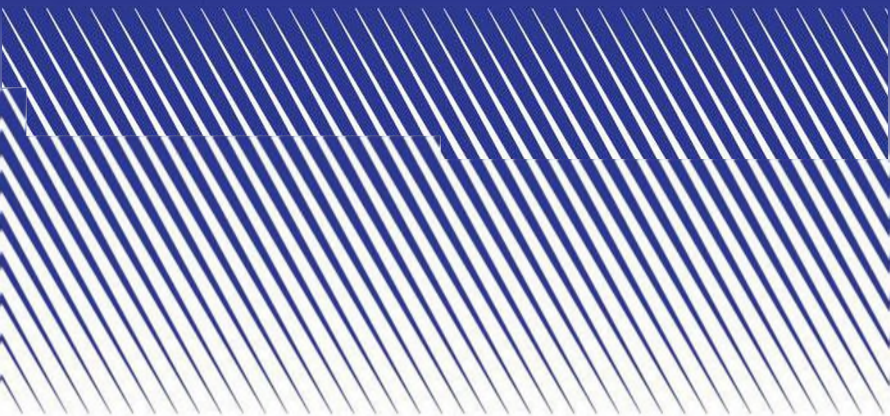




ИНСТИТУТ
КОСМИЧЕСКИХ
ИССЛЕДОВАНИЙ
РАН



ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ В ЕСТЕСТВЕННЫХ НАУКАХ МЕТОДЫ СУПЕРКОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ

СБОРНИК ТРУДОВ

1–3 октября 2014 года, Россия, Таруса

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

Серия «Механика, управление и информатика»

МОСКВА
ИКИ РАН
2014

**Вычислительные технологии в естественных науках.
Методы суперкомпьютерного моделирования**

Сборник трудов

Россия, Таруса, 1–3 октября 2014 г. Под ред. Р. Р. Назирова, Л. Н. Шура

Статьи, вошедшие в настоящий сборник, основаны на докладах, представленных на расширенном семинаре «Методы суперкомпьютерного моделирования», который прошёл 1–3 марта 2014 г. в Тарусе на базе гостиницы «Интеркосмос» Государственного бюджетного учреждения науки Института космических исследований Российской академии наук (ИКИ РАН). Это очередной семинар из серии расширенных семинаров, посвящённых вычислительным технологиям в естественных науках. На первом из них 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 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 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. Шура. 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 пятом 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. Шура. М.: ИКИ РАН, 2011, 148 с.). Шестой семинар был посвящён исследованиям в области перспективных компьютерных систем с акцентом на обсуждении концептуальных подходов, разработке новых методов и устройств (Вычислительные технологии в естественных науках. Перспективные компьютерные системы: устройства, методы и концепции / Под ред. Р. Р. Назирова, Л. Н. Шура. М.: ИКИ РАН, 2011, 192 с.). На седьмом семинаре обсуждались алгоритмы и методы в области вычислительной физики, причём особый акцент делался на обсуждении полученных новых физических результатов. Материалы этого семинара можно найти на страницах семинара <http://www.iki.rssi.ru/seminar/>. Восьмой семинар получил название «Инфокоммуникационные технологии в научных исследованиях» и его материалы были изданы в двойном номере журнала «Информационное общество» (вып. 1-2, 2013). Важная особенность проводимых нами расширенных семинаров состоит в привлечении молодёжи к активному участию в качестве докладчиков. Как правило, более половины докладов делаются молодыми исследователями. Семинар поддержан грантом Российского научного фонда (РНФ) (проект № 14-21-00158) «Разработка алгоритмов и методов для задач математического моделирования на суперкомпьютерных системах, включая гибридные».

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

Редактор: *Корниленко В. С.*
Компьютерная вёрстка: *Комарова Н. Ю.*
Дизайн обложки: *Захаров А. Н.*

Мнение редакции не всегда совпадает с точкой зрения авторов статей.
Отдельные статьи даны в авторской редакции.
Качество иллюстраций соответствует предоставленному авторами материалу.

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

СОДЕРЖАНИЕ

Математическое моделирование на суперкомпьютерах (Вместо предисловия) (<i>Назирова Р. Р., Шур Л. Н.</i>)	5
<i>Алешин С. В.</i> Оценка инвариантных числовых показателей аттракторов систем дифференциальных уравнений с запаздыванием	10
<i>Бараш Л. Ю., Халатников И. М.</i> Эффективная проводимость двумерных замощений плоскости: сравнение аналитических и численных результатов	18
<i>Бараш Л. Ю., Шур Л. Н.</i> Использование технологии multi-GPU для многомерного численного интегрирования методом Монте-Карло.	25
<i>Биктимиров М. Р., Жижченко А. Б., Овсянников А. П., Шер А. А., Климов П. А.</i> Организация эффективного сетевого доступа к ресурсам дата-центра в распределённой инфокоммуникационной инфраструктуре	33
<i>Васенин В. А., Роганов В. А., Зензинов А. А.</i> Гибридный подход к моделированию средств обеспечения информационной безопасности GRID/CLOUD-систем.	41
<i>Глызин С. Д.</i> Инвариантные числовые характеристики аттрактора и диффузионный хаос в уравнении Гинзбурга-Ландау с гантелеобразной пространственной областью.	54
<i>Демидов Д. Е.</i> Реализация генератора вычислительных ядер OpenCL в библиотеке VexCL	67
<i>Денисенко М. В., Сатанин А. М.</i> Мезоскопические флуктуации населённости кубита в поле управляющего сигнала	81
<i>Дружинин О. А., Троицкая Ю. И., Зилитинкевич С. С.</i> Прямое численное моделирование турбулентного устойчиво стратифицированного воздушного потока над взволнованной водной поверхностью	91
<i>Карбачинский И. О.</i> Проблемы построения систем защиты от спама в Интернете	104
<i>Меньшутин А. Ю., Шур Л. Н.</i> Многомерное обобщение модели Diffusion Limited Aggregation (DLA)	110

<i>Моисеенко С. Г., Бисноватый-Коган Г. С., Арделян Н. В.</i> Моделирование магниторотационных процессов в коллапсирующих сверхновых и развитие магнито-дифференциально-вращательной неустойчивости.	121
<i>Ноздрин Д. А.</i> Сведение задачи оптимизации потоков в интеллектуальной инженерной сети к задаче, решаемой симплекс-методом.	136
<i>Старостин Н. В., Панкратова М. А.</i> Архитектурно-зависимая декомпозиция в методиках суперкомпьютерного моделирования.	146
<i>Шикота С. К.</i> К задаче разработки методики построения распределённой инфокоммуникационной системы для научной и научно-учебной деятельности.	154
<i>Шур Л. Н.</i> Алгоритм Ванга-Ландау: случайное блуждание по спектру энергии.	160

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ НА СУПЕРКОМПЬЮТЕРАХ (ВМЕСТО ПРЕДИСЛОВИЯ)

Читателю предлагаются избранные статьи, написанные на основе приглашённых докладов на очередном, 9-м, расширенном семинаре по вычислительным методам в естествознании. Семинар под названием «Методы суперкомпьютерного моделирования» был проведён 1–3 октября 2014 г. на базе гостиницы «Интеркосмос» в Тарусе Института космических исследований Российской академии наук (ИКИ РАН) и был посвящён 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 состоит в активном участии молодых научных сотрудников и аспирантов, число которых составляет примерно половину участников. Также в его рамках проводились научно-учебные семинары с приглашёнными студентами, которые делают первые шаги в научных исследованиях по тематике семинара. Проведение семинара поддержано Российским научным фондом (РНФ) (проект № 14-21-00158 «Разработка алгоритмов и методов для задач математического моделирования на суперкомпьютерных системах, включая гибридные», руководитель Щур Л. Н.).

Представленные на семинаре доклады были разбиты на пять секций: вычислительная физика (восемь докладов); математическое моделирование (четыре доклада); информационные системы для науки (пять докладов); методы и алгоритмы (два доклада); гибридные вычисления (три доклада).

Научная программа конференции открылась секцией «Вычислительная физика», на которой Моисеенко Сергей Григорьевич (ИКИ РАН, Москва) представил доклад «Моделирование магниторотационных процессов в коллапсирующих сверхновых и развитие магнито-дифференциально-ротационной неустойчивости» [8]. Особое внимание докладчик уделил методу вычислений с перестройкой расчётной сетки с учётом локального выполнения законов сохранения. В докладе Дружинина Олега Александровича (Институт прикладной физики Российской академии наук (ИПФ РАН),

Нижний Новгород) «Прямое численное моделирование турбулентного воздушного потока над взволнованной водной поверхностью» [9] представлен разработанный автором подход к проблеме с интересным граничным условием в виде периодической поверхности воды, а также результаты моделирования. В докладе Бараша Льва Юрьевича (Институт теоретической физики им. Л. Д. Ландау Российской академии наук (ИТФ РАН), Научный центр РАН в Черногоровке) «Эффективная проводимость двумерных замощений плоскости: сравнение аналитических и численных результатов» [10] был изложен метод вычисления проводимости композитных систем, основанный на решении соответствующих дифференциальных уравнений на проводимость. На основе метода были проведены вычисления для тех случаев, которые исследовались ранее аналитически. Доклад Сатанина Аркадия Михайловича (Нижегородский государственный университет им. Н. И. Лобачевского, г. Нижний Новгород) «Динамика туннелирования Ландау-Зинера» [11] была изложена теория амплитудной спектроскопии. В докладе Щура Льва Николаевича (ИТФ РАН, Научный центр РАН в Черногоровке) «Алгоритм Ванга-Ландау: случайное блуждание по спектру энергии и эффективная параллелизация» [12] обсуждались тонкости алгоритма вычисления плотности состояний в спектре энергии спиновых систем. Меньшутин Антон Юрьевич (ИТФ РАН, Научный центр РАН в Черногоровке) «Многомерное обобщение модели DLA» [13] изложил пути расширения построенного ранее алгоритма моделирования структур роста на большие размерности. Доклад Муныева Вячеслава Олеговича (Нижегородский государственный университет им. Н. И. Лобачевского (ННГУ)) «Генерация субгармоник в квантовых джозефсоновских цепях» [14] был посвящён резонансным состояниям в джозефсоновских цепях, в частности, была предложена схема интерференции в три канала. Бобков Сергей Алексеевич (Национальный исследовательский центр (НИЦ) «Курчатовский институт», Москва) представил доклад «Методы классификации дифракционных изображений для эксперимента XFEL» [15], в котором обсуждались алгоритмы обработки экспериментальных дифракционных изображений: эти изображения будут получать на строящейся в Германии установке по генерации монохроматического рентгеновского пучка.

Секцию «Математическое моделирование» открыл доклад Глызина Сергея Дмитриевича (Ярославский государственный университет им. П. Г. Демидова (ЯрГУ), Научный центр РАН в Черногоровке) «Численные методы оценки ляпуновских экспонент и ляпуновской размерности аттрактора уравнения «реакция-диффузия» в гантелеобразной области» [16]. Его молодой коллега Алешин Сергей Владимирович (ЯрГУ, Научный центр РАН в Черногоровке) изложил в докладе «Параллельные алгоритмы в задаче об оценке инвариантных числовых показателей аттракторов систем дифференциальных уравнений с запаздыванием» [17] эвристические методы оценки ляпуновских показателей для систем с запаздыванием. Два других доклада секции были посвящены практическим задачам. Ноздрин Дмитрий Алексеевич (Национальный исследовательский университет Высшая школа экономики (НИУ ВШЭ), Москва) в докладе «Алгоритм управления распределением потоков мощностей в замкнутой энергосистеме электросети» [18] подробно изложил методы оптимизации электросиловых систем, а Меньшутин Антон Юрьевич (ИТФ РАН, Научный центр РАН в Черногоровке) «Анализ больших данных сотовых операторов» [19] предло-

жил использование анонимизированных данных сотовых операторов для оптимизации транспортных потоков.

Секция «Информационные системы для науки» открылась докладом Долгова Бориса Вячеславовича (Академинформ, Москва) «Использование IT-технологий и инфраструктур для агрегации научной информации. Опыт Канады, Нидерландов, Германии» [20], в котором были изложены принципы мониторинга научных сетей с целью оптимизации инфраструктуры для научной информации. Доклад Карбачинского Игоря Олеговича (НИУ ВШЭ) «Проблемы построения систем защиты от спама в Интернете» [21] изложил способы анализа потоков данных из сети Интернет для построения эффективной системы защиты информации. Доклад Зензинова Андрея Александровича (Московский государственный университет (МГУ) имени М. В. Ломоносова) «Гибридный подход к моделированию средств обеспечения информационной безопасности Grid/Cloud-систем» [22] был посвящён построению систем имитационного моделирования узлов компьютерных систем и анализу её устойчивости к неправомерным действиям. Климов Павел Алексеевич (Вычислительный центр им. А. А. Дородницына Российской академии наук (ВЦ РАН), Москва) представил доклад «Организация эффективного сетевого доступа к ресурсам дата-центра в распределённой инфокоммуникационной инфраструктуре» [23]. Роганов Владимир Александрович (МГУ) в докладе «Динамическое распараллеливание программ: новые архитектуры, методы и приложения» [24] поделился опытом динамического распараллеливания программ.

Небольшая, но важная и интересная секция «Методы и алгоритмы» состояла из двух докладов. Панкратова Маргарита Александровна (Нижегородский государственный университет им. Н.И. Лобачевского (ННГУ)) в докладе «Архитектурно-зависимая декомпозиция в методиках суперкомпьютерного моделирования» [25] изложила итерационный метод для решения задачи оптимизации декомпозиции параллельных расчётов. Кривчиков Максим Александрович (МГУ) представил доклад «К формальной верификации программных комплексов суперкомпьютерного моделирования с использованием предметно-ориентированных языков» [26].

Внимание слушателей была повышенным в ходе работы секции «Гибридные вычисления», на котором было представлено три доклада. В докладе «Использование Multi-GPU для многомерного численного интегрирования методом Монте-Карло» [27] Бараш Лев Юрьевич (ИТФ РАН, Научный центр РАН в Черноголовке) изложил результаты тестирования разработанного программного обеспечения для классической задачи численного интегрирования методом Монте-Карло, которые показали хорошую масштабируемость метода вплоть до доступного числа графических узлов на отечественном суперкомпьютере «Ломоносов». Глызин Дмитрий Сергеевич (ЯрГУ, Научный центр РАН в Черноголовке) в докладе «Использование гибридных кластерных вычислений в численном решении задачи о фазовом переходе» [28] представил подход для моделирования процессов переноса тепла при строительстве фундаментов сооружений в условиях многолетней мерзлоты. Демидов Денис Евгеньевич (Казанский филиал Межведомственного суперкомпьютерного центра Российской академии наук (КазФ МСЦ РАН)) в докладе «Реализация генератора вычислительных ядер OpenCL/CUDA в библиотеке VexCL» [29] представил разработанную им библиотеку для оптимизации программирования на языке OpenCL.

Особое событие семинара – это двухчасовые воспоминания академика Исаака Марковича Халатникова, в которых он рассказывал о жизни и судьбе своего учителя, величайшего физика-теоретика, лауреата Нобелевской премии по физике, академика Льва Давидовича Ландау. Исаак Маркович, известный физик-теоретик, которому 17 октября 2014 г. исполнилось 95 лет, был ближайшим учеником и соратником Ландау. Воспоминания Халатникова поражали точностью дат, имён и чёткостью оценок. Одна из участниц семинара заметила, что никогда и не мечтала встретить современника легендарного Ландау. Мы планируем издать воспоминания академика Халатникова отдельно, после завершения работы по переводу аудиоматериала в текстовый [30].

Мультимедийные версии докладов можно найти по ссылке [31].

Литература

- [1] Труды семинара по вычислительным технологиям в естественных науках. Вып. 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. Вып. 3. Вычислительная химия и биология / Под ред. Р. Р. Назирова, Л. Н. Шура. 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. Шура. М.: ИКИ РАН, 2011.
- [6] Вычислительные технологии в естественных науках. Перспективные компьютерные системы: устройства, методы и концепции: Труды семинара / Под ред. Р. Р. Назирова, Л. Н. Шура. М.: ИКИ РАН, 2011.
- [7] Инфокоммуникационные технологии в научных исследованиях // Информационное общество. 2013. № 1-2.
- [8] *Моисеенко С. Г., Бисноватый-Коган Г. С.* «Моделирование магниторотационных процессов в коллапсирующих сверхновых и развитие магнито-дифференциально-ротационной неустойчивости»
- [9] *Дружинин О. А., Троицкая Ю. И., Зилитинкевич С. С.* «Прямое численное моделирование турбулентного воздушного потока над взволнованной водной поверхностью»
- [10] *Бараш Л. Ю., Халатников И. М.* «Эффективная проводимость двумерных замощений плоскости: сравнение аналитических и численных результатов»
- [11] *Денисенко М. В., Сатанин А. М.* «Динамика туннелирования Ландау-Зинера»
- [12] *Шур Л. Н.* «Алгоритм Ванга-Ландау: случайное блуждание по спектру энергии и эффективная параллелизация»
- [13] *Меньшутин А. Ю.* «Многомерное обобщение модели Diffusion Limited Aggregation (DLA)»
- [14] *Денисенко М. В., Муняев В. О., Сатанин А. М.* «Генерация субгармоник в квантовых джозефсоновских цепях»
- [15] *Бобков С. А., Теслюк А. Б., Горбцов О. Ю., Ефанов О. М., Голосова М. В., Вартанянц И. А., Ильин В. А.* «Методы классификации дифракционных изображений для эксперимента XFEL»

- [16] *Глызин С. Д.* «Численные методы оценки ляпуновских экспонент и ляпуновской размерности аттрактора уравнения «реакция-диффузия» в гантелеобразной области»
- [17] *Алешин С. В.* «Параллельные алгоритмы в задаче об оценке инвариантных числовых показателей аттракторов систем дифференциальных уравнений с запаздыванием»
- [18] *Ноздрин Д. А.* «Алгоритм управления распределением потоков мощностей в замкнутой энергосистеме электросети»
- [19] *Меньшиутин А. Ю.* «Анализ больших данных сотовых операторов»
- [20] *Биктимиров М. Р., Глебский В. Л., Долгов Б. В., Полицарнов С. А.* «Использование IT-технологий и инфраструктур для агрегации научной информации. Опыт Канады, Нидерландов, Германии»
- [21] *Карбачинский И. О.* «Проблемы построения систем защиты от спама в Интернете»
- [22] *Васенин В. А., Роганов В. А., Зензинов А. А.* «Гибридный подход к моделированию средств обеспечения информационной безопасности Grid/Cloud-систем»
- [23] *Биктимиров М. Р., Жиженко А. Б., Шер А. А., Овсянников А. П., Климов П. А.* «Организация эффективного сетевого доступа к ресурсам дата-центра в распределённой инфокоммуникационной инфраструктуре»
- [24] *Роганов В. А.* «Динамическое распараллеливание программ: новые архитектуры, методы и приложения»
- [25] *Старостин Н. В., Панкратова М. А.* «Архитектурно-зависимая декомпозиция в методиках суперкомпьютерного моделирования»
- [26] *Васенин В. А., Кривчиков М. А.* «К формальной верификации программных комплексов суперкомпьютерного моделирования с использованием предметно-ориентированных языков»
- [27] *Бараш Л. Ю., Шур Л. Н.* «Использование Multi-GPU для многомерного численного интегрирования методом Монте-Карло»
- [28] *Глызин Д. С.* «Использование гибридных кластерных вычислений в численном решении задачи о фазовом переходе»
- [29] *Демидов Д. Е.* «Реализация генератора вычислительных ядер OpenCL/CUDA в библиотеке VexCL»
- [30] *Халатников И. М.* «Воспоминания о Ландау»
- [31] Мультимедийные версии докладов. URL: <http://iki.rssi.ru/seminar/2014100103/>

Р. Р. Назиров,
доктор технических наук, ИКИ РАН
Л. Н. Шур
профессор, доктор физико-математических наук,
ИТФ РАН, Научный центр РАН в Черноголовке

УДК 517.929

ОЦЕНКА ИНВАРИАНТНЫХ ЧИСЛОВЫХ ПОКАЗАТЕЛЕЙ АТТРАКТОРОВ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ЗАПАЗДЫВАНИЕМ

С. В. Алешин

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

Среди инвариантных характеристик динамических систем большую роль играют ляпуновские показатели и ляпуновская размерность. Анализ спектра показателей Ляпунова широко применяется для исследования сложной динамики в системах обыкновенных дифференциальных уравнений и в моделях, сводящихся к отображениям. В конечномерном случае, по теореме Оселедца, линеаризованная на аттракторе система обыкновенных дифференциальных уравнений всегда является правильной по Ляпунову, и, тем самым, верхний предел может быть заменён обычным, что позволяет эффективно вычислять показатели Ляпунова. В статье рассматривается вопрос вычисления показателей Ляпунова для систем дифференциальных уравнений с запаздывающим аргументом, для которого данная теорема, вообще говоря, не работает. Приводятся результаты тестирования разработанного алгоритма для уравнения Хатчинсона. Отмечено хорошее совпадение результатов численного моделирования с утверждениями аналитического плана. Иллюстрируется применение алгоритма к некоторым задачам. В частности, рассматривается система уравнений Ланга-Кобаяши на возможность получения режима гиперхаоса.

Исследование выполнено при поддержке Российского научного фонда (РНФ) (проект № 14-21-00158).

Ключевые слова: показатели Ляпунова, системы дифференциальных уравнений с запаздывающим аргументом, уравнение Хатчинсона, модель Ланга-Кобаяши.

ВВЕДЕНИЕ

Определение показателей Ляпунова можно найти в книге Б. Ф. Былова и др. (1966). Для линейной системы из n уравнений, записанной в векторной форме,

$$\frac{dx}{dt} = A(t)x, \quad (1)$$

где $x \in \mathbb{R}^n$ — вектор действительных чисел длины n , а $A(t)$ — $n \times n$ матрица, показатель решения определяется формулой

$$\chi = \overline{\lim}_{t \rightarrow \infty} \frac{1}{t} \ln |x(t)|.$$

Случаи, когда показатели удаётся найти аналитически, исключительно редки. Для вычисления старшего показателя обычно применяют метод Бенеттина [Benettin et al., 1980]. Дальнейшее развитие данный метод получил в работе [Wolf et al., 1985]. В ней в вычислительный алгоритм авторы добавили перенормировку начальных условий по алгоритму Грамма-Шмита [Гантмахер, 1966], что позволило вычислять спектр показателей Ляпунова.

В конечномерном случае, по теореме Оселедеца [Оселедец, 1968], линеаризованная на аттракторе система вида (1) всегда является правильной по Ляпунову, и, тем самым, верхний предел может быть заменён на обычный, что позволяет эффективно вычислять показатели Ляпунова. В случае уравнений с запаздывающим аргументом и краевых задач такую теорему доказать не удаётся. Поэтому при разработке алгоритмов вычисления ляпуновских экспонент важно иметь модельное уравнение с запаздыванием, для которого спектр этих экспонент может быть вычислен каким-либо другим способом. Наличие такой задачи позволяет протестировать разработанный алгоритм и убедиться в его работоспособности.

В настоящей статье описывается расширение стандартного алгоритма для вычисления нескольких первых показателей Ляпунова для систем дифференциальных уравнений с запаздывающим аргументом вида:

$$\frac{dx}{dt} = F(t, x(t), x(t - \tau)),$$

где $F(t, x(t), x(t - \tau))$ — достаточно гладкая функция.

Приводятся результаты тестирования указанного алгоритма на примере уравнения Хатчинсона [Hutchinson, 1948]. Иллюстрируется применение разработанного алгоритма для вычисления спектра ляпуновских экспонент системы Ланга-Кобаяши [Глазков, 2008].

Оцениваемые инвариантные характеристики отображений не являются в строгом смысле именно показателями Ляпунова. Но для случая равновесия в уравнении Хатчинсона, как будет показано ниже, получены результаты, иллюстрирующие «близость» оцениваемых характеристик и ляпуновских экспонент.

АЛГОРИТМ ПОСТРОЕНИЯ СПЕКТРА ПОКАЗАТЕЛЕЙ ЛЯПУНОВА ДЛЯ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ЗАПАЗДЫВАЮЩИМ АРГУМЕНТОМ

Опишем алгоритм получения первых K показателей Ляпунова для систем дифференциальных уравнений с запаздывающим аргументом следующего вида:

$$\frac{dx_0}{dt} = F(t, x_0(t), x_0(t - \tau)), \quad (2)$$

где для $\forall t x_0(t) \in \mathbb{R}^J$; J — размерность системы.

В качестве фазового пространства примем пространство непрерывных на отрезке $[-\tau, 0]$ функций в \mathbb{R}^J , а именно $C([-\tau, 0]; \mathbb{R}^J)$.

Для решения системы (2) численными методами требуется перейти от непрерывного времени к дискретному. Для этого выберем $N + 1 \geq K$ точек, с помощью которых разобьём отрезок длины τ на N равных частей. Длина каждой части будет равна $\delta = \tau/N$.

В качестве численного метода для решения системы (2) с начальными условиями

$$x_0(\varphi) = f_0(\varphi), \quad \varphi \in [-\tau, 0], \quad f_0(\varphi) \in \mathbb{R}^J \quad (3)$$

выберем метод Дормана-Принса восьмого порядка (DOPRI853) [Hairer et al., 2008] с постоянным шагом (в качестве длины шага выберем δ).

Будем решать данную систему (2) с соответствующим начальным условием (3) выбранным методом до момента времени θ — времени достаточного приближения траектории решения к аттрактору. При этом на промежутке $t \in [\theta - \tau, \theta]$ получим решение $x_0^0(t) \in \mathbb{R}^J$, которое в дальнейшем будет выступать в качестве начального условия. Используя его, получим решение $x_*(t) \in \mathbb{R}^J$, на котором будем оценивать ляпуновские показатели.

Дополним систему уравнений (3) с начальным условием $x_0^0(t)$ следующими K идентичными системами

$$\dot{x}_i(t) = Ax_i(t) + Bx_i(t - \tau), \quad (4)$$

где

$$A = \{a_{ij}\}; \quad a_{ij} = \frac{\partial F_i(t, x_*(t), x_*(t - \tau))}{\partial x_j(t)}; \quad B = \{b_{ij}\}; \quad b_{ij} = \frac{\partial F_i(t, x_*(t), x_*(t - \tau))}{\partial x_j(t - \tau)};$$

$$i = 1, \dots, K, \quad j = 1, \dots, J;$$

$$F(t, x_*(t), x_*(t - \tau)) = \begin{bmatrix} F_1(t, x_1(t), \dots, x_j(t), x_1(t - \tau), \dots, x_j(t - \tau)), \dots, \\ F_J(t, x_1(t), \dots, x_j(t), x_1(t - \tau), \dots, x_j(t - \tau)) \end{bmatrix}^T.$$

Они представляют собой линеаризованные на решении $x_*(t)$ системы уравнений (2). Для каждого уравнения из $K \times J$ уравнений системы

$$x_{i,j}^0(\varphi) = \sqrt{\frac{KJ}{N}}, \quad (5)$$

при $\varphi \in \left[(\theta - \tau) + \frac{N(iJ + j - 1)}{KJ}, (\theta - \tau) + \frac{N(iJ + j)}{KJ} \right], \quad i = 1, \dots, K, j = 1, \dots, J$ и

$$x_{i,j}^0(\varphi) = 0 \quad (6)$$

— в противоположном случае.

Таким образом:

- 1) евклидова норма начального условия i -й системы уравнений из (4) равна единице:

$$x_i^0(t) = 1 = \sqrt{\sum_{k=0}^N \sum_{j=1}^J x_{i,j,k}^2}, \quad i = 1, \dots, K; \quad (7)$$

- 2) скалярное произведение начального условия i -й системы уравнений на s -ю равно нулю $i \neq s$:

$$\langle x_i^0, x_s^0(t) \rangle = \sum_{k=0}^N \sum_{j=1}^J x_{i,j,k}^2 x_{s,j,k}^2 = \delta_{i,s}, \quad (8)$$

где $\delta_{i,s}$ — символ Кронекера; $x_{i,j,k}^2$ — квадрат значения функции $x_{i,j}^0(\varphi)$ в точке $(\theta - 1) + kN/K + \nu N/(KJ), k = 1, \dots, N, \nu = 1, \dots, J$.

Решая совместно систему (2) с начальным условием $x_0(\varphi) = x_0^0(\varphi)$ и системы уравнений (4) с начальными условиями (5), (6) ($\varphi \in [\theta - \tau, \theta]$) на проме-

жутке $t \in [\theta, \theta + T]$ получаем для каждой из систем решения $x_i^1(t) \in \mathbb{R}^J$ ($i = 0, \dots, K$).

Учитывая, что величины x_i^k ведут себя экспоненциальным образом, необходимо время от времени их перенормировывать. Отметим, что проблему представляет как неограниченный рост, так и стремление решения к нулю. Выбор времени перенормировки T можно осуществлять двумя различными способами: через равные промежутки времени или динамически [Глызин, 2007].

Далее на промежутке $t \in [\theta + T - \tau, \theta + T]$ ортонормируем полученные решения $x_i^1(t)$, $i = 1, \dots, K + 1$ методом Грама-Шмидта [Гантмахер, 1966]. При этом после процедуры ортогонализации и до начала процедуры нормирования вычисляем и запоминаем величины

$$\xi_i^1 = x_{\text{опт}}^1(t). \quad (9)$$

Затем повторно решаем систему (2), (4), где в качестве начальных условий используем полученные ортонормированные решения.

Посчитав

$$\lambda_i = \lim_{M \rightarrow \infty} \frac{\sum_{k=1}^M \xi_i^k}{TM}, \quad i = 1, \dots, K + 1, \quad (10)$$

получаем оценку показателей Ляпунова.

Перейдём к результатам тестирования алгоритма.

РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ ПРИВЕДЁННОГО АЛГОРИТМА НА ПРИМЕРЕ УРАВНЕНИЯ ХАТЧИНСОНА

Подробные результаты тестирования алгоритма приведены в работе [Алешин, 2012]. Повторим кратко основные результаты. Вычислительные эксперименты проводились для уравнения Хатчинсона, динамика которого была ранее изучена в работе [Глызин, 2007]. Оно имеет следующий вид:

$$\dot{x}_0 = rx_0(t)(1 - x_0(t-1)), \quad (11)$$

где $r > 0$.

Ненулевые решения уравнения (11) асимптотически устойчивы при $r \in \left[0, \frac{\pi}{2}\right]$, причём, при $r \in [0, e^{-1}]$ монотонно, а при $r \in \left[e^{-1}, \frac{\pi}{2}\right]$ решение стремится к единице колебательным образом. В этом случае показатели Ляпунова для уравнения Хатчинсона совпадают с действительными частями решений следующей системы уравнений:

$$\begin{cases} \tau + r \cos \omega e^{-\tau} = 0, \\ \omega + r \sin \omega e^{-\tau} = 0. \end{cases} \quad (12)$$

Действительные части τ , решения данной системы при r , равных 1,5 и 1,0, представлены во вторых столбцах табл. 1 и 2. Будем называть их эталонными значениями.

Для уравнения (11) линеаризованные уравнения имеют следующий вид:

$$\dot{x}_i = r(1 - x_0(t-1))x_i(t) - rx_0(t)x_i(t-1). \quad (13)$$

Для всех опытов использовались следующие параметры:

- количество вычисляемых показателей Ляпунова $K = 10$;
- время выхода на аттрактор $\theta = 150$;
- время перенормировки вектора начальных условий $T = 4$;
- количество пересчётов показателей Ляпунова $M = 5000$;
- начальное условие $x_0(\varphi) = 1/2 \sin \varphi + 1$, где $\varphi \in [-1, 0]$.

Таблица 1. Первые десять показателей Ляпунова для уравнения Хатчинсона при $R = 1,5$ и их разность с эталонными значениями

i	τ_i	λ_i	σ_i
1	-0,0328	-0,0331	0,0003
2	-0,0328	-0,0332	0,0004
3	-1,6509	-1,6521	0,0012
4	-1,6509	-1,6521	0,0012
5	-2,2447	-2,2462	0,0015
6	-2,2447	-2,2462	0,0015
7	-2,6130	-2,6145	0,0015
8	-2,6130	-2,6145	0,0015
9	-2,8811	-2,8827	0,0016
10	-2,8811	-2,8827	0,0016

Таблица 2. Первые десять показателей Ляпунова для уравнения Хатчинсона при $R = 1$ и их разность с эталонными значениями

i	τ_i	λ_i	σ_i
1	-0,3181	-0,3186	0,0005
2	-0,3181	-0,3187	0,0006
3	-2,0623	-2,0637	0,0014
4	-2,0623	-2,0637	0,0014
5	-2,6532	-2,6548	0,0016
6	-2,6532	-2,6548	0,0016
7	-3,0202	-3,0219	0,0017
8	-3,0202	-3,0219	0,0017
9	-3,2878	-3,2895	0,0017
10	-3,2878	-3,2895	0,0017

Перейдём к результатам численного моделирования.

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Численное моделирование проводилось для системы уравнений Ланга-Кобаяши [Глазков, 2008] динамики полупроводникового лазера, учитывающей воздействие отражённого излучения на резонатор:

$$\begin{cases} \frac{dE}{dt} = \nu(1 + i\alpha)EZ + \gamma e^{i(\omega - \omega_0)t} E(t-h), \\ \frac{dZ}{dy} = Q - Z - (1 + Z)|E|^2, \end{cases} \quad (14)$$

где $E(t)$ — комплексная амплитуда электрического поля; $Z(t)$ — инверсия носителей, $\gamma > 0$ — интенсивность внешнего излучения; ω и ω_0 — оптические частоты задающего и синхронизируемого лазера соответственно; Q — превышение током накачки первой пороговой величины; ν — отношение времени затухания инверсии носителей и инверсии фотонов в резонаторе; α — коэффициент уширения линии, отвечающий за нелинейное взаимодействие между амплитудой и фазой поля.

Для проведения вычислений перепишем систему уравнений (14) относительно вещественной и мнимой частей, и сделав замену $E(t) = E(t)e^{i\omega t}$, полагая $E = x_0 + iy_0$, имеем

$$\begin{cases} \dot{x}_0 = -\omega y_0 + \nu z_0(x_0 - \alpha y_0) + \\ \quad + \gamma [x_0(t-h)\cos((\omega - \omega_0)h) - y_0(t-h) - y_0(t-h)\sin((\omega - \omega_0)h)]; \\ \dot{y}_0 = \omega x_0 + \nu z_0(y_0 + \alpha x_0) + \\ \quad + \gamma [y_0(t-h)\cos((\omega - \omega_0)h) + x_0(t-h)\sin((\omega - \omega_0)h)]; \\ \dot{z}_0 = Q - z_0 - (1 - z_0)(x_0^2 + y_0^2). \end{cases} \quad (15)$$

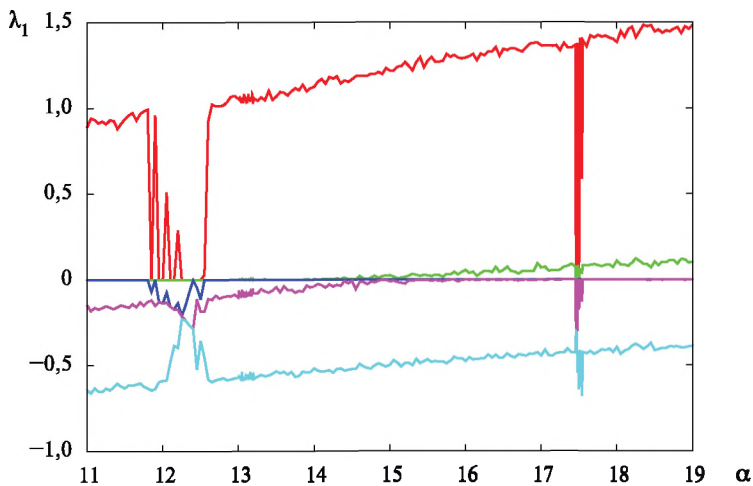
Линеаризованная на решениях $x_{i,*}, y_{i,*}, z_{i,*}$ система уравнений Ланга-Кобаяши имеет следующий вид:

$$\begin{cases} \dot{x}_i = \nu z_0 x_i - y_i(\nu z_0 \alpha + \omega) + \nu z_i(x_0 - \alpha y_0) + \\ \quad + \gamma [\cos((\omega - \omega_0)h)x_i(t-h) - y_i(t-h)\sin((\omega - \omega_0)h)]; \\ \dot{y}_i = \nu z_0 y_i + x_i(\nu z_0 \alpha + \omega) + \nu z_i(y_0 - \alpha x_0) + \\ \quad + \gamma [x_i(t-h)\sin((\omega - \omega_0)h) + y_i(t-h)\cos((\omega - \omega_0)h)]; \\ \dot{z}_i = -2(1 + z_0)x_0 x_i + 2(1 + z_0)y_0 y_i + z_i [1 + (x_0^2 + y_0^2)]. \end{cases} \quad (16)$$

На рисунке изображена зависимость первых пяти показателей Ляпунова от параметра α . Его значения изменялись в промежутке от 11 до 19. Красной линией нанесён график старшего показателя Ляпунова (λ_1). Остальные показатели, занумерованные в порядке возрастания (сверху вниз), изображены кривыми, нанесёнными, соответственно, зелёной, синей, фиолетовой и голубой линиями.

При значениях параметра α от 11 до 12 старший показатель Ляпунова системы Ланга-Кобаяши положительный, второй и третий — близки к нулю, третий и четвёртый — отрицательны. На этом промежутке установлен хаотический колебательный режим.

На промежутке от 12 до 12,6 хаотические колебания многократно сменяются периодическими, что видно из резких провалов значений старшей ляпуновской экспоненты до нуля. Далее при росте значений α до 14,7, как и в первом случае, есть один положительный показатель, два нулевых и два отрицательных.



Зависимость первых пяти показателей Ляпунова уравнения (15) от параметра α при $Q = 10$, $\nu = 2$, $\gamma = 2$, $\omega = 0,3$, $\omega_0 = 1$, $h = 1$

На промежутке от 14,7 до 17,46 присутствуют два положительных, два нулевых и один отрицательный показатель. Наличие двух положительных показателей Ляпунова указывает на появление режима гиперхаоса [Kaplan, Yorke, 1979].

Далее, при α от 17,46 до 17,5 находится промежуток, на котором хаотический режим сменяется периодическим, но при дальнейшем росте параметра вновь возвращается режим гиперхаоса.

ЗАКЛЮЧЕНИЕ

В статье рассмотрен вопрос вычисления показателей Ляпунова для систем дифференциальных уравнений с запаздывающим аргументом. Описан алгоритм получения нескольких первых ляпуновских экспонент. Приведён тестовый пример. Исследована система уравнений Ланга-Кобаяши на возможность получения режима гиперхаоса.

ЛИТЕРАТУРА

- [Алешин, 2012] Алешин С. В. Вычисление спектра показателей Ляпунова для дифференциальных уравнений с запаздывающим аргументом // Заметки по информатике и математике. 2012. № 4. С. 7–12.
- [Былов и др., 1966] Былов Б. Ф., Виноградов Р. Э., Гробман Д. М., Немыцкий В. В. Теория показателей Ляпунова и её приложения к вопросам устойчивости. М.: Наука, 1966.
- [Гантмахер, 1966] Гантмахер Ф. Р. Теория матриц, 2-е изд., М.: Наука, 1966.
- [Глазков, 2008] Глазков Д. В. Особенности динамики модели Ланга-Кобаяши в одном критическом случае // Моделирование и анализ информационных систем. 2008. Т. 15. № 2. С. 36–45.
- [Глызин, 2007] Глызин С. Д. Учёт возрастных групп в уравнении Хатчинсона // Моделирование и анализ информационных систем. 2007. Т. 14. № 3. С. 29–42.

- [Глызин и др., 2005] *Глызин Д. С., Глызин С. Д., Колесов А. Ю., Розов Н. Х.* Метод динамической перенормировки для нахождения максимального ляпуновского показателя хаотического аттрактора // Дифференциальные уравнения. 2005. Т. 41. № 2. С. 268–273.
- [Кашенко, Майоров, 2009] *Кашенко С. А., Майоров В. В.* Модели волновой памяти. М.: Кн. дом «ЛИБРОКОМ», 2009.
- [Оседедец, 1968] *Оседедец В. И.* Мультипликативная эргодическая теорема. Характеристические показатели Ляпунова динамических систем // Тр. Моск. матем. об-ва. 1968. Т. 19. С. 197–231.
- [Benettin et al., 1980] *Benettin G., Galgani L., Giorgilli A., Strelcyn J.-M.* Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: A method for computing all of them. Part I: Theory. Part II: Numerical application // Meccanica. 1980. V. 15. P. 9–30.
- [Hairer et al., 2008] *Hairer E., Wanner G., Norsett S. P.* Solving Ordinary Differential Equations 1 (Springer Series in Computational Mathematics): Nonstiff Problems. 2nd ed., revised. Springer, 2008.
- [Hutchinson, 1948] *Hutchinson G. E.* Circular causal in ecology // Annals of the New York Academy of Sciences. 1948. N. 50. P. 221–246.
- [Kaplan, Yorke, 1979] *Kaplan J. L., Yorke J. A.* Chaotic Behavior of Multidimensional Difference Equations, in Functional Differential Equations and Approximations of Fixed Points (Springer Lecture Notes in Mathematics. V. 730) / Ed. H.-O. Peitgen, H.-O. Walter. Springer, Berlin, 1979. P. 228–237.
- [Wolf et al., 1985] *Wolf A., Swift J. B., Swinney H. L., Vastano J. A.* Determining Lyapunov exponents from a time series // Physica D. 1985. V. 16. P. 285.

THE NUMERICAL EVALUATION OF ATTRACTORS EXPONENTS OF DELAY DIFFERENTIAL EQUATIONS SYSTEM

S. V. Aleshin

P.G. Demidov Yaroslavl State University (P.G. Demidov YSU)
Science Center in Chernogolovka, Russian Academy of Sciences

Lyapunov exponents and Lyapunov dimension play an important role among the invariant characteristics of dynamic systems. Analysis of the spectrum of Lyapunov exponents are widely used to study the complex dynamics in systems of ordinary differential equations and models that can be reduced to the maps. As follows from the Oseledets theorem in the finite-dimensional case the linearized on attractor system of the ordinary differential equations is always Lyapunov proper and thus the upper limit may be replaced by a normal limit, that allowing to effectively compute Lyapunov exponent. In this article we considered the question of numerical evaluation of Lyapunov exponents for delay differential equations. The Oseledets theorem not proved in this case. We tested our developed algorithm on Hutchinson equation with known Lyapunov spectra. Then algorithm was apply for Lang-Kobayashi model and we got the hyper chaos regime.

Keywords: Lyapunov exponents, delay differential equations, Hutchinson equation, Lang-Kobayashi model.

ЭФФЕКТИВНАЯ ПРОВОДИМОСТЬ ДВУМЕРНЫХ ЗАМОЩЕНИЙ ПЛОСКОСТИ: СРАВНЕНИЕ АНАЛИТИЧЕСКИХ И ЧИСЛЕННЫХ РЕЗУЛЬТАТОВ

Л. Ю. Бараш, И. М. Халатников

Институт теоретической физики им. Л. Д. Ландау Российской академии наук (ИТФ РАН), Научный центр РАН в Черноголовке

Представлен численный метод для получения эффективной проводимости двумерных замощений плоскости многоугольниками. Для периодических структур используется сеточный метод с релаксацией для решения уравнения Лапласа с соответствующими условиями сшивки на границах между областями компонент с разными проводимостями и с соответствующими периодическими граничными условиями. Метод позволяет определить эффективную проводимость с высокой точностью как в области применимости теории возмущений, так и в случаях, когда проводимости компонент существенно различны. Полученные результаты хорошо согласуются с имеющимися аналитическими результатами.

Работа выполнена при поддержке Российского научного фонда (РНФ) (проект № 14-21-00158).

Ключевые слова: композитные материалы, эффективная проводимость, теория макроскопически неоднородных сред, уравнение Лапласа.

ВВЕДЕНИЕ

Теория проводимости макроскопически неупорядоченных сред насчитывает более ста лет, её возникновение связано с именами Максвелла и Рэля [Rayleigh, 1892]. В первых теоретических работах были рассмотрены слабонеоднородные среды и системы с малой концентрацией включений сферической формы (см., напр., [Ландау, Лифшиц, 2005]). В дальнейшем для вычисления проводимости были предложены различные приближённые аналитические подходы, наиболее удачным из которых является теория эффективной среды Брюггемана (см, напр., обзор [Kirkpatrick, 1973]). Значительный интерес к таким системам также связан с изучением фазового перехода металл-диэлектрик по концентрации включений [Efros, Shklovskii, 1976]. В двумерном случае для проводимости бинарных систем был получен ряд важных точных аналитических результатов, таких как соотношение взаимности [Дыхне, 1970; Keller, 1964]. Из соотношения взаимности следует, что для случайно-неоднородной двумерной среды с равными концентрациями компонент (а также для системы со структурой шахматной доски) эффективная проводимость равна среднему геометрическому из проводимостей компонент. Двумерные случайные системы и периодические решётки относятся к одному классу универ-

Бараш Лев Юрьевич — младший научный сотрудник ИТФ им. Л. Д. Ландау РАН, научный сотрудник Научного центра РАН в Черноголовке, кандидат физико-математических наук, barash@itp.ac.ru

Халатников Исаак Маркович — почётный директор ИТФ им. Л. Д. Ландау РАН, главный научный сотрудник Научного центра РАН в Черноголовке, академик РАН, доктор физико-математических наук, профессор

сальности вблизи точки фазового перехода [Дыхне, 1970; Емец, 1989]. Аналитическое решение задачи о проводимости было также найдено для ряда двумерных двойкопериодических систем, в основном, с диэлектрическими или идеальнопроводящими включениями [Емец, 1986]. Для структуры шахматной доски известна точная зависимость эффективной проводимости от параметра $h = \sigma_2 / \sigma_1$ (здесь σ_1 и σ_2 — проводимости компонент композита), а также асимптотика электрического поля рядом с углами решётки, приводящая к нарушению линейного режима протекания тока при небольших значениях h [Дыхне, 1970; Емец, 1989; Сатанин и др., 1996]. Имеется весьма обширная литература, посвящённая нахождению различных приближений для эффективной проводимости для гетерогенных систем (см., напр., [Torquato, 2002]). Отметим также развитие метода двухточечных корреляционных функций для определения эффективной проводимости стохастических систем [Li et al., 2006], успешные применения математических методов из теории случайных блужданий для определения эффективной проводимости стохастических систем [Fel, Khanin, 2002]. В последние 10 лет теория эффективной проводимости активно применяется для исследования задач нанотехнологий: для определения эффективной проводимости композитных материалов с углеродными нанотрубками, с графеновыми нанолитами, для нахождения эффективной теплопроводности коллоидных наножидкостей и во многих других исследованиях [Bagchi, Nomura, 2006; Prasher, 2006; Xie et al., 2008].

В 2000 г. была разработана диаграммная техника, на основе которой были получены эффективные проводимости двумерных композитных систем в случаях, когда различие в проводимостях компонент композита мало по сравнению с самими проводимостями [Khalatnikov, Kamenshchik, 2000]. В настоящей работе представлен численный метод, который позволяет определить эффективную проводимость с высокой точностью как в области применимости теории возмущений, так и в случаях, когда проводимости компонент существенно различны. Ниже проведено тестирование развитого подхода для двумерной структуры шахматной доски с двухцветными и трёхцветными раскрасками и сравнение с имеющимися аналитическими результатами. В дальнейшем планируется рассмотрение более сложных систем.

МЕТОД ЧИСЛЕННОГО ОПРЕДЕЛЕНИЯ ЭФФЕКТИВНОЙ ПРОВОДИМОСТИ

Закон Ома для изотропной среды имеет вид $\mathbf{j} = -\sigma \nabla \varphi$. Для стационарного распределения токов имеет место уравнение непрерывности $\operatorname{div} \mathbf{j} = 0$. Следовательно, $\Delta \varphi + \nabla \ln \sigma \cdot \nabla \varphi = 0$. Представим потенциал φ в виде $\varphi = -\mathbf{E} \cdot \mathbf{r} + \psi(x, y)$, где среднее поле $\mathbf{E} = -\langle \nabla \varphi \rangle$, т.е. $\langle \nabla \psi \rangle = 0$. Угловые скобки обозначают здесь усреднение по всей плоскости. Итак, получаем уравнение для величины $\psi(x, y)$:

$$\Delta \psi + \nabla \ln \sigma \cdot (-\mathbf{E} + \nabla \psi) = 0. \quad (1)$$

Для изотропного распределения проводимости эффективная проводимость σ_{eff} определяется соотношением $\mathbf{J} = \sigma_{\text{eff}} \mathbf{E}$, где $\mathbf{J} = \langle \mathbf{j} \rangle$. Зная функ-

цию $\psi(x, y)$, можно определить эффективную проводимость, используя соотношение

$$\sigma_{\text{eff}} \mathbf{E} = \langle \mathbf{j} \rangle = \langle -\sigma(-\mathbf{E} + \nabla\psi) \rangle = \langle \sigma \rangle \mathbf{E} - \langle \sigma \nabla\psi \rangle. \quad (2)$$

Заметим, что если $\psi(x, y)$ — периодическая функция обеих координат — x и y , то $\langle \nabla\psi \rangle = 0$. Действительно, $\left\langle \frac{\partial\psi}{\partial y} \right\rangle = \frac{1}{S} \int \frac{\partial\psi}{\partial y} dx dy = \frac{1}{S} \int \left[\int \frac{\partial\psi}{\partial y} dy \right] dx = 0$, где интегрирование производится по элементарной ячейке площади S . Точно так же, $\langle \partial\psi/\partial x \rangle = 0$. Верно и обратное утверждение: если $\langle \nabla\psi \rangle = 0$ для задачи с периодическим замощением плоскости, то $\psi(x, y)$ — периодическая функция на плоскости. Действительно, при сдвиге на период замощения в любом направлении задача не меняется. Следовательно, ψ меняется на константу, но ψ остаётся прежним, так как иначе бы не выполнялось условие $\langle \nabla\psi \rangle = 0$. Для нахождения эффективной проводимости в такой задаче достаточно найти величину $\langle \nabla\psi \rangle$ в каждой области постоянной проводимости и применить соотношение (2).

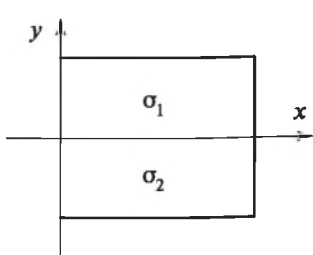
Для задачи с периодическим замощением плоскости уравнение (1) сводится к уравнению Лапласа $\Delta\psi = 0$ в каждой области постоянного σ и соответствующим условиям сшивки между областями. Выведем условие сшивки для случая горизонтальной границы между областями (рис. 1). Без ограничения общности граница проходит по линии $y = 0$, рядом с которой выполняется $\sigma(y) = \sigma_2 + (\sigma_1 - \sigma_2)\theta(y)$, где $\theta(y)$ — функция Хевисайда. Тогда $\partial\sigma/\partial y = (\sigma_1 - \sigma_2)\delta(y)$. Следовательно,

$$\int_{-\epsilon}^{\epsilon} \left[\Delta\psi + \frac{(\sigma_1 - \sigma_2)\delta(y)}{\sigma(y)} \left(-E_y + \frac{\partial\psi}{\partial y} \right) \right] dy = 0,$$

$$\frac{\partial\psi}{\partial y} \Big|_{-\epsilon}^{\epsilon} + \frac{2(\sigma_1 - \sigma_2)}{\sigma(y)} \left(-E_y + \frac{1}{2} \left[\left(\frac{\partial\psi}{\partial y} \right)_{+0} + \left(\frac{\partial\psi}{\partial y} \right)_{-0} \right] \right) = 0,$$

$$\sigma_1 \left(\frac{\partial\psi}{\partial y} \right)_{+0} - \sigma_2 \left(\frac{\partial\psi}{\partial y} \right)_{-0} = (\sigma_1 - \sigma_2) E_y. \quad (3)$$

Таким образом, на границе между областями, изображёнными на рис. 1, величина $\psi(x, y)$ непрерывна, а $\partial\psi/\partial y$ имеет скачок, определяемый формулой (3). Для решения уравнения (1) применим в каждой области сеточный метод с релаксацией, с соответствующими условиями сшивки на границах между областями и с соответствующими периодическими граничными условиями. Имеем в каждой области уравнение



$$\frac{1}{\kappa} \frac{\partial\psi}{\partial t} = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2}. \quad (4)$$

Разностные представления для вторых производных на равномерной сетке имеют вид

Рис. 1. Сшивка величины ψ на границе между областями с проводимостями σ_1 и σ_2

$$\hat{\delta}_x^2 \psi_{i,j} = \frac{1}{h_x^2} (\psi_{i+1,j} - 2\psi_{i,j} + \psi_{i-1,j}),$$

$$\hat{\delta}_y^2 \psi_{i,j} = \frac{1}{h_y^2} (\psi_{i,j+1} - 2\psi_{i,j} + \psi_{i,j-1}).$$

Как известно, неявный метод с переменной направлений для решения уравнения состоит из двух частей. На первом шаге метода производная по x учитывается неявно, и разностным представлением уравнения (4) является

$$\frac{\psi_{i,j}^{n+1/2} - \psi_{i,j}^n}{\kappa h_t / 2} = \hat{\delta}_x^2 \psi_{i,j}^{n+1/2} + \hat{\delta}_y^2 \psi_{i,j}^n.$$

Это выражение можно переписать в виде

$$c'_i \psi_{i-1,j}^{n+1/2} + \left(d'_i - \frac{2}{\kappa h_t} \right) \psi_{i,j}^{n+1/2} + e'_i \psi_{i+1,j}^{n+1/2} = c''_j \psi_{i,j-1}^n + \left(d''_j - \frac{2}{\kappa h_t} \right) \psi_{i,j}^n + e''_j \psi_{i,j+1}^n. \quad (5)$$

Здесь $c'_i = e'_i = 1/h_x^2$; $d'_i = -2/h_x^2$; $c''_j = e''_j = -1/h_y^2$; $d''_j = 2/h_y^2$. Правая часть выражения (5) известна, поскольку ψ на шаге n известно. Для каждого j применяем прогонку для решения системы уравнений (5), и получаем ψ на шаге $n + 1/2$.

На втором шаге метода производная по y берётся неявно, и разностным представлением для уравнения (4) становится

$$\frac{\psi_{i,j}^{n+1} - \psi_{i,j}^{n+1/2}}{\kappa h_t / 2} = \hat{\delta}_x^2 \psi_{i,j}^{n+1/2} + \hat{\delta}_y^2 \psi_{i,j}^{n+1}.$$

Это выражение можно переписать в виде

$$c''_j \psi_{i,j-1}^{n+1} + \left(d''_j + \frac{2}{\kappa h_t} \right) \psi_{i,j}^{n+1} + e''_j \psi_{i,j+1}^{n+1} = c'_i \psi_{i-1,j}^{n+1/2} + \left(d'_i + \frac{2}{\kappa h_t} \right) \psi_{i,j}^{n+1/2} + e'_i \psi_{i+1,j}^{n+1/2}. \quad (6)$$

Правая часть выражения (6) известна, поскольку ψ на шаге $n + 1/2$ известно. Для каждого i применяем прогонку для решения системы уравнений (6), и получаем ψ на шаге $n + 1$.

Для периодических граничных условий вместо трёхдиагональной матрицы A линейного уравнения, решаемого методом прогонки, имеем матрицу вида

$$A = \begin{pmatrix} d_1 & e_1 & 0 & 0 & \dots & c_1 \\ c_2 & d_2 & e_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & c_{n-1} & d_{n-1} & e_{n-1} \\ e_n & \dots & 0 & 0 & c_n & d_n \end{pmatrix}.$$

Как показано в работе [Karniadakis, Kirby, 2003], метод решения такой системы уравнений сводится к обычному алгоритму прогонки, применённому для усечённой матрицы

$$A^c = \begin{pmatrix} d_1 & e_1 & 0 & 0 & \dots & 0 \\ c_2 & d_2 & e_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & c_{n-2} & d_{n-2} & e_{n-2} \\ 0 & \dots & 0 & 0 & c_{n-1} & d_{n-1} \end{pmatrix}.$$

РЕЗУЛЬТАТЫ ВЫЧИСЛЕНИЙ

Для двумерной двухцветной шахматной доски результат численного определения эффективной проводимости совпадает с формулой Келлера-Дыхне. Это проверено численно для многих пар значений σ_1, σ_2 .

Для двумерной трёхцветной шахматной доски результат вычислений эффективной проводимости при $|\sigma_1 - \sigma_2| \ll \sigma_1$ совпадает с результатами работы [Khalatnikov, Kamenshchik, 2000] для всех рассмотренных раскрасок и порядков теории возмущений. На рис. 2 изображены замощения плоскости, рассмотренные в работе [(Khalatnikov, Kamenshchik, 2000)]. Отметим, что для замощения плоскости, приведённого на последнем рисунке, вычисленная эффективная проводимость совпадает с результатом, полученным в работе [Khalatnikov, Kamenshchik, 2000] при помощи теории возмущений, но не совпадает с приближённым значением эффективной проводимости, полученным при помощи теории среднего поля Брюгеммана, предполагающей круглую форму включений.

