

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ,
ИНСТИТУТ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК

**ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ
В ЕСТЕСТВЕННЫХ НАУКАХ
МЕТОДЫ СУПЕРКОМПЬЮТЕРНОГО
МОДЕЛИРОВАНИЯ**

Часть 2

СБОРНИК ТРУДОВ

21–23 апреля 2015 г., Россия, Таруса

Под редакцией
Р.Р. Назирова, Л.Н. Щура

Серия «Механика, управление и информатика»

МОСКВА
ИКИ РАН
2015

Вычислительные технологии в естественных науках.

Методы суперкомпьютерного моделирования

Часть 2

Сборник трудов

Россия, Таруса, 21–23 апреля 2015 г. Под редакцией Р.Р. Назирова, Л.Н. Щура

Статьи, вошедшие в настоящий сборник, основаны на докладах, представленных на расширенном семинаре «Методы суперкомпьютерного моделирования», который проходил 21–23 апреля 2015 г. в г. Таруса на базе гостиницы «Интеркосмос» Государственного бюджетного учреждения науки Института космических исследований РАН (ИКИ РАН). Это очередной семинар из серии расширенных семинаров, посвящённых вычислительным технологиям в естественных науках. На первом семинаре рассматривались вопросы компьютерного моделирования актуальных задач физики и механики, его труды изданы в 2009 г. в выпуске 1 (Труды семинара по вычислительным технологиям в естественных науках. Вып. 1. Вычислительная физика / Под ред. Р.Р. Назирова. М.: КДУ, 2009. 288 с.); второй семинар был посвящён обсуждению физических и медицинских аспектов исследования сложных нелинейных физических процессов в организме человека (Методы нелинейного анализа в кардиологии и онкологии: Физические подходы и клиническая практика. Вып. 2 / Под ред. Р.Р. Назирова. М.: КДУ, 2010. 206 с.); на третьем семинаре обсуждались проблемы компьютерного моделирования задач биологии и химии (Труды семинара по вычислительным технологиям в естественных науках. Вып. 3. Вычислительная химия и биология / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: КДУ, 2010. 118 с.); на четвёртом — проблемы технического зрения (Техническое зрение в системах управления мобильными объектами-2010: Труды научно-технической конференции-семинара. Вып. 4. / Под ред. Р.Р. Назирова. М.: КДУ, 2011, 328 с.). На четвертом семинаре обсуждались вопросы моделирования событий глобальных масштабов (Труды семинара по вычислительным технологиям в естественных науках. Системы глобального масштаба / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: ИКИ РАН, 2011, 148 с.). Пятый семинар был посвящён теме исследований в области перспективных компьютерных систем, с акцентом на обсуждение концептуальных подходов, разработке новых методов и устройств (Труды семинара Вычислительные технологии в естественных науках. Перспективные компьютерные системы: устройства, методы и концепции. / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: ИКИ РАН, 2011, 192 с.). На шестом семинаре были обсуждены алгоритмы и методы в области вычислительной физики, причем особый акцент делался на обсуждении полученных новых физических результатов. Материалы этого семинара можно найти на страницах семинара <http://www.iki.rssi.ru/seminar/>. Седьмой семинар получил название «Инфокоммуникационные технологии в научных исследованиях» и его материалы были изданы в двойном номере журнала «Информационное общество» (выпуски 1–2, 2013). Восьмой семинар, как и нынешний, получил название «Методы суперкомпьютерного моделирования». Это связано с тем, проведение семинаров с такой тематикой поддержано в рамках выполнения научных исследований по гранту РНФ 14-21-00158 «Разработка алгоритмов и методов для задач математического моделирования на суперкомпьютерных системах, включая гибридные» Важная особенность расширенных семинаров состоит в привлечении молодежи к активному участию в качестве докладчиков. Как правило, более половины докладов делаются молодыми исследователями.

Ключевые слова: математическое моделирование, гибридные компьютерные системы, моделирование на суперкомпьютерах, методы и алгоритмы параллельных вычислений.

Редактор:

Компьютерная верстка: *Титова А.Ю.*

Дизайн обложки: *Захаров А.Н.*

Электронная версия сборника размещена на сайте ИКИ РАН <http://www.iki.rssi.ru/print.htm>.

Отдельные материалы даны в авторской редакции.

СОДЕРЖАНИЕ

<i>Назиров Р.Р., Щур Л.Н.</i> Введение «Методы суперкомпьютерного моделирования»	5
<i>Алешин С.В.</i> Пространственно неоднородные решения уравнения Колмогорова-Петровского-Пискунова с запаздыванием	11
<i>Антонов А.С., Теплов А.М.</i> Анализ масштабируемости параллельных программ на основе технологий суперкомпьютерного кодизайна	18
<i>Бараш Л.Ю., Щур Л.Н., Боровский М.</i> Алгоритм «Population Annealing» и его применение к двумерной модели Изинга	29
<i>Белов А.В., Пашков Д.В.</i> Синтез алгоритмов построения доверенных информационных систем с трехзвенной архитектурой	43
<i>Богданов А.В., Дегтярев А.Б.</i> Суперкомпьютинг без суперкомпьютеров: что мы можем, а что не можем? . . .	61
<i>Гергель В.П., Баркалов К.А.</i> Достижение эксафлопсной производительности в задачах глобальной оптимизации	78
<i>Гергель В.П., Денисенко М.В., Линёв А.В., Пашин Д.С., Сатанин А.М.</i> Квантовые траектории и томография состояний связанной системы кубит-измерительный прибор	86
<i>Досаев А.С., Трошцкая Ю.И.</i> Численное моделирование динамики нелинейных волн на глубокой воде . . .	100
<i>Дружинин О.А., Островский Л.А., Фэнь Ж.</i> Прямое численное моделирование взаимодействия мелкомасштабной турбулентности с внутренними волнами	110
<i>Ивановский Л.И., Самсонов С.О.</i> Динамика одного двумерного отображения и устойчивые режимы сингулярно возмущенной системы нейронного типа	121

<i>Козелков А.С., Шагалев Р.М., Курулин В.В., Ялозо А.В., Лашкин С.В.</i> Проблемы использования суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в индустриальных приложениях	133
<i>Моисеенко С.Г., Бисноватый-Коган Г.С., Арделян Н.В., Космачевский К.В.</i> Моделирование коллапса холодного быстровращающегося протозвездного облака	151
<i>Рыбакин Б.П.</i> Моделирование 3D задач нестационарной газодинамики на многопроцессорных ЭВМ и графических ускорителях	175

ВВЕДЕНИЕ МЕТОДЫ СУПЕРКОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ»

Читателю предлагаются избранные статьи, написанные на основе приглашенных докладов на очередном 9-м расширенном семинаре по вычислительным методам в естествознании. Под названием «Методы суперкомпьютерного моделирования» семинар был проведен 21–23 апреля 2015 года на базе «Интеркосмос» Института космических исследований Российской академии наук. Как и предыдущий, семинар был посвящен методам и практике разработки и применения методов моделирования на суперкомпьютерах для решения задач в различных областях естественных наук, а также вопросам применения новых информационных и телекоммуникационных технологий при проведении исследований в области естественных наук. Особое внимание было уделено новым разработкам, идеям и методам, а также ярким результатам, добиться которых без применения современных методов проведения моделирования и без использования новых информационных и телекоммуникационных технологий было бы затруднительно или невозможно.

Особенность проводимых нами научных семинаров [1–8] состоит в активном участии молодых научных сотрудников и аспирантов, число которых составляет примерно половину участников. Другая особенность проводимых нами расширенных семинаров состоит в том, что в их рамках проводились научно-учебные семинары с приглашенными студентами, которые делают первые шаги в научных исследованиях по тематике семинара. Еще одна особенность семинаров, начиная с восьмого, состоит в том, что их проведение поддержано грантом РНФ 14-21-00158 «Разработка алгоритмов и методов для задач математического моделирования на суперкомпьютерных системах, включая гибридные» (руководитель Щур Л.Н.)

На семинаре был представлен 21 устный доклад и 5 стендовых докладов.

Сотрудник Института космических исследований РАН (г. Москва) Моисеенко Сергей Григорьевич представил доклад «Моделирование коллапса холодного быстровращающегося протозвездного облака» [9]. Плотность облака меняется на много порядков, что потребовало использования особой реализации разностной схемы. Применялась полностью консервативная неявная операторно-разностная схема в Лагранжевых координатах и треугольная сетка переменной структуры.

Были исследованы нестационарные магниторотационные процессы в замагниченном облаке. Этот результат получен в развитие доложенной на предыдущем семинаре работы [32].

Доклад Рыбакина Виктора Петровича [10], сотрудника МГУ им. М.В. Ломоносова (г. Москва), был посвящен моделированию трехмерных газодинамических задач на многопроцессорных гибридных компьютерах.

Кураченко Анатолий Викторович представил работу группы сотрудников научно-исследовательского центра «Курчатовский институт», в которой изложен развиваемый группой подход по применению облачных систем для решения задач секвенирования [11].

Два доклада были посвящены другой актуальной теме – эффективности обработки сейсмических данных и моделированию процессов сеймики. Это работа коллектива Казанского государственного университета (г. Казань), которую представил Биряльцев Евгений Васильевич [12] и работа коллектива Московского физико-технического института (г. Долгопрудный), которую представил Хохлов Николай Игоревич [13].

Пашков Денис Витальевич (Московский институт электроники и математики Национального исследовательского университета «Высшая школа экономики», г. Москва) изложил подход по синтезу алгоритмов построения доверенных информационных систем [14].

Первый день работы конференции завершился традиционным докладом-ужином, который представил тандем из Санкт-Петербургского университета, Богданов Александр Владимирович и Дягтярев Александр Борисович. В докладе «Суперкомпьютерные технологии без суперкомпьютеров: какие задачи мы можем, а какие пока не можем решить» [15] был изложен взгляд на проблему решения задач, требующих большого объема вычислений, с историческим дискурсом и изложением личного опыта. Этот доклад был особо интересен молодым исследователям, поскольку некоторые изложенные докладчиками истории произошли задолго до рождения молодых слушателей.

С практического решения ряда затронутых в докладе-ужине проблем началась утренняя сессия. В докладе руководителя старейшего в России факультета вычислительной математики и кибернетики (Нижегородский государственный университет им. Н.И. Лобачевского, г. Нижний Новгород) Гергеля Виктора Павловича была изложена методика решения задач глобальной оптимизации, которая имеет перспективы использования на компьютерах следующего десятилетия [16].

Козелков Андрей Сергеевич (Российский федеральный ядерный центр ВНИИЭФ, г. Саров) представил работу коллектива известного ядерного исследовательского центра по разработке программного обеспечения для решения индустриальных задач, требующих гигантских объемов вычисления с применением моделирования гидродинамических процессов [17].

Коллектив исследователей Ярославского государственного университета им. П.Г. Демидова, состоящий как из маститых ученых, так и юных исследователей, поделился успехами в численном исследовании задач, имеющих как чисто

теоретическое, так и сугубо практическое значение. Руководитель коллектива Глызин Сергей Дмитриевич изложил разрабатываемый им подход к исследованию неклассических релаксационных колебаний, возникающих в небесной механике [18].

Доклады его молодых коллег, Алешина Сергея Владимировича [19] и Фролова Даниила Александровича [20], были посвящены исследованию поведения динамических систем с запаздыванием и автоматизации моделирования задач диффузии на гибридных вычислительных системах.

В докладах следующей сессии речь шла о задачах пограничных течений жидкости. Гайдуков Роман Константинович (Высшая школа экономики, г. Москва) исследовал задачу течения вблизи поверхности с неровностями [21].

Дружинин Олег Александрович (Институт прикладной физики РАН, г. Нижний Новгород) изложил исследование важной практической задачи численного моделирования мелкомасштабной турбулентности и внутренних волн в жидкости [22].

Развитию численных методов решения задач была посвящена секция. В докладе Бараша Льва Юрьевича [23] (Научный центр РАН в Черноголовке и ИТФ им. Л.Д. Ландау РАН, г. Черноголовка) был изложен новый метод исследования задач релаксации спиновых моделей и его реализация на гибридных вычислительных системах.

В докладе Теслюка Антона Борисовича [24] (НИЦ «Курчатовский институт», г. Москва) обсуждались проблемы анализу и моделированию экспериментальных данных по расшифровке структуры макромолекул за счет рассеяния лазерного излучения. Эта работа является подготовкой к анализу данных со строящейся экспериментальной установки.

На вечерней сессии молодыми исследователями было представлено пять стендовых докладов:

В.П. Гергель, М.В. Денисенко, А.В.Линев, Д.С.Пашин (докладчик), А.М. Сатанин (Нижегородский государственный университет им. Н.И. Лобачевского, г. Нижний Новгород) «Квантовые траектории и томография состояний связанной системы кубит-измерительный осциллятор» [25].

Сиднев А.А., (Нижегородский государственный университет, г. Нижний Новгород) «Макромодульный подход по разработке сложного программного обеспечения».

Ивановский Л.И., Самсонов С.О. (Ярославский государственный университет им. П.Г. Демидова, г. Ярославль) «Фазовые перестройки одной двумерной динамической системы с импульсным воздействием» [26].

Досаев А.С. (Институт прикладной физики РАН, г. Нижний Новгород) «Численное моделирование динамики нелинейных гравитационных волн на глубокой воде» [27].

Зиганурова Л.Ф., (Высшая школа экономики, Москва) «Моделирование профиля времен в локальном консервативном алгоритме ПМДС».

Третий день конференции продолжился представлением исследования широкого спектра задач.

Три доклада были посвящены исследованию свойств потенциальных устройств для квантовых вычислений. Молодой сотрудник Муняев Вячеслав Олегович (Нижегородский государственный университет им. Н.И. Лобачевского, г. Нижний Новгород) изложил исследование динамики джозефсоновского осциллятора [28], а его научный руководитель Сатанин Аркадий Михайлович представил результаты оригинальных исследований генерации перепутанных состояний в волноводе с таким осциллятором [29]. Коллективом из Института прикладной физики (г. Нижний Новгород) было разработано программное обеспечение для исследования таких задач с использованием гибридных ЭВМ [30].

Анализ масштабируемости параллельных приложений на основе технологий суперкомпьютерного кодизайна был темой доклада сотрудника МГУ им. М.В. Ломоносова (г. Москва) Теплова Алексея Михайловича.

Завершающий доклад об использовании новых командных расширений процессоров Intel для повышения эффективности генерации псевдослучайных чисел сделала студентка 3-го курса Московского института электроники и математики Гуськова Мария Сергеевна.

С полными мультимедийными версиями докладов можно ознакомиться на странице семинара [33].

ЛИТЕРАТУРА

- [1] Труды семинара по вычислительным технологиям в естественных науках. Выпуск 1. Вычислительная физика / Под ред. Р.Р. Назирова. М.: КДУ, 2009.
- [2] Методы нелинейного анализа в кардиологии и онкологии: Физические подходы и клиническая практика. Вып.2 / под ред. Р.Р. Назирова. М. КДУ, 2010
- [3] Труды семинара по вычислительным технологиям в естественных науках. Выпуск 3. Вычислительная химия и биология / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: КДУ, 2010.
- [4] Техническое зрение в системах управления мобильными объектами, 2010: Труды научно-технической конференции-семинара. Вып. 4. / Под ред. Р.Р. Назирова. М. : КДУ, 2011. 328 с.
- [5] Вычислительные технологии в естественных науках. Системы глобального масштаба. Труды семинара / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: КДУ, 2011.
- [6] Вычислительные технологии в естественных науках. Перспективные компьютерные системы: устройства, методы и концепции. Труды семинара / Под ред. Р.Р. Назирова, Л.Н. Щура. М.: ИКИ РАН, 2011.
- [7] Инфокоммуникационные технологии к научным исследованиям. Информационное общество. № 1-2, 2013.
- [8] Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Труды семинара / Под ред. Р.Р. Назирова, Л.Н. Щура. М: ИКИ РАН, 2014.
- [9] *Моисеенко С.Г., Бисноватый-Коган Г.С., Арделян Н.В., Космачевский К.В.* «Моделирование коллапса холодного быстровращающегося протозвездного облака». Вычислительные технологии в естественных науках. Методы суперкомпьютерно-

- го моделирования. Часть 2. Сборник трудов, стр. 151-174/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [10] *Рыбакин Б.П.* «Моделирование 3D задач нестационарной газодинамики на многопроцессорных ЭВМ и графических ускорителях». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 175-183 / Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [11] *Кураченко А.В., Ильин В., Пойда А., Теслик А., Недолужко А.* «Облачные системы для анализа данных секвенирования».
- [12] *Биряльцев Е.В.* «Новые методы сейсморазведки на основе высокопроизводительных вычислений».
- [13] *Петров И.Б., Хохлов Н.И.* «Применение современных высокопроизводительных технологий для моделирования задач сейсмологии и геофизики».
- [14] *Белов А.В., Пашков Д.В.* «Синтез алгоритмов построения доверенных информационных систем с трехзвенной архитектурой». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 43-60 / Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [15] *Богданов А.В., Дегтярев А.Б.* «Суперкомпьютинг без суперкомпьютеров: что мы можем, а чего не можем?». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 61-77 / Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [16] *Гергель В.П., Баркалов К.А.* «Достижение эксафлопсной производительности в задачах глобальной оптимизации». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 78-85 / Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [17] *Козелков А.С., Шагалиев Р.М., Курулин В.В., Ялозо А.В., Лашкин С.В.* «Проблемы использования суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в промышленных приложениях». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 133-150 / Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [18] *Глызин С.Д.* «О неклассических релаксационных колебаниях в сингулярно возмущенных динамических системах».
- [19] *Алешин С.В., Глызин С.Д.* «Пространственно неоднородные решения уравнения Колмогорова-Петровского-Пискунова с запаздыванием». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 11-17/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [20] *Глызин Д.С., Фролов Д.А.* «Автоматизация конечно-разностного моделирования диффузионных задач на гибридных вычислительных кластерах».
- [21] *Данилов В.Г., Гайдуков Р.К.* «Течения в пограничных слоях вдоль поверхностей с малыми неровностями».
- [22] *Дружинин О.А., Островский Л.А., Фэнь Ж.* «Прямое численное моделирование взаимодействия мелкомасштабной турбулентности с внутренними волнами». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 110-120/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.

- [23] *Бараи Л.Ю., Щур Л.Н., Боровский М.* Алгоритм «Population Annealing» и его применение к двумерной модели Изинга». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 29-42/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [24] *Теслюк А.Б., Бобков С.А., Ильин В.А.* НИЦ «Курчатовский институт», г. Москва, «Моделирование и обработка данных для экспериментов по рассеянию излучения лазеров на свободных электронах на макромолекулярных объектах».
- [25] *Гергель В.П., Денисенко М.В., Линева А.В., Пашин Д.С., Сатанин А.М.* «Квантовые траектории и томография состояний связанной системы кубит-измерительный осциллятор». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 86-99/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [26] *Ивановский Л.И., Самсонов С.О.* «Динамика одного двумерного отображения и устойчивые режимы сингулярно возмущенной системы нейронного типа». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 121-132/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [27] *Досаев А.С., Троицкая Ю.И.* «Численное моделирование динамики нелинейных волн на глубокой воде». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 100-109/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [28] *Денисенко М.В., Муняев В.О., Сатанин А.М.* «Генерация перепутанных (entangled) состояний фотонов в волноводной линии с встроенным джозефсоновским осциллятором».
- [29] *Мартынов В.О., Миронов В.А., Смирнов Л.А.* «Численное моделирование диссипативной динамики многочастичных открытых квантовых систем с использованием технологий вычисления на GPU».
- [30] *Антонов А.С., Теплов А.М.* «Анализ масштабируемости параллельных приложений на основе технологий суперкомпьютерного кодизайна». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Часть 2. Сборник трудов, стр. 18-28/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2015.
- [31] *Гуськова М.С.* «AVX2 и эффективная генерация псевдослучайных чисел».
- [32] *Моисеенко С.Г., Бисноватый-Коган Г.С.* «Моделирование коллапса холодного быстроевращающегося протозвездного облака». Вычислительные технологии в естественных науках. Методы суперкомпьютерного моделирования. Сборник трудов, стр. 121-135/ Под ред. Р.Р. Назирова, Л.Н. Щура. М. – М: ИКИ РАН, 2014.
- [33] Мультимедийные версии докладов. URL: <http://iki.rssi.ru/seminar/2015042123/>

ОРГАНИЗАТОРЫ СЕМИНАРА

Назирова Р.Р.,

д.т.н., профессор

Институт космических исследований РАН

Щура Л.Н.,

д.ф.-м.н., профессор

Научный центр РАН в Черноголовке

ПРОСТРАНСТВЕННО НЕОДНОРОДНЫЕ РЕШЕНИЯ УРАВНЕНИЯ КОЛМОГОРОВА-ПЕТРОВСКОГО-ПИСКУНОВА С ЗАПАЗДЫВАНИЕМ

С.В. Алешин

Ярославский государственный университет им. П.Г. Демидова (ЯрГУ им. П.Г. Демидова)
Научный центр РАН в Черногоровке

Рассматривается задача распространения волны плотности в логистическом уравнении с запаздыванием и диффузией (уравнение Фишера-Колмогорова-Петровского-Пискунова с запаздыванием). Проведен численный анализ процесса распространения волны. Выделены значения запаздывания, для которых характерно появления различных пространственно неоднородных режимов.

Исследование выполнено за счет гранта Российского научного фонда (проект № 14-21-00158).

Ключевые слова: уравнение Фишера-Колмогорова-Петровского-Пискунова, диффузия, запаздывание, численные методы.

ВВЕДЕНИЕ

Уравнением Колмогорова-Петровского-Пискунова (далее уравнением КПП) называют логистическое уравнение с диффузией

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1 - u), \quad (1)$$

Здесь $t > 0$ – временная переменная, а $x \in (-\infty, \infty)$ – пространственная. В основополагающих работах о процессе распространения генной волны А.Н. Колмогорова [Колмогоров и др., 1937] и Р.А. Фишера [Fisher, 1937] рассмотрен вопрос о таких неотрицательных решениях $u(t, x)$, для которых $t > 0$ при каждом выполнении условия: $u(t, x) \rightarrow 1$ при $x \rightarrow \infty$ и $u(t, x) \rightarrow 0$ при $x \rightarrow -\infty$. Показано, что при $t \rightarrow \infty$ функция $u(t, x)$ принимает определенную форму. Обзор результатов для уравнения (1) приведен в работах Данилова [Danilov et al., 1995], Мюррея [Murray, 2001] и Вольперта [Volpert et al., 2000].

Для любых классических граничных условий (например, для условий Неймана: $\frac{\partial u}{\partial x} \Big|_{x=a} = \frac{\partial u}{\partial y} \Big|_{x=b} = 0$; периодических граничных условий $u(t, x + T) \equiv u(t, x)$

и др.) уравнение (1) имеет только один аттрактор – однородное состояние равновесия $u_0 \equiv 1$. Все остальные состояния равновесия неустойчивы.

Здесь будут рассматриваться уравнения КПП с запаздыванием [Алешин и др., 2015]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1 - u(t - \tau, x)), \quad (2)$$

Алешин Сергей Владимирович – ассистент, старший лаборант-исследователь, ЯрГУ им. П.Г. Демидова, Научный центр РАН в Черногоровке, fktiby@yandex.ru

где $\tau > 0$ – величина запаздывания, $u(t, x)$ – достаточно гладкая по x и t функция при $t > 0$ и $x \in (-\infty, \infty)$. Кроме этого будем предполагать, что $u(t, x) \equiv 0$ при $a < x < b$. При этом разницу $|a-b|$ будем считать достаточно большой для того, чтобы увидеть распространение волны до момента встречи фронта с границами a или b .

Устойчивые неоднородные по пространственной переменной решения уравнения (2) с периодическими краевыми условиями и условиями Неймана были численно проанализированы в статьях [Глызин, 2009; Glyzin, 2013] и [Глызин и др., 2010].

Перейдем к описанию вычислительной процедуры. В уравнении (2) вторую производную по пространственной переменной $\frac{\partial^2 u}{\partial x^2}$ заменим второй разделенной разностью. Для этого отрезок $[a, b]$ разобьем на N равных частей и построим сетку узлов с шагом $h = \frac{b-a}{N}$: $x_j = a + jh$, где $j = 0, \dots, N-1$. Обозначим через $u_j(t)$ значение функции $u(t, x)$ в соответствующих узлах сетки. В итоге получим следующую систему обыкновенных дифференциальных уравнений:

$$\dot{u}_j = \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + u_j (1 - u_j(t - \tau)). \quad (3)$$

Для учета краевых условий полагаем $u_{-1}(t) \equiv u_N(t) \equiv 0$.

Моделирование проводилось на вычислительном кластере ЯрГУ. Одновременно решалось от $N = 0,9 \times 10^5$ до $N = 1,8 \times 10^5$ уравнений. Особенностью численного решения такого большого количества дифференциальных уравнений с запаздыванием является повышенная требовательность к объемам доступной памяти для хранения всего решения на промежутке запаздывания. В связи с чем решение приходилось сохранять на жестком диске кластера, что существенно замедляло процедуру расчета.

Для вычислений использовался метод Дормана-Принца пятого порядка с переменной длиной шага интегрирования [Hairer et al., 2008]. Абсолютная и относительная погрешность алгоритма была установлена в значение 10^{-12} . Начальный шаг интегрирования равен 10^{-3} .

Начальные условия выбраны в виде прямоугольного импульса высоты 0,1 и единичной ширины, расположенного в центре отрезка $[a, b]$ для всех $-\tau \leq t \leq 0$. В частности, для случая $x \in [0, 1800]$:

$$u_j(t) = \begin{cases} 0,1, & j \in [89950, 90050], \\ 0, & \text{иначе,} \end{cases} \quad (4)$$

где $t \in [-\tau, 0]$. Для последующего анализа полученные данные прореживались.

Перейдем к результатам моделирования. Описание поведения уравнения КПП с запаздыванием будем проводить в сравнении с классическим уравнением КПП без запаздывания. На рис. 1 представлено распространение волны постоянной высоты от начального всплеска единичной ширины и высоты 0,1. Отметим, что скорость распространения волны равна углу наклона профиля волны на соответствующих рисунках.

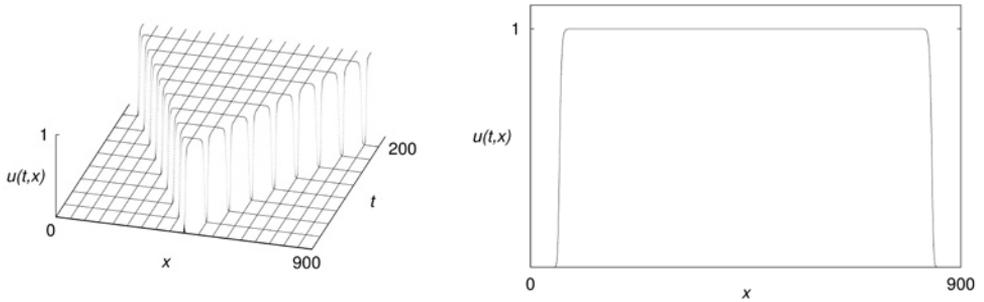


Рис. 1. Решение системы (3) при $\tau = 0$. Слева распространение волны на плоскости (x, t) . Справа разрез при $t = 200$

Рассмотрим теперь систему (3) с ненулевым запаздыванием. При увеличении параметра τ можно выделить несколько этапов качественно различного поведения решений системы. Точные значения границ промежутков запаздывания приведены в работе [Алешин и др., 2015].

При относительно малом значении τ поведение решения практически идентично поведению системы КПП без запаздывания.

При увеличении τ примерно до $\frac{\pi}{2}$ распространение волны начинает сопровождаться некоторым количеством осцилляций в районе ее фронта. Так при $\tau = 1$ на рис. 2 показана такая волна.

В этом случае максимальная амплитуда всплеска равна примерно 1,2, при том, что после прохождения фронта волны значение $u(t, x)$ устанавливается равным единице.

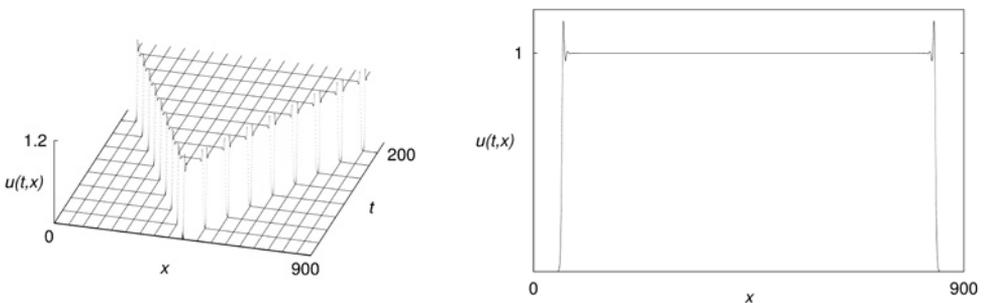


Рис. 2. Решение системы (3) при $\tau = 1$. Слева распространение волны на плоскости (x, t) . Справа разрез при $t = 200$

Следующее существенное изменение в распространении фронта волны задачи (3) наблюдается при $\tau \geq \frac{\pi}{2}$. Отметим, что при этом условии уравнение $\dot{u} = u(1 - u(y - \tau))$ имеет устойчивое периодическое решение. В этом случае фронт распространения волны содержит довольно большое число интенсивных осцилляций, распространяющихся с той же скоростью, что и фронт волны у стандартного уравнения КПП. После некоторого количества интенсивных осцилляций

после фронта волны величина $u(t,x)$ быстро сходится к значению единица за исключением центральной области, в которой появляется след от начального прямоугольного импульса, распространяющийся существенно медленней, чем фронт волны (см. рис. 3).

Для того, чтобы прояснить характер распространения волны в этом случае, на рисунке 3 даны разрезы решения при $x = 900$ и $t = 400$ соответственно. Из рисунка 3 видно, что колебательный режим в центральной области близок, но не тождественен колебаниям решения логистического уравнения с запаздыванием и диффузией и периодическими краевыми условиями. Полученные результаты хорошо согласуются с тем, что локальная динамика данного уравнения в окрестности состояния равновесия описывается комплексным уравнением Гинзбурга-Ландау [Алешин и др., 2015]. Причем это уравнение не имеет устойчивых автономных периодических решений.

При дальнейшем увеличении τ характер поведения средней части распространяющейся волны резко меняется.

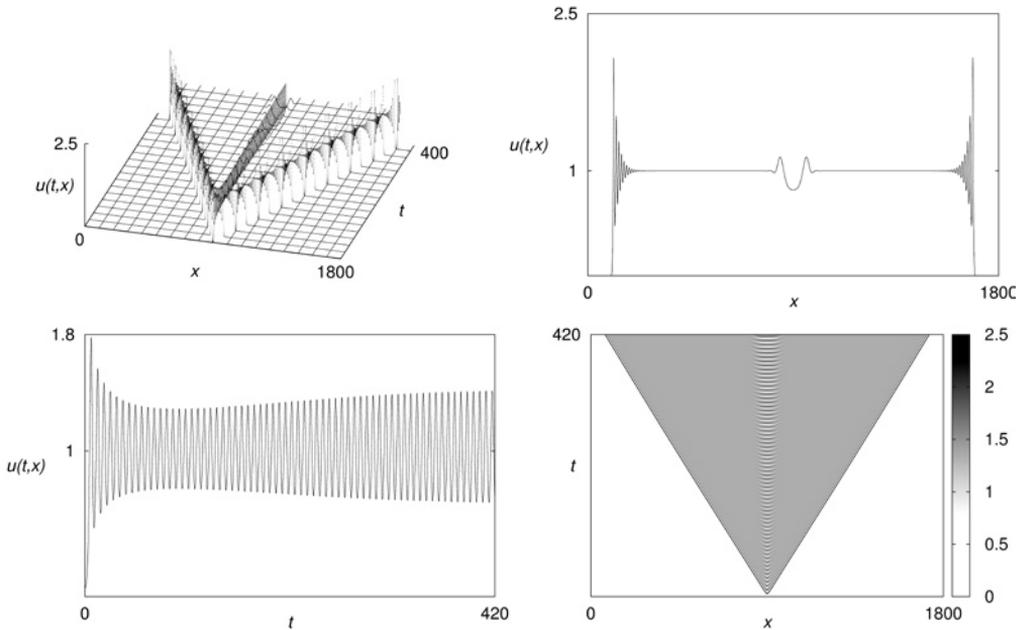


Рис. 3. Решение системы (3) при $\tau = 1,6$. Слева сверху распространение волны на плоскости (x,t) . Справа сверху разрез при $t = 200$. Слева внизу разрез при $x = 900$. Справа снизу плотность распределения $u(t,x)$ в оттенках серого

Так на рисунке 4 при $\tau = 1,8$ мы можем наблюдать относительно медленное распространение колебательной структуры, находящейся в центральной части распространяющейся волны на протяжении от 0 до примерно 300. А на промежутке $t > 300$ распространение колебательной структуры имеет уже скорость близкую к скорости распространения фронта основной волны и равной двум. Общее представление о характере распространяющегося волнового фрон-

та можно получить из продольного и поперечного разреза, представленного на рисунке 4, а также трехмерной картины участка сложного поведения системы с границами $[400, 1400] \times [400, 420]$.

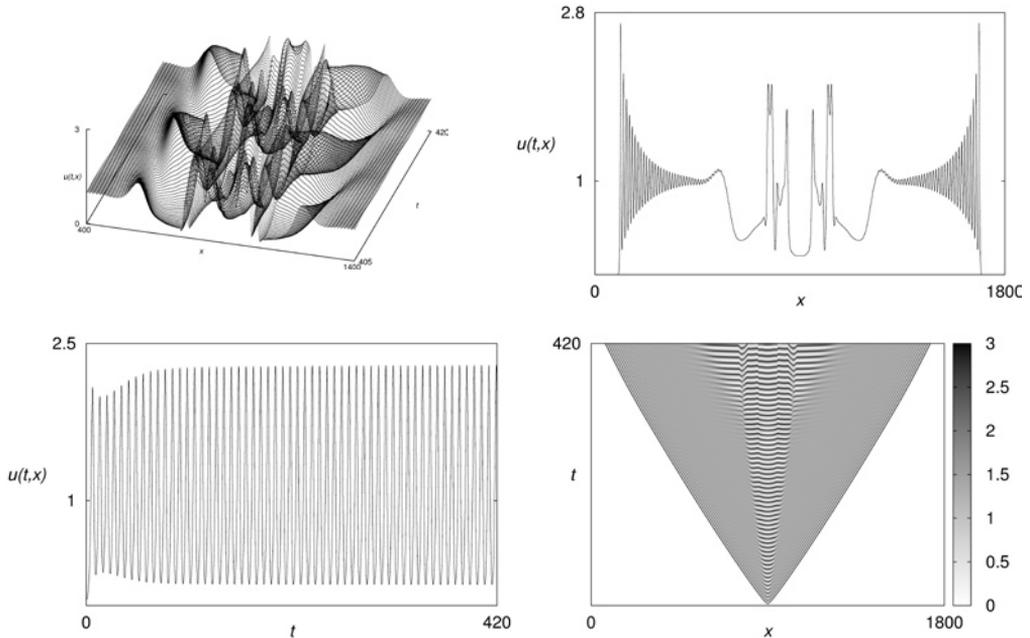


Рис. 4. Решение системы (3) при $\tau = 1,8$. Слева сверху фрагмент ($x \in [400, 1400]$, $t \in [405, 420]$) графика распространения волны. Справа сверху разрез при $t = 200$. Слева внизу разрез при $x = 900$. Справа снизу плотность распределения $u(t,x)$ в оттенках серого

Обратим внимание на следующие особенности, вытекающие из анализа графиков. Нетрудно видеть, что у границы распространения основной волны имеется сходящийся к единице набор высокоамплитудных всплесков, которые потом переходят в сложное пространственно-неоднородное распространение решения в средней части волны. В отличие от предыдущих случаев здесь уже нет промежутков, на котором значение $u(t,x)$ сходится к единице. Рисунок 4 показывает сильную изрезанность пространственно-временного распределения функции $u(t,x)$.

При $\tau = 2$ наблюдается еще один новый эффект в распространении волны возмущения (рис. 5) для системы (3). В этом случае, так же как и ранее у границы волнового фронта, имеется некоторое число высокоамплитудных колебаний, которые при приближении к средней части профиля $u(t,x)$ сходятся не к единице, как ранее, а некоторому высокочастотному колебательному режиму, который разрушается только в средней части, где расположен след от начального возмущения.

Отметим также то обстоятельство, что структура решения сохраняется при внешних воздействиях. Так добавление разового единичного импульса на уча-

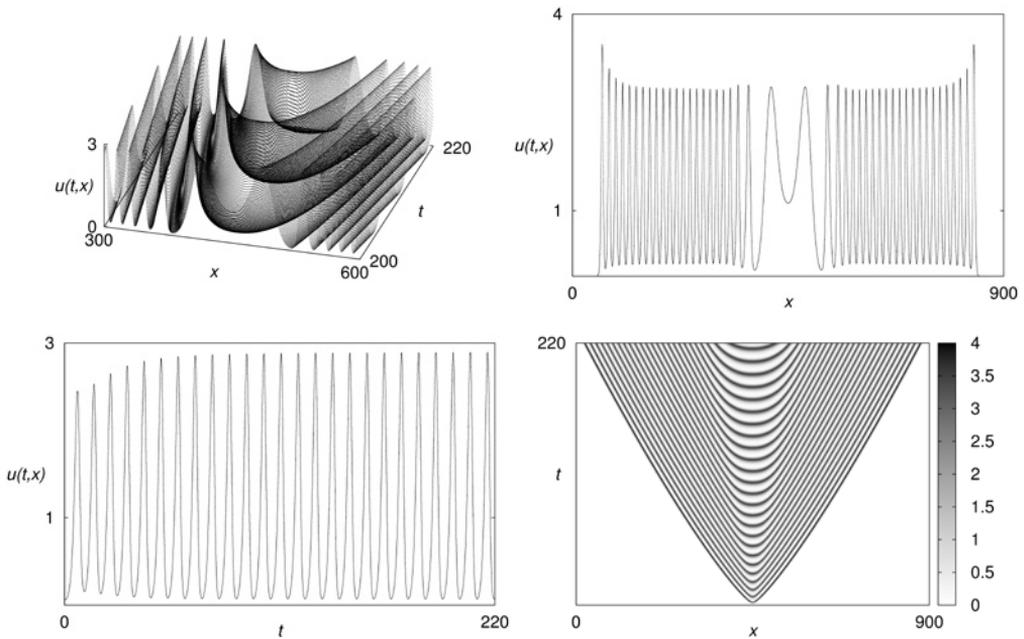


Рис. 5. Решение системы (3) при $\tau = 2$. Слева сверху фрагмент ($x \in [300, 600]$, $t \in [200, 220]$) графика распространения волны. Справа сверху разрез при $t = 200$. Слева внизу разрез при $x = 900$. Справа снизу плотность распределения $u(t, x)$ в оттенках серого

сток единичной ширины, расположенный в центральной части изучаемой области не оказал существенного влияния (опыт проводился при $\tau = 1,6$ и $t = 200$). Через промежуток времени не более 10 единиц вид решения с внешним воздействием и без него существенно не отличались.

ЗАКЛЮЧЕНИЕ

Численный анализ процесса распространения волны позволил выделить четыре промежутка значений запаздывания. При достаточно малых значениях запаздывания данное уравнение имеет решения, близкие к решениям стандартного уравнения КПП. Увеличение параметра запаздывания приводит сначала к появлению затухающей колебательной составляющей в пространственном распределении решения. Дальнейший рост данного параметра приводит к разрушению бегущей волны. Это выражается в том, что в окрестности участка начального возмущения сохраняются незатухающие по времени и медленно распространяющиеся по пространству колебания, близкие к решениям соответствующей краевой задачи с периодическими граничными условиями. Наконец, если значение запаздывания достаточно велико, то во всей области распространения волны наблюдаются интенсивные пространственно-временные колебания.

Автор выражает благодарность Глызину С.Д и Кащенко С.А. за постановку задачи и обсуждение полученных результатов.

ЛИТЕРАТУРА

- [Алешин и др., 2015] *Алешин С.В., Глызин С.Д., Кащенко С.А.* Уравнение Колмогорова-Петровского-Пискунова с запаздыванием // Моделирование и анализ информационных систем. 2015. Т. 22, № 2. С. 304-321.
- [Глызин, 2009] *Глызин С.Д.* Разностные аппроксимации уравнения «реакция-диффузия» на отрезке // Моделирование и анализ информационных систем. 2009. Т. 16, № 3. С. 96-116.
- [Глызин и др., 2010] *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Конечномерные модели диффузионного хаоса // Журнал вычислительной математики и математической физики. 2010. Т. 50, № 5. С. 860-875.
- [Колмогоров и др., 1937] *Колмогоров А.Н., Петровский И.Г., Пискунов Н.С.* Исследование уравнения диффузии, соединенной с возрастанием вещества, и его применение к одной биологической проблеме // Бюл. МГУ, сер. Математика и механика 1 (6): 1-26, 1937.
- [Danilov et al., 1995] *Danilov V.G., Maslov V.P., Volosov K.A.* Mathematical Modelling of Heat and Mass Transfer Processes // Kluwer: Dordrecht, 1995.
- [Fisher, 1937] *Fisher R.A.* The Wave of Advantageous Genes // Ann. Eugenics 7, 355-369, 1937.
- [Glyzin, 2013] *Glyzin S.D.* Dimensional Characteristics of Diffusion Chaos // Automatic Control and Computer Sciences. 2013. Vol. 47, No. 7. P. 452-469. DOI: 10.3103/S0146411613070031.
- [Hairer et al., 2008] *Hairer E., Wanner G., Norsett S.P.* Solving Ordinary Differential Equations 1 (Springer Series in Computational Mathematics): Nonstiff Problems // 2 ed., revised, Springer, 2008, 539 p.
- [Murray, 2001] *Murray J.D.* Mathematical Biology. I. An Introduction (Third Edition) // Berlin, 2001, 576 p.
- [Volpert et al., 2000] *Volpert A., Volpert V., Volpert V.* Traveling Wave Solutions of Parabolic Systems // American Mathematical Society, 2000, 453 p.

THE SPATIALLY INHOMOGENEOUS SOLUTION IN THE FISHER-KOLMOGOROV EQUATION WITH DELAY

S.V. Aleshin

P.G. Demidov Yaroslavl State University

(P.G. Demidov YSU)

Science Center in Chernogolovka, Russian Academy of Sciences

We considered the problem of density wave propagation of logistic equation with delay and diffusion (Fisher-Kolmogorov equation with delay). We performed a numerical analysis of wave propagation and identified the values of delay which is characterized by the appearance of different spatially inhomogeneous regimes.

Keywords: Fisher-Kolmogorov equation, diffusion, delay, numerical analysis.

АНАЛИЗ МАСШТАБИРУЕМОСТИ ПАРАЛЛЕЛЬНЫХ ПРОГРАММ НА ОСНОВЕ ТЕХНОЛОГИЙ СУПЕРКОМПЬЮТЕРНОГО КОДИЗАЙНА

А.С. Антонов, А.М. Теплов

Научно-исследовательский вычислительный центр Московского государственного университета имени М.В. Ломоносова (НИВЦ МГУ имени М.В. Ломоносова)

В работе изложен подход к анализу масштабируемости параллельных приложений с использованием технологий суперкомпьютерного кодизайна, основанный на использовании объединенных и структурированных данных о работе параллельной программы. Предложенный подход позволяет изучать свойства приложения на основе собранных данных о различных динамических характеристиках его работы. Рассматриваются данные с различных уровней исследования работы приложения, начиная с аппаратного уровня. Такой подход позволяет делать обоснованные выводы о потенциальных причинах снижения эффективности параллельных приложений.

Работа выполняется при поддержке гранта РФФИ 13-07-00790 а.

Ключевые слова: Масштабируемость, суперкомпьютерный кодизайн, эффективность, параллельные вычисления, подход к анализу масштабируемости.

ВВЕДЕНИЕ

Одним из важнейших свойств параллельной программы является ее масштабируемость. В научных публикациях можно найти достаточно большое число определений этого свойства. Эти определения в большинстве своем указывают на тот факт, что на программы при изменении каких-либо параметров их работы начинают оказывать большее или меньшее влияние различные факторы [Антонов, Теплов, 2014а], изменяющие качество их работы. Большое число определений говорит о сложности понятия масштабируемости и важности его изучения при анализе работы параллельных приложений.

Важнейшим недостатком большинства подходов к анализу масштабируемости является то, что это понятие рассматривается авторами только с точки зрения изменения какого-то одного параметра, и только в редких определениях встречается два параметра. При проведении исследований такой подход приведет к большому разбросу результатов, потому что параметров, оказывающих влияние на характеристики выполнения программы, значительно больше. Кроме того, рассмотрение какой-то одной из характеристик не отражает полной картины свойств исследуемой программы.

Технологии суперкомпьютерного кодизайна позволяют использовать для разработки высокоэффективных программ информацию об особенностях выполнения программ, начиная с анализа использования аппаратных ресурсов вычислительной системы и до уровня алгоритмической структуры программы.

Антонов Александр Сергеевич – НИВЦ МГУ имени М.В. Ломоносова, вед.н.с., asa@parallel.ru

Теплов Алексей Михайлович – НИВЦ МГУ имени М.В. Ломоносова, м.н.с., alex-teplov@yandex.ru

Согласование особенностей всех уровней анализа работы программы позволяет разрабатывать программы с очень высокой эффективностью. Анализ масштабируемости на основе информации об этих особенностях работы программы позволяет анализировать, каким образом изменяется эффективность программ при изменении условий выполнения и делать выводы о причинах снижения масштабируемости и путях ее улучшения. Потому анализ масштабируемости с использованием технологий кодизайна является перспективным направлением изучения свойств параллельных программ.

ОБОБЩЕНИЕ ПОНЯТИЯ МАСШТАБИРУЕМОСТЬ

Свойство масштабируемости параллельных программ требует значительно более глубокого анализа [Теплов, 2014] и рассмотрения большого и разнообразного числа характеристик [Антонов и др., 2013], отражающих качество выполнения этих программ. Об этом свидетельствует и тот факт, что в различных определениях параллельной программы сделан упор на изменение различных характеристик выполнения программы на вычислительной системе.

Параметром запуска параллельной программы будем называть любой параметр, оказывающий влияние на работу параллельной программы и не влияющий на реализацию алгоритма работы программы (например: размерность матрицы с данными – a , число процессов – b , размер блока матрицы – c , число процессов на узел – d и т.п.). Параметрами запуска параллельной программы не будут являться те параметры, которые не влияют на выполнение программы (например: объем комментариев в тексте программы, именование переменных в программе и т.п.) – все то, что может изменяться, но при этом не будет влиять ни на порядок выполняемых операций в программе, ни на время их выполнения.

Параметры запуска по типу можно разделить на:

1. **Аппаратные** – относящиеся к параметрам аппаратуры.
2. **Системные** – относящиеся к операционной системе, компиляторам, используемым библиотекам и системным приложениям и их настройкам.
3. **Программные** – влияющие на логику работы программы или на вычислительную сложность задачи, а также различные параметры, предусмотренные разработчиком.

Динамическими характеристиками параллельной программы будем называть такие характеристики ее работы, которые определяют качество работы и могут быть измерены или вычислены только при наличии данных о реальном выполнении этой программы.

Динамические характеристики можно разделить на группы несколькими способами. К примеру, можно поделить их по способу получения значений на две группы.

К первой группе можно отнести **измеряемые характеристики**. Примерами таких характеристик могут быть время выполнения, количество кэш-промахов, количество переданных за время работы байт по сети или по шине данных, количество обращений в память за время работы и т.п.

Ко второй группе можно отнести **вычисляемые характеристики**, значения которых можно получить из измеряемых характеристик. Такие характеристики используются для лучшего понимания качества работы программы и для лучшего представления результатов. Примерами могут быть ускорение, производительность, скорость передачи данных, кэш-промахов в секунду, моментальная и средняя эффективность и т.п.

Для понимания свойств параллельной программы и для понимания причин найденных особенностей масштабируемости важно рассматривать все доступные изменяемые параметры запуска и все доступные динамические характеристики. Нужно рассматривать масштабируемость как более общее свойство параллельной программы, включающее в себя как зависимость значений динамических характеристик от различных параметров запуска, так и рассмотрение связей между изменениями значений различных динамических характеристик.

Поэтому **масштабируемостью** [Антонов, Теплов, 2014б] целесообразно называть *свойство параллельной программы, характеризующее зависимость изменения всей совокупности динамических характеристик работы этой программы от множества параметров ее запуска.*

Таким образом, в рамках предложенного подхода к анализу масштабируемости параллельных программ должно рассматриваться влияние большого числа параметров запуска всех уровней. Анализ изменений динамических характеристик должен сводиться к анализу изменений всех характеристик выполнения, начиная с использования аппаратных ресурсов вычислительной системы и заканчивая интегральными показателями работы.

Технологии суперкомпьютерного кодизайна направлены на создание высокоэффективных программ, учитывающих и согласующих особенности взаимодействия аппаратных свойств вычислительной системы, свойств реализации программы в коде и свойств алгоритмической основы программы. Потому использование для анализа таких технологий позволит собрать и совместно проанализировать как данные об аппаратных особенностях, влияющих на работу программу, так и необходимость внесения изменений в алгоритм, реализацию или аппаратную платформу для улучшения эффективности работы программы. Это позволит не только найти причины, снижающие эффективность работы программы, но и оценить их влияние на масштабируемость. Но для этого важно рассматривать как все параметры, так и характеристики совместно, а не по отдельности каждый фактор.

Например, если провести исследование масштабируемости теста Linpack, при сборке которого использовалась библиотека MKL, используя в качестве параметров запуска теста число процессоров и размер задачи, то соберем данные о ее масштабируемости в этих координатах, соответствующих выбранным параметрам запуска. Область значений параметров запуска для решения этой задачи будет соответствовать всем рассмотренным размерностям задачи и числу процессоров. Если же добавить еще одно значение числа процессоров, то для проведения исследования требуется выполнить тест на всех размерах задачи из области значений этого параметра запуска.

Если мы захотим снова исследовать масштабируемость теста, но при его сборке с другой библиотекой (например, ATLAS или GotoBLAS), мы вынуждены добавить и другой параметр запуска – «библиотека BLAS». Для проведения полноценного исследования потребуется проведение экспериментов уже на всех значениях всех остальных параметров запуска, включая размер задачи и число процессоров.

АНАЛИЗ МАСШТАБИРУЕМОСТИ НА ПРАКТИКЕ

Покажем на практике, как можно использовать предложенный подход к анализу масштабируемости приложений для поиска причин снижения эффективности работы приложений.

Условия проведения эксперимента

Для исследования были выделены два параметра запуска: число процессов при работе теста и размерность матрицы.

Область изменения параметров запуска: число процессов – [8, 128] с шагом в 8 процессов, размерность матрицы – [1000, 100 000] с шагом в 1000.

Размер вычислительного блока принят константным и равным 224, что является значением по умолчанию для HPL. Решетка процессов увеличивалась по формуле $8 \cdot N$, что позволило равномерно распределить процессы по узлам с полным заполнением процессами.

Для запуска использовали суперкомпьютер «Ломоносов» [Воеводин и др., 2012] с использованием узлов с двумя четырехядерными процессорами Intel Xeon 5570 и 8 ГБ оперативной памяти. Запускали по 8 процессов на узел, чтобы на каждое физическое ядро процессора распределялось по одному процессу.

Полученные результаты экспериментов

Представим результаты в виде поверхности, в которой координаты X и Y будут соответствовать числу процессоров и размерности матрицы соответственно, а координата Z – эффективности работы теста. Под эффективностью будем понимать эффективность реализации, равной отношению полученной производительности к максимальной теоретически возможной для данного типа процессора.

При рассмотрении полученной поверхности (рисунок 1) нужно отметить достаточно высокую эффективность работы теста на использованном оборудовании. На большей части рассмотренных значениях параметров запуска эффективность близка к 80%. Если рассмотреть область значений параметров, соответствующих большому размеру задачи и малому числу процессов, то можно заметить резкое снижение эффективности до 0%. Эта область соответствует экспериментам, в которых достигнут предел размера матрицы, уместяющейся в доступный объем оперативной памяти.

Также на полученной поверхности можно заметить плавно снижающуюся эффективность работы при росте числа процессов и фиксированном размере задачи. Если рассмотреть область значений параметров, соответствующую небольшой размерности задачи, то можно обратить внимание на очень быстрое снижение эффективности выполнения теста при росте числа процессоров на задачах небольшой размерности (с размером матрицы порядка 15000 и менее).

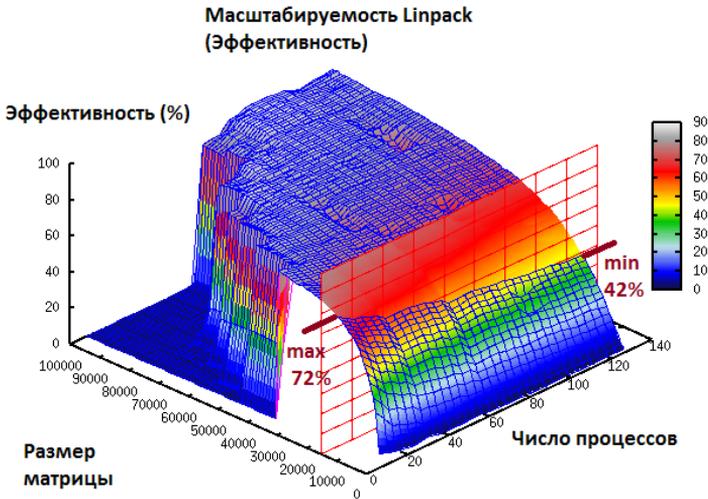


Рис. 1. Масштабируемость Linpack benchmark. Данные о эффективности выполнения программы. Плоскость показывает сечение графика при размере матрицы в 15000. Обозначены максимум в 72% и минимум в 42% значений эффективности в сечении

Эффективность выполнения уменьшается со значения, близкого к 72%, до значения около 42% при увеличении числа процессоров от 4 до 128. Падение эффективности в 30% для данного размера задачи является очень существенным, и для определения его причин можно обратиться к данным системного мониторинга каждого запуска. Плоскостью обозначено сечение полученной фигуры при фиксированном размере задачи в 15000. Пересечение плоскости с фигурой позволяет увидеть как уменьшается эффективность выполнения с ростом числа процессов.

Попробуем выяснить причины столь сильного снижения эффективности и рассмотрим более детально работу программы при фиксированном размере задачи. Для каждого запуска программы нужно собрать данные системного мониторинга, чтобы понять, с изменением каких динамических характеристик аппаратного уровня коррелирует такое изменение значения эффективности.

Наибольший интерес представляют собой данные с датчиков, поступающие в момент вычислений в программе. Самый простой способ – взять среднее значение за весь промежуток работы программы – не всегда подходит для анализа масштабируемости, особенно при анализе программ с малым временем работы. Это связано с тем, что системы мониторинга собирают данные с датчиков аппаратуры с какой-то периодичностью, и потому точность собираемой информации

повышается с временем работы программы. С другой стороны при постановке программы и завершении ее работы происходят различные системные процессы инициализации и завершения, время которых возрастает с ростом числа использованных процессов. При работе программ с небольшими объемами данных это время может быть сравнимо с временем работы программы. Это будет накладывать существенные искажения на среднее значение характеристики.

Если рассматривать поведение характеристик системного уровня [Адинец, 2013], то мы увидим, что в процессе выполнения программы они не постоянны и меняются во времени по сложному закону. На рисунке 2 показан пример работы короткой программы. На нем видно, что время работы невычислительной части и завершение программы сравнимо с временем работы вычислительной части.



Рис. 2. Пример поведения динамических характеристик скорости передачи данных по сети InfiniBand в пакетах в секунду и МБ/сек при работе достаточно короткой программы

На рисунке 2 в завершающей части работы программы видно изменение профиля работы программы и сильное снижение показаний характеристик. Эти участки работы программы вносят искажение в значение среднего значения характеристик и потому должны исключаться из рассмотрения.

Для получения значений, определяющих значение динамических характеристик при выполнении всей программы, интегральные значения динамических характеристик могут быть выбраны различным способом. Для анализа работы теста Linpack будем использовать среднее значение характеристики за время выполнения вычислительной части программы.

Рассмотрим полученные зависимости для работы теста Linpack при размерности задачи в 15000. Рассмотрим участок с наибольшей интенсивностью снижения эффективности работы программы, на котором число процессов изменяется в пределах от 16 до 48. На этом участке эффективность работы теста снижается на 14%, что достаточно существенно, чтобы увидеть изменения в поведении динамических характеристиках работы аппаратного уровня и существенно сократить объем анализируемых данных.

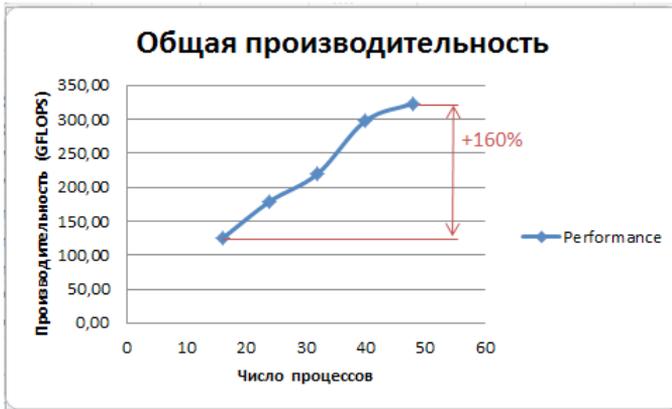


Рис. 3. Изменение производительности теста Linpack при размере матрицы 15 000 и изменении числа процессов

Из графика производительности (рисунок 3) видно, что на рассмотренной области значений числа процессов программа увеличивает общую производительность на 160%, хотя общая эффективность работы, как отношение к реальной производительности к пиковой, снижается на 14%. Это указывает на наличие каких-то накладных расходов. Проанализируем собранные и усредненные данные системного мониторинга.

При выполнении теста при фиксированном размере задачи на рассмотренном участке число кэш-промахов в секунду стабильно снижается (рисунок 4). Это указывает на положительный эффект от декомпозиции данных – данные стали лучше помещаться в кэш. Важно отметить, что почти на 20% снизилось число более дорогих с точки зрения производительности промахов в секунду L3 и на 15% произошло снижение среднего числа промахов в секунду по всем вычислительным узлам. Меньший уровень снижения числа промахов в секунду L1 может объясняться возросшей производительностью: промахов становится меньше, но они происходят чаще, потому как общее число операций в секунду возросло. Наблюдаемое снижение числа кэш-промахов должно вести к росту производительности, и потому из данных системного мониторинга следует, что работа кэш-памяти не является причиной снижения эффективности.

На рисунке 5 показано изменение интенсивности работы с памятью. При увеличении числа процессов на данном участке по данным системного мониторинга следует, что число чтений из памяти в секунду остается практически неизменным. Рост максимального и среднего по всем узлам числа операций записи в память может объясняться возросшей производительностью программы. Поэтому программе чаще необходимо записывать в память полученные результаты. Безусловно, возросшее на 25% максимальное число операций записи в секунду будет увеличивать накладные расходы. Однако численные значения далеки от предельных для использованного оборудования, и потому маловероятно, что они станут основной причиной снижения общей эффективности программы.

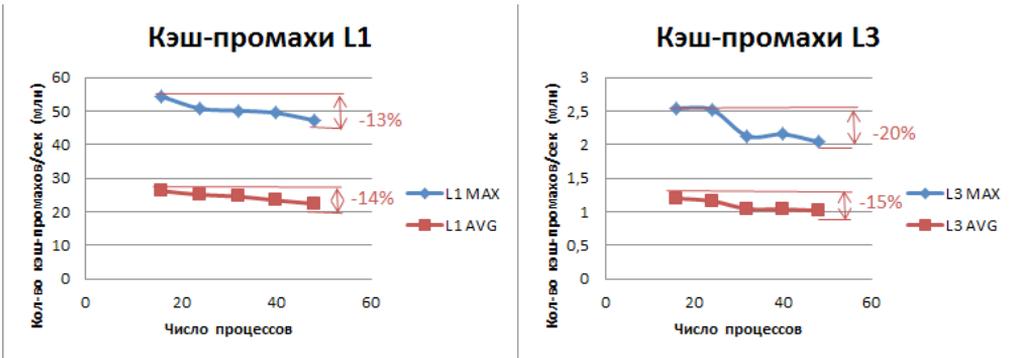


Рис. 4. Изменение максимального и среднего по всем узлам числа кэш-промахов в секунду уровня L1 и L3 при выполнении теста Linpack при размере матрицы 15 000 и изменении числа процессов

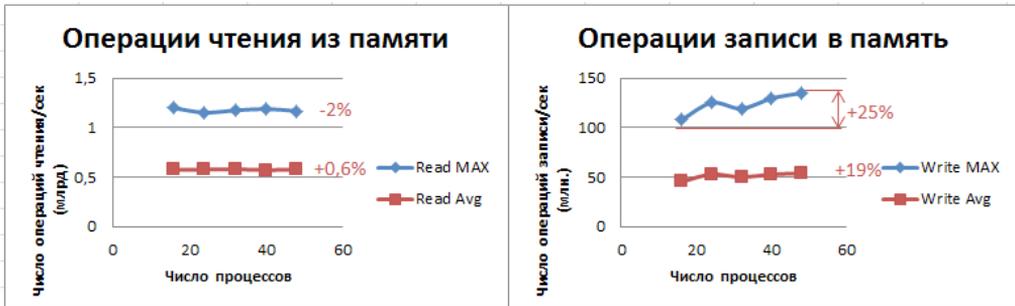


Рис. 5. Изменение максимального и среднего по всем узлам числа операций чтения и записи в память в секунду при выполнении теста Linpack при размере матрицы 15 000 и изменении числа процессов

На рисунке 6 показано изменение средней и максимальной по всем узлам скорости передачи данных при выполнении теста. Из полученных данных следует, что увеличение числа процессов приводит к значительному увеличению интенсивности обмена данными по коммуникационной сети Infiniband. Крайне важным моментом является то, что средняя скорость (увеличилась на 35%) по узлам растет быстрее максимальной (увеличилась на 19%) скорости, выраженной в байтах в секунду, что указывает на увеличение интенсивности передачи данных всеми узлами. Другим важным моментом является тот факт, что на всех проведенных экспериментах средняя и максимальная скорость передачи выраженная в пакетах в секунду почти неотличима. Это указывает на то, что число сообщений между узлами одно и то же, однако их длина различна, а передача более коротких сообщений сопряжена с большими накладными расходами.

Важно обратить внимание на то, что скорость, выраженная в пакетах в секунду, возрастает более чем в два раза более интенсивно, чем скорость выраженная в мегабайтах в секунду. При увеличении с 16 процессоров до 48 средняя и максимальная скорость, выраженная в пакетах в секунду, возросла на 77%.

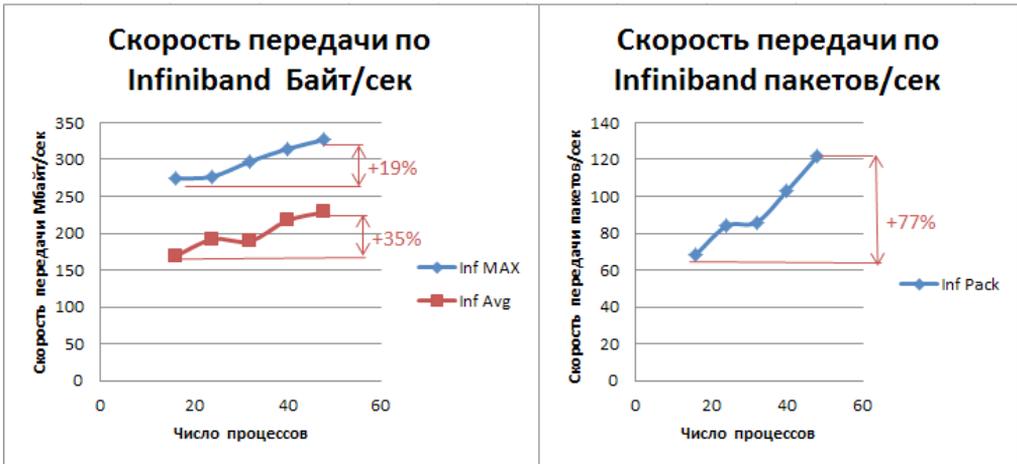


Рис. 6. Изменение максимальной и средней по всем узлам скорости передачи данных по сети Infiniband. В байт/сек и пакетах/сек при выполнении теста Linpack при размере матрицы 15 000 и изменении числа процессов

Это говорит о значительном снижении максимальной и средней длины пакета, а также о значительном увеличении числа пакетов, одновременно находящихся в состоянии передачи по сети Infiniband.

