

Параллельная реализация алгоритма прямого метода Монте-Карло для моделирования стационарного течения одноатомного газа

Л.В. Павленко, С.А. Маякова

В данной работе рассматривается схема параллелизации алгоритма прямого метода Монте-Карло для моделирования столкновений молекул одноатомного газа. Протестированы два варианта параллельного алгоритма: алгоритм распараллеливания по независимым статистическим испытаниям и алгоритм распараллеливания по данным. Исследованы ускорение и эффективность этих алгоритмов на примере моделирования стационарного течения. Проведено сравнение эффективностей каждого из подходов к распараллеливанию.

1. Введение

Большое научное и прикладное значение, которое в настоящее время имеет динамика разреженных газов, объясняется практической важностью решения широкого круга задач, связанных с современным этапом освоения космоса, развитием вакуумной технологии, лазерной техники и других отраслей научно-технического прогресса. Методы, развитые в динамике разреженных газов, широко применяются для решения задач, не связанных с разреженностью среды – теории гомогенных и гетерогенных процессов, теории испарений и адсорбционных процессов, неравновесных течений, обоснования и установления граничных условий и коэффициентов переноса в механике сплошных сред. Необходимость качественного и количественного анализа явлений динамики разреженных газов, сложность и многомерность уравнений, с которыми приходится иметь дело, стимулировали развитие эффективных и оригинальных численных методов.

В настоящее время многие задачи динамики разреженных газов решаются численными статистическими методами Монте-Карло [1-2]. Эти методы выбираются потому, что сложная нелинейная структура интеграла столкновения и большое количество переменных создают существенные трудности для анализа, в том числе и численного и практически исключают конечно-разностный подход для серьезных задач. И, в тоже время, многомерность и вероятностная природа кинетических процессов создают естественные предпосылки для применения методов Монте-Карло. Метод прямого моделирования Монте-Карло (ПММК) состоит в имитации реального течения газа с помощью моделирующих частиц. При этом сравнительно небольшое число молекул моделирует поведение гораздо большего числа реальных молекул в течениях разреженного газа.

2. Математическая модель

Динамика разреженных газов описывается известным интегродифференциальным кинетическим уравнением Больцмана [3]

$$\frac{\partial f}{\partial t} + \vec{\xi} \nabla f = \int (f' \cdot f'_1 - f \cdot f_1) \cdot \vec{g} \cdot b \cdot db \cdot d\varepsilon \cdot d\vec{\xi}_1;$$

$f = f(t, x, y, z, \xi_x, \xi_y, \xi_z)$ – функция распределения молекул по времени, координатам и скоростям;

f', f'_1 – функции распределения, соответствующие скоростям молекул после столкновения $\vec{\xi}', \vec{\xi}'_1$;

\vec{g} – относительные скорости молекул при парных столкновениях $\vec{g} = \vec{\xi} - \vec{\xi}_1 = \vec{\xi}' - \vec{\xi}'_1$;

b, ε – прицельное расстояние и азимутальный угол при столкновениях частиц.

3. Метод решения

К настоящему времени безусловный приоритет в динамике разреженного газа принадлежит методу Бёрда [4], модификации которого трудами отечественных исследователей О. Белоцерковского, В. Яницкого, М. Иванова, В. Перепухова, А. Ерофеева позволили буквально на порядки повысить эффективность метода. Суть метода заключается в том, что эволюция системы на малом промежутке времени Δt рашепляется на два ясных физических процесса:

- свободно-молекулярный перенос $\frac{\partial f}{\partial t} = -\bar{\xi} \nabla f$;
- релаксацию в соответствии с оператором столкновений $J(f)$ в кинетическом уравнении $\frac{\partial f}{\partial t} = J(f)$.

Это хорошо известная схема расщепления первого порядка по Δt для любого операторного уравнения, но в данном случае она подкупает тем, что рашепляет динамику такой сложной кинетической системы на два ясных физических процесса. Функция распределения моделируется N частицами, которые на первом этапе в течении времени Δt перелетают на расстояния $\bar{\xi}_j \Delta t$, а на втором в каждой ячейке между собой сталкиваются в соответствии с частотой столкновения на протяжении Δt .