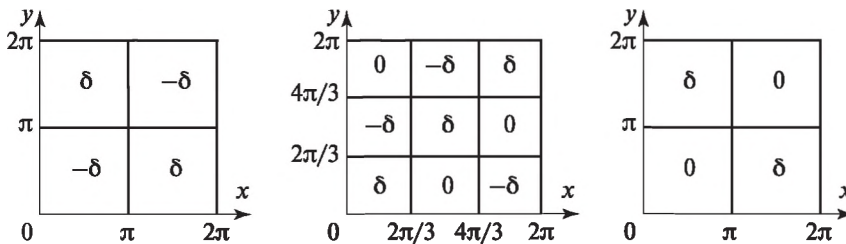


Рис. 2. Замощения плоскости, рассмотренные в работе [Khalatnikov, Kamenshchik, 2000]

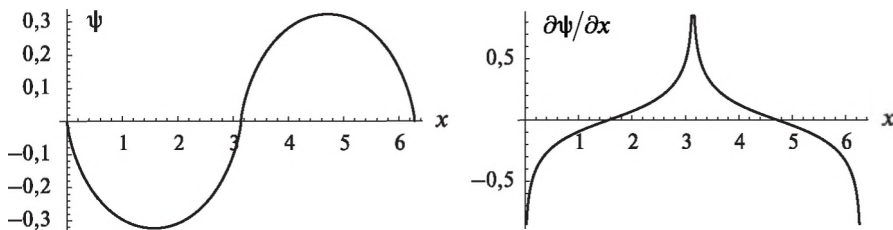


Рис. 3. Вычисленные значения величин ψ и $\partial\psi/\partial x$ вдоль горизонтальной границы для структуры двухцветной шахматной доски с $\sigma_1 = 1,3$, $\sigma_2 = 0,7$ и периодом 2π

Компьютерное моделирование подтверждает интегрируемую расходимость в локальном электрическом поле вблизи углов границ областей со степенной зависимостью, приведённой в работе [Сатанин и др., 1996]. На рис. 3 приведены графики величин ψ и $\partial\psi/\partial x$ вдоль горизонтальной границы, которые иллюстрируют расходимость локального электрического поля вблизи углов границ структуры двухцветной шахматной доски с $\sigma_1 = 1,3$, $\sigma_2 = 0,7$ и периодом $2l$. Расходимости не влияют на построение теории возмущений для эффективной проводимости по малому параметру $(\sigma_1 - \sigma_2)/\sigma_1 = 1 - h$, поскольку они интегрируемы и показатель степени $2 \operatorname{arctg} \frac{1-h}{2\sqrt{h}}$ в этом случае мал. В силу интегрируемости особенностей локального поля, расходимость не возникает в самой величине ψ , определяемой численно.

Проведённые расчёты показывают, что развитый метод хорошо работает для периодических замощений плоскости со структурой шахматной доски с двухцветными и трёхцветными раскрасками и может быть применён к более сложным случаям.

ЛИТЕРАТУРА

- [Дыхне, 1970] Дыхне А. М. Проводимость двумерной двухфазной системы // Журн. эксперим. и теор. физики (ЖЭТФ). 1970. Т. 59. С. 110.
- [Емец, 1986] Емец Ю. П. Электрические характеристики композиционных материалов с регулярной структурой. Киев: Наукова думка, 1986. 192 с.
- [Емец, 1989] Емец Ю. П. Преобразование симметрии двумерной двухкомпонентной электропроводной системы // Журн. эксперим. и теор. физики (ЖЭТФ). 1989. Т. 96. С. 701–711.
- [Ландау, Лифшиц, 2005] Ландау Л. Д., Лифшиц Е. М. Теоретическая физика: уч. пособие. Т. 8. Электродинамика сплошных сред. М.: Физматлит, 2005. 649 p.
- [Сатанин и др., 1996] Сатанин А. М., Скузоваткин В. В., Хорьков С. В. Нарушение линейного режима протекания тока в периодических структурах // Письма в ЖЭТФ. 1996. Т. 64(7). С. 495–499.
- [Bagchi, Nomura, 2006] Bagchi A., Nomura S. On the effective thermal conductivity of carbon nanotube reinforced polymer composites // Composites Science and Technology. 2006. V. 66. P. 1703–1712.
- [Efros, Shklovskii, 1976] Efros A. L., Shklovskii B. I. Critical behavior of conductivity and dielectric constant near metal-non-metal transition threshold // Physica Status Solidi B: Basic Research. 1976. V. 76. N. 2. P. 475–485.
- [Fel, Khanin, 2002] Fel L. G., Khanin K. M. On Effective Conductivity on \mathbb{Z}^d Lattice // J. Statistical Physics. 2002. V. 108. P. 1015–1031.
- [Karniadakis, Kirby, 2003] Karniadakis G. E., Kirby R. M. Parallel Scientific Computing in C++ and MPI. Cambridge: Cambridge University Press, UK, 2003. 628 p.
- [Keller, 1964] Keller J. B. Theorem on conductivity of composite medium // J. Mathematical Physics. 1964. V. 5. P. 548–549.
- [Khalatnikov, Kamenshchik, 2000] Khalatnikov I. M., Kamenshchik A. Yu. A Diagram Technique For Perturbation Theory Calculations Of The Effective Conductivity Of Two-Dimensional Systems // Журн. эксперим. и теор. физики (ЖЭТФ). 2000. Т. 118. Вып. 6. С. 1456–1462 [JETP. 2000. V. 91(6). P. 1261–1267].

- [Kirkpatrick, 1973] *Kirkpatrick S.* Percolation and conduction // *Reviews of Modern Physics*. 1973. V. 45. N. 4. P. 574–588.
- [Li et al., 2006] *Li D. S., Sahelia G., Khaleelb M., Garmestani H.* Quantitative Prediction of Effective Conductivity in Anisotropic Heterogeneous Media Using Two–point Correlation Functions // *Computational Materials Science*. 2006. V. 38. Iss. 1. P. 45–50.
- [Prasher et al., 2006] *Prasher R, Evans W, Fish J.* et al. Effect of aggregation on thermal conduction in colloidal nanofluids // *Applied Physics Letters*. 2006. V. 89. P. 143119. doi: 10.1063/1.2360229.
- [Rayleigh, 1892] *Lord Rayleigh.* On the Instability a Cylinder of Viscous Liquid under Capillary Force // *Philosophical Magazine and J. Sciences*. 1892. Ser. 5. V. 34. N. 207. P. 145–155.
- [Torquato, 2002] *Torquato S.* Random heterogeneous materials: microstructure and macroscopic properties. Springer, 2002. 701 p.
- [Xie et al., 2008] *Xie S. H., Liu Y. Y., Li J. Y.* Comparison of the effective conductivity between composites reinforced by graphene nanosheets and carbon nanotubes // *Applied Physics Letters*. 2008. V. 92. Iss. 24. P. 243121.

EFFECTIVE CONDUCTIVITY OF THE PLANE CONSISTING OF PIECES OF DIFFERENT CONDUCTIVITIES: COMPARING ANALYTICAL AND NUMERICAL RESULTS

L. Yu. Barash, I. M. Khalatnikov

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

The numerical method for calculating effective conductivity of the plane consisting of pieces of different conductivities is presented. For periodic structures we use a finite difference method with relaxation for solving Laplace equation with corresponding matching conditions between components and corresponding periodic boundary conditions. The method allows one to obtain effective conductivity with high accuracy both when perturbation theory is applicable and when the conductivities of components are essentially different. The results obtained are in good agreement with available analytical results.

Keywords: composites, effective conductivity, macroscopically disordered media, Laplace equation.

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

Khalatnikov Isaak Markovich — honorary director Landau Institute for Theoretical Physics RAS; principal researcher, Science Center in Chernogolovka RAS, Academician of Russian Academy of Sciences, Doctor of Science, professor

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ MULTI-GPU ДЛЯ МНОГОМЕРНОГО ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ МЕТОДОМ МОНТЕ-КАРЛО

Л. Ю. Бараш, Л. Н. Щур

Институт теоретической физики им. Л. Д. Ландау Российской академии наук (ИТФ РАН), Научный центр РАН в Черноголовке

Задача многомерного численного интегрирования методом Монте-Карло имеет многочисленные приложения. Применение современных методов вычислений с использованием гибридных вычислительных систем с графическими процессорами позволяет значительно увеличить производительность при выполнении задач интегрирования. Вычисление многомерного интеграла методом Монте-Карло с использованием одиночного графического процессора даёт повышение производительности в 40–100 раз по сравнению с использованием центрального процессора. Также продемонстрировано линейное увеличение производительности вычислений Монте-Карло при увеличении числа узлов гибридной вычислительной системы с графическими процессорами.

Работа выполнена при частичной поддержке РФФИ (проекты № 12-07-13121 и 13-07-00570), а также с использованием ресурсов суперкомпьютерного комплекса МГУ имени М. В. Ломоносова.

Ключевые слова: статистические методы, метод Монте-Карло, случайные числа, графический процессор, численное интегрирование.

ВВЕДЕНИЕ

Численные методы Монте-Карло активно применяются для научно-прикладных расчётов и прикладных целей, а также в задачах вычислительной статистики, в операционных исследованиях, вычислительной физике, квантовой физике, физической химии, структурной механике, вычислительных финансах и в других областях фундаментальных и прикладных исследований. Например, это класс задач компьютерного моделирования свойств новых материалов, в том числе наноматериалов, и моделирования квантовых статистических систем. Это задачи прохождения потока элементарных частиц в твёрдых, жидких, газообразных и плазменных средах. Это класс задач протекания, например, проблема повышения эффективности извлечения нефти и газа из нефтегазоносных месторождений или задача распространения инфекции в различных средах, таких как ботанический сад или компьютерные сети.

При моделировании по методу Монте-Карло производительность вычислений может быть повышена на много порядков при использовании параллельных вычислительных систем, а особенно при использовании гибридных вычислительных систем с ускорителями вычислений на основе графических процессоров.

В настоящей статье это показано на примере использования гибридных суперкомпьютеров для решения классической задачи многомерного численного

Бараш Лев Юрьевич — младший научный сотрудник ИТФ им. Л. Д. Ландау РАН, научный сотрудник Научного центра РАН в Черноголовке, кандидат физико-математических наук, barash@itp.ac.ru

Щур Лев Николаевич — заведующий отделом, доктор физико-математических наук, профессор, shchur@chg.ru

интегрирования методом Монте-Карло. Эта задача имеет приложения в физике частиц, где многомерное численное интегрирование широко применяется для обработки данных с ускорителей элементарных частиц. Например, при обработке данных с Большого адронного коллайдера обнаружение новых событий распада привело к экспериментальному подтверждению существования бозона Хиггса, за предсказание которого была присуждена Нобелевская премия двум теоретикам — П. Хиггсу и Ф. Энглеру — в октябре 2013 г. В банковском секторе многомерное численное интегрирование используется как часть программного обеспечения, применяемого при построении прогнозов. Многомерное численное интегрирование — это классическая задача использования метода Монте-Карло, которая выбрана эталоном для проверки эффективности программно-аппаратных решений, методов и алгоритмов.

Рассмотрим простейший пример задачи, в которой требуется многомерное интегрирование методом Монте-Карло. Предположим, что известна зависимость от координаты и скорости полной энергии десяти взаимодействующих частиц. В трёхмерном пространстве каждая частица имеет три компоненты скорости и три компоненты координаты. Если рассматривать движение относительно центра масс и относительные координаты, то в данном примере полная энергия является функцией 54 координат. Следовательно, для вычисления средней энергии на одну частицу необходимо вычислить интеграл в 54-мерном пространстве.

Известно, что зависимость ошибки от числа n точек интегрирования для обычного численного метода вычисления интегралов убывает как n^{-a} в одномерном случае (значение величины a зависит от выбранного метода). В размерности d ошибка ведёт себя как $n^{-a/d}$, т.е. убывает тем медленнее, чем больше размерность пространства. Напротив, при вычислении многомерных интегралов по методу Монте-Карло ошибка убывает как $n^{-1/2}$ вне зависимости от размерности интеграла. Заметим, что вычислительное время приблизительно пропорционально n и в классических методах интегрирования, и в методе Монте-Карло. Поэтому для малой размерности d детерминированные методы численного интегрирования, такие как метод Симпсона, являются предпочтительными, если область интегрирования не слишком сложна. Численное интегрирование по методу Монте-Карло даёт преимущества для вычисления многомерных интегралов. Для вычисления интегралов большой размерности метод Монте-Карло в настоящее время не имеет альтернативы.

ЧИСЛЕННОЕ ИНТЕГРИРОВАНИЕ МЕТОДОМ МОНТЕ-КАРЛО

Рассмотрим, например, двумерный интеграл

$$F = \int_R f(x, y) dx dy, \quad (1)$$

где R обозначает область интегрирования. Для его вычисления простым методом Монте-Карло следует построить прямоугольник, покрывающий область R , и сгенерировать n случайных точек (x_i, y_i) , равномерно распределённых внутри прямоугольника. Для оценки интеграла следует использовать формулу

$$F_n = \frac{A}{n} \sum_{i=1}^n f(x_i, y_i) H(x_i, y_i), \quad (2)$$

где A — площадь прямоугольника; $H(x, y)$ — функция Хэвисайда, т.е. $H(x, y) = 1$, если $(x, y) \in R$, и $H(x, y) = 0$, если это не так. Выражение для F_n представляет площадь области R , умноженную на среднее значение функции $f(x, y)$ в R .

Чтобы оценить ошибку вычислений, полезно вычислить выборочную дисперсию $\tilde{\sigma}^2$, определяемую как

$$\tilde{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (f(x_i) - \langle f \rangle)^2, \quad (3)$$

где $\langle f \rangle = \frac{1}{n} \sum_{i=1}^n f(x_i)$. Число $(n-1)$ записано в знаменателе по следующим соображениям. Чтобы вычислить выборочную дисперсию, нужно посчитать значения функции в n точках и вычислить по ним f , следовательно, останется только $n-1$ — независимая переменная для вычисления $\tilde{\sigma}^2$. При больших значениях n выборочная дисперсия стремится к дисперсии $\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2$, где $\langle f^2 \rangle = \frac{1}{n} \sum_{i=1}^n f(x_i)^2$.

Можно показать, что реальная ошибка вычисления интеграла для простого метода Монте-Карло при больших n ведёт себя как

$$\sigma_m = \frac{\sigma}{\sqrt{n-1}} \approx \frac{\sigma}{\sqrt{n}}. \quad (4)$$

Следовательно, имеется всего два способа уменьшить ошибку интегрирования по методу Монте-Карло: либо увеличить число пробных точек, либо уменьшить дисперсию.

Один из способов уменьшения дисперсии заключается в следующем: вместо равномерного заполнения области интегрирования пробными точками, более плотно заполнять пробными точками те области, где значение функции $f(x)$ является бóльшим по величине или быстро меняется. Такие выборки точек называются выборками по значимости.

Строго говоря, в области интегрирования вводится положительная функция распределения $p(x)$, такая что $\int_R p(x) dx = 1$, тогда интеграл можно переписать в виде $F = \int_R \left| \frac{f(x)}{p(x)} \right| p(x) dx$, выборку пробных точек производить не равномерно, а согласно распределению вероятности $p(x)$, а интеграл вычислять при помощи суммы

$$F_n = \frac{1}{n} \sum_{i=1}^n \frac{f(x_i)}{p(x_i)}. \quad (5)$$

Идея заключается в таком выборе $p(x)$, чтобы дисперсия отношения $f(x)/p(x)$ была минимальной. Для этого форма $p(x)$ выбирается так, чтобы

она повторяла форму $f(x)$ везде, где можно, особенно в областях, где величина $f(x)$ велика. В результате отношение $f(x)/p(x)$ является медленно меняющейся функцией, имеющей малую дисперсию.

Другой способ уменьшения дисперсии — так называемая стратифицированная выборка. Эта техника заключается в разбиении полной области интегрирования на подобласти, интегрирование по методу Монте-Карло в каждой подобласти и, затем, суммирование результатов. Разобьём область интегрирования $M = [0, 1]^d$ на k областей M_j , где $j = 1, 2, \dots, k$. В каждой области M_j производится интегрирование по методу Монте-Карло с количеством точек, равным N_j . Для интеграла получаем оценку

$$F_N = \sum_{j=1}^k \frac{\text{vol}(M_j)}{N_j} \sum_{n=1}^{N_j} f(x_{j,n}), \quad (6)$$

а вместо дисперсии $\sigma^2(f(x))/N$ имеем теперь выражение

$$\sum_{j=1}^k \frac{\text{vol}(M_j)}{N_j} \sigma^2(f(x)) \Big|_{M_j}, \quad (7)$$

где

$$\sigma^2(f(x)) \Big|_{M_j} = \left[\frac{1}{\text{vol}(M_j)} \int_{M_j} dx f(x)^2 \right] - \left[\frac{1}{\text{vol}(M_j)} \int_{M_j} dx f(x) \right]^2. \quad (8)$$

Точный выбор подобластей и числа точек в каждой подобласти может, по сравнению с простым методом Монте-Карло, привести к резкому уменьшению дисперсии, в то же время неправильный их выбор может привести и к увеличению дисперсии. Как правило, общая дисперсия минимизируется, если числа точек в каждой подобласти пропорциональны $\sigma(f(x)) \Big|_{M_j}$.

РЕАЛИЗАЦИЯ АЛГОРИТМА ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ МЕТОДОМ МОНТЕ-КАРЛО

Для проведения расчётов многомерного численного интегрирования было создано программное обеспечение при помощи технологий MPI и CUDA. В процессе разработки были использованы программные коды из статей [Hahn, 2007; Kanzaki, 2011], также использовались разработанные нами библиотеки случайных чисел RNGSSELIB [Barash, Shchur, 2013] и PRAND [Barash, Shchur, 2014]. В результате разработано прикладное программное обеспечение GINT, которое реализует метод Монте-Карло для многомерного численного интегрирования для параллельных гибридных систем с ускорителями на основе графических процессоров.

Программа многомерного интегрирования имеет следующую структуру. Для каждой переменной диапазон её значений разбивается на интервалы, таким образом, область интегрирования разбивается на прямоугольную сетку, состоящую из гиперкубов. В каждом из них производится интегрирование по методу Монте-Карло, а дисперсии, вычисленные в этих гиперкубах, ис-

пользуются для определения нового разбиения диапазона значений каждой переменной на интервалы на следующей итерации, в частности, для определения оптимального числа пробных точек в каждом из гиперкубов на следующей итерации. Дисперсия для полного интеграла уменьшается от итерации к итерации. Схема вычислений выглядит следующим образом.

- 1) инициализация параметров;
- 2) генерация N пробных точек, каждая из которых является вектором из k случайных чисел, если размерность равна k ;
- 3) вычисление подынтегральной функции на множестве пробных точек;
- 4) суммирование значений подынтегральной функции и их квадратов для всей области интегрирования, а также в каждом гиперкубе; вычисление средних величин и дисперсий;
- 5) оптимизация сетки и других параметров после вычисления N значений;
- 6) повторение шагов 2–5 вплоть до M итераций.

При этом на всех итерациях, кроме первой, используется выборка по значимости с плотностью распределения вероятности, равной

$$p(x) = \frac{|f(x)|}{\int |f(x)| dx}.$$

Чтобы не хранить гигантский объём данных о функции $p(x)$, для каждой переменной используется своя одномерная функция распределения, которую ищем таким образом, чтобы выполнялось условие

$$p(u_1, \dots, u_d) = p_1(u_1)p_2(u_2) \dots p_d(u_d).$$

После нескольких итераций оптимальная сетка будет, в конце концов, найдена. Таким образом, в программе используются хорошо известные методы уменьшения дисперсии, выборка по значимости и стратифицированная выборка.

Наиболее существенную часть компьютерного времени (около 99 %) занимают шаги 2 и 3, и именно эти шаги вычисления можно значительно ускорить при помощи гибридной вычислительной системы с графическими процессорами. Для этого на каждой итерации управляющий узел (master) раздаёт задания (наборы гиперкубов) остальным узлам (slave). Каждый из остальных узлов производит вычисления с помощью процессора GPU (генерацию пробных точек при помощи генераторов псевдослучайных чисел, вычисления подынтегральной функции на множестве пробных точек и средних величин и дисперсий в каждом гиперкубе) и возвращает обобщённые вычисленные данные управляющему узлу. Этот узел затем окончательно выполняет шаги 4 и 5 для всей области интегрирования, вычисляет значение интеграла и погрешность интегрирования и переходит к следующей итерации в том случае, если точность всё ещё не достигает требуемой.

Подынтегральная функция задаётся на языке программирования C++ в виде `float func (float* vars, float weight)`, где `vars` содержит массив значений всех переменных, а `weight` — весовая функция. Также задаются следующие параметры вычисления: `ncall` — число пробных точек на каждой итерации;

itmx — максимальное число итераций; acc — требуемая точность в течение каждой итерации; nBlockSize — число нитей в каждом блоке нитей в GPU. Для itmx и nBlockSize значения используются напрямую, в то же время значения ncall и acc вычисляются из заданных пользователем значений как $ncall=1024*ncall0$, $acc=0.00001*acc0$.

Программа компилируется в три этапа: сначала запускаются компиляторы trisxx и nvcc в режиме компиляции (-c) для отдельных частей программы, соответствующих использованию MPI и CUDA, а затем компоновщик trisxx компоует все части программы в общий запускаемый файл.

Программа создаёт следующие выходные данные: 1) значение вычисленного интеграла; 2) значение точности вычисленного результата, вычисленное при помощи «хи-квадрат»-статистики.

РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ ПРОГРАММЫ

Тесты показали, что вычисление многомерного интеграла методом Монте-Карло с использованием одиночного графического процессора даёт повышение производительности в 40–100 раз по сравнению с использованием центрального процессора. При этом на системе, состоящей из большого количества вычислительных узлов, скорость расчёта увеличивается пропорционально количеству узлов. Для тестирования программы использовалось одновременно от 1 до 750 GPU. Время выполнения задачи на одном узле составило 21 700 с.

На рис. 1 и 2 приведены графики масштабируемости, т. е. графики эффективности параллелизации $k=T(1)/T(N)$ в зависимости от N . Здесь $T(N)$ — время вычисления многомерного интеграла с заданной точностью (не зависящей от N) при запуске вычисления в режиме использования N -mpi-узлов.

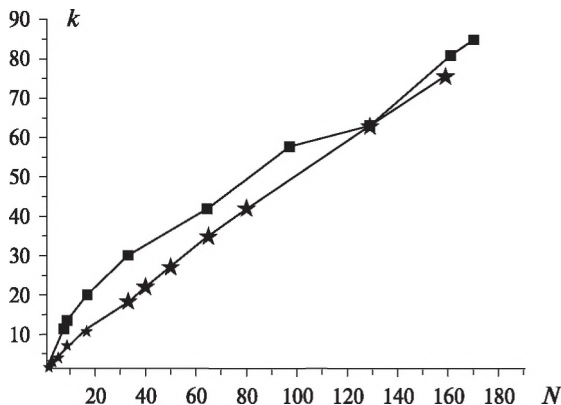


Рис. 1. Зависимость эффективности параллелизации $k = T(1) / T(N)$ от N , где $T(N)$ — время вычисления многомерного интеграла с заданной точностью (не зависящей от N) при запуске вычисления в режиме N -mpi-узлов. Звёздочки обозначают вычисления на суперкомпьютере «Ломоносов» Московского государственного университета имени М. В. Ломоносова (МГУ) [Воеводин, 2012], квадратики — вычисления на суперкомпьютере К-100 Института прикладной математики им. М. В. Келдыша Российской академии наук (ИПМ им. М. В. Келдыша РАН)

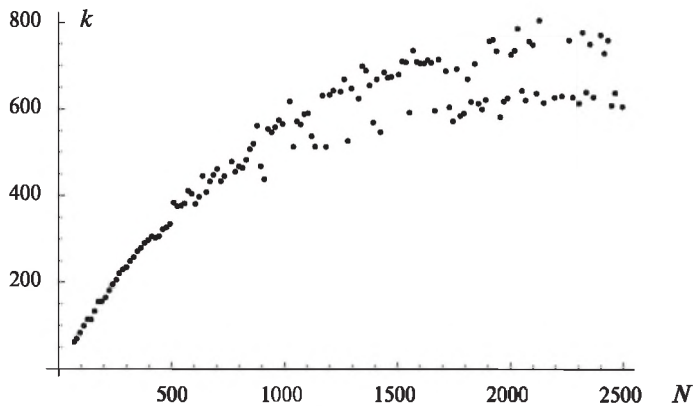


Рис. 2. То же, что на рис. 1. Вычисления были проведены на суперкомпьютере «Ломоносов» МГУ [Воеводин, 2012]. Число задействованных графических процессоров вычисляется как одна четвёртая числа mpi -узлов

Число mpi -узлов может превышать число графических процессоров суперкомпьютера в связи с тем, что каждый mpi -узел выполняется на одном ядре CPU, следовательно, каждый графический процессор обрабатывает вычислительные запросы от нескольких mpi -узлов. Это соответствует архитектуре GPU-раздела суперкомпьютера. Видно, что для N , меньших некоторого N_0 , зависимость, по существу, линейна. Это особенно хорошо видно на рис. 1. На рис. 2 для N , больших N_0 , эффективность параллелизации выходит на насыщение, а также присутствуют флуктуации. Тестирование показывает, что величина N_0 зависит от величины $T(1)$. При увеличении требуемой точности расчёта увеличивается время расчёта на одиночном узле $T(1)$, а также величина N_0 . Это объясняется тем, что при запуске малых задач на большом числе узлов доля расходов на синхронизацию вычислений становится большей по сравнению с расходами на вычисление интеграла. Это соответствует отмеченному выше насыщению кривой эффективности параллелизации при $N > N_0$. При увеличении $T(1)$ увеличиваются расходы на вычисление интеграла, поэтому для заданного N доля расходов на синхронизацию падает, следовательно, величина N_0 растёт, что отодвигает переход от линейной зависимости к насыщению вправо.

ЛИТЕРАТУРА

- [Воеводин, 2012] Воеводин Вл. В., Жуматий С. А., Соболев С. И., Антонов А. С., Брызгалов П. А., Никитенко Д. А., Стефанов К. С., Воеводин Вад. В. Практика суперкомпьютера «Ломоносов» // Открытые системы. 2012. № 7. С. 36–39.
- [Barash, Shchur, 2013] Barash L. Yu., Shchur L. N. RNGSSELIB: Program library for random number generation. More generators, parallel streams of random numbers and Fortran compatibility // Computer Physics Communications. 2013. V. 184. N. 10. P. 2367–2369.
- [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 // Computer Physics Communications. 2014. V. 185. P. 1343–1353.

[Hahn, 2007] *Hahn T.* Cuba — a Library for Multidimensional Numerical Integration // Computer Physics Communications. 2007. V. 176. N. 1. P. 712.

[Kanzaki, 2011] *Kanzaki J.* Monte Carlo integration on GPU // European Physical J. C. 2011. V. 71. P. 1559. URL: <http://arxiv.org/abs/1010.2107>.

APPLYING MULTI-GPU FOR MONTE-CARLO NUMERICAL INTEGRATION

L.Yu. Barash, L. N. Shchur

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

Multidimensional numerical integration with Monte-Carlo method has a number of important applications. Applying modern methods of calculations and massive parallelism can significantly improve performance of the integration. Monte-Carlo integration run about 40–100 times faster on the GPU than on the CPU. Also demonstrated is a linear increase in performance of the Monte Carlo calculations with increasing the number of GPUs.

Keywords: statistical methods, Monte Carlo, pseudorandom numbers, GPGPU, numerical integration.

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

Shchur Lev Nikolaevich — head of department, professor, shchur@chg.ru

ОРГАНИЗАЦИЯ ЭФФЕКТИВНОГО СЕТЕВОГО ДОСТУПА К РЕСУРСАМ ДАТА-ЦЕНТРА В РАСПРЕДЕЛЁННОЙ ИНФОКОММУНИКАЦИОННОЙ ИНФРАСТРУКТУРЕ

М. Р. Биктимиров¹, А. Б. Жижченко¹, А. П. Овсянников², А. А. Шер³, П. А. Климов³

¹ Вычислительный центр им. А. А. Дородницына Российской академии наук (ВЦ РАН)

² Межведомственный суперкомпьютерный центр Российской академии наук (МСЦ РАН)

³ Закрытое акционерное общество «Академинформ»

Проведён анализ информационного обмена информационно-вычислительных ресурсов, объединённых опорной телекоммуникационной сетью Российской академии наук в Московском регионе. На его основе определены требования для разработки телекоммуникационных решений для доступа к ресурсам дата-центра. Предложены архитектурные решения для организации телекоммуникационной компоненты дата-центра с целью обеспечения надёжной работы и доступа к ресурсам дата-центра из российских научных сетей. Создан прототип телекоммуникационного узла доступа дата-центра.

Работа выполнена в рамках Проекта «Исследование вопросов эффективного сетевого доступа к оборудованию дата-центра в распределённой инфокоммуникационной инфраструктуре» Программы фундаментальных исследований № 14 Президиума РАН «Проблемы создания распределённой информационно-вычислительной среды на основе развития GRID-технологий, облачных вычислений и современных телекоммуникационных сетей».

Ключевые слова: научные телекоммуникации, научные сети, инфокоммуникационная инфраструктура, дата-центры.

ВВЕДЕНИЕ

Современные тенденции развития информационно-коммуникационных технологий (Information and Communication Technologies) (ИКТ): облачные вычисления, обработка больших данных, суперкомпьютерные вычисления — требуют, с одной стороны, концентрации вычислительных ресурсов в дата-центрах, а с другой — роста внешнего информационного обмена дата-центров. Так, концепция облачных вычислений подразумевает аутсорсинг ИТ-сервисов и, как следствие, прямую их концентрацию в «облаках», т.е. специализированных дата-центрах. Обработка больших данных изначально (Data-Grid Большого адронного коллайдера) строилась на идее распределённого хранения данных и распределённых же вычислений, однако выполняемых на совокупности «вычислительных ферм», каждая из которых может рассматриваться как отдельный дата-центр. А суперкомпьютерный центр есть дата-центр, использующий суперкомпьютерные ресурсы. Облачные вычисления требуют огромного информационного обмена с клиентами,

Биктимиров Марат Рамилевич — директор по инновациям, кандидат технических наук, marat@ras.ru

Жижченко Алексей Борисович — заведующий отделом, доктор физико-математических наук, профессор, академик РАН, om@ras.ru

Овсянников Алексей Павлович — заведующий отделом, apo@jscs.ru

Шер Александр Алексеевич — технический директор, sher@academinform.ru

Климов Павел Алексеевич — системный администратор, pklimov@academinform.ru

а резервирование облачного сервиса в нескольких географически разнесённых дата-центрах обуславливает консолидированный информационный обмен между ними для репликации данных. Технология обработки больших данных предполагает интенсивный информационный обмен между «вычислительными фермами». Суперкомпьютерные центры должны принимать и отдавать большие объёмы исходных данных или результатов. Таким образом, для эффективной работы вышеупомянутых технологий телекоммуникационная компонента дата-центра имеет критическое значение. Целью настоящей работы является исследование телекоммуникационных решений для обеспечения надёжной работы и высокопроизводительного доступа к ресурсам дата-центра.

Первая часть настоящей работы анализирует трафик опорной телекоммуникационной сети РАН и определяет факторы, которые следует принимать во внимание при проектировании подключения дата-центра к телекоммуникационной сети. Во второй части определяются требования к интеграции дата-центров и телекоммуникационных сетей, рассматриваются схемы подключения дата-центров, обеспечивающие их надёжную и эффективную работу.

АНАЛИЗ ИНФОРМАЦИОННОГО ОБМЕНА ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ И ДАТА-ЦЕНТРОВ ОПОРНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ РОССИЙСКОЙ АКАДЕМИИ НАУК В МОСКОВСКОМ РЕГИОНЕ

Источником данных для настоящего исследования являлась опорная телекоммуникационная сеть Российской академии наук (ОТС РАН) в Московском регионе, объединяющая значительное число информационно-вычислительных ресурсов. В качестве базы для анализа была использована статистика, собираемая по протоколу NetFlow [Cisco Systems..., 2004] с основных маршрутизаторов ОТС РАН. NetFlow — сетевой протокол, разработанный для учёта сетевого трафика компанией Cisco Systems, он имеет три основных компонента:

- сенсор — датчик, осуществляющий фоновый процесс на устройстве (например, маршрутизаторе), собирающий информацию о проходящем трафике и направляющий её коллектору;
- коллектор — серверная программа, принимающая данные от сенсоров и сохраняющая их в файлах для дальнейшей обработки;
- система обработки и представления данных, которая читает файлы, созданные коллектором, и выдаёт отчёт в удобной для человека форме.

Анализ собранных данных выявил, что примерно 44 % трафика приходится на протокол http [Fielding et al., 1999], 8 % — на https [Rescorla, 2000], 5...6 % — на SSH [Ylonen, 2006] и столько же — на Remote Desktop Protocol [MS-RDPBCGR, 2014] (рис. 1).

Трафик http/https может быть порождён как посещениями пользователей сайтов, так и приложениями облачных дата-центров, причём различить их по собираемым Netflow данным невозможно. Поэтому основное внимание было посвящено протоколу ssh, который используется и для доступа к вычислительным ресурсам, как суперкомпьютерным, так и систем обработки боль-

ших данных, и для защищённого трансфера данных между файловыми системами или хранилищами данных таких ресурсов.

В частности, ssh является основным протоколом для доступа к суперкомпьютерам Межведомственного суперкомпьютерного центра РАН, одного из крупнейших суперкомпьютерных центров в стране.

В ходе исследований проведён анализ трафика головного узла суперкомпьютера МВС100К. Головной узел используется пользователями для постановки заданий в очередь и обмена данными с локальной файловой системой. Доступ пользователей осуществляется по протоколу SSH (tcp/22).

Была сделана выборка потоков по порту 22 протокола TCP, адресом назначения или отправления которых, являлся головной узел кластеров суперкомпьютера МВС100К в МСЦ РАН.

Потоком в протоколе NetFlow считается набор пакетов, передаваемый с определённого порта одного хоста на другой определённый порт другого хоста в течение некоторого интервала времени. Информация о потоках собиралась с маршрутизаторов в коллектор с интервалом 5 мин. Число потоков можно рассматривать в качестве меры времени, которое пользователи используют для интерактивной работы с суперкомпьютером, включающей передачу данных, постановку задач в очередь, наблюдение за прохождением задач и прочие действия.

Статистика потоков была отфильтрована для исключения сетевых атак, направленных на подбор пароля по протоколу SSH. В качестве критерия фильтрации использовалось условие превышения потоком некоторого минимального значения трафика, выбранного равным 100 байтам.

Произведено следующее разделение потоков и трафика по адресатам:

- внутри ОТС РАН,
- с российскими научными сетями,
- с зарубежными научными сетями,
- с коммерческими интернет сервис-провайдерами.

Для определения организационной принадлежности адресов использовались базы данных whois [Daigle, 2004].

На научно-образовательные сети приходится 40 % потоков, из них 18 % — на институты ОТС РАН, остальные 60 % — обычный Интернет (рис. 2), т. е. большую часть времени пользователи работают с суперкомпьютером из сетей коммерческих провайдеров (из домашних сетей).

Совершенно другая картина для трафика (рис. 3). На научно-образовательные сети — 79 % входящего и 92 % исходящего трафика, из которых 53 % входящего и 76 % исходящего трафика приходится на институты ОТС РАН. На обычный Интернет приходится 21 % входящего и 8 % исходящего трафика. Можно объяснить это тем, что пользователи предпочитают хранить данные и результаты работы на серверах своих институтов, в то время как собственно интерактивное взаимодействие с машиной производится из домашних сетей.

Следует также отметить, что объём исходящего трафика (88 %) почти в 7,5 раз превышает объём входящего (12 %), что может объясняться спецификой задач моделирования, преимущественно решаемых на суперкомпьютерах.

Относительно малым (1,5 % входящего и 2,5 % исходящего трафика) является обмен с зарубежными научно-образовательными сетями.

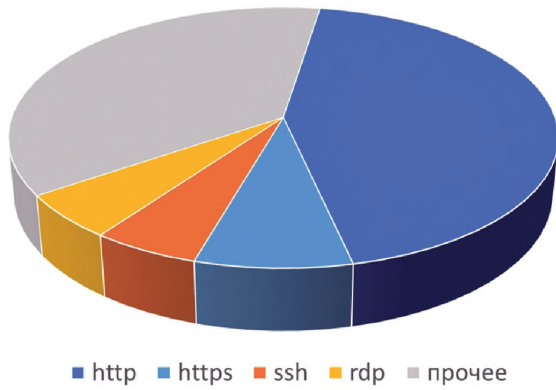


Рис. 1. Соотношение трафика ОТС РАН между протоколами

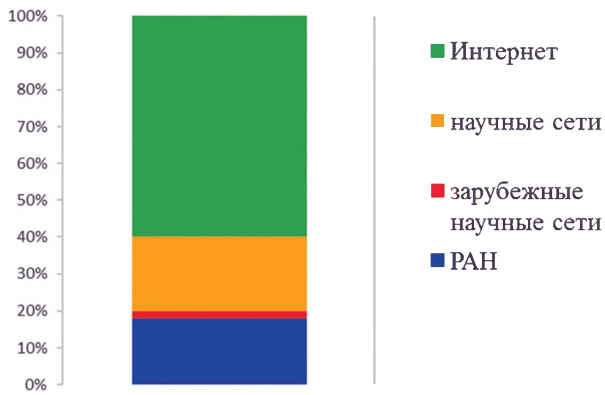


Рис. 2. Соотношение потоков суперкомпьютера между научными сетями и Интернет

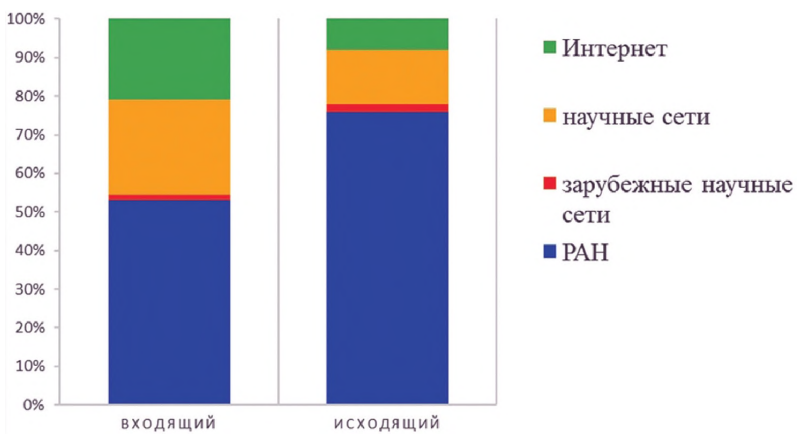


Рис. 3. Соотношение трафика суперкомпьютера между научными сетями и Интернет

Таким образом, высокопроизводительная вычислительная система в сети является значительным источником данных, причём наиболее интенсивный информационный обмен осуществляется с институтами пользователей суперкомпьютера. Кроме того, значительную долю трафика суперкомпьютера составляет обмен данными с Институтом прикладной математики (ИПМ) им. Келдыша (13 %) и Московским государственным университетом имени М. В. Ломоносова (МГУ) (8 %), которые также располагают мощными суперкомпьютерными ресурсами. Известно, что у этих систем значительное число общих пользователей, и они предпочитают использовать для своих задач все доступные ресурсы для решения своей задачи. Это требует дублирования результатов и исходных данных на локальных файловых системах разных суперкомпьютеров, что объясняет приведённые цифры информационного обмена между суперкомпьютерными ресурсами.

Средняя скорость передачи данных с головного узла суперкомпьютера составляет 57 Мбит/с, однако пиковая скорость достигает 749 Мбит/с. Таким образом, обмен данными с головным узлом обладает значительной неравномерностью и для обеспечения качества сервиса следует предусмотреть соответствующий запас пропускной способности.

Таким образом, при проектировании телекоммуникационных сетей для подключения суперкомпьютеров следует учитывать:

- интенсивность обмена данными с институтами, использующими суперкомпьютер, научно-образовательными сетями и другими высокопроизводительными вычислительными системами;
- тенденцию к преобладанию исходящего трафика над входящим;
- интенсивность и неравномерность информационного обмена с головным узлом суперкомпьютера.

Большинство требований могут быть удовлетворены интеграцией дата-центра с телекоммуникационными узлами научно-образовательной сети, в которой эти узлы являются концентраторами трафика и объединены высокоскоростными магистральными каналами.

ПОДКЛЮЧЕНИЕ ДАТА-ЦЕНТРА К НАУЧНОЙ СЕТИ

Высокопроизводительная научная телекоммуникационная сеть представляет собой совокупность узлов и точек подключения, объединённых оптическими линиями связи и общим управлением. Узел представляет собой систему активного телекоммуникационного оборудования, осуществляющего функции транзитной передачи данных. Точка подключения разграничивает ответственность между телекоммуникационной сетью и институтом, организацией или иными сетями и может также включать активное сетевое оборудование под общим управлением. Узлы могут также выполнять роль точек подключения институтов, а также внешних сетей.

К телекоммуникационным сетям предъявляются требования производительности, надёжности, безопасности, гибкости управления и масштабируемости.

На физическом уровне требование высокой производительности удовлетворяется использованием волоконно-оптических линий связи, потенциально обеспечивающих пропускную способность в десятки и сотни гигабит в секунду. В сетевой инфраструктуре РАН используется оборудование, обеспечивающее скорость передачи данных 10 Гбит/с по одному волокну. При использовании аппаратуры волнового уплотнения возможно использование нескольких длин волн, что позволяет достичь пропускной способности одного волокна в десятки и сотни гигабит в секунду. Волновое уплотнение эффективно на дальних магистральных линиях длиной в сотни километров, в городских условиях при развитой системе коллекторов телефонной канализации использование оптических кабелей с большим количеством волокон пока является наиболее дешёвым решением. При использовании агрегации линий связи на канальном уровне можно получить пропускную способность в десятки гигабит в секунду на одном канале. В настоящее время на рынке представлены решения интерфейсов 40 Гбит/с и появляются решения 100 Гбит/с (Alcatel Lucent, Cisco Systems), обеспечивающие пропускную способность 100 Гбит/с в одном трансивере. Эти решения более эффективны, чем агрега-

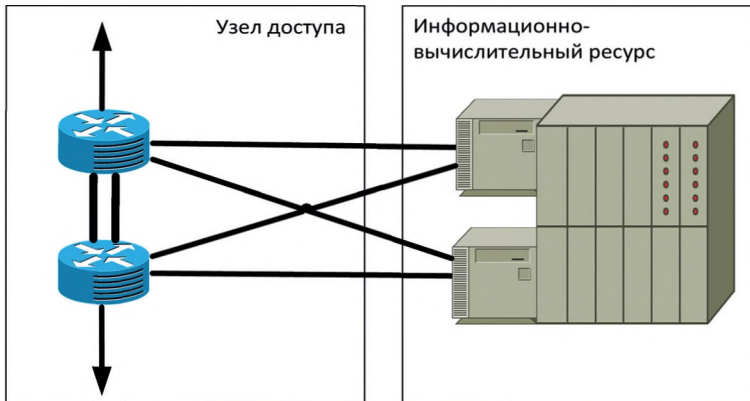


Рис. 4. Схема узла доступа с резервированием на основе маршрутизаторов

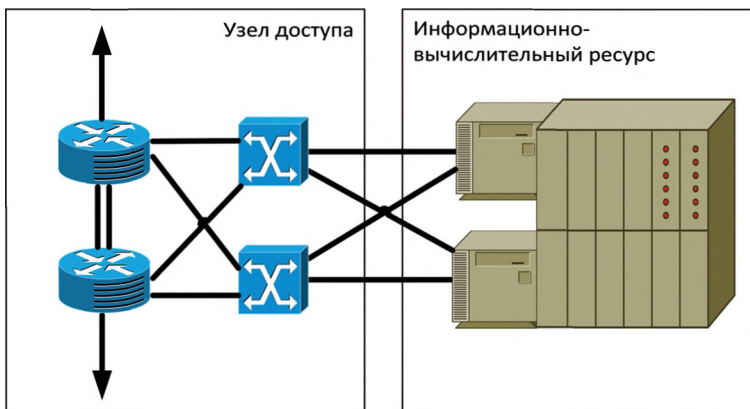


Рис. 5. Схема узла доступа с резервированием с использованием коммутаторов с протоколом MLAG

ция 10-гигабитных каналов, однако их стоимость весьма высока, хотя и сравнима с решениями на основе волнового уплотнения.

Для обеспечения надёжности необходимо использовать резервирование как линий связи, так и активного оборудования. Для магистральных линий связи целесообразно использование кольцевой топологии или циклических графов. В узлах подключения суперкомпьютеров обязательным является дублирование маршрутизаторов (коммутаторов) и подключение интерфейсного узла или узлов хранилищ данных двумя интерфейсами к каждому из них. Для обеспечения непрерывной передачи данных при отказе канала/порта/узла/маршрутизатора должны использоваться протоколы, обеспечивающие горячее резервирование или непрерывный информационный обмен (HSRP, VRRP [Li et al., 1998], динамическая маршрутизация).

Для обеспечения высокого уровня надёжности подключение ресурсов сверхвысокой производительности следует осуществлять к узлам сети (узлам доступа), отвечающим следующим требованиям:

- каждый узел доступа соединяется каналами связи, как минимум, с двумя другими узлами доступа;
- каналы связи между узлами доступа не имеют общих участков, т. е. не используют разные порты одного устройства (коммутаторов, оборудования волнового уплотнения и т. д.) и волокна в одной кабеле;
- в состав каждого узла доступа входят два маршрутизатора;
- каждый маршрутизатор связан хотя бы одним каналом с другим узлом доступа;
- маршрутизаторы узла доступа связаны друг с другом двумя каналами;
- каждый интерфейсный или головной узел дата-центра (или его важнейших информационно-вычислительных ресурсов) подключён к каждому из маршрутизаторов узла доступа, при этом на них и маршрутизаторах узлов доступа настроена динамическая маршрутизация.

На рис. 4 и 5 приведены возможные варианты схем узла доступа дата-центра.

Схема, приведённая на рис. 5, позволяет обеспечить более эффективное использование интерфейсов информационно-вычислительных ресурсов за счёт использования многокоммутаторного протокола агрегирования MLAG [Shefer, 2014].

ЗАКЛЮЧЕНИЕ

Проведён анализ статистики информационного обмена опорной телекоммуникационной сети Российской академии наук и её информационно-вычислительных ресурсов, определены требования для разработки телекоммуникационных решений для доступа к дата-центрам. Предложены архитектурные решения для высокопроизводительных научных телекоммуникационных сетей, обеспечивающие надёжную работу и доступ к высокопроизводительным дата-центрам и информационно-вычислительным ресурсам. Реализован прототип телекоммуникационного узла для дата-центра Вычислительного центра им. А. А. Дородницына Российской академии наук, подключённого к опорной телекоммуникационной сети Российской академии наук в Москве.

ЛИТЕРАТУРА

- [Cisco Systems..., 2004] Cisco Systems NetFlow Services Export Version 9 = Протокол экспорта NetFlow версии 9 / Ed. В. Claise; Пер. на рус. Н. Малых. IETF, 2004. 15 с. URL: <http://www.protocols.ru/WP/wp-content/uploads/2014/03/rfc3954.pdf>.
- [Daigle, 2004] *Daigle L.* WHOIS Protocol Specification. IETF, 2004. URL: <https://tools.ietf.org/html/draft-daigle-rfc954bis-01>.
- [Fielding et al., 1999] *Fielding R., Gettys T. J., Mogul J. C., Nielsen H. F., Masinter L., Leach P. J., Berners-Lee T.* Hypertext Transfer Protocol — HTTP/1.1, IETF, 1999. URL: <http://tools.ietf.org/pdf/rfc2616.pdf>.
- [Li et al., 1998] *Li T., Cole B., Morton P., Li D.* Cisco Hot Standby Router Protocol (HSRP). IETF, 1998. URL: <http://tools.ietf.org/html/rfc2281>.
- [MS-RDPBCGR, 2014] [MS-RDPBCGR]: Remote Desktop Protocol: Basic Connectivity and Graphics Remoting. Microsoft Corporation, 2014. URL: <http://msdn.microsoft.com/en-us/library/cc240445.aspx>.
- [Rescorla, 2000] *Rescorla E.* HTTP Over TLS. IETF, 2000. URL: <http://tools.ietf.org/html/rfc2818>.
- [Shefer, 2014] *Shefer A.* Open MLAG: The Road to the Open Ethernet Switch System. Mellanox, 2014. URL: <http://www.mellanox.com/blog/2014/07/open-mlag-the-road-to-the-open-ethernet-switch-system>.
- [Ylonen, 2006] *Ylonen T.* The Secure Shell (SSH) Transport Layer Protocol. IETF, 2006. URL: <http://tools.ietf.org/html/rfc4253>.

EFFICIENT NETWORK CONNECTIVITY FOR THE DATA-CENTER IN THE DISTRIBUTED ICT INFRASTRUCTURE

M. R. Biktimirov¹, A. B. Zhizhchenko¹, A. P. Ovsyannikov², A. A. Sher³, P. A. Klimov³

¹ Dorodnicyn Computing Centre of Russian Academy of Sciences (CC RAS)

² Joint Supercomputer Center of the Russian Academy of Sciences (JSCC RAS)

³ Academinform CJSC

The article describes the results of the analysis of the traffic of the informational and computational resources connected to the Moscow region research telecommunication network of the Russian Academy of Sciences. Based on the result of the analysis the demands for data-center connectivity were considered. Reliable and high-performance network architecture solutions were proposed. The prototype of the telecommunication node for data-center connectivity is developed in the Computational Center of the Russian Academy of Sciences.

Keywords: scientific telecommunications, scientific networks, communication infrastructure, data centers.

Biktimirov Marat Ramilevich — deputy director of innovation, PhD, marat@ras.ru

Zhizhchenko Aleksey Borisovich — head of department, doctor of physical and mathematical sciences, professor, academician RAN, om@ras.ru

Ovsyannikov Aleksey Pavlovich — head of department, apo@ras.ru

Sher Aleksandr Alekseevich — technical Director, sher@academinform.ru

Klimov Pavel Alekseevich — system administrator, pklimov@academinform.ru

ГИБРИДНЫЙ ПОДХОД К МОДЕЛИРОВАНИЮ СРЕДСТВ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ GRID/CLOUD-СИСТЕМ

В. А. Васенин¹, В. А. Роганов¹, А. А. Зензинов^{1,2}

¹ Научно-исследовательский институт механики Московского государственного университета имени М. В. Ломоносова (НИИ механики МГУ)

² Институт автоматизации проектирования Российской академии наук (ИАП РАН)

Эффективное и безопасное использование распределённых информационных систем на базе технологий облачных и грид-вычислений требует оценки и оптимизации средств обеспечения их безопасности. В этих целях представляется целесообразным проводить моделирование исследуемых систем и сценариев поведения при различных условиях их эксплуатации. В настоящее время широко распространена дискретно-событийная симуляция, которая не подходит для тестирования реальных программных средств. В этой связи авторы предлагают подход к проведению натурального, имитационного, виртуального и аналитического моделирования, а также их объединения в рамках гибридного режима моделирования. Такой подход позволяет исследовать поведение распределённых систем с разных сторон, учитывая особенности архитектуры, программного обеспечения (ПО), назначения систем и условий их эксплуатации. В статье также представлен прототип программной среды с поддержкой гибридного режима моделирования.

Ключевые слова: грид-вычисления, облачные вычисления, распределённые вычислительные системы, информационная безопасность, моделирование, натурное моделирование, имитационное моделирование, виртуальное моделирование, гибридное моделирование.

ВВЕДЕНИЕ

Разработка и поддержка программного обеспечения (ПО) для больших распределённых компьютерных систем требует большого внимания к его качеству и безопасности. К настоящему времени для обеспечения безопасности ресурсов грид- (Grid) и облачных (Cloud) систем создано большое число разнородных методов и средств. Вместе с тем, обеспечение надёжной защиты таких систем от деструктивных информационных воздействий является очень сложной задачей. Для её эффективного решения необходимо оценивать и корректировать уровень защищённости распределённых систем на разных архитектурных уровнях. Это обстоятельство побудило авторов начать работы по проектированию и реализации программной среды для гибридного моделирования Grid- и Cloud-систем. При помощи такой среды в рамках одной модели могут быть одновременно эффективно представлены и опробованы сразу несколько уровней системы при небольших затратах на сам процесс исследования.

Подход к гибриднему моделированию уже хорошо зарекомендовал себя и активно используется в научных исследованиях и в промышленных отраслях. В частности, такой подход можно встретить в космической отрасли, где он позволяет изучать сразу большое число различных режимов поведения

Васенин Валерий Александрович — заведующий лабораторией, доктор физико-математических наук, vassenin@msu.ru

Роганов Владимир Александрович — старший научный сотрудник, var@msu.ru

Зензинов Андрей Александрович — аспирант, andrey.zenzinov@gmail.com

сложной системы без реального запуска космического аппарата. В случае Grid- и Cloud-систем для построения гибридной модели были выбраны Натурное, Имитационное, Виртуальное и Аналитическое моделирование, что и дало, по первым буквам перечисленных слов, сводное название предлагаемой среде моделирования — НИВА.

Среда моделирования НИВА не предполагает высокой квалификации и навыков программирования её пользователей, обладает графическим интерфейсом и визуализацией процесса моделирования. Включая механизмы поддержки различных подходов, НИВА предоставляет удобный инструментарий для проведения исследований программных и архитектурных решений на направлении безопасности Grid/Cloud-систем.

1. СУЩЕСТВУЮЩИЕ ПОДХОДЫ

Сама по себе тема моделирования распределённых вычислительных систем не нова и имеет многолетнюю историю. Вскоре после появления распределённых вычислений были созданы средства для их изучения на основе дискретно-событийного моделирования, которое, по-видимому, является наиболее универсальным подходом и прекрасно зарекомендовало себя в области изучения сетей передачи данных.

На настоящее время наибольшую известность и широкое распространение получили специализированные программные средства SimGrid [Casanova et al., 2008], GridSim [Вууа, Mushred, 2002], CloudSim [Calheiros et al., 2011]. При помощи этих средств было проведено немало научных, а также практических исследований, призванных выявить и объяснить наличие «узких мест» в конкретных распределённых системах, например [Yaser Mansouri et al., 2013]. Среди российских средств аналогичного назначения следует отметить систему, разработанную в российском Институте системного программирования РАН [Грушин, Поспелов, 2010]. В этих средах распределённая система, как правило, представляется совокупностью входящих в неё объектов верхнего уровня, т. е. вычислительных узлов, межузловых соединений и так далее.

Для решения поставленной задачи не было необходимости создавать новый дискретно-событийный имитатор (симулятор). В качестве базового симулятора для гибридной системы моделирования было принято решение использовать SimGrid. Его роль и функциональные возможности в качестве базового симулятора будут описаны ниже. Отметим, что средства GridSim и CloudSim написаны на языке программирования Java, в то время как SimGrid реализован на языке C. С этих позиций SimGrid выглядит более подходящим решением для неинтерактивного проведения больших серий экспериментов. Кроме этого, SimGrid позволяет моделировать как Grid-системы, так и Cloud, в то время как GridSim ориентирована на моделирование именно Grid-систем, а CloudSim — Cloud-систем. Следует также отметить, что у SimGrid существует графический интерфейс для задания конфигураций распределённых систем, реализованный в виде плагина для среды разработки Eclipse.

С широким распространением поддержки виртуализации в аппаратуре связан и активно развивается подход на основе виртуального моделирования. В этом случае в исследуемой модели при помощи системы виртуальных ма-

шин, связанных между собой виртуальными сетями, имитируется аппаратный уровень.

На настоящее время существуют работы, представляющие результаты исследований вложенных облачных систем [Spillner et al., 2012]. При таком подходе происходит запуск виртуальных систем на виртуальных машинах, входящих в инфраструктуру облачной системы. Конечно, есть и другие подходы, например аналитическое и натурное моделирование. Если при моделировании ставится цель охватить большое число узлов (например, моделировать большие сегменты Интернет), и при этом интересуют макроуровень, то можно попытаться использовать аналитические методы из теории систем массового обслуживания, сети Петри и т.д. В отличие от них натурное моделирование поможет адекватно изучать даже такие тонкие явления в программном обеспечении как уязвимости типа Race Condition. Каждый из перечисленных подходов хорош на своём уровне, однако ни один из них не способен, во всей полноте, эффективно охватить весь спектр возможных взаимодействий, которые имеют место в больших распределённых системах. Подходы, обеспечивающие полноценное, комплексное моделирование Grid/Cloud-технологий с использованием методов и средств виртуализации, изучены пока слабо.

Принимая во внимание изложенные выше соображения, представляется целесообразным попытаться объединить все эти подходы. При этом следует иметь в виду, что всё перечисленное выше программное обеспечение для моделирования Grid- и Cloud-систем требует высокой квалификации программиста для определения и анализа интересующих активностей. Вместе с тем, большинство подобных средств не имеет даже механизмов для графического представления сетевой конфигурации, что на практике очень важно. По этой причине главными приоритетами при разработке среды НИВА были поддержка гибридного режима моделирования и удобство работы пользователя.

2. ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ

Среда моделирования Grid/Cloud-систем, именуемая НИВА, предоставляет пользователю средства для натурального, имитационного, виртуального и аналитического моделирования распределённых вычислительных систем как в интерактивном, так и в неинтерактивном режимах, с возможностью визуализации и специализированной постобработки полученных результатов. Сочетание различных моделей позволяет проводить комплексное исследование больших распределённых систем, а также погружать их в близкий к реальному контекст глобальной сети. В этом случае пользователю предоставляются возможности:

- при помощи натурального моделирования можно оценить реальную вычислительную эффективность счётных узлов и используемого интерконнекта на реальном оборудовании;
- используя технологии виртуального моделирования, провести необходимые эксперименты с дистрибутивом, оценить совместимость ПО, провести эффективное тестирование на известные уязвимости;
- применяя имитационную модель, смоделировать поведение значительного количества узлов Grid/Cloud, а также последствия возможных

атак и динамику распространения скрыто внедрённых программ типа вирусов;

- реализую аналитические описания, моделировать Grid-систему, как составную часть Интернета, задав при помощи формул вероятность внешних воздействий и интенсивность приходящих из глобальной сети запросов.

