

На правах рукописи

Давыдов Александр Александрович

**Численное моделирование задач газовой
динамики на гибридных вычислительных
системах**

05.13.18 – Математическое моделирование, численные методы и
комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2012

Работа выполнена в Институте прикладной математики
имени М.В. Келдыша Российской академии наук

Научный

руководитель:

доктор физико-математических наук,
Луцкий Александр Евгеньевич

Официальные

оппоненты:

доктор физико-математических наук,
профессор Елизарова Татьяна Геннадьевна

кандидат физико-математических наук,
Семенов Илья Витальевич

Ведущая

организация: Институт теоретической и прикладной механики
им. С.А. Христиановича Сибирского отделения РАН

Защита состоится «___» _____ 2012 г. в ___ часов на заседании
Диссертационного совета Д 002.024.03 при Институте прикладной математи-
ки им. М.В. Келдыша РАН по адресу: 125047, г. Москва, Миусская пл., д. 4.

С диссертацией можно ознакомиться в библиотеке ИПМ им. М.В. Келдыша
РАН.

Автореферат разослан «__» _____ 2012 г.

Ученый секретарь Диссертационного совета,
доктор физико-математических наук

Н.В. Змитренко

Общая характеристика работы

Актуальность темы. Диссертационная работа посвящена актуальной задаче создания прикладного программного обеспечения для гибридных в вычислительных систем с графическими процессорами.

Гибридный суперкомпьютер - это параллельная вычислительная система, в которой, наряду с универсальными процессорами, используются различные типы ускорителей, или сопроцессоров, обладающих нестандартной архитектурой (векторной, мультитредовой, реконфигурируемой под конкретную задачу и т.п.). Основу гибридного суперкомпьютера составляет многопроцессорная вычислительная система из многоядерных процессорных узлов традиционной архитектуры. Каждый из вычислительных узлов снабжается небольшим (обычно 1-4) числом вычислителей нетрадиционной архитектуры, используемых в качестве сопроцессоров для ускорения наиболее трудоемких фрагментов приложения.

В последние несколько лет в области высокопроизводительных вычислений отмечается быстро растущий интерес к нетрадиционным архитектурам вычислителей. Причина этого явления — совершенно конкретные технические факторы, имеющие долговременный, системный характер.

Развитие традиционных процессоров в сторону усложнения их внутренней структуры привело к феномену доминирования вспомогательных операций во времени исполнения программ вычислительного характера (см. напр. [1]). Простые по внутренней структуре процессоры 20-30-летней давности медленно выполняли арифметические операции над вещественными числами, на их фоне вспомогательные действия по вычислению адресов этих чисел и выборке их из памяти были мало заметны. Усложнение внутренней структуры процессоров с целью ускорения операций над вещественными числами было магистральным путем развития аппаратной базы высокопроизводительных вычислений.

Современные процессоры в результате многолетнего развития по этому пути выполняют «полезные» операции с плавающими числами так же быстро, как «вспомогательные» операции по вычислению адресов этих чисел, и в десятки раз быстрее, чем эти числа выбираются из памяти. В итоге вспо-

могательные действия доминируют во времени исполнения программы. Рост скорости выполнения операций над вещественными числами перестает приводить к значительному ускорению счета. Кроме того, тактовые частоты процессоров практически достигли своего предела 5-7 лет назад.

Один из способов решения проблемы состоит в использовании вычислительных устройств, в которых сам характер и набор требуемых вспомогательных действий - иной, чем в традиционном процессоре общего назначения. Такие устройства и называются вычислителями с нетрадиционной архитектурой.

Вычислительные устройства с новыми нефоннеймановскими архитектурами позволяют получить существенный выигрыш в производительности при решении задач математической физики. Однако для достижения высоких результатов требуется хорошо разбираться как в особенностях новых архитектур, так и особенностях применяемых алгоритмов.

Цели и задачи диссертационной работы.

- Исследование применимости графических процессоров (GPU) для решения задач газовой динамики.
- Исследование возможности эффективного использования для расчетов большого числа GPU и обоснование целесообразности создания вычислительных систем на основе графических процессоров.
- Разработка высокоэффективного параллельного программного комплекса для решения задач газовой динамики на гибридных вычислительных системах.
- Математическое моделирование прикладных задач с использованием разработанного программного комплекса.

