всплесков и активных ядер галактик. Этой гипотезе и сейчас отдает предпочтение большинство авторов. Действительно, такие ударные волны имеют две очевидные привлекательные черты: они могут эффективно ускорять электроны посредством диффузионного механизма ускорения и могут производить собственное магнитное поле, необходимое для генерации синхротронного излучения энергичных электронов.

Однако внимательное изучение простейших моделей бесстолкновительных ударных волн выявляет серьезные противоречия. С одной стороны, генерируемое магнитное поле имеет небольшой пространственный масштаб и затухает слишком быстро для того, чтобы позволить ускорен-

ным электронам потерять свою энергию через излучение. С другой стороны, даже довольно слабое вмороженное магнитное поле в набегающем потоке останавливает диффузионное ускорение электронов [1].

Еще одним аргументом против моделей с бесстолкновительными ударными волнами (точнее, против синхротронного излучения как основного источника наблюдаемых фотонов во время основного импульса гамма-всплесков) является недостаток мягкого рентгеновского излучения в спектрах некоторых гамма-всплесков. В режиме быстрого синхротронного охлаждения и для однородной излучающей зоны фотонный индекс (т. е. степенной показатель функции распределения фотонов по энергиям) в области низких энергий не может быть больше, чем $\alpha = -3/2$. Большинство гамма-всплесков выходит за эту границу (см. положение левой вертикальной линии на рис. 1). Переход к преимущественно комптоновскому быстрому охлаждению позволяет объяснить фотонный индекс, меньший, чем $\alpha = -1$ [2], что немного улучшает ситуацию (см. положение центральной вертикальной линии на рис. 1). Наконец, асимптотика спектра отдельной частицы имеет фотонный индекс $\alpha = -2/3$ (см. положение правой вертикальной линии на рис. 1), к которому можно приблизиться, если излучение преимущественно комптоновское, а зона излучения неоднородна [3], или если электроны ускоряются и излучают в то же время [4, 5]. Очевидно, что синхротронное излучение не может объяснить все наблюдаемые спектры даже в доведенных до предела моделях.

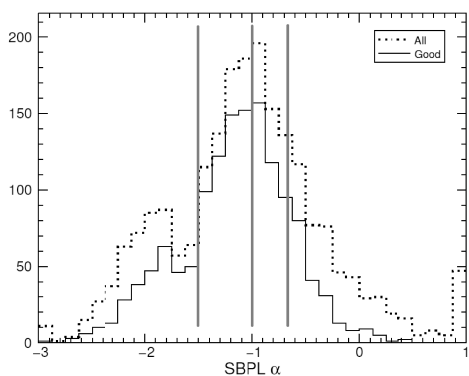


Рис. 1. Распределение наблюдаемых спектральных индексов в области низких энергий для спектральной модели в виде степенного спектра со сглаженным изломом [6]

В попытках решить проблему недостатка рентгеновского излучения было предложено несколько моделей. Наиболее популярной является модель радиационно-доминированной ударной волны (см., например, [7, 8]). Как подсказывает название, структура такой ударной волны определяется

излучением, захваченным внутри нее (таким образом, радиационно-доминированные ударные волны могут существовать только под фотосферой). Ширина фронта ударной волны определяется непрозрачностью среды и примерно равна средней длине свободного пробега фотонов. В радиационно-доминированных ударных волнах электроны не ускоряются, поскольку профиль скорости течения слишком гладкий. Вместо этого фотоны отбирают энергию непосредственно от потока из-за комптонизации на градиенте скорости. В некотором смысле можно рассматривать этот процесс как диффузионное ускорение фотонов на ударной волне.

Радиационно-доминированные ударные волны позволяют решить проблему недостатка рентгеновского излучения, взамен создавая проблему избытка рентгеновского излучения. Теперь фотонный индекс в области низких энергий не может быть меньше, чем $\alpha = -1$, так как при комптонизации сохраняется число фотонов. Более мягкие спектры возможны, если есть истинное излучение, но фотоны должны рождаться глубоко под фотосферой, что снижает эффективность. Кроме того, фотоны с энергиями выше $\Gamma m_e c^2 \approx 100$ МэВ не могут получаться при комптонизации тепловыми электронами из-за отдачи.

Еще одна возможность объяснить излучение в фазе основного импульса гамма-всплесков — магнитное пересоединение в джетах с преобладанием электромагнитной энергии. Модели этого типа все еще находятся на ранней стадии развития, но они похожи на модель «перемежающегося» пульсарного ветра [9], если применить соответствующее масштабирование. Исходя из этой аналогии можно ожидать довольно жесткой функции распределения ускоренных частиц, которая плохо согласуется с наблюдениями.

Наконец, отметим, что эффективное преобразование кинетической энергии в излучение возможно даже для стационарного релятивистского течения. Значительную часть вещества в джетах гамма-всплесков могут составлять свободные нейтроны [10]. На некотором расстоянии от центрального источника, ниже фотосферы джета, столкновения становятся слишком редкими для поддержания одинаковых гидродинамических скоростей протонов и ионов, т. е. происходит расщепление нейтронной компоненты и плазмы. Расщепление приводит к неупругим столкновениям протонов и нейтронов и рождению в них пионов, которые в итоге распадаются на фотоны высокой энергии, электроны, позитроны и нейтрино (мощный нейтринный сигнал свидетельствует в пользу этой модели). Этот механизм излучения надежен и эффективен [10, 11], но едва ли способен воспроизвести наблюдаемые спектры, если в модель не включать другие процессы излучения.

Можно сделать вывод, что каждый механизм излучения имеет свои преимущества и недостатки, но синхротронное излучение кажется предпочтительным выбором. Оно может объяснить спектры главных импуль-

сов для большинства гамма-всплесков, и его вклад представляется необходимым даже для тех всплесков, чьи спектры явно выходят за пределы синхротронного механизма. Кроме того, синхротронное излучение от бесстолкновительных ударных волн хорошо объясняет спектры послесвечения гамма-всплесков. Далее мы сосредоточимся именно на модели релятивистских бесстолкновительных ударных волн с синхротронным излучением.

Что нужно иметь в виду

Прежде чем обсуждать возможную модель, которая могла бы объяснить все наблюдения гамма-всплесков, перечислим некоторые факты, дающие представление о физике излучающей зоны. Если предположить, что лоренц-фактор гидродинамического движения $\Gamma \sim 300$, а плотность энергии магнитного поля одного порядка с плотностью энергии плазмы, то можно оценить лоренц-фактор электронов, ответственных за генерацию синхротронного излучения. В сопутствующей системе отсчета этот лоренц-фактор порядка нескольких сотен во время основного импульса и порядка нескольких тысяч на ранней стадии послесвечения.

Из этих цифр можно сделать вывод: 1) электроны быстро (по сравнению с временным масштабом расширения) теряют энергию на излучение и 2) комптонизация типичных (отвечающих пику спектрального распределения энергии) фотонов происходит на границе перехода в режим Клейна — Нишины (т. е. режим, когда квантовые эффекты становятся определяющими и сечение рассеяния зависит от энергии фотона). Последнее, как будет показано позже, — это не просто совпадение, а прямое следствие из модели модифицированной релятивистской ударной волны.

Рассмотрим теперь по шагам путь преобразования кинетической энергии потока в излучение (рис. 2). Во-первых, на фронте ударной волны генерируется магнитное поле. Этот шаг не является обязательным, можно исходить из конфигурации с уже существующим магнитным полем. Во-вторых, происходит диссипация магнитной энергии, а заряженные частицы в этом процессе нагреваются. По крайней мере половина диссипируемой

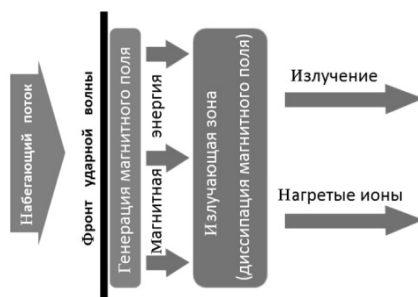


Рис. 2. Шаги преобразования энергии в излучающей зоне гамма-всплесков

энергии идет на нагрев легких частиц (электронов и, возможно, позитронов). Эта доля может быть и больше, если плазма в основном состоит из электронно-позитронных пар. В-третьих, электроны излучают всю свою энергию (так как они находятся в режиме быстрого радиационного остывания).

Из приведенного выше рассмотрения немедленно следует, что для стационарного решения поток магнитной энергии в излучающую область одного порядка с потоком энергии излучения из нее. Это значит, что $\frac{e_B c}{3} \sim \frac{e_r c}{2}$, где e_B — плотность энергии магнитного поля, а e_r — плотность энергии излучения. Таким образом, плотность энергии магнитного поля одного порядка с плотностью энергии излучения (в соответствии с нашими предыдущими оценками эффект Клейна — Нишины является умеренным и не влияет на вывод).

В режиме Клейна — Нишины сечение комптонизации (электрон-фотонного рассеяния) и поглощения фотонов (двухфотонное рождение электрон-позитронной пары) близки по величине, так что времена жизни электронов высоких энергий и фотонов примерно одинаковы. Если электроны быстро теряют энергию на излучение, то и произведенные ими комптоновские фотоны эффективно поглощаются.

Модели излучающей области релятивистских ударных волн должны включать обратное комптоновское излучение в качестве важной своей части, а также должны учитывать поглощение комптоновских фотонов как в отходящем потоке, так и в набегающем потоке.

Конверсионный механизм ускорения

Существует механизм ускорения, специфичный для таких условий, когда присутствуют релятивистские гидродинамические движения и одновременно поглощение фотонов высоких энергий является эффективным [12]. Рабочий цикл этого (конверсионного) механизма ускорения может быть описан следующим образом (рис. 3). В процессе комптоновского рассеяния электрон порождает высокоэнергетический фотон (метка 1 на рисунке), этот фотон обгоняет ударную волну и рождает электрон-позитронную пару в набегающем потоке (метка 2). Ударная волна доходит до изотропизованных частиц и повышает их энергию (метка 3).

Следует отметить, что существует цикл конверсионного ускорения для протонов (преобразование протонов в нейтроны и обратно), который требует более жестких условий и здесь не обсуждается.

В каждом цикле конверсионного ускорения энергия частицы увеличивается на фактор порядка Γ^2 , а эффективность ускорения (отношение суммарной энергии частиц, которые прошли через цикл, к энергии частиц,

которые были в начале) можно оценить как $p_c \Gamma^2$. Вероятность прохождения через весь цикл p_c является произведением вероятности выхода фотона из отходящего потока ($\approx 1/3$), относительной эффективности комптоновского излучения ($= y/(1+y)$) и эффективности радиационного остывания.

В зависимости от эффективности конверсионного ускорения возможны три ситуации. Если $p_c \Gamma^2 \ll 1$, то данный механизм не играет существенной роли; этот случай реализуется в нерелятивистских ударных волнах или крайне неэффективных (с точки зрения излучения) релятивистских ударных волнах. Если $p_c \Gamma^2 \sim 1$, то механизм удобен для построения моделей излучающей области [13]; этот случай реализуется в активных ядрах галактик с их умеренно релятивистскими джетами. Наконец, и именно этот случай реализуется в источниках гамма-всплесков, если $p_c \Gamma^2 \gg 1$, то результатом конверсионного ускорения является передача почти всей доступной энергии частицам с максимально возможным гирорадиусом. Они затем быстро излучают, порождая электромагнитный каскад и приводя в конечном счете к формированию плоских спектров без особенностей, которые никогда не наблюдаются у гамма-всплесков.

Можно представить два способа уменьшить эффективность конверсионного ускорения до разумных пределов. Во-первых, можно снизить эффективность излучения ниже $1/\Gamma$, что вряд ли можно считать хорошим решением для гамма-всплесков с их чрезвычайно большими наблюдаемыми светимостями. Во-вторых, можно уменьшить скачок лоренц-фактора на фронте ударной волны и сделать его умеренно релятивистским. Это приемлемое решение, но оно требует построения модели модифицированной релятивистской ударной волны.

Модифицированные релятивистские ударные волны

В модели модифицированных релятивистских ударных волн есть три ключевых момента. Во-первых, она должна учитывать комптонизацию излучения и поглощение комптонизированных фотонов наравне с синхротронным излучением и ускорением частиц (обыкновенно учитываются только два последних эффекта). Во-вторых, она должна объяснить, почему конверсионное ускорение ограничено и не приводит к перекачке всей энергии в наиболее энергичные частицы. В-третьих, она должна решить проблему несоизмеримости пространственных масштабов, суть которой можно изложить следующим образом.

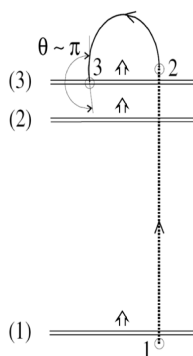


Рис. 3. Шаги конверсионного ускорения для электронов и позитронов [12]

Длина нарастания магнитного поля в бесстолкновительной ударной волне обычно принимается равной толщине плазменного скин-слоя, $l_s \sim 10$ см в фазе основного импульса гамма-всплеска. Длина радиационного остывания электронов в этой фазе составляет $l_c \sim 10^5$ см. Наконец, длина затухания магнитного поля l_d физически связана с длиной нарастания, но должна быть сравнима с длиной радиационного остывания, которая по меньшей мере на четыре порядка больше.

Модель, которая была предложена в работе [5], согласованным образом рассматривает как ускорение частиц, так и их излучение (в том числе обратное комптоновское) и решает проблему несоизмеримости масштабов.

В этой модели основная излучающая зона находится в отходящем потоке у фронта ударной волны. Там электроны (и позитроны) нагреваются затухающим магнитным полем и генерируют синхротронное и обратное комптоновское излучение, причем их мощности примерно равны, когда ударная волна находится в режиме быстрого остывания электронов или близко к нему. Столкновения испущенных в этой зоне высокоэнергичных комптоновских фотонов с низкоэнергичными синхротронными фотонами рождают электрон-позитронные пары в набегающем течении и модифицируют эту область (рис. 4). Рожденные пары затем сносятся потоком к фронту ударной волны и увеличивают свою энергию. Процесс носит общий характер и будет иметь место в любой релятивистской ударной волне.



Рис. 4. Перенос энергии и импульса в модифицированной релятивистской ударной волне

По оценкам, энергия излучающих электронов близка к значению, где эффект Клейна — Нишины становится важным. Это является естественным результатом эволюции параметров ударных волн. Если ударную волну запустить с такими параметрами, что комптонизация происходит в режиме Клейна — Нишины, то комптоновские фотоны эффективно поглощаются и количество вторичных электрон-позитронных пар быстро превысит число первичных электронов. Так как мощность нагрева фиксирована затуханием магнитного поля, то большее число излучающих частиц означает,

что типичная частица становится менее энергичной и производит менее энергичные комптоновские фотоны, которые поглощаются менее эффективно. Эта обратная связь быстро приводит параметры ударной волны к таким, что энергия типичного комптоновского фотона оказывается чуть ниже порога рождения пар при взаимодействии с наиболее распространенными синхротронными фотонами. Непрозрачность для комптоновского излучения падает с темпом рождения пар. Таким образом, темп рождения пар стабилизируется на таком уровне, что он точно восполняет потери частиц через отходящий поток. Полученное состояние представляет собой аттрактор в пространстве параметров, который существует в широком диапазоне физических условий. Интересно, что для типичных физических условий это устойчивое решение правильно предсказывает положение пика в спектре излучения гамма-всплесков.

Еще одна ключевая составляющая модели — модификация магнитного поля электрон-позитронными парами. Пары, которые рождаются в набегающем потоке, имеют анизотропную функцию распределения, что, в свою очередь, приводит к предварительному усилению магнитного поля. Поле постепенно нарастает на длине на несколько порядков больше, чем толщина плазменного скин-слоя. При этом магнитная энергия сконцентрирована в крупномасштабных модах, которые медленно растут, но и затухают в отходящем потоке в течение длительного времени. Пространственный масштаб как роста, так и затухания магнитного поля устанавливается длиной поглощения комптоновского излучения и, следовательно, оказывается одного порядка с длиной радиационного остывания электронов. Эта картина качественно подтверждается недавними результатами численного моделирования [14] (рис. 5).

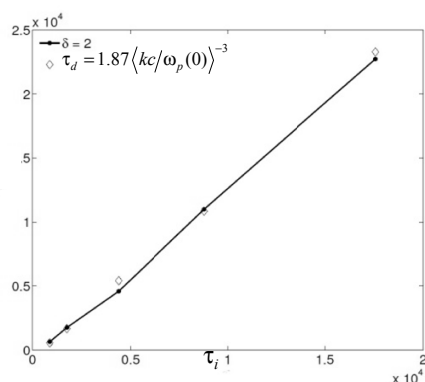


Рис. 5. Временной масштаб затухания магнитного поля в зависимости от длительности инжекции. Ромбами показано время затухания, теоретически предсказанное для среднего волнового числа магнитных возмущений, измеренных при моделировании в конце инжекции [14]

Энергия, выделяемая при затухании магнитного поля, неизбежно переходит к заряженным частицам, а затем в синхротронное и обратное комптоновское излучение. Локальная функция распределения лептонов

(и, следовательно, локальный спектр излучения) зависит от темпа диссипации магнитного поля и по мере сноса плазмы отходящим течением меняется вместе с меняющимся магнитным полем. По сравнению с однозонными моделями синхротронного излучения ударной волны такая структура излучающей области обеспечивает дополнительную гибкость в формировании спектров. Это также приводит к единому масштабу длины для затухания магнитного поля, ускорения частиц и их радиационного остывания. Таким образом, для модели не стоит проблема несоизмеримости масштабов.

Существует три различных излучающих зоны. Основным (самым мощным) источником синхротронного и комптоновского излучения является область затухающего магнитного поля в отходящем течении. Наиболее энергичные пары, которые образовались в набегающем потоке, могут иметь достаточно времени, чтобы излучить свою энергию еще до достижения фронта ударной волны. Они ответственны за формирование дополнительной излучающей зоны в набегающем потоке. Для типичных параметров эта зона дает небольшой вклад в излучение по сравнению с основной излучающей зоной в отходящем потоке. Однако при определенных условиях ситуация может измениться. Набегающий поток перехватывает значительную долю излучения, испущенного в отходящем потоке, и потенциально часть этой энергии может переизлучаться в самом набегающем потоке.

Третья компонента излучения связана с энергичными лептонами из набегающего потока, которые при пересечении фронта ударной волны оказываются в сильном магнитном поле и быстро теряют свою энергию, в основном из-за синхротронного излучения. Эту компоненту излучения можно назвать переходной, так как она обязана своим происхождением частицам, которые, переходя из набегающего потока в отходящий поток, излучают в непосредственной близости от фронта ударной волны. Вклад переходной излучающей области в спектр синхротронного излучения представляет собой степенной хвост, простирающийся в область высоких энергий.

Из-за передачи импульса набегающему потоку комптоновскими фотонами эта область течения ускоряется. Основное изменение лоренц-фактора происходит в набегающем потоке, по мере того как он постепенно нагружается вторичными парами. Чем больше параметр обратной связи \tilde{a} (т. е. часть энергии потока, которая излучается в отходящем потоке, а затем поглощается в набегающем потоке), тем меньше величина скачка лоренц-фактора на фронте ударной волны (рис. 6). Если параметр обратной связи является достаточно большим, $\tilde{a} \gtrsim 0,14$, то скачок исчезает совсем.

Для типичных параметров скачок лоренц-фактора на фронте ударной волны умеренно релятивистский, даже если сама ударная волна ультрарелятивистская (так что разница между внутренними и внешними ударными

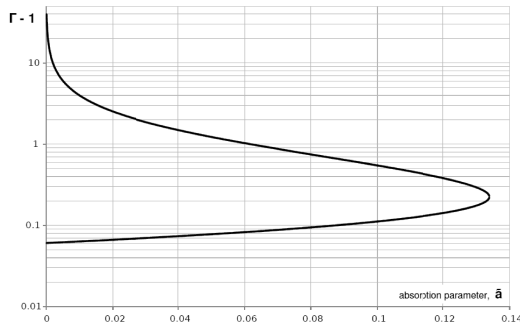


Рис. 6. Лоренц-факторы течения у фронта ударной волны (в системе отсчета фронта) как функции параметра обратной связи \bar{a} : верхняя ветвь для набегающего потока, нижняя ветвь для отходящего потока. Лоренц-фактор ударной волны принят бесконечно большим [5]

волнами в источниках гамма-всплесков не столь значительная, как принято полагать). В то же время скорость отходящего потока по отношению к фронту ударной волны заметно превосходит величину $c/3$, вычисляемую для немодифицированных ударных волн. Большая скорость отходящего потока вместе с относительно небольшим скачком лоренц-фактора на фронте ударной волны в значительной степени снижают эффективность диффузионного механизма ускорения, так что он, вероятно, не играет существенной роли. Тем не менее степенной хвост функции распределения электронов при высоких энергиях все-таки формируется благодаря конверсионному механизму ускорения, который также подавлен и имеет ограниченную эффективность.

Одна из наиболее важных особенностей модели — сопоставимые эффективности синхротронного и обратного комптоновского излучения. Это является неотъемлемым свойством представленной модели, так как комптоновское излучение ответственно за самонастройку параметров ударной волны. Нет никакого способа избавиться от комптоновского пика в спектре излучения, оставаясь в рамках модели. Излучение частично скрыто от наблюдения двухфотонным поглощением внутри самой ударной волны, но низкоэнергетическая часть комптоновского излучения покидает ударную волну без существенного ослабления.

В заключение можно сказать, что стандартная картина бесстолкновительных ударных волн должна быть изменена, с тем чтобы учесть нелокальную взаимосвязь между набегающим и отходящим потоками. Эта связь влияет на структуру ударной волны и регулирует ее; она возникает из-за поглощения в набегающем потоке комптоновских фотонов, испущенных в отходящем потоке. В результате ударная волна имеет естественный механизм самонастройки (путем регулирования множественности

рождения вторичных пар), так что параметры ударной волны эволюционируют в сторону аттрактора, где пик спектра синхротронного излучения приходится на такую энергию, что для типичного излучающего электрона комптонизация происходит в режиме, близком к режиму Клейна — Нишины.

Выводы

Есть много правдоподобных механизмов для преобразования кинетической энергии потока в излучение в релятивистских ударных волнах: синхротронное излучение ускоренных электронов, комптонизация в радиационно-доминированных ударных волнах, пересоединение в джетах с преобладанием электромагнитной энергии, преобразование кинетической энергии потока в момент расщепления протонной и нейтронной компонент. Каждый механизм имеет свои привлекательные черты, но ни один из них, взятый в отдельности, не может воспроизвести все наблюдаемые спектры. На данный момент синхротронное излучение видится предпочтительным выбором. Оно является единственным вариантом для объяснения спектров излучения послесвечения гамма-всплесков, но и для объяснения спектров излучения в момент всплеска требуется некоторый вклад синхротронного излучения.

Наблюдение гамма-всплесков в спектральной области 100 ГэВ должно дать ключ к выбору правильного механизма. Модели синхротронного излучения неизбежно предсказывают измеримый пик обратного комптоновского излучения в спектре. Его обнаружение позволит наконец дать прямую оценку напряженности магнитного поля и энергии излучающих электронов. Другие механизмы предсказывают вклад в излучение на высоких энергиях с отличающейся величиной и/или с отличающимися спектрами, так что это излучение является полезным инструментом, когда необходимо отличить одну модель от другой.

Детектирование нейтринного сигнала (или получение верхнего предела) в области энергий 10—100 ГэВ в момент всплеска также поможет сделать выбор в пользу той или иной модели. Такой сигнал несет информацию о количестве нейтронов в джетах и может способствовать измерению лоренц-фактора джета и эффективности его излучения.

Неустранимые противоречия в существующих моделях подталкивают нас к пересмотру парадигмы релятивистских бесстолкновительных ударных волн. В рамках новой парадигмы ударные волны являются сильно модифицированными, так что скачок лоренц-фактора на разрыве оказывается лишь умеренно релятивистским. В таких ударных волнах диффузионный механизм ускорения подавлен и заменяется (ограниченным) конверсионным ускорением. Благодаря обратной связи между набегающим и отходящим потоками временные масштабы нарастания и диссипа-

ции магнитного поля оказываются сравнимыми со временем радиационного остывания электронов. Тем самым устраняется проблема слишком быстрого спада магнитного поля в излучающей области. Наконец, обратное комптоновское излучение является существенной частью физической картины модифицированных релятивистских ударных волн. Его обнаружение (или несомненное доказательство отсутствия) позволит значительно продвинуть наше понимание физики релятивистских ударных волн.

Литература

1. *Sironi, L.* Relativistic Shocks: Particle Acceleration and Magnetization / L. Sironi, U. Keshet, M. Lemoine // *Space Science Rev.* — 2015. — V. 191. — P. 519—544.
2. *Derishev, E. V.* Physical parameters and emission mechanism in gamma-ray bursts / E. V. Derishev, V. V. Kocharovskiy, V. V. Kocharovskiy // *Astron. Astrophys.* — 2001. — V. 372. — P. 1071—1077.
3. *Derishev, E. V.* Synchrotron emission in the fast cooling regime: which spectra can be explained? // *Astrophys. Sp. Sci.* — 2007. — V. 309. — P. 157—161.
4. *Poutanen, J.* Gamma-ray burst spectra from continuously accelerated electrons / J. Poutanen, B. E. Stern // *Nuovo Cimento C.* — 2005. — V. 28. — P. 443.
5. *Derishev, E. V.* Particle acceleration, magnetization and radiation in relativistic shocks / E. V. Derishev, T. Piran // *MNRAS.* — 2016. — V. 460. — P. 2036—2049.
6. *Goldstein, A.* The BATSE 5B Gamma-Ray Burst Spectral Catalog / A. Goldstein, R. D. Preece, R. S. Mallozzi, M. S. Briggs, G. J. Fishman, C. Kouveliotou, W. S. Paciesas, J. M. Burgess // *Astrophys. J. Suppl. Ser.* — 2013. — V. 208. — P. 21.
7. *Levinson, A.* Observational Signatures of Sub-photospheric Radiation-mediated Shocks in the Prompt Phase of Gamma-Ray Bursts // *Astrophys. J.* — 2012. — V. 756. — P. 174.
8. *Beloborodov, A. M.* Sub-photospheric shocks in relativistic explosions [Электронный ресурс] // *ArXiv.org / Cornell University.* — [Ithaca], 2016. — arXiv:1604.02794. — Режим доступа: <https://arxiv.org/abs/1604.02794>. — (Дата обращения: 22.02.17).
9. *Lyubarsky, Y.* Reconnection in a Striped Pulsar Wind / Y. Lyubarsky, J. G. Kirk // *Astrophys. J.* — 2001. — V. 547. — P. 437—448.
10. *Derishev, E. V.* The Neutron Component in Fireballs of Gamma-Ray Bursts: Dynamics and Observable Imprints / E. V. Derishev, V. V. Kocharovskiy, V. V. Kocharovskiy // *Astrophys. J.* — 1999. — V. 521. — P. 640—649.
11. *Beloborodov, A. M.* Collisional mechanism for gamma-ray burst emission // *MNRAS.* — 2010. — V. 407. — P. 1033—1047.
12. *Derishev, E. V.* Particle acceleration through multiple conversions from a charged into a neutral state and back / E. V. Derishev, F. A. Aharonian, V. V. Kocharovskiy, V. V. Kocharovskiy // *Phys. Rev. D.* — 2003. — V. 68. — P. 043003.
13. *Stern, B. E.* A photon breeding mechanism for the high-energy emission of relativistic jets / B. E. Stern, J. Poutanen // *MNRAS.* — 2006. — V. 372. — P. 1217—1226.
14. *Garasev, M.* Impact of continuous particle injection on generation and decay of the magnetic field in collisionless shocks / M. Garasev, E. Derishev // *MNRAS.* — 2016. — V. 461. — P. 641—646.

НЕЛИНЕЙНЫЕ ВОЛНОВЫЕ ЯВЛЕНИЯ В ПЛАЗМЕННО-ПЫЛЕВЫХ СИСТЕМАХ

С. И. Попель

Институт космических исследований РАН, г. Москва
Московский физико-технический институт, г. Долгопрудный
Национальный исследовательский университет «Высшая школа экономики», г. Москва

Введение

Пылевая плазма (*dusty plasma*) представляет собой ионизованный газ, содержащий электроны, ионы, нейтральные частицы, а также заряженные твердые частицы и/или жидкие капли, которые либо самопроизвольно образуются в плазме в результате различных процессов, либо вводятся в плазму извне. В лабораторных условиях пылевая плазма была впервые обнаружена И. Лэнгмюром еще в 1920-х годах [1]. Однако ее активное исследование началось лишь в последние десятилетия в связи с целым рядом приложений, таких как электрофизика и электродинамика продуктов сгорания ракетных топлив, электрофизика рабочего тела магнетогидродинамических генераторов на твердом топливе, физика пылегазовых облаков в атмосфере. Существенным направлением исследований является физика серебристых облаков и полярных мезосферных радиоотражений в летней полярной мезосфере Земли [2, 3]. Пыль и пылевая плазма широко распространены в космосе [4—7]. Они обнаружены в планетных кольцах, приповерхностных областях безатмосферных космических тел, хвостах комет, в межпланетных и межзвездных облаках. Пылевая плазма обнаружена также вблизи искусственных спутников Земли [8] и в пристеночной области термоядерных установок с магнитным удержанием [9]. Повышенный интерес к изучению пылевой плазмы связан с рядом приложений, таких как технологии плазменного напыления и травления в микроэлектронике, а также с производством тонких пленок и наночастиц. Наличие пылевых частиц в плазме не только приводит к загрязнению поверхности полупроводникового элемента и тем самым к увеличению выхода дефектных элементов, но и возмущает плазму, как правило, непредсказуемым образом. Уменьшение или предотвращение этих негативных эффектов невозможно без изучения свойств пылевой плазмы. Бурный рост плазменно-пылевых исследований, продолжающийся и по сей день, начался в середине 1990-х годов, после того как в двух лабораториях на Тайване [10] и в Германии [11] удалось наблюдать формирование плазменно-пылевых кристаллических структур. Возможность существования пылевой плазмы в различных фазах позволила по аналогии с так называемыми *комплексными жидкостями* ввести новый термин для ее описания — *комплексная плазма*.

Практически в любых экспериментальных исследованиях, в которых проявляются неустойчивости плазмы, необходим анализ нелинейных эф-

фектов. Сам факт наличия неустойчивости говорит о неприменимости линейной теории для описания плазмы, поскольку экспоненциальный рост возмущений очень быстро делает линейное приближение непригодным. В пылевой плазме нелинейные эффекты имеют свою специфику и представляют важное направление исследований. Однако их рассмотрение представляет существенно более сложную задачу, чем в обычной плазме, без пыли. В связи с этим на настоящем этапе имеются лишь отдельные попытки рассмотреть нелинейные плазменно-пылевые задачи, и до сих пор не построено систематической теории, описывающей нелинейную физику пылевой плазмы. В настоящей статье обсуждается ряд нелинейных плазменно-пылевых задач, показывающих важность нелинейных исследований в пылевой плазме с точки зрения объяснения экспериментов и природных явлений.

Ударные волны

Важным направлением исследований в пылевой плазме являются ударные волны. Ударные волны — один из собственных плазме типов нелинейных волн. Существуют различные возможности достижения «нелинейного уровня» волнами в пылевой плазме. Кроме традиционных для обычной плазмы возможностей (развитие неустойчивостей, ввод энергии в плазму), в пылевой плазме важным оказывается коллективный процесс, аналогичный процессу генерации ударных волн в газовой динамике.

Дисперсия пылевых звуковых и пылевых ионно-звуковых волн такова, что части волнового фронта, отвечающие более высокой плотности, движутся быстрее, чем соответствующие части меньшей плотности. В результате фронт волны делается круче и возможно образование разрыва на профиле возмущения [12]. Однако существует принципиальное различие между ростом крутизны фронта ударной волны в газовой динамике и в пылевой плазме. В нейтральном газе рост крутизны фронта может быть остановлен диссипативными процессами, связанными с вязкостью газа, которая обусловлена столкновениями частиц газа. В пылевой плазме существенна аномальная диссипация, связанная с зарядкой пылевых частиц. Этот процесс через воздействие на электрическое поле в волне обеспечивает существование стационарных пылевых ионно-звуковых ударно-волновых структур в пылевой плазме в отсутствие внешнего электромагнитного излучения в предположении постоянства полного числа электронов и ионов (что может быть обеспечено, например, рекомбинацией электронов и ионов, попадающих на пылевую частицу, в нейтральные атомы, которые в дальнейшем отделяются от пылевой частицы и ионизируются) [13, 14]. Именно этот процесс приводит к тому, что возмущение, например в виде пылевого ионно-звукового солитона, может эволюционировать в пылевую ионно-звуковую ударно-волновую структуру [15]. Отмеченный вид пылевых ионно-звуковых ударных волн принципиально отличается от бесстолкновительных ударных волн в плазме, которые связаны с взаимо-

действием типа «волна — частица» [16]. Пылевые ионно-звуковые ударные волны бесстолкновительны в том смысле, что эффект столкновений между электронами и ионами пренебрежимо мал. Однако в отличие от классических бесстолкновительных ударных волн аномальная диссипация вызывается взаимодействием электронов и ионов с пылевыми частицами в виде микроскопических токов, текущих на частицу.

Таким образом, представляется, что на ионно-звуковых временных масштабах именно ударно-волновые структуры являются наиболее важными нелинейными структурами для пылевой плазмы. Именно на ионно-звуковых временных масштабах удалось выполнить первые успешные эксперименты по ударно-волновым структурам в пылевой плазме [17, 18]. Ниже приводится краткое описание этих экспериментов.

Эксперименты в Японии

Эксперименты в Японии в Institute of Space and Astronautical Science проводились на двойном плазменном приборе, отличающемся от обычного наличием дозатора пыли ([17], см. также [19]). Продольный размер установки составлял 90 см, а ее диаметр — 40 см. В плазму вводились стеклянные пылевые частицы со средним радиусом $a \approx 4,4$ мкм.

Параметры плазмы в экспериментах [17] принимали следующие значения: парциальное давление газа нейтралов (аргон) находилось в диапазоне $P_{Ar} = (3 \div 6) \cdot 10^{-4}$ торр, температура электронов $T_e \approx (1 \div 1,5)$ эВ, температура ионов $T_i < 0,1$ эВ, значение концентрации пылевых частиц n_d изменялось в пределах от 0 до $\sim 10^5$ см $^{-3}$, а концентрации электронов n_e — от $\sim 10^8$ до $\sim 10^9$ см $^{-3}$. Для возбуждения ионно-звуковой ударной волны к управляющему аноду (driver anode) прилагался трапециевидный, быстро достигающий амплитуды импульс напряжения (ramp voltage) с амплитудой 2 В и длительностью приблизительно 10 мкс.

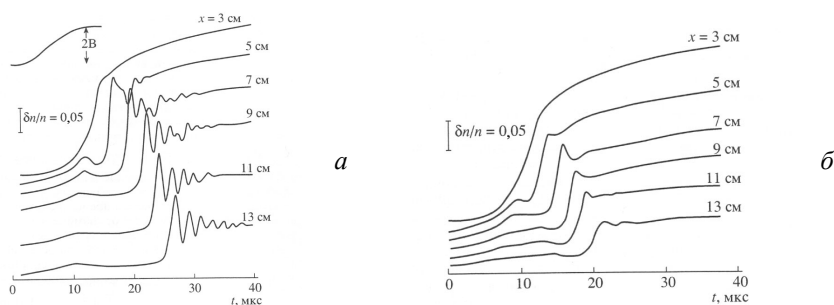


Рис. 1. Проявление осцилляторной структуры ударно-волнового фронта в отсутствие пыли (а) и ее подавление при наличии пыли с концентрацией $n_d = 5 \cdot 10^4$ см $^{-3}$ (б) в экспериментах на двойном плазменном приборе [17]

В экспериментах [17] в качестве основной выделялась следующая особенность ионно-звуковых ударно-волновых возмущений в пылевой плазме. В окрестности фронта на профиле ударных волн в отсутствие пыли имеются осцилляции, обусловленные разделением зарядов (электронов и ионов). Наличие пыли подавляет эти осцилляции. Данное утверждение хорошо иллюстрируется рис. 1, на котором показаны ударно-волновые профили на разных расстояниях от сетки, разделяющей установку на две части, — в отсутствие пыли и при ее наличии с концентрацией $n_d = 5 \cdot 10^4 \text{ см}^{-3}$.

Отметим, что на той же самой установке проводились эксперименты по наблюдению пылевых ионно-звуковых солитонов [20]. При этом наиболее существенным отличием от экспериментов по ударным волнам было использование положительной части синусоидального импульса напряжения длительностью 20 мкс.