Цикл работы оператора состоит из следующей последовательности действий: выбора/редактирования конфигурации; выбора/редактирования сценария в соответствующих графических редакторах; загрузки конфигурации перед запуском сценария (их связывание происходит динамически по именам узлов); запуска процесса моделирования с визуализацией, либо без таковой, с возможным последующим детальным анализом полученных результатов. Процесс интерактивного моделирования может быть в любой момент приостановлен оператором (например, для детального изучения состояния моделируемой системы), выполняться по шагам (в терминах квантов виртуального времени), либо досрочно прекращён (например, для редактирования конфигурации). Существует также возможность запрограммировать автоматическую остановку процесса моделирования при возникновении определённых условий.

Результатом серии проведённых запусков в неинтерактивном режиме является отчёт. На основании отчёта может быть принято решение о пригодности или непригодности предлагаемых средств и стратегий защиты, а также могут быть выработаны рекомендации для дальнейшего повышения надёжности и защищённости разрабатываемых или реальных Grid/Cloud-систем.

Разумеется, любой среде моделирования присущ и ряд ограничений. Ограничения на размер моделируемых систем диктуются доступным объёмом оперативной памяти, а также быстродействием используемой аппаратной платформы. Однако преимущества гибридного подхода позволяют иногда обходить отмеченный недостаток, предоставляя возможность для аналитического моделирования отдельных подсистем на основе имеющихся о них эмпирических или статистических данных. Как следствие, гибридный подход позволяет охватить практически все случаи рассматриваемых Grid/Cloud-систем ценой потери точности.

3. ГИБРИДНОЕ МОДЕЛИРОВАНИЕ

Общим знаменателем между различными симуляторами в данном случае выступает имитационная модель, как наиболее универсальная. Такой подход избавляет от необходимости отвечать на сложный вопрос о прямом взаимодействии множества разнородных симуляторов. Он требует лишь минимальных усилий по созданию «имитационных представителей» для тех сущностей, которые должны взаимодействовать напрямую, но располагаются в разных симуляторах. В качестве базового симулятора для имитационного моделирования используется современная, активно развивающаяся среда разработки SimGrid с открытым исходным кодом. Она дополнена разработанными авторами средствами для интерактивной визуализации, графическим редакто-

ром сценариев и библиотекой шаблонов, реализующих типовые активности в Grid/Cloud-системах.

Гибридное моделирование позволяет задействовать одновременно несколько симуляторов. Организовать взаимодействие симуляторов можно по-разному. В рассматриваемой системе было принято решение сделать это на базе имитационной модели.

С позиции архитектуры системы взаимодействие симуляторов выглядит следующим образом.

1. Один из имитационных симуляторов (в нашем случае SimGrid) является главным, и все моделируемые сущности (за исключением моделируемых аналитически) имеют представление в виде узлов в имитационной модели. Вместе с тем, некоторые из сущностей также имеют представление в виде реальных процессов на виртуальных машинах или на реальных узлах, в случае виртуального или натурального моделирования, соответственно.
2. Если узел помечен как выполняющийся на внешнем симуляторе, то вместо типовой логики вызова действий SimGrid занимается лишь переназначением задач во внешний симулятор и наоборот. При этом внешний симулятор общается с SimGrid по несложному коммуникационному протоколу.
3. Внешние симуляторы получают полную конфигурацию системы с назначенными действиями. По этой причине они либо обращаются к другим узлам «натурным» образом (т.е. проводят реальные DDoS-атаки, пересылают реальный вредоносный код), либо пересылают сообщение в SimGrid, если целевой узел оказался лишь имитацией.
4. Поскольку в случае натурального и виртуального моделирования действия не имитируются, а выполняются реально (физически), поддержка этих действий требует реализации сопрягающей логики. Это означает, что при получении, например, сообщения с вредоносным кодом от имитационной модели в задачу натурального симулятора входит её преобразование в реальный вредоносный код. Как следствие, во многих случаях может оказаться проще проводить независимое моделирование в этих моделях и затем сравнивать полученные результаты для повышения уровня их достоверности.

4. КЛЮЧЕВЫЕ ПРОГРАММНЫЕ КОМПОНЕНТЫ

Структурно программная среда НИВА представляет собой связку ключевых компонентов для решения следующих задач:

- определение конфигурации и сценария использования Grid/Cloud-систем;
- преобразование (трансляция) конфигурации и сценария в низкоуровневую программу на языке C для запуска под управлением моделирующего ядра;
- запуск результирующей программы в составе одного или нескольких симуляторов;
- визуализация и протоколирование существенных происходящих событий;

- пост-обработка результатов и выполнение последующих команд оператора (интерактивный режим) или управляющего командного файла(неинтерактивный режим).

4.1. Графический редактор конфигураций Eclipse/SimGrid

Редактор конфигураций Grid/Cloud-систем входит в состав используемого базового пакета SimGrid и позволяет задавать и модифицировать типовые конфигурации Grid/Cloud-систем при работе в интегрированной среде разработки Eclipse. Среда Eclipse традиционно используется в качестве оболочки во многих средствах моделирования, а также включает в себя средства для коллективной работы разработчиков (имеет поддержку для работы с сетевыми репозиториями). На рис. 1 и 2 показаны примеры задания конфигураций систем, соблюдающих топологию «звезда» и «дерево», соответственно.

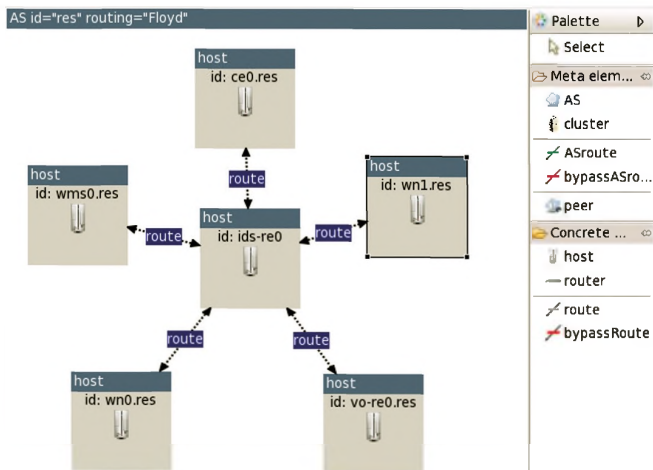


Рис. 1. Интерфейс редактора конфигураций Eclipse/SimGrid («Звезда»)

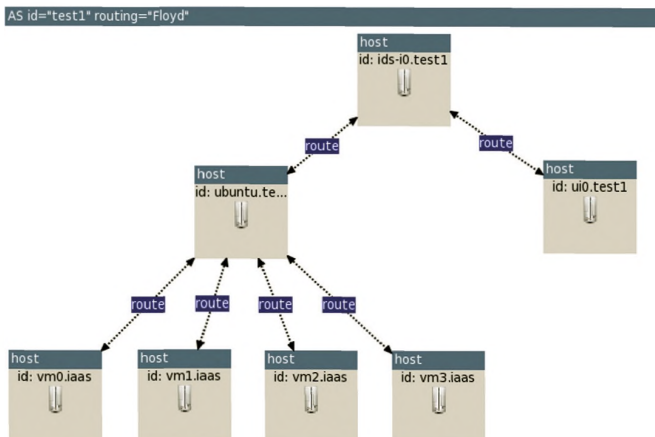


Рис. 2. Интерфейс редактора конфигураций Eclipse/SimGrid («Дерево»)

4.2. Графический редактор сценариев

Редактор сценариев (средство определения активностей в Grid/Cloud-среде) имеет простой графический интерфейс и позволяет определить сценарий использования распределённой системы как совокупность параллельно работающих распределённых активностей. Под активностью здесь и далее понимается совокупность взаимодействующих процессов в распределённой системе, сообща решающих какую-либо задачу. Примерами активностей могут служить взаимодействующие пользовательские процессы в ходе выполнения вычислительного задания, а также совокупность вредоносных действий злоумышленника при проведении спланированной атаки. На рис. 3 левая часть представляет собой список существующих активностей, правая — список шаблонов действий (составных частей активностей), редактирование же выбранной активности осуществляется в центральной части.

Активности состоят из действий. Некоторые действия являются подготовительными, т.е. выполняются перед началом моделирования. Это очень важный момент, существенно упрощающий работу оператора, так как вместо частого изменения конфигурации достаточно видоизменить (или деактивировать) активность, незначительно модифицирующую конфигурацию распределённой системы перед стартом.

Графический интерфейс редактора сценариев позволяет также осуществлять просмотр развёрнутой конфигурации (рис. 4), не дожидаясь запуска на моделирование.

Графический редактор сценариев реализован в виде программы на языке Java при помощи набора стандартных технологий — Swing, JavaBeans и с использованием хорошо зарекомендовавших себя библиотек. Для упрощения работы оператора многие операции по редактированию сценария выполняются перетаскиванием объектов при помощи мыши.

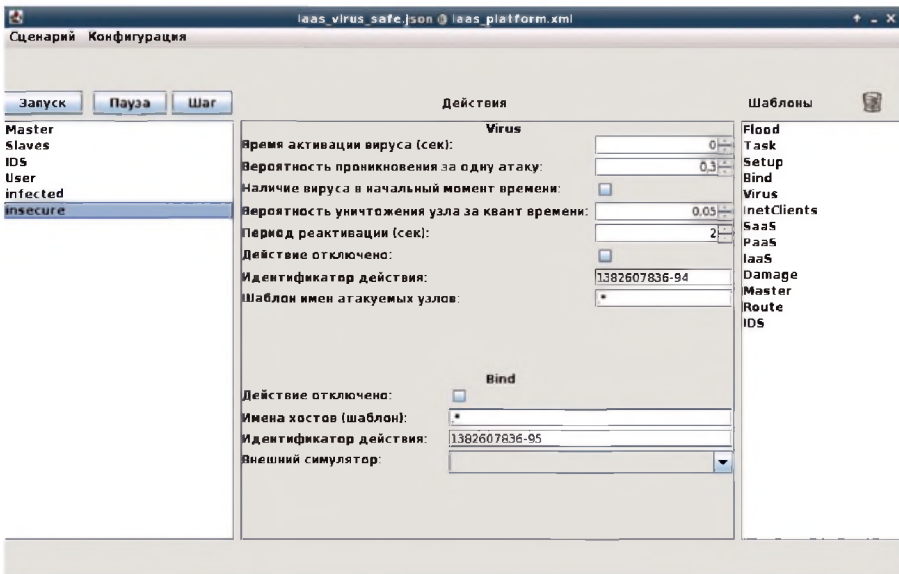


Рис. 3. Параметры активности в редакторе сценариев

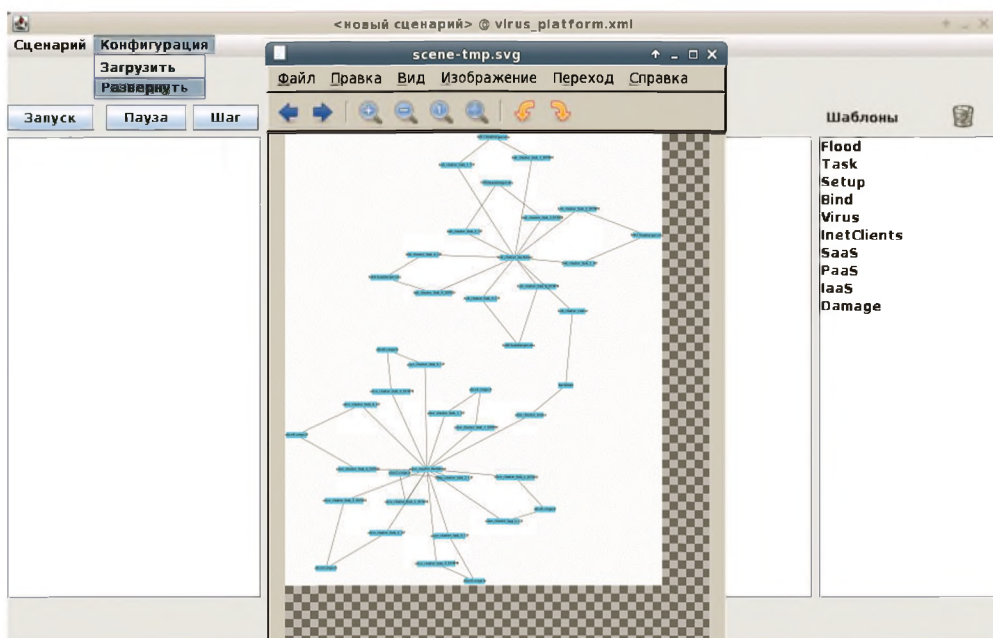


Рис. 4. Просмотр развёрнутой конфигурации

4.3. Сервер визуализации «Сцена»

Для интерактивной визуализации процесса моделирования разработан сервер «Сцена», выступающий как универсальное средство отображения состояния моделируемой системы. Используемый векторный формат SVG поддерживает масштабирование и позволяет отображать сложные совокупности графических элементов с возможностью их выборочной детализации (рис. 5).

Сервер «Сцена» поддерживает концепцию «визуальных переменных». К их числу относятся параметры, которые изменяются во время моделирования задействованными симуляторами и отображаются Сценой непосредственно при их изменениях (загруженность/доступность узлов, поражение узла вирусом и т. д.). Ещё одной особенностью Сцены является возможность её запуска на отдельном мониторе (проекторе) в качестве серверного процесса с возможностью параллельного обслуживания нескольких симуляторов.

Сервер визуализации «Сцена» реализован на языке Java и позволяет группе процессов отображать состояние моделируемой системы. При этом один из этих процессов заявляет себя как главный и устанавливает начальную картинку, передавая URI файла с графикой в формате SVG, в котором может присутствовать логика обработки запросов на отрисовку изменений на языке JavaScript. Каждый процесс, участвующий в ходе моделирования, будь то процесс-симулятор или реальный процесс в Grid/Cloud среде, может посылать запросы на отрисовку отдельных графических признаков (изменение цвета, надписей и так далее), а также получать информацию об изменениях на «Сцене». Последнее может быть также востребовано в ходе более продвинутой интерактивной визуализации.

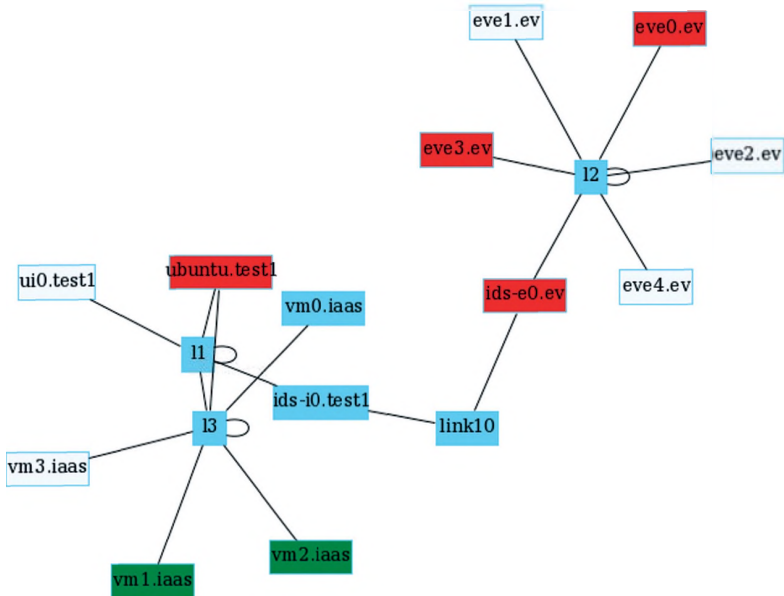


Рис. 5. Сервер визуализации «Сцена»

4.4. Система виртуального моделирования

Виртуальное моделирование основано на работе с виртуальными объектами, такими как виртуальные машины и виртуальные сети. Для управления этими объектами используется хорошо зарекомендовавшая себя кроссплатформенная библиотека libvirt.

Система развёртывания [Zenzinov, 2013] представляет собой комплекс программ, написанных на языке Python. Она реализует следующие функции:

- развёртывает виртуальные машины по заданной конфигурации;
- запускает виртуальные машины с сетевым сервисом для выполнения заданий;
- по завершении выполнения сценария система развёртывания производит останов виртуальных машин и удаление временных файлов.

Входными данными для создания набора виртуальных машин является общая конфигурация для виртуальных машин, конфигурация виртуальных сетей и набор шаблонов дисковых образов с предустановленными операционными системами и необходимым ПО. Конфигурационные файлы в данном случае имеют формат JSON и создаются редактором сценариев автоматически.

Используемый в системе сетевой сервис позволяет в режиме виртуального или гибридного моделирования выполнять на виртуальных машинах различные задания. Такой подход предоставляет возможность использовать в процессе моделирования реальное ПО, проводить более детализированные сценарии работы распределённой системы. Для использования какого-либо дополнительного ПО необходимо изменить исходные шаблоны дисковых образов.

4.5. Расширение функциональных возможностей среды моделирования

Поскольку технологии, используемые в создании распределённых компьютерных систем, наряду со средствами и методами их защиты, постоянно совершенствуются, то и среда моделирования Grid/Cloud-систем должна быть расширяемой. Программный интерфейс среды НИВА позволяет расширять её функциональные возможности с относительно небольшими временными и ресурсными затратами. Доступны следующие возможности расширения:

- создание новых шаблонов действий;
- расширение сервера визуализации «Сцена»;
- в случае гибридного моделирования — реализация совместной работы имитационного моделирования и реальных программ, запускаемых во время моделирования на реальных или виртуальных узлах.

5. ЭКСПЕРИМЕНТЫ

Для иллюстрации работы среды моделирования были проведены следующие эксперименты.

5.1. Гибридное моделирование штатного режима работы Grid-системы

В этом эксперименте Grid-система состоит из пользовательского узла, узла распределения заданий (мастер-узел) и вычислительных узлов. Типичный сценарий штатной работы такой системы следующий. Пользователь посылает в систему задачу, мастер-узел распределяет работу между вычислительными узлами. Последние в данном эксперименте были смоделированы виртуально, в то время как остальные моделировались имитационно.

В качестве типовой вычислительной задачи было выбрано численное вычисление интеграла $\int_0^{10} 3x^2 dx$ методом Симпсона. На мастер-узле вычислительное задание представлено в виде небольшой программы на языке Python. Мастер-узел рассылает его виртуальным узлам через сетевой сервис, а затем они выполняют задание.

При помощи утилиты `virt-manager` для управления виртуальными машинами можно следить за ходом выполнения задач, а именно — за получением задачи, запуском программы и выводом результата.

6.2. Вирусная атака на облачную систему типа «инфраструктура как сервис»

Второй эксперимент представляет собой моделирование вирусной атаки из внешней сети против облачной системы вида «инфраструктура как сервис» (IaaS). Экспериментальная система оснащена системой обнаружения вторжений (СОВ), а также узлом, отвечающим за облачную инфраструктуру. В настоящих облачных системах такой узел управляет виртуальной инфраструктурой. В рамках данного эксперимента виртуальные машины моделируются имитационным образом.

Эксперимент проводился в двух режимах: COB включена; COB выключена. В первом случае COB успешно блокирует распространение вируса по информационной системе. В другом случае можно наблюдать заражение системы в соответствии с заданными параметрами вероятности заражения. В ходе эксперимента некоторые заражённые узлы уничтожаются. Вероятность уничтожения также задаётся отдельным параметром.

Следует отметить, если узел, отвечающий за облачную инфраструктуру, будет уничтожен, то приписанные к нему виртуальные машины также станут недоступными и прекратят свою активность. Это поведение хорошо согласуется с поведением реальных облачных систем.

5.3. Неинтерактивный запуск серии экспериментов

Чтобы продемонстрировать возможность неинтерактивного запуска серии экспериментов были выбраны сценарии с распространением вируса и атакой вида «отказ в обслуживании» (DDoS) на Grid/Cloud-системы в шести различных конфигурациях. Эти конфигурации включают в себя три вида облачных систем: IaaS («инфраструктура как сервис»); PaaS («платформа как сервис»); SaaS («программное обеспечение как сервис»). Они отличаются между собой архитектурными элементами и способом распределения вычислительных заданий. Кроме облачных систем, рассматривались Grid-системы с тремя различными сетевыми топологиями: «кольцо»; «звезда»; «линия». Как и в предыдущем разделе, рассматривались два случая: COB включена; COB выключена. Суммарно это даёт 24 различные конфигурации.

По завершении выполнения экспериментов система выдаёт пользователю отчёт, который содержит информацию о произошедших в ходе эксперимента событиях, а также итоговые результаты: число успешно выполненных задач; число невыполненных задач; число «упавших» узлов; доля успешных заданий для каждой конфигурации системы.

В таблице представлен пример таких результатов в случае вирус-атаки при отключённой COB. Эти результаты полностью зависят от заданных параметров.

Пример результатов серии экспериментов с вирус-атакой

Конфигурация информационной системы					
IaaS	PaaS	SaaS	Ring	Star	Line
<i>Число выполненных заданий</i>					
23	33	51	15	47	5
<i>Число невыполненных заданий</i>					
10	5	0	13	1	18
<i>Доля успешных заданий, %</i>					
69,70	86,84	100	53,57	97,92	21,74
<i>Число «упавших» узлов</i>					
2	1	2	1	4	1

ЗАКЛЮЧЕНИЕ

В результате проделанной работы был разработан прототип программной среды для моделирования Grid/Cloud-систем с поддержкой нескольких режимов симуляции. В дополнение к широко используемому дискретно-событийному моделированию использование технологий виртуализации позволяет оценивать эффективность средств обеспечения безопасности. Такой подход позволяет принимать во внимание особенности реальных программ, которые могут содержать в себе уязвимости. В сочетании с остальными режимами моделирования это даёт новые уникальные возможности. Разрабатываемая система может использоваться как инструментарий для исследований таких задач как оценка безопасности программного обеспечения информационных систем, поиск оптимальных конфигураций для защиты систем от конкретных видов атак. Следует отметить, что среда НИВА может использоваться и в задачах, не связанных с тематикой информационной безопасности.

Кроме практической значимости исследования, связанные с разработкой представленной программной среды, имеют теоретическое значение. Возникает целый ряд вопросов, обусловленных гибридной природой моделирования распределённых компьютерных систем и систем безопасности. Предстоит ответить на вопросы, связанные с взаимодействием симуляторов, найти и изучить возможные эффекты, возникающие в процессе такого взаимодействия.

ЛИТЕРАТУРА

- [Грушин, Поспелов, 2010] *Грушин Д. А., Поспелов А. И.* Система моделирования Grid: реализация и возможности применения // Тр. Ин-та системного программирования РАН. 2010. Т. 18. С. 243–260.
- [Buyya, Murshed, 2002] *Buyya R., Murshed M.* Gridsim: A toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing // *Concurrency and Computation: Practice and Experience (CCPE)*. 2002. V. 14. P. 1175–1220.
- [Calheiros et al., 2011] *Calheiros R. N., Ranjan R., Beloglazov A., de Rose C. A. F., Buyya R.* Cloudsim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms // *Software: Practice and Experience*. 2011. V. 41. P. 0038–0644.
- [Casanova et al., 2008] *Casanova H., Legrand A., Quinson M.* SimGrid: a generic framework for large-scale distributed experiments // *Proc. 10th Intern. Conf. Computer Modeling and Simulation (UKSIM'08)*. 2008. P. 126–131.
- [Spillner et al., 2012] *Spillner J., Brito A., Brasileiro F., Schill A.* () Analysis of overhead and profitability in nested cloud environments // *Cloud Computing and Communications (LATINCLOUD)*. 2012. P. 13–18.
- [Yaser Mansouri et al., 2013] *Yaser Mansouri, Adel Nadjaran Toosi, Rajkumar Buyya.* Broker-ing algorithms for optimizing the availability and cost of cloud storage services // *Proc. 5th IEEE Intern. Conf. Cloud Computing Technology and Science*. 2013.
- [Zenzinov, 2013] *Zenzinov A.* Automated deployment of virtualization-based research models of distributed computer systems // *Proc. 7th Spring/Summer Young Researchers Colloquium on Software Engineering (SYRCoSE'2013)*. 2013. P. 128–132.

HYBRID APPROACH TO MODELING OF SECURITY TOOLS OF GRID/CLOUD-SYSTEMS

V. A. Vasenin¹, V. A. Roganov¹, A. A. Zenzinov^{1,2}

¹ Institute of Mechanics Lomonosov Moscow State University (IMEC MSU)

² Institute of Computer Aided Design of the Russian Academy of Sciences (ICAD RAS)

Efficient and safe use of distributed information systems based on Grid and Cloud-computing technologies requires an assessment and optimization of their security level. For this purpose it seems appropriate to model the studied systems and their behavior under different working conditions. In recent time discrete-event simulation is widely used, but it is not convenient to testing real software solutions. Therefore, the authors propose an approach to modeling distributed systems in full-scale, simulation, virtual and analytical sense, and its combination — hybrid modeling. Such approach allows to investigate the behavior of distributed systems from different perspectives, taking into account features of the architecture, software, purposes of these systems and operation conditions. The paper also describes the mechanisms of the developed modeling software.

Keywords: grid-computing, cloud-computing, distributed computer systems, information security, modeling, full-scale modeling, virtual modeling, simulation, hybrid modeling.

Vasenin Valery Aleksandrovich — head of laboratory, doctor of physical-mathematical Sciences, vasenin@msu.ru

Roganov Vladimir Aleksandrovich — senior research scientist, var@msu.ru

Zenzinov Andrei Aleksandrovich — postgraduate student, andrey.zenzinov@gmail.com.

ИНВАРИАНТНЫЕ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ АТТРАКТОРА И ДИФФУЗИОННЫЙ ХАОС В УРАВНЕНИИ ГИНЗБУРГА-ЛАНДАУ С ГАНТЕЛЕОБРАЗНОЙ ПРОСТРАНСТВЕННОЙ ОБЛАСТЬЮ

С. Д. Глызин

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

Рассматривается феномен многомодового диффузионного хаоса, одним из признаков которого является увеличение ляпуновской размерности аттрактора распределённых эволюционных динамических систем при уменьшении коэффициента диффузии. Для иллюстрации этого эффекта рассматривается краевая задача типа «реакция-диффузия» в области, состоящей из двух прямоугольных частей, связанных между собой перемычкой. Ширину перемычки примем бифуркационным параметром задачи, она меняется так, что мера области сохраняется. Изучены условия возникновения хаотических колебаний и построена зависимость инвариантных характеристик аттрактора задачи от ширины перемычки. Параметр диффузии при этом выбран так, что за счёт уменьшения ширины перемычки может появиться пространственно неоднородный хаотический аттрактор. Для полученного аттрактора вычисляются ляпуновские экспоненты и ляпуновская размерность. Показано, что величина размерности при уменьшении параметра растёт до некоторого предела, кроме того выяснилось, что увеличение размерности связано с усложнением распределения по пространственной переменной устойчивых режимов системы.

Исследование выполнено при поддержке Российского научного фонда (проект № 14-21-00158).

Ключевые слова: диффузионный хаос, аттрактор, ляпуновская размерность, уравнение Гинзбурга-Ландау, бифуркация.

ВВЕДЕНИЕ

Формирование пространственно неоднородных структур в однородных средах обычно связано с изучением модельных краевых задач, имеющих общее название уравнения «реакция-диффузия». Данный феномен был впервые описан математически Аланом Тьюрингом в работе [Turing, 1952], посвящённой явлению диффузионной неустойчивости пространственно однородных решений. Теорема Тьюринга-Пригожина [Nicolis, Prigogine, 1977] описывает ситуацию, когда при определённом изменении диффузионного параметра системы пространственно однородное решение теряет устойчивость, а в его окрестности появляются устойчивые пространственно неоднородные режимы.

Рассмотрим параболическую краевую задачу вида

$$\frac{\partial u}{\partial t} = \nu D \Delta u + F(u), \quad \left. \frac{\partial u}{\partial \mathbf{n}} \right|_{\partial \Omega} = 0, \quad (1)$$

Глызин Сергей Дмитриевич — доктор физико-математических наук, профессор, заведующий кафедрой компьютерных сетей ЯрГУ, ведущий научный сотрудник ОПСи Научного центра в Черноголовке РАН, glyzin@uniyar.ac.ru

где Δ — оператор Лапласа; $u \in \mathbb{R}^k$; $k \geq 2$; $D = \text{diag}\{d_1, \dots, d_k\}$; $d_j > 0$; $j = 1, \dots, k$; $\nu > 0$ — параметр, отвечающий за пропорциональное уменьшение коэффициентов диффузии; \mathbf{n} — внешняя нормаль к достаточно гладкой границе $\partial\Omega$ ограниченной области $\Omega \subset \mathbb{R}^m$; $m \geq 1$; $F(u)$ — гладкая вектор-функция. Краевую задачу (1) принято называть системой «реакция-диффузия», она служит математической моделью многих биофизических и экологических процессов (см., напр., [Ахромеева и др., 2007; Мищенко и др., 2005; Nicolis, Prigogine, 1977]).

Под термином «диффузионный хаос» будем понимать странный аттрактор краевой задачи (1), нетривиально зависящий от пространственной переменной. В настоящее время существуют две концепции диффузионного хаоса — маломодовый и многомодовый хаос. Первый из них может возникать в системе (1) при «средних» значениях параметра ν , а второй — при $\nu \rightarrow 0$.

Интерес к маломодовому хаосу инициирован известными работами Э. Лоренца [Lorenz, 1963], а также Д. Рюэля и Ф. Такенса [Ruelle, Takens, 1971], а затем Й. Курамото [Kuramoto, 1978], в которых был поставлен общий вопрос: можно ли связать стохастические режимы в распределённой системе, имеющей бесконечно много степеней свободы, с наличием странного аттрактора в системе небольшого числа обыкновенных дифференциальных уравнений. (Обычно эта система представляет собой каким-то образом упрощённую модель исходной системы.) В ряде случаев это действительно удаётся сделать. Для примера сошлёмся на известное уравнение Гинзбурга — Ландау, для которого сформулированный вопрос был решён в работах С. П. Курдюмова, Г. Г. Малинецкого и Т. С. Ахромеевой (см. обзор в книге [Ахромеева и др., 2007]) для некоторых типов краевых условий. Точнее говоря, в этих работах численными методами был обнаружен странный аттрактор в трёхмерной системе, получающейся из уравнения Гинзбурга — Ландау на отрезке с граничными условиями Неймана в результате двухмодовой галеркинской аппроксимации. Другой пример — анализ странных аттракторов простейших конечно-разностных аппроксимаций краевых задач вида (1) на отрезке (см. [Глызин, 1997]).

Многомодовый диффузионный хаос сначала был теоретически описан в статье [Колесов, 1988], посвящённой исследованию динамики нелинейных осцилляторов, слабо связанных через диффузию. Численный анализ данной модели проделан в работах автора [Глызин, 2009; 2013], из результатов которых вытекает важное следствие: если в системе (1) на отрезке при $\nu \rightarrow 0$ наблюдается диффузионный хаос, то его ляпуновская размерность неограниченно растёт. Отметим также численный анализ уравнения Гинзбурга-Ландау, приведённый в работе [Гапонов-Грехов, 1987] для периодических краевых условий.

Для уточнения понятия многомодового диффузионного хаоса воспользуемся формулировками из статьи [Глызин и др., 2010]. Будем предполагать, что отвечающая системе (1) точечная модель, т. е. система обыкновенных дифференциальных уравнений

$$\dot{u} = F(u), \quad (2)$$

имеет экспоненциально орбитально устойчивый цикл

$$u = u_0(t), \quad \frac{du_0}{dt} \neq 0 \quad (3)$$

некоторого периода $T_0 > 0$. Понятно, что такая ситуация является типичной. Учитывая, что цикл (3) сохраняется и в распределённой модели (1), для изучения его устойчивости линеаризуем уравнение из (1) на нём и применим к получившейся линейной краевой задаче метод Фурье по системе собственных функций оператора Лапласа. В результате приходим к системе

$$\dot{h} = [A_0(t) - zD]h, \quad (4)$$

где $A_0(t) = F'(u)|_{u=u_0(t)}$; параметр z принимает дискретные значения $\nu\lambda_k$, $k = 0, 1, \dots$; $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$ — занумерованные в порядке возрастания собственные значения оператора $-\Delta$ в области Ω с граничными условиями Неймана. Всюду ниже для удобства будем считать параметр z в (4) непрерывно меняющимся на полуоси $z \geq 0$. Обозначим, далее, через μ_s , $s = 1, \dots, k$, мультипликаторы системы (4) и положим

$$\alpha(z) = \max_{1 \leq s \leq k} \left\{ \frac{1}{T_0} \operatorname{Re} \ln \mu_s(z) \right\}. \quad (5)$$

Заметим, что $\alpha(0) = 0$, так как при $z = 0$ система (4) (в силу экспоненциальной устойчивости цикла (3) для точечной модели (2)) имеет простой единичный мультипликатор, которому отвечает решение Флоке $h = \dot{u}_0(t)$, а все остальные её мультипликаторы лежат в круге $\{\mu \in \mathbb{C} : |\mu| < 1\}$.

В дальнейшем нам потребуется следующее определение, введённое в статье [Глызин и др., 2010]. Следует отметить, что аналог данного определения впервые был введён в работе [Колесов, 1985].

Определение. Будем говорить, что параболическая краевая задача (1) имеет многомодовый диффузионный аттрактор, если выполняются следующие ограничения:

- а) точечная модель (2) имеет экспоненциально орбитально устойчивый цикл (3);
- б) найдутся такие $0 \leq z_1 < z_2$, что функция (5) строго положительна на интервале $z_1 < z < z_2$;
- в) при всех достаточно малых $\nu > 0$ динамическая система, порождаемая краевой задачей (1) в фазовом пространстве $C(\Omega; \mathbb{R}^k)$, $\bar{\Omega} = \Omega \cup \partial\Omega$, имеет хаотический аттрактор A_ν , ляпуновская размерность $d_L(A_\nu)$ которого стремится к $+\infty$ при $\nu \rightarrow 0$.

Как уже отмечалось выше, условие а) типично для краевых задач вида (1), возникающих в различных биофизических и экологических приложениях. Требование $\alpha(z) > 0$ при $z \in (z_1, z_2)$, фигурирующее в условии б), обеспечивает неустойчивость цикла (3) в рамках распределённой модели (1) при всех достаточно малых значениях параметра ν . Что же касается условия в), являющегося самым главным из всех трёх, то оно гарантирует реализуемость при $\nu \rightarrow 0$ феномена многомодового диффузионного хаоса. Учитывая, что термин «хаотический аттрактор» допускает различные интерпретации, будем для определённости придерживаться понятия хаоса, предложенного в работе [Колесов, Розов, 2009]. Отметим, что упомянутую в условии в) ляпуновскую размерность $d_L(A_\nu)$ будем вычислять через характеристические показатели аттрактора A_ν по известной формуле Каплана-Йорке [Frederickson et al., 1983].

Опираясь на приведённое определение, в статье [Глызин и др., 2010] сформулирована концепция диффузионного хаоса, заключающаяся в установлении самого факта существования динамических систем с многомодовым диффузионным аттрактором (см. [Глызин и др., 2010], а также [Глызин, 2013], где собраны несколько примеров многомодового диффузионного хаоса). Численное обоснование данного явления нетрудно получить для простых областей, например, как в упомянутой выше работе для отрезка. В случае плоских невыпуклых областей ситуация существенно сложнее. В данной работе рассматривается задача об устойчивых режимах динамической системы типа «реакция-диффузия» в ситуации, когда при сохранении меры области определения меняется её граница. В частности, подробно рассмотрен случай гантелеобразной области. Краевые задачи параболического типа в такой области были рассмотрены ранее в статьях Хосе Арриета [Arrieta, 1995a, b] и серии более поздних его статей с соавторами [Arrieta et al., 2006; 2009]. В этих публикациях, однако, ставилась несколько иная задача, заключавшаяся в исследовании ситуации, когда ширина перемычки, соединяющей области, стремится к нулю.

С целью более осмысленного подбора параметров задачи сделаем некоторые дополнительные предположения. В первую очередь, область определения пространственной переменной для простоты выберем в виде двух прямоугольных частей, связанных между собой перемычкой. Ширину перемычки примем бифуркационным параметром задачи, причём максимальное значение этого параметра соответствует исходной прямоугольной области. Далее, предполагается, что параметры точечной (не зависящая от пространственной переменной) системы находятся вблизи бифуркации Андронова-Хопфа и, тем самым, задача имеет пространственно однородный цикл. Следующее предположение состоит в том, что в задаче возможна диффузионная потеря устойчивости однородного цикла и параметр диффузии для указанной выше прямоугольной области выбран вблизи, но больше критического значения параметра диффузии.

Таким образом, основная решаемая в работе задача состоит в выяснении, при каком значении ширины перемычки пространственно однородный режим теряет устойчивость и какие при этом появляются устойчивые режимы.

В случае появления неупорядоченных колебаний численно определяется набор инвариантных числовых характеристик аттрактора (ляпуновских экспонент) с тем, чтобы по ним судить о сложности бифурцирующих пространственно неоднородных режимов.

Одним из основных результатов, полученных на этом пути, является тот факт, что увеличение ляпуновской размерности аттрактора тесно связано с усложнением пространственного распределения аттрактора. Иначе говоря, в более сложном случае имеется более изрезанный профиль распределения переменных по пространственной области, что позволяет и в данной ситуации говорить о многомодовом диффузионном хаосе.

ВЫБОР ПАРАМЕТРОВ ЗАДАЧИ НА ОСНОВЕ АСИМПТОТИЧЕСКИХ МЕТОДОВ

Рассмотрим краевую задачу (1) в ситуации, когда соответствующая ей точечная задача близка к бифуркации Андронова-Хопфа:

$$\dot{u} = \varepsilon \nu D \Delta u + (A_0 + \varepsilon A_1)u + F_2(u, u) + F_3(u, u, u) + F_*(u), \quad \left. \frac{\partial u}{\partial \bar{n}} \right|_{\partial \Omega} = 0, \quad (6)$$

как и выше, $u = u(t, x) : \mathbb{R}_+ \times \mathbb{R}^2 \rightarrow \mathbb{R}^k$, $0 < \varepsilon \ll 1$ — положительный малый параметр, $k \times k$ матрица A_0 имеет пару собственных чисел на мнимой оси $\pm i\omega$, причём все остальные собственные числа данной матрицы лежат в левой комплексной полуплоскости, $k \times k$ матрица A_1 такова, что собственные числа матрицы $A_0 + \varepsilon A_1$ переходят в правую комплексную полуплоскость при $\varepsilon > 0$. Вектор-функции векторного аргумента $F_2(u, u)$, $F_3(u, u, u)$ линейны по каждому аргументу, гладкая функция $F_*(u)$ имеет порядок малости по u выше третьего. Введём собственный вектор a , соответствующий собственному числу $i\omega$ так, что $(A_0 - i\omega E)a = 0$, в свою очередь собственный вектор сопряжённой задачи $(A_0^T + i\omega E)b = 0$ выберем при дополнительном условии $(a, b) = 1$, где $(*, *)$ — скалярное произведение в унитарном пространстве.

Хорошо известно, что система (6) с помощью замены

$$u(t, x, \varepsilon) = \sqrt{\varepsilon} (\xi(s, x) \exp(i\omega t) \cdot a + \bar{\xi}(s, x) \exp(-i\omega t) \cdot \bar{a}) + \varepsilon u_2(s, x) + \varepsilon^{3/2} u_3(s, x) + \dots, \quad (7)$$

где $s = \varepsilon t$, в результате приравнивания коэффициентов при степенях $\sqrt{\varepsilon}$ на третьем шаге приводится к укороченной квазинормальной форме (см. [Васильева и др., 1985]):

$$\frac{\partial \xi}{\partial s} = \nu(Da, b) \Delta \xi + (A_1 a, b) \xi + d |\xi|^2 \xi, \quad \left. \frac{\partial \xi}{\partial \bar{n}} \right|_{\partial \Omega} = 0, \quad (8)$$

где $d = d_0 + ic_0$ — первая ляпуновская величина, определяемая нелинейностью задачи. Нетрудно найти

$$d_0 + ic_0 = \left[\begin{array}{l} F_2(a, w_0) + F_2(w_0, a) + F_2(\bar{a}, w_1) + F_2(w_1, \bar{a}) + \\ + F_3(a, a, \bar{a}) + F_3(a, \bar{a}, a) + F_3(\bar{a}, a, a), b \end{array} \right], \quad (9)$$

где $w_0 = -A_0^{-1}(F_2(\bar{a}, a) + F_2(a, \bar{a}))$; $w_1 = (2i\omega - A_0)^{-1}F_2(a, a)$.

Основное предположение, принимаемое на этом этапе, состоит в том, что $d_0 < 0$. Учтём, кроме того, что собственные числа матрицы $A_0 + \varepsilon A_1$ переходят при $\varepsilon > 0$ в правую комплексную полуплоскость, что гарантирует положительность величины $\text{Re}(A_1 a, b)$. В этой ситуации, выполняя в системе (8) замены $\xi = \eta \sqrt{-\text{Re}(A_1 a, b)/d_0} \cdot \exp(\text{Im}(A_1 a, b)s)$, $\tau = \text{Re}(A_1 a, b)s$, получим уравнение Гинзбурга-Ландау вида

$$\frac{\partial \eta}{\partial \tau} = \nu(1 - ic_1) \Delta \eta + \eta - (1 + ic_2) |\eta|^2 \eta, \quad \left. \frac{\partial \eta}{\partial \bar{n}} \right|_{\partial \Omega} = 0, \quad (10)$$

где $\nu \text{Re}(Da, b)/\text{Re}(A_1 a, b)$ снова обозначено ν ; $c_1 = -\text{Im}(Da, b)/(\text{Re}(A_1 a, b) \times \text{Re}(Da, b))$; $c_2 = -c_0/d_0$. Относительно уравнения (10) известно, что его грубым режимам соответствуют, при достаточно малых ε , грубые режимы исходной системы (6) той же устойчивости [Васильева и др., 1985].

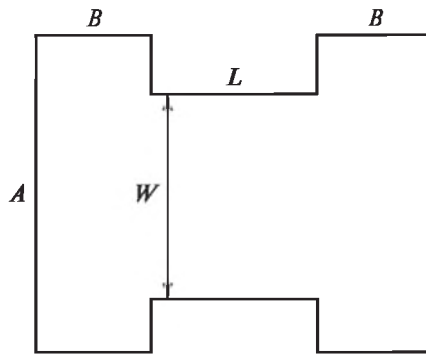


Рис. 1. Область определения Ω

Уравнение (10) имеет однородный цикл

$$\eta = \exp(-ic_2\tau + i\alpha), \quad \alpha = \text{const.} \quad (11)$$

В точечной модели (в отсутствие диффузионного слагаемого) цикл (11) является орбитально асимптотически устойчивым. Исследование устойчивости этого цикла для задачи (10) ($\nu \neq 0$) даёт следующее условие: цикл (11) устойчив (неустойчив) в рамках краевой задачи (10) при $\nu - \nu_* > 0$ (< 0), где

$$\nu_* = \frac{2(c_1c_2 - 1)}{\lambda_1(1 + c_1^2)} > 0, \quad (12)$$

Учитывая положительность параметра ν и положительность величины λ_1 , из соотношения (12) следует, что для появления диффузионной неустойчивости цикла необходимо выполнение условия

$$c_1c_2 > 1. \quad (13)$$

Предпринятый анализ позволяет уточнить постановку задачи для численного эксперимента. Рассмотрим уравнение Гинзбурга-Ландау (10) в области Ω (рис. 1). Будем считать, что недеформированная область Ω представляет собой квадрат со стороной единица. Зафиксируем параметры $c_1 = 0,5$, $c_2 = 8$ так, что выполнено неравенство (13).

В этой ситуации величина $\lambda_1 = -\pi^2$ и, тем самым, критическое значение ν , вычисленное по формуле (12), даёт $\nu_* = 6/(\pi^2 \cdot 1,25) \approx 0,4863$. Выберем $\nu = 2$, заведомо превосходящим порог ν_* , и рассмотрим процесс потери устойчивости однородного цикла при уменьшении ширины полосы, связывающей прямоугольники в Ω .

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО СЧЁТА

Предположим, что у краевой задачи (10) величина ν относительно мала, однако больше, чем критическое значение ν_* , при котором однородный цикл теряет устойчивость. Выберем область определения пространственной переменной Ω в виде двух соединённых перемычкой прямоугольников (см. рис. 1).

Будем считать, что прямоугольники имеют фиксированную длину A и ширину B , а перемычка имеет длину L и ширину W . Теперь будем уменьшать W и, соответственно, увеличивать L таким образом, чтобы произведение LW оставалось неизменным. Рассмотрим характер потери устойчивости однородного цикла при указанном изменении Ω и определим инвариантные числовые характеристики возникающих при этом аттракторов задачи.

Для численной оценки решений краевой задачи (10) будем использовать разностные аппроксимации по пространственным переменным. Построим разностную модель, аппроксимирующую (10), для этого отнесём начало координат на плоскости изменения пространственной переменной в левый нижний угол области Ω , затем разобьём область Ω на прямоугольники размером не более чем $\delta x \times \delta y$, где $\delta x = A/N$, $\delta y = (L + 2B)/M$. Далее, узлами сетки будем считать точки (x_i, y_j) с координатами $x_i = (i - 0,5)\delta x$, $y_j = (j - 0,5)\delta y$, $i = 1, \dots, N - 1$, $j = 1, \dots, M - 1$. Для точек $(x_i, y_j) \in \Omega$, вместо краевой задачи (10), заменяя оператор Лапласа его разностным аналогом, выпишем систему обыкновенных дифференциальных уравнений

$$\frac{\partial \eta_{i,j}}{\partial \tau} = v(1 - ic_1) \left[\frac{\eta_{i-1,j} + \eta_{i+1,j} - 2\eta_{i,j}}{(\delta x)^2} + \frac{\eta_{i,j-1} + \eta_{i,j+1} - 2\eta_{i,j}}{(\delta y)^2} \right] + \eta_{i,j} - (1 + ic_2) |\eta_{i,j}|^2 \eta_{i,j}, \quad (14)$$

где $(i, j) \in Ind$, а множество индексов Ind определяется следующим образом:

$$Ind = \left\{ (i, j) \mid i \in \{1, \dots, N - 1\}, j \in \{1, \dots, M - 1\}, (x_i, y_j) \in \Omega \right\}. \quad (15)$$

Граничные условия непроницаемости моделируются стандартным образом: на внешней границе области Ω считаем, что

$$\left. \begin{aligned} \eta_{0,j} &= \eta_{1,j}, & \eta_{N,j} &= \eta_{N-1,j}, & j &= 1, \dots, M - 1, \\ \eta_{i,0} &= \eta_{i,1} \left((x_i, y_1) \in \Omega \right), & \eta_{i,M} &= \eta_{i,M-1} \left((x_i, y_{M-1}) \in \Omega \right), & i &= 1, \dots, N - 1. \end{aligned} \right\} \quad (16)$$

На границе перемычки и на границах прямоугольников, к ней примыкающих, приблизим краевые условия по аналогии с (16), учитывая с помощью линейных аппроксимаций тот факт, что граница проходит не посередине расстояния между узловыми точками. Получившуюся систему обыкновенных дифференциальных уравнений будем решать методами Дормана-Принса 8(5) порядка (см. [Dormand, Prince, 1980]).

Поведение решений системы (14) с уменьшением ширины перемычки довольно быстро усложняется, поэтому возникает потребность в вычислении некоторых инвариантных характеристик аттрактора задачи, определяющих его хаотичность. Алгоритм определения спектра ляпуновских экспонент и ляпуновской размерности базируется на известной мультипликативной эргодической теореме Оселедца [Оседец, 1968], согласно которой для конечномерной динамической системы величина старшего ляпуновского показателя не зависит от начальной точки на аттракторе (точнее говоря, он одинаков для почти всех начальных значений по некоторой инвариантной мере). В нашем случае бесконечномерная задача (10) заменена конечномерной системой обыкновенных дифференциальных уравнений (14) и их ляпуновские экспоненты могут, вообще говоря, не совпадать. В связи с этим дополнительный интерес представляет вопрос об адекватности модели (14).

Для вычисления ляпуновских экспонент системы (14) будем пользоваться методом Беннетина [Benettin et al., 1976; Wolf et al., 1985] (точнее, его вариантом, изложенным в работе [Глызин и др., 2005]). Для оценки ляпуновской размерности будем пользоваться формулами Каплана-Йорке (см. [Frederickson et al., 1983]).

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО ЭКСПЕРИМЕНТА

Перейдём к описанию численного эксперимента, относящегося к одному из основных результатов статьи. Вычисления для краевой задачи (10) в области Ω выполнялись при следующих значениях параметров: $c_1 = 0,5$, $c_2 = 8$, $\nu = 2$, выбор которых обоснован выше. Размеры области Ω предполагались равными $A = 1$, $B = 1/3$, площадь перемычки при этом не меняется ($LW = 1/3$), а параметры разбиения принимались равными $N = 200$, $M = 200$. Введём бифуркационный параметр

$$\mu = \frac{W}{A}. \quad (17)$$

Система (14)–(16) решалась численно при различных μ в промежутке от 1 до 0,2. Нижняя граница $\mu = 0,2$ выбрана из соображений сохранения точности применяемой разностной схемы.

Выполненные расчёты показали, что при уменьшении μ на промежутке от 1 до 0,62 остаётся устойчивым пространственно однородный цикл. Иллюстрацией этому служит рис. 2а, на котором изображена проекция фазовой кривой устойчивого режима системы (14) на плоскость ($\text{Re } \eta_{5,5}$, $\text{Re } \eta_{150,55}$). Близость кривой к биссектрисе первого координатного угла как раз и означает, что колебания в различных точках области Ω синхронизированы. При $\mu \approx 0,62$ пространственно однородный цикл мягко теряет устойчивость, при этом рождается пространственно неоднородный цикл. Значение $\mu \approx 0,6$ является следующим критическим, поскольку при $\mu < 0,6$ колебательные режимы, устанавливающиеся в системе (14), становятся неупорядоченными.

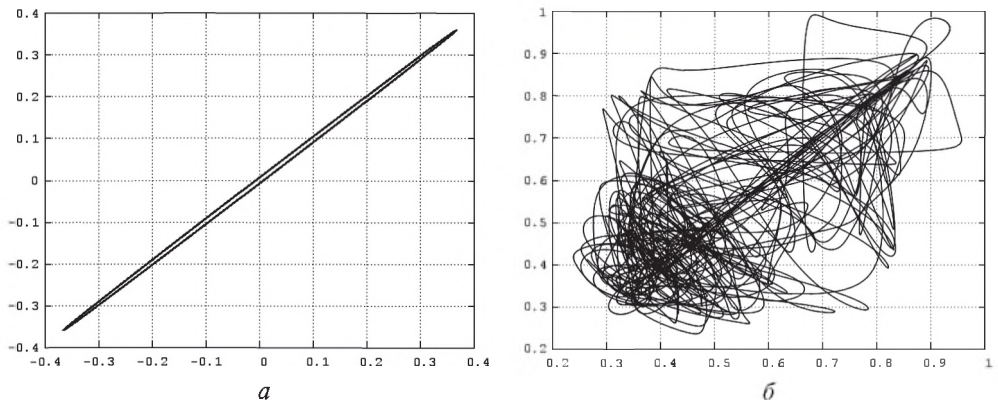


Рис. 2. Проекция фазового портрета системы (14): а — $\mu = 0,62$, пространственно однородный цикл; б — $\mu = 0,59$, пространственно неоднородный цикл

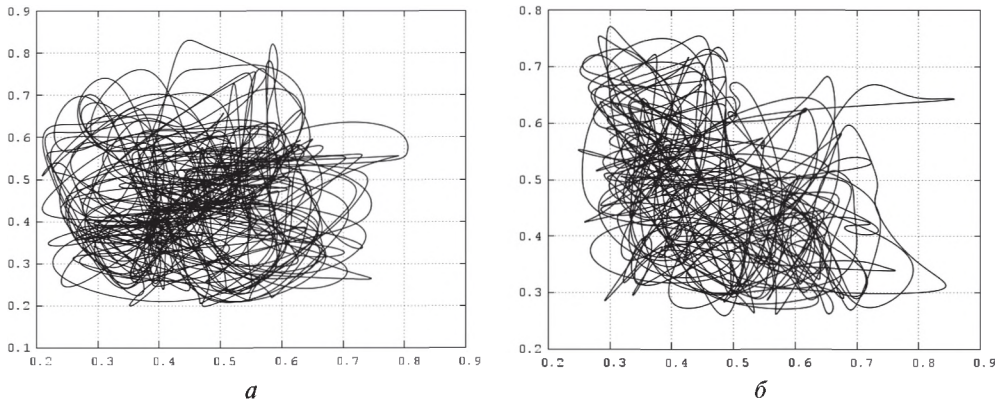


Рис. 3. Проекция фазовых портретов решений системы (14): а — $\mu = 0,45$; б $\mu = 0,22$

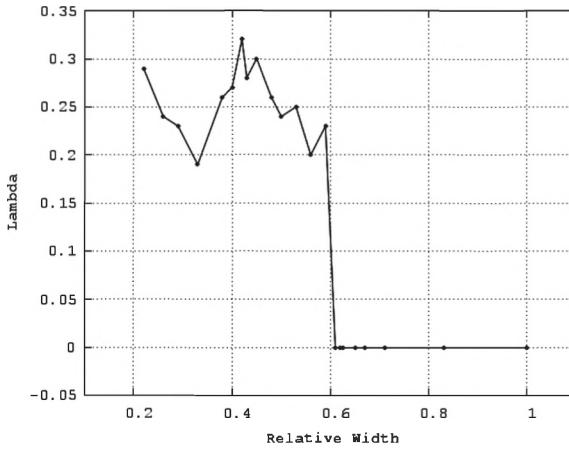


Рис. 4. График зависимости старшего ляпуновского показателя $\lambda_1(\mu)$ аттрактора системы (14) от относительной толщины перемычки μ

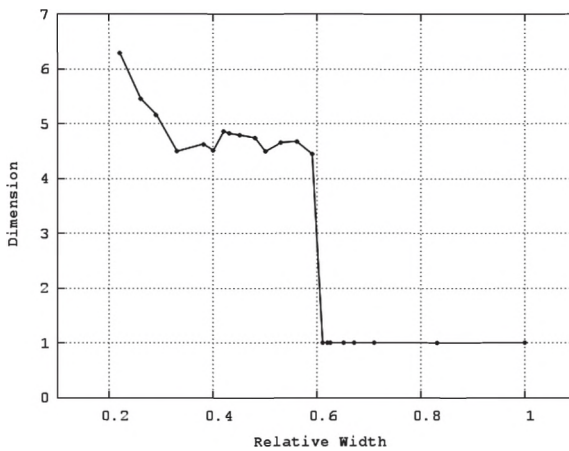


Рис. 5. График зависимости ляпуновской размерности $d_L(\mu)$ аттрактора системы (14) от относительной толщины перемычки μ

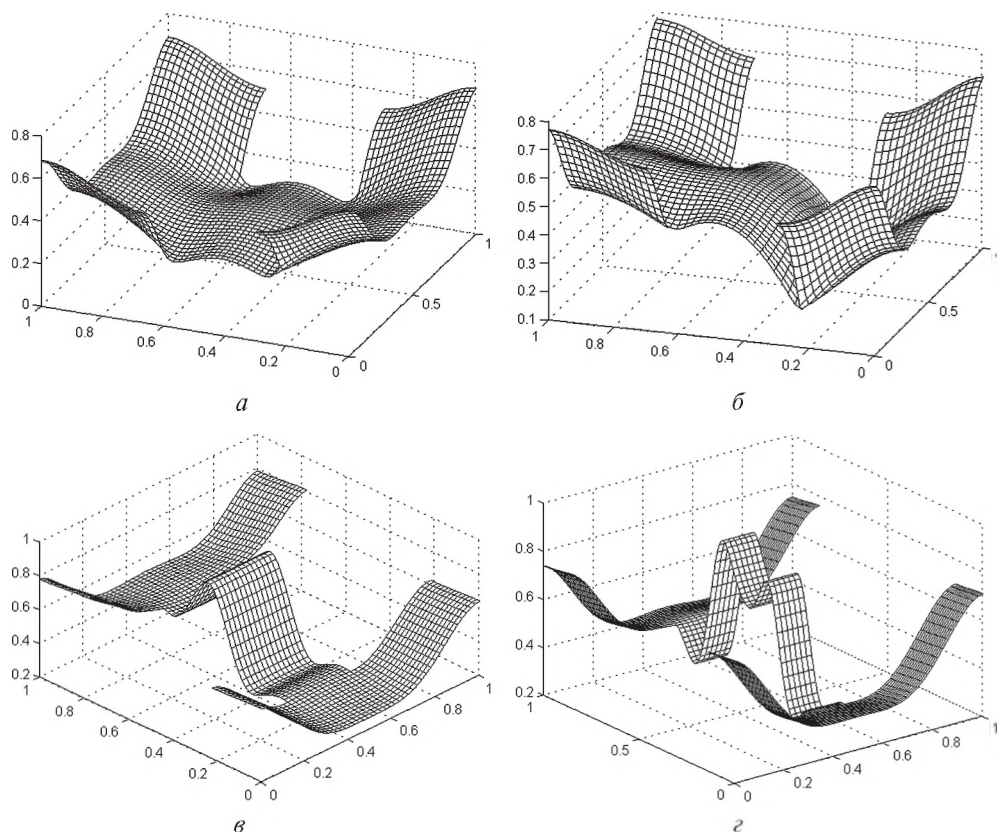


Рис. 6. Пространственное распределение амплитуды решения системы (14) при различных значениях относительной толщины перемычки и фиксированном t : a — $\mu = 0,61$; $б$ — $\mu = 0,59$; $в$ — $\mu = 0,38$; $г$ — $\mu = 0,22$

При значениях, близких к $\mu = 0,6$ (рис. 2б, построенный для $\mu = 0,59$), колебательный режим устроен так, что на протяжении длительных промежутков времени значения в различных точках области синхронизированы, это приводит к тому, что на указанном рисунке в окрестности биссектрисы первого координатного угла наблюдается уплотнение траекторий системы.