Моделирование столкновения сводится к статистической реализации эволюции в течении времени Δt не уравнения Больцмана, а модели Каца [5]

$$\frac{\partial \varphi_1(t, \bar{\xi}_1)}{\partial t} = \frac{N-1}{N} \int (\varphi_2[t, \bar{\xi}', \bar{\xi}'_2] - \varphi_2(t, \bar{\xi}_1, \bar{\xi}'_2]) \cdot g_{12} \cdot d\sigma_{12} \cdot d\bar{\xi}_2$$

где φ_1, φ_2 - одно и двух частичные функции распределения.

Время столкновения в модели Каца рассчитывается в соответствии со статистикой столкновения в идеальном газе по схеме Бернулли [6]. Эта схема позволяет использовать существенно меньшее число частиц в ячейке и более мелкий шаг расчетной сетки.

Для получения информации о поле течения моделируемый объем физического пространства разбивается на ячейки, которые приближенно определяют пространственные положения молекул в определенные моменты времени. Результатом моделирования течения газа методом ПМК являются макроскопические параметры газа, определяемые параметрами частиц в ячейках. При численном моделировании значений этих параметров неизбежно возникает статистическая погрешность. Для уменьшения статистической погрешности определения параметров потока увеличивают объем выборки (проводят N_{MC} статистически независимых испытаний).

В работе рассматриваются следующие параметры ($\langle \cdot \rangle$ - усреднение по независимым статистическим испытаниям):

- $n_0 = E(n) \approx \langle n \rangle$ - среднее число молекул в ячейке, усредненное по независимым статистическим испытаниям;
- $c_0 = E(c) \approx \langle c \rangle$ - средняя скорость, приходящаяся на одну молекулу в ячейке, усредненная по независимым статистическим испытаниям;
- $T_0 \approx \frac{m}{3k} (\langle c^2 \rangle - \langle c \rangle^2)$ - средняя температура в ячейке;

- $T = T_0^{срeднее}$ - средняя температура всего моделируемого объема;
- $\sigma = \sqrt{D(T)} \approx \sqrt{(\langle T^2 \rangle - \langle T \rangle^2)}$ - среднеквадратичное отклонение температуры, необходимое для оценки статистической погрешности $\varepsilon = \frac{\sigma \cdot 100\%}{\sqrt{N_{MC}}}$.

4. Последовательный алгоритм ПММК и результаты моделирования

Алгоритм ПММК [4] состоит из двух основных циклов. Во вложенном цикле проводится моделирование одной реализации процесса. На каждом временном шаге проводится последовательное выполнение процедур моделирования. Внешний цикл повторяет реализации n раз для получения требуемого объема выборки. Каждая новая реализация повторяется независимо от предыдущих. Независимость результатов в различных реализациях обеспечивается за счет использования разных последовательностей псевдослучайных чисел.