На рисунке 7 показана зависимость средней и максимальной длины пакета, передаваемого по сети Infiniband. Из полученного графика можно сделать вывод о том, что при увеличении числа процессов коммуникационная сеть становится сильнее загружена большим количеством пакетов малой длины. Так как длина сообщений снижается почти на треть, а их число возрастает на 70%, можно сделать вывод, что это может являться основной причиной снижения эффективности работы программы.

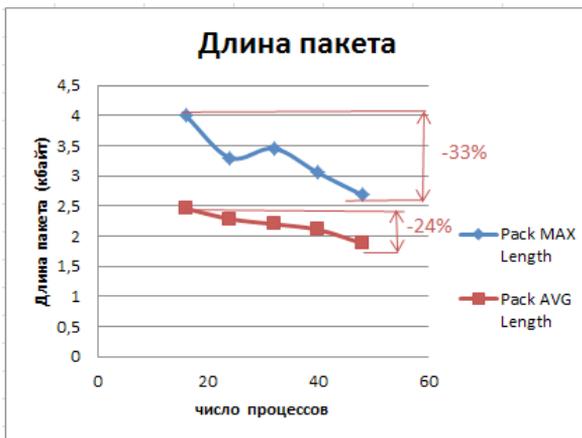


Рис. 7. Изменение максимальной и средней по всем узлам длины пакета передачи данных по Infiniband при выполнении теста Linpack при размере матрицы 15 000 и изменении числа процессов

В рассмотренном примере предложенного подхода к анализу масштабируемости параллельных программ практическая сложность проведения исследований заключается в очень большом объеме данных, необходимых для анализа. Объем данных растет очень быстро как при расширении области значений уже рассматриваемых параметров запуска, так и при добавлении новых параметров. При рассмотрении нескольких динамических характеристик необходимо собирать и анализировать очень большой объем информации о них, а также часто проводить дополнительный анализ введенных новых динамических характеристик. Это необходимо для выбора значений, образующих взаимозависимость между значениями, которые становятся показательными.

ЗАКЛЮЧЕНИЕ

Предложенный подход к анализу масштабируемости использует большое число данных о работе программы, позволяя проанализировать изменения в работе программы на разных уровнях ее работы. Технологии суперкомпьютерного кодизайна, используемые для анализа масштабируемости, позволяют изучать особенности программного и аппаратного уровня и делать выводы о путях улучшения масштабируемости программы: внесения изменений в алгоритмическую часть, программную реализацию или аппаратную платформу. Сложность предложенного подхода к изучению понятия масштабируемости высока за счет использования большого числа характеристик, но это компенсируется возможностью получения значительно большего объема информации о свойствах масштабируемости параллельных программ и получения выводов, обоснованно указывающих на причины снижения эффективности программ. Описанный подход может быть использован как в целях описания свойств параллельных программ [Фролов, 2015; Антонов и др., 2015], так и для повышения уровня общей эффективности работы больших вычислительных комплексов в целом за счет увеличения эффективности работы каждого запуска, используя полученную информацию о масштабируемости работающих на комплексе программ.

ЛИТЕРАТУРА

- [Адинец, 2013] Адинец А., Брызгалов П., Воеводин В. и др. Jobdigest – подход к исследованию динамических свойств задач на суперкомпьютерных системах // Вестник Уфимского государственного авиационного технического университета. — 2013. — Т. 17, № 2 (55). — С. 131–137.
- [Антонов и др., 2013] Антонов А.С., Жуматий С.А., Никитенко Д.А., Стефанов К.С., Теплов А.М., Швец П.А. Исследование динамических характеристик потока задач суперкомпьютерной системы // Вычислительные методы и программирование: Новые вычислительные технологии (Электронный научный журнал). 2013. Т. 14, № 2. С. 104–108.
- [Антонов, Теплов, 2014а] Антонов А.С., Теплов А.М. Использование данных системного мониторинга для определения факторов, уменьшающих масштабируемость приложения // Известия ЮФУ. Технические науки. 2014. № 12 (161). С. 90–101.

- [Антонов, Теплов, 2014б] Антонов А., Теплов А. О практической сложности понятия масштабируемости параллельных программ // Высокопроизводительные параллельные вычисления на кластерных системах (НПС 2014): Материалы XIV Международной конференции (г. Пермь, 10-12 ноября 2014 г.). Издательство Пермского национального исследовательского политехнического университета Пермь, 2014. С. 20–27.
- [Антонов и др., 2015] Антонов А.С., Воеводин Вад. В., Воеводин Вл. В., Теплов А.М., Фролов А.В. Первая версия Открытой энциклопедии свойств алгоритмов. // Параллельные вычислительные технологии (ПаВТ'2015): труды международной научной конференции (31 марта – 2 апреля 2015 г., г. Екатеринбург), Издательский центр ЮУрГУ Челябинск, 2015. С. 31–42.
- [Воеводин и др., 2012] Воеводин В.В., Жуматий С., Соболев С. и др. Практика суперкомпьютера “Ломоносов” // Открытые системы. 2012. № 7. С. 36–39.
- [Теплов, 2014] Теплов А.М. Об одном подходе к сравнению масштабируемости параллельных программ // Вычислительные методы и программирование: Новые вычислительные технологии (Электронный научный журнал). 2014. Т. 15, № 4. С. 697–711.
- [Фролов, 2015] Фролов А.В., Воеводин В.В., Коньшин И.Н., Теплов А.М. Исследование структурных свойств алгоритма разложения Холецкого: от давно известных фактов до новых выводов. // Параллельные вычислительные технологии (ПаВТ'2015): труды международной научной конференции (31 марта – 2 апреля 2015 г., г. Екатеринбург), Издательский центр ЮУрГУ Челябинск, 2015. С. 320–331.

ANALYSIS OF THE PARALLEL PROGRAMS SCALABILITY BASED ON SUPERCOMPUTER CO-DESIGN TECHNOLOGIES

A.S. Antonov, A.M. Teplov

Research Computing Center M.V. Lomonosov Moscow State University
(RCC Lomonosov Moscow State University)

An approach to the analysis of parallel applications scalability using supercomputer co-design technologies based on the use of integrated and structured data on the parallel program. The proposed approach allows us to study the properties of the application on the basis of the data collected on the various dynamic characteristics of its work. Data is observed from the different levels of application runtime research, starting with the hardware level. This approach allows making informed judgments about potential reasons for the parallel applications efficiency degradation.

Keywords: Scalability, supercomputer co-design, efficiency, parallel computing, approach to scalability analysis.

Antonov Alexander Sergeyeovich – leading researcher, asa@parallel.ru

Teplov Alexey Mikhaylovich – junior researcher, alex-teplov@yandex.ru

АЛГОРИТМ «POPULATION ANNEALING» И ЕГО ПРИМЕНЕНИЕ К ДВУМЕРНОЙ МОДЕЛИ ИЗИНГА

Л.Ю. Бараш¹, Л.Н. Щур¹, М. Боровский²

¹ Институт теоретической физики им. Л.Д. Ландау РАН
(ИТФ им. Л.Д. Ландау РАН), Научный центр РАН в Черногоровке

² Университет Павла Йозефа Шафарика в г. Кошице

Проведен анализ недавно предложенного алгоритма изучения моделей статистической физики. Разработана реализация алгоритма для графических процессоров. Изучается влияние параметров алгоритма (числа шагов Монте-Карло, размера популяции, числа температурных шагов, числа независимых запусков) на точность вычислений. Продемонстрировано отсутствие критического замедления алгоритма «Population Annealing» для двумерной модели Изинга.

Работа выполнена при финансовой поддержке гранта РФФ 14-21-00158, а также с использованием ресурсов суперкомпьютерного комплекса МГУ им. М.В. Ломоносова.

Ключевые слова: Population Annealing, модель Изинга, метод Монте-Карло, графический процессор.

ВВЕДЕНИЕ

Одним из сложных вопросов вычислительной статистической физики является моделирование низкотемпературного равновесия систем со сложным спектром свободной энергии, например, таким, какой возникает в спиновых стеклах и биомолекулах. Известные варианты методов Монте-Карло плохо применимы для таких систем. Например, при их применении моделирование может привести систему в локальный минимум энергии и находиться в нем бесконечное время. Таким образом, большинство методов оказывается неспособно корректно произвести выборку из равновесного распределения. Другой важной задачей является необходимость эффективной реализации алгоритмов Монте-Карло на современных вычислительных системах, которые, как правило, являются гетерогенными системами с центральным процессором и с графическими ускорителями на каждом узле. Для решения этих проблем в последнее время предлагаются и используются усовершенствованные алгоритмы Монте-Карло. Одним из таких алгоритмов является алгоритм Population Annealing (PA). Идея метода была предложена в работе [Hukushima, Iba, 2003], в современном усовершенствованном виде он был опубликован значительно позднее в работе [Machta, 2010]. Свойства этого алгоритма, однако, недостаточно изучены. Очевидно, что прежде чем изменять алгоритм к сложным системам, таким, как спиновые стекла, необходимо

Бараш Лев Юрьевич – н.с. ИТФ им. Л.Д. Ландау РАН; н.с. Научного центра РАН в Черногоровке; к.ф.-м.н., barash@itp.ac.ru

Щур Лев Николаевич – заведующий отделом, д.ф.-м.н., профессор, ИТФ им. Л.Д. Ландау РАН, Научный центр РАН в Черногоровке, shchur@chg.ru

Боровский Михал – университет Павла Йозефа Шафарика в г. Кошице, Словакия, аспирант

изучить свойства алгоритма на известных задачах, например, на двумерной модели Изинга, для которой хорошо известно как ее точное решение, так и изучено поведение других алгоритмов Монте-Карло. Это позволит нам провести изучение особенностей метода и влияния параметров алгоритма на эффективность применения метода и точность вычисления термодинамических величин.

ФОРМУЛИРОВКА АЛГОРИТМА

Схема алгоритма РА изображена на нижеследующем рисунке 1.

Инициализируем R_0 реплик для $\beta = 0$.

for k = K to 1 step -1 do

 вычисляем отношение стат. сумм $Q(\beta_k, \beta_{k-1})$

 for all $i \leq \tilde{R}_{\beta_k}$ do

 Вычисляем весовую функцию $\tau_i(\beta_k, \beta_{k-1})$

 Перевыборка: создаем $M[(R_{\beta'}/\tilde{R}_{\beta})\tau_i(\beta, \beta')]$ копий реплики i

 end for

 Вычисляем новый размер популяции $\tilde{R}_{\beta_{k-1}}$

 for all $i \leq \tilde{R}_{\beta_{k-1}}$ do

 Уравновешивание реплики i : производим θ_{k-1} проходов Монте-Карло

 end for

 Вычисляем наблюдаемые величины и свободную энергию для β_{k-1}

end for

Рис. 1. Псевдокод алгоритма «Population annealing»

Мы начинаем с инициализации R_0 реплик системы, то есть ее копий. Каждая реплика имеет одинаковый набор параметров, описывающих моделируемую систему. Начальная температура выбирается с бесконечным значением $\beta = 0$, и каждая реплика находится в случайно выбранном состоянии фазового пространства. Следующим шагом алгоритма является перевыборка. На этом шаге мы одновременно понижаем температуру для всего множества реплик. Алгоритм производит понижение температуры в заранее известной последовательности: $T_k = 1/\beta_k$, $T_{k-1} = 1/\beta_{k-1}$, ..., $T_0 = 1/\beta_0$, где $\beta_k = 0$ и $\beta_k < \beta_{k-1} < \dots < \beta_0$. Рассмотрим перевыборку из температуры $1/\beta$ в температуру $1/\beta'$ для \tilde{R}_{β} реплик, где $\beta' > \beta$. В работе [Machta, 2010] используется метод пуассоновской перевыборки, в котором число копий состояния i в новой популяции есть $n_i = M[\mu_i(\beta, \beta')]$, где

$$\mu_i(\beta, \beta') = (R_{\beta'}/\tilde{R}_{\beta})\tau_i(\beta, \beta'),$$

$$\tau_i(\beta, \beta') = \exp[-(\beta' - \beta)E_i]/Q(\beta, \beta'),$$

$$Q(\beta, \beta') = \frac{1}{\tilde{R}_{\beta}} \sum_{i=1}^{\tilde{R}_{\beta}} \exp[-(\beta' - \beta)E_i].$$

Здесь $M[a]$ – случайная целая величина с пуассоновским распределением и средним значением a . Если $n_i = 0$, то конфигурация i уничтожается. Размер новой популяции будет равен $\tilde{R}_{\beta'} = \sum_i n_i$. Следовательно, среднее отклонение величины $\tilde{R}_{\beta'}$ от заранее заданной величины $R_{\beta'}$ равно нулю, а среднеквадратичное отклонение равно $\sqrt{R_{\beta'}}$, оно много меньше $R_{\beta'}$, если $R_{\beta'} \geq 100$. Таким образом, размер популяции будет слегка отличаться от заранее заданных значений $\{R_{\beta'}\}$.

Наряду с пуассоновской перевыборкой существуют версии алгоритма с мультиномиальной перевыборкой [Machta, Ellis, 2011], остаточной перевыборкой [Douc, Carré, 2005] и перевыборкой методом ближайшего целого [Wang et al., 2015].

Для мультиномиальной перевыборки размер популяции $\tilde{R}_{\beta'} = \sum_i n_i$ совпадает с заранее заданным значением $R_{\beta'}$, а количества копий $n_1, n_2, \dots, n_{R_{\beta'}}$ выбираются совместно из мультиномиального распределения $f[n_1, n_2, \dots, n_{R_{\beta'}}; R_{\beta'}, \tau_1/R_{\beta'}, \tau_2/R_{\beta'}, \dots, \tau_{R_{\beta'}}/R_{\beta'}]$.

Для остаточной перевыборки также выполняется $\tilde{R}_{\beta'} = R_{\beta'}$. Числа копий в методе остаточной перевыборки выбираются совместно при помощи мультиномиального распределения $f[n_1 - \lfloor \mu_1 \rfloor, n_2 - \lfloor \mu_2 \rfloor, \dots, n_{R_{\beta'}} - \lfloor \mu_{R_{\beta'}} \rfloor; R - T, \frac{\mu_1 - \lfloor \mu_1 \rfloor}{R - T}, \frac{\mu_2 - \lfloor \mu_2 \rfloor}{R - T}, \dots, \frac{\mu_{R_{\beta'}} - \lfloor \mu_{R_{\beta'}} \rfloor}{R - T}]$, где $R = R_{\beta'}$, $T = \sum_{i=1}^{R_{\beta'}} \lfloor \mu_i \rfloor$. Здесь $\lfloor x \rfloor$ обозначает наибольшее целое число, не превосходящее x .

Для перевыборки методом ближайшего целого размер популяции $\tilde{R}_{\beta'}$ может отличаться от заранее заданной величины $R_{\beta'}$, при этом с вероятностью $\mu_i - \lfloor \mu_i \rfloor$ выбираем $n_i = 1 + \lfloor \mu_i \rfloor$, в противном случае выбираем $n_i = \lfloor \mu_i \rfloor$.

Итак, во всех используемых методах перевыборки числа n_i , являясь целыми неотрицательными числами, выбираются из некоторого дискретного распределения вероятности, такого, что среднее значение n_i равно μ_i , дисперсия величины n_i не превышает μ_i , а среднеквадратичное отклонение величины $\tilde{R}_{\beta'} = \sum_i n_i$ от $R_{\beta'}$ мало по сравнению с $R_{\beta'}$. При выполнении этих требований конкретный выбор метода не является существенно важным.

Следующий шаг алгоритма называется уравниванием. На этом шаге для каждой реплики производится θ проходов решетки алгоритмом Монте-Карло, основанным на марковской цепи, например, алгоритмом Метрополиса. Наконец, затем вычисляются наблюдаемые величины как средние значения по популяции. Поскольку величина $Q(\beta, \beta')$ с хорошей точностью приближает отношение статистических сумм при температурах $1/\beta$ и $1/\beta'$, то свободную энергию \tilde{F} тоже можно получать для каждой температуры в процессе вычислений при помощи выражения

$$-\beta_k \tilde{F}(\beta_k) = \sum_{l=k}^{k+1} \ln Q(\beta_l, \beta_{l-1}) + \ln \Omega,$$

где Ω – размер фазового пространства системы. Например, $\Omega = 2^N$ для модели Изинга, содержащей N спинов. Вычисление наблюдаемых величин как средних

по популяции является важным отличием алгоритма «Population annealing» от методов Монте-Карло, основанных на марковской цепи, где наблюдаемые величины обычно вычисляются как среднее по марковскому процессу.

Перевыборка и уравнивание производятся до тех пор, пока не завершится заранее заданная последовательность температур (см. рисунок 1).

РЕАЛИЗАЦИЯ АЛГОРИТМА ДЛЯ GPU

Нами была разработана реализация алгоритма для модели Изинга с использованием технологии CUDA и возможностей графических ускорителей Nvidia. Вычислительные ядра, выполняемые на графическом процессоре, разработаны для следующих частей программы: 1) процесс уравнивания; 2) параллельное вычисление энергии, намагниченности и их моментов всех реплик системы; 3) вычисление величины $Q(\beta, \beta')$; 4) вычисление числа копий каждой реплики для пуассоновской переборки.

Для реализации параллельного алгоритма для процесса уравнивания мы использовали методы, приведенные в работах [Weigel, 2011; Weigel, 2012]. В этих работах предложен метод «декомпозиция двойной шахматной доски»: исходная решетка из $N \times N$ узлов (спинов) рассматривается как $B \times B$ блоков, каждый из которых является решеткой из $T \times T$ узлов, которая, в свою очередь, делится на четные и нечетные узлы в шахматном порядке. Каждый блок нитей обрабатывает только узлы одной из подрешеток размером $T \times T$. Для одновременного использования и обновления конфигурации спинов такой подрешетки при помощи множества нитей блока можно использовать быструю разделяемую память, если значение T не слишком велико, например, для $T=16$. Поскольку необходимо параллельно и независимо производить обновление спинов решетки алгоритмом Метрополиса, то сначала обрабатываются нитями блока $T^2/2$ четных узлов подрешетки, а затем – $T^2/2$ нечетных узлов подрешетки. По той же причине блоками нитей сначала параллельно и независимо обрабатываются $B^2/2$ «четных» подрешеток размера $T \times T$, затем – $B^2/2$ «нечетных».

Важным является избавление от вычисления экспоненты при каждом обновлении каждого спина, поскольку вычисление экспоненты не является быстрой однопиковой операцией. Это несложно реализовать, поскольку возможных значений величины ΔE в алгоритме Метрополиса всего лишь небольшое конечное число. Например, для двумерной модели Изинга $\Delta E = 2J s_i \sum_{\text{ближ.сосед.}} s_j + 2H s_i$, где J – константа связи, а H – магнитное поле. Величина $\sum_{\text{ближ.сосед.}} s_j$ может принимать одно из пяти значений $-4, -2, 0, 2, 4$, а спин s_i – одно из двух значений ± 1 , поэтому величина ΔE может принимать только одно из десяти значений. Использование таблицы, в которой хранятся десять значений величины $\min(1, \exp(-\Delta E/T))$, помещенной в быструю текстурную или константную память графической карты, значительно ускоряет вычисление.

Наиболее быстрым способом генерации случайных чисел для выполнения уравнивания является заполнение перед каждым проходом Монте-Карло

массива в памяти GPU случайными числами. Размер такого массива равен полному числу обрабатываемых спинов (т.е. размеру популяции, умноженному на число спинов решетки). Для этого в библиотеке PRAND [Barash, Shchur, 2014] есть эффективные функции с названиями `rng_generate_gpu_array_`, где `rng` должно быть заменено на название одного из генераторов случайных чисел. Возможность легко менять тип генератора случайных чисел в рамках библиотеки PRAND позволяет тестировать зависимость деталей алгоритма от способа генерации случайных чисел, что позволяет избежать возможных корреляций с особенностями алгоритма [Shchur, 1999].

В алгоритме «Population Annealing» уравнивание производится сразу для большого количества реплик системы, поэтому каждый блок нитей обрабатывает подрешетку размера $T \times T$ одной из реплик. Количество реплик, как правило, достаточно велико, поэтому можно последовательно обрабатывать одним блоком нитей все B^2 подрешеток одной реплики (сначала четные подрешетки, затем нечетные), и это не приводит к дополнительному замедлению по сравнению с их параллельной обработкой. Для такой реализации число блоков нитей будет в точности равно размеру популяции.

Параллельное вычисление энергии, намагниченности и их моментов для всех реплик системы производится следующим образом. Каждый блок нитей последовательно обрабатывает B^2 подрешеток размера $T \times T$ одной из реплик системы при помощи T^2 параллельных нитей. Предполагаем, что $T = 2^n$. Для нахождения наблюдаемой величины необходимо вычислить сумму по всем узлам решетки. На первом шаге вычисления каждая нить вычисляет свою часть каждой суммы, соответствующую одному узлу решетки, при этом все вычисленные величины записываются в быструю разделяемую память. Таким образом, для каждой вычисляемой величины имеем массив в разделяемой памяти, состоящий из T^2 величин, которые необходимо сложить. Для этого каждая из первых $T^2/2$ нитей прибавляет к своему элементу массива номер i элемент массива номер $i + T^2/2$, затем каждая из первых $T^2/4$ нитей прибавляет к своему элементу массива номер i элемент массива номер $i + T^2/4$, и так далее. Это позволяет произвести необходимые суммирования быстрее по сравнению как с последовательным алгоритмом, так и с параллельным алгоритмом, использующем атомарные операции.

Параллельное вычисление величины $Q(\beta, \beta')$ сводится к суммированию большого количества слагаемых, каждое из которых может быть вычислено отдельной вычислительной нитью. Такую сумму можно разбить на куски размера T^2 , каждый из которых можно оптимизировать при помощи использования разделяемой памяти таким же образом, как и рассмотренное выше вычисление энергий.

Вычисление числа копий в пуассоновской перевыборке сводится к независимому вычислению величин τ_i , μ_i и n_i для каждой реплики, где i – номер реплики. Чтобы произвести это вычисление параллельно с использованием графического ускорителя, достаточно поручить каждой нити вычисление для одной из реплик.

Дальнейшая оптимизация алгоритма может быть произведена при помощи применения современной версии техники «Asynchronous multi-spin coding» (AMSC), которая позволяет ускорить вычисления в несколько раз [Lulli et al., 2015]. Идея метода AMSC состоит в том, чтобы хранить в одном элементе массива спины нескольких (например, восьми) реплик, находящиеся в одном и том же узле системы. При этом формулируется версия алгоритма Метрополиса, которая позволяет обновить все спины, находящиеся в одном элементе массива, при помощи только одного случайного числа и небольшого количества быстрых логических операций, таких как XOR. Таким образом, можно сохранить использование метода «двойной шахматной доски» для процесса уравнивания, но обрабатывать по 8 реплик за раз. При этом возникают корреляции между репликами, однако влияние таких корреляций можно минимизировать.

Для применения техники AMSC к алгоритму «Population Annealing» важно убедиться, чтобы не возникло ситуации, такой, что в одном элементе массива хранятся спины реплик, которые были созданы как копии одной и той же системы во время предыдущей перевыборки. Ясно, что такая ситуация может привести к значительным корреляциям между репликами и к ошибочным результатам вычислений. Для исключения такой ситуации можно использовать следующий метод. Предположим, что у нас имеется реплик ($8R$), спины которых записаны в трехмерном массиве размером $R \times N \times N$, т.е. в элементе массива с индексом (i, k, l) записаны спины восьми реплик с номерами $8i, \dots, 8i+7$, находящиеся в одном и том же узле (k, l) решетки. Тогда последовательность создания копий реплик в течение процесса перевыборки должна быть организована, например, таким образом, чтобы сначала создавались реплики с номерами $8i$, где $i = 0, 1, 2, \dots, R'$, затем – реплики с номерами $(8i+1)$, $i = 0, 1, 2, \dots, R'$, и т.д. Важно проследить, чтобы номера копий одной и той же реплики имели один и тот же остаток от деления на 8. Такая последовательность создания реплик исключит возможность кодирования в одном элементе массива одинаковых спинов, находящихся в копиях одной и той же реплики на одном и том же узле и предотвратит возникновение значительных корреляций.

ВЛИЯНИЕ КОЛИЧЕСТВА ШАГОВ МОНТЕ-КАРЛО

На рисунках 2 и 3 приведены результаты вычислений двумерной модели Изинга на решетке 32×32 алгоритмом «Population Annealing» с разным числом θ проходов Монте-Карло. Использовалась программа со следующими параметрами: число реплик $R = 10^3$, шаг по обратной температуре $\Delta\beta = 10^{-3}$, число независимых запусков $M = 200$. Для сравнения на рисунке 2 приведен график точного решения для теплоемкости в двумерной модели Изинга, который проще всего построить, используя выражения из работы [Beale, 1996]. Как видно из рисунка 2, результат вычисления теплоемкости близок к точному решению. Из рисунков видно, что начиная с некоторой сравнительно небольшой величины θ реплики можно считать уравновешенными. Дальнейшее увеличение θ не дает каких-либо

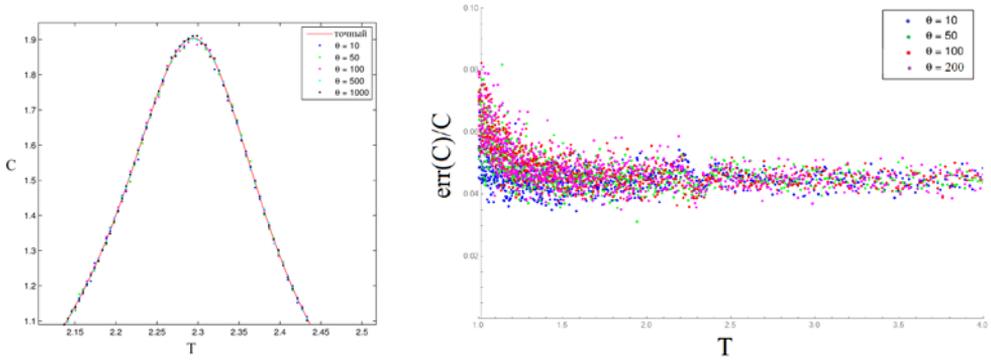


Рис. 2. Слева: зависимость теплоемкости от температуры – результат вычислений для разных значений θ , а также точное решение. Справа: стандартное отклонение вычисленных 200 значений теплоемкости, деленное на теплоемкость, в зависимости от температуры, для разных значений θ

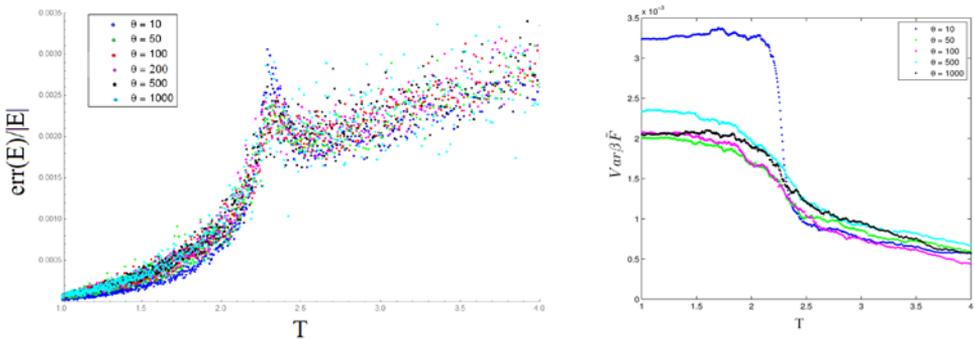


Рис. 3. Слева: стандартное отклонение вычисленных 200 значений внутренней энергии, деленное на внутреннюю энергию, в зависимости от температуры. Справа: дисперсия обезразмеренной свободной энергии в зависимости от температуры

преимуществ, поскольку и ошибка вычисления и дисперсия свободной энергии остаются практически неизменными. Алгоритму «Population annealing» необходимо для уравнивания значительно меньшее значение θ , чем алгоритмам Монте-Карло, основанным на марковской цепи.

ВЛИЯНИЕ РАЗМЕРА ПОПУЛЯЦИИ

На рисунках 4 и 5 приведены результаты вычислений двумерной модели Изинга на решетке 32×32 алгоритмом «Population Annealing» с разным размером популяции R . Использовалась программа со следующими параметрами: число проходов Монте-Карло $\theta = 100$, шаг по обратной температуре $\Delta\beta = 10^{-3}$, число независимых запусков $M = 200$. Вычисления подтверждают, что стандартная ошибка пропорциональна $1/\sqrt{R}$, где R – размер популяции.

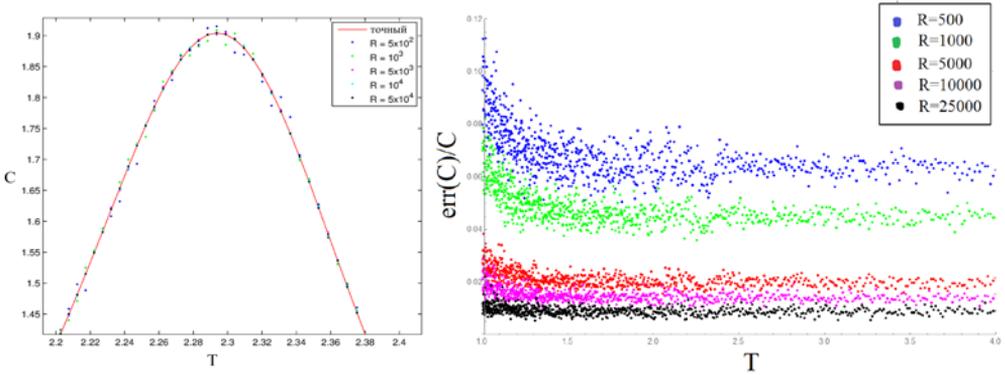


Рис. 4. Слева: зависимость теплоемкости от температуры, а также точное решение. Справа: стандартное отклонение 200 значений теплоемкости, деленное на теплоемкость, в зависимости от температуры. Приведены результаты вычислений для разных значений R

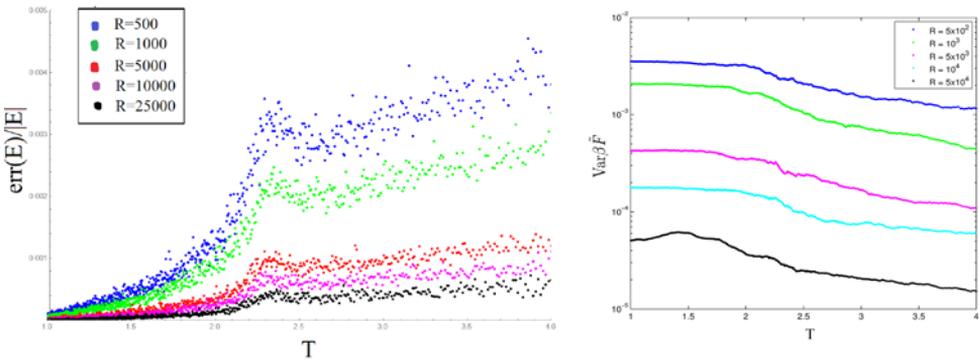


Рис. 5. Слева: стандартное отклонение 200 значений внутренней энергии, деленное на внутреннюю энергию, в зависимости от температуры. Справа: дисперсия обезразмеренной свободной энергии в зависимости от температуры. Приведены результаты вычислений для разных значений R

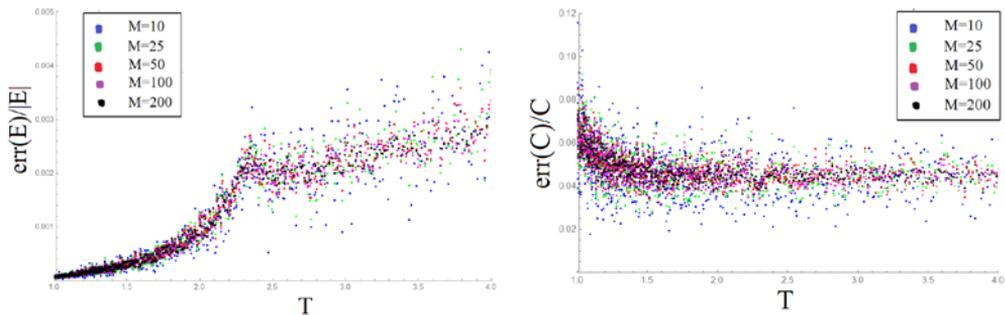


Рис. 6. Слева: стандартное отклонение M значений внутренней энергии, деленное на внутреннюю энергию, в зависимости от температуры. Справа: стандартное отклонение M значений теплоемкости, деленное на теплоемкость, в зависимости от температуры. Приведены результаты вычислений для разных значений M

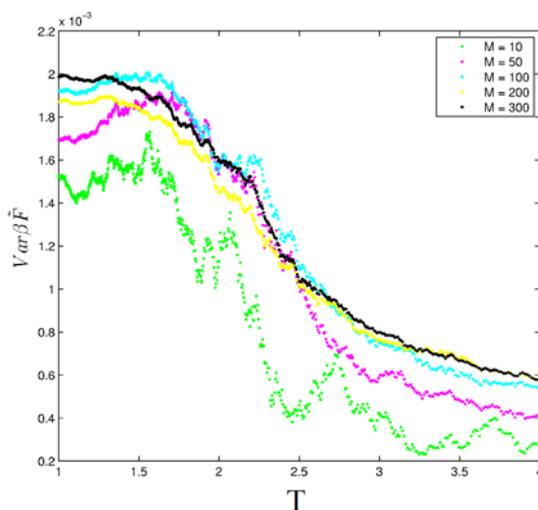


Рис. 7. Дисперсия безразмерной свободной энергии в зависимости от температуры. Приведены результаты вычислений для разных значений M

ВЛИЯНИЕ ЧИСЛА НЕЗАВИСИМЫХ ЗАПУСКОВ

На рисунках 6 и 7 приведены результаты вычислений двумерной модели Изинга на решетке 32×32 алгоритмом «Population Annealing» с разным количеством независимых запусков M . Использовалась программа со следующими параметрами: число реплик $R = 10^3$, число проходов Монте-Карло $\theta = 100$, шаг по обратной температуре $\Delta\beta = 10^{-3}$, число независимых запусков $M = 200$. Рисунки показывают, что в среднем стандартное отклонение не зависит от M , однако разброс величин стандартного отклонения уменьшается при увеличении M . Конечно, большее значение M позволяет более точно оценить ошибку вычислений.

МЕТОД ВЗВЕШЕННЫХ СРЕДНИХ

В работе [Machta, 2010] предлагается метод вычисления средних при помощи взвешенного усреднения результатов независимых запусков программы. Именно, предлагается в качестве взвешенного среднего для наблюдаемой функции A вычислять величину $\bar{A}(\beta) = \sum_{r=1}^M \bar{A}_r(\beta) \omega_r(\beta)$, где

$$\omega_r(\beta) = \exp(-\beta \bar{F}_r(\beta)) / \sum_{r=1}^M \exp(-\beta \bar{F}_r(\beta)).$$

Здесь $\bar{F}_r(\beta)$ – значение свободной энергии, вычисленное для температуры $1/\beta$ при выполнении запуска программы номер r .

На рисунке 8 приведено для двумерной модели Изинга на решетке 64×64 нормированное отличие теплоемкости (слева) и энергии (справа), вычисленных методом взвешенных средних для $M = 200$ запусков и вычисленных при

помощи простого усреднения тех же M значений. Использовалась программа со следующими параметрами: размер популяции $R = 5000$, число проходов Монте-Карло $\theta = 100$, шаг по обратной температуре $\Delta\beta = 10^{-3}$, число независимых запусков $M = 200$. Из рисунков видно, что для двумерной модели Изинга для достаточно большого размера популяции отличие между величиной, полученной методом взвешенного усреднения и величиной, полученной методом простого усреднения, очень мало.

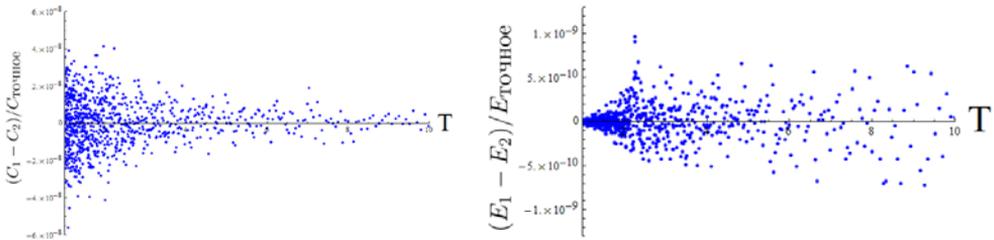


Рис. 8. Нормированное отличие теплоемкости (слева) и энергии (справа), вычисленных методом взвешенных средних для $M = 200$ запусков, и вычисленных при помощи простого усреднения тех же значений

ВЛИЯНИЕ КОЛИЧЕСТВА ТЕМПЕРАТУРНЫХ ШАГОВ

На рисунке 9 представлено стандартное отклонение 200 значений теплоемкости, деленное на теплоемкость, для $N_T = 50$, $N_T = 100$, $N_T = 200$, $N_T = 500$, $N_T = 1000$ и $N_T = 2000$, шагов по обратной температуре при расчете двумерной модели Изинга на решетке 32×32 . При этом использовались следующие минимальная и максимальная обратные температуры: $\beta_{min} = 0$, $\beta_{max} = 3$. Другие параметры программы: размер популяции $R = 500$, число проходов Монте-Карло $\theta = 100$, число независимых запусков $M = 200$. Видно, что число температурных шагов влияет незначительно на точность вычислений. Также из рисунка видно, что при небольших значениях температуры точность вычислений теплоемкости падает, как и следовало ожидать. Для увеличения точности следует увеличивать размер популяции для небольших температур.

Для решеток большего размера число температурных шагов должно превышать некоторое критическое значение (зависящее от величины θ) для того, чтобы точность вычислений не возрастала в критической области.

ОТСУТСТВИЕ КРИТИЧЕСКОГО ЗАМЕДЛЕНИЯ

На рисунках 10 и 11 изображена зависимость относительного отклонения теплоемкости и внутренней энергии от точного решения для решеток 16×16 , 32×32 , 48×48 , 64×64 в области температур $1 \leq T \leq 10$. Планки погрешностей изображают стандартное отклонение 200 значений. Использовалась программа со следующими параметрами: размер популяции $R = 5000$, число проходов Монте-Карло $\theta = 100$, шаг по обратной температуре $\Delta\beta = 10^{-3}$, число независимых запусков

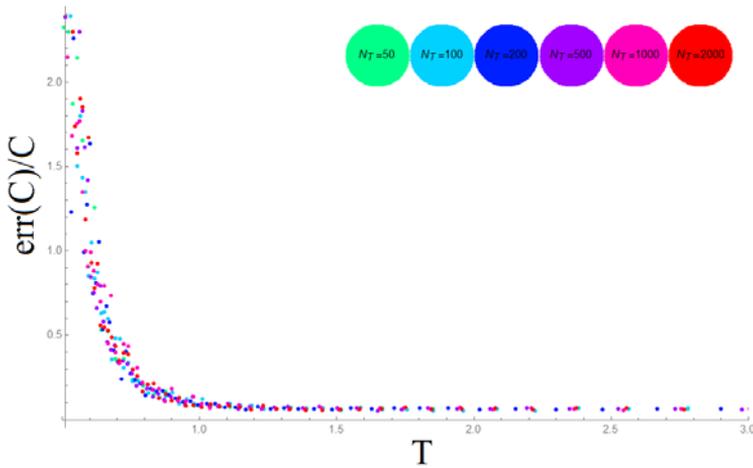


Рис. 9. Стандартное отклонение 200 значений теплоемкости, деленное на теплоемкость, для разного количества температурных шагов

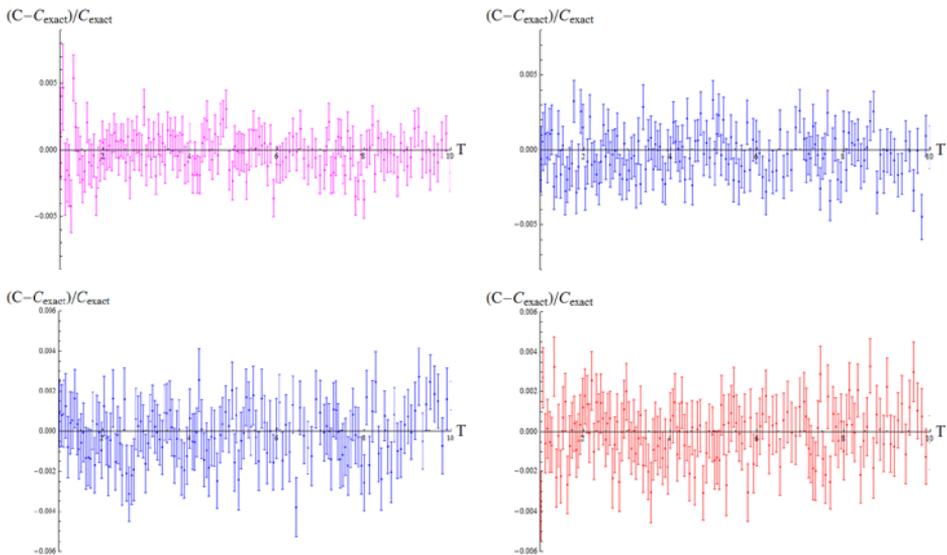


Рис. 10. Температурная зависимость относительного отклонения теплоемкости от точного решения для решеток 16×16 , 32×32 , 48×48 , 64×64 . Планки погрешностей изображают стандартное отклонение 200 значений

$M = 200$. Для многих алгоритмов Монте-Карло, основанных на марковской цепи, критическое замедление приводит к росту ошибок при увеличении размеров системы в связи с ростом корреляционного времени в критической области. Из рисунков видно, что в данном случае ошибка вычислений энергии и теплоемкости не увеличивается при увеличении размера системы. Таким образом, из результатов вычислений не видно, чтобы алгоритм «Population Annealing» был подвержен критическому замедлению.

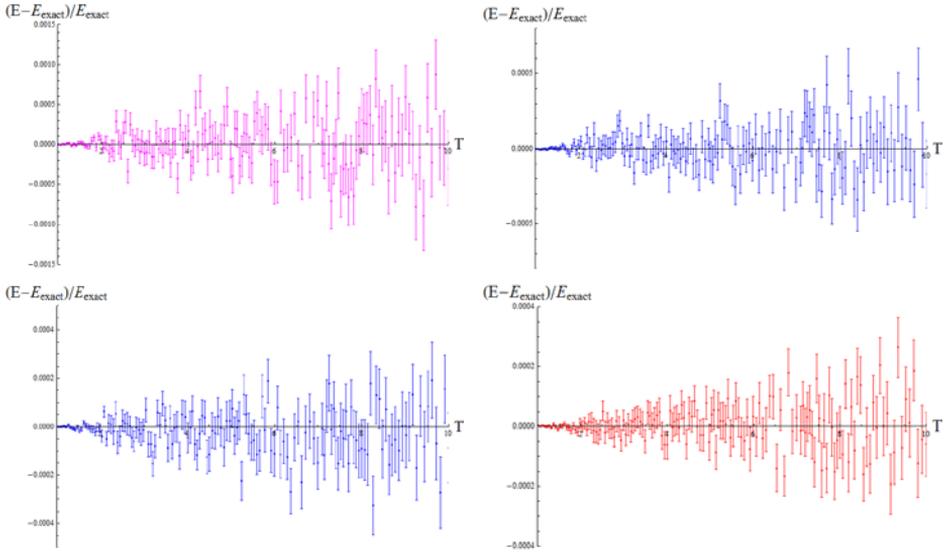


Рис. 11. Температурная зависимость относительного отклонения внутренней энергии от точного решения для решеток 16×16 , 32×32 , 48×48 , 64×64 . Планки погрешностей изображают стандартное отклонение 200 значений

КУМУЛЯНТ БИНДЕРА И ОШИБКА ЕГО ВЫЧИСЛЕНИЯ

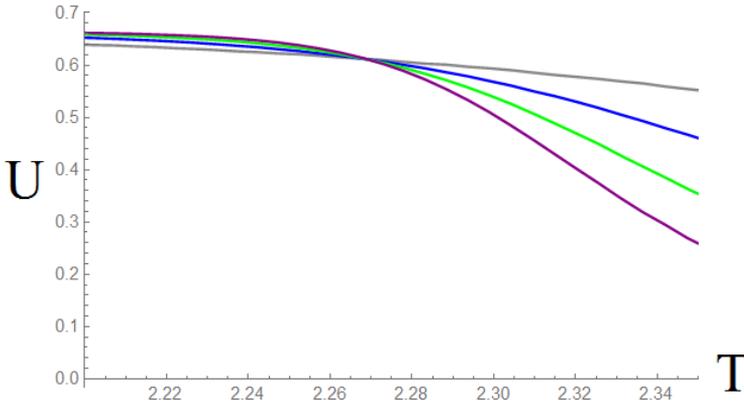


Рис. 12. Кумулянт биндера для двумерной модели Изинга для решеток 16×16 (серый график), 32×32 (синий график), 48×48 (зеленый график), 64×64 (фиолетовый график)

На рисунке 12 изображена температурная зависимость кумулянта Биндера [Binder, 1981a; Binder, 1981b]

$$U = 1 - \frac{\langle M^4 \rangle}{3 \langle M^2 \rangle^2}$$

для решеток 16×16 (серый график), 32×32 (синий график), 48×48 (зеленый график), 64×64 (фиолетовый график). Как известно, для фазового перехода второго рода кумулянты Биндера разных размеров сетки пересекаются в критической

точке. Среднеквадратичное отклонение для вычисленных значений кумулянта в критической точке равно 0,0037 для решетки 16×16 , 0,0034 для решетки 32×32 , 0,0039 для решетки 48×48 , 0,0040 для решетки 64×64 . Таким образом, существенного увеличения ошибки вычислений кумулянта с увеличением размера сетки не наблюдается, что также является косвенным свидетельством отсутствия эффекта критического замедления для рассматриваемого алгоритма.

ЗАКЛЮЧЕНИЕ

Проведенное детальное тестирование алгоритма «Population Annealing» на двумерной модели Изинга подтверждает, что алгоритм является перспективным алгоритмом вычислительной статистической физики, свободным от недостатков классических методов Монте-Карло, таких как критическое замедление, а также хорошо подходящим для использования на параллельных системах с графическими ускорителями.

Авторы благодарны Мартину Вейгелю за полезные обсуждения.

Реализация алгоритма для GPU была выполнена при частичной поддержке гранта РФФИ 13-07-00570.

ЛИТЕРАТУРА

- [Воеводин и др., 2012] Воеводин В.В., Жуматий С.А., Соболев С.И., Антонов А.С., Брызгалов П.А., Никитенко Д.А., Стефанов К.С., Воеводин В.В. Практика суперкомпьютера “Ломоносов” // Открытые системы. - Москва: Издательский дом “Открытые системы”, N 7, 2012. С. 36-39.
- [Barash, Shchur, 2014] *Barash L. Yu., Shchur L.N.* PRAND: GPU accelerated parallel random number generation library: Using most reliable algorithms and applying parallelism of modern GPUs and CPUs // *Comput. Phys. Commun.*, 185(4), pp. 1343-1353 (2014).
- [Beale, 1996] *Beale P.D.* Exact Distribution of Energies in the Two-Dimensional Ising Model // *Phys. Rev. Lett.* 76 (1), pp.78-81 (1996).
- [Binder, 1981a] *Binder K.* Finite size scaling analysis of ising model block distribution functions // *Zeitschrift für Physik B Condensed Matter*, Vol. 43, Issue 2, pp. 119-140 (1981).
- [Binder, 1981b] *Binder K.* Critical Properties from Monte Carlo Coarse Graining and Renormalization // *Phys. Rev. Lett.* 47, 693 (1981).
- [Douc, Cappé, 2005] *Douc R., Cappé O.* Comparison of resampling schemes for particle filtering // *Proceedings of 4th International Symposium on Image and Signal Processing and Analysis ISPA 2005*, pp. 64–69 (2005).
- [Hukushima, Iba, 2003] *Hukushima K., Iba Y.* Population Annealing and Its Application to a Spin Glass // *AIP Conf. Proc.* 690, 200 (2003).
- [Lulli et al., 2015] *Lulli M., Bernaschi M., Parisi G.* Highly optimized simulations on single- and multi-GPU systems of the 3D Ising spin glass model // *Comp. Phys. Commun.*, 196, pp. 290-303 (2015).
- [Machta, 2010] *Machta J.* Population annealing with weighted averages: A Monte Carlo method for rough free-energy landscapes // *Phys. Rev. E* 82, 026704 (2010), 3 pages.

- [Machta, Ellis, 2011] *Machta J., Ellis E.S.* Monte Carlo Methods for Rough Free Energy Landscapes: Population Annealing and Parallel Tempering // *Journal of Statistical Physics*, 144 (3), pp. 541-553 (2011).
- [Shchur, 1999] *Shchur L.N.* On the quality of random number generators with taps // *Comput. Phys. Commun.*, 121-122, 83-85 (1999).
- [Wang et al., 2015] *Wang W., Machta J., Katzgraber H.G.* Comparing Monte Carlo methods for finding ground states of Ising spin glasses: Population annealing, simulated annealing, and parallel tempering // *Phys. Rev. E* 92, 013303 (2015).
- [Weigel, 2011] *Weigel M.* Simulating spin models on GPU // *Comp. Phys. Commun.* 182(9), 1833-1836 (2011).
- [Weigel, 2012] *Weigel M.* Performance potential for simulating spin models on GPU // *Journal of Computational Physics*, Volume 231, Issue 8, pp. 3064-3082 (2012).

POPULATION ANNEALING ALGORITHM AND ITS APPLICATION TO THE TWO-DIMENSIONAL ISING MODEL

*L.Yu. Barash*¹, *L.N. Shchur*¹, *M. Borovsky*²

¹ Landau Institute for Theoretical Physics, Russian Academy of Sciences
Science center in Chernogolovka, Russian Academy of Sciences

² University of Pavol Jozef Šafárik in Košice

The analysis of the recently proposed algorithm for computational statistical physics is carried out. The realization of the algorithm for GPU is developed. We study influence of parameters of the algorithm (such as number of Monte Carlo sweeps, population size, number of temperature steps, number of independent runs) on the accuracy of the calculations. It is shown that there is no critical slowing down in the application of Population Annealing to the two-dimensional Ising model.

Keywords: Population Annealing, Ising model, Monte Carlo, GPU.

Barash Lev Yuryevich – researcher, Landau Institute for Theoretical Physics; researcher, Science Center in Chernogolovka RAS. PhD, barash@itp.ac.ru

Shchur Lev Nikolaevich – head of department, doctor of physical and mathematical sciences, professor, shchur@chg.ru

Michal Borovský – postgraduate student, University of Pavol Jozef Šafárik in Košice, Slovakia

СИНТЕЗ АЛГОРИТМОВ ПОСТРОЕНИЯ ДОВЕРЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ С ТРЕХЗВЕННОЙ АРХИТЕКТУРОЙ

А.В. Белов, Д.В. Пашков

Московский институт электроники и математики

Научно-исследовательского университета Высшая школа экономики (МИЭМ НИУ ВШЭ)

В работе рассмотрены вопросы построения доверенных информационных систем (ИС). В соответствии с рассмотренными моделями угроз и нарушителя обоснована необходимость использования трехзвенной архитектуры при создании защищенных КИС. Предложены алгоритмы построения доверенных ИС, отвечающих повышенным требованиям к информационной безопасности.

Ключевые слова: КИС, защита КИС, доверенные системы, защита баз данных, угрозы ИС, трехзвенная архитектура, построение доверенных систем.

Защита корпоративных информационных систем (КИС) является одной из самых сложных задач, стоящих перед специалистами, отвечающими за обеспечение информационной безопасности. С одной стороны, для работы с КИС необходимо предоставлять доступ к данным всем сотрудникам, кто должен осуществлять сбор и обработку конфиденциальных данных. С другой стороны, требуется ограничить круг лиц, имеющих доступ к базе данных, являющейся ядром КИС, сохранять конфиденциальность, целостность и доступность данных. При этом в открытой печати практически отсутствуют материалы, посвященные методике комплексного решения задачи защиты данных в КИС, которую можно было бы применять при проектировании КИС различного назначения. Решения по защите данных КИС могут быть разными, они могут иметь различную архитектуру, использовать различные программные компоненты.

Необходимо отметить, что большинство применяемых решений являются коммерческими, часто не имеющими российских сертификатов, то есть не являются доверенными. Главной особенностью таких решений является их закрытость, все разработки являются коммерческой тайной и их исходные коды не могут быть полностью проанализированы с точки зрения информационной безопасности. Например, проверка на недеklarированные возможности не проводится по причине отказа поставщиков решений предоставить исходные коды в специальные сертифицирующие органы, такие, как ФСТЭК, ФСБ и другие. Соответственно, данные продукты не являются доверенными.

Современные корпоративные информационные системы являются документоориентированными системами. Основными информационными объектами

Белов Александр Владимирович – руководитель департамента прикладной математики Московского института электроники и математики Научно-исследовательского университета Высшая школа экономики (МИЭМ НИУ ВШЭ), к.т.н., AVBelov@hse.ru

Пашков Денис Витальевич – выпускник Московского института электроники и математики Научно-исследовательского университета Высшая школа экономики (МИЭМ НИУ ВШЭ), 2015 г., специальность «Компьютерная безопасность», denis__pashkov@mail.ru

КИС являются документы, транзакции и пользователи системы. Это означает, что при проектировании доверенных КИС должны быть учтены основные свойства хранимой информации – конфиденциальность, целостность, доступность. Таким образом, необходимо обеспечить доступ к информации полномочных пользователей, ограничить доступ для остальных, сохранить доступность и целостность данных, система должна удовлетворять требованиям безопасности в условиях модели нарушителя и модели угроз, отвечать требованиям к функционалу.

СТАТИСТИКА:

В ежегодном отчете компании Verizon об инцидентах, связанных с ЗИ, сообщается, что в среднем 30% атак совершается против БД, при этом при успешной атаке компрометируется 75% данных. Для сравнения, при успешных атаках на другие системы – 10% [DBIR, 2015].



Рис. 1. Количество компрометируемой информации

Для любых организаций наибольшее значение имеют, непосредственно, сами данные. Чем сложнее структура КИС компании и чем больше объем данных, циркулирующих в ней, тем выше риски потерь данных. При этом имеется множество решений по защите от внешнего нарушителя, и многие думают, что, используя данные системы защиты, данные их КИС в безопасности. Но это не так. Системы открыты для внутреннего нарушителя, у которого есть доступ к данным и каналам передачи информации.

Так как построить доверенную систему, используя недоверенные элементы, невозможно, разработано решение на основе трехзвенной архитектуры, позволяющее строить защищенные КИС.

Идея данного решения заключается в использовании доверенного клиента и сервера, являющегося посредником между клиентским приложением и сервером хранения и обработки данных. В данной схеме прямого взаимодействия между клиентом и недоверенной СУБД не допускается. Все запросы тщательно проверяет и обрабатывает сервер-посредник. Такой подход позволяет избавиться от известных уязвимостей СУБД.

Так же достоинствами применения трехзвенной архитектуры является:

1. Более высокий уровень защищенности данных.
Обусловлено тем, что сервер-посредник КИС может включать в себя подсистемы СОА, подсистемы фильтрации трафика, более детальный и глубокий анализ пакетов, проверки пользователей на основе стандартизированных криптографических протоколов и т.д.
2. Масштабируемость.
3. Возможность балансировки нагрузки.
4. Прозрачность для пользователя.

КИС с двухзвенной архитектурой, использующая доверенное клиентское ПО, взаимодействующее с недоверенным ПО, также не будет доверенной.

Именно поэтому с точки зрения безопасности трехзвенная архитектура имеет преимущество перед двухзвенной, так как вся логика работы в данном случае содержится на сервере приложений, являющемся доверенным, а СУБД выступает в роли хранилища данных.

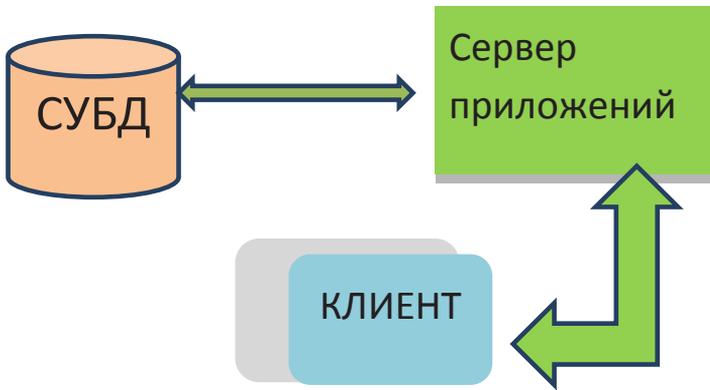


Рис. 2. Трехзвенная архитектура

Трехзвенная архитектура обладает явными преимуществами перед двухзвенной, по этой причине синтез алгоритмов построения доверенных информационных систем основан на данной архитектуре.

Одним из самых надежных методов защиты конфиденциальной информации является шифрование данных. При этом выделяют несколько подходов: шифрование всей базы данных, шифрование отдельных таблиц, строк или ячеек. Таким образом, уровень защищенности данных будет зависеть от стойкости криптографического алгоритма шифрования.

Применение данного метода защиты на практике, к сожалению, ограничено. Используя шифрование в качестве защиты данных, инструменты для поиска и обработки данных СУБД, такие как запросы, индексы, DML операторы не применимы. Поэтому для поиска и обработки информации необходимо постоянно выполнять шифрование/расшифровку данных. Данный процесс является трудоемким и снижает до минимума производительность КИС, особенно при обра-

ботке больших массивов данных. Шифрование применяется для защиты от подмены клиента, а также для защиты передаваемого трафика по каналам связи.

Другим надежным методом обеспечения конфиденциальности данных является разграничение доступа. Выделяется три основных вида разграничения доступа: мандатное, дискреционное, ролевое. На практике также применяются комбинации данных моделей.

Модели разграничения основаны на атрибутах пользователей КИС. Мандатное разграничение основано на метках доступа, дискреционное – на матрице доступа и прав пользователей. Таким образом, разграничение доступа в КИС построено на проверке предъявляемых пользователями идентификаторов.

В работе рассматривается один из подходов к проектированию доверенных корпоративных информационных систем.

ПОСТАНОВКА ЗАДАЧИ

В связи с тем, что проектируемые КИС должны хранить и обрабатывать разнокатегорийную информацию, необходимо построить модели угроз и потенциальных нарушителей. Обосновать выбор информационной архитектуры КИС, разработать для выбранной архитектуры модель данных и алгоритмы взаимодействия приложений, обеспечивающих требуемый уровень безопасности КИС в условиях защиты от внутреннего нарушителя, кем является пользователь КИС.

КЛАССЫ УГРОЗ:

1. Игнорирование введенных ограничений и нарушение правил пользования системой при эксплуатации системы.
2. Нарушение выделенных привилегий.
3. Выделение чрезмерных полномочий пользователям.
4. Эксплуатация уязвимостей СУБД.
5. Недостаточно полное логирование в журналах аудита.
6. Выполнение запросов заведомо приносящих ущерб БД (SQL-инъекции) [Проскурин, 2009].
7. Отказ в обслуживании.

ОСНОВНЫЕ УГРОЗЫ КИС:

1. Непреднамеренное удаление или искажение информации.
2. Игнорирование введенных ограничений при эксплуатации системы.
3. Нарушение правил пользования системой.
4. Предоставление третьим лицам идентификационных данных и доступа к ЭВМ.
5. Выявление логина/пароля.
6. Несанкционированный доступ к информации.
7. Понижение уровня доступа информации.

8. Модификация, искажение, удаление информации.
9. Регистрация в системе от имени другого пользователя.
10. Несанкционированное выполнение запросов к БД.
11. Неправомерный доступ к служебной информации.
12. Изменение роли, уровня доступа, группы пользователя.
13. Несанкционированные действия в клиентской программе.
14. Подмена клиента.
15. Работа различных пользователей под одним логином.

К БЕЗОПАСНОСТИ КИС ПРЕДЪЯВЛЯЮТСЯ СЛЕДУЮЩИЕ ТРЕБОВАНИЯ:

1. Использование атрибутов безопасности.
2. Генерация журнала аудита.
3. Просмотр журнала аудита.
4. Хранение журнала аудита.
5. Регистрация пользователей в системе.
6. Обеспечение конфиденциальности данных.
7. Ролевое разграничение пользователей.
8. Разграничение доступа к информации.
9. Управление доступом к информации.
10. Аутентификация пользователей.
11. Управление данными.
12. Управление потоками информации.
13. Импорт данных.
14. Экспорт данных.
15. Контроль целостности данных.

МОДЕЛЬ УГРОЗ И НАРУШИТЕЛЕЙ

Непреднамеренные действия:

1. Непреднамеренное удаление или искажение информации.
2. Игнорирование введенных ограничений при эксплуатации системы.
3. Нарушение правил пользования системой.
4. Предоставление третьим лицам идентификационных данных и доступа к ЭВМ.

Преднамеренные действия:

1. Несанкционированный доступ к информации.
2. Несанкционированное получение логина/пароля с/без доступа к ЭВМ.
3. Регистрация в системе от имени другого пользователя.
4. Подмена клиента.
5. Выполнение запросов заведомо приносящих ущерб БД.
6. Модификация, искажение, удаление информации.
7. Изменение роли, уровня доступа, группы пользователя.

8. Понижение уровня доступа информации.
9. Несанкционированное выполнение запросов к БД.
10. Неправомерный доступ к служебной информации.
11. Несанкционированные действия в клиентской программе.
12. Работа различных пользователей под одним логином.

Модель нарушителя:

1. Незарегистрированный пользователь БД, имеющий доступ к программе:
 - Не имеет идентификационных данных.
2. Зарегистрированный пользователь БД, имеющий доступ к программе:
 - Имеет идентификационные данные.
 - Обладает всеми возможностями своей роли в системе, группы и уровня доступа.
 - Имеет роль в информационной системе.
 - Не имеет доступа к служебным таблицам и информации.
3. Администратор КИС:
 - Имеет доступ к программе.
 - Имеет идентификационные данные.
 - Обладает всеми возможностями своей роли в системе, группы и уровня доступа.
 - Имеет роль администратора в информационной системе.
 - Имеет доступ к служебным таблицам и информации.
 - Имеет доступ к журналу аудита.
4. Администратор безопасности КИС:
 - Имеет доступ к программе.
 - Имеет идентификационные данные.
 - Обладает всеми возможностями своей роли в системе, группы и уровня доступа.
 - Имеет роль администратора в информационной системе.
 - Имеет доступ к служебным таблицам и информации.
 - Имеет доступ к журналу аудита.
 - Обладает всеми привилегиями и всеми уровнями доступа.
 - Может добавлять новых пользователей.
 - Устанавливает уровень доступа к информации.

МОДЕЛЬ ДАННЫХ ЯДРА СИСТЕМЫ

1. **Таблица пользователей** содержит данные пользователей, является служебной таблицей, к которой имеет доступ только администратор системы.

Содержит поля:

- Логин.
- Хэш – в БД хранится свертка пароля.
- Метка уровня доступа – указывает уровень доступа пользователя к информации.

- Метка группы доступа – является указателем в таблице групп доступа на свою группу.
- Системная роль – является указателем в таблице ролей, указывает роль пользователя в системе.
- Время жизни учетной записи.

2. **Таблица данных** содержит данные-документы, является рабочей таблицей, к которой имеют доступ все пользователи системы.

Содержит поля:

- Имя (название документа).
- Информация (текст).
- Группа доступа – отражает группу, к которой должен принадлежать пользователь для доступа к документу.
- Контрольная сумма – хэш значение, вычисленное по имени документа и содержания (текста).
- Дата создания.
- Уровень доступа – отражает уровень доступа.
- Владелец – указывает на имя пользователя, создавшего документ.

3. **Журнал аудита** содержит записи всех действий всех пользователей системы, является служебной таблицей, к которой имеют доступ только администратор безопасности и администратор системы.