Дальнейшее уменьшение μ приводит к усложнению колебательных режимов (рис. 3). Степень их сложности находит своё отражение в построенных для аттрактора системы (14) графиках зависимости от μ старшего ляпуновского показателя λ_1 и ляпуновской размерности d_L . Нетрудно видеть (рис. 4, 5), что после возникновения при $\mu < 0,6$ у системы хаотического аттрактора величина $\lambda_1(\mu)$ колеблется в относительно узких пределах, а значение $d_L(\mu)$ после промежутка от $\mu \approx 0,6$ до $\mu \approx 0,32$, где оно почти не меняется, при $\mu < 0,32$ растёт.

Усложнение аттрактора системы можно проиллюстрировать и с помощью графиков пространственного распределения амплитуды решения системы (14) при различных значениях относительной толщины перемычки и фиксированном t . На рис. 6 в узлах решётки области Ω построены величины

$|\eta_{i,j}(t)|$, где $(i, j) \in \text{Ind}$. Нетрудно видеть, что режимы, получающиеся при малых μ (см. рис. 6в, г, построенные при $\mu = 0,33$ и $\mu = 0,22$ соответственно), оказываются существенно более изрезанными по сравнению с режимами, возникающими при значениях μ , близких к критическому значению, при котором теряет устойчивость пространственно однородный цикл (см. рис. 6а, б, построенные при $\mu = 0,61$ и $\mu = 0,59$ соответственно).

ЗАКЛЮЧЕНИЕ

Подводя итог, следует отметить, что исследуемое уравнение Гинзбурга-Ландау (10) может обладать многомодовым хаотическим аттрактором не только за счёт уменьшения коэффициента диффузии, но и в результате изменения геометрии области определения пространственной переменной. Один из основных результатов численного анализа состоит в том, что увеличение ляпуновской размерности аттрактора тесно связано с усложнением его пространственного распределения. Это означает, что в более сложном случае имеется более изрезанный профиль распределения переменных по пространственной области, что позволяет говорить о многомодовом диффузионном хаосе.

Задача о многомодовом диффузионном хаосе и его инвариантных характеристиках пока ещё далека от полного решения. Поставленную в статье задачу необходимо решать в более сложных областях и при меньших значениях параметра диффузии. Это потребует, с одной стороны, сеток с большим числом узлов для адекватного представления аттрактора со сложным пространственным распределением, и с другой, вычисления большего числа ляпуновских экспонент для оценки ляпуновской размерности. На этом пути необходимо использовать широкие возможности, открывающиеся в результате применения гибридных параллельных вычислительных систем.

ЛИТЕРАТУРА

- [Ахромеева и др., 2007] *Ахромеева Т. С., Курдюмов С. П., Малинецкий Г. Г., Самарский А. А.* Структуры и хаос в нелинейных средах. М.: Физматлит, 2007.
- [Васильева и др., 1986] *Васильева А. Б., Кащенко С. А., Колесов Ю. С., Розов Н. Х.* Бифуркация автоколебаний нелинейных параболических уравнений с малой диффузией // Математический сборник. 1986. Т. 130(172). № 4(8). С. 488–499.
- [Гапонов-Грехов, Рабинович, 1987] *Гапонов-Грехов А. В., Рабинович М. И.* Автоструктуры. Хаотическая динамика ансамблей // Нелинейные волны. Структуры и бифуркации. М.: Наука, 1987. С. 7–44.
- [Глызин и др., 2005] *Глызин Д. С., Глызин С. Д., Колесов А. Ю., Розов Н. Х.* Метод динамической перенормировки для нахождения максимального ляпуновского показателя хаотического аттрактора // Дифференциальные уравнения. 2005. Т. 41. № 2. С. 268–273.
- [Глызин, 1997] *Глызин С. Д.* Динамические свойства простейших конечноразностных аппроксимаций красовой задачи «реакция-диффузия» // Дифференциальные уравнения. 1997. Т. 33. № 6. С. 805–811.

- [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 // *Журн. вычислительной математики и математической физики.* 2010. Т. 50. № 5. С. 860–875.
- [Глызин, 2013] *Глызин С. Д.* Размерностные характеристики диффузионного хаоса // *Моделирование и анализ информационных систем.* 2013. Т. 20. № 1. С. 30–51.
- [Колесов, 1988] *Колесов А. Ю.* Описание фазовой неустойчивости системы гармонических осцилляторов, слабо связанных через диффузию // *Докл. АН СССР.* 1988. Т. 300. № 1. С. 831–835.
- [Колесов, Розов, 2009] *Колесов А. Ю., Розов Н. Х.* К вопросу об определении хаоса // *Успехи математических наук.* 2009. Т. 64. Вып. 4(388). С. 125–172.
- [Колесов, 1985] *Колесов Ю. С.* Проблема адекватности экологических уравнений. Ярославль, 1985. Деп. ВИНТИ. 1985. № 1901-85.
- [Мищенко и др., 2005] *Мищенко Е. Ф., Садовничий В. А., Колесов А. Ю., Розов Н. Х.* Автоволновые процессы в нелинейных средах с диффузией. М.: Физматлит, 2005.
- [Оседедец, 1968] *Оседедец В. И.* Мультипликативная эргодическая теорема. Характеристические показатели Ляпунова динамических систем // *Тр. Моск. мат. об-ва.* Т. 19. М.: Изд-во Моск. ун-та, 1968. С. 179–210.
- [Arrieta, 1995a] *Arrieta J. M.* Neumann eigenvalue problems on exterior perturbations of the domain // *J. Differential Equations.* 1995. V. 118. N. 1. P. 54–103.
- [Arrieta, 1995b] *Arrieta J. M.* Rates of Eigenvalues on a Dumbbell Domain. Simple Eigenvalue Case // *Trans. American Mathematical Society.* 1995. V. 347. N. 9. P. 3503–3531.
- [Arrieta et al., 2006] *Arrieta J. M., Carvalho A. N., Lozada-Cruz G.* Dynamics in dumbbell domains I. Continuity of the set of equilibria // *J. Differential Equations.* 2006. V. 231. N. 2. P. 551–597.
- [Arrieta et al., 2009] *Arrieta J. M., Carvalho A. N., Lozada-Cruz G.* Dynamics in dumbbell domains II. The limiting problem // *J. Differential Equations.* 2009. V. 247. N. 1. P. 174–202.
- [Benettin et al., 1976] *Benettin G., Galgani L., Strelcyn J. M.* Kolmogorov entropy and numerical experiments // *Physical Review.* 1976. V. A14. P. 2338–2345.
- [Dormand, Prince, 1980] *Dormand J. R., Prince P. J.* A Family of Embedded Runge — Kutta Formulae // *J. Computational and Applied Mathematics.* 1980. V. 6. P. 19–26.
- [Frederickson et al., 1983] *Frederickson P., Kaplan J., Yorke J.* The Lyapunov dimension of strange attractors // *J. Differential Equations.* 1983. V. 49. N. 2. P. 185–207.
- [Kuramoto, 1978] *Kuramoto Y.* Diffusion-Induced Chaos in Reaction Systems // *Progress of Theoretical Physics Supplement.* 1978. N. 64(1978). P. 346–367.
- [Lorenz, 1963] *Lorenz E. N.* Deterministic nonperiodic flow // *J. Atmospheric Sciences (JAS).* 1963. V. 20. P. 130–141.
- [Nicolis, Prigogine, 1977] *Nicolis G., Prigogine I.* Self-Organization in Non-Equilibrium Systems: From Dissipative Structures to Order through Fluctuations. Wiley, 1977. 512 p.
- [Ruelle, Takens, 1971] *Ruelle D., Takens F.* On the nature of turbulence // *Communications in Mathematical Physics.* 1971. V. 20. P. 167–192.
- [Turing, 1952] *Turing A. M.* The Chemical Basis of Morphogenesis // *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.* 1952. V. 237. N. 641. P. 37–72.
- [Wolf et al., 1985] *Wolf A., Swift J. B., Swinney H. L., Vastano J. A.* Determining Lyapunov exponents from a time series // *Physica D.* 1985. V. D16. P. 285–317.

INVARIANT DIMENSIONAL CHARACTERISTICS OF ATTRACTOR AND DIFFUSION CHAOS IN GINZBURG-LANDAU EQUATION WITH THE DUMBBELL DOMAIN

S. Glyzin

P. G. Demidov Yaroslavl State University (P. G. Demidov YSU)
Science Center in Chernogolovka, Russian Academy of Sciences

The phenomenon of multimode diffusion chaos is considered. For a number of examples it is shown, that the Lyapunov dimension of the attractor of a distributed dynamical system increases as the diffusion coefficient tends to 0. Then we consider a boundary problem of reaction-diffusion type in the domain consisting of two rectangular areas connected by a bridge. The bridge width is a bifurcation parameter of the problem and is changed in such way that the measure of the domain is preserved. The conditions on chaotic oscillations emergence were studied and the dependence of invariant characteristics of the attractor on the bridge width was constructed. By decreasing the bridge width the homogeneous cycle loses stability and then the spatially inhomogeneous chaotic attractor emerges. For the obtained attractor we compute Lyapunov exponents and Lyapunov dimension and notice that the dimension grows as the parameter decreases but is bounded. We show that the dimension growth is connected with the growing complexity of stable solutions distribution with respect to the space variable.

Keywords: diffusion chaos, attractor, Lyapunov dimension, Ginzburg-Landau equation, bifurcation

Glyzin Sergey — doctor of physical and mathematical sciences, professor, YSU, Leading Researcher at Scientific Center in Chernogolovka RAS . glyzin@uniyar.ac.ru

РЕАЛИЗАЦИЯ ГЕНЕРАТОРА ВЫЧИСЛИТЕЛЬНЫХ ЯДЕР OPENCL В БИБЛИОТЕКЕ VEXCL

Д. Е. Демидов

Казанский филиал Межведомственного суперкомпьютерного центра
Российской академии наук (КазФ МСЦ РАН)

Представлена методика, используемая в библиотеке с открытым исходным кодом VexCL для автоматической генерации вычислительных ядер OpenCL из выражений языка программирования C++. Данная методика позволяет значительно упростить написание программ, использующих графические ускорители для повышения эффективности научных расчётов. Показано преимущество такого подхода по сравнению с использованием библиотек с фиксированным программным интерфейсом.

Работа выполнена при поддержке РФФИ (проекты № 12-07-0007 и 12-01-00333а), а также за счёт средств субсидии, выделенной в рамках государственной поддержки Казанского федерального университета в целях повышения его конкурентоспособности среди ведущих мировых научно-образовательных центров.

Ключевые слова: C++, OpenCL, GPGPU, VexCL.

ВВЕДЕНИЕ

Применение видеокарт для научных расчётов в последнее время приобретает всё большую популярность. Основными программными платформами здесь являются NVIDIA CUDA и OpenCL. CUDA — это проприетарная программно-аппаратная архитектура, разрабатываемая компанией NVIDIA и поддерживающая вычисления общего назначения на видеокартах NVIDIA. OpenCL — это открытый стандарт для разработки кроссплатформенных приложений, использующих современные многоядерные процессоры, видеокарты и аппаратные ускорители. Обе платформы имеют как достоинства, так и недостатки. CUDA предлагает более полный набор библиотек для научных расчётов, но работает только с аппаратным обеспечением NVIDIA. OpenCL поддерживает гораздо более широкий набор оборудования, но низкоуровневый программный интерфейс OpenCL требует больших усилий со стороны разработчика. Ещё одно важное отличие CUDA от OpenCL заключается в принципе работы с вычислительными ядрами (функциями, исполняющимися на вычислительных устройствах; *англ.* computational kernels). Платформа компании NVIDIA компилирует вычислительные ядра в код PTX во время компиляции основной программы. PTX — это псевдоассемблер, который во время выполнения преобразуется в машинный код для конкретного устройства, на котором выполняется вычислительное ядро. Так как PTX — это низкоуровневый язык, то окончательный этап компиляции в машинный код занимает незначительное время. В стандарте OpenCL вычислительные ядра компилируются во время выполнения программы из высокоуровневого C-подобного языка. Это ведёт к увеличению времени инициализации программ, использующих OpenCL, и обычно считается недостатком OpenCL. Однако данная особенность позволяет генерировать во время выполнения программы вычислительные ядра, оптимизированные под конкретную вы-

числительную задачу. Последняя особенность активно используется в библиотеке с открытым исходным кодом VexCL (<https://github.com/ddemidov/vexcl>), разработанной для упрощения применения технологии OpenCL в вычислительных программах C++.

В данной работе представлена методика, используемая в VexCL для автоматической генерации вычислительных ядер OpenCL из выражений языка C++. На примере параметрического исследования системы Лоренца показана эффективность данного подхода по сравнению с библиотеками NVIDIA CUBLAS и Thrust [Bell, Hoberock, 2011; <http://thrust.github.com>].

БИБЛИОТЕКА VEXCL

VexCL — это библиотека с открытым исходным кодом, разработанная в КазФ МСЦ РАН и предназначенная для упрощения разработки C++ приложений, использующих технологию OpenCL. Библиотека предоставляет удобную и интуитивную нотацию для записи векторных выражений и позволяет использовать вычислительные мощности современных видеокарт и арифметических ускорителей. В работах [Ahnert et al., 2014; Demidov et al., 2013] представлено сравнение удобства использования и производительности VexCL с существующими аналогами.

Базовым классом библиотеки VexCL является контейнер `vex::vector<T>`, инкапсулирующий массив элементов типа T в памяти вычислительного устройства. Экземпляры данного класса можно комбинировать с помощью широкого набора выражений языка C++, таких как арифметические и логические операторы, встроенные и определённые пользователем функции. Также предоставляются операции умножения разреженной матрицы на вектор, генерации случайных чисел, сортировки и префиксного сканирования массивов, быстрого преобразования Фурье и др. Каждое векторное выражение приводит к автоматической генерации, компиляции и исполнению вычислительного ядра OpenCL. Так, выражение

```
x = 2 * y - sin(z);
```

где `x`, `y` и `z` — экземпляры класса `vex::vector<double>`, соответствует следующему вычислительному ядру:

```
kernel void vexcl_vector_kernel
(
    ulong n,
    global double * prm_1,           // x
    int prm_2,                       // 2
    global double * prm_3,          // y
    global double * prm_4           // z
)
{
    for(ulong idx=get_global_id(0); idx<n; idx+=get_global_size(0))
    {
        prm_1[idx] = ( ( prm_2 * prm_3[idx] ) - sin( prm_4[idx] ) );
    }
}
```

При реализации данной функциональности была использована техника шаблонов выражений C++ [Veldhuizen, 1995], описанная ниже.

ШАБЛОНЫ ВЫРАЖЕНИЙ

Техника шаблонов выражений была впервые реализована в библиотеке классов C++ для научных расчётов Blitz++. Использование этой техники позволило получить производительность, сравнимую с Фортраном 77/90 [Veldhuizen, 1995]. Для иллюстрации основной идеи рассмотрим следующий пример. Пусть требуется реализовать небольшой предметно-ориентированный язык, упрощающий выполнение арифметических действий над векторами. Так, выражение

```
x = y + z;
```

где x , y и z — экземпляры класса `vector`, должно быть эквивалентно и равно по производительности следующему циклу:

```
for(int i = 0; i < n; ++i)
    x[i] = y[i] + z[i];
```

На первый взгляд, очевидное решение заключается в перегрузке оператора сложения:

```
vector operator+(const vector &a, const vector &b) {
    vector tmp( a.size() );
    for(int i = 0; i < a.size(); ++i)
        tmp[i] = a[i] + b[i];
    return tmp;
}
```

К сожалению, такой подход, несмотря на свою простоту, имеет существенные недостатки: требуется выделение динамического временного массива и выполнение дополнительных операций обращения к памяти при записи и чтении промежуточных результатов. Так, при выполнении операции $a = x + y + z$ потребуется два временных вектора и вдвое больше операций обращения к памяти, чем необходимо. Это приводит к значительному снижению эффективности программы.

Альтернативная идея заключается в том, чтобы отложить фактический расчёт результата до операции присваивания. Для этого, во-первых, объявим структуру `vsum`, которая будет хранить ссылки на два вектора — левый и правый аргументы оператора сложения, а во-вторых, перегрузим оператор сложения двух векторов так, чтобы он возвращал структуру `vsum`:

```
struct vsum {
    const vector &lhs;
    const vector &rhs;

    double operator[](int i) const { return lhs[i] + rhs[i]; }
};
```

```
vsum operator+(const vector &lhs, const vector &rhs) {
    return vsum{lhs, rhs};
}
```

Определённый таким образом оператор сложения чрезвычайно легковесен. Фактическая же работа по сложению двух векторов выполняется в операторе присваивания вектора:

```
const vector& vector::operator=(const vsum &s) {
    for(int i = 0; i < size; ++i)
        data[i] = s[i];
    return *this;
}
```

Здесь используется оператор индексации [], определённый в структуре `vsum`. Как видно, отпала необходимость создания временного вектора и дополнительных операций обращения к памяти. Однако всё ещё есть ограничения по операции сложения и нельзя складывать более двух векторов.

Для того чтобы обобщить решение на случай произвольной арифметической операции, создадим для каждой из таких операций вспомогательную структуру. Так, операция сложения будет соответствовать структуре `plus`:

```
struct plus {
    static double apply(double a, double b) { return a + b; }
};
```

Статический метод `apply` здесь возвращает результат сложения двух своих аргументов. Аналогичным образом определяются остальные арифметические операции. Далее, переименуем структуру `vsum` в `binary_op`. Эта структура теперь является шаблоном; аргументы шаблона — это типы левого и правого операндов и операция, которую необходимо применить к операндам. Оператор сложения теперь принимает аргументы произвольных типов и возвращает структуру `binary_op` с соответствующими параметрами шаблона:

```
template <class LHS, class OP, class RHS>
struct binary_op {
    const LHS &lhs;
    const RHS &rhs;

    double operator[](int i) const {
        return OP::apply(lhs[i], rhs[i]);
    }
};

template <class LHS, class RHS>
binary_op<LHS, plus, RHS> operator+(const LHS &lhs, const RHS
&rhs)
{
    return binary_op<LHS, plus, RHS>{lhs, rhs};
}
```

```
for (size_t i = 0; i < n; ++i)
```

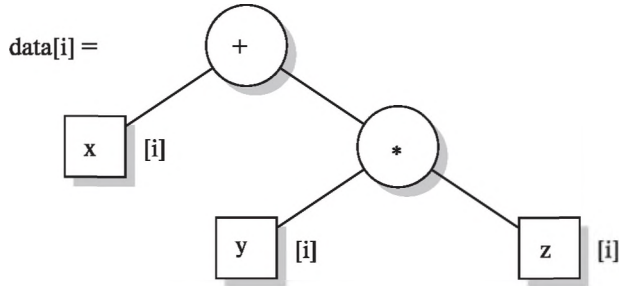


Рис. 1. Дерево, соответствующее выражению $x + y * z$

Вся работа по-прежнему выполняется в операторе присваивания вектора, который теперь принимает произвольное выражение в качестве аргумента:

```
template <class Expr>
const vector& vector::operator=(const Expr &expr) {
    for(int i = 0; i < size; ++i)
        data[i] = expr[i];
    return *this;
}
```

Заметим, что *шаблон выражения* `binary_op`, определённый таким образом, имеет рекурсивную древовидную структуру. Так, выражение $x + y * z$ имеет тип

```
binary_op<vector, plus, binary_op<vector, multiply, vector>>
```

и соответствует дереву выражения, представленному на рис. 1. Структура выражения автоматически определяется компилятором в соответствии с приоритетами операций языка C++. Работа по *вычислению* выражения в данном случае выполняется оператором индексации `[]` и фактически соответствует симметричному обходу дерева выражения.

ГЕНЕРАЦИЯ ВЫЧИСЛИТЕЛЬНЫХ ЯДЕР OPENCL

Библиотека VexCL использует технику, аналогичную описанной в предыдущем разделе, для автоматической генерации, компиляции и выполнения вычислительных ядер OpenCL. Шаблоны выражений при этом формируются аналогичным образом; меняется только процедура *вычисления* выражения. Общая последовательность расчётов с использованием технологии OpenCL выглядит следующим образом:

- вычислительное ядро компилируется во время выполнения основной программы из исходного кода C99, хранящегося в символьном массиве;
- аргументы ядра задаются при помощи вызовов функций OpenCL API;
- ядро ставится в очередь на выполнение вычислительным устройством.

Исходный код ядра чаще всего либо считывается из текстового файла, распространяющегося вместе с основной программой, либо хранится в статической строковой переменной. Однако исходный код может быть и *сгенерирован* во время исполнения.

Далее опишем методику генерации исходного кода OpenCL, применяющуюся в библиотеке VexCL. Вспомним, что выражение вида $a = x + y - z$ должно приводить к генерации следующего вычислительного ядра:

```
kernel void vexcl_vector_kernel(
    ulong n,
    global double * prm_1,
    global double * prm_2,
    global double * prm_3,
    global double * prm_4
)
{
    for(size_t idx=get_global_id(0); idx<n; idx+=get_global_
size(0)) {
        prm_1[idx] = ( ( prm_2[idx] + prm_3[idx] ) - prm_4[idx] );
    }
}
```

Данное ядро имеет чётко определённую структуру. От конкретного выражения зависят лишь строки 3–6, в которых задаются параметры ядра, и строка 10, выполняющая непосредственный расчёт. При этом параметрами ядра становятся терминалы (листья) дерева выражения, а при формировании тела цикла в строке 10 используется выражение в целом. Таким образом, задачу генерации и исполнения вычислительного ядра можно разбить на следующие шаги: формирование списка аргументов ядра; формирование строкового представления выражения; генерация ядра; задание аргументов ядра и постановка ядра в очередь на исполнение. Ниже эти шаги рассматриваются подробнее.

Формирование списка аргументов ядра. Напомним, что техника шаблонов выражений, описанная в предыдущем разделе, позволяет получить дерево, соответствующее выражению C++ на этапе компиляции программы. Для обхода дерева во время исполнения программы использовался оператор индексации [], определённый как в структуре `binary_op`, так и в классе `vector`. В данном случае добавим в класс `vector` статический метод `declare_params`. Метод получает ссылки на выходной текстовый поток, в котором формируется исходный код ядра, и на текущую позицию в списке параметров. Каждый терминал «знает», параметры какого типа и в каком количестве ему необходимы. Так, в случае вектора в список параметров добавляется указатель на глобальную память. Структура `binary_op` рекурсивно делегирует работу по формированию списка параметров своим операндам.

```
/*static*/
void vector::declare_params(std::ostream &src, unsigned &pos) {
    src << «,\n global double * prm_» << ++pos;
}

template <class LHS, class OP, class RHS>
```

```

/*static*/
void binary_op::declare_params(std::ostream &src, unsigned &pos)
{
    LHS::declare_params(src, pos);
    RHS::declare_params(src, pos);
}

```

Формирование строкового представления выражения. На данном шаге нам потребуется строковое представление бинарных операций. Для этого добавим в структуры `plus`, `minus` и др. статический метод `to_string`:

```

struct plus {
    static std::string to_string(std::ostream &src) { src << « +
<< ; }
};

```

Теперь каждый терминал может сформировать свою часть строкового представления выражения, а структура `binary_op` может рекурсивно объединить эти части в единое целое:

```

/*static*/
void vector::to_string(std::ostream &src, unsigned &pos) {
    src << «prm_» << ++pos << «[idx]»;
}

template <class LHS, class OP, class RHS>
/*static*/
void binary_op::to_string(std::ostream &src, unsigned &pos) {
    src << «( «;
    LHS::to_string(src, pos);
    OP::to_string(src);
    RHS::to_string(src, pos);
    src << « )»;
}

```

Генерация ядра. На данном этапе есть всё необходимое для формирования полного исходного текста вычислительного ядра. Шаблон функции `kernel_source()`, приведённый ниже, возвращает сгенерированный код для заданного выражения C++. Параметрами шаблона являются типы выражений слева и справа от знака присваивания.

```

template <class LHS, class RHS>
std::string kernel_source() {
    std::ostringstream src;

    src << «kernel void vexcl_vector_kernel(\n  ulong n»;

    unsigned pos = 0;
    LHS::declare_params(src, pos);
    RHS::declare_params(src, pos);
    src <<
        «)\n{\n»
        « for(size_t idx = get_global_id(0); idx < n; idx +=
get_global_size(0)) {\n»
        «     «;
}

```



```

pos = 0;
LHS::to_string(src, pos); src << « = «;
RHS::to_string(src, pos); src << «;\n»;

src << «  }\n}\n»;

return src.str();
}

```

Задание аргументов ядра. До сих пор пользовались фактом, что исходный код ядра целиком определяется типом выражения. Это позволяло использовать статические методы для генерации исходного кода ядра. Для задания аргументов ядра потребовались конкретные значения терминалов. По-прежнему, непосредственная работа по вызову функций OpenCL API выполняется терминалами выражения, а само выражение делегирует работу своим операндам.

```

void vector::set_arguments(cl::Kernel &krn, unsigned &pos) const
{
    krn.setArg(pos++, buffer);
}

template <class LHS, class RHS>
void binary_op::set_arguments(cl::Kernel &krn, unsigned &pos)
const {
    lhs.set_arguments(krn, pos);
    rhs.set_arguments(krn, pos);
}

```

Постановка ядра в очередь на исполнение. Как и в примере из раздела «Шаблоны выражений», основная работа по вычислению выражения проводится в операторе присваивания вектора. Оператор присваивания по-прежнему является шаблонным методом с типом выражения в качестве параметра шаблона. Так как код вычислительного ядра полностью определяется типом выражения, можно использовать локальную статическую переменную для кэширования скомпилированного ядра. Это позволит выполнить дорогостоящую работу по генерации и компиляции единственный раз.

```

template <class Expr>
const vector& vector::operator=(const Expr &expr) {
    static cl::Kernel krn = build_kernel(kernel_source<This,
Expr>());

    unsigned pos = 0;
    krn.setArg(pos++, size);

    this->set_arguments(krn, pos);
    expr.set_arguments(krn, pos);

    queue.enqueueNDRangeKernel(krn, cl::NullRange, size,
cl::NullRange);

    return *this;
}

```

Рассмотренный пример в достаточной мере описывает процесс генерации исходного кода, используемого в библиотеке VexCL. Однако фактическая реализация VexCL оказывается намного более сложной. Это объясняется тем, что, помимо векторов, определяются и другие типы терминалов, такие как скаляры, константы, порядковые номера элементов и пр. Кроме бинарных, поддерживаются унарные, тернарные, n -арные операции, допускается использование встроенных функций OpenCL и функций, определённых пользователем. На время подготовки данной статьи основной исходный код библиотеки (помимо сопутствующих примеров и тестов) содержит более 28 тысяч строк. Для упрощения работы с шаблонами выражений используется библиотека Boost.Proto [Niebler, 2007].

ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ

Рассмотрим эффективность описанной методики на примере параметрического исследования системы Лоренца [Spagow, 1982]. Система Лоренца описывает хаотическую траекторию частицы $\mathbf{p} = (x, y, z)^T \in \mathbb{R}^3$:

$$\dot{x} = \sigma(y - x), \quad \dot{y} = Rx - y - xz, \quad \dot{z} = xy - bz, \quad (1)$$

где $\sigma, R, b \in \mathbb{R}$ — параметры системы. Будем одновременно искать решения системы (1) для большого числа различных значений параметра R . Для решения будем использовать Boost.odeint [Ahnert, Mulansky, 2011] — современную библиотеку шаблонов языка C++, предоставляющую набор абстрактных алгоритмов решения обыкновенных дифференциальных уравнений (ОДУ), реализованных независимо от пользовательских структур данных. Это позволяет применять Boost.odeint для решения ОДУ с использованием графических ускорителей [Ahnert et al., 2014; Demidov et al., 2013].

Библиотека Boost.odeint позволяет решать ОДУ следующего общего вида:

$$\dot{x} = f(x, t), \quad x(0) = x_0, \quad (2)$$

где x — переменная состояния; t чаще всего обозначает время; $f(x, t)$ — системная функция, определяющая ОДУ. Boost.odeint поддерживает работу с библиотеками Thrust и VexCL. Поэтому для решения системы Лоренца остаётся только определить соответствующую системную функцию.

Версия системного функтора для библиотеки Thrust представлена ниже. Структура `thrust_lorenz` содержит параметры модели и перегружает оператор вызова функции `()` с сигнатурой, требуемой Boost.odeint. Для хранения переменных состояния используется класс `thrust::device_vector<double>`. Вектор делится на три равные части, использующиеся для хранения компонент x , y и z . Оператор вызова функции использует стандартный для библиотеки Thrust подход: несколько итераторов упаковываются в единственный `thrust::zip_iterator` и передаются для обработки в алгоритм `thrust::for_each` вместе с функтором `lorenz_func`.

```
struct thrust_lorenz {
    typedef thrust::device_vector<double> State;
    int n;
```

```

double sigma, b;
const State &R;

thrust_lorenz(int n, double sigma, double b, const State &R)
    : n(n), sigma(sigma), b(b), R(R) {}

struct lorenz_functor {
    double sigma, b;

    lorenz_functor(double sigma, double b) : sigma(sigma),
b(b) {}

    template <class T>
    __host__ __device__ void operator()(T t) const {
        double r = thrust::get<0>(t);
        double x = thrust::get<1>(t);
        double y = thrust::get<2>(t);
        double z = thrust::get<3>(t);

        thrust::get<4>(t) = sigma * (y - x);
        thrust::get<5>(t) = r * x - y - x * z;
        thrust::get<6>(t) = -b * z + x * y;
    }
};

void operator()(const State &p, State &dpdt, double t) const
{
    BOOST_AUTO(start,
        thrust::make_zip_iterator( thrust::make_tuple(
            R.begin(),
            p.begin(), p.begin() + n, p.begin() + 2 *
n,
            dpdt.begin(), dpdt.begin() + n,
            dpdt.begin() + 2 * n
        ) )
    );

    thrust::for_each(start, start+n, lorenz_functor(sigma,
b));
}
};

```

В рамках данной работы была также проведена адаптация Boost.odeint для работы с библиотекой CUBLAS (высокоэффективная реализация программного интерфейса BLAS от компании NVIDIA). При этом вызовы CUBLAS использовались только во внутренних алгоритмах Boost.odeint, а в качестве системного функтора применялась приведённая выше структура `thrust_lorenz`.

Версия системного функтора для библиотеки VexCL оказывается намного компактнее благодаря использованию векторных выражений. Для хранения переменных состояния используется трехкомпонентный вектор `vex::multivector<double, 3>`.

```

struct vexcl_lorenz {
    typedef vex::multivector<double, 3> State;
    double sigma, b;

```

```

const vex::vector<double> &R;

void operator()(const State &p, State &dpdt, double t) const
{
    dpdt = std::tie(
        sigma * (p(1) - p(0)),
        R * p(0) - p(1) - p(0) * p(2),
        p(0) * p(1) - b * p(2) );
}
};

```

На рис. 2 представлено время расчёта для библиотек Thrust, CUBLAS и VexCL. В расчётах использовалась видеокарта NVIDIA Tesla K40c. На графике слева для каждого из вариантов расчёта показано число секунд, затраченное на выполнение 10 000 шагов по времени в зависимости от числа различных значений параметра R . На графике справа отображено ускорение относительно библиотеки CUBLAS. Библиотека Thrust также позволяет выполнять расчёты на центральном процессоре с использованием технологии OpenMP. Соответствующие результаты также представлены на рис. 2 для процессора Intel Core i7 930.

Как видно на рис. 2, для достаточно больших задач библиотека CUBLAS оказывается на 50...60 % медленнее вариантов Thrust и VexCL(v1). Это объясняется тем, что линейные комбинации вида $x_0 = \sum_{i=1,n} \alpha_i x_i$, используемые в алгоритмах Boost.odeint, приходится реализовывать в виде такой последовательности вызовов CUBLAS:

```

cublasDset(...);           //  $x_0 = 0$ 
cublasDaxpy(...);         //  $x_0 = x_0 + \alpha_1 x_1$ 
...
cublasDaxpy(...);         //  $x_0 = x_0 + \alpha_n x_n$ 

```

Промежуточные значения при этом хранятся в глобальной памяти устройства, что и приводит к такому замедлению работы.

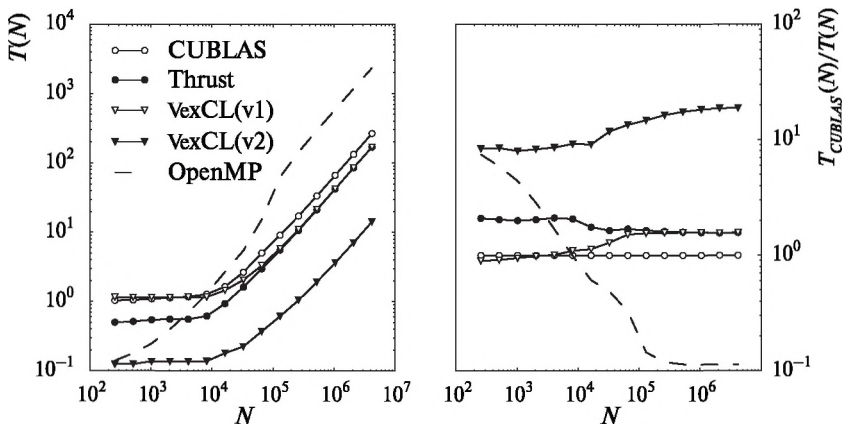


Рис. 2. Сравнение производительности нескольких GPGPU-библиотек при одновременном решении большого числа систем Лоренца

Варианты Thrust и VexCL(v1) показывают одинаковую производительность на достаточно больших задачах. С уменьшением числа параметров R производительность VexCL падает. Это объясняется тем, что технология OpenCL имеет большее время инициализации по сравнению с архитектурой CUDA. Именно ненулевым временем инициализации (как для OpenCL, так и для CUDA) объясняется характерное поведение вариантов с использованием GPU при малых значениях N на левом графике (см. рис. 2). Впрочем, это отличие OpenCL от CUDA представляется незначительным, если учесть, что на малых задачах использование видеокарт оказывается в принципе неэффективным и вариант OpenMP показывает наилучшую производительность. С ростом же N центральный процессор оказывается в 10–15 раз медленнее видеокарты.

Заметим, что варианты CUBLAS, Thrust и VexCL(v1) имеют один и тот же недостаток: в качестве переменных состояния используются массивы, расположенные в глобальной памяти. Там же хранятся и промежуточные переменные, используемые в алгоритмах Boost.odeint. Более того, каждая операция приводит к запуску отдельного вычислительного ядра. Запуск ядра на выполнение имеет ненулевую стоимость как в CUDA, так и в OpenCL, но, что более важно, каждое вычислительное ядро должно считывать и записывать промежуточное состояние в медленную глобальную память. Так как рассматриваемая задача ограничена по производительности пропускной способностью памяти, это ведёт к заметному снижению скорости расчётов. Фактически, это та же проблема использования временных переменных, которая рассматривалась в начале раздела «Шаблоны выражений» и при анализе варианта CUBLAS.

Данную проблему можно было бы обойти написанием специального монолитного ядра, которое за один запуск выполняло бы один шаг интегрирования. Но при этом пришлось бы повторить работу, уже проведённую авторами Boost.odeint. Кроме того, для каждой новой задачи или при смене алгоритма интегрирования пришлось бы полностью переписывать такое ядро. VexCL позволяет автоматически генерировать монолитное ядро при помощи специальных символьных переменных. Экземпляры класса `vex::symbolic<T>` выводят в заданный текстовый поток все арифметические операции, выполняемые с ними. Например, в следующем примере две символьные переменные объявляются и участвуют в арифметическом выражении:

```
vex::generator::set_recorder(std::cout);
vex::symbolic<double> x = 6, y = 7;
x = sin( x * y );
```

При выполнении этого кода будет сгенерирован следующий текст:

```
double var1 = 6;
double var2 = 7;
var1 = sin( (var1 * var2) );
```

Используя символьные переменные в качестве переменных состояния и выполнив единственный шаг по времени, можно получить текстовое представление алгоритма интегрирования Boost.odeint (в данном случае применяется метод Рунге-Кутты четвёртого порядка). С помощью функционала би-

библиотеки VexCL можно сгенерировать монолитное вычислительное ядро из такого представления. Такое применение символьных переменных позволяет использовать существующие алгоритмы Boost.odeint вместо написания специализированного монолитного ядра. Как видно на рис. 2, вариант VexCL(v2) оказывается примерно на порядок быстрее вариантов Thrust и VexCL(v1) для больших значений N . На малых задачах наблюдается характерное для OpenCL снижение производительности, но даже здесь вариант VexCL(v2) оказывается самым эффективным.

ЗАКЛЮЧЕНИЕ

Представлена методика, применяющаяся в библиотеке с открытым исходным кодом VexCL для автоматической генерации вычислительных ядер OpenCL из выражений языка программирования C++. Данная методика позволяет значительно упростить написание программ, использующих графические карты для ускорения расчётов. Также описана возможность использования обобщённых (шаблонных) алгоритмов C++ для автоматической генерации монолитных вычислительных ядер OpenCL.

На примере параметрического исследования системы Лоренца показана гибкость и эффективность описанного подхода. За счёт уменьшения объёма промежуточных результатов, хранящихся в глобальной памяти вычислительного устройства, удаётся ускорить расчёты от 1,5 до 10 раз по сравнению с библиотеками CUBLAS и Thrust.

Кроме того, использование библиотеки VexCL позволяет существенно нивелировать недостатки технологии OpenCL по сравнению с NVIDIA CUDA: библиотека скрывает от разработчика низкоуровневые детали OpenCL API, позволяя сконцентрироваться на смысловой части задачи. В то же время применение OpenCL позволяет использовать весь спектр оборудования, поддерживаемого этой технологией.

ЛИТЕРАТУРА

- [Ahnert, Mulansky, 2011] *Ahnert K., Mulansky M.* Odeint — Solving Ordinary Differential Equations in C++ // IP Conf. Proc. 2011. V. 1389. P. 1586–1589.
- [Ahnert et al., 2014] *Ahnert K., Demidov D., Mulansky M.* Solving ordinary differential equations on GPUs // Numerical Computations with GPUs / Ed. V. Kindratenko. Springer International Publishing, 2014. P. 125–157.
- [Bell, Hoberock, 2011] *Bell N., Hoberock J.* Thrust: A productivity-oriented library for CUDA // GPU Computing Gems. 2011. V. 7.
- [Demidov et al., 2013] *Demidov D., Ahnert K., Rupp K., Gottschling P.* Programming CUDA and OpenCL: A case study using modern C++ libraries // SIAM J. Scientific Computing. 2013. V. 35. N. 5. P. C453–C472.
- [Niebler, 2007] *Niebler E.* Proto: A compiler construction toolkit for DSEs // Proc. 2007 Symp. Library-Centric Software Design / ACM. 2007. P. 42–51.
- [Sparrow, 1982] *Sparrow C.* The Lorenz equations: bifurcations, chaos, and strange attractors. N. Y.: Springer-Verlag, 1982. V. 41.
- [Veldhuizen, 1995] *Veldhuizen T.* Expression Templates // C++ Report. 1995. V. 7. N. 5. P. 26–31.

IMPLEMENTATION OF OPENCL KERNEL GENERATOR IN VEXCL LIBRARY

D. E. Demidov

Kazan Branch of Joint Supercomputer Center, Russian Academy of Sciences

The technique used in the opensource VexCL library for automatic OpenCL kernel generation from C++ expressions is presented. The technique allows to greatly reduce the effort of creation of computational programs that use graphical accelerators. The advantages of automatic code generation with respect to libraries with fixed programming interfaces are shown.

Keywords: C++, OpenCL, GPGPU, VexCL.

Demidov Denis Evgenievitch — senior researcher, PhD, dennis.demidov@gmail.com

МЕЗОСКОПИЧЕСКИЕ ФЛУКТУАЦИИ НАСЕЛЁННОСТИ КУБИТА В ПОЛЕ УПРАВЛЯЮЩЕГО СИГНАЛА

М. В. Денисенко, А. М. Сатанин

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

Изучаются мезоскопические флуктуации населённости джозефсоновского кубита в переменном поле, которое представляет собой суперпозицию электромагнитных импульсов большой амплитуды. Показано, что относительная фаза импульсов ответственна за темп переходов Ландау-Зинера и, соответственно, за частоту переходов между адиабатическими состояниями. Поскольку в коаксиальных линиях на кубит поступают случайные последовательности импульсов, длительность которых контролируется с точностью до периода поля, то это приводит к сильным мезоскопическим флуктуациям населённости кубита. При этом влияние относительной фазы импульсов на населённость подобно магнитному полю в мезоскопике, что ведёт к разрушению интерференционно картины населённости. Изучено влияние длительности импульса и шума на обнаруженные флуктуационные эффекты.

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

Ключевые слова: кубит, переходы Ландау-Зинера, мезоскопические флуктуации.

ВВЕДЕНИЕ

Как хорошо известно, эффекты слабой локализация и мезоскопические флуктуации полной проводимости (кондактанса) обусловлены фазовой когерентностью при прохождении электронных волн через образцы, размеры которых малы по сравнению с характерной длиной неупругого рассеяния [Гантмахер, 2013; Lee et al., 1986]. В условиях квазиклассичности большой размер траекторий электронов приводит к тому, что рассеяние электронных волн на хаотически расположенных отдельных примесях является сильно коррелированным процессом. Особенно чувствителен к конфигурации примесей кондактанс одномерной проволоки, когда достаточно сдвинуть одну примесь на расстояние порядка длины волны электрона, чтобы заметно изменить интерференцию электронных волн; в этом случае возвратные траектории электронов всегда проходят одни и те же рассеиватели, что и обуславливает сильную чувствительность проводимости к виду потенциала [Сатанин, 1993]. При этом квантовые интерференционные поправки неупорядоченных систем весьма чувствительны к изменению тех параметров, которые нарушают симметрию относительно обращения времени, например, к результату действия магнитного поля или спин-орбитального взаимодействия. При измерениях кондактанса управляющими параметрами обычно служат концентрация электронов, магнитное поле, частота поля, размеры системы и др. В то же время конфигурацию примесного потенциала на эксперименте трудно

Денисенко Марина Валерьевна — младший научный сотрудник, mar.denisenko@gmail.com
Сатанин Аркадий Михайлович — доктор физико-математических наук, профессор кафедры теоретической физики, sarkady@mail.ru

контролировать. Между тем, в недавних модельных экспериментах было показано, как можно менять реализации поля рассеивателей и изучать интересные эффекты в теории слабой локализации и квантового хаоса [Штокман, 2004].

Другая возможность открывается в связи с изучением временной динамики тока в сверхпроводниковых рамках микронных размеров со встроенными в них джозефсоновскими переходами (слабыми связями). При низких температурах, когда энергия теплового движения намного меньше расстояния между низшими уровнями, движение тока в такой рамке описывается квантовой механикой: можно возбуждать как токи вращающиеся в противоположных направлениях, так и ток, который является их квантовой суперпозицией. Иными словами, сверхпроводниковая рамка со слабыми связями, при учёте только пары нижних уровней, ведёт себя как двухуровневая система (кубит) [You, Nori, 2005, 2011]. Интересно отметить, что изменение магнитного потока, пронизывающего рамку, приводит к периодической зависимости расстояния между уровнями, что совершенно аналогично зависимости энергии блоховского электрона от квазиимпульса в зонной теории твёрдых тел. Если рамку с током поместить в переменное поле, то динамика населённостей будет меняться во времени, причём приложенное поле будет вести себя подобно потенциалу, приложенному к блоховскому электрону, тогда как время будет играть роль пространственной координаты.

В данной работе изучается динамика переходов между уровнями кубита в сильном переменном поле, которое формируется путём сложения двух импульсов, имеющих разные частоты и характеризующихся относительной фазой, играющей роль управляющего параметра. Мы покажем, что в сильном поле населённость кубита эволюционирует адиабатически, кроме как в непосредственной близости от квазипересекающихся уровней, где наблюдаются квантово-когерентные переходы Ландау-Зинера [Landau, 1932; Zener, 1932] (см. обзор С. Н. Шевченко [Shevchenko et al., 2010]). Данные переходы происходят в моменты времени, соответствующие сближениям уровней и ними аналогичны процессам рассеяния на примесных центрах. При этом количество актов рассеяния соответствует числу квазипересечений на периоде внешнего поля. Зависимости населённостей кубита от времени могут сильно различаться в повторяющихся актах измерений из-за потерь в коаксиальных линиях передач, которые обуславливают сбой фазы сигнала как целое. Относительная фаза сигналов даёт дополнительный сдвиг фазы в процессе эволюции, а также приводит к нарушению симметрии относительно обращения времени. Данный факт позволяет говорить об аналогии с мезоскопическими системами, согласно которой флуктуирующая величина (вероятность перехода между основным и возбуждённым уровнями кубита в единицу времени) ведёт себя подобно проводимости квантовых проволок. В свою очередь, населённость кубита как функция длительности импульса флуктуирует подобно проводимости проволоки в зависимости от её длины, а относительная фаза двух импульсов, подобная влиянию магнитного поля в проводниках, ответственна за изменение конфигурации рассеивателей на периоде внешнего поля. Фазовый шум также влияет на населённости уровней и флуктуационные эффекты. Мы продемонстрируем здесь те принципиальные различия, которые имеют место между случайной временной динамикой населённостей кубита и мезоскопическими флуктуациями в проволоках.

ДИНАМИКА КУБИТА В СИЛЬНОМ ПОЛЕ

Будем изучать систему близкую к той, которая реализована в экспериментах [Berns et al., 2006, 2008; Gustavsson et al., 2013; Oliver et al., 2005]. Речь пойдёт о сверхпроводниковом кубите, представляющем собой ниобиевую рамку, «разорванную» тремя джозефсоновскими переходами (так называемый 3JJ кубит [Orlando et al., 1999]). В отсутствие внешних полей, когда эффективный профиль системы представляет собой симметричный двухъямный потенциал, в контуре могут циркулировать незатухающие сверхтоки по часовой стрелке и против неё $\pm I_p$, т. е. кубит может находиться в двух адиабатических состояниях $|0\rangle = (0, 1)^T$ и $|1\rangle = (1, 0)^T$ (собственные векторы матрицы Паули σ_z). Положения уровней кубита зависят от внешнего стационарного магнитного потока Φ , пронизывающего сверхпроводящий контур, туннельного расщепления Δ и определяются выражением: $E_{\pm} = \pm \sqrt{\epsilon_0^2 + \Delta^2}$, где $\epsilon_0 = 2I_p(\Phi/\Phi_0 - 1/2)$; $\Phi_0 = h/2e$ — квант магнитного потока [Orlando et al., 1999]. Для манипулирования состояниями кубита введём переменное высокочастотное поле $\epsilon_-(t)$ большой амплитуды, подаваемое на кубит по коаксиальным линиям (см., напр., работу [Gustavsson et al., 2013]). Динамика кубита описывается уравнением Шрёдингера $i \frac{\partial}{\partial t} |\psi(t)\rangle = H(t) |\psi(t)\rangle$ с гамильтонианом

$$H(t) = -\frac{1}{2} \begin{bmatrix} \epsilon(t) & \Delta \\ \Delta & -\epsilon(t) \end{bmatrix}, \quad (1)$$

где $\epsilon(t) = \epsilon_0(t) + \epsilon_-(t)$ — суммарное приложенное поле. В качестве примера будем рассматривать импульс конечной длительности вида

$$\epsilon_-(t) = A(\cos(\omega t + \theta) - \gamma \cos(2\omega t)), \quad (2)$$

где A — амплитуда внешнего переменного поля; γ и θ — относительные амплитуда и фаза смешиваемых сигналов. В условиях эксперимента [Gustavsson et al., 2013] импульсы, подаваемые с генератора, могут испытать при прохождении коаксиальных линий потери и сдвиги фаз, что, несомненно, будет отражаться на времени прихода импульса на кубит. Для учёта этого эффекта, введём случайное время прихода импульса на кубит t_0 , или соответствующую случайную фазу $\varphi = t_0\omega$. Как и в работе [Gustavsson et al., 2013], нас будет интересовать поведение населённости при большой амплитуде поля $A \gg \sqrt{\Delta^2 + \epsilon_0^2}$, однако для начала полезно продемонстрировать особенности поведения системы в слабом поле (в приближении Раби).

В случае монохроматического сигнала ($\gamma = 0$) и пределе слабого поля, когда $A \ll \sqrt{\epsilon_0^2 + \Delta^2}$, вероятность перехода между уровнями кубита осциллирует на частоте Раби [Скалли, Зубайри, 2003]. При этом фаза φ входит в аргумент квадрата синуса, а усреднение вероятности приводит к тому, что населённости верхнего и нижнего уровней кубита в среднем равны.

В случае сильного поля, когда A сравнима или больше расстояния между уровнями в кубите $\sqrt{\Delta^2 + \epsilon_0^2}$, можно использовать переход во вращающуюся систему координат [Скалли, Зубайри, 2003]. Эффективный гамильтониан системы для управляющего поля в форме (2), можно записать в виде:

$$H_{eff}(t) = -\frac{1}{2} \sum_{n,m=-\infty}^{\infty} \begin{pmatrix} 0 & \Delta_{nm} \exp(i(\epsilon_0 + (n-2m)\omega)t) \\ \Delta_{nm}^* \exp(-i(\epsilon_0 + (n-2m)\omega)t) & 0 \end{pmatrix}, \quad (3)$$

где

$$\Delta_{nm} = \Delta J_n \left(\frac{A}{\omega} \right) J_m \left(\frac{\gamma A}{2\omega} \right) \exp(in(\theta + \varphi) - 2im\varphi) \cdot \exp \left[-i \frac{A}{\omega} \sin(\theta + \varphi) + i \frac{\gamma A}{2\omega} \sin(2\varphi) \right]$$

($J_n(x)$ — функция Бесселя первого порядка типа n).

В отсутствие внешнего воздействия, когда $\epsilon_0 = 0$ и $A = 0$, гамильтониан (3) отражает резонансную динамику системы с частотой Раби $\Omega = \Delta$. В присутствии только постоянного смещения ($A = 0$, $\epsilon_0 \neq 0$) Раби-динамика сохранится, но ϵ_0 будет играть роль отстройки от резонанса. При увеличении сдвига ϵ_0 собственные функции системы стремятся к состояниям $|0\rangle$, $|1\rangle$, однако осцилляции между ними подавлены из-за отсутствия резонанса, и в случае $\epsilon_0 \ll \Delta$ максимальная населённость пропорциональна $(\Delta/\epsilon_0)^2$ [Ashhab et al., 2007].

В резонансном приближении в соотношении (3) можно пренебречь быстро осциллирующими слагаемыми за исключением тех, для которых выполняется условие:

$$\epsilon_0 + (n - 2m)\omega = 0. \quad (4)$$

В этом случае осцилляции между уровнями характеризуются обобщённой частотой Раби $\Omega = \left| \sum_{n,m} \Delta_{nm} \right|$.

Протокол измерения, используемый в работах [Berns et al., 2006, 2008; Gustavsson et al., 2013; Oliver et al., 2005], позволяет различать вероятности нахождения кубита в состояниях с определённым значением тока, в частности, выделить состояние с $+I_p$, которое соответствует возбуждённому адиабатическому уровню кубита. Учитывая сказанное, будем интересоваться вероятностью перехода $P_{\alpha \rightarrow \beta}(t)$ кубита из состояния

$$|\alpha\rangle = \frac{\Delta}{\sqrt{\Delta^2 + (\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0)^2}} \begin{pmatrix} 1 \\ \frac{\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0}{\Delta} \end{pmatrix} \quad (5)$$

в состояние

$$|\beta\rangle = |1\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (6)$$

(на уровень с энергией $E_- = -\epsilon_0/2$). Следовательно, выражение для вероятности межуровневых переходов имеет вид:

$$P_{\alpha \rightarrow \beta}(t) = \frac{\Delta (\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0)}{\Delta^2 + (\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0)^2} \left[\frac{\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0}{2\Delta} + \frac{4\Delta^2 - (\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0)^2}{2\Delta (\sqrt{\epsilon_0^2 + \Delta^2} - \epsilon_0)} \times \right. \\ \left. \times \cos^2(\Omega t) + \sin(2\Omega t) \cdot \sin(\chi_{nm}) \right], \quad (7)$$

где $\chi_{nm} = \text{Arg}(\Delta_{nm})$.

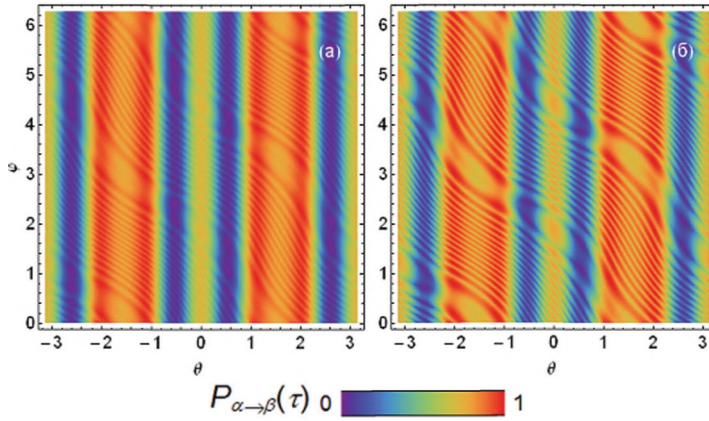


Рис. 1. Зависимость населённости $P_{\alpha \rightarrow \beta}(\tau)$ от относительной фазы θ и общей фазы φ , рассчитанная в резонансном приближении (а) и численным решением уравнения для матрицы плотности (б). Параметры системы: $\Delta/\omega = 0,5$; $A/\omega = 22$; $\varepsilon_0/\omega = 2$; $\gamma = 0,25$; $\tau = 10T$

Выражение (7) показывает, что вероятность перехода зависит как от времени запаздывания импульсов (общей фазы φ), так и относительной фазы θ импульсов. На рис. 1 изображена интерференционная картина вероятности перехода кубита из основного состояния на верхний уровень $P_{\alpha \rightarrow \beta}(\tau)$ в зависимости от фаз импульсов в момент времени $\tau = 10T$ (T — период сигнала), равный длительности подаваемого импульса. Красные области означают резонансные пики, где $P_{\alpha \rightarrow \beta} \approx 1$, а сиреневые — «провалы», когда населённость падает до нуля ($P_{\alpha \rightarrow \beta} \approx 0$). Заметим, что резонансная теория (см. рис. 1(а)) даёт хорошее согласование с прямым численным расчётом (см. рис. 1(б)), что позволяет качественно объяснить формирование интерференционной картины при различных значениях импульса. Как следует из соотношения (4), положение резонансных пиков определяется условием $\varepsilon_0 + (n - 2m)\omega = 0$ (при $\omega \gg \Delta$), где n и m — номера соответствующих функций Бесселя, определяющих частоту обобщённого раби-резонанса Ω . Отметим, что при малых амплитудах $A \ll \omega$, ε_0 , Δ импульса наблюдается обычная раби-динамика системы и вероятность перехода не чувствительна к относительной фазе и запаздыванию импульса. Однако при увеличении амплитуд частота Раби, пропорциональная сумме произведений функций Бесселя $J_n\left(\frac{A}{\omega}\right)$ и $J_m\left(\frac{\gamma A}{2\omega}\right)$, взятых

с разными фазами: $\exp\left[-i\left|\frac{A}{\omega}\sin(\theta + \varphi) - n(\theta + \varphi)\right|\right]$ и $\exp\left[i\left|\frac{\gamma A}{2\omega}\sin(2\varphi) - 2m\varphi\right|\right]$,

оказывается чувствительной к значению относительной разности фаз θ . Однако по отношению к общей фазе φ импульса частота Раби не чувствительна, так как мы учитываем запаздывание импульса как целого, т.е зависимость от задержки на время $t_0 = \varphi/\omega$ проявляется лишь в набеге фаз частоты Раби $\chi_{nm} = \text{Arg}(\Delta_{nm})$, что продемонстрировано на рис. 1(а). Также имеется связь симметрии расположения максимумов и минимумов на резонансной картине фаз, которая устанавливается преобразованием: $\varphi \rightarrow \varphi + \pi$ и $\theta \rightarrow \theta + \pi$, где

s — любое целое число, что следует из формы импульса (2). Отметим, что развёртка по амплитудам сигналов и магнитному потоку (описываемому параметром ε_0) также имеет интересную резонансную структуру [Satanin et al., 2014].

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Как следует из качественного анализа, проведённого выше, населённость кубита, в случае воздействия на него импульса сложной формы, чувствительна как к относительной, так и к абсолютной фазам смешиваемых импульсов. Однако в реальной ситуации необходимо принять во внимание фазовый и энергетический шум. Оператор плотности (матрицу плотности) кубита ρ в марковском приближении можно записать в виде:

$$\frac{\partial \hat{\rho}}{\partial t} = -i[\hat{\rho}, \hat{H}] + \frac{\Gamma_f}{2}(\hat{\sigma}_z \hat{\rho} \hat{\sigma}_z - \hat{\rho}) + \frac{\Gamma_e}{2}(2\hat{\sigma}_- \hat{\rho} \hat{\sigma}_+ - \hat{\sigma}_+ \hat{\sigma}_- \hat{\rho} - \hat{\rho} \hat{\sigma}_+ \hat{\sigma}_-), \quad (8)$$

где скорость Γ_f характеризует процесс затухание фазы (дефазировка, dephasing), параметр Γ_e отвечает за темп потери энергии. Согласно экспериментальным работам [Berns et al., 2006, 2008; Gustavsson et al., 2013; Oliver et al., 2005], процессы термического возбуждения (спонтанного возбуждения системы за счёт температуры) в данной работе не учитываются, так как они на три порядка меньше двух других механизмов релаксации.