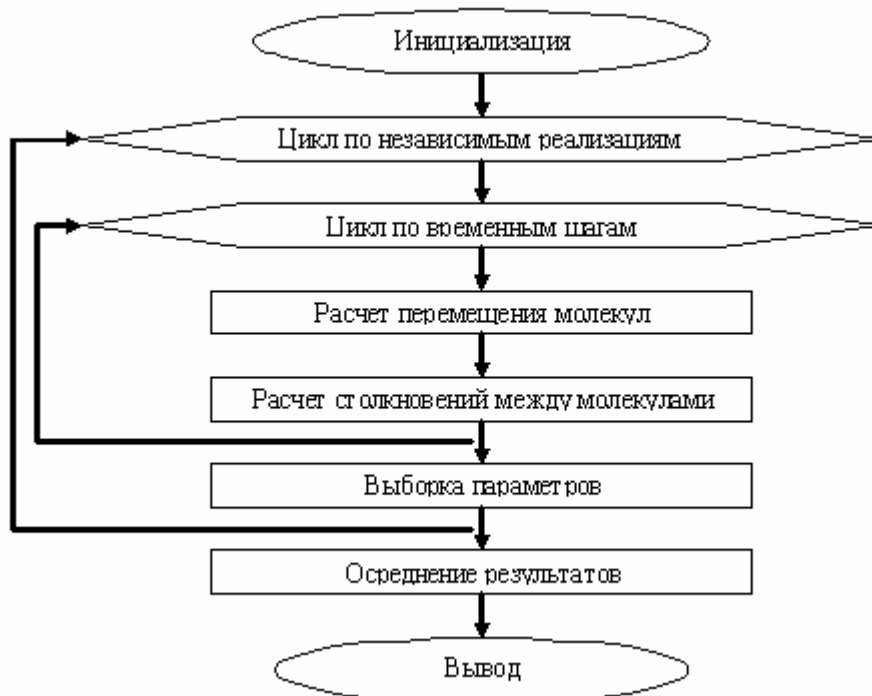


Рис. 1. Укрупненная блок-схема последовательного алгоритма ПММК

При программной реализации описанной математической модели не использовалось никаких специальных математических пакетов. Была разработана программа на языке Си, реализующая алгоритм, приведенный на рис. 1, и проведено моделирование стационарного течения одноатомного газа с использованием ПММК. При этом рассматривалось поведение одноатомного газа с плотностью 10^{21} молекул на единицу объема в течение 0.0025 с. Для моделирования использовалось 10000 моделирующих частиц, пространство разбивалось на 100 ячеек, было проведено 1000 независимых статистических испытаний. Получены следующие результаты.

Таблица 1. Результаты моделирования.

Число столкновений	3056227
Средняя температура	301.039072
Погрешность вычисления температуры	0.039101%
Соотношение экспериментального и теоретического числа столкновений	1.000612

5. Оптимизация алгоритма ПММК

Для достижения большой точности моделирования приходится использовать не менее 10000 моделирующих частиц в области и проводить по крайней мере 1000 независимых статистических испытаний. Это приводит к тому, что последовательный алгоритм ПММК работает достаточно долго. Поэтому для сокращения времени работы программы применяются алгоритмы распараллеливания ПММК.

5.1. Алгоритм распараллеливания по независимым статистическим испытаниям

Независимость отдельных реализаций друг от друга позволяет выполнить их параллельно на разных процессорах [1]. Применение такого подхода на p -процессорной системе приводит к сокращению в p раз числа итераций внешнего цикла метода ПММК. Обмен данными происходит только после завершения всех расчетов на процессорах. Обработка результатов ведется последовательно на одном процессоре после завершения обмена данными. Область эффективного применения такого алгоритма $p \leq N_{MC}$. Для достижения оптимальных значений ускорения и эффективности p должно быть кратно p . При $p > N_{MC}$ возможности ускорения, предоставляемые вычислительной системой, являются избыточными.