Содержит поля:

- Код выполняемой операции – один из кодов фильтра запросов.
- Запрос от клиента – описание запроса или сам запрос от клиента.
- Статус запроса.
- Имя запросившего пользователя.
- Сеанс выполняемого запроса.
- Дата и время запроса.

4. **Таблица сеансов** содержит данные о всех зарегистрированных пользователях системы, после успешной аутентификации клиента, он регистрируется в системе, клиенту выдается специальный идентификатор – номер сеанса. Таблица является служебной, доступ к которой имеют только администратор безопасности и администратор системы.

Содержит поля:

- Имя пользователя.
- Состояние сеанса – указывает, открыт или закрыт сеанс.
- Время создания сеанса.
- Время закрытия сеанса.

5. **Фильтр запросов** содержит информацию, на основе которой осуществляется фильтрация запросов клиентов к БД, содержит идентификаторы всего множества запросов, является служебной таблицей.

Содержит поля:

- Код запроса – специальный идентификатор запроса.
- Метка роли – роль, которой должен обладать клиент для выполнения.

6. **Таблица групп** содержит имена групп и их идентификаторы (числовые значения), является служебной таблицей.

Содержит поля:

- Имя группы.
- Идентификатор (права группы).

7. **Таблица ролей** содержит имена ролей и их идентификаторы в системе, является служебной таблицей.

Содержит поля:

- Имя роли.
- Идентификатор.

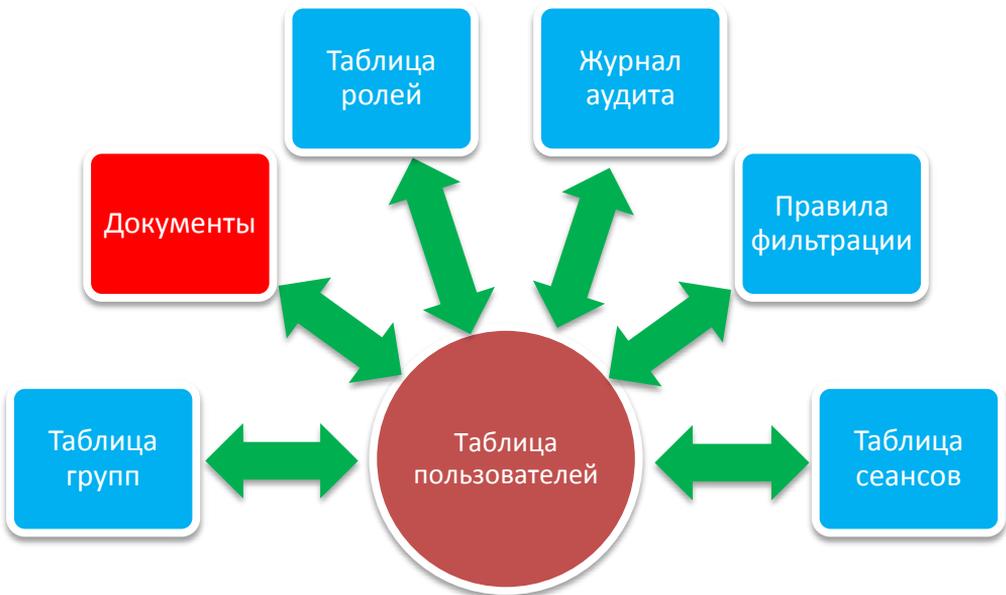


Рис. 3. Связь таблиц КИС

Как видно на схеме, все таблицы напрямую связаны с таблицей пользователей, это логично, ведь вся система разграничения доступа к разнокатегорийной информации построена на проверке атрибутов пользователя.

Данная модель документоориентированной КИС предназначена для хранения и обработки разнокатегорийной информации, т.е. информации разделенной на категории уровня доступа, доступ к которым происходит на основе мандатной и ролевой модели разграничения доступа.

ПОДСИСТЕМА ОБЕСПЕЧЕНИЯ ИБ



Рис. 4. Подсистема обеспечения ИБ

РАЗГРАНИЧЕНИЕ ДОСТУПА. АУТЕНТИФИКАЦИЯ.

Обеспечение конфиденциальности данных:

Конфиденциальности данных обеспечивается путем ограничения доступа к информации различных пользователей, при этом пользователю не известно ничего о недоступных для него данных, в том числе существуют ли они в БД или нет.

Разграничение доступа:

Каждый клиент ИС в соответствие с предъявленными требованиями обладает следующими атрибутами:

- Логин.
- Пароль.
- Метка уровня доступа.
- Метка группы доступа.
- Системная роль.
- Время жизни учетной записи.

Клиент получает доступ к данным БД и функционалу ИС на основе своих атрибутов. Разграничение доступа к данным основано на модели мандатного разграничения, используются метки доступа, группы доступа к информации.

Информация может быть трех уровней:

1. Открытая информация.
2. Конфиденциальная информация второго уровня.
3. Конфиденциальная информация первого уровня.

При этом каждая категория информации делится на группы. Таким образом, пользователю, имеющему достаточный уровень доступа к информации, необходимо принадлежать к группе, имеющей доступ к этой информации.

В данной информационной системе имеется три системных роли:

- Пользователь – клиент БД, работающий с доступной информацией.
- Администратор безопасности – клиент БД, имеющий доступ к управлению сеансами пользователей и к журналу регистрации событий. Следит за работой клиентов, регистрируемой в журнале аудита.
- Администратор системы – клиент, управляющий всеми пользователями, всей информацией системы, проводит модерацию новых данных.

АУТЕНТИФИКАЦИЯ ПОЛЬЗОВАТЕЛЕЙ:

Проверка подлинности, предъявляемых идентификаторов пользователем ведутся на основе криптографического алгоритма *Стрибог ГОСТ Р 34.11-2012* на секретном ключе сервера приложений.

АДМИНИСТРИРОВАНИЕ. АУДИТ:

Журнал аудита:

ИС регистрирует клиентов в системе, а также их действия в журнал, который обладает следующими атрибутами:

- Код выполняемой операции.
- Запрос от клиента.
- Статус запроса.
- Имя запросившего пользователя.
- Сеанс выполняемого запроса.
- Дата и время запроса.

ПОДСИСТЕМА ПРОВЕРКИ И ФИЛЬТРАЦИИ ОПЕРАЦИЙ:

Каждая операция имеет свой код, на основе которого система принимает решение, может ли ее выполнять отдельный пользователь. Данная проверка выполняется на основе таблицы «Фильтр запросов».

Запрос запускается в специальной тестовой среде, после этого, проводится анализ результата на предмет возможного нанесения вреда КИС.

Производится контроль целостности критически важной информации.

КОНТРОЛЬ ЦЕЛОСТНОСТИ ДАННЫХ:

Основан на криптографическом алгоритме *Стрибог ГОСТ Р 34.11-2012*.

ПОДСИСТЕМА УПРАВЛЕНИЯ ИНФОРМАЦИЕЙ:

Позволяет управлять импортом/экспортом информации.

Производит модерацию добавленных данных, изменение атрибутов информации.

ИМПОРТ/ЭКСПОРТ ИНФОРМАЦИИ:

Предусмотрен функционал ввода/вывода информации, вывод производится на основе политики доступа к информации, после ввода информации документ проходит модерацию суперпользователем, после чего только становится доступным.

ПОДСИСТЕМА АДМИНИСТРИРОВАНИЯ:

Включает в себя управление пользователями (добавление новых, удаление, изменение атрибутов пользователей).

Возможность управления сеансами и подключениями, а также блокирования хостов.

Мониторинг подсистемы аудита, реакция на события.

Роли системы:

Пользователи ИС для работы с разнокатегорийной информацией обладают следующими возможностями:

- Доступ к определенной информации.
- Возможность выполнения запросов к БД.
- Поиск необходимой информации по различным параметрам.
- Добавление новых данных в БД.
- Просмотр доступных к просмотру данных.

Администратор безопасности обладает следующими возможностями:

- Включает возможности обычного пользователя.
- Имеет доступ к журналу аудита, может выбирать данные из журнала по различным параметрам.
- Имеет доступ к части служебной информации (о сеансах).
- Управления сеансами пользователей.

Администратор системы (суперпользователь):

- Включает возможности пользователя и администратора безопасности.
- Доступна вся информация о системе.
- Доступны все функции панели администрирования.
- Возможность управления пользователями.
- Модерация данных в БД.

Администратор системы выполняет функции управления доступом к информации, он может менять роль, уровень доступа, принадлежность к группе пользователя, также осуществляет управление данными ИС: может изменить уровень доступа или группу доступа данных (документов), осуществляет модерацию данных. Администратор системы создает резервные копии данных.

СТРУКТУРНО-ФУНКЦИОНАЛЬНАЯ СХЕМА ВЗАИМОДЕЙСТВИЯ
(ОБЩИЙ АЛГОРИТМ)

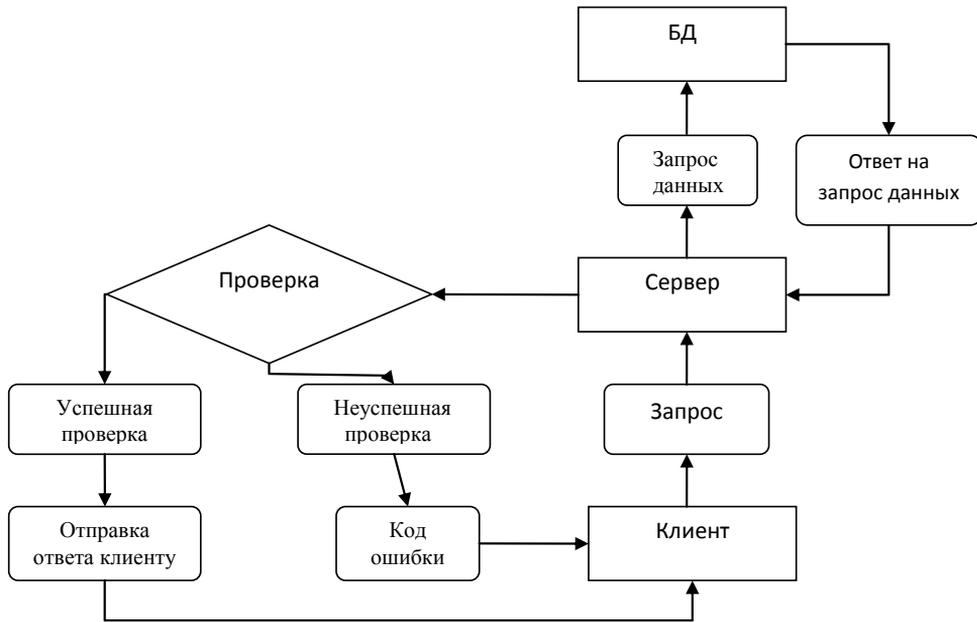


Рис. 5. Схема взаимодействия клиента и сервера

Каждый клиент для работы в КИС обладает логином и паролем.

Для входа в систему ему необходимо пройти авторизацию, после успешно пройденного процесса аутентификации, клиенту выдается уникальный идентификатор приложения и сеанса.

Клиент отправляет запрос на выполнение операции серверу, который обрабатывает его.

На этом этапе в работу подключается подсистема обеспечения ИБ:

1. Выполняется проверка подлинности клиентского приложения на основе криптографического алгоритма ГОСТ Р 34.11-2012 «Стрибог» на секретном ключе сервера.
2. Проверка клиента на основе атрибутов пользователя системы и криптографического алгоритма и сеанса на предмет блокирования.
Если клиент совершает несколько неудачных попыток, хост, с которого отправлялись запросы, заносится в черный список, блокируется.
3. Далее обработка передается в фильтр запросов, который на основе кода операции и атрибутов пользователя проверяет полномочия пользователя на выполнение.
4. Следующим этапом выполняется проверка запрос на предмет угрозы целостности и нарушения конфиденциальности информации.

У каждой записи в системе есть метка, если в результате запроса окажется строка с недопустимой меткой для пользователя, проверка считается не пройденной.

5. После успешно пройденных проверок запрос выполняется и ответ отправляется клиенту.

АНАЛИЗ РЕШЕНИЯ ОТНОСИТЕЛЬНО УГРОЗ БЕЗОПАСНОСТИ.

Рассмотрим процесс аутентификации:

Для того, чтобы пройти аутентификацию, клиенту необходимо знать логин/пароль доступа к серверу приложений и логин/пароль доступа к ИС. Если у него их нет, войти в систему не удастся.

Если нарушитель решит заниматься подбором логина/пароля, ему также необходимо будет знать ключ шифрования трафика, так как на неверном ключе сервер просто «не поймет» запросы клиента, к тому же можно настроить систему реагирования на блокирование хоста, с которого поступает слишком много запросов на аутентификацию.

Пароль не хранится в чистом виде, он хранится только в виде свертки, нарушителю придется решить очень сложную задачу по восстановлению свертки *ГОСТ Р 34.11-2012 «Стрибог»*.

Нарушителю не удастся обойти аутентификацию, все его действия сведутся к проблеме получения секретного ключа шифрования трафика, а при наличии такового – к проблеме анализа криптографической хэш-функции.

В результате правильно реализованного алгоритма аутентификации сведения до минимума угроза, связанная с регистрацией клиента от имени другого пользователя.

Угроза выявления логина/пароля в данных условиях скорее решается организационными мерами. Но в рамках данной работы, даже имея доступ к компьютеру легального клиента, нарушитель сможет получить только служебную информацию по настройкам подключения.

Несанкционированный доступ к информации:

Угроза сведена к минимуму за счет использования мандатного и ролевого разграничения доступа. Клиент может выполнять запросы, пользоваться функционалом, отведенным только к его конкретной роли, иметь доступ к документам только своего уровня доступа. Это осуществляется с помощью алгоритмов проверки данных пользователя, возможности и правомерности запроса клиента, а также фильтра запросов.

Работа различных пользователей под одним логином:

Клиенты не смогут работать под одним логином, после аутентификации клиента ему предоставляется идентификатор открытого сеанса, если сервер получит запрос на открытие еще одного сеанса от имени данного клиента, ему будет отказа-

но, так как сеанс клиента уже открыт. Закрыть сеанс может либо сам легальный клиент, либо принудительно администратор.

Таким образом, данная угроза исключена.

Понижение уровня доступа информации:

Данная угроза решается с помощью ввода обработки, проверки всех новых документов в системе. Таким образом, если нарушитель скопирует документ с более высоким уровнем доступа и сохранит его, то новый документ унаследует все атрибуты владельца, а так как доступ нарушителя к информации более высокого уровня исключен, документ сохранится с теми же атрибутами. После чего документ будет проверен администратором, без чьего подтверждения документ не окажется доступным.

Модификация, искажение, удаление информации:

Все действия, связанные с каким либо изменением данных, не доступны обычному клиенту, они доступны только администратору. В случае возникших проблем администратор решит ее с помощью восстановления из резервной копии.

Несанкционированное выполнение запросов к БД:

Проблема решена с помощью проверки подлинности клиента, проверки прав на выполнение, фильтра запросов. А также введением ролевого разграничения.

Весь возможный функционал системы описан, и коды запросов хранятся в БД, если нарушитель решит подделать запрос, то помимо проблем схожих при аутентификации, нарушителю придется угадывать идентификатор сеанса.

Без идентификатора, то есть если клиент не прошел аутентификацию, запрос не будет выполняться.

Если клиент решит подменить код операции, имитировать нажатие на недоступную для него кнопку, то данный запрос не будет выполнен, так как он не пройдет проверки на возможность и право на выполнение от имени клиента.

Неправомерный доступ к служебной информации:

Проблема решена с помощью проверки подлинности клиента, проверки прав на выполнение, фильтра запросов. А также введением ролевого разграничения.

Выполнение запросов, заведомо приносящих ущерб БД:

Проблема решена с помощью проверки подлинности клиента, проверки прав на выполнение, фильтра запросов. А также введением ролевого разграничения.

Подмена клиента:

Проблема решена с помощью использования специальных идентификаторов приложения и потоков между клиентом и сервером, сервер хранит идентификаторы клиента, его адрес. Поэтому получив запрос от клиента, сервер отправит ответ именно тому приложению, от которого пришел запрос.

Остальные угрозы решаются с помощью организационных мер, а также их решения входят частично в описанные выше решения.

Таким образом, трехзвенная архитектура является решением, отвечающим всем требованиям с точки зрения функционала и безопасности, можно разрабатывать полноценную защищенную систему, учитывая дальнейшие замечания.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТАНДАРТНЫХ РЕШЕНИЙ И ПРЕДЛАГАЕМОГО.

Аутентификация.

Наиболее часто применяемые протоколы – Kerberos, RADIUS и NTS, так называемые протоколы AAA. (*Authentication, Authorization, Accounting – аутентификация, авторизация, управление [Кабанов и др., 2012]*).

Выдается сеансовый ключ, протоколы основаны на криптографических алгоритмах

AES, MD5 (Message-Digest algorithm 5), SHA-2.

AES – подозрение на алгоритмические закладки, за пределами США во всех протоколах. Основанных на данном алгоритме ключ 56 бит.

MD5 – алгоритм выработки хэш-значения был разработан в 1991 году, как более надёжный вариант предыдущего алгоритма MD4. В 1993 году Берт ден Бур – первые сообщения о криптоанализе.

В конце 2008 года US-CERT призвал разработчиков программного обеспечения, владельцев веб-сайтов и пользователей прекратить использовать MD5 в любых целях, так как исследования продемонстрировали ненадёжность этого алгоритма.

Данные алгоритмы широко применяются в коммерческих продуктах в то время, как вышли более надежные алгоритмы, например, алгоритм выработки хэш-значения SHA-3.

Алгоритм построения коллизий представлен в работе [Ekerå, 2009].

Предлагаются стандарты ГОСТ 28147-89, ГОСТ 34.11, об успешных попытках криптоанализа неизвестно.

2. Контроль целостности.

Контроль целостности в КИС всегда основан на выработке и проверке хэш-значения, поэтому для данного пункта справедливы замечания из предыдущего пункта.

3. Перехват аутентификационных данных.

Для того, чтобы перехватить и получить открытые данные было максимально сложно, необходимо шифровать трафик. По поводу алгоритма блочного шифрования было сказано в п.1. Рекомендуется использовать отечественный алгоритм ГОСТ 28147-89.

4. SQL-инъекции.

К данному типу угроз потенциально уязвима любая КИС. Как устроена защита от данной угрозы в коммерческих системах – коммерческая тайна.

Пример:

Поиск в системе документа по имени.

```
Select * from DOCUMENTS where NAME='График дежурств' and LEVEL= LOW
```

Запрос возвращает документ с низким уровнем доступа.

```
Select * from DOCUMENTS where NAME=" union select * from DOCUMENTS — and LEVEL= LOW
```

Решение:

Обнаруживать аномальные SQL-выражения.

Как?

- Грамматический анализ или регулярные выражения
- Белые списки или Черные списки

5. Модификация, искажение, удаление информации

Подсистема модерации не позволит удалить или изменить документ без подтверждения, не допустит целенаправленного удаления и искажения и позволит уменьшить потери при ошибках пользователей.

6. Исключается нарушение чтение вверх, запись вниз.

Пользователь не прочитает документы с высокой меткой доступа – не позволит система разграничения доступа, не сможет скопировать документ и сохранить его с низким уровнем – не позволит система модерации.

7. Реакция системы защиты на события

Настройка правил реакции позволит реагировать на те или иные события, например:

Признать подозрительным вход в систему в необычно позднее время, вход в систему под пользователем из другого региона, реагировать на многократный ошибочный ввод пароля.

Соответствующая реакция: логирование, блокировка хоста, введение таймаута для ввода идентификаторов.

8. Надежность и доверенность построенной системы.

Система может быть полностью проанализирована, определить реакцию системы на события, можно протестировать на все уязвимости, выполнить доработку.

Проанализировать коммерческий продукт в полной мере не представляется возможным.

ЗАКЛЮЧЕНИЕ

Результатом исследования является метод построения доверенных систем, отвечающих требованиям к информационной безопасности. Данный метод основан на использовании в трехуровневой модели доверенного клиента и сервера-посредника. КИС, в которой будут реализованы механизмы, перечисленные

в требованиях к системе, будет надежной. При этом качество и уровень защищенности системы напрямую зависит от того, насколько качественно и полно эти механизмы применяются, насколько ответственно относится к вопросам защиты персонал, разрабатывающий, администрирующий и эксплуатирующий систему.

Предложенное решение является моделью для построения информационных систем, к которым предъявляются требования к информационной безопасности. Описанная подсистема безопасности отвечает предъявляемым к системе требованиям. Она позволяет избавиться от известных уязвимостей таких, как SQL инъекции, к которым заведомо уязвима любая СУБД.

За счет использования шифрования передаваемых данных стойким алгоритмом шифрования становится невозможным считывание трафика из сети. Использование криптографической хэш-функции позволяет обеспечить надежную проверку и контроль целостности данных, а также атрибутов пользователей.

За счет использования подсистемы аудита можно установить пользователя, выполнявшего запрос, тип, время, что позволит быстрее разобраться в инцидентах, связанных с функционированием системы, а также на основе журналов аудита возможна настройка реакции системы на события.

Использование подсистем ограничения доступа позволяет обеспечить конфиденциальность данных КИС.

Подсистема администрирования позволит обеспечить гибкое управление пользователями КИС.

Фильтрация и проверка запросов не позволит выполнять запросы, заведомо нарушающие конфиденциальность и целостность информации.

Трехзвенная архитектура позволяет обеспечить масштабируемость, возможность балансировки нагрузки, прозрачность для пользователя.

Решение, способствует повышению уровня защищенности системы. Также стоит отметить, что данный класс решений может быть реализован на различных СУБД, ОС и т.д., т.е. модель применима для целого класса программ, имеющих трехзвенную архитектуру.

ЛИТЕРАТУРА

- [Кабанов и др., 2012] *Кабанов А.С., Лось А.Б., Першаков А.С.*, Теоретические основы компьютерной безопасности, М: РИО МИЭМ, 2-е изд., 2012 г., с. 16-30.
- [Проскурин, 2009] *Проскурин В.Г.*, Защита программ и данных, ИД «Академия», 2-е изд., 2012 г., с. 44-55.
- [DBIR, 2015] 2015 Data Breach Investigations Reports [Электронный ресурс] , Verizon, № 1, 2015 г., с.31-55. URL: <http://www.verizonenterprise.com/DBIR/2015/>
- [Ekerå, 2009] *М. Ekerå* Differential Cryptanalysis of MD5 [Электронный ресурс], Master of Science Thesis Stockholm, Sweden, 2009 г., с. 65-69. URL: https://www.nada.kth.se/utbildning/grukth/exjobb/rapportlister/2009/rapporter09/ekera_martin_09008.pdf.

SYNTHESIS ALGORITHMS OF THE DEVELOPMENT TRUSTED INFORMATION SYSTEMS WITH A 3-TIER ARCHITECTURE

A.V.Belov, D.V.Pashkov

Moscow Institute of Electronics and Mathematics

National Research University Higher School of Economics (MIEM HSE)

The work is devoted to building trusted information systems with 3-tier architecture, which satisfy high requirements of information security, reviewed main information security vulnerabilities and information system threatening, offered algorithms for constructing that systems, grounded selection of 3-tier architecture to construct trusted system. Those algorithms allow to operate with rate category information.

Keywords: CIS, CIS security, trusted systems, database security, information system threatening, 3-tier architecture, trusted systems developing.

Belov Alexander Vladimirovich –Head of Department of Applied Mathematics Moscow Institute of Electronics and Mathematics in National Research University Higher School of Economics (MIEM NRU HSE), AVBelov@hse.ru

Pashkov Denis Vitalyevich– Postgraduate student of Moscow Institute of Electronics and Mathematics in National Research University Higher school of Economics (MIEM NRU HSE), denis__pashkov@mail.ru

СУПЕРКОМПЬЮТИНГ БЕЗ СУПЕРКОМПЬЮТЕРОВ: ЧТО МЫ МОЖЕМ, А ЧТО НЕ МОЖЕМ?

А.В. Богданов, А.Б. Дегтярев

Санкт-Петербургский государственный университет (СПбГУ)

В статье поднимаются проблемы решения комплексных задач, требующих использования суперкомпьютерной техники, на многопроцессорном кластерном ресурсе, доступном в настоящее время большинству исследователей. Показано, что без серьезного предварительного математического исследования эффективная реализация комплексной задачи на кластерной технике невозможна. Предложены пути решения подобных задач.

Ключевые слова: кластер, виртуальный суперкомпьютер.

ВВЕДЕНИЕ

Понятие суперкомпьютера, как вычислительного инструмента, стало нарицательным после появления списка задач так называемого «большого вызова» (“Grand challenges”). Это произошло в конце 1980-х годов в целях обозначения необходимости финансирования исследований в области высокопроизводительных вычислений и коммуникаций. Как известно, эти годы оказались переломными в области IT в силу появления и завоевания рынка первыми персональными компьютерами, глобальной сетью и «всемирной паутиной» (WWW). Встал вопрос о целесообразности развития вычислительной техники в том ключе, как это было предыдущие 30 лет. Был выделен круг фундаментальных и прикладных проблем, жизненно важных для развития человечества, эффективное решение которых возможно только с использованием сверхмощных вычислительных ресурсов – суперкомпьютеров (в отличие от персональных компьютеров и рабочих станций).

С тех пор конкретный список задач много раз видоизменялся и конкретизировался, а сам термин “Grand challenges” стал нарицательным. В список входят комплексные задачи физики, нанотехнологий, аэронавтики, биологии, национальной безопасности, науки о Земле, энергетики, окружающей среды и т.д.

Если постараться понять, какие вычислительные средства необходимы для решения задач «большого вызова», то очень обобщенно можно заключить, что имеем дело с тремя принципиальными конфигурациями, которые направлены на решение [Богданов и др., 2015]

1. задач обработки больших и сверхбольших объемов данных;
2. большого количества слабо связанных задач;
3. одной сильно связанной задачи, требующей большого объема памяти и производительности вычислительных мощностей.

Богданов Александр Владимирович – профессор кафедры компьютерного моделирования и многопроцессорных систем, факультет Прикладной математики – процессов управления, д-р физ.-мат. наук, Санкт-Петербургский государственный университет (СПбГУ), a.v.bogdanov@spbu.ru

Дегтярев Александр Борисович – профессор кафедры компьютерного моделирования и многопроцессорных систем, факультет Прикладной математики – процессов управления, д-р техн. наук, Санкт-Петербургский государственный университет (СПбГУ), a.degyarev@spbu.ru

В любом случае, особенно в третьем варианте, такие задачи можно рассматривать как комплексные, т.е. (а) в рассматриваемой задаче присутствуют процессы принципиально разных масштабов изменчивости; (б) характер процессов оказывается существенно нелинейным.

Расчет комплексных задач, особенно нелинейных, всегда был вызовом для вычислительных наук. Этот вызов всегда приводил к замечательным взлетам и очень обидным падениям. Причем падения, на наш взгляд, тесно связаны с деградацией архитектур вычислительных систем, которые для сохранения эффективности оказываются многоуровневыми с очень сложной коммуникацией потоков данных. В настоящее время появление многопоточных и графических процессоров привело к появлению проблем того же типа.

Нам представляется, что решение этих проблем должно быть комплексным и состоять

- а) из подготовки специализированных вычислительных комплексов;
- б) разработки алгоритмов, учитывающих специфику современных вычислительных комплексов;
- в) применения методов более эффективного управления потоками данных в процессе вычислений.

Что бы понять возникающие проблемы, рассмотрим один из типичных подходов численной реализации комплексной задачи. В качестве примера можно рассмотреть:

- метод сеток,
- решение комплексной задачи методом итераций,
- применение метода разложения (представление результата в виде ряда по известным решениям), и т.п.

Все эти методы приводят задачу к матричной системе большой размерности (до миллионов строк и столбцов). При параллельном расчете потоков в этой матрице, при большой разнице масштабов процессов, возникают неоправданные затраты вычислительного ресурса при расчете медленных процессов. Если отношение масштабов не зависит от параметров, то можно существенно упростить задачу, используя асимптотические методы. Однако в большинстве случаев это не так. К сожалению, самые интересные эффекты возникают там, где характерные масштабы описываемых процессов сравниваются. С точки зрения теории динамических систем [Арнольд, 1971] появление больших недиагональных членов матрицы связано с неправильным выбором базиса. Однако выбор правильного базиса, особенно для нелинейных систем, представляет собой очень сложную задачу. Существует элегантный подход, связанный с использованием нелинейных интегральных преобразований [Сумецкий, Фельштын, 1990], которые для очень многих случаев позволяют диагонализировать соответствующие матрицы. Центральная проблема этого подхода состоит в том, что для таких преобразований, как правило, не существует обратных преобразований. В тех случаях, когда их удается найти, мы получаем известные результаты. В качестве рабочей гипотезы мы можем высказать предположение, что обратное преобразование существует только для вполне интегрируемых систем [Bogdanov, 1986]. Тем более,

очень большое значение имеют случаи, когда требуемый результат получается без применения обратного преобразования (например, когда мы считаем средние значения каких-то величин или интегральные характеристики используются непосредственно для решения других задач).

Однако эту идею можно применить следующим образом. В качестве ядра разложения использовать вполне интегрируемую систему, а результат ее квази-диагонализации подставить в остаток.

Эта общая идея для своей реализации требует нескольких шагов, которые, в том числе, используют особенности современных вычислительных устройств. Анализ этого вопроса посвящена настоящая статья.

1. ПРОБЛЕМЫ СОВРЕМЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ ПЛАТФОРМ

Большинство современных расчетов проводится на кластерных системах, т.е. на системах с высокопроизводительными процессорами, соединенных относительно медленными коммутаторами и с использованием передачи сообщений в качестве основного способа межпроцессного взаимодействия. В этом случае даже в одноранговой системе эта схема приводит к большим задержкам в передаче данных и проблемам при масштабировании.

Ускорение кластерной системы может быть представлено следующим образом [Degtyarev, 2003; Gankevich et al., 2015]:

$$S_N = \frac{T_0}{T_N} = \frac{N}{1 - \alpha + \alpha N + \beta \gamma N^3} \quad (1)$$

где N – количество вычислительных узлов,

α – доля последовательных вычислений,

β – коэффициент, учитывающий диаметр системы,

γ – коэффициент, представляющий отношение производительности вычислительного узла (процессора) к производительности коммутатора (скорость линка).

Вид зависимости (1) приведен на рис. 1.

Отличительной чертой этой зависимости является наличие для конкретной задачи оптимального числа процессоров, зависящего преимущественно от параметра γ . При наличии явных накладных расходов с ростом количества используемых процессоров ускорение не растет, а спадает до нуля после достижения этого оптимального числа процессоров.

Практически это приводит к тому, что для задач с большой передачей данных на каждом такте (например, инженерные приложения) масштабируемость ограничивается десятком, редко двумя десятками вычислительных узлов.

В действительности ситуация обстоит еще хуже. В современных процессорах существует многоуровневая система памяти с большим количеством ядер и внутренней коммуникацией между ними. При этом пользователь чаще всего не может управлять отдельными ядрами и осуществлять между ними балансировку

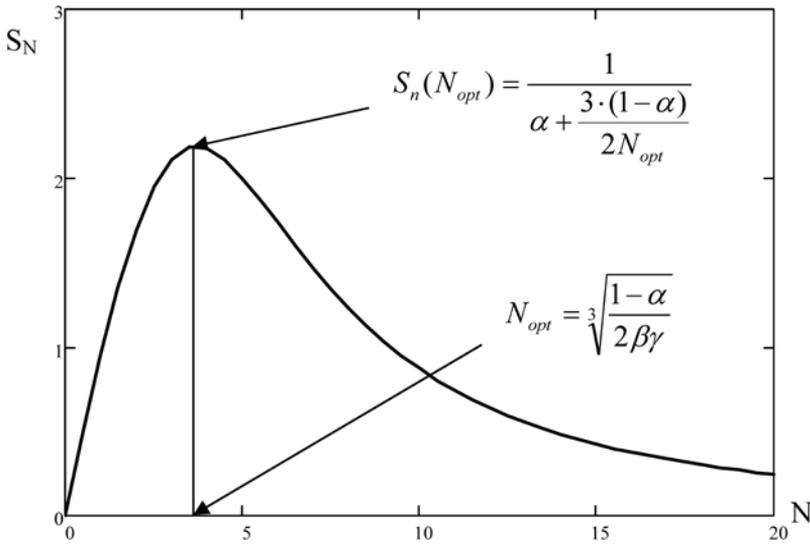


Рис. 1. Ускорение кластерной системы

нагрузки. Это приводит еще к более сложной формуле, чем (1) и более резкому снижению масштабируемости [Шошмина, Богданов, 2007]. Поэтому Linux-кластеры нельзя рассматривать в качестве машин общего назначения.

Существуют, однако, очень показательные примеры, когда на базе Linux-кластера делали очень эффективные специализированные системы. Связано это с адаптацией архитектуры кластера под вычислительный алгоритм и оптимизацией вычислительных процессов на всех этапах. К сожалению, такую адаптацию/перестройку невозможно производить каждый раз, когда возникает новая задача. Современные вычислительные комплексы очень трудно поддаются такой перестройке, и дело это столь дорогое, что в практических приложениях не может быть использовано. Связано это со знаменитой проблемой отображения алгоритма на вычислительную архитектуру. Каждый вычислительный алгоритм имеет свои особенности, связанные с характером обменов между вычислительными процессами. Подготовить вычислительную архитектуру, которая годилась бы на все случаи жизни, невозможно.

История вычислительной техники, однако, показывает, что все недостатки вычислительных систем во все времена компенсировались эффективным программным обеспечением. Нам представляется, что на современном этапе при помощи технологий (а) метакомпьютинга; (б) сервис ориентированной архитектуры и (в) виртуализации вполне можно получить подобное решение.

Основная идея нашего подхода состоит в виртуализации всех компонентов вычислительной системы, создания виртуального вычислительного комплекса, на который наилучшим способом отображается вычислительный алгоритм. Использование виртуальной общей памяти лежит в основе реализации парадигмы работы с разделяемой памятью. Это позволяет обойти указанную проблему ухудшения масштабируемости вычислительной системы.

2. РЕАЛИЗАЦИЯ ВИРТУАЛЬНОГО СУПЕРКОМПЬЮТЕРА

В качестве основных принципов реализации виртуального суперкомпьютера отметим следующие положения [Bogdanov, 2012].

1. Облако полностью определяется его API. Это очевидно с точки зрения пользователя, но то же самое верно и с точки зрения взаимодействия различных видов облаков (см. п. 2.).

2. Операционная среда должна быть типа UNIX. Одной из главных проблем вычислительного GRIDа является балансировка нагрузки. Это очень трудная задача, так как пользователь отрезан от ресурсов. Частично эта проблема решается с помощью организации проблемно ориентированной среды (PSE). Тем не менее, для успешной реализации такого подхода многие стандартные инструменты UNIX должны быть введены в API.

3. Облако использует протоколы, совместимые с популярными общественными облаками. Общественные облака не слишком приспособлены для решения сложных проблем и причина этого ясна – более сложная проблема, которую вы решаете, требует более надежных инструментов в использовании. Универсальные средства не могут быть использованы для сложных проблем. Вот почему для их решения должны быть построены специализированные частные облака. Однако если ресурсов частного облака оказывается не достаточно, то дополнительные ресурсы могут быть получены из публичного облака.

4. Облако обрабатывает данные на основе распределенных файловых систем. Основная проблема в публичном облаке для обработки данных происходит по той причине, что на каждом компьютере в таком облаке используется своя файловая система. Это препятствует как обработке больших объемов данных, так и масштабированию при решении задач. Чтобы преодолеть это, распределенная файловая система должна использоваться в частном облаке, тип которого определяется природой решаемой задачи. Если добавить сюда три способа, обеспечивающие согласованность данных (теорема Брюера [Brewer, 2010]), мы можем увидеть, что есть много возможностей организации обработки данных, из которых используются лишь немногие.

5. Консолидация данных достигается путем распределенной федеративной БД. Есть три уровня консолидации – сервера, данные и ресурсы. Более или менее понятно, как произвести консолидацию серверов. Консолидация данных более сложный процесс, а консолидация ресурсов является реальной проблемой для поставщика облачных услуг. Мы предполагаем, что самый естественный путь решения этой проблемы является использование инструментария федеративных БД. До сих пор нам удалось сделать это, используя инструменты IBM DB2, однако мы считаем, что и возможности последних версий PostgreSQL позволяют выработать бесплатный инструментарий для поставленной цели.

7. Обработка больших объемов данных осуществляется с помощью общей виртуальной памяти. На самом деле весь предыдущий опыт показывает, что единственным способом комфортной обработки больших объемов данных является использование SMP системы. Теперь мы можем эффективно использовать инструменты общей памяти (например, Open CL) в гетерогенной среде и таким образом построить виртуальную SMP систему. Тот же инструмент используется

для распараллеливания. Необходимо отметить, что очень эффективными являются возможности использования единого образа операционной среды.

8. Облако использует комплексный grid, как механизм безопасности. Одна из проблем облаков – это вопросы безопасности, однако можно надеяться, что правильное сочетание инструментов безопасности GRID с технологиями доступа Cloud даст эффективное решение проблемы.

3. ТЕСТИРОВАНИЕ ВИРТУАЛЬНОГО РАСПРЕДЕЛЕННОГО ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА

При создании виртуальных кластеров, ориентированных на решение крупномасштабных задач, приемлемыми для использования оказываются только легкие технологии виртуализации. Причина кроется в том, что при масштабировании задачи на большое число узлов даже малая доля накладных расходов службы виртуализации не является приемлемой (это следует непосредственно из формулы (1)). В случае виртуальных кластеров масштабируемые накладные расходы происходят от виртуализации процессоров. Это означает, что никакие средства пара- и полной виртуализации не подходят для больших виртуальных кластеров. В этом случае в качестве объекта для дальнейшего исследования и применения можно рассматривать технологии контейнеров. Другой проблемой является динамическое создание и удаление виртуальных кластеров.

Для проведения экспериментов нами была создана система, которая включает в себя множество стандартных компонентов, являющихся общими в области высокопроизводительных вычислений. Среди этих компонент необходимо отметить распределенную параллельную файловую систему, которая сохраняет домашние каталоги с входными и выходными данными эксперимента, планировщик ресурсов кластера, который распределяет ресурсы для рабочих мест и клиентских программ для пре- и пост- обработки данных и т.д. Для обеспечения контейнеризации использовалась технология LXC (Linux Container technology), GlusterFS использовалась для обеспечения параллельной файловой системы, а система TORQUE для обеспечения планирования задач. Подробные характеристики аппаратного и программного обеспечения, номера версий перечислены в таблице 1

Таблица 1. Аппаратные и программные компоненты системы

Компонент	Описание	Компонент	Описание
CPU model	Intel Xeon E5440	Операционная система	CentOS 7
CPU clock rate (GHz)	2.83	Версия ядра	3.10
Кол. ядер на CPU	4	LXC версия	1.0.5
Кол. CPUs на узел	2	GlusterFS версия	3.5.1
RAM (GB)	4	TORQUE версия	5.0.0
Disk model	ST3250310NS	OpenMPI версия	1.6.4
Disk speed (rpm)	7200	IMB версия	4.0
Количество узлов	12	OpenFOAM версия	2.3.0
Скорость коммутатора (Gbps)	1		

Создание виртуального кластера в такой среде требует следующих шагов. Во-первых, клиент передает задание, требующее определенное количество ядер. Тогда, в соответствии с распределением этих ядер среди вычислительных узлов, на каждом узле из списка запускается контейнер, внутри которого в качестве единственной программы работает управляющий демон SSH. Здесь есть два варианта: либо запустить контейнеры с виртуализацией сети (используя macvlan или сеть типа LXC) и сгенерировать достаточное количество IP-адресов для кластера или использовать пространство сетевых имен для хоста (с сетью не типа LXC) и генерировать только номер порта для запуска SSH демона. Следующим шагом является копирование (возможно измененного) файла узла с хоста в первый контейнер, и запуск скрипта уже внутри него. При завершении работы демона SSH в каждом контейнере все контейнеры будут уничтожены.

Для корректной работы этого алгоритма, прежде чем запускать скрипт, внутри контейнера должна быть смонтирована клиентская домашняя директория. Кроме того, так как некоторые программы MPI для корректной работы требуют временных каталогов на каждом узле, корневая файловая система контейнера должна быть смонтирована в режиме копирования при записи, так что все изменения в файлах и все новые файлы записываются во временный каталог хоста. Все файлы, которые не менялись, используются только для чтения из файловой системы, смонтированной на виртуальной сети; это можно реализовать на базе Union или аналогичной файловой системы, так что задачи, работающие на кластере, вообще не обращаются к контейнерам.

Подводя итог, отметим, что только стандартные инструменты Linux используются для построения системы, в этой конфигурации нет непрозрачных образов виртуальных машин, нет сложных технологий полной виртуализации и нет тяжелых стеков облачных вычислений.

Чтобы оценить эффективность построенного вычислительного комплекса были использованы тесты OpenMPI и Intel MPI Benchmarks (IMB) для измерения пропускной способности сети, а пакет OpenFOAM был использован для измерения общей производительности на реальных приложениях.

Первый эксперимент был связан с созданием виртуального кластера, запуска на нем пустой (с /bin/true в виде исполняемого файла) программы MPI и сравнением времени исполнения с работой обычного физического кластера. В этом эксперименте в контейнере была установлена та же операционная система и версия OpenMPI, как и в хост-машине. При этом виртуализация сети не использовалась, каждый прогон повторялся несколько раз, и средние результаты показаны на графике (рис. 2).

Результаты показывают, что к каждому LXC прогону после 8 ядра добавляются постоянные накладные расходы в 1,5 секунды. Одну секунду можно отнести за счет отсутствия кэша внутри контейнера с файлами конфигурации SSH, ключевых файлов и библиотек в нем, а другую половину секунды следует списать на создание контейнеров, как показано на рис. 3. Скачек после 8 ядра отмечает границы одной машины, что подразумевает использование сети, а не общей памяти, для передачи данных.

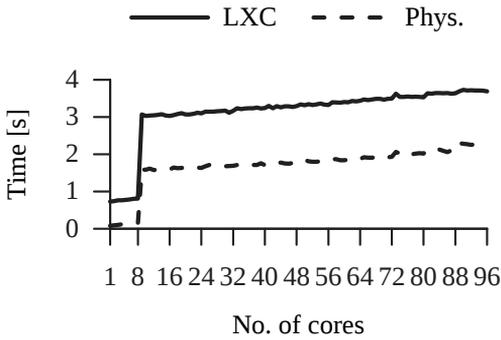


Рис. 2. Сравнение производительности LXC и физического кластера, запускаемого под MPI

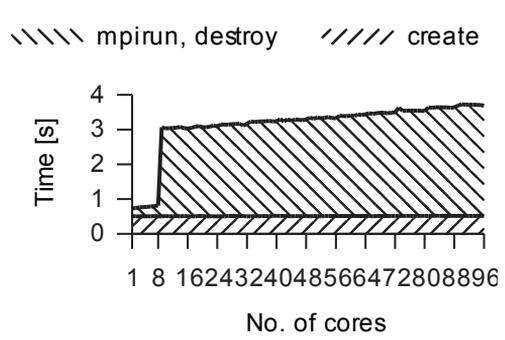


Рис. 3. Обвал виртуальной сети (LXC)

Создание контейнеров – это полностью параллельный процесс, и он занимает примерно одно и то же время независимо от числа узлов. Накладные расходы по уничтожению контейнеров оказались незначительными и были включены `mpirun`. Таким образом, использование контейнеров Linux добавляет некоторые постоянные накладные расходы на запуск параллельной задачи в зависимости от конфигурации системы, которые делятся между созданием контейнеров и резервным копированием.

Второй эксперимент состоял в измерении, с использованием набора ИМВ, характеристик различных типов LXC сетей. Было обнаружено, что выбор типа виртуализации сети существенно влияет на производительность. Как и в предыдущем тесте, контейнер был создан с той же операционной системой и тем же набором ИМВ, что исполнялись и на хост-машине.

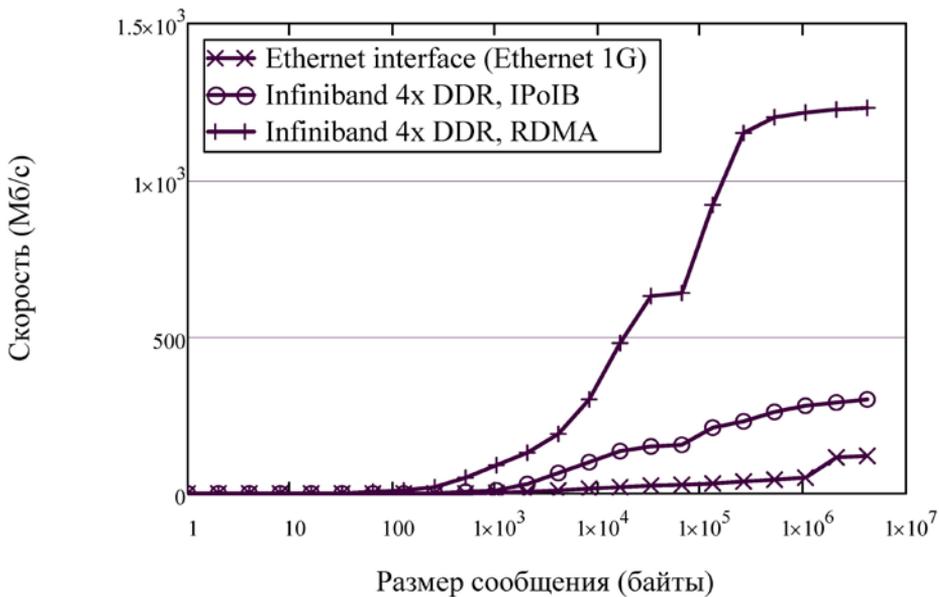


Рис.4. Тестирование производительности сети 1 (кластер 1)

Пропускная способность сети была определена с помощью теста транспозиций и отображается на графике (рис. 4). Из графика видно, что, пока размер сообщения меньше 214 байт, производительность является приблизительно одинаковой для всех типов сети, однако, при больших объемах данных существует провал при реализации виртуального Ethernet. Трудно судить, почему это происходит: некоторые исследования [James, 2004] показывают, что при высокой нагрузке падает эффективность многосвязных сетей (а виртуализация всегда связана с многосвязностью), но IMB не приводит к большой нагрузке на систему. Кроме того, эксперимент показал, что, как и ожидалось, пропускная способность уменьшается с увеличением числа ядер из-за накладных расходов на синхронизацию (рис. 5).

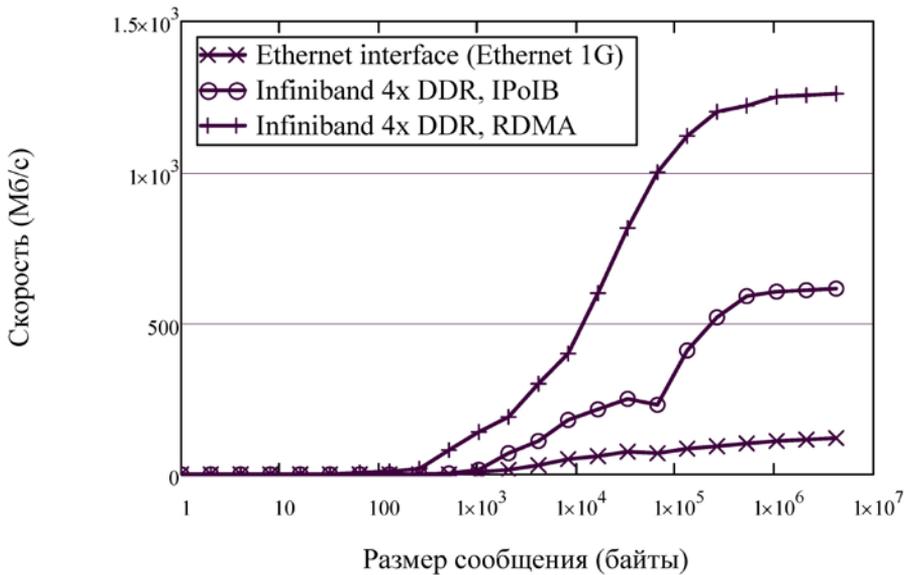


Рис. 5. Тестирование производительности сети 2 (гибридный кластер)

Третий и последний эксперимент призван изучить производительность реальных приложений. Для этой роли был выбран OpenFOAM в качестве комплексной параллельной задачи с участием большого количества сетевых коммуникаций, ввода/вывода с диска и высокой загрузки процессора, а также классическая задача перемножения матриц (результаты тестирования приведены на рис. 8). Случай прорыва плотины (модуль RAS) был запущен с разным количеством ядер (общее количество ядер определяется как квадрат числа ядер на узле) и LXC сетей различных типов. Средние по нескольким прогонам представлены на графике (рис. 6). Измерения для 4 и 9 ядер были отброшены, потому что наблюдались значительные изменения времени выполнения для этих чисел на физических машинах. Из графика видно, что низкая производительность виртуальной сети Ethernet снизила реальную производительность OpenFOAM примерно на 5–10%, тогда как результаты с macvlan и без виртуализации близки к произ-

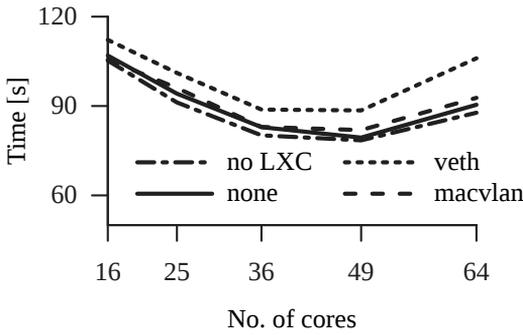


Рис. 6. Усредненная производительность работы OpenFOAM с различными сетями LXC

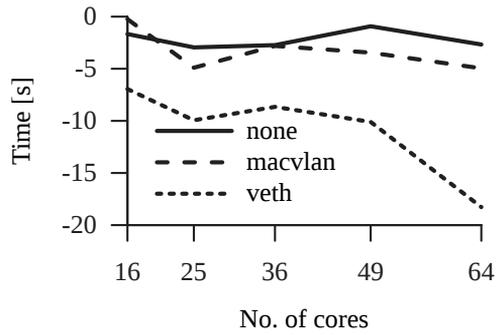


Рис. 7. Разница в производительности работы OpenFOAM на физическом и виртуальном кластерах. Отрицательные числа показывают замедление виртуального кластера

водительности физического кластера (рис. 7). Таким образом, выбор типа сети является основным фактором, влияющим на производительность параллельных приложений, работающих на виртуальных кластерах, и их накладные расходы могут быть снижены с помощью сетей типа macvlan сети или путем отказа от виртуализации сети.

На рис. 9 приведено комплексное сравнение работы модуля пакета OpenFOAM на виртуальном и физическом кластерах.

Подводя итог, можно сказать, что есть два основных типа накладных расходов при использовании виртуальных кластеров. Накладные расходы на создание виртуального кластера, которые постоянны и малы по сравнению со средним временем типичного параллельного задания. Сетевые накладные расходы, которые могут быть устранены, не используя виртуализацию сети или применяя специальные протоколы виртуализации.

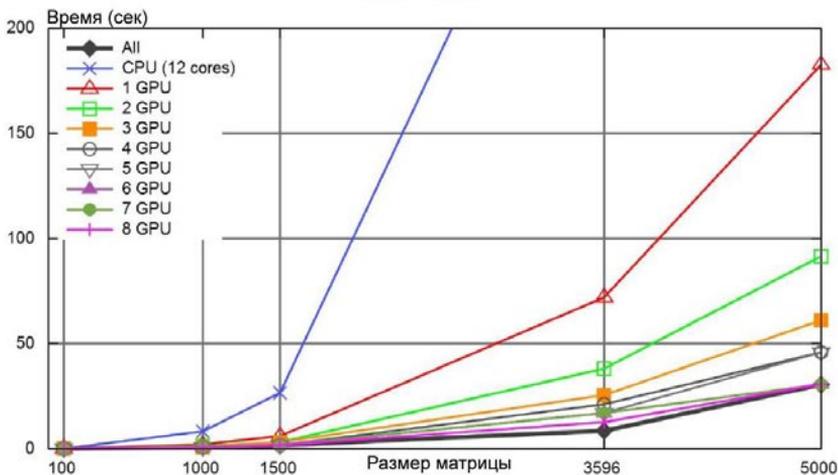


Рис. 8. Решение тестовой тригонометрической задачи на гибридном кластере в различных конфигурациях

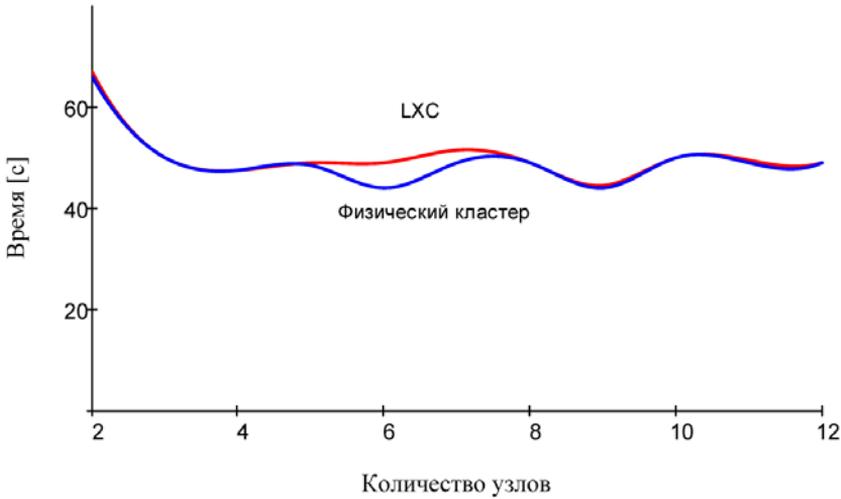


Рис. 9. Выполнение модуля OpenFOAM на виртуальном и физическом кластере

Представленный подход для создания виртуальных кластеров из контейнеров Linux оказался очень эффективным, и его производительность сравнима с производительностью обычного физического кластера. Использование контейнеров не приводит к существенным накладным расходам на виртуализацию процессора, но и сетевые накладные расходы виртуализации могут быть существенно уменьшены, если используется пространство сетевых имен для хост-компьютера и через сеть автоматически передаются только те файлы, которые необходимы сетевой файловой системе, а не целые образы. С точки зрения системного администратора, хранение каждого приложения НРС в собственном контейнере делает управление версиями и связями легко осуществляемым, а их конфигурация не мешает конфигурации хост-машины и других контейнеров.

Проведенное тестирование показало достаточную эффективность виртуального кластера, и в дальнейшем мы можем перейти к реализации программы, описанной во введении.

4. ПРИМЕРЫ РЕАЛИЗАЦИИ АЛГОРИТМОВ ДЛЯ ВИРТУАЛЬНЫХ КЛАСТЕРОВ

В последнее время широкое применение в математической физике получили методы функционального интегрирования. Функциональный интеграл является очень эффективным инструментом, поскольку в замкнутой и компактной форме позволяет записать решение дифференциального уравнения в частных производных. Он же может служить основой получения высокоэффективных параллельных алгоритмов.

Конечно, основная проблема, с которой приходится считаться, это отсутствие математически корректного обоснования функционального интегриро-

вания в общем случае. Для многих практически важных случаев, однако, можно либо получить такое обоснование, либо обосновать следствия, полученные в рамках такого подхода. Не претендуя на общий случай, укажем важную для приложений возможность. Так для уравнения в частных производных второго порядка для оператора эволюции легко получить фейнмановское представление [Bogdanov, 1986]

$$G(X, X') = \int DX \exp(i / h S(X, X')),$$

расходимость которого легко устанавливается анализом меры DX в таком интеграле. Еще Фейнман указывал на удобство для решения этой проблемы увеличения размерности и переходе от конфигурационного пространства X к фазовому пространству (X, P) . Оказывается, что и интеграл по фазовому пространству для пропагатора $G(X, X')$ расходится. Однако сходится интеграл в смешанном представлении $G(P, X)$.

Поскольку все объекты, интересные для приложений, являются квадратичными формами типа

$$A = \int dX \cdot G^+(P', X)G(X, P),$$

именно такой объект и может использоваться для реализации процедуры, описанной во введении.

Например, для уравнения Шредингера представление пропагатора имеет вид

$$G(X, P) = \int DPDX / (2\pi\hbar)^s \exp\left[i / h \int xdP - i / h \int Hd\alpha + i\phi\right]$$

В духе сформулированного подхода можно сделать каноническое (не меняющее меры интегрирования) преобразование переменных интегрирования $(P, X) \rightarrow (Q, Y)$, обращающее в нуль Гамильтониан H .

Остающийся квадратичный функциональный интеграл легко вычисляется и для пропагатора получается представление

$$G(P^* X) = \exp\left[-i / h \cdot P^* X^* + i / h \cdot F_1 |^* + i / h \cdot Q^*(Y_0 - Y^*)\right]$$

Где F_1 это генератор преобразования, являющийся решением уравнения первого порядка в частных производных, эквивалентного системе обыкновенных дифференциальных уравнений.

Таким образом, исходная задача сводится к обычному интегралу размерности, равной количеству степеней свободы, в каждой точке подынтегрального выражения которого решается система обыкновенных дифференциальных уравнений. Поскольку начальные данные для системы в каждой точке независимы, они могут решаться параллельно, и предлагаемый подход дает нам абсолютно параллельный алгоритм.

2. Существует замечательное применение идеи, изложенной во введении, в физической и химической кинетике. Дело в том, что уравнения кинетики с помощью нелинейного преобразования приводятся к квазидиогональной форме уравнения Рикатти. Поскольку представляющие интерес физические величины

могут быть также эффективно представлены в преобразованных координатах, как и в исходных заселенностях, проблема обратного преобразования здесь не стоит.

Система кинетических уравнений, описывающих релаксацию в молекулярных системах, состоит в реальных случаях из тысяч уравнений типа

$$Dc(\mathbf{j},t)/Dt = I(\mathbf{j}-\mathbf{1},t) - I(\mathbf{j},t), \quad (2)$$

где \mathbf{j} мульти-индекс, описывающий молекулярные состояния, и I молекулярный поток в индексном пространстве формы $I(\mathbf{j},t) = K(\mathbf{j},\mathbf{j}+\mathbf{1})c(\mathbf{j},t) - K(\mathbf{j}+\mathbf{1},\mathbf{j})c(\mathbf{j}+\mathbf{1},t)$, где K – константа скорости молекул.

Проблемы уравнения (2), на самом деле, происходят из двух факторов. Во-первых, имеются большие недиагональные члены, соответствующие важным физическим переходам. Во-вторых, значения $I(\mathbf{j},t)$ очень велики, а выражение в правой части уравнения (2) является малой разностью близких величин. Ситуация становится драматичной, если мы начинаем интегрирование из состояния теплового равновесия, когда все I равны.

Чтобы преодолеть эти трудности, полезно ввести новые переменные $f(\mathbf{j},t) = c(\mathbf{j}+\mathbf{1},t)/c(\mathbf{j},t)a(\mathbf{j},t)$, где $a(\mathbf{j},t)$ является отношением констант скорости $K(\mathbf{j},\mathbf{j}+\mathbf{1})$ и $K(\mathbf{j}+\mathbf{1},\mathbf{j})$. f является так называемой медленной переменной, которая становится постоянной в условиях равновесия. В этом случае уравнение для f принимает вид

$$df(\mathbf{j},t)/dt = \tilde{R}(f(\mathbf{j},t)) + H(\mathbf{j},t)f(\mathbf{j},t) + S(\mathbf{j})f(\mathbf{j},t) \quad (3)$$

где \tilde{R} релаксационный член, который диагонален по \mathbf{j} , H – член, пропорциональный гидродинамическому градиенту, а S является источником изменения заселенности. Это единственный член, который является недиагональным в $f(\mathbf{j},t)$. Основным преимуществом уравнения (3) является то, что не только сумма трех слагаемых в правой части (3) мала, но все они малы по отдельности. При этом их относительные величины легко определить заранее. Кроме того, важно, что основной вклад в правой части (4) вносят диагональные элементы, что открывает интересные возможности для параллельного алгоритма.

3. Очень важный пример нам дает теория нелинейных уравнений. Задачи эти хорошо изучены и многие классы вполне интегрируемых систем хорошо описаны [Уизем, 1977]. На практике, однако, приходится иметь дело с системами, которые не являются вполне интегрируемыми. В духе подхода, изложенного во введении, мы будем рассматривать их, как возмущения вполне интегрируемых систем. Одна из реализаций этого подхода может быть связана с разложением искомого решения по решениям вполне интегрируемых уравнений. Тогда вся нелинейность уходит именно в эти решения, а для невязки мы получаем простые дифференциальные уравнения, как в обычной теории возмущений. В качестве реализации этой идеи укажем пример обобщенного уравнения КДВБ.

$$v_t + vv_x + \alpha v_{xx} + \beta v_{xxx} = \gamma \mathcal{I}(v) \quad (4)$$

где α является мерой эффектов диссипации, β является мерой дисперсии и выражается через транспортные коэффициенты и время релаксации, γ является мерой

межфазного взаимодействия и выражается с помощью интегральных скобок и времени релаксации, а $I(v)$ – интегральный оператор, который в первом приближении может быть принят линейным и выражен следующим образом

$$I(v) = -\int_0^t G(t, \tau) v_\tau d\tau \quad (5)$$

где G для различных моделей взаимодействия представляется в экспоненциальном или обратном степенном виде.

Применяя разложения по решениям уравнения КДВ, получаем

$$\frac{da}{d\tau} = -\Phi(\kappa)a$$

где

$$\Phi(\kappa) = 2 \int_0^\infty \exp(-\kappa\theta') \int_0^\infty \cosh^{-2}(\theta + \gamma\theta') \cosh^{-3}\theta \sinh\theta d\theta' d\theta$$

Помимо упрощения процедуры вычислений анализ этого уравнения позволяет делать и качественные выводы, например, предсказывать появление предвестника.

5. ПРИМЕРЫ ЗАДАЧ, РЕШАЕМЫХ В РАМКАХ РАЗВИВАЕМОГО ПОДХОДА

Укажем лишь некоторые важные задачи, которые были решены таким образом:

1. Численное моделирование классической траекторной задачи.

Квантовый подход строится на генерации трубок траекторий, которые позволяют учесть влияние классической неинтегрируемости динамической квантовой системы [Bogdanov et al., 2008]. Когда объем классического хаоса в фазовом пространстве не превышает размеров квантовой ячейки, в соответствующей квантовой системе генерируется квантовый хаос. Успех в численном исследовании в этом случае был достигнут только лишь благодаря предлагаемому подходу.

2. Прямой вычислительный эксперимент в гидроаэродинамике.

Прямые численные эксперименты в механике сплошных сред с использованием цифровых компьютеров основаны на ограниченном наборе числовых объектов, которые интерполируют параметры состояния физических полей во времени. Предлагаемый подход позволяет исключить из рассмотрения математические модели механики жидкости в виде дифференциальных уравнений в частных производных. Вычислительный эксперимент проводится на основе фундаментальных законов сохранения. Дуализм корпускулярных и континуальных моделей сплошной среды позволил представить процедуру вычислений в виде трех последовательных этапов, сочетающих подходы Эйлера и Лагранжа. Такое разделение направлено на обеспечение эффективной процедуры вычислений, особенно в условиях многопроцессорной вычислительной среды, поскольку позволяет использовать явные численные схемы [Богданов и др., 2015; Дегтярев, Храмушин, 2014].

3. Виртуальный полигон.

Под виртуальным полигоном понимается программно-аппаратный комплекс, обеспечивающий всестороннее моделирование взаимодействия динамических объектов друг с другом и окружающей средой. Цель такой системы перенести в виртуальную среду постановку сложных, дорогостоящих и зачастую опасных экспериментов. Поскольку создать в этом случае одну универсальную математическую модель, в равной мере описывающей все компоненты системы, невозможно, приходится согласовывать и гармонизировать разные модели, предназначенные для описания поведения отдельных компонентов. В данном случае возникают проблемы отображения каждой конкретной системы на распределенную вычислительную архитектуру [Bogdanov et al., 2005; Gankevich, Degtyarev, 2013].

4. Описание физико-химических процессов в реакторе PCVD.

Описанный в работе [Korkhov et al., 2008] виртуальный плазменно-химический реактор может рассматриваться в качестве примера виртуального полигона.

5. Моделирование психологически политического развития.

