

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра дискретной математики и информационных технологий

**ЛОКАЛИЗАЦИЯ РАСПРЕДЕЛЕНИЯ СОСТОЯНИЙ,
ВОЗБУЖДАЕМЫХ ИМПУЛЬСОМ ВНЕШНЕГО ПОЛЯ В
СИСТЕМЕ БЕЗМАССОВЫХ ФЕРМИОНОВ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

Студентки 2 курса 271 группы
направления 09.04.01 — Информатика и вычислительная техника
факультета КНиИТ
Ульяновой Анастасии Алексеевны

Научный руководитель
доцент, к. ф.-м. н.

А. Д. Панферов

Заведующий кафедрой
доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2023

ВВЕДЕНИЕ

В настоящее время актуальны поиски новых материалов, которые позволили бы продолжить прогресс в быстродействии вычислительных и коммуникационных устройств. Среди перспективных материалов выделяются двумерные структуры различного химического состава. В качестве самого известного примера можно отметить графен, получение образцов которого позволило разработать и экспериментально подтвердить ставшую уже классической модель безмассовых фермионов [1, 2]. Использование этой модели позволяет прогнозировать поведение этого материала в различных условиях. Особенно актуально уметь воспроизводить его отклик на действие внешних электрических полей [3–8]. Для решения таких задач был разработан подход на основе квантового кинетического уравнения [9–11], основные элементы которого представлены в первой главе ВКР. Его использование позволяет получать точные результаты для широкого диапазона параметров внешнего поля, однако требует большого объёма вычислений. Необходимые вычисления включают численное решение дифференциальных уравнений, определяющих поведение модели, и обработку полученных массивов данных на этапе получения физически наблюдаемых характеристик.

Целью представляемой работы является уменьшение необходимых для такого моделирования ресурсов на этапе построения сетки воспроизводимых значений заселённости возбуждённых состояний системы. Критерием выступает минимальное количество используемых для покрытия заданной области двумерного пространства узлов при точном воспроизведении интегральных характеристик моделируемого распределения. Для достижения поставленной цели разработана итерационная процедура поэтапного построения сетки, на каждом этапе которой выполняется обработка и анализ результатов моделирования предыдущих этапов и дополнение сетки новыми узлами. Представленные результаты обладают научной новизной, а целевая область их применения определяет научную значимость проделанной работы.

Необходимость работы с массивами промежуточных данных потребовало изучения современных подходов к работе с большими данными (Big Data) и их использования.

В ходе выполнения работы были решены следующие задачи:

1. Изучены особенности работы с большими массивами данных: типы дан-

ных, способы их хранения и обработки, инструменты для обработки и анализа больших объёмов данных по материалам курса «Big data» Cisco Networking Academy.

2. Изучены типовые инструменты для работы с большими данными на примере языка программирования Python и библиотек для него:
 - библиотека обработки данных, представленных в форме многомерных массивов NumPy;
 - библиотека анализа данных Pandas;
 - библиотека визуализации данных Matplotlib.
3. Реализованы стандартные процедуры обработки массивов данных для результатов численного моделирования физических процессов в системе безмассовых фермионов в условиях действия на неё импульсов внешнего поля.
4. Изучены и реализованы в качестве метода локализации алгоритмы AAVB и отслеживания изолиний скалярной функции.
5. Изучен и реализован метод локализации, основанный на выборочном поэтапном уплотнении сетки.
6. Выполнена программная реализация адаптации вычислительной сетки при произвольном числе областей локализации функции распределения.

Магистерская работа состоит из введения, 4 глав: «Модель безмассовых фермионов», «Большие данные и работа с ними средствами Python», «Локализация возбуждённых состояний модели на языке программирования Python», «Практическая реализация с использованием инструментов Wolfram Mathematica», заключения и списка использованных источников. Общий объём работы — 75 страниц, включая 45 рисунков и 2 таблицы, список использованных источников содержит 38 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В **первой главе** «Модель безмассовых фермионов» представлена используемая модель рассматриваемого физического процесса и сформулированы основные проблемы, требующие нахождения решения в рамках данной работы.