Научная новизна.

1. Исследована возможность ускорения расчета задач газовой динамики при помощи графических процессоров NVIDIA.
2. Показана эффективность параллельного расчета задач газовой динамики на большом числе графических процессоров.

3. Разработан и реализован параллельный комплекс программ для задач газовой динамики на гибридных вычислительных системах с графическими процессорами NVIDIA. В процессе расчетов получено подтверждение высокой работоспособности и параллельной эффективности комплекса.

Практическая значимость. Реализованный в работе программный комплекс применим для решения широкого круга прикладных задач аэрогазодинамики. Универсальность комплекса позволяет использовать его как на новейших гибридных вычислительных системах, так и на широко распространенных традиционных кластерах. При помощи разработанного комплекса решены важные практические задачи.

Апробация работы. Основные результаты, представленные в диссертационной работе, докладывались и обсуждались на многих российских и международных конференциях:

1. Исследование возможностей ускорения расчета задач аэрогазодинамики с помощью векторных сопроцессоров. V Всероссийская конференция молодых специалистов, СпГУ ИТМО, Санкт-Петербург, 04.2008.
2. Acceleration of a CFD solver using commodity graphics hardware. XIV International conference of the methods of aerophysical research. Russia, Novosibirsc, July 2008.
3. Применение графических процессоров для расчета задач аэродинамики. XVII Всероссийская конференция «Теоретические основы и конструирование численных алгоритмов для решения задач математической физики с приложением к многопроцессорным системам», посвященная памяти К.И. Бабенко. Дюрсо, 2008 год.
4. Макет гибридного суперкомпьютера «МВС-экспресс». XVII Всероссийская конференция «Теоретические основы и конструирование численных алгоритмов для решения задач математической физики с приложением к многопроцессорным системам», посвященная памяти К.И. Бабенко. Дюрсо, 2008 год. Соавторы: С. С. Андреев, С. А. Дбар, А. О. Лацис, Е. А. Плоткина

5. Численное моделирование задач аэро-газодинамики на гибридном суперкомпьютере «МВС-Экспресс». Международная научная конференция «Параллельные вычислительные технологии 2009», Нижний Новгород, 30 марта - 1 апреля 2009 года. Сборник тезисов.
6. Программный комплекс для расчета задач газовой динамики на гибридном суперкомпьютере «МВС-Экспресс». XVIII Всероссийская конференция «Теоретические основы и конструирование численных алгоритмов для решения задач математической физики с приложением к многопроцессорным системам», посвященная памяти К. И. Бабенко. Дюрсо, 2010 год.
7. Numerical Simulation of the Continuous Media Mechanics Problems on the Hybrid Supercomputer «MVS-Express». XV International conference of the methods of aerophysical research. Russia, Novosibirsc, November 2010.
8. О моделях и технологиях программирования суперкомпьютеров с нетрадиционной архитектурой. Научный сервис в сети Интернет: суперкомпьютерные центры и задачи: Труды Международной суперкомпьютерной конференции (20-25 сентября 2010 г., г. Новороссийск). - М.: Изд-во МГУ, 2010. - 694 с. Соавторы: С. С. Андреев, С. А. Дбар, А. О. Лацис, Е. А. Плоткина

Структура работы. Диссертация состоит из введения, 5 глав, заключения и списка цитируемой литературы.

Содержание работы

Во **введении** обосновывается актуальность работы, ставятся цели и задачи исследования, раскрывается его научная новизна и практическая ценность, дается краткое содержание диссертации по главам, излагаются основные положения, выносимые на защиту.

В **первой главе** приводится описание физического устройства графического процессора NVIDIA и аппаратно программных средств CUDA (Compute Unified Device Architecture). Представлена реализация численного метода С.

К. Годунова для системы трехмерных уравнений Эйлера для архитектуры CUDA.