Традиционный подход использования математических методов в социальных науках чаще всего ограничивается применением статистических методов разной степени сложности и теории игр. Использование, в первую очередь, таких инструментов обусловлено комплексностью рассматриваемых проблем и их плохой формализованностью. В работе [Bogdanov et al., 2015] предпринята попытка распространить описанный подход к решению этой проблемы.

ЗАКЛЮЧЕНИЕ

Проведенный анализ показал, что на новых вычислительных ресурсах можно решать комплексные нелинейные задачи. Однако это требует больших усилий. Наиболее сложная часть состоит в подготовке виртуального распределенного вычислительного комплекса. К сожалению, пока эту задачу формализовать не удастся. Она требует высокой квалификации, как системных программистов, так и пользователей. Что касается разработки алгоритмов, то с появлением виртуальных кластеров эта задача, скорее всего, упрощается, и на первый план выходит именно математическая постановка задач. Нам представляется, что следует ожидать значительного прогресса в обоих этих направлениях.

ЛИТЕРАТУРА

- [Арнольд, 1971] *Арнольд В.И.* Обыкновенные дифференциальные уравнения, М.: Наука, 1971, 240 с.
- [Богданов и др., 2015] *Богданов А.В., Дегтярев А.Б., Храмушин В.Н.* Высокопроизводительные вычисления на гибридных системах: будут ли решены «задачи боль-

- шого вызова»? //Компьютерные исследования и моделирование. 2015. Т.7. № 3. С.429-438.
- [Дегтярев, Храмушин, 2014] *Дегтярев А.Б., Храмушин В.Н.* Проектирование и построение вычислительных экспериментов в гидромеханике с использованием явных численных схем и алгоритмов тензорной математики. //Математическое моделирование. 2014. Т.26. № 11. С.4-17.
- [Шошмина, Богданов, 2007] *Шошмина И.В., Богданов А.В.* Использование grid-технологий для проведения вычислений. //Вестн. С.-Петербур. ун-та. Сер. 4: Физика, химия. 2007. Вып. 3. С.130–137.
- [Сумецкий, Фельштын, 1990] *Сумецкий, М.Ю., Фельштын М.Л.* Резонансное туннелирование электронов через двух-и трехмерную наноструктуры. (препринт) – Л. : НТО АН СССР, 1990, 31 с.
- [Уизем, 1977] *Уизем Дж.* Линейные и нелинейные волны, М.:Мир,1977,622 с.
- [Bogdanov, 1986] *Bogdanov A.V.*, Calculation of the quantum-mechanical inelastic scattering amplitude via solution of the classical dynamical problems. *Sov. Phys. Tech. Phys.* 1986. 31 (7). P. 833-835.
- [Bogdanov et al., 2005] *Bogdanov A., Degtyarev A., Nechaev Yu.* Parallel Algorithms for Virtual Testbed. //Proc. of V Int. Conf. on Computer Science and Information Technologies (CSIT'2005), Yerevan, Armenia, P. 393-398.
- [Bogdanov, et al., 2008] *Bogdanov A., Gevorkyan A., Nyman G.* Regular and Chaotic Quantum Dynamics in Atom-Diatom Reactive Collisions. //Physics of Atomic Nuclei. Springer. N 5. P.876-883.
- [Bogdanov, 2012] *Bogdanov A.* Private Cloud vs Personal Supercomputer. //Proc. of the 5th Int. Conf. "Distributed Computing and Grid Technologies in Science and Education". Dubna: JINR, P.57-59.
- [Bogdanov et al., 2015] *Bogdanov A., Burikova I., Degtyarev A., Konovalova M., Yuriev A.* Mathematical Model of Psychology-Political Classification of Political Parties. //European Physical Journal (в печати).
- [Brewer, 2010] *Brewer E.A.* A Certain Freedom: Thoughts on the CAP Theorem. //Proc.of the XXIX ACM SIGACT-SIGOPS symposium on Principles of distributed computing, N.Y.: ACM. 29. N 1. P.335-336.
- [Degtyarev, 2003] *Degtyarev A.* High performance computer technologies in shipbuilding. // in book L.Birk, S.Harries (eds.) "Optimistic. Optimization in Marine Design", Mensch & Buch Verlag, Berlin, 2003, P.245-258
- [Gankevich, Degtyarev, 2013] *Gankevich I., Degtyarev A.* Model of Distributed Computations in Virtual Testbed. //Proc. of IX Int. Conf. on Computer Science and Information Technologies (CSIT'2013), Yerevan, Armenia, P. 240-244,.
- [Gankevich et al., 2015] *Gankevich I., Tipikin Yu., Degtyarev A., Korkhov V.* Novel Approaches for Distributing Workload on Commodity Computer Systems. (Springer LNCS V. 9158)/ Ed. Gervasi O., et al. Springer, N.Y., 2015. P.259-271.
- [James, 2004] *James T Yu.* Performance evaluation of Linux Bridge. In Telecommunications System Management Conference, 2004.
- [Korkhov et al., 2008] *Korkhov V., Krzhizhanovskaya V., Sloot P.* A Grid Based Virtual Reactor: Parallel Performance and Adaptive Load Balancing. //Journal of Parallel and Distributed Computing. 68 (5). P.596-608.

SUPERCOMPUTING WITHOUT SUPERCOMPUTERS: WHAT WE CAN AND WHAT WE CAN NOT?

A.V. Bogdanov, A.B. Degtyarev

Saint Petersburg State University (SPSU)

The article discusses the possibilities of solving of complex problems requiring the use of supercomputer technology on a multiprocessor cluster resource currently available to most researchers. It is shown that without serious preliminary mathematical research effective implementation of complex tasks on cluster technique is not possible. The ways of solution of such problems are illuminated.

Keywords: cluster, virtual supercomputer.

Bogdanov Alexander – Professor for computer simulation and multiprocessor systems, Faculty of Applied Mathematics – Control Processes, Dr. Sci., a.v.bogdanov@spbu.ru

Degtyarev Alexander Borisovich – Professor for computer simulation and multiprocessor systems, Faculty of Applied Mathematics – Control Processes, Dr. Sc., a.degtyarev@spbu.ru

ДОСТИЖЕНИЕ ЭКЗАФЛОПСНОЙ ПРОИЗВОДИТЕЛЬНОСТИ В ЗАДАЧАХ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

В.П. Гергель, К.А. Баркалов

Нижегородский государственный университет им. Н.И. Лобачевского (ННГУ)

Эффективное использование огромных вычислительных возможностей экзафлопсных систем представляет собой глобальную проблему «вызова» всему спектру вычислительных наук. Условиями для успешного достижения экзафлопсной производительности являются: значительная вычислительная трудоемкость решаемых задач, высокий запас параллелизма вычислений, низкая информационная зависимость вычислений, устойчивость вычислений к аппаратным сбоям вычислителей. В работе показывается, что все перечисленные условия достижимы для задач глобальной оптимизации.

Исследование выполнено при поддержке РФФИ, грант № 11-01-00682-а, а также Министерства образования и науки Российской Федерации, соглашение № 14.В37.21.0393

Ключевые слова: высокопроизводительные вычисления, глобальная оптимизация, редукция размерности, параллельные алгоритмы.

1. ПРОБЛЕМА ОРГАНИЗАЦИИ ЭФФЕКТИВНЫХ ЭКЗАФЛОПСНЫХ ВЫЧИСЛЕНИЙ

Колоссальный вычислительный потенциал современных суперкомпьютерных систем позволяет приступить к решению многих сложнейших научно-прикладных проблем, анализ которых ранее находился за гранью возможного. Динамика развития производительности суперкомпьютеров опережает все самые смелые ожидания специалистов. Так, в 2008 г. был преодолен петафлопсный (10^{15} операций с плавающей запятой в сек.) рубеж скорости вычислений, в 2018–2020 гг. ожидается достижение экзафлопсного (10^{18}) уровня производительности. По оценкам специалистов, подобные суперкомпьютерные системы с рекордными вычислительными показателями будут существенно многопроцессорными (до миллиарда вычислительных ядер), гибридными с разными типами вычислительных устройств с многоуровневой структурой организации вычислений (распределенные вычислительные устройства - вычислительные узлы с общей разделяемой памятью – многоядерные процессорные элементы – ускорители вычислений) (International Exascale Software Project. URL: <http://www.exascale.org>). Эффективное использование огромных вычислительных возможностей экзафлопсных систем представляет собой глобальную проблему «вызова» всему спектру вычислительных наук. Условиями для успешного преодоления этой проблемы являются:

- значительная **вычислительная трудоемкость** решаемых задач (свыше 10^{18} операций);

Гергель Виктор Павлович – д.т.н., проф., Нижегородский государственный университет им. Н.И. Лобачевского, gergel@unn.ru

Баркалов Константин Александрович – к.ф.-м.н., доц., Нижегородский государственный университет им. Н.И. Лобачевского, barkalov@vmk.unn.ru

- высокий запас параллелизма (**масштабируемость**) вычислений (вплоть до использования 10^9 процессоров);
- низкая интенсивность информационных взаимодействий (**локальность**) параллельно выполняемых вычислений;
- **устойчивость** вычислений к аппаратным сбоям вычислителей, которые неизбежно будут происходить при столь больших количествах вычислительных элементов.

В работе показывается, что все перечисленные условия достижимы для задач глобальной оптимизации.

2. ОЦЕНКА ВЫЧИСЛИТЕЛЬНОЙ ТРУДОЕМКОСТИ ЗАДАЧ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

Прежде всего, следует отметить, что задачи оптимизации имеют важное прикладное значение и чрезвычайно широкую область приложений – практически каждая задача проектирования новых устройств, изделий, систем включает в себя этап выбора оптимальных вариантов. В наиболее сложных ситуациях проблема выбора формулируется как задача глобальной оптимизации, когда критерии оптимальности могут быть многоэкстремальными, т.е. иметь множество локально-оптимальных решений, среди которых необходимо выделить наилучшие варианты. Допустимость многоэкстремальности существенно повышает сложность решения оптимизационной задачи, ибо, если для подтверждения локального минимума достаточно исследования локальной окрестности, то глобальный минимум является интегральной характеристикой решаемой оптимизационной задачи и требует исследования всей области глобального поиска. Как результат, поиск глобального оптимума сводится к построению некоторого покрытия (сетки) в области вариации параметров. Поэтому на сложность решения задач рассматриваемого класса решающее влияние оказывает размерность: они подвержены “проклятию размерности”, состоящему в экспоненциальном росте вычислительных затрат при увеличении размерности. Так, если вычисление одного значения оптимизируемой функции требует m операций, а количество узлов сетки при решении одномерной задачи равно k , то общая оценка общего объема вычислений для N -мерной задачи глобальной оптимизации имеет вид

$$T = mk^N.$$

Приняв, например, $m = 10^3$, $k = 10$, $N = 20$, общий объем вычислений получается равным $T = 10^{23}$, что требует порядка суток вычислений на суперкомпьютере эксафлопского уровня производительности.

Снижение объема вычислений может быть достигнуто при построении существенно неравномерных покрытий областей глобального поиска – эти сетки должны быть достаточно плотными в окрестности глобального оптимума и более разреженными вдали от искомого решения. Построение таких оптимальных покрытий обеспечивается при повышении сложности самих численных методов глобального поиска.

Применение неравномерных покрытий позволяет повысить размерность решаемых задач глобальной оптимизации в 2–3 раза, что является критичным в приложениях. Получаемые оценки размерности (25–50) позволяют охватить большинство научно-прикладных задач, поскольку, как правило, количество варьируемых параметров, по которым проявляется многоэкстремальность оптимизируемых критериев, ограничено.

3. ОСНОВНЫЕ ПРИНЦИПЫ ОРГАНИЗАЦИИ ВЫЧИСЛЕНИЙ В ЗАДАЧАХ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

3.1. Класс решаемых задач

Задача многомерной многоэкстремальной оптимизации может быть определена как проблема поиска наименьшего значения действительной функции $\varphi(y)$ ¹

$$\varphi(y^*) = \min \{ \varphi(y) : y \in D \}, \quad D = \{ y \in R^N : a_i \leq y_i \leq b_i, 1 \leq i \leq N \}, \quad (1)$$

где $a, b \in R^N$ есть заданные векторы.

Численное решение задачи (1) сводится к построению оценки $y_e^* \in D$, отвечающей тому или иному понятию близости к точке y^* на основе конечного числа k значений функционалов задачи, вычисленных в точках области D .

Относительно класса рассматриваемых задач предполагается выполнение двух важных принципиальных требований:

- процедуры проведения *испытаний* (вычисления значений целевой функции) является *вычислительно-трудоемкими*,
- функция $\varphi(y)$ должна удовлетворять *условию Липшица* (ограниченность изменения значений функции при ограниченной вариации параметров).

Следует отметить, что именно выполнимость условия Липшица позволяет строить численные оценки искомого глобального решения по конечному набору вычисленных значений функции $\varphi(y)$.

Отметим также, что все далее излагаемые результаты получены в рамках информационно-статистической теории многоэкстремальной оптимизации [Стронгин, 1978; Strongin, Sergeyev, 2000], являющейся одним из наиболее активно развиваемых направлений исследований в области глобального поиска.

3.2. Редукция размерности

Для снижения сложности алгоритмов глобальной оптимизации, формирующих эффективные покрытия области поиска, в теории и практике многоэкстремальной оптимизации широко используются различные схемы редукции размерности, которые позволяют свести решение многомерных оптимизационных задач к семейству задач одномерной оптимизации. Редукция размерности позволяет существенно снизить сложность разрабатываемых алгоритмов глобального по-

¹ Для уменьшения сложности изложения материала задача глобальной оптимизации в данной работе рассматривается при отсутствии функциональных ограничений.

иска. Кроме того, данный подход позволяет задействовать весь имеющийся аппарат одномерной многоэкстремальной оптимизации для построения эффективных многомерных методов глобального поиска.

Один из наиболее общих методов редукции размерности состоит в применении разверток единичного отрезка вещественной оси на гиперкуб. Роль таких разверток играют непрерывные однозначные отображения типа кривой Пеано, называемые также кривыми, заполняющими пространство.

Использование развертки Пеано $y(x)$, однозначно отображающей единичный отрезок вещественной оси на единичный гиперкуб, позволяет свести многомерную задачу минимизации в области D к одномерной задаче условной минимизации на отрезке $[0, 1]$

$$\varphi(y(x^*)) = \min \{ \varphi(y(x)) : x \in [0, 1] \}. \quad (2)$$

Вопросы численного построения отображений типа кривой Пеано и соответствующая теория подробно рассмотрены в [Стронгин, 1978; Strongin, Sergeyev, 2000].

3.3. Общая схема алгоритмов глобального поиска

Успехи в разработке одномерных алгоритмов глобального поиска в рамках информационно-статистической теории глобального поиска позволили получить чрезвычайно значимый результат – к формулированию общей схемы представления алгоритмов, в рамках которой с единых позиций можно провести анализ свойств методов и даже ставить задачу построения алгоритмов с заданными свойствами.

Данная общая схема, получившая наименование *схемой характеристической представимости алгоритмов глобального поиска* (см., например, [Стронгин, 1978]), состоит в следующем.

Алгоритм глобального поиска называется *характеристическим*, если, начиная с некоторого шага поиска $k_0 \geq 1$, выбор точки x^{k+1} очередной итерации (*решающее правило алгоритма*) включает выполнение следующего набора действий:

1. Упорядочить поисковую информацию в порядке возрастания значений точек ранее выполненных итераций (новый порядок расположения данных в поисковой информации отражается при помощи нижнего индекса):

$$x_1 \leq x_2 \leq \dots \leq x_k,$$

2. Вычислить для каждого интервала (x_{i-1}, x_i) , $1 < i \leq k$, величину $R(i)$, называемую далее *характеристикой* интервала.

3. Определить интервал (x_{i-1}, x_i) , которому соответствует максимальная характеристика $R(i)$, т.е.

$$R(t) = \max \{ R(i) : 1 < i \leq k \},$$

4. Вычислить точку x^{k+1} очередной итерации глобального поиска в интервале с максимальной характеристикой в соответствии с некоторым правилом

$$x^{k+1} = s(t) \in (x_{i-1}, x_i).$$

4. ПАРАЛЛЕЛЬНЫЕ ВЫЧИСЛЕНИЯ ДЛЯ ПОИСКА ГЛОБАЛЬНО-ОПТИМАЛЬНЫХ РЕШЕНИЙ

4.1. Основной принцип организации параллельных вычислений

Прежде всего, следует отметить, что в случае многомерной многоэкстремальной оптимизации многие широко используемые подходы для распараллеливания или не обеспечивают должного эффекта или вообще не применимы. В частности, один из основных способов организации параллельных вычислений состоит в разделении области решения задачи между процессорами и параллельного решения задачи в этих подобластях. В случае решения оптимизационных задач такой подход обладают низкой эффективностью, поскольку при разделении области поиска только небольшая часть процессоров будет обладать подобластью с искомым глобальным минимумом, все же остальные процессоры будут работать в подобластях, в которых отсутствует глобальный минимум исходной задачи.

Плохо применим в данном случае и подход, связанный с распараллеливанием непосредственно алгоритма решения задачи, ибо в силу сформулированных предположений о классе рассматриваемых задач основной объем вычислений в процессе глобального поиска занимают расчеты, связанные с вычислением значений функционалов оптимизационной задачи. В этом плане, целесообразным представляется распараллеливание процедур расчета значений функционалов, однако такой подход не обладает общностью и требует проведения действий по распараллеливанию для каждой оптимизационной задачи в отдельности.

С учетом выполненного обсуждения, в качестве основного принципа построения параллельных методов многоэкстремальной оптимизации в диссертации будет применяться схема параллельных вычислений, в которой будет обеспечиваться возможность одновременного (параллельного) расчета значений функционалов оптимизационной задачи в нескольких разных точках области поиска (см., например, [Strongin, Sergeyev, 2000; Стронгин, 2009]). Такой подход характеризуется *эффективностью* (распараллеливается именно та часть вычислительного процесса, в котором выполняется основной объем вычислений) и *общностью*, поскольку применим для всех вычислительно-трудоемких задач многоэкстремальной оптимизации.

4.2. Параллельные вычисления для систем с распределенной памятью

Новый подход для организации параллельных вычислений в задачах глобальной оптимизации состоит в использовании для редукции размерности множества отображений

$$Y_L(x) = \{y^1(x), \dots, y^L(x)\}$$

вместо применения единственной кривой Пеано $y(x)$ [Strongin, Sergeyev, 2000; Стронгин, 2009]. Использование множества отображений $Y_L(x)$ приводит к формированию соответствующего множества одномерных многоэкстремальных задач

$$\min \{ \varphi(y^l(x)) : x \in [0, 1] \}, \quad 1 \leq l \leq L. \quad (3)$$

Каждая задача из данного набора может решаться независимо, при этом любое вычисленное значение $z = \varphi(y^i(x^i))$ функции $\varphi(y)$ в i -й задаче может интерпретироваться как вычисление значения $z = \varphi(y^s(x^s))$ для любой другой s -й задачи без повторных трудоемких вычислений функции $\varphi(y)$. Подобное информационное единство позволяет решать исходную задачу (1) путем параллельного решения L задач вида (4) на наборе отрезков $[0, 1]$. Каждая одномерная задача решается на отдельном процессоре. Результаты вычислений в точке x^k , полученные конкретным процессором для решаемой им задачи, интерпретируются как результаты вычислений во всех остальных задачах (в соответствующих точках x^{k1}, \dots, x^{kL}). При таком подходе вычисления в точке $x^k \in [0, 1]$, осуществляемые в s -й задаче, состоят в последовательности действий:

1. Определить образ $y^k = y^s(x^k)$ при соответствии $y^s(x)$.
2. Проинформировать остальные процессоры о начале проведения испытания в точке y^k (блокирование точки y^k).
3. Вычислить значение оптимизируемой функции $z^k = \varphi(y^s(x^k))$.
4. Определить прообразы $x^{kl} \in [0, 1]$, $1 \leq l \leq L$, точки y^k , и интерпретировать значение функции $j(y)$, вычисленное в точке $y^k \in D$, как одновременные вычисления в L точках

$$x^{k1}, \dots, x^{kL},$$

с одинаковыми результатами

$$\varphi(y^1(x^{k1})) = \dots = \varphi(y^L(x^{kL})) = z^k.$$

5. Проинформировать остальные процессоры о результатах испытания в точке y^k .

Каждый процессор при таком подходе должен содержать свою копию программных средств, реализующих вычисление функций задачи, и решающее правило алгоритма. Для организации взаимодействия на каждом процессоре создается L очередей, в которые процессоры помещают информацию о выполненных итерациях (точки очередных итераций и значения оптимизируемой функции).

Рассмотренный подход позволяет использовать для решения задач глобальной оптимизации порядка нескольких тысяч процессоров – ограничением возможного количества используемых вычислительных элементов является трудоемкость информационного взаимодействия при организации обменов результатов итераций глобального поиска между процессорами.

4.3. Параллельные вычисления для систем с общей разделяемой памятью

Ранее рассмотренная схема применима и для систем с общей разделяемой памятью, тем более, что в этом случае будут отсутствовать затраты на обмен результатов итераций глобального поиска между процессорами. Вместе с этим, для организации параллельных вычислений может быть применена и другая схема, при которой параллельно могут решаться и отдельные задачи семейства (3). Для обоснования подхода можно отметить, что используемые в алгоритмах характе-

ристики интервалов (см. выше общую схему методов глобального поиска) могут рассматриваться как некоторые меры важности интервалов на предмет содержания в интервалах искомого решения оптимизационной задачи. Следуя данному пониманию, для организации параллельных вычислений после выбора точки вычисления значения функции для первого процессора в точном соответствии с последовательным алгоритмом для второго процессора точку очередной итерации целесообразно выбирать из следующего по важности интервала (т.е. из интервала со следующей по порядку максимальной характеристикой) и т.д.

На данной основе может быть предложена общая схема *параллельных асинхронных характеристических методов*, определяемая следующим образом [Стронгин и др., 2012].

Выбор точек очередных итераций глобального поиска, начиная с некоторого шага поиска $k_0 \geq 1$ включает выполнение следующего набора действий:

1) При отсутствии свободных процессоров алгоритм находится в состоянии ожидания. Если закончили проведение испытаний $1 \leq q = q(k) \leq p(k)$ процессоров, то устанавливается номер итерации $k = k + q$, и осуществляется переход к следующему правилу общей схемы.

2) Упорядочить поисковую информацию в порядке возрастания значений точек ранее выполненных итераций (новый порядок расположения данных в поисковой информации отражается при помощи нижнего индекса):

$$0 = x_0 < x_1 < \dots < x_{\tau-1} < x_k = 1.$$

3) Вычислить для каждого интервала (x_{i-1}, x_i) , $1 < i \leq k$, величину $R(i)$, называемую далее *характеристикой* интервала,.

4) Упорядочить характеристики $R(i)$, $1 < i \leq \tau$, в порядке убывания:

$$R(t_1) \geq R(t_2) \geq \dots \geq R(t_{k-1}) \geq R(t_k) \quad (4)$$

5) Выбрать в последовательности (4) q наибольших характеристик с номерами t_j , $1 \leq j \leq q$, и в интервалах, соответствующих этим характеристикам, выполнить испытания в точках

$$\bar{x}^j = d(t_j) \in (x_{t_j-1}, x_{t_j}), \quad 1 \leq j \leq q.$$

Как можно заметить, рассмотренная общая схема практически полностью соответствует схеме характеристической представимости алгоритмов глобального поиска.

Отметим также, рассмотренная схема асинхронных параллельных вычислений полностью применима и для случая, когда вычислительные узлы с общей разделяемой памятью содержат ускорители вычислений вида современных графических процессоров. Отличие состоит лишь в том, что определяемые точки итераций глобального поиска должны передаваться для использования ускорителям вычислений.

Предполагая наличие количество вычислительных элементов (ядер) на общей памяти порядка 100, использование параллельных алгоритмов в рамках рассмотренных схем параллельного глобального поиска позволяет довести воз-

можное число одновременно используемых вычислительных устройств суперкомпьютерных систем экзафлопсного уровня производительности до нескольких сотен тысяч.

ЛИТЕРАТУРА

- [Стронгин, 1978] *Стронгин Р.Г.* Численные методы в многоэкстремальных задачах. М.: Наука, 1978, 240 с.
- [Стронгин, 2009] *Стронгин Р.Г., Гергель В.П., Баркалов К.А.* Параллельные методы решения задач глобальной оптимизации // Известия высших учебных заведений. Приборостроение. 2009. Т. 52. № 10. С. 25-33.
- [Стронгин и др., 2012] *Стронгин Р.Г., Гергель В.П., Гришагин В.А., Баркалов К.А.* Параллельные вычисления в задачах глобальной оптимизации. – М.: МГУ, 2013, 280 с.
- [Strongin, Sergeyev, 2000] *Strongin R.G., Sergeyev Ya.D.* Global optimization with non-convex constraints. Sequential and parallel algorithms. Kluwer Academic Publishers, Dordrecht, 2000, 704 p.

ACHIEVING EXAFLOPS PERFORMANCE IN GLOBAL OPTIMIZATION PROBLEMS

V.P. Gergel, K.A. Barkalov

Lobachevsky State University of Nizhni Novgorod (NNGU)

Effective use of huge computing potential of exascale systems is the great challenge for all computational sciences. Conditions for successful achievement of exaflops performance are large computational complexity of problems, high reserve of computation parallelism, low information dependence of computations, stability of calculations relative to hardware failures. The paper shows that all mentioned conditions are achievable for global optimization problems.

Keywords: high-performance computing, global optimization, dimension reduction, parallel algorithms.

Gergel Victor Pavlovich – doctor of technical sciences, professor, gergel@unn.ru

Barkalov Konstantin Aleksandrovich – candidate of physico-mathematical sciences, associate professor, barkalov@vmk.unn.ru

КВАНТОВЫЕ ТРАЕКТОРИИ И ТОМОГРАФИЯ СОСТОЯНИЙ СВЯЗАННОЙ СИСТЕМЫ КУБИТ-ИЗМЕРИТЕЛЬНЫЙ ПРИБОР

В.П. Гергель, М.В. Денисенко, А.В. Линёв, Д.С. Пашин, А.М. Сатанин

Нижегородский государственный университет им. Н.И. Лобачевского (ННГУ)

Разработан алгоритм, основанный на стохастическом методе квантовых траекторий (квантовом методе Монте-Карло), для моделирования диссипативной динамики многоуровневой квантовой системы. На основе развитого метода проведено моделирование процесса взаимодействия кубита с квантовым нелинейным осциллятором. Показано, что в результате измерения средней энергии (числа квантов) осциллятора удается провести томографию состояний кубита, как в среднем по ансамблю, так и в единичных траекториях. При этом квантовый метод Монте-Карло позволяет осуществить эффективное распараллеливание на гетерогенном высокопроизводительном вычислительном кластере при расчете экспериментально наблюдаемых величин в многоуровневых квантовых системах с учетом процессов релаксации. Описаны основные принципы распараллеливания алгоритма для решения стохастического уравнения Шредингера, которые были реализованы в программном комплексе, позволяющем осуществлять высокопроизводительные параллельные расчеты с использованием GPU. Продемонстрирована эффективность работы графических ускорителей, масштабируемость метода и описывается возможность взаимодействия нескольких графических ускорителей с применением технологии MPI.

Работа поддержана РФФИ (проект № 14-07-00582), программой СИ7 02.В.49.21.0003 МОН РФ и ННГУ, некоммерческим фондом «Династия».

Ключевые слова: кубит, неразрушаемые измерения, метод квантовых траекторий (квантовый метод Монте-Карло), GPU, CUDA.

ВВЕДЕНИЕ

Двухуровневая система – кубит – является базовым элементом квантового компьютера, который в настоящее время активно разрабатывается во многих лабораториях мира [Нильсен, Чанг, 2006; Ladd et al., 2010; You, Nori, 2011]. Важную роль в условиях экспериментов с квантовыми системами играют процессы инициализации состояний, процессы измерений и шумовое воздействие на исследуемую систему.

Гергель Виктор Павлович – Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина 23, 603950, Н. Новгород, Россия. Д.т.н., профессор кафедры МО ЭВМ, dean@vmk.unn.ru

Денисенко Марина Валерьевна – Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина 23, 603950, Н. Новгород, Россия. М.н.с., mar.denisenko@gmail.com

Линёв Алексей Владимирович – Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина 23, 603950, Н. Новгород, Россия. Зав лабораторией при кафедре ПИ, alin@unn.ru

Пашин Дмитрий Сергеевич – Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина 23, 603950, Н. Новгород, Россия. Магистрант, agent150@rambler.ru

Сатанин Аркадий Михайлович – Нижегородский государственный университет им. Н.И. Лобачевского, пр. Гагарина 23, 603950, Н. Новгород, Россия. Д.ф.-м. н., профессор кафедры теоретической физики, sarkady@mail.ru

двумя системами. Для инициализации и измерений необходимо привести регистр кубитов во взаимодействие с измерительным прибором, который добавляет дополнительные степени свободы и шумовое влияние, что необходимо принимать во внимание при моделировании.

В данной работе рассматривается простейшая модель многоуровневой квантовой системы: кубит, связанный с квантовым нелинейным осциллятором, который используется для проективных измерений состояний кубита [Siddiqi et al., 2004, 2006], что было продемонстрировано экспериментально в недавней работе [Vijay et al., 2011]. Суть эксперимента состояла в следующем: кубит был помещен в квантовый сверхпроводящий резонатор, куда через входные порты для инициализации начального состояния подавались Раби-импульсы заданной длительности. Сигнал с кубита через выходные порты подавался на шумоустойчивый параметрический усилитель [Hatridge et al., 2011] – нелинейный джозефсоновский осциллятор, показания которого оказываются чувствительны к состояниям кубита. Процесс манипуляции состояниями кубитов в условиях шума можно было наблюдать во времени, исследуя как динамику системы в случае единичной реализации, так и генерируя временной ансамбль систем, обладающих одинаковыми начальными условиями.

Численное моделирование влияния окружения на простейшую систему «кубит + осциллятор» имитируется путем их взаимодействия с бозонным термостатом [Скалли, Зубайри, 2003]. Измерение производится за счет считывания среднего числа квантов на осцилляторе (средней энергии нелинейного джозефсоновского осциллятора). При условии слабого обратного воздействия на термостат и локального по времени возмущения (марковское приближение), динамика системы может быть описана уравнением для матрицы плотности [Скалли, Зубайри, 2003]. Число состояний связанной системы равно $2N$, где N – число уровней нелинейного осциллятора. Для выполнения измерений важно, чтобы измерительный прибор, производящий дискриминирование состояний кубита, работал в квазиклассическом режиме, т.е. число активных состояний N было велико. Решение связанных дифференциальных уравнений для элементов матрицы плотности масштабируется как $(2N)^4$. Квантовый метод Монте-Карло основан на введении для расчета элементов матрицы плотности вспомогательной волновой функции, которая подчиняется стохастическому дифференциальному уравнению, по форме напоминающему уравнение Шредингера со случайными коэффициентами (винеровским случайным процессом) [Molmer et al., 1992; Plenio, Knight, 1998]. В случае системы «кубит-осциллятор» необходимо решать $2N$ уравнений для стохастической волновой функции, а наблюдаемые величины определять затем усреднением билинейной комбинаций монте-карловских волновых функций по большому числу реализаций M (число реализаций определяется необходимой точностью). Решение независимых дифференциальных уравнений хорошо распараллеливается, поэтому можно ожидать значительного ускорения процесса вычислений усредненных величин, аналогичным тем, что получают при решении матрицы плотности.

В представленной работе описаны принципы распараллеливания и численного моделирования стохастического уравнения Шредингера, а также программный комплекс, позволяющий осуществлять высокопроизводительные параллельные расчеты на GPU. Показана эффективность применения квантового метода Монте-Карло для численного моделирования томографии состояний кубитов, которые были возбуждены короткими Раби-импульсами заданной длительности.

МОДЕЛЬ И ОСНОВНЫЕ УРАВНЕНИЯ

Хотя наши рассуждения могут быть отнесены к любой многоуровневой квантовой системе, здесь мы будем интересоваться динамикой джозефсоновского осциллятора, взаимодействующего с зарядовым кубитом специального вида (так называемый «квантроний кубит» – *quantronium qubit*) [Vion et al., 2002]. Фактически этот прибор состоит из расщепленного зарядового кубита, когда два сверхпроводящих перехода (перечеркнутые квадраты на рис. 1), функционирующие в режиме кулоновской блокады, имеют полную емкость C_j , каждый из них – джозефсоновскую энергию $E_j/2$, которые «смещены» постоянным потенциалом V_g и могут возбуждаться дополнительно микроволновым излучением через емкость C_g (см. рис. 1).

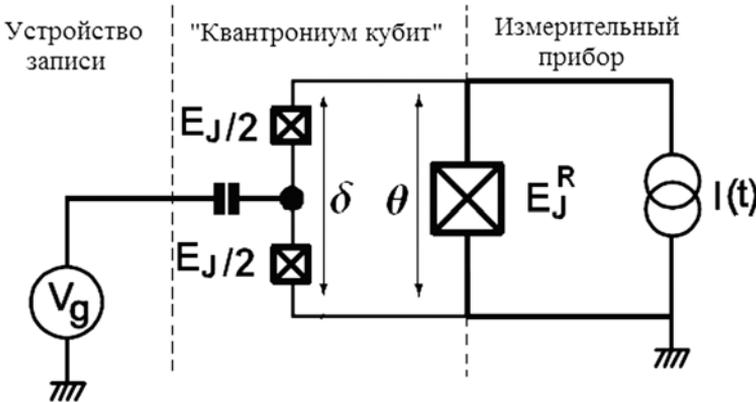


Рис. 1. Схема кубита (центр рисунок), связанного с устройством записи (левая часть рисунок) и взаимодействующего со слабодиссипативным измерительным джозефсоновским осциллятором (правая часть рисунок). Перечеркнутыми квадратиками обозначены джозефсоновские переходы

Используя уравнения Джозефсона и законы Кирхгофа, можно записать гамильтониан островка и переходов в виде $H = H_{QO} + H_j$, где гамильтониан расщепленного куперовского островка («черный кружок», разделяющий джозефсоновские переходы на рис. 1) H_{QO} включает в себя электростатическую энергию островка и джозефсоновскую энергию двух маленьких сверхпроводящих переходов (каждый с энергией $E_j/2$):

$$H_{QO} = 4E_c \left(\hat{N} - \frac{1}{2} - \frac{C_g V(t)}{2e} \right)^2 - E_j \cos \frac{\theta}{2} \cos \gamma \quad , \quad (1)$$

а гамильтониан джозефсоновского осциллятора имеет вид:

$$H_J = \frac{Q^2}{2C} - E_J^R \cos\theta - \frac{\hbar^2}{2e} I(t)\theta, \quad (2)$$

где \hat{N} и $Q/2e$ обобщенные импульсы, сопряженные к фазам γ и θ , соответственно куперовского и джозефсоновского переходов (подробности вывода гамильтониана системы можно найти в работе [Vion et al., 2002]). Как видно, связь двух подсистем осуществляется за счет общего участка цепи, обуславливающего появление фазы θ в потенциальной энергии (1). Управление состояниями кубита осуществляется переменным напряжением $V(t) = V_g + V_{rf}(t)\cos\omega t$, которое меняет заряд на емкости C_g , а переменный ток $I(t)$ служит для управления измерительным осциллятором.

Вблизи точки кулоновского вырождения, когда энергии островка для N и $N+1$ частиц равны, можно использовать базисные состояния $|N\rangle$ и $|N+1\rangle$ оператора числа частиц $\hat{N}|N\rangle = N|N\rangle$. В этом подпространстве: $|N\rangle = (1, 0)^T$ and $|N+1\rangle = (0, 1)^T$ гамильтониан аппроксимируется выражением

$$H \cong 2E_C \frac{C_g V_{rf}(t)}{e} \sigma_x - E_J \cos \frac{\theta}{2} \sigma_z + \frac{Q^2}{2C} - E_J^R \cos\theta - \frac{\hbar^2}{2e} I(t)\theta, \quad (3)$$

где σ_x и σ_z – спиновые матрицы Паули.

При относительно малом токе $I(t)$ джозефсоновский осциллятор работает в слабо нелинейном режиме. Тогда косинусы фазы в выражение (3) можно разложить по степеням θ :

$$H_{eff} \cong 2E_C \frac{C_g V_{rf}(t)}{e} \sigma_x - E_J \sigma_z + \frac{Q^2}{2C} + E_J (1 + \lambda \sigma_z) \frac{\theta^2}{2} - E_J \left(1 + \frac{\lambda}{4} \sigma_z\right) \frac{\theta^4}{4!} - \frac{\hbar^2}{2e} I(t)\theta, \quad (4)$$

где $\lambda = E_J / 4E_J^R$. Вводя операторы рождения a^+ и уничтожения a для параметризации координаты $\theta = (a + a^+) (2E_C^R / E_J)^{1/4}$ и импульса осциллятора $\hat{N} = -i(a - a^+) (E_C^R / 2E_J)^{-1/4}$ и учитывая, что $[\hat{\theta}, \hat{N}] = i$, можно получить эффективный гамильтониан связанной системы «кубит+нелинейный осциллятор» в расширенном гильбертовом пространстве размерности $2N$:

$$H_{eff} = H_q \otimes E^{(N)} + E^{(2)} \otimes H_{osc} + H_{int}, \quad (5)$$

где единичные матрицы $E^{(2)}$ и $E^{(N)}$ (размерностями, соответственно, 2×2 и $N \times N$), гамильтониан кубита, записан в стандартном виде:

$$H_q = \frac{\hbar}{2} (\omega_q \sigma_z + \varepsilon(t) \sigma_x). \quad (6)$$

Можно видеть, что расстояние между уровнями кубита ($\hbar\omega_q = -2E_J$) определяется частотой $\omega_q/2\pi$ (~ 10 ГГц); а управляющая функция $\varepsilon(t) = 4E_C \frac{C_g V_{rf}(t)}{e}$ будет индуцировать переходы между состояниями кубита $|\sigma\rangle$ ($\sigma_z |\sigma\rangle = \sigma |\sigma\rangle$, $\sigma = \pm 1$). Предполагается, что управление кубитом будет осуществляется Раби-импульсами переменного поля $\varepsilon(t) \sim V(t) = A \cos \omega t$ определенной длительности.

Измерение состояний кубита предполагается производить нелинейным осциллятором, где для простоты учтены только резонансные слагаемые в отличие от исходного уравнения (4):

$$H_{osc} = \hbar\omega_J a^+ a - \mu(a^+ a)^2 + f(t)(a + a^+), \quad (7)$$

где $\omega_J = \sqrt{\frac{2eI_c}{\hbar C}}$ – собственная частота осциллятора, $\mu = E_c^R / 12\hbar$ – параметр нелинейности, $f(t) \square I(t) / I_c = f_0 \cos\Omega t$ – возбуждающее поле осциллятора ($f_0 = \frac{\hbar}{2e} \left(\frac{2E_c^R}{E_J} \right)^{1/4}$). Практически в процессе измерения джозефсоновский осциллятор может быть захвачен в нелинейный резонанс током «накачки» $I(t)$. Такая процедура осуществлена экспериментально в [Vion et al., 2002].

Согласно (4) взаимодействие кубита с измерительным осциллятором описывается выражением:

$$H_{int} = \lambda \left(\hbar\omega_J a^+ a - \frac{1}{4}\mu(a^+ a)^2 \right) \sigma_z, \quad (8)$$

где λ – константа связи кубита и осциллятора. Отметим, если управляющее поле отсутствует, $\varepsilon(t) = 0$, то оператор взаимодействия (8) коммутирует с гамильтонианом кубита (6), поэтому осциллятор производит так называемое «неразрушающее измерение» [Braginsky, Khalili, 1995]. В случае же действия поля ($\varepsilon(t) \neq 0$) состояния подсистем являются запутанными (entanglement states).

МЕТОД КВАНТОВЫХ ТРАЕКТОРИЙ

При рассмотрении динамики системы существенное влияние оказывают процессы декогерентности. Так, например, для сверхпроводникового кубита, таковыми процессами являются: флуктуации заряда на джозефсоновских контактах, квази-частицы на островках сверхпроводимости, взаимодействие с ядерными спинами в подложке, затухание при взаимодействии с электромагнитным управляющим полем и измерительным прибором. Отметим, что эксперименты проводятся при криогенных температурах, поэтому вклад шума фермионных возбуждений мал и не учитывается в данной работе. Гамильтониан, соответствующий вкладу бозонных возбуждений имеет вид:

$$H_{noise} = F_z \sigma_z + F_x \sigma_x + R a^+ + R^+ a, \quad (9)$$

где эрмитовы операторы F_z , F_x резервуара отвечают за продольную и поперечную релаксацию кубита (линейно зависят от операторов резервуара), а оператор $R(t)$ отвечает за релаксацию измерительного прибора. Таким образом, полный гамильтониан, описывающий релаксационную динамику связанной системы «кубит+осциллятор», имеет вид

$$H = H_{eff} + H_{noise}. \quad (10)$$

Для системы «кубит+осциллятор» уравнение для оператора плотности ρ , где исключены переменные резервуара F_z , F_x , R в форме Линдблада, в борн-марковском приближении [Скалли, Зубайри, 2003] записывается в виде:

$$\begin{aligned} \frac{\partial \rho}{\partial t} = & \frac{1}{i\hbar} [H, \rho] + \gamma_\varphi (\sigma_z \rho \sigma_z - \rho) + \frac{\gamma_\varepsilon}{2} (2 \sigma_- \rho \sigma_+ - \sigma_+ \sigma_- \rho - \rho \sigma_+ \sigma_-) \\ & + \frac{\gamma}{2} (2 a \rho a^\dagger - a^\dagger a \rho - \rho a^\dagger a), \end{aligned} \quad (11)$$

где $\sigma_\pm = (\sigma_x \pm i\sigma_y)/2$, а γ_φ и γ_ε – соответственно фазовая и энергетическая скорости релаксации кубита, γ – параметр релаксации осциллятора. Отметим, что данные параметры выражаются стандартным образом через корреляционными функциями бозонного термостата ($\langle F_z(t) F_z^\dagger \rangle$, $\langle R(t) R^\dagger \rangle$ и т.д.) (см. вывод в [Скалли, Зубайри, 2003]).

Уравнение для матрицы плотности (11) дает усредненную по ансамблю динамику системы. Чтобы проследить динамику кубита и осциллятора в каждом акте измерения, обратимся к эквивалентному языку описания динамики – методу квантовых траекторий или, как его ещё называют, квантовому методу Монте-Карло [Molmer et al., 1992; Plenio, Knight, 1998]. Согласно данному подходу эволюция системы может быть переформулирована на эквивалентном языке динамики состояний – «квантовых траекторий».

Для рассматриваемой системы суть метода квантовых траекторий сводится к следующему. Применяя конечно-разностную аппроксимацию для производной по времени в уравнении (11) перепишем его в виде:

$$\rho(t + \Delta t) = U \rho(t) U^\dagger + \Delta t (\gamma_\varphi \sigma_z \rho \sigma_z + \gamma_\varepsilon \sigma_- \rho \sigma_+ + \gamma a \rho a^\dagger), \quad (12)$$

где первое слагаемое в (12) определяется оператором эволюции $U = e^{-iH_{dis}\Delta t/\hbar}$ и отвечает за диссипативную динамику системы с гамильтонианом Вигнера-Вайскопфа (см. детали в [Скалли, Зубайри, 2003])

$$H_{dis} = H_q - i\hbar \frac{\gamma_\varphi}{4} \sigma_z^\dagger \sigma_z - i\hbar \frac{\gamma_\varepsilon}{2} \sigma_+^\dagger \sigma_- - i\hbar \frac{\gamma}{2} a^\dagger a. \quad (13)$$

Второе слагаемое в уравнении (12) можно связать со спонтанными переходами (скачками) в системе за счет взаимодействия с бозонным термостатом. Попробуем придать последнему слагаемому вероятностный смысл, имея в виду, что с уравнением (12) можно связать некоторый случайный процесс, которые будет имитировать относительный вклад диссипативной динамики и спонтанных переходов.

Формальное решение уравнения (12) сводится к следующему. Рассмотрим некоторую реализацию чистого состояния $\rho_i(t) = |\psi_i(t)\rangle\langle\psi_i(t)|$ (индекс i , нумерующий реализации, будем опускать в тех местах, где это не вызывает недоразумений). Стохастическая волновая функция связанной системы $|\psi_i(t)\rangle = \sum_{n,\sigma} C_{n,\sigma}(t) |n\sigma\rangle$ может быть разложена по базису $|n\sigma\rangle = |n\rangle \otimes |\sigma\rangle$. Следовательно, она имеет $2N$ компонент. Из уравнения (12) можно увидеть, что изменение матрицы плотности обусловлено двумя возможными вкладами, сопровождающимися изменением коэффициентов $C_{n,\sigma}(t)$ стохастической волновой функции:

1) либо с вероятностью $1 - \Delta P$ согласно

$$C_{n,\sigma}(t + \Delta t) = \frac{e^{-iH_{dis}\Delta t/\hbar} C_{n,\sigma}(t)}{\sqrt{1 - \Delta P}}, \quad (14)$$

происходит изменение волновой функции («траектории») за счет диссипативной динамики, здесь введены следующие обозначения: $\Delta P = \Delta t \bar{\gamma}$,

$$\bar{\gamma} = \gamma_{\varphi} + \frac{\gamma_{\varepsilon}}{2} \left(1 + \sum_{n,\sigma} \sigma |C_{n,\sigma}(t)|^2 \right) + \gamma \sum_{n,\sigma} n |C_{n,\sigma}(t)|^2;$$

2) либо с вероятностью ΔP изменение «траектории» сопровождается одним из скачков:

$$\text{фазовый для кубита} - C_{n,\sigma}^{\varphi}(t + \Delta t) = \sigma C_{n,\sigma}(t),$$

$$\text{энергетический для кубита} - C_{n,\sigma}^e(t + \Delta t) = \frac{\sum_{\sigma'} (\sigma_{-})_{\sigma\sigma'} C_{n,\sigma'}(t)}{\sqrt{\frac{1}{2} \left(1 + \sum_{n,\sigma} \sigma |C_{n,\sigma}(t)|^2 \right)}},$$

$$\text{спонтанный для осциллятора} - C_{n,\sigma}^{\gamma}(t + \Delta t) = \frac{\sqrt{n+1} C_{n+1,\sigma}(t)}{\sum_{n,\sigma} n |C_{n,\sigma}(t)|^2},$$

взятыми соответственно с весами

$$p_{\varphi} = \frac{\gamma_{\varphi}}{\bar{\gamma}}, \quad p_{\varepsilon} = \frac{\gamma_{\varepsilon}}{2\bar{\gamma}} \left(1 + \sum_{n,\sigma} \sigma |C_{n,\sigma}(t)|^2 \right) \quad \text{и} \quad p_{\gamma} = \frac{\gamma}{\bar{\gamma}} \sum_{n,\sigma} n |C_{n,\sigma}(t)|^2.$$

Итак, за время наблюдения эволюции системы мы получаем квантовую траекторию, представляющую одну реализацию мысленного эксперимента – диссипативную динамику системы. Учитывая, что процесс релаксации является случайным, каждая траектория уникальна. Для нахождения усредненной динамики (аналог того, что получаем в результате решения уравнения (11)) следует получить набор единичных реализаций $M \sim 1000-10000$ и далее усреднить величину по ним

$$\rho = \frac{1}{M} \sum_{i=1}^M |\psi_i(t)\rangle \langle \psi_i(t)|. \quad (15)$$

Кроме того, можно вычислить ожидаемое значение любой наблюдаемой квантовой системы. Например, среднее число квантов осциллятора (средняя энергия) для двух состояний кубитов (основного « \downarrow » и возбужденного « \uparrow »):

$$\bar{n}_{\downarrow} = \frac{1}{M} \sum_{i=1}^M \sum_{n=0}^{2N} n |C_{2n,\downarrow}^i(t)|^2$$

$$\bar{n}_{\uparrow} = \frac{1}{M} \sum_{i=1}^M \sum_{n=0}^{2N} n |C_{2n+1,\uparrow}^i(t)|^2 \quad (16)$$

Метод квантовых траекторий, в отличие от аналогичного метода для матрицы плотности, имеет ряд преимуществ. Во-первых, это возможность исследовать релаксационные процессы квантовых систем в единичных реализациях, а не только усредненные по ансамблю квантовые траектории. Во-вторых, это эффективность распараллеливания метода. Уравнение для матрицы плотности приводит к системе линейных уравнений, включающих $2N^2$ комплексных переменных, где $2N$ – это размерность гильбертова пространства («кубит+осциллятор»). Квантовый же метод Монте-Карло требует лишь рассмотрения $2N$ комплексных

переменных, характеризующих вектор состояний. Реализации статистически независимы, поэтому есть возможность генерировать каждую реализацию в отдельном потоке (на отдельном процессоре), собирая затем данные и усредняя.

МОДЕЛИРОВАНИЕ ТОМОГРАФИИ СОСТОЯНИЙ КУБИТА

Опишем численный эксперимент, имитирующий процессы, происходящие в открытой квантовой системе «кубит+нелинейный осциллятор». Как известно, классический нелинейный осциллятор демонстрирует бифуркационное поведение (бифуркация Андронова-Хопфа) при изменении амплитуды внешнего переменного поля [Siddiqi et al., 2004]. Выбирая соответствующим образом частоту и амплитуду внешней силы (накачки), можно осуществить колебание осциллятора вблизи сепаратрисы, отделяющей в фазовом пространстве положения равновесия осциллятора (разумеется, это осуществимо на временах, существенно меньших времени релаксации осциллятора). Однако в случае измерений состояний кубита работа осциллятора осуществляется в квантовом режиме, когда осциллятор захватывается в квантовый нелинейный резонанс. Поскольку частота осциллятора будет зависеть от состояния кубита, то он может «перебрасываться» в то или иное положение равновесия, зависящее от состояний кубита. Такой перебор может быть зарегистрирован, а процедура измерения осуществляет процесс «проектирования» квантового состояния. Эффективность использования бифуркационного джозефсоновского осциллятора в нелинейном режиме для квантовых измерений продемонстрирована экспериментально в работах [Siddiqi et al., 2004, 2006].

Численный эксперимент качественно описывает схему эксперимента [Vijaya et al., 2011] по измерению состояния кубита после воздействия коротких Раби-импульсов: π и 2π импульсов. В начальный момент времени кубит находится в основном состоянии, за счет π -импульса происходит возбуждение на высший уровень, а 2π -импульс оставляет кубит в основном состоянии. Далее воздействие на кубит прекращается ($A=0$ в выражении (6)), а, следовательно, гамильтониан кубита (6) коммутирует с гамильтонианом осциллятора (7), что позволяет выделить два подпространства, отвечающие за состояния кубита – основное (« \downarrow ») и возбужденное (« \uparrow »):

$$H_{\downarrow} = \hbar \omega_J (1 - \lambda) a^+ a - \mu \left(1 - \frac{\lambda}{4}\right) (a^+ a)^2 + f_0 \cos(\Omega t) (a + a^+), \quad (17)$$

$$H_{\uparrow} = \hbar \omega_J (1 + \lambda) a^+ a - \mu \left(1 + \frac{\lambda}{4}\right) (a^+ a)^2 + f_0 \cos(\Omega t) (a + a^+).$$

Таким образом, за счет эффективного сдвига собственной частоты измерительного осциллятора $\omega_{\downarrow} = \omega_J (1 - \lambda)$ и $\omega_{\uparrow} = \omega_J (1 + \lambda)$, а также параметра нелинейности, произойдет изменение состояний кубита, что позволит осуществить «дискриминацию» (томографию) состояний кубита.

Предварительно приведем результат тестирования программы. На рис. 2 приводится характерный график измерения средних энергий осциллятора (16)

\bar{n}_\downarrow (красная кривая) и \bar{n}_\uparrow (черная кривая) на временах, когда характерное время энергетической релаксации много больше времени действия считывающего импульса $\tau: 1/\gamma_\varepsilon \gg \tau = 50T$ ($T = 2\pi/\Omega$). Зависимости на рис. 2 усреднены по $M = 10000$ траекториям, что дает хорошее согласование с решением на основе прямого решения дифференциального уравнения (11). Погрешность за счет использования стохастического метода квантовых траекторий составляет $\sim 1/\sqrt{M} = 1\%$.

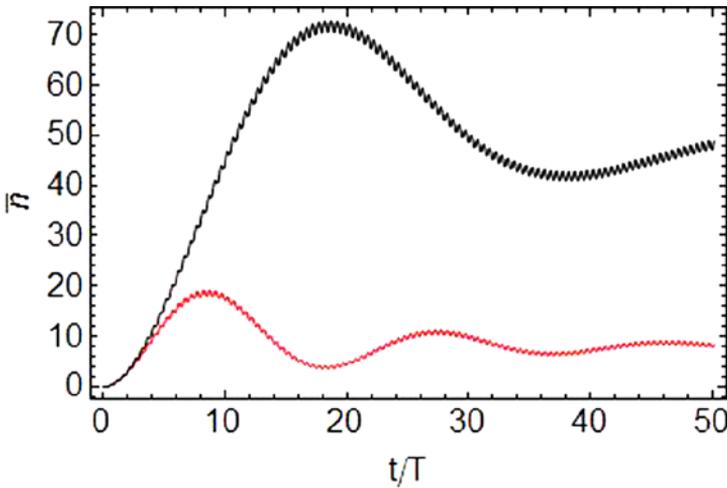


Рис. 2. Средняя энергия осциллятора в зависимости от состояний кубита: красная кривая – после действия на кубит 2π Раби-импульса, черная кривая – после воздействия на кубит π Раби-импульса. Параметры системы: $\omega_j = 1,03$, $\Omega = 1$, $\mu = 10^{-4}$, $f_0 = 0,3$, $\lambda = 0,02$, $\gamma = 0,01$ (в безразмерных единицах)

Подобные эксперименты над кубитами проводились в работе [Siddiqi et al., 2006] для калибровки измерительного прибора, а также для определения диапазона разрешения состояний кубита. Аналогичные зависимости разброса средних энергий осциллятора в зависимости от состояний кубита были промоделированы нами (см. рис. 3). Прямоугольные столбики на гистограмме рис. 3 отражают, какое количество «квантовых траекторий» (аналог одного акта измерения системы на опыте [Siddiqi et al., 2006]) попало в выделенный интервал энергий в момент прекращения действия считывающего импульса $t^* = 50T$ (конец кривых на рис. 2), а сплошные кривые – аппроксимирующие функции распределения, построенные по статистическим данным. За счет того, что состояния хорошо разнесены по энергиям (функции распределения не перекрываются), то во время энергетического скачка кубита энергия осциллятора изменится на значительную величину $\delta\bar{n} = \bar{n}_\uparrow - \bar{n}_\downarrow \sim 40$ квантов, что и регистрируется на опыте. Следовательно, квантовый метод Монте-Карло позволяет напрямую промоделировать результаты экспериментов [Siddiqi et al., 2006; Vijay et al., 2011], а также подобрать рабочую точку и характеристики импульсов для измерительного прибора и кубита.

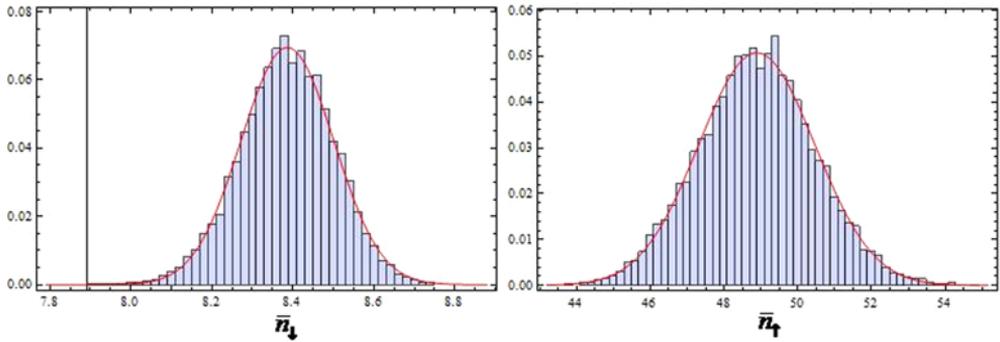


Рис. 3. Гистограмма распределения средних энергий (б) \bar{n}_\downarrow (черная кривая на рис. 2) и (а) \bar{n}_\uparrow (красная кривая на рис. 2) осциллятора в момент времени $t^* = 50T$ для $M = 10000$. Параметры системе аналогичны тем, что представлены на рис. 2

СХЕМА РАСПАРАЛЛЕЛИВАНИЯ И МАСШТАБИРУЕМОСТЬ

Программная реализация описанного выше метода квантовых траекторий выполнена на языке программирования C использованием технологий MPI [Гергель, 2010] и CUDA [Боресков, Харламов, 2010; Сандерс, Кэндрот, 2011]. Основные используемые структуры данных – массив для хранения коэффициентов гамильтониана связанной системы «кубит+осциллятор», массивы для хранения характеристик усредненной траектории. Распределение вычислительной нагрузки между узлами кластера выполнено с использованием технологии MPI. На каждом узле кластера для расчета траекторий используются возможности графических ускорителей. Поскольку расчет отдельных квантовых траекторий является независимым, каждая из них рассчитывается отдельным потоком GPU, при этом один поток GPU может последовательно рассчитать несколько траекторий. Незначительная доля последовательного кода не оказывает существенного влияния на время выполнения.

На первом шаге алгоритма выполняются инициализация MPI, выделение памяти на GPU, расчет коэффициентов гамильтониана связанной системы «кубит+осциллятор». Для хранения данных на GPU в текущей реализации используется только глобальная память.

На втором шаге алгоритма выполняется расчет на GPU отдельных траекторий и сбор статистических характеристик траекторий на каждом GPU. При расчете диссипативной динамики системы использовался метод Рунге-Кутты 4 порядка. Формирование массива случайных чисел, необходимого для реализации скачков, выполнялось посредством алгоритма MCG59. Независимость расчета отдельных траекторий позволяет равномерно загрузить все используемые для вычислений GPU.

На третьем шаге производится сбор статистических данных на нулевой процесс, их усреднение и вывод, а также освобождение выделенной памяти и деинициализация MPI.

Вычислительные эксперименты проводились на кластере «Лобачевский» Нижегородского государственного университета им. Н.И.Лобачевского. Мы использовали до 64 узлов следующей конфигурации: 2x Intel Xeon E5-2660 CPU (8 cores, 2.2 GHz), 64 GB RAM, NVIDIA Kepler K20X (2 или 3), OS Linux (CentOS 6.4), Intel C/C++ Compiler, Intel MPI, CUDA 6.4 Toolkit. На каждом узле использовались два графических сопроцессора.

При проведении экспериментов рассчитывалась система с $N = 50$ уровнями, время расчета $t^* = 50T$, шаг по времени при расчете скачков $dt = 0,02$, шаг по времени при расчете методом Рунге-Кутты $\Delta t_{RK} = 0,000667$, число траекторий M от 4 000 до 128 000.

Таблица 1. Время расчета и ускорение алгоритма для расчета физических величин при усреднении по $M = 128\ 000$ квантовых траекторий.

Число GPU устройств	Время, с	Ускорение относительно 4 GPU
4	5898	1
8	3841	1,54
16	1291	4,57
32	1163	5,07
64	1054	5,60
128	1017	5,80

В таблице 1 приведены данные по ускорению в зависимости от использования нескольких графических устройств. Отметим, что при использовании большого числа графических ускорителей недостаточное количество траекторий на GPU не позволяет эффективно использовать его возможности, что наглядно видно из рис. 4, и начиная уже с 32 GPU устройств ускорения при вычислениях не наблюдается.

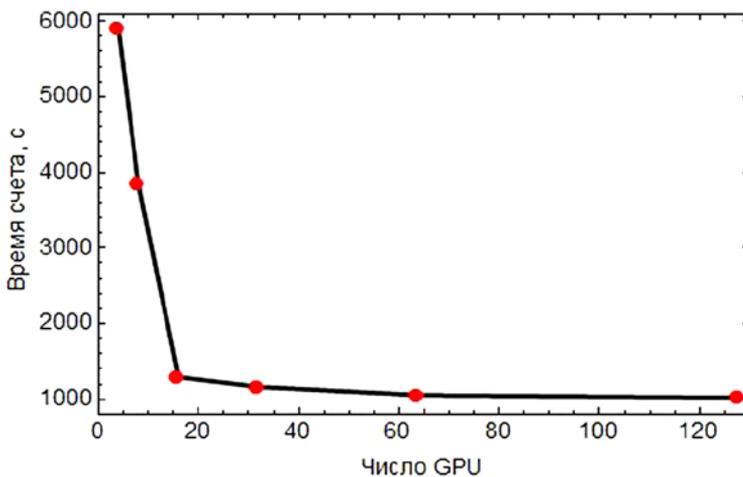
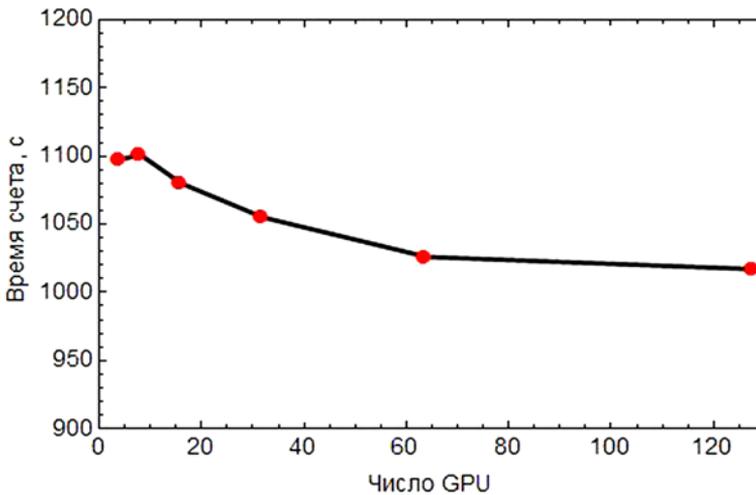


Рис.4. Время работы программы при фиксировании суммарного числа квантовых траекторий $M = 128\ 000$

Таблица 2. Время расчета при распределении по $M = 1\,000$ квантовых траекторий на каждый GPU ускоритель.

Число GPU устройств	Время, с
4	1097,25
8	1101,05
16	1080
32	1055
64	1026
128	1017

**Рис. 5.** Время работы программы при фиксированном числе квантовых траекторий $M = 1\,000$ на каждом GPU

Данные таблицы 2, изображенные на рис. 5, иллюстрируют абсолютную слабую масштабируемость реализации, ожидаемую при использовании метода Монте-Карло с независимыми испытаниями.

ЗАКЛЮЧЕНИЕ

В работе проведено численное моделирование диссипативной динамики джозефсоновского кубита, взаимодействующего с измерительным осциллятором, с использованием метода квантовых траекторий. Данный подход позволяет промоделировать напрямую результаты «однократных» неразрушающих измерений (single-shot measurement) [Neumann et al., 2010; Vamivakas et al., 2010; Vijay et al., 2011], то есть производить мониторинг состояний системы в реальном времени для одного акта измерения. Приведены результаты исследования возбуждения системы двумя типами Раби-импульсов: коротким и длительным в единичных реализациях и для усредненной по ансамблю динамики. Показано, что поведение кубита в этих случаях существенно различаются. В случае усреднен-

ной динамики для длительного Раби-импульса населенности уровней системы выходят на насыщение $P \sim 0,5$, в отличие от единичных реализаций, где можно проследить несколько типов квантовых скачков: фазовый, энергетический и термический (взаимодействие с термостатом). Результаты численного моделирования качественно согласуются с результатами эксперимента [Vijay et al., 2011]. Если проводить аналогию с измерениями в квантовой оптике, например, измерением резонансной флуоресценции атомов, то метод квантовых траекторий позволяет получить информацию о системе не только в среднем по ансамблю, но и промоделировать поведение каждого атома в отдельности.

В силу того, что расчет квантовых траекторий выполняется независимо, данный метод допускает естественное распараллеливание, позволяющее реализовать расчеты траекторий в виде отдельных потоков. Основываясь на этом свойстве, моделирование диссипативной динамики в данной работе производился с использованием графических ускорителей и технологии CUDA, что позволило достичь заметного ускорения при расчете усредненных характеристик.