Эксперименты в США

Эксперименты в США в Университете штата Айова (the University of Iowa) проводились на Q-машине, модифицированной для того, чтобы обеспечить присутствие пыли в установке [18]. Плазма паров цезия (ионы Cs^+) в этом эксперименте создавалась методом поверхностной ионизации. Ионы цезия при этом возникали в результате облучения пучком атомов цезия нагретой до температуры порядка (2000 ÷ 2500) К горячей пластины в торцевой части установки и поверхностной ионизации атомов цезия на ней.

Параметры плазмы в экспериментах [18] принимали следующие значения: $T_e \approx T_i \approx 0,2 \text{ эВ}$, $n_i \sim (10^6 \div 10^7) \text{ см}^{-3}$, $a \sim (0,1 \div 1) \text{ мкм}$. Параметр $\epsilon Z_d \equiv n_d Z_d / n_i$ менялся в пределах от 0 до 0,95 (здесь Z_d — зарядовое число пылевой частицы). Исследовалась эволюция начального возмущения — «ступеньки» ионной концентрации, созданной в установке с помощью сетки, на которой поддерживался электростатический потенциал порядка -6 В по отношению к потенциалу горячей пластины. Ширина начального возмущения равнялась приблизительно 25 см. Профили концентрации плазмы определялись на основе измерений ионного тока насыщения, падающего на зонд Ленгмюра.

Вот основные результаты экспериментов [18].

1. Пылевые ионно-звуковые ударные волны формируются лишь при достаточно больших значениях концентрации пыли (для данных [18] при $\epsilon Z_d \geq 0,75$). При этом вывод о формировании ударной волны основывается на росте со временем крутизны фронта возмущения. В случае достаточно малых концентраций пыли роста крутизны не происходит. Наоборот, фронт уширяется. Данный результат хорошо иллюстрируется рис. 2, на котором изображены профили концентрации плазмы на различных расстояни-

ях от сетки в отсутствие пыли и при наличии пыли ($\epsilon Z_d \geq 0,75$) в зависимости от времени [18]. Наклонными прямыми линиями на рис. 2 показано продвижение переднего и заднего фронтов импульса. Расхождение наклонных прямых линий на рис. 2, *а* соответствует уширению, а их схождение на рис. 2, *б* — росту крутизны профиля импульса.

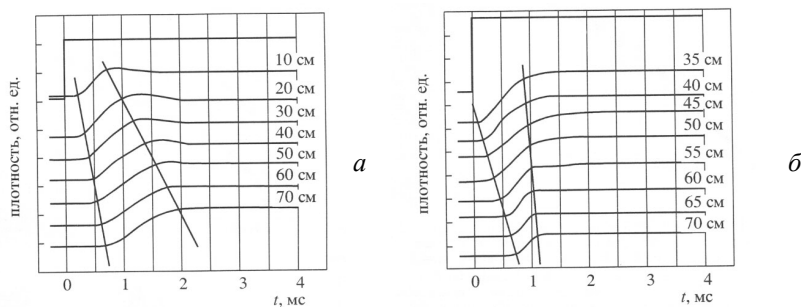


Рис. 2. Профили концентрации плазмы n на различных расстояниях от сетки в отсутствие пыли (*а*) и при наличии пыли ($\epsilon Z_d \geq 0,75$) (*б*) в зависимости от времени t в экспериментах на Q-машине [18]. В верхней части рисунков показан прямоугольный импульс напряжения, приложенного к сетке

2. Скорость пылевых ионно-звуковых волн существенно повышается с увеличением ϵZ_d .

Теоретические исследования

Начало теоретическим работам по пылевым ионно-звуковым ударно-волновым структурам положила статья [13]. В ней описывался новый вид ударных волн, связанных с аномальной диссипацией, природа которой вытекает из процесса зарядки пылевых частиц. Продолжением этой статьи послужили работы [21, 22], где была развита методология исследования ударно-волновых структур в пылевой космической плазме и предложены возможные применения исследований ударных волн в пылевой плазме, а именно ударные волны и процессы формирования звезд и планет, ударные волны при взрывах сверхновых звезд, ускорение частиц в ударных волнах, фазовые переходы в ударных волнах в пылевой плазме. В работе [15] показана возможность превращения пылевого ионно-звукового солитона в пылевую ионно-звуковую ударно-волновую структуру. После проведения экспериментов [17, 18] по пылевым ионно-звуковым ударным волнам основным направлением теоретических исследований стали попытки их описать. Для этого была разработана модель ионизирующего источника [14, 23, 24], на основе которой удалось объяснить основные результаты экспериментов. Так, например, в работе [24] наряду с основ-

ными результатами экспериментов [18], перечисленными выше, выделялось следующее свойство наблюдаемых ударных волн.

В случае формирования ударно-волновой структуры ширина фронта $\Delta\xi$ соответствует теоретической оценке, основанной на модели [13]:

$$\Delta\xi \sim Mv_s / v_{ch}, \quad (1)$$

где Mv_s — скорость ударно-волновой структуры, M — число Маха, v_s — скорость ионного звука, v_{ch} — частота зарядки пылевых частиц.

Оценка (1) указывает на то, что ударные волны в экспериментах [18] связаны с аномальной диссипацией, природа которой вытекает из процесса зарядки пылевых частиц. Для параметров экспериментов [18] $T_e \approx T_i \approx 0,2$ эВ, $n_i \approx 10^7$ см⁻³, $a \approx 1$ мкм и $M \approx 1,94$ (см. [14]) теоретическое значение $\Delta\xi \approx 80$ см, полученное на основе (1), согласуется с наблюдаемой величиной (40 см). Учет эффектов рассеяния ионов на пылевых частицах при вычислении значения ширины ударно-волнового фронта уменьшает это значение. Численный расчет в рамках модели ионизирующего источника, учитывающий наряду с эффектами зарядки пылевых частиц, в частности, эффекты рассеяния ионов на пылевых частицах, дает величину $\Delta\xi \approx 20$ см [14].

В работах [17, 20], кроме описания экспериментов, предпринималась попытка их теоретического исследования. Однако при этом использовалось модифицированное уравнение Кортевега — де Вриза — Бюргерса (КдВБ), в котором численные коэффициенты выбирались для обеспечения качественного согласия теоретических и экспериментальных результатов. В данном подходе учитывалась кинематическая вязкость ионов при их взаимодействии с пылевыми частицами. Однако, как было показано в работе [25], в условиях экспериментов [17, 20] средняя длина свободного пробега ионов при столкновениях с пылевыми частицами превосходит характерную длину пылевой ионно-звуковой волны, и, следовательно, вязкость должна полагаться нулевой. Учет вязкости, однако, может оказаться необходим при рассмотрении сильностолкновительной пылевой плазмы, когда длина свободного пробега ионов оказывается меньше длины волны пылевой звуковой волны или сопоставима с ней.

Покажем, что теоретические методы, развитые в работах [14, 23, 24], позволяют описать все основные результаты лабораторных экспериментов по генерации пылевых ионно-звуковых ударных волн [17, 18]. Как уже отмечалось, в экспериментах [17], выполненных на двойном плазменном приборе, в качестве основной выделялась следующая особенность ударных ионно-звуковых волн в пылевой плазме. В окрестности фронта на профиле ударных волн в отсутствие пыли имеются осцилляции, обусловленные разделением зарядов (электронов и ионов). Наличие пыли подав-

ляет эти осцилляции. Вычисления в рамках модели ионизирующего источника проводились для разных значений концентрации пыли в ситуации, когда $a = 4,4$ мкм, $T_e = T_i = 1,5$ эВ,

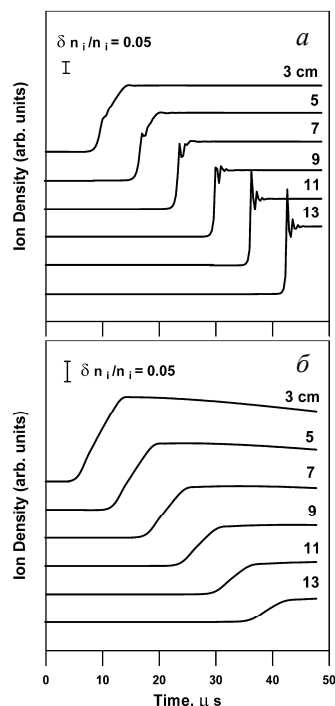


Рис. 3. Теоретическая зависимость от времени концентрации ионов на разных расстояниях от сетки для данных эксперимента [17] и $n_{d0} = 0$ (а); $n_{d0} = 1,46 \cdot 10^4 \text{ см}^{-3}$ (б)

$n_{i0} = 2,3 \cdot 10^8 \text{ см}^{-3}$ (фоновая концентрация ионов бралась фиксированной для всех расчетов; здесь и ниже фоновые величины имеют индекс «0»). Ширина «ступеньки» ($\Delta x \approx 20$ см), характеризующей начальное возмущение в экспериментах [17], и ее форма были получены самосогласованно. Учитывалось также сравнение теоретических и экспериментальных значений разности потенциалов между пылевой частицей и плазмой в зависимости от параметров плазмы в той же установке и практически в тех же экспериментальных условиях, в которых проводились эксперименты [17]. Сравнение показало, что (хотя и имеет место соотношение $T_e \gg T_i$) лучшее соответствие получалось для теоретической кривой, вычисленной при $T_e = T_i$ [26]. Поэтому в вычислениях использовались равные значения T_i и T_e .

На рис. 3, построенном по аналогии с рис. 1, представлена зависимость от времени концентрации ионов на разных расстояниях от сетки для значений $n_{d0} = 0$ (концентрация электронов при этом $n_{e0} = 2,3 \cdot 10^8 \text{ см}^{-3}$) и $n_{d0} = 1,46 \cdot 10^4 \text{ см}^{-3}$ (концентрация электронов при этом $n_{e0} = 4,6 \cdot 10^8 \text{ см}^{-3}$). В расчетах, как и в экспериментах, присутствие пыли подавляет осцилляции в профиле ударной волны, вызванные разделением зарядов (электронов и ионов). Теоретическое значение ширины фронта ударной волны (во временных единицах) ~ 5 мкс соответствует экспериментальным результатам.

Кроме того, рис. 3, а, полученный в рамках модели ионизирующего источника, качественно правильно отражает зарождение в экспериментах осцилляций в профиле ударной волны, вызванных разделением зарядов, что не удается получить, например, в рамках модели, использующей уравнение КдВБ [17]. На рисунке 3, а показано, что на расстоянии 3 см от сет-

ки возникает возмущение на фронте ударной волны, которое на больших расстояниях от сетки развивается и принимает вид осцилляций на фронте, которые связывают с разделением зарядов. Такая картина находится в полном соответствии с экспериментальной. Таким образом, в рамках модели ионизирующего источника удается описать все основные особенности экспериментов [17].

Аналогично вычисления, проводимые на основе модели ионизирующего источника, позволяют объяснить все основные результаты экспериментов [18], выполненных на Q-машине.

1. Пылевые ионно-звуковые ударные волны формируются лишь при достаточно больших значениях концентрации пыли (для данных [18] при $\epsilon Z_{d0} \equiv n_{d0} Z_{d0} / n_{i0} \geq 0,75$).

2. В случае формирования ударно-волновой структуры ширина фронта $\Delta\xi$ соответствует теоретической оценке (1).

3. Имеется существенный рост скорости пылевых ионно-звуковых волн с ростом ϵZ_{d0} .

При вычислениях использовались следующие параметры плазмы: $T_e = T_i = 0,2$ эВ, фоновая концентрация ионов $n_{i0} = 1,024 \cdot 10^7 \text{ см}^{-3}$ бралась одинаковой для всех расчетов, $a = 0,1$ мкм, ширина «ступеньки», характеризующей начальное возмущение, $\Delta x \approx 25$ см, начальное превышение ионной концентрации в «ступеньке» над фоновой (в остальной части установки) (см. рис. 2).

На рисунке 4, построенном по аналогии с рис. 2, представлена зависимость от времени концентрации ионов на разных расстояниях от сетки для значений $\epsilon Z_{d0} = 0$ и $\epsilon Z_{d0} = 0,75$. Жирными линиями показаны профили плотности ионов. Тонкие линии показывают уширение волнового фронта при $\epsilon Z_{d0} = 0$ и рост крутизны при $\epsilon Z_{d0} = 0,75$, что находится в соответствии с экспериментальными данными [18].

Численное значение ширины ударного фронта $\Delta\xi / Mv_s \sim 0,3$ мс (см. рис. 4, б) соответствует экспериментальному (см. рис. 2, б) и полученному аналитически (1) в рамках теоретической модели [13].

Эволюция возмущения происходит таким образом, что приблизительно через 1 мс после начала распространения возмущения по фоновой плазме скорость (фронта) возмущения V_p становится близкой к постоянной [24]. На рисунке 5 приведена зависимость скорости возмущения, нормированной к ее значению в отсутствие пыли ($\epsilon Z_{d0} = 0$), от ϵZ_{d0} . Рисунок построен с использованием экспериментальных точек [18], которые показаны крестиками. Данные вычислений изображены кружками. Видно хорошее соответствие теории и эксперимента.

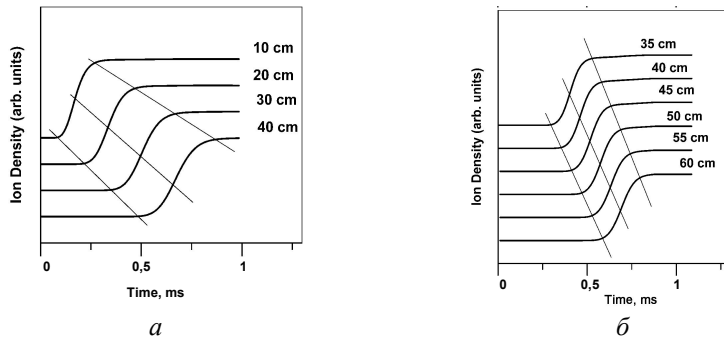


Рис. 4. Теоретическая зависимость от времени концентрации ионов на разных расстояниях от сетки для данных эксперимента [18] и $\epsilon Z_{d0} = 0$ (а); $\epsilon Z_{d0} = 0,75$ (б)

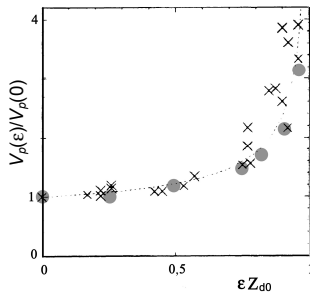


Рис. 5. Зависимость скорости возмущения, нормированной к ее значению в отсутствие пыли, от ϵZ_{d0}

Отметим также, что ударно-волновые структуры в экспериментах [17, 18] слабо затухают, что хорошо видно из теоретических (см. рис. 3, 4) и экспериментальных (см. рис. 1, 2) графиков. Указанное затухание обусловлено диссипацией, связанной с процессами поглощения ионов пылевыми частицами и кулоновских столкновений между ионами и пылевыми частицами. Таким образом, в условиях лабораторных экспериментов [17, 18] не удастся получить квазистационарные пылевые ионно-звуковые нелинейные структуры. Существование таких квазистационарных структур возможно в ситуации, когда присутствует внешнее электромагнитное излучение и фотоэффект приводит к положительным зарядам пылевых частиц. При этом для спектра солнечного излучения в окрестности Земли можно пренебречь эффектами изменения числа ионов и количества их движения из-за их взаимодействия с пылевыми частицами, что обуславливает малость диссипации в пылевой плазме, связанной с этими эффектами. Однако эффекты зарядки пылевых частиц в отмеченных нелинейных структурах в присутствии электромагнитного излучения, как и в рассматриваемых здесь структурах, играют существенную роль.

В настоящее время в природе известно по крайней мере одно явление, в котором оказываются важными ударно-волновые явления, связанные с процессом зарядки пылевых частиц. Это — формирование головной ударной волны при взаимодействии солнечного ветра с пылевой комой кометы [6]. Оказывается, что при достаточно плотной пылевой коме (с концентрацией пыли у поверхности ядра кометы, превосходящей 10^6 см^{-3}) головная ударная волна может быть связана с аномальной диссипацией, обусловленной зарядкой пылевых частиц. При этом ширина $\Delta\xi$ ее фронта удовлетворяет соотношению (1), свойственному для пылевых ионно-звуковых ударных волн.

Солитоны

Как известно, в отсутствие диссипации (или если она мала на характерных временах динамики системы) баланс между нелинейными и дисперсионными эффектами может приводить к образованию симметричной уединенной волны — солитона, представляющего собой нелинейную волну, которая сохраняет свою форму и скорость при движении и взаимодействии с себе подобными уединенными волнами, т. е. является устойчивым образованием. При этом имеет место определенный сдвиг фаз (смещение максимума солитона после взаимодействия по сравнению с индивидуально движущимся солитоном). Как известно, в пылевой плазме происходит видоизменение формы солитоноподобного возмущения и превращение его в ударную волну. Таким образом, вопрос о существовании солитонов в пылевой плазме является не очевидным. Оказывается (см., например, [27—30]), что удастся дать обоснование существованию так называемых слабодиссипативных пылевых ионно-звуковых солитонов сжатия и разрежения. Их форма определяется стационарным «консервативным» солитонным решением (полученным в пренебрежении диссипацией) с соответствующим числом Маха. После взаимодействия друг с другом каждый из солитонов восстанавливает форму солитона, распространяющегося без взаимодействия. Однако в отличие от обычных солитонов их полная энергия и полный импульс уменьшаются со временем. В солитонах сжатия значительной является роль захваченных электронов, следовательно изучение наблюдаемых свойств может использоваться для диагностики микроскопических свойств этих электронов. В случае положительно заряженных пылевых частиц возмущения затухают медленнее и распространяются на большие расстояния, чем в случае отрицательно заряженных пылевых частиц. Кроме того, существенным отличием случая положительно заряженных пылевых частиц от случая отрицательно заряженных пылевых частиц является отсутствие в первом пылевых ионно-звуковых солитонов разрежения.

В лабораторной пылевой плазме исследование солитонов возможно в установках типа двойной плазменный прибор и Q-машина. Полезными

наблюдения за солитонами могут оказаться в ионосферной и космической плазме, что, в частности, представляет интерес для выявления знака заряда и диагностики вещества пылевых частиц.

Модуляционное взаимодействие

В обычной плазме (без пыли) нелинейное модуляционное взаимодействие, связанное с усилением фазовых корреляций и уменьшением стохастичности в системе, является одним из основных процессов, приводящих к самоорганизации, т. е. переходу системы в более упорядоченное состояние [31]. В пылевой плазме имеются другие каналы самоорганизации — например, приводящие к формированию плазменно-пылевого кристалла. Вместе с тем модуляционное взаимодействие остается весьма важным процессом, определяющим свойства плазменной системы и имеющим свои особенности. Так, при развитии модуляционного взаимодействия в пылевой плазме происходит уменьшение микроскопического тока электронов и увеличение микроскопического тока ионов, стекающих на пылевую частицу [32]. Электростатическая энергия, выделяемая в результате уменьшения заряда пылевой частицы, вносит свой вклад в развитие модуляционных процессов, что и приводит к усилению модуляционных процессов по сравнению со случаем плазмы без пыли. Важная особенность модуляционной неустойчивости в пылевой плазме состоит в увеличении диапазона волновых векторов, внутри которого возможен эффективный перенос энергии при развитии модуляционных процессов. Энергия в один шаг (по терминологии так называемой каскадной перекачки энергии [33, 34]) перекачивается в диссипативную область, где важным процессом является затухание Ландау. Происходит подавление возникновения областей сильной нелинейности, которое часто сопровождает развитие модуляционного взаимодействия в плазме без пыли.

Модуляционная неустойчивость в пылевой плазме может развиваться в тех областях плазмы, которые в отсутствие пыли устойчивы относительно раскачки модуляционных возмущений. Примером служит случай запыленной ионосферы. Развитие модуляционной неустойчивости в запыленной ионосфере обусловлено джоулевым нагревом, пондеромоторной силой, а также процессами, связанными с зарядкой пылевых частиц и их динамикой [35]. Возможность развития модуляционной неустойчивости в этом случае связана с пылевыми частицами, поскольку в их отсутствие в плазме ионосферы на высотах 80—120 км вследствие амбиполярной диффузии происходит подавление модуляционных процессов.

В 1999—2001 и 2003 годах проводились эксперименты по изучению излучения ионосферы у поверхности Земли, использующие модуляционный радиометр с маломощным усилителем на вводе [36, 37]. Оказалось, что во время выпадения высокоскоростных метеорных потоков Персеиды,

Леониды, Ориониды, Геминиды, кроме известных типов помех, присутствуют максимумы (так называемые *пылевые* линии) в диапазоне частот 12—50 Гц, которые меняют свою частоту и амплитуду, возле них появляются и исчезают сателлиты, меняется форма пика и т. п. Причиной этого эффекта служит модуляционное взаимодействие электромагнитных и пылевых звуковых волн [38].

Во время выпадения высокоскоростных метеорных потоков Персеиды, Леониды, Ориониды, Геминиды происходит испарение метеороидов на высотах 80—120 км в зависимости от размеров каждого из них и их начальной скорости. Пары метеорного вещества конденсируются, приводя к появлению наномасштабных (пылевых) частиц космического происхождения. В результате нескомпенсированного электронного и ионного токов, а также действия фотоэффекта, возбужденного действием солнечного излучения, частицы приобретают электрические заряды. При распространении электромагнитной волны в запыленной нижней ионосфере Земли в результате модуляционного взаимодействия возбуждаются электростатические низкочастотные колебания с характерными частотами, близкими к частотам пылевых звуковых волн, что может приводить к модуляции электромагнитных волн. Именно низкочастотная составляющая промодулированной волны на фоне ионосферных шумов и фиксируется у поверхности Земли [36, 37]. Наблюдения низкочастотных флуктуаций радишумов ионосферы, включающих устойчивые спектральные *пылевые* линии, могут служить косвенным подтверждением существования пылевой звуковой моды в запыленной ионосфере Земли [38].

Эффект возбуждения специфических низкочастотных флуктуаций радишумов ионосферы во время выпадения высокоскоростных метеорных потоков может сопровождаться генерацией достаточно интенсивных инфразвуковых волн, достигающих поверхности Земли, а также акустико-гравитационных вихрей, которые приводят к усилению относительной интенсивности зеленого излучения ночного неба [38]. Кроме того, в результате развития модуляционного взаимодействия при воздействии на ионосферу электромагнитного излучения от мощных нагревных стенов возможно возбуждение на высотах 80—100 км достаточно интенсивных неоднородностей электронной и ионной концентраций (вплоть до ~ 5 %) [39].

Заключение

В этой статье затронута лишь небольшая часть исследований по нелинейной волновой физике пылевой плазмы. Например, практически незатронутой оказалась проблема фазовых переходов в ударных волнах [40]. Нелинейные волновые процессы представляются важными при взаимодействии хвоста магнитосферы Земли с пылевой плазмой у поверхности Луны [41] и т. д.

Фактически, результаты по нелинейной физике пылевой плазмы (не смотря на их значимость и весьма активное развитие направления физики пылевой плазмы в целом) до сих пор довольно фрагментарны. Не существует систематического описания нелинейных плазменно-пылевых процессов, хотя и имеется весьма развернутое описание (см., например, [16, 31, 42] и ссылки в этих работах) нелинейных процессов в обычной плазме (без пыли), которое, на первый взгляд, могло бы служить примером и для пылевой плазмы. Существуют, по-видимому, две причины такого состояния дел. Прежде всего, следует отметить сложность решения нелинейных плазменно-пылевых проблем по сравнению со случаем обычной плазмы. Вместе с этим создание нелинейной теории обычной плазмы осуществлялось в тот период, когда исследования по физике плазмы находились на подъеме и многие ведущие ученые принимали участие в разработке основ этой науки. При этом развитие нелинейной физики плазмы рассматривалось как одно из центральных плазменных направлений. Сейчас исследования по физике плазмы не столь интенсивны. Надо полагать, что дальнейшее комплексное экспериментальное и теоретическое рассмотрение нелинейных явлений и эффектов позволит более глубоко понять физические процессы, протекающие в пылевой плазме, и природу этого интереснейшего состояния вещества.

Участие автора в работе научной школы «Нелинейные волны — 2016» осуществлено при поддержке в рамках программы фундаментальных исследований президиума РАН «Экспериментальные и теоретические исследования объектов Солнечной системы и планетных систем звезд», а также РФФИ (проект 15-02-05627-а).

Литература

1. *Langmuir, I.* A new type of electric discharge: the streamer discharge / I. Langmuir, C. G. Found, A. F. Dittmer // *Science*. — 1924. — V. 60. — P. 392—394.
2. *Кдумов, Б. А.* Формирование структур в запыленной ионосфере / Б. А. Кдумов, Г. Е. Морфилл, С. И. Попель // *ЖЭТФ*. — 2005. — Т. 127, вып. 1. — С. 171—185.
3. *Дубинский, А. Ю.* Формирование и эволюция плазменно-пылевых структур в ионосфере / А. Ю. Дубинский, С. И. Попель // *Письма в ЖЭТФ*. — 2012. — Т. 96, вып. 1. — С. 22—28.
4. *Альвен, Х.* Эволюция Солнечной системы / Х. Альвен, Г. Аррениус. — М.: Мир, 1979. — 511 с.
5. *Goertz, C. K.* Dusty plasmas in the solar system // *Rev. Geophys.* — 1989. — V. 27, № 2. — P. 271—292.
6. *Popel, S. I.* Charged dust and shock phenomena in the solar system / S. I. Popel, A. A. Gisko // *Nonlin. Processes. Geophys.* — 2006. — V. 13. — P. 223—229.
7. *Popel, S. I.* The effect of microscopic charged particulates in space weather / S. I. Popel, S. I. Kopnin, M. Y. Yu, J. X. Ma, Feng Huang // *J. Phys. D: Applied Phys.* — 2011. — V. 44. — P. 174036-1—174036-7.
8. *Whipple, E. C.* Potentials of surfaces in space // *Rep. Prog. Phys.* — 1981. — V. 44. — P. 1197—1250.

9. Цытович, В. Н. Пыль в установках управляемого термоядерного синтеза / В. Н. Цытович, Дж. Вингер // УФН. — 1998. — Т. 168, вып. 8. — С. 899—907.
10. Chu, J. H. Direct observation of coulomb crystals and liquids in strongly coupled rf dusty plasmas / J. H. Chu, Lin I // Phys. Rev. Lett. — 1994. — V. 72, № 25. — P. 4009—4012.
11. Thomas, H. Plasma crystal: Coulomb crystallization in a dusty plasma / H. Thomas, G. E. Morfill, V. Demmel, J. Goree, B. Feuerbacher, D. Möhlmann // Phys. Rev. Lett. — 1994. — V. 73, № 5. — P. 652—655.
12. Ландау, Л. Д. Гидродинамика / Л. Д. Ландау, Е. М. Лифшиц. — М. : Наука, 1986. — 736 с.
13. Popel, S. I. Shock waves in plasmas containing variable-charge impurities / S. I. Popel, M. Y. Yu, V. N. Tsytovich // Phys. Plasmas. — 1996. — V. 3, № 12. — P. 4313—4315.
14. Popel, S. I. Evolution of perturbation in charge-varying dusty plasmas / S. I. Popel, A. P. Golub', T. V. Losseva, R. Bingham, S. Benkadda // Phys. Plasmas. — 2001. — V. 8, № 5. — P. 1497—1504.
15. Попель, С. И. Формирование ударно-волновых структур в пылевой плазме / С. И. Попель, А. П. Голубь, Т. В. Лосева, Р. Бингхэм, С. Бенкадда // Физика плазмы. — 2001. — Т. 27, № 6. — С. 483—490.
16. Сагдеев, Р. З. Коллективные процессы и ударные волны в разреженной плазме // Вопросы теории плазмы. — М. : Атомиздат, 1964. — Вып. 4. — С. 20—80.
17. Nakamura, Y. Observation of ion-acoustic shocks in a dusty plasma / Y. Nakamura, H. Bailung, P. K. Shukla // Phys. Rev. Lett. — 1999. — V. 83, № 8. — P. 1602—1605.
18. Luo, Q.-Z. Experimental study of shock formation in a dusty plasma / Q.-Z. Luo, N. D'Angelo, R. L. Merlino // Phys. Plasmas. — 1999. — V. 6, № 9. — P. 3455—3458.
19. Nakamura, Y. Experiments on ion-acoustic shock waves in a dusty plasma // Phys. Plasmas. — 2002. — V. 9, № 2. — P. 440—445.
20. Nakamura, Y. Observation of ion-acoustic solitary waves in a dusty plasma / Y. Nakamura, A. Sarma // Phys. Plasmas. — 2001. — V. 8, № 9. — P. 3921—3926.
21. Popel, S. I. Shock structures in plasmas containing variable-charge macro particles / S. I. Popel, V. N. Tsytovich, M. Y. Yu // Astrophys. Space Sci. — 1998. — V. 256, № 1/2. — P. 107—123.
22. Popel, S. I. Shocks in space dusty plasmas / S. I. Popel, V. N. Tsytovich // Astrophys. Space Sci. — 1999. — V. 264, № 1—4. — P. 219—226.
23. Popel, S. I. Formation of shocks related to dust particle charging in complex plasmas / S. I. Popel, A. P. Golub', T. V. Losseva, R. Bingham // Письма в ЖЭТФ. — 2001. — Т. 73, № 5. — С. 258—262.
24. Попель, С. И. Пылевые ионно-звуковые ударно-волновые структуры: теория и лабораторные эксперименты / С. И. Попель, А. П. Голубь, Т. В. Лосева // Письма в ЖЭТФ. — 2001. — Т. 74, № 7. — С. 396—401.
25. Томас, Х. Комплексная плазма. III. Эксперименты по сильной связи и дальним корреляциям / Х. Томас, Г. Е. Морфилл, В. Н. Цытович // Физика плазмы. — 2003. — Т. 29, № 11. — С. 963—1030.
26. Nakamura, Y. A dusty double plasma device / Y. Nakamura, H. Bailung // Rev. Sci. Instrum. — 1999. — V. 70, № 5. — P. 2345—2348.

27. *Popel, S. I.* Weakly-dissipative dust-ion-acoustic solitons / S. I. Popel, A. P. Golub', T. V. Losseva, A. V. Ivlev, S. A. Khrapak, G. Morfill // *Phys. Rev. E.* — 2003. — V. 67, № 5. — P. 056402-1—056402-5.
28. *Losseva, T. V.* Evolution of weakly-dissipative hybrid dust ion-acoustic solitons in complex plasmas / T. V. Losseva, S. I. Popel, A. P. Golub', P. K. Shukla // *Phys. Plasmas.* — 2009. — V. 16, № 9. — P. 093704-1—093704-5.
29. *Losseva, T. V.* Weakly dissipative dust-ion-acoustic solitons in complex plasmas and the effect of electromagnetic radiation / T. V. Losseva, S. I. Popel, A. P. Golub', Yu. N. Izvekova, P. K. Shukla // *Phys. Plasmas.* — 2012. — V. 19, № 1. — P. 013703-1—013703-6.
30. *Лосева, Т. В.* Ионно-звуковые солитоны в пылевой плазме / Т. В. Лосева, С. И. Попель, А. П. Голубь // *Физика плазмы.* — 2012. — Т. 38, № 9. — С. 792—806.
31. *Попель, С. И.* Генерация регулярных полей и модуляционное взаимодействие в природных и технологических плазменных системах / С. И. Попель. — М. : МФТИ, 2009. — 308 с.
32. *Benkadda, S.* Modulational instability in dusty plasmas / S. Benkadda, V. N. Tsytovich // *Phys. Plasmas.* — 1995. — V. 2, № 8. — P. 2970—2974.
33. *Hasegawa, A.* Self-organization processes in continuous media // *Adv. Phys.* — 1985. — V. 34, № 1. — P. 1—42.
34. *Грач, С. М.* Динамические характеристики плазменной турбулентности ионосферы, инициированной воздействием мощного коротковолнового радиоизлучения / С. М. Грач, Е. Н. Сергеев, Е. В. Мишин, А. В. Шиндин // *УФН.* — 2016. — Т. 186, вып. 2. — С. 1189—1228.
35. *Копнин, С. И.* Модуляционное возбуждение низкочастотных пылевых звуковых колебаний в нижней ионосфере Земли / С. И. Копнин, С. И. Попель, Минг Ю // *Физика плазмы.* — 2007. — Т. 33, № 4. — С. 323—336.
36. *Мусатенко, С. И.* Пылевая плазма в среднеширотной ионосфере в периоды метеорных потоков / С. И. Мусатенко, Ю. С. Мусатенко, Е. В. Курочка, А. В. Ласточкин, В. Я. Чолий, О. И. Максименко, А. С. Слипченко // *Геомагнетизм и аэронавигация.* — 2006. — Т. 46, № 2. — С. 182—192.
37. *Мусатенко, С. И.* Фликкер-шум среднеширотной ионосферы в метровом диапазоне 11—12 августа 1999 г. / С. И. Мусатенко, Ю. С. Мусатенко, Е. В. Курочка, М. М. Медведский, А. А. Сухой, А. С. Слипченко, В. Я. Чолий // *Геомагнетизм и аэронавигация.* — 2001. — Т. 41, № 6. — С. 812—819.
38. *Kopnin, S. I.* Phenomena associated with complex (dusty) plasmas in the ionosphere during high-speed meteor showers / S. I. Kopnin, S. I. Popel, M. Y. Yu // *Phys. Plasmas.* — 2009. — V. 16, № 6. — P. 063705-1—063705-7.
39. *Копнин, С. И.* К вопросу о модуляционном возбуждении неоднородностей в плазме запыленной ионосферы / С. И. Копнин, С. И. Попель, Т. И. Морозова // *Физика плазмы.* — 2015. — Т. 41, № 2. — С. 188—194.
40. *Samsonov, D.* Shock melting of a two-dimensional complex (dusty) plasma / D. Samsonov, S. K. Zhdanov, R. A. Quinn, S. I. Popel, G. E. Morfill // *Phys. Rev. Lett.* — 2004. — V. 92, № 25. — P. 255004-1—255004-4.
41. *Попель, С. И.* Волновые процессы при взаимодействии хвоста магнитосферы Земли с пылевой плазмой у поверхности Луны / С. И. Попель, Т. И. Морозова // *Физика плазмы.* — 2017. — Т. 43, № 5. — С. 474—484.
42. *Цытович, В. Н.* Теория турбулентной плазмы / В. Н. Цытович. — М. : Атомиздат, 1971. — 424 с.

Нелинейные процессы в нейросистемах

ПРЕДСКАЗАНИЕ НЕПРЕДСКАЗУЕМОГО: НОВЫЕ ВОЗМОЖНОСТИ ПОДАВЛЕНИЯ АБСАНС-ЭПИЛЕПСИИ С ИСПОЛЬЗОВАНИЕМ ИНТЕРФЕЙСА «МОЗГ — КОМПЬЮТЕР»

*А. Е. Храмов¹, А. А. Короновский², Ж. ван Лютаалар³,
А. Лютжоханн⁴, В. А. Максименко¹, В. А. Макаров¹*

¹ Саратовский государственный технический университет им. Ю. А. Гагарина

² Саратовский государственный университет им. Н. Г. Чернышевского

³ Дондеровский центр сознания

университета им. св. Радбоды Утрехтского, г. Наймеген, Нидерланды

⁴ Институт физиологии Вестфальского университета им. Вильгельма, г. Мюнстер, Германия

Введение

Потенциальная стратегия лечения резистентных к стандартной терапии пациентов с эпилепсией (а их доля превышает 30 % от более чем 50 миллионов человек) — это различные формы электрической стимуляции мозга, осуществляемой либо заранее заданным графиком стимуляции (так называемая разомкнутая (open loop) схема), либо с использованием обратной связи (close-loop-схема) [1—3]. Использование обратной связи помогает автоматически обнаруживать эпилептические припадки, что запускает электрическую стимуляцию головного мозга, которая прерывает эпилептический разряд [4]. До сих пор такое экспериментальное лечение было в состоянии только уменьшить судороги, в то время как конечной целью является их полная отмена [5—6].

Возможным направлением для достижения этой цели является разработка системы стимулирования с обратной связью (интерфейс «мозг — компьютер»), который бы не прерывал, а не допускал судороги [7]. Такой подход возможен только в случае решения проблемы прогнозирования эпилептических приступов, а она, несмотря на некоторые успехи для височной эпилепсии, остается практически нерешенной для абсанс-эпилепсии, отличающейся вторичными генерализованными приступами, формирование которых всегда ранее рассматривалось как внезапное и непредсказуемое [8—11]. Типичные абсансы продолжаются несколько секунд и характеризуются, во-первых, нарушением сознания и, во-вторых, генерализованным, т. е. наблюдающимся во всех электроэнцефалографических отведениях из головного мозга, пик-волновым разрядом с частотой от 3 до 4 Гц (у человека) и 8—12 Гц у животных моделей абсанс-эпилепсии. Пик-волновой разряд можно рассматривать как гиперсинхронное состояние всей таламо-кортикальной сети головного мозга.