Несмотря на то, что в целом архитектура CUDA хорошо подходит для реализации явных сеточных методов, применимость ее для таких методов как метод С. К. Годунова и алгоритмов на его основе не вполне очевидна. В действительности, в методе С. К. Годунова на каждой грани расчетной сетки решается задача о распаде произвольного разрыва. В классическом подходе эта задача решается итерационным способом, и, скорость сходимости итерационного процесса меняется от грани к грани. Архитектура CUDA устроена таким образом, что параллельные потоки разделяются на так называемые варпы (warps). Внутри варпа все потоки выполняют идентичный набор инструкций. Это означает, что если итерационный процесс сходится с разной скоростью в разных потоках одного варпа, то время выполнения варпа будет определяться временем выполнения самого медленно сходящегося потока. Тем не менее, на практике оказывается, что даже в случае расчетов течений с большими градиентами физических величин, ускорение расчетов на GPU относительно CPU падает незначительно. Это связано, по видимому, тем, что увеличение числа итераций не требует дополнительного чтения из медленной глобальной памяти, а все вычисления производятся с данными из сверхбыстрой регистровой памяти.

На основании проведенных в главе 1 исследований можно сделать вывод, что аппаратно–программные средства CUDA могут быть успешно использованы для решения задач газовой динамики. Так, при расчете трехмерных уравнений Эйлера методом С. К. Годунова были получены ускорения расчета вплоть до 70 раз по сравнению с одним ядром универсального процессора (Рис. 1).

Вторая глава посвящена описанию программного комплекса «Express 3D». Программный комплекс «Express 3D» представляет собой структуру памяти для хранения многоблочных индексных сеток, набор методов для обмена данными между узлами вычислительной системы, набор методов для обмена данными между графическими ускорителями и универсальными процессорами. В комплекс «Express 3D» также входят утилиты для подготовки начальных данных, оптимального разбиения блоков между процессорами.

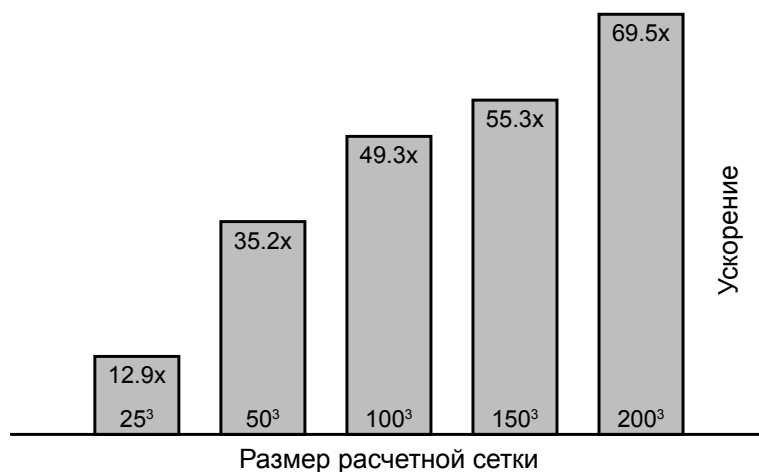


Рис. 1. Зависимость ускорения расчета на GPU по сравнению с одним ядром CPU зависимости от размерности сетки

Модульная структура комплекса позволяет в короткие сроки добавлять к нему новые математические модели и алгоритмы. Реализованные в первой главе алгоритмы для решения уравнений Эйлера, а так же алгоритмы решения квазигазодинамических и модифицированных квазигазодинамических уравнений, о которых речь пойдет в последующих главах естественным образом были включены в состав комплекса в виде отдельных модулей.

Эффективное использование любой многопроцессорной системы не возможно без качественной балансировки загрузки. Это в еще большей степени относится и к гибридным системам. При вычислениях на современных многопроцессорных системах один из основных вкладов в снижение эффективности параллельного приложения несет не обмен данными между узлами, а именно плохая балансировка загрузки.

Время обработки блока на универсальном процессоре практически линейно зависит от его размерности (общего количества элементов). Это позволяет балансировать загрузку процессоров по объему обрабатываемых данных. Фактически это можно делать до начала расчета. Напротив, ускорение расчета на графическом ускорителе по сравнению с универсальным процессором существенно зависит от объема обрабатываемых ядром данных и о некоторых других параметров, таких как размерности блока по направлениям. В таблице 1 приведено ускорение расчета блоков характерного размера при ис-

Размерность блока сетки	4050	11070	12690	19215	37515	43005	52521	101199	117547	197579
Время GPU, с.	11	18	18	17	23	23	23	44	46	69
Время CPU, с.	29	79	96	142	283	323	407	810	978	1700
Ускорение	2,6	4,4	5,3	8,4	12,3	14,0	17,7	18,4	21,3	24,6

Таблица 1 : Ускорение расчета блока на GPU по сравнению с одним ядром CPU в зависимости от размерности блока сетки

пользовании графического процессора. Таким образом, время обработки блока на GPU зависит от размерности блока существенно нелинейно. Более того, заранее трудно спрогнозировать с какой скоростью будет обрабатываться тот или иной блок. Все это приводит к необходимости балансировать загрузку не по объему данных, а по временам обработки каждого блока.