Для численных расчётов матрицу плотности удобно представить в виде $\hat{\rho} = \frac{1}{2}(\hat{\mathbf{I}} + \boldsymbol{\sigma} \mathbf{R})$, где $\hat{\mathbf{I}}$ — единичная матрица; $\boldsymbol{\sigma} = \{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ — матрица Паули. Уравнение (8) для матрицы плотности эквивалентно уравнению для вектора $\mathbf{R} = \text{Tr}(\boldsymbol{\sigma} \rho(t))$:

$$\left. \begin{aligned} \frac{\partial R_x}{\partial t} &= \varepsilon(t)R_y - 2\Gamma_f R_x - \frac{\Gamma_e}{2} R_x, \\ \frac{\partial R_y}{\partial t} &= -\varepsilon(t)R_x + \Delta R_z - 2\Gamma_f R_y - \frac{\Gamma_e}{2} R_y, \\ \frac{\partial R_z}{\partial t} &= -\Delta R_y - \Gamma_e R_z - \Gamma_e. \end{aligned} \right\} \quad (9)$$

Начальное состояние (5) запишется в виде:

$$\mathbf{R}(0) = \left[\frac{2\Delta\sqrt{\varepsilon_0^2 + \Delta^2}}{\Delta^2 + \left(\sqrt{\varepsilon_0^2 + \Delta^2} - \varepsilon_0\right)^2}, 0, -1 + \frac{2\Delta^2}{\Delta^2 + \left(\sqrt{\varepsilon_0^2 + \Delta^2} - \varepsilon_0\right)^2} \right], \quad (10)$$

а вероятность перехода в состояние (6) будет определяться соотношением: $P_{\alpha \rightarrow \beta}(t) = 1 + R_z(t)$. Численное решение (9) с начальными условиями (10) выполнялось методом Рунге-Кутты четвёртого порядка. При этом был рассчитан логарифм скорости межуровневых переходов

$$\Gamma(\tau) = -\frac{\ln[P_{\alpha \rightarrow \beta}(\tau)]}{\tau} \quad (11)$$

в зависимости от относительной фазы импульсов θ , длительности импульса τ , а также от параметров, ответственных за фазовую и энергетическую релаксацию. Случайный ансамбль формировался численным генерированием систем с различным временем прихода импульса (равновероятным на периоде импульса). Флуктуации характеризовались среднеквадратичным отклонением $\sigma = \sqrt{\langle \Gamma^2 \rangle - \langle \Gamma \rangle^2}$, где $\langle \cdot \rangle$ означает усреднение по всем возможным временам прихода импульса.

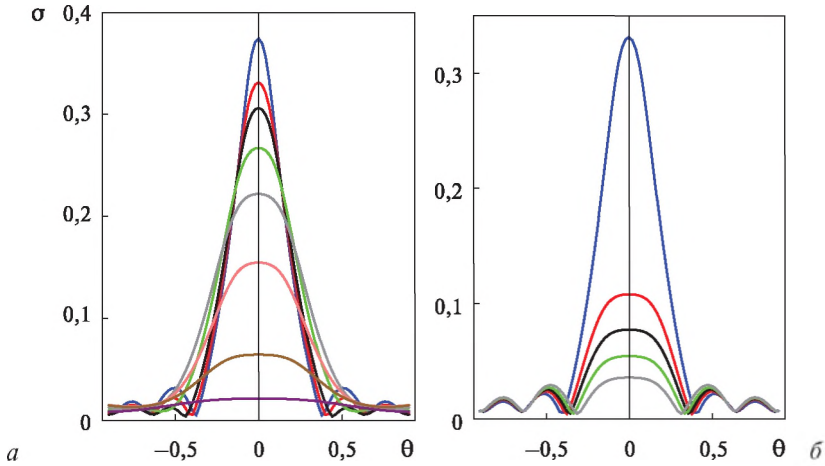


Рис. 2. Зависимости флуктуаций σ от относительной фазы импульсов θ (а) при $\Gamma_e/\omega = 0,0001$ и различных значений фазового шума: $\Gamma_f/\omega = 0,001$ (синяя линия); $0,075$ (красная); $0,005$ (чёрная); $0,0025$ (зелёная); $0,01$ (серая); $0,025$ (розовая); $0,05$ (коричневая) и $0,075$ (бордовая). Влияние энергетического шума на флуктуации в зависимости от θ (б) при $\Gamma_e/\omega = 0,001$ (синяя кривая); $0,002$ (красная); $0,003$ (чёрная); $0,004$ (зелёная); $0,005$ (серая). Параметры системы: $\Gamma_f/\omega = 0,005$; $\Delta/\omega = 0,5$; $A/\omega = 29,83$; $\varepsilon_0/\omega = 5$; $\gamma = 0,25$

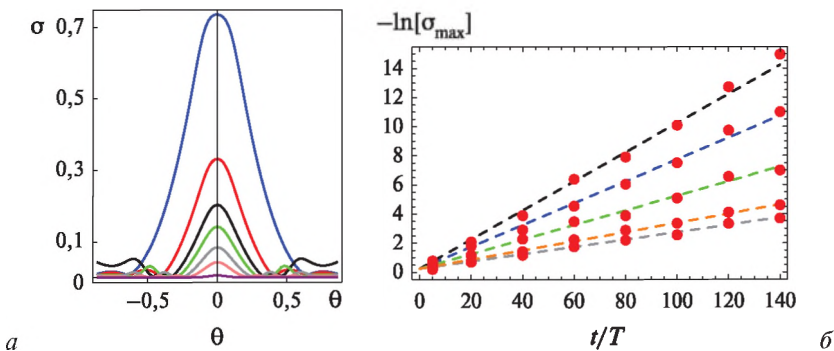


Рис. 3. Флуктуация σ от относительной фазы импульса θ при $\Gamma_f/\omega = 0,005$ (для $\tau = 5T$ (синяя кривая); $10T$ (красная); $20T$ (чёрная); $30T$ (зелёная); $50T$ (серая); $100T$ (розовая); $200T$ (бордовая) (а). Зависимость убывания пика дисперсии σ_{\max} от времени действия импульса (б). Красные точки — рассчитанные значения, а пунктирные линии — качественный закон убывания пика дисперсии. Чёрная линия — при $\Gamma_f/\omega = 0,1$; синяя — $0,075$; зелёная — $0,05$; оранжевая — $0,025$; серая — $0,005$. Параметры системы аналогичны тем, что представлены на рис. 2а

На рис. 2 приведены зависимости отношения флуктуаций логарифма населённости σ от относительной фазы сигналов. Как следует из (1) и (2), при $\theta \neq 0$ гамильтониан неинвариантен относительно преобразования $t \rightarrow -t$. Это означает, что фазы волн, распространяющихся вперёд и назад по времени, будут отличаться, поэтому фазовая когерентность будет подавлена. На рис. 2а видно, что вблизи $\theta = 0$ флуктуации максимальны, а уменьшение характерного времени энергетической релаксации $T_e = 1/\Gamma_e$ приводит к уменьшению флуктуаций (см. рис. 2б). Можно сказать, что параметр T_e играет роль, аналогичную длине неупругой релаксации.

Как следует из работ [Berns et al., 2006, 2008; Gustavsson et al., 2013; Oliver et al., 2005], в проводимых экспериментах время энергетической релаксации на несколько порядков больше времени фазовой релаксации. Приведём зависимости флуктуаций от параметра дефазировки Γ_f . В этом случае (рис. 3а) также наблюдается пик флуктуаций при $\theta = 0$, который уширяется и резко падает с ростом Γ_f . В этом смысле параметр Γ_f аналогичен обратной длине упругого рассеяния. Отметим, что подобные зависимости наблюдались в недавнем эксперименте с джозефсоновским кубитом [Gustavsson, 2013].

На рис. 3б приведено поведение максимума среднеквадратичного отклонения (σ_{\max}) в зависимости от длительности импульса t . Максимум величины убывает по следующему закону: $\sigma_{\max} \approx \exp(-\Gamma_f t / T)$, следовательно, логарифм является линейной функцией длительности сигнала.

ВЫВОДЫ

Таким образом, на временах, меньших T_e , динамика населённости кубита во времени демонстрирует поведение, которое аналогично поведению контактанса проволоки в зависимости от длины. В случае кубита управляющее переменное поле создаёт квазислучайный потенциал, а нарушающая обратимость по времени фаза смешиваемых сигналов действует подобно магнитному полю. При этом переходы Ландау-Зинера между состояниями кубита имитируют процессы, аналогичные процессам рассеяния электрона на примесных центрах. Изменение параметров сигнала приводит к сильным мезоскопическим флуктуациям населённости кубита. Следует отметить, что «реализацию» импульса наперёд заданной формы, а также смену реализации поля легко осуществить путём программирования импульсного генератора [Bylander et al., 2009], поэтому во временной области проще наблюдать различные локализационные и мезоскопические эффекты. Например, в двухчастотном поле с несоизмеримыми периодами легко исследовать локализацию в квазипериодическом потенциале.

Отметим теперь, почему в рассматриваемой системе нет явления типа андерсоновской локализации. Физически это обусловлено тем, что населённость не может превышать единицы, поэтому она остаётся ограниченной. Математически же это объясняется следующим геометрическим фактором: оператор эволюции в динамической задаче принадлежит компактной группе $SU(2)$, а в модели Андерсона — некомпактной группе $SU(1,1)$. Поэтому системы ведут себя аналогично на масштабах меньших длины пробега (времени декогерентности), динамика на таких группах мало отличается, тогда как на больших масштабах поведение систем сильно различается.

Один из авторов (А. М. Сатанин) весьма признателен Ю. К. Лозовику за полезные замечания.

ЛИТЕРАТУРА

- [Гантмахер, 2013] *Гантмахер В. Ф.* Электроны в неупорядоченных средах. М.: Физматлит, 2013. 288 с.
- [Сатанин, 1993] *Сатанин А. М.* Чувствительность сопротивления одномерного проводника к изменениям реализаций случайного потенциала // Журн. экспериментальной и теоретической физики (ЖЭТФ). 1993. Т. 104. С. 3759.
- [Скалли, Зубайри, 2003] *Скалли М. О., Зубайри М. С.* Квантовая оптика. М.: Физматлит, 2003. 512 с.
- [Штокман, 2004] *Штокман Х. Ю.* Квантовый хаос: введение / Пер. с англ. А. И. Малышева. М.: Физматлит, 2004. 376 с.
- [Ashhab et al., 2007] *Ashhab S., Johansson J. R., Zagoskin A. M., Nori F.* Two-level systems driven by large-amplitude fields // *Physical Review A*. 2007. V. 75. P. 063414.
- [Berns et al., 2006] *Berns D. M., Oliver W. D., Valenzuela S. O., Shytov A. V., Berggren K. K., Levitov L. S., Orlando T. P.* Coherent Quasiclassical Dynamics of a Persistent Current Qubit // *Physical Review Letters*. 2006. V. 97. P. 150502.
- [Berns et al., 2008] *Berns D. M., Rudner M. S., Valenzuela S. O., Berggren K. K., Oliver W. D., Levitov L. S., Orlando T. P.* Amplitude spectroscopy of a solid-state artificial atom // *Nature*. 2008. V. 455. P. 51.
- [Bylander et al., 2009] *Bylander J., Rudner M. S., Shytov A. V., Valenzuela S. O., Berns D. M., Berggren K. K., Levitov L. S., Oliver W. D.* Pulse imaging and nonadiabatic control of solid-state artificial atoms // *Physical Review B*. 2009. V. 80. P. 220506.
- [Gustavsson et al., 2013] *Gustavsson S., Bylander J., Oliver W. D.* Time-Reversal Symmetry and Universal Conductance Fluctuations in a Driven Two-Level System // *Physical Review Letters*. 2013. V. 110. P. 016603.
- [Landau, 1932] *Landau L. D.* On the theory of transfer of energy at collisions II // *Physikalische Zeitschrift der Sowjetunion*. 1932. V. 2. P. 46–51.
- [Lee et al., 1986] *Lee P. A., Stone A. D., Fukuyama H.* Universal conductance fluctuations in metals: Effects of finite temperature, interactions, and magnetic field // *Physical Review B*. 1986. V. 35. P. 1039.
- [Oliver et al., 2005] *Oliver W. D., Yu Y., Lee J. C., Berggren K. K., Levitov L. S., Orlando T. P.* Mach-Zehnder Interferometry in a Strongly Driven Superconducting Qubit // *Science*. 2005. V. 310. P. 1653.
- [Orlando et al., 1999] *Orlando T. P., Mooij J. E., Tian L., van der Wal C. H., Levitov L. S., Lloyd S., Mazo J. J.* Superconducting Persistent-Current Qubit // *Physical Review B*. 1999. V. 60. P. 15398.
- [Satanin et al., 2014] *Satanin A. M., Denisenko M. V., Gelman A. I., Nori F.* Amplitude and Phase Effects in Josephson Qubits Driven by a Biharmonic Electromagnetic Field // *Physical Review B*. 2014. V. 90. P. 104516.
- [Shevchenko et al., 2010] *Shevchenko S. N., Ashhab S., Nori F.* Landau-Zener-Stückelberg interference // *Physics Reports*. 2010. V. 492.
- [You, Nori, 2005] *You J. Q., Nori F.* Superconducting circuits and quantum information // *Physics Today*. 2005. V. 58(11). P. 42.
- [You, Nori, 2011] *You J. Q., Nori F.* Atomic physics and quantum optics using superconducting circuits // *Nature*. 2011. V. 474. P. 589.
- [Zener, 1932] *Zener C.* Non-Adiabatic Crossing of Energy Levels // *Proc. Royal Society A*. 1932. V. 137. P. 696.

MESOSCOPIC FLUCTUATIONS OF A QUBIT POPULATION DRIVEN BY CONTROL PULSES

М. V. Denisenko, A. M. Satanin

Lobachevsky State University of Nizhny Novgorod (NNGU)

Mesoscopic fluctuations of a Josephson qubit population driven by an alternating field, which is a superposition of electromagnetic pulses of strong amplitude, are studied. It is shown that the relative phase of the pulses is responsible for the Landau-Zener transition rate and, respectively, for the transitions sequence between the adiabatic states. Since in the coaxial lines on the qubit a random sequence of pulses is arrived whose durations are controlled with accuracy of the field's period, that leads to strong mesoscopic fluctuations of the qubit population.

Because the influence of the relative phase of the pulses on the qubits population similar the magnetic field in the mesoscopic systems that leads to the destruction of the interference pattern of the population. We investigate the influence of the pulse duration and noise on the discovered fluctuations effects.

Keywords: qubits, Landau-Zener transitions, mesoscopic fluctuations.

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

Satanin Arkady Mikhailovich — professor, loctor of physical and mathematical sciences, sarkady@mail.ru

ПРЯМОЕ ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ТУРБУЛЕНТНОГО УСТОЙЧИВО СТРАТИФИЦИРОВАННОГО ВОЗДУШНОГО ПОТОКА НАД ВЗВОЛНОВАННОЙ ВОДНОЙ ПОВЕРХНОСТЬЮ

О. А. Дружинин¹, Ю. И. Троицкая¹, С. С. Зилитинкевич²

¹ Институт прикладной физики Российской академии наук (ИПФ РАН)

² Финский метеорологический институт, Хельсинки, Финляндия

Параметризация потоков импульса и тепла в турбулентном, устойчиво стратифицированном пограничном слое над взволнованной водной поверхностью необходима для крупномасштабных численных климатических моделей и прогноза погоды. При этом связь между крупномасштабными полями скорости и температуры и турбулентными потоками на границе раздела океан-атмосфера даётся относительно простыми функциями, включающими так называемые «балковые» коэффициенты. Значения этих коэффициентов в различных моделях климата меняются в довольно широких пределах, и их конкретизация является довольно сложной задачей натурального и численного эксперимента. Мы используем прямое численное моделирование как инструмент для детального исследования пограничного слоя на границе раздела океан-атмосфера, чтобы далее уточнять и верифицировать параметризацию, используемую в крупномасштабных климатических моделях. Особое внимание уделяется эффектам, обусловленным устойчивой стратификацией воздушного потока.

Работа выполнена при поддержке РФФИ (проекты № 14-05-00367, 13-05-91175).

Ключевые слова: численное моделирование, турбулентный ветер, двухмерные волны на воде.

ВВЕДЕНИЕ

Детальное понимание механизма взаимодействия поверхностных волн с ветром важно для корректной параметризации турбулентных потоков импульса и тепла, определяющих процессы обмена между атмосферой и океаном. В современных моделях прогноза погоды при параметризации процессов обмена используются так называемые балк-формулы, основывающиеся на обобщении известных лабораторных и натуральных данных. Коэффициенты обмена, содержащиеся в балк-формулах, зависят от деталей процесса взаимодействия ветра с волнами, которые и по сей день остаются до конца не выясненными [Fairall et al., 2003]. Особый интерес представляет влияние стратификации воздуха на ветро-волновое взаимодействие.

Экспериментальные измерения в пограничном слое над взволнованной водной поверхностью сложны, особенно в тонком слое вблизи водной поверхности, содержащем вязкий подслои и буферную зону, и в подошвах волн. Типичная ширина этого слоя порядка миллиметра и обычно мала по сравнению с амплитудой поверхностных волн. Контактные измерения, в том числе датчиками, отслеживающими волновую поверхность (wave-following probes),

Дружинин Олег Александрович — ведущий научный сотрудник, доктор физико-математических наук, druzhinin@hydro.appl.sci-nnov.ru

Троицкая Юлия Игоревна — заведующая отделением, доктор физико-математических наук, yuliya@hydro.appl.sci-nnov.ru

Зилитинкевич Сергей Сергеевич — заведующий лабораторией, доктор физико-математических наук, sergej.zilitinkevich@fmi.fi

не позволяют детально исследовать свойства воздушного потока в этой области [Donelan et al., 2005; Hsu, Hsu, 1983; Hsu et al., 1981]. Методы оптической анемометрии (PIV) [Adrian, 1991] дают возможность измерять среднюю скорость ветра на расстоянии порядка 1 мм от водной поверхности, достаточном для исследования буферной зоны и прилегающей области вязкого подслоя для ветров с динамической скоростью (трения) менее 10 см/с [Reul et al., 1999; Veron et al., 2007; Troitskaya et al., 2011]. Наличие стратификации воздуха и необходимость одновременного измерения профилей температуры и скорости значительно усложняют задачу эксперимента. С другой стороны, детальное понимание свойств воздушного потока вблизи водной поверхности чрезвычайно важно для моделирования взаимодействия волн с ветром.

Численный эксперимент представляет собой альтернативу лабораторному и натурному экспериментам. Ранние попытки численного моделирования процесса взаимодействия турбулентного ветра с поверхностными волнами основывались на осреднённых по ансамблю стационарных двумерных уравнениях Навье-Стокса (RANS) [Al-Zanadi, Hui, 1984; Chalikov, 1986; Gent, Taylor, 1976]. При этом неизвестные турбулентные напряжения Рейнольдса выражаются через градиенты среднего поля скорости с использованием коэффициента турбулентной вязкости. Несмотря на привлекательность RANS-моделей и их способность при определённых условиях качественно предсказывать структуру среднего поля скорости ветра над волнами, эти модели имеют ряд недостатков. Пожалуй, наиболее существенным недостатком является неспособность предсказывать свойства воздушного потока вблизи водной поверхности, в вязком подслое и буферной зоне.

В современных численных исследованиях используются вихреразрешающие (Large Eddy Simulations, LES) модели (см., напр., [Sullivan et al., 2008]). В этих моделях рассматриваются трёхмерные нестационарные уравнения Навье-Стокса с разрешением, позволяющим отслеживать динамику крупномасштабных вихревых структур в воздушном потоке над водной поверхностью. Однако и эти модели не позволяют исследовать свойства потока вблизи водной поверхности, поскольку первый узел сетки в вертикальном направлении обычно находится в логарифмическом слое.

По-видимому, единственным методом, способным обеспечить полное детальное разрешение течения в погранслое над водной поверхностью является прямое численное моделирование (Direct Numerical Simulation, DNS). Впервые этот метод был использован в работе [Sullivan et al., 2000] для исследования воздушного нестратифицированного по плотности потока с балковым (основанным на балковой скорости потока, длине поверхностной волны и кинематической вязкости воздуха) числом Рейнольдса $Re = 8800$ и крутизной волны $ka = 0,1$ (где k и a — волновое число и амплитуда волны). Позднее в работе [Yang, Shen, 2010] используется DNS для исследования воздушного потока над волнами с крутизной до $ka = 0,25$ и числом Рейнольдса $Re \approx 10\,000$. Метод DNS даёт полное разрешение течения вплоть до масштаба вязкой диссипации. Недостаток этого подхода — ограничение сверху по числу Рейнольдса (порядка 10^5). В работе [Druzhinin et al., 2012] метод DNS использовался для численного моделирования нестратифицированного потока с числом Рейнольдса $Re = 15\,000$ и максимальной крутизной волн $ka = 0,2$. Таким образом, число Рейнольдса здесь почти вдвое больше Re в работе [Sullivan et al., 2000] ($Re = 8800$) и на 50 % больше $Re = 9943$ в работе

[Yang, Shen, 2010], в то время как крутизна близка к максимальной крутизне $ka = 0,25$, задаваемой в работе [Yang, Shen, 2010].

В данной статье рассматривается устойчиво стратифицированный воздушный поток с балковым числом Рейнольдса в интервале от $Re = 15\ 000$ до $Re = 80\ 000$ и максимальной крутизной волны $ka = 0,2$ при различных числах Ричардсона (основанного на балковой разности температур на поверхности воды и воздуха, балковой скорости потока и длине волны). Насколько нам известно, до сих пор прямое численное моделирование воздушного потока над волнами при условиях устойчивой стратификации не проводилось.

1. ФОРМУЛИРОВКА ЗАДАЧИ

Проводится прямое численное моделирование (DNS) устойчиво стратифицированного турбулентного течения Куэтта над взволнованной водной поверхностью. Схема численного эксперимента аналогична схеме, рассматриваемой в работе [Druzhinin et al., 2012], где проводилось DNS турбулентного нестратифицированного погранслоя (рис. 1). Рассматривается декартова система координат с осью x , ориентированной в направлении среднего воздушного потока, осью z , направленной вертикально вверх (против ускорения свободного падения g), и осью y в поперечном направлении, параллельном фронту волны. Рассматривается двумерная поверхностная волна с амплитудой a , длиной λ и фазовой скоростью c , периодическая в направлении x . В данной работе рассматривается крутизна волны от нуля до максимального значения $ka = 2\pi a/\lambda = 0,2$. DNS проводится в системе отсчёта, движущейся с фазовой скоростью волны, т.е. горизонтальная координата x в этой системе отсчёта связана с координатой в лабораторной системе отсчёта x' соотношением $x = x' - ct$. Таким образом, в движущейся системе отсчёта нижняя граница области счёта, совпадающая с волной, стационарна. На этой границе ставится условие прилипания, т.е. x - и z -компоненты скорости воздушного потока здесь совпадают с орбитальными скоростями частиц жидкости в волне и не зависят от времени. Рассматривается область счёта с размерами $L_x = 6\lambda$, $L_y = 4\lambda$ и $L_z = \lambda$ в x -, y -, и z -направлениях. Течение предполагается периодическим в направлениях x и y .

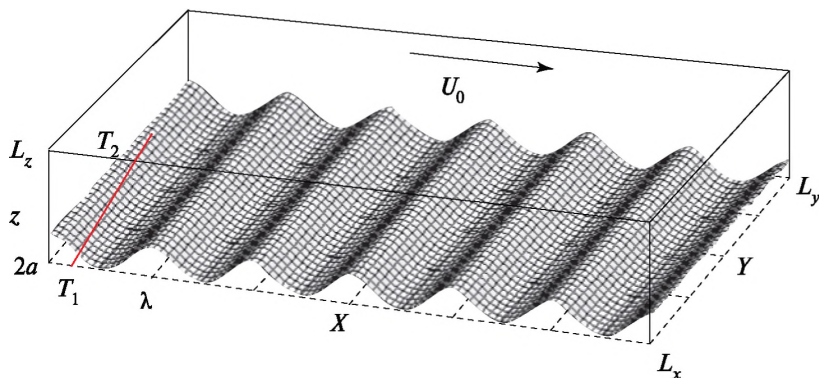


Рис. 1. Схема численного эксперимента

Ставится условие прилипания на верхней (плоской) границе, которая движется с балковой скоростью U_0 в направлении x . Благодаря этому условию создаётся внешний источник потока импульса за счёт действия вязкого напряжения сдвига, который компенсирует вязкую диссипацию кинетической энергии турбулентности и обеспечивает стационарность течения. Для того чтобы обеспечить устойчивую стратификацию воздушного потока, задаются граничные условия для температуры в виде $T = T_1$ на нижней границе (совпадающей с поверхностной волной) и $T = T_2$ на верхней (плоской) границе при $z = L_z$, где $T_1 < T_2$.

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

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

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

$$\frac{\partial \tilde{T}}{\partial t} + \frac{\partial(U_j \tilde{T})}{\partial x_j} + U_z = \frac{1}{\text{PrRe}} \frac{\partial^2 \tilde{T}}{\partial x_j \partial x_j}, \quad (3)$$

где U_i ($i = x, y, z$) — компоненты скорости; P — давление; \tilde{T} — отклонение температуры от исходного линейного профиля $(1 + z)$. Таким образом, безразмерная температура равна $T = 1 + z + \tilde{T}$. Переменные в уравнениях (1)–(3) нормированы на длину волны λ , балковую скорость U_0 и разность температур $(T_2 - T_1)$, давление нормировано на ρU_0^2 (где ρ — плотность воздуха). Балковые числа Рейнольдса и Ричардсона записываются в виде

$$\text{Re} = \frac{U_0 \lambda}{\nu}, \quad (4)$$

$$\text{Ri} = g \frac{T_2 - T_1}{T_1} \frac{\lambda}{U_0^2}, \quad (5)$$

где ν — кинематическая вязкость воздуха; g — ускорение свободного падения. Число Прандтля $\text{Pr} = \nu/\kappa$ (где κ — теплопроводность) полагается равным $\text{Pr} = 0,7$. Фактор f в силе Архимеда (последнем слагаемом) в правой части соотношения (1) определяется как

$$f(t) = 1 - \exp\left(-\frac{t}{100}\right), \quad (6)$$

где t — безразмерное время. Этот фактор «включает» силу Архимеда и связанный с ней стабилизирующий эффект стратификации таким образом, что до момента времени $t < 100$ успевает развиваться турбулентное нестратифицированное течение Куэтта.

Для того чтобы избежать эффектов, связанных с сильной геометрической нелинейностью при интегрировании (1)–(3) из-за волнообразной ниж-

ней границы (см. рис. 1), используется конформное преобразование координат, переводящее плоскость (x, z) в декартовых координатах в плоскость (ξ, η) в криволинейных координатах, а именно:

$$x = \xi - a \exp(-k\eta) \sin k\xi, \quad (7)$$

$$z = \eta + a \exp(-k\eta) \cos k\xi. \quad (8)$$

При этом преобразовании волновая граница $z_b(x) = a \cos k\xi(x)$ переходит в плоскую границу $\eta = 0$. В дополнение используется также отображение по вертикальной координате η , обеспечивающее достаточное разрешение больших градиентов полей скорости и температуры благодаря сгущению узлов сетки вблизи нижней и верхней границ [Druzhinin et al., 2012].

Уравнения (1)–(3) дискретизируются в прямоугольной области с размерами $0 < \xi < 6$, $0 < \eta < 4$ и $0 < \eta < 1$ при помощи метода конечных разностей второго порядка на разнесённой сетке, состоящей из $360 \times 240 \times 180$ узлов. Интегрирование по времени осуществляется с использованием метода Адамса-Башфорфа второго порядка точности в две стадии на каждом временном шаге n .

На нижней и верхней границах накладываются условия прилипания (Дирихле) для скорости воздушного потока и нулевое условие для отклонений температуры. Таким образом, скорость воздушного потока на нижней границе совпадает со скоростью частиц в поверхностной волне, а на верхней границе — со скоростью $(1 - c)$ в направлении движения границы. На боковых границах, при $x = 0, 6$ и $y = 0, 4$, на все поля накладываются периодические граничные условия.

Течение инициируется как слабо возмущённое ламинарное течение Куэтта, $U_i = z\delta_{ix} + u_{if}$ ($i = x, y, z$) где u_{if} — бездивергентное изотропное, однородное случайное поле с широким спектром и амплитудой порядка 0,05. Начальное отклонение температуры полагается равным нулю. Интегрирование во времени сначала проводится без набора статистики до окончания переходных процессов, связанных с выбором начального условия. По достижении статистически стационарного режима течения проводится расчёт осреднённых по времени и координате y полей. Для усреднения по времени берутся поля скорости и температуры в дискретные моменты t_k , $k = 1, \dots, 1000$ с интервалом $t_{k+1} - t_k = 0,2$. Помимо осреднения по y и t рассчитываются также фазовые (оконовые) средние по длине волны.

Таким образом, для данного поля $f(x, y, z, t)$ определяются фазовое среднее и дисперсия:

$$\langle f \rangle(x, z) = \frac{1}{6N_t N_y} \sum_{j=1}^{N_y} \sum_{k=1}^{N_t} \sum_{n=0}^5 f(x + n\lambda, y_j, z, t_k), \quad (9)$$

$$\langle f^2 \rangle(x, z) = \frac{1}{6N_t N_y} \sum_{j=1}^{N_y} \sum_{k=1}^{N_t} \sum_{n=0}^5 f^2(x + n\lambda, y_j, z, t_k), \quad (10)$$

где $N_y = 240$, $N_t = 500$ и $0 < x < 1$. Далее флуктуация поля f определяется как:

$$f'(x, z) = \left(\langle f^2 \rangle(x, z) - \langle f \rangle^2(x, z) \right)^{1/2}. \quad (11)$$

(В случае расчёта потока, например $\langle U_x U_z \rangle$, в соотношение (10) подставляется произведение $(U_x U_z)$ вместо f^2 , и в (11) — разность $(\langle U_x U_z \rangle - \langle U_x \rangle \langle U_z \rangle)$.)

Средний вертикальный профиль поля f , $[f](z)$, рассчитывается осреднением фазового среднего по продольной координате:

$$[f](z) = \frac{1}{N_x} \sum_{i=1}^{N_x/6} \langle f \rangle(x_i, z). \quad (12)$$

Таким образом, вертикальные турбулентные потоки импульса и тепла, $\tau_m(z)$ и $\tau_h(z)$, определяются в виде

$$\tau_m(z) = [\langle U_x U_z \rangle - \langle U_x \rangle \langle U_z \rangle], \quad (13)$$

$$\tau_h(z) = [\langle U_x \tilde{T} \rangle - \langle U_x \rangle \langle \tilde{T} \rangle]. \quad (14)$$

Динамическая скорость и температура, u_* и T_* , определяются из асимптотических значений потоков импульса и тепла на достаточно большом расстоянии от поверхности в виде

$$u_* = \sqrt{\tau_m}, \quad T_* = \tau_h / u_*. \quad (15)$$

Численный алгоритм верифицировался путём сравнения результатов DNS для устойчиво стратифицированного турбулентного погранслоя над плоской поверхностью при различных балковых числах Ричардсона и числах Рейнольдса до $Re = 80\,000$. Результаты DNS сравнивались с предсказанием теории подобия Монино-Обухова, при этом проверялся баланс бюджета потоков импульса и тепла и кинетической и потенциальной энергии. Результаты также сравнивались с результатами предыдущих исследований с применением DNS [Flores, Riley, 2011].

4. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Численное моделирование устойчиво стратифицированного погранслоя (УСПС) над взволнованной водной поверхностью проводилось для крутизны волны в интервале $0 < ka < 0,2$, отношения фазовой скорости к динамической скорости (возраста волны) в интервале $0 < c/u_* < 10$, и балковых числах Рейнольдса и Ричардсона в интервалах $15\,000 < Re < 80\,000$ и $0 < Ri < 0,3$. Результаты DNS показывают, что режим течения коренным образом зависит от чисел Re , Ri и крутизны волны ka . Для заданных Re и ka существует пороговое число Ричардсона, Ri_c , такое, что при $Ri < Ri_c$ устанавливается стационарный турбулентный режим, а при $Ri > Ri_c$ турбулентные флуктуации затухают благодаря стабилизирующему эффекту стратификации и вязкой диссипации.

На рис. 2 можно проследить поведение амплитуд флуктуаций скорости и температуры по результатам DNS УСПС для балкового числа Рейнольдса $Re = 15\,000$ при изменении числа Ричардсона при фиксированной крутизне волны $ka = 0,2$, а также при изменении ka при фиксированном Ri . На рис. 2 видно, что турбулентный режим подавляется при достаточно больших Ri и малых ka .

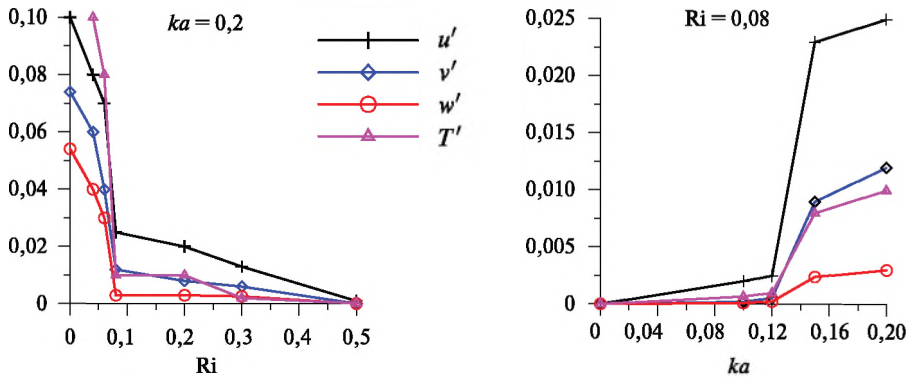


Рис. 2. Зависимость амплитуды флуктуации скорости и температуры в устойчиво стратифицированном погранслое от числа Ричардсона Ri при фиксированной крутизне волны ($ka = 0,2$) (слева) и от крутизны волны при фиксированном $Ri (= 0,08)$. Балковое число Рейнольдса $Re = 15\,000$

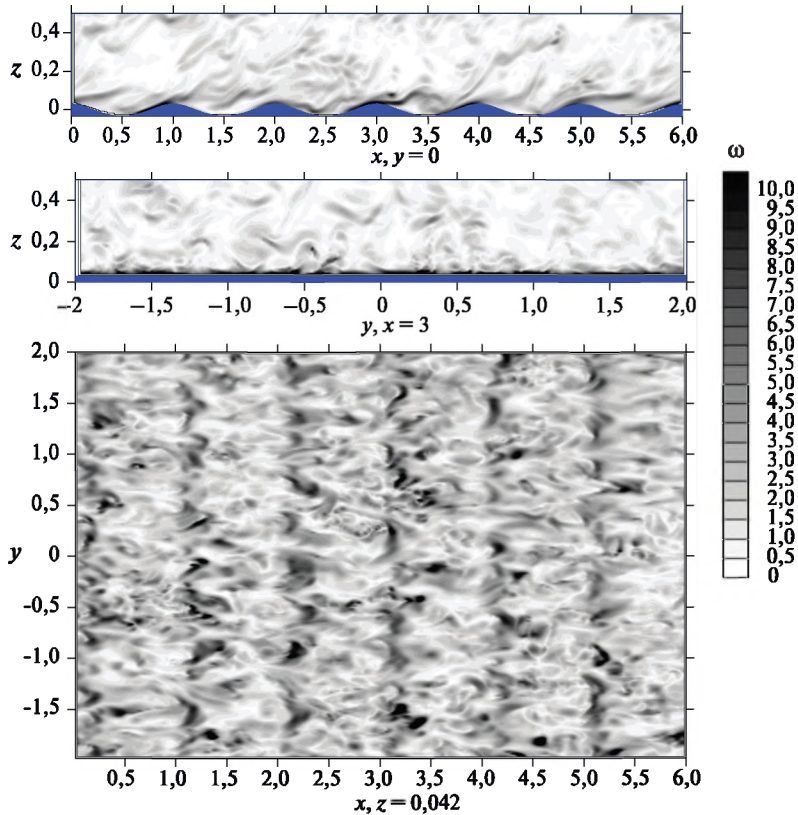


Рис. 3. Распределение модуля завихрённости в центральных плоскостях (x, z) и (y, z) (вверху и в середине) и в плоскости (x, y) при $z = 0,042$ ($u_* z / \nu \approx 18$, внизу) в DNS УСПС над поверхностной волной с крутизной $ka = 0,2$ и фазовой скоростью $c = 0,05$ ($c/u_* \approx 2$) в момент времени $t = 1000$; балковые числа $Re = 15\,000$ и $Ri = 0,04$. Турбулентный режим

Необходимо отметить, что эффект подавления турбулентности в устойчиво-стратифицированном погранслое уже наблюдался в численном и лабораторном моделировании [Flores, Riley, 2011] для случая плоской поверхности. Эти результаты показывают, что подавление турбулентности не происходит, и погранслоем остаётся турбулентным при условии, что число Рейнольдса, базирующееся на масштабе Обухова (L), (где $\kappa = 0,4$ — постоянная Кармана), динамической скорости (u_*) и кинематической вязкости ($\nu = \text{Re}^{-1}$) должно быть достаточно большим.

$$L = \frac{u_*^2}{\kappa \text{Ri} T_*}. \quad (16)$$

$$\text{Re}_L = \frac{Lu_*}{\nu} = \frac{\text{Re} u_*^3}{\kappa \text{Ri} T_*} > 100. \quad (17)$$

Результаты нашего исследования подтверждают критерий (17) для случая плоской границы и волновой границы. На рис. 3 показано поле модуля завихрённости в стационарном турбулентном УСПС над поверхностной волной с фазовой скоростью волной с крутизной $ka = 0,2$ и фазовой скоростью $c = 0,05$ ($c/u_* \approx 2$) в момент времени $t = 1000$; балковые числа $\text{Re} = 15\,000$ и $\text{Ri} = 0,04$. Для рассматриваемых параметров DNS оценка (17) даёт $\text{Re}_L \approx 460$, и наблюдается стационарный турбулентный режим. На рис. 4 приведены вертикальные профили турбулентных потоков импульса и тепла, τ_m и τ_h , для тех же параметров потока и поверхностной волны с различной крутизной ($ka = 0; 0,1$ и $0,2$). На рис. 4 видно, что с ростом амплитуды волны турбулентные потоки растут.

При достаточно больших числах Ричардсона, когда критерий (17) не выполняется, флуктуации скорости и температуры в УСПС подавляются стратификацией и диссипацией. Результаты DNS, однако, показывают, что в рассматриваемом случае погранслоя над водной поверхностью флуктуации скорости и температуры остаются конечными вблизи волновой поверхности, даже при достаточно сильной стратификации (при $\text{Ri} > \text{Ri}_c$ и невыполнении условия (17)). Мы предлагаем называть этот режим течения в устойчиво стратифицированном погранслое «режимом волновой накачки».

На рис. 5 приведено мгновенное распределение модуля завихрённости в DNS УСПС над поверхностной волной с теми же параметрами фазовой скорости и крутизны ($ka = 0,2$, $c = 0,05$) и балковым числом Рейнольдса $\text{Re} = 15\,000$, но при числе Ричардсона $\text{Ri} = 0,08$. Показаны центральные плоскости (x, z) и (y, z) и плоскость (x, y) при $z = 0,12$ в момент времени $t = 1000$. В этом случае реализуется режим волновой накачки: течение реламинаризируется в области, достаточно далёкой от поверхностной волны, и остаётся квазитурбулентным вблизи волновой поверхности. На рис. 5 видно, что под действием ненулевых возмущений скорости, обусловленных граничным условием на поверхности волны, развиваются двумерные волновые возмущения в воздушном потоке. Эти возмущения «подпитывают» развитие вторичной неустойчивости, приводящей к формированию поперечных структур. Таким образом, динамика УСПС в режиме волновой накачки является результатом нелинейных волновых взаимодействий.

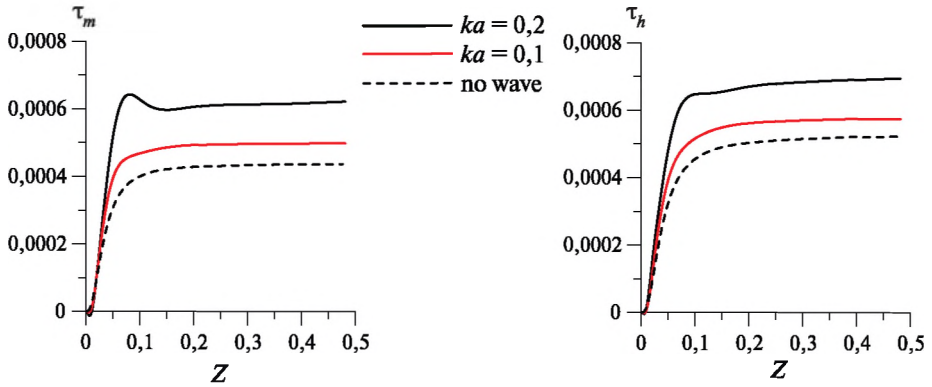


Рис. 4. Вертикальные профили турбулентных потоков импульса τ_m (слева) и тепла τ_h (справа) в DNS УСПС над поверхностной волной с фазовой скоростью $c = 0,05$, балковыми числами $Re = 15\,000$ и $Ri = 0,04$ при различной крутизне волн ($ka = 0; 0,1$ и $0,2$)

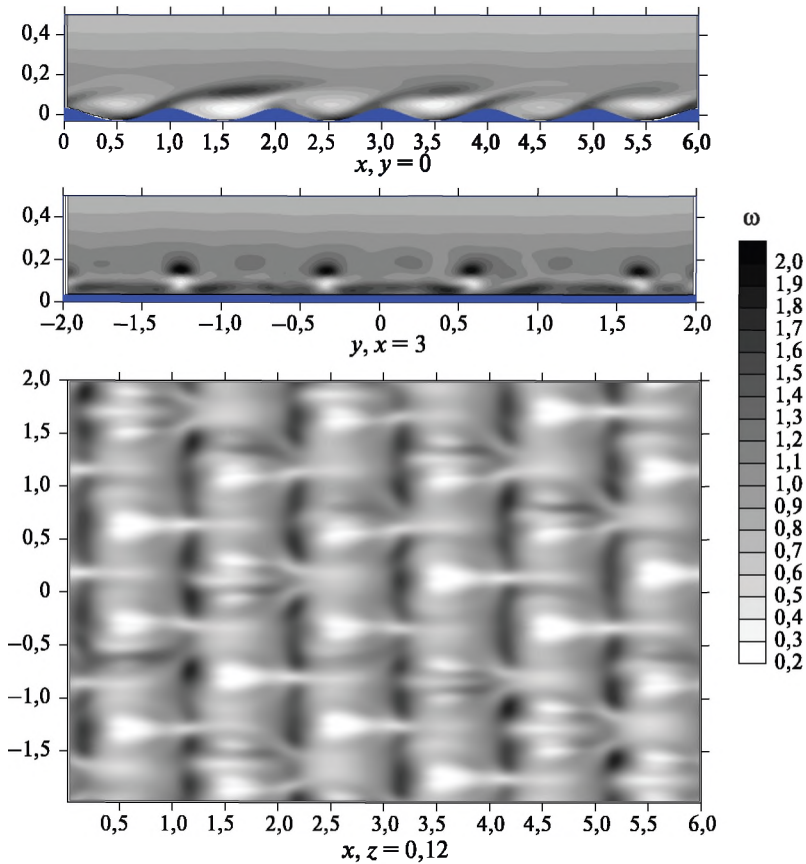


Рис. 5. Мгновенное поле модуля завихренности в режиме волновой накачки устойчиво-стратифицированного пограничного слоя. Показаны центральные плоскости (x, z) и (y, z) (вверху и посередине) и плоскость (x, y) при $z = 0,12$ (внизу) в момент времени $t = 1000$ при $Re = 15\,000$, $Ri = 0,08$. Крутизна волны $ka = 0,2$, фазовая скорость $c = 0,05$. Режим волновой накачки

Амплитуды двумерных возмущений, подпитываемые поверхностной волной за счёт граничного условия (прилипания) на водной поверхности, могут быть вычислены аналитически с использованием теоретической модели УСПС над взволнованной водной поверхностью, приведённой в статье [Troitskaya et al., 2013a, b]. Модель базируется на системе уравнений (1)–(3) осреднённых по координате y и переписанных в криволинейных координатах (7), (8). Соответствующая двумерная система уравнений переписывается в переменных «завихрённость – функция тока» (χ - ψ) в виде [Troitskaya et al., 2013a, b]:

$$ik(\psi_{0\eta}\chi_1 - \psi_1\chi_{0\eta}) - ikRiT_1 + ikaRi \frac{dT_0}{d\eta} e^{-k\eta} = \frac{1}{\text{Re}} \left[\frac{d^2\chi_1}{d\eta^2} - k^2\chi_1 \right], \quad (18)$$

$$\psi_{1\eta\eta} - k^2\psi_1 = \chi_1 - 2ka e^{-k\eta}\chi_0, \quad (19)$$

$$ik(\psi_{0\eta}T_1 - \psi_1T_{0\eta}) = \frac{1}{\text{RePr}} \left[\frac{d^2T_1}{d\eta^2} - k^2T_1 \right]. \quad (20)$$

Здесь завихрённость χ , функция тока ψ , и температура T представлены в виде суммы средних полей и возмущений в виде:

$$\chi = \chi_0(\eta) + \text{Re} \left[\chi_1(\eta) e^{ik\xi} \right]; \quad \chi_0(\eta) = U_{0\eta}(\eta), \quad (21)$$

$$\psi = \psi_0(\eta) + \text{Re} \left[\psi_1(\eta) e^{ik\xi} \right]; \quad \psi_0(\eta) = \int (U_0(\eta) - c) d\eta, \quad (22)$$

$$T = T_0(\eta) + \text{Re} \left[T_1(\eta) e^{ik\xi} \right]. \quad (23)$$

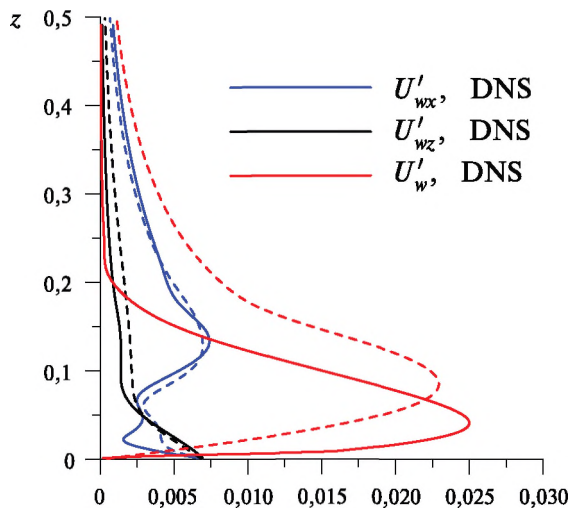


Рис. 6. Возмущения полей скорости и температуры, индуцированные поверхностной волной в устойчиво стратифицированном погранслое в режиме волновой накачки. Результаты DNS (сплошные линии) сравниваются с предсказанием теоретической модели (пунктирные линии)

Система (21)–(23) решается численно с граничными условиями в виде:

$$\psi_1|_{\eta=0} = 0, \quad \frac{\partial \psi_1}{\partial \eta} \Big|_{\eta=0} = 2cka, \quad T_1|_{\eta=0} = 0. \quad (24)$$

Численное решение модельных уравнений (21)–(23) для вертикальных профилей возмущений скорости и температуры сравнивается с результатами DNS УСПС на рис. 6 для значений параметров $Re = 15\,000$, $Ri = 0,08$, $ka = 0,2$, $c = 0,05$. Профили возмущений средних полей, индуцированных поверхностной волной в DNS, рассчитываются как

$$U'_{iw} = [\langle U_i \rangle^2] - [\langle U_i \rangle]^2, \quad T'_w = [\langle T \rangle^2] - [\langle T \rangle]^2, \quad (25)$$

где $i = x, z$. На рис. 6 видно хорошее согласие предсказаний модели (21)–(23) с результатами DNS. Наблюдающиеся различия, по-видимому, вызваны эффектами, связанными с развитием вторичных неустойчивостей течения, обуславливающих неоднородность возмущений по фронту волны, в y -направлении, которые модель (21)–(23) не учитывает.

ЗАКЛЮЧЕНИЕ

Выполнено прямое численное моделирование (DNS) устойчиво стратифицированного пограничного слоя (УСПС) над взволнованной водной поверхностью. Результаты показывают, что режим течения коренным образом зависит от балковых чисел Рейнольдса (Re) и Ричардсона (Ri) и от крутизны поверхностной волны (ka). Для заданного Re существует пороговое число Ричардсона, Ri_c , такое, что при $Ri < Ri_c$ устанавливается стационарный турбулентный режим, а при $Ri > Ri_c$ турбулентные флуктуации затухают в области, достаточно далёкой от водной поверхности, благодаря стабилизирующему эффекту стратификации и вязкой диссипации. Однако вблизи поверхности амплитуды флуктуаций скорости и температуры остаются конечными (ненулевыми). Мы называем этот режим течения в УСПС «режимом волновой накачки». Как в случае стационарного турбулентного режима, так и в случае режима волновой накачки потоки импульса и тепла и амплитуды флуктуаций скорости и температуры растут с ростом крутизны волны ka . Результаты DNS хорошо согласуются с предсказанием теоретической модели течения УСПС, где решения для возмущений полей скорости и температуры в режиме волновой накачки находятся как решение двумерной линеаризованной системы с неоднородным граничным условием, обусловленным наличием поверхностной волны.

ЛИТЕРАТУРА

- [Монин, Яглом, 1992] Монин А. С., Яглом А. М. Статистическая гидромеханика. Т. 1. СПб.: Гидрометеиздат, 1992.
- [Adrian, 1991] Adrian R. J. Particle Imaging techniques for experimental fluid dynamics // Annual Review of Fluid Mechanics. 1991. V. 23. P. 261–304.

- [Al-Zanaidi, Hui, 1984] *Al-Zanaidi M. A., Hui W. H.* Turbulent airflow over water waves — a numerical study // *J. Fluid Mechanics*. 1984. V. 148. P. 225–246.
- [Chalikov, 1986] *Chalikov D. V.* Numerical simulation of the boundary layer above water waves // *Boundary Layer Meteorology*. 1986. V. 34. P. 63–98.
- [Donelan et al., 2005] *Donelan M. A., Babanin A. V., Young I. R.* et al. Wave follower field measurements of the wind input spectral function. Pt. I: Measurements and calibrations // *J. Atmospheric and Oceanic Technology*. 2005. V. 22. P. 799–813.
- [Druzhinin et al., 2012] *Druzhinin O. A., Troitskaya Yu. I., Zilitinkevich S. S.* Direct numerical simulation of a turbulent wind over a wavy water surface // *J. Geophysical Research*. 2012. V. 117. C00J05. doi: 10.1029/2011JC007789.
- [Fairall et al., 2003] *Fairall C. W., Bradley E. F., Hare J. E.* Bulk parameterization of air-sea fluxes: updates and verification for the COARE algorithm // *J. Climate*. 2003. V. 16. P. 571–591.
- [Flores, Riley, 2011] *Flores O., Riley J. J.* Analysis of turbulence collapse in the stably stratified surface layer using direct numerical simulation // *Boundary-Layer Meteorology*. 2011. V. 139. P. 241–259. doi: 10.1007/s10546-011-9588-2.
- [Gent, Taylor, 1976] *Gent P. R., Taylor P. A.* A numerical model of the air flow above water waves // *J. Fluid Mechanics*. 1976. V. 77. P. 105–128.
- [Hsu, Hsu, 1983] *Hsu C. T., Hsu E. Y.* On the structure of turbulent flow over a progressive water wave: theory and experiment in a transformed wave-following coordinate system. Pt. 2 // *J. Fluid Mechanics*. 1983. V. 131. P. 123–153.
- [Hsu et al., 1981] *Hsu C. T., Hsu E. Y., Street R. L.* On the structure of turbulent flow over a progressive water wave: theory and experiment in a transformed, wave-following co-ordinate system // *J. Fluid Mechanics*. 1981. V. 105. P. 87–117.
- [Reul et al., 1999] *Reul N., Branger H., Giovanangeli J.-P.* Air flow separation over unsteady breaking waves // *Physics of Fluids*. 1999. V. 11. P. 1959–1961.
- [Sullivan et al., 2000] *Sullivan P. P., McWilliams J. C., Moeng C.-H.* Simulation of turbulent flow over idealized water waves // *J. Fluid Mechanics*. 2000. V. 404. P. 47–85.
- [Sullivan et al., 2008] *Sullivan P. P., Edson J., Hristov T., McWilliams J. C.* Large-eddy simulations and observations of atmospheric marine boundary layers above nonequilibrium surface waves // *J. Atmospheric Sciences (JAS)*. 2008. V. 65. P. 1225–1245.
- [Troitskaya et al., 2011] *Troitskaya Yu., Sergeev D., Ermakova O., Balandina G.* Statistical Parameters of the Air Turbulent Boundary Layer over Steep Water Waves Measured by the PIV Technique // *J. Physical Oceanography*. 2011. V. 41. P. 1421–1454.
- [Troitskaya et al., 2013a] *Troitskaya Yu. I., Ezhova E. V., Zilitinkevich S. S.* Momentum and buoyancy transfer in atmospheric turbulent boundary layer over wavy water surface. Pt. 1: Harmonic wave // *Nonlinear Processes in Geophysics*. 2013. V. 20. doi: 10.5194/npg-20-1-2013.
- [Troitskaya et al., 2013b] *Troitskaya Yu. I., Ezhova E. V., Sergeev D. A., Kandaurov A. A., Baidakov G. A., Vdovin M. I., Zilitinkevich S. S.* Momentum and buoyancy transfer in atmospheric turbulent boundary layer over wavy water surface. Pt. 2. Wind wave spectra // *Nonlinear Processes in Geophysics*. 2013. V. 20. doi: 10.5194/npg-20-1-2013.
- [Veron et al., 2007] *Veron F., Saxena G., Misra S. K.* Measurements of the viscous tangential stress in the airflow above wind waves // *Geophysical Research Letters*. 2007. V. 34. L19603. doi: 10.1029/2007GL031242.
- [Yang, Shen, 2010] *Yang D., Shen L.* Direct-simulation-based study of turbulent flow over various waving boundaries // *J. Fluid Mechanics*. 2010. V. 650. P. 131–180.

DIRECT NUMERICAL SIMULATION STUDY OF A TURBULENT STABLY STRATIFIED AIR FLOW ABOVE THE WAVY WATER SURFACE.

O. A. Druzhinin¹, Yu. I. Troitskaya¹, S. S. Zilitinkevich²

¹ Institute of Applied Physics Russian Academy of Sciences (IAP)

² Finnish Meteorological Institute (FMI), Helsinki, Finland

Parameterization of turbulent momentum and heat fluxes in a turbulent stably stratified boundary layer over a wavy water surface is important for large-scale climate modeling and weather forecast. The relation between mean large-scale velocity and temperature fields and turbulent fluxes at the ocean-atmosphere interface includes the so-called bulk coefficients. These coefficients vary in a wide range in different models, and their evaluation constituted a complicated task for field and laboratory experiments and numerical simulation. We employ direct numerical simulation (DNS) as an instrument to study in detail the dynamics of the ocean-atmosphere boundary layer in order to further develop and verify parameterization procedure for large-scale climate modeling. Especially we are interested in the effects caused by the stable stratification of the air flow.

Keywords: numerical simulation, turbulent wind, two-dimensional water waves.

Druzhinin Oleg Aleksandrovich — leading researcher, doctor of physical and mathematical sciences, druzhinin@hydro.appl.sci-nnov.ru

Troitskaya Yuliya Igorevna — head of the department, doctor of physical and mathematical sciences, yuliya@hydro.appl.sci-nnov.ru

Zilitinkevich Sergei Sergeevich — head of the laboratory, doctor of physical and mathematical sciences, sergej.zilitinkevich@fmi.fi

ПРОБЛЕМЫ ПОСТРОЕНИЯ СИСТЕМ ЗАЩИТЫ ОТ СПАМА В ИНТЕРНЕТЕ

И. О. Карбачинский

Национальный исследовательский университет «Высшая школа экономики»
(НИУ ВШЭ)

Трафик в поисковых системах играет одну из важнейших ролей в онлайн экономике. Поэтому неудивительно, что за последние годы количество спама в Интернете неуклонно растёт. Ежедневно в сети появляются тысячи новых сайтов, содержащих некачественный контент. В статье детально рассматривается устройство защиты от спама в современных информационно-поисковых системах по интернет-пространству (веб). Рассматриваются проблемы построения спам-классификатора, проблемы переобучения и отбора признаков. Рассматриваются способы применения методов уменьшения размерности пространства для улучшения качества классификатора, а также различные способы визуализации многомерных данных. Особое внимание уделено отбору информативных признаков и оптимизации скорости работы классификатора.

Ключевые слова: веб-спам, информационный поиск, критерий MRMR, уменьшение размерности пространства, Метод главных компонент, классификация спама

Ежедневно в сети появляются тысячи новых сайтов, содержащих спам и некачественный контент. Существует множество различных видов спама. Ежедневно мы сталкиваемся с бесполезными рассылками в наших почтовых ящиках, смс-спамом (sms — Short Message Service, служба коротких сообщений) и спамом в выдаче поисковых систем. В данной работе проблема борьбы со спамом будет детально рассмотрена на примере спама в информационно-поисковых системах. Под спамом будут подразумеваться веб-страницы в выдаче поисковой системы, содержащие вредоносные, некачественные и бесполезные для пользователя данные. Система защиты от спама, независимо от типа спама, как правило, представляет собой бинарный классификатор, обученный на некотором обучающем множестве, в нашем случае множестве веб-документов. Обучающее множество представляет собой размеченный набор веб-страниц, где каждая страница отмечена маркером принадлежности к классу спама, или не спама. Составить обучающее множество можно несколькими способами. Самый очевидный способ — сделать это вручную. Для составления обучающего множества в поисковых системах существуют специальные люди, называемые ассессорами. Их работа позволяет составить качественное обучающее множество веб-документов, которое впоследствии может использоваться для обучения классификатора.