ЛИТЕРАТУРА

- [Боресков, Харламов, 2010] *Боресков А.В., Харламов А.А.* Основы работы с технологией CUDA // М.: ДМК Пресс, 2010. 232 с.
- [Гергель, 2010] *Гергель В.П.* Теория и практика параллельных вычислений // М.: Интернет-Университет Информационных Технологий; Бином. Лаборатория знаний, 2010, 423 с.
- [Нильсен, Чанг, 2006] *Нильсен М., Чанг И.* Квантовые вычисления и квантовая информация // М.: Мир, 2006. 824 с.
- [Сандерс, Кэндрот, 2011] *Сандерс Д., Кэндрот Э.* Технология CUDA в примерах. Введение в программирование графических процессов: Пер. с англ. Слинкина А.А. // М.: ДМК Пресс, 2011. 232 с.
- [Скалли, Зубайри, 2003] *Скалли М.О., Зубайри М.С.* Квантовая оптика // М.: Физматлит, 2003. 512 с.
- [Braginsky, Khalili, 1995] *Braginsky V.B., Khalili F.Ya.* Quantum measurement // Cambridge University Press, 1995. 191 p.
- [Hatridge et al., 2011] *Hatridge M., Vijay R., Slichter D.H., Clarke J., Siddiqi I.* Dispersive magnetometry with a quantum limited SQUID parametric amplifier // Phys. Rev. B. 2011. V. 83. P. 134501.
- [Ladd et al., 2010] *Ladd T.D., Jelezko F., Laflamme R., Nakamura Y., Monroe C., O'Brien J.L.* Quantum computers // Nature. 2010. V. 464. P. 45.
- [Molmer et al., 1992] *Molmer K., Castin Y., Dalibard J.* Wave-function approach to dissipative processes in quantum optics // Phys. Rev. Lett. 1992. V. 68. P. 580.
- [Plenio, Knight, 1998] *Plenio M.B. and Knight P.L.* The quantum-jump approach to dissipative dynamics in quantum optics // Rev. Mod. Phys. 1998. V. 70. P. 101.
- [Neumann et al., 2010] *Neumann P., Beck J., Steiner M., Rempp F., Fedder H., Hemmer P., Wrachtrup J., Jelezko F.* Single-shot readout of a single nuclear spin // Science. 2010. V. 329. P. 542.
- [Siddiqi et al., 2004] *Siddiqi I., Vijay R., Pierre F., Wilson C.M., Metcalfe M., Rigetti C., Frunzio L., Devoret M.H.* RF-Driven Josephson Bifurcation Amplifier for Quantum Measurement // Phys. Rev. Lett. 2004. V. 93. P. 207002.

- [Siddiqi et al., 2006] *Siddiqi I., Vijay R., Metcalfe M., Boaknin E., Frunzio L., Schoelkopf R.J., Devoret M.H.* Dispersive measurements of superconducting qubit coherence with a fast latching readout // *Phys. Rev. B.* 2006. V. 73. P. 054510.
- [Vamivakas et al., 2010] *Vamivakas A.N., Lu C.-Y., Matthiesen C., Zhao Y., Fält S., Badolato A., Atatüre M.* Observation of spin-dependent quantum jumps via quantum dot resonance fluorescence // *Nature.* 2010. V. 467. P. 297.
- [Vijay et al., 2011] *Vijay R., Slichter D.H., Siddiqi I.* Observation of Quantum Jumps in a Superconducting Artificial Atom // *Phys. Rev. Lett.* 2011. V. 106. P. 110502.
- [Vion et al., 2002] *Vion D., Aassime A., Cottet A., Joyez P., Pothier H., Urbina C., Esteve D., Devoret M.H.* Manipulating the Quantum State of an Electrical Circuit // *Science.* 2002. V. 296. P. 886.
- [You, Nori, 2011] *You J.Q., Nori F.* Atomic Physics and Quantum Optics using Superconducting circuits // *Nature.* 2011. V. 474. P. 589.

QUANTUM TRAJECTORIES AND TOMOGRAPHY STATES OF THE COUPLED SYSTEM QUBIT-MEASURING DEVICE

V.P. Gergel, M.V. Denisenko, A.V. Liniov, D.S. Pashin, A.M. Satanin

Lobachevsky State University of Nizhny Novgorod (NNGU), Nizhny Novgorod

The algorithm based on quantum trajectory technique (Monte Carlo wave-function method) for modeling dissipative dynamics of multi-level quantum system was developed. The simulation of the interaction process a qubit with a quantum nonlinear oscillator based on the developed method has been conducted. It is shown that the measuring result for the average energy (number of photons) of the oscillator gives a possibility to carry out separation (tomography) of qubit states both as an average over an ensemble and as for a single trajectory. Monte Carlo wave-function method allows for effective parallelization on heterogeneous high performance computing clusters in calculation of experimentally observed values in multi-level quantum systems with relaxation processes. The basic principles of parallelism algorithm for solving stochastic Schrödinger equation, which were implemented by us in program complex, allowing to carry out high performance parallel calculations using GPU have been described. The efficiency of graphics accelerators, scalability of the method was demonstrated and possibility of interconnection of multiple graphics cards using MPI-technology was described.

Keywords: qubit, non-destructive measurements, quantum trajectory technique (Monte Carlo wave-function method), GPU, CUDA.

Gergel Victor Pavlovich – doctor of technical sciences, professor, dean@vmk.unn.ru

Denisenko Marina Valerievna – junior researcher, mar.denisenko@gmail.com

Liniov Aleksey Vladimirovich – chief of laboratory, alin@unn.ru

Pashin Dmitry Sergeevich – undergraduate, agent150@rambler.ru

Arkady Mikhailovich – professor, doctor of physical and mathematical sciences, sarkady@mail.ru

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ НЕЛИНЕЙНЫХ ВОЛН НА ГЛУБОКОЙ ВОДЕ

А.С. Досаев, Ю.И. Троицкая

Институт прикладной физики Российской академии наук (ИПФ РАН)

Подробная информация о зависимости характеристик коротких капиллярных волн от характеристик более крупномасштабного волнения необходима для интерпретации данных дистанционного зондирования поверхности океана. В данной работе методами численного моделирования изучается процесс генерации дециметровыми волнами «паразитной» капиллярной ряби, а также исследуется механизм каскадной модуляции ряби длинными волнами. Численная модель основана на полностью нелинейных уравнениях движения для двумерного потенциального течения со свободной поверхностью и использует отслеживающие форму границы жидкости конформные координаты.

Работа выполнена при поддержке гранта РФФИ № 14-17-00667 (Троицкая Ю.И.), а также грантов РФФИ №№ 13-05-00865, 14-05-91767 АФ_а, 14-05-31415 мол_а, и гранта Президента МК. 3550.2014.5 (Досаев А.С.).

Ключевые слова: численное моделирование, волны на воде, gprgr.

ВВЕДЕНИЕ

При построении радиоизображения поверхности океана полезный сигнал определяется характеристиками поверхностных волн сантиметрового диапазона, поэтому понимание механизмов модуляции распределения коротких волн в поле течений длинных волн важно для интерпретации данных дистанционного зондирования [Hasselmann et al., 1985]. Кроме того, передача импульса от атмосферы океану происходит в значительной мере за счет взаимодействия ветра с коротковолновыми компонентами волнения [Miles, 1957; Филипс, 1980]. Существование механизма, в результате действия которого распределение коротких волн становится зависимым от фазы длинной волны, может играть важную роль для процесса генерации длинных волн [Longuet-Higgins, 1969].

Трансформация дециметровых волн на неоднородном течении удовлетворительно описывается в рамках уравнений кинематической теории и закона сохранения волнового действия [Bretherton, Garrett, 1968]. Такая «двухмасштабная» (т.е. учитывающая лишь взаимодействие длинной и короткой волн) модель поверхностного волнения используется для расчета модуляционной передаточной функции в [Keller, Wright, 1975]. Двухмасштабная модель, однако, не позволяет объяснить особенности модуляции сантиметровых волн; в связи с этим в [Lyzenga, Bennett, 1988] предлагается учитывать наличие в системе промежуточ-

Досаев Александр Сергеевич – Институт прикладной физики Российской академии наук, младший научный сотрудник, dosaev@hydro.appl.sci-nnov.ru

Троицкая Юлия Игоревна – Институт прикладной физики Российской академии наук, заведующая отделом, доктор физико-математических наук, yuliya@hydro.appl.sci-nnov.ru

ного масштаба – волн с длиной больше брэгговской, но тем не менее коротких по сравнению с длинной волной. Использованная для расчета модуляционной передаточной функции в [Romeiser et al., 1994] «трехмасштабная» модель учитывает модуляцию и асимметричное распределение брэгговских волн на волнах промежуточной длины.

В работе [Ермаков, Салашин, 1994] экспериментально, а в [Charnotskii et al., 2002] численно в рамках приближенной теории Лонге-Хиггинса [Longuet-Higgins, 1995] исследуется влияние крупномасштабного течения на распределение капиллярной ряби, возникающей на склонах дециметровых волн. Известно, что крутые дециметровые волны на поверхности жидкости могут генерировать на своем переднем склоне квазистационарную относительно гребня картину «паразитной» капиллярной ряби. В работах Лонге-Хиггинса [Longuet-Higgins, 1963; Longuet-Higgins, 1995] описан механизм этого явления и дано детальное объяснение формы генерируемого цуга капиллярных волн. С увеличением амплитуды волны Стокса на гребне возникает узкая область, где кривизна поверхности значительно превышает значения кривизны для остальных участков профиля. Учитывая действие сил поверхностного натяжения по теории возмущений, мы можем считать, что к поверхности жидкости приложено давление $p = T\kappa$, где T – коэффициент поверхностного натяжения, κ – кривизна поверхности волны Стокса. Тогда связанное с гребнем узкое распределение давления можно рассматривать как движущийся источник, который при наличии фазового синхронизма может излучать волны. Так как групповая скорость капиллярных волн превышает фазовую, капиллярная рябь генерируется на переднем склоне основной волны. Амплитуда генерируемой ряби чувствительна к изменениям амплитуды основной волны. Поскольку цуг гравитационно-капиллярных волн на поверхности длинной волны модулируется таким образом, что на гребне длинной волны короткие волны становятся круче, чем у подошвы, амплитуда ряби также зависит от фазы длинной волны.

Настоящая работа посвящена исследованию трехмасштабного или «каскадного» механизма модуляции коротких поверхностных волн в присутствии длинных волн методами численного моделирования в рамках точной системы уравнений гидродинамики, выраженных в криволинейных конформных координатах. Программа для численного интегрирования уравнений движения была написана с использованием интерфейса OpenCL и проводит вычисления на графическом процессоре.

2. ОПИСАНИЕ МОДЕЛИ

Для описания динамики нелинейных поверхностных волн на глубокой воде будем использовать уравнения эволюции двумерного потенциального течения невязкой жидкости со свободной поверхностью. В области течения поле скоростей определяется потенциалом, который в силу несжимаемости жидкости удовлетворяет уравнению Лапласа:

$$\dot{V} = \nabla \varphi, \Delta \varphi = 0.$$

На границе жидкости должно выполняться динамическое граничное условие

$$\varphi_t + \frac{|\nabla \varphi|^2}{2} + p + gy = const \tag{1}$$

где x, y – горизонтальная и вертикальная декартовы координаты, p – давление, а также кинематическое граничное условие, требующее, чтобы граница жидкости была жидкой поверхностью. Если возвышение поверхности $\eta(x, t)$ не является многозначной функцией горизонтальной координаты x , кинематическое граничное условие может быть записано как

$$\eta_t + \varphi_x \eta_x = \varphi_y. \tag{2}$$

Состояние системы полностью определяется формой свободной поверхности и поверхностным распределением потенциала.

Распространенным методом решения задачи со свободной поверхностью является использование конформных координат, «выпрямляющих» границу. При этом область течения в каждый момент времени представляется конформным образом нижней полуплоскости комплексного переменного $w = u + iv$:

$$x + iy = z(u + iv, t). \tag{3}$$

В конформных координатах (u, v) потенциал φ по-прежнему удовлетворяет уравнению Лапласа; решение этого уравнения в нижней полуплоскости тривиально.

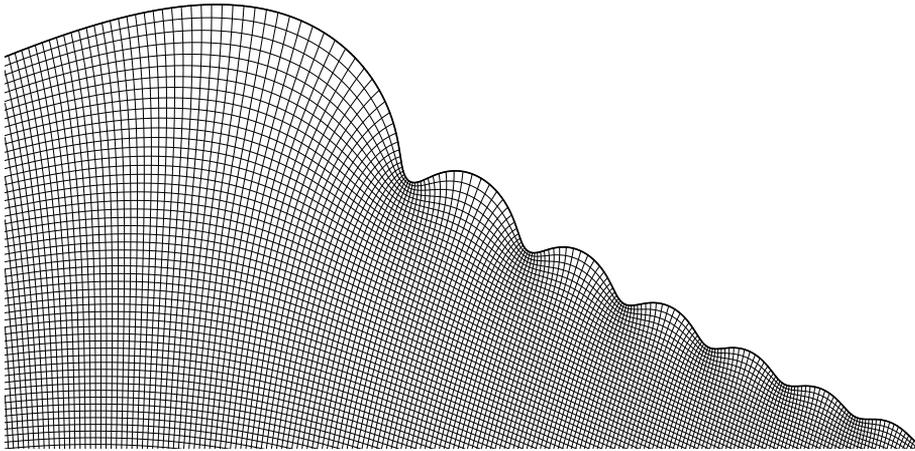


Рис. 1. Рябь на переднем склоне гравитационно-капиллярной волны и конформная координатная сетка

Действительная и мнимая части на действительной оси связаны преобразованием Гильберта \hat{H} :

$$z(u) = u + \tilde{x}(u) + iy(u), \tilde{x}(u) = -\hat{H}y(u).$$

Похожим образом конструируется комплексный потенциал θ :

$$\theta = \varphi + i\xi, \xi(u) = \hat{H}\varphi(u)$$

– получаемая в соответствии с этим определением функция $\theta(u)$ имеет аналитическое продолжение на всю область течения и удовлетворяет условию затухания связанных с волновым движением компонент скорости на большой глубине.

Уравнения движения, описывающие изменение во времени функций $z(u)$ и $\theta(u)$, могут быть переписаны в переменных Дьяченко

$$R = \frac{1}{z'}, \quad V = \frac{i\theta'}{z'}, \quad (4)$$

что приводит к системе [Zakharov et al., 2002]

$$\begin{aligned} R_t &= i(FR' - F'R), \\ V_t &= i(FV' - \hat{P}'(|V|^2 + 2p)) + g(R - 1), \end{aligned} \quad (5)$$

где

$$F = \hat{P}(\bar{R}V + R\bar{V}), \quad P = \frac{1}{2}(1 + i\hat{H}).$$

Как показано в [Dyachenko, 2004], эта форма уравнений более устойчива, чем уравнения в переменных z и θ , и больше подходит для численного интегрирования.

Капиллярное слагаемое в поверхностном давлении выражается через переменные Дьяченко как

$$p_c = 2T \operatorname{Im}(\bar{Q}Q'), \quad Q = \sqrt{R},$$

где T – коэффициент поверхностного натяжения. Эффект вязкого затухания волн моделируется введением добавки к давлению

$$p_v = -4\nu \frac{\partial^2 \varphi}{\partial l^2} = -4\nu |R| \frac{\partial}{\partial u} \operatorname{Im} \left(\frac{|R|}{R} V \right),$$

где ν – коэффициент вязкости, l – натуральный параметр на свободной поверхности.

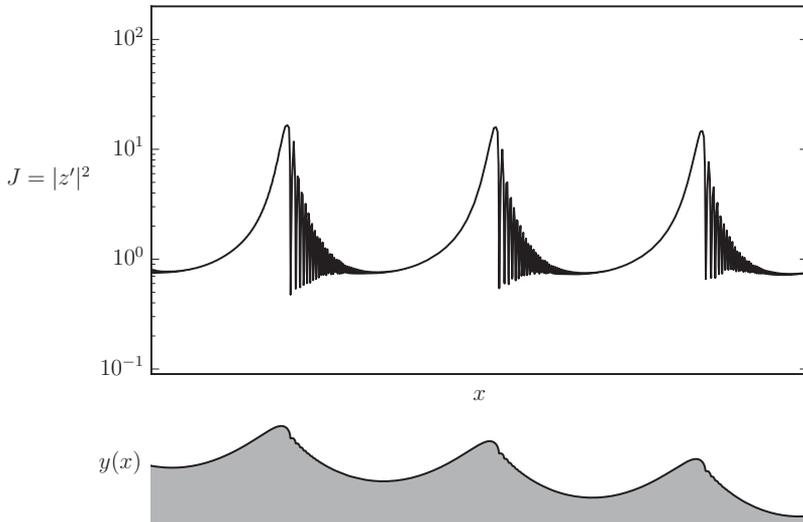


Рис. 2. Значения якобиана конформного отображения в различных точках поверхности

Форма данных уравнений движения такова, что вычисление правых частей (5) сводится к последовательности преобразований Фурье и локальных операций над функциями в фурье- либо в координатном представлении.

На однородной по конформной координате координатной сетке можно использовать алгоритм быстрого преобразования Фурье (БПФ). Недостатком такого подхода является то, что точки сетки, однородной по u , неоднородно распределены по длине профиля поверхности; точки «разбегаются» с гребней волн и собираются у подошвы – это заметно из рис. 1 и наглядно иллюстрируется графиком якобиана отображения $z(w)$ на рис. 2. При моделировании очень крутых и обрушивающихся волн авторами наблюдалось еще более сильное растяжение сетки, с $J \sim 10^4$, что соответствует падению пространственного разрешения на отдельных участках в $\sqrt{J} \sim 10^2$ раз. В рассматриваемой задаче подобных ситуаций не возникало, и необходимость компенсировать растяжение увеличением числа узлов сетки не приводила к серьезному росту времени вычислений. С другой стороны, БПФ эффективно распараллеливается и позволяет в полной мере использовать преимущества архитектуры графических процессоров.

3. ГЕНЕРАЦИЯ КАПИЛЛЯРНОЙ РЯБИ КРУТОЙ ВОЛНОЙ СТОКСА

Проиллюстрируем описанные во введении особенности генерации ряби результатами численных экспериментов. В ходе серии экспериментов моделировалась эволюция волн Стокса длиной $\lambda = 10$ см различной начальной амплитуды. В отсутствие вязкого затухания и эффектов поверхностного натяжения эти волны распространялись бы без изменения формы. Действие сил поверхностного натяжения приводит к тому, что на переднем склоне волны начинает формироваться картина капиллярной ряби (см. рис. 3). Декремент вязкого затухания для ряби значительно выше, чем для основной волны; откачка энергии паразитными капиллярными волнами продолжается до тех пор, пока основная волна не становится слишком полой, чтобы эффективно генерировать рябь.

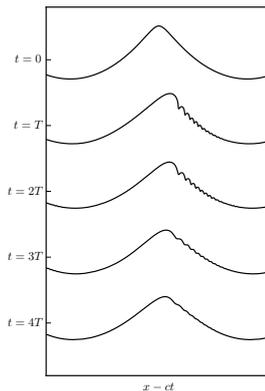


Рис. 3. Эволюция волны Стокса длиной 10 см при включении капиллярных эффектов (T – период волны Стокса, вычисляемый по линейному дисперсионному соотношению, c – фазовая скорость)

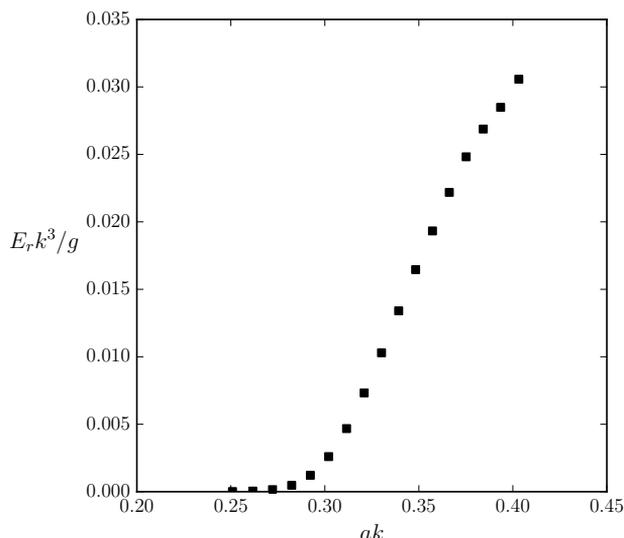


Рис. 4. Безразмерная энергия капиллярной ряби в зависимости от начальной крутизны волны Стокса (k – волновое число волны Стокса)

Характерная длина капиллярных волн в данных экспериментах составляла порядка 0,5 см. Будем определять запасенное в капиллярных волнах количество энергии как разность между полной энергией системы и энергией первых 10 гармоник основной волны:

$$E_r = E - E_{10}$$

На рис. 4 приведена зависимость максимального наблюдавшегося за время моделирования количества энергии капиллярных волн в зависимости от крутизны начальной волны Стокса. Как можно видеть, существует пороговое значение крутизны (в данном случае – $ak \sim 0,3$), ниже которого капиллярные волны практически не генерируются.

4. МОДЕЛИРОВАНИЕ КАСКАДНОГО МЕХАНИЗМА МОДУЛЯЦИИ РЯБИ

Для экспериментального наблюдения каскадной модуляции ряби начальные условия задавались в виде длинной и пологой стационарной волны ($\lambda_0 = 6$ м, $a_0 k_0 \sim 0,06$), по поверхности которой распространяется цуг коротких и крутых гравитационно-капиллярных волн ($\lambda_s \sim 15$ см, $a_s k_s \sim 0,3$). Распределения волнового числа и амплитуды коротких волн в начальный момент времени задавались таким образом, чтобы в ходе дальнейшей эволюции амплитуда и волновое число оставались приблизительно постоянными в фиксированной фазе длинной волны. Эти стационарные распределения находились из условий постоянства вдоль поверхности длинной волны 1) частоты коротких волн, наблюдаемой в связанной с длинной волной системе отсчета:

$$\Omega + k_s U = const \tag{6}$$

и 2) потока волнового действия:

$$(c_g + U) \frac{g^* a_s^2}{2\Omega} = const. \quad (7)$$

Здесь U – скорость течения на поверхности длинной волны в связанной с волной системе отсчета, c_g – групповая скорость коротких волн,

$$\Omega(k_s, l) = \sqrt{g^* k_s} \quad (8)$$

– внутренняя частота коротких волн (приближенно задается линейным дисперсионным соотношением),

$$g^*(l) = g \cos \alpha - \kappa U^2 \quad (9)$$

– эффективное ускорение свободного падения [Phillips, 1981], α – угол наклона невозмущенной поверхности длинной волны, l – натуральный параметр на невозмущенном профиле длинной волны.

После того, как из уравнений (6), (7) найдены зависимости $k_s(l)$, $a_s(l)$, к координатам свободной поверхности добавляется возмущение, связанное с волной Стокса:

$$z = z_0 + \hat{P}z_s(l(u)),$$

где

$$z_s(l) = \sum_{n \geq 0} \frac{A_n(a_s k_s)}{k_s} \exp(-in \int k_s dl).$$

Применение оператора проекции \hat{P} мотивировано тем, что после наложения на координаты возмущения z_s функция $z(w)$ должна остаться аналитической в нижней полуплоскости. Зависимость безразмерных амплитуд A_n гармоник волны Стокса от крутизны $a_s k_s$ находится из заранее составленной таблицы. Похожие изменения вносятся в поле скоростей.

Решение уравнений (6), (7) таково, что короткие волны короче и круче на гребне длинной волны, где они сжимаются течением, чем у подошвы – это утверждение остается верным также и для дальнейшей эволюции системы. На установление картины ряби требуется время порядка 0,2 с. На рис. 5 приведен оконный спектр возвышения поверхности $y(k, x)$ на момент времени 1 с после начала счета (использовалось гауссово окно с шириной, равной периоду длинной волны λ_0 , и среднеквадратичным отклонением $\sigma = 0,05\lambda_0$). Капиллярной ряби на спектре соответствуют волновые числа $k/k_0 \sim 10^3$ и больше. В этот и последующие моменты времени распределение ряби следует за гребнем длинной волны, медленно затухая по мере уменьшения амплитуды дециметровых волн.

В спектре на рис. 5 (во всяком случае в длинноволновой области) различимы полосы, соответствующие гармоникам волны Стокса. Поскольку период волны Стокса изменяется в зависимости от фазы длинной волны, подсчет коэффициента модуляции гармоник непосредственно по формуле

$$I(k) = \frac{a_{\max}(k) - a_{\min}(k)}{a_{\max}(k) + a_{\min}(k)}$$

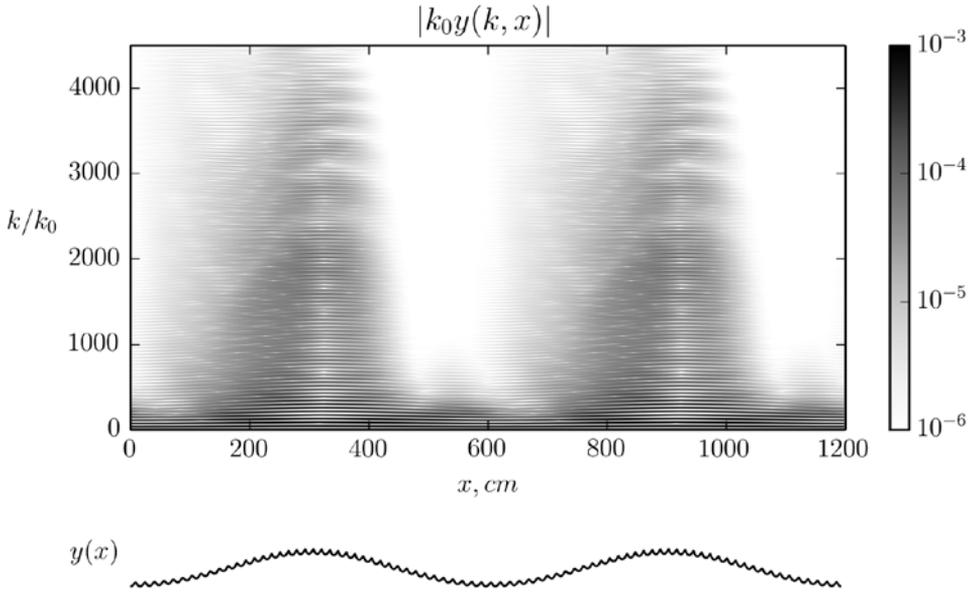


Рис 5. Оконный спектр возвышения поверхности

где a_{\min} , a_{\max} – минимальное и максимальное вдоль всей поверхности длиной волны значения $|y(x, k)|$ при фиксированном заданном k , дает близкие к единице значения практически для всех волновых чисел. Для малых k большие значения коэффициента модуляции будут обусловлены тем, что полосы, соответствующие гармоникам волны Стокса $k = nk_s$, значительно отклоняются от любой линии $k = \text{const}$. Чтобы оценить коэффициент модуляции с учетом трансформации волны Стокса в поле течений длиной волны, амплитуды предварительно осредняются по прямоугольному окну:

$$a(k, x) = \sqrt{\frac{1}{2\Delta k / k_0} \sum_{k' \in [k - \Delta k, k + \Delta k]} |y(k', x)|^2}, \quad \Delta k = 0,2k.$$

Ширина окна выбрана таким образом, чтобы окно целиком охватывало диапазон изменения волнового числа в ближайшей полосе. На рис. 6 приведен график коэффициента модуляции, рассчитанного по усредненным значениям амплитуд.

В то время как основная гармоника $k = \langle k_s \rangle$ модулирована слабо (коэффициент модуляции порядка крутизны длиной волны $a_0 k_0$), коэффициент модуляции для высоких гармоник волны Стокса и вынужденных волн капиллярной ряби приближается к единице. Это соответствует представлению о том, что амплитуда генерируемой ряби зависит от значений кривизны на гребне волны Стокса, которая в свою очередь определяется амплитудами высоких гармоник. Резкое увеличение коэффициента модуляции в интервалах между $k / \langle k_s \rangle = 1, 2, 3, 4, 5$ можно объяснить тем, что при малых значениях полезного сигнала на коэффициент модуляции влияют преимущественно шумы.

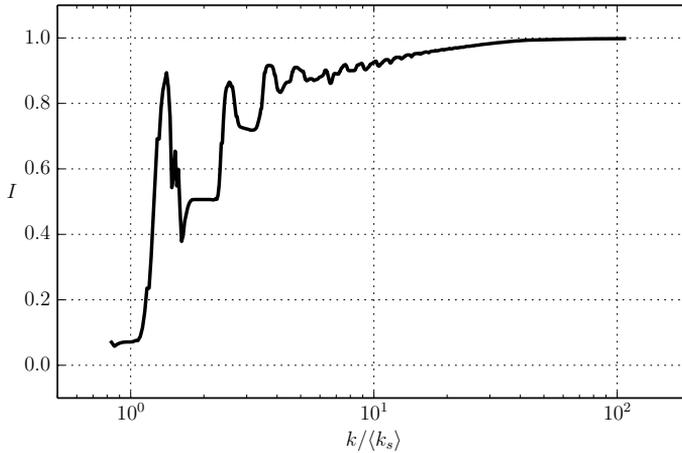


Рис. 6. Коэффициент модуляции в зависимости от номера гармоники волны Стокса

5. ЗАКЛЮЧЕНИЕ

Проведено численное моделирование процесса генерации «паразитных» капиллярных волн крутыми дециметровыми волнами. Продемонстрировано наличие порога по крутизне основной волны, ниже которого капиллярная рябь на ней практически не генерируется. С использованием точных уравнений движения воспроизведено действие каскадного механизма модуляции капиллярной ряби, в том числе показано, что даже слабая модуляция цуга дециметровых волн пологой длинной волной может приводить к возникновению сильного контраста между гребнем и подошвой длинной волны в распределении капиллярной ряби.

ЛИТЕРАТУРА

- [Ермаков, Салашин, 1994] *Ермаков С.А., Салашин С.Г.* Об эффекте сильной модуляции капиллярно-гравитационной ряби внутренними волнами // Доклады Академии наук. 1994. Т. 337. № 1. С. 108-111.
- [Филипп, 1980] *Филипп О.М.* Динамика верхнего слоя океана. Л.: Гидрометеиздат, 1980, 320 с.
- [Bretherton, Garrett, 1968] *Bretherton F.P., Garrett C.J.R.* Wavetrains in inhomogeneous moving media // Proc. R. Soc. Lond. A. 1968. V. 302. P. 529-554.
- [Charnotskii et al., 2002] *Charnotskii M., Naugolnykh K., Ostrovsky L., Smirnov A.* On the cascade mechanism of short surface modulation // Nonlinear Processes in Geophysics. 2002. V. 9. P. 281-288.
- [Dyachenko, 2004] *Dyachenko A.I.* // Proceedings of the II International conference “Frontiers of Nonlinear Physics”, N.Novgorod-St.Petersburg, Russia, July 5-12, 2004.
- [Hasselmann et al., 1985] *Hasselmann K., Raney R.K., Plant W.J., Alpers W., Shuchman R.A., Lyzenga D.R., Rufenach C.L., Tucker M.J.* Theory of synthetic aperture radar ocean imaging: a MARSEN view // J. Geophysical Research. 1985. V. 90. N. C3. P. 4659-4686.

- [Keller, Wright, 1975] *Keller W.C., Wright J.W.* Microwave scattering and the straining of wind-generated waves // *Radio Science*. 1975. V. 10. N. 2. P. 139-147.
- [Longuet-Higgins, 1963] *Longuet-Higgins M.S.* The generation of capillary waves by steep gravity waves // *J. Fluid Mechanics*. 1963. V. 16. P. 138-159.
- [Longuet-Higgins, 1969] *Longuet-Higgins M.S.* A nonlinear mechanism for the generation of sea waves // *Proc. R. Soc. Lond. A*. 1969. V. 311. P. 371-389.
- [Longuet-Higgins, 1995] *Longuet-Higgins M.S.* Parasitic capillary waves: a direct calculation // *J. Fluid Mechanics*. 1995. V. 301. P. 79-107.
- [Lyzenga, Bennett, 1988] *Lyzenga D.R., Bennett J.R.* Full-spectrum modeling of synthetic aperture radar internal wave signatures // *J. Geophysical Research*. 1988. V. 93. N. C10. P. 12.345-12.354.
- [Miles, 1957] *Miles J.W.* On the generation of surface waves by shear flows // *J. Fluid Mech*. 1957. V. 3. N. 2. P. 185-204.
- [Phillips, 1981] *Phillips O.M.* The dispersion of short wavelets in the presence of a dominant long wave // *J. Fluid Mech*. 1981. V. 107. P. 465-485.
- [Romeiser et al., 1994] *Romeiser R., Schmidt A., Alpers W.* A three-scale composite surface model for the ocean wave-radar modulation transfer function // *J. Geophysical Research*. 1994. V. 99. N. C5. P. 9785-9801.
- [Zakharov et al., 2002] *Zakharov V.E., Dyachenko A.I., Vasilyev O.A.* New method for numerical simulation of a nonstationary potential flow of incompressible fluid with a free surface // *Eur. J. Mech. B/Fluids*. 2002. V. 21. P. 283-291.

NUMERICAL SIMULATION OF NONLINEAR DEEP WATER WAVES DYNAMICS

A.S. Dosaev, Y.I. Troitskaya

Institute of Applied Physics of the Russian Academy of Sciences (IAP RAS)

Detailed knowledge of how short capillary waves characteristics depend on parameters of underlying long waves is important for making a correct interpretation of ocean surface remote sensing data. "Parasitic" capillary ripples generation by decimeter waves, as well as the cascade mechanism of capillary ripples modulation by long waves is studied by means of numerical simulation. Numerical model used is based on a fully nonlinear system of equations describing dynamics of a two-dimensional potential flow with a free surface.

Keywords: numerical modeling, water waves, gpgpu

Dosaev Aleksandr Sergeevich – Institute of Applied Physics of the Russian Academy of Sciences, junior researcher, dosaev@hydro.appl.sci-nnov.ru

Troitskaya Yuliya Igorevna – Institute of Applied Physics of the Russian Academy of Sciences, head of department, doctor of physical and mathematical sciences, yuliya@hydro.appl.sci-nnov.ru

ПРЯМОЕ ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ВЗАИМОДЕЙСТВИЯ МЕЛКОМАСШТАБНОЙ ТУРБУЛЕНТНОСТИ С ВНУТРЕННИМИ ВОЛНАМИ

О.А. Дружинин¹, Л.А. Островский², Ж. Фэнь³

¹ Институт прикладной физики Российской академии наук (ИПФ РАН)

² Лаборатория исследований окружающей среды NOAA, г. Булдер, Колорадо, США

³ Китайский Научно-Исследовательский Центр Кораблестроения, г. Вукси, Жианг-Су, КНР

Взаимодействие мелкомасштабной турбулентности с внутренними волнами (ВВ) играет важную роль в перемешивании и динамике верхнего слоя океана. Среди многочисленных аспектов этого взаимодействия можно выделить как эффект затухания ВВ под воздействием турбулентности, так и возможность усиления турбулентности в поле достаточно сильных ВВ. В данной работе эти процессы исследуются с помощью метода прямого численного моделирования.

Работа выполнена при поддержке РФФИ (№№ 14-05-00367, 15-05-02430).

Ключевые слова: численное моделирование, внутренние волны, мелкомасштабная турбулентность, верхний слой океана.

1. ВВЕДЕНИЕ

Взаимодействие мелкомасштабной турбулентности с внутренними волнами (ВВ) оказывает непосредственное влияние на динамику сезонного пикноклина и перемешивание в верхнем слое океана [Phillips, 1977]. Среди многих аспектов этого взаимодействия можно выделить как эффект затухания ВВ под действием турбулентности, так и возможность усиления турбулентности в поле сильных ВВ.

Эффект затухания ВВ под действием турбулентности исследовался в нашей предыдущей работе [Druzhinin et al., 2013] с помощью прямого численного моделирования. Результаты показывают, что если амплитуда ВВ достаточно мала (более чем в два раза меньше) по сравнению с амплитудой турбулентных пульсаций, происходит эффективное подавление ВВ турбулентностью. Однако оказалось также, что ВВ большой амплитуды слабо затухают на турбулентности. В этом случае, оценки декремента затухания ВВ с помощью общеизвестных параметризаций [Ostrovsky, Zaborskikh, 1996] дают существенно завышенные (по порядку величины) значения.

В данной работе рассматривается случай, когда исходная амплитуда ВВ больше или порядка амплитуды турбулентности, и исследуется возможность усиления мелкомасштабной турбулентности в поле ВВ.

Дружинин Олег Александрович – Институт прикладной физики РАН, в.н.с., д.ф.-м.н., druzhinin@hydro.appl.sci-nnov.ru

Островский Лев Аронович – Лаборатория исследований окружающей среды NOAA, профессор, г. Булдер, Колорадо, США, lev.a.ostrovsky@noaa.gov

Жао Фэнь – Китайский Научно-Исследовательский Центр Кораблестроения, профессор, г. Вукси, Жианг-Су, КНР, zhaofeng@cssrc.com.cn

2. ФОРМУЛИРОВКА ЗАДАЧИ

Рассматривается устойчиво-стратифицированная жидкость со ступенчатым вертикальным профилем плотности $\rho_0(z)$ (т.е. пикноклином, рис. 1). Исходная турбулентность локализована в слое над пикноклином. В начальный момент времени в пикноклине возбуждается первая мода монохроматической внутренней волны, распространяющейся слева направо вдоль пикноклина. Рассматриваются периодические граничные условия в горизонтальной плоскости (по координатам x и y) и условие Неймана (отсутствие потока) по вертикальной координате z . Толщина пикноклина и частота плавучести в

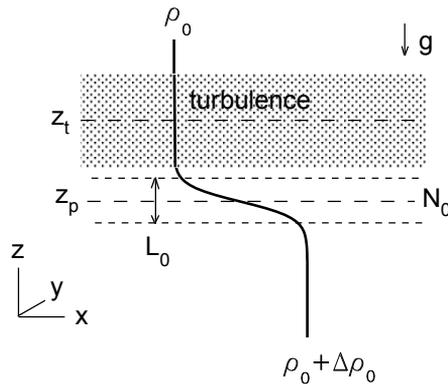


Рис. 1. Схема численного эксперимента

центре пикноклина $N_0 = \left(-\frac{g}{\rho_0} \frac{d\rho_0}{dz} \right)^{1/2}$, (где g – ускорение свободного падения), берутся в качестве нормировочных масштабов длины и времени, L_0 и $T_0=1/N_0$, которые используются для приведения уравнений динамики жидкости к безразмерному виду.

Численное моделирование основано на интегрировании полных трехмерных уравнений Навье-Стокса для несжимаемой жидкости, записанных в приближении Буссинеска. Уравнения записываются в безразмерных переменных [Монин, Яглом, 1992]:

$$\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{1}{\text{Re}} \frac{\partial^2 U_i}{\partial x_j^2} - \text{Ri} \delta_{iz} \rho \quad (1)$$

$$\frac{\partial U_j}{\partial x_j} = 0 \quad (2)$$

Уравнение для плотности имеет вид:

$$\frac{\partial \rho}{\partial t} + U_j \frac{\partial \rho}{\partial x_j} - U_z N_{ref}^2(z) = \frac{1}{\text{Re Pr}} \frac{\partial^2 \rho}{\partial x_j^2} \quad (3)$$

В (1)–(3) $U_i (i = x, y, z)$ – компоненты скорости, P и ρ – отклонения давления и плотности от гидростатических значений. Числа Рейнольдса и Ричардсона определяются как

$$\text{Re} = \frac{U_0 L_0}{\nu}, \text{Ri} = \left(\frac{L_0 N_0}{U_0} \right)^2 \quad (4)$$

Число Прандтля (соотношение кинематической вязкости и коэффициента молекулярной диффузии) $\text{Pr} = \nu/\kappa$ полагается равным единице. Переменные в (1)–(3) нормированы на масштабы длины и времени, L_0 и T_0 , и масштаб скорости $U_0 = L_0/T_0$. При такой нормировке число Ричардсона тождественно равно единице, $\text{Ri} = 1$. Отклонение плотности нормировано на скачок плотности в пикноклине, $\Delta\rho_0$ (рис. 1).

Исходный профиль частоты плавучести, $N_{ref}(z)$, задается равным

$$N_{ref}(z) = \frac{1}{\cosh 2(z - z_p)} \quad (5)$$

где z_p определяет положение центра пикноклина, и профиль безразмерной плотности $\rho_{ref}(z)$ определяется как:

$$\rho_{ref}(z) = \rho_{ref}(-\infty) - \int_{-\infty}^z N_{ref}^2(z) dz = \rho_{ref}(-\infty) - 0.5 \tanh 2(z - z_p) \quad (6)$$

с произвольной константой $\rho_{ref}(-\infty)$. Для удобства значение $\rho_{ref}(-\infty)$ задается равным 1,5. Т.о., исходный профиль плотности принимает вид:

$$\rho_{ref}(z) = 1 + 0.5 [1 - \tanh 2(z - z_p)] \quad (7)$$

Полное значение плотности равно сумме $\rho_{ref}(z)$ и отклонения плотности ρ .

Уравнения (1)–(3) интегрируются в прямоугольной области с размерами $0 \leq x \leq 40$, $-10 \leq y \leq 10$ и $0 \leq z \leq 20$ с помощью метода конечных разностей второго порядка точности на разнесенной («шахматной») сетке, состоящей из $400 \times 200 \times 200$ узлов по координатам x , y и z . Интегрирование проводится с использованием метода Адамса-Башфорфа с шагом по времени $\Delta t = 0,01$. Уравнение Пуассона для давления решается с помощью быстрого преобразования Фурье по координатам x и y , и метода Гаусса по координате z [Druzhinin et al., 2013]. Рассматривается число Рейнольдса $\text{Re} = 20000$.

3. ИНИЦИАЛИЗАЦИЯ ВНУТРЕННИХ ВОЛН

Начальные поля скорости и плотности задаются в виде решения для первой моды внутренней волны с длиной λ (и волновым числом $k = 2\pi/\lambda$) и частотой ω . Это решение получается из линеаризованной системы уравнений (1)–(3) для волны, распространяющейся слева направо вдоль пикноклина (оси x) [Phillips, 1977]:

$$U_x^{IW}(x, z, t) = -\frac{1}{k} \frac{dW(z)}{dz} \sin(kx - \omega t) \quad (8)$$

$$U_z^{IW}(x, z, t) = W(z) \cos(kx - \omega t) \quad (9)$$

$$\rho_z^{IW}(x, z, t) = \frac{W(z)}{\omega} \frac{d\rho_{ref}}{dz} \sin(kx - \omega t) \quad (10)$$

Начальные условия берутся при $t = 0$. Функция $W(z)$ определяется с помощью численного решения краевой задачи на собственные значения для уравнения Тэйлора-Гольдштейна [Phillips, 1977]:

$$\frac{d^2 W}{dz^2} + \left(\frac{N^2}{\omega^2} - 1 \right) k^2 W = 0 \quad (11)$$

с условиями $W(z) \rightarrow W_0 \exp[k(z - z_p)]$ при $z \ll z_p$, и $W(z) \rightarrow W_0 \exp[-k(z - z_p)]$ при $z \gg z_p$, где W_0 – амплитуда ВВ при $z = z_p$. Задача (11) решалась с помощью метода «пристрелки» со шшивкой при $z = z_p$ [Hazel, 1972]. Профиль скорости $W(z)$ и дисперсионное соотношение $\omega(k)$ для 1-й моды ВВ с длиной волны $\lambda = 10$ приведены на рис. 2а.

DNS проводилось для начальных условий (8)–(10) при $t = 0$, соответствующих ВВ с $\lambda = 10$ (частота $\omega = 0,489$, период $T \approx 13$). Амплитуда ВВ задавалась равной $W_0 = 0,1$. Исходное поле турбулентности при этом

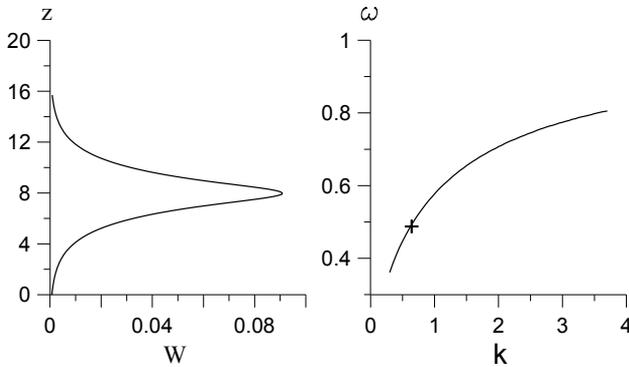


Рис. 2а. Собственная функция $W(z)$ краевой задачи (11) для 1-й моды ВВ с длиной $\lambda = 10$ (слева) и дисперсионное соотношение

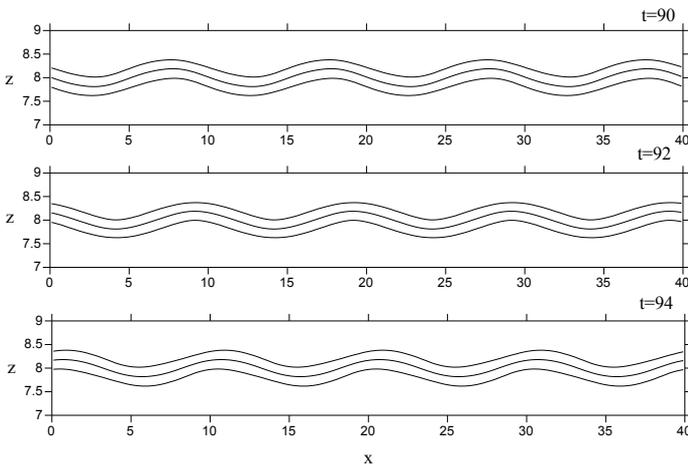


Рис. 2б. Изопикны в центральной плоскости при $y = 0$ в поле ВВ с длиной $\lambda = 10$ и амплитудой $W_0 = 0,1$

отсутствовало. Рисунок 2b показывает изопикны в различные моменты времени. Амплитуда смещения изопикн равна $a \approx 0,2$, т.о. крутизна волны равна $ka = 2\pi a/\lambda \approx 0,12$. Это значение может рассматриваться достаточно малым, что позволяет пренебречь нелинейными эффектами при распространении ВВ.

4. ИНИЦИАЛИЗАЦИЯ И ХАРАКТЕРИСТИКИ ТУРБУЛЕНТНОСТИ

Положение турбулизованного слоя задавалось равным $z_t = 10$, таким, что эффекты турбулентного перемешивания и генерации внутренних волн турбулентностью оказываются пренебрежимо малыми.

Исходное турбулентное поле скорости задавалось равным:

$$U_i(x, y, z) = U_{i0} U_i^f(x, y, z) \exp\left[-0.5(z - z_t)^2\right], \quad (12)$$

где $i = x, y, z$. $U_i^f(x, y, z)$ – однородное, изотропное случайное поле с заданным спектром энергии в виде

$$E(k) = E_0 k \exp\left(-\frac{k}{k_f}\right) \quad (13)$$

Волновое число k_f в (13) определяет положение максимума энергии и масштаб турбулентности. Фактор E_0 задается таким, что абсолютный максимум поля $U_i^f(x, y, z)$ равен единице. Таким образом, параметр U_{i0} в (12) определяет амплитуду турбулентной скорости; параметры U_{i0} и k_f задавались равными $U_{i0} = 0,1$ и $k_f = 1$. Расчеты показывают, что в этом случае эффекты турбулентного перемешивания и генерации ВВ турбулентностью пренебрежимо малы.

Вертикальные профили средней скорости и плотности, $\langle U_i \rangle(z)$ и $\langle \rho \rangle(z)$, получались осреднением по горизонтальной плоскости (x, y) для различных z . Среднеквадратичные отклонения (флуктуации) определялись как

$$U'_i = \left(\langle U_i^2 \rangle - \langle U_i \rangle^2\right)^{1/2}, \quad \rho' = \left(\langle \rho^2 \rangle - \langle \rho \rangle^2\right)^{1/2} \quad (14)$$

и кинетическая энергия равнялась

$$E = \frac{1}{2} \sum_{i=x,y,z} U_i'^2 \quad (15)$$

Вычислялось также градиентное число Ричардсона, Ri_g [Phillips, 1977], которое является критерием, разделяющим режимы слабой и сильной стратификации турбулентности. В рассматриваемом случае (при отсутствии среднего сдвига скорости) это число определяется отношением квадрата частоты плавучести и среднеквадратичного сдвига скорости, пропорционального диссипации турбулентной кинетической энергии ε [Thorpe, 2007]:

$$Ri_g = -\frac{Ri}{\varepsilon Re} \frac{d \langle \rho \rangle}{dz} = \frac{N^2}{\varepsilon Re} \quad (16)$$

где ε вычисляется в виде [Phillips, 1977]:

$$\varepsilon = \frac{1}{Re} \left\langle \left(\frac{\partial U'_i}{\partial x_j} \right)^2 \right\rangle \quad (17)$$

5. УСИЛЕНИЕ ТУРБУЛЕНТНОСТИ В ПОЛЕ ВВ

Расчеты динамики турбулентности в присутствии ВВ проводись для начальных параметров, обсуждаемых выше (в разделах 3 и 4).

На рисунке 3 представлены вертикальные профили флуктуаций скорости U'_x , U'_y , U'_z и средней плотности $\langle \rho \rangle$ (слева), и градиентное число Ричардсона Ri_g (справа), полученные в DNS при $t = 100$ (вверху) и $t = 400$ (внизу). На рисунке приводятся также профили флуктуаций U'_y в отсутствие ВВ и профили U'_x и U'_z во внутренней волне, распространяющейся в пикноклине без исходной турбулентности.

Рисунок 3 показывает, что профили флуктуаций скорости U'_x и U'_z во внутренней волне, распространяющейся в пикноклине в присутствии и в отсутствие тур-

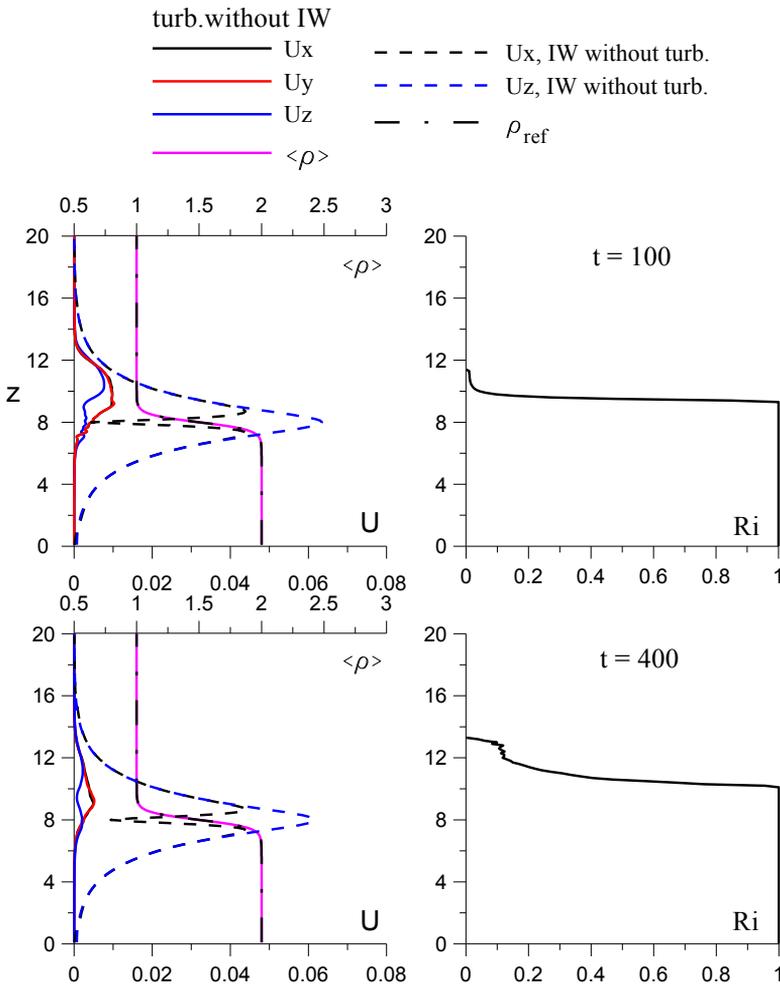


Рис. 3. Профили флуктуаций скорости U'_x , U'_y , U'_z и средней плотности $\langle \rho \rangle$ (слева), и градиентное число Ричардсона Ri_g (справа), полученные в DNS при $t = 100$ (вверху) и $t = 400$ (внизу)

булентности, практически совпадают. Это означает, что турбулентность слабо влияет на ВВ, и эффект затухания ВВ в рассматриваемом случае пренебрежимо мал. Это наблюдение согласуется с предыдущими результатами [Druzhinin et al., 2013], показывающими что ВВ затухают на турбулентности, если только их амплитуда мала (как минимум, в два раза меньше) по сравнению с амплитудой турбулентных пульсаций. В рассматриваемом случае исходная амплитуда ВВ равна амплитуде турбулентности.

Из рисунка 3 видно также, что на больших временах ($t = 400$) флуктуации у-компоненты скорости U'_y значительно возрастают в окрестности пикноклина, при $z \approx 8$, под действием поля ВВ. Важно отметить, что эта компонента скорости является чисто турбулентной, поскольку поле ВВ характеризуется лишь x и z компонентами, и, следовательно, не дает прямого вклада в y компоненту скорости. Качественно похожее усиление турбулентности наблюдалось в лабораторном эксперименте [Матусов и др., 1989]. В эксперименте создавались стратификация с пикноклином. Турбулизированный слой над пикноклином создавался с помощью осциллирующей решетки. В пикноклине с помощью волнопродуктора возбуждалась сильная, необрушающаяся стоячая внутренняя волна. Результаты показали, что при прекращении колебаний решетки турбулентность затухала во всей области, за исключением области вблизи пикноклина, где флуктуации скорости поддерживались на некотором уровне полем ВВ. Это наблюдение находится в качественном согласии с результатами расчетов на рис. 3.

На рисунке 4 представлены временные зависимости флуктуаций поперечной компоненты скорости U'_y в центре пикноклина (при $z = 8$), полученных в DNS как для свободно затухающей турбулентности (без инициализации ВВ), так и под действием ВВ. Из рисунка видно, что под воздействием ВВ U'_y растет со временем и при $t = 400$ на порядок превышает турбулентные флуктуации свободно затухающей турбулентности.

Мгновенные поля у-компоненты завихренности ($\omega_y = \partial_z U_x - \partial_x U_z$) и плотности ($\rho + \rho_{ref}(z)$), полученные в DNS при $t = 100$ и 400 , представлены на рис. 5. Распределение завихренности при $t = 100$ характеризуется наличием двух различных областей – $7 < z < 9$ и $9 < z < 12$ – со слабой и сильной стратификацией, соответственно. В области с сильной стратификацией, при $7 < z < 9$, где градиентное число Ричардсона велико, $Ri_g > 1$, распределение завихренности характеризуется наличием локальных минимумов и максимумов, расположенных вблизи подошв и гребней ВВ. В области $9 < z < 12$, стратификация слабая, т.к. $Ri_g < 1$, поэтому распределение завихренности однородно в горизонтальной плоскости. На больших временах (при $t = 400$) завихренность сконцентрирована в узком горизонтальном слое вблизи центра пикноклина (при $z \approx 8$). При этом $Ri_g > 1$, и эффекты стратификации значительны. В области, достаточно удаленной от пикноклина (при $z > 9$), завихренность исчезающе мала.

Таким образом, исходя из данных на рисунках 4 и 5, приходим к выводу, что на больших временах турбулентность локализована в узком слое в окрестности центра пикноклина, где происходит ее генерация в поле ВВ. Этот вывод согласуется с экспериментальными данными [Матусов и др., 1989].

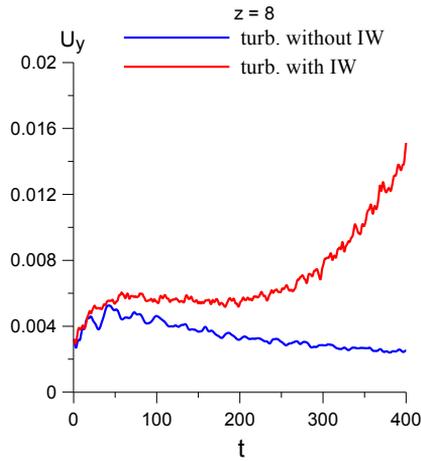


Рис. 4. Временные зависимости флуктуаций поперечной компоненты скорости U'_y в центре пикноклина (при $z = 8$) для свободно затухающей турбулентности (синий цвет) и для турбулентности при распространении ВВ в пикноклине (красный)

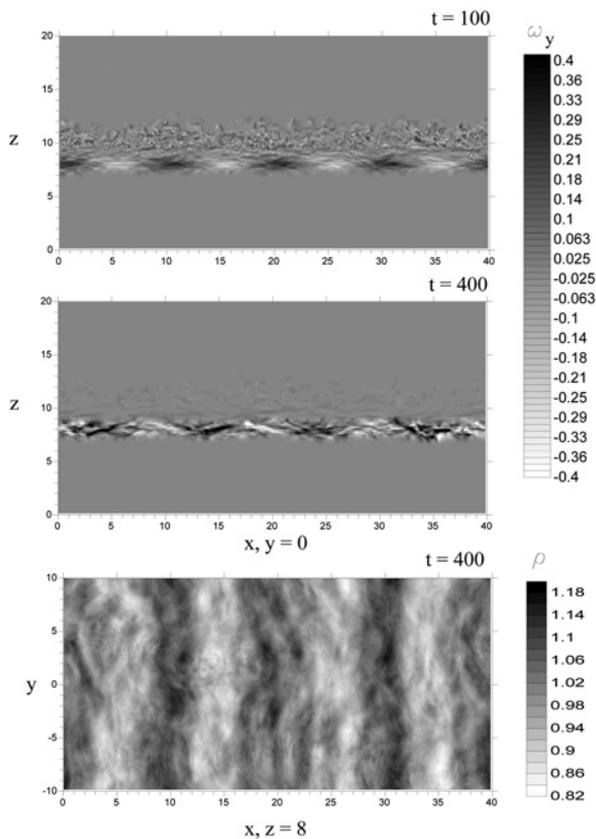


Рис. 5. Распределения завихренности (тонировка) и плотности (линии уровней 1.3, 1.5, 1.7) в центральной плоскости (x,z) в моменты времени $t = 100$ (вверху) и $t = 400$ (в середине), и распределение плотности в плоскости (x,y) на горизонте пикноклина $z = 8$ при $t = 400$ (внизу)

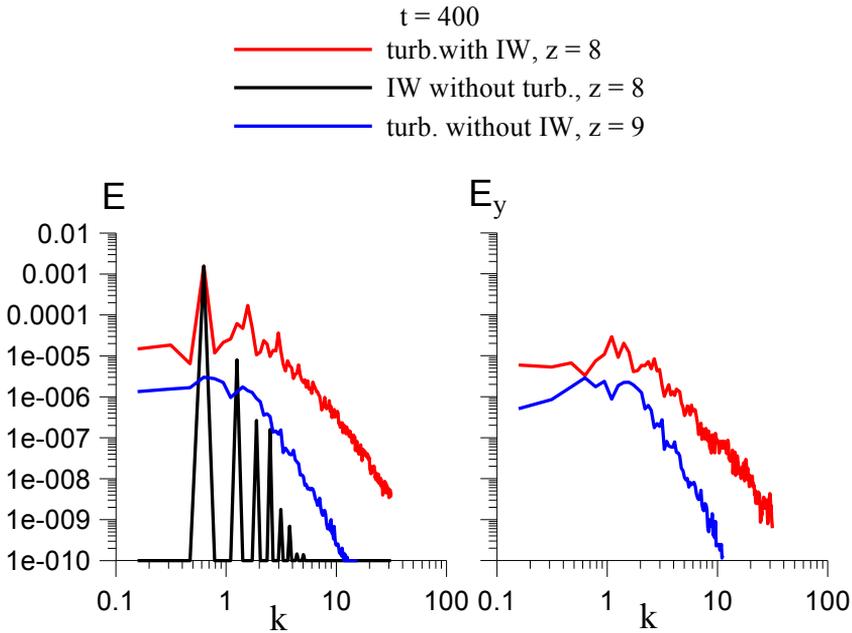


Рис. 6. Спектры кинетической энергии турбулентности и ВВ в момент времени $t = 400$.

Красным цветом показаны спектры турбулентности на горизонте пикноклина при распространении в нем внутренней волны. Синим цветом обозначены спектры свободно затухающей турбулентности на горизонте максимума энергии. В черном цвете представлен спектр ВВ. Справа показаны спектры y -компоненты скорости турбулентности

Рисунок 5 (нижняя панель) показывает также, что на больших временах фронт ВВ существенно искажается турбулентностью. Эта рефракция ВВ вызвана возрастанием турбулентности и может приводить к более существенному затуханию ВВ по мере того, как растет амплитуда турбулентных пульсаций [Druzhinin et al., 2013].

На рисунке 6 показаны спектры кинетической энергии турбулентности E и спектры y -компоненты скорости турбулентности E_y , как свободно затухающей, так и эволюционирующей в поле ВВ, распространяющейся в пикноклине. На рисунке приведен также спектр внутренней волны, распространяющейся в пикноклине в отсутствие турбулентности. Рисунок показывает, что под действием ВВ спектр энергии турбулентности возрос на порядок на всем интервале волновых чисел. Пик в спектре $E(k)$ при $k = 2\pi/10$ обусловлен непосредственным вкладом ВВ в поле скорости турбулентности. Этот вклад отсутствует в спектре E_y , т.к. ВВ характеризуется лишь x и z компонентами скорости и не влияет напрямую на y -компоненту скорости.

Из рисунка 6 видно также, что высшие гармоники в спектре ВВ имеют амплитуды, много меньшие по сравнению с амплитудой основной гармоники (при $k = 2\pi/10$). Таким образом, можно сделать вывод о том, что нелинейные эффекты при распространении ВВ пренебрежимо малы.

6. ЗАКЛЮЧЕНИЕ

Методом прямого численного моделирования проведены расчеты динамики турбулентности, изначально локализованной в слое над пикноклином в стратифицированной жидкости, и исследованы эффекты воздействия на турбулентность внутренних волн (ВВ), распространяющихся в пикноклине. Показано, что ВВ, амплитуда которых порядка или превышает амплитуду турбулентных пульсаций, слабо затухают на турбулентности и способны существенно усиливать турбулентность в окрестности горизонта пикноклина. Расчеты показывают, что спектр турбулентности, находящейся под воздействием ВВ, усиливается на порядок по сравнению со спектром свободно затухающей турбулентности.

Приведем некоторые оценки масштабов турбулентности и внутренних волн, реализуемых в лабораторных и натуральных условиях, на которые распространяются обсуждаемые выше результаты. В данной работе используются масштабы скорости, длины и времени – U_0 , L_0 и $T_0 = L_0 / U_0$ – при нормировке физических переменных. Для воспроизведения лабораторных условий положим $L_0 = 20$ см (длина ВВ – $\lambda = 2$, толщина пикноклина – 20 см), тогда для рассматриваемого числа Рейнольдса $Re = 20000$ и кинематической вязкости воды $\nu = 0,01$ см²/с получаем масштаб скорости $U_0 = Re \nu / L_0 = 10$ см/с. Т.о. исходные амплитуды турбулентности и ВВ $0,1 U_0 = 1$ см/с, временной масштаб $T_0 = 2$ с и частота плавучести $N_0 = 0,5$ рад/с. Экстраполируя на натурные условия океана, положим частоту плавучести равной $N_0 = 0,01$ рад/с [Phillips, 1977], тогда временной масштаб равен $T_0 = 100$ с и масштаб длины $L_0 = 20$ м, что соответствует длине ВВ 200 м и толщине пикноклина 20 м. Масштаб скорости при этом оказывается равным $U_0 = 20$ см/с, и исходные амплитуды скорости турбулентности и внутренней волны равны $0,1 U_0 = 2$ см/с. Приведенные масштабы можно рассматривать как типичные и реализуемые как в лабораторном эксперименте (см. напр. постановку эксперимента Матусова и др. [Матусов и др., 1989]), так и в натуральных условиях океана [Phillips, 1977; Thorpe, 2007].