В работах [9–11] при описании процессов образования и эволюции носителей заряда в двумерном монослое графена в условиях действия внешнего электрического поля был реализован метод, основанный на использовании квантового кинетического уравнения. Данный метод достаточно универсален и может использоваться при рассмотрении других аналогичных материалов. При таком подходе подсистема слабосвязанных электронов материала рассматривается как квантовая система с двумя энергетическими уровнями. Конкретный вид закона дисперсии (зависимости энергии состояний от импульса) может быть различным. В простейшем случае он соответствует приближению безмассовых фермионов, но может использоваться и строгая модель взаимодействия ближайших соседей. Кинетическое уравнение определяет эволюцию функции распределения $f(p_1, p_2, t)$ носителей по возбуждённым состояниям в условиях внешнего воздействия и может быть представлено в виде системы обыкновенных дифференциальных уравнений 1, определяемой через гамильтониан модели.

$$\begin{cases} \dot{f}(\bar{p}, t) = \frac{1}{2}\lambda(\bar{p}, t)u(\bar{p}, t), \\ \dot{u}(\bar{p}, t) = \lambda(\bar{p}, t)(1 - 2f(\bar{p}, t)) - \frac{2\epsilon(\bar{p}, t)}{\hbar}v(\bar{p}, t), \\ \dot{v}(\bar{p}, t) = \frac{2\epsilon(\bar{p}, t)}{\hbar}u(\bar{p}, t), \end{cases} \quad (1)$$

где $u(\bar{p}, t)$ и $v(\bar{p}, t)$ вспомогательные функции (детали можно посмотреть в работе [11]). Поведение моделируемой системы фактически определяется коэффициентами $\epsilon(\bar{p}, t)$ и $\lambda(\bar{p}, t)$.

Основной сложностью при использовании данного метода является то, что система уравнений, лежащая в его основе, может решаться только численными методами и для каждой рассматриваемой точки импульсного пространства $\{p_1, p_2\}$ решение необходимо искать независимо [12]. Это вызывает высокую вычислительную сложность процедуры получения наблюдаемых характеристик. Ещё одной проблемой не имеющей эффективного решения было предварительное определение местоположения области локализации макси-

мальных значений. При моделировании процессов в рассматриваемых системах было обнаружено, что локализацию областей концентрации возбуждённых состояний (где требуется использовать плотные сетки) удобнее выполнять по данным о $f(p_1, p_2, t)$ не после завершения внешнего воздействия, а во время его [13].

В работе [14] выполнены исследования и сопоставление поведения в импульсном пространстве функции распределения в конечном состоянии $f(p_1, p_2, t \rightarrow \infty)$, в момент максимума внешнего воздействия $f(p_1, p_2, t_{E_{max}})$, а также максимальных значений функции распределения в выбранной точке импульсного пространства за всё время действия внешнего поля $f_{max}(p_1, p_2)$ и средних значений функции распределения за время действия внешнего поля $f_{aver}(p_1, p_2)$. Это позволило разработать итерационный метод построения разномасштабных регулярных сеток, обеспечивающий в рамках унифицированной процедуры локализацию области генерации возбуждённых состояний с охватом всех рассматриваемых значений (p_1, p_2) (первой зоны Бриллюэна) и формирование в ней плотных сеток разного масштаба, представляемый в настоящей работе.

Вторая глава «Большие данные и работа с ними средствами Python» посвящена проблематике работы с большими данными, языку программирования Python, библиотекам анализа и обработки данных. Некоторые типовые процедуры обработки данных демонстрируются применительно к исследуемой предметной области.

Основными характеристиками больших данных являются 4 V: объём (volume), скорость (velocity), разнообразие (variety) и достоверность (veracity). Несмотря на то, что до недавнего времени только первые три V являлись главными характеристиками больших данных, в современных условиях возникла необходимость расширения до семи V: ценность (value), изменчивость (variability), визуализация (visualization) [15].

В data science и области больших данных встречается много разных типов данных, для каждого из которых требуются свои инструменты и методы. Основные категории данных: структурированные, неструктурированные, на естественном языке, машинные, графовые, аудио, видео и графика, а также потоковые.

Существует множество программ, которые используются для форма-

тирования данных, их очистки, анализа и визуализации. Многие компании и организации обращаются к инструментам с открытым исходным кодом для обработки, агрегирования и обобщения своих данных. Язык программирования Python стал широко используемым инструментом для обработки и манипулирования данными.

Python — это один из самых распространённых языков программирования. Возможности языка Python могут быть расширены за счёт использования большого количества сторонних библиотек, что позволяет решать различные задачи [16]. Пакеты позволяют форматировать, структурировать, анализировать и визуализировать данные. Использование таких библиотек упрощает и ускоряет процесс создания инструментов для обработки данных.