Для обучения классификатора, вне зависимости от используемого алгоритма, необходимо из каждого элемента обучающего множества выделить некоторый набор информативных признаков. Данные признаки позволяют численно охарактеризовать каждый документ в обучающем множестве. Можно придумать очень много таких признаков. Однако не все они будут полезны для обучения классификатора. Выбирать признаки следует так, чтобы их значения по возможности отличались на множествах спам- и не спам-документов. Чем лучше множество признаков будет разделять обучающее множество на два класса, тем проще будет построить классификатор. Очевидно, что интуитивное составление множества признаков не может дать

гарантированно хороший результат. Поэтому в данной работе мы рассмотрим несколько алгоритмов, позволяющих осуществить отбор наиболее релевантных признаков. Подобные алгоритмы позволяют автоматизировать процесс отбора признаков, гарантированно давая хороший результат. Итогом работы подобных алгоритмов, как правило, является упорядоченный по «релевантности» список признаков, что позволяет выбрать необходимое количество лучших признаков.

Составив обучающее множество и выделив необходимые признаки, следует разделить размеченное обучающее множество на два подмножества. Первое из них будет использоваться для обучения классификатора, а второе для проверки качества его работы. Существует множество различных алгоритмов обучения. Каждый из них имеет свои преимущества и недостатки. Выбирать алгоритм стоит исходя из характера обучающего множества и распределения значений информативных признаков на нём. В простых задачах, как правило, достаточно применения одного из известных алгоритмов (Knn, SVM, линейная регрессия). Для более сложных задач, как правило, совместно используются сразу несколько алгоритмов, например, с применением технологии бустинга. Для построения нашего классификатора мы используем бустинг деревьев решений и двухслойный персептрон, при размере обучающего множества, состоящего из нескольких сотен тысяч документов и 150 информативных признаков.

Построив классификатор, необходимо оценить его качество. Для оценки бинарного классификатора используется специальная таблица, называемая Confusion Matrix [Fawcett, 2006]. Таблица предоставляет четыре возможных исхода работы классификатора: TP — документ оказался спам документом, и классификатор правильно определил данный документ; FP — документ оказался хорошим, но классификатор посчитал его спамом; FN — документ оказался спамом, но классификатор посчитал его хорошим; TN — документ оказался хорошим, и классификатор правильно определил его класс. В зависимости от того, какой классификатор мы строим, необходимо по-разному оценивать ошибки классификации. Например, в задачах классификации почтового спама любой FP-случай может крайне негативно сказаться на репутации почтового сервиса, так как пользователь может не получить важное письмо, не являющееся спамом. В то же время большое количество FN-случаев для поисковой системы означает, что по большей части, выдача этой поисковой системы состоит из спама. На основе Confusion Matrix (см. рисунок) вводятся основные числовые критерии, позволяющие оценить качество классификатора: точность (P), полнота (R) и F1-мера ($F1$), являющаяся обобщённой характеристикой полноты и точности.

Задача построения спам-классификатора в интернете, как правило, осложняется большим объёмом обучающей выборки и большим количеством информативных признаков. Как правило, это приводит к очень медленной скорости обучения. Порой процесс построения модели классификатора может длиться около суток. Как правило, подобные ограничения не позволяют быстро построить качественный классификатор и экспериментировать с выбором обучающего алгоритма. Если удалось построить классификатор с приемлемым качеством с первого раза, — это, скорее, везение. Приходится делать не одну итерацию обучения и отбора признаков, чтобы добиться должного качества.

		Реальный класс	
		Позитивный (1)	Негативный (0)
Предсказанный класс	Позитивный (1)	True positive (1; 1)	False positive (1; 0)
	Негативный (0)	False negative (0; 1)	True negative (0; 0)

Рис. 1: Таблица Confusion Matrix:

$$\text{точность } F = \frac{TP}{TP + FP}; \text{ полнота } R = \frac{TP}{TP + FN}; F1\text{-мера } FI = \frac{2PR}{P + R}$$

Причинами плохого качества могут являться как плохо составленное обучающее множество или плохой алгоритм обучения, так и просто плохо подобранные признаки. Рассмотрим подробнее проблему отбора признаков. Самым очевидным способом отбора признаков является прямой перебор. Если построить модели на всех всевозможных подмножествах множества признаков и измерить их качество, то останется просто выбрать то подмножество признаков, на котором достигается наибольшее качество классификатора. Но в реальных задачах, где процесс построения модели длится довольно долго, а размер множества признаков достигает сотни признаков, использование подобного подхода становится невозможным. Существует множество различных оптимизаций данного подхода, позволяющих значительно сократить количество необходимых построений модели. Однако даже такие оптимизации не позволяют быстро и эффективно отобрать оптимальное множество признаков в условиях большого количества признаков.

Для отбора признаков существует целый класс алгоритмов, объединённых общим названием Feature Selection. Как правило, все эти алгоритмы итеративные и разделяются на три класса: Wrapper, Filter и Embedded-методы. Wrapper-методы на каждой своей итерации заново обучают классификатор, и в условиях длительного процесса обучения не применимы. Embedded методы — алгоритмы, являющиеся непосредственно частью процесса обучения. Например, деревья решений используют в процессе построения встроенный механизм отбора признаков. Filter-методы — эвристические алгоритмы, использующие аппарат теории информации и математической статистики для определения «полезности» конкретного признака, в контексте распределения его значений на элементах обучающего множества. В условиях нашей задачи наиболее подходящими видятся Filter-алгоритмы, так как они не требуют построения модели на каждой итерации и, как следствие, работают очень быстро.

Одним из наиболее эффективных алгоритмов данного класса является критерий MRMR (Maximum Relevance Minimum Redundancy) [Acid et al., 2011]. Данный критерий позволяет найти наиболее релевантные и в то же время наименее избыточные признаки. Главный принцип, лежащий в основе MRMR, заключается в том, что мы не должны использовать признаки, сильно коррелирующие между собой, и в то же время, должна быть принята во внимание их избыточность. Поэтому должны быть отобраны наиболее разнородные и не похожие друг на друга признаки. Основные понятия, лежащие в основе данного критерия, заимствованы из теории информации. Таким образом, избыточность $W_I(U)$ произвольного подмножества признаков $U \in X$, где X — множество всевозможных признаков, используемых в модели, определяется:

$$W_I(U) = \frac{1}{|U|^2} \sum_{x_i, x_j \in U} MI(x_i, x_j),$$

где $x_k \in U$ — вектор значений k -го признака для каждого документа из обучающего множества, $MI(x_i, x_j)$ — взаимная информация признаков x_i и x_j , а релевантность $V_I(U)$ подмножества $U \in X$ относительно класса C определяется:

$$V_I(U) = \frac{1}{|U|} \sum_{x_i \in U} MI(C, x_i).$$

В задаче классификации спама множество C состоит всего из двух классов: спам и не спам. Таким образом, согласно критерию MRMR, наиболее релевантное и при этом наименее избыточное подмножество $\bar{U} \in X$ признаков можно получить согласно следующему правилу:

$$(V_I(U) - W_I(U)).$$

Однако на практике используют модификацию критерия MRMR, позволяющую избежать полного перебора всевозможных подмножеств множества признаков X . Модифицированный критерий позволяет получить упорядоченный по убыванию эффективности список всех признаков. Таким образом, для обучения можно выбрать любое необходимое количество наиболее качественных признаков.

Было проведено несколько экспериментов с классификатором спама поисковой системы go.mail.ru. Обучающее множество представляет собой несколько сотен тысяч веб-документов. Качество классификатора составляло порядка 97 %, и была поставлена задача увеличить качество без увеличения количества признаков, участвующих в обучении, так как увеличение числа признаков может сильно сказаться как на скорости построения модели, так и на скорости классификации. Модель обучается более чем на 150 признаках. Большой объём обучающего множества и довольно долгая скорость обучения не позволяли применить отбор признаков с помощью Wrapper-методов. Поэтому был применён критерий MRMR.

Во множестве из 150 признаков, участвующих в обучении, выделялся наиболее спорный и проблемный класс из 40 признаков. Признаки данного класса представляют собой частоты определённых слов-маркеров в документе. Данные слова являются наиболее характерными для классов спама и не

спама. Например, такие слова, как «скачать», «бесплатно» и «знакомства». Однако эти слова были подобраны вручную. Поэтому было решено составить новый список из 40 слов, которые бы гарантированно были полезны для обучения. Для этого из обучающего множества были отобраны 1000 наиболее частотных слов и посчитаны их частоты для каждого документа из обучающего множества. После этого к полученным данным был применён критерий MRMR, и были отобраны 40 первых слов из полученного рейтинга MRMR. Полученные слова частично пересекались с исходными, однако большинство из них были новыми. Обучив модель с новыми признаками, качество классификации повысилось на 1,2 %. Это заметное улучшение было достигнуто без увеличения количества признаков или изменения алгоритма или параметров обучения.

Существует множество других подходов к оптимизации множества информативных признаков. Например, вместо того, чтобы отбирать признаки, можно попробовать сжать их. Если преобразовать множество признаков большой размерности в множество меньшей размерности, при этом максимально сохранив информацию о расстояниях между объектами в искомом пространстве, можно получить заметный прирост качества классификации. Существует множество подобных алгоритмов. Наиболее известными из них являются алгоритмы PCA [Jolliffe, 1986] и Random Projection [Dasgupta, 2000]. Но, к сожалению, видимый эффект они дают при очень большом размере множества признаков, превышающего десятки тысяч признаков. Поэтому, как правило, такие алгоритмы применяются в задачах классификации и распознавания изображений и видео [Dasgupta, Gupta, 2002; William, Lindenstrauss, 1984].

В данной работе был показан общий подход к построению бинарного классификатора на примере построения спам-классификатора поисковой системы. Построение классификатора — сложный процесс, и не существует универсального алгоритма, который бы гарантированно давал хороший результат. Существует множество различных алгоритмов обучения, и каждый из них может хорошо работать с одними данными и совершенно не работать с другими. Также имеет место проблема переобучения. Очень часто классификатор показывает очень хорошие результаты на обучающем множестве, но, когда его начинают применять к реальным данным, его качество стремительно уменьшается. Для борьбы с переобучением существуют специальные методы регуляризации, налагающие штраф за сложность модели в процессе обучения. Таким образом, регуляризация не позволяет алгоритму чрезмерно «подгоняться» под данные, используемые для обучения. Качество построенной модели зависит от множества факторов, начиная с построения правильного обучающего множества и заканчивая выбором хороших информативных признаков. Добавление новых признаков может также дать как заметное увеличение качества модели, так и не дать вовсе ничего. Рассмотренные алгоритмы Feature selection позволяют оптимизировать процесс добавления новых признаков и выявить, насколько полезным будет новый признак и не будет ли он избыточным в контексте решаемой задачи.

ЛИТЕРАТУРА

- [Acid et al., 2011] *Acid S., Silvia A., de Campos L. M.* Minimum redundancy maximum relevancy versus score-based methods for learning // *Intelligent Systems Design and Applications (ISDA)*. 11th Intern. Conf. 2011. P. 619–623, doi: 10.1109/ISDA.2011.6121724.
- [Dasgupta, 2000] *Dasgupta S.* Experiments with random projection // *Proc. 16th Conf. Uncertainty in Artificial Intelligence (UAI)*. 2000. P. 143–151.
- [Dasgupta, Gupta, 2002] *Dasgupta S., Gupta A.* An elementary proof of the Johnson-Lindenstrauss Lemma // *Random Structures and Algorithms*. 2002. V. 22. N. 1. P. 60–65.
- [Fawcett, 2006] *Fawcett T.* An Introduction to ROC Analysis // *Pattern Recognition Letters*. 2006. V. 27(8). P. 861–874, doi: 10.1016/j.patrec.2005.10.010.
- [Jolliffe, 1986] *Jolliffe I. T.* *Principal Component Analysis*. N. Y.: Springer-Verlag, 1986.
- [William, Lindenstrauss, 1984] *William B. J.; Lindenstrauss J.* Extensions of Lipschitz mappings into a Hilbert space // *Contemporary mathematics*. 1984. V. 26. Iss. 189–206. P. 1–1.1.

PROBLEMS OF BUILDING WEB SPAM PROTECTION SYSTEMS

I. Karbachinsky

National Research University Higher School of Economics

Today web Traffic plays a major role in the online economy. It is not surprising that in recent years the amount of internet spam is growing steadily. Every day appear thousands of new web sites containing low-quality content. This paper provides a detailed analysis of anti-spam protection systems in modern information retrieval systems. We consider the problem of constructing a web spam classifier, the problem of overfitting and feature selection algorithms analysis.

Keywords: distributed information and communication system, free software, scientific infrastructure

Karbachinsky Igor — aspirant, igorkarbachinsky@mail.ru

МНОГОМЕРНОЕ ОБОБЩЕНИЕ МОДЕЛИ DIFFUSION LIMITED AGGREGATION (DLA)

А. Ю. Меньшутин, Л. Н. Щур

Институт теоретической физики им. Л. Д. Ландау Российской академии наук
(ИТФ РАН), Научный центр РАН в Черноголовке

В работе описывается численный алгоритм для построения многомерных структур роста в рамках модели агрегации, ограниченной диффузией — Diffusion Limited Aggregation (DLA). Обсуждаются способы ускорения процесса численного моделирования, методы организации работы с памятью, методы и особенности симуляции случайного блуждания в пространстве размерностью больше 2, а также другие элементы численного алгоритма. Для описания численного эксперимента используется предметно-ориентированный язык.

Работа поддержана Российским научным фондом (РНФ) (проект № 14-21-00158).

Ключевые слова: алгоритм, случайное блуждание, модель DLA, предметно-ориентированный язык.

ВВЕДЕНИЕ

Равновесные критические явления — достаточно хорошо изученная область статистической физики [Stanley, 1999]. Многочисленные теории позволяют вычислять критические индексы, которые характеризуют поведение системы вблизи критической точки. Эти теории также позволяют понять, почему различные системы ведут себя одинаковым образом при приближении к точке фазового перехода. Идея универсальности подобных явлений подтверждается как подобными теориями, так и различными экспериментами.

Неравновесные критические явления, напротив, всё ещё мало изучены. Подобные явления можно разделить на два класса. Первый — это явления, связанные с различными динамическими свойствами [Hohenberg, Halperin, 1977] обычных термодинамических систем, например, скорость релаксации в магнитной системе. Второй — это достаточно обширный класс явлений, встречающихся в природе и проявляющих фрактальные свойства и свойства скейлинга.

В 1981 г. Виттен и Сандер предложили модель, впоследствии названную моделью агрегации, ограниченной диффузией (DLA) [Witten, Sander, 1981], которая описывает образование двумерной структуры посредством простых правил. Правила, которые описывают данную модель — это броуновское случайное блуждание частиц в присутствии зародыша и их последующая агрегации к нему. Повторение этих правил приводит к образованию очень сложных фрактальных объектов. Причём, данная модель не имеет параметра, который требуется варьировать для получения критических свойств, в отличие, например, от перколяционных моделей, где вероятность перколяции p определяет близость системы к критической точке. Отличие такого явного параметра — ключевая особенность моделей второго типа, когда критические свойства

Меньшутин Антон Юрьевич — младший научный сотрудник, кандидат физико-математических наук, may@itp.ac.ru

Щур Лев Николаевич — заведующий отделом, доктор физико-математических наук, профессор, shchur@chg.ru

проявляются в результате некоторого динамического процесса. Подобные объекты относятся к классу самоорганизующихся систем с критическим поведением (SOC) [Bak et al., 1987]. Считается, что критическая точка в таких объектах является аттрактором, и конечные предельные свойства системы не зависят от начальных условий. Изучение подобных систем — важная задача теоретической физики.

В наших ранних работах [Menshutin, Shchur, 2006; Menshutin et al., 2007, 2008] были разработаны эффективный двумерный алгоритм, а также различные методы построения и анализа кластеров в рамках DLA. Основная сложность при изучении подобных объектов заключается в необходимости генерации чрезвычайно больших ансамблей кластеров размерами в десятки миллионов частиц. Большая статистика и большие размеры кластеров необходимы для получения корректных значений показателей, например, фрактальной размерности. В случае, если размерность пространства больше 2, представляет интерес как изучение асимптотических значений фрактальной размерности кластеров, так и оценка влияния анизотропных правил прилипания частиц на форму генерируемых структур. Теоретические предсказания Халси и Либига [Halsey, Leibig, 1992] о стремлении фрактальной размерности к величине $D = d - 1$, где d — размерность пространства, могут также изучаться численно. Таким образом, цель данной статьи — дать описание математически корректного алгоритма построения кластеров DLA для размерностей пространства больше $d > 2$.

АЛГОРИТМ

Описание алгоритма многомерной модели DLA будет выполнено как обобщение двумерного алгоритма. Элементы алгоритма, специфичные для многомерных обобщений, будут описаны отдельно.

Исходный алгоритм, предложенный в оригинальной работе [Witten, Sander, 1981], состоит из следующих шагов:

- 1) на квадратной двумерной решётке в начале координат находится частица-зародыш (занятая клетка);
- 2) вдали от кластера (от зародыша) рождается новая частица;
- 3) новая частица блуждает случайным образом;
- 4) если частица подходит вплотную к занятой клетке, то она прилипает;
- 5) если частица уходит достаточно далеко от кластера, она уничтожается;
- 6) повтор начиная с шага 3 до тех пор, пока частица не прилипнет или не погибнет, после чего запуск новой частицы (шаг 2).

Как и большинство объектов класса SOC, кластеры DLA описываются в виде алгоритма. Можно сказать, что процесс образования случайной структуры описан исходя из «первых принципов», при этом выделен минимально необходимый набор правил, приводящий к нетривиальным результатам. Простота формулировки правил роста подобных структур в виде указанного алгоритма приводит к тому, что построение теоретических моделей в более общих терминах, описывающих характеристики готовой структуры (например, в терминах корреляционной функции плотность-плотность) — чрезвычайно сложная задача.

В форме, изложенной выше, данный алгоритм требует доопределения, а именно, необходимо определить области, где происходит рождение и уничтожение частиц. Чаще всего их задают окружностями радиуса R_b (окружность рождения) и R_d (окружность уничтожения), причём необходимо, чтобы $R_d \gg R_b$. Новые частицы рождаются в случайном месте на окружности рождения. В свою очередь, частицы, вышедшие за R_d , — уничтожаются. Для формулировки алгоритма с корректным учётом физических свойств случайного блуждания необходимо вспомнить свойства случайного блуждания на плоскости, а именно свойство возвратности случайного блуждания в 2D. В терминах модели DLA данное свойство может быть сформулировано в следующем виде: любая частица, где бы она ни находилась, рано или поздно прилипнет к кластеру.

Условие $R_d \gg R_b$ не следует явным образом из свойства возвратности. При этом правило алгоритма, допускающее уничтожение частиц, находится, на первый взгляд, в явном противоречии с тем фактом, что все частицы рано или поздно добавляются к кластеру.

Рассмотрим следующую задачу. Пусть имеется окружность радиуса R_b и в начальный момент времени частица находится в точке $(0, r)$, причём $r > R_b$. Тогда можно вычислить вероятность первого пересечения границы окружности R_b в точке $(R_b \cos \varphi, R_b \sin \varphi)$ при случайном блуждании частицы:

$$P(\varphi) = \frac{x^2 - 1}{2\pi(x^2 - 2x \cos \varphi + 1)}, \quad (1)$$

где $x = r/R_b$.

В случае, если $R_d \gg R_b$, вероятность вырождается и угловая зависимость пропадает. Таким образом, те частицы, которые в процессе случайного блуждания отошли от кластера достаточно далеко, на окружность рождения будут возвращаться с постоянной вероятностью, не зависящей от угла. Поэтому правило уничтожения частиц и последующего рождения новой частицы в случайном месте на окружности рождения для частиц, находящихся далеко от кластера, эквивалентно возврату частицы к кластеру.

Анализ показывает, что ошибка определения вероятности, связанная с заменой выражения (1) на константу, падает с ростом x как $1/x$. Поэтому для достижения точности 1 % отношение R_d к R_b должно быть больше 100. Это означает, что необходимо поддерживать достаточно большое пустое пространство вокруг кластера, что существенно снижает эффективность алгоритма.

Важной особенностью модели DLA в виде, предложенном Виттенем и Сандером, является необходимость использования решётки, которая вносит значительные искажения в свойства кластеров. Наиболее естественный шаг — это переход к безрешёточной модели, в которой частицы — это окружности единичного радиуса. При этом случайное блуждание моделируется как шаг конечной длины в случайном направлении, а окружность рождения R_b — это окружность, описывающая кластер, т. е. минимальная из возможных окружностей.

Важная модификация, базирующаяся на равномерности распределения случайного блуждания по углу, заключается в возможности увеличения длины шага, если частица находится вдали от кластера. Если расстояние от частицы до кластера равно Δr , то можно вместо единичного шага взять

шаг, незначительно меньший Δr , и переместить частицу сразу на большое расстояние.

В свою очередь, правило уничтожения частиц в случае отсутствия решётки может быть заменено правилом возврата частиц в соответствии с распределением (1): частицы, отошедшие от кластера на некоторое расстояние (за окружность R_d , которая может быть немного больше R_b), возвращаются обратно на R_b . Существуют поправки к формуле (1), связанные с конечностью шага, который совершают частицы при случайном блуждании. В работе [Ziff, 2009] для общей вероятности возврата частицы в случае пространства размерности $D=3$ данная поправка аналитически вычислена. Чтобы избежать необходимости её учёта, шаг частицы может уменьшаться при приближении к кластеру, при этом поправка будет пренебрежимо мала.

Указанные выше правила корректно описывают процесс случайного блуждания на плоскости и воспроизводят все его свойства, однако для практической реализации алгоритма необходимо эффективно реализовать все его шаги. Наибольшую сложность представляют задача организации хранения координат частиц в памяти, задача поиска размеров свободной области вокруг двигающейся частицы и задача определения факта столкновения частицы с кластером. Для решения указанных задач используется следующий метод.

Для определения факта столкновения необходимо выполнить попарное сравнение координат двигающейся частицы с координатами частиц кластера. Для минимизации количества необходимых операций естественно проводить сравнения не со всеми частицами кластера, а с частицами, которые находятся по соседству. Для этого вся область моделирования разбивается на ячейки размером 32 (за единичный размер принимается радиус 1 частицы) и координаты частиц кластера хранятся в этих ячейках. Данная структура реализует пространственный индекс, что позволяет проводить сравнение для проверки факта столкновения только с частицами, лежащими в текущей и в соседних с текущей ячейками.

Для определения размеров свободной области Болл и Брэди [Ball, Brady, 1985] предложили алгоритм иерархических сеток. На нижнем уровне иерархии находятся ячейки, помечаемые как занятые или свободные. Занятые ячейки — это те ячейки, в которых находятся частицы кластера. Следующий уровень иерархии — это ячейки, размер которых в 2 раза больше. Если, как минимум, одна из подячеек нижележащего уровня занята — то ячейка текущего уровня также помечается как занятая. Далее этот процесс повторяется для высшего уровня иерархии. Таким образом, для определения размера свободной области ищется сводная ячейка (на максимальном уровне иерархии), которая не занята и которая содержит в себе движущуюся частицу. После этого частица перемещается на границу свободной ячейки по заранее вычисленному вероятностному распределению, которое эквивалентно случайному блужданию в данной квадратной области при наличии поглощающих границ. Иллюстрация данного алгоритма (с различными уровнями иерархии) показана на рис. 1. Недостаток данного метода — низка точность определения размера свободной области, особенно в случае, если частица находится около границы ячейки. Дополнительно, генерация с высокой точностью точек с необходимым распределением, симулирующим большой шаг на границу квадрата, — достаточно сложная вычислительная задача.

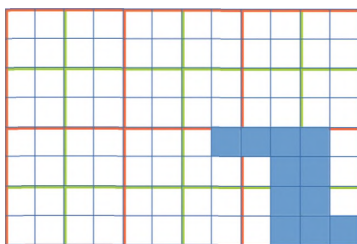


Рис. 1. Иерархическая сетка для определения размера свободной области по алгоритму Болла и Брэди. Синие ячейки — ячейки нижней иерархии, содержащие частицы кластера

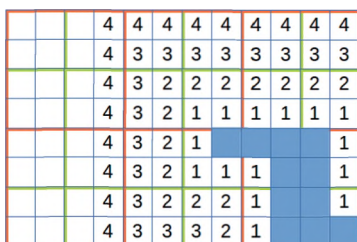


Рис. 2. Сетка для определения размера свободной области по предложенному алгоритму. Синие ячейки — ячейки нижней иерархии, содержащие частицы кластера. Цифрами показан размер свободной области вокруг каждой ячейки сетки

Для повышения эффективности алгоритма нами предложен альтернативный метод, в котором каждой ячейке сетки приписан размер вписанного квадрата в виде целого числа в единицах размера одной ячейки. Данный алгоритм показан на рис. 2.

Значение 1 означает, что какая-то из соседних ячеек занята и, соответственно, частица не может сделать большой шаг, значение 2 — что безопасно сделать большой шаг размером в 1 ячейку сетки. Соответственно, размер возможного шага равен числу, записанному в ячейке, за вычетом 1, и умноженному на размер одной ячейки сетки.

Пересчёт размеров свободных областей вокруг кластера выполняется только в случае, если ранее пустая ячейка становится занятой, что происходит достаточно редко. При этом пересчёт выполняется только по тем ячейкам, которые являются соседями занятой ячейки или соседями ячеек, изменивших состояние в процессе такого пересчёта. Область пересчёта может распространяться до бесконечности, поэтому для корректной работы её необходимо ограничить сверху некоторой величиной. По умолчанию считается, что размер свободной области не более 15 ячеек для всех ячеек сетки. Таким образом, когда пересчёт достигает ячеек со значением 15, — он автоматически останавливается.

Для больших кластеров их линейные размеры намного превышают размер 15 ячеек, что приводит к падению производительности алгоритма. Для компенсации этого фактора используется ещё один уровень иерархии ячеек, которые в несколько раз больше ячеек нижнего уровня иерархии. Соответственно, поиск размера свободной области начинается с ячеек высокого уровня иерархии (самых больших ячеек).

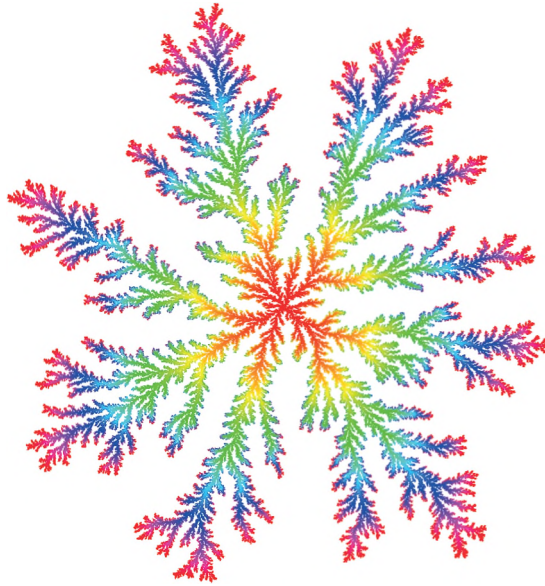


Рис. 3. Типичный кластер DLA. Размер кластера — 50 млн частиц. Разный цвет соответствует разному возрасту частиц

Если частица находится в области рядом с кластером, и, например, текущая ячейка помечена значением 1 (см. рис. 2), результатом работы алгоритма является запрет большого шага, так как в этом случае размер свободной области считается нулевым.

Однако используемый алгоритм является очень грубым и во многих случаях расстояние до кластера оказывается всё равно достаточно большим. В этом случае выполняется точный поиск размера свободной области путём поиска минимального расстояния до всех частиц кластера, находящихся в окрестности текущего положения новой частицы.

Типичный кластер, построенный по этому алгоритму, показан на рис. 3.

МНОГОМЕРНОЕ ОБОБЩЕНИЕ МОДЕЛИ

Многомерный алгоритм похож на описанный ранее двумерный аналог, за исключением следующих элементов.

В многомерном пространстве вероятность возврата частиц на сферу рождения больше не равна единице и зависит от отношения расстояния, на котором находится частица, к радиусу сферы, к которой происходит прилипание. Таким образом, для частиц, отошедших за окружность возврата, генерируется случайное число в диапазоне от 0 до 1. Затем оно сравнивается с величиной $(r/R_b)^{2-d}$ и, если оно меньше, то частица возвращается на окружность рождения с распределением:

$$P(\varphi) = \frac{x^2 - 1}{\Omega_d R_b^{d-1} (x^2 - 2x \cos \varphi + 1)^{d/2}}, \quad (2)$$

в противном случае частица уничтожается.

Для генерации случайных частиц с распределением (1) существует преобразование, которое переводит случайные числа с равномерным распределением на отрезке $[0;1)$ в числа с распределением (1). К сожалению, в 3-мерном случае невозможно выразить данное преобразование в виде элементарных функций. В многомерном случае необходимо использовать интегральную форму и численный метод, позволяющий провести вычисления. С ростом размерности пространства число случаев возврата падает, но при этом число событий, когда частица вышла за окружность возврата, — растёт. Это может привести к росту времени выполнения алгоритма из-за необходимости генерации случайных чисел с распределением (2). В этом случае метод численно-интегрирования может быть успешно реализован на базе GPU.

Дополнительно к необходимости генерации случайных чисел для выявления факта возврата на окружность и определения координат точки возврата, на каждом шаге необходимо генерировать случайные числа для нахождения направления случайного шага. Для многомерного случая методы генерации подобных чисел описаны в работе [Marsaglia, 1972]. Первый метод — это генерация случайных чисел, равномерно распределённых в гиперкубе, отброс точек, которые находятся вне сферы. Таким образом, генерируются точки, равномерно заполняющие сферу. Затем они проецируются на поверхность N -мерного шара с использованием масштабного преобразования. Недостаток этого метода — необходимость отброса точек, которые попали вне сферы на начальном шаге. Это, в свою очередь, приводит к низкой производительности алгоритма, поскольку в многомерном пространстве с ростом N число таких событий будет стремительно расти.

Метод, который наиболее применим, основан на использовании случайных чисел с нормальным распределением. Аналогично предыдущему методу необходимо генерировать N чисел, а затем выполнить их проекцию на поверхность сферы с использованием масштабного преобразования. В свою очередь, количество операций для генерации случайных чисел в многомерном случае достаточно велико, что приведёт к существенному замедлению алгоритма в многомерном случае. Соответственно, данные числа могут быть заранее сгенерированы с использованием GPU и сохранены в таблице в памяти, откуда они потом могут выбираться центральным процессором.

Организация памяти в многомерном случае, а также метод определения размера максимального шага с использованием ранее описанного алгоритма напрямую переносится из двумерного случая, не требуя каких-то особых модификаций.

ПРЕДМЕТНО-ОРИЕНТИРОВАННЫЙ ЯЗЫК ДЛЯ МОДЕЛИ DLA*

Параллельно с развитием модели и её обобщением на многомерный случай нами ведётся работа по развитию программно-аппаратного комплекса для выполнения указанного моделирования. Задача моделирования подобных объектов требует изучения большого числа кластеров, что может быть выпол-

* Раздел выполнен в рамках выполнения проекта РФФИ № 12-07-00693-а.

нено на базе многоядерной вычислительной системы, состоящей из нескольких вычислительных узлов.

Кроме того, изучаемая нами модель имеет большое множество подвариантов, например, анизотропные модели, получаемые небольшой модификацией правила прилипания частиц.

Можно сказать, что модель, которую приходится изучать, состоит из описания:

- 1) правил и методов генерации кластера,
- 2) правил генерации ансамбля кластеров,
- 3) методов анализа свойств ансамбля кластеров.

Простейшая реализация выполняет 1-й пункт с использованием заранее заданных веток алгоритмов, выбор которых происходит на этапе компиляции с использованием возможностей препроцессора языка C. Препроцессор используется для повышения общей производительности программы для отказа от условных операторов, которые существенно бы влияли на производительность модели для выбора различных блоков кода в процессе выполнения. Другая группа параметров модели, например, размер кластера, задаётся с использованием аргументов командной строки.

Выбор метода анализа свойств кластера осуществляется также с использованием аргументов командной строки, при этом, учитывая количество доступных методов анализа, перечень таких аргументов может быть достаточно большим.

Для автоматизации генерации ансамбля кластеров используются программы на языке Perl, которые по заданному шаблону генерируют набор задач для системы очередей Torque и осуществляют их запуск. Аналогичным образом выполняются задачи анализа ансамбля кластеров.

Разработанный нами программный комплекс достаточно сложен для использования, поэтому для повышения его качества был разработан программно-аппаратный комплекс с использованием программного обеспечения (ПО) Hubzero. Данная система позволяет посредством веб-интерфейса получить доступ к графическому интерфейсу пользователя (Graphical user interface, GUI), который является дополнительной надстройкой над ранее разработанным комплектом ПО. Данный интерфейс в более простой для пользователя форме показывает перечень всех параметров, которые он должен задать для запуска модели с заданными параметрами. Пример интерфейса пользователя показан на рис. 4.

Проблема при использовании интерфейса пользователя — сложность описания сложных комплексных моделей, включающих генерацию ансамбля кластеров и их последующий анализ.

Для модели DLA данный метод применим только для простых задач, например, для обучения студентов, поскольку анализ новых вариантов модели требует существенной доработки интерфейса пользователя, который оказывается чрезмерно перегруженным.

Для практического применения, построения и анализа моделей необходим другой подход. Наиболее естественным кажется использование предметно-ориентированного языка (DSL), который охватывает все необходимые 3 аспекта — описание одного кластера, описание ансамбля кластеров, описание анализа ансамбля кластеров.

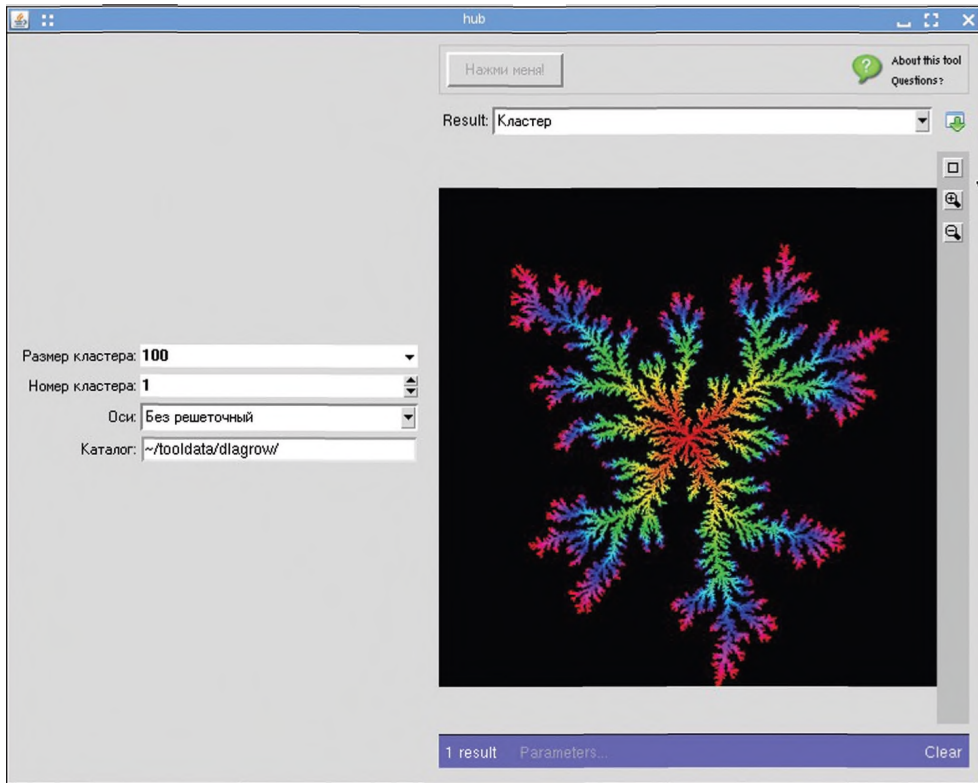


Рис. 4. Пример графического интерфейса пользователя для модели DLA платформы Hubzero

Как мы видели ранее, описание модели DLA в виде алгоритма состоит всего из 6 шагов. При этом основные отличия между вариантами моделей находятся в реализации шага прилипания частиц. Все они могут быть описаны на специализированном языке, который достаточно понятен и без дополнительных пояснений.

Описание ансамбля кластеров и методов анализа заключается в задании размеров ансамбля, выбора способов и очередности анализа, а также задания мест хранения промежуточных файлов и результатов моделирования.

Итоговая цель нашей работы — создание DSL для описания всего численного эксперимента. Данный DSL будет выполнять следующие функции — генерация текста программы на языке C++ для её последующей компиляции и генерация заданий для вычислительного кластера. Для создания DSL будет использована среда Eclipse с расширениями Xtext/Xtend для его описания и реализации.

Выводы

В работе подробно описаны процесс создания многомерной версии алгоритма модели DLA, а также отдельные шаги, которые необходимо выполнить для повышения производительности программы. Предлагаемые методы активно

используют средства GPU. Дополнительно рассмотрены основные элементы предметно-ориентированного языка, развиваемого для упрощения описания моделей и для упрощения проведения численных экспериментов за счёт автоматизации запуска задач на вычислительном кластере.

ЛИТЕРАТУРА

- [Bak et al., 1987] *Bak P., Tang C., Wiesenfeld K.* Self-organized criticality: An explanation of $1/f$ noise // *Physical Review Letters*. 1987. V. 59. P. 381–384.
- [Ball, Brady, 1985] *Ball R. C., Brady R. M.* Large scale lattice effect in diffusion-limited aggregation // *J. Physics A*. 1985. V. 18. P. L809.
- [Ball et al., 1985] *Ball R. C., Brady R. M., Rossi G., Thomson B. R.* Anisotropy and Cluster Growth by Diffusion-Limited Aggregation // *Physical Review Letters*. 1985. V. 55. P. 1406.
- [Halsey, Leibig, 1992] *Halsey T. C., Leibig M.* Theory of branched growth // *Physical Review A*. 1992. V. 46. P. 7793.
- [Hohenberg, Halperin, 1977] *Hohenberg P. C., Halperin B. I.* Theory of dynamic critical phenomena // *Reviews of Modern Physics*. 1977. V. 49. P. 435.
- [Marsaglia, 1972] *Marsaglia G.* Choosing a Point from the Surface of a Sphere // *Annals of Mathematical Statistics*. 1972. P. 645–646. doi: 10.1214/aoms/1177692644.
- [Meakin, 1985] *Meakin P.* The structure of two-dimensional Witten-Sander aggregates // *J. Physics A*. 1985. V. 18. P. L661.
- [Menshutin, Shchur, 2006] *Menshutin A. Yu., Shchur L. N.* Test of mutiscaling in a diffusion-limited-aggregation model using an off-lattice killing-free algorithm // *Physical Review E*. P. 2006. V. 73. P. 011407.
- [Menshutin et al., 2007] *Menshutin A. Yu., Shchur L. N., Vinokur V. M.* Probing surface characteristics of diffusion-limited aggregation clusters with particles of variable size // *Physical Review E*. 2007. V. 75. Art. 010401(R).
- [Menshutin et al., 2008] *Menshutin A. Yu., Shchur L. N., Vinokur V. M.* Finite size effect of harmonic measure estimation in a DLA model: Variable size of probe particles // *Physica A*. 2008. V. 387. P. 6299.
- [Stanley, 1999] *Stanley H. E.* Scaling, universality, and renormalization: Three pillars of modern critical phenomena // *Reviews of Modern Physics*. 1999. V. 71. P. 358.
- [Witten, Sander, 1981] *Witten T. A., Sander L. M.* Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon // *Physical Review Letters*. 1981. V. 47. P. 1400.
- [Ziff et al., 2009] *Ziff R. M., Majumdar S. N., Comtet A.* Capture of particles undergoing discrete random walks // *J. Chemical Physics*. 2009. V. 130. P. 204104–5.

GENERALIZATION OF DIFFUSION LIMITED AGGREGATION (DLA) MODEL FOR MULTI-DIMENSIONS

A. Yu. Menshutin, L. N. Shchur

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

We describe numerical algorithm for Diffusion Limited Aggregation (DLA) clusters generation in multi space dimensions. Techniques for numerical simulation speed-up, for memory organization, for random walk simulation in space with dimensions greater than 2, and

other numerical algorithms are discussed. For numerical experiment description we propose a Domain Specific Language (DSL).

Keywords: algorithm, random walk, DLA model, domain-specific language.

Menshutin Anton Yurievich — junior researcher, PhD, may@itp.ac.ru

Shchur Lev Nikolaevich — head of department, doctor of physical and mathematical sciences, professor, shchur@itp.ac.ru

МОДЕЛИРОВАНИЕ МАГНИТОРОТАЦИОННЫХ ПРОЦЕССОВ В КОЛЛАПСИРУЮЩИХ СВЕРХНОВЫХ И РАЗВИТИЕ МАГНИТОДИФФЕРЕНЦИАЛЬНО-ВРАЩАТЕЛЬНОЙ НЕУСТОЙЧИВОСТИ

С. Г. Моисеенко¹, Г. С. Бисноватый-Коган^{1,2}, Н. В. Арделян³

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)

Сверхновые с коллапсирующим ядром сопровождаются формированием нейтронной звезды или чёрной дыры. Гравитационная энергия трансформируется в энергию взрыва, что наблюдается во взрывах сверхновых второго типа и типов Ib,c (SN II, SNIb, SNIc). Представлены результаты двумерного магнитогазродинамического (МГД) моделирования магниторотационного взрыва коллапсирующей сверхновой, где источником энергии является энергия вращения, а магнитное поле играет роль «передаточного ремня», трансформирующего вращательную энергию в энергию взрыва. Тороидальная часть магнитной энергии на начальном этапе эволюции растёт линейно со временем из-за дифференциального вращения. Когда тороидальная компонента магнитного поля существенно превышает полоидальную компоненту магнитного поля, возникает магнито-дифференциально-вращательная неустойчивость (МДРН), приводящая к быстрому росту магнитной энергии. На конечной стадии развития процесса формируется быстрая МГД ударная волна, приводящая к взрыву сверхновой. В случае, когда начальное магнитное поле является полем дипольного типа, возникает слабо коллимированный струйный выброс (джет). При очень больших значениях магнитного поля развитие МДРН не обнаружено.

Работа Г. С. Бисноватого-Когана и С. Г. Моисеенко была частично поддержана РФФИ (проекты № 14-02-00728, 14-29-06045), грантом Президента РФ поддержки ведущих научных школ НШ-261.2014.2, программой президиума РАН П21/3. Работа С. Г. Моисеенко была также частично поддержана РФФИ (проект № 14-29-06055).

Ключевые слова: сверхновые МГД, численное моделирование.

ВВЕДЕНИЕ

Сверхновая представляет собой один из самых мощных взрывов во Вселенной, при котором высвобождается около 10^{51} эрг в виде кинетической и радиационной энергии. Взрыв сверхновой представляет собой конечную стадию эволюции массивной звезды с массой, большей $\sim 8M_{Sun}$. Взрыв сверхновой типа Ia представляет собой термоядерный взрыв вырожденного углеродно-кислородного ядра с полным разрушением звезды, при этом начальная масса звезды не превышает $\sim 12M_{Sun}$, а электроны вырождены в С-О ядре. Для

Моисеенко Сергей Григорьевич — заведующий отделом, доктор физико-математических наук, moiseenko@iki.rssi.ru

Бисноватый-Коган Геннадий Семёнович — главный научный сотрудник, доктор физико-математических наук, профессор

Арделян Николай Васильевич — ведущий научный сотрудник ВМиК МГУ, доктор физико-математических наук

больших начальных масс эволюция звезды продолжается до образования железного ядра и звезда коллапсирует из-за потери гидродинамической устойчивости. При коллапсе и образовании нейтронной звезды происходит высвобождение $\sim 6 \cdot 10^{53}$ эрг, которая уносится нейтрино. Первый механизм, предложенный в работе [Colgate, White, 1966] для объяснения взрыва сверхновой с коллапсирующим ядром, связан с нейтринной депозицией. Огромный поток энергии, уносимой нейтрино, разогревает схлопывающиеся внутренние слои материи, приводит к изменению движения вещества, образованию ударной волны, что должно приводить к взрыву сверхновых типа II и Ib,c. В дальнейшем более точные расчёты показали, что получаемая при моделировании энергия взрыва слишком мала для объяснения результатов наблюдений. В последнее время был предложен ряд модификаций нейтринной модели, однако проблема объяснения взрыва сверхновой с помощью нейтринного механизма до сих пор не решена. Обзор по этой теме представлен в работе [Bisnovatyι-Kogan, 2011]. Как показали результаты наших недавних расчётов (статья готовится к печати), для начального магнитного поля $H_0 = 10^{12}$ Гс взрыв сверхновой происходит сразу после коллапса, формируется коллимированный джет так же как и в статье [Takiwaki et al., 2004]. Для случая, когда магнитное поле слабее ($H_0 \approx 10^9$ Гс) показано, что после линейного роста тороидального магнитного поля из-за дифференциального вращения, оно растёт экспоненциально со временем из-за магниторотационной неустойчивости (МРН) типа Тейлера [Tauleg, 1973]. Мы назвали неустойчивость типа Тейлера при наличии дифференциального вращения магнито-дифференциально-вращательной неустойчивостью (МДРН). Истечение при взрыве в наших расчётах слабо коллимировано, однако в работе [Takiwaki et al., 2004] была получена сильная коллимация истечения при той же постановке задачи. В описанных в данной работе результатах моделирования начальное магнитное поле однородно и направлено вдоль оси вращения, так же как и в статье [Takiwaki et al., 2004].

1. МАГНИТОРОТАЦИОННЫЙ МЕХАНИЗМ ВЗРЫВА

В магниторотационном (МР) механизме взрыва сверхновой преобразование вращательной энергии нейтронной звезды в энергию взрыва происходит при помощи магнитного поля [Бисноватый-Коган, 1970]. Нейтронные звезды вращаются и обладают магнитным полем до 10^{13} Гс и более. Часто при взрыве сверхновой происходит одностороннее или двустороннее истечение вещества, что свидетельствует о том, что взрывы являются несферическими. В дифференциально вращающихся формирующихся нейтронных звёздах радиальное магнитное поле закручивается, магнитное давление сильно возрастает, что приводит к формированию МГД ударной волны, при помощи которой вращательная энергия трансформируется в энергию взрыва.

Расчёты МР-взрыва сверхновой в рамках одномерной нестационарной модели приведены в статье [Bisnovatyι-Kogan et al., 1976] для случая цилиндрической симметрии. Источником энергии взрыва была вращательная энергия системы (нейтронной звезды и оболочки). Как показали расчёты, в процессе эволюции происходит расслоение оболочки, основная её часть «прилипает» к нейтронной звезде и вращается вместе с ней практически

твёрдотельно. Внешняя часть оболочки улетает с большой скоростью, унося существенную часть вращательной энергии и углового момента. Эффективность МР-механизма составляет около 10% от вращательной энергии, эжектируемая масса $\sim 0,1 M_{Sun}$, энергия взрыва $\sim 10^{51}$ эрг. Эжектируемая масса и энергия взрыва слабо зависят от значения параметра $\alpha = E_{mag0}/E_{grav0}$, где E_{mag0} — магнитная энергия в начальный момент времени; E_{grav0} — гравитационная энергия в начальный момент времени. Время развития взрыва зависит от параметра α как $t_{expl} \approx 1/\sqrt{\alpha}$. Использование явных разностных схем затруднено для малых значений параметра α из ограничения Куранта на шаг по времени. При малых значениях параметра α система решаемых МГД-уравнений становится жёсткой, с параметром жёсткости α .

2. ДВУМЕРНЫЕ РАСЧЁТЫ

Численный метод, использованный в наших расчётах, основан на неявной операторно-разностной полностью консервативной схеме в лагранжевых переменных на треугольной сетке переменной структуры. Неявность схемы позволяет проводить расчёты с существенно бóльшим шагом по времени, чем в случае применения явной схемы. Применение неявной схемы важно при проведении расчётов задач с существенно различными временными масштабами: малый масштаб определяется большой скоростью звука в центральных областях звезды, большой масштаб связан с характерным временем эволюции магнитного поля. Применяемый метод приведён в статьях [Арделян, Черниговский, 1984; Ardeljan, Kosmachevskii, 1995; Ardeljan et al., 1996]. В этих же работах была исследована его устойчивость. Схема является полностью консервативной, что означает выполнение законов сохранения массы, импульса и закона сохранения энергии в различных формах. Метод был детально описан в работе [Ardeljan et al., 2000].

Моделирование МР-взрыва сверхновой в двумерной постановке было приведено в работе [Ardeljan et al., 2005], уравнения идеальной МГД с самогравитацией были численно решены при помощи операторно-разностного метода, описанного выше. Задача моделировалась в приближении аксиальной симметрии ($\partial/\partial\phi = 0$) и симметрии по отношению к экваториальной плоскости ($z = 0$). В начальный момент времени задавался тороидальный ток J_ϕ (отсчёт времени начинался с момента получения стационарной конфигурации вращающейся нейтронной звезды) по которому, при помощи закона

Био-Савара, вычислялись компоненты H_r, H_z ; $\mathbf{H} = \frac{1}{c} \int_V \frac{\mathbf{J} \times \mathbf{R}}{R^3} dV$. Начальное

магнитное поле квадрупольного типа симметрии получалось при использовании токов, ориентированных в противоположных направлениях в разных полусферах. Нейтринное охлаждение учитывалось при помощи варианта метода ограничения потоков [Ardeljan, et al., 2005].

Магнитное поле усиливается благодаря дифференциальному вращению и последующему развитию магниторотационной неустойчивости. Распределение тороидальной компоненты магнитного поля (рис. 1) для случая $\alpha = 10^{-6}$ в момент времени, когда тороидальная магнитная энергия достигает максимума, представлено на рис. 1. Максимальное значение тороидального

магнитного поля, полученное в расчётах, составляет $H_{\phi} = 2,5 \cdot 10^{16}$ Гс. Магнитное поле на поверхности нейтронной звезды после взрыва $H = 4 \cdot 10^{12}$ Гс. Изменение со временем вращательной и кинетической полоидальной энергии при развитии МР-взрыва представлено на рис. 2. Практически вся гравитационная энергия, переходящая во внутреннюю энергию при коллапсе, уносится слабо взаимодействующими нейтрино. Энергия, выбрасываемая в виде полоидальной кинетической энергии, составляет $\sim 0,6 \cdot 10^{51}$ эрг, эжектируемая масса $\sim 0,14 M_{Sun}$.

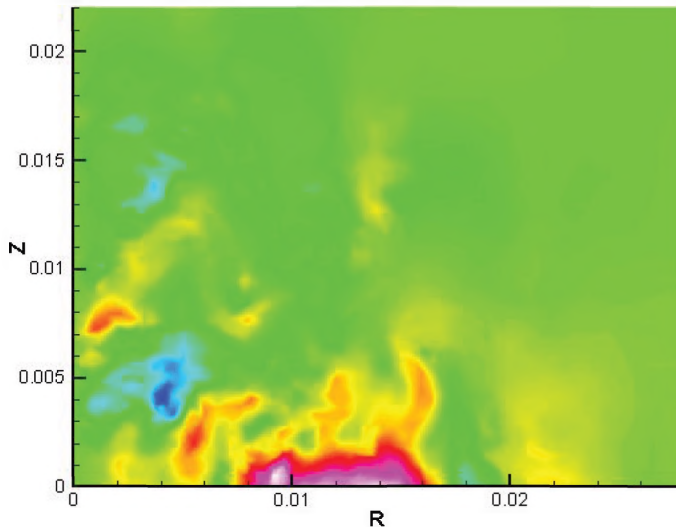


Рис. 1. Распределение тороидального магнитного поля в момент максимума тороидальной магнитной энергии при начальном магнитном поле квадрупольного типа

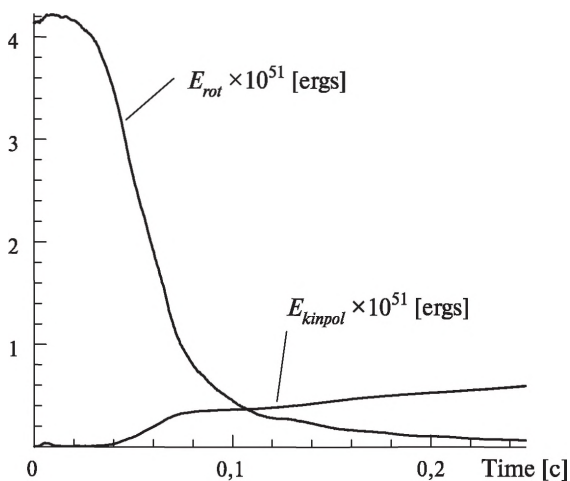


Рис. 2. Изменение со временем вращательной и полоидальной кинетической энергии во время МР-взрыва для начального поля дипольной конфигурации из работы [Moiseenko et al., 2006]

Были проведены расчёты для начальных конфигураций полоидального магнитного поля квадрупольного [Ardeljan, et al., 2005] и дипольного [Moiseenko et al., 2006] типов симметрии. Перед коллапсом отношение вращательной энергии к гравитационной и внутренней энергии к гравитационной составляло: $E_{rot}/E_{grav} = 0,0057$, $E_{int}/E_{grav} = 0,727$, где E_{rot} — вращательная энергия; E_{int} — внутренняя энергия; E_{grav} — гравитационная энергия. Начальное магнитное поле «включалось» после коллапса. Отношение между начальной магнитной и гравитационной энергиями было равно 10^{-6} . Начальное полоидальное магнитное поле в центре в момент начала эволюции тороидального магнитного поля было $\sim 3,2 \cdot 10^{13}$ Гс.

Магнитное поле играет роль поршня для возникающей МГД ударной волны. Зависимость от времени эжектируемой массы и энергии МР-взрыва представлено на рис. 3, 4. При МР-взрыве выбрасывается $\sim 0,14 M_{Sun}$ массы и $\sim 0,6 \cdot 10^{51}$ эрг ($\sim 10\%$ от вращательной энергии).

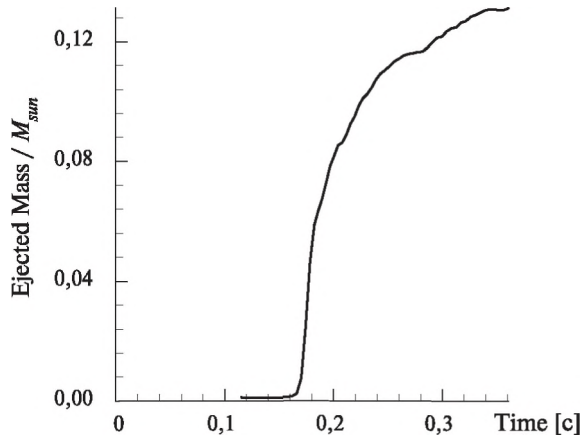


Рис. 3. Зависимость от времени эжектируемой массы при МР взрыве для случая начального магнитного поля дипольного типа из работы [Moiseenko et al., 2006]

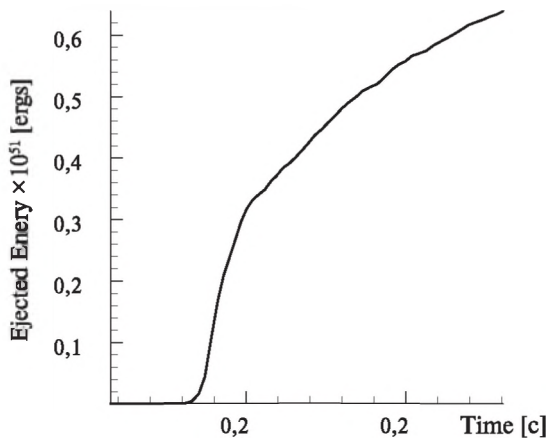


Рис. 4. Зависимость от времени энергии взрыва при МР взрыве для случая начального магнитного поля дипольного типа из работы [Moiseenko et al., 2006]

Результаты моделирования МР-взрыва сверхновой для различных начальных масс ядра и различных начальных энергий представлены в работе [Bisnovatyi-Kogan et al., 2008]. Начальная масса ядра варьировалась от $1,2$ до $1,7M_{\odot}$, начальная удельная вращательная энергия E_{rot}/M_{core} , где E_{rot} — вращательная энергия; M_{core} — масса ядра, варьировалась от $0,19 \cdot 10^{19}$ до $0,4 \cdot 10^{19}$ эрг/г. Энергия взрыва растёт с увеличением массы ядра и начальной энергии вращения. Полученная энергия МР-взрыва сверхновой $(0,5...2,6) \cdot 10^{51}$ эрг досточна для объяснения взрыва сверхновой типов II и Ib. Энергия взрыва сверхновых типа Ic может быть больше.

3. МАГНИТОРОТАЦИОННАЯ НЕУСТОЙЧИВОСТЬ

Магниторотационная неустойчивость вызывает экспоненциальный рост магнитного поля. Различные типы МРН приведены в работах [Велихов, 1959; Spruit, 2002]. МРН начинает развиваться, когда отношение тороидальной и полоидальной магнитных энергий становится большим. В одномерном МРН не возникает, поскольку количество степеней свободы системы ограничено, время развития МР-взрыва растёт с уменьшением параметра α по следующему закону: $t_{expl} \approx 1/\sqrt{\alpha}$, $\alpha = E_{mag0}/E_{grav0}$. Благодаря развитию МРН, зависимость времени МР-взрыва от параметра α существенно слабее. МР-взрыв происходит, когда магнитная энергия становится порядка внутренней энергии, хотя бы в некоторых местах. Начальная магнитная энергия линейно зависит от α , МРН приводит к экспоненциальному росту магнитной энергии. Общее время развития МР-взрыва в двумерных расчётах можно представить *логарифмически* растёт с уменьшением α , $t_{expl} \approx -\log \alpha$. Результаты одномерного моделирования [Bisnovatyi-Kogan et al., 1976] и двумерного [Ardeljan, et al., 2005; Moiseenko et al., 2006] отражают эти зависимости. В одномерном случае (в безразмерных величинах) имеем (t_{expl} в произвольных единицах): $\alpha = 0,001$, $t_{expl} = 10$; $\alpha = 10^{-12}$, $t_{expl} = 10^6$; в двумерном случае: $\alpha = 0,001$, $t_{expl} = 10$ и $\alpha = 10^{-6}$, $t_{expl} \approx 6$, $\alpha = 10^{-12}$, $t_{expl} \approx 12$. Зависимость времени МР-взрыва представлена на графиках для квадрупольной [Moiseenko et al., 2005] и дипольной конфигурации магнитного поля [Moiseenko et al., 2006]. Качественное описание и модельный пример МРН, иллюстрирующий возможность экспоненциального роста магнитного поля в двумерном случае, представлены в статьях [Ardeljan et al., 2005; Moiseenko et al., 2006].