ЛИТЕРАТУРА

- [Матусов и др., 1989] Матусов П.А., Островский Л.А., Цимринг Л.Ш. Усиление мелкомасштабной турбулентности внутренними волнами. // ДАН, 307, 1989. С. 979-984.
- [Монин, Яглом, 1992] Монин А.С., Яглом А.М. Статистическая гидромеханика // Т.1., Гидрометеиздат, С.-Петербург. 1992. 695 с.
- [Druzhinin et al., 2013] Druzhinin O.A., Ostrovsky L.A., Zilitinkevich S.S. The study of the effect of small-scale turbulence on internal gravity waves propagation in a pycnocline // *Nonlin. Processes Geophys.*, 20, 2013. Pp. 977-986.
- [Hazel, 1972] Hazel P. Numerical studies of the stability of inviscid stratified shear flows // *J. Fluid Mech.*, 51, 1972. Pp. 39-61.
- [Ostrovsky, Zaborskikh, 1996] Ostrovsky L.A., Zaborskikh D.V. Damping of internal gravity waves by small-scale turbulence // *J. Phys. Oceanogr.*, 26, 1996. Pp. 388-397.
- [Phillips, 1977] Phillips O.M. The dynamics of the upper ocean // 2nd ed., Cambridge. 1977. 336 p.
- [Thorpe, 2007] Thorpe S.A. An introduction to ocean turbulence // Cambridge. 2007. 293 p.

DIRECT NUMERICAL SIMULATION STUDY OF THE INTERACTION OF SMALL-SCALE TURBULENCE WITH INTERNAL WAVES

O.A. Druzhinin¹, L.A. Ostrovsky², Z. Feng³

¹ Institute of Applied Physics RAS

² NOAA Environmental Science Research Lab. Boulder, CO, USA

³ China Ship Scientific Research Center, Wuxi, Jiangsu, China

The interaction between small-scale turbulence and internal waves (IW) plays an important role in the dynamics and mixing of the upper ocean. Among many aspects of this interaction are the effects of IW damping by turbulence on the one hand, and the possibility of the enhancement of turbulence by sufficiently strong IWs, on the other hand. In the present work, we study this interaction by performing direct numerical simulation.

Keywords: numerical simulation, internal waves, small-scale turbulence, upper ocean.

Druzhinin Oleg Aleksandrovich – Institute of Applied Physics RAS, druzhinin@hydro.appl.sci-nnov.ru
Ostrovsky L. A. – NOAA Environmental Science Research Lab. Boulder, CO, USA, lev.a.ostrovsky@noaa.gov
Feng Z. – China Ship Scientific Research Center, Wuxi, Jiangsu, China

ДИНАМИКА ОДНОГО ДВУМЕРНОГО ОТОБРАЖЕНИЯ И УСТОЙЧИВЫЕ РЕЖИМЫ СИНГУЛЯРНО ВОЗМУЩЕННОЙ СИСТЕМЫ НЕЙРОННОГО ТИПА

Л.И.Ивановский, С.О.Самсонов

Ярославский государственный университет им. П.Г. Демидова
(ЯрГУ им. П.Г. Демидова)

Рассматривается математическая модель нейронной системы, представляющая собой цепочку из трех диффузионно связанных сингулярно возмущенных нелинейных дифференциальных уравнений с запаздыванием. Изучаются вопросы существования, устойчивости и асимптотического представления релаксационных периодических решений этой системы на основании бифуркационного анализа специального двумерного отображения. Особое внимание уделяется числу сосуществующих устойчивых режимов, которое удастся определить, сочетая асимптотические и численные методы. Для нахождения всех устойчивых состояний равновесия модельного отображения используются алгоритмы параллельных вычислений.

Исследование выполнено за счет гранта Российского научного фонда (проект № 14-21-00158).

Ключевые слова: релаксационные колебания, устойчивые режимы, бифуркации.

ВВЕДЕНИЕ

Для цепочки сингулярно возмущенных осцилляторов с запаздыванием, моделирующей слабое электрическое взаимодействие нейронов, в статье [Глызин С.Д. и др., 2011а; Глызин С.Д. и др., 2011б] предложена основанная на модели из книги [Кашенко С.А., Майоров В.В., 2009] система дифференциальных уравнений с запаздыванием

$$\dot{u}_j = d(u_{j+1} - 2u_j + u_{j-1}) + \lambda [-1 + \alpha f(u_j(t-1)) - \beta g(u_j)] \overline{u_{j,j}} = 1, m, \quad (1)$$

где $m \geq 2$, $u_0 = u_1$, $u_3 = u_4$, $\lambda \gg 1$, $\beta > 0$, $\alpha > 1 + \beta$, а гладкие функции $f(u)$ и $g(u)$ удовлетворяют условиям $f(0) = g(0) = 1$, $f(u)$, $g(u)$, $uf'(u)$, $ug'(u) = O(1/u)$, при $u \rightarrow +\infty$. При сформулированных ограничениях система (1) допускает синхронный цикл $u_1 = u_2 = u_*(t, \lambda)$, где функция $u_*(t, \lambda)$ удовлетворяет уравнению

$$\dot{u} = \lambda [-1 + \alpha f(u(t-1)) - \beta g(u)] u.$$

В [Глызин С.Д. и др., 2011а; Глызин С.Д. и др., 2011б] показано, что периодическое решение $u_*(t, \lambda)$ с периодом $T_*(\lambda)$ асимптотически близко при $\lambda \gg 1$ к функции $\exp(\lambda x_0)$, где

Ивановский Леонид Игоревич – студент 5 курса факультета Информатики и вычислительной техники Ярославского государственного университета им. П.Г. Демидова, лаборант-исследователь ОПСи НЦЧ РАН, leon19unknown@gmail.com

Самсонов Сергей Олегович – студент 5 курса факультета Информатики и вычислительной техники Ярославского государственного университета им. П.Г. Демидова, therealfac@gmail.com

$$x_0(t) = \begin{cases} (\alpha - 1)t, & \text{при } 0 \leq t \leq 1, \\ \alpha - t, & \text{при } 1 \leq t \leq \alpha, \\ -(1 + \beta)(t - \alpha), & \text{при } \alpha \leq t \leq \alpha + 1, \\ (\alpha - \beta - 1)(t - \alpha - 1) - 1 - \beta, & \text{при } \alpha + 1 \leq t \leq T_0, \end{cases}$$

$T_0 = \alpha + 1 + (\beta + 1)/(\alpha - \beta - 1)$, причем справедливо предельное равенство $\lim_{\lambda \rightarrow \infty} T^*(\lambda) = T_0$.

С помощью замен вида

$$u_1 = \exp\left(\frac{x}{\varepsilon}\right), u_j = \exp\left(\frac{x}{\varepsilon} + \sum_{k=1}^{j-1} y_k\right), j = \overline{2, m}, \varepsilon = \frac{1}{\lambda} \ll 1,$$

где x близко к x_0 , систему (1) можно свести к следующей системе:

$$\dot{x} = \varepsilon d(e^{y_1} - 1) + F(x, x(t-1), \varepsilon),$$

$$\dot{y}_j = d(e^{y_{j+1}} + e^{-y_j} - e^{y_j} - e^{-y_{j-1}}) + G_j(x, x(t-1), y_1, \dots, y_j, y_1(t-1), \dots, y_j(t-1), \varepsilon), j = \overline{1, m-1},$$

где $y_0 = y_m = 0$, а функции имеют вид

$$F(x, u, \varepsilon) = -1 + \alpha f\left(\exp\left(\frac{u}{\varepsilon}\right)\right) - \beta g\left(\exp\left(\frac{x}{\varepsilon}\right)\right),$$

$$G_j(x, u, y_1, \dots, y_j, v_1, \dots, v_j, \varepsilon) =$$

$$= \frac{\alpha}{\varepsilon} \left[f\left(\exp\left(\frac{u}{\varepsilon} + \sum_{k=1}^j v_k\right)\right) - f\left(\exp\left(\frac{u}{\varepsilon} + \sum_{k=1}^{j-1} v_k\right)\right) \right] +$$

$$+ \frac{\beta}{\varepsilon} \left[g\left(\exp\left(\frac{x}{\varepsilon} + \sum_{k=1}^{j-1} y_k\right)\right) - g\left(\exp\left(\frac{x}{\varepsilon} + \sum_{k=1}^j y_k\right)\right) \right], j = \overline{1, m-1}.$$

В статьях [Глызин С.Д. и др., 2011а; Глызин С.Д. и др., 2011б; Глызин С.Д. и др., 2012а; Глызин С.Д. и др., 2012б; Глызин С.Д. и др., 2013] выполнено сведение этой системы к системе обыкновенных дифференциальных уравнений с импульсным воздействием. Рассмотрим данную задачу в случае трех осцилляторов ($m = 3$). При этом (см. [Ивановский Л.И., Самсонов С.О., 2014]) получаем следующую систему:

$$\begin{cases} \dot{y}_1 = d(e^{y_2} + e^{-y_1} - e^{y_1} - 1) \\ \dot{y}_2 = d(1 + e^{-y_2} - e^{y_2} - e^{-y_1}). \end{cases} \quad (2)$$

$$y_j(+0) = \frac{\alpha - 1}{\alpha - \beta - 1} y_j(-0), y_j(1 + 0) = y_j(1 - 0) - \frac{\alpha}{\alpha - 1} y_j(+0).$$

$$y_j(\alpha + 0) = (1 + \beta) y_j(\alpha - 0),$$

$$y_j(\alpha + 0) = (1 + \beta) y_j(\alpha - 0),$$

Функции $y_1(t)$ и $y_2(t)$ характеризуют фазовые сдвиги между компонентами системы (1). Ниже будем изучать следующее отображение:

$$\Phi(z): \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \rightarrow \begin{pmatrix} y_1(T^*) \\ y_2(T^*) \end{pmatrix}, \quad (3)$$

где $(y_1(t), y_2(t))^T$ – решения системы (2) с импульсным воздействием и следующими начальными условиями: $y_1(-0) = z_1, y_2(-0) = z_2$.

В статьях [Глызин С.Д. и др., 2011а; Глызин С.Д. и др., 2011б; Глызин С.Д. и др., 2012а; Глызин С.Д. и др., 2012б] доказано, что экспоненциально устойчивым неподвижным точкам отображения (3) соответствуют орбитально асимптотически устойчивые циклы системы (2) и, в свою очередь, системы (1). Асимптотический анализ отображения (3) позволяет показать, что при достаточно малых значениях параметра d оно имеет как минимум три устойчивые неподвижные точки. Заметим, что нулевое состояние равновесия устойчиво при любых значениях d , при этом ему соответствует однородный (синхронный) цикл системы (1). Наша задача состоит в определении таких значений α и β , при которых отображение (3) имеет наибольшее число устойчивых неподвижных точек. Поскольку сделать это в полной мере с использованием одних лишь асимптотических методов довольно затруднительно, поиск устойчивых состояний осуществлялся с помощью специально разработанного приложения. Вычисление координат неподвижных точек в этой программе осуществлялось параллельно, на независимых потоках центрального процессора.

ПОИСК НЕПОДВИЖНОЙ ТОЧКИ РЕЛАКСАЦИОННОГО ОТОБРАЖЕНИЯ

Расчёт координат конечной точки отображения (3) идёт следующим образом: вычисления начинаются из точки $(y_1(-0), \dots, y_{m-1}(-0))^T$ координаты которой представляют собой начальные условия системы (2). Начальным моментом времени считается величина $t = -0$. Далее, итерационно, с шагом h , при помощи метода Рунге-Кутты подсчитывались координаты следующей точки отображения (3). Этот метод имеет четвёртый порядок точности, т.е. ошибка на каждом интервале подсчёта будет иметь порядок $O(h^5)$. Координаты следующей точки отображения (3) подсчитываются по итерационной формуле:

$$y_j(t+h) = y_j(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), j = \overline{1, m-1}$$

где $y_j(t+h), y_j(h)$ – координаты точек, полученных на текущей и предыдущей итерациях, h – величина шага сетки t по для метода Рунге-Кутты, а коэффициенты k_1, k_2, k_3, k_4 рассчитываются по следующим формулам:

$$k_1 = d(e^{y_{j+1}} + e^{-y_j} - e^{y_j} - e^{-y_{j-1}})$$

$$k_2 = d(e^{y_{j+1}+hk_1/2} + e^{-y_j+hk_1/2} - e^{y_j+hk_1/2} - e^{-y_{j-1}+hk_1/2})$$

$$k_3 = d(e^{y_{j+1}+hk_2/2} + e^{-y_j+hk_2/2} - e^{y_j+hk_2/2} - e^{-y_{j-1}+hk_2/2})$$

$$k_4 = d(e^{y_{j+1}+hk_3} + e^{-y_j+hk_3} - e^{y_j+hk_3} - e^{-y_{j-1}+hk_3})$$

Вычисления такого рода будут продолжаться до тех пор, пока не будет получено конечное состояние в значении времени T_0 . Однако в точках $0, 1, \alpha, \alpha+1$ переключения релаксационной системе придавался импульс. Координаты точки в этом случае подсчитывались по формулам:

$$y_j(+0) = \frac{\alpha - 1}{\alpha - \beta - 1} y_j(-0), \quad y_j(1 + 0) = y_j(1 - 0) - \frac{\alpha}{\alpha - 1} y_j(+0),$$

$$y_j(\alpha + 0) = (1 + \beta) y_j(\alpha - 0),$$

$$y_j(\alpha + 1 + 0) = y_j(\alpha + 1 - 0) - \frac{\alpha}{1 + \beta} y_j(\alpha + 0),$$

$$j = 1, 2, y_0 = y_3 = 0.$$

Далее происходила проверка начальной и конечной точек отображения (3) на эквивалентность. Если их координаты были достаточно близки друг к другу, это означало, что найдено устойчивое состояние равновесия. В противном же случае недавно посчитанную точку $(y_1(T_0), \dots, y_{m-1}(T_0))^T$ считали начальной и производили для неё ту же последовательность операций, которая была описана ранее.

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО ИССЛЕДОВАНИЯ

На координатной плоскости параметров (α, β) можно выделить области A_1, A_2, A_3 и кривые l_0, \dots, l_4 . Графическая визуализация данных множеств приведена на рис. 1.

Признак, по которому произведено разбиение на области, – это различие максимального количества обнаруживаемых устойчивых неподвижных точек у отображения.

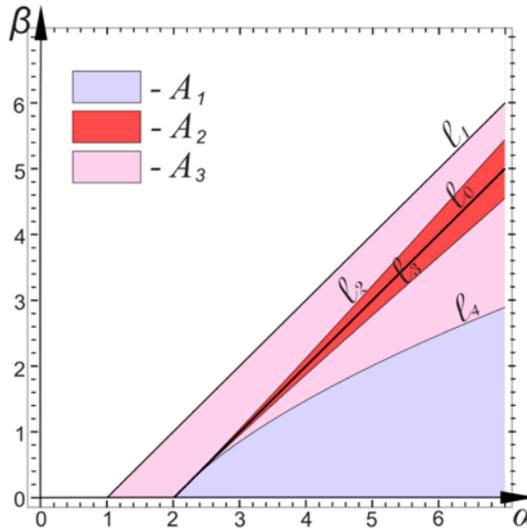


Рис. 1. Разбиение плоскости параметров на области с одинаковыми сценариями фазовых перестроек

Для значений параметров α и β из области A_1 возможно единовременное наличие у отображения пяти устойчивых неподвижных точек. В области A_2 обнаруживаются ситуации максимум с семью устойчивыми положениями. Области A_3 соответствуют наиболее сложные случаи, максимальное число устойчивых неподвижных точек в которых достигает шести.

Наиболее важным элементом, служащим для построения введенных областей, является прямая $l_0 = \{(\alpha, \beta) : \beta = \alpha - 2\}$.

Симметричным образом относительно ее проведены кривые l_2 и l_3 , касающиеся l_0 в точке $(2, 0)$. Эти кривые являются границами области A_2 :

$$A_2 = \{(\alpha, \beta) : \beta > l_3, \beta < l_2\}.$$

Также в точке $(2, 0)$ проведена касательная l_4 кривая к прямой l_0 , в совокупности с осью абсцисс образующая границы области A_1 :

$$A_1 = \{(\alpha, \beta) : \beta < l_4, \beta > 0\}.$$

Двусвязная область A_3 представляется в виде:

$$A_3 = \{(\alpha, \beta) : \beta > l_2, \beta < l_1, \beta > l_4, \beta < l_3\},$$

где прямая $l_1 = \{(\alpha, \beta) : \beta = \alpha - 1\}$.

Далее рассмотрим типичные бифуркации для каждой из введенных областей (см. также заметку [Ивановский Л.И., Самсонов С.О., 2014]).

Основные бифуркации в области A_1 . При изменении параметра d для любых фиксированных значений пары $(\alpha, \beta) \in A_1$ в фазовом пространстве отображения наблюдается один и тот же сценарий бифуркационных перестроек. Для определенности возьмем $\alpha = 5,0$ и $\beta = 0,4$. Далее начнем менять значение параметра d .

Подобный численный анализ отображения (3) позволяет получить следующую последовательность бифуркаций:

1) При $d < d_1$; $d_1 = 0,019$ отображение имеет пять устойчивых неподвижных точек и шесть неустойчивых. Схематическое изображение фазового портрета отображения Φ в данном случае можно увидеть на рис. 2а. Буквами S_j обозначены устойчивые, а U_j – неустойчивые неподвижные точки.

2) При $d = d_1$ два симметричных друг другу седла U_5 и U_6 сливаются с самосимметричной точкой S_4 и отбирают у нее устойчивость, образуя седло U' .

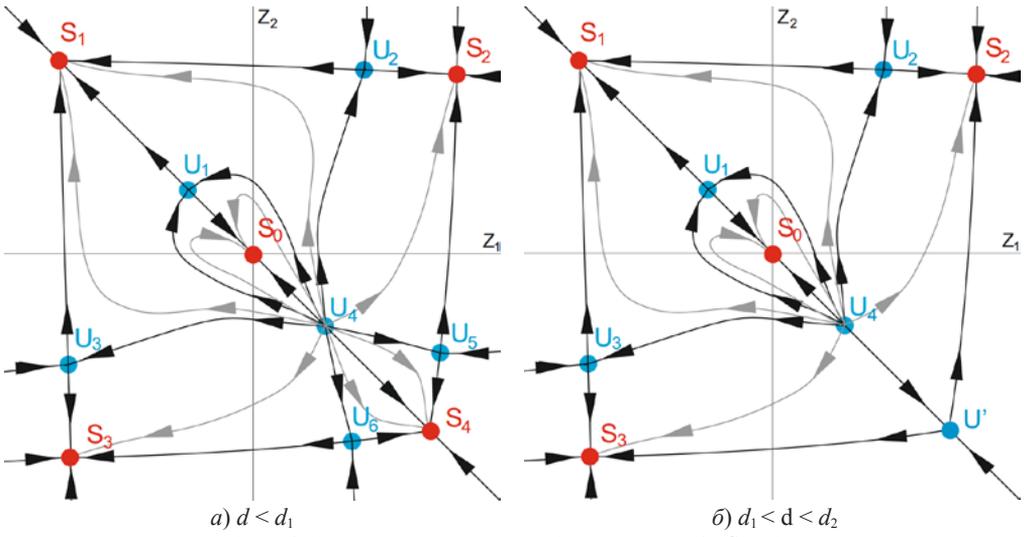
3) При $d_1 < d < d_2$; $d_2 = 0,031$ отображение имеет четыре устойчивые неподвижные точки и пять неустойчивых. Фазовый портрет в данном случае имеет изображенный на рис. 2б вид.

4) При $d = d_2$ неустойчивый узел U_4 и седло U_7 , сливаясь, пропадают.

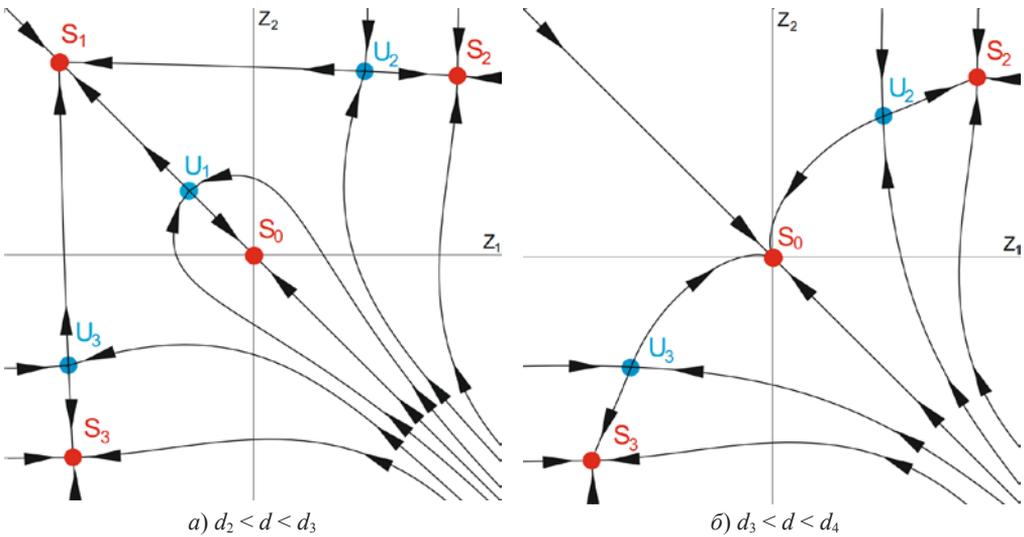
5) При $d_2 < d < d_3$; $d_3 = 0,059$ отображение имеет четыре устойчивые неподвижные точки и три неустойчивые. Схема фазового пространства представлена на рис. 3а.

6) При $d = d_3$ устойчивый узел S_1 и седло U_1 , сливаясь друг с другом, пропадают.

7) При $d_3 < d < d_4$; $d_4 = 0,127$ отображение имеет три устойчивые неподвижные точки и две неустойчивые. Фазовый портрет изображен на рис. 3б.



а) $d < d_1$ б) $d_1 < d < d_2$
 Рис. 2. Фазовые портреты отображения при $(\alpha, \beta) \in A_1$



а) $d_2 < d < d_3$ б) $d_3 < d < d_4$
 Рис. 3. Фазовые портреты отображения при $(\alpha, \beta) \in A_1$

8) Последняя бифуркация происходит при $d = d_4$. Симметричные друг другу устойчивые узлы S_2 и S_3 сливаются с симметричными седлами U_2 и U_3 и исчезают. Тем самым, при $d > d_4$ отображение имеет единственное нулевое устойчивое состояние.

Полученный сценарий фазовых перестроек остается верным для любых $(\alpha, \beta) \in A_1$. Отличия от рассмотренного случая состоят лишь в числовых выражениях бифуркационных значений параметра d .

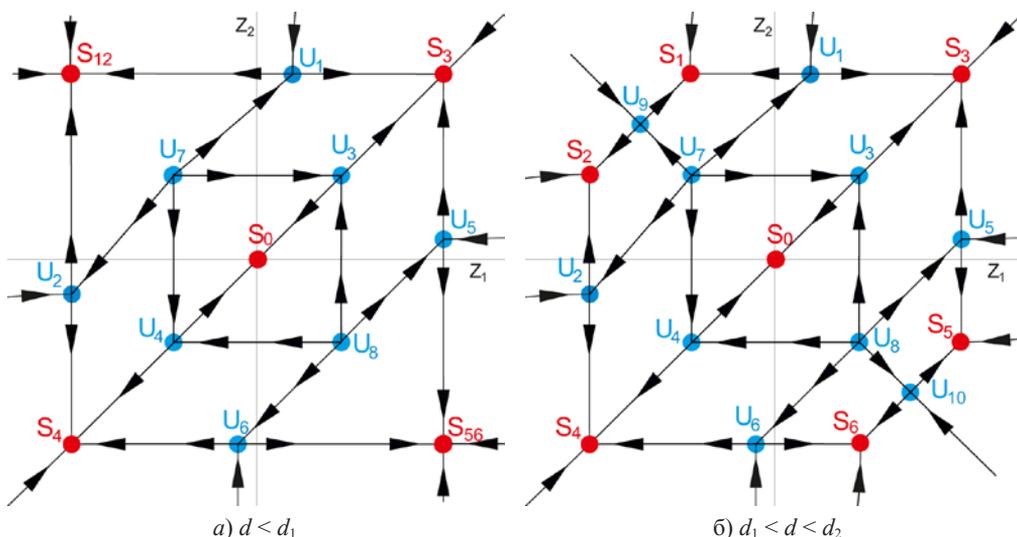
Основные бифуркации в области A_2 . При изменении параметра d для любых фиксированных значений пары $(\alpha, \beta) \in I_0$ в фазовом пространстве отображения наблюдается один и тот же сценарий бифуркационных перестроек. Для

удобства зафиксируем величины $\alpha = 5,0$ и $\beta = 3,0$ и будем менять значение параметра d . В результате получим следующую последовательность бифуркаций:

1) При $d < d_1$; $d_1 = 0,047$ отображение имеет пять устойчивых неподвижных точек и восемь неустойчивых. Схематическое изображение фазового портрета отображения Φ в данном случае можно увидеть на рис. 4а.

2) При $d = d_1$ от устойчивого самосимметричного узла S_{12} ответвляется пара симметричных друг другу устойчивых узлов S_1 и S_2 и самосимметричная седловая точка U_9 . Одновременно с этой бифуркацией происходит подобная с устойчивым узлом S_{56} . Он также распадается на два устойчивых узла S_5 и S_6 и седло U_{10} .

3) При $d_1 < d < d_2$; $d_2 = 0,056$ отображение имеет семь устойчивых неподвижных точек. Фазовый портрет в данном случае имеет изображенный на рис. 4б вид.



а) $d < d_1$ б) $d_1 < d < d_2$
Рис. 4. Фазовые портреты отображения при $(\alpha, \beta) \in I_0$

4) При $d = d_2$ неустойчивые узлы U_7 и U_8 сливаются соответственно с седловыми точками U_9 и U_{10} .

5) При $d_2 < d < d_3$; $d_3 = 0,058$ отображение имеет семь устойчивых неподвижных точек. Фазовый портрет, изображенный на рис. 5а при $d = 0,057$, позволяет проиллюстрировать эту ситуацию.

6) При $d = d_3$ симметричные друг другу устойчивые узлы S_1 и S_2 сливаются с симметричными седлами U_1 и U_2 и исчезают. Одновременно с этим проходит аналогичная бифуркация, в результате которой исчезают устойчивые узлы S_5 и S_6 и седла U_5 и U_6 .

7) При $d_3 < d < d_4$; $d_4 = 0,172$ отображения имеет три устойчивые и две неустойчивые неподвижные точки. Фазовый портрет изображен на рис. 5б.

8) Наконец, при $d = d_4$ симметричные друг другу устойчивые узлы S_3 и S_4 сливаются с симметричными седлами U_3 и U_1 и исчезают. При отображении имеет единственное нулевое устойчивое состояние.

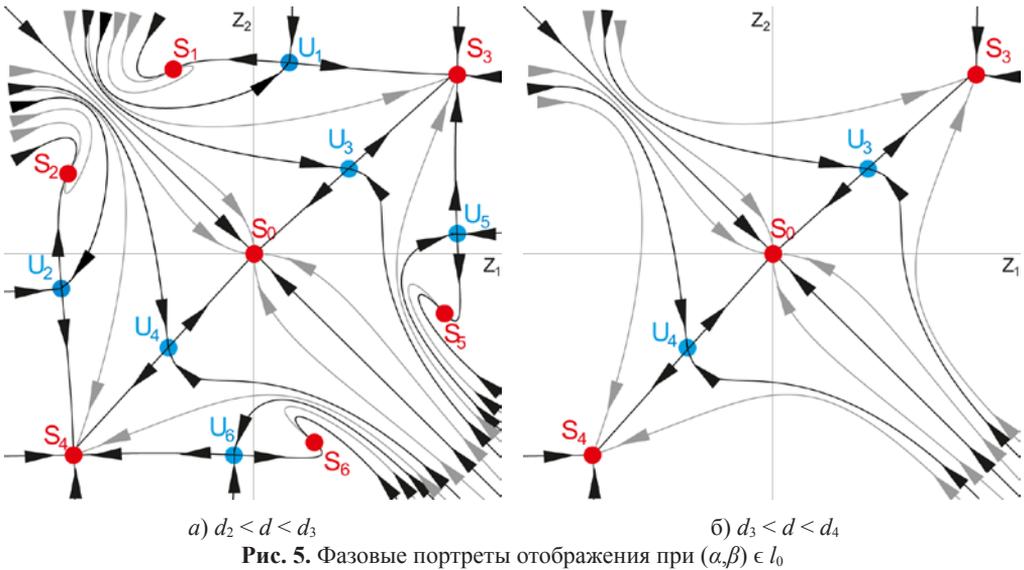


Рис. 5. Фазовые портреты отображения при $(\alpha, \beta) \in l_0$

Далее программно было исследовано поведение отображения в подобластях:

$$A'_2 = \{(\alpha, \beta) : \beta > l_0, \beta < l_2\} \text{ и } A''_2 = \{(\alpha, \beta) : \beta > l_3, \beta < l_0\}.$$

Эти области образованы разделением области A_2 прямой l_0 на две компоненты.

Обнаружилось, что в введенных областях A'_2 и A''_2 нарушается одновременность протекания бифуркаций распада устойчивых узлов S_{12} и S_{56} , а также последующих бифуркаций слияния их компонент. При этом другие качественные отличия со сценарием при $(\alpha, \beta) \in l_0$ отсутствуют.

Для области A'_2 характерен сдвиг бифуркации распада устойчивого узла S_{12} по d вперед. Она происходит теперь при большем d , чем раньше. Также бифуркация распада устойчивого узла S_{56} сдвигается по d и происходит при меньшем d , чем раньше.

Области A''_2 наоборот свойственен сдвиг бифуркации распада S_{12} по d назад, а узла S_{56} – вперед.

Основные бифуркации в области A_3 . Поскольку область A_3 – двусвязная, введем в рассмотрение следующие ее подобласти:

$$A'_3 = \{(\alpha, \beta) : \beta > l_2, \beta < l_1\} \text{ и } A''_3 = \{(\alpha, \beta) : \beta > l_4, \beta < l_3\}.$$

С помощью уже хорошо знакомого метода можно получить последовательность фазовых перестроек для значений $(\alpha, \beta) \in A'_3$. Например, в случае $\alpha = 4,0$ и $\beta = 2,3$ наблюдаются следующие бифуркации:

1) При $d < d_1$; $d_1 = 0,040$ отображение имеет пять устойчивых неподвижных точек и восемь неустойчивых. Схематическое изображение фазового портрета отображения Φ в данном случае можно увидеть на рис. 6а.

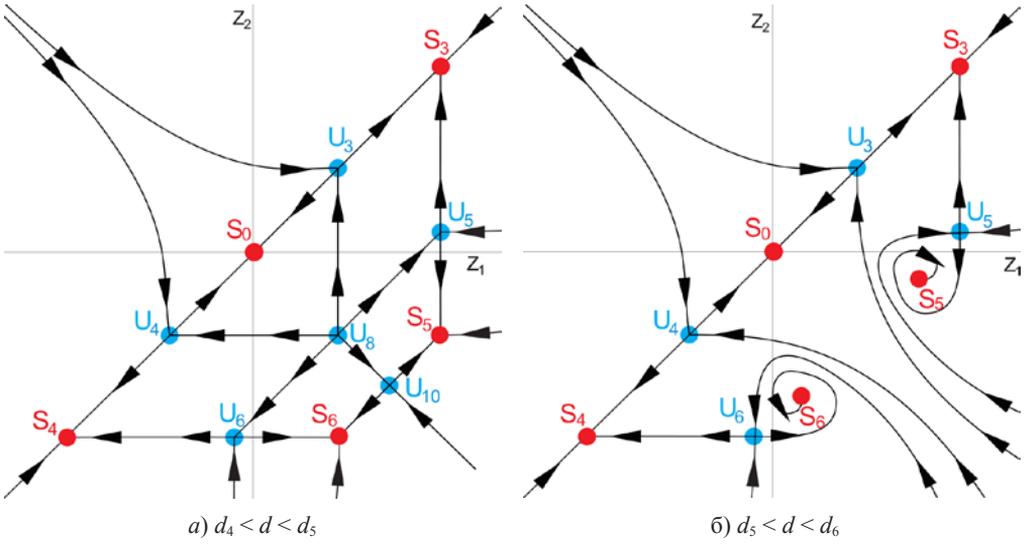


Рис. 8. Фазовые портреты отображения при $(\alpha, \beta) \in A_3$

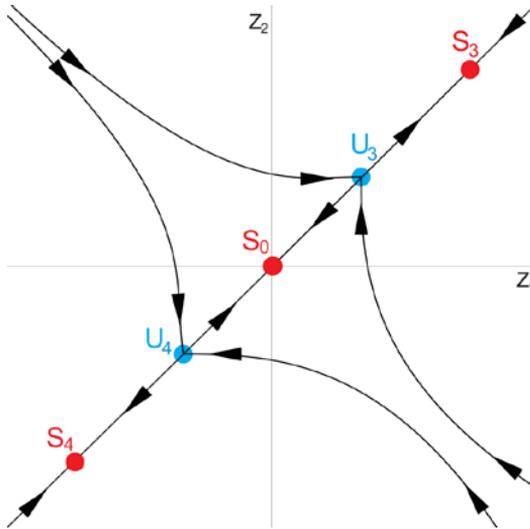


Рис. 9. Фазовый портрет отображения при $d_6 < d < d_7$

5) При $d_2 < d < d_3$; $d_3 = 0,053$ отображение имеет четыре устойчивые неподвижные точки. Фазовый портрет изображен на рис. 7а.

6) При $d = d_3$ неустойчивый узел U_7 сливается с седловой точкой U_9 .

7) При $d_3 < d < d_4$; $d_4 = 0,055$ отображение по-прежнему имеет четыре устойчивые неподвижные точки. Фазовый портрет в данном случае имеет изображенный на рис. 7б вид.

8) При $d = d_4$ от устойчивого самосимметричного узла S_{56} ответвляется пара симметричных друг другу устойчивых узлов S_5 и S_6 и самосимметричная седловая точка U_{10} .

9) При $d_4 < d < d_5$; $d_5 = 0,057$ отображение имеет пять устойчивых неподвижных точек. Фазовый портрет в данном случае изображен на рис. 8а.

10) При $d = d_5$ неустойчивый узел U_8 сливается с седловой точкой U_{10} . Одновременно с этим симметричные устойчивые неподвижные точки S_5 и S_6 меняют свой тип и становятся устойчивыми фокусами.

11) При $d_5 < d < d_6$; $d_6 = 0,069$ отображение имеет пять устойчивых неподвижных точек. Фазовый портрет в данном случае имеет изображенный на рис. 8б вид.

12) При $d = d_6$ симметричные друг другу устойчивые фокусы S_5 и S_6 сливаются с симметричными седлами U_5 и U_6 и исчезают.

13) При $d_6 < d < d_7$; $d_7 = 0,170$ отображение имеет три устойчивые неподвижные точки и две неустойчивые. Фазовый портрет в данном случае изображен на рис. 9.

14) Наконец, при $d = d_6$ симметричные друг другу устойчивые узлы S_3 и S_4 сливаются с симметричными седлами U_3 и U_4 и исчезают. При $d > d_6$ отображение имеет единственное нулевое устойчивое состояние.

Полученный сценарий фазовых перестроек характерен для любых $(\alpha, \beta) \in A'_3$.

Отличия от рассмотренного случая могут состоять в очередности бифуркации исчезновения симметричных друг другу устойчивых узлов S_1 и S_2 и бифуркации слияния неустойчивого узла U_7 с седлом U_9 . Аналогичные изменения порядка возможны с узлами S_5 и S_6 и неустойчивыми точками U_8 и U_{10} .

Для области A'_3 присуще образование пары устойчивых фокусов из точки S_{12} , а из S_{56} наоборот – пары устойчивых узлов. Также для данной области характерны различные сценарии исчезновения точек типа фокус. Были выявлены более сложные случаи возникновения вокруг точек типа фокус устойчивого и неустойчивого инвариантных кривых. Тем не менее, описанные отличия не влияют на количество устойчивых неподвижных точек – их число с увеличением параметра d по-прежнему следует согласно описанному сценарию: $5 \rightarrow 6 \rightarrow 4 \rightarrow 5 \rightarrow 3 \rightarrow 1$.

ЗАКЛЮЧЕНИЕ

На координатной плоскости параметров были выделены области, соответствующие различным бифуркационным сценариям. Благодаря численному анализу модельного отображения с помощью приложения, осуществляющего поиск устойчивых режимов на основе алгоритмов параллельных вычислений, в каждой из областей были подробно рассмотрены основные перестройки, происходящие в фазовом пространстве отображения. Также были установлены множества значений параметров отображения (3), при которых возможно сосуществование семи, шести или пяти устойчивых неподвижных точек.

ЛИТЕРАТУРА

[Глызин С.Д. и др., 2011а] Глызин С.Д., Колесов А.Ю., Розов Н.Х. Релаксационные автоколебания в нейронных системах I // Дифференциальные уравнения, 2011. Т. 47. № 7. С. 919 – 932.

- [Глызин С.Д. и др., 2011б] *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Релаксационные автоколебания в нейронных системах II // Дифференциальные уравнения, 2011. Т. 47. № 12. С. 1675 – 1692.
- [Глызин С.Д. и др., 2012а] *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Релаксационные автоколебания в нейронных системах III // Дифференциальные уравнения, 2012. Т. 48. № 2. С. 155 – 170.
- [Глызин С.Д. и др., 2012б] *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Дискретные автоволны в нейронных системах // Журнал вычислительной математики и математической физики. 2012. Т. 52, № 5. С. 840–858.
- [Глызин С.Д. и др., 2013] *Глызин С.Д., Колесов А.Ю., Розов Н.Х.* Моделирование эффекта взрыва в нейронных системах // Математические заметки. 2013. Т. 93, № 5. С. 682–699.
- [Ивановский Л.И., Самсонов С.О., 2014] *Ивановский Л.И., Самсонов С.О.* Фазовые перестройки одной двумерной динамической системы с импульсным воздействием // Модел. и анализ информ. систем, 2014. Т. 21, № 6. С. 179 – 181.
- [Кащенко С.А., Майоров В.В., 2009] *Кащенко С.А., Майоров В.В.* Модели волновой памяти // М.: Либроком, 2009. 288 с.

DYNAMICS OF TWO-DIMENSIONAL MAPPING AND STABLE REGIMES OF SINGULARLY PERTURBED NEURON SYSTEM

L.I. Ivanovsky, S.O. Samsonov

P.G. Demidov Yaroslavl State University (P.G. Demidov YSU)

Let us consider a mathematical model of neuron system, which is presented as a chain of three diffused connected and singularly perturbed nonlinear differential equations. In the further text researched the questions of existence, stability and asymptotic forms of relaxation periodic solutions of this system due to a bifurcation analysis of two-dimensional mapping. Also the special attention is paid to the number of coexisting stable regimes, which were detected due to asymptotic and computational methods. For searching all stable regimes of model mapping used algorithms of parallel calculations.

This work was supported by the Russian Science Foundation (project nos. № 14-21-00158).

Keywords: relaxation oscillations, stable regimes, bifurcations.

Ivanovsky Leonid Igorevich – 5th year student of IT Faculty of P.G. Demidov Yaroslavl State University, laboratory researcher at Scientific Center in Chernogolovka RAS, leon19unknown@gmail.com

Samsonov Sergey Olegovich – 5th year student of IT Faculty of P.G. Demidov Yaroslavl State University, therealfac@gmail.com

ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ СУПЕРКОМПЬЮТЕРОВ ДЛЯ МАСШТАБИРУЕМОГО ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ ЗАДАЧ ГИДРОДИНАМИКИ В ИНДУСТРИАЛЬНЫХ ПРИЛОЖЕНИЯХ

А.С. Козелков, Р.М. Шагалиев, В.В. Курулин, А.В. Ялозо, С.В. Лашкин

Федеральное государственное унитарное предприятие
Российский федеральный ядерный центр
Всероссийский научно-исследовательский институт экспериментальной физики
(ФГУП «РФЯЦ-ВНИИЭФ»)

В работе проводится анализ двух основных проблем на пути эффективного применения методов вычислительной гидродинамики в промышленных приложениях – моделирование турбулентности и ускорение расчетов. Приводятся результаты исследования потенциала применения вихреразрешающих подходов для моделирования турбулентности в промышленных приложениях при использовании произвольных неструктурированных расчетных сеток. Предложены алгоритмы ускорения масштабируемых высокопроизводительных вычислений, основанные на применении многосеточных технологий.

Работа выполнена при финансовой поддержке РФФИ – проект «офи_м» № 13-0712079.

Ключевые слова: численное моделирование, суперкомпьютеры, вычислительная гидродинамика, турбулентность, многосеточные технологии, промышленные приложения.

ВВЕДЕНИЕ

Быстрое развитие в последние десятилетия различных областей науки и техники явилось следствием колоссального прогресса в применении численных методов в исследовании гидродинамики жидкостей, как теоретической, так и экспериментальной. Вычислительная гидродинамика по своей полноте, разнообразию задач и, что особенно актуально, практическому применению выделилась в отдельную самостоятельную науку. Практическому внедрению вычислительной гидродинамики способствует бурный рост производительности вычислительной техники, введение в строй суперкомпьютеров терафлопного и даже петафлопного класса, а также серьезные успехи, достигнутые в последние годы в области построения эффективных численных методов. Однако, несмотря на существенный прогресс в этой области, точное описание развития и предсказания характеристик многих гидродинамических процессов остается скорее исключением,

Козелков Андрей Сергеевич – к.ф.-м.н., начальник научно-исследовательской лаборатории, askozelkov@mail.ru

Шагалиев Рашит Мирзагалиевич – д.ф.-м.н., начальник математического отделения, первый заместитель директора, R.M.Shagaliev@vniief.ru

Курулин Вадим Викторович – математик, kurulin@mail.ru

Ялозо Андрей Владимирович – научный сотрудник, A.V.Yalozo@itmf.vniief.ru

Лашкин Сергей Викторович – начальник группы, S.V.Lashkin@itmf.vniief.ru

чем правилом. В первую очередь к таким процессам относятся турбулентность, горение, течения в анизотропных пористых средах, многофазные течения и другие. Усугубляет ситуацию и тот факт, что широкое практическое внедрение даже уже разработанных и математически обоснованных методов вычислительной гидродинамики возможно лишь при использовании в расчетах «качественных» расчетных сеток, состоящих из многогранников правильной формы.

Совершенствование вычислительных технологий идет по пути перехода от анализа отдельных конструктивных элементов к все более и более всеобъемлющему трехмерному анализу поведения технологических комплексов в целом. Достижения вычислительной гидродинамики позволяют уйти от расчета обособленных частей и включить в рассмотрение смежные области, а также специфические технологические особенности, которые влияют на эффективность работы промышленного изделия в целом. Требования, предъявляемые к процессу математического моделирования поведения промышленных изделий, не могли не сказаться на развитии методов вычислительной гидродинамики. В настоящее время все больший и больший акцент делается на построении эффективных методов расчета на произвольных неструктурированных сетках, состоящих из многогранников произвольной формы.

Длительный опыт применения методов вычислительной гидродинамики в интересах решения индустриальных задач показал, что из всех нерешенных проблем наиболее важными с фундаментальной точки зрения являются турбулентность и время проведения расчета. Обе эти проблемы достаточно тесно связаны друг с другом.

В представленной статье детально анализируются обе проблемы и возможность их эффективного решения при использовании масштабируемых высокопроизводительных вычислений.

ТУРБУЛЕНТНОСТЬ

Численное моделирование турбулентности опирается на несколько подходов, основанных на описании вихревых структур различных масштабов, среди которых выделяют три основных, а именно: прямое численное моделирование DNS (Direct Numerical Simulation) [Mozer et al., 1999], моделирование крупных вихрей LES (Large Eddy Simulation) [Lesieur, 2008] и решение осредненных по Рейнольдсу уравнений Навье-Стокса RANS (Reynolds-Averaged Navier-Stokes) [Волков, Емельянов, 2010].

Наиболее применимыми на практике являются RANS модели турбулентности, которые характеризуются устойчивым итерационным процессом и приемлемыми результатами для большинства практически важных типов течений [Волков, Емельянов, 2010]. Однако они не являются универсальными и подходящими для решения широкого круга прикладных задач, и это налагает определенные требования к их применимости на практике. Данного недостатка лишен «свободный от эмпиризма» DNS подход, применение которого на практике ограничено

в виду потребности в огромных вычислительных ресурсах [Spalart, 2010]. Альтернативой DNS является «практически свободный от эмпиризма» LES подход, применение которого требует определенного качества дискретных моделей, что проявляется в существенно более мелких сетках по сравнению с RANS. Особенно это характерно для областей вблизи твердых стенок, где вихревые структуры имеют довольно малые размеры и требования к сеткам для LES довольно значительны [Spalart, 2001].

Понимание того, что применение LES подхода возможно только в очень отдаленной перспективе, послужило толчком для создания альтернативных моделей. Эти модели базируются на постулате о том, что в течение ближайших десятилетий на практике будут применяться RANS модели, а для моделирования крупных отрывных вихревых структур вдали от стенок, в силу намного меньших требований к вычислительным ресурсам по сравнению с DNS, можно использовать LES модели. Такие модели получили название гибридных RANS-LES моделей [Spalart, 2001; Travin et al., 2002].

Исторически первой гибридной моделью является метод отсоединенных вихрей (DES) [Travin et al., 2002]. Исследованию применения DES подхода для моделирования турбулентных течений посвящено уже достаточно много работ [Spalart, 2001; Travin et al., 2002; Revell et al., 2008; Козелков и др., 2014б]. Практика его применения в большей своей части сводится к расчетам на многоблочных структурированных сетках с применением схем высокого порядка точности вплоть до седьмого [Козелков и др., 2014б]. Использование сеток такого класса вызывает определенные проблемы при моделировании течений в промышленных конструкциях сложной геометрической конфигурации, где построение блочно-структурированной сетки возможно, но это неоправданно дорого и крайне неэффективно. При расчетах прикладных задач вычислительной гидродинамики все большее предпочтение отдается неструктурированным расчетным сеткам, состоящим из многогранников произвольной формы с выделением призматических пограничных слоев. Использование произвольной неструктурированной сетки приводит к ряду трудностей, одна из которых заключается в том, что на неструктурированной сетке заметно сужается круг доступных схем для дискретизации конвективных потоков. От свойств выбранной схемы дискретизации напрямую зависит качество DES моделирования: схема должна иметь малую диссипацию и обеспечивать устойчивость счета на произвольной неструктурированной сетке. Этот аспект является основополагающим для применения DES моделей в интересах решения индустриальных задач.

Исследование применимости схем для использования DES-моделей можно провести на задаче о течении за обратным уступом, которая является достаточно универсальным и общепризнанным тестом для оценки корректности моделирования турбулентных течений [Зайков и др., 1996; Vogel, Eaton, 1985]. Исследуемое течение содержит отрыв потока от поверхности канала и его последующее присоединение к нижней стенке, а применение DES для решения этой задачи приводит к заметному улучшению результатов по сравнению с RANS [Gritskevich et al., 2011]. Технология моделирования для сеток всех типов одинаковая: сначала

осуществляется стационарный расчет с использованием модели SST, далее производится нестационарный расчет задачи с использованием модели IDDES. Начальными условиями для нестационарного расчета выступают результирующие поля стационарного расчета.

Общее время моделирования нестационарного течения составляет 1 секунду, при этом в течение первых 0,5 секунд течение выходит на статистически установившейся режим, после чего в течение следующих 0,5 секунд накапливаются статистические данные для определения осредненных характеристик течения, по которым определяется коэффициент трения C_f вдоль нижней стенки.

Результаты расчета на блочно-структурированной сетке представлены на рисунке 1. Результаты IDDES по отношению к RANS заметно ближе к эксперименту, с хорошей точностью предсказано как положение точки присоединения вихря, так и минимальное и максимальное значение коэффициента трения. Этот результат будем считать эталонным и применять для сравнения с результатами, полученными на неструктурированных сетках.

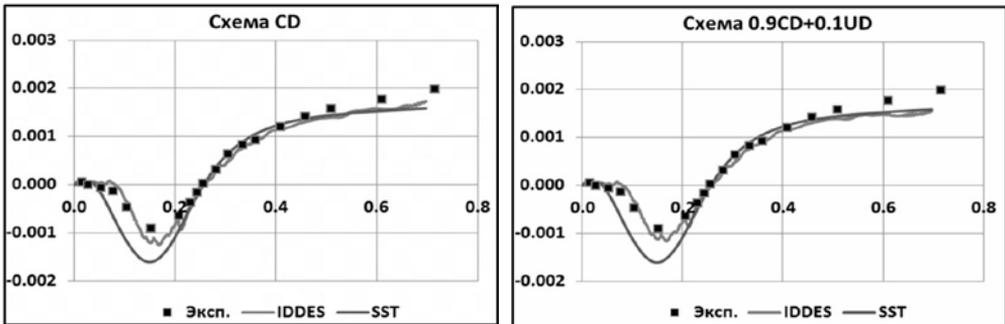


Рис. 1. Коэффициент трения на блочно-структурированной сетке

Для анализа эффективности DES-моделирования на полиэдральных и тетраэдральных сетках используются 7 расчетных сеток: 3 полиэдральных, 2 тетраэдральных и 2 сетки из треугольных призм. В таблице 1 приведены их основные характеристики.

Таблица 1. Характеристики неструктурированных сеток

№	имя	тип ячеек	кол-во ячеек	размер ячеек в области LES, м
1	P1	полиэдры	1,2млн.	0,036
2	P2	полиэдры	2,0млн.	0,027
3	P3	полиэдры	3,2млн.	0,022
4	PR1	треугольные призмы	2,0 млн.	0,036
5	PR2	треугольные призмы	4,0 млн.	0,025
6	T1	тетраэдры	3,8 млн.	0,034
7	T2	тетраэдры	6,2 млн.	0,021

Результаты моделирования для неструктурированных сеток представлены на рисунке 2. На полиэдральной расчетной сетке P1 хороших результатов получить не удалось – положение точки присоединения вихря сильно сдвинуто вниз по потоку. Данная сетка имеет характерный размер ячеек, сопоставимый с блочно-структурированной сеткой, однако полиэдральные элементы заметно повышают диссипативность схемы, и это плохо сказывается на точности моделирования вихрей в области LES. Эта проблема частично устраняется путем измельчения сетки: на сетке P2 положение точки присоединения ближе к его экспериментальному значению. На самой мелкой сетке P3 результаты лучше и приближаются к результатам, полученным на блочно-структурированной сетке.

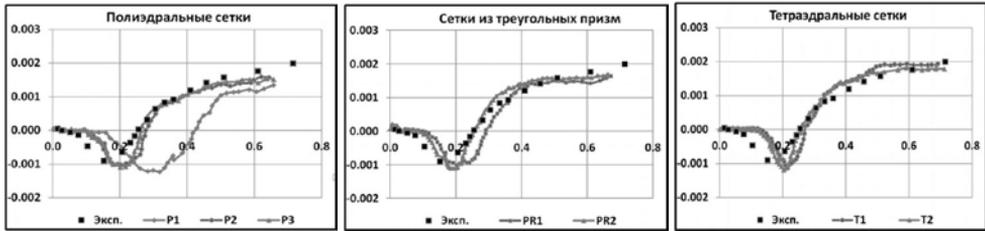


Рис. 2. Коэффициент трения – полиэдральные сетки P1-P3 (слева), сетки из треугольных призм PR1-PR2 (по центру) и тетраэдральные сетки (справа)

Использование сеток, составленных из треугольных призм, дает схожую картину. Разрешения сетки PR1 не хватает для правильного описания течения в основной области, в то время как результаты, полученные на сетке PR2, состоящей из 4 млн. ячеек и имеющей в 1,5 раза меньший характерный размер ячеек, схожи с результатами, полученными с помощью блочно-структурированной сетки в 1,05 млн. ячеек (рис. 2 (по центру)).

При использовании тетраэдральных сеток приемлемый результат получается на сетке T2, которая содержит 6 млн. ячеек с характерным размером в 1,7 раза меньшим, чем на блочно-структурированной сетке (рис. 2 (справа)). Результаты показывают, что использование в основной области неструктурированных сеток требует уменьшения характерного размера ячеек в основной области, что заметно повышает необходимое общее число ячеек. Для получения приемлемых результатов на полиэдральной сетке число ячеек в основной области должно быть как минимум в 2–3 раза больше, чем на гексагональной сетке, а характерный размер полиэдральных ячеек должен быть в 1,3 раза меньше, чем характерный размер шестигранников. Для сетки, составленной из треугольных призм, приемлемый результат был получен лишь на сетке, содержащей в 4 раза больше ячеек с характерным размером 0,025 м, что в 1,4 раза меньше, чем характерный размер ячеек блочно-структурированной сетки. Использование тетраэдральной сетки требует уменьшения характерного размера в 1,6 раза до 0,021 м. Чтобы обеспечить такой характерный размер тетраэдров в основной области, LES требуется около 6 млн. расчетных ячеек, что в 6 раз больше, чем число ячеек на блочно-

структурированной сетке. Более детальную оценку эффективности схем дискретизации конвективных потоков, а также рекомендуемые значения констант для DES-моделей можно найти в [Козелков и др., 2014а].

Необходимо сказать, что в данной задаче рассматривались лишь осредненные характеристики течения, а именно – осредненный коэффициент трения. В случае расчета нестационарных характеристик течения использование неструктурированных сеток может потребовать дальнейшего уменьшения характерного размера ячеек. Расчетная область данной задачи не содержала регионов с «плохой» геометрией (острые кромки, плохо обтекаемые элементы) и регионов с сильно изменяющимися градиентами скоростей, поэтому схема $0.9CD+0.1UD$ была вполне приемлемой для устойчивого счета. В других случаях ее применение не гарантирует устойчивости с такими коэффициентами смещения. Увеличение в формулировке разностной схемы доли «противопоточности» (до 20–30%) не решает проблему, т.к. схема становится сильно диссипативной.

Основываясь на данном исследовании, можно сказать, что при использовании вихреразрешающих подходов и конечно-объемной технологии дискретизации оправданным является использование лишь преимущественно шестигранных расчетных сеток. Оценка необходимого характерного размера расчетных ячеек для моделирования конкретного течения с использованием модели DES является нетривиальной задачей [Spalart, 2001]. В случае задач внутренней гидродинамики характерные размеры ячеек можно оценить из условия разрешения пограничного слоя. В работе [Piomelli, Balaras, 2002] показано, что безразмерная величина ячейки вдоль потока Δx^+ и поперек потока Δy^+ должны удовлетворять следующим требованиям: $\Delta x^+ < 60$, $\Delta y^+ < 25$ и вблизи твердой стенки $\Delta y_{wall}^+ < 1$, где:

$$\Delta x^+ = \frac{\rho u_t \Delta x}{\nu}, \quad \Delta y^+ = \frac{\rho u_t \Delta y}{\nu}. \quad (1)$$

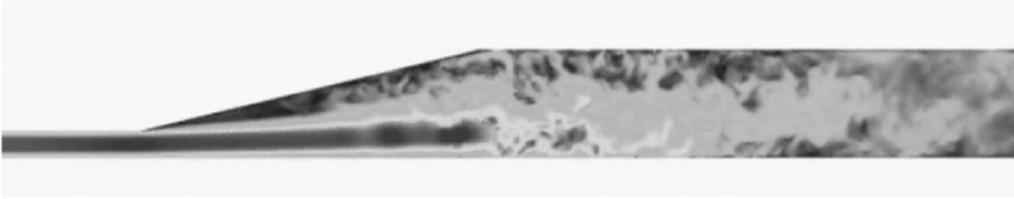
Здесь ρ – плотность среды, ν – кинематическая вязкость, u_t – безразмерная скорость в ближайшей точке твердой поверхности, Δx и Δy – характерный размер ячеек вдоль и поперек потока. Выполнение такого требования гарантирует получение правильного решения в пограничном слое и является общеупотребимым в задачах с использованием вихреразрешающих моделей турбулентности [Piomelli, Balaras, 2002].

Для типичной задачи внутренней гидродинамики – задача о смешении разнотемпературных теплоносителей [Odemark et al., 2009] при числе Рейнольдса $Re = 1,5 \cdot 10^4$ – такие оценки дают примерно $15 \cdot 10^6$ ячеек расчетной сетки и $3 \cdot 10^4$ шагов по времени. Для турбулентного течения теплоносителя в трубе с охлаждением с числом $Re = 5 \cdot 10^5$ уже потребуется более $1 \cdot 10^9$ ячеек и не менее $1 \cdot 10^5$ шагов по времени.

При наличии больших зон отрыва (обтекание препятствий, уступов, крыловых профилей), а также в сложных геометрических конфигурациях подобный критерий может дать лишь очень грубую оценку необходимого сеточного разрешения, и в этом случае для качественного использования вихреразрешающих

моделей требуются расчеты на нескольких последовательно сгущающихся расчетных сетках [Spalart, 2001]. Практика использования модели DES показывает, что для получения хороших результатов в задаче обтекания упрощенной модели автомобиля [Aljume et al., 2014] необходимо порядка $5 \cdot 10^6$ ячеек. В случае обтекания реальной сеточной модели автомобиля и измерения аэроакустических характеристик приемлемая расчетная сетка может легко превысить порог в миллиард ячеек. Для расчета задач такого класса нужны высокоэффективные численные методы решения систем линейных алгебраических уравнений и алгоритмы распараллеливания вычислений. Кроме того, что для использования метода DES нужны сетки достаточной дискретизации, еще одной сложностью его использования в промышленных приложениях является то, что в нем область модели LES определяется автоматически по наличию турбулентных составляющих течения и характерному размеру сеточного разрешения [Gritskevich et al., 2011]. Этот метод не содержит механизмов создания искусственных турбулентных пульсаций в поле скорости при переходе от RANS решения к LES, данные пульсации образуются естественными условиями в процессе течения жидкости. Это приводит к появлению переходных зон, где уже функционирует модель LES, но характерного турбулентного контента в ней еще недостаточно [Spalart, 2009]. При расчете течений с большими отрывными областями такие зоны малы и не приводят к заметному ухудшению решения, однако в задачах с небольшими областями отрыва, характерными, например, для внутренних течений, наличие переходных зон может существенно ухудшить качество DES-моделирования [Spalart, 2009].

Данного недостатка лишен зонный RANS-LES подход с применением явной, искусственной генерации турбулентных флуктуаций скорости на входных границах LES области [Jarrin et al., 2009; Адамьян, 2011]. Параметрами для генерации подобных флуктуаций выступают скорость и компоненты тензора напряжений Рейнольдса, полученные из RANS решения, а в качестве генератора используются специализированные алгоритмы [Jarrin et al., 2009; Адамьян, 2011; Menter, 1994]. В большинстве работ в качестве базовой используется модель RANS SST [Menter, 1994]. В задачах с существенно ассиметричными зонами такие модели дают существенную погрешность в определении тензора рейнольдсовских напряжений [Menter et al, 2009], что на границе областей RANS-LES приводит к погрешностям в генерации турбулентных пульсаций. Недостаточная точность в генерации пульсаций, в свою очередь, приводит к увеличению за границей RANS-LES переходной зоны, что при расчете задач требует дополнительного увеличения области LES и, следовательно, увеличения количества ячеек в расчетной сетке. Выходом здесь является использование моделей класса RSM [Menter et al, 2009]. В работе [Козелков и др., 2014a] приводится алгоритм, основанный на модели EARSM, позволяющий увеличить точность расчета для данного подхода. На рисунке 3 показаны мгновенные поля скорости, полученные по модели EARSM-LES и модели DES при расчете течения в прямоугольном несимметричном диффузоре, в котором появляется анизотропия гидродинамических характеристик.



а) Использование DES;



б) Использование EARSМ-LES с генератором синтетической турбулентности

Рис. 3. Мгновенное поле скорости

Из рисунков видно, что модель DES затягивает турбулизацию потока, что приводит к большой погрешности в профиле осредненной скорости. Кроме того, вследствие отрицательного градиента давления в канале образуется обширная отрывная зона, нахождение положения и формы которой, в свою очередь, является непростой задачей. В работе [Козелков и др., 2014а] показано, что при использовании явной генерации турбулентного контента структура течения меняется в нужную сторону и профиль скорости становится ближе к экспериментальным данным (рис. 4). Кроме того, увеличивается точность расчета благодаря использованию в качестве базовой модели EARSМ.

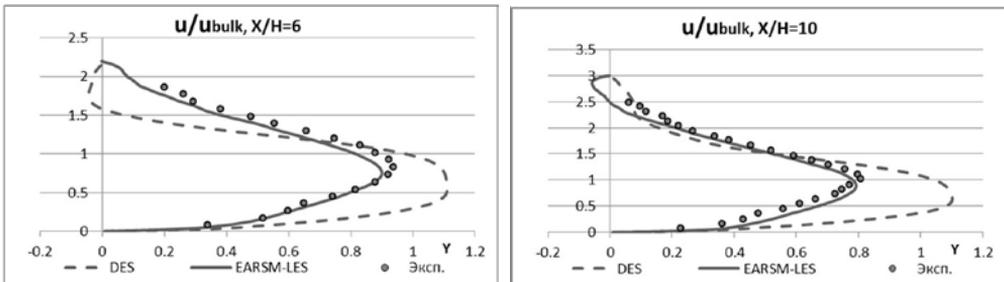


Рис.4. Профили продольной скорости

Зонный RANS-LES подход оказывается достаточно эффективным при расчете с любым размером области отрыва [Козелков и др., 2014в]. Использование данной технологии в промышленных расчетах позволит выделить требуемые для уточнения расчетных характеристик зоны и применить в них генератор синтетической турбулентности, который совместно с моделью EARSМ позволит при гораздо меньших вычислительных ресурсах более точно промоделировать рассматриваемое течение.

СКОРОСТЬ СЧЕТА

Средняя расчетная сетка, применяемая в настоящее время на практике для решения задач вычислительной гидро-, аэродинамики в промышленных приложениях с использованием RANS моделей турбулентности, составляет порядка 50–100 млн. ячеек, а среднее время расчета до установления в зависимости от задачи – более суток. Для использования численных методов при проектировании современных промышленных изделий потребуется решать несколько постановок задач с несколькими десятками вариантов параметров потока. Скорость счета здесь играет очень большую роль. И, естественно, ситуацию здесь усугубляет тот факт, что для разрешения вихревых структур различных масштабов с использованием LES и DES моделей турбулентности потребуются не только более качественные сетки, но и более продолжительное время счета.

Большая часть течений, моделируемых на практике, решаются в «квази-стационарном» режиме, для которого предпочтительно использование неявных схем дискретизации. Применение неявных схем снимает ограничения на шаг интегрирования по времени, который выбирается, исходя из соображений необходимой точности вычислений. Неявная дискретизация уравнений Навье–Стокса порождает систему разностных уравнений, которая решается тем или иным итерационным методом. Выбор подходящего итерационного метода и его реализация во многом определяют общую эффективность вычислительного алгоритма.

В большинстве практических задач дискретизация уравнений Навье–Стокса дает систему разностных уравнений с плохо обусловленной матрицей. Число обусловленности для большинства задач составляет 10^7 – 10^8 , а для некоторых течений приближается к 10^{10} [Козелков и др., 2014в]. Это является причиной того, что наиболее затратным этапом является решение матрицы СЛАУ, которое при использовании классических итерационных методов занимает около 90% вычислительного времени расчетного шага. При этом классические итерационные методы либо перестают работать, либо дают очень медленную скорость сходимости. Одним из универсальных методов решения систем разностных уравнений является многосеточный метод [Stuben, Trottenberg, 1982], основанный на использовании последовательности вложенных сеток и операторов перехода от одной сетки к другой. Данный метод в будущем будет основным при решении промышленных задач, однако при решении плохо обусловленных матриц он требует аккуратной и оптимальной реализации. В литературе описывается достаточно большое число реализаций алгебраических многосеточных методов, многие из которых имеют открытый исходный код как, например, Boomer AMG. Обзор алгебраических многосеточных методов приводится в работе [Волков и др., 2013]. Однако его реализация на системах петафлопсного класса требует специфических алгоритмов распараллеливания, особенно в многосвязных областях.

При распараллеливании многосеточного метода огрубление матриц при переходе от уровня к уровню происходит независимо на каждом MPI-процессе

[Козелков и др., 2014в]. Процесс огрубления в параллельном режиме порождает две проблемы. Во-первых, огрубление останавливается в случае, если на каждом процессе осталось по одной ячейке. Во-вторых, на грубых уровнях, где размерность матриц невелика, время, затрачиваемое на межпроцессорные обмены, из-за латентности коммуникационной среды начинает многократно превосходить время, затрачиваемое на вычисления. Для решения данных проблем в [Козелков и др., 2014в] предложено выполнять сбор всех матриц небольшого размера на одном процессе, формировать из них глобальный уровень и продолжать огрубление и решение в последовательном режиме (рис. 5 слева).