С использованием реальных данных, представляющих результат физического процесса на основе модели безмассовых фермионов и представленных библиотек из экосистемы языка программирования Python, выполнена обработка массивов с целью сортировки, выборки данных в заданных диапазонах и поиска экстремумов. Проиллюстрированы различные форматы графической визуализации такой обработки.

В третьей главе «Локализация возбуждённых состояний модели на языке программирования Python» рассмотрены варианты решения поставленной задачи с использованием алгоритмов AABB, отслеживания изолиний и масштабирования сеток.

Для обработки результатов моделирования, необходимо осуществлять выборку данных и построение графиков, данная задача решается при помощи алгоритма AABB и построения изолиний. Axis-aligned bounding box (AABB — «параллельный осям ограничивающий параллелепипед») — это параллелепипед со сторонами, параллельными осям координат, ограничивающий некоторый геометрический объект в пространстве [17]. Чаще за размеры параллелепипеда берётся модуль максимальной разности проекций на выбранную ось между двумя точками. Построение изолиний можно осуществлять следующими методами: метод статистического окна и метод шагающих квадратов. Рассмотрим эти методы подробнее.

Метод статистического окна применяется при большом числе значений функций. Заключается в построении изолиний поверхности по средним групповым отметкам. На координатную плоскость с точками наносят квадратную

сетку со сторонами a и b . В каждой ячейке сетки определяют среднее арифметическое значение отметок всех попавших в неё точек и подписывают это среднее в центре статистического окна. Окно перемещают на половину своего размера сначала по одной оси, а затем по другой, или на полный его размер [18]. Если точки группируются в одной какой-либо части ячейки, то среднее подписывают в их геометрическом центре. Пограничные точки учитывают в обоих соседних окнах. Затем, выбираем значение изолинии и по отметкам средних строим её. Для уменьшения вычислительных операций при построении изолинии, заданного уровня, выбираем значение функции методом ААВВ. В результате получаем сглаженный контур.

Метод шагающих квадратов является интерпретацией метода шагающих кубов для плоскости [19]. На вход алгоритм получает регулярную сетку, в каждом узле которой известно значение анализируемой функции. Выходная сетка может иметь меньшее разрешение (в этом случае теряется точность, но уменьшается ступенчатость). Далее для каждого узла выходной сетки проверяется, выше ли значение в нём, чем на изоповерхности. Всем узлам, которые выше, приписывается «+», остальным «-». Далее рассматриваются квадраты выходной сетки, вершины которых лежат в отмеченных узлах. Всего получается 16 различных случаев, которые с учётом симметрий и поворотов можно свести к четырём:

- Случай 1: все вершины имеют один знак.
- Случай 2: у одной вершины знак отличается.
- Случай 3: вершины с одинаковыми знаками имеют общее ребро.
- Случай 4: вершины с одинаковыми знаками не имеют общего ребра.

В четвёртом случае невозможно однозначно определить форму сегмента изолинии, поэтому дополнительно просматривается значение в центре квадрата, если входные данные это позволяют. При невозможности узнать значение в центре квадрата принятое решение может повлиять на связность изолинии.

Метод локализации с использованием масштабирования сеток был разработан на основе изучения поведения функции распределения, рассматриваемой модели [13–14]. Параметры для генерации базовой регулярной сетки и последующей её адаптации, физические параметры задачи, а также промежуточные и итоговые результаты работы процедуры задаются в текстовых

файлах. Описание итераций метода и графические результаты его апробации представлены в работе.

В четвертой главе «Практическая реализация с использованием инструментов Wolfram Mathematica» приведено программное решение для локализации областей максимальных значений функции распределения и результаты тестирования системы. В главе представлен пошаговый алгоритм процедуры локализации и его программная реализация средствами системы компьютерной алгебры Wolfram Mathematica.