Нами был разработан интерфейс «мозг — компьютер», направленный на прогнозирование и предотвращение пик-волновых разрядов — электроэнцефалографического паттерна приступов абсанс-эпилепсии — для

крыс линии WAG/Rij, отличающихся генетической предрасположенностью к абсанс-эпилепсии. Реализованный интерфейс «мозг — компьютер» предсказывал приблизительно 50 % пик-волновых разрядов, а остальные разряды детектировались в самом начале. Итогом применения интерфейса «мозг — компьютер» с замкнутой цепью обратной связи для запуска электрической стимуляции мозга стало достижение 72 % снижения продолжительности судорожной активности.

Интерфейс «мозг — компьютер» для предсказания и подавления эпилептических разрядов

Генерализованные пик-волновые разряды — электроэнцефалографический маркер абсанс-эпилепсии — производятся таламо-кортикальной нейронной сетью [8—14]. Долгое время преобладало мнение, что пик-волновые разряды непредсказуемы и спонтанно формируются на фоновой электроэнцефалограмме (ЭЭГ) [10]. Однако анализ имеющихся наборов данных ЭЭГ с использованием современных радиofизических методов диагностики нестационарных сигналов показал, что имеются некоторые изменения в связи между корой и ядрами таламуса, которые могут быть зафиксированы в интервале не больше 2 с до начала эпилептического разряда [15, 16], а сами пик-волновые разряды предваряются дельта- и альфа / тета-предшественниками в коре головного мозга и таламусе [17]. Здесь обсуждаются новые возможности автоматического предсказания эпилептических событий, которые основаны на оценке степени синхронизации между структурами мозга в режиме реального времени.

Для изучения эффективности предсказания пик-волновых разрядов исследовались ЭЭГ шести крыс линии WAG/Rij, демонстрирующие около нескольких сотен спонтанно возникающих эпилептических разрядов в течение дня, которые показали высокую надежность разработанного подхода [18]. ЭЭГ-сигналы, записанные от двух корковых и таламической структуры 1—3, усиливались многоканальным усилителем и регистрировались аналого-цифровым преобразователем в системе сбора данных WinDAQ, передавались в компьютер, где они анализировались в реальном времени с помощью разработанного алгоритма предсказания, основанного на оценке уровня синхронизации каналов (рис. 1). Всякий раз, когда уровень синхронизации превышал заданное пороговое значение, специальный маркер устанавливался в дополнительном канале интерфейса (4 на рис. 1, а). После детектирования и прогноза возникновения пик-волновых разрядов результаты были проверены квалифицированным электрофизиологом, и в итоге были оценены чувствительность и специфичность алгоритма прогнозирования эпилептических разрядов. Оказалось, что в среднем 45 % пик-волновых разрядов правильно предсказаны, а остальные пик-волновые разряды были детектированы в самом начале (на первой 0,5 с) форми-

рования, при этом ни один пик-волновой разряд не остался недетектированным (рис. 2, *а*). Следует также отметить, что у большинства животных уровень прогноза пик-волновых разрядов был примерно 50—66 % и только два животных демонстрировали уровень ниже 45 % (рис. 2, *в*). Можно предположить, что для этих животных уровень точности предсказания можно повысить, если подобрать параметры алгоритма индивидуально.

Метод предсказания пик-волновых разрядов положен в основу интеллектуальной компоненты разработанного интерфейса «мозг — компьютер» для обработки ЭЭГ-сигналов и стимуляции в реальном времени, представляющей собой последовательность коротких импульсов длительностью 1 с и малой амплитуды, с частотой 130 Гц, которые запускались в момент, когда алгоритм предсказания формировал маркер эпилептического разряда (см. рис. 1, *б*). Ранее было показано, что такой тип стимуляции

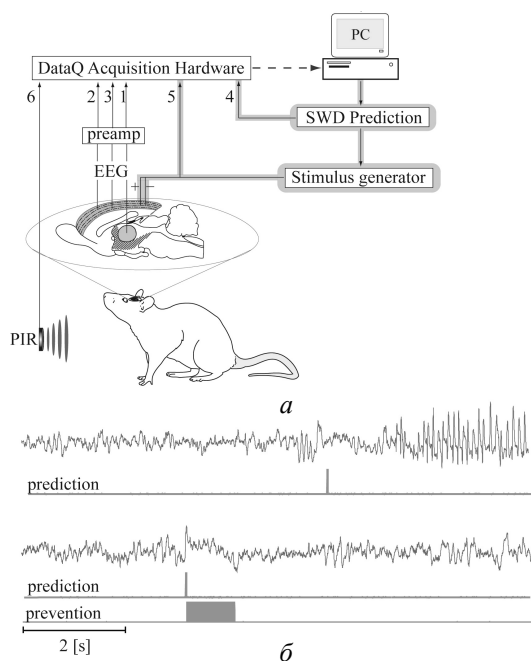


Рис. 1. Схема разработанного и экспериментально проверенного интерфейса «мозг — компьютер» (*а*) и иллюстрация работы интерфейса в случае только предсказания (*вверху*) и предсказания и подавления эпилептических разрядов путем воздействия серией импульсов длительностью 1 с (*внизу*) (*б*). Набор аналоговых входов 1—6 системы сбора данных соответствует записи ЭЭГ сигналов 1—3, маркеру предсказанного пик-волнового разряда 4, последовательности стимулирующих мозг импульсов длительностью 1 с 5 и сигналу от инфракрасного монитора движения (PIR) 6. Штриховая линия соответствует цифровому входу ПК, обратная связь показана линиями с тенью [18]

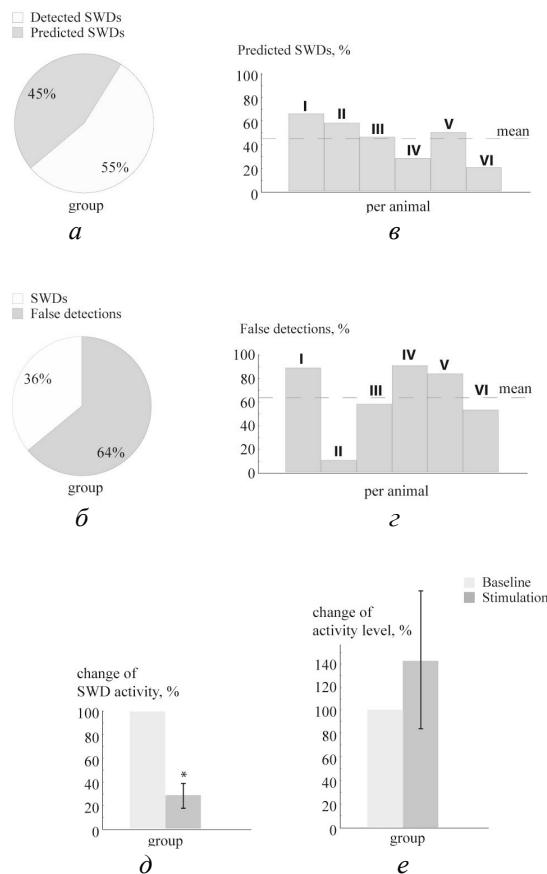


Рис. 2. Средний процент корректно предсказанных и корректно детектированных пик-волновых разрядов (а); средний процент корректно и ошибочно предсказанных пик-волновых разрядов (б); процент корректно и ошибочно предсказанных пик-волновых разрядов для каждого животного (в, з); длительность эпилептических разрядов и уровня активности животных в течение контрольного периода и периода стимуляции, усредненные по группе животных (д, е) [18]

достаточно эффективен и снижает на 90 % длительность пик-волновых разрядов [19]. При сравнении эпилептической активности в контрольный период (длительностью 1 ч), когда стимуляция не проводилась, с работой в течение 1 ч интерфейса, направленной на предсказание и подавление эпилептической активности, было показано, что длительность пик-волновых разрядов в целом снизилась на $(72 \pm 10)\%$ (рис. 2, д). Уменьшение длительности эпилептической активности зависит от комбинации операций «предсказание / предотвращение» (в 45 % случаев) и «детекти-

рование / подавление» пик-волновых разрядов. Обнаруженный эффект уменьшения длительности эпилептической активности не связан с эффективностью метода детектирования событий на ЭЭГ. Так, у двух животных длительность эпилептической активности снизилась соответственно на 98 и 100 %, т. е. благодаря предсказанию и стимуляции мозга удалось полностью избавиться от пик-волновых разрядов.

Однако алгоритм предсказаний демонстрирует большое число ложных срабатываний (рис. 2, б, з). Поэтому можно было бы ожидать, что в течение работы интерфейса «мозг — компьютер» в режиме стимуляции на головной мозг крысы подается большое число дополнительных стимулирующих импульсов, что может привести к изменению поведения животного. Мы сравнивали уровень активности крысы между контрольным периодом и периодом стимуляции с использованием инфракрасного датчика движения (PIR) (канал б на рис. 1). Изменений в активности крысы при стимуляции обнаружено не было (рис. 2, е), т. е. уменьшение длительности эпилептической активности из-за стимуляции нельзя объяснить одновременным увеличением двигательной активности животного. Более того, анализ записей ЭЭГ не показал других аномальных форм активности головного мозга, что позволяет сделать вывод: низкоамплитудная стимуляция в течение коротких интервалов времени может рассматриваться как падающая и не изменяющая поведение животного лечебная стратегия.

Нейрофизиологические эксперименты

Все эксперименты проводились в Дондеровском центре сознания Университета Наймегена им. св. Радбода Утрехтского (Нидерланды) и были одобрены этическим комитетом по экспериментам на животных университета (RU-DEC). В качестве экспериментальных животных использовались самцы крыс линии WAG/Rij в возрасте 6—7 месяцев и весом около 320—370 г [18]. До вживления электродов крысы содержались парами, после — поодиночке в условиях 12:12 светового цикла (темная фаза соответствовала времени 8:30—20:30). Каждому животному были имплантированы 2 кортикальных и один таламический электрод для регистрации ЭЭГ и 2 кортикальных электрода для стимуляции. Точность расположения электродов обеспечивалась использованием гистологического контроля *in vitro* в соответствии с атласом мозга крыс [20].

Все эксперименты проводились на свободно движущихся животных между 9:00 и 16:00. На первом этапе экспериментов записывались контрольные данные, по которым подбирался индивидуальный для каждого животного параметр алгоритма предсказания (пороговый уровень), далее находилась индивидуальная для каждого животного интенсивность стимуляции путем определения минимальной амплитуды импульсов, при которой три подряд идущих пик-волновых разряда прерывались воздейст-

вием импульсного сигнала длительностью 1 с на вживленные электроды. Регистрируемые сигналы ЭЭГ усиливались многоканальным усилителем, фильтровались в диапазоне 1—49,5 и 50,5—100 Гц. Далее аналоговые сигналы оцифровывались системой сбора данных WINDAQ (DATAQ-Instruments Inc., Akron, OH, USA) с постоянной частотой дискретизации 500 Гц и передавались в компьютер. Пассивная система регистрации движения на основе инфракрасного датчика (PIR) записывала двигательную активность животного (RK2000DPC LuNAR PR Ceiling Mount, Rokonet RISCO Group S.A., Drogenbos, BE). Все эксперименты (как контрольные периоды, так и периоды стимуляции) проводились в течении одного дня (длительность каждого периода — 1 час).

Метод предсказания эпилептических разрядов

Алгоритм, оценивающий степень синхронизации между структурами головного мозга и предсказывающий увеличение степени синхронности перед началом пик-волнового разряда, основан как на анализе синхронизации электрической активности нейронов в окрестности одного электрода (локальная синхронизация), так и на анализе синхронизации между нейронными ансамблями коры и таламуса (глобальная синхронизация). Последнее достигается изучением многоканальных записей ЭЭГ из двух слоев коры и таламуса.

Каждый сигнал многоканальной записи рассматривается как макроскопическая характеристика ансамбля взаимодействующих нервных клеток близи окрестности регистрирующего электрода. Полученные временные сигналы ЭЭГ $X_i(t)$ обрабатываются с использованием непрерывного вейвлетного преобразования [21]

$$A_i(s, t) = \frac{1}{\sqrt{s}} \int_{t-s}^{t+s} X_i(t') \varphi^* \left(\frac{t-t'}{s} \right) dt' \quad (1)$$

со специально разработанной материнской функцией:

$$\varphi(\eta) = \pi^{1/4} \exp[i\omega_0 \eta] \exp \left[\frac{-10\eta^4}{2} \right], \quad (2)$$

где $s = 1/f$ — временной масштаб вейвлетного преобразования [с], f — линейная частота сигнала [Гц]. Материнская функция (2) является модификацией хорошо известного вейвлета Морле [21], который характеризуется лучшим временным разрешением, чем стандартный морле-вейвлет, и поэтому более подходит для системы реального времени в связи с лучшим разрешением во времени локальных особенностей сигналов [22].

Вейвлетный спектр мощности $W_i(s, t) = |A_i^2(s, t)|$ в режиме реального времени рассчитывался в окне длительностью 600 мс в частотном диапа-

зоне $\Delta f \in [3, 20]$ Гц. Интеллектуальная система интерфейса с применением рассматриваемого алгоритма предсказания была реализована в среде Borland Delphi 7 с установленным пакетом DATAQ ActiveX Control Library и работала на персональном компьютере (Intel Core2Quad, 4.0 Gb RAM, Win 7, 64 bit). Конфигурация используемого компьютера и программное обеспечение позволили осуществить вейвлетное преобразование в реальном времени с шагом дискретизации $\Delta t = 5 \cdot 10^{-3}$ с, что было достаточно для обеспечения качественной частотно-временной декомпозиции сигнала в интересующем нас частотном диапазоне Δf .

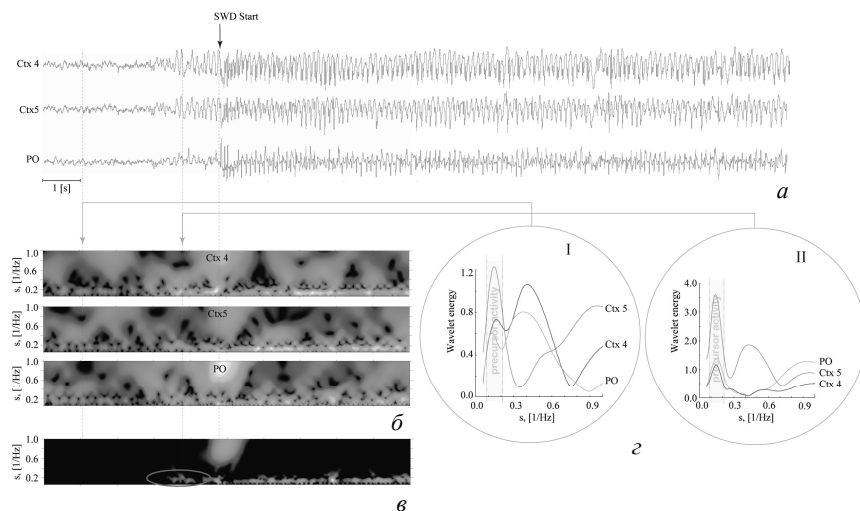


Рис. 3. Многоканальная ЭЭГ с тремя отведениями из двух слоев (слой 4 (Ctx4) и 5 (Ctx5)) соматосенсорной коры и постериорного ядра таламуса (PO) (а) и вейвлетные спектры мощности для этих каналов в координатах «время — временные масштабы» (б). Результирующая поверхность $W(s, t)$ (в); паттерн на вейвлетной поверхности, рассматриваемый как предшественник, выделен овалом. Мгновенные распределения энергии по масштабам, взятые за 4 с (I) и 1 с (II) до начала пикового разряда (г)

Анализ вейвлетных спектров мощности (рис. 3, б) отдельных каналов (рис. 3, а) показал, что первые признаки синхронизации, развивающейся в каждом нейронном ансамбле, появляются за несколько секунд до возникновения пик-волнового разряда и выражаются в локальном увеличении энергии вейвлетного спектра в диапазоне частот 5—10 Гц (см. рис. 3, г, на котором видно, что за 1 с (II) до начала пик-волнового разряда основные спектральные компоненты начинают синхронизоваться). Одновременно с локальной синхронизацией наблюдается увеличение взаимодействия меж-

ду соматосенсорной корой и таламусом, которое удобно анализировать, рассматривая функцию, представляющую собой произведение вейвлетных спектров каждого отдельного канала $W_i(s, t)$:

$$W(s, t) = \prod W_i(s, t). \quad (3)$$

На рис. 3, в показан результат расчета функции (3), на котором хорошо виден изолированный паттерн (выделен овалом), определяемый увеличением степени глобальной синхронизации в таламо-кортикальной сети и рассматриваемый нами как прогностический критерий последующего формирования пик-волнового разряда.

Для автоматического детектирования предшественников пик-волновых разрядов и уменьшения числа ошибочных срабатываний алгоритма при возможном обнаружении осцилляторных паттернов, отличных от предшественников в альфа-диапазоне, рассматривались три частотных интервала, которые соответствуют типичным паттернам синхронной нейронной активности крыс WAG/Rij: Δs_1 (диапазон сонных веретен, 7—20 Гц), Δs_2 (диапазон альфа/тета-предшественников, 5—10 Гц), Δs_3 (диапазон низкочастотных колебаний (дельта-предшественники), 3—5 Гц). Для каждого из этих диапазонов рассчитывалась величина средней энергии $W_{\Delta s_j}(t)$ путем усреднения вейвлетного спектра мощности $W(s, t)$ в соответствующем диапазоне временных масштабов Δs_j :

$$W_{\Delta s_j}(t) = \frac{1}{\Delta s_j} \int_{s \in \Delta s_j} \frac{1}{\tau} \int_{t-\tau}^t W(s, t_0) ds dt_0, \quad j = \overline{1, 3}, \quad (4)$$

где $\tau = 500$ мс было выбрано эмпирически, с учетом минимальной длительности обнаруженного предшественника.

На рис. 4 показаны результаты расчетов усредненных величин в соответствии с выражением (4), откуда видно, что средняя энергия в альфа-диапазоне $W_{\Delta s_2}$, полученная в ходе наблюдения предшественника, становится больше, чем величины $W_{\Delta s_{1,3}}$, и, более того, существенно превышает аналогичную величину $W_{\Delta s_2}$ при регистрации фоновой активности.

Таким образом, вводя индивидуально для каждого животного пороговое значение W_{th} , возможно автоматически детектировать паттерн предшественника пик-волновых разрядов с использованием трех основных условий:

$$\text{а) } W_{\Delta s_2}(t) > W_{th}; \quad \text{б) } W_{\Delta s_2}(t) > W_{\Delta s_1}(t); \quad \text{в) } W_{\Delta s_2}(t) > W_{\Delta s_3}(t).$$

Условия «б» и «в» были введены для точной диагностики отличий предшественника от сонных веретен и низкочастотной дельта-активности.

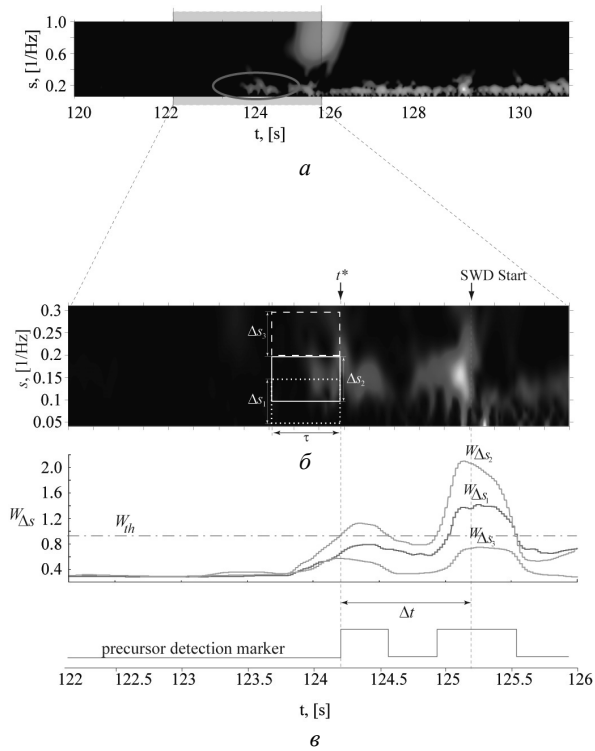


Рис. 4. Вейвлетный спектр мощности $W(s, t)$, отвечающий периоду, предшествующему пик-волновому разряду, и началу эпилептического приступа (а). Увеличенный фрагмент вейвлетного спектра, соответствующего формированию пик-волнового разряда (б). Рамки отмечают области параметров (s, t) , для которых величина $W(s, t)$ усреднялась. Зависимости $W_{\Delta s_i}$ от времени (см. формулу (4)) при усреднении в частотных диапазонах Δs_i (в): W_{th} — пороговое значение энергии, используемое для предсказания пик-волнового разряда; Δt — время между детектированием предшественника и началом эпилептического разряда

Подобно эпилептическому разряду данные типы активности также связаны с синхронной динамикой нейронного ансамбля головного мозга, но характеризуются более высокими (более 12 Гц) и более низкими (менее 5 Гц) частотами соответственно. При выполнении условий «а» — «в» интерфейс «мозг — компьютер» формирует маркер предшественника (показан на рис. 4, в), запускающий стимулирующий импульс, подаваемый в мозг и прерывающий аномальную гиперсинхронную активность, которая приводит к формированию эпилептического разряда.

Заключение

В настоящей статье показано, что, несмотря на общепринятое мнение, пик-волновые эпилептические разряды могут быть с определенной степенью предсказаны, а разработанный алгоритм предсказания эпилептических событий может быть эффективно включен как обеспечивающий, по существу, уменьшение суммарной длительности эпилептической активности интеллектуальный элемент интерфейса «мозг — компьютер», работающего в реальном времени по принципу обратной связи. Данный эффект достигается в комбинации «предсказания — предотвращения» и «детектирования — подавления» пик-волновых эпилептических разрядов. Полученные результаты открывают новые пути для исследований оптимальных параметров стимуляции головного мозга для предотвращения эпилептических приступов, а также развития и оптимизации соответствующего интерфейса «мозг — компьютер» с высокой чувствительностью и малым числом ошибок распознавания предшественников — системы профилактики эпилептических разрядов, которая сможет объединить интеллектуальные системы распознавания и прогнозирования эпилептических приступов с замкнутой системой с обратной связью для глубокой стимуляции мозга.

Работа поддержана РФФИ (грант 15-02-00624).

Литература

1. *Berenyi, A.* Closed-loop control of epilepsy by transcranial electrical stimulation / A. Berenyi, M. Belluscio, D. Mao, G. Buzsaki // *Science*. — 2012. — V. 337. — P. 735—737.
2. *Osorio, I.* Automated seizure abatement in humans using electrical stimulation / I. Osorio [et al.] // *Ann. Neurol.* — 2005. — V. 57. — P. 258—268.
3. *Paz, J. T.* Closed-loop optogenetic control of thalamus as a tool for interrupting seizures after cortical injury / J. T. Paz [et al.] // *Nature Neuroscience*. — 2013. — V. 16. — P. 64—70.
4. *Fridley, J.* Brain Stimulation for the Treatment of Epilepsy / J. Fridley, J. G. Thomas, J. C. Navarro, D. Yoshor // *Neurosurg. Focus*. — 2012. — V. 32, № 3. — Art. e13.
5. *Elliott, J. O.* Comorbidity, health screening, and quality of life among persons with a history of epilepsy / J. O. Elliott, B. Lu, B. Shneker, C. Charyton, J. Layne Moore // *Epilepsy & Behavior*. — 2009. — V. 14. — P. 125—129.
6. *Walker, M. C.* The impact of new antiepileptic drugs on the prognosis of epilepsy: seizure freedom should be the ultimate goal / M. C. Walker, J. W. Sander // *Neurology*. — 1996. — V. 46. — P. 912—914.
7. *Stacey, W. C.* Technology insight: neuroengineering and epilepsy-designing devices for seizure control / W. C. Stacey, B. Litt // *Nat. Clin. Pract. Neurol.* — 2008. — V. 4. — P. 190—201.

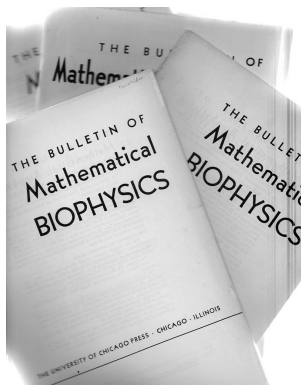
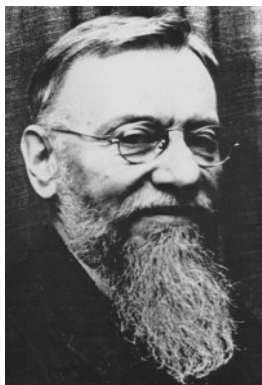
8. *Feldwisch-Drentrup, H.* Joining the benefits: combining epileptic seizure prediction methods / H. Feldwisch-Drentrup [et al.] // *Epilepsia*. — 2010. — V. 51. — P. 1598—1606.
9. *Jobst, B.* Brain stimulation for surgical epilepsy // *Epilepsy Research*. — 2010. — V. 89. — P. 154—161.
10. *Lopes da Silva, F. H. L.* Dynamical diseases of brain systems: different routes to epileptic seizures / F. H. L. Lopes da Silva [et al.] // *IEEE Trans. Biomedical Engineering*. — 2003. — V. 50. — P. 540—548.
11. *Mormann, F.* Seizure prediction: the long and winding road / F. Mormann, R. G. Andrzejak, C. E. Elger, K. Lehnertz // *Brain : a journal of neurology*. — 2007. — V. 130. — P. 314—333.
12. *Blumenfeld, H.* From molecules to networks: cortical/subcortical interactions in the pathophysiology of idiopathic generalized epilepsy // *Epilepsia*. — 2003. — V. 44, Suppl. 2. — P. 7—15.
13. *Huguenard, J. R.* Thalamic synchrony and dynamic regulation of global fore-brain oscillations / J. R. Huguenard, D. A. McCormick // *Trends in Neurosciences*. — 2007. — V. 30. — P. 350—356.
14. *Pinault, D.* Cellular and network mechanisms of genetically-determined absence seizures / D. Pinault, T. J. O'Brien // *Thalamus & Related Systems*. — 2005. — V. 3. — P. 181—203.
15. *Luttjohann, A.* The dynamics of cortico-thalamo-cortical interactions at the transition from pre-ictal to ictal LFPs in absence epilepsy / A. Luttjohann, G. van Luijtelaar // *Neurobiology of Disease*. — 2012. — V. 47. — P. 49—60.
16. *Luttjohann, A.* Dynamics of networks during absence seizure's on- and offset in rodents and man / A. Luttjohann, G. van Luijtelaar // *Frontiers in Physiology*. — 2015. — V. 6. — P. 16.
17. *van Luijtelaar, G.* Spike-wave discharges in WAG/Rij rats are preceded by delta and theta precursor activity in cortex and thalamus / G. van Luijtelaar, A. Hramov, E. Sitnikova, A. Koronovskii // *Clinical Neurophysiology : official journal of the International Federation of Clinical Neurophysiology*. — 2011. — V. 122. — P. 687—695.
18. *Depaulis, A.* Genetic Models of Absence Epilepsy in the Rat / A. Depaulis, G. van Luijtelaar // *Models of Seizures and Epilepsy* / eds A. Pitkänen, P. A. Schwartzkroin, S. L. Moshé. — San Diego [etc.] : Elsevier, Academic Press, 2006. — Chap. 18. — P. 233—248.
19. *van Heukelum, S.* Timing of high-frequency cortical stimulation in a genetic absence model / S. van Heukelum, J. Kelderhuis, P. Janssen, G. van Luijtelaar, A. Luttjohann // *Neuroscience*. — 2016. — V. 324. — P. 191—201.
20. *Paxinos, G.* The Rat Brain in Stereotaxic Coordinates / G. Paxinos, C. Watson. — San Diego : Academic Press, 1998. — 237 p.
21. *Hramov, A. E.* Wavelets in Neuroscience / A. E. Hramov, A. A. Koronovskii, V. A. Makarov, A. N. Pavlov, E. Sitnikova. — Berlin ; Heidelberg : Springer-Verlag, 2015. — 318 p. — (Springer Series in Synergetics).
22. *van Luijtelaar, G.* Methods of automated absence seizure detection, interference by stimulation, and possibilities for prediction in genetic absence models / G. van Luijtelaar, A. Luttjohann, V. V. Makarov, V. A. Maksimenko, A. A. Koronovskii, A. E. Hramov // *J. Neurosci. Methods*. — 2016. — V. 260. — P. 144—158.

НЕКОТОРЫЕ ИСТОРИЧЕСКИЕ АСПЕКТЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ЖИВЫХ СИСТЕМ (ДОКОМПЬЮТЕРНАЯ ЭРА)

Д. И. Трубецков

Саратовский государственный университет им. Н. Г. Чернышевского

Начало статьи выглядит несколько неожиданно: создателем математической биологии, по крайней мере на Западе, считается Николай Рашевский. Фамилия мало кому знакома.



Николай Рашевский (1899—1972) основал Общество математической биологии и стал редактором первого журнала по этой дисциплине [1, с. 12; 2]

Начало XX века. Россия потерпела поражение в Русско-японской войне 1904—1905 годов. 1914 год — начало Первой мировой войны. В стране глубокий экономический кризис и нарастающее социальное недовольство. Грядёт Октябрьская революция. Всё это заставило физика-теоретика украинского происхождения Николая Рашевского покинуть Россию вместе с супругой Эмилией. Сменив несколько государств, в 1924 году они оказались в США, где Рашевский занялся изучением деления клеток в исследовательской лаборатории «Вестингауз». Рассмотрение деления клеток с позиции физики и математики считалось тогда передовым фронтом науки, как любят писать журналисты.

В 1934 году Николай Рашевский (так его называли в США) стал работать в Чикагском университете на должности старшего преподавателя кафедры физиологии. Два дальнейших события, связанных с его деятельностью, имели большое значение для развития математической биологии.

В 1938 году вышла его работа *Mathematical Biophysics*, ставшая знаменитой, поскольку в ней впервые была сделана попытка изложить физи-

ко-математические основы биологии [3]. А в 1939 году Рашевски создал научный журнал *The Bulletin of Mathematical Biophysics* («Вестник математической биофизики»). Первоначальное название журнала позднее сменилось на *The Bulletin of Mathematical Biology* («Вестник математической биологии»). Это первый научный журнал, посвящённый исследованиям в математической биологии. Приведём длинную цитату из книги «Математика жизни. Численные модели в биологии и экологии».

«В последующие годы Рашевски занимался теоретической работой и применил теорию множеств и логику высказываний в исследованиях биологических систем. Он изучал различные общества и способы организации живых существ, а также иерархии, которые они образуют. Сегодня ответы на стоявшие перед ним вопросы кажутся очевидными: рассмотрим, к примеру, последовательность молекулы → клетки → ткани → органы → системы → индивид → популяция. Рашевски создал теорию биологических отношений — реляционную биологию, а также ввёл понятие “множество организмов”. Все эти открытия до недавнего времени оставались незамеченными большинством биологов, которых в основном интересовали полевые исследования или работа в лаборатории. Подлинный размах и возможности теоретического аппарата, терпеливо выстроенного Рашевски, стали очевидны лишь с возникновением так называемой биологии сложных систем. А развитие этой дисциплины, в свою очередь, было бы невозможным без распространения компьютеров.

Наверное, одной из важнейших особенностей первого этапа развития математической биологии, который мы будем дальше называть этапом зарождения биоматематики, стало влияние на неё физики. Это неудивительно, если учесть, что в 1921 году Рашевски преподавал теорию относительности в Праге. Подобно Эйнштейну, посвятившему последние годы жизни работе над “единой теорией”, в 1960-е Рашевски пытался создать единую теорию биологии. Он мечтал выразить на языке математики биологические принципы, описывающие жизнь во всех её проявлениях, будь то растения, животные или микроорганизмы» [1, с. 12—13].

На русском языке, похоже, вышла только одна из книг Рашевского, в которой представлен обзор некоторых математических проблем биологии, тесно связанных с медициной [4]. В книге обсуждаются вопросы оседания частиц в дыхательных путях, движения крови по системе кровообращения, действие лекарственных препаратов, патология желёз внутренней секреции, деятельность центральной нервной системы. Проблемы теоретической биологии рассмотрены в отдельной главе. На мой взгляд, самостоятельный интерес представляет введение к книге, которое имеет название «Цели и методы теоретических исследований». Оно начинается так: «У французов есть поговорка “*Savoir c'est prévoir*” — “знать — значит предвидеть”. Другими словами, знать — значит предсказать. Критерий предсказания можно использовать для оценки системы научных знаний,

полученной на основании фактов или из чисто теоретических рассуждений. Может показаться, что любое знание фактов представляет неоспоримую научную ценность, однако это не всегда так» [4, с. 7]. И далее следует ряд примеров к тезису «не всегда так».

В разделе «Методы теоретического исследования» Рашевский задаётся вопросом: «Каким образом создаётся теория?»

Ещё одна длинная цитата, поскольку уж очень хорошо написано: «В сущности, основой любой теории является совокупность предположений или гипотез относительно интимного, не видимого глазом механизма совокупности исследуемых явлений. Любое явление, будь то физическое или биологическое, если оно достойно теоретического рассмотрения, обычно достаточно сложно, причём в биологии эта сложность больше, чем в физике. Но даже и в физике простота является исключением. Исследователям в ньютоновские времена все явления казались проще, так как все сложности не были ещё известны.

Рассматривая многочисленные примеры из истории науки, можно прийти к выводу, что ни одна теория не описывает явления во всей сложности и что не следует признавать теорию негодной лишь потому, что она лишь приближённо объясняет определённые явления. Из этого вывода вытекает другой, из которого следует, что в наше время, когда техника эксперимента настолько точна, что может дать целый ряд второстепенных данных относительно интересующих нас явлений, теоретик должен умышленно абстрагировать существенные черты явления. Но каким образом он должен узнавать именно существенные черты?

На этот вопрос, вероятно, следует ответить так: способным и преуспевающим теоретиком является именно тот, кто верно угадывает существенное и умело использует абстрактное. Окончательное подтверждение пользы этой абстракции может быть получено иногда через годы и даже десятилетия...

Теория хороша, когда она приводит к заключениям, согласующимся с экспериментальными данными, если она внутренне непротиворечива и если она даёт возможность предсказывать новые факты. Хорошая теория в некоторых отношениях может терять связь с реальностью.

Требование неразрывной связи теории с реальностью по всем пунктам само по себе не имеет связи с реальностью. Это требование противоречит самому духу теоретического мышления» [4, с. 12—14].

Рашевский, как и Шрёдингер, задавался вопросом: «Что такое жизнь?» — но ответа на вопрос не дал, как, впрочем, и все другие после него. В 1972 году он умер от сердечного приступа.

В книге «Математика жизни» сказано, что во времена Рашевского были работы Карла Людвиг фон Берталанфи, носящие прикладной характер [1, с. 13]. Берталанфи родился в Австрии в 1901 году, учился в университетах Инсбрука и Вены, работал в Лондонском университете и различных

канадских институтах. Свою карьеру закончил в Университете штата Нью-Йорк. Умер он внезапно от сердечного приступа в 1972 году, как и Рашевский. Основной вклад Бергаланфи в науку — общая теория систем, но были работы и по математической биологии. Наиболее знаменита его формулировка (в 1938 году) уравнения роста, которое применяется в рыбоводческих хозяйствах. Учёный показал, что размер рыбы $L(t)$ с её возрастом t подчиняется уравнению $L(t) = L_k - (L_k - L_0)e^{-kt}$, где L_k — максимальный размер, k — постоянная роста.

Ещё одним важным событием для математической биологии стало создание Альфредом Джеймсом Лоткой в 1925 году и годом позже Вито Вольтеррой известной модели «хищник — жертва». Это одна из самых прозрачных моделей в математической биологии и, пожалуй, самая популярная модель в экологии. Широта её применения поражает (см., например, [5]).

С момента выхода книги Чарльза Дарвина «Происхождение видов» в центре важнейших тем биологии была эволюция. В 1930-е годы ей уделяется большое внимание и в математической биологии. Под эволюцией понимают физиологические и другие изменения, претерпеваемые живыми существами с течением времени. Согласно Дарвину, живые существа — будь то растения, животные или микроорганизмы — реализуют различные решения задач адаптации к окружающей среде. И чем лучше «решение», которое представляет биологический вид, тем лучше он приспособлен, тем больше его шансы на выживание и, следовательно, на достижение репродуктивного возраста. Если организму удалось размножиться, гены будут переданы следующему поколению. После публикации книги Дарвина в Великобритании возникла английская биометрическая школа, к которой принадлежали такие видные учёные, как Фрэнсис Гальтон и Карл Пирсон, впервые применившие в биологии методы статистики.