4. ОБРАЗОВАНИЕ ДЖЕТОВ ПРИ МАГНИТОРОТАЦИОННОМ ВЗРЫВЕ

Образование джетов при МР-взрыве происходит, когда начальное магнитное поле является дипольным. Двумерные расчёты с начальным магнитным полем дипольного типа дают практически такую же энергию взрыва $\sim 0,5 \cdot 10^{51}$ эрг и эжектированную массу $\sim 0,14M_{Sun}$, что и в случае начального магнитного поля квадрупольного типа. Однако, в случае диполя, выброс слабо коллимирован вдоль оси вращения [Moiseenko et al., 2006] (рис. 5).

Моделирование МР-взрыва сверхновой было проведено с уравнением состояния из статьи [Shen et al., 1998]. Сравнение результатов наших расчётов

с начальным однородным магнитным полем на лагранжевой сетке с результатами аналогичных расчётов из работ [Takiwaki et al., 2004, 2009], где использовалась эйлерова схема при одинаковых начальных и граничных условиях, показало хорошее согласие результатов для сильного начального магнитного поля ($H_0 = 10^{12}$ Гс), однако для более слабого начального магнитного поля ($H_0 = 10^9$ Гс) в наших расчётах был получен слабо коллимированный джет (см. также [Burrows et al., 2007]), результаты этих расчётов будут опубликованы. МДРН развивается в случае слабого начального магнитного поля и отсутствует при сильном начальном магнитном поле (рис. 6, 10).

Проведены расчёты для следующих значений начального однородного поля: $H_0 = 10^9, 10^{12}$ Гс. Отношение вращательной энергии к абсолютной величине гравитационной энергии в начальный момент времени было $E_{rot0}/E_{grav0} = 1\%, 2\%$. В случае, когда начальное магнитное поле не очень большое ($H_0 = 10^9$ Гс) и $E_{rot0}/E_{grav0} = 1\%$, развивается МДРН, что означает экспоненциальный рост всех компонент магнитного поля (см. рис. 6).

На рис. 7 представлена лагранжева треугольная сетка и отношение тороидальной магнитной энергии к полоидальной E_{tor}/E_{pol} (заливкой) для случая $H_0 = 10^9$ Гс, и $E_{rot0}/E_{grav0} = 1\%$. Тороидальная магнитная энергия больше полоидальной магнитной энергии в значительной части области формирования нейтронной звезды. МДРН достаточно хорошо разрешается на нашей треугольной сетке.

В случае сильного начального магнитного поля ($H_0 = 10^{12}$ Гс) и $E_{rot0}/E_{grav0} = 1\%$ превышение тороидальной магнитной энергии E_{tor} над полоидальной E_{pol} в этой области отсутствует.

На рис. 8 представлено изменение со временем вращательной, магнитной полоидальной и тороидальной энергий при МР-взрыве для случая $H_0 = 10^9$ Гс.

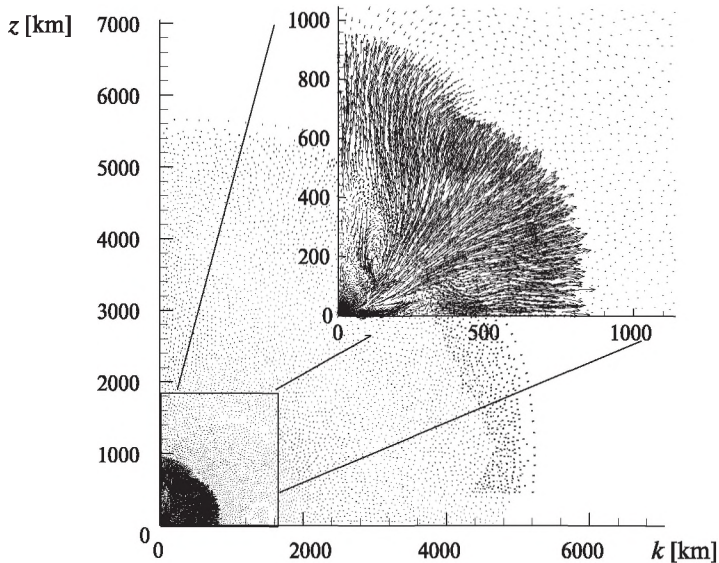


Рис. 5. Поле скоростей (выброс) на момент времени $t = 0,075$ с из работы [Moiseenko et al., 2006]

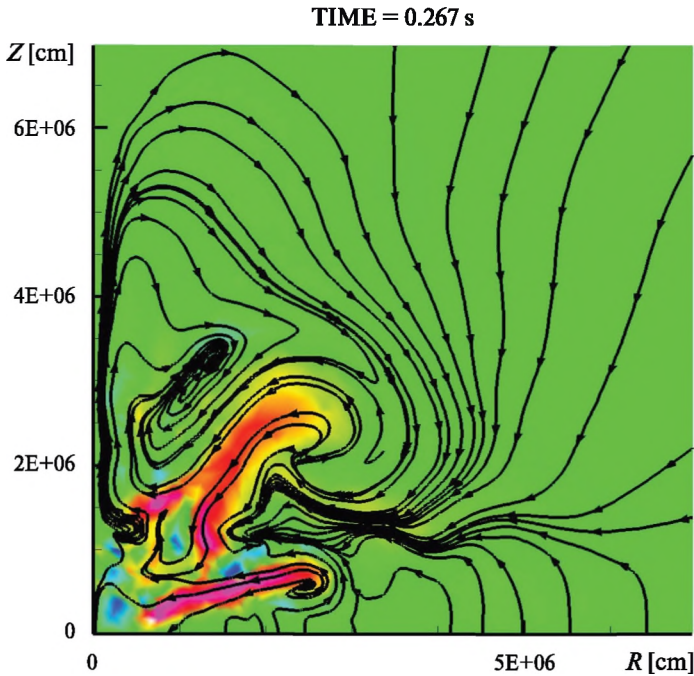


Рис. 6. Развитая МДРН благодаря конвекции и МРН/тейлоровской неустойчивости на момент времени $t = 267$ мс для случая $H_0 = 10^9$ Гс, $E_{rot0}/E_{grav0} = 1\%$ (линии уровня — тороидальное магнитное поле, линии со стрелками — силовые линии полоидального магнитного поля)

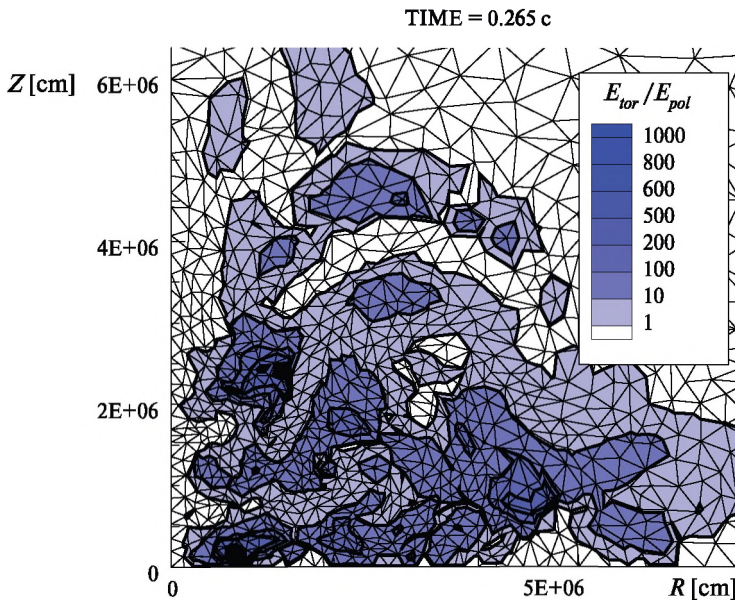


Рис. 7. Лагранжева треугольная сетка и отношение тороидальной магнитной энергии E_{tor} к полоидальной магнитной энергии E_{pol} , E_{tor}/E_{pol} в момент времени $t = 265$ мс для случая $H_0 = 10^9$ Гс, $E_{rot0}/E_{grav0} = 1\%$

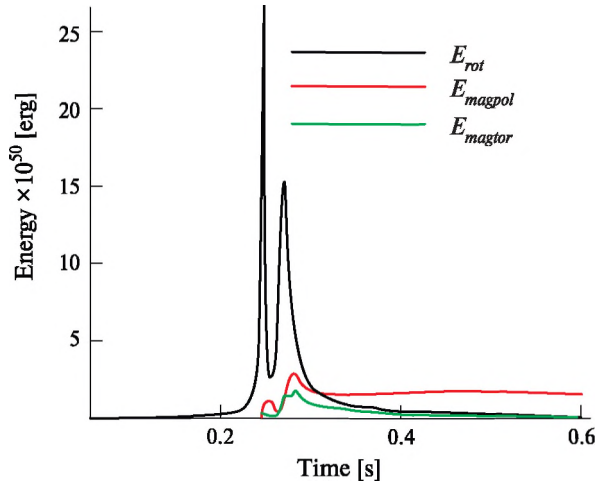


Рис. 8. Зависимость от времени вращательной E_{rot} (чёрная линия), магнитной полоидальной энергии E_{magpol} (красная линия) и магнитной тороидальной энергии E_{magtor} (зелёная линия) для случая $H_0 = 10^9$ Гс, $E_{rot0}/E_{grav0} = 1\%$

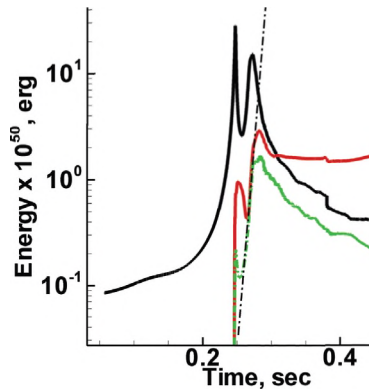


Рис. 9. Увеличенная часть рис. 8, вертикальная ось отложена в логарифмических координатах. Зависимость от времени вращательной энергии E_{rot} (чёрная линия), магнитной полоидальной энергии E_{magpol} (красная линия) и магнитной тороидальной энергии E_{magtor} (зелёная линия) для случая $H_0 = 10^9$ Гс, $E_{rot0}/E_{grav0} = 1\%$. Прямая штрихпунктирная линия показывает экспоненциальный рост тороидальной и полоидальной магнитных энергий со временем

Рисунок 9 представляет собой увеличенную центральную часть рис. 8, его вертикальная ось отложена в логарифмическом масштабе. Прямая штрихпунктирная линия на рис. 9 показывает, что тороидальная и полоидальная магнитные энергии растут экспоненциально со временем из-за развития МРДН.

Вращательная энергия имеет два максимума. Первый максимум связан с тем, что благодаря сохранению углового момента при сжатии происходит сильный рост вращательной энергии, максимум которой совпадает с первым максимумом плотности. Первое сжатие и последующий отскок вещества происходят, когда магнитное поле растёт медленно и потери углового момента ядра звезды малы. Развитие магниторотационной неустойчивости приводит

к быстрому росту компонент магнитного поля, большому потоку углового момента от ядра, что вызывает вторичное сжатие. Вторичное сжатие не переходит в расширение, поскольку происходит быстрая потеря вращательной энергии благодаря сильному потоку углового момента из ядра.

Полоидальная энергия магнитного поля растёт при первичном сжатии до момента времени $t \approx 0,225$ с. Затем она медленно убывает со временем из-за образования ударной волны отскока и её движении наружу. Тороидальная магнитная энергия растёт как квадратичная функция благодаря накручиванию магнитных силовых линий (тороидальная компонента магнитного поля растёт линейно со временем). Начиная с момента времени $t \approx 0,3$ с и полоидальная, и тороидальная компоненты магнитной энергии начинают расти экспоненциально со временем из-за МДРН. В момент времени $t \approx 0,36$ с происходит насыщение роста этих видов энергии. Возникает и развивается МГД ударная волна, что приводит к МР-взрыву. МР-взрыв происходит во всех направлениях, не формируя коллимированный поток.

Магниторотационный взрыв для экстремально сильного начального магнитного поля ($H_0 = 10^{12}$ Гс) развивается качественно по-другому. Начальное магнитное поле настолько велико, что оно сильно растёт при первичном сжатии и взрыв возникает ещё до развития МДРН (см. рис. 10). На рис. 11 представлена зависимость вращательной, тороидальной и полоидальной магнитных энергий от времени. Вращательная энергия имеет один экстремум при $t \approx 0,32$ с, что соответствует моменту максимального сжатия и сопровождается ростом тороидальной и полоидальной магнитных энергий. Полоидальная магнитная энергия растёт из-за сжатия и сохранения потока магнитного поля, тороидальная магнитная энергия растёт благодаря дифференциальному вращению и сжатию ядра.

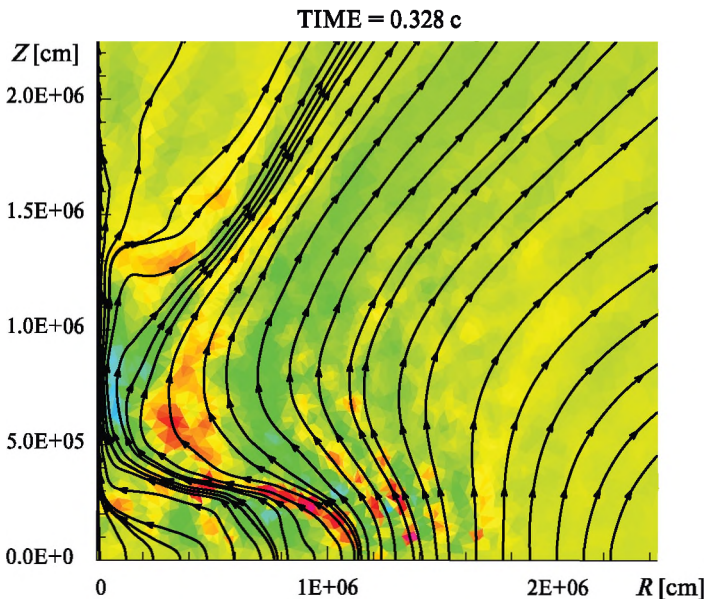


Рис. 10. Отсутствие МДРН в момент времени $t = 328$ мс для случая $H_0 = 10^{12}$ Гс, $E_{rot0}/E_{grav0} = 1\%$ (линии уровня — тороидальная компонента магнитного поля, линии со стрелками — силовые линии полоидального магнитного поля)

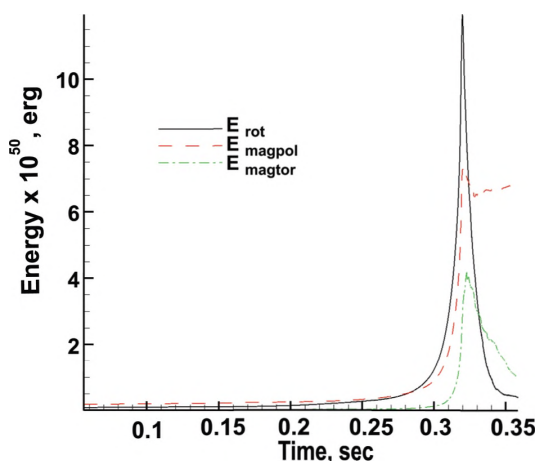


Рис. 11. Зависимость от времени вращательной энергии E_{rot} (сплошная линия), магнитной полоидальной энергии E_{magpol} (штриховая линия) и магнитной тороидальной энергии E_{magtor} (штрихпунктирная линия) для случая $H_0 = 10^{12}$ Гс, $E_{rot0}/E_{grav0} = 1\%$

При сильном начальном магнитном поле происходит быстрая потеря углового момента ядра уже на стадии сжатия. Центробежная сила уменьшается и первое сжатие не сопровождается отскоком. Магниторотационный взрыв происходит сразу после сжатия, в центре остаётся медленно вращающаяся нейтронная звезда.

Силовые линии магнитного поля играют роль «рельс». Вещество движется вдоль силовых линий. Магнитное давление преобладает на периферии ядра. Магниторотационный взрыв развивается преимущественно вдоль оси вращения и происходит формирование коллимированного потока (протоджета). В этом случае МР-взрыва образуется коллимированный джет. Степень коллимации джета примерно такая же, как и в работе [Takiwaki et al., 2004]. Для случая $E_{rot0}/E_{grav0} = 1\%$ и $H_0 = 10^9$ Гс энергия МР-взрыва сверхновой $\sim 4 \cdot 10^{50}$ эрг, для случая $H_0 = 10^{12}$ Гс энергия МР-взрыва сверхновой $\sim 8 \cdot 10^{50}$ эрг. Энергия взрыва, полученная в наших расчётах лагранжевым методом, примерно соответствует энергии МР-взрыва, полученной при использовании эйлерова метода [Takiwaki et al., 2004] (за исключением случая $H_0 = 10^{12}$ Гс и $E_{rot0}/E_{grav0} = 1\%$).

5. АСИММЕТРИЯ ВЗРЫВА

Как показывают наблюдения, взрывы сверхновых с коллапсирующим ядром асимметричны. Наши расчёты показывают, что МР-взрыв сверхновой происходит после развития МРДН. Развитие МРДН стохастический процесс, из чего следует, что форма взрыва сверхновой может меняться. МР-механизм взрыва сверхновой может приводить к различной форме взрыва сверхновой. Важно отметить, что МР-механизм взрыва сверхновой всегда приводит к асимметричным выбросам.

МР-взрыв сверхновой, результаты численного моделирования которого представлены в данной работе, симметричен относительно экваториаль-

ной плоскости, в силу постановки задачи. Однако в реальности экваториальная симметрия может нарушаться из-за МРДН, одновременного наличия дипольного и квадрупольного магнитных полей [Wang et al., 1992] и начального тороидального магнитного поля [Bisnovatyi-Kogan, Moiseenko, 1992]. Нарушение симметрии по отношению к экваториальной плоскости может привести к эффекту отдачи и образованию быстролетающих радиопульсаров. Скорость отскока при отдаче, направленная вдоль оси вращения, сформированная магнитогидродинамическими процессами при асимметрии магнитного поля, как показывают оценки [Bisnovatyi-Kogan, 1993], не превышает 300 км/с.

Когда магнитная ось и ось вращения не совпадают, процесс МР-взрыва является трёхмерным. Однако накручивание магнитного поля происходит вокруг оси вращения, следовательно, можно ожидать, что направление скорости отскока нейтронной звезды будет сильно коррелировано с направлением оси вращения.

При развитии МР-взрыва регулярная компонента магнитного поля может на короткий промежуток времени превышать 10^{16} Гс [Ardeljan, et al., 2005; Moiseenko et al., 2006]. При таких величинах магнитного поля сечение взаимодействия нейтрино зависит от величины магнитного поля. Величина анизотропии магнитного поля по отношению к плоскости, перпендикулярной оси вращения [Mikami et al., 2008], может достигать 50 %, что может привести к сильной анизотропии потока нейтрино. Скорость отскока из-за анизотропии нейтринного потока может достигать нескольких тысяч километров в секунду [Bisnovatyi-Kogan, 1993], что позволяет объяснить существование наиболее быстро движущихся радиопульсаров [Vlemmings et al., 2005]. Вместе с этим, благодаря стохастической природе МДРН, уровень анизотропии должен сильно меняться, что приводит к большому разбросу скоростей нейтронных звёзд. Это свойство МР-взрыва отличает его от моделей взрыва сверхновых, связанных с сильной нейтринной конвекцией, в которых направление отскока может быть произвольным [Bugrows et al., 1995; Janka, Müller, 1995]. В работе [Dessart et al., 2006] было отмечено, что конвекция в протонейтронной звезде является скорее вторичным, а не определяющим фактором для успешного взрыва сверхновой с коллапсирующим ядром.

Анализ наблюдений пульсаров показывает, что ось вращения и вектор скорости у пульсаров совпадают, как это и предсказывается МР-механизмом взрыва сверхновой. Это совпадение было впервые приведено в статье [Smirnova et al., 1996] и подтверждается со всё большей вероятностью в настоящее время в работах [Johnston et al., 2005, 2007; Noutsos et al., 2012]. Нарушение параллельности этих векторов может быть в случае, когда сверхновая взрывается в двойной системе.

ЗАКЛЮЧЕНИЕ

Эффективность трансформации вращательной энергии в энергию взрыва при магниторотационном механизме взрыва составляет ~10 %. Магнитодифференциально-вращательная неустойчивость значительно ускоряет развитие магниторотационного взрыва, при малых значениях начального магнитного поля. В случае, когда начальное магнитное поле является полем

дипольного типа, возможно формирование джета при магниторотационном взрыве сверхновой. Энергия магниторотационного взрыва слабо зависит от деталей уравнения состояния вещества, модели нейтринного переноса и выбора численного метода. Наблюдаемое совпадение оси вращения и вектора скорости пульсара, образующегося при взрыве сверхновой, естественным образом объясняется в рамках магниторотационного механизма взрыва сверхновой.

ЛИТЕРАТУРА

- [Арделян, Черниговский, 1984] *Арделян Н. В., Черниговский Г. С.* О сходимости разностных схем для двумерных уравнений газовой динамики с учётом гравитации в акустическом приближении // Дифференциальные уравнения. 1984. Т. 20. С. 1119–1127.
- [Бисноватый-Коган, 1970] *Бисноватый-Коган Г. С.* О механизме взрыва вращающейся звезды, как сверхновой // *Астрономический журн.* 1970. Т. 47. С. 813–816.
- [Велихов, 1959] *Велихов Е. П.* Устойчивость течения идеально проводящей жидкости между вращающимися цилиндрами в магнитном поле // *Журн. экспериментальной и теоретической физики (ЖЭТФ).* 1959. Т. 36, Вып. 5. С. 1399.
- [Ardeljan, et al., 1996] *Ardeljan N. V., Bisnovatyi-Kogan G. S., Kosmachevskii K. V., Moiseenko S. G.* An implicit Lagrangian code for the treatment of nonstationary problems in rotating astrophysical bodies // *Astronomy and Astrophysics Supplement Series.* 1996. V. 115. P. 573–594.
- [Ardeljan, et al., 2000] *Ardeljan N. V., Bisnovatyi-Kogan G. S., Moiseenko S. G.* Nonstationary magnetorotational processes in a rotating magnetized cloud // *Astronomy and Astrophysics (A&A).* 2000. V. 355. P. 1181–1190.
- [Ardeljan, et al., 2005] *Ardeljan N. V., Bisnovatyi-Kogan G. S., Moiseenko S. G.* Magnetorotational supernovae // *Monthly Notices of the Royal Astronomical Society (MNRAS).* 2005. V. 359. P. 333–344.
- [Ardeljan, Kosmachevskii, 1995] *Ardeljan N. V., Kosmachevskii K. V.* Implicit Free-Lagrange Method for Computing Two-Dimensional Magnetogasdynamic Flows // *Computational Mathematics and Modeling.* 1995. V. 6. P. 209–224.
- [Ardelyan et al., 1979] *Ardelyan N. V., Bisnovatyi-Kogan G. S., Popov Y. P.* The investigation of a magnetorotational explosion of a supernova in a cylindrical model // *Soviet Astronomy.* 1979. V. 23. P. 705–711 (= *Арделян Н. В., Бисноватый-Коган Г. С., Попов Ю. П.* Исследование магниторотационного взрыва сверхновой в цилиндрической модели // *Астрономический журн.* 1979. Т. 56. № 6. С. 1244–1256).
- [Bisnovatyi-Kogan et al., 1976] *Bisnovatyi-Kogan G. S., Popov I. P., Samokhin A. A.* The magnetohydrodynamic rotational model of supernova explosion // *Astrophysics and Space Science.* 1976. V. 41. P. 287–320.
- [Bisnovatyi-Kogan, 1993] *Bisnovatyi-Kogan G. S.* Asymmetric neutrino emission and formation of rapidly moving pulsars // *Astronomy and Astrophysics (A&A) Transact.* 1993. V. 3. N. 4. P. 287–294.
- [Bisnovatyi-Kogan, 2011] *Bisnovatyi-Kogan G. S.* *Stellar Physics. 2: Stellar Evolution and Stability.* Berlin; Heidelberg: Springer-Verlag, 2011.
- [Bisnovatyi-Kogan, Moiseenko, 1992] *Bisnovatyi-Kogan G. S., Moiseenko S. G.* // *Soviet Astronomy.* 1992. V. 36. P. 285 (= *Бисноватый-Коган Г. С., Моисеенко С. Г.* Нарушение

- зеркальной симметрии магнитного поля во вращающихся звёздах и возможные астрофизические проявления // *Астрономический журн.* 1992. Т. 69. С. 563–571).
- [Bisnovatyi-Kogan et al., 2008] *Bisnovatyi-Kogan G. S., Moiseenko S. G., Ardelyan N. V.* Different Magneto-rotation Supernovae // *Astronomy Reports.* 2008. V. 52. N. 12. P. 997–1008 (= *Бисноватый-Коган Г. С., Моисеенко С. В., Арделян Н. В.* Различные магниторотационные сверхновые // *Астрономический журн.* 2008. Т. 85. № 12. С. 1109–1121).
- [Burrows et al., 1995] *Burrows A., Hayes J., Fryxell B. A.* On the Nature of Core-Collapse Supernova Explosions // *Astrophysical J. (ApJ).* 1995. V 450. P. 830–850.
- [Burrows et al., 2007] *Burrows A., Dessart L., Livne E., Ott C. D., Murphy J.* Simulations of magnetically driven supernova and hypernova explosions in the context of rapid rotation // *Astrophysical J. (ApJ).* 2007. V. 664. P. 416.
- [Colgate, White, 1966] *Colgate S. A., White R. H.* The hydrodynamic behavior of supernovae explosions // *Astrophysical J. (ApJ).* 1966. V. 143. P. 626.
- [Dessart et al., 2006] *Dessart L., Burrows A., Livne E., Ott C. D.* Multidimensional Radiation/Hydrodynamic Simulations of Proto-Neutron Star Convection // *Astrophysical J. (ApJ).* 2006. V. 645. P. 534–550.
- [Janka, Müller, 1995] *Janka H.-T., Müller E.* The First Second of a Type II Supernova: Convection, Accretion, and Shock Propagation // *Astrophysical J. Letters.* 1995. V. 448. P. L109.
- [Johnston et al., 2005] *Johnston S., Hobbs G., Vigeland S., Kramer M., Weisberg J. M., Lyne A. G.* Evidence for alignment of the rotation and velocity vectors in pulsars // *Monthly Notices of the Royal Astronomical Society (MNRAS).* 2005. V. 364. P. 1397.
- [Johnston et al., 2007] *Johnston S., Kramer M., Karastergiou A., Hobbs G., Ord S., Wallman J.* Evidence for alignment of the rotation and velocity vectors in pulsars. II. Further data and emission heights // *Monthly Notices of the Royal Astronomical Society (MNRAS).* 2007. V. 381. P. 1625–1637.
- [Mikami et al., 2008] *Mikami H., Sato Y., Matsumoto T., Hanawa T.* Three-dimensional Magnetohydrodynamical Simulations of a Core-Collapse Supernova // *Astrophysical J. (ApJ).* 2008. V. 683. P. 357.
- [Moiseenko et al., 2005] *Moiseenko S. G., Bisnovatyi-Kogan G. S., Ardelyan N. V.* Magnetorotational Supernova Simulations // 1604–2004: ASP Conference Series. V. 342. Proc. Conf. held 15–19 June, 2004, Padua, Italy / Ed. M. Turatto, S. Benetti, L. Zampieri, W. Shea. San Francisco: Astronomical Society of the Pacific, 2005. P. 190.
- [Moiseenko et al., 2006] *Moiseenko S. G., Bisnovatyi-Kogan G. S., Ardelyan N. V.* Magnetorotational core collapse model with jets // *Monthly Notices of the Royal Astronomical Society (MNRAS).* 2006. V. 370. P. 501–512.
- [Noutsos et al., 2012] *Noutsos A., Kramer M., Carr P., Johnston S.* Pulsar spin–velocity alignment: further results and discussion // *Monthly Notices of the Royal Astronomical Society (MNRAS).* 2012. V. 423. P. 2736–2752. doi: 10.1111/j.1365-2966.2012.21083.x.
- [Shen et al., 1998] *Shen H., Toki H., Oyamatsu K., Sumiyoshi K.* Relativistic equation of state of nuclear matter for supernova and neutron star // *Nuclear Physics A.* 1998. V. 637. P. 435.
- [Smirnova et al., 1996] *Smirnova T. V., Shishov V. I., Malofeev V. M.* The Spatial Structure of Pulsar Emission Sources Determined Using Interstellar Scintillation // *Astrophysical J. (ApJ).* 1996. V. 462. P. 289. doi: 10.1086/177150.
- [Spruit, 2002] *Spruit H. C.* Dynamo action by differential rotation in a stably stratified stellar interior // *Astronomy and Astrophysics (A&A).* 2002. V. 381. P. 923–932.
- [Takiwaki et al., 2004] *Takiwaki T., Kotake K., Nagataki S., Sato K.* Magneto-driven Shock Waves in Core-Collapse Supernovae // *Astrophysical J. (ApJ).* 2004. V. 616. P. 1086. doi: 10.1086/424993

- [Takiwaki et al., 2009] *Takiwaki T., Kotake K., Sato K.* Special Relativistic Simulations of Magnetically-dominated Jets in Collapsing Massive Stars // *Astrophysical J. (ApJ)*. 2009. V. 691. P. 1360–1379. doi: 10.1088/0004-637X/691/2/1360
- [Taylor, 1973] *Taylor R.* The stability of stars containing magnetic fields // *Monthly Notices of the Royal Astronomical Society (MNRAS)*. 1973. V. 161. P. 365.
- [Vlemmings et al., 2005] *Vlemmings W. H. T., Chatterjee S., Brisken W. F.* et al. Pulsar Astrometry at the Microarcsecond Level // *Mem. SAI*. 2005. V. 77. P. 531 (astro-ph/0509025).
- [Wang et al., 1992] *Wang J. C. L., Sulkanen M. E., Lovelace R. V. E.* Intrinsically Asymmetric Astrophysical Jets // *Astrophysical J. (ApJ)*. 1992. V. 390. P. 46–65.

SIMULATION OF MAGNETOROTATIONAL PROCESSES IN CORE-COLLAPSED SUPERNOVAE AND DEVELOPMENT OF MAGNETO-DIFFERENTIAL-ROTATIONAL INSTABILITY

*S. G. Moiseenko*¹, *G. S. Bisnovatyi-Kogan*^{1,2}, *N. V. Ardeljan*³

¹ Space Research Institute of Russian Academy of Sciences (IKI RAN)

² National Research Nuclear University MEPhI (Moscow Engineering Physics Institute) (NRNU MEPhI)

³ Faculty of Computational Mathematics and Cybernetics of Lomonosov Moscow State University (CMC MSU)

Core-collapse supernovae are accompanied by formation of neutron stars. The gravitation energy is transformed into the energy of the explosion, observed as SN II, SN Ib,c type supernovae. We present results of 2-D MHD simulations, where the source of energy is rotation, and magnetic field serves as a «transition belt» for the transformation of the rotation energy into the energy of the explosion. The toroidal part of the magnetic energy initially grows linearly with time due to differential rotation. When the twisted toroidal component strongly exceeds the poloidal field, magneto-rotational instability develops, leading to a drastic acceleration in the growth of magnetic energy. Finally, a fast MHD shock is formed, producing a supernova explosion. Mildly collimated jet is produced for dipole-like type of the initial field. At very high initial magnetic field no MRI development was found.

Keywords: supernovae, MHD, numerical simulation.

Moiseenko Sergey Grigorievich — head of department, doctor of physical and mathematical sciences, moiseenko@iki.rssi.ru

Bisnovatyi-Kogan Gennsdiy Semenovich — chief researcher, doctor of physical and mathematical sciences, professor

Ardeljan Nikolai Vasilievich — leading researcher CMC MSU, doctor of physical and mathematical sciences

СВЕДЕНИЕ ЗАДАЧИ ОПТИМИЗАЦИИ ПОТОКОВ В ИНТЕЛЛЕКТУАЛЬНОЙ ИНЖЕНЕРНОЙ СЕТИ К ЗАДАЧЕ, РЕШАЕМОЙ СИМПЛЕКС-МЕТОДОМ

Д. А. Ноздрин

Московский институт электроники и математики Научно-исследовательского учреждения Высшая школа экономики (МИЭМ НИУ ВШЭ)

Работа посвящена рассмотрению балансных уравнений в интеллектуальных инженерных сетях, в частности, в электросетях, и задаче оптимизации потоков в условиях сетевой архитектуры и деления общей сети на зоны. Важной задачей, встающей перед многими комплексами, оказывается задача просчёта оптимального распространения потоков, такого, чтобы выбранная целевая функция стремилась к максимуму (минимуму). Данная задача актуальна для большинства интеллектуальных инженерных сетей с различными накладываемыми условиями. Если ввести ограничения по пропускной способности и/или ограничения на управляющие воздействия, то эта задача становится задачей управления потоками, которая включает в себя разные типы потоковых задач. Такими задачами могут быть задачи определения существования потока, транспортные задачи, задачи построения топологии и пр. Варьируя условия и ограничения задачи, можно адаптировать алгоритмы под сети газо-, водоснабжения, электросети, транспортные сети, сети сотовой связи. В данной статье задача рассматривается в контексте электросетей.

Ключевые слова: уравнение, сети, инженерия, сетевая архитектура, функция, условие, поток, управление, алгоритм, оптимизация.

ВВЕДЕНИЕ

Применительно к энергетическим системам России многие из алгоритмов применимы с существенными ограничениями, поскольку все элементы энергосистем имеют множество параметров, зависящих от различных явлений. Особенности, возникающие ограничения и невозможность влиять в некоторые моменты на управляющие воздействия при решении задач ещё сильнее сужают применимость существующих алгоритмов [Кармен и др., 2007; Tarjan, 1983]. Однако задача, в силу территориальных особенностей и особенностей государственного регулирования, имеет принципиально большое значение для энергетики России [Вайнзихер, 2008; Дорофеев, Макаров, 2009].

Метод, описанный в статье, позволяет находить решение задачи при следующих условиях [Ноздрин, 2013].

- Система отвечает уравнениям баланса (входящий поток в систему равен исходящему потоку). В случае рассмотрения не узлов, а целиком электросетей мощность потребления не превышает мощность генераторов (с учётом динамических толчков системы).
- Каждая линия электропередачи может быть как однонаправленной (ток течёт от установленного начала линии к концу), так и реверсивной (ток может протекать в любом направлении).
- В общем случае, в узлах системы может происходить поступление электрического тока в систему. Если это — не тупиковая ветвь (напри-

мер, установлены аккумуляторные батареи или промежуточная котельная), для простоты рассмотрения данные источники рассматриваются как установленные на ребре с «нулевой» ценой передачи.

- Линии электропередач обладают определённой пропускной способностью (допустимой нагрузкой, максимальным значением силы тока).
- Вся сеть делится на меньшие части, «зоны», расчёты некоторых величин ведутся для каждой «зоны» отдельно.
- Метод позволяет переформулировать задачу с введёнными ограничениями таким образом, чтобы решение можно было найти, используя стандартные методы, такие как симплекс-метод.

ПОСТАНОВКА ЗАДАЧИ

Математическая модель будет составлена из системы ограничений, роль которой играет система линейных уравнений и неравенств, и целевой функции, которая будет минимизироваться.

Систему уравнений будем рассматривать для каждой зоны (в частном случае — вся сеть). Уравнение баланса для одиночного узла j выглядит следующим образом:

$$\sum_i (p_{ij} + g_{ij}) = \sum_k p_{jk} + f_j. \quad (1)$$

Здесь i и j — вершины входящих в расчётный узел рёбер; j и k — вершины рёбер, которые выходят из узла; p — поток (мощность) на входящих в узел линиях и выходящих из узла линиях; g — потребление или распределение на входящих линиях; f — потоки, входящие в данную зону из других зон и выходящие из неё наружу.

Замечание: если для некоторых индексов i и j линии не существуют, то суммирование по таким величинам не производится ($p = 0$).

Перепишем формулу в следующем виде:

$$\sum_i p_{ij} - \sum_k p_{jk} = f_j - \sum_i g_{ij}. \quad (2)$$

Задача, которую необходимо решить, — построить распределение потоков в электросети, при условии минимизации потерь. Пусть c_{ij} — это проводимость линии электропередач, учитывающая длину линии, соединяющей узел i с узлом j . Целевая функция в таком случае будет следующей:

$$\sum_{i,j} c_{ij} p_{ij}. \quad (3)$$

Минимизируемая целевая функция, а также уравнения составляют постановку задачи линейного программирования. Поставленная задача линейного программирования будет решаться при помощи симплекс-метода [Кармен и др., 2007]. Такая задача в стандартной постановке выглядит следующим образом: даны m чисел b_1, b_2, \dots, b_m , таких что $b_x \in R$ (действительные числа), даны n действительных чисел c_1, c_2, \dots, c_n , таких что $c_x \in R$, и $m \times n$ чисел a_{ij} , где $a_x \in R, j = 1, 2, \dots, n$ и $\sum_{j=1}^n a_{ij} p_j \leq b_j$ [Карманов, 2004]. Необходимо найти n чисел $p_1, p_2, \dots, p_n, p_x \in R$, максимизирующих функцию:

$$\sum_{j=1}^n c_j p_j \tag{4}$$

в условиях

$$\sum_{j=1}^n a_{ij} p_j \leq b_j, \tag{5}$$

где $i = 1, 2, \dots, m$,

$$p_j \geq 0, \tag{6}$$

где $j = 1, 2, \dots, n$.

Матричный вид данной задачи будет выглядеть следующим: $\max \langle c, x \rangle$, $Ax \leq b$, $x \geq 0$, где $p = (p_1, p_2, \dots, p_n)^T$; $c = (c_1, c_2, \dots, c_n)$; $b = (b_1, b_2, \dots, b_m)^T$; $A = [a_{ij}]$ — матрица размера $m \times n$, содержащая коэффициенты, знаки неравенств выполняются для векторов по координатам, скобки $\langle \cdot \rangle$ обозначают скалярное произведение векторов.

СВЕДЕНИЕ УРАВНЕНИЙ БАЛАНСА К ЗАДАЧЕ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Для решения задачи (2)–(3) необходимо перейти от системы нумерации переменных по двум индексам к системе нумерации с одним индексом, при этом обеспечить взаимно однозначное соответствие между ними [Рыбников, 2010].

Пусть m — количество узлов графа, а n — количество линий. Если присвоить узлам графа уникальные номера — от 1 до m , то ориентированный граф в этом случае — это множество узлов $Y = (1, 2, \dots, m)$ и множество пар узлов $A_{ij} = (i, j) \in Y \times Y$, $i \neq j$.

Для переменной p_{ij} индекс ij потока уникален, число таких индексов равно n . Построим биекцию пары (i, j) на множество чисел $(1, 2, \dots, n)$. Это отображение примет вид некой «модифицированной» матрицы смежности размером $m \times n$. Делать это будем следующим образом.

Если i — номер узла, из которого он вытекает, а j — номер узла, куда он течёт, тогда в клетку на пересечении i -й строки и j -го столбца проставим номер линии l . Для преобразования в обратную сторону необходима матрица дешифратор размером $2n$, где в строке l каждому индексу линии l на первом месте поставим индекс i начала этой линии, а на втором — индекс j окончания линии. Таким образом, будет реализовано взаимно однозначное отображение между двойной и одинарной индексацией линий.

При помощи полученного соотношения переформулируем задачу для применения симплекс-метода. Если обозначить полученное отображение буквой T , то $T: p_{ij} \leftrightarrow p_i$, $T: g_{ij} \leftrightarrow g_i$, $T: c_{ij} \leftrightarrow c_i$, где $i, j = 1, 2, \dots, m$; $l = 1, 2, \dots, n$; $T(z_{ij}) = Z_{T(ij)} \leftrightarrow z_l$, где z может быть как переменной, так и заданной величиной.

С учётом смены индексации задача (2)–(3) будет записана так:

$$\sum_i p_{T(ij)} - \sum_k p_{T(jk)} = f_j - \sum_i g_{T(ij)}, \tag{7}$$

$$\sum_{T(ij)} c_{T(ij)} p_{T(ij)} \rightarrow \min, \tag{8}$$

$$p_{T(ij)} \geq 0. \quad (9)$$

Учитывая, что T — взаимно однозначное отображение, условие можно записать следующим образом:

$$F(x) = \sum_l c_l p_l \rightarrow \min. \quad (10)$$

Чтобы решить задачу линейного программирования, необходимо составить матрицу A , вектор b (ограничения) и вектор c , содержащий коэффициенты линейной формы, после чего симплекс-методом можно будет вычислить решение, которое будет удовлетворять первоначальным условиям задачи. Преобразование переменных величин в двойной индексации в переменные в одинарной происходит следующим образом: если переменная задана двойным индексом, её номер находится в клетке матрицы смежности, по этому номеру находится значение переменной с одним индексом. Если переменная задана одинарным индексом, ищем номер в матрице-«дешифраторе», и по нему находим два значения индекса.

Матрица $A = |a_{kl}|$ заполняется значениями при помощи «модифицированной» матрицы смежности и имеет гипербинарные значения. Индекс $j = 1, 2, \dots, m$ обозначает строки матрицы. Если значение в клетке ij матрицы смежности равно l , тогда $a_{ij} = 1$ (входящие в узел линии). Если значение в клетке ji матрицы смежности равно l , тогда $a_{ij} = -1$ (выходящие из узла линии).

Вектор $b = (b_j)^T$ строится следующим образом. Индекс $j = 1, 2, \dots, m$ нумерует координаты данного вектора b . По умолчанию вектор инициализирован значением $b: b_j = f_j$. Для каждого j перебираем значения $i = 1, 2, \dots, m$. В клетках, где значение равно l , вычитаем из полученного ранее значения b_j величину $g_{T(ij)} = g_l$.

Похожим образом строим вектор $c = (c_j)^T$. Проходим по значениям переменных $i = 1, 2, \dots, m; j = 1, 2, \dots, m$. Если значение в клетке равно l , то $g_{T(ij)} = g_l$. Для того чтобы поставить задачу не в виде (2) а в виде задачи (5) (в виде неравенств), необходимо задать вектор $b' = -b$ и матрицу $A' = -A$. Таким образом, матрица с условиями A'' и вектор b'' с ограничениями будут выглядеть так:

$$A'' = \begin{pmatrix} A \\ A' \end{pmatrix} \text{ и } b'' = \begin{pmatrix} b \\ b' \end{pmatrix}.$$

Необходимо найти максимум линейной формы $\left[\sum_l c_l p_l \right]$, чтобы определить минимум $F(x)$. Для этого заменим вектор c на вектор c'' , где $c'' = -c$.

УЧЁТ ТРЕБОВАНИЙ КОНТРАКТА И ДРУГИХ ВОЗМОЖНЫХ ОГРАНИЧЕНИЙ

Такой учёт означает, что поток на конце линии должен оставаться неотрицательным и должны полностью выполняться условия распределения на линии. В формулах это можно записать следующим образом:

$$p_{ij} + g_{ij} \geq 0, \quad (11)$$

$$-P_{ij} \leq q_{ij}. \quad (12)$$

Эти условия должны выполняться только для таких i и j , для которых существуют линии.

Используя ранее введённое соответствие индексов T :

$$T : g_{ij} = g_{T(ij)} \rightarrow g_l, \quad (13)$$

получаем:

$$-P_l \leq g_l. \quad (14)$$

В виде матриц это будет выглядеть следующим образом: $-Ep = g$, где E — единичная матрица размерности n , умноженная на -1 .

Для обозначения ограничения по максимальной мощности (току) введём переменную q_{ij} . Эта величина будет обозначать, какой максимальный поток может войти в линию, которая начинается в i и заканчивается в j .

Тогда запишем ограничение на максимальный ток в линии в следующем виде:

$$P_{ij} \leq q_{ij}. \quad (16)$$

Используя введённое ранее соответствие индексов:

$$T : q_{ij} = q_{T(ij)} \rightarrow q_l, \quad (17)$$

получаем:

$$P_l \leq q_l. \quad (18)$$

Записываем в матричном виде

$$Ep = q. \quad (19)$$

Вектор q рассчитывается следующим образом. Перебираем в цикле $l = 1, \dots, m$. Воспользовавшись дешифратором, определяем индекс ij . По нему найдём q_{ij} и присвоим его координате l . Аналогично определяем и вектор g .

Далее строим матрицы и векторы:

$$A''' = \begin{bmatrix} A \\ A' \\ -E \\ E \end{bmatrix}, \quad b''' = \begin{bmatrix} b \\ b' \\ g \\ q \end{bmatrix}, \quad c''' = c.$$

Таким образом, задача в матричном виде (4)–(6) поставлена. Решением этой задачи будет являться решение задачи распределения и оптимизации потоков с учётом требований удовлетворения контракта и потребления/распределения в узлах и линиях сети.

УЧЁТ ПОТРЕБЛЕНИЯ НА РЕВЕРСИВНЫХ ЛИНИЯХ

В рассмотрении в сети могут присутствовать реверсивные линии. Это означает, что в сети возможны потоки с противоположным направлением (отрицательные): $P_{ij} = -P_{ji}$.

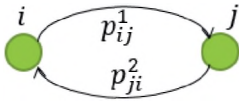


Рис. 1. Реверсивные линии без потребления

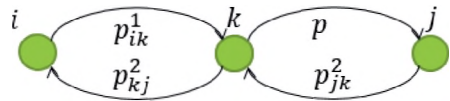


Рис. 2. Реверсивные линии с потреблением

Допустим, есть реверсивная линия ij без потребления на ней. Тогда можно заменить её двумя линиями ij и ji с противоположными направлениями, как показано на рис. 1.

Становится понятно, что два потока — p_{ij}^1 и p_{ji}^2 — моделируют поток на линии ij . Учитывая знак, получаем:

$$p_{ij} = p_{ij}^1 - p_{ji}^2, \quad (20)$$

где p_{ij}^1 и p_{ji}^2 неотрицательны, а p_{ij} может быть отрицательным.

В случае линии с потреблением на ней самой ситуация более сложная. Возможно следующее решение задачи. Пусть M — количество реверсивных линий, $M \geq 1$. Для каждой реверсивной линии ij введём новую «фиктивную» вершину $k = k(ij)$, где $k = m + 1, \dots, m + M$ (рис. 2).

После введения фиктивной вершины заменим каждую линию на 2 новые реверсивные линии, k будем считать большим m (числа вершин в сети), поэтому новая нумерация вершин не будет противоречить старой и новые линии не будут совпадать со старыми. Будем считать, что теперь потребление или поступление электроэнергии g_{ij} на линии ij будет происходить в вершине k и обозначим его $f_k = f_{k(ij)} = -g_{k(ij)}$. Поскольку полученные линии реверсивны, то они подлежат замене на две линии каждая: линии ik и ki для линии ik , и линии jk и kj для линии kj .

Введение нового узла означает, что в матрице смежности появится новая строка и столбец. В их пересечении и будет ячейка с новым номером линии. Сделаем также важное допущение, согласно которому потребления или распределение на новой линии сосредоточенно в середине линии:

$$c_{ik} = c_{jk} = c_{kj} = c_{ki} = \frac{1}{2} c_{ij}.$$

Произведённые ранее операции над линиями сети приводят к модификации матрицы шифратора/дешифратора и матрицы смежности. Необходимо построить «новую» матрицу смежности. Её размерностью будет $(m + M) \times (m + M)$, где M — число реверсивных линий; m — исходное число вершин графа.

Для $k = m + 1$ для всех линий изначального графа — если линия l не реверсивная, то, воспользовавшись «старым» дешифратором, запишем в «новую» матрицу значение с тем же номером, который был в «старой».

Если линия l реверсивна, тогда линия имеет двойной индекс: $ij = ij(l)$. Возьмём новый индекс фиктивной вершины $k = m + 1, \dots, m + M$ и запишем 1 в клетки с индексами ik, ki, kj, jk , после чего вычисляется «новый» вектор распределения потока в узле k : $f'_k = f'_{k(ij)} = -g_{k(ij)}$. В матрицу «биекции» размера $m \times m$ в клетку с номером ij пишется значение $k = k(ij)$. Матрица-

«дешифратор» теперь имеет размер $M \times 2$. В «дешифратор» в строку k на первое место пишется номер i , на второе — номер j . Нумерация начинается с $m + 1$ и каждый раз увеличиваем k на 1. Алгоритм повторяется над всеми реверсивными линиями до линии n .

Если искомая матрица построена, тогда необходимо найти новую сквозную индексацию линий. Их общее количество равно $n + 3M = n - M + 4M$. Присвоим $r = 1$. Перебираем все клетки матрицы по $i = 1, \dots, m + M$; $j = 1, \dots, m + M$. Если в клетке с номером ij значение, не равное 0, тогда клетке приписываем значение r и прибавляем к r единицу. Пройдя по всем клеткам, получаем новую сквозную индексацию линий $r = 1, \dots, n + 3M$ дополненного графа и «новую» матрицу смежности. После этого необходимо также использовать «дешифратор». Его новая размерность будет равна $(n + 3M) \times 2$. Точно также для каждого номера r , перебирая все клетки, на первое место в строке с номером r записывается индекс i -й строки, а на второе место — индекс j -го столбца.

В результате построено новое взаимно однозначное отображение T' :

$$T': ij \leftrightarrow r, \quad (21)$$

где $i, j = 1, 2, \dots, m + M$; $r = 1, \dots, n + 3M$.

После этого составим «новый» вектор притока/распределения f'_j . Для индексов $t = 1, 2, \dots, m$ компоненты вектора равны «старым». Для индексов $k = m + 1, \dots, m + M$ подставляется вычисленный ранее «новый» вектор f'_j .

Составим «новый» вектор c'_s . Проходим все значения $l = 1, 2, \dots, n$. Если линия l не является реверсивной, то, воспользовавшись «старым» шифратором, найдём двойной индекс $ij = ij(l)$. По этому индексу в «новой» матрице смежности в клетке с индексом ij получаем r , после присваивается $c'_r = c_l$. В результате будет пройдено $n - M$ «старых» линий и $n - M$ новых. Если линия реверсивная, тогда, воспользовавшись двойным индексом $ij = ij(l)$, по матрице отображения находится индекс $k = k(ij)$ фиктивной вершины. В получившейся «новой» матрице смежности берём последовательно значения r из клеток с индексами ik, ki, kj, jk . после чего присваиваем $c'_r = \frac{1}{2}c_l$. В итоге были заданы все $n + 3M$ координат для вектора c .

Составим «новый» вектор $g'_{ij} = g'_{ij(l)}$. Перебираем все значения по $l = 1, 2, \dots, n$. В случае, если линия в «старой» системе номеров не реверсивна, тогда, воспользовавшись «старым» шифратором, найдём двойной индекс $ij = ij(l)$ и в «новой» матрице смежности в клетке с индексом ij получим r , после этого присвоим $g'_r = g_l$. В результате будет перебрано $n - M$ «старых» линий и $n - M$ новых линий. Потребления на оставшихся $4M$ линиях нет, так как с реверсивных линий их значения потребления перенесли на фиктивные вершины. В итоге определены все $n + 3M$ координат для вектора g' .

Составим «новый» вектор $q'_{ij} = q'_{ij(l)}$. Перебираем все значения по $l = 1, 2, \dots, n$. Если линия в «старой» системе номеров не реверсивна, тогда, воспользовавшись «старым» шифратором, найдём двойной индекс $ij = ij(l)$ и в «новой» матрице смежности в клетке с индексом ij получим r , после чего присвоим $q'_r = q_l$. В результате перебрано $n - M$ «старых» линий и $n - M$ новых линий. Если в старой системе номеров линия реверсивна, тогда, воспользовавшись двойным индексом $ij = ij(l)$, по матрице отображения найдём индекс $k = k(ij)$ фиктивной вершины. Из клеток с номерами ik, kj в «новой» матрице

смежности возьмём последовательно значения r и присвоим $q'_r = q_r$. Для клеток с индексами jk, ki поступаем таким же образом. В итоге все $n + 3M$ координат для вектора q' заданы.

В итоге задана новая постановка задачи, приведённая к виду условий (5)–(7), которая может решаться методом без реверсивных линий (описанным выше). Решение примет вид $\{p'_r\}, r = 1, \dots, n + 3M$, после этого будет необходимо перевести решение в систему старых переменных.

Допустим, есть некоторое решение выше обозначенной задачи. В задаче изначально находилось n «старых» значений, получилось $n + 3M$ «новых» значений, среди которых $n - M$ не реверсивных и $4M$ принадлежали к реверсивным линиям. Для нереверсивных линий (переменных) «новая» переменная означает поток в начале линии. Поток в конце линии получается из начального потока вычитанием потребления на линии. В случае реверсивных линий поступаем следующим образом:

Допустим, $k = k(ij)$ — индекс фиктивной вершины. Чтобы получить поток в начале линии ij , сложим значения p_{ik} и $-p_{ki}$, причём, если результат положителен, то поток вытекает из вершины i , если результат отрицателен — поток втекает в вершину i . Таким образом, получим поток в начале линии. Далее, чтобы получить поток в начале линии ij , сложим значения p_{ik} и $-p_{ki}$, причём, если результат положителен, поток вытекает из вершины j в обратную сторону. Таким образом, получим поток в конце линии.

Уточним число переменных после этой операции. В результате для линий с отсутствием реверса получим $n - M$ переменных и для линий с реверсом — $2M = 4M/2$ переменных. Для линий с отсутствием реверса удвоим число $n - M$ переменных и сложим его с $2M$ переменных для реверсивных линий. Получается по n переменных в начале потока и потока в конце, в итоге $2n$ переменных. Запишем «таблицу решений» задачи. Все значения таблицы по умолчанию инициализируются 0. Поток в начале линии обозначается первой строкой, поток в конце линии — второй строкой. Индекс столбца равен индексу линии — переменной.

Перебираем индексы $r = 1, \dots, n + 3M$. Воспользовавшись «новой» матрицей-«дешифратором», получаем двойной индекс $i'j' = i'j'(r)$. В случае, когда оба индекса $i', j' \leq m$, линия не реверсивная и $i = i', j = j'$. В «старой» матрице смежности находим «старый» индекс линии $l = l(ij)$ и $g_l = g_{l(ij)}$. Присвоив $p_l = p'_r$, рассчитываем $y_l = p_l + g_l$, учитывая, что отток мощности имеет знак «-». В столбец с индексом l «таблицы решений» ставится p_l и y_l в первую и во вторую строку, соответственно. Если один из номеров — i' или j' — больше m , то это означает, что $i'j'$ есть «часть» изначально реверсивной линии. По построению у нас возможны только два случая, разберём их.

Случай 1: $j' > m$. Переобозначим $k = j'$, тогда получаем линию $i'k$. Устанавливаем двойной индекс по специальной матрице-«дешифратору» $ij = ij(k)$. Найдём в «старой» матрице смежности индекс линии $l = l(ij)$. Случай 1 подразделяется на случаи 1.1 и 1.2.

Случай 1.1: если $i' = i$ совпало с началом линии. В таком случае к значению в клетке пересечения столбца l и первой строки в «таблицу решений» прибавляем p'_r (см. рис. 2, линия ik).

Случай 1.2: если $i' = j$ совпало с концом линии. В таком случае к значению в клетке пересечения столбца l и второй строки в «таблицу решений» прибавляем $-p'_r$ (см. рис. 2, линия jk).

Случай 2: $i' > t$. Переобозначим $k = j'$, тогда получаем линию kj' . Устанавливаем двойной индекс по специальной матрице-«дешифратору» $ij = ij(k)$. Найдём в «старой» матрице смежности индекс линии $l = l(ij)$. Случай 2 подразделяется на случаи 2.1 и 2.2.

Случай 2.1: если $j' = j$ совпало с концом линии. В таком случае к значению в клетке пересечения столбца l и второй строки в «таблицу решений» прибавляем p'_r (поток в начале линии kj' ki' , по сути, совпадает с потоком в начале линии ij) (см. рис. 2, линия kj).

Случай 2.2: если $j' = i$ совпало с началом линии. В таком случае к значению в клетке пересечения столбца l и первой строки в «таблицу решений» прибавляем $-p'_r$. (поток в начале линии ki' , по сути, совпадает с потоком в начале линии ij) (см. рис. 2 линия ki).

Задача оптимизации и распределения потоков в электросети с учётом требований контрактов и возможности наличия реверсивных линий, а также возможности потребления и генерации мощности (тока) на линиях и в вершинах была решена.

При решении данной задачи с реверсивными линиями пришлось менять целевую функцию, а также матрицы и вектора ограничений и условий. Однако верна следующая теорема, которая обосновывает практическую применимость метода для вычислений.

Теорема. При возможности существования в системе хотя бы одной реверсивной линии метод даёт решение не хуже (в рамках подсчёта целевой функции), чем при полном запрете на существование реверсивных линий.