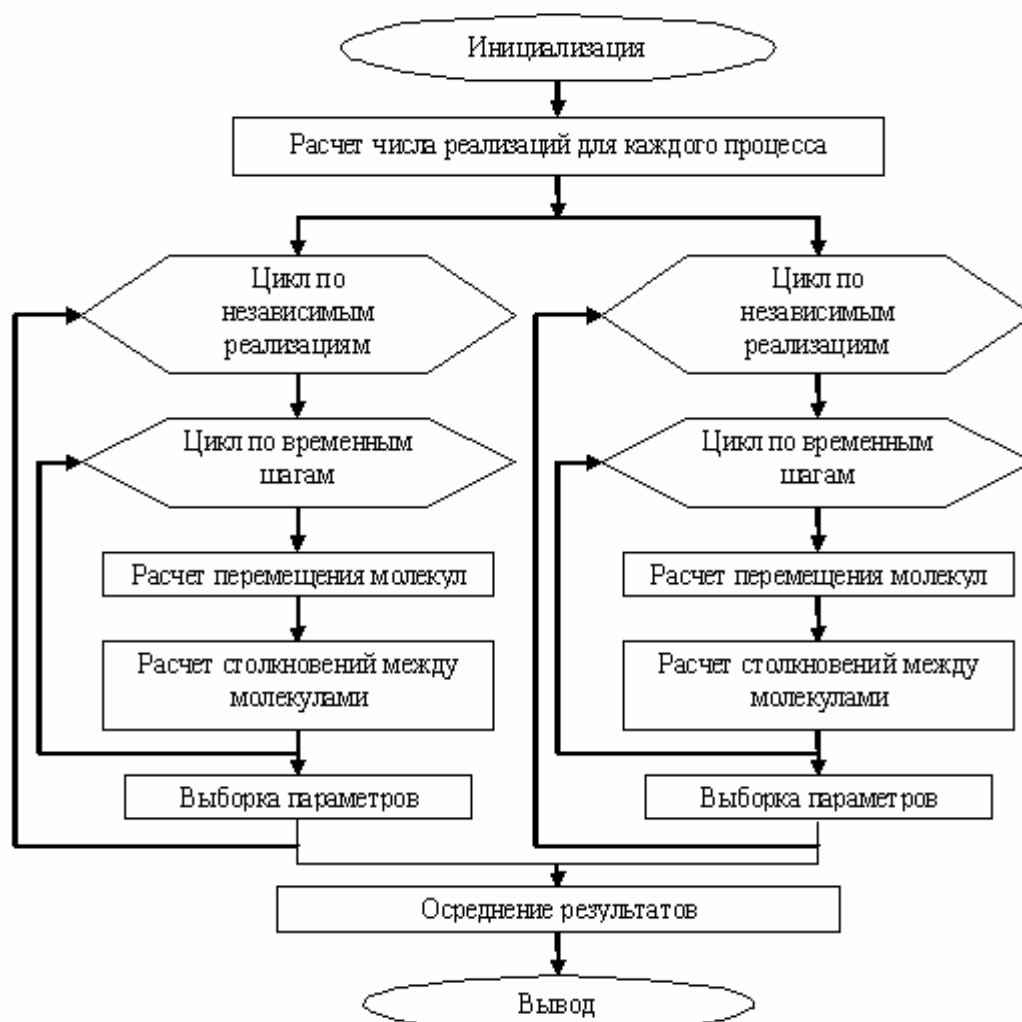


Рис. 2. Укрупненная блок-схема алгоритма распараллеливания по независимым статистическим испытаниям

При моделировании рассмотренной выше задачи с использованием алгоритма распараллеливания по независимым статистическим испытаниям были получены следующие результаты.

Таблица 2. Время работы алгоритма распараллеливания по независимым статистическим испытаниям на различном числе ядер (для вычислительной платформы со следующими характеристиками: два двухъядерных 2 x Intel Xeon (2.00 GHz), 4 Mb Cache, 4Gb RAM).

Время работы последовательного алгоритма	4018 с
Время работы параллельного алгоритма на 2 ядрах	2038 с
Время работы параллельного алгоритма на 3 ядрах	1373 с
Время работы параллельного алгоритма на 4 ядрах	1014 с

Таким образом, получено ускорение близкое к линейному и эффективность приблизительно равна 1.

5.2. Алгоритм распараллеливания по данным

Время расчета задачи методом ПММК определяется временем выполнения внутреннего цикла. Длительность этого цикла зависит от числа моделирующих частиц в области и числа ячеек. Поэтому для сокращения времени работы программы применяется алгоритм распараллеливания по данным, предложенный авторами статьи. При этом ячейки моделируемого объема равномерно распределяются между процессорами. На каждом временном шаге процедуру перемещения каждый процессор выполняет для всех молекул, а процедуру расчета столкновений – только для молекул, расположенных в ячейках, обрабатываемых данным процессором. Затем происходит обмен данными между процессорами и переход к следующему временному шагу. После завершения всех расчетов обработка результатов ведется последовательно на одном процессоре. Ускорение и эффективность этого алгоритма зависят от числа моделирующих частиц в области и числа ячеек – чем больше частиц и ячеек, тем больше времени занимает расчет столкновений относительно процедуры перемещения молекул, тем больше ускорение и эффективность. Область эффективного применения такого алгоритма на p -процессорной системе: число процессоров p не должно превышать числа ячеек.