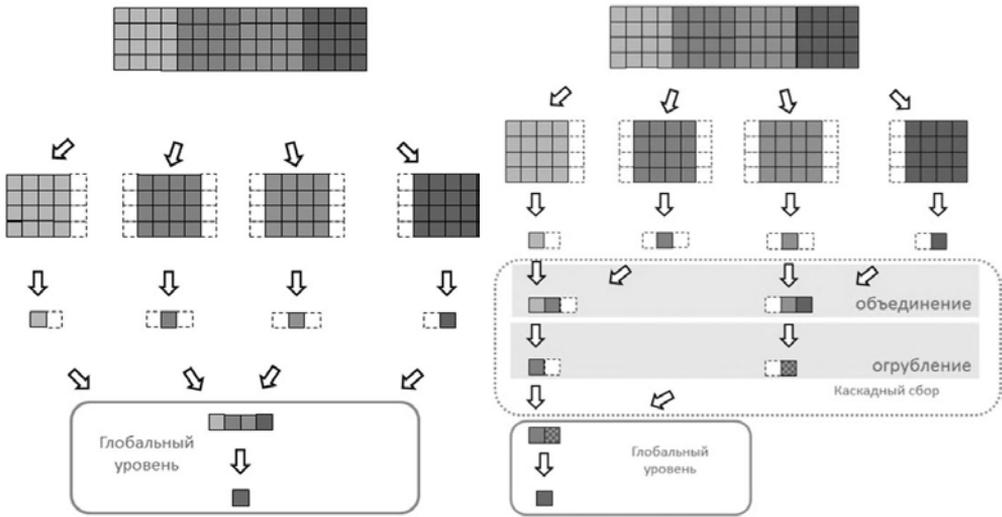


Рис. 5. Формирование глобального уровня (слева) и его каскадный сбор (справа)

Использование глобального уровня позволяет не только завершить процесс огрубления полностью в параллельном случае, но и, из-за хранения глобального уровня на одном процессе, удастся избежать накладных расходов из-за избыточных межпроцессорных обменов при обработке уровней с самыми грубыми сетками.

Для анализа эффективности предложенного алгоритма рассмотрим стационарное изотермическое турбулентное течение вязкой несжимаемой жидкости в прямолинейной гладкой круглой трубе, соответствующее числу Рейнольдса 10000 на трех различных сеточных моделях, соответствующих 40000, 800000 ячеек и 1500000 ячеек. Для каждой из представленных моделей задача решалась в двух постановках с использованием двух методов решения СЛАУ: предобусловленным методом сопряженных градиентов (PCG) и алгебраическим многосеточным методом с агрегативным способом огрубления (AMG). Решение задачи проводилось в параллельном режиме, и для всех вариантов количество процессоров равнялось двенадцати. Результаты решения для матриц давления представлены в таблице 2.

Таблица 2. Полное время решения задачи (в секундах) и число итераций

Решатель СЛАУ	Сетка 1 (40000)		Сетка 2 (800000)		Сетка 3 (1500000)	
	Число итераций	Время решения	Число итераций	Время решения	Число итераций	Время решения
AMG	461	20.66	590	565	1168	3828
PCG	489	62.57	549	16113	1168	158742

Из представленных результатов видно, что время решения задачи при использовании метода сопряженных градиентов для самой грубой сеточной модели увеличилось в 3 раза. Время решения данной задачи для самой подробной сеточной модели с использованием метода сопряженных градиентов составило более чем в 40 раз больше времени по сравнению с AMG. Отметим здесь, что эти цифры не являются характерными для данного класса задач, а зависят от многих факторов, таких, как структура течения, сложность сеточной модели, накладываемые граничные условия и т.д. Опыт практического применения многосеточного метода для решения задач вычислительной гидродинамики показывает, что среднее время увеличения скорости счета по сравнению с методом сопряженных градиентов составляет 4–6 раз, хотя и эти цифры не являются характерными и зависят от типа рассматриваемой задачи.

Для того чтобы показать преимущество использования глобального уровня, было выбрано три различных постановки, отличающихся между собой по количеству используемых MPI-процессов. Для численных экспериментов было выбрано 6, 12 и 24 процессов. Сеточная модель решаемой задачи представлена на рисунке 7 – средняя. Задача решалась в двух вариациях с использованием решателя AMG без использования глобального уровня (AMG1) и с использованием глобального уровня (AMG2). Время решения и число итераций до сходимости решения задачи представлено в таблице 3.

Таблица 3. Полное время решения задачи (секунды) и число итераций

РешательСЛАУ	6 процессов		12 процессов		24 процесса	
	Число итераций	Время решения	Число итераций	Время решения	Число итераций	Время решения
AMG1	1425	647	1377	365	1397	238
AMG2	1469	552	1402	233	1442	104

Из представленных результатов видно, что время решения при использовании решателя с глобальным уровнем уменьшается на 24 процессах более чем в два раза, хотя на 6 процессах это значение равно 1,2 раза. Видна тенденция нелинейного увеличения физического времени счета с увеличением числа процессов. Это обусловлено, в первую очередь, временем, необходимым для передачи сообщений межпроцессорного взаимодействия MPI. Такая тенденция прослеживается во всех задачах гидродинамики, и единственным способом решения данной проблемы является сбор глобальной матрицы и решение ее в последовательном режиме.

Наиболее ярко различие между применением классических и многосеточных технологий решения СЛАУ проявляется при решении задач, составляющих десятки миллионов счетных ячеек, где время счета каждой отдельной задачи может различаться на несколько часов, а то и десятков часов. В качестве демонстрации такого преимущества приведем решение задачи течения в канале за обратным уступом, представленной в разделе «Турбулентность». Результаты решения задачи на разном количестве процессов на сетке в 30 млн. ячеек представлены в таблице 4.

Таблица 4. Результаты решения задачи обтекания обратного уступа

Количество процессов	AMG2		AMG1		PCG	
	Время решения, с	Число итераций	Время решения, с	Число итераций	Время решения, с	Число итераций
100	42701	10295	176973	10179	177169	10414
250	17458	9886	126147	10240	83461	10719
500	11199	9882	72190	10211	54799	9998
1000	12352	10368	75130	10355	23787	10009
2000	15848	10001	95201	9611	19858	10172

Из результатов расчетов следует, что использование глобального уровня в многосеточном методе позволяет значительно сократить время решения задачи. Кроме того, многосеточный метод с глобальным уровнем по времени решения является более предпочтительным, чем метод сопряженных градиентов. Тем не менее, следует отметить, что метод сопряженных градиентов лучше масштабируется на большом числе процессов. Это связано с тем, что решение на глобальном уровне является этапом, выполняющимся в последовательном режиме, и, следовательно, ухудшает масштабируемость. Этот аспект также скажется на ограниченности алгоритма при использовании масштабных сеток в сотни миллионов ячеек. Это хорошо видно из таблицы 5 для расчета задачи на 2000 процессорах. В этом расчете время для AMG и PCG практически сопоставимы, и эта тенденция будет сохраняться при увеличении количества процессоров, но все же эффективность распараллеливания будет зависеть от сложности геометрии и конкретной физической постановки задачи.

На наш взгляд на ближайшую перспективу 3–5 лет данный алгоритм закроет все потребности в масштабируемых численных вычислениях в интересах решения промышленных задач. Его можно эффективно использовать на сетках вплоть до 500 млн. ячеек. Однако при перспективе расширения класса решаемых задач вычислительной гидродинамики, включая усложняющиеся модели для горения и многофазности с использованием свободных от эмпиризма моделей турбулентности, этого алгоритма будет явно недостаточно.

Дискретные модели большой размерности, состоящие из десятков и сотен миллионов ячеек, необходимо решать с использованием тысяч или нескольких десятков тысяч процессоров. Поэтому данный алгоритм становится самым узким местом ввиду плохой масштабируемости: при определенном размере ре-

шаемой задачи, памяти узла, на котором расположен основной процессор, может просто не хватить для построения глобального уровня. Выходом здесь является применение алгоритма каскадных уровней для постепенного уменьшения числа процессов, участвующих в счете, при сохранении при этом преимуществ глобального уровня. Алгоритм каскадных уровней представлен в [Козелков и др., 2014в]. Он включает в себя этап разбиения всего исходного набора остаточных после огрубления матриц на пары и выполнение их попарного объединения. На втором этапе операция повторяется до тех пор, пока не будет получена общая матрица грубого уровня (рис. 5 справа).

Само по себе объединение уровней, помимо формирования информации о новом уровне, содержит в себе процедуру переопределения информации о межпроцессорных обменах, которая так же требует затрат какой-то части процессорного времени. Кроме того, параллельная процедура огрубления каскадного уровня требует введения дополнительных межпроцессорных обменов, которые в случае со скалярной реализацией сбора глобального уровня – отсутствуют. Основным преимуществом каскадной схемы суммирования является наличие масштабируемости алгоритма. При этом фактически снимается ограничение на максимальный размер решаемой задачи из-за возможной нехватки памяти узла, на котором производится формирование глобального уровня. Кроме того, увеличивается скорость построения глобального уровня, поскольку его составные части формируются и огрубляются независимо друг от друга. Результаты применения данного алгоритма для расчета турбулентного течения в круглой трубе представлены в таблице 5.

Таблица 5. Результаты вычислительного эксперимента в круглой трубе

Количество процессов	Скалярная реализация сбора глобального уровня		Каскадный сбор глобального уровня	
	Время построения, с	Время решения, с	Время построения, с	Время решения, с
100	0,003	0,019	0,001	0,026
250	0,025	0,021	0,002	0,104
500	0,099	0,069	0,003	0,137
1000	0,294	0,128	0,004	0,198
1500	0,573	0,550	0,005	0,226

На небольшом процессорном поле время построения глобального уровня составляет незначительную часть от общего времени шага, и, несмотря на то, что каскадное построение во всех случаях работает быстрее, это ускорение нивелируется на фоне общего времени решения СЛАУ. В то же время при увеличении процессорного поля до нескольких сотен видно, что каскадный сбор дает существенное ускорение общего решения СЛАУ, поскольку в данном случае скалярный алгоритм сгенерирует относительно большую матрицу глобального

уровня, огрубление и решение которой проводится в последовательном режиме и требует значительных затрат времени.

Решение задачи обтекания обратного подогреваемого уступа на расчетной сетке 14 млн. дает цифры, представленные в таблице 6.

Таблица 6. Течение за обратным уступом на сетке в 14 млн. ячеек

Количество процессов	Скалярная реализация сбора глобального уровня		Каскадный сбор глобального уровня	
	Время решения, с	Число итераций	Время решения, с	Число итераций
100	13378	6113	11964	5397
250	3977	4484	4466	5581
500	2418	4629	2635	4452
1000	3657	5006	1171	4485
1500	6189	4592	1359	5675

По результатам расчетов, во-первых, можно сделать вывод, что использование процедуры построения глобального уровня дает значительное преимущество по скорости сходимости многосеточного метода, а следовательно и общему времени решения задачи, причем данный эффект усиливается с увеличением процессорного поля. Так, для тестовых расчетов на 1000 и 1500 процессорах использование скалярной реализации алгоритма построения глобального уровня позволило уменьшить общее время решения задачи примерно в 3 раза. Во-вторых, каскадный алгоритм построения глобального уровня дает лучшую производительность по сравнению со скалярной процедурой построения глобального уровня на большом количестве процессоров. Для наглядности, ниже на графике приводятся только результаты, полученные с использованием скалярной версии алгоритма сбора глобального уровня, и каскадной версии, реализованной в рамках данной работы. Для сетки в 30 млн. результаты сведены в таблице 7.

Таблица 7. Течение за обратным уступом на сетке в 30 млн. ячеек

Количество процессов	Скалярная реализация сбора глобального уровня		Каскадный сбор глобального уровня	
	Время решения, с	Число итераций	Время решения, с	Число итераций
125	31238	10482	33127	10828
250	17420	10853	18591	11057
500	11500	11241	11227	10142
1000	9382	9144	6263	9043
2000	17124	8463	8184	11240

Сравнение производительности скалярного и каскадного алгоритмов построения глобального уровня приводятся на рис. 6.

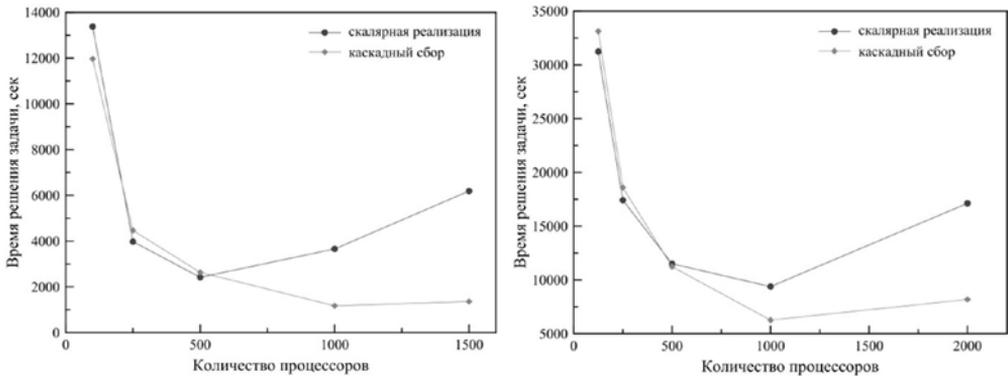


Рис. 6. Сравнение производительности скалярного и каскадного алгоритмов на задаче размерностью 14 млн. ячеек (слева) и 30 млн. ячеек (справа)

Как видно из полученных результатов, использование глобального уровня значительно увеличивает скорость сходимости задачи с увеличением количества процессоров. При использовании 1000 процессоров скалярная процедура построения глобального уровня дает общее ускорение решения задачи примерно в 2 раза. Алгоритм каскадного сбора обеспечивает существенное ускорение процедуры построения глобального уровня. При измельчении задачи менее чем по ~60 тыс. расчетных ячеек на процессор (использовании более 500 процессоров) имеем значительное и возрастающее преимущество по производительности алгоритма каскадного сбора глобально уровня.

Использование глобального уровня положительно сказывается на общей производительности многосеточного метода. При этом использование каскадного алгоритма дает еще большую производительность по сравнению со скалярной процедурой построения глобального уровня. Причем получаемое ускорение растет с увеличением количества процессоров, используемых для решения задачи. Однако в ряде случаев алгоритм каскадного сбора дает незначительное замедление производительности по сравнению со скалярной процедурой построения глобального уровня. Это обусловлено наличием дополнительных межпроцессорных обменов при построении и последующем огрублении иерархии каскадных уровней.

ЗАКЛЮЧЕНИЕ

В статье дана оценка применимости современных подходов к моделированию турбулентных течений в промышленных приложениях. Описана технология применения зонного RANS-LES подхода совместно с моделью EARSM, позволяющая при гораздо меньших вычислительных ресурсах более точно моделировать турбулентные течения. Для эффективного применения многосеточных тех-

нологий предложены специализированные алгоритмы ускорения, основанные на «каскадном» сборе данных с процессоров и глобальном уровне.

ЛИТЕРАТУРА

- [Адамьян, 2011] *Адамьян Д.Ю.* Эффективный метод генерации синтетической турбулентности на входных границах LES области в рамках комбинированных RANS-LES подходов к расчету турбулентных течений // Математическое моделирование, том 23, № 7, 2011, с. 3–19.
- [Волков, Емельянов, 2010] *Волков К.Н., Емельянов В.Н.* Течения и теплообмен в каналах и вращающихся полостях. // М.: Физматлит, 2010, 480 с.
- [Волков и др., 2013] *Волков К.Н., Дерюгин Ю.Н., Емельянов В.Н., Карпенко А.Г., Козелков А.С., Тетерина И.В.* Методы ускорения газодинамических расчетов на неструктурированных сетках // Москва: Физматлит, 2013. 536 с., ISBN 978-5-9221-1542-1.
- [Зайков и др., 1996] *Зайков Л.А., Стрелец М.Х., Шур М.Л.* Сравнение возможностей дифференциальных моделей турбулентности с одним и двумя уравнениями при расчете течений с отрывом и присоединением. Течение в каналах с обратным уступом // Теплофизика высоких температур, 1996, том 34, № 35, с. 724-736.
- [Козелков и др., 2013] *Козелков А.С., Дерюгин Ю.Н., Лашкин С.В., Силаев Д.П., Симонов П.Г., Тятюшкина Е.С.*, Реализация метода расчета вязкой несжимаемой жидкости с использованием многосеточного метода на основе алгоритма SIMPLE в пакете программ ЛОГОС // Журнал ВАНТ, сер. Математическое моделирование физических процессов, 2013, вып. 4, стр. 44-56.
- [Козелков и др., 2014а] *Козелков А.С., Курулин В.В., Тятюшкина Е.С., Пучкова О.Л.*, Моделирование турбулентных течений вязкой несжимаемой жидкости на неструктурированных сетках с использованием модели отсоединенных вихрей // Матем. моделирование, 26:8 (2014), 81–96.
- [Козелков и др., 2014б] *Козелков А.С., Курулин В.В., Пучкова О.Л., Лашкин С.В.* Моделирование турбулентных течений с использованием алгебраической модели рейнольдсовых напряжений с универсальными пристеночными функциями // Вычислительная механика сплошных сред. – 2014. – Т. 7, № 1. – С. 40-51.
- [Козелков и др., 2014в] *Козелков А.С., Шагалиев Р.М., Дерюгин Ю.Н., Курулин В.В., Ялозо А.В., Лашкин С.В.* Исследование потенциала суперкомпьютеров для масштабируемого численного моделирования задач гидродинамики в промышленных приложениях // Тезисы конференции «Супервычисления», г. Саров, Нижегородская обл., 2014, с. 82-83.
- [Aljure et al., 2014] *Aljure D.E. et al.* Flow and turbulent structures around simplified car models // Computers & Fluids 96 (2014), pp. 122–135.
- [Gritskevich et al., 2011] *Gritskevich M.S., Garbaruk A.V., Schütze J., Menter F.R.* Development of DDES and IDDES Formulations for the $k-\omega$ Shear Stress Transport Model // Flow Turbulence Combust, 2011, Vol. 88 (3), 2012, pp. 431-449.
- [Jarrin et al., 2009] *Jarrin N., Prosser R., Uribe J. et al.* Reconstruction of turbulent fluctuations for hybrid RANS/LES simulations using a Synthetic-Eddy Method // International Journal of Heat and Fluid Flow. 2009. Vol. 30, no. 3. Pp. 435–442.
- [Lesieur, 2008] *Lesieur M.* Turbulence in Fluids // Springer, ISBN 978-1-4020-6434-0, 2008, 558 p.

- [Menter, 1994] *Menter F.R.* Two-equation eddy-viscosity turbulence models for engineering applications // *AIAA Journal*, Vol. 32, No. 8, pp. 1598-1605, 1994.
- [Menter et al., 2009] *Menter F.R., Garbaruk A.V., Egorov Y.* Explicit algebraic Reynolds stress models for anisotropic wall-bounded flows // *Proc. of 3rd European Conference for Aero-Space Sciences (EUCASS)*, Versailles, July 6-9th, 2009, 14 p. URL: http://cfd.spbstu.ru/agarbaruk/c/document_library/DLFE-6688.pdf
- [Mozer et al., 1999] *Mozer D., Kim J. & Mansour N.N.* DNS of Turbulent Channel Flow // *Phys. Fluids*, 1999, vol. 11, pp. 943-945.
- [Odemark et al., 2009] *Odemark Y. et al.* High-Cycle Thermal Fatigue in Mixing Tees: New Large-Eddy Simulations Validated Against New Data Obtained by PIV in the Vattenfall Experiment // *Proceedings of the 17th International Conference on Nuclear Engineering*, Brussels, Belgium, 2009, pp. 775-785.
- [Piomelli, Balaras, 2002] *Piomelli U., Balaras E.* Wall-Layer Models for Large-Eddy Simulations // *Annual Review of Fluid Mechanics*. 2002. Vol. 34, no. 1. Pp. 349-374.
- [Revell et al., 2008] *Revell A., Craft T., Laurence D.* Turbulence Modelling of Strongly Detached Unsteady Flows: The Circular Cylinder // *Advances in Hybrid RANS-LES Modelling. Notes on Numerical Fluid Mechanics and Multidisciplinary Design 97*, Springer Berlin Heidelberg, 2008, vol. 97, pp. 279-288.
- [Spalart, 2001] *Spalart P.R.* Young-Person's guide to detached-eddy simulation grids // *Tech. Rep. NASA/CR-2001-211032*, NASA, Langley Research Center, 2001.
- [Spalart, 2009] *Spalart P.R.* Detached-Eddy Simulation // *Annual Rev. Fluid Mech.* 2009. 41:181-202.
- [Spalart, 2010] *Spalart P.R.* Strategies for turbulence modeling and simulations // *Heat Fluid Flow*, 2000, v. 21, pp. 252-263.
- [Stuben, Trottenberg, 1982] *Stuben K., Trottenberg, U.*, Multigrid methods: fundamental algorithms, model problem analysis and applications. *Multigrid Methods, Lecture Notes in Mathematics 960* (eds W. Hackbusch and U. Trottenberg). 1-176. Springer, Berlin, 1982.
- [Travin et al., 2002] *Travin A., Shur, M., Strelets, M., Spalart, P.R.* Physical and numerical upgrades in the detached-eddy simulation of complex turbulent flows, // *Proceedings of Euromech Coll. Les of complex transitional and turbulent flows*, Munich, Germany. Kluwer, Dordrecht, 2002, vol. 65, pp. 239-254.
- [Vogel, Eaton, 1985] *Vogel, J.C. and Eaton, J.K.* Combined Heat Transfer and Fluid Dynamic Measurements Downstream of a Backward-Facing Step // *Journal of Heat Transfer*, 1985, vol. 107, pp. 922 - 929.

CHALLENGES ASSOCIATED WITH SUPERCOMPUTERS AS APPLIED TO HYDRODYNAMIC SCALABLE NUMERICAL SIMULATIONS IN INDUSTRIAL APPLICATIONS

A.S. Kozelkov, R.M. Shagaliyev, V.V. Kurulin, A.V. Yalozo, S.V. Lashkin

RUSSIAN FEDERAL NUCLEAR CENTER

All-Russian Research Institute Of Experimental Physics(RFNC-VNIIEF), Sarov

The paper discusses two major bottlenecks in efficient use of computational hydrodynamics methods in industrial applications: turbulence simulations and speedup. Results of the analysis of turbulence simulation capabilities of eddy-resolving methods in industrial applications on

arbitrary unstructured grids are presented. Speedup algorithms for scalable high-performance computations based on multi-grid technologies are presented.

The work has been funded by the Russian Foundation for Basic Research (RFFI), Project «ofi_m» No. 13-0712079.

Keywords: numerical simulations, supercomputers, computational hydrodynamics, turbulence, multi-grid technologies, industrial applications.

Kozelkov Andrey Sergeyeovich – candidate of physical and mathematical sciences (PhD), chief of research laboratory, askozelkov@mail.ru

Shagaliyev Rashit Mirzagaliyevich – doctor of physical and mathematical sciences, chief of mathematical department, first deputy director, R.M.Shagaliyev@vniief.ru

Kurulin Vadim Viktorovich – mathematician, kurulin@mail.ru

Yalozo Andrey Vladimirovich – research associate, A.V.Yalozo@itmf.vniief.ru

Lashkin Sergey Viktorovich – chief of group, S.V.Lashkin@itmf.vniief.ru

МОДЕЛИРОВАНИЕ КОЛЛАПСА ХОЛОДНОГО БЫСТРОВАЩАЮЩЕГОСЯ ПРОТОЗВЕЗДНОГО ОБЛАКА

С.Г. Моисеенко¹, Г.С. Бисноватый-Коган^{1,2}, Н.В. Арделян³, К.В. Космачевский³

¹ Институт космических исследований Российской академии наук (ИКИ РАН)

² Национальный исследовательский ядерный университет «МИФИ» (Московский инженерно-физический институт, государственный университет) (НИЯУ МИФИ)

³ Московский государственный университет имени М.В. Ломоносова, факультет вычислительной математики и кибернетики (ВМиК МГУ)

Рассмотрена задача о коллапсе холодного быстровращающегося протозвездного облака. Численное двумерное моделирование задачи проведено при помощи полностью консервативной неявной операторно-разностной схемы на треугольной сетке переменной структуры. Показано, что в результате коллапса формируется распределение плотности в экваториальной плоскости в виде диска.

Работа ГСБК и СГМ была частично поддержана грантами РФФИ 14-02-00728 и 14-29-06045, грантом Президента РФ поддержки ведущих научных школ НШ-261.2014.2, и программой президиума РАН П21/3. Работа СГМ была также частично поддержана грантом РФФИ 14-29-06055.

Ключевые слова: протозвездные облака, численное моделирование.

1. ПОСТАНОВКА ЗАДАЧИ О КОЛЛАПСЕ ВРАЩАЮЩЕГОСЯ ХОЛОДНОГО ГАЗОВОГО ОБЛАКА

Задача о коллапсе вращающегося холодного протозвездного газового облака описывается системой уравнений газовой динамики с гравитацией [Ландау, Лифшиц, 1988]:

$$\begin{aligned} \frac{dx}{dt} &= \mathbf{u}, \quad \frac{d\rho}{dt} + \rho \operatorname{div} \mathbf{u} = 0, \\ \rho \frac{d\mathbf{u}}{dt} &= -\operatorname{grad} p - \rho \operatorname{grad} \Phi, \\ \rho \frac{d\varepsilon}{dt} + p \operatorname{div} \mathbf{u} &= 0, \quad \eta = \frac{1}{\rho} = \frac{T\mathfrak{R}}{p}, \quad \varepsilon = \frac{T\mathfrak{R}}{\gamma - 1}, \\ \Delta \Phi &= 4\pi G \rho. \end{aligned} \tag{1}$$

Моисеенко Сергей Григорьевич – заведующий отделом, доктор физико-математических наук, ИКИ РАН, moiseenko@iki.rssi.ru

Бисноватый-Коган Геннадий Семенович – главный научный сотрудник, доктор физико-математических наук, профессор, ИКИ РАН, НИЯУ МИФИ, gkogan@iki.rssi.ru

Арделян Николай Васильевич – ведущий научный сотрудник, доктор физико-математических наук, ВМиК МГУ

Космачевский Константин Викторович – старший научный сотрудник, кандидат физико-математических наук, ВМиК МГУ

где $\frac{d}{dt}$ – субстанциональная производная по времени t , $x = (r, z)$, $\mathbf{u} = (u^r, u^\varphi, u^z)$ – вектор скорости, p – давление, ε – внутренняя энергия, Φ – гравитационный потенциал, ρ – плотность, T – температура, G – постоянная тяготения, \mathfrak{R} – универсальная газовая постоянная, γ – показатель адиабаты. Задача решается в предположении осевой симметрии и симметрии относительно экваториальной плоскости.

В начальный момент времени предполагается, что облако представляет собой однородно вращающийся шар постоянной плотности со следующими параметрами:

$$\begin{aligned} \rho &= 1,492 \cdot 10^{-17} \text{ г/см}^3, \quad p = 1,548 \cdot 10^{-10} \text{ дин/см}^2, \\ r &= 3,81 \cdot 10^{16} \text{ см}, \quad \omega = 2,008 \cdot 10^{-12} \text{ рад/с}, \\ M &= 1,73 M_{\text{sun}} = 3,457 \cdot 10^{33} \text{ г}, \quad \gamma = 5/3, \quad u^r = u^z = 0. \end{aligned} \quad (2)$$

На внешней границе облака давление равно малой постоянной величине ($p = 0,87 \cdot 10^{-13}$ дин/см²). На этой границе значения гравитационного потенциала Φ определяются при помощи интегральной формулы, являющейся аппроксимацией формулы для объемного потенциала [Тихонов, Самарский, 1972].

Для приведения системы уравнений (1) к безразмерному виду все размерные величины в (1) были представлены в виде: $F = F_0 \tilde{F}$, где F_0 – размерный масштабный множитель, \tilde{F} – безразмерная функция. Характерные масштабы были выбраны следующим образом:

$$\rho_0 = 1,492 \cdot \frac{10^{-17} \text{ г}}{\text{см}^3}, \quad r_0 = z_0 = 3,81 \cdot 10^{16} \text{ см},$$

$$t_0 = 5 \cdot 10^{11} \text{ с}, \quad p_0 = \rho_0 t_0^{-2} r_0^2,$$

$$u_0^r = u_0^z = r_0 t_0^{-1}, \quad \Phi_0 = 4\pi G \rho_0 r_0^2,$$

$$\omega_0 = 1/t_0, \quad T_0 = u_0^2 / \mathfrak{R}, \quad \varepsilon_0 = u_0^2.$$

В безразмерных переменных система уравнений (1) записывается в следующем виде (знак \sim опущен):

$$\frac{dx}{dt} = \mathbf{u}, \quad \frac{d\rho}{dt} + \rho \text{div} \mathbf{u} = 0,$$

$$\rho \frac{d\mathbf{u}}{dt} = -\text{grad} p - q \rho \text{grad} \Phi, \quad q = 4\pi G \rho_0 t_0^2 \approx 3.127, \quad (3)$$

$$\rho \frac{d\varepsilon}{dt} + p \text{div} \mathbf{u} = 0, \quad \eta = \frac{1}{\rho} = \frac{T}{p}, \quad \varepsilon = \frac{T}{\gamma - 1},$$

$$\Delta \Phi = \rho.$$

Система уравнений (3) решается при следующих граничных условиях:

$$u^z = 0 \text{ при } z = 0;$$

$$u^r = 0 \text{ при } r = 0$$

на внешней границе $p = 10^{-6}$, значения гравитационного потенциала Φ задаются, как и в [Арделян и др., 1987б] при помощи интегральной формулы.

Начальные условия в безразмерных переменных записываются в следующем виде:

$$\rho = 1, \omega = 1,004, p = 1,78 \cdot 10^{-3}$$

в круге $r^2 + z^2 \leq 1$.

2. МЕТОД РАСЧЕТА

Для моделирования задачи о коллапсе быстро вращающегося холодного протозвездного облака использовалась неявная полностью консервативная схема в Лагранжевых переменных на треугольной сетке переменной структуры [Арделян, 1979; Арделян, 1981; Арделян, 1983а; Арделян, 1983б; Арделян, Космачевский, 1993; Арделян и др., 1987а; Ardeljan et al., 1996].

3. РАСЧЕТ ГРАВИТАЦИОННОГО ПОТЕНЦИАЛА НА ВНЕШНЕЙ ГРАНИЦЕ ОБЛАКА И ГРАВИТАЦИОННОЙ СИЛЫ ВБЛИЗИ ОСИ Z

Значения гравитационного потенциала на внешней границе облака могут быть найдены при помощи формулы для объемного потенциала [Тихонов, Самарский, 1972]:

$$\Phi(x_0) = -\frac{1}{4\pi} \int_V \frac{\rho dV}{R_{xx_0}}, \quad (4)$$

где R_{xx_0} – расстояние от точки x_0 до точки x , ρ – плотность и V – объем облака. Сеточный аналог формулы (4) был получен в [Арделян и др., 1987б] и может быть записан в следующей форме (обозначения см. [Арделян и др., 1987а]):

$$\Phi_q = -\frac{1}{\pi} \sum_{\Delta_i \in \omega_i} m_i \lambda_{iq}, \quad \bar{x}_q \in \gamma_j, \quad (5)$$

где

$$\begin{aligned} \lambda_{iq} &= a_1^{-1} F\left(\frac{\pi}{2}, \kappa_1\right) + a_2^{-1} F\left(\frac{\pi}{2}, \kappa_2\right), \\ a_1^2 &= \tilde{r}_i^2 + r_q^2 + (\tilde{z}_i - z_q)^2 + 2\tilde{r}_i r_q, \\ a_2^2 &= \tilde{r}_i^2 + r_q^2 + (\tilde{z}_i + z_q)^2 + 2\tilde{r}_i r_q, \\ \kappa_1^2 &= 4\tilde{r}_i r_q / a_1^2, \quad \kappa_2^2 = 4\tilde{r}_i r_q / a_2^2, \end{aligned} \quad (6)$$

$$\tilde{r}_i = \frac{1}{3} \sum_{j=1}^3 r_j, \quad \tilde{z}_i = \frac{1}{3} \sum_{j=1}^3 z_j$$

и

$$F\left(\frac{\pi}{2}, \kappa\right) = \int_0^{\frac{\pi}{2}} (1 - \kappa^2 \sin^2 t)^{-1/2} dt \quad (7)$$

полный эллиптический интеграл первого рода. Индекс i – номер ячейки, j – номер узла и q – номер граничного узла. Нетрудно показать, что аппроксимирует (4) с точностью $O(h)$.

Гравитационную силу $\vec{F}=(F_r, F_z)$ можно найти, применяя разностный оператор *grad* к сеточной функции гравитационного потенциала. Применение указанного приема в узлах сетки вблизи оси *z* может привести к большим ошибкам в значениях гравитационной силы \vec{F} .

В начальный момент времени градиент давления в задаче о коллапсе достаточно мал и сравним по величине с численной ошибкой, присутствующей в значениях гравитационной силы, что приводит к ошибочной картине течения. Например, при расчете задачи о коллапсе невращающегося облака с начальными данными из [Арделян и др., 1987б] ($\omega = 0$) распределение плотности, скорости и других физических величин становится несферическим вскоре после начала расчета, хотя задача является сферически симметричной и искажение сферической симметрии связано с численными ошибками.

Для преодоления указанной трудности был предложен следующий способ: вычисление значений гравитационной силы $\vec{F}=(F_r, F_z)$ в узлах сетки, вблизи оси *z* используя только значения сеточной функции плотности. Для этого выражение для объемного потенциала в непрерывной форме (4) было продифференцировано по координатам *r, z*. Разностные аналоги для выражения градиента от объемного потенциала записываются в следующей форме:

$$(\nabla_r \Phi)_k \approx \frac{1}{\pi} \sum_{\Delta_i \in \omega_i} m_i \eta_{ik}, \quad \bar{x}_k \in \omega_x, \tag{8}$$

где

$$\eta_{ik} = \frac{(r_k + r_i)\kappa_1^2 - 2r_i}{a_1^3 \kappa_1^2} \Pi(-\kappa_1^2, \kappa_1, \frac{\pi}{2}) + \frac{2r_i}{a_1^3} F(\frac{\pi}{2}, \kappa_1) + \frac{(r_k + r_i)\kappa_2^2 - 2r_i}{a_2^3 \kappa_2^2} \Pi(-\kappa_2^2, \kappa_2, \frac{\pi}{2}) + \frac{2r_i}{a_2^3} F(\frac{\pi}{2}, \kappa_2), \tag{9}$$

и

$$\Pi(h, k, \frac{\pi}{2}) = \int_0^{\frac{\pi}{2}} \frac{dt}{(1 + h \sin^2 t) \sqrt{1 - k^2 \sin^2 t}}$$

– полный эллиптический интеграл третьего рода;

$$(\nabla_z \Phi)_k \approx \frac{1}{\pi} \sum_{\Delta_i \in \omega_i} m_i \zeta_{ik}, \quad \bar{x}_k \in \omega_x, \tag{10}$$

где

$$\zeta_{ik} = \frac{z_k - z_i}{a_1^3} \Pi(-\kappa_1^2, \kappa_1, \frac{\pi}{2}) + \frac{z_k + z_i}{a_2^3} \Pi(-\kappa_2^2, \kappa_2, \frac{\pi}{2}).$$

Остальные обозначения такие же, как и в (6).

Нетрудно показать, что формулы (8, 10) аппроксимируют свои непрерывные аналоги с точностью $O(h)$. Непосредственное применение формул (8, 10) является достаточно дорогостоящим, так для вычисления значений гравитационной силы в *n* узлах сетки требуется $O(nN)$ операций, где *N* общее число ячеек сетки. Машинное время, необходимое для вычисления значений гравитационной силы по формулам (8, 10) во всех узлах, находящихся вблизи оси *z*, превышает время, требуемое на все остальные вычисления на каждом временном шаге.

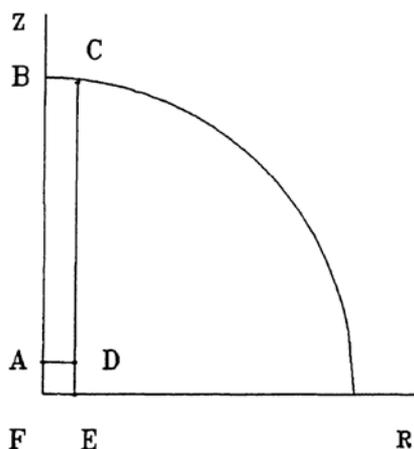


Рис. 1. FBCE – часть расчетной области, в которой корректируются значения гравитационной силы. В области ABCD используется процедура интерполяции (экстраполяции). В области ADEF применяются формулы (8, 10)

Для уменьшения затрат машинного времени используется следующий прием. На каждом временном шаге задача Дирихле для гравитационного потенциала решается для всей расчетной области. Гравитационная сила находится как $grad \Phi$, где Φ – гравитационный потенциал. Затем значения гравитационной силы корректируются вблизи оси z . Внутри области ABCD (Рис. 1) значения z компоненты гравитационной силы F_z экстраполируются в область ABCD и на часть оси z – АВ из части расчетной области, находящейся слева от CD. Из предположения осевой симметрии следует, что $F_r = 0$ на оси z и функция F_r интерполируется внутри ABCD по ее значениям на АВ и CD. В прямоугольнике ADEF значения гравитационной силы определяются по формулам (8, 10). Предложенный прием существенно уменьшает затраты машинного времени и позволяет сохранить достаточную точность при вычислении гравитационной силы.

4. РЕЗУЛЬТАТЫ РАСЧЕТОВ КОЛЛАПСА БЫСТРОВРАЩАЮЩЕГОСЯ ГАЗОВОГО ОБЛАКА

В начальный момент времени в облаке:

$$\alpha = \frac{E_{in}}{|E_{gr}|} = 0,00425, \quad \beta = E_{rot}/|E_{gr}| = 0,324,$$

где E_{in} – внутренняя энергия, E_{gr} – гравитационная энергия, и E_{rot} – вращательная энергия. Выбранные значения параметров α, β лежат существенно в той зоне, где в работе [Boss, 1980] было получено кольцеобразное распределение плотности (т.е. максимум плотности находился на периферии экваториальной плоскости). На Рис. 2 и 3 представлена треугольная сетка и распределение плотности на начальной стадии расчета ($t = 0,04$). На Рис. 4 представлено распределение плотности вдоль оси z в момент времени $t = 0,04$.

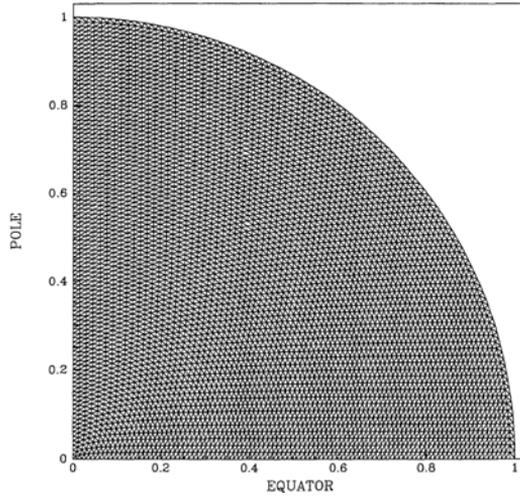


Рис. 2. Треугольная сетка в момент времени $t = 0,04$

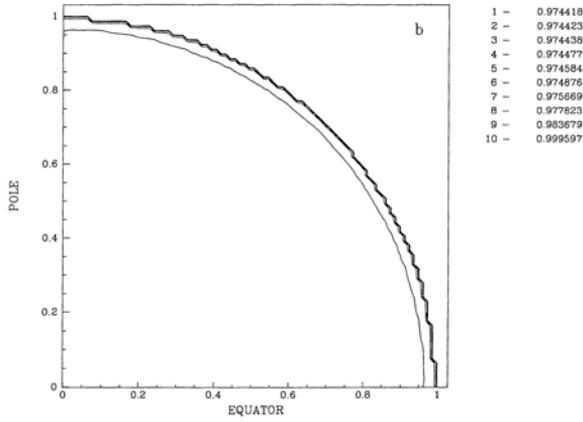


Рис. 3. Линии уровня плотности в момент времени $t = 0,04$

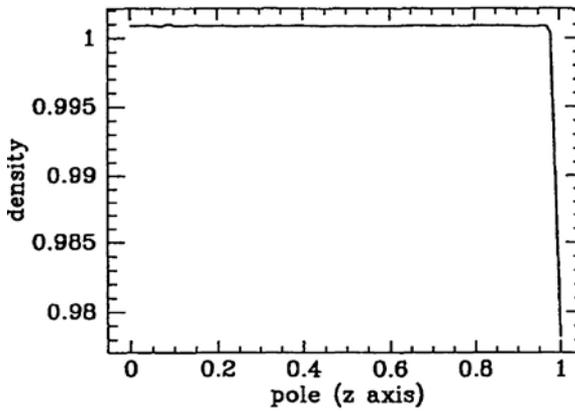


Рис. 4. Распределение плотности вдоль оси вращения z в момент времени $t = 0,04$

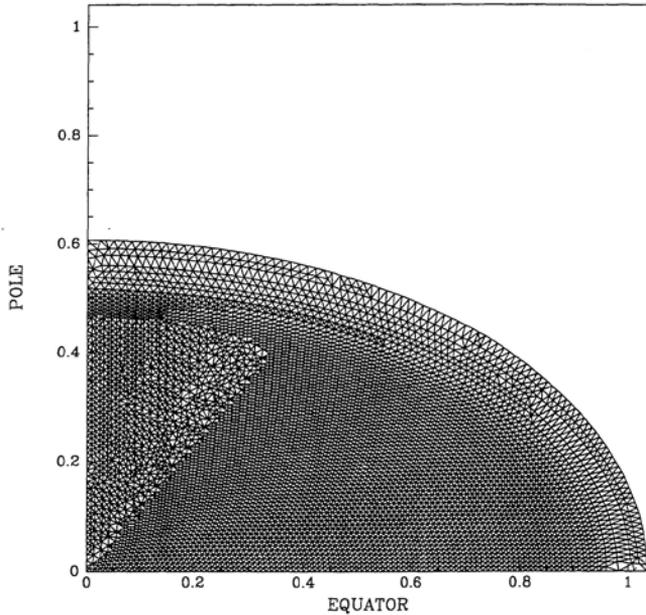


Рис. 5. Треугольная сетка в момент времени $t = 0,94$

На начальной стадии коллапса облако сжимается существенно сильнее вдоль оси z , чем вдоль экваториальной плоскости, что связано с достаточно низким начальным давлением в облаке и быстрым вращением. Влияние гравитационной силы компенсируется в r направлении действием центробежных сил. Следует отметить, что в начальные моменты времени коллапса «гантелевидная» структура в распределении плотности $\rho(z)$ не возникает (Рис. 4), как это имело место в работе [Арделян и др., 19876]. Происхождение этой «гантелевидной» структуры было вызвано численными ошибками при вычислении гравитационной силы. На Рис. 5 представлена треугольная сетка в момент времени $t = 0,94$. Распределение плотности в центральной части облака остается близким к константе для моментов времени $t = 0,94$ (Рис. 6) и $t = 1,18$ (Рис. 8, 9).

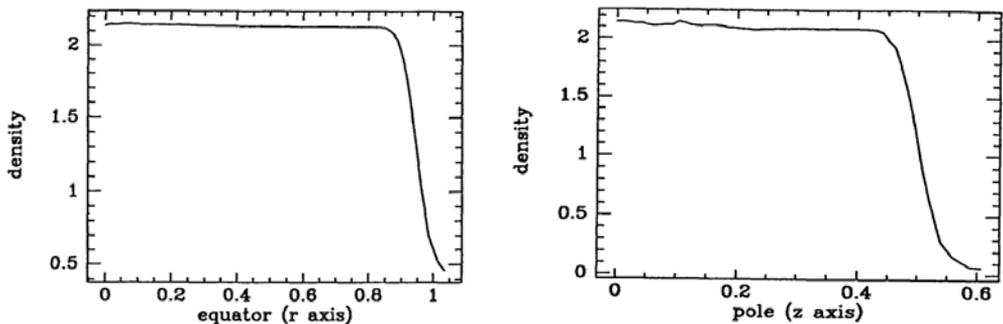


Рис. 6. Распределение плотности вдоль оси r (левый график) и вдоль оси z (правый график) в момент времени $t = 0,94$

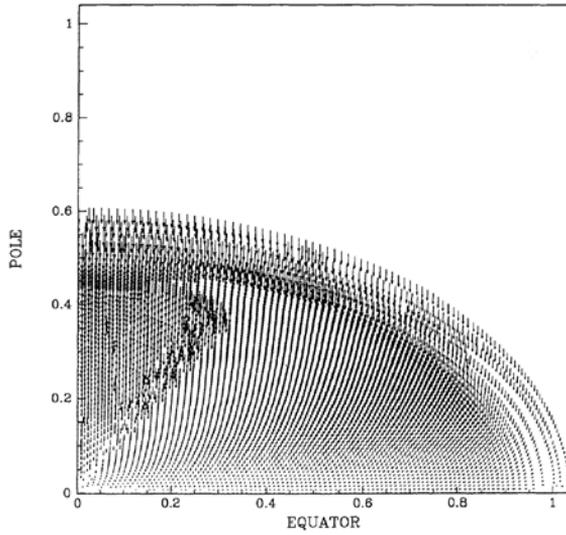


Рис. 7. Поле скоростей в момент времени $t = 0,94$

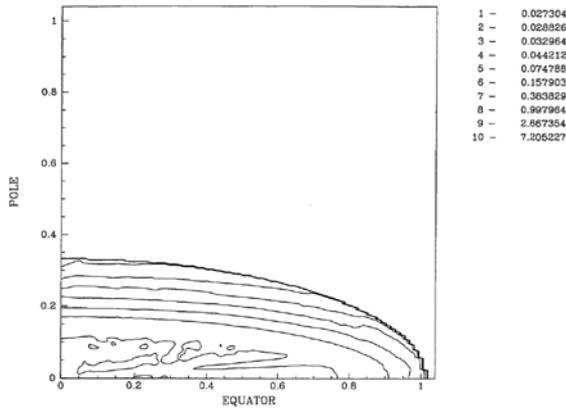


Рис. 8. Уровни плотности в момент времени $t = 1,18$

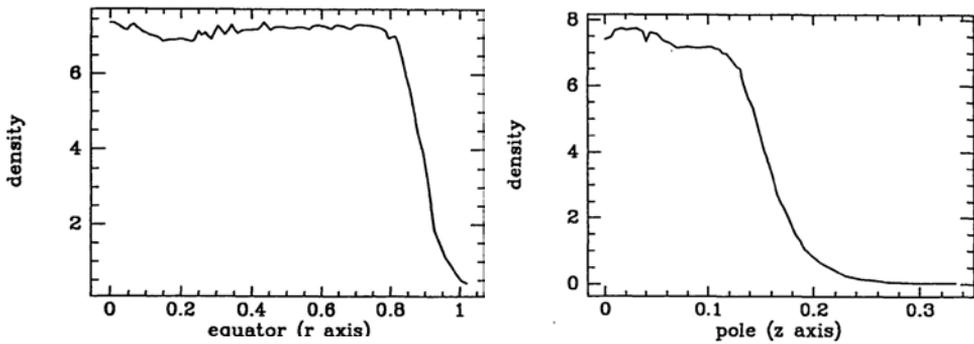


Рис. 9. Распределение плотности вдоль оси r (левый график) и вдоль оси z (правый график) в момент времени $t = 1,18$

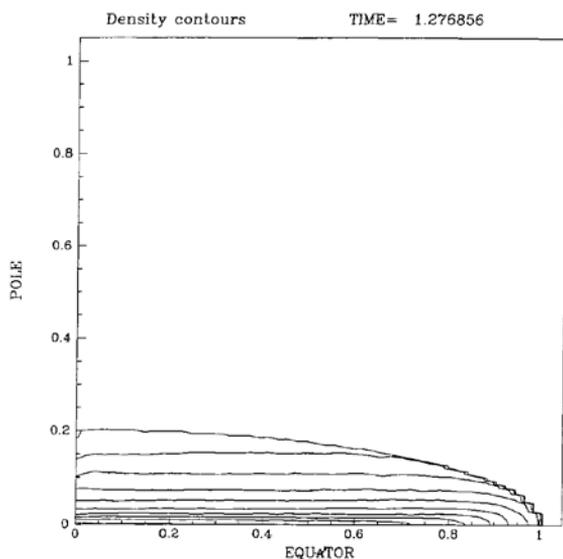


Рис. 10. Уровни плотности в момент времени $t = 1,27$

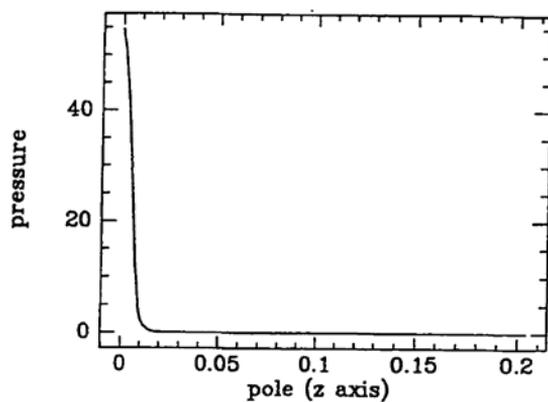


Рис. 11. Распределение давления вдоль оси z в момент времени $t = 1,27$

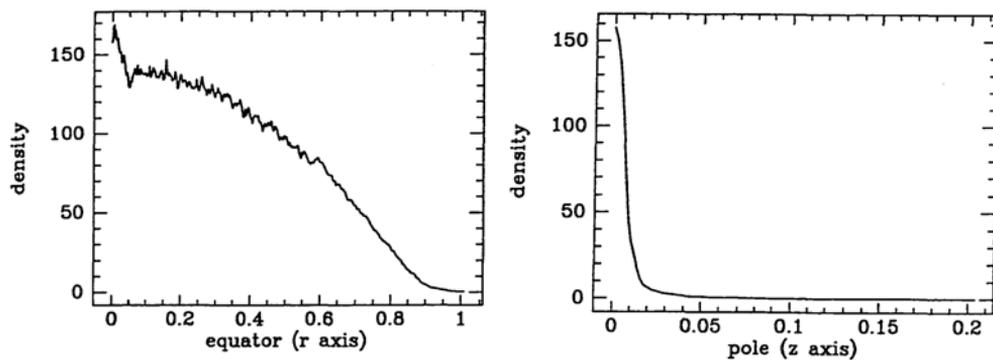


Рис. 12. Распределение плотности вдоль оси r (левый график) и вдоль оси z (правый график) в момент времени $t = 1,27$

«Численный шум», который виден на Рис. 9, 12, связан с использованием процедуры перестройки сетки. Возникновение этого «шума», известный эффект при использовании свободно-лагранжевых методов [Trease et al., 1991]. При отсутствии процедуры перестройки сетки «численный шум» сглаживается схемной и искусственной вязкостью. При использовании перестройки сетки небольшие численные осцилляции в решении имеют место, однако они не приводят к неустойчивости численного расчета.

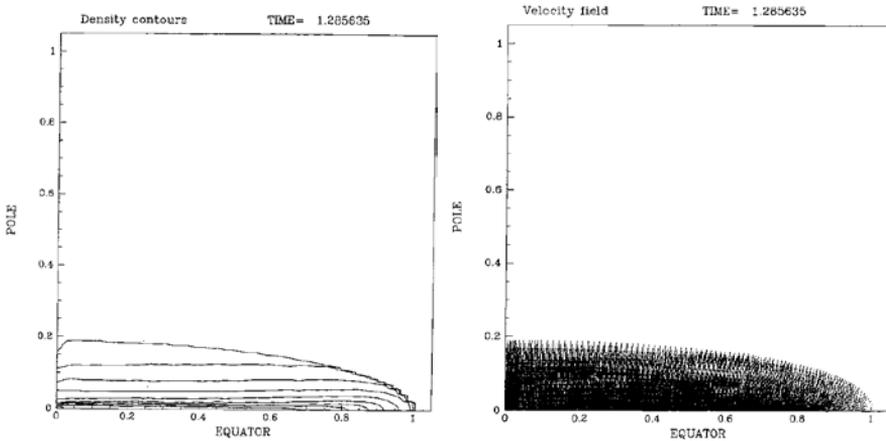


Рис. 13. Уровни плотности (левый график) и поле скоростей (правый график) в момент времени $t = 1,285$

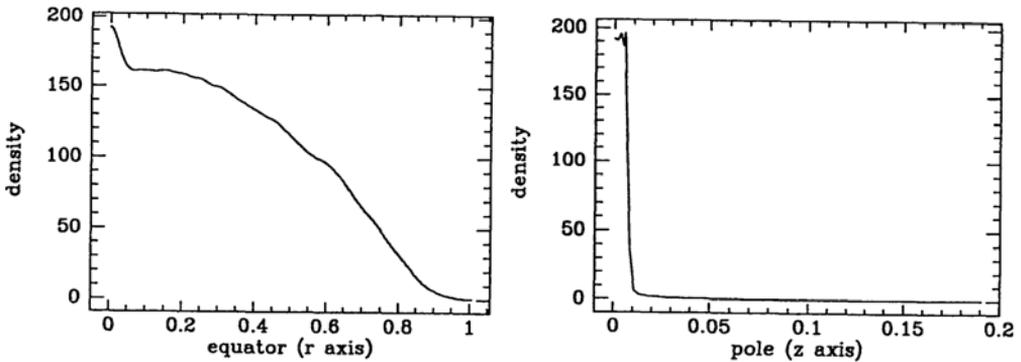


Рис. 14. Распределение плотности вдоль оси r (левый график) и вдоль оси z (правый график) в момент времени $t = 1,285$

Поле скоростей (Рис. 7) показывает, что практически во всей расчетной области вещество движется к экваториальной плоскости. Газ движется медленно вдоль экваториальной плоскости лишь вблизи периферийной части экватора. Облако продолжает сжиматься, уплотняться, и к моменту времени $t = 1,276856$ значение плотности в центре $\rho \approx 150\rho_0$ (Рис. 12), давление в центре $p \approx 6 \cdot 10^4 p_0$ (Рис. 11). Следует отметить, что, несмотря на то, что для численного моделирования использовалась неявная схема, решение исходной системы уравнений (3)

на данной стадии течения менялось со временем настолько быстро, что для того, чтобы избежать потери точности в численном решении, величина используемого шага по времени соответствовала числу Куранта [Рихтмайер, Мортон, 1972], меньшему единицы. Изменение формы протозвездного облака со временем представлено на Рис. 37. Темная вертикальная полоса, заметная на этих рисунках (см. например правый верхний график на Рис. 37, $t = 1,356753$), связана со сгущением сетки вблизи оси z . Это сгущение было сделано «вручную» для повышения точности расчетов вблизи оси z . Во внутренней части облака плотность растет монотонно до момента времени $t = 1,285635$ (момент максимального сжатия) и равна $\rho \approx 195\rho_0$, где ρ_0 – значение плотности в начальный момент времени (Рис. 13, 14).

Возможность перестраивания сетки особенно важна на данной стадии расчетов, поскольку сильное и существенно неоднородное сжатие приводит к искажению треугольной лагранжевой сетки и последующему перехлесту ячеек сетки.

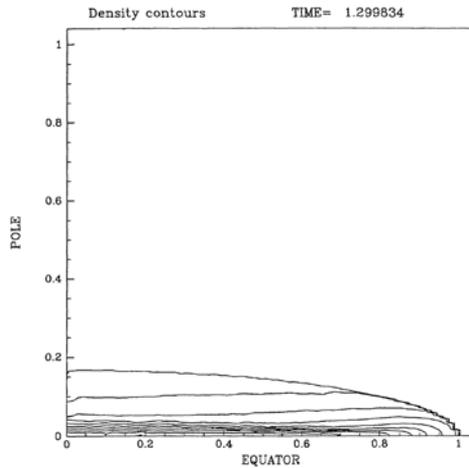


Рис. 15. Уровни плотности в момент времени $t \approx 1,3$

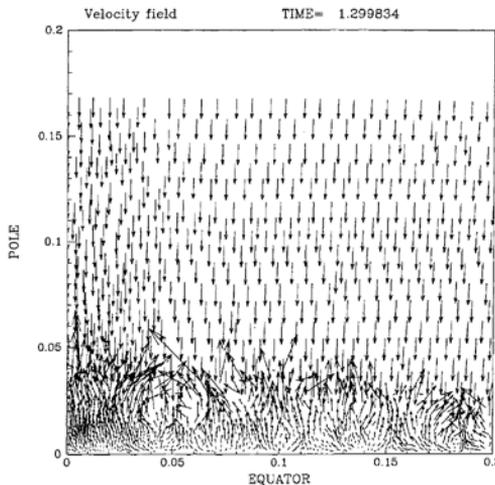


Рис. 16. Поле скоростей вблизи оси вращения в момент времени $t \approx 1,3$

На конечной стадии первого сжатия $t = 1,285635$ облако принимает форму эллипсоида с соотношением полуосей $\sim 1:5$. Однако внутренняя часть облака, составляющая до 90% массы облака, содержится в довольно уплощенном эллипсоиде с соотношением полуосей $\sim 1:100$ (левый график на Рис. 13). В произвольном сечении расчетной области (четверти эллипса), параллельном оси z , число узлов сетки на данной стадии коллапса было не менее 15. В момент максимального сжатия $t = 1,285635$ вблизи центра облака возникает ударная волна. Графики для момента времени $t = 1,299834$ (Рис. 15, 16, 17) описывают стадию коллапса, когда ударная волна уже сформировалась и начинается ее отражение от экваториальной плоскости. Фронт ударной волны ясно виден на Рис. 16, где приведена центральная часть поля скоростей.

Как видно из Рис. 16, поле скоростей за ударной волной имеет квазипериодическую вихреобразную структуру; кроме того, вблизи экваториальной плоскости плотность имеет волнообразное распределение. Размер этого квазипериода составляет примерно десять узлов в r направлении, из чего мы заключаем, что это физический, а не численный эффект, связанный с неустойчивостью, возникающей за ударной волной, движущейся в гравитационном поле.

Аналогичная волновая структура за ударной волной при наличии гравитационного поля была исследована аналитически в одномерном акустическом приближении в работе [Lamb, 1909] и для случая одномерной гравитационной газовой динамики в работе [Косовичев, Попов, 1979]. Основной причиной возникновения таких колебаний является изменение дисперсионных свойств сплошной среды при наличии гравитационного поля. Двумерный эффект, полученный в наших расчетах, возникает, по-видимому, по тем же самым причинам. Кроме того, возникновение такой вихреобразной структуры может быть связано с движением наклонной ударной волны по неоднородному фону. Квазипериодическая вихреобразная структура поля скоростей, аналогичная нашей, была получена при расчетах начальной стадии коллапса газового облака при аналогичных начальных данных при помощи РРМ-метода (кусочно-параболический метод) на эйлеровой сетке, проведенных в институте им. Макса Планка по астрофизике (Гархинг, ФРГ).

Образование ударной волны и начало ее отражения завершает стадию первого сжатия. Образовавшаяся ударная волна существенно сильнее вблизи оси z , чем на периферии и движется быстрее вблизи оси вращения. За фронтом ударной волны плотность быстро спадает. Фронт ударной волны достигает внешней границы облака вблизи оси z в момент времени $t = 1,314921$, плотность в центральных ячейках в этот момент времени равна $\rho \sim 93\rho_0$ (Рис. 18). Распределение плотности вдоль оси r – экватора $\rho(r)$ (Рис. 20) представляет собой кольцо, т.е. максимум в распределении плотности вдоль экватора расположен на периферии экваториальной плоскости, а не в ее центре. На этой промежуточной стадии распределение плотности вблизи экватора напоминает распределение плотности, полученное на конечной стадии расчетов, выполненных на основе эйлеровых разностных схем при аналогичных начальных условиях [Boss, 1980]. На этой стадии коллапса лагранжевы и эйлеровы подходы дают качественно аналогичные результаты (т.е. кольцевидное распределение плотности). После образования ударной волны скорость газа вблизи экваториальной плоскости меняет свой знак из-за отражения газа от экваториальной плоскости. Несмотря на расширение ве-

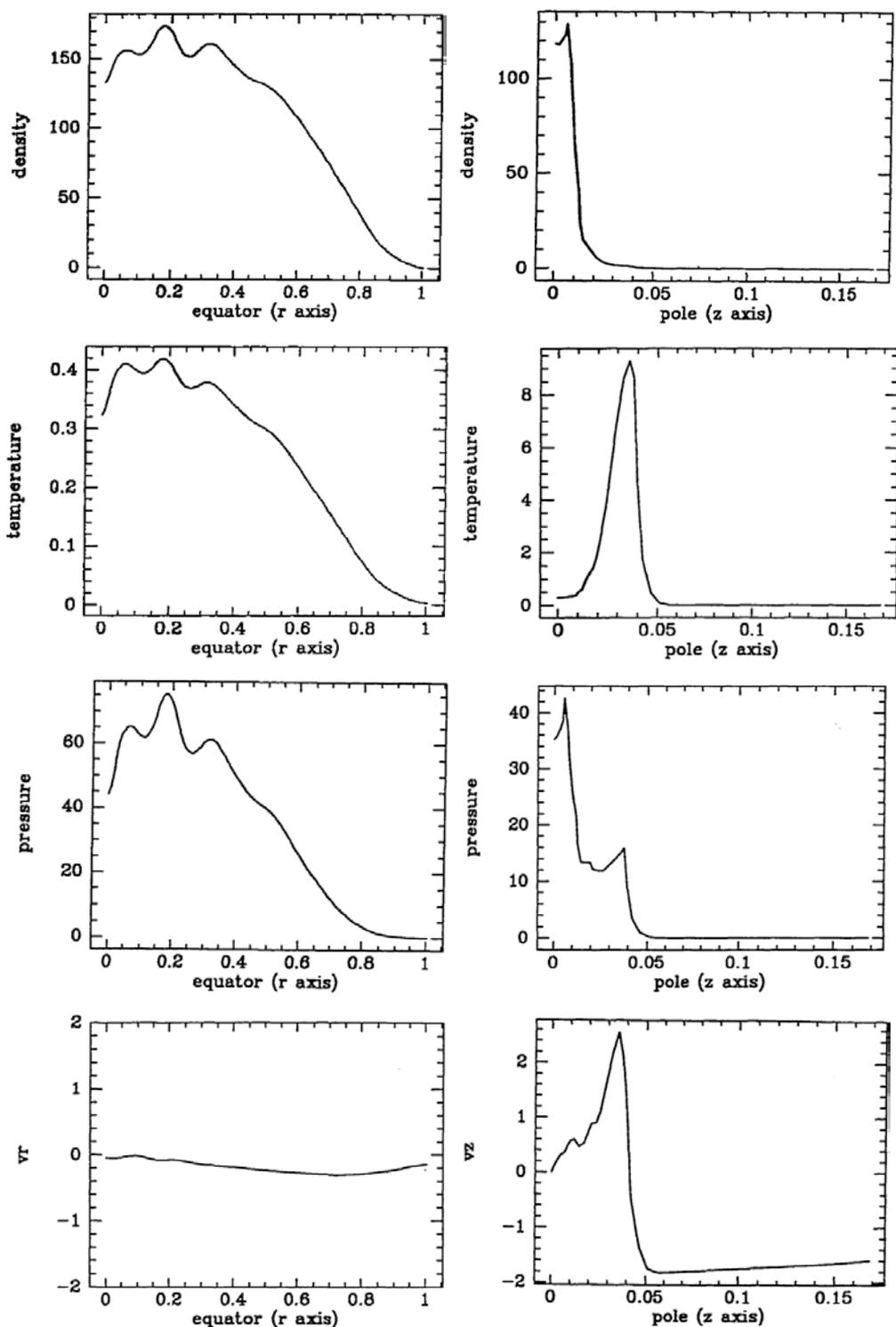


Рис. 17. Распределение параметров вдоль осей r и z в момент времени $t \approx 1,3$

щества во внутренних частях облака за фронтом ударной волны, оболочка облака продолжает сжиматься до тех пор, пока ударная волна не достигнет свободной поверхности облака в момент времени $t = 1,314921$ (Рис. 19). Число Маха, характеризующее интенсивность ударной волны, равно $M \approx 30$ в центральной части облака, вскоре после формирования ударной волны и начала отражения газа от экваториальной плоскости. Число Маха для фронта ударной волны возрастает, так как ударная волна движется во внешние части облака по быстро ниспадающему профилю плотности. В точке соприкосновения ударной волны и свободной поверхности возникает неустойчивость, которая может привести к опрокидыванию части оболочки облака вблизи оси z . Хотя этот эффект не оказывает существенного влияния на процессы, происходящие внутри ядра облака, он не позволяет продолжать расчет из-за нелокального перехлеста сетки. Для преодоления указанной трудности и для уменьшения кривизны свободной границы скорости граничных узлов были изменены путем введения в операторно-разностную схему оператора, действующего только в граничных узлах [Ardeljan, Kosmachevskii, 1993]. Максимум плотности на данной стадии коллапса все еще находится на периферии экваториальной плоскости, и распределение плотности является кольцевидным. Начиная с $t = 1,314921$, облако расширяется, и к моменту времени $t = 1,356753$ его форма напоминает начальную (шар единичного радиуса) (Рис. 37, верхний график в правой колонке, Рис. 22, 23), но с существенно неоднородным распределением плотности. С этого момента времени облако начинает расширяться также и в радиальном направлении. В момент времени, когда точка соприкосновения свободной границы облака и фронта ударной волны достигает экваториальной плоскости, ударная волна исчезает, но облако продолжает расширяться и при $t = 1,404643$ имеет форму эллипсоида, вытянутого вдоль оси z (Рис. 25, 26). Результаты расчетов на этой стадии показывают, что применение на внешней границе облака условия постоянного объема, как это делается в работах [Аксенов, Имшенник, 1994; Boss, 1980], использующих эйлеров подход, привело бы к существенному искажению картины течения.