В 1930 году Рональд Эйmlер Фишер сформулировал основную теорему естественного отбора: при определённых условиях и за определённое время t ритм, или скорость, с которой повышается средняя приспособленность конкретного вида, равна разнообразию возможных значений генов [1, с. 15—16]: $\Delta W = (\sigma_w)^2 / W$. Здесь ΔW — средний рост приспособленности, W — средняя приспособленность, $(\sigma_w)^2$ — множество возможных значений генов (генных вариаций).

Рональд Эйmlер Фишер родился в Лондоне в 1890 году. Окончил в 1912 году Кембриджский университет, в котором изучал физику и математику. В 1919 году



Рональд Эйmlер Фишер
(1890—1962)

стал сотрудником Ротамстедской экспериментальной станции (до этого он преподавал и занимался статистикой), где столкнулся с проблемой статистической обработки массовых опытов по селекции сельскохозяйственных культур и генетике. Именно тогда он написал свою известную книгу «Методы статистики для научных работников». На протяжении всей жизни Фишер разрабатывал, пропагандировал методы математической статистики и внедрял их в биологию и другие области экспериментальных исследований. Он занимал гальтоновскую кафедру в Лондоне, затем — кафедру генетики в Кембридже. После своей отставки в 1957 году Фишер переехал в Австралию; умер в Аделаиде в 1962 году.

Фишер сформулировал основные концепции популяционной генетики, что было важно для объяснения механизма эволюционного процесса. Его книга «Генетическая теория естественного отбора» (1930) стала классическим сочинением, с которого началось развитие популяционной генетики.

Фишер одновременно с А. Н. Колмогоровым, И. Г. Петровским и Н. С. Пискуновым исследовал уравнение

$$\frac{\partial U}{\partial t} = k(1-U)U + D \frac{\partial^2 U}{\partial x^2}$$

(частный случай уравнения Колмогорова — Петровского — Пискунова).

Имя Эрвина Шрёдингера уже упоминалось в связи с попытками Николая Рашевского ответить на вопрос: «Что такое жизнь?» Но упоминания мало, поскольку книга Э. Шрёдингера [6] занимает особое место в науке. В ней Шрёдингер попытался прояснить с точки зрения физики вопросы, интересовавшие человечество с древних времён.

«Большой, важный и очень часто обсуждаемый вопрос состоит в следующем: как физика и химия смогут объяснять те явления в пространстве и времени, которые происходят внутри живого организма?» [6, с. 11].

Кратко о самом Эрвине Шрёдингере [7, 8].



Эрвин Шрёдингер (1887—1961)

12 августа 1887 года в семье Рудольфа Шрёдингера и Георгины Бауэр появился единственный сын — Эрвин. Родился он в Вене, где и учился в гимназии, затем в университете, который окончил в 1910 году, получив степень доктора после защиты диссертации на тему «Электрическая проводимость на поверхности изоляторов во влажном воздухе».

В 1926 году Шрёдингер сформулировал знаменитое уравнение, которое носит его имя, и в шести статьях обосновал

волновую версию квантовой механики, а также доказал эквивалентность своей теории и матричной механики Гейзенберга. В ноябре 1933 года ему была присуждена Нобелевская премия, которую он разделил с Полем Дираком. Любопытно, что Поль Дирак так писал о Шрёдингере: «Из всех физиков, которых я знаю, мне кажется, Шрёдингер больше всех похож на меня. ...Думаю, это потому, что мы оба безнадежно увлечены математической красотой, которая и определяет нашу работу» (цит. по [1, с. 12]).

В этом отношении интересны высказывания двух великих учёных.

Э. Шрёдингер: «Я пропустил красоту вперёд науки».

П. Дирак: «Физический закон должен обладать математическим изяществом» («The physical law should have the mathematical beauty»).

В 1935 году Шрёдингер опубликовал статью «Современное состояние квантовой механики»; в ней он описал «парадокс кота» и ввёл термин «запутанность» (см., например, [8, гл. 4]). Нас больше всего интересует 1944 год, в котором Шрёдингер опубликовал одну из наиболее важных работ ушедшего XX века — книгу «Что такое жизнь? Физический аспект живой клетки» [6]. Умер Эрвин Шрёдингер в Вене 4 января 1961 года.

Из изложенного выше может сложиться впечатление о гладкой, успешной жизни Шрёдингера, но это не так. Приведём начало введения к книге «Шрёдингер. Квантовые парадоксы. На волне Вселенной»: «Эрвин Шрёдингер — видный представитель исчезнувшего сегодня вида. Он носитель великой культуры Центральной Европы, которая долгое время развивалась на берегах Дуная и была рассеяна в результате мировых войн и деятельности политических экстремистов прошедшего века. Учёный жил и работал в те же годы, что и Фрейд, Климт, Шёнберг, Витгенштейн — в Вене, а Манн, Гросс и Брехт — в Берлине. Он словно впитывал в себя яркий свет современников, оставаясь в их тени. Шрёдингер участвовал в Первой мировой войне в звании лейтенанта артиллерии, знал нищету и невзгоды послевоенного времени, своими глазами видел, как изгоняли евреев, и дважды становился эмигрантом: сначала — когда Гитлер пришёл к власти в Германии, а затем — после аннексии Австрии. Он наблюдал закат двух великих эпох, сформировавших его, — исторической и научной (эпохи классической физики).

Единственный сын в богатой семье, он получил образование и с детства мог воспользоваться всеми преимуществами, которые дают человеку деньги и просвещенное окружение. Однако первоначальный комфорт сменили значительные трудности, в которых Шрёдингер проявил себя как страстная, увлечённая личность, не чуждая и противоречий. Мятежный характер учёного помог ему строить семейную жизнь, отойдя от лекал традиционного буржуазного брака, однако в области политики и особенно науки он проявил себя как убеждённый консерватор. < ... > «В отношениях с окружающими учёный проявлял себя как опытный обольститель, это в равной мере справедливо и для женщин, и для студенческой аудитории, и для слушателей его докладов на конференциях. Однако обаяние учёного

порой приносило ему одни только неприятности, и его бурная личная жизнь иногда становилась уж слишком наэлектризованной. Известно, что у Шрёдингера было три внебрачные дочери от трёх разных женщин и при этом один брак на всю жизнь. Однажды он приехал в Оксфорд сразу с двумя партнёрами, и это событие вызвало настоящий скандал. И всё же, несмотря на весь свой авантюризм и настойчивые ухаживания за некоторыми женщинами, Шрёдингер вовсе не был донжуаном. Он просто любил саму любовь, для него эта вспышка страсти была главной движущей силой, которая стимулировала его прогресс в исследованиях, — по крайней мере, так считал сам учёный. И действительно, его постоянство и непостоянство наложило отпечаток и на научную карьеру: исследователь всё время стремился к новым горизонтам» [8, с. 8—9].

Широта исследований Шрёдингера поражает: диэлектрические материалы, магнетизм, термодинамика, спектроскопия, квантовая механика, общая теория относительности, радиоактивность, космические лучи, поверхностное натяжение, сверхпроводимость, учение о цвете, единые теории поля. Он поставил также множество экспериментов. При этом он работал один (в этом он был подобен Дираку).

«В научной деятельности, как и вообще в жизни, я никогда не придерживался какой-либо генеральной линии, не следовал руководящей программе, рассчитанной на длительные сроки. Хотя я очень плохо умею работать в коллективе, в том числе, к сожалению, с учениками, мои труды никогда не были совершенно самостоятельными, поскольку мой интерес к какому-либо вопросу всегда зависит от интереса, проявляемого к этому же вопросу другими» [8, с. 11].

Но вернёмся к книге «Что такое жизнь? Физический аспект живой клетки», которая продолжает играть важную роль в современной науке. К проблемам, затронутым в книге, внимание Шрёдингера привлекли работы Макса Дельбрюка, который разработал модель гена, основанную на предположении о квантовой природе элементарного процесса передачи наследственных признаков. Модель объясняла устойчивость и изменчивость молекулярных сил. Шрёдингер, используя эти идеи, попытался распространить их на живую клетку, в деятельности которой главную роль играют гены. Он рассматривал вопрос о тесной связи мутаций и отбора с квантовыми скачками (вопрос о тесной связи фундаментальных биологических и физических явлений) и хотел количественно описать хорошо известные биологам мутации, т. е. найти формулы для относительных частот различного рода мутаций.

Ю. А. Данилов в предисловии к третьему русскому изданию книги пишет следующее: «Идеи Э. Шрёдингера условно можно отнести к двум главным темам — “порядок из порядка” и “порядок из беспорядка”».

К первому кругу идей относятся размышления автора о том, каким образом организмы передают информацию из поколения в поколение, как

ген, имеющий микроскопически малые размеры, противостоит тепловым флуктуациям и удерживает наследственную информацию (шрёдингеровская модель аperiodического кристалла, в структуре которого изолирован “план” организма, перекликается с современными представлениями о мозаиках Пенроуза и квазикристаллах).

Полвека развития молекулярной биологии, неравновесной термодинамики и нелинейной физики подтвердили пророческий характер и плодотворность идей Шрёдингера, относящихся к проблеме “порядок из беспорядка”».

Остановимся немного подробнее на двух главных темах, указанных выше, следуя тексту книги Шрёдингера и комментируя отдельные места.

Концепция аperiodического кристалла сыграла важную роль в биологии, поскольку она используется в формулировке принципа хранения генетического кода. В разделе 46 «Аperiodическое твёрдое тело» Шрёдингер пишет следующее: «Небольшую молекулу можно назвать “зародышем твёрдого тела”. Исходя из такого твёрдого зародыша, очевидно, возможно представить себе два различных пути построения всё больших и больших ассоциаций. Один — это сравнительно однообразный путь повторения снова и снова одной и той же структуры в трёх направлениях. Так растёт кристалл. Если периодичность установилась, то уже нет определённого предела для размера такого агрегата. Другой путь — построение всё более и более увеличивающегося агрегата без скучного механизма повторения. Это случай всё более и более сложной органической молекулы, в которой каждый атом, каждая группа атомов играет индивидуальную роль, не вполне равнозначную роли других атомов и групп атомов. Мы можем совершенно точно назвать это образование аperiodическим кристаллом или твёрдым телом и выразить нашу гипотезу словами: мы полагаем, что ген, или, возможно, целая хромосомная нить представляет собой аperiodическое твёрдое тело» [6, с. 66—67].

Раздел 47 книги носит название «Разнообразное содержание, сжатое до миниатюрного кода» и начинается так: «Часто задают вопрос, как такая крошечная частичка вещества — ядро оплодотворённого яйца — может вместить сложный шифровальный код, включающий в себя всё будущее развитие организма? Хорошо упорядоченная ассоциация атомов, наделённая достаточной устойчивостью для длительного сохранения своей упорядоченности, представляется единственно мыслимой материальной структурой, в которой разнообразие возможных (“изомерных”) комбинаций достаточно велико, чтобы заключать в себе сложную систему детерминации в пределах минимального пространства» [6, с. 67].

На простом числовом примере, основанном на азбуке Морзе, Шрёдингер поясняет, что даже при небольшом числе элементарных «кирпичиков» можно получить почти неограниченное число возможных расположений. Он пишет: «Этим мы хотели проиллюстрировать, что, представив ген в

виде молекулы, мы не можем считать немислимым точное соответствие миниатюрного шифровального кода чрезвычайно сложному, специфическому плану развития организма. Мы не можем также считать немислимым и содержание в нём факторов, реализующих этот план» [6, с. 67].

Во времена Шрёдингера уже существовало представление о генах как о больших молекулах. Отсюда его определение аperiodического кристалла как составляющей структуры молекулы, несущей наследственную информацию. Когда такие структуры повторяются в наследственной молекуле, то они напоминают некий алфавит — код, несущий генетическую информацию. Но тогда можно принять термин «аperiodический кристалл» как описывающий двойную спираль.

Как удачно указано в книге «Шрёдингер. Квантовые парадоксы. На волне Вселенной»: «Конфигурация молекулы не только включала инструкции для организма, программу жизни, но и способствовала исполнению этой программы: “Это одновременно и проект архитектора, и рабочие руки строителя”» [8, с. 146].

Структуру ДНК, как известно, открыли Розалинда Франклин, Морис Уилкинс, Джеймс Уотсон и Фрэнсис Крик. Некоторых из них на эти исследования толкнула книга Шрёдингера «Что такое жизнь?». В частности, Нобелевский лауреат Джеймс Уотсон писал впоследствии: «Мне понравилась идея, согласно которой жизнь могла продолжиться благодаря учебнику с инструкциями, записанными в виде конкретного кода» (цит. по [8, с. 146]).

Шрёдингер в своей книге одним из первых описал открытые системы, показав, что в них могут происходить процессы самоорганизации и самовоспроизведения. Эти идеи, высказанные в середине сороковых годов, были предтечей современной теории образования структур в живых системах. Вот как звучит его определение жизни: «Жизнь — это упорядоченное и закономерное поведение материи, основанное не только на одной тенденции переходить от упорядоченности к неупорядоченности, но и частично на существовании упорядоченности, которая поддерживается всё время» [6, с. 73].

Дать ответ на вопрос «Что такое жизнь?» пытались многие (см., например, [9]). Приведём два примера, которые хорошо соответствуют времени их написания.

Первый пример — статья академика В. Гольданского [10] (обратите внимание, что она опубликована в журнале «Коммунист» в разделе «Горизонты науки» в 1986 году). Начинается статья так: «Знаменитое определение Энгельса: “Жизнь есть способ существования белковых тел, существенным моментом которого является постоянный обмен веществ с окружающей их внешней природой, причём с прекращением этого обмена веществ прекращается и жизнь...” (К. Маркс и Ф. Энгельс. Соч. Т. 20, с. 616) — намного опередило своё время, указав на самое принципиальное в характеристике сущности живого... Прекрасно понимая, что наука

о сущности жизни делала ещё только первые шаги, он писал: «Наша дефиниция жизни, разумеется, весьма недостаточна, поскольку она далека от того, чтобы охватить все явления жизни...» (Там же, с. 84)».

Далее во вводной части статьи В. Гольданского изложены краткие сведения об основных молекулах жизни — биополимерах: белках (полипептидах) и ДНК и РНК (полинуклеотидах). Указано, что основной признак жизни — самовоспроизведение, самообновление белковых тел, в основе которых лежит саморепликация, состоящая в удвоении молекулы ДНК с передачей рождающейся клетке генетической информации. Наконец, следует определение: «...жизнь есть форма существования биополимерных тел (систем), способных к саморепликации в условиях постоянно-го обмена веществом и энергией с окружающей средой».

Второй пример — статья [11], опубликованная в 2010 году, содержит следующее определение: «Жизнь с точки зрения физики можно определить как результат процесса игры взаимодействий материи, в которой одна её часть приобрела свойство запоминать вероятности появления удач и неудач в предыдущих раундах этой игры, повышая тем самым свой шанс на существование в последующих раундах. Эту часть материи и называют сегодня живой природой».

Удивительно современно в книге Шрёдингера выглядит содержание главы VI «Упорядоченность, неупорядоченность, энтропия», которая содержит следующие разделы: 56. Упорядоченность, основанная на «упорядоченности»; 57. Живая материя избегает перехода к равновесию; 58. Питание «отрицательной энтропией»; 59. Что такое энтропия? 60. Статистическое значение энтропии; 61. Организация, поддерживаемая извлечением упорядоченности из окружающей среды [6, с. 72—78].

Закончим рассказ об Эрвине Шрёдингере цитатой: «Хотя Шрёдингер не основал ни одной школы и не собрал вокруг себя последователей, он написал одно из самых вдохновляющих научных произведений XX века — сборник “Что такое жизнь?”, куда вошёл цикл его лекций, прочитанных в Тринити-колледже в Дублине в 1943 году. Эта книга убедила целое поколение учёных в том, что физика содержит уникальные возможности для изучения живых существ. Шрёдингер предвидел структурные особенности, выражающиеся в том, что наследственность связана с хромосомами, и вывел современную концепцию генетического кода».

Одна из наиболее выдающихся частей наследия Шрёдингера имеет отношение к его языку, к его способности находить для описания экспериментальных ситуаций яркие образы, которые сразу же подхватывают даже его научные оппоненты. Например, все знают о коте, который носит имя учёного и стал символом загадок в квантовой механике» [8, с. 11].

Пожалуй, изложенное охватывает все основные события в математической биологии до конца Второй мировой войны.

Ситуация меняется с появлением компьютеров. Возникает вычислительная биоматематика [1, с. 15—22].

В 1951 году в университете Манчестера был создан компьютер Ferranti Mark I, на котором начал работать Алан Тьюринг, используя его для математического моделирования биологических задач. Его интересовали две задачи морфогенеза (образования форм):

1) Как объяснить, почему живые организмы обретают конечную форму (почему ветви деревьев обретают именно такую структуру, почему членистоногие состоят как бы из отдельных кусочков, а кольчатые черви — из колец)?

2) Почему на коже некоторых позвоночных есть узоры (полосы у зебр, круглые пятна у леопарда)?

Тьюринг предложил первую математическую теорию морфогенеза, используя для анализа формы листьев последовательность Фибоначчи $a_n = a_{n-1} + a_{n-2}$. Если принять первое число Фибоначчи a_1 за 0, a_2 положить равным 1, то последовательность будет такой: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144 и т. д. Она описывает количество лепестков цветов, расположение листьев растений и чешуек шишек. Интересно, что если лист соперничает с другими за доступ к солнечному свету, то оптимальное расположение листьев даёт последовательность Фибоначчи. Такая особенность растений получила название филлотаксиса Фибоначчи. В ставшей классической работе [12] Тьюринг показал, что периодическое в пространстве и стационарное во времени распределение концентраций может установиться в первоначально однородной системе, где химические реакции сочетаются с диффузией.

1950—1960-е годы — время проводимых с помощью компьютеров математических исследований живых существ и жизни в предположении, что растения, животные и микроорганизмы находятся в стационарном состоянии, которое возможно благодаря механизмам саморегуляции, или гомеостаза. Изучались линейные системы. Их изучение было связано с уже упоминавшимся системным анализом Карла Людвиг фон Берталанфи (в 1968 году он опубликовал книгу *General system theory: foundations, development, applications*). По Берталанфи, система есть множество реально существующих объектов (частей или элементов системы), связанных и взаимодействующих между собой.

«Важный момент теории систем фон Берталанфи заключается в том, что части системы взаимодействуют между собой, а сами системы являются незамкнутыми и взаимодействуют с окружающей средой. При этом из среды в систему поступает входная информация, результатом обработки или преобразования которой является ответ системы, или выходная информация, поступающая обратно в среду. Такие понятия, как саморегулирование и обратная связь, баланс и гомеостаз, в этой модели возникают естественным образом» [1, с. 24].

Заметим, что эти годы — годы кибернетики. Любопытно упомянуть, что в это же время Джон фон Нейман и Алан Тьюринг писали: «Жизнь — это информация».

1970-е годы — годы Ильи Романовича Пригожина, время теории диссипативных структур — систем, непрерывно обменивающихся с окружающей средой и функционирующих благодаря тому, что они далеки от равновесного состояния.

1980—1990-е годы — годы изучения неравновесных, нелинейных систем, особенно динамического хаоса в них. Сошлёмся опять на книгу «Математика жизни»: «В середине 1980-х учёные объединили нелинейные, хаотические и диссипативные системы в одно целое — сложные системы, изучению которых в биологии уделяется наибольшее внимание. К таким системам относятся, например, муравейники, мозг, иммунная система, клетка, морфогенез и экосистемы» [1, с. 28—29].

Появляются и эвристические методы изучения сложных систем, например метод клеточных автоматов.

На этом закончим краткое описание лишь некоторых исторических аспектов математического моделирования живых систем. И вновь уместна цитата: «Обратите внимание, что не только зарождение, но и последующее развитие этой дисциплины неизменно находилось под большим влиянием преобладающих на этот момент физических интерпретаций жизни... Сегодня одним из самых важных достижений математической биологии являются математические модели, позволяющие проводить с помощью компьютера сложные эксперименты» [1, с. 29].

Поэтому в конце статьи укажем на несколько моделей, которые описывают явления, чрезвычайно важные для человечества, но о которых не часто упоминают в литературе по нелинейной динамике (см. [1, с. 58; 8, с. 210; 13]). К ним относится математическое изучение раковых заболеваний, СПИДа, свиного гриппа и различных эпидемий (см. [1, с. 62—64; 4, § 10; 15, с. 344]).

Работа выполнена при поддержке гранта президента РФ для ведущих научных школ России (грант НШ-828.2014.2) и грантов РФФИ № 14-02-00329, 16-02-00238, 17-02-00061.

Литература

1. *Лаос-Бельтра, Р.* Математика жизни. Численные модели в биологии и экологии / Р. Лаос-Бельтра ; пер. с исп. — М. : DeAgostini, 2014. — 164 с. — (Мир математики : в 45 т. ; т. 28).

2. Physics & Maths Catalog [Электронный ресурс] // JF Ptak Science Books : Blog Bookstore / J. F. Ptak. — [USA, 2008—]. — Режим доступа: <http://longstreet.typepad.com/books/2014/05/catalog-1-2014-physics-math.html>. — Загл. с экрана. — (Дата обращения: 01.03.17).

3. *Rashevsky, N.* Mathematical Biophysics : Physicomathematical Foundations of Biology / N. Rashevsky. — Chicago : University of Chicago Press, 1938. — 340 p.
4. *Рашевски, Н.* Некоторые медицинские аспекты математической биологии : пер. с англ. / Н. Рашевски ; под ред. и с предисл. акад. В. В. Парина. — М. : Медицина, 1966. — 243 с.
5. *Трубецков, Д. И.* Феномен математической модели Лотки — Вольтерры и сходных с ней // Изв. вузов. Приклад. нелинейн. динамика. — 2011. — Т. 19, № 2. — С. 69—88.
6. *Шрёдингер, Э.* Что такое жизнь? Физический аспект живой клетки / Э. Шрёдингер. — Ижевск : Ижев. респ. тип., 1999. — 96 с.
7. *Хоффман, Д.* Эрвин Шрёдингер : пер. с нем. / Д. Хоффман. — М. : Мир, 1987. — 96 с.
8. Шрёдингер. Квантовые парадоксы. На волне Вселенной : пер. с фр. — М. : DeAgostini, 2015. — 168 с. — (Наука. Величайшие теории ; вып. 5).
9. *Реутов, В. П.* Как в XX веке физики, химики и биологи отвечали на вопрос: «Что есть жизнь?» / В. П. Реутов, А. Н. Шехтер // УФН. — 2010. — Т. 180. — С. 393.
10. *Гольданский, В.* Возникновение жизни с точки зрения физики // Коммунист. — 1986. — № 1. — С. 86.
11. *Иваницкий, Г. Р.* XXI век: что такое жизнь с точки зрения физики // УФН. — 2010. — Т. 180. — С. 337.
12. *Turing, A. M.* Chemical foundations of morphogenesis // Philos. Trans. Roy. Soc. London, Ser. B : Biological Sciences. — 1952. — V. 237, № 641. — P. 37.
13. *Эрроусмит, Д.* Обыкновенные дифференциальные уравнения. Качественная теория с приложениями / Д. Эрроусмит, К. Плейс. — М. : Мир, 1986. — 243 с.
14. *Трубецков, Д. И.* Введение в синергетику. Хаос и структуры / Д. И. Трубецков. — М. : Книжный дом «ЛИБРОКОМ», 2014. — 240 с.
15. *Ильин, И. В.* Моделирование нелинейной динамики глобальных процессов / под ред. И. В. Ильина, Д. И. Трубецкова. — М. : Изд-во Моск. ун-та, 2010. — 412 с.

**Современные проблемы
теории нелинейных колебаний
и волн**

СВЯЗАННЫЕ ХИМИЧЕСКИЕ ОСЦИЛЛЯТОРЫ

В. К. Ванаг

Балтийский федеральный университет им. И. Канта, г. Калининград

Изучение двух связанных осцилляторов началось с Х. Гюйгенса, который заметил синхронные колебания маятников двух часов, висящих рядом на одной стене [1]. Интерес к синхронизации подобных механических систем не исчез и в наши дни, спустя 340 лет после открытия Х. Гюйгенса [2]. Этот «механический» интерес стал частным случаем общего интереса к эффектам синхронизации осцилляторов различной природы: химических, физических и биологических [3]. В качестве примеров феномена синхронизации можно привести коллективную динамику нейронов мозга человека, согласованные во времени световые вспышки многих тысяч светлячков, которые обычно собираются в огромных количествах на берегах рек в Южной Азии [4], синхронное деление клеток на ранних стадиях эмбрионального развития, резонансы в радиотехнике, подстройку частот аплодисментов в концертном зале, которая приводит к овациям, и многое другое.

В данном обзоре мы будем говорить о синхронизации химических осцилляторов, наиболее ярким представителем которых является реакция Белоусова — Жаботинского (БЖ) [5]. Ее механизм, получивший название механизма ФКН (Филда — Кёреша — Нойеса) или FKN (Field — Körös — Noyes) [6], хорошо известен. Эта реакция может идти в закрытом реакторе несколько часов с периодом колебаний около одной минуты. За БЖ-реакцией удобно следить как оптическими методами, так и при помощи электродов (Pt-электрод или ион-селективный электрод). Исходными реагентами БЖ-реакции являются бромат (BrO_3^-) и малоновая кислота (МА) — $\text{CH}_2(\text{COOH})_2$; реакция между ними происходит в кислой среде (обычно для подкисления берется серная кислота) в присутствии катализатора, в качестве которого часто используется ферроин ($\text{Fe}^{2+}(\text{phen})_3$) или металлокомплексы рутения ($\text{Ru}^{2+}(\text{bpy})_3$). Упрощенная модель БЖ-реакции, ортогонализатор [7], записывается так:



где $\text{A} = \text{BrO}_3^-$, $\text{X} = \text{HBrO}_2$ — активатор, т. е. некая частица, которая автокаталитически размножается (в реакции (3)), $\text{Y} = \text{Br}^-$ — ингибитор, частица, которая препятствует автокатализу (в реакции (2)) ингибитор уничтожает активатор, Z — окисленная форма катализатора, например ферриин

(если в качестве катализатора используется ферроин), $P = \text{HOBr}$ — продукт реакции; B — броммалоновая кислота, которая образуется по ходу реакции из малоновой кислоты. Стехиометрический коэффициент g указывает на то, сколько ионов бромидов образуется в реакции (5), которая играет роль отрицательной обратной связи (производство ингибитора) (обычно g меняется от 0,5 до 1). Реакция (3) — это положительная обратная связь, реакция (4) — это обрыв автокатализа, а реакция (1) — это инициация молекул активатора.

Модель (1)–(5) описывается системой дифференциальных уравнений с двумя переменными, $x = [X]$ и $z = [Z]$:

$$\varepsilon dx/dt = x - x^2 - fz(x - q)/(x + q), \quad (6)$$

$$dz/dt = x - z. \quad (7)$$

Быстрая переменная $y = [Y]$ исключена через квазиравновесие, т. е. скорость изменения концентрации частиц Y (которая составляется из реакций (1)–(5)) полагается равной нулю. Три параметра, $\varepsilon \ll 1$, $f = 2g$ и $q \ll 1$, определяют все динамические режимы, включая и релаксационные колебания. Модель (6)–(7) называют моделью орегонатора, которая качественно (но никак не количественно) описывает колебания в реакции БЖ.

Реакцию БЖ можно проводить в качественно разных реакторах. Наиболее известен макрореактор, представляющий собой бюкс (или кювету) средних размеров, от 3 до 30 мл, в который равномерно подаются все реагенты. Смесь непрерывно перемешивается (магнитной мешалкой), и работает вакуумный насос, откачивающий избыток жидкости, чтобы поддерживать постоянный объем реакционной смеси. Это так называемый ПРПП (проточный реактор постоянного перемешивания). Такие макрореакторы используются для изучения динамики небольшого числа связанных осцилляторов [8–10].

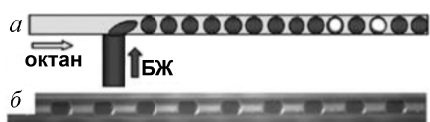
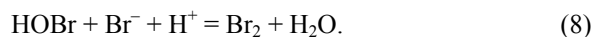


Рис. 1. Схематичное (а) и реальное (б) представление капилляра с БЖ-микрокапельками. Белые капли в схематичном представлении находятся в окисленном состоянии

Относительно недавно были придуманы микрореакторы для БЖ-реакции. Это микрокапельки воды (или водного раствора с БЖ-реакцией), образующиеся при смешивании в микроканалах (в англоязычной литературе говорят о *microfluidic devices* или *microfluidic chips*) водной фазы с несмешивающейся жидкостью (маслом) — так называемые БЖ-капельки (рис. 1) [11–13]. Размер БЖ-капелек можно варьировать в диапазоне от нескольких десятков до сотен микрометров. Ансамбль БЖ-капелек (как в одномерной, так и в двумерной конфигурациях) идеально подходит для исследования синхронных режимов большого числа связанных осцилляторов.

Линейный размер реактора можно уменьшить и до размеров порядка единиц нанометров, если проводить БЖ-реакцию в водных нанокляпях обращенной микроэмульсии АОТ (коммерческое название широко применяемого поверхностно-активного вещества для приготовления микроэмульсии «вода в масле») [14]. Однако в этом случае индивидуальность отдельного реактора теряется из-за интенсивного массообмена между нанокляпями. БЖ-АОТ-систему можно рассматривать как непрерывную (микрорегетогенную) среду с существенно различными коэффициентами диффузии активатора и ингибитора, причем в качестве последнего в данном случае выступает Br_2 , не присутствующий в схеме орегонатора. Молекулы Br_2 находятся практически в равновесии с ионами Br^- благодаря быстрой обратимой реакции [15]:



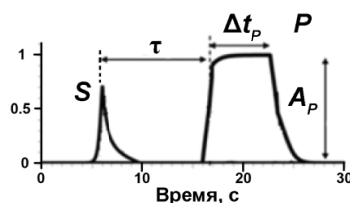
Гидрофобные молекулы Br_2 являются «переносчиком информации» между БЖ-микрокаплями.

Можно выделить следующие общие типы связи между осцилляторами:

1) Связь диффузионного типа, которая относится к локальным связям. Сила такой связи пропорциональна разности концентраций диффундирующего вещества в соседних реакторах.

2) Импульсная связь, которая может быть как локальной, так и далекодействующей. Говоря о такой связи, мы можем представлять себе синхронические связи между нейронами или СМС-сообщения по сотовой связи. Импульсная связь может характеризоваться также и некоторым временем задержки τ после какого-либо события S в реакторе i , и импульсным воздействием на реактор j с заранее определенной амплитудой A_p (рис. 2).

Рис. 2. Импульсная связь. Импульс P в реакторе j возникает спустя τ секунд после того, как некая переменная S в реакторе i достигает некоего критического значения. Почти прямоугольный импульс P характерен длительностью Δt_p и амплитудой A_p



3) Связь между реакторами может быть ингибиторной (путем воздействия ингибитора) или активаторной (путем воздействия активатора). Про чисто ингибиторную (или активаторную) связь мы будем говорить, что связь установлена по ингибитору (или по активатору). Оба типа связи могут иметь как импульсную, так и диффузионную (постоянную) природу. Смешанная связь также возможна.

4) Если каждый элемент сети осцилляторов связан с каждым (all-to-all coupling), то говорят о глобальной связи, которая может реализоваться через усредненную концентрацию некоего мессенджера (так называемое среднее поле).

Работы по изучению двух связанных химических осцилляторов начались в 1975 году в Праге [10]. Наиболее отчетливый результат по динамическим режимам двух БЖ-осцилляторов, связанных массообменом, был получен в группе И. Эпштейна в Бостоне [8]. Эти авторы показали, что существуют три режима, зависящие от силы связи между реакторами (т. е. скорости массообмена) и начальных условий (обнаруживается эффект бистабильности или гистерезиса): синфазные колебания, колебания в противофазе и стационарный режим, когда все колебания подавлены.

Более интересные режимы возникают для двух одинаковых микрокапель, диффузионно связанных только по ингибитору или только по активатору через разделительную гидрофобную среду, даже если эти капли, будучи изолированными (например, в капилляре содержится только одна капля), находятся в стационарном состоянии (рис. 3) [16].

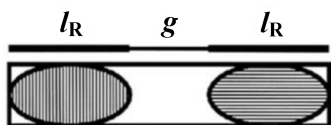


Рис. 3. В находящиеся в капилляре каплях длиной l_R протекает гипотетическая реакция, описываемая, например, моделью бресселятора в стационарном состоянии

Если между каплями может диффундировать только ингибитор, то при малой длине зазора g (т. е. когда сила связи велика) возникают стационарные структуры Тьюринга, которые в данном случае выглядят как неомогенное стационарное состояние системы с совершенно разными значениями переменных бресселятора в левой и правой каплях. Напомним, что модель бресселятора описывает систему, состоящую из следующих четырех реакций: $A \rightarrow X$; $B + X \rightarrow Y + D$; $2X + Y \rightarrow 3X$; $X \rightarrow E$, где переменная X — это активатор (автокаталитически размножающаяся частица), а переменная Y — это ингибитор. Если зазор g достаточно велик, но не чрезмерно, то могут возникнуть высокоамплитудные синфазные релаксационные колебания при симметричном возбуждении исходного стационарного состояния. В этом случае зазор умеренной длины играет роль временной задержки, при которой и возможны такие синфазные колебания, что их период зависит от величины g .

Если же в зазоре между каплями может диффундировать только молекула активатора, то в некотором диапазоне длин зазора g в системе возникают синусоидальные противофазные колебания, природа которых объясняется волновой неустойчивостью [17, 18].

Рассмотрим теперь те же две изолированные капли, но находящиеся в колебательном состоянии, и будем считать, что колебания в каплях описываются моделью орегонатора [19]. Связь между каплями осуществляется диффузионно посредством некоего мессенджера, который линейно и равновесно связан с ингибитором внутри капель. Так же как и в случае с макрореакторами, можно найти устойчивые противофазные колебания,

в то время как синфазные колебания хоть и существуют, но неустойчивы к малым, но конечным возмущениям. Помимо этих двух симметричных мод, существуют три несимметричные моды: стационарная структура Тьюринга (малый зазор g), колебания только одного осциллятора при подавленных колебаниях второго (большой зазор g) и хаотические бѣстинговые колебания (еще больший зазор g), когда наблюдаются пачки колебаний, разделенные большими (и нерегулярными) отрезками времени, в течение которых нет никакой активности.

Теперь мы подошли к случаю, когда между каплями могут диффундировать как молекулы активатора, так и молекулы ингибитора. Для этого более реалистичного случая рассмотрим и более реалистичную модель БЖ-реакции с четырьмя переменными (модель Ванга — Эпштейна) [20]:

$$dx/dt = -k_1xy + k_2y - 2k_3x^2 + k_4x(c_0 - z)/(c_0 - z + c_{\min}), \quad (9)$$

$$dy/dt = -3k_1xy - 2k_2y - k_3x^2 + k_7u + k_9z, \quad (10)$$

$$dz/dt = 2k_4x(c_0 - z)/(c_0 - z + c_{\min}) - k_9z - k_{10}z, \quad (11)$$

$$du/dt = 2k_1xy + k_2y + k_3x^2 - k_7u, \quad (12)$$

где $u = [Br_2]$, переменные x , y и z — те же, что и в модели орегонатора, c_0 — это общая концентрация катализатора (сумма концентраций окисленной и восстановленной формы), $c_{\min} \ll c_0$, k_i — константы скоростей реакций, некоторые из которых зависят от $[MA]$, $[A]$ и $h (= [H^+])$ [21]. Связь между двумя соседними осцилляторами осуществляется посредством диффузии в зазоре g молекул S (это те же молекулы Br_2 , но в масле), линейно связанных с ингибитором Br_2 в каплях, и молекул W (это те же молекулы X , но в масляном зазоре g), линейно связанных с активатором X в каплях, и выражается дифференциальными уравнениями для $s_{n,n+1}$ и $w_{n,n+1}$ (концентрации частиц S и W в зазоре между каплями n и $n+1$) [21]:

$$ds_{n,n+1}/dt = C_U r_V (u_n + u_{n+1} - 2s_{n,n+1}/P_U), \quad (13)$$

$$dw_{n,n+1}/dt = C_X r_V (x_n + x_{n+1} - 2w_{n,n+1}/P_X), \quad (14)$$

где P_U и P_X — равновесные коэффициенты распределения молекул U и X соответственно между масляной и водной фазами ($P_U = 2,5$; $P_X < 0,1$); $r_V = l_R/g$, $k_f \cong 2D/l_R^2$, D — коэффициент диффузии, который приблизительно одинаков как для молекул S и W , так и для молекул U и X . Сила связи по ингибитору $C_U = k_f r_V P_U / (1 + r_V P_U)$, а по активатору $C_X = k_f r_V P_X / (1 + r_V P_X)$. При этом уравнения (9) и (12) модифицируются следующим образом:

$$dx_n/dt = -k_1x_ny_n + k_2y_n - 2k_3x_n^2 + k_4x_n(c_0 - z_n)/(c_0 - z_n + c_{\min}) + C_X(w_{n,n+1}/P_X + w_{n-1,n}/P_X - 2x_n), \quad (15)$$

$$du_n/dt = 2k_1x_ny_n + k_2y_n + k_3x_n^2 - k_7u_n + C_U(s_{n-1,n}/P_U + s_{n,n+1}/P_U - 2u_n). \quad (16)$$

Как и раньше, обнаруживаются следующие режимы: синфазные колебания (IP), которые устойчивы в данном случае, противофазные колебания (AP) и стационарные структуры Тьюринга (T). Если эти режимы помес-

тить на параметрическую плоскость, например на плоскость (k_f, r_V) , то обнаруживается тройная точка, в которой сходятся все три области различных мод (рис. 4).