Существенным для этого утверждения является то, что для замены одной реверсивной линии выбраны четыре новые линии, каждая из которых имеет половину длины реверсивной (симметрия), и происходящий при такой замене перенос потребления мощности с линии на новый «фиктивный» узел. Актуальность этого утверждения происходит от того, что признак реверсивности конкретной линии может задаваться пользователем дополнительно при практическом использовании вышеописанного метода [Кармен и др., 2007]. Утверждается, что такие изменения задачи не сделают хуже результат расчёта системы, а при практическом применении часто улучшает (по сравнению значений целевой функции).

ЗАКЛЮЧЕНИЕ

Вышеописанный метод решения задачи распределения потоков в сложной инженерной системе позволяет свести решение задачи об оптимизации потоков с учётом различных условий (реверсивных линий, присутствия внешних источников/потребителей) к задаче, решаемой стандартными методами линейного программирования. В электросетях это имеет большое значение, так как методы линейного программирования хорошо подлежат параллелизации, в учёте перехода энергосистем на сетевое управление это имеет решающее значение.

ЛИТЕРАТУРА

- [Вайнзихер, 2008] *Вайнзихер Б. Ф.* Электроэнергетика России 2030: Целевое видение. М.: Альпина Бизнес Бук, 2008.
- [Дорофеев, Макаров, 2009] *Дорофеев В. В., Макаров А. А.* Активно-адаптивная сеть — новое качество ЕЭС России // Энергоэксперт. 2009. № 4(15).
- [Карманов, 2004] *Карманов В. Г.* Математическое программирование: учеб. пособие. 5-е изд., стереотип. М.: Физматлит, 2004. 264 с.
- [Кармен и др., 2007] *Кармен Т. Х., Лайзерсон Ч. Е., Ривест Р. Л., Штейн К.* Алгоритмы: построение и анализ. 2-е изд. М.: ИД «Вильямс», 2007. 1296 с.
- [Ноздрин, 2013] *Ноздрин Д. А.* Интеллектуальное управление инженерными сетями // Управление большими системами: Материалы 10-й Всерос. шк.-конф. молодых учёных. Уфа: УГАТУ, 2013.
- [Рыбников, 2010] *Рыбников К. К.* Введение в дискретную математику и теорию решения экстремальных задач на конечных множествах: уч. пособие для студентов вузов. М.: Гелиос АРВ, 2010. 320 с.
- [Tarjan, 1983] *Tarjan R.* Data Structures and Network Algorithms. Society for Industrial and Applied Mathematics, 1983.

REDUCING THE TASK OF OPTIMIZATION IN INTELLIGENT ENGINEERING NETWORKS TO THE PROBLEM TO BE SOLVED BY THE SIMPLEX METHOD

D. A. Nozdrin

Moscow Institute of Electronics and Mathematics in National Research University
Higher School of Economics (MIEM NRU HSE)

The work is devoted to consideration of the balance equations in intelligent engineering networks, in particular, in networks and the problem of optimization of flows in terms of network architecture and dividing the network into zones. An important task facing many complexes task is the calculation of the optimal distribution of flows, such that the selected target function is sought to the maximum (minimum). This issue is relevant for most intelligent utility networks with different imposed conditions. If we impose restrictions on bandwidth and/or constraints on control actions, this task becomes a task flow management, which includes different types of ow problems. Such tasks can be the problem of determining the existence of flow, transport problems, the task of building topology and other Varying conditions and restrictions tasks can be adapted algorithms under the gas distribution networks, water supply, electricity, transport network, a cellular network. In this article, the problem is considered in the context of electricity.

Keywords: equation, network engineering, network architecture, function, condition, the flow control algorithm, optimization.

Nozdrin Dmitriy Alexeevich — postgraduate student, dmitriy.nozdrin@gmail.com

АРХИТЕКТУРНО-ЗАВИСИМАЯ ДЕКОМПОЗИЦИЯ В МЕТОДИКАХ СУПЕРКОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ

Н. В. Старостин, М. А. Панкратова

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

Численное моделирование физических процессов в параллельной среде предполагает распределение аппроксимационной сетки по процессорам. Результат распределения во многом определяет количество межпроцессорных коммуникаций, коллизий во время пересылок информации, а, следовательно, и время выполнения параллельной программы. В работе приводится постановка задачи архитектурно-зависимой декомпозиции, предлагаются пути решения задачи.

Ключевые слова: архитектурно-зависимая декомпозиция, метод рекурсивной бисекции.

ВВЕДЕНИЕ

Многие методики численного моделирования физических процессов основываются на аппроксимационных сетках, которые моделируются с помощью графов. На практике сеточные структуры имеют большие порядки ($10^6 \dots 10^9$ узлов), что приводит к необходимости их распределения по узлам параллельной вычислительной сети. Эффективность работы программы в параллельной среде во многом зависит от интенсивности межпроцессорных коммуникаций, которую можно сократить за счёт удачного распределения расчётной сетки по узлам вычислительной сети. В работе приводится постановка задачи архитектурно-зависимой декомпозиции расчётной сетки, предлагаются пути решения задачи, основанные на рекурсивной бисекции графов.

1. ПОСТАНОВКА ЗАДАЧИ

Взвешенный граф $G(V, F, w, u)$ моделирует расчётную сетку, $V = \{v_1, \dots, v_n\}$ — множество узлов сетки, $E \subseteq V^{(2)}$ — связи между узлами, веса вершин $w(v_i) \in N$, $i = \overline{1, n}$ — затраты на расчёт, веса рёбер $u(v_i, v_j) \in i$, $i, j = \overline{1, n}$ — интенсивность коммуникаций.

Матрица $T = (t_{ij})_{k \times k}$ моделирует вычислительную сеть. Элемент матрицы $t_{ij} \in N$ соответствует оценке затрат на передачу данных между парой вычислителей.

Решением задачи является вектор $x = \{x_1, \dots, x_n\}$, $x_i \in \{1, \dots, k\}$, $i = \overline{1, n}$, где i -й элемент соответствует узлу расчётной сетки, а значение x_i определяет узел вычислительной сети, на который он распределён. Для ускорения процесса параллельного расчёта на распределённой аппроксимационной сетке важно сбалансировать загрузку процессоров, чтобы время работы каждого процессора было максимально одинаковым. Пусть

$$W = \sum_{i=1}^n w(v_i)$$

— суммарный вес узлов сетки,

$$W_j = \sum_{\{i: x_i=j, i=1, n\}} w(v_i), \quad j = \overline{1, k}$$

— загрузка (оценка относительного времени работы) вычислительного узла j ,

$$\tilde{W} = \frac{W}{k}$$

— идеальная загрузка вычислительных узлов. Зададим $\varepsilon > 0$ — коэффициент отклонения от идеальной загрузки. Тогда неравенства

$$\left| \frac{W_j}{\tilde{W}} - 1 \right| \leq \varepsilon, \quad j = \overline{1, k} \quad (1)$$

задают ограничения по балансировке нагрузки вычислительных узлов.

Зададим функцию

$$\beta(x, v_i, v_j) = u(v_i, v_j) t_{x, x_j},$$

которая определяет оценку затрат на коммуникационный обмен между парой узлов вычислительной сети.

Проблема моделирования коммуникационных обменов в программе, работающей в параллельной вычислительной сети, связана с тем, что очень сложно, а зачастую практически невозможно спрогнозировать: точное время каждой коммуникации; возникновение коллизий на уровне сетевых устройств; снижение пропускной способности каналов в связи с повышением интенсивности обменов; маршруты, по которым происходит передача пакетов. Всё это значительно снижает адекватность применения модели в терминах планирования коммуникационных обменов. С практической точки зрения удобнее оперировать оценками затрат на всю совокупность коммуникационных обменов [Pellegrini, 2012]. Нижняя оценка предполагает случай, когда все обмены происходят одновременно и друг с другом не конфликтуют. В данном случае имеет место минимаксный критерий:

$$F_1(x) = \max_{(v_i, v_j) \in V} \beta(x, v_i, v_j) \rightarrow \min. \quad (2)$$

Верхняя оценка предполагает, что все обмены друг с другом конфликтуют и происходят последовательно. В данном случае имеет место аддитивный критерий:

$$F_2(x) = \sum_{(v_i, v_j) \in V} \beta(x, v_i, v_j) \rightarrow \min. \quad (3)$$

2. ПРОБЛЕМА ОРГАНИЗАЦИИ ИСПОЛНЕНИЯ ПАРАЛЛЕЛЬНОЙ ПРОГРАММЫ

Как правило, при организации исполнения параллельной программы заранее известны только структура расчётной сетки и общее число вычислительных узлов, которое будет выделено для расчёта. Структура выделенного участка

вычислительной сети определяется только в момент начала расчёта, когда планировщик выделяет очередной задаче требуемое количество вычислителей. С другой стороны, в силу того, что размер исходных данных очень велик и не «умещается» в памяти одного процессора, аппроксимационная сетка должна быть декомпозирована ещё до начала расчёта. В результате в начале расчёта исходные данные оказываются уже распределёнными по выделенным процессорам, т. е. каждый процессор владеет только частью исходных данных. Таким образом, на этапе декомпозиции данных информация о топологии выделенного участка вычислительной сети не известна. Поэтому в задачах (1), (2) и (1), (3) элементы матрицы принимают значения $t_{ij} = \text{const}$, $i \neq j$, $t_{ii} = 0$, $i, j = \overline{1, k}$. В этой постановке задача (1), (2) соответствует задаче минимизации ширины ленты графа [Джордж, Лю, 1984], а задача (1), (3) — классической задаче k -разбиения графа [Батищев, Старостин, 2000]. Инструменты для решения обозначенных задач реализованы с помощью многоуровневых подходов в рамках библиотеки MATRUZ [Старостин, Филимонов, 2013].

3. МЕТОД РЕКУРСИВНОЙ БИСЕКЦИИ ДЛЯ РЕШЕНИЯ ПОСТАВЛЕННОЙ ЗАДАЧИ

На основе вышесказанного можно предложить два подхода к решению поставленной задачи. В случае, если мы располагаем информацией о топологии участка вычислительной сети, выделенного для расчёта данной сетки, до запуска расчёта, можно решать непосредственно поставленную задачу архитектурно-зависимой декомпозиции. Предлагается подход, основанный на рекурсивной бисекции графов расчётной сетки и вычислительной сети. Такой подход будем называть одноэтапным. Ниже приведён псевдокод алгоритма, реализующего данный подход.

В том случае, когда мы заранее не располагаем информацией о топологии выделенного участка вычислительной сети, необходимо решать задачу в два этапа. На первом этапе необходимо декомпозировать расчётную сетку на заданное количество частей, определяемое степенью распараллеливания расчёта. Предлагается осуществлять декомпозицию многоуровневым итерационным методом. На втором этапе необходимо перераспределить уже декомпозированную расчётную сетку по узлам выделенного участка вычислительной сети (информация о нём становится известной после запуска расчёта) с целью снижения затрат на межпроцессорные коммуникации. Предлагается осуществлять перераспределение методом рекурсивной бисекции графов декомпозированной расчётной сетки и вычислительной сети. Такой подход будем называть двухэтапным. Ниже приведён псевдокод алгоритма, реализующего данный подход.

4. ОДНОЭТАПНЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ

На листинге 1 приведён псевдокод алгоритма, реализующий одноэтапный подход к решению поставленной задачи, основанный на рекурсивной бисекции графов.

```

1. Run1(graph G, matrix T, out int[] X)
2. {
3.   if(size(T) == 1)
4.     map(G, T, out X)
5.   else
6.     {
7.       bisect(G) -> graph G1, graph G2;
8.       spectrbisect(T, G1, G2) -> matrix T1, matrix T2;
9.       Run1(G1, T1, out X);
10.      Run1(G2, T2, out X);
11.    }
12. }

```

Листинг 1. Одноэтапный алгоритм

Алгоритм принимает на вход граф, моделирующий расчётную сетку, матрицу, моделирующую участок вычислительной сети, и сохраняет полученное решение в массиве (строка 1). В основной части алгоритма (строки 6–11) происходит бисекция расчётной сетки многоуровневым итерационным алгоритмом (строка 7), бисекция вычислительной сети спектральным методом [Якобовский, 2004; Fiedler, 1975a, b] пропорционально произведённой бисекции расчётной сетки (строка 8), и алгоритм рекурсивно повторяется на полученных подграфах (строки 9, 10). В том случае, если размер вычислительной сети оказался равным 1, то бисекция не происходит, а вершины соответствующего подграфа расчётной сетки назначаются на данный вычислительный узел (строки 3, 4).

5. ДВУХЭТАПНЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ

На листинге 2 приведён псевдокод алгоритма, реализующий двухэтапный подход к решению поставленной задачи, основанный на декомпозиции расчётной сетки и рекурсивной бисекции графов.

Алгоритм принимает на вход граф, моделирующий расчётную сетку; матрицу, моделирующую участок вычислительной сети, и сохраняет полученное решение в массиве (строка 1). В первой части алгоритма (строки 3, 4) происходит декомпозиция расчётной сетки на количество частей, соответствующее размеру вычислительной сети (строка 3). Декомпозиция осуществляется многоуровневым итерационным алгоритмом. Затем строится суперграф, вершины которого соответствуют частям декомпозированной сетки, а ребро между вершинами существует в том случае, если в графе расчётной сетки между соответствующими частями существовало хотя бы одно ребро (строка 4).

```

1. Run2(graph G, matrix T, out int[] X)
2. {
3.   partition p = decomp(G, size(T));
4.   graph g = coars(G, p);
5.   Run1(g, T, out X);
6. }

```

Листинг 2. Двухэтапный алгоритм

Заметим, что в данной части алгоритма не требуется информация о топологии вычислительной сети. Во второй части алгоритма, когда становится доступной информация о топологии выделенного участка вычислительной сети, происходит перераспределение частей декомпозированной сетки по узлам вычислительной сети одноэтапным методом (строка 5).

6. МНОГОУРОВНЕВЫЙ ИТЕРАЦИОННЫЙ АЛГОРИТМ ДЕКОМПОЗИЦИИ ГРАФА

Кратко опишем основные идеи алгоритма. Ключевая идея многоуровневого алгоритма [Старостин, Филимонов, 2013] разбиения состоит в следующем: вместо того, чтобы строить k -разбиение непосредственно для исходного графа, сначала строится ряд его приближений (загрублений). Каждое загрубление уменьшает (редуцирует) размерность исходной задачи. Процесс редукции продолжается до тех пор, пока порядок графа не снизится до сотен или даже десятков. На графе таких размерностей и отыскивается так называемое начальное разбиение. Полученное решение используется для построения решения для исходной задачи. Это достигается путём выполнения серии переносов решения задачи меньшей размерности на задачу большей размерности с последующим улучшением перенесённого решения. Таким образом, работу многоуровневого алгоритма можно разделить на три фазы: первая — фаза загрузки, когда производится ряд последовательных редукций размерности задачи, вторая — фаза поиска начального разбиения и третья — фаза восстановления решения, когда производится серия последовательных переносов решения задачи меньшей размерности на менее редуцированную задачу с последующим улучшением спроецированного решения. Работа алгоритма для задачи k -разбиения показана на рис. 1.

Основным недостатком описанного алгоритма является то, что он не обладает «обратной связью» — не используют информацию о предыдущих итерациях своей работы. Предлагается многоуровневый итерационный алгоритм задачи разбиения [Старостин, Филимонов, 2013]. Алгоритм на вход получает не только исходные данные задачи, но и одно из допустимых решений задачи.

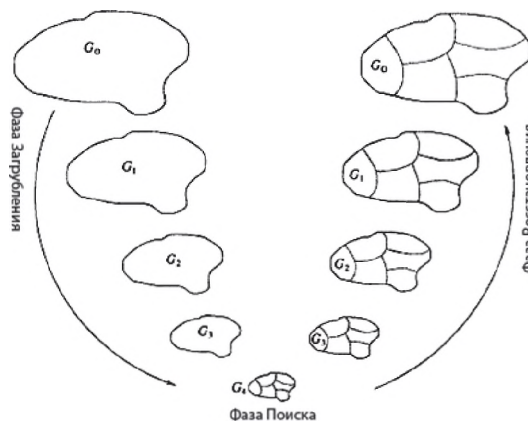


Рис. 1. Многоуровневый алгоритм k -разбиения графа



Рис. 2. Многоуровневый итерационный алгоритм k -разбиения графа

Цель алгоритма — исследование областей вокруг начального решения и поиск лучшего решения. Схема итерационного алгоритма заключается в следующем: строится начальное решение многоуровневым алгоритмом, каждая следующая итерация использует решение, полученное на предыдущей итерации для редукции графа в многоуровневом алгоритме. Работа итерационного алгоритма показана на рис. 2.

7. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Целью данной работы было выяснить, как влияет предварительная декомпозиция расчётной сетки на качество получаемого решения. Был проведён вычислительный эксперимент на топологии вычислительной сети типа Fat tree (рис. 3).

Примеры расчётных сеток были взяты из библиотек университета Флориды (<http://www.cise.ufl.edu/research/sparse/matrices/>) и Гринвича (<http://staffweb.cms.gre.ac.uk/~wc06/partition>). Для эксперимента использован разработанный академический программный код на платформе $c\#$.NET Framework. В табл. 1 показаны типовые результаты работы предложенных методов решения поставленной задачи.

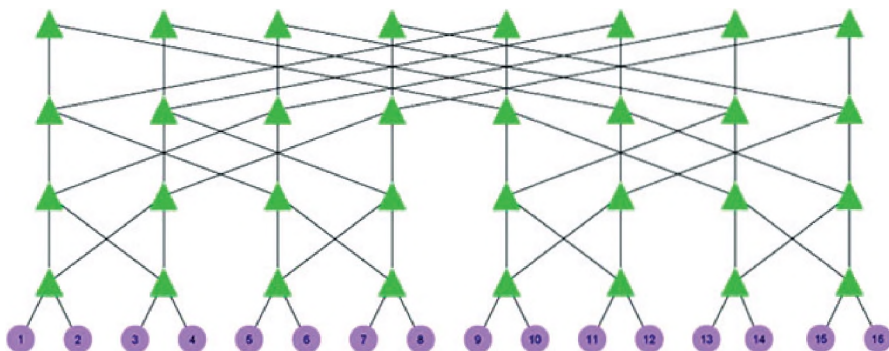


Рис. 3. Топология вычислительной сети Fattree

Таблица 1. Вычислительный эксперимент. Использование рекурсивной бисекции

Сетка	Размер	BC	Размер	1-этапный	Время [с]	2-этапный	Время [с]
gr_30_30	900	fattree16	16	882	0,54	772	2,91
bcsttk29	13992	fattree16	16	39043	7,387	33817	86,442
bcsttk32	44609	fattree16	16	65070	11,946	57498	129,319
bcsttk33	8738	fattree16	16	92670	13,993	84455	174,572
barth5	15606	fattree16	16	1452	1,451	1399	11,673
vibrobox	12328	fattree16	16	63050	10,714	63512	84,799
ef_sphere	16386	fattree16	16	2959	1,579	3232	12,919
ef_4elt2	11143	fattree16	16	1813	1,202	1524	9,433
fe_rotor	99617	fattree16	16	37971	15,919	31652	132,22

Также были получены результаты с помощью библиотеки Scotch [Pellegrini, 2012]. В табл. 2 приведены результаты, полученные с помощью данной библиотеки и с использованием предварительной декомпозиции расчётной сетки (аналог двухэтапного метода).

Таблица 2. Вычислительный эксперимент. Использование Scotch

Сетка	Размер	BC	Размер	Scotch	Время [с]	Scotch + Decomp	Время Scotch [с]	Время Decomp [с]
gr_30_30	900	fattree16	16	530	0,004	442	0,001	2,872
bcsttk29	13992	fattree16	16	36319	0,092	20427	0,001	86,59
bcsttk32	44609	fattree16	16	81906	0,624	40289	0,001	130,337
bcsttk33	8738	fattree16	16	77552	0,168	51877	0,001	171,051
barth5	15606	fattree16	16	2022	0,042	925	0,001	11,472
vibrobox	12328	fattree16	16	57269	0,159	43693	0,001	85,819
ef_sphere	16386	fattree16	16	3045	0,052	1705	0,001	12,778
ef_4elt2	11143	fattree16	16	1209	0,032	921	0,001	9,34
fe_rotor	99617	fattree16	16	92742	0,461	21007	0,001	131,148

Следует отметить, что в результатах, полученных в эксперименте, алгоритм, реализованный в библиотеке Scotch, не всегда использует все узлы вычислительной сети для отображения расчётной сетки. Однако в реализации предложенного алгоритма жёстко учитывается балансное ограничение (1). В силу этого можно сделать вывод о разных математических постановках, используемых при исследовании проблемы. Таким образом, результаты, полученные алгоритмом Scotch и предложенным алгоритмом, объективно сравнивать нельзя.

Тем не менее, по результатам обоих проведённых экспериментов можно сказать, что предварительная декомпозиция расчётной сетки даёт возможность получить решение лучшего качества. Это объясняется тем, что качественная декомпозиция позволяет уменьшить общее число внешних связей, влияющих на значение критерия.

ЗАКЛЮЧЕНИЕ

В работе была поставлена задача архитектурно-зависимой декомпозиции. Были рассмотрены особенности применения данной задачи на практике в качестве инструмента, повышающего эффективность параллельных расчётов в суперкомпьютерном моделировании. Предложены два метода решения поставленной задачи, применение которых зависит от технологии проведения параллельного расчёта. На численном эксперименте показаны преимущества метода с предварительной декомпозицией расчётной сетки.

ЛИТЕРАТУРА

- [Батищев, Старостин, 2000] *Батищев Д. И., Старостин Н. В.* *k*-разбиение графов // Вестник ННГУ «Математическое моделирование и оптимальное управление». Н. Новгород, 2000. С. 37–25.
- [Джордж, Лю, 1984] *Джордж А., Лю Дж.* Численное решение больших разреженных систем уравнений / Пер. с англ. М.: Мир, 1984. 333 с.
- [Старостин, Филимонов, 2013] *Старостин Н. В., Филимонов А. В.* Разработка и реализация параллельного многоуровневого алгоритма равномерного разбиения нераспределенного графа // Материалы 13-й Всерос. конф. Н. Новгород, 14–16 ноября 2013. Н. Новгород: Изд-во Нижегород. госун-та, 2013. С. 243–247.
- [Якобовский, 2004] *Якобовский М. В.* Обработка сеточных данных на распределённых вычислительных системах // Вопросы атомной науки и техники. Сер. Мат. моделирование физ. процессов. 2004. Вып. 2. С. 40–53.
- [Fiedler, 1975a] *Fiedler M.* Eigenvectors of Acyclic Matrices // *Czechoslovak Mathematical J.* 1975. V. 25(100). P. 607–618.
- [Fiedler, 1975b] *Fiedler M.* A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory // *Czechoslovak Mathematical J.* 1975. 25(100). P. 619–633.
- [Pellegrini, 2012] *Pellegrini F.* Scotch and LibScotch 6.0 User's guide. France, University of Bordeaux 1 & LaBRI, 2012.

ARCHITECTURE DEPENDENT DECOMPOSITION IN SUPERCOMPUTING SIMULATION

N. V. Starostin, M. A. Pankratova

Lobachevsky State University of Nizhni Novgorod

Physical processes numerical simulation uses mapping of the approximation grid to the computing system. Number of interprocessor communications, communication conflicts, execution time depend on the quality of mapping. The paper presents the mathematical model of the architecture dependent decomposition problem and methods to solve this problem.

Keywords: architecture dependent decomposition, recursive bisection method.

Starostin Nikolay Vladimirovich — PhD, nvstar@mail.ru

Pankratova Margarita Alexandrovna — postgraduate student, e-mail: pankratova_margarita@mail.ru

К ЗАДАЧЕ РАЗРАБОТКИ МЕТОДИКИ ПОСТРОЕНИЯ РАСПРЕДЕЛЁННОЙ ИНФОКОММУНИКАЦИОННОЙ СИСТЕМЫ ДЛЯ НАУЧНОЙ И НАУЧНО-УЧЕБНОЙ ДЕЯТЕЛЬНОСТИ

С. К. Шикота

Научный центр Российской академии наук в Черноголовке

Ставится задача по разработке методики построения распределённой инфокоммуникационной системы для обеспечения научной и научно-учебной деятельности творческого коллектива. Проводится анализ существующих инфокоммуникационных систем — порталов и платформ — с целью выявления общих компонентов, которые используются для управления данными, метаданными, вычислительными ресурсами, правами доступа, добавлением заданий и мониторингом, а также для обработки результатов моделирования.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 14-07-00552-а).

Ключевые слова: распределённая инфокоммуникационная система, свободно распространяемое программное обеспечение, научная инфраструктура.

ВВЕДЕНИЕ

Высокий темп обмена научной информацией, повышение темпа научных исследований, увеличение числа международных проектов, которые играют определяющую роль в создании распределённых виртуальных научных коллективов, достаточно жёсткие требования финансирующих организаций к результативности научных исследований — эти факторы характеризуют научные исследования в XXI веке. Как следствие, можно наблюдать значительное усиление конкуренции в научной среде.

Для того чтобы оставаться конкурентноспособным, коллективу необходимо использовать в своей работе современные информационно-коммуникационные технологии. На сегодняшний день предпочтительно используемыми технологиями и средствами, позволяющими хранить и обрабатывать большие массивы данных, совместно работать над документами участникам распределённых коллективов, расположенных, зачастую, на различных континентах, проводить видеосеминары, читать лекции и т.д. являются технология грид, облачные технологии, а также такие средства как веб-порталы.

ПОСТАНОВКА ЗАДАЧИ

Для выработки тактики разработки и внедрения новых информационно-коммуникационных технологий (ИКТ) и систем требуется проведение глубоких исследований по составу и характеру научной активности коллектива с целью выявления общих черт деятельности для дальнейшего использования результатов анализа при создании востребованных компонентов. Анализ подходов,

развиваемых в российских научных центрах, представлен в статье [Шикота, 2013]. При этом важно принимать во внимание уже сложившийся стиль научной работы коллектива, который зависит от многих факторов, таких как область исследований, законодательная база, существующие административные подходы, культурные особенности коллектива [Kenna, Berche, 2011]. Подробный анализ составляющих научной деятельности приведён в статье [Щур, 2012]. Не стоит забывать, что зачастую пользователям необходим не только доступ к разным вычислительным ресурсам, но и возможность контролировать весь процесс выполнения их задач на всех доступных ресурсах, получать информацию о месте выполнения, состоянии использования приложений, о заданных параметрах и используемых файлах.

Поставленная задача достаточно сложна, она включает в себя как необходимость технологических решений, так и усложняется человеческим фактором. Системы разрабатывают специалисты по ИКТ, а предназначены они для учёных-предметников, которые, во-первых, не являются профессионалами в области ИКТ, а во-вторых, достаточно консервативны и неохотно меняют стиль работы. Поэтому очень важно изучить реальные потребности научного сотрудника, студента и аспиранта в таких технологиях. Разработчикам важно терпеливо изучить и учесть пожелания исследователя к конфигурации рабочего места, обработке результатов исследований и стилю общения с коллегами. Очень важно постараться вовлечь научного сотрудника в процесс разработки системы. Причастность к работе важна как технологически, так и психологически.

АНАЛИЗ ТЕХНОЛОГИЙ ДЛЯ РАСПРЕДЕЛЁННОЙ НАУЧНОЙ ИНФРАСТРУКТУРЫ

Для наших целей наибольший интерес представляют такие информационно-вычислительные среды для проведения научных исследований, доступ к которым осуществляется через интуитивно понятный пользовательский интерфейс. Также представляется целесообразным разработка такого инструментария, который позволил бы пользователю сравнительно легко интегрировать доступные ему вычислительные средства и программное обеспечение в необходимую для выполнения конкретного проекта программно-аппаратную среду.

Наш подход направлен конкретно на организацию работы научной и научно-учебной группы. В качестве модели мы берём лабораторию по проведению исследований в области прикладных сетевых технологий и организацию научной и учебной работы по обучению магистров и аспирантов в области прикладных инфокоммуникационных систем.

Наиболее распространённая на сегодняшний день технология для проведения совместных исследований распределёнными научными коллективами — технология грид.

Концепция технологии грид изложена в статьях [Foster, 2001, 2002]. Грид — это среда, которая позволяет пользователям видеть все ресурсы грид так, как если бы они были расположены непосредственно в его лаборатории. Хотя, на самом деле, они могут находиться в любой точке земного шара. При этом для незарегистрированных пользователей эти ресурсы не видны совсем.

В последнее время растёт интерес к «облачным» технологиям [Керес, 2011], которые позволяют пользователю получить в своё распоряжение некоторый информационный ресурс, расположенный у сервис-провайдера. В отличие от грид-технологии, предусматривающей возможность для пользователя предоставлять свои ресурсы другим участникам виртуального коллектива, облачные технологии позволяют получать ресурсы у сервис-провайдера, который обеспечивает функционирование ресурсов собственными силами, а пользователь получает доступ к этим ресурсам посредством сети Интернет.

Существует несколько уровней доступа к «облачным» ресурсам:

- уровень IaaS (Infrastructure as a service) — провайдер предоставляет пользователю доступ к физическим ресурсам: серверам, хранилищам данных, сетям и т. д.;
- уровень SaaS (Software as a service) — провайдер предоставляет пользователю доступ к готовому приложению, запущенному в «облаке»;
- уровень PaaS (Platform as a service) — пользователь получает доступ к платформе, на которой он может быстрым и удобным способом запускать свои собственные задачи.

Для решения поставленной задачи наиболее подходящим представляется использование ресурсов в рамках концепции PaaS — когда имеется установленная операционная система и набор средств для разработки приложений под данную платформу, что позволит упростить взаимодействие пользователей и поставщиков услуг.

Ещё одним важным направлением развития инфокоммуникационных систем для совместной работы распределённых коллективов является создание различных веб-ориентированных сервисов, использующих концепцию Web 2.0 [Shang Shari, 2011]. Под Web 2.0 принято понимать системы, наполнением которых занимаются сами пользователи: форумы, блоги, wiki, социальные сети и т. д. Подобные системы являются прекрасным инструментом взаимодействия исследователей, позволяющим достаточно оперативно обмениваться информацией.

Использование концепции Web 2.0 в рамках поставленной задачи позволит, например, решить проблему оперативного взаимодействия исследователей, а также упростить работу со студентами.

АНАЛИЗ СУЩЕСТВУЮЩИХ ПЛАТФОРМ И ПОРТАЛОВ

Такой анализ позволяет выделить общие компоненты, которые используются для управления данными, метаданными, вычислительными ресурсами, правами доступа, добавлением заданий и мониторингом, а также для просмотра результатов моделирования.

Ниже приведён анализ таких систем для совместной работы как: AccessGrid (<http://www.accessgrid.org/>), TGViz [Binns, 2007], SIDGrid [Levow, 2007], OLSGW [Wu, 2008] и UniHUB (<https://unihub.ru/about>).

Среда AccessGrid представляет собой инструментарий, позволяющий пользователям не только обмениваться аудио- и видеoinформацией, но и совместно просматривать презентации, веб-страницы, обмениваться файлами, совместно редактировать документы и т. д.

Средство коллективной работы AccessGrid использует распределённую модель обработки информации, что позволяет легко наращивать мощность и возможности системы простым включением дополнительных компонент, как аппаратных, так и программных [Щур, 2010].

Ещё одно средство для совместной работы — TeraGrid Visualization Portal (TGViz), разработанный в рамках проекта TeraGrid (<http://en.wikipedia.org/wiki/TeraGrid>), портал, который объединил высокопроизводительные вычислительные ресурсы, хранилища данных и экспериментальные установки. Проект длился с 2004 по 2011 г. Портал обеспечивал простой доступ к ресурсам и сервисам проекта TeraGrid. Для визуализации результатов использовалась система удалённого доступа к рабочему столу VNC, а также приложение для виртуализации ParaView. Портал TGViz построен на системах uPortal (<http://www.jasig.org/uportal/>) и GridSphere (<http://www.vlab.msi.umn.edu/gridsphere/gridsphere>). Для запуска задач на TeraGrid используется система GRAM (<http://dev.globus.org/wiki/GRAM>). Такой подход предусматривает возможность реализации удалённого просмотра каталогов, а также удалённого запуска приложений.

Портал TGViz обеспечивает простую и удобную возможность визуализировать большие наборы данных участниками распределённых коллективов. Использование интуитивно понятного веб-портала позволяет «спрятать» процесс взаимодействия на программно-аппаратном уровне за одной кнопкой, таким образом, пользователю не нужно обладать какими-то специальными знаниями в области программирования, чтобы запустить процесс визуализации.

В проекте SIDGrid (Social Informatics Data Grid) были разработаны портал по информатике и дата-центр для научных сотрудников в области социологии и исследования социального поведения. Система SIDGrid позволяет загружать в неё большие объёмы данных, а также предоставляет возможность масштабирования анализа данных до размеров вычислительных ресурсов TeraGrid. Портал SIDGrid демонстрирует эффективные результаты по обработке данных (например, поиск по тегам и образам), по спектру представляемых веб-сервисов (посредством SOAP) и по возможностям интеграции рабочих станций. Система SIDGrid написана с использованием PHP/Drupal (<http://drupal.org/>) и Python (<http://www.python.org/>) для интерфейса пользователя и использует стандарт Swift (Wilde, 2011) для выполнения вычислительных заданий.

Среда OLSGW — OpenSocial Life Science Gateway — интегрированная вычислительная среда, которая обеспечивает возможность научного сообщества, занимающегося наукой о жизни, использовать ресурсы TeraGrid как для вычислений, так и для управления данными. Основанная на сервис-ориентированной платформе, OLSGW объединяет приложения по биоинформатике, а также коллекции данных в веб-ориентированном портале, что позволяет исследователям создавать научные группы для проведения совместных вычислений, обмена данными и процессами. Портал разрабатывался с использованием Java и ODGE.

Портал UniHUB, разработанный в Институте системного программирования РАН на базе свободно распространяемого программного обеспечения HUBzero (<https://hubzero.org/about>), предлагает пользователям интерактивные инструменты, которые доступны из веб-браузера. В качестве системы

управления содержимым портала используется свободно распространяемый программный продукт Joomla (<http://www.joomla.org/>). Работа с программами происходит в «облачной» инфраструктуре без необходимости скачивания или компилирования какого-либо программного кода. UniHUB также предоставляет доступ к обучающим курсам, тематическим статьям, инструкциям по использованию программ и другой необходимой документации с помощью технологии Web 2.0.

Как показывает проведённый анализ, в большинстве случаев при разработке порталов для проведения совместных исследований используется свободно распространяемое программное обеспечение, которое, с одной стороны, позволяет экономить средства на лицензионных отчислениях, с другой — создавать сложные, многокомпонентные системы, в полном объёме отвечающие потребностям конечного пользователя — сотрудника, аспиранта, студента научного центра.

ЗАКЛЮЧЕНИЕ

В свете последних событий представляется актуальным проведение работ по разработке методики создания национального лицензионно независимого программного обеспечения для построения распределённой инфокоммуникационной системы для научной и научно-учебной деятельности. Для выполнения таких работ требуется провести дополнительные исследования по выбору адекватной программно-аппаратной платформы, разработать набор сценариев деятельности научной группы и научно-учебной группы, провести работы по выбору программных модулей для осуществления отдельных элементов сценария и разработать интерфейс пользователя.

ЛИТЕРАТУРА

- [Шикота, 2013] Шикота С. К. К задаче развития инфокоммуникационной инфраструктуры научного центра // Информационное об-во. 2013. № 1-2. С. 102–111.
- [Шур и др., 2010] Шур Л. Н., Крашаков С. А., Меньшутин А. Ю., Шикота С. К., Григорьева М. В. Средство общения виртуальных коллективов AccessGrid — реализация для научных семинаров и рабочих групп // Тр. 12-й Всерос. науч. конф. RCDL'2010 «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Казань, 13–17 окт. 2010. 2010. С. 519–523.
- [Шур, 2012] Шур Л. Н. Роль инфокоммуникационных технологий в развитии процесса глобализации научных исследований // Информационное об-во. 2012. № 5. С. 16–24.
- [Binns et al., 2007] Binns J., DiCarlo J., Insley J. A., Leggett T., Lueninghoener C., Navarro J.-P., Papka M. E. Enabling Community Access to TeraGrid Visualization Resources // Concurrency and Computation: Practice and Experience. Special Issue: Science Gateways — Common Community Interfaces to Grid Resources. 2007. V. 19. Iss. 6. P. 783–794.
- [Foster et al., 2001] Foster I., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations // Intern. J. High Performance Computing Applications. 2001. V. 15(3). P. 200–222.

- [Foster et al., 2002] *Foster I., Kesselman C., Nick J., Tuecke S.* The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration. 2002. URL: <http://www.globus.org/alliance/publications/papers/ogsa.pdf>.
- [Kenna, Berche, 2011] *Kenna R., Berche B.* Critical mass and the dependency of research quality on group size // *Scientometrics*. 2011. V. 86. P. 527–540.
- [Kepes, 2011] *Kepes B.* Understanding the cloud computing stack: SaaS, PaaS, IaaS. 2011. URL: http://www.rackspace.com/knowledge_center/whitepaper/understanding-the-cloud-computing-stack-saas-paas-iaas.
- [Levow et al., 2007] *Levow G.-A., Waxmonsky S.* et al. SIDGrid: A Framework for Distributed, Integrated multimodal Annotation, Archiving, and Analysis // 7th SIGDial Workshop on Discourse and Dialogue. 2007.
- [Shang Shari et al., 2011] *Shang Shari S. C., Li Eldon Y., Wu Ya-Ling* et al. Understanding Web 2.0 service models: A knowledge-creating perspective // *Information and Management*. 2011. V. 48(4-5). P. 178–184.
- [Wild et al., 2011] *Wilde M., Hategan M., Wozniak J. M., Clifford B., Katz D. S., Foster I.* Swift: A language for distributed parallel scripting // *Parallel Computing*. 2011.
- [Wu et al., 2008] *Wu W., Papka M. E., Stevens R.* Toward an OpenSocial Life Science Gateway // *Grid Computing Environments Workshop (GCE'08)*. 2008. P. 1–6.

TO THE PROBLEM OF DEVELOPMENT A TECHNIQUE FOR CREATION THE DISTRIBUTED INFOCOMMUNICATION SYSTEM FOR SCIENTIFIC AND SCIENTIFIC-EDUCATIONAL ACTIVITY

S. Shikota

Science Center in Chernogolovka, Russian Academy of Sciences

The goal is to develop a technique for creation distributed information and communication system to ensure scientific and scientific-educational activity of the creative team. The analysis of existing information and communication systems, portals and platforms in order to identify common components that are used for data management, metadata, computing resources, access rights, adding jobs and monitoring, and also for processing of simulation results.

Keywords: distributed information and communication system, free software, scientific infrastructure.

Shikota Svetlana Kalikstovna —programmer, sveta@chg.ru

АЛГОРИТМ ВАНГА-ЛАНДАУ: СЛУЧАЙНОЕ БЛУЖДЕНИЕ ПО СПЕКТРУ ЭНЕРГИИ

Л. Н. Щур

Институт теоретической физики им. Л. Д. Ландау Российской академии наук
(ИТФ РАН), Научный центр РАН в Черноголовке

Проведён анализ точности вычисления плотности состояний методом Ванга-Ландау. На примере вычисления плотности состояний двумерной модели Изинга показано, что стандартный метод ведёт к конечной точности вычислений. Проанализирована матрица переходов по спектру энергий и показано, что она является стохастической. Это наблюдение даёт возможность замены эвристического требования равномерности гистограммы на физическое требование регулирования точности заполнения столбцов матрицы переходов. Ставится вопрос о модификации стандартного метода Ванга-Ландау. Проанализированы пути эффективной параллелизации метода для возможности его реализации на массивно-параллельных суперкомпьютерных системах.

Работа выполнена при поддержке Российского научного фонда (РНФ) (проект № 14-21-00158).

Ключевые слова: метод Монте-Карло, плотность состояний спектра энергий, метод Ванга-Ландау, матрица переходов, точность вычислений, параллельные алгоритмы.

ВВЕДЕНИЕ

Для моделирования современных задач в области математической физики необходимо применение методов, которые обладают такими свойствами как хорошее масштабирование при параллельных вычислениях, возможность повышения точности вычислений, а также преодоления критического замедления и т. п. Например, в такой области как исследование свойств термодинамических систем, наиболее употребим и универсален метод Метрополиса [Metropolis et al., 1953]. Он обладает свойствами хорошей масштабируемости при параллельных вычислениях и возможностью повышения точности вычислений. Его применение, однако, становится затруднительным в критической области, области вблизи фазового перехода. В случае фазового перехода второго рода он проявляет свойства критического замедления, т. е. замедления скорости релаксации, что делает его применение затруднительным для систем большого размера. Кроме того, его применение не всегда ведёт к вопросу об ответе на вопрос о типе фазового перехода, т. е. дифференциации фазового перехода первого и второго рода.

Кластерные методы Свендсена-Ванга [Swendsen, Wang, 1987] и Вольфа [Wolff, 1988] позволяют преодолеть критическое замедление. Это связано с тем, что метод Метрополиса является локальным, в то время как размер корреляций растёт по мере приближения к критической точке. Кластерные методы работают именно на корреляционной длине, поэтому замедление, если и имеет место, практически незначительно. Недостатком кластерных методов является принципиальная сложность их параллелизации, поскольку

Щур Лев Николаевич — заведующий отделом, доктор физико-математических наук, профессор, shchur@chg.ru

в процессе их применения строится кластер, распространяющийся сложным образом по всему размеру моделируемой системы. Этот метод так же, как и метод Метрополиса, сложно применить к ответу на вопрос о типе фазового перехода. Кластерные методы основаны на представлении статистической суммы в виде разложения по кластерам [Hoshen, Kopelman, 1976].

Метод Ванга-Ландау [Wang, Landau, 2001] основан на представлении статистической суммы в виде разложения по плотности состояний энергии (см. следующий раздел). Метод свободен от критического замедления и, как будет обсуждено ниже, хорошо масштабируется при параллельных вычислениях. Метод позволяет ответить на вопрос о типе перехода.

В настоящей работе выявлено, что существующая реализация метода Ванга-Ландау имеет ограниченную точность. Известно также, что существующая реализация ведёт к неправильным результатам в случае непрерывного спектра и в случае смешанного взаимодействия с конкурирующими членами. В работе обсуждаются пути построения реализации метода, свободного от указанных недостатков.

МЕТОДЫ ВЫЧИСЛЕНИЯ СТАТИСТИЧЕСКОЙ СУММЫ

Для анализа свойств статистической системы необходимо вычислить её статистическую сумму Z . Метод Метрополиса основан на представлении статистической суммы в виде суммы по номеру состояния системы i

$$z = \sum_i \exp\left(\frac{-E_i}{kt}\right),$$

где E_i — энергия системы в состоянии с номером i ; t — температура; k — постоянная Больмана.

Методы в этой работе мы будем обсуждать в применении к модели Изинга [Ising, 1925]. Это простейшая модель ферромагнетика, в которой в узлах регулярной гиперкубической решётки помещаются переменные s_i , принимающие значения $+1$ и -1 , а на рёбрах — константы J_{ij} . Строится функция

$$E = -\frac{1}{2} \sum_{ij} J_{ij} s_i s_j - h \sum_i s_i,$$

где сумма производится по всем парам ближайших соседей i и j . Переменные s_i называются спинами, h играет роль внешнего магнитного поля и E является энергией системы. В случае нулевой температуры эта система классических спинов упорядочена, но значение энергии двукратно вырождено — спины могут принимать значение $+1$ (смотреть всё вверх) или -1 (смотреть всё вниз). В случае одномерной решётки, как показал Е. Изинг, фазовый переход отсутствует. Модель Изинга на квадратной решётке и в нулевом магнитном поле решил Л. Онзагер [Beale, 1996], который обнаружил фазовый переход второго рода (особенность во второй производной логарифма от статистической суммы). Его точное решение используется для проверки методов и алгоритмов, поскольку позволяет точно вычислить значения термодинамических величин на прямоугольной решётке любого конечного размера и при любой величине температуры. Мы будем проверять точность метода именно путём сравнения результатов моделирования с точным решением Онзагера.

Кластерные методы моделирования [Swendsen, Wang, 1987; Wolff, 1988] основаны на представлении статистической суммы q -компонентной модели Поттса в виде разложения по кластерам [Hoshen, Kopelman, 1976]:

$$Z = \sum_{\{G\}} v^{e\{G\}} \cdot q^{k\{G\}},$$

где G — один из кластеров; $e\{G\}$ — число элементов в этом кластере G ; $k\{G\}$ — число таких кластеров.

Метод Ванга-Ландау основан на представлении статистической суммы в виде разложения по числу уровней $g(E)$ с данной энергией E :

$$z = \sum_E g(E) \cdot \exp\left(\frac{-E}{kt}\right).$$

В англоязычной литературе $g(E)$ именуется Density of States, или DoS.

АЛГОРИТМ ВАНГА-ЛАНДАУ

В оригинальной работе Ф. Ванга и Д. П. Ландау [Wang, Landau, 2001] предложен следующий алгоритм, который для определённости будем применять для получения плотности состояний модели Изинга.

Процесс инициализации алгоритма Ванга-Ландау состоит из следующих шагов:

- I1 задаётся начальное состояние значений спинов s_i ;
- I2 вычисляется значение энергии начального состояния E_0 ;
- I3 присваиваются начальные значения плотности состояний $g(E) = 1$ для всех значений энергии E ;
- I4 устанавливаются начальные значения вспомогательной гистограммы $H(E) = 0$ для всех значений E ;
- I5 фиксируется начальное значение параметра $f = 2,718281828$.

Основной цикл алгоритма таков:

- C1 выбирается случайно спин s_i (для этого выбирается случайно узел i на решётке);
- C2 вычисляется энергия E_{k+1} нового состояния решётки с (виртуально) перевёрнутым спином $s_i \rightarrow -s_i$;
- C3 если $g(E_{k+1}) < g(E_k)$, то принимается новое состояние;
- C4 если нет, то принимается новое состояние с вероятностью $g(E_k)/g(E_{k+1})$;
- C5 если не произошло событий C4 и C5, то состояние остаётся неизменённым.

Основной цикл алгоритма повторяется MN раз, где N — это число спинов в системе; M — параметр алгоритма, некоторое большое число, имеющее значение, например, 10^4 . После этого проверяется свойство «ровности» диаграммы. Эмпирически предлагается повторять процесс до тех пор, пока элементы вспомогательного алгоритма будут отличаться не более чем на 5%. Если гистограмма недостаточно «ровная», то шаги основного цикла C1–C5 повторяются ещё MN раз.

Если желаемый уровень «ровности» диаграммы достигнут, то делают такие операции нормировки (сравнение с операциями инициализации I3–I5):

- S1 нормируются значения плотности состояний $g(E)$ для всех значений энергии E так, чтобы была выполнена нормировка на единицу значения плотности состояний энергии основного состояния моделируемой системы $g(E_{gs}) = 1$;
- S2 устанавливаются значения вспомогательной гистограммы $H(E) = 0$ для всех значений E ;
- S3 изменяется текущее значение параметра $f = f^{1/2}$.

Процессы основного цикла и нормировки повторяем до тех пор, пока значение параметра f не приблизится (близко) к 1, например $\log f = 10^{-9}$.

Принятие нового состояния на шагах основного цикла C3 и C4 состоит из таких операций:

- A1 переворот спина $s_i := -s_i$;
- A2 изменение значения DoS с текущим значением E_{k+1} энергии системы, $g(E_{k+1}) := fg(E_{k+1})$;
- A3 изменение значения вспомогательной гистограммы $H(E_{k+1}) := H(E_{k+1}) + 1$.

Непринятие нового состояния на шаге основного цикла C5 состоит из таких операций:

- N1 изменение значения DoS с текущим значением E_k энергии системы, $g(E_k) := fg(E_k)$;
- N2 изменение значения вспомогательной гистограммы $H(E_k) := H(E_k) + 1$.

ТОЧНОСТЬ АЛГОРИТМА ВАНГА-ЛАНДАУ

На рис. 1 приведено сравнение точного результата для плотности состояний энергии и вычисленного по алгоритму Ванга-Ландау для двумерной модели Изинга на квадратной решётке с линейным размером $L = 16$.

Отметим замечательное совпадение расчётных значений плотности состояний на масштабах рисунка. На самом деле, совпадение результатов не столь замечательно выглядит на следующем рисунке (рис. 2), где мы отложили по вертикали относительное отклонение вычисленных двумя способами значений плотности состояний, т. е. величины $1 - \log g(E_{Beale}) / \log g(E_{WL})$.

Если относительная точность вычисления в центре распределения достигает значений на порядок лучше процента, то на краях она велика, в несколько процентов. Более того, точность вычисления слабо зависит от выбора параметров. Можно менять степень «ровности» гистограммы, выбирать различную энергию для нормировки плотности состояний для осуществления нормировки в операции S1 или способ приближения параметра f к единице в операции S3 — всё это мало влияет на точность вычисления плотности состояний. Типичное поведение относительной точности приведено на рис. 3, где показано изменение точности вычислений по мере итераций основного процесса — точность выходит на насыщение.

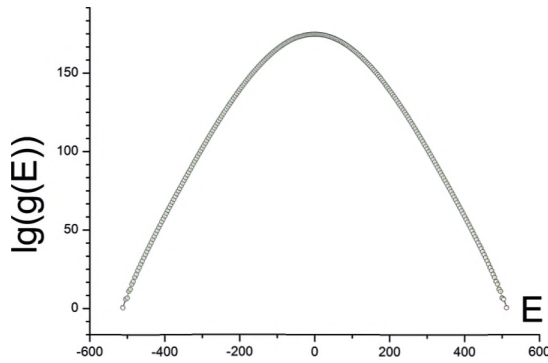


Рис. 1. Логарифм плотности состояний энергии двумерной модели Изинга на квадратной решётке с числом спинов 16×16 . Сплошная линия — точный аналитический результат, вычисленный по формулам из работы [Beale, 1996]. Кружки — расчёт по алгоритму Ванга-Ландау [Wang, Landau, 2001] с параметрами, указанными в описании алгоритма

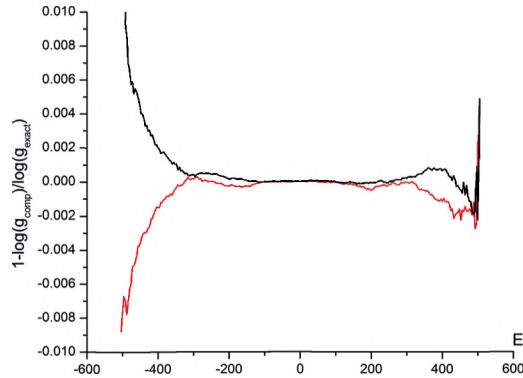


Рис. 2. Относительная точность вычисления плотности состояний для двух способов вычисления — по точным аналитическим выражениям Биля [Beale, 1996] и по алгоритму Ванга-Ландау [Wang, Landau, 2001]. Цветами выделены две реализации

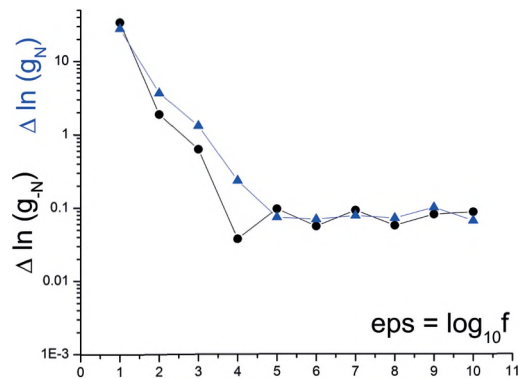


Рис. 3. Изменение усреднённой по состояниям относительной точности вычисления плотности состояний. Параметры моделирования — модель Изинга с числом спинов 8×8 , параметр $M = 10^7$. Цветами выделены две реализации

ПУТИ МОДИФИКАЦИИ АЛГОРИТМА ВАНГА-ЛАНДАУ

Наше предложение по модификации метода Ванга-Ландау основано на наблюдении в статистике переходов по спектру энергий. Ограничимся дискретным спектром значений энергии системы. Пронумеруем состояния дискретного спектра. Построим матрицу M переходов между состояниями с номерами k и k' . Оказывается, эта матрица является стохастической при приближении параметра f к 1. Это свойство можно использовать для модификации алгоритма Ванга-Ландау.

В описанном выше оригинальном алгоритме Ванга-Ландау заменим гистограмму H матрицей переходов M . Именно, сделаем следующие изменения в алгоритме (обозначим их добавочной буквой m в описании алгоритма):

- I4m установим начальные значения матрицы переходов $H(k, k') = 0$ между всеми номерами уровней энергии E ;
- S2m установим значения матрицы переходов $H(k, k') = 0$ между всеми номерами уровней;
- A3 увеличим значения элемента матрицы переходов $M(k, k+1) := M(k, k+1) + 1$;
- A3 увеличим значения элемента матрицы переходов $M(k, k) := M(k, k) + 1$.

Критерий «ровности» вспомогательной гистограммы заменяем на критерий выполнения условия стохастичности матрицы M , т. е. точности выполнения условия того, что сумма по столбцам матрицы M должна быть единицей. Таким образом, мы заменяем эмпирическое требование «плоскости» гистограммы на естественное требование стохастичности матрицы, что присуще самому процессу блуждания по спектру энергий.

НЕУСТОЙЧИВОСТЬ ФУНКЦИИ DOS В МЕТОДЕ ВАНГА-ЛАНДАУ

Для проверки устойчивости функции плотности состояний энергии был проведён следующий численный эксперимент. В качестве начального состояния (шаг I3 алгоритма Ванга-Ландау) мы выбирали точное аналитическое решение Биля [Beale, 1996]. Алгоритм применялся с различными начальными значениями параметра f (шаг I5), брались значения от $\log(f) = 1$ до $\log(f) = 10^{-9}$. Во всех случаях предельное значение ошибки стремилось приблизительно к тому, которое показано на рис. 3.

ЗАКЛЮЧЕНИЕ

В работе проанализирован алгоритм Ванга-Ландау вычисления плотности состояний спиновых систем. На примере двумерной модели Изинга без магнитного поля обнаружено, что алгоритм не позволяет вычислить плотность состояний в произвольной точностью. Напротив, обнаружено, что применение алгоритма ведёт к отходу от точного состояния. Обнаружено также, что алгоритм обеспечивает случайное блуждание по состояниям энергии таким образом, что матрица переходов между энергиями является стохастической.

Это свойство позволяет заменить эмпирическое условие «ровности» гистограммы на естественное условие точности выполнения условия стохастичности матрицы переходов, что позволяет оптимизировать время вычисления. Тем не менее, по-прежнему это не позволяет повысить точность вычисления. Необходимо проведение дополнительных исследований по поиску естественного параметра взамен параметра алгоритма f .

ЛИТЕРАТУРА

- [Beale, 1996] *Beale P. D.* Exact Distribution of Energies in the Two-Dimensional Ising Model // *Physical Review Letters*. 1996. V. 76. Iss. 1. P. 78.
- [Hoshen, Kopelman, 1976] *Hoshen J., Kopelman R.* Percolation and cluster distribution. I. Cluster multiple labeling technique and critical concentration algorithm // *Physical Review B*. 1976. V. 14. Iss. 8. P. 3438.
- [Ising, 1925] *Ising E.* Beitrag zur Theorie des Ferromagnetismus // *Zeitschrift für Physik*. 1925. V. 31. P. 253–258.
- [Metropolis et al., 1953] *Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E.* Equation of State Calculations by Fast Computing Machines // *J. Chemical Physics*. 1953. V. 21. N. 6. P. 1087–1092.
- [Swendsen, Wang, 1987] *Swendsen R. H., Wang J.-S.* Nonuniversal critical dynamics in Monte Carlo simulations // *Physical Review Letters*. 1987. V. 58. N. 2. P. 86.
- [Wang, Landau, 2001] *Wang F., Landau D. P.* Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States // *Physical Review Letters*. 2001. V. 86. P. 2050–2053.
- [Wolff, 1988] *Wolff U.* Collective Monte Carlo Updating for Spin Systems // *Physical Review Letters*. 1988. V. 62. P. 361.

ALGORITHM WANG-LANDAU: RANDOM WALK IN ENERGY SPECTRUM

L. N. Shchur

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

Analysis of the accuracy of density of state estimation using Wang-Landau method was done. It is demonstrated using example of two-dimensional Ising model, that method leads to the finite accuracy of estimations. Matrix of transitions in energy spectrum were constructed and appears to be the stochastic one. These observations bring possibility to substitute heuristic requirement of the histogram flatness with the physical requirement of tuning accuracy of columns in the transition matrix. We formulate question on the modification of Wang-Landau method. The possible approach to the effective parallelization were analyzed and provide possibility for realizations on the massively parallel supercomputers. Supported by Russian Scientific Foundation through the grant 14-21-00158.

Keywords: Monte Carlo method, density of states, Wang-Landau method, transition matrix, accuracy of computations, parallel algorithms.

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

**Computer Technologies in Sciences.
Methods of Simulations on Supercomputers
Proceedings**

Russia, Tarusa, October 1–3, 2014. 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 1–3 October 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 (Computing Technology in Science. Global scale systems / Eds. R. R. Nazirov, L. N. Shchur. M.: IKI RAN, 2011. 148 p.). Sixth 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 RAN, 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/>. На седьмом семинаре обсуждались алгоритмы и методы в области вычислительной физики, причём особый акцент делался на обсуждении полученных новых физических результатах. Материалы этого семинара можно найти на страницах семинара <http://www.iki.rssi.ru/seminar/>. Eighth workshop was transformed into the conference „Infocommunication technologies in scientific research“ with the papers published in the double volume issue of journal Information Society. 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. Seminar was supported through the grant by Russian Scientific Foundation 14-21-00158 „Algorithms and methods for mathematical simulations on supercomputing systems, including hybrid ones“.

Keywords: simulations, hybrid computing systems, simulations on supercomputers, methods and algorithms of parallel computing.

055(02)2

Ротапринт ИКИ РАН
117997, Москва, Профсоюзная, 84/32
Подписано к печати 20.11.2014

Заказ 3341

Формат 70×100/16

Тираж 100

13,65 усл.-печ. л.