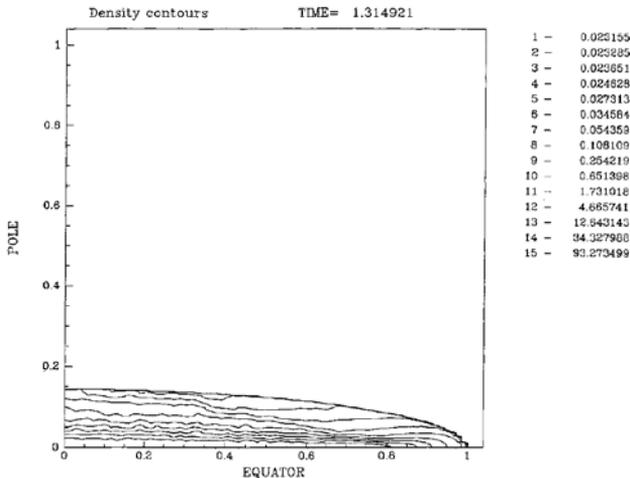


Рис. 18. Уровни плотности в момент времени $t \approx 1,315$

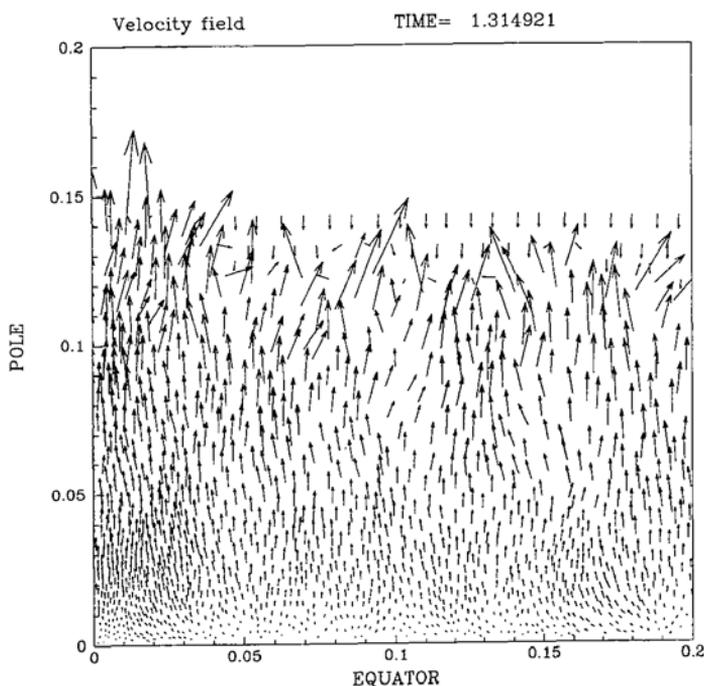


Рис. 19. Поле скоростей вблизи оси вращения в момент времени $t \approx 1,315$

На протяжении стадии первого сжатия и последующего разлета вращение облака является практически однородным, лишь часть оболочки облака возле оси z вращается с меньшей угловой скоростью. Картина линий уровня плотности при $t = 1,499228$ (Рис. 28, 29, 30) показывает, что облако состоит из небольшого плотного ядра и оболочки, существенно вытянутой в $-z$ направлении. Вращение облака становится дифференциальным: центральная часть облака вращается быстрее, чем его периферия. Максимум плотности продолжает оставаться на периферии экваториальной плоскости. Начиная с момента времени $t = 1,499228$, и далее максимум плотности расположен в центре облака, и распределение плотности представляет собой диск. Значение плотности в центре продолжает падать и при $t = 1,499228$ достигает своего минимума $\rho \sim 10\rho_0$. С момента времени $t = 1,499228$ начинается второе сжатие, которое происходит только в центральной части облака. Значение плотности в центре начинает плавно расти. Сильно вытянутая оболочка облака, представляющая собой газ с малой плотностью, продолжает разлетаться. Часть вещества оболочки ($\sim 5\%$ массы облака) имеет кинетическую энергию большую, чем ее потенциальная энергия, и, следовательно, за бесконечное время эта часть облака улетит бесконечно далеко от облака. В момент времени $t = 1,774253$ расчет был закончен. Распределение плотности и поле скоростей при $t = 1,774253$ представлено на Рис. 31 и Рис. 32 соответственно. На Рис. 33 представлена центральная часть поля скоростей вблизи оси вращения. Видно, что в облаке происходит вторичное сжатие. Значение плотности в центре облака в этот момент времени $\rho \approx 32\rho_0$ (Рис. 34).

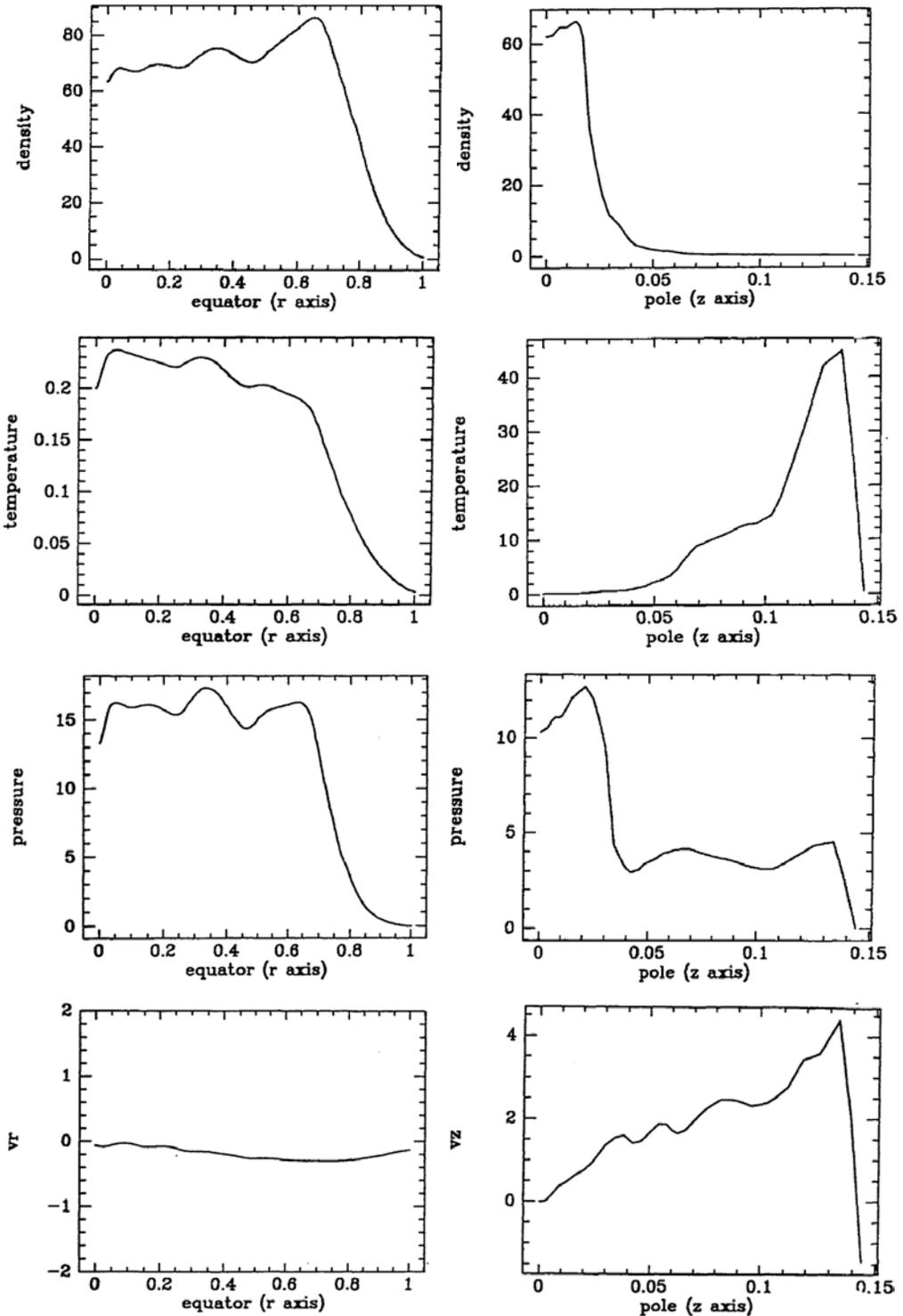


Рис. 20. Распределение параметров вдоль осей r и z в момент времени $t \approx 1,315$

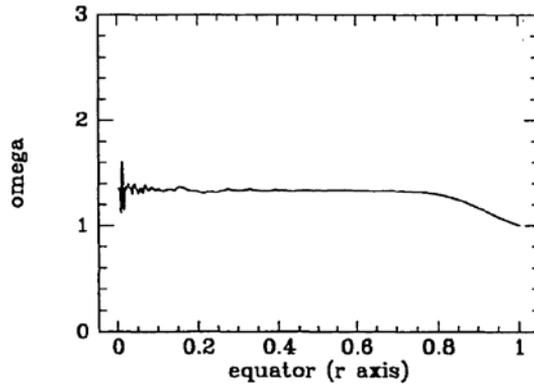


Рис. 21. Распределение угловой скорости вдоль оси r в момент времени $t \approx 1,315$

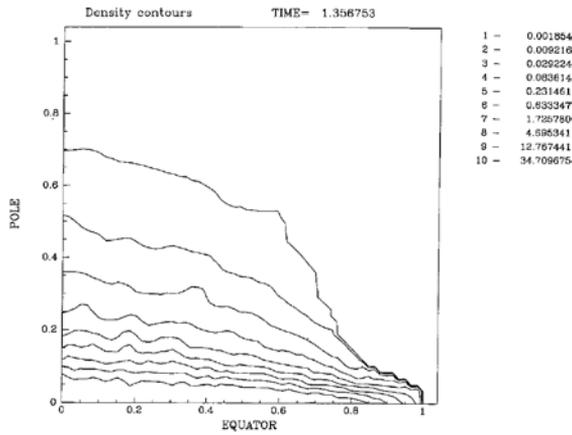


Рис. 22. Уровни плотности в момент времени $t \approx 1,356$

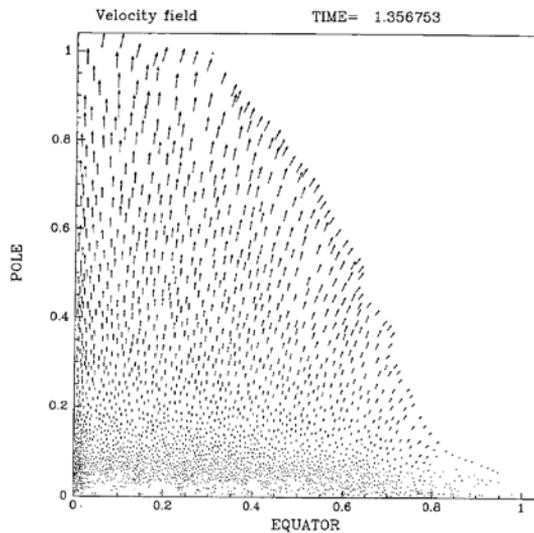


Рис. 23. Поле скоростей вблизи оси вращения в момент времени $t \approx 1,356$

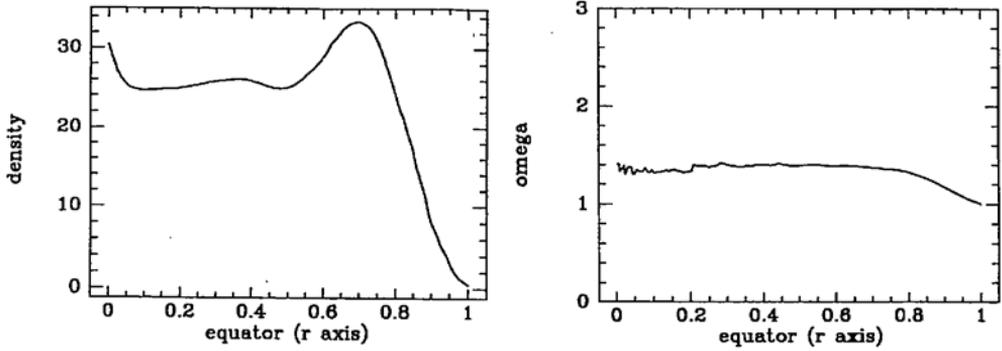


Рис. 24. Распределение плотности (левый график) и угловой скорости (правый график) вдоль оси r в момент времени $t \approx 1,356$

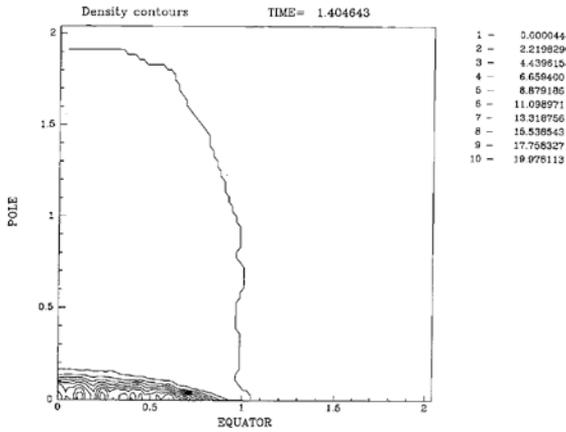


Рис. 25. Уровни плотности в момент времени $t \approx 1,405$

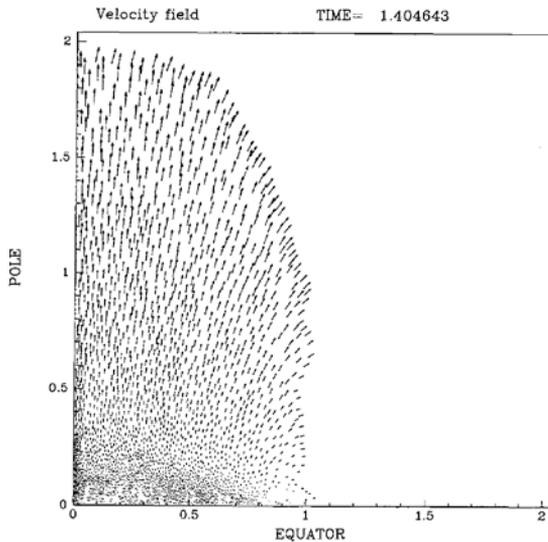


Рис. 26. Поле скоростей вблизи оси вращения в момент времени $t \approx 1,405$

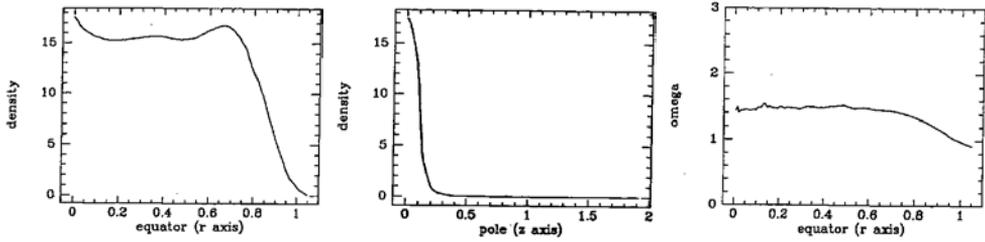


Рис. 27. Распределение плотности вдоль оси r (левый график), вдоль оси z (средний график) и угловой скорости вдоль оси r (правый график) в момент времени $t \approx 1,405$

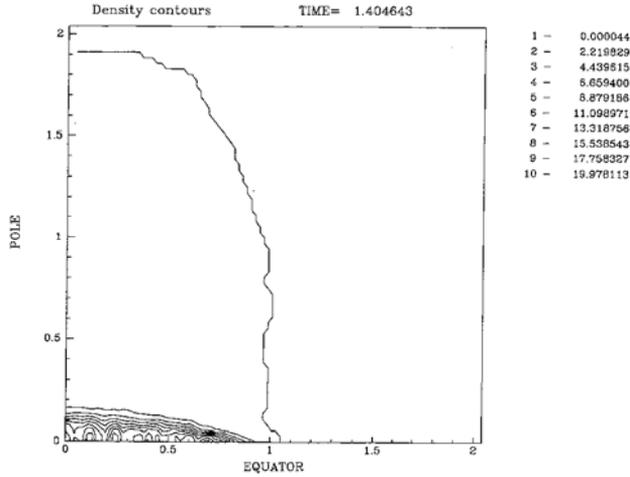


Рис. 28. Уровни плотности в момент времени $t \approx 1,499$

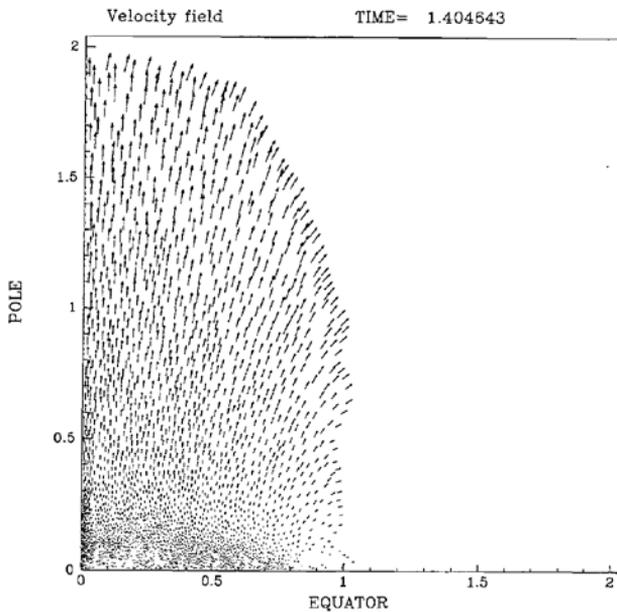


Рис. 29. Поле скоростей в момент времени $t \approx 1,499$

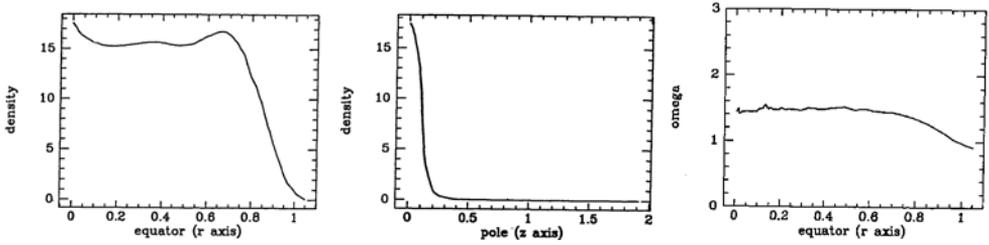


Рис. 30. Распределение плотности вдоль оси r (левый график), вдоль оси z (средний график) и угловой скорости вдоль оси r (правый график) в момент времени $t \approx 1,499$

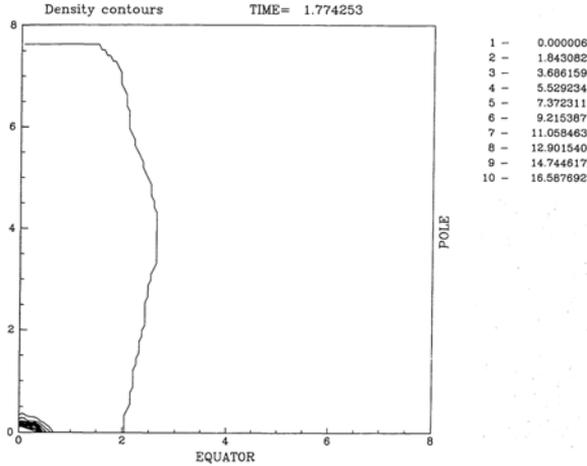


Рис. 31. Уровни плотности в момент времени $t \approx 1,774$

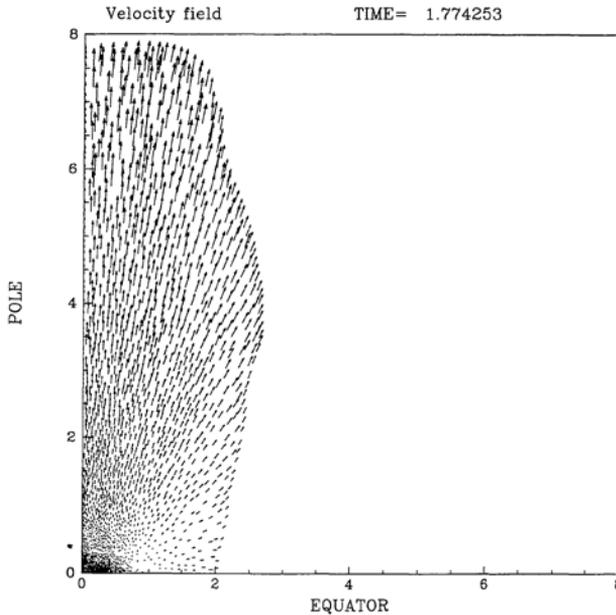


Рис. 32. Поле скоростей вблизи оси вращения в момент времени $t \approx 1,774$

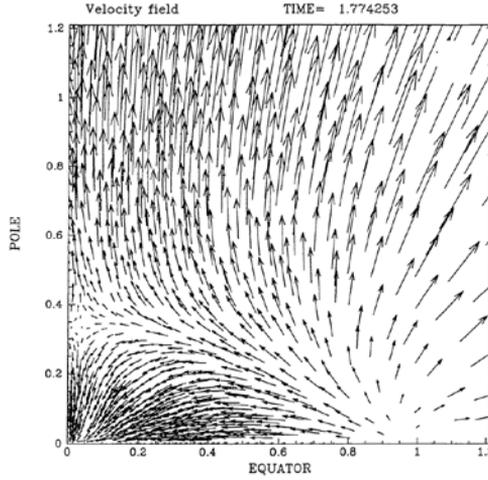


Рис. 33. Центральная часть поля скоростей вблизи оси вращения в момент времени $t \approx 1,774$

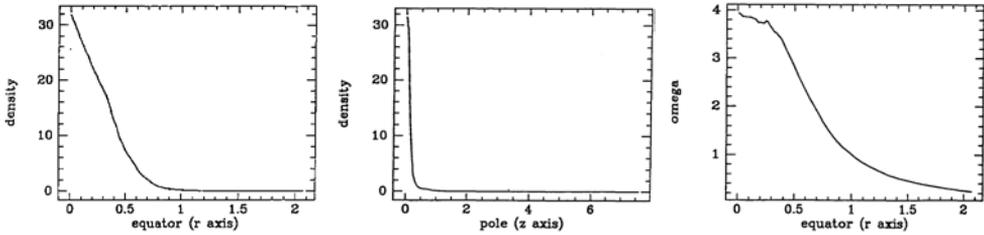
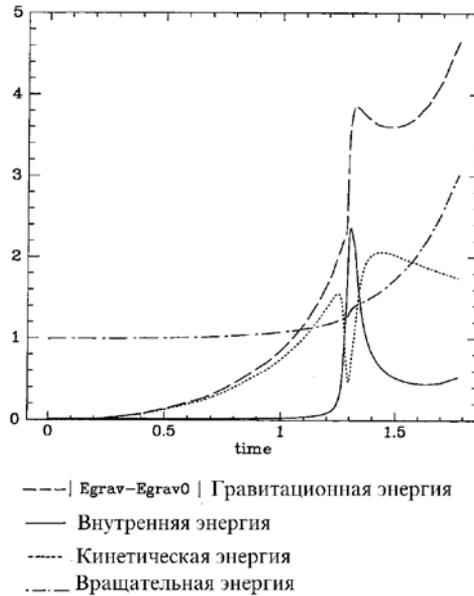


Рис.34. Распределение плотности вдоль оси r (левый график), вдоль оси z (средний график) и угловой скорости вдоль оси r (правый график) в момент времени $t \approx 1,774$



- | $E_{grav} - E_{grav0}$ | Гравитационная энергия
- | — | Внутренняя энергия
- ...| ... | Кинетическая энергия
- .-.-| -.-.- | Вращательная энергия

Рис. 35. Изменение различных видов энергии со временем

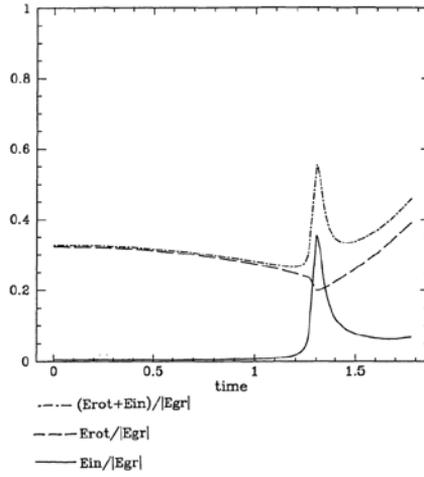


Рис. 36. Изменение коэффициентов $\alpha = E_{in}/|E_{gr}|$, $\beta = E_{rot}/|E_{gr}|$ и $\alpha+\beta$ со временем

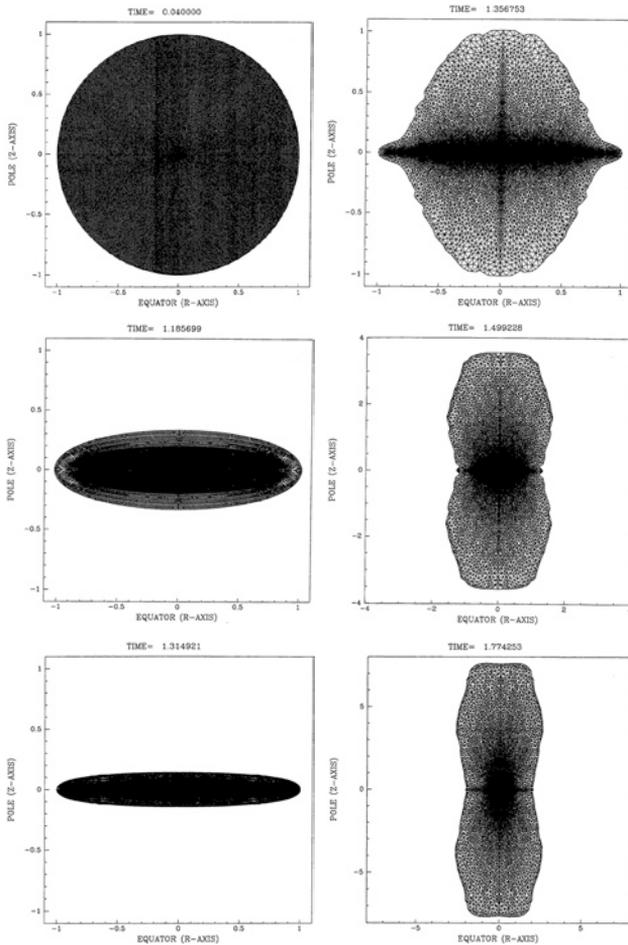


Рис. 37. Изменение формы протозвездного облака со временем в процессе коллапс

Максимальная относительная ошибка в величине полной энергии в конце расчета составила 4,9%. На Рис. 35 приведены графики изменения во времени различных видов энергии. На Рис. 36 представлено изменение во времени следующих параметров: $\alpha = E_{in}/|E_{gr}|$, $\beta = E_{rot}/|E_{gr}|$ и $\alpha+\beta$.

Второе сжатие, начинается в момент времени $t = 1,499228$ при следующих параметрах

$$\alpha \approx 0,08, \beta \approx 0,34 \quad (11)$$

Оно происходит существенно более плавно. Следует отметить, что даже в работе [Boss, 1980] при начальных значениях параметров, близких к (11) в равновесном состоянии, был получен диск, и мы можем заключить, что второе и последующие сжатия облака приведут в итоге к формированию дискообразного тела.

ЛИТЕРАТУРА

- [Аксенов, Имшенник, 1994] *Аксенов А.Г., Имшенник В.С.* Численное исследование устойчивости быстровращающейся нейтронной звезды (аксиально-симметричная модель) // Письма в АЖ, 1994, 20, с. 32-49.
- [Арделян, 1979] *Арделян Н.В.* Об устойчивости разностных схем для многомерных уравнений акустики. // Вестн. Моск. ун-та. Сер 15, Вычисл. матем. и кибернетика, 1979, N.2, с.65-69.
- [Арделян, 1981] *Арделян Н.В.* О сеточных аналогах основных дифференциальных операторов на нерегулярной треугольной сетке. // Разностные методы математической физики. - М.: Изд-во МГУ, 1981. с.49-58.
- [Арделян, 1983а] *Арделян Н.В.* Сходимость разностных схем двумерных уравнений акустики и Максвелла // ЖВМиМФ, 1983, т.23, N.6, с.1168-1176.
- [Арделян, 1983б] *Арделян Н.В.* Об использовании итерационных методов при реализации неявных разностных схем двумерной магнитной гидродинамики. // ЖВМиМФ, 1983, т.23, N.6, с.1417-1426.
- [Арделян, Космачевский, 1993] *Арделян Н.В., Космачевский К.В.* Неявный свободно-лагранжевый метод для расчета двумерных магнитогазодинамических течений. // В кн. «Математическое моделирование». М.: МГУ, 1993, с. 25-44.
- [Арделян и др., 1987а] *Арделян Н.В., Космачевский К.В., Черниговский С.В.* Вопросы построения и исследования полностью консервативных разностных схем магнитной газовой динамики. // М.: МГУ, 1987. 111 с.
- [Арделян и др., 1987б] *Арделян Н.В., Бисноватый-Коган Г.С., Попов Ю.П., Черниговский С.В.* Расчет коллапса вращающегося газового облака на лагранжевой сетке.// АЖ, 1987, 64, с. 495-508.
- [Косовичев, Попов, 1979] *Косовичев А.Г., Попов Ю.П.* К расчету одномерных нестационарных задач гравитационной газовой динамики. // Журн. вычислит. матем. и мат. физики. 1979, 19, с. 1253-1261.
- [Ландау, Лифшиц, 1988] *Ландау Л.Д., Лифшиц Е.М.* Гидродинамика // М.: Наука, 1988. - 733 с.
- [Рихтмайер, Мортон, 1972] *Рихтмайер Р., Мортон К.* Разностные методы решения краевых задач //М.: Мир, 1972. 418 с.

- [Тихонов, Самарский, 1972] *Тихонов А.Н., Самарский А.А.* Уравнения математической физики // М.: Наука, 1972. 736 с.
- [Ardeljan, Kosmachevskii, 1993] *Ardeljan N.V., Kosmachevskii K.V.* A Conservative Stable Free Lagrange Method for the Solution of 2D Magnetohydrodynamic Flows. // Preprint of Max Plank Institut of Astrophysics (Garching), No.716, 1993. Pp. 1-28.
- [Ardeljan et al., 1996] *Ardeljan N.V., Bisnovaty-Kogan G.S., Kosmachevskii K.V., Moiseenko S.G.* An implicit Lagrangian code for the treatment of nonstationary problems in rotating astrophysical bodies. // *Astron. & Astrophys. Suppl. Ser.*, 1996a, 115, pp. 573-594.
- [Boss, 1980] *Boss A.P.* Collapse and equilibrium of Rotating, Adiabatic Clouds. // *ApJ*, 1980, 242, pp. 699-709.
- [Lamb, 1909] *Lamb H.* Proc. London Math. Soc., 1909, 7, 122.
- [Trease et al., 1991] *Trease H.E., Fritts M.J., Crowley W.P. (Eds.)* Advances in the Free-Lagrange Method., Berlin, 1991. 326 p.

SIMULATION OF THE COLLAPSE OF COLD RAPIDLY ROTATING PROTOSTELLAR CLOUD

S.G. Moiseenko¹, G.S. Bisnovaty-Kogan^{1,2}, N.V. Ardeljan³, K.V. Kosmachevskii³

¹ Space Research institute of Russian Academy of Sciences (IKI RAS)

¹ National Research Nuclear University MEPhI (Moscow Engineering Physics Institute) (NRNU MEPhI)

¹ Faculty of Computational Mathematics and Cybernetics of Lomonosov Moscow State University (CMC MSU)

We consider the problem of collapse of cold rapidly rotating protostellar cloud. Numerical 2D simulation was done using completely conservative implicit operator-difference scheme on triangular grid of variable structure. It was shown that collapse of the cloud lead to the density distribution of disk shape.

Keywords: protostellar clouds, numerical simulations.

Moiseenko Sergey Grigorievich – head of department, doctor of physical and mathematical sciences, IKI RAS, moiseenko@iki.rssi.ru

Bisnovaty-Kogan Gennadiy Semenovich – chief researcher, doctor of physical and mathematical sciences, professor, IKI RAS, NRNU MEPhI, gkogan@iki.rssi.ru

Ardeljan Nikolai Vasilievich – leading researcher, doctor of physical and mathematical sciences, professor, CMC MSU

Kosmachevskii Konstantin Viktorovich – senior researcher, candidate of physical and mathematical sciences, CMC MSU

МОДЕЛИРОВАНИЕ 3D ЗАДАЧ НЕСТАЦИОНАРНОЙ ГАЗОДИНАМИКИ НА МНОГОПРОЦЕССОРНЫХ ЭВМ И ГРАФИЧЕСКИХ УСКОРИТЕЛЯХ

Б.П. Рыбакин

Федеральное государственное учреждение Федеральный научный центр
Научно-исследовательский институт системных исследований Российской академии наук
(ФГУ ФНЦ НИИСИ РАН),
Московский государственный университет имени М.В. Ломоносова (МГУ им. Ломоносова)

В работе приведены результаты математического моделирования трехмерных газодинамических течений, индуцированных сильной ударной волной. Исследованы процессы взаимодействия ударной волны, распространяющейся в межзвездной среде, с облаком молекулярного газа. Построен параллельный алгоритм и программа для проведения вычислений на гибридных вычислительных системах с помощью технологий OpenMP и CUDA.

Данная работа была выполнена при поддержке гранта Российского Фонда поддержки Фундаментальных Исследований № 14-29-06055.

Ключевые слова: Сверхзвуковые течения, ударные волны, параллельные вычисления, CUDA.

ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Математическое моделирование нестационарных многомерных задач газовой динамики позволяет проводить детальные исследования в различных областях знаний. Зачастую проведение натуральных экспериментов затруднено, дорогостояще или попросту невозможно, поэтому только численный эксперимент дает возможность получить необходимую информацию. Проведение численных экспериментов в трехмерной постановке на достаточно подробных сетках требует больших вычислительных ресурсов. Для проведения расчетов был разработан параллельный алгоритм решения задач нестационарной газодинамики [Rybakin, 2013]. Для ускорения расчетов были использованы технологии OpenMP для CPU и CUDA для графических ускорителей. Визуализация возникающих многомерных течений осуществлялась с помощью графического комплекса HDVIS [Rybakin, Goryachev, 2014].

Рассмотрим процесс взаимодействия ударной волны большой интенсивности с молекулярным облаком. Взрывы сверхновых звезд и дальнейшее распространение остатков этих взрывов сопровождается образованием ударных волн большой интенсивности. Такие ударные волны оказывают большое влияние на межзвездную среду (МЗС) и на облака молекулярного газа. Взаимодействие ударных волн от сверхновых (УВС) с молекулярными облаками (МО) вносят заметный вклад в их фрагментацию и приводит к турбулизации среды. Механизм сжатия, фрагментации и образование турбулентных течений является объектом исследований в данной работе.

Рыбакин Борис Петрович – ведущий научный сотрудник, д.ф.-м.н., Федеральное Государственное Учреждение Федеральный Научный Центр Научно Исследовательский Институт Системных Исследований РАН, МГУ им. Ломоносова, rybakin@vip.niisi.ru

Ударная волна взаимодействует с молекулярным облаком, которое находится в межзвездной среде. В работе используется однофазная модель газа, которая обеспечивает достаточно детальную физическую картину. Невозмущенная МЗС состоит из сравнительно теплого вещества (~ 10000 К) и небольших неоднородно распределенных холодных облаков (~ 100 К), которые обладают большой плотностью [McKee, 1988; Klein et al., 1994; Melioli et al., 2005]. Изначально облака находятся в динамическом равновесии с фоновым газом [Poludnenko et al., 2003]. Теплопроводностью и радиационными потерями при взаимодействии будем пренебрегать. В работе используется уравнение состояния идеального газа с $\gamma = 5/3$.

Плотность межзвездной среды принимается равной $\rho = 2,15 \cdot 10^{-25}$ г/см³, температура $T = 10^4$ К, $u = 0,0$. Плотность в облаке $\rho = 1,075 \cdot 10^{-22}$ г/см³, температура $T = 100$ К, $u = 0,0$. Параметры газа за ударной волной определяются из уравнений Ренкина – Гюгонно. В наших расчетах число Маха ударной волны равно $M = 7$, плотность $\rho = 8,6 \cdot 10^{-22}$ г/см³, температура $T = 1,5 \cdot 10^5$ К, скорость 10^4 км/с [Stone, Norman, 1992; Truelove et al., 1998]. Толщина фронта ударной волны достаточно большая $\sim 2\text{--}5$ парсеков, что намного больше радиуса облака. Радиус облака равен $r = 0,05$ парсека, время прохождения ударной волной всего облака равно 960 лет.

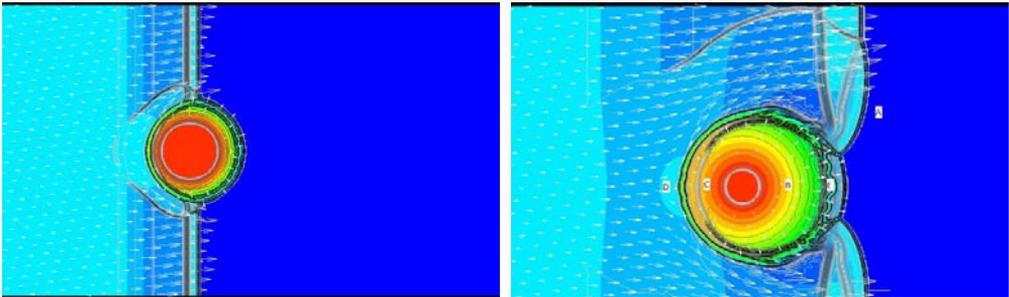


Рис. 1. Векторное поле течения в сечении $z = 1/2$ для двух моментов безразмерного времени

Для проведения расчетов было проведено обезразмеривание всех величин, все результаты приведены в безразмерном виде. Задача решается на трехмерной эйлеровой сетке. Размеры ячеек по x , y , z в приведенных расчетах одинаковы ($dx = dy = dz$). Возникающее трехмерное движение описывает система уравнений Эйлера, которая состоит из законов сохранения массы, количества движения и энергии. В расчетах используется уравнение состояния идеального газа. Более подробно математическая постановка задачи приведена в [Rybakin, 2013; Rybakin et al., 2014].

МЕТОДИКА И РЕЗУЛЬТАТЫ РАСЧЕТОВ

Для решения поставленной задачи была использована программа, которая основана на явной трехмерной схеме типа TVD второго порядка точности [Harten,

1983; Toro, 1997]. Данная схема сохраняет консервативность относительно массы, импульса и энергии газа, а также хорошо описывает ударные волны и контактные разрывы.

Для верификации программного кода было проведено большое количество различных двух и трехмерных тестов [Rybakin, 2013; Rybakin et al., 2014]. Проведенное тестирование показало хорошее совпадение результатов расчетов с аналитическими решениями и экспериментальными данными. Фронты ударных волн размазывались на величину 3–4 ячейки, контактные разрывы на 3–5 ячеек. Тестовые расчеты проводились на сетках различного размера от 128^3 до 1024^3 .

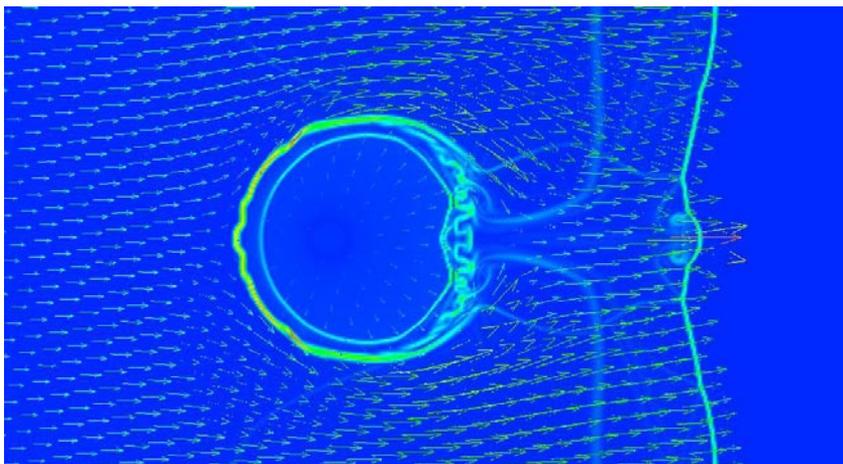


Рис. 2. Векторное поле скорости и ширен-изображение плотности

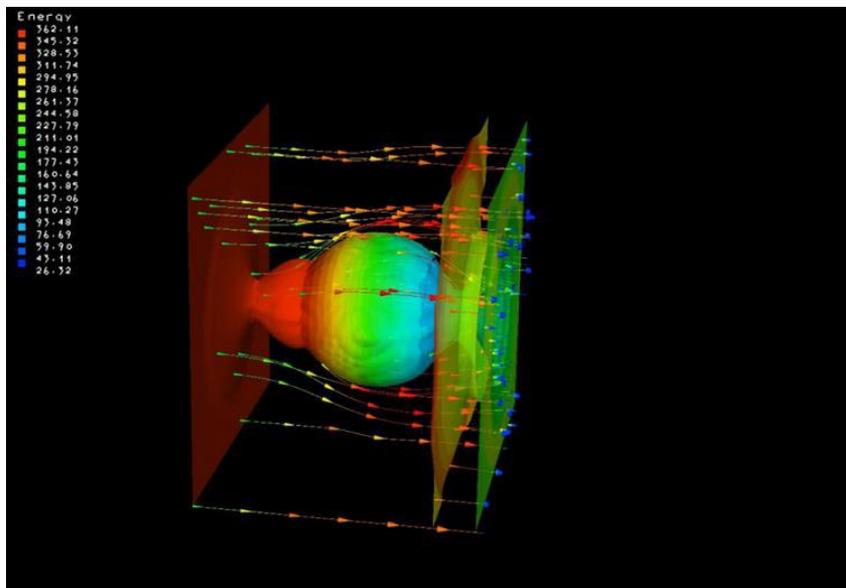


Рис. 3. Изоповерхность плотности для $\rho = 3,25$. На ней цветом отображена энергия. Стрелками показаны линии тока

Размер расчетной области выбирался из такого условия, чтобы радиус сферического облака содержал не меньше 64 ячеек. Меньшее количество ячеек отрицательно сказывается на разрешении происходящих процессов. В начальный момент времени ударная волна примыкает к левой границе облака и в момент времени $t = 0$ начинает взаимодействовать с облаком.

При взаимодействии ударной волны с облаком проходит несколько характерных стадий взаимодействия. На начальной стадии под действием падающей ударной волны облако сжимается. Разность величин в МЗС, молекулярном облаке и за фронтом ударной волны меняется на несколько порядков. Поэтому на приведенных графиках используется логарифм плотности, энергии и температуры. На рис. 1 (левый рисунок) приведены результаты расчета для логарифма плотности на момент безразмерного времени $t = 7$. На правом рисунке приведен график логарифма плотности в момент времени, когда ударная волна прошла все облако.

После прохождения ударной волной $l = 3/4$ диаметра начинается процесс фрагментирования облака и начинается перемешивание МО с ударной волной и межзвездной средой [McKee, 1988; Stone, Norman, 1992; Klein et al., 1994; Truelove et al., 1998]. На рис. 1 (правый рисунок) буквами отмечены некоторые характеристики возникающего течения.

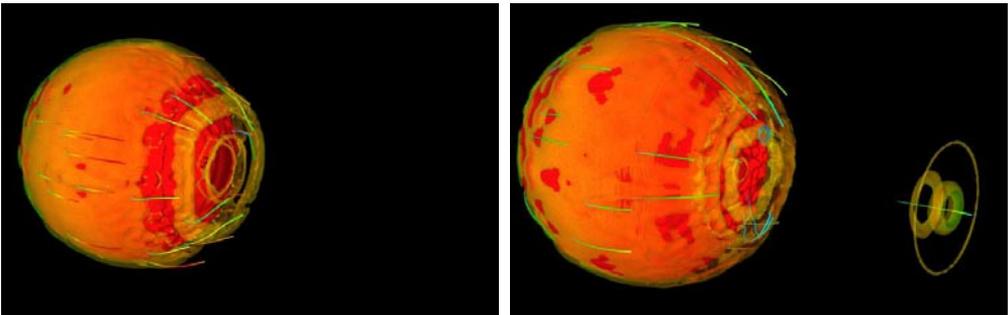


Рис. 4. Образование поверхностных возмущений. Для двух моментов времени представлены изоповерхности логарифма плотности $\log(\rho)=3$ (красным цветом) и $\log(\rho)=2$ – зеленым. Приведены линии тока

Буквой А отмечено положение ударной волны, которая огибает облако. Внутри адиабатического облака по плотному веществу распространяется прямая (В) и обратная (С) волны. Вещество внутри облака сжимается проходящей через облако волной (В). Скорость звука в облаке на порядок меньше скорости звука в межзвездной среде и значительно меньше скорости самой ударной волны. На рис. 1 можно заметить, что в начальные моменты времени, когда головная ударная волна огибает облако, на левой границе образуется волна, которая движется навстречу потоку. За счет этого на левой границе облака начинает формироваться первичное возмущение (D). На правой границе возникает неустойчивость Рихтмаейра-Мешкова (E). Следующая стадия характеризуется расширением облака (рис. 2) и образованием сложной структуры волн. За фронтом прошедшей ударной волны формируется вихрь, который развивается с течением времени (рис. 5). На поверхности облака виден процесс абляции (уноса вещества).

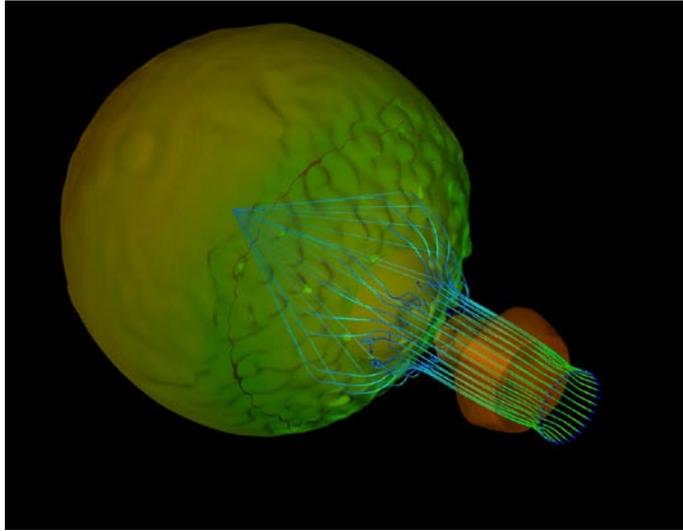


Рис. 5. График логарифма плотности $\rho = 1,9$ в $t = 27,03$, приведены линии тока

После прохождения ударной волной сферического облака профиль ударной волны (УВС) становится вогнутым, на передней кромке видны тройные ножки Маха (F) (рис. 1, правый). Падающая ударная волна распадается на проходящую и отраженную ударные волны. На поздней стадии после прохождения ударной волной всего облака на оси тыльной поверхности облака в области отраженной ударной волны, формируется область высокого давления. Эта область высокого давления образует ударную волну, которая начинает распространяться навстречу падающей волне. На рис. 3. приведены изоденсы для значения плотности $\rho = 3,25$. На этой поверхности цветом отображено значение энергии, приведены линии тока для выделенных частиц. Они изображаются линиями со стрелками. На левой стороне облака по оси X хорошо видна область высокого давления.

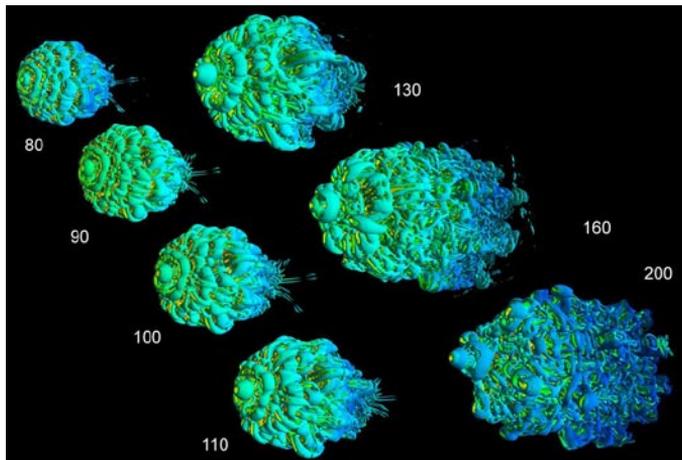


Рис. 6. Эволюция по времени Q – критерия, изоповерхность $Q = 10$

Прошедшая ударная волна «сдувает» возмущения с поверхности молекулярного облака. На рис. 4 можно отметить образование поверхностных возмущений. На этом рисунке для двух моментов времени представлены изоповерхности логарифма плотности $\log(\rho)=3$ (красным цветом) и $\log(\rho)=2$ – зеленым. Зеленым цветом показаны также линии тока для выделенных частиц.

Неустойчивость Рихтмаейра-Мешкова (Р-М) является частным случаем неустойчивости Рэлея-Гейлора и возникает из-за нарастания возмущения на границе раздела двух сред с различными плотностями, подвергнутых воздействию ударной волны. На рис. 5 видно формирование неустойчивости Р-М на правой границе. На правой половине сферы за фронтом ударной волны видно сформировавшееся вихревое кольцо, которое с течением времени развивается и образует вихревую структуру.

На рис. 5 показана поверхность $\log(\rho) = 1,9$. На эту изоповерхность отображается распределение энергии. Начиная с момента безразмерного времени $t = 27,03$ появляются линии тока, которые проходят через точки на окружности, с центром на оси $X = 6,1$ и радиусом $R = 0,25$. Треки начинают появляться, когда волна достигает этой точки. Видны линии тока, образующиеся в циркуляционной зоне у дальней (по ходу движения волны) точке.

Фрагментация облака начинается на поверхности облака (рис. 2) под действием проходящей и отраженной ударных волн. Этот процесс начинается в момент безразмерного времени $t = 20,07$ после прохождения ударной волной всего облака. Возникающие возмущения приводят к образованию завихренности над поверхностью облака и уносу (абляции) его вещества. Это хорошо видно на рис. 6. На этом рисунке представлена визуализация второго инварианта тензора скорости – Q-критерий [Hunt et al., 1988; Hejazialhosseini et al., 2012; Hejazialhosseini et al., 2013; Korneev, Levchenko, 2014]. Вихри образуют кольцевые структуры на поверхности облака. На рис. 6. приведены результаты расчетов в «поздние» моменты времени, начиная с $t = 60$ и до $t = 200$. Зарождающиеся вихри превращаются в вихревые дорожки и инициируют вихревые шпильки. Соседние вихри с противоположным направлением вращения образуют систему первичных и вторичных вихревых структур. Эта система хорошо видна на этом рисунке, начиная с момента времени $t = 130$. Со временем эти вихревые структуры растягиваются, изгибаются и меняют форму, образуя вторичные, третичные и т.д. системы.

Цветовая палитра (от бирюзового к темно-голубому) для изоповерхностей отображает значение логарифма плотности облака от 2 до 0,2. Анализ полученных графиков позволяет сделать заключение, что в процессе взаимодействия вихревые структуры фрагментируют молекулярное облако.

ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ И ВИЗУАЛИЗАЦИЯ

Математическое моделирование трехмерных задач нестационарной газовой динамики на достаточно больших сетках требуют больших вычислительных ресурсов. Для сокращения времени расчетов был построен параллельный алгоритм и программа, которая позволяет проводить вычисления на гибридной вычислительной системе.

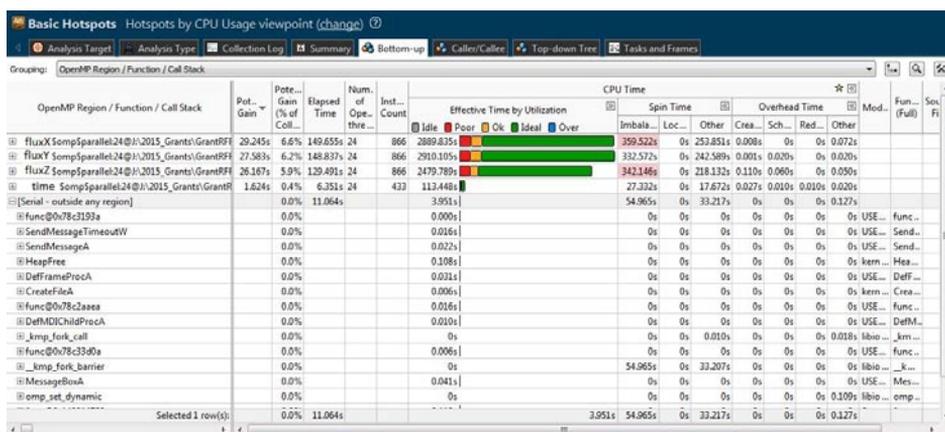


Рис. 7. Результаты распараллеливания программы на процессорах Intel Xeon E2630

Для профилирования программы был использован Intel VTune Amplifier XE. На рис. 7 приведены результаты проведенного анализа по распараллеливанию программы на CPU. На этом рисунке показано «качество» построенного параллельного алгоритма. Более 80% общего времени занимает расчет четырех подпрограмм. После оптимизации этих подпрограмм их степень распараллеливания стала достаточно высокой. На рисунке зеленым цветом показано «идеальное» распараллеливание, желтым – хорошее, красным – плохое. Плохое распараллеливание означает, что эффективность в этом случае меньше 50%. Некоторые подпрограммы рассчитывались на графическом процессоре. Более подробная информация о деталях распараллеливания приведена в [Rybakin, 2013; Rybakin et al., 2014]. Визуализация трехмерных расчетов проводилась с помощью программы HDVIS [Rybakin, Goryachev, 2014].

ВЫВОДЫ

Было проведено численное моделирование процесса взаимодействия ударной волны с молекулярным облаком. Подробно изучен процесс абляции с поверхности облака и образование вихревых структур. В начальные моменты времени на левой границе облака начинает формироваться первичное возмущение. На правой границе возникает неустойчивость Рихтмаейра-Мешкова. Следующая стадия характеризуется расширением облака и образованием сложной волновой структуры. За фронтом прошедшей ударной волны формируются вихри, которые развиваются с течением времени. На поверхности облака виден процесс абляции (уноса вещества). В моменты безразмерного времени, начиная с $t \sim 80$, образуется система вихрей. Вихри образуют кольцевые структуры на поверхности облака. С течением времени они превращаются в вихревые дорожки и иницируют вихревые шпильки. Соседние вихри с противоположным направлением вращения образуют систему первичных и вторичных вихревых структур. Со временем эти вихревые структуры растягиваются, изгибаются и меняют форму, образуя вторичные, третичные и т.д. системы.

ЛИТЕРАТУРА

- [Harten, 1983] *Harten A.* High resolution schemes for hyperbolic conservation laws. // *J. Comp. Phys.* 1983. – V49. p. 357-393.
- [Hejazialhosseini et al., 2012] *Hejazialhosseini B., Rossinelli D., Conti C., Koumoutsakos P.* (2012) High throughput software for direct numerical simulations of compressible two-phase flows // In: *Proceeding of Int. Conf. for High Performance Computing, Networking, Storage and Analysis, Salt Lake City, USA*, pp. 1-12.
- [Hejazialhosseini et al., 2013] *Hejazialhosseini B., Rossinelli D., Koumoutsakos P.* (2013) Vortex dynamics in 3D shock-bubble interaction // *Physics of Fluids*, 25, 110816 (2013), URL: <http://dx.doi.org/10.1063/1.4819345>.
- [Hunt et al., 1988] *Hunt J., Wray A., Moin P.* (1988) Eddies, streams, and convergence zones in turbulent flows // *Proc. Summer Program Center for Turbulence Research (NASA Ames/Stanford Univ.)*, 1988, pp. 193-208.
- [Klein et al., 1994] *Klein R.I., McKee C.F., Colella P.* On the hydrodynamics interaction of shock waves with interstellar clouds, 1. Nonradiative shocks in small clouds // *Astrophys. J.*, vol. 420, 1994, pp. 213-236.
- [Korneev, Levchenko, 2014] *Korneev B.A., Levchenko V.D.* Effective numerical simulation of the gas bubble-shock interaction problem using the RKDG numerical method and the Diamond Torre algorithm of the implementation flows. // *Keldysh Institute preprints*, 2014, № 097, Moscow, Russia. 2014. 12 p.
- [McKee, 1988] *McKee C.* Supernova remnant shocks in an inhomogeneous interstellar medium // In *Supernova Remnants and the Interstellar Medium*, IAU Coll. 101, eds. R.S. Roger & T.L. Landecker, Cambridge University Press, 1988. pp. 205-222.
- [Melioli et al., 2005] *Melioli C., de Gouveia Dal Pino E, Raga A.* Multidimensional hydrodynamical simulations of radiative cooling SNRs-clouds interactions: an application to starburst environments // *Astronomy & Astrophysics*, vol. 443, 2005, pp. 495-508.
- [Poludnenko et al., 2003] *Frank A., Poludnenko A., Gardiner T., Lebedev V., Drake R.* Stellar outflows with new tools: advanced simulations and laboratory experiments // *RevMexAA (Serie de Conferencias)*, 15, 2003, pp. 85-91.
- [Rybakin, 2013] *Rybakin B.* Modeling of III-D Problems of Gas Dynamics on Multiprocessing Computers and GPU. *Computers and Fluids* // vol. 80, 2013. Pp. 403–407, doi: 10.1016/j.compfluid.2012. 01.016, 31-Jan-2012.
- [Rybakin, Goryachev, 2014] *Rybakin B., Goryachev V.* The supersonic shock wave interaction with low-density gas bubble // *Acta Astronautica*, Vol. 94, Issue 2, 2014, pp. 749–753.
- [Rybakin et al., 2014] *Rybakin B.P., Stamov L.I., Egorova E.V.* Accelerated solution of problems of combustion gas dynamics on GPUs // *Computers & Fluids* 90, 2014. Pp. 164–171.
- [Stone, Norman, 1992] *Stone J.M., Norman M.L.* The three-dimensional interaction of a supernova remnant with an interstellar cloud // *APJL*, vol. 390, 1992. Pp. 17–19.
- [Toro, 1997] *Eleuterio F. Toro.* *Riemann Solver sand Numerical Methods for Fluid Dynamics* // Springer-Verlag Berlin Heidelberg, 1997, p. 727.
- [Truelove et al., 1998] *J. Kelly Truelove, Richard I. Klein, Christopher F. Mckee, John H. Holliman, Louish H. Howell, Jeffrey A. Greenough, D. Tod Woods.* Self-Gravitational Hydrodynamics With Three-Dimensional Adaptive Mesh Refinement: Methodology And Applications To Molecular Cloud Collapse And Fragmentation// *The Astrophysical Journal*, 495(2), pp. 821-852, 1998.

MODELING 3D PROBLEMS OF NON-STATIONARY GAS DYNAMICS ON MASSIVELY PARALLEL COMPUTERS AND GRAPHICS ACCELERATORS

B.P. Rybakin

Scientific Research Institute of System Analysis, Russian Academy of Science (SRISA RAS),
Lomonosov Moscow State University (Lomonosov MSU)

The paper presents the results of mathematical modeling of three-dimensional gas-dynamic flows induced by a strong shock wave. The processes of interaction of a shock wave propagating in the interstellar medium with a molecular gas cloud. Parallel algorithm and program for carrying out calculations on hybrid computing systems using OpenMP and CUDA technologies was built.

Keywords: Supersonic Flows, Shock Waves, Parallel Calculations, CUDA

Rybakin Boris Petrovich—Doctor of physics and mathematics, leading researcher, rybakin@vip.niisi.ru

**Computer Technologies in Sciences.
Methods of Simulations on Supercomputers.
Part 2
Proceedings**

Russia, Tarusa, April 21–23, 2015. Eds. R.R. Nazirov, L.N. Shchur

This issue is the collection of the papers based on the talks presented at the extended workshop “Methods of simulations on supercomputers”, which was held at 21–23 April 2014, in Tarusa hotel Interkosmos of the RAS Space Research Institute. It is the workshop in the series of extended workshops, devoted to the computer technologies in natural sciences. The first workshop was devoted to computer simulations of the actual problems in physics and mechanics and papers was published in the issue 1 (Proceedings of the seminar on Computing Technology in Science. Iss. 1. Computational Physics / Ed. R.R. Nazirov. M.: KDU, 2009. 288 p.), the second workshop was devoted to the discussion of the physical and medical aspects of the research of the complex nonlinear processes in the human body (Methods of Non-linear Analysis in Cardiology and Oncology: Physicist’s Approaches and Clinical Practice Iss. 2. / Ed. R.R. Nazirov. M.: KDU, 2010. 206 p.), the third workshop was devoted to the computer simulations in biology and chemistry (Materials of the seminar on Computing Technology in Science. Iss. 3. Computational Chemistry and Biology / Eds. R.R. Nazirov, L.N. Shchur. M.: KDU, 2010. 118 p.), the fourth was devoted to the discussion of the computer vision in control systems of the mobile objects (Computer vision in control systems of the mobile objects: Proceedings of the conference-workshop-2010. Iss. 4 / Ed. R.R. Nazirov. M.: KDU, 2011. 328 p.). The fifth workshop was devoted to the simulation of events of global scale (Proceedings of the seminar on Computing Technology in Science. Global scale systems / Eds. R.R. Nazirov, L.N. Shchur. M.: IKI, 2011, 148 p.). Fifth workshop was devoted to the investigations of promising computing systems, with emphasis on the conceptual approaches and on the development of new methods and tools (Advanced computing: systems, methods and concepts. Proceedings of the seminar on Computer Technology in Science / Eds. R.R. Nazirov, L.N. Shchur. M.: IKI, 2011, 192 p.). Algorithms and methods in computational physics was discussed at the sixth workshop with the special accents on the physical results. Materials may be found at the seminar web-pages

<http://www.iki.rssi.ru/seminar/> . Seventh workshop was transformed into the conference «Infocommunication technologies in scientific research» with the papers published in the double volume issue of journal Information Society». Eighth workshop, as a current one, named “Methods of simulations on supercomputers”. The name is connected with the fact that workshops are supported within the research program the grant by Russian Scientific Foundation 14-21-00158 «*Algorithms and methods for mathematical simulations on supercomputing systems, including hybrid ones*». Important highlight of the workshop is the active participation of the young scientists, and more than half of the talks were done by the young speakers.

Keywords: simulations, hybrid computing systems, simulations on supercomputers, methods and algorithms of parallel computing.