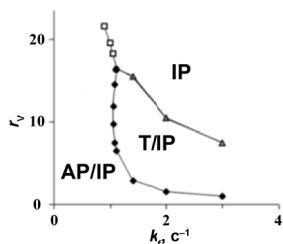


Рис. 4. $P_X = 0,003$, $k_{op} = 0,07 \text{ M}^{-1} \cdot \text{c}^{-1}$, $k_9 = k_{op}[\text{MA}]$, $[\text{MA}] = 0,05 \text{ M}$. Обозначения AP/IP и T/IP означают, что в данной области могут существовать два различных режима (моды) в зависимости от начальных условий

Отметим, что структуры Тьюринга (которые также получаются и в численном эксперименте для соответствующей реакционно-диффузионной системы) образуются при одинаковых коэффициентах диффузии молекул активатора и ингибитора, но при разных коэффициентах распределения P_U и P_X . Это расширяет наши знания об условиях существования структур Тьюринга (да и других диссипативных структур) на случай гетерогенных сред.

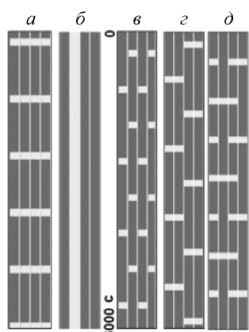


Рис. 5. Образующиеся при моделировании системы (9)—(16) динамические режимы четырех диффузионно связанных по кругу БЖ-осцилляторов: *a* — синфазные колебания (IP), P_X — любое; *б* — стационарные структуры Тьюринга, $P_X = 0 \div 0,0045$; *в* — противофазные колебания (AP), $P_X = 0 \div 0,003$; *г* — синфазные-противофазные колебания (IPAP), $P_X = 0,001 \div 0,005$; *д* — режим «3 + 3», $P_X = 0,005 \div 0,006$. Параметры модели: $P_U = 2,5$, $r_V = 10$, $k_f = 0,45 \text{ c}^{-1}$, $k_{op} = 0,07 \text{ M}^{-1} \cdot \text{c}^{-1}$, $k_9 = k_{op}[\text{MA}]$, $[\text{MA}] = 0,02 \text{ M}$, $h = 0,15 \text{ M}$, $A = 0,3 \text{ M}$

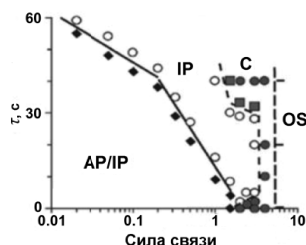
Для большего числа диффузионно связанных по кругу микроосцилляторов можно найти немало количество различных режимов синхронизации, в которых осцилляторы объединяются в кластеры (кластер — это группа осцилляторов, колеблющихся синфазно) [13, 21]. В качестве примера на рис. 5 (в координатах «пространство — время») показаны режимы для четырех осцилляторов. Горизонтальные белые черточки-прямоугольнички обозначают спайки БЖ-осцилляторов. Спайки представляют собой кратковременный и быстрый (автокаталитический) переход катализатора из восстановленного состояния (темный цвет на вертикальных колонках, соответствующих динамике осцилляторов 1, 2, 3 и 4) в окисленное

(белый цвет на вертикальных колонках). Синфазный режим (рис. 5, *a*), стационарные структуры Тьюринга, которые в данном случае несимметричны (рис. 5, *b*), и противофазные колебания (рис. 5, *в*) аналогичны таким же модам для двух связанных осцилляторов. Новыми являются моды IPAP (in-phase-anti-phase) (рис. 5, *г*), в которой два соседних осциллятора (пара 1—2 и пара 3—4) колеблются синфазно (отсчет ведется слева направо), а пары 1—2 и 3—4 колеблются в противофазе, и мода «3 + 3» (рис. 5, *д*), в которой можно видеть два кластера, колеблющихся в противофазе, но осцилляторы 1 и 3 принадлежат обоим кластерам, что увеличивает частоту их колебаний вдвое по сравнению с осцилляторами 2 и 4.

Рассмотрим теперь динамические режимы двух импульсно связанных БЖ-осцилляторов [9, 22—24]. В отличие от диффузионно связанных осцилляторов, кратковременная связь между осцилляторами осуществляется только благодаря спайку в одном из осцилляторов (см. рис. 2). Эта связь может быть однонаправленной и совершенно несимметричной. Если БЖ-реакция находится в стационарном состоянии, то связи между осцилляторами нет. Соответственно, импульсно связанные осцилляторы не могут порождать структуры Тьюринга.

Для двух идентичных осцилляторов, симметрично связанных по ингибитору, была найдена диаграмма всех мод, которая на плоскости «сила связи — время задержки τ » показана на рис. 6. Режимы AP и IP ничем не отличаются от найденных ранее для диффузионной связи.

Рис. 6. Диаграмма динамических режимов двух БЖ-осцилляторов, импульсно связанных по ингибитору: AP — противофазные колебания, IP — синфазные колебания, C — комплексный режим, OS — режим, в котором один из осцилляторов подавлен



Новые режимы C и OS представлены на рис. 7. Комплексными (C) мы называем такие режимы, в которых разные осцилляторы дают разное число спайков за один глобальный период.

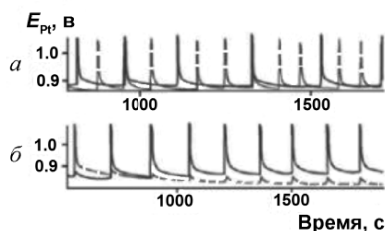


Рис. 7. Режимы C (*a*) и OS (*б*). Колебания регистрировались по потенциалу E платинового электрода

В случае симметричной импульсной связи по активатору образуются другие режимы синхронизации, показанные на рис. 8. Отметим, что в эксперименте в реакторы импульсно впрыскивался раствор ионов серебра в качестве активатора, так как Ag^+ выводит из реакции ингибитор Br^- . Для режимов AIP и FAP время между ближайшими спайками приблизительно равно τ . Для режима OS (oscillations-suppression) подавленный осциллятор находится в окисленном состоянии, в то время как для связи по ингибитору — в восстановленном состоянии.

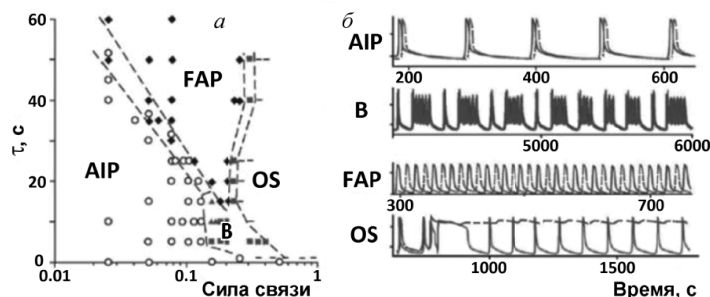


Рис. 8. Диаграмма в плоскости «сила связи — время задержки τ » (а) и примеры динамических режимов БЖ-осцилляторов (сплошная и штриховая линии), импульсно связанных по активатору (б): AIP — почти синфазные колебания, В — бёстинг, FAP — быстрые противофазные колебания, OS — режим, в котором один из осцилляторов подавлен

В случае, если связь между одинаковыми осцилляторами несимметрична, или осцилляторы обладают разными собственными периодами (T_1 и T_2), наблюдается появление различных резонансных режимов типа $n : m$, где n и m — это число спайков второго и первого осцилляторов соответственно за один глобальный период. В качестве примера на рис. 9 показана диаграмма режимов в плоскости (T_2/T_1 , τ) для активаторной связи [24].

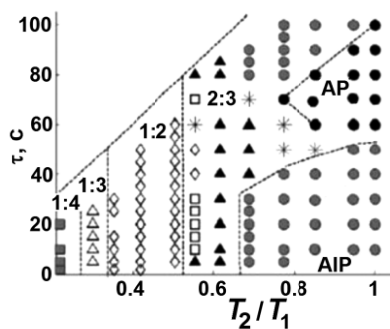
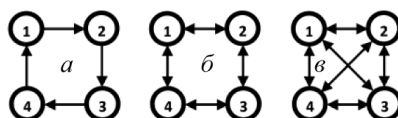


Рис. 9. Диаграмма резонансов при постоянной силе связи. Область резонанса 1 : 1 расщепляется на две подобласти: AIP и AP (= FAP). Области резонансов 2 : 3, 1 : 2, 1 : 3 и 1 : 4 показаны на диаграмме. Звездочки обозначают более сложные режимы

Рассмотрим теперь четыре осциллятора, импульсно связанных ингибиторной связью [25]. Такая сеть осцилляторов используется иногда для моделирования аллюров четвероногих [26]. Можно перечислить следующие семь аллюров: шаг (walk), рысь (trot), иноходь (pace), легкий галоп (canter), галоп (gallop), карьер (bound) и прыжки одновременно на всех четырех ногах (pronk). Интересно, что некоторые из них можно предсказать наиболее распространенными модами двух связанных осцилляторов, а именно IP и AP. Действительно, обозначим эти моды комбинациями двух стрелок: $\uparrow\uparrow$ для IP и $\uparrow\downarrow$ для AP. Теперь будем комбинировать эти две пары стрелок в разных сочетаниях. $IP + IP = \uparrow\uparrow\uparrow\uparrow$, что соответствует аллюру «прыжки» (pronk). $AP + AP = \uparrow\downarrow\downarrow\downarrow$, что соответствует рыси. Другая комбинация $AP + AP$ или $IP + IP$ дает $\uparrow\uparrow\downarrow\downarrow$, что соответствует иноходи или карьере. Комбинация $IP + AP$ дает кластер «3 + 1» ($\uparrow\uparrow\uparrow\downarrow$), которому не соответствует ни один нормальный аллюр. А особой комбинации $AP + AP$ в виде $\uparrow\rightarrow\downarrow\leftarrow$ можно поставить в соответствие шаг (walk), где горизонтальные стрелки в AP-паре $\rightarrow\leftarrow$ указывают на то, что ее фаза колебаний сдвинута на четверть периода по отношению к фазе колебаний в AP-паре $\uparrow\downarrow$.

Посмотрим теперь, как же реализуются все эти моды в реальном эксперименте для трех типов связности одинаковых осцилляторов, показанных на рис. 10 [25].

Рис. 10. Три типа связности одинаковых осцилляторов: *a* — однонаправленная связь по кругу, *b* — двунаправленная (или обоюдная) связь по кругу, *в* — «каждый с каждым»



Конфигурация (или связность) с однонаправленными связями по кругу (рис. 10, *a*) дает 4 основных регулярных ритма (рис. 11, слева направо): шаг (W), противофазные колебания (AP) (соответствует аллюру «рысь»), обратный шаг (WR) и синфазные колебания (IP) (pronk), причем эти моды появляются в системе последовательно по мере увеличения времени задержки τ .



Рис. 11. Регулярные ритмы для однонаправленной ингибиторной импульсной связи между четырьмя БЖ-осцилляторами

Расположение этих ритмов на плоскости «сила связи C_{inh} — время задержки τ » представлено на рис. 12, *a*, из которого видно, что в системе существует мультистабильность: в некоторых областях диаграммы могут

присутствовать до трех различных режимов. Также видно, что один и тот же режим, например W (шаг), находится в разных областях диаграммы.

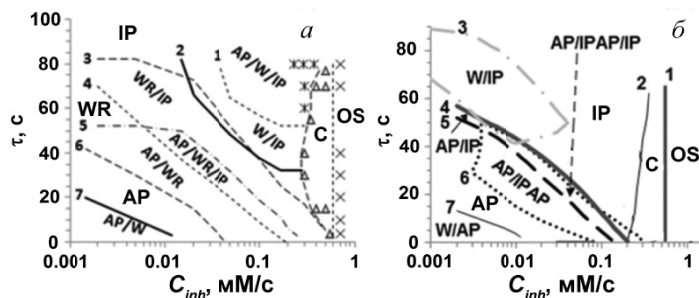


Рис. 12. Диаграмма динамических режимов четырех однонаправленно (а) и двунаправленно (б) связанных БЖ-осцилляторов на плоскости «сила связи C_{inh} — задержка τ ». Режимы: W — шаг, AP — противофазные колебания (два кластера), WR — обратный шаг, IP — синфазные колебания (один кластер), C — комплексные колебания, OS — хотя бы один осциллятор подавлен

В случае двунаправленной импульсной связи по кругу (см. рис. 10, б) к четырем регулярным режимам, показанным на рис. 11, добавляется еще один режим, названный нами IPAP (рис. 13, а). Это режим, соответствующий упомянутому выше иноходи или карьере. Заметим, что для двунаправленной связи уже нет различия между режимами W и WR, так как исчезает выделенное направление. Диаграмма всех режимов показана на рис. 12, б. Как и в случае однонаправленной связи, можно видеть мультистабильность и появление одного и того же режима в разных областях диаграммы.

В качестве примера одного из сложных периодических режимов (область C на диаграмме, представленной на рис. 12, б) на рис. 13, б показан режим «2/3», в котором число спайков (обозначенных черточками) за глобальный период для осциллятора 2 равно двум, а для осциллятора 3 — трем (осцилляторы 2 и 4 синфазны, а осцилляторы 1 и 3 в некотором смысле противофазны).

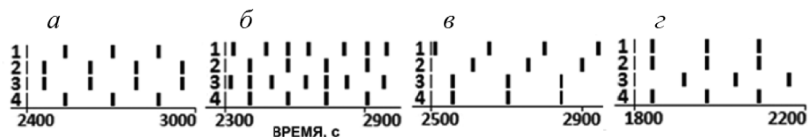


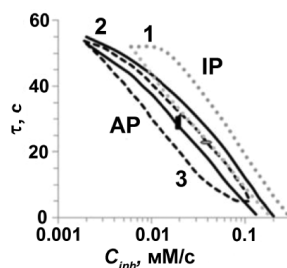
Рис. 13. Графики «пространство — время» для режимов IPAP (а), «2/3» (б), «3C1» (в), «3 + 1» (z)

В случае связности «каждый с каждым» (см. рис. 10, в) добавляются еще два новых режима, которые мы назвали «3C1» (рис. 13, в) и «3 + 1» (рис. 13, z). Таким образом, мы можем сказать, что с увеличением числа

связей растет и число возможных режимов. Несимметричные моды 3С1 и «3 + 1» интересны также и тем, что они возникают в полностью симметричной системе, где ни один осциллятор не отличается по своим характеристикам от другого.

Режимы С (комплексный) и OS, возникающие при достаточно больших силах связи, весьма разнообразны [25]. Остановимся на другом классе сложных режимов (С-II), которые возникают внутри областей регулярных режимов (т. е. при умеренных силах связи), причем только в таких областях, в которых существует тристабильность. Небольшие области, в которых обнаружены режимы С-II, помечены на рис. 14 двумя короткими и толстыми штрихами. Вертикальный штрих соответствует случаю связности «каждый с каждым», а горизонтальный — двунаправленной связности. Режим С-II для случая однонаправленных связей не найден.

Рис. 14. Области тристабильности для трех типов связности. В области, ограниченной кривой 1 (пунктирная линия, случай однонаправленной связи), существуют AP, IP и WR-моды; ограниченной кривой 2 (жирная сплошная линия, случай двунаправленной связи) — AP, IP и IPAP-моды; ограниченной кривой 3 (штриховая линия, случай связи «каждый с каждым») — моды AP, IP и «3 + 1»



При начальных условиях, которые занимают промежуточное положение между всеми тремя регулярными модами (возможными в выделенных областях), появляется новая устойчивая мода С-II, в которой числа спаек N_i (за глобальный период) i -х осцилляторов отличаются друг от друга. Так, например, для моды на рис. 15, а $N_1 : N_2 : N_3 : N_4 = 4 : 3 : 4 : 4$, а для моды на рис. 15, б $N_1 : N_2 : N_3 : N_4 = 2 : 3 : 2 : 3$. Для моды на рис. 15, а осциллятор 2 дает синфазные спайки поочередно с осцилляторами 4, 1, 3, опять 4, 1, 3 и т. д.

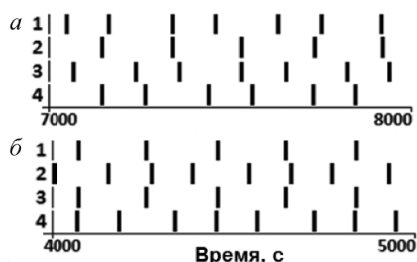


Рис. 15. Примеры сложных режимов С-II для связей «каждый с каждым» (а) и двунаправленных связей (б)

Если рассматривать осцилляторы 4, 1, 3 как кластеры, состоящие из одного элемента, то этот режим может быть интерпретирован как периодическое переключение (или перескоки) осциллятора 2 от одного кластера к другому, т. е. перескоки между следующими квазиравновесными трехкластерными состояниями системы: $(2 + 4, 1, 3) \rightarrow (4, 2 + 1, 3) \rightarrow (4, 1, 2 + 3) \rightarrow (2 + 4, 1, 3)$ и т. д., где знак «+» в скобках между 2 и n ($n = 4, 1$ или 3) обозначает синхронное появление спайков в соответствующих осцилляторах. Это напоминает гетероклинические петли, описанные М. Рабиновичем [27].

Подводя итог рассмотрения четырех импульсно связанных (по ингибитору) одинаковых осцилляторов, можно заключить, что общее число регулярных мод (таких как W, WR, AP, 3Cl, IPAP, 3 + 1, IP), сложных мод и OS-мод (которые мы здесь не рассматривали) велико. Если осцилляторы или связи между ними начинают различаться между собой, то число новых мод еще возрастает. Если эти моды несут в себе какую-то информацию (как в случае нервной системы), то должны быть считывающие элементы (интегральные нейроны), которые понимали бы смысл каждой моды. Можно предположить, что число нейронов головного мозга, которые ответственны за создание синхронных паттернов (мод) не так уж и велико. Однако число паттернов (ритмов), как мы видели на примере сети из четырех осцилляторов, многократно превышает число нейронов, генерирующих эти паттерны. Если для каждого такого паттерна нужен хотя бы один интегральный (или анализирующий) нейрон, то число нейронов, выполняющих функцию узнавания динамических паттернов, должно многократно превышать число ритмогенерирующих нейронов. Анализирующие же нейроны для выполнения своих функций должны встраиваться в обратные связи, которые приводят к моторным действиям или когнитивным процессам.

Оставляя в стороне философские размышления, вернемся к связанным осцилляторам и сравним импульсную связь с диффузионной на примере четырех осцилляторов, связанных по кругу. Для диффузионной связи образующиеся моды показаны на рис. 5, а для импульсной связи — на рис. 12, б (в виде диаграммы), рис. 11 и 13, а. Видно, что моды IP, AP и IPAP одинаковы для обоих типов связи. Моды «Тьюринг» (рис. 5, б) и комплексная мода «3 + 3» (рис. 5, д) не найдены для импульсной связи, а мода W (шаг) не обнаруживается для диффузионной связи. При увеличении силы диффузионной связи система переходит в стационарную тьюринговскую моду, а при увеличении силы импульсной связи сначала образуются комплексные колебания, а затем OS-моды с увеличивающимся числом подавленных осцилляторов, но стационарный тьюринговский режим невозможен.

Нейрофизиологи отмечают, что для функционирования нервной системы даже простейших существ (беспозвоночных) важны оба типа связи —

химические (импульсные) синапсы и электрические синапсы (аналогичные диффузионной связи между соседними клетками) [28]. Однако даже простейшая система из двух связанных осцилляторов, но с двумя типами связи, импульсной и диффузионной, пока еще не исследовалась.

Остановимся кратко на методах изучения большого числа связанных осцилляторов. Если не пользоваться суперкомпьютерами, то применение полных химических моделей в данном случае слишком затратно. Удобнее пользоваться методом фазовых осцилляторов, которые описывают любой осциллятор как вращение точки по кругу с одной переменной — фазой ϕ (задающей положение точки на круге). Полагают, что если осциллятор делает полный круг и достигает фазы 2π или 1 (для простоты), то он генерирует спайк, который может воздействовать на другие осцилляторы. Импульсное воздействие со стороны другого осциллятора рассматривается как смещение фазы — вперед для активаторной связи и назад для ингибиторной связи. Зависимости величины смещения фазы от фазы самого осциллятора называются кривыми переустановки фазы (или *phase response curves*, PRC).

Впервые описание ансамбля связанных осцилляторов фазовыми осцилляторами предложил А. Винфри (A. Winfree) в 1967 году. Однако представленные им уравнения были весьма сложны. Й. Курамото в 1975 году упростил уравнения Винфри. А спустя еще 11 лет Х. Сакагучи и Й. Курамото ввели в уравнения Курамото фазовый сдвиг α [29]:

$$\dot{\phi}_k = \omega_k + \frac{\varepsilon}{N} \sum_{j=1}^N \sin(\phi_j - \phi_k + \alpha). \quad (17)$$

В 1993 году было показано, что динамика системы (17) при $\omega_k = \omega$ может быть описана всего тремя переменными и $N - 3$ константами движения [30]. И, наконец, в 2008 году Э. Отт и Т. Антонсен нашли решение для параметра порядка задачи Курамото [31].

В больших ансамблях взаимодействующих осцилляторов возникают такие интересные явления, как *quorum sensing* — коллективный отклик системы при увеличении плотности частиц (осцилляторов) [32], или как динамический режим «Химера», когда одна часть популяции колеблется синхронно, а другая часть — асинхронно или даже хаотически [33]. Полагают, что для появления режима «Химера» важны нелокальные связи. Упомянем также эффект полихронизации [34], отличающийся от синхронизации тем, что группы осцилляторов, дающих спайки в разные моменты времени (не синхронно), могут коллективно активировать другие осцилляторы (нейроны), на которые их воздействие приходит одновременно из-за различных времен задержки импульсных связей.

Сети осцилляторов могут выполнять такие функции, как запоминание, обучение, распознавание образов, адаптация и принятие решений [35—

39]. Всё это делает сети связанных осцилляторов привлекательными для разработки искусственного интеллекта и химического компьютера. Последний может быть создан, например, на базе рассмотренных выше БЖ-капелек. При создании такого нового параллельного компьютера (который иногда называют *soft computer*) необходимо решить целый ряд задач, к которым можно отнести следующие:

1) Исследовать роль обратных связей в сетях осцилляторов, способных к выработке адаптивного поведения и обеспечивающих переключения между разными ритмами.

2) Разработать алгоритмы не только для переключения между различными ритмами, но и для переключения между разными типами связей, например импульсными и диффузионными или активаторными и ингибиторными (в нейронных сетях эту работу выполняют гормоны).

3) При исследовании роли обратных связей особо следует рассмотреть случаи, когда наряду с функцией обратной связи выполняется еще и логическая операция типа AND или OR. Такое может происходить, если сигнал, используемый для создания обратной связи, сравнивается с другим (например, внешним) сигналом. Результатом такого сравнения (которое может происходить по алгоритму логической функции AND или OR или другой) может стать, например, ингибиторное или активаторное воздействие на исходную ритмогенерирующую систему.

4) Различные роли разных осцилляторов неизбежно поднимают вопрос об иерархичности таких сетей, как сенсорные сети, сети-процессоры или генераторы ритмов, считывающие или анализирующие сети, сети, управляющие механикой (моторные нейроны).

5) Так как сеть микроосцилляторов является гетерогенной средой, то возникает потребность разработки теории диффузионных и импульсных нестабильностей в гетерогенных системах и дискретных сетях.

Работа выполнена при финансовой поддержке РФФИ (грант 15-07-01726).

Литература

1. *Pikovsky, A.* Synchronization. A universal concept in nonlinear sciences / A. Pikovsky, M. Rosenblum, J. Kurths. — Cambridge : University Press, 2003. — P. 357—361.
2. *Dilão, R.* Antiphase and in-phase synchronization of nonlinear oscillators : the Huygens's clocks system // *Chaos*. — 2009. — V. 19. — Art. 023118.
3. *Acebron, J. A.* The Kuramoto model : a simple paradigm for synchronization phenomena / J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, R. Spigler // *Rev. Mod. Phys.* — 2005. — V. 77. — P. 137—185.
4. *Strogatz, S. H.* Coupled Oscillators and Biological Synchronization / S. H. Strogatz, I. Stewart // *Sci. Am.* — 1993. — V. 269. — P. 102—109.

5. Белоусов, Б. П. Периодически действующая реакция и ее механизм // Сб. рефератов по радиационной медицине за 1958 г. — М. : Медгиз, 1959. — С. 145.
6. Field, R. J. Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in bromate-cerium-malonic acid system / R. J. Field, E. Körös, R. M. Noyes // J. Am. Chem. Soc. — 1972. — V. 94. — P. 8649—8664.
7. Field, R. J. Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction / R. J. Field, R. M. Noyes // J. Chem. Phys. — 1974. — V. 60. — P. 1877—1884.
8. Crowley, M. F. Experimental and Theoretical Studies of a Coupled Chemical Oscillator : Phase Death, Multistability, and In-Phase and Out-of-Phase Entrainment / M. F. Crowley, I. R. Epstein // J. Phys. Chem. — 1989. — V. 93. — P. 2496—2502.
9. Horvath, V. Pulse-Coupled Chemical Oscillators with Time Delay / V. Horvath, P. L. Gentili, V. K. Vanag, I. R. Epstein // Angew. Chem. Int. Ed. — 2012. — V. 51. — P. 6878—6881.
10. Marek, M. Synchronization in Two Interacting Oscillatory Systems / M. Marek, I. Stuchl // Biophys. Chem. — 1975. — V. 3. — P. 241—248.
11. Toiya, M. Diffusively coupled chemical oscillators in a microfluidic assembly / M. Toiya, V. K. Vanag, I. R. Epstein // Angew. Chem. Int. Ed. — 2008. — V. 47. — P. 7753—7755.
12. Toiya, M. Synchronization of chemical micro-oscillators / M. Toiya, H. O. Gonzalez-Ochoa, V. K. Vanag, S. Fraden, I. R. Epstein // J. Phys. Chem. Lett. — 2010. — V. 1. — P. 1241—1246.
13. Ваняг, В. К. Диссипативные структуры в системах диффузионно связанных химических нано- и микроосцилляторов // Рос. хим. ж-л (Ж-л Рос. хим. об-ва им. Д. И. Менделеева). — 2009. — Т. 53, № 6. — С. 16—24.
14. Ваняг, В. К. Волны и динамические структуры в реакционно-диффузионных системах. Реакция Белоусова — Жаботинского в обращенной микроэмульсии // УФН. — 2004. — Т. 174. — С. 991—1010.
15. Field, R. J. Oscillations and Traveling Waves in Chemical Systems / R. J. Field, M. Burger. — N. Y. : Wiley, 1985. — Chapter 3.
16. Vanag, V. K. Diffusive Instabilities in Heterogeneous Systems / V. K. Vanag, I. R. Epstein // J. Chem. Phys. — 2003. — V. 119. — P. 7297—7307.
17. Turing, A. M. The chemical basis of morphogenesis // Philos. Trans. R. Soc. London, Ser. B. — 1952. — V. 237. — P. 37—72.
18. Vanag, V. K. Packet Waves in a Reaction-Diffusion System / V. K. Vanag, I. R. Epstein // Phys. Rev. Lett. — 2002. — V. 88. — Art. 088303.
19. Vanag, V. K. Periodic perturbation of one of two identical chemical oscillators coupled via inhibition / V. K. Vanag, I. R. Epstein // Phys. Rev. E. — 2010. — V. 81. — Art. 066213.
20. Vanag, V. K. A model for jumping and bubble waves in the Belousov-Zhabotinsky-aerosol OT system / V. K. Vanag, I. R. Epstein // J. Chem. Phys. — 2009. — V. 131. — Art. 104512.
21. Vanag, V. K. Excitatory and inhibitory coupling in a one-dimensional array of Belousov-Zhabotinsky micro-oscillators : theory / V. K. Vanag, I. R. Epstein // Phys. Rev. E. — 2011. — V. 84. — Art. 066209.

22. *Lavrova, A. I.* Two pulse-coupled non-identical, frequency-different BZ oscillators with time delay / A. I. Lavrova, V. K. Vanag // *Phys. Chem. Chem. Phys.* — 2014. — V. 16. — P. 6764—6772.
23. *Проскуркин, И. С.* Динамические режимы двух разночастотных химических осцилляторов, связанных ингибиторной импульсной связью с задержкой / И. С. Проскуркин, В. К. Ваняг // *ЖФХ.* — 2015. — Т. 89, № 2. — С. 340—344.
24. *Proskurkin, I. S.* Inhibitory and excitatory pulse coupling of two frequency-different chemical oscillators with time delay / I. S. Proskurkin, A. I. Lavrova, V. K. Vanag // *Chaos.* — 2015. — V. 25. — Art. 064601.
25. *Vanag, V. K.* Dynamical regimes of four almost identical chemical oscillators coupled via pulse inhibitory coupling with time delay / V. K. Vanag, P. S. Smelov, V. V. Klinshov // *Phys. Chem. Chem. Phys.* — 2016. — V. 18. — P. 5509—5520.
26. *Golubitsky, M.* Nonlinear dynamics of networks : the groupoid formalism / M. Golubitsky, I. Stewart // *Bull. Am. Math. Soc.* — 2006. — V. 43. — P. 305—364.
27. *Rabinovich, M.* Neuroscience – transient dynamics for neural processing / M. Rabinovich, R. Huerta, G. Laurent // *Science.* — 2008. — V. 321. — P. 48—50.
28. *Bargmann, C. I.* From the connectome to brain function / C. I. Bargmann, E. Marder // *Nature Methods.* — 2013. — V. 10. — P. 483—490.
29. *Sakaguchi, H.* A soluble active rotator model showing phase transition via mutual entrainment / H. Sakaguchi, Y. Kuramoto // *Prog. Theor. Phys.* — 1986. — V. 76. — P. 576—581.
30. *Watanabe, S.* Integrability of Globally Coupled Oscillator Array / S. Watanabe, S. H. Strogatz // *Phys. Rev. Lett.* — 1993. — V. 70. — P. 2391—2394.
31. *Ott, E.* Low dimensional behavior of large systems of globally coupled oscillators / E. Ott, T. M. Antonsen // *Chaos.* — 2008. — V. 18. — Art. 037113.
32. *Taylor, A. F.* Dynamical Quorum Sensing and Synchronization in Large Populations of Chemical Oscillators / A. F. Taylor, M. R. Tinsley, F. Wang, Z. Huang, K. Showalter // *Science.* — 2009. — V. 323. — P. 614—617.
33. *Tinsley, M. R.* Chimera and phase-cluster states in populations of coupled chemical oscillators / M. R. Tinsley, S. Nkomo, K. Showalter // *Nature Physics.* — 2012. — V. 8. — P. 662—665.
34. *Izhikevich, E. M.* Polychronization : computation with spikes // *Neural Computation.* — 2006. — V. 18. — P. 245—282.
35. *Hohmann, W.* Pattern Recognition by Electrical Coupling of Eight Chemical Reactors / W. Hohmann, M. Kraus, F. W. Schneider // *J. Phys. Chem. A.* — 1999. — V. 103. — P. 7606—7611.
36. *Johansson, C.* Towards cortex sized artificial neural systems / C. Johansson, A. Lansner // *Neural Networks.* — 2007. — V. 20. — P. 48—61.
37. *Buzsáki, G.* Neural Syntax: Cell Assemblies, Synapsembles, and Readers // *Neuron.* — 2010. — V. 68. — P. 362—385.
38. *Adamatzky, A.* On architectures of circuits implemented in simulated Belousov-Zhabotinsky droplets / A. Adamatzky, J. Holley, P. Dittrich, J. Gorecki, B. De Lacy Costello, K.-P. Zauner, L. Bull // *BioSystems.* — 2012. — V. 109. — P. 72—77.
39. *Huepe, C.* Adaptive-network models of swarm dynamics / C. Huepe, G. Zschaler, A.-L. Do, T. Gross // *New J. Phys.* — 2011. — V. 13. — Art. 073022.

СОЛИТОННЫЙ ГАЗ: ВЗАИМОДЕЙСТВИЯ, ТУРБУЛЕНТНОСТЬ И ВОЛНЫ-УБИЙЦЫ

Е. Н. Пелиновский^{1,2}, Е. Г. Шургалина^{1,2}, А. В. Слюняев^{1,2}

¹ Институт прикладной физики РАН, г. Нижний Новгород

² Нижегородский государственный технический университет им. Р. Е. Алексеева

Введение

Солитоны являются важной частью современной нелинейной физики. В интегрируемых системах, таких как уравнение Кортевега — де Вриза (КдВ) и нелинейное уравнение Шрёдингера (НУШ), существует довольно много подходов (метод обратной задачи рассеяния, преобразования Дарбу и Бэклунда, билинейный метод Хироты и т. д.), позволяющих получить точные решения для описания взаимодействия солитонов. Солитонная турбулентность является частью волновой турбулентности, для которой кинетические уравнения описывают параметры соответствующей задачи рассеяния, а не амплитуды волн или амплитуды гармоник Фурье. В 1971 году В. Е. Захаров первым показал фундаментальную роль парных взаимодействий в рамках уравнения Кортевега — де Вриза [1]. Затем кинетическая теория для солитонного газа была развита Г. Элем с соавторами [2—5]. Эта теория способна описывать различные свойства солитонных ансамблей. Солитонная турбулентность вырождена в каком-то смысле, потому что солитоны (а точнее, дискретные значения ассоциированной спектральной задачи в методе обратной задачи рассеяния) сохраняются в процессе взаимодействия. В работах [6—8] представлено так называемое нелинейное преобразование Фурье, являющееся вариантом метода обратной задачи рассеяния для периодических граничных условий. Применительно к морским волнам на мелкой воде (в рамках уравнения Кортевега — де Вриза) оно позволяет выяснить «состав» наблюдаемых случайных волн. Эти компоненты волнового поля (аналоги кноидальных волн или солитонов) сохраняются во времени, но их суперпозиция ведет к случайной динамике морской поверхности. Случайные ансамбли солитонов уже были исследованы в литературе в рамках уравнения КдВ, главным образом численно [9—14], а также в рамках модифицированного уравнения КдВ [15], в том числе авторами. В этой статье будет дан обзор основных результатов.

1. Математическая модель

Наш интерес связан с исследованием солитонного газа в интегрируемых системах, описываемых уравнениями типа Кортевега — де Вриза, такими как каноническое уравнение Кортевега — де Вриза

$$\frac{\partial u}{\partial t} + 6u \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0 \quad (1)$$

и модифицированное уравнение Кортевега — де Вриза (мКдВ)

$$\frac{\partial u}{\partial t} + 6u^2 \frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3} = 0. \quad (2)$$

Оба уравнения имеют точные решения в виде солитонов. Решение для КдВ имеет вид

$$u(x, t) = A \operatorname{sech}^2 \left(\sqrt{\frac{A}{2}} (x - ct - x_0) \right), \quad c = 2A, \quad (3)$$

а для мКдВ —

$$u(x, t) = sA \operatorname{sech} \left(A(x - ct - x_0) \right), \quad c = A^2, \quad (4)$$

параметр $s = \pm 1$ определяет полярность солитона.

Как видно из (3) и (4), в случае уравнения КдВ солитоны могут быть только одной полярности, в то время как в уравнении мКдВ существуют солитоны обеих полярностей.

2. Двухсолитонное взаимодействие

Как было показано В. Е. Захаровым в 1971 году для уравнения КдВ, точная динамика мультисолитонных волновых полей описывается через попарные взаимодействия. Поэтому для определения роли солитонных взаимодействий в общей динамике солитонного газа мы рассматриваем сначала особенности взаимодействия именно двух солитонов в рамках уравнений КдВ и мКдВ. Этот анализ основан на известном точном двухсолитонном решении.