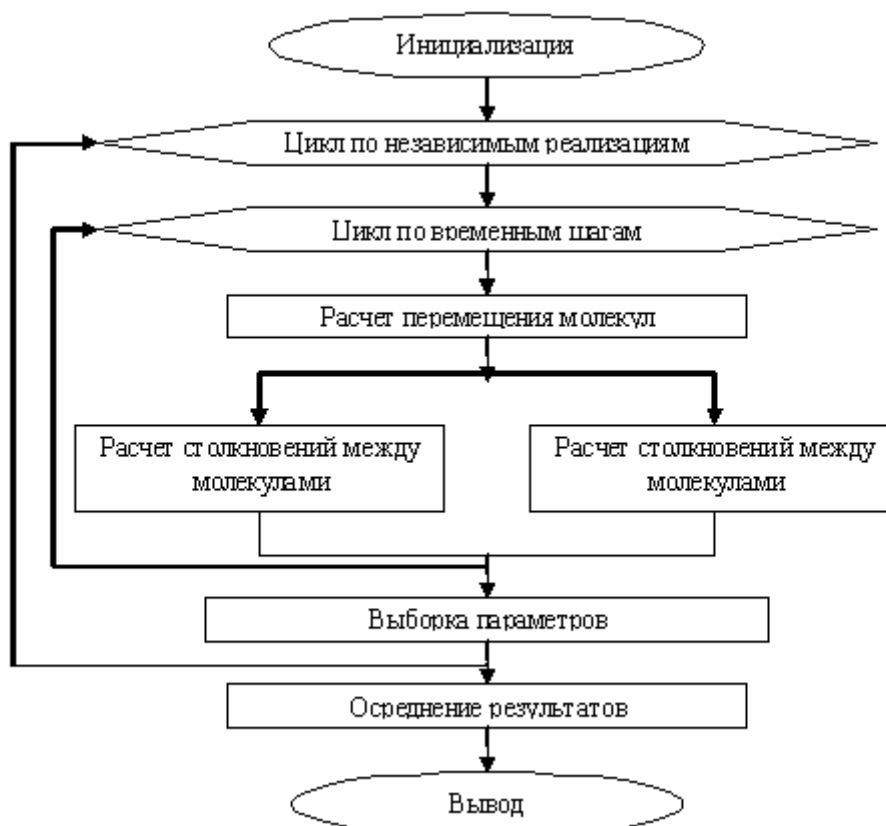


Рис. 3. Укрупненная блок-схема алгоритма распараллеливания по данным

При моделировании рассмотренной выше задачи с использованием алгоритма распараллеливания по данным были получены следующие результаты.

Таблица 3. Время работы алгоритма распараллеливания по данным на различном числе ядер (для вычислительной платформы со следующими характеристиками: два двухъядерных 2 x Intel Xeon (2.00 GHz), 4 Mb Cache, 4Gb RAM).

Время работы последовательного алгоритма	4018 с
Время работы параллельного алгоритма на 2 ядрах	2274 с
Время работы параллельного алгоритма на 3 ядрах	1666 с
Время работы параллельного алгоритма на 4 ядрах	1341 с

Таким образом, получены несколько меньшие ускорение и эффективность, чем при использовании алгоритма распараллеливания по независимым статистическим испытаниям.

5.3. Сравнение методов распараллеливания

В результате проведенной работы можно сделать следующие выводы:

- алгоритм распараллеливания по независимым статистическим испытаниям дает лучшие ускорение и эффективность по сравнению с алгоритмом распараллеливания по данным;
- при достаточно большом числе моделирующих частиц в области и числе ячеек эффективность алгоритма распараллеливания по данным приближается к эффективности алгоритма распараллеливания по независимым статистическим испытаниям;
- алгоритм распараллеливания по независимым статистическим испытаниям имеет ограниченную область эффективного применения ($p \leq N_{MC}$), поэтому при выходе за пределы этой области его выгодно дополнять алгоритмом распараллеливания по данным.

6. Заключение

В ходе работы оба подхода к распараллеливанию тестировались на вычислительной платформе со следующими характеристиками: два двухъядерных 2 x Intel Xeon (2.00 GHz), 4 Mb Cache, 4Gb RAM. В дальнейшем планируется провести расчеты на кластерной системе с большим числом процессоров. При этом алгоритм распараллеливания по независимым статистическим испытаниям скорее всего покажет ускорение близкое к линейному, что следует из самого алгоритма. Ускорение алгоритма распараллеливания по данным может носить нелинейный характер. В дальнейшем эти предположения будут проверены.

Литература

1. Богданов А.В., Гришин И.А., Захаров В.В. и др. Алгоритмы двухуровневой параллелизации для решения нестационарных задач молекулярной газовой динамики // Математическое моделирование. – 2000.–Т.12, №6. –С. 95-101.
2. Мемнонов В.П. Метод прямого статистического моделирования: другой способ соединения расщепляемых процессов // Математическое моделирование. – 1999. –Т.11, №3. –С. 77-82.
3. Базаров И. П., Геворкян Э. В., Николаев П. Н. Неравновесная термодинамика и физическая кинетика. –МГУ, 1989. – 240 с.
4. Bird G.A. Molecular gas dynamics and the direct simulation of gas flows. – Clarendon press, Oxford, 1994. – 458 p.
5. Кац М. Вероятность и смежные вопросы в физике. –Едиториал УРСС, 2003. – 273 с.
6. Ширяев А.Н. Вероятность. –МЦНМО, 2004. – 520 с.