Процедура локализация области максимальных значений функции распределения является ключевой для построения мультимасштабной сетки. Опишем её алгоритм:

1. Проверяем не исчерпан ли лимит количества шагов масштабирования `grid_options.txt`.
2. Перед началом процедуры оптимизации проверяем не осталось ли не посчитанных узлов сетки.
3. Определяется максимальное значение функции распределения из всех посчитанных (ММЗФР).
4. Проверяется, не превышает ли ММЗФР назначенный верхний порог. Если не превышает переходим к шагу 3, если превышает, то к шагу 5.
5. Проверяется, превышает ли ММЗФР значение нижний порог рассматриваемых значений МЗФР/диапазон МЗФР, используемый на этапе локализации области максимальных значений до выхода за верхний порог. Данное условие определяет достаточен ли диапазон значений МЗФР для корректного анализа на первом шаге масштабирования сетки и имеет смысл только при анализе первого поколения. Если не превышает, то работа с сеткой прерывается. Необходимо увеличить k_1 . Выводится сообщение, что начальное разбиение недостаточно представительно. Если превышает, то переходим к шагу 4.
6. Выбираются все узлы со значением МЗФР в диапазоне от ММЗФР* (диапазон МЗФР, используемый на этапе локализации области максимальных значений до выхода за верхний порог) до ММЗФР. Проверяется, не превышает ли их количество ограничивающее значение (максимальное количество масштабируемых узлов, при котором используется коэффициент масштабирования k_2). Если ограничение превышено,

процедура прерывается, выводится соответствующее сообщение. Если ограничение не нарушено, выбранные узлы помечаются к разбиению.

7. Если ММЗФР превысил верхний порог, подсчитывается количество узлов сетки с МЗФР более чем верхнее пороговое значение МЗФР. Если превышено ограничение (максимальное количество масштабируемых узлов, при котором используется коэффициент масштабирования k_2), процедура прерывается, выводится соответствующее сообщение. Если это количество менее чем максимальное количество масштабируемых узлов, при котором используется коэффициент масштабирования k_2 , но более чем максимальное количество масштабируемых узлов, при котором используется коэффициент масштабирования k_1 , они помечаются к разбиению. Если количество узлов с МЗФР более верхнего порога менее чем максимальное количество масштабируемых узлов, при котором используется коэффициент масштабирования k_1 , выполняется процедура, описанная на шаге 4.

Алгоритм был реализован в виде системы блокнотов Wolfram Mathematica.

Далее было проведено тестирование системы с различными параметрами процедуры локализации и физическими параметрами модели для нахождения оптимальных параметров локализации при адаптации сетки. Для нахождения оптимальных параметров локализации были проведены эксперименты и протестирована система с различными начальными значениями коэффициента масштабирования k_1 , который изменяли от 2 до 6 и значениями коэффициента масштабирования k_2 на этапе оптимизации сетки, который был равен 1 или 2. При сопоставимой плотности и приемлемом для анализа числе узлов расчётной сетки при начальном коэффициенте масштабирования k_1 равном 3 и коэффициенте масштабирования k_2 равном 1 на этапе оптимизации, получаем более быстрое время выполнения программы при меньшем количестве шагов. Таким образом, проанализировав полученные результаты, оптимальным значением для начального коэффициента масштабирования k_1 является 3, а для коэффициента масштабирования k_2 на этапе оптимизации сетки — 1. Результаты тестирования представлены в форме таблиц, качество построенных сеток иллюстрируется в графической форме.

ЗАКЛЮЧЕНИЕ

В работе представлены решения для адаптивной поэтапной локализации областей с максимальными значениями заселённости возбуждённых состояний в графене в условиях действия на него внешнего электрического поля. Такая локализация позволяет строить расчётную сетку переменной плотности, размещая большую часть её узлов в областях с высокой концентрацией возбуждённых состояний, дающих определяющий вклад в наблюдаемые параметры моделируемой системы. Это обеспечило достижение цели, заключавшейся в минимизации необходимых для моделирования ресурсов.

Задачи, поставленные на этапе планирования содержания и хода выполнения работы, были полностью решены.

Промежуточные результаты представлялись в докладе на национальном суперкомпьютерном форуме 2022 года [14].