Двухсолитонное решение уравнения КдВ имеет вид [16, 17]

$$\begin{aligned} u(x, t) = & 8(A_1 - A_2) \left(A_1 \cosh^2 \left(\sqrt{\frac{A_2}{2}} (x - 2A_2 t) \right) + A_2 \sinh^2 \left(\sqrt{\frac{A_1}{2}} (x - 2A_1 t) \right) \right) \times \\ & \times \left((\sqrt{2A_1} - \sqrt{2A_2}) \cosh \left(\sqrt{\frac{A_1}{2}} (x - 2A_1 t) + \sqrt{\frac{A_2}{2}} (x - 2A_2 t) \right) + \right. \\ & \left. + (\sqrt{2A_1} + \sqrt{2A_2}) \cosh \left(\sqrt{\frac{A_1}{2}} (x - 2A_1 t) - \sqrt{\frac{A_2}{2}} (x - 2A_2 t) \right) \right)^{-2}, \end{aligned} \quad (5)$$

при этом предполагается, что $A_1 > A_2$. В этом решении опущены дополнительные свободные параметры, соответствующие начальным положениям солитонов (которые всегда могут быть исключены соответствующим сдвигом координаты и времени). Запись (5) обладает определенной симметрией, делающей ее удобной для анализа.

Двухсолитонное решение уравнения мКдВ имеет следующую структуру [18, 19]:

$$\begin{aligned}
u(x, t) = & 2\gamma(s_1 A_1 \cosh(A_2(x - A_2 t)) + s_2 A_2 \cosh(A_1(x - A_1 t))) \times \\
& \times (s_1 s_2 (\gamma^2 - 1) + \gamma^2 \cosh(A_1(x - A_1 t) - A_2(x - A_2 t)) + \\
& + \cosh(A_1(x - A_1 t) + A_2(x - A_2 t)))^{-1}, \\
& \gamma = \frac{A_1 + A_2}{A_1 - A_2} > 1,
\end{aligned} \tag{6}$$

где мы опять пренебрегли индивидуальными фазами солитонов, которые могут быть исключены преобразованием времени и координаты.

При удалении солитонов друг от друга решения (5) и (6) представляются суммой двух невзаимодействующих солитонов:

$$u(x, t) = u_1(x, t) + u_2(x, t), \tag{7}$$

где $u_{1,2}$ есть односолитонные решения (3) или (4) соответствующих уравнений (1) или (2) с амплитудами $A_{1,2}$. В обоих уравнениях взаимодействие однополярных солитонов ведет к нелинейному смещению более быстрого (наибольшего) солитона вперед, а более медленного — назад. Интересно, что выражение для сдвига фаз, записанное через скорости взаимодействующих солитонов, будет одинаковым для обоих уравнений:

$$\Delta x_{1,2} = \pm \frac{2}{\sqrt{c_{1,2}}} \ln \left(\frac{\sqrt{c_2} + \sqrt{c_1}}{\sqrt{c_2} - \sqrt{c_1}} \right), \tag{8}$$

где c_1, c_2 — скорости взаимодействующих солитонов, а сдвиг Δx положителен для большего солитона и отрицателен для меньшего. Отметим, что такой сдвиг фаз характерен и для обобщенного уравнения, содержащего квадратичную и кубическую нелинейность одновременно (уравнение Гарднера [20, 21]).

Двухсолитонные решения (5), (6) инвариантны по отношению к одновременной смене знака координаты и времени, а также к смене знака времени в точке $x = 0$ и смене знака координаты в момент $t = 0$. Потому $t = 0$ соответствует моменту максимального взаимодействия, когда образуется симметричный профиль волны. Значение второй производной по координате от двухсолитонного решения в момент $t = 0$ в точке $x = 0$ определяет качественно форму этого профиля. Так, для случая взаимодействия однополярных солитонов (как для КдВ, так и для мКдВ) существуют два режима взаимодействия солитонов — обмен и обгон в зависимости от отношения амплитуд (рис. 1, *a, б*). В обоих случаях амплитуда результирующего импульса уменьшается в момент взаимодействия. В случае же взаимодействия разнополярных солитонов (что возможно лишь в случае мКдВ) реализуется только один случай взаимодействия — сценарий «поглощение-испускание», когда более быстрый солитон поглощает солитон с меньшей скоростью и возникает кратковременное увеличение амплитуды (рис. 1, *в*).

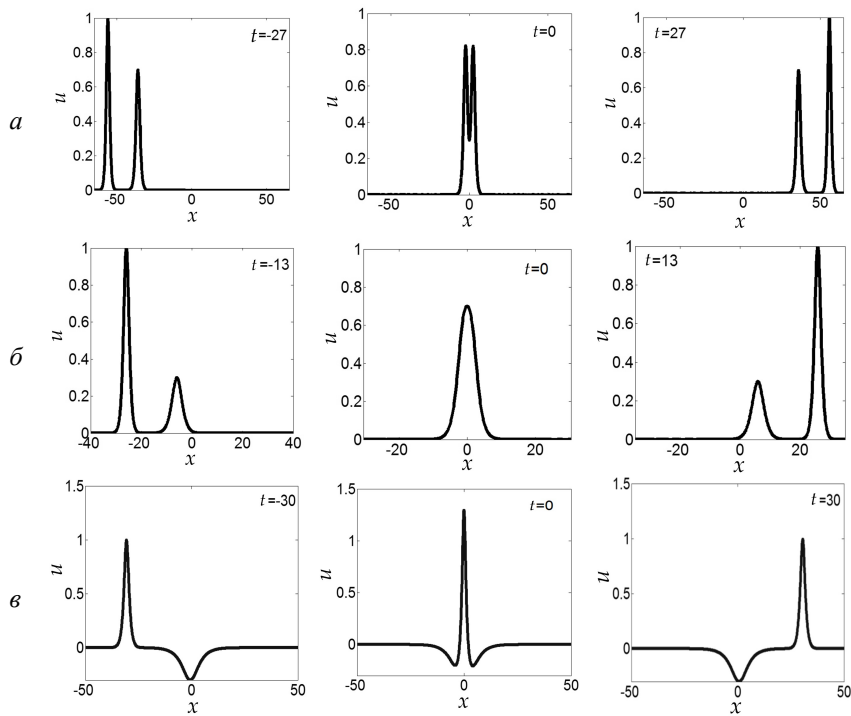


Рис. 1. Разные виды взаимодействия двух солитонов:
a — обменный, *б* — обгонный, *в* — поглощение-испускание

3. Численное моделирование динамки мультисолитонных полей

Динамика нерегулярных солитонных ансамблей изучается численно с помощью псевдоспектрального метода для периодических граничных условий (детали описаны в работе [14]). Если ограничиваться случаем солитонного газа малой плотности, начальное условие удобно задавать в виде суперпозиции разнесенных солитонов

$$u(x, 0) = \sum_{i=1}^N A_i \operatorname{sech}^2 \left[\sqrt{\frac{A_i}{2}} (x - i\Delta x) \right] \quad (9)$$

для уравнения КдВ и аналогично

$$u(x, 0) = \sum_{i=1}^N s_i A_i \operatorname{sech} \left[A_i (x - i\Delta x) \right] \quad (10)$$

для мКдВ, где N — число солитонов в расчетной области (в наших экспериментах $N = 100$). Начальные положения солитонов выбраны так, что они

хорошо отделены друг от друга, $\Delta x = 20$. В наших экспериментах задан случайным образом набор амплитуд солитонов и реализации различаются между собой только расположением солитонов в расчетной области. В начальный момент времени амплитуды солитонов (по модулю) принимают равновероятное значение в пределах от $A_{\min} = 1$ до $A_{\max} = 3$. Две реализации таких полей представлены на рис. 2.

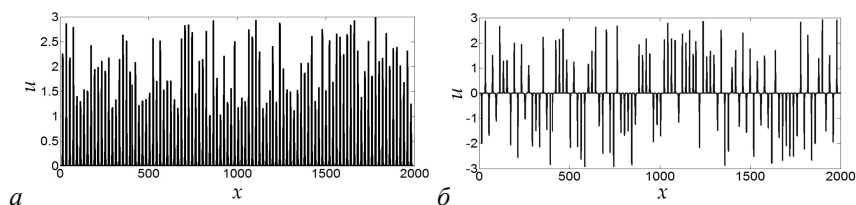


Рис. 2. Начальные солитонные поля: *a* — однополярное, *b* — разнополярное

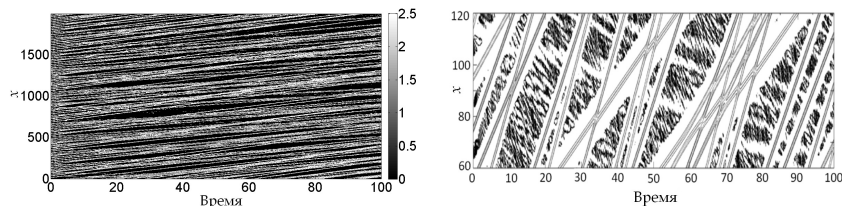


Рис. 3. Пространственно-временная диаграмма солитонных однополярных мКДВ-полей (*справа* — увеличенный фрагмент рисунка)

Солитоны имеют различные амплитуды и, соответственно, различные скорости, так что с течением времени они изменяют взаимное расположение и взаимодействуют. Рисунок 3 демонстрирует эволюцию солитонного поля, изображенного на рис. 2, *a*, в плоскости xt . Различные наклоны траекторий соответствуют различным скоростям солитонов. Траектории солитонов до и после взаимодействия не ложатся на одну прямую линию, демонстрируя сдвиг фаз в результате нелинейного взаимодействия солитонов. В среднем за 100 единиц расчетного времени каждый солитон участвует в 10—20 взаимодействиях. За это время усредненные характеристики солитонного газа достигают стационарного значения (см. раздел 6).

4. Эффект отрицательной скорости солитона

Отличительной особенностью рассматриваемых моделей (1) и (2) является то, что одиночные солитоны движутся в одном направлении. Однако если солитоны взаимодействуют между собой, то при усреднении они могут изменять скорость движения вплоть до изменения направления

движения на обратное (в физической задаче этому соответствует движение со скоростью меньше линейной скорости длинных волн). Пример такой динамики показан на рис. 4, где построены траектории солитонов со случайными амплитудами в рамках уравнения мКдВ.

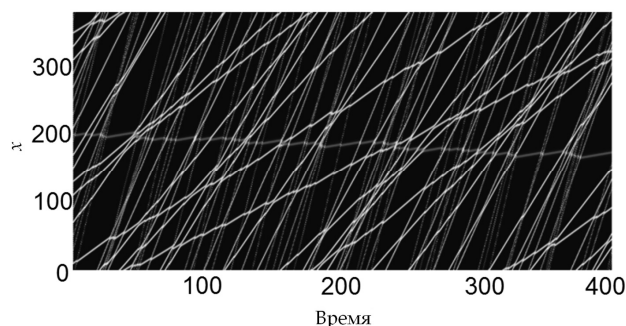


Рис. 4. Пространственно-временная диаграмма поля солитонов одной полярности в рамках уравнения мКдВ, содержащего малый солитон с эффективной отрицательной скоростью

На рисунке 4 малый солитон (выходящий из $x = 200$) движется в обратную сторону. Такая динамика оказывается возможной в результате отрицательного сдвига фазы малого солитона во время нелинейного взаимодействия с большими солитонами, о чем упоминалось еще в статье [22] в рамках уравнения КдВ. В недавней работе [23] было найдено выражение для средней скорости солитона в «холодном» газе солитонов в рамках кинетической теории для уравнения КдВ, из которого также следует этот эффект, хотя он и не отмечался. Так как сдвиг фаз солитонов (8) универсален для рассматриваемых интегрируемых уравнений типа КдВ, критерий возникновения эффективной отрицательной скорости также оказывается универсальным для уравнений (1) и (2).

Оценка для условий возникновения эффекта обратного движения малого солитона может быть легко получена из кинематических соображений сравнением расстояния, которое малый солитон проходит до взаимодействия с большим солитоном, со сдвигом фазы вследствие взаимодействия. В случае движения малого солитона («дефекта») в решетке, состоящей из больших солитонов, двигающихся со скоростью c и отделенных друг от друга расстоянием L , условие эффективной нулевой скорости дефекта принимает простую форму

$$c_d = \frac{4\sqrt{c}}{L}, \quad (11)$$

где c_d есть скорость «свободного» малого солитона.

В случае движения в газе солитонов со случайными амплитудами и фазами малый солитон также может двигаться в обратном направлении, как показано на рис. 4. Здесь необходимо рассматривать большой временной интервал, учитывая взаимодействия больших солитонов не только с дефектом, но и между собой. Если, однако, плотность солитонного газа достаточно мала, то вклад взаимодействий больших солитонов между собой должен нивелироваться (часть солитонов ускоряется, а часть тормозится), в то время как отрицательный набег фазы при взаимодействии пробного солитона со всеми большими солитонами накапливается. Поэтому, если пренебречь сдвигом фаз при взаимодействии больших солитонов, критерий нулевой скорости дефекта в солитонном газе примет вид

$$c_d = \frac{4 \langle 1/\sqrt{c} \rangle}{L \langle 1/c \rangle}, \quad (12)$$

где угловыми скобками обозначено усреднение по ансамблю. В случае периодической решетки равенство (12) переходит в (11). Обе формулы справедливы как для уравнения Кортевега — де Вриза, так и для модифицированного уравнения Кортевега — де Вриза.

5. Экстремумы волновых полей и функции распределения амплитуд

Некоторые выводы о динамике волновых полей можно сделать, анализируя графики экстремумов волновых полей. Поскольку парные взаимодействия однополярных солитонов ведут к уменьшению амплитуды результирующего импульса [24], то максимальное значение экстремума волнового поля, состоящего из однополярных солитонов, не превосходит значения амплитуды наибольшего солитона в реализации (в нашем случае $A_{\max} = 3$) (рис. 5, *a*). Большие волны не возникают в ансамбле однополярных солитонов, что также справедливо для однополярного солитонного газа уравнения мКдВ [14].

Разнополярный солитонный газ ведет себя абсолютно по-другому из-за того, что взаимодействие разнополярных солитонов приводит к увеличению амплитуды результирующего импульса, что уже отмечалось выше (см. также [25, 26]). В таких полях возможно появление волн, амплитуды которых превышают амплитуду наибольшего солитона в начальной реализации (рис. 5, *b*) [14, 15].

Данные процессы влияют на функции распределения волновых полей. В начальный момент времени амплитуды солитонов распределены по равномерному закону (рис. 6). С течением времени функция распределения амплитуд волн (локальных максимумов волнового поля) меняется в каждой реализации. Рисунок 6 демонстрирует качественно разный характер изменения функций распределений волновых полей, состоящих из

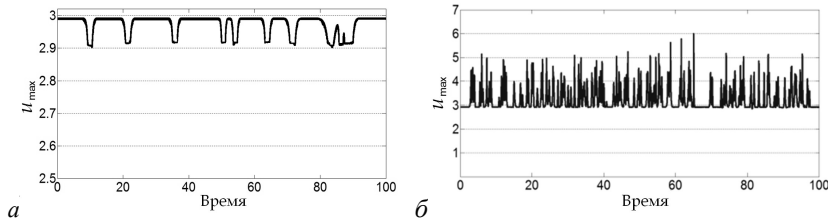


Рис. 5. Изменение во времени максимума волновых полей:

a — однополяренный солитонный газ, *б* — разнополяренный солитонный газ

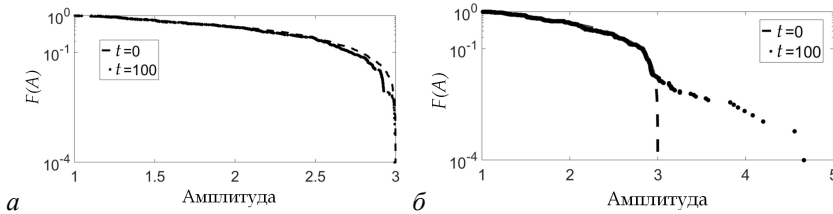


Рис. 6. Функция распределения амплитуд солитонов в момент времени $t = 100$ и в начальный момент времени: *a* — однополяренный солитонный газ, *б* — разнополяренный солитонный газ

однополяренных и разнополяренных солитонов. Так, в случае однополяренного газа число малоамплитудных импульсов возрастает, а число больших волн уменьшается (рис. 6, *a*). В случае же знакопеременных полей эффект противоположен: доля малоамплитудных волн уменьшается, а волн с большими амплитудами — увеличивается. В результате хвосты функции распределения могут значительно возрастать (рис. 6, *б*).

6. Статистические моменты волновых полей

Статистические моменты разреженного солитонного газа могут быть вычислены аналитически, например в начальный момент времени, когда солитоны находятся на достаточном расстоянии друг от друга и взаимодействие солитонов полностью не учитывается [13, 14].

В частности, для уравнения КдВ первые четыре статистических момента (среднее, дисперсия, асимметрия и эксцесс) в начальный момент времени легко вычисляются:

$$\langle u(t=0) \rangle = 2\sqrt{2}\rho \langle A^{1/2} \rangle, \quad (13)$$

$$\sigma^2(t=0) = \frac{8}{3\sqrt{2}}\rho \langle A^{3/2} \rangle - 8\rho^2 \langle A \rangle, \quad (14)$$

$$Sk(t=0) \approx \frac{\sqrt{3}\sqrt[4]{2} \langle A^{5/2} \rangle}{5\sqrt{\rho} \langle A^{3/2} \rangle^{3/2}}, \quad (15)$$

$$Kur(t=0) \approx \frac{9\sqrt{2} \langle A^{7/2} \rangle}{35\rho \langle A^{3/2} \rangle^2}. \quad (16)$$

Коэффициент ρ определяется соотношением N/L и имеет ясный физический смысл плотности солитонного газа. Отметим, что из положительности дисперсии волнового поля появляется ограничение на плотность газа солитонов уравнения КдВ [5, 14]:

$$\rho < \rho_{\text{кр}} = \frac{\langle A^{3/2} \rangle}{3\sqrt{2} \langle A \rangle}. \quad (17)$$

Важно отметить, что первые два момента (среднее и дисперсия) являются инвариантами уравнения Кортевега — де Вриза («масса» и «энергия» соответственно), и они сохраняются во времени. Другие статистические моменты — асимметрия (15) и эксцесс (16) — изменяются в процессе взаимодействия солитонов. В данном случае формулы (15) и (16) являются асимптотическими, справедливыми для очень разреженного газа ($\rho \rightarrow 0$).

Аналогичные вычисления могут быть сделаны и для газа солитонов уравнения мКдВ [15]:

$$\langle u \rangle = \frac{1}{L} \int_0^L u(x,t) dx = \pi\rho \langle s \rangle, \quad (18)$$

$$\sigma^2 = 2\rho \langle A \rangle - \langle u \rangle^2, \quad (19)$$

$$Sk \approx \frac{\pi\rho}{2\sigma^3} \langle sA^2 \rangle, \quad (20)$$

$$Kur \approx \frac{4\rho}{3\sigma^4} \langle A^3 \rangle. \quad (21)$$

Из (19) следует, что плотность однополярного газа солитонов уравнения мКдВ также ограничена критической плотностью:

$$\rho < \rho_{\text{кр}} = 2 \langle A \rangle / \pi, \quad (22)$$

в то время как на разнополярный газ такого ограничения нет.

На рис. 7 представлена временная эволюция третьего и четвертого моментов солитонного газа, рассчитанных для одной реализации, а также результат усреднения по 20 реализациям. С течением времени усредненные по реализациям коэффициенты асимметрии и эксцесса уменьшаются и после нескольких столкновений выходят на почти стационарные значения, в то время как в одной реализации они, естественно, остаются сильно варьирующимися. Причина спада статистических моментов — в характере взаимодействия солитонов одной полярности, приводящего толь-

ко к уменьшению третьего и четвертого моментов. Соответственно, такое поведение характерно для однополярных солитонных полей в рамках уравнений КдВ и мКдВ.

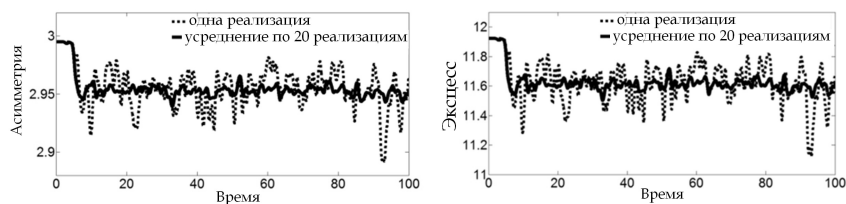


Рис. 7. Временная эволюция асимметрии и эксцесса в случае однополярного солитонного газа

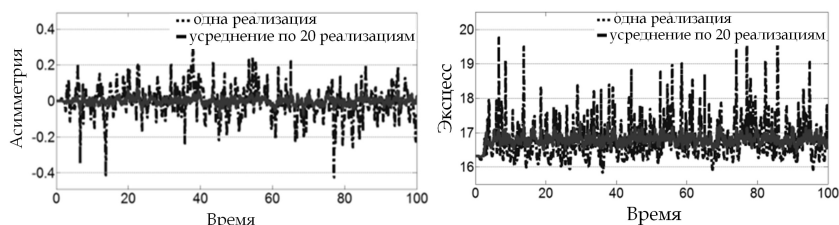


Рис. 8. Временная эволюция асимметрии и эксцесса в случае разнополярного солитонного газа

В отличие от статистических моментов однополярного солитонного газа в случае разнополярного солитонного газа мКдВ эксцесс преимущественно увеличивается по сравнению с начальным состоянием (рис. 8), указывая на рост хвостов функции распределения (см. рис. 6, б). Асимметрия может принимать как отрицательные, так и положительные значения в разные моменты времени, но ее значение в среднем близко к нулю вследствие баланса между солитонами положительной и отрицательной полярности. Таким образом, нелинейное взаимодействие разнополярных солитонов способно приводить к возникновению anomalously больших волн в сравнении с типичной амплитудой солитонов.

7. Роль мультисолитонных взаимодействий в появлении волн-убийц

С введением коэффициента усиления волн AI в момент t в виде

$$AI = \frac{\max_x |u(x, t)|}{\max_x |u(x, t_{\text{инд}})|}, \quad (23)$$

где $t_{\text{инд}}$ — момент, когда солитоны могут считаться независимыми (для рассмотренных выше численных примеров $t_{\text{инд}} = 0$), из приведенных выше

результатов следует, что в случае солитонного газа однополярных солитонов (КдВ или мКдВ) $AI \leq 1$. При взаимодействии двух разнополярных солитонов (случай мКдВ) $AI < 2$ (стремится к 2 в случае солитонов близких амплитуд, может равняться 2 для «дышащих волновых пакетов» — солитонов нелинейного уравнения Шрёдингера или бризеров мКдВ, о которых здесь не рассказывалось). В то же время общепринятой характеристикой так называемых волн-убийц (популяция особенно больших волн) является выполнение критерия $AI > 2$ [27, 28].

В работах [29, 30] для анализа «состава» гипотетических волн-убийц (заданных с помощью модельного профиля) был использован подход формального обращения времени в рамках уравнений КдВ и нелинейного уравнения Шрёдингера, что привело к довольно неожиданному выводу (который, впрочем, согласуется с приведенным выше обзором). Волна-убийца может быть «собрана» не более чем из одного солитона, остальная часть должна состоять из квазилинейного диспергирующего цуга. Иными словами, доминирующую роль в формировании волн, характеризуемых параметром $AI > 2$, должна играть дисперсионная фокусировка с учетом нелинейных поправок; большое количество солитонов препятствует их объединению в большую волну.

Как выяснилось позднее [31], вывод исследований [29, 30] оказался сильно ограниченным выбором ожидаемого профиля аномально высокой волны. Выводы работ [29, 30] верны в случае однополярных солитонов уравнений КдВ и мКдВ, но существенно неполны при рассмотрении разнополярных солитонов уравнения мКдВ. Мультисолитонные взаимодействия солитонов разной полярности могут приводить к эффективному сложению их амплитуд, при этом волна в фокусе обладает сложной структурой. Для анализа сфокусированной волны удобно воспользоваться N -солитонным решением мКдВ, записанным с использованием преобразования Дарбу в форме, предложенной в работе [18]:

$$u(x, t) = -i \frac{\partial}{\partial x} \ln \frac{W_N(\Psi_{1x}, \Psi_{2x}, \dots, \Psi_{Nx})}{W_N(\Psi_1, \Psi_2, \dots, \Psi_N)}. \quad (24)$$

Здесь W_N — вронскианы для пробных функций Ψ_j и их производных по координате, где

$$\Psi_j = \exp(\mu_j x - 4\mu_j^3 t) + i \exp(-\mu_j x + 4\mu_j^3 t). \quad (25)$$

Действительные числа $\mu_j, j = 1, \dots, N$ могут иметь разный знак и определяют амплитуды $|2\mu_j|$ и полярности солитонов (существенно, что полярность солитонов связана со знаком μ_j нетривиальным образом). В этом виде решение обладает той же удобной симметрией, что и в случае решений (5) и (6). Для его анализа требуется воспользоваться формальной записью определителя в виде

$$\det A = \sum_{(\alpha_1, \alpha_2, \dots, \alpha_N)} (-1)^{n(\alpha_1, \alpha_2, \dots, \alpha_N)} a_{\alpha_1 1} a_{\alpha_2 2} \cdot \dots \cdot a_{\alpha_N N}, \quad (26)$$

где a_{jk} — элементы матрицы A , суммирование производится по всем возможным перестановкам в наборе $(\alpha_1, \alpha_2, \dots, \alpha_N)$, задаваемом последовательностью натуральных чисел $(1, 2, \dots, N)$ (всего $N!$ перестановок), а $n(\alpha_1, \alpha_2, \dots, \alpha_N)$ — число парных перестановок элементов, необходимое для преобразования набора $(1, 2, \dots, N)$ в набор $(\alpha_1, \alpha_2, \dots, \alpha_N)$.

Оказывается возможным показать три важных свойства N -солитонного решения уравнения мКдВ (24): 1) решение обладает симметрией по отношению к изменению знака времени и координаты, а потому точка $(x = 0, t = 0)$ соответствует локальному экстремуму, и ее будем называть точкой фокуса; 2) значение в этой точке определяется линейной суперпозицией парциальных амплитуд солитонов

$$u(0, 0) = (-1)^{N+1} \sum_{j=1}^N 2\mu_j; \quad (27)$$

3) полярность солитона совпадает с μ_1 в случае $N = 1$, а в общем случае полярность s_m солитона номер m определяется комбинацией параметров всех солитонов:

$$s_m = \text{sgn} \left[\mu_m \prod_{j=1, j \neq m}^N (\mu_j^2 - \mu_m^2) \right]. \quad (28)$$

На основании этих свойств можно определить условия эффективной фокусировки множества солитонов. Оптимальным (с точки зрения максимума $AI = |u(0, 0)| / \max_j |\mu_j|$ при заданных полной энергии поля и числе солитонов) является выбор параметров μ_j с примерно одинаковой величиной (и одинаковым знаком). Тогда учет (28) приводит к выводу о чередовании полярности солитонов для больших времен, $t \rightarrow \pm\infty$, когда они упорядочены по скоростям (и амплитудам). Знак $u(0, 0)$ совпадает с полярностью наиболее быстрого солитона. Цуг из однополярных солитонов соответствует набору μ_j с разными знаками, что в силу (27) не приводит к усилению в точке фокуса.

Для обеспечения оптимальной фокусировки множества солитонов требуется «приготовить» упорядоченный цуг из солитонов, когда более быстрые солитоны расположены позади медленных. Положения солитонов должны обеспечивать одновременное схождение их траекторий в одной точке (с учетом сдвигов в процессе взаимодействия). Кроме того, важен выбор полярности солитонов: линейная суперпозиция парциальных амплитуд солитонов мКдВ достигается при чередующейся полярности солитонов (рис. 9, момент $t = -100$). Правильность построенной картины фокусировки множества солитонов проверена непосредственно с помо-

щью точного решения (24). В примере, приведенном на рис. 9, в фокусировке участвует 7 солитонов с примерно одинаковыми амплитудами, а максимум волны в фокусе ($t = 0$) равен их сумме, $AI = 6,79$.

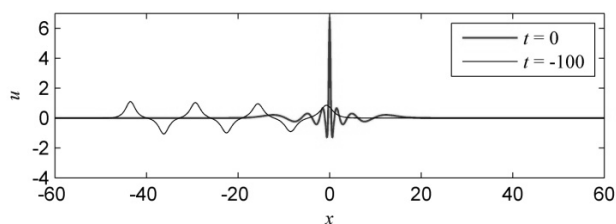


Рис. 9. Огромная волна, образовавшаяся в результате фокусировки семи солитонов мКдВ

Примененный подход и полученный результат могут быть обобщены на случай бризеров мКдВ [31], а также на другие интегрируемые уравнения (уравнение Гарднера, нелинейное уравнение Шрёдингера).

Заключение

Исследование солитонного газа в рамках уравнения Кортевега — де Вриза и модифицированного уравнения Кортевега — де Вриза показало значительное отличие в динамике солитонных полей, состоящих из однополярных и разнополярных солитонов. Картина взаимодействия солитонов существенно влияет на статистические характеристики волновых полей. Так, в случае однополярного солитонного газа асимметрия и эксцесс уменьшаются при взаимодействии солитонов, также как и хвосты функций распределения. В случае разнополярного солитонного газа эксцесс увеличивается, что ведет к увеличению доли волн с большими амплитудами. Продемонстрировано, что аномально большие волны (волны-убийцы) могут возникать в солитонных ансамблях, составленных из солитонов разной полярности.

Работа выполнена при финансовой поддержке грантов РФФИ 16-32-60012, 16-35-00175, 16-02-00167, 16-05-00049, 16-55-52019, 14-02-00983, гранта президента РФ для государственной поддержки ведущих научных школ РФ НШ-6637.2016.5, а также фонда «Фольксваген».

Литература

1. Захаров, В. Е. Кинетическое уравнение для солитонов // ЖЭТФ. — 1971. — Т. 60. — С. 993—1000.
2. El, G. A. Kinetic equation for a dense soliton gas / G. A. El, A. M. Kamchatnov // Phys. Rev. Lett. — 2005. — V. 95. — Art. 204101.

3. *El, G. A.* Soliton turbulence as the thermodynamic limit of stochastic soliton lattices / G. A. El, A. L. Krylov, S. A. Molchanov, S. Venakides // *Physica D.* — 2005. — № 152/153. — P. 653—664.
4. *El, G. A.* Kinetic equation for a soliton gas and its hydrodynamic reductions / G. A. El, A. M. Kamchatnov, M. V. Pavlov, S. A. Zykov // *J. Nonlinear Science.* — 2011. — V. 21. — P. 151—191.
5. *El, G. A.* Critical density of a soliton gas // *Chaos.* — 2016. — V. 26, № 2. — P. 1—7.
6. *Osborne, A. R.* Nonlinear Ocean Waves and the Inverse Scattering Transform / A. R. Osborne. — Oxford : Academic Press, 2010. — 976 p.
7. *Osborne, A. R.* Soliton basis states in shallow-water ocean surface waves / A. R. Osborne, E. Segre, G. Boffetta, L. Cavaleri // *Phys. Rev. Lett.* — 1991. — V. 67. — P. 592—595.
8. *Osborne, A. R.* Solitons, cnoidal waves and nonlinear interactions in shallow-water ocean surface waves / A. R. Osborne, M. Serio, L. Bergamasco, L. Cavaleri // *Physica D.* — 1998. — V. 123. — P. 64—81.
9. *Salupere, A.* On the KdV soliton formation and discrete spectral analysis / A. Salupere, G. A. Maugin, J. Engelbrecht, J. Kalda // *Wave Motion.* — 1996. — V. 123. — P. 49—66.
10. *Salupere, A.* Long-time behaviour of soliton ensembles. Part 1 : Emergence of ensembles / A. Salupere, P. Peterson, J. Engelbrecht // *Chaos Soliton. Fract.* — 2002. — V. 14. — P. 1413—1424.
11. *Salupere, A.* Long-time behaviour of soliton ensembles. Part 2 : Periodical patterns of trajectories / A. Salupere, P. Peterson, J. Engelbrecht // *Chaos Soliton. Fract.* — 2003. — V. 15. — P. 29—40.
12. *Salupere, A.* Long-time behavior of soliton ensembles / A. Salupere, P. Peterson, J. Engelbrecht // *Math. Comput. Simulat.* — 2003. — V. 62. — P. 137—147.
13. *Dutykh, D.* Numerical simulation of a solitonic gas in KdV and KdV–BBM equations / D. Dutykh, E. Pelinovsky // *Phys. Lett. A.* — 2014. — V. 378. — P. 3102—3110.
14. *Шургалина, Е. Г.* Динамика ансамбля нерегулярных волн в прибрежной зоне / Е. Г. Шургалина, Е. Н. Пелиновский. — Н. Новгород : НГТУ, 2015. — 179 с.
15. *Shurgalina, E. G.* Nonlinear dynamics of a soliton gas: Modified Korteweg – de Vries equation framework / E. G. Shurgalina, E. N. Pelinovsky // *Phys. Lett. A.* — 2016. — V. 380, № 24. — P. 2049—2053.
16. *Lax, P. D.* Integrals of nonlinear equations of evolution and solitary waves // *Communications on Pure and Applied Mathematics.* — 1968. — V. 21. — P. 467—490.
17. *Dimakis, A.* KdV soliton interactions: a tropical view / A. Dimakis, F. Muller-Hoissen // *J. Physics: Conference Series.* — 2014. — V. 482. — P. 012010.
18. *Anco, S. C.* Interaction properties of complex mKdV solitons / S. C. Anco, N. T. Ngatat, M. Willoughby // *Physica D.* — 2011. — V. 240. — P. 1378—1394.
19. *Куликов, М. Ю.* *N*-солитонные решения модифицированного уравнения Кортевега — де Вриза : препринт № 404 / М. Ю. Куликов, Г. М. Фрайман ; ИПФ РАН. — Н. Новгород, 1996.
20. *Слюняев, А. В.* Динамика солитонов большой амплитуды / А. В. Слюняев, Е. Н. Пелиновский // *ЖЭТФ.* — 1999. — Т. 116. — С. 318—335.

21. *Слюняев, А. В.* Динамика локализованных волн большой амплитуды в слабодиспергирующей среде с квадратичной и положительной кубической нелинейностью // ЖЭТФ. — 2001. — Т. 119. — С. 606—612.
22. *Кузнецов, Е. А.* Устойчивость стационарных волн в нелинейных средах со слабой дисперсией / Е. А. Кузнецов, А. В. Михайлов // ЖЭТФ. — 1974. — Т. 67, № 5. — С. 1717—1727.
23. *Carbone, F.* Macroscopic dynamics of incoherent soliton ensembles: Soliton gas kinetics and direct numerical modelling / F. Carbone, D. Dutykh, G. A. El // EPL. — 2016. — V. 113, № 3. — P. 30003—30008.
24. *Pelinovsky, E. N.* Two-soliton interaction as an elementary act of soliton turbulence in integrable systems / E. N. Pelinovsky, E. G. Shurgalina, A. V. Sergeeva, T. G. Talipova, G. A. El, R. H. J. Grimshaw // Phys. Lett. A. — 2013. — V. 377, № 3/4. — P. 272—275.
25. *Пелиновский, Е. Н.* Двухсолитонное взаимодействие в рамках модифицированного уравнения Кортевега — де Вриза / Е. Н. Пелиновский, Е. Г. Шургалина // Изв. вузов. Радиофизика. — 2014. — Т. 17, № 10. — С. 825—833.
26. *Пелиновский, Е. Н.* Формирование волн-убийц в солитонном газе, описываемом модифицированным уравнением Кортевега — де Вриза / Е. Н. Пелиновский, Е. Г. Шургалина // ДАН. — 2016. — Т. 470, № 1. — С. 26—29.
27. *Куркин, А. А.* Физика волн-убийц в океане / А. А. Куркин, Е. Н. Пелиновский, А. В. Слюняев // Нелинейные волны' 2004. — Н. Новгород : ИПФ РАН, 2005. — С. 37—51.
28. *Kharif, C.* Rogue Waves in the Ocean / C. Kharif, E. Pelinovsky, A. Slunyaev. — Berlin ; Heidelberg : Springer-Verlag, 2009. — 216 p.
29. *Pelinovsky, E.* Nonlinear-dispersive mechanism of the freak wave formation in shallow water / E. Pelinovsky, T. Talipova, C. Kharif // Physica D. — 2000. — V. 147. — P. 83—94.
30. *Kharif, C.* Focusing of nonlinear wave groups in deep water / C. Kharif, E. Pelinovsky, T. Talipova, A. Slunyaev // Письма в ЖЭТФ. — 2001. — Т. 73. — С. 190—195.
31. *Slunyaev, A. V.* Role of multiple soliton interactions in the generation of rogue waves : The modified Korteweg – de Vries framework / A. V. Slunyaev, E. N. Pelinovsky // Phys. Rev. Lett. — 2016. — V. 117, iss. 18. — P. 214501.

ФИЗИКА ВИХРЕЙ В СВЕРХПРОВОДНИКАХ КАК ПРИМЕР ЗАДАЧИ НЕЛИНЕЙНОЙ ФИЗИКИ

А. С. Мельников

Институт физики микроструктур РАН, г. Нижний Новгород

Исследование структуры и динамики вихревых систем в сверхпроводниках второго рода и разнообразных искусственных структурах является одной из ключевых проблем физики конденсированных сред.

Возникновение вихревых состояний в сверхпроводниках, как известно, непосредственно связано с разрушением однородной сверхпроводящей фазы под действием приложенного магнитного поля \mathbf{H} . В зависимости от соотношения между двумя микроскопическими масштабами, характеризующими сверхпроводник (глубина проникновения магнитного поля λ и длина когерентности ξ), разрушение сверхпроводимости в магнитном поле происходит путем фазового перехода первого ($\lambda/\xi < 2^{-1/2}$) или второго ($\lambda/\xi > 2^{-1/2}$) рода. В последнем случае возникает область промежуточных магнитных полей (рис. 1), в которой сверхпроводимость оказывается разрушенной частично, и магнитное поле проникает в образец в виде трубок, несущих квантованный магнитный поток $\Phi_0 = hc/2e$, или вихрей.

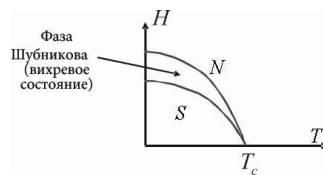


Рис. 1. Фазовая диаграмма сверхпроводника второго рода: S — однородная сверхпроводящая фаза, N — несверхпроводящая фаза

Экспериментально сверхпроводники 2-го рода (название соответствует роду перехода в магнитном поле) обнаружены в пионерских работах Л. В. Шубникова в 1937 году [1]. В 1957 году теоретическое объяснение этого феномена дал А. А. Абрикосов [2]. В 2003 году его теория была отмечена Нобелевской премией вместе с феноменологической теорией Гинзбурга — Ландау, которая позволяет описывать и понимать огромный круг явлений в сверхпроводниках как 1-го, так и 2-го рода.

Структура вихревых решений. Теория Гинзбурга — Ландау

Функционал свободной энергии сверхпроводника, согласно теории Гинзбурга — Ландау, имеет вид

$$\tilde{F} = F_n + \int \left(\frac{\hbar^2}{4m} \left| \left(\nabla - \frac{2ie}{\hbar c} \mathbf{A} \right) \Psi \right|^2 + a |\Psi|^2 + \frac{b}{2} |\Psi|^4 + \frac{\mathbf{B}^2}{8\pi} - \frac{\mathbf{B}\mathbf{H}}{4\pi} \right) dV.$$

Здесь Ψ — это комплексный параметр порядка сверхпроводника, который может рассматриваться как волновая функция сверхпроводящих носителей заряда (куперовских пар).

Несмотря на разнообразие описанных ниже вихревых структур их волновые функции обладают общим свойством, а именно наличием ненулевой циркуляции фазы волновой функции $\Psi = fe^{i\varphi}$ при обходе по контуру, охватывающему ось вихревой линии: $\oint \nabla\varphi d\mathbf{l} = 2\pi N$. Целое число N здесь можно называть завихренностью. Именно это общее свойство обеспечивает топологическую защищенность вихрей на плоскости: мы можем их создать или уничтожить, лишь порождая или аннигилируя решения с противоположной завихренностью (или вводя/выводя вихри с края образца).

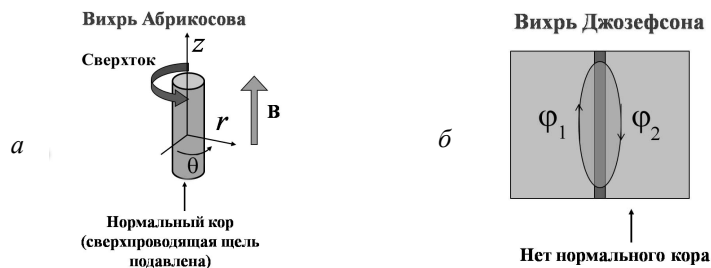


Рис. 2. Структура абрикосовских (а) и джозефсоновских (б) вихрей

Уравнения для параметра порядка и для магнитного поля получаются варьированием функционала по Ψ и векторному потенциалу \mathbf{A} . Возникающая нелинейная задача весьма сложна для аналитического решения и в общем случае может быть решена лишь с использованием численных подходов. Но существует ряд важных предельных случаев, которые допускают построение приближенных асимптотических или основанных на пробных функциях решений. Одним из таких предельных случаев является задача об изолированном вихре (рис. 2, а), в котором можно искать волновую функцию и магнитное поле в аксиально-симметричном виде: $\Psi = f(\rho)e^{i\theta}$, $\mathbf{B} = B(\mathbf{r})z_0$, где (ρ, θ, z) — цилиндрическая система координат с осью z , выбранной вдоль магнитного поля (совпадающего в этом случае с осью вихря). Модуль параметра порядка оказывается подавленным в области так называемого кора (или ядра) вихревой линии ($\rho < \xi$) и обращается точно в ноль при $\rho = 0$. В предельном случае сверхпроводников сильно 2-го рода ($\lambda/\xi \gg 1$) магнитное поле и создающие его сверхтекучие токи затухают на масштабе λ от оси вихревой линии. В том же пределе взаимодействие вихрей определяется взаимодействием их токов: потенциал отталкивания спадает логарифмически при расстояниях между вихрями $r < \lambda$ и спадает экспоненциально быстро при $r > \lambda$.

Существенные отклонения от описанной картины возникают лишь в магнитных полях, близких к верхнему критическому полю подавления сверхпроводимости $H_{c2} = \Phi_0/2\pi\xi^2$, когда расстояние между вихрями становится сравнимо с размерами их коров (керна). Эта ситуация была рассмотрена в работе А. А. Абрикосова 1957 года, где решение нелинейных уравнений Гинзбурга — Ландау удалось получить с использованием определенного типа теории возмущений по малому параметру $(1 - H/H_{c2})$. Возникающая при этом вихревая структура представляла собой периодическую решетку вихрей (рис. 3, а) с одним квантом магнитного потока на элементарную ячейку (энергетически выгодная структура решетки в дальнейшем оказалась правильной треугольной или гексагональной).

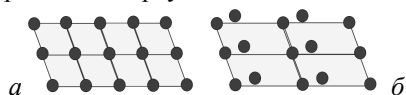


Рис. 3. Структура решетки вихрей: а — с одним квантом потока на элементарную ячейку, б — с двумя квантами потока на элементарную ячейку

Предположение о количестве квантов потока, приходящемся на элементарную ячейку, вообще-то, не является очевидным: изменение числа квантов (рис. 3, б) может возникать в сверхпроводящих соединениях с более сложным параметром порядка [3] или для высокой степени анизотропии кристаллической структуры [4]. Такое изменение структуры решетки может быть понято и как следствие качественного изменения парного потенциала взаимодействия вихрей: в определенном диапазоне расстояний между вихрями они начинают притягиваться друг к другу. В случае сильно анизотропного сверхпроводника модификация потенциала взаимодействия возникает вследствие существенного изменения самой структуры вихревой линии в магнитных полях, ориентированных не по осям анизотропии. Так, рассматривая слоистые сверхпроводящие системы (сверхпроводники, в которых сверхпроводящие носители заряда локализованы на двумерных атомных плоскостях с джозефсоновской связью между ними), мы можем получить вихревую

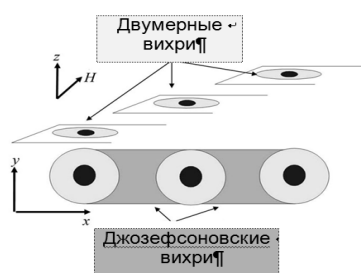
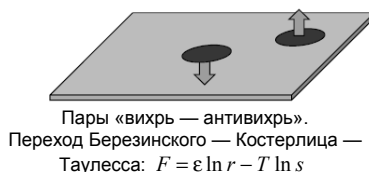


Рис. 4. Структура вихревой линии в слоистом сверхпроводнике

линию, состоящую из серии двумерных вихрей (панкейков), которые соединены между собой отрезками джозефсоновских вихрей, параллельных сверхпроводящим слоям (рис. 4). Здесь джозефсоновским вихрем мы называем вихрь без кора (керна), напоминающий решение с ненулевой завихренностью в джозефсоновском контакте между двумя сверхпроводящими берегами (см. рис. 2, б).

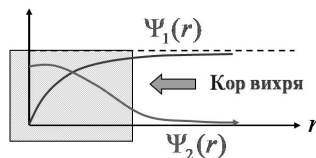
Заметим, что задачи о вихрях в слоистых системах отнюдь не являются отвлеченной схоластической проблемой, поскольку именно такие структуры естественным образом возникают в высокотемпературных купратных сверхпроводниках (и не только) в приложенных магнитных полях. Именно открытие в конце 1980-х годов высокотемпературных сверхпроводящих соединений типа BSCCO стимулировало теоретический и экспериментальный интерес к упомянутым выше вихревым конфигурациям. Сильная анизотропия делает актуальным еще один важный механизм генерации вихрей, который работает и в отсутствие приложенного магнитного поля. Дело в том, что в двумерных системах панкейки с противоположной завихренностью могут возникать флуктуационным образом (рис. 5), подобно тому как это происходит с магнитными вихрями в XY-модели при переходе Березинского — Костерлица — Таулесса (заметим, что работы последних двух авторов на данную тему были отмечены Нобелевской премией 2016 года). Проблема наблюдения упомянутого перехода в сверхпроводниках имеет ряд интересных особенностей, возникающих, например, вследствие экранировки.

Рис. 5. Флуктуационная пара вихрей с противоположной завихренностью. Свободная энергия пары вихрей



Альтернативную возможность реализовать притяжение между вихрями представляют системы, которые могут быть описаны в рамках теории Гинзбурга — Ландау параметром порядка, состоящим из нескольких комплексных компонент, например Ψ_1 и Ψ_2 . Эти компоненты могут иметь несовпадающие координаты особенностей [5], и в результате в коре (крене) вихря компонента Ψ_1 может возникнуть компонента Ψ_2 , не имеющая завихренности (рис. 6). Дополнительные комплексные компоненты в теории Гинзбурга — Ландау могут возникать из-за того, что в системе есть несколько сортов куперовских пар, либо отличающихся по симметричной классификации, либо отвечающих различным листам поверхности Ферми в зоне Бриллюэна [6].

Рис. 6. Кор (крен) вихря в сверхпроводнике с двухкомпонентным параметром порядка



Отметим кратко, что притяжение вихрей может возникать и в системах с сочетанием нескольких типов упорядочения, таких как магнетизм и сверхпроводимость. Примером могут служить ферромагнитные сверхпроводники или слоистые структуры «сверхпроводник — ферромагнетик». Притяжение вихрей может при этом приводить к формированию, например, страйп-доменной структуры промежуточного состояния [7], состоящего из вихревых полос (рис. 7).

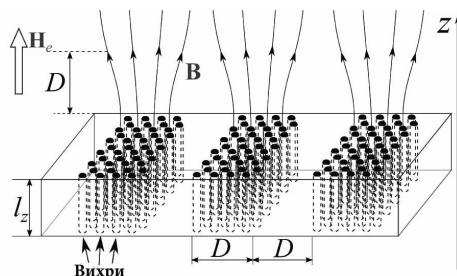


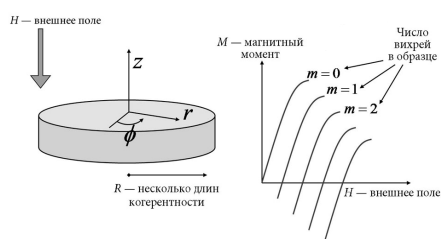
Рис. 7. Страйп-доменная структура промежуточного состояния, состоящего из вихревых полос

Образование вихревых кластеров может происходить и в обычном сверхпроводнике, если удержание вихрей в кластере обусловлено либо пиннингом на неоднородностях образца, либо границами системы и соответствующими краевыми экранирующими токами. Вблизи границ образца решетка может оказаться существенно искаженной в результате взаимодействия вихрей с краевыми экранирующими токами, индуцированными приложенным магнитным полем. Очевидно, что наиболее сильными такие искажения окажутся в достаточно малых образцах. Критерием малости образца является отношение его геометрических размеров L к характерным длинам в сверхпроводнике (лондоновская глубина проникновения и сверхпроводящая длина когерентности). В литературе к сверхпроводникам таких малых размеров обычно применяется термин «мезоскопические сверхпроводники». Силы, действующие со стороны краевых токов, сжимают вихревую систему, расстояния между отдельными вихрями становятся при этом меньше характерных расстояний в массивном образце. Такое сжатие вихревой системы оказывается особенно существенным, когда размер образца становится порядка нескольких длин когерентности. Расстояние между отдельными вихревыми линиями может обратиться в нуль, и в результате сформируются так называемые многоквантовые (или гигантские) вихри. В массивных образцах многоквантовые вихри неустойчивы, тем самым мезоскопические системы предоставляют уникальную возможность реализовать такие объекты экспериментально.

Необычные свойства вихревой фазы малых образцов привлекали интерес как теоретиков, так и экспериментаторов довольно давно: первые работы на эту тему относятся к 60-м годам XX века. Однако только со-

временное развитие технологии и измерительной техники сделало возможным детальное экспериментальное исследование систем такого типа, что стимулировало и теоретические работы по изучению свойств вихрей в мезоскопических сверхпроводниках. В результате за последнее время в теоретическом понимании особенностей такой вихревой фазы достигнут значительный прогресс: построено описание разнообразных фазовых переходов в системе вихрей, рассчитаны вихревые конфигурации как при слабых, так и при сильных магнитных полях, исследовано влияние геометрии образца на симметрию вихревых конфигураций и обнаружен новый тип смешанного состояния из вихрей и antivихрей, развита теория электронной структуры вихревых фаз и квантового транспорта в таких системах [8]. Думается, что математические подходы и опыт, накопленный при решении теоретических задач в модельной системе, интересны как сами по себе, так и для возможных приложений в смежных областях (сверхтекучести, физике ультрахолодных газов в ловушках). Кроме того, выяснение физических свойств простых вихревых систем бывает важным для понимания поведения вихревых кластеров в массивных образцах. Экспериментально переходы между различными вихревыми конфигурациями могут наблюдаться с использованием различных методик, чувствительных к магнитному моменту образца (рис. 8), или туннельных измерений.

Рис. 8. Вихри в мезоскопических дисках и зависимость их магнитного момента от магнитного поля

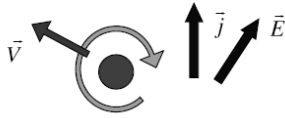


Динамика вихрей

Помимо интереса к фундаментальной задаче исследования собственно вихревых структур существует еще одна крайне важная причина, объясняющая, почему необходимо понимать физику вихрей. Именно устройство вихрей и их подвижность определяют диссипативные свойства сверхпроводников в магнитных полях. В этом плане задача о вихрях становится актуальной прикладной проблемой. Управление пиннингом вихрей, их положениями и подвижностью позволяет контролировать характеристики транспорта заряда и тепла. Важной базовой задачей является простейший расчет подвижности вихревой линии в обтекающем ее транспортном токе и соответствующего сопротивления в резистивном состоянии в рамках нестационарной теории Гинзбурга — Ландау с релаксационным уравнением для параметра порядка [9]

$$\gamma \left(\frac{\partial}{\partial t} - 2ie\Phi \right) \Psi = \left(\frac{1}{4m} \left(\nabla - \frac{2ie}{\hbar c} \mathbf{A} \right)^2 + a \right) \Psi - b |\Psi|^2 \Psi.$$

Задача была решена в 1971 году Л. П. Горьковым и Н. Б. Копниным [10] с помощью ставшей с тех пор стандартной техники «солитон + поле», использующейся по настоящее время в самых разных задачах о движении солитонов. В рамках этого подхода решение для параметра порядка выбирается в виде $\Psi = \Psi_0(\mathbf{r} - \mathbf{V}t) + \Psi_1$, где Ψ_0 — решение для статического вихря, а Ψ_1 рассматривается как малая поправка, удовлетворяющая линейному уравнению типа



$\hat{L} \Psi_1 = G(\Psi_0)$. Требование нерасходимости поправки Ψ_1 приводит к некоторому соотношению ортогональности для правой части приведенного уравнения, что дает нам уравнение движения вихревой линии $\eta \mathbf{V} = \mathbf{F}$. Здесь сила, действующая на вихрь, определяется суммой так называемой силы Лоренца, пропорциональной транспор-

Рис. 9. Вязкое движение вихрей в транспортном токе. $\rho = \rho_n \frac{H}{H_{c2}}$

тному току \mathbf{j} , и силы пиннинга: $\mathbf{F} = \Phi_0/c [\mathbf{j}, \mathbf{n}] + \mathbf{F}_{pin}$. Вязкость вихря определяется как диссипацией нормальных квазичастиц в коре (керне), так и релаксацией параметра порядка при перемещении кора. Электрическое поле и сопротивление могут быть найдены из соотношения $\mathbf{E} = 1/c [\mathbf{V}, \mathbf{B}]$ (рис. 9).

Модификация уравнения динамики параметра порядка или переход к более последовательной микроскопической теории дают несколько модифицированное уравнение движения, включающее увлечение вихрем потоком и массовый член: $m d\mathbf{V}/dt + \eta \mathbf{V} + \alpha[\mathbf{n}, \mathbf{V}] = \mathbf{F}$. Несмотря на весьма упрощенный характер использованной нестационарной теории Гинзбурга — Ландау, ее результаты оказываются весьма полезными для качественного, а зачастую и количественного анализа характеристик резистивного состояния, включая диссипативные и холловские эффекты, а также зависимость СВЧ-импеданса сверхпроводника от магнитного поля.

Квантовая механика квазичастиц в вихрях Абрикосова

Несмотря на необычайную эффективность описанных выше феноменологических подходов, их применение часто оказывается недостаточным, в особенности при низких температурах или при расчете характеристик сверхпроводящего состояния, чувствительных к спектру и пространственному распределению волновых функций квазичастиц в корах (кернах) вихрей. Еще в 1964 году в работе [11] было сделано важное теорети-

ческое предсказание о спектре квазичастиц, локализованных в корах вихрей вследствие андреевского отражения от стенок кора. Энергии этих состояний лежат в подщелевой области спектра $\varepsilon < \Delta_0$, а межуровневое расстояние легко оценить следующим образом:

$$\varepsilon_{\min} \sim \frac{\hbar^2}{m\xi^2} \sim \frac{\hbar^2 \Delta_0}{m\hbar v_F \xi} \sim \frac{\Delta_0}{k_F \xi},$$

где k_F — импульс Ферми. Именно эти андреевские состояния определяют подвижность вихревой линии в микроскопической теории, а переходы между ними отвечают за диссипацию при движении вихря.

Более аккуратное описание квантовой механики квазичастиц можно построить на основе уравнений Боголюбова — де Жена

$$\left(-\frac{\hbar^2 \left(\nabla - \frac{ie}{\hbar c} \mathbf{A} \right)^2}{2m} - \frac{\hbar^2 k_F^2}{2m} + U \right) u + \Delta v = (\varepsilon \pm \beta H) u,$$

$$\left(\frac{\hbar^2 \left(\nabla + \frac{ie}{\hbar c} \mathbf{A} \right)^2}{2m} + \frac{\hbar^2 k_F^2}{2m} - U \right) v + \Delta^* u = (\varepsilon \pm \beta H) v,$$

записанных для электронной (u) и дырочной (v) компонент волновой функции квазичастиц. Ядро оператора щели $\hat{\Delta} u = \int \Delta(\mathbf{r}_1, \mathbf{r}_2) u(\mathbf{r}_2) d\mathbf{r}_2$ само определяется этими волновыми функциями, и в этом смысле мы получаем достаточно сложную нелинейную и нелокальную задачу. Для ее решения используется ряд эффективных квазиклассических методов теории сверхпроводимости, основанных на малости фермиевской длины волны электронов по сравнению с длиной когерентности.

Очевидно, что с увеличением магнитного поля приближение изолированной вихревой линии оказывается недостаточным для расчета спектра квазичастиц в вихревой решетке. При этом весьма важно учитывать туннелирование квазичастиц между соседними вихрями. Величину этого эффекта можно оценить с помощью стандартной квантово-механической теории возмущений в вырожденном спектре. Результаты такой оценки проиллюстрированы на рис. 10. Здесь изображено, как вырожденные уровни пары вихрей, находящихся на расстоянии a , расщепляются с учетом экспоненциально малого туннелирования. Сравнивая масштаб туннельного расщепления с межуровневым расстоянием, мы получаем критическое расстояние a_c между вихрями, на котором туннелирование должно приводить к существенной перестройке спектра с дальнейшим переходом

его к спектру типа Ландау для нормального металла. Соответствующую модификацию плотности состояний [12] можно обнаружить, используя сканирующую туннельную спектроскопию.

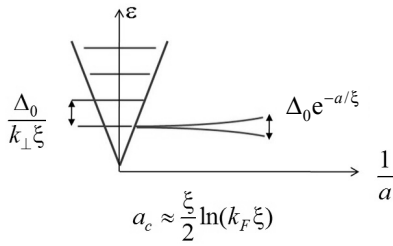


Рис. 10. Расщепление спектра Кароли — де Жена — Матрикона под влиянием тунnelирования квазичастиц между вихрями

Спектры квазичастиц в вихрях и вихревых решетках оказываются весьма чувствительны к типу сверхпроводящего спаривания, и в этом плане спектроскопия вихревого состояния представляет собой эффективный тест на определение устройства сверхпроводящей щели на поверхности Ферми. Упомянутая туннельная методика не единственный способ определения свойств квазичастиц, весьма эффективными в этом отношении могут считаться измерения теплопроводности или теплоемкости. Отметим также, что роль вихрей как ям для квазичастиц весьма важна для процессов релаксации в неравновесных состояниях, в результате вихри оказывают значительное влияние на работу разнообразных мезоскопических устройств, таких как холодильники на основе контактов «нормальный металл — изолятор — сверхпроводник» или одноэлектронные турникеты [13].

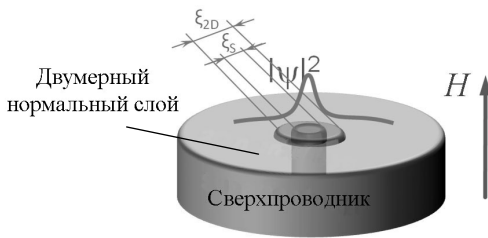


Рис. 11. Два масштаба кора вихря в системе с индуцированной сверхпроводимостью

В заключение остановимся на одном современном аспекте физики вихрей, исследование которого направлено на определение возможностей использования состояний квазичастиц в корах вихрей для квантовых вычислений. При этом для создания специального типа состояний квазичастиц (майорановских состояний) оказывается необходимым осуществлять инженерию сверхпроводящего параметра порядка в вихрях, что становится возможным в гибридных системах, состоящих из обычного сверхпроводника с *s*-типом спаривания и несверхпроводящей системы с нетриви-

альным устройством спектра электронов, такой как топологический изолятор. Обмен электронами между подсистемами приводит к так называемому эффекту близости и появлению индуцированного параметра порядка в исходно несверхпроводящей части структуры. Кор вихря в получающейся структуре оказывается двухмасштабным [14] (рис. 11), а мини-щель в спектре квазичастиц может быть подавлена, что позволяет реализовать состояние майорановского типа.

Литература

1. Шубников, Л. В. Магнитные свойства сверхпроводящих металлов и сплавов / Л. В. Шубников, В. И. Хоткевич, Ю. Д. Шепелев, Ю. Н. Рябинин // ЖЭТФ. — 1937. — Т. 7, вып. 2. — С. 221—237.
2. Абрикосов, А. А. О магнитных свойствах сверхпроводников второй группы // ЖЭТФ. — 1957. — Т. 32, вып. 6. — С. 1442—1452.
3. Мельников, А. С. Фазовые переходы в вихревых решетках гексагональных экзотических сверхпроводников // ЖЭТФ. — 1992. — Т. 101. — С. 1978—2000.
4. Samokhvalov, A. V. Attraction between pancake vortices and vortex molecule formation in the crossing lattices in thin films of layered superconductors / A. V. Samokhvalov, A. S. Mel'nikov, A. I. Buzdin // Phys. Rev. B. — 2012. — V. 85. — P. 184509.
5. Бараш, Ю. С. Внутренняя структура вихрей в экзотических сверхпроводниках вблизи нижнего критического поля / Ю. С. Бараш, А. С. Мельников // ЖЭТФ. — 1991. — Т. 100. — С. 307—322.
6. Babaev, E. Type-1.5 superconductivity in multicomponent systems / E. Babaev, J. Carlström, M. Silaev, J. M. Speight // Physica C: Superconductivity and its Applications. — 2017. — V. 533. — P. 20—35.
7. Beshpalov, A. A. Clustering of vortex matter in superconductor-ferromagnet superlattices / A. A. Beshpalov, A. S. Mel'nikov, A. I. Buzdin // Europhys. Lett. — 2015. — V. 110, iss. 3. — Art. 37003.
8. Mel'nikov, A. S. Mesoscopic superconductor as a ballistic quantum switch / A. S. Mel'nikov, V. M. Vinokur // Nature. — 2002. — V. 415. — P. 60—62.
9. Kopnin, N. B. Theory of Nonequilibrium Superconductivity / N. B. Kopnin. — N. Y. : Oxford University press, 2001. — 328 p.
10. Горьков, Л. П. Вязкое течение потока в сверхпроводниках с парамагнитными примесями / Л. П. Горьков, Н. Б. Копнин // ЖЭТФ. — 1971. — Т. 60. — С. 2331—2343.
11. Caroli, C. Bound Fermion states on a vortex line in a type II superconductor / C. Caroli, P. G. de Gennes, J. Matricon // Phys. Lett. — 1964. — V. 9. — P. 307.
12. Mel'nikov, A. S. Intervortex quasiparticle tunneling and electronic structure of multi-vortex configurations in type-II superconductors / A. S. Mel'nikov, M. A. Silaev // Письма в ЖЭТФ. — 2006. — Т. 83. — С. 675—680.
13. Taupin, M. Tunable quasiparticle trapping in Meissner and vortex states of mesoscopic superconductors / M. Taupin, I. M. Khaymovich, M. Meschke, A. S. Mel'nikov, J. P. Pekola // Nature Communications. — 2016. — V. 7. — Art. 10977. — doi:10.1038/ncomms10977.
14. Kopnin, N. B. Predicted Multiple Cores of a Magnetic Vortex Threading a Two-Dimensional Metal Proximity Coupled to a Superconductor / N. B. Kopnin, I. M. Khaymovich, A. S. Mel'nikov // Phys. Rev. Lett. — 2013. — V. 110. — P. 027003.

ОПРОКИДЫВАНИЕ ВИХРЕВЫХ ЛИНИЙ КАК ПРЕДТЕЧА КОЛМОГОРОВСКОЙ ТУРБУЛЕНТНОСТИ

Д. С. Агафонцев^{1,2}, Е. А. Кузнецов^{2,3,4}, А. А. Майлыбаев⁵

¹ Институт океанологии им. П. П. Ширшова РАН, г. Москва

² Новосибирский государственный университет

³ Физический институт им. П. Н. Лебедева РАН, г. Москва

⁴ Институт теоретической физики им. Л. Д. Ландау РАН, г. Москва

⁵ Национальный институт теоретической и прикладной математики,
г. Рио-де-Жанейро, Бразилия

Введение

Несмотря на то что со времени классических работ Л. Ф. Ричардсона [1], А. Н. Колмогорова [2] и А. М. Обухова [3], посвященных гидродинамической турбулентности при больших числах Рейнольдса, $Re \gg 1$, прошло уже более 75 лет и достигнуто значительное понимание ее природы, проблема развитой турбулентности остается все еще до конца не решенной. Главная причина состоит в том, что развитая гидродинамическая турбулентность ни в каком варианте не может быть изучена с помощью теории возмущений, в отличие, например, от волновой турбулентности.

В волновой турбулентности есть два основополагающих фактора, определяющих динамику волновой системы: линейная дисперсия волн и нелинейность. Если дисперсионные эффекты преобладают над нелинейными, то в этом случае каждая волна с частотой Ω и волновым вектором \mathbf{k} движется долгое время свободно и только на больших расстояниях $L \gg k^{-1}$ начинает испытывать влияние других волн из-за нелинейности. Это основание для применения к такому ансамблю волн статистического описания на основе теории возмущений. В результате мы приходим к теории слабой (волновой) турбулентности, описывающей поведение ансамбля волн на языке кинетических уравнений для волнового действия, представляющего собой классический предел чисел заполнения. На этом пути удается продвинуться достаточно далеко, в частности найти спектры турбулентности как точные решения кинетических уравнений, получившие названия спектров Колмогорова — Захарова (см., например, [4]).

Эти решения характеризуются постоянным потоком энергии, числа частиц и т. д. Существенно, что данные решения не могут быть рассмотрены как термодинамически равновесные — они, как и спектр Колмогорова — Обухова для развитой гидродинамической турбулентности, реализуются в инерционном интервале — промежуточной области между накачкой и затуханием — и характеризуются конечным значением потока того или иного интеграла движения. С ростом амплитуды (т. е. с усилением нелинейности) главную роль начинают играть когерентные структуры в виде солитонов, бризеров, вихрей, для которых нелинейность компенси-

руется дисперсионными эффектами. Такие объекты иногда оказываются устойчивыми, часто — для интегрируемых моделей типа КДВ или НУШ. При этом турбулентность в интегрируемых моделях, получившая название интегрируемой, обладает рядом особенностей (см. недавнюю работу [5]). В случае неустойчивости когерентных структур типичным сценарием их нелинейного развития является коллапс — образование особенности за конечное время (см. обзор [6] и цитируемую там литературу). Классические примеры — самофокусировка света в средах с керровской нелинейностью и опрокидывание волн звукового типа.

Если говорить о гидродинамике Эйлера, которая описывает турбулентность в инерционном интервале, то она ни в каком пределе не может быть изучена пертурбативно. Ее гамильтониан, совпадающий с полной кинетической энергией жидкости, представляет собой гамильтониан взаимодействия (см., например, [7]). Таким образом, гидродинамика Эйлера есть система с предельно сильным нелинейным взаимодействием. Следует ожидать поэтому, что для развитой гидродинамической турбулентности когерентные структуры должны играть более существенную, а возможно — определяющую роль, нежели для волновой турбулентности. Нужно отметить, что проблема взаимодействия когерентных структур и хаотических компонент — одна из центральных в современной теории турбулентности.

Известно, что теория Колмогорова — Обухова [2, 3] описывает развитую гидродинамическую турбулентность в инерционном интервале масштабов — промежуточной области между длинноволновой накачкой и вязким затуханием. Размер этой области растет пропорционально $Re^{3/4}$. В основе теории Колмогорова — Обухова лежат два важных предположения:

- турбулентность в инерционном интервале однородна и изотропна;
- нелинейное взаимодействие между флуктуациями является локальным.

Последнее означает, что взаимодействие между масштабами одного порядка превосходит взаимодействие между флуктуациями с сильно отличающимися масштабами. В случае стационарной турбулентности поведение системы определяется (постоянным) потоком энергии P из области накачки в область затухания (по сути, основано на свойстве локальности). Спектр турбулентности — так называемый колмогоровский спектр — имеет универсальное поведение, определяемое единственной величиной P . По этой причине с точностью до константы C_K (константы Колмогорова) спектр $E(k)$ может быть получен исходя из соображений размерности. По размерности спектр $E(k)$ представляет собой плотность энергии в фазовом пространстве, умноженную на $4\pi k^2$:

$$E(k) = 4\pi k^2 \frac{\rho c^2}{k^3} F\left(\frac{P}{\rho c^2 \cdot kc}\right),$$

где ρ — плотность (в дальнейшем равна 1), c — скорость света (введена из соображений удобства), а F — функция от безразмерного параметра $P/(\rho c^2 \cdot kc)$. Очевидно, что спектр $E(k)$ не должен зависеть от скорости света, отсюда функция $F(\xi) = C\xi^{2/3}$. В результате мы приходим к колмогоровскому спектру

$$E(k) = C_K P^{2/3} k^{-5/3}, \quad C_K = 4\pi C.$$

Данный вывод принадлежит Р. З. Сагдееву, в этом выводе нет ничего, кроме соображений размерности. Из этих же соображений следует, что время перекачки энергии из энергосодержащих масштабов L в область диссипации конечно и определяется только L и P : $T \sim L^{2/3} P^{-1/3}$. Флуктуации скорости и завихренности $\omega = [\nabla \times \mathbf{v}]$ на масштабе ℓ даются соответственно следующими соотношениями:

$$\langle \delta v \rangle \sim P^{1/3} \ell^{1/3}, \quad \delta \omega \sim P^{1/3} \ell^{-2/3}.$$

Таким образом, для флуктуаций ω мы имеем особенность при $\ell \rightarrow 0$ и конечном времени перекачки T , что указывает на возможность коллапса в трехмерной гидродинамике Эйлера.

1. Представление вихревых линий

В сжимаемой гидродинамике (газодинамике) появление особенности за конечное время (коллапс) хорошо известно, оно обязано опрокидыванию. В этом случае одна «жидкая» частица догоняет другую частицу, в результате чего происходит формирование бесконечных градиентов для характеристик газа — плотности и скорости. Главная причина опрокидывания связана со сжимаемостью газа. С математической точки зрения этот процесс представляет собой формирование складки, что может быть описано на языке отображений, в данном случае отвечающих переходу от эйлерова описания к лагранжеву. Особенность возникает в точке обращения якобиана данного отображения в ноль. В несжимаемой гидродинамике — гидродинамике Эйлера, казалось бы, нет причины для опрокидывания, поскольку якобиан преобразования от эйлерова описания к лагранжеву в силу несжимаемости тождественно равен 1. Несмотря на это, в несжимаемой гидродинамике есть сжимаемые объекты — непрерывно распределенные вихревые линии, что следует из простого наблюдения.

Рассмотрим уравнения движения завихренности $\omega = [\nabla \times \mathbf{v}]$ идеальной жидкости, так называемые уравнения Гельмгольца:

$$\frac{\partial \omega}{\partial t} = \text{rot} [\mathbf{v} \times \omega], \quad \text{div} \mathbf{v} = 0. \quad (1)$$

В силу векторного произведения только компонента скорости, нормальная к вихревой линии, \mathbf{v}_n , может изменить ω . При этом в ситуации общего положения $\text{div} \mathbf{v}_n \neq 0$, что является причиной сжимаемости непрерывно рас-

пределенных вихревых линий [9, 10]. Компонента скорости, параллельная завихренности, не меняет завихренность, обеспечивая несжимаемость для полной скорости, $\operatorname{div} \mathbf{v} = 0$.

Отметим, что уравнение (1) часто называют уравнением вмороженности. На самом деле вмороженность есть свойство этого уравнения, которое гласит, что всякая жидкая частица приклеена к своей вихревой линии и движется вместе с ней.

У частицы, таким образом, имеется только одна «свобода» — движение вдоль вихревой линии, которое, очевидно, не меняет завихренность. Поэтому \mathbf{v}_n есть скорость самой вихревой линии. Это утверждение имеет простое геометрическое объяснение. Всякая деформация вдоль произвольной линии не изменяет ее — только поперечные деформации приводят к ее перемещению. Поэтому движение вихревой линии определяется скоростью \mathbf{v}_n , находится из решения системы обыкновенных дифференциальных уравнений для «новых» лагранжевых траекторий:

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}_n(\mathbf{r}, t) \text{ при } \mathbf{r}(t=0) = \mathbf{a}, \quad (2)$$

и это решение задает сжимаемое отображение $\mathbf{r} = \mathbf{r}(\mathbf{a}, t)$. Последнее есть следствие формулы Лиувилля, примененной к этому уравнению,

$$\frac{dJ}{dt} = \operatorname{div} \mathbf{v}_n \cdot J, \quad (3)$$

где J — якобиан отображения. В силу $\operatorname{div} \mathbf{v}_n \neq 0$ на величину J не накладывается никаких ограничений: якобиан может принимать произвольные значения, в том числе и нулевые. Уравнение (1) в терминах этого отображения допускает интегрирование:

$$\boldsymbol{\omega}(\mathbf{r}, t) = \frac{(\boldsymbol{\omega}_0(\mathbf{a})\nabla_{\mathbf{a}})\mathbf{r}}{J}, \quad (4)$$

где $\boldsymbol{\omega}_0(\mathbf{a})$ — начальное значение завихренности, которое имеет смысл инвариантов Коши (см., например, [7]).