Основные источники информации:

- 1 Two-dimensional gas of massless Dirac fermions in graphene / K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, and A.A. Firsov // *Nature*. – 2005. – Vol. 438, № 7065.– Pp. 197–200.
- 2 The electronic properties of graphene / A.H.C. Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, and A.K. Geim // *Reviews of Modern Physics*. – 2009. – Vol. 81, № 1. – Pp. 109–162.
- 3 High frequency electric field induced nonlinear effects in graphene M.M. Glazov, S.D. Ganichev // *Physics Reports*.– 2014.– Vol. 535.– Pp.101-138.
- 4 Ultrafast terahertz response of multilayer graphene in the nonperturbative regime P. Bowlan, E. Martinez-Moreno, K. Reimann, T. Elsaesser, and M. Woerner // *Physical Review B*. – 2014.– Vol. 89.– 041408.
- 5 Ultrafast nonlinear optical response of Dirac fermions in graphene M. Baudisch, A. Marini, J.D. Cox, T. Zhu, F. Silva, S. Teichmann, M. Massicotte, F. Koppens, L.S. Levitov, F.J. García de Abajo, and J. Biegert // *Nature Communications*.– 2018.– Vol. 9.– 1018.
- 6 Circularly polarized extreme ultraviolet high harmonic generation in graphene Zi-Yu Chen, Rui Gin // *Optics Express*.– 2019.– Vol. 27.– P. 3761.
- 7 Terahertz-induced high-order harmonic generation and nonlinear charge

- transport in graphene W. Mao, A. Rubio, and S.A. Sato // Physical Review B. – 2022.– Vol. 106.– 024313.
- 8 Gate-tunable quantum pathways of high harmonic generation in graphene Soonyoung Cha, Minjeong Kim, Youngjae Kim, Shinyoung Choi, Sejong Kang, Hoon Kim, Sangho Yoon, Gunho Moon, Taeho Kim, Ye Won Lee, Gil Young Cho, Moon Jeong Park, Cheol-Joo Kim, B. J. Kim, JaeDong Lee, Moon-Ho Jo, and Jonghwan Kim // Nature Communications.– 2022.– Vol. 13.– 6630.
 - 9 Residual currents generated from vacuum by an electric field pulse in 2+1 dimensional QED models / S.A. Smolyansky, D.V. Churochkin, V.V. Dmitriev, A.D. Panferov, and B. Kämpfer // EPJ Web of Conferences. – 2017. – Vol. 138. – Pp.6004.
 - 10 Comparing two different descriptions of the I-V characteristic of graphene: theory and experiment / A.D. Panferov, S.A. Smolyansky, D.B. Blaschke, N.T. Gevorgyan // EPJ Web of Conferences. – 2019. – Vol. 204. – Pp.6008.
 - 11 Nonperturbative kinetic description of electron-hole excitations in graphene in a time dependent electric field of arbitrary polarization S.A. Smolyansky, A.D. Panferov, D.B. Blaschke, and N.T. Gevorgyan // Particles.– 2019.– Vol. 2.– P. 208-230.
 - 12 Использование адаптивной сетки на основе квадродерева для моделирования конечного состояния квантово-полевой системы при импульсном внешнем воздействии А.Д. Панферов, А.В. Маханьков, А.А. Трунов // Программные системы: теория и приложения.– 2020.– Vol. 11.– С. 79-92.
 - 13 Моделирование поведения графена во внешних электрических полях А.Д. Панферов, Н.А. Новиков, А.А. Трунов // Программные системы: теория и приложения.– 2021.– Vol. 12.– С. 3-19.
 - 14 Моделирование поведения двухуровневой квантовой системы с использованием масштабируемых регулярных сеток А.Д. Панферов, Н.В. Поснова, А.А. Ульянова // Национальный суперкомпьютерный форум-2022 Россия, Переславль-Залесский, 29 ноября – 2 декабря 2022 г.
 - 15 Материалы курса Cisco Networking Academy «Big data» [Электронный ресурс] URL: <https://lms.netacad.com/course/view.php?id=562884>
 - 16 Python.org [Электронный ресурс] Python 3.9.9 documentation URL:

<https://www.python.org/> (дата обращения — 15.11.2021) Загл. с экрана.
Яз. англ.

17 Википедия — свободная энциклопедия [Электронный ресурс] ААВВ URL:
<https://ru.wikipedia.org/wiki/ААВВ> (дата обращения — 10.02.2022) За-
гл. с экрана. Яз. рус.

18 ГОСы ГМ 3 [Электронный ресурс] Построение изолиний топографиче-
ской поверхности URL: <http://gosgm3.blogspot.com/p/33.html> (дата об-
ращения — 25.02.2022) Загл. с экрана. Яз. рус.

19 Marching cubes: a high resolution 3d surface construction algorithm / W.E.
Lorenson, H.E. Cline // Computer Graphics. — 1987. Volume 21, P. 163-169.