Уравнения (4) и (2) вместе с условием несжимаемости $\operatorname{div} \mathbf{v} = 0$ образуют замкнутую систему представления вихревых линий (ПВЛ), впервые введенного в работе [9] (см. также [8, 10]).

Позднее стало ясно, что существование сжимаемых распределений для бездивергентных полей присуще всем вмороженным полям (магнитное поле в идеальной МГД [15], ротор завихренности для двумерных течений — di-vorticity [16] и т. д.).

Поскольку завихренность в (4) в знаменателе содержит J , который принимает произвольные значения, то вероятность коллапса обеспечивает обращение якобиана J в ноль, что в ситуации общего положения может возникнуть сначала в отдельной точке. Такой сценарий оказался возможным для трехмерной интегрируемой гидродинамики [9, 11]. Трехмерная

интегрируемая гидродинамика получается из идеальной гидродинамики Эйлера в приближении локальной индукции. Эти уравнения интегрируются с помощью ПВЛ, в результате чего выясняется, что каждая вихревая линия представляет собой автономный объект, не взаимодействующий со всеми другими, но со своей нелинейной динамикой, которая описывается с помощью интегрируемого уравнения Ландау — Лифшица. Являясь свободной, каждая вихревая линия может догонять другую линию, т. е. происходит опрокидывание вихревых линий. Как результат, в некоторой точке за конечное время якобиан ПВЛ обращается в ноль, что приводит к появлению особенности для завихренности. Опрокидывание вихревых линий в трехмерной интегрируемой гидродинамике происходит благодаря сжимаемости вихревых линий, несмотря на несжимаемость потока [11].

Отметим, что до сих пор вопрос о коллапсе для гидродинамики Эйлера остается дискуссионным, несмотря на большое количество численных и точных аналитических результатов (см. обзоры [12, 13] и цитируемую там литературу). От решения проблемы «есть коллапс или его нет» зависит наше понимание природы развитой гидродинамической турбулентности.

Основной вывод, который можно сделать исходя из сказанного, состоит в том, что в гидродинамике Эйлера существуют сжимаемые объекты — это поле непрерывно распределенных вихревых линий. Интуитивно ясно, что сжимаемость вихревых линий должна обеспечивать появление структур типа ударных волн, которые вначале возникают из-за опрокидывания в отдельной точке, а затем область опрокидывания расширяется, приводя к формированию каустики. Именно такие структуры — структуры блинного типа — наблюдаются в наших численных экспериментах [17—19]. Следует отметить, что впервые структуры такого типа наблюдались в численных экспериментах М. Браше с соавторами (1992) [20].

Данная статья представляет собой обзор последних результатов, полученных нами в этом направлении. Основное внимание в статье уделяется численным результатам по моделированию процессов зарождения развитой ($Re \gg 1$) турбулентности в рамках уравнений Эйлера.

2. Численное моделирование

Вначале нужно сказать несколько слов об уравнениях, решаемых численно, и применяемых численных схемах. В численных экспериментах мы использовали прямое интегрирование уравнений Эйлера в терминах завихренности, т. е. уравнений (1), а также интегрирование уравнений Эйлера в представлении вихревых линий — формулы (2) и (4), которые были переписаны в эйлеровом представлении. Последнее означает, что рассматривалось не прямое отображение $\mathbf{r} = \mathbf{r}(\mathbf{a}, t)$, а обратное $\mathbf{a} = \mathbf{a}(\mathbf{r}, t)$, для которого уравнения движения в соответствии с (2) записываются в виде

$$\frac{\partial \mathbf{a}}{\partial t} + (\mathbf{v}_n \nabla) \mathbf{a} = 0. \quad (5)$$

Соотношение (4) в терминах $\mathbf{a}(\mathbf{r}, t)$ переписывается как

$$\boldsymbol{\omega}(\mathbf{r}, t) = \frac{1}{2} \varepsilon_{\alpha\beta\gamma} \omega_{0\alpha}(\mathbf{a}) [\nabla a_\beta \times \nabla a_\gamma]. \quad (6)$$

Уравнения (1), а также (5), (6) численно интегрировались в периодическом ящике $\mathbf{r} = (x, y, z) \in [-\pi, \pi]^3$ с использованием псевдоспектрального метода Рунге — Кутты четвертого порядка точности. Чтобы избежать влияния неустойчивости на больших волновых числах (bottle-neck instability), использовалась фильтрация в фурье-пространстве на каждом шаге по времени с помощью функции отсекающей, предложенной в работе [14],

$$\rho(\mathbf{k}) = \exp\left(-36 \sum_{j=x,y,z} (k_j / K_{\max}^{(j)})^{36}\right), \quad (7)$$

где $\mathbf{k} = (k_x, k_y, k_z)$ — волновой вектор, $K_{\max}^{(j)} = N_j/2$ — максимальные волновые числа и N_j — числа узлов решетки вдоль направлений $j = x, y, z$, причем $|k_j| \leq K_{\max}^{(j)}$. Функция (7) отсекает примерно 20 % спектра по краям спектрального пространства в каждом из трех направлений. Обратный оператор ротора в уравнении (1) вычислялся в фурье-пространстве (см., например, [17]). Адаптивный шаг по времени определялся с помощью критерия устойчивости CFL (Куранта — Фридрихса — Леви) с числом Куранта 0,5.

Мы использовали адаптивную анизотропную прямоугольную решетку, которая является однородной в каждом направлении и подстраивается независимо вдоль каждой из трех координат. Алгоритм адаптации происходит от стандартного правила деалиасинга, оптимизированного для квадратичной нелинейности уравнений Эйлера. На малых временах фурье-спектр решения сосредоточен на малых гармониках, тогда как высшие гармоники содержат численный шум. Мы отслеживали границу «сигнал — шум» до тех пор, пока она не достигнет $2K_{\max}^{(j)}/3$ по любому из трех направлений $j = x, y, z$ (детали см. в [17]). Затем решетку переопределяли вдоль соответствующего направления, используя фурье-интерполяцию, которая имеет ошибку, сравнимую с ошибкой округления. Пока моделирование выполняется таким образом, ошибка алиасинга исключается, а влияние фурье-фильтрации (7) на «содержащие сигнал» гармоники пренебрежимо мало.

Все наши численные эксперименты начинались на кубической решетке 128^3 . Когда общее число узлов достигало 1024^3 , мы продолжали эксперимент на фиксированной решетке до тех пор, пока фурье-спектр завихренности на $2K_{\max}^{(j)}/3$ не превысит 10^{-13} от своего максимального значения вдоль любого из трех направлений j (см. [17]).

Начальные условия для всех численных экспериментов выбирались в виде сдвигового течения $v_x = \sin z$; $v_y = \cos z$ (вырожденное ABC-течение)

и случайных возмущений, спектр которых был сосредоточен в области малых волновых чисел. Мы изменяли соотношение между регулярной (*ABC*-течение) и случайной компонентами. В первых двух экспериментах [17] это соотношение (по энергии) было порядка единицы. В последующих экспериментах [18] это соотношение изменялось в широких границах.

3. Численные результаты

Во всех численных экспериментах мы наблюдали формирование структур блинного типа (вихревых слоев), толщины которых значительно меньше продольных размеров блина. Все эти структуры представляют собой области повышенной завихренности. Для каждой такой структуры можно определить ее размеры с помощью разложения вблизи локального максимума завихренности ω_{\max} :

$$|\omega| = \omega_{\max} + \frac{1}{2} H_{ij} \Delta X_i \Delta X_j, \quad (8)$$

где $\Delta \mathbf{X} = \mathbf{x} - \mathbf{x}_{\max}$. В этом случае собственные значения λ_n матрицы вторых производных $H_{ij} = \partial_i \partial_j |\omega|$, вычисленных в локальном максимуме завихренности, будут определять характерные размеры структуры: $\ell_n = \sqrt{2\omega_{\max} / |\lambda_n|}$. Максимальное значение $|\lambda_1|$ ($|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$) задает толщину блина, а собственный вектор, соответствующий этому собственному значению, — направление нормали к структуре типа блина. В собственных осях H_{ij} разложение (8) записывается в виде

$$|\omega| = \omega_{\max} \left(1 - \sum_{n=1,2,3} \frac{a_n^2}{\ell_n^2} \right).$$

Следует отметить, что численное определение локальных максимумов завихренности является нетривиальной задачей из-за специфической геометрии блинообразных структур (соответствующие методы описаны в [17]).

В данном разделе мы приведем результаты численных экспериментов только для одних начальных условий (н. у.), помеченных в работе [17] как I_1 . Для этих н. у. соотношение между регулярной (*ABC*-течение) и случайной компонентами было порядка единицы. В работе [17] были представлены результаты формирования структур блинного типа для максимального времени счета $t_{m1} = 6,89$ с максимальной решеткой $486 \times 1024 \times 2048$. В работе [19] для тех же начальных условий достигнуто максимальное время $t_{m2} = 7,75$ с решеткой $972 \times 2048 \times 4096$, в восемь раз большей по сравнению с решеткой в работе [17]. Для каждого из двух экспериментов разрешение решетки соответствовало 10—12 точкам по толщине самого тонкого блина (на уровне полумаксимума).

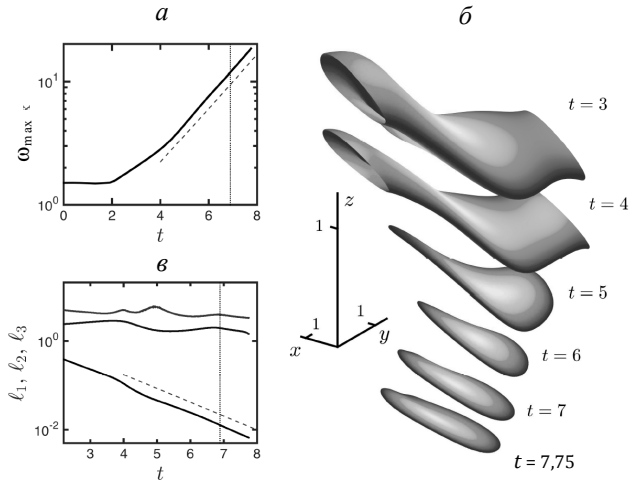


Рис. 1. Глобальный максимум завихренности как функция времени (а): по вертикали логарифмическая шкала; пунктирная линия показывает наклон, пропорциональный $e^{\beta_2 t}$, с $\beta_2 = 0,5$; тонкая вертикальная линия отвечает максимальному времени моделирования $t_{m1} = 6,89$ [17]. Поведение модуля завихренности $\omega \geq 0,85\omega_{\max}$ в разные моменты времени (б); различные оттенки черного цвета представляют собой величины завихренности в средней плоскости блина: от темного для $0,85\omega_{\max}(t)$ до светлого для ω_{\max} ; структуры сдвинуты по вертикали для лучшей визуализации. Характерные пространственные масштабы ℓ_1 (верхняя кривая), ℓ_2 (средняя кривая) и ℓ_3 (нижняя кривая) как функции времени (в); пунктирная линия указывает на наклон, пропорциональный $e^{-\beta_1 t}$, с $\beta_1 = 0,74$

На рисунке 1 представлена эволюция структуры, соответствующей глобальному максимуму завихренности. Из этих данных можно видеть, что область повышенной завихренности (рис. 1, б) представляет собой сжимающуюся в одном направлении структуру типа блина с растущей завихренностью. Толщина блина уменьшается экспоненциально с декрементом β_1 (рис. 1, в). Другие два размера изменяются незначительно. Максимальная завихренность растет экспоненциально со временем с инкрементом β_2 (рис. 1, а). Отношение показателей $\beta_2/\beta_1 \approx 2/3$. Таким образом, между максимальной завихренностью ω_{\max} и толщиной блина ℓ_1 возникает следующее соотношение колмогоровского типа:

$$\omega_{\max}(t) \sim \ell_1(t)^{-2/3}.$$

Этот результат был вначале получен численно в работе [17], а затем подтвержден для большего интервала времени в работе [19].

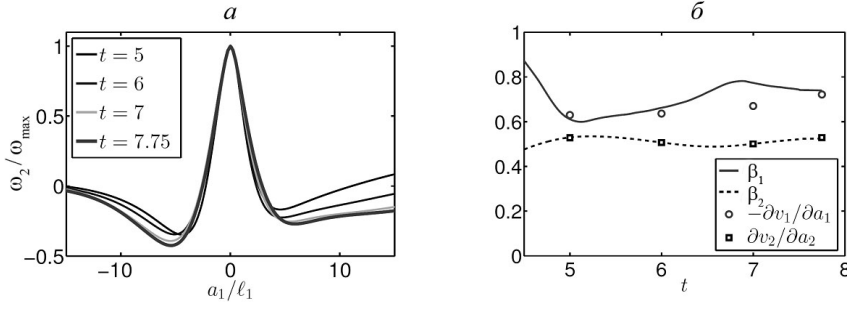


Рис. 2. Нормированная компонента завихренности ω_2 / ω_{\max} как функция автомодельной переменной $\xi = a_1 / \ell_1$ в разные моменты времени (а): t у кривых растет сверху вниз. Сравнение логарифмических производных $\beta_1 = -\dot{\ell}_1 / \ell_1$ и $\beta_2 = \dot{\omega}_{\max} / \omega_{\max}$ с градиентами скорости $-\partial v_1 / \partial a_1$ и $\partial v_2 / \partial a_2$, полученными численно в точке глобального максимума завихренности (б)

Численное моделирование показывает, что вектор завихренности в основном лежит в плоскости блина ($\|\mathbf{n}_2$). Другие компоненты завихренности оказываются малыми. Рисунок 2, а дает представление, как изменяется при прохождении глобального максимума завихренности компонента ω_2 , нормированная на максимальное значение завихренности ω_{\max} , в зависимости от координаты a_1 вдоль нормали к блину. С хорошей точностью эта зависимость демонстрирует автомодельное поведение.

3.1. Точное решение уравнений Эйлера как модель эволюции блинообразных структур

Для описания эволюции блинных структур было найдено точное решение уравнений Эйлера [19]. Пусть в декартовой системе $\mathbf{a} = a_1 \mathbf{n}_1 + a_2 \mathbf{n}_2 + a_3 \mathbf{n}_3$ завихренность зависит только от a_1 и имеет одну компоненту, параллельную оси a_2 . Тогда уравнение (1) допускает следующее решение:

$$\boldsymbol{\omega}(\mathbf{a}, t) = \omega_2 \mathbf{n}_2, \quad \omega_2 = \Omega(t) f' \left(\frac{a_1}{\ell_1(t)} \right), \quad (9)$$

$$\mathbf{v}(\mathbf{a}, t) = -\Omega(t) \ell_1(t) f \left(\frac{a_1}{\ell_1(t)} \right) \mathbf{n}_3 + \begin{pmatrix} -\beta_1(t) a_1 \\ \beta_2(t) a_2 \\ \beta_3(t) a_3 \end{pmatrix}. \quad (10)$$

Здесь $\Omega(t)$ есть амплитуда завихренности, $\ell_1(t)$ — толщина блина, $f(\xi)$, зависящая от автомодельной переменной $\xi = a_1 / \ell_1(t)$, является произ-

вольной функцией. Параметры $\beta_1(t)$, $\beta_2(t)$ и $\beta_3(t)$, определяющие растягивающее поле скорости при больших a , имеют вид

$$\beta_1 = -\dot{\ell}_1/\ell_1, \quad \beta_2 = \dot{\Omega}/\Omega, \quad -\beta_1 + \beta_2 + \beta_3 = 0, \quad (11)$$

где точка означает дифференцирование по времени. Скорость (10) представляет собой суперпозицию сдвигового течения $-\Omega\ell_1 f(a_1/\ell_1)\mathbf{n}_3$ и асимметричного потенциального растягивающего течения $(-\beta_1 a_1, \beta_2 a_2, \beta_3 a_3)$. Давление в этом случае задается выражением, определяемым только параметрами β_i :

$$p = p_0 + (\dot{\beta}_1 - \beta_1^2)\frac{a_1^2}{2} - (\dot{\beta}_2 + \beta_2^2)\frac{a_2^2}{2} - (\dot{\beta}_3 + \beta_3^2)\frac{a_3^2}{2}. \quad (12)$$

Данное решение представляет собой обобщение решения, найденного в частном случае в работе [20]: $\Omega = \ell_1^{-1} = e^{t/T}$, $\beta_{1,2} = 1/T$, $\beta_3 = 0$.

Следует отметить, что распределение (9)—(12) имеет бесконечную энергию в \mathbf{R}^3 . В частности, ввиду произвольности $\Omega(t)$ и $\ell_1(t)$ легко находится решение, приводящее к появлению особенности за конечное время. Найденное новое решение, помимо произвольной функции $f(\xi)$, характеризуется безразмерным параметром асимметрии растягивающего потока

$$\sigma = \frac{\beta_2 - \beta_3}{\beta_2 + \beta_3} = \frac{2\beta_2}{\beta_1} - 1. \quad (13)$$

В наших численных экспериментах наблюдается почти экспоненциальное поведение для $\Omega(t)$ и для толщины $\ell_1(t)$ (см. рис. 1, а, в), что соответствует постоянным значениям β_1 , β_2 и β_3 . В этом случае параметр асимметрии определяет степенной закон между амплитудой завихренности и толщиной $\ell_1(t)$:

$$\Omega(t) \propto \ell_1(t)^{-\zeta}, \quad \zeta = \frac{\beta_2}{\beta_1} = \frac{\sigma + 1}{2}. \quad (14)$$

В поперечном направлении для данного решения вариация скорости в плоскости блина на масштабе ℓ_1 ведет себя как $\delta v_3 \propto \Omega \ell_1 \propto \ell_1^{1-\zeta}$. Эта вариация стремится к нулю при $\ell_1 \rightarrow 0$ и $\zeta < 1$.

Следует отметить, что анзац (9)—(12) может быть применен и в случае уравнения Навье — Стокса, если считать функцию f зависящей также от времени. В этом случае f определяется из уравнения $f_t - \frac{\nu}{\ell_1^2} f_{\zeta\zeta} = 0$, где ν — коэффициент вязкости. После простой замены это уравнение сводится к уравнению теплопроводности. В случае осесимметричных течений

($\beta_2 = \beta_3 = \beta_1/2$) данное решение переходит в решение в виде вихря Лундгрена [21].

В работе [19] было проверено, что решение (9)—(12) хорошо согласуется с результатами численного моделирования. Мы провели несколько проверок. Как это видно из рис. 2, *a*, автомодельная зависимость $f(\xi)$ (вдоль оси a_1) хорошо согласуется с решением. Область автомодельности при больших ξ должна сшиваться с основным течением. Важно отметить, что функция $f(\xi)$ не является универсальной, она меняется при переходе от одного блина к другому, а также при изменении начальных условий.

Другой тест связан со сравнением градиентов скорости. Для решения (9)—(12)

$$[\partial v_i / \partial a_j] = \begin{pmatrix} -\beta_1 & 0 & 0 \\ 0 & \beta_2 & 0 \\ -\omega_2 & 0 & \beta_3 \end{pmatrix}. \quad (15)$$

На рисунке 2, *b* представлено сравнение при разных временах логарифмических производных $\beta_1 = -\dot{\ell}_1 / \ell_1$ и $\beta_2 = \dot{\omega}_{\max} / \omega_{\max}$ с градиентами скорости $-\partial v_1 / \partial a_1$ и $\partial v_2 / \partial a_2$, найденными в точке глобального максимума завихренности. Можно видеть очень хорошее общее согласие, за исключением большего отклонения для $-\partial v_1 / \partial a_1$ при $t = 7$. Это отклонение может быть связано с изменением (до 20 %) производной скорости в пределах толщины блина. В заключительный момент времени $t_{m2} = 7,75$ в точке глобального максимума завихренности численно градиент скорости равен

$$[\partial v_i / \partial a_j]_{a=0} = \begin{pmatrix} -0,72 & 0,04 & -0,03 \\ -0,11 & 0,53 & -0,09 \\ -18,42 & -0,04 & 0,19 \end{pmatrix}, \quad (16)$$

откуда видно, что имеется единственная большая компонента (3, 1), $\partial v_3 / \partial a_1 \approx -\omega_{\max}$. Диагональные элементы очень хорошо согласуются с параметрами $-\beta_1 = -0,74$, $\beta_2 = 0,53$ и $\beta_3 = 0,21$, в то время как все оставшиеся компоненты: (1, 2), (1, 3), (2, 1), (2, 3) и (3, 2) — оказываются малыми.

Таким образом, в области максимума завихренности численное моделирование дает очень хорошее согласие с точным решением. При больших расстояниях вдоль вихревого слоя его нормаль поворачивается (см. рис. 1, *b*) и согласие с точным решением нарушается. В действительности в каждой точке вихревого слоя при введении соответствующей системы координат, ориентированной по нормали, точное решение локально хорошо согласуется с численным результатом моделирования.

Решение (9)—(12) не определяет связи между параметрами β_1 и β_2 , которые в численном эксперименте приводят к зависимости колмогоровского типа между максимальной завихренностью и толщиной вихревого слоя,

$$\omega_{\max}(t) \sim \ell_1(t)^{-\zeta}, \quad \zeta \approx 2/3. \quad (17)$$

По нашему мнению, параметр ζ должен определяться исходя из сшивки точного решения (9)—(12) с областью, по размеру значительно превышающей толщину вихревого слоя, т. е. нелокальным образом. Этот вопрос пока остается открытым.

К сказанному следует добавить, что поведение завихренности во всех локальных максимумах удовлетворительно укладывается на зависимость $2/3$ для наклона на логарифмической плоскости $(\omega_{\max}(t), \ell_1(t))$ (рис. 3).

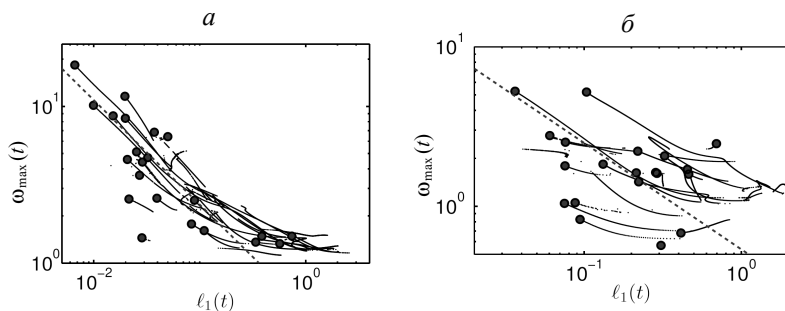


Рис. 3. *a* — Соотношение между локальными максимумами завихренности $\omega_{\max}(t)$ и соответствующими характерными толщинами $\ell_1(t)$. Линии показывают эволюцию максимумов, жирные точки соответствуют времени t_{m2} . Пунктирная линия соответствует степенному закону $\omega_{\max} \propto \ell_1^{-2/3}$. *б* — Такой же график для других случайных начальных условий с финальной решеткой $1152 \times 972 \times 864$ (с меньшим пространственным разрешением)

3.2. Сжимаемость отображения ПВЛ

В данном разделе мы приведем результаты численного моделирования уравнений Эйлера в представлении вихревых линий, т. е. в рамках уравнений (5) и (6).

Как отмечалось ранее, представление вихревых линий указывает на сжимаемость вихревых линий. Причем в силу (4) поведение завихренности существенно зависит от якобиана J . Особенно эта корреляция должна проявляться при уменьшении J и, соответственно, росте завихренности в ее максимальной точке. На рисунке 4 представлена зависимость от времени минимума якобиана (с хорошей точностью соответствует максимуму завихренности) для н. у. I_2 [17]. Логарифмическая шкала показывает, что

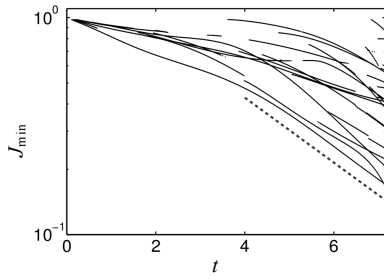


Рис. 4. Эволюция в логарифмической шкале локальных минимумов якобиана. Пунктирная линия соответствует наклону, пропорциональному e^{-t/T_J} , $T_J = 2,9$

$J_{\min} \propto e^{-t/T_J}$ с $T_J \approx 2,9$. Это значение практически совпадает со временем нарастания максимальной завихренности $T_\omega \approx 2,7$. Вычисления в рамках ПВЛ также хорошо соотносятся с прямым интегрированием уравнений Эйлера. На рисунке 5 представлены две зависимости — логарифмической производной максимальной завихренности $d/dt[\log \omega_{\max}]$ и вычисленной в точке максимума завихренности дивергенции нормальной компоненты скорости, взятой со знаком минус, $-\text{div } \mathbf{v}_n$. Согласно (3) последняя величина есть логарифмическая производная якобиана с соответствующим знаком. Обе эти зависимости сильно взаимно коррелируют. Данные численные результаты показывают, что формирование структур блинного типа обязано сжимаемости вихревых линий. При этом минимуму якобиана соответствует максимум завихренности.

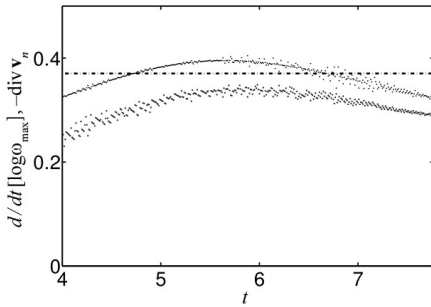


Рис. 5. Эволюция логарифмической производной $d/dt[\log \omega_{\max}]$ (верхняя линия) и $-\text{div } \mathbf{v}_n$ (нижняя линия), вычисленных в точке глобального максимума завихренности. Штрихпунктирная прямая линия соответствует инкременту $1/T_\omega$ роста ω_{\max} , где $T_\omega \approx 2,7$

3.3. Спектры турбулентности

В настоящее время в численных экспериментах мы наблюдаем формирование структур блинного типа для всех рассматриваемых начальных условий (более 30) [17—19]. В ходе эволюции всех таких структур с высокой завихренностью выполняется в той или иной степени соотношение колмогоровского типа $\omega_{\max}(t) \sim \ell(t)^{-2/3}$. В нашей первой работе [17] пока-

зано, что блины в x -пространстве порождают сильно анизотропные распределения в k -пространстве в виде джетов, вытянутых в направлениях, перпендикулярных блинам (рис. 6). Поскольку при эволюции блинов их толщина уменьшается, то, соответственно, анизотропия джетов растет со временем. Из-за этого мгновенный спектр турбулентности оказывается сильно изрезанным и анизотропным. Число структур блинного типа увеличивается, и, соответственно, растет число джетов. Каждый из джетов имеет свое угловое распределение порядка $(kL)^{-1}$, где L — характерный продольный размер блина. С увеличением количества джеты начинают перекрываться в k -пространстве. Когда таких перекрытий достаточно много, в этих областях происходит формирование колмогоровского спектра $E(k) \sim k^{-5/3}$ (рис. 7). Начальные условия IC_1 [18] не содержат никакой начальной анизотропии. По этой причине после усреднения по углу колмогоровский спектр не наблюдается. Начальные условия IC_2 и IC_3 представляют собой суперпозицию сдвигового течения (вырожденное ABC -течение)

Рис. 6. Изоповерхности модуля Фурье амплитуд завихренности (нормированных на максимальное значение) $|\tilde{\omega}(\mathbf{k})| = 0,2$ в k -пространстве при t_m . Сплошные линии показывают максимальные k -векторы (нормированные на ℓ_1^{-1} для всех джетов)

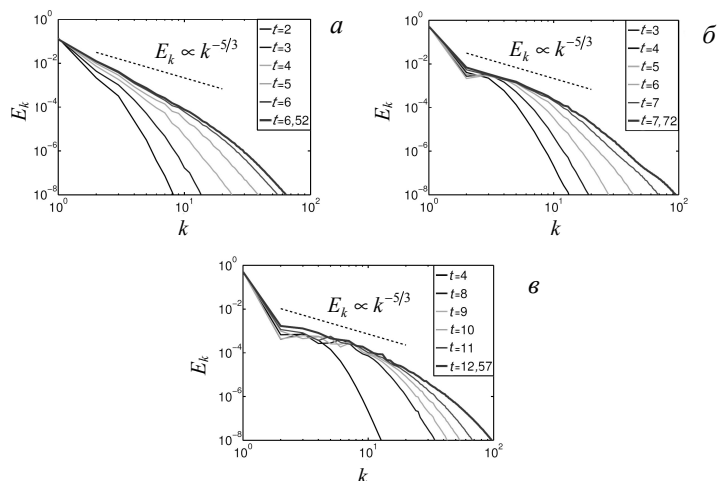
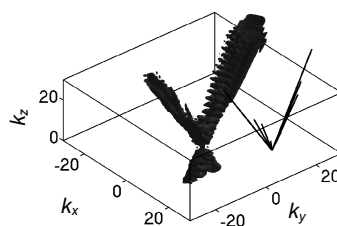


Рис. 7. Спектр энергии $E_k(t)$ в разные моменты времени для н. у. [18] IC_1 (тип I) (а), IC_2 (тип II) (б) и IC_3 (тип III) (в). Время у кривых растет снизу вверх

и изотропного гауссова шума. Анизотропия, порождаемая сдвиговым течением, обуславливает появление джетов с сильным угловым перекрытием, благодаря чему формируется колмогоровский спектр. Следует подчеркнуть, что неусредненный по углу спектр оказывается сильно анизотропным. Таким образом, несмотря на колмогоровское поведение спектра, турбулентность в условно инерционном интервале далека от изотропной и однородной, по крайней мере для времен численного эксперимента.

Заключение

В данной работе на основе представления вихревых линий, точного решения и прямого численного интегрирования уравнений Эйлера показано следующее:

- на стадии возбуждения спектр турбулентности в инерционном интервале далек от изотропного;
- основной вклад в спектр связан с появлением когерентных структур в виде блинов, которым в спектре турбулентности соответствуют джеты с растущей со временем анизотропией;
- максимальная завихренность в блине и его ширина ℓ связаны универсальным соотношением колмогоровского типа $\omega_{\max} \sim \ell^{-2/3}$, при этом рост завихренности и сужение структур блинного типа зависят от времени экспоненциально, без какой-либо тенденции к взрывному поведению;
- найденное точное решение уравнений Эйлера хорошо согласуется с численным моделированием в области блина;
- зарождение и развитие структур блинного типа является следствием сжимаемости вихревых линий;
- наблюдаемое увеличение числа таких структур приводит к формированию колмогоровского спектра для полностью невязких течений.

Данная работа была поддержана Российским фондом фундаментальных исследований (грант 17-01-00622). Численные эксперименты выполнялись в Новосибирском суперкомпьютерном центре (НГУ) и информационном центре Института фундаментальной и прикладной математики (ИМРА, Рио-де-Жанейро). Д. С. Агафонцев благодарен за поддержку со стороны ИМРА во время визитов в Бразилию.

Литература

1. *Richardson, L. F.* Atmospheric diffusion shown on a distance-neighbor graph // Proc. R. Soc. London. Ser. A. — 1926. — V. 110. — P. 709—737.
2. *Колмогоров, А. Н.* Локальная структура турбулентности в несжимаемых вязких жидкостях при очень больших числах Рейнольдса // ДАН СССР. — 1941. — Т. 30. — С. 299—303.
3. *Обухов, А. М.* Спектральное распределение энергии в турбулентном потоке // ДАН СССР. — 1941. — Т. 32. — С. 22—24.

4. *Zakharov, V. E.* Kolmogorov Spectra of Turbulence. I. Wave Turbulence / V. E. Zakharov, V. S. Lvov, G. E. Falkovich. — Berlin ; Heidelberg : Springer-Verlag, 1992. — (Series in nonlinear dynamics).
5. *Agafontsev, D. S.* Integrable turbulence and formation of rogue waves / D. S. Agafontsev, V. E. Zakharov // Nonlinearity. — 2015. — V. 28. — P. 2791.
6. *Захаров, В. Е.* Солитоны и коллапсы — два сценария эволюции нелинейных волновых систем / В. Е. Захаров, Е. А. Кузнецов // УФН. — 2012. — Т. 182. — С. 569—592.
7. *Захаров, В. Е.* Гамильтоновский формализм для нелинейных волн / В. Е. Захаров, Е. А. Кузнецов // УФН. — 1997. — Т. 167. — С. 1137—1168.
8. *Kuznetsov, E. A.* Vortex line representation for the hydrodynamic type equations // J. Nonl. Math. Phys. — 2006. — V. 13. — P. 64—80.
9. *Кузнецов, Е. А.* Гамильтонова динамика вихревых линий в системах гидродинамического типа / Е. А. Кузнецов, В. П. Рубан // Письма в ЖЭТФ. — 1998. — Т. 67. — С. 1015—1020.
10. *Кузнецов, Е. А.* Представление вихревых линий для идеальных и вязких жидкостей // Письма в ЖЭТФ. — 2002. — Т. 76. — С. 406—410.
11. *Кузнецов, Е. А.* Коллапс вихревых линий в гидродинамике / Е. А. Кузнецов, В. П. Рубан // ЖЭТФ. — 2000. — Т. 118. — С. 853—876.
12. *Chae, D.* Incompressible Euler Equations: the blow-up problem and related results // Handbook of Differential Equations: Evolutionary Equation / Eds. C. M. Dafermos and M. Pokorný. — Elsevier, 2008. — V. 4. — P. 1—55.
13. *Gibbon, J. D.* The three-dimensional Euler equations: Where do we stand? // Physica D. — 2008. — V. 237, № 14—17. — P. 1894—1904.
14. *Hou, T. Y.* Computing nearly singular solutions using pseudo-spectral methods / T. Y. Hou, R. Li // J. Comp. Phys. — 2007. — V. 226, № 1. — P. 379—397.
15. *Kuznetsov, E. A.* Compressible dynamics of magnetic field lines for incompressible MHD flows / E. A. Kuznetsov, T. Passot, P. L. Sulem // Phys. Plasmas. — 2004. — V. 11. — P. 1410—1415.
16. *Kuznetsov, E. A.* Effects of sharp vorticity gradients in two-dimensional hydrodynamic turbulence / E. A. Kuznetsov, V. Naulin, A. H. Nielsen, J. J. Rasmussen // Phys. Fluids. — 2007. — V. 19. — P. 105110.
17. *Agafontsev, D. S.* Development of high vorticity structures in incompressible 3D Euler equations / D. S. Agafontsev, E. A. Kuznetsov, A. A. Mailybaev // Phys. Fluids. — 2015. — V. 27. — Art. 085102 ; arXiv:1502.01562.
18. *Агафонцев, Д. С.* Развитие структур развитой завихренности в несжимаемых трехмерных уравнениях Эйлера: влияние начальных условий / Д. С. Агафонцев, Е. А. Кузнецов, А. А. Майлыбаев // Письма в ЖЭТФ. — 2016. — Т. 104. — С. 695—700.
19. *Agafontsev, D. S.* Asymptotic solution for high vorticity regions in incompressible 3D Euler equations / D. S. Agafontsev, E. A. Kuznetsov, A. A. Mailybaev // J. Fluid Mech. — 2017. — V. 813. — R1-1—R1-10. — doi:10.1017/jfm.2017.1 ; arXiv:1609.07782.
20. *Brachet, M. E.* Numerical evidence of smooth self-similar dynamics and possibility of subsequent collapse for three-dimensional ideal flows / M. E. Brachet, M. Meneguzzi, A. Vincent, H. Politano, P. L. Sulem // Phys. Fluids. — 1992. — V. A 4. — P. 2845—2854.
21. *Lundgren, T. S.* Strained spiral vortex model for turbulent fine structure // Phys. Fluids. — 1982. — V. 25. — P. 2193.

Научное издание

НЕЛИНЕЙНЫЕ ВОЛНЫ' 2016

Утверждено к печати ФИЦ «Институт прикладной физики РАН».
603950, Нижний Новгород, ул. Ульянова, 46

Редактор *И. А. Кокорина*
Верстка *М. В. Башевой, А. А. Ереминой*
Технический редактор *Д. П. Семенова*

Подписано к печати 25.04.2017 г.
Формат 60 × 90 ¹/₁₆. Бумага офсетная № 1.
Усл. печ. л. 20,0. Уч.-изд. л. 19,6. Темплан 2017 г. Поз. 1. Тираж 200 экз.

Отпечатано в полном соответствии с качеством
предоставленных диапозитивов в ГУП ППП «Типография «Наука»», АИЦ РАН
121099 Москва, Шубинский пер., 6

Заказ №