В комплексе «Экспресс-3D» применяется следующий алгоритм распределения блоков по процессорам (или по графическим процессорам) обеспечивающий равномерность загрузки: сначала производится разбиение блоков по объему данных на некоторое количество процессоров (ускорителей). Это может быть и один процессор (ускоритель), главное, чтобы объема оперативной памяти хватило для размещения данных задачи. Далее проводится расчет некоторого числа шагов с целью определения среднего времени обработки каждого блока. После чего производится разбиение блоков на процессоры (графические процессоры) уже по временам обработки блоков с использованием алгоритма оптимального распределения камней по ящикам предложенного А. Шараховым [11], который является переработанным и дополненным алгоритмом решения "Задачи о куче камней"[9].

Современная элементная база, из которой собираются гибридные вычислительные системы, такова, что на один графический ускоритель приходится несколько ядер универсального процессора. Для управления каждым графическим процессором отводится одно ядро универсального процессора, а остальные, как правило, не задействованы.

Комплекс «Экспресс-3D» позволяет использовать для расчета не только

графические процессоры, но и ядра универсального процессора. Так, если узел вычислительной системы состоит из 8 ядер универсального процессора и 2 графических ускорителей (например, СК «Ломоносов», МГУ), то работу следует разбивать на 8 частей - на 6 универсальных ядер и на 2 пары CPU+GPU. Конечно, сложность балансировки загрузки при этом существенно возрастает.

В работе предложен эффективный алгоритм для распределения блоков между процессорами и ускорителями. Основная идея алгоритма состоит в том, чтобы мелкие блоки, которые не очень хорошо ускоряются на GPU, но число которых, как правило, велико, считать на CPU. Применение такого подхода позволяет на 30-40% повысить эффективность узла вычислительной системы (по сравнению с вычислениями только на ускорителях). В таблице 2 приведены времена расчета и ускорение расчета на одном и двух узлах установки К-100, работающей в ИПМ им. М. В. Келдыша РАН [12], при использовании только GPU и при расчете на GPU и CPU одновременно.

	1 узел К-100		2 узла К-100	
	3GPU	3GPU+8CPU	6GPU	6GPU+16CPU
Время	65.98	47.72	38.8	28.15
ускорение		1.38		1.37

Таблица 2 : Сравнение времени расчета на GPU и на CPU+GPU.

При решении задач газовой динамики на многопроцессорных вычислительных системах возникает необходимость обмена информацией между процессорами. Количество информации которой необходимо обмениваться зависит от применяемого численного метода. В то же время частота обменов, и, как следствие, нагрузка на коммуникационную среду зависит от скорости выполнения вычислений между обменами. С применением для вычислений графических процессоров скорость расчета, как показано в главе 1, возрастает в десятки раз. Соответственно, соразмерно возрастает и нагрузка на коммуникационную среду.

Для получения высокоэффективного параллельного приложения требуется не только использование современного оборудования, но и моделей программирования соответствующим этому оборудованию.

В комплексе «Express 3D» в качестве коммуникационной среды и модели параллельного программирования была использована технология Shmem [13].

Модель программирования shmem [15] подразумевает взаимодействие независимых процессов, каждый – со своей локальной памятью, занумерованных по порядку от нуля, и в этом она похожа на модель программирования MPI [6]. Отличие от MPI состоит в том, что обмены данными между процессами являются односторонними. Чтобы данные были переданы из процесса А в процесс Б, требуется не согласованная активность обоих процессов, как в MPI, а лишь «желание» одного из участников. Например, процесс А может «насильно» послать данные процессу Б, без какой-либо ответной активности с его стороны. Процесс Б, в свою очередь, может «насильно» прочитать данные из процесса А. В англоязычных описаниях этой модели ее принято называть «модель put/get», в отличие от принятой в MPI «модели send/receive».

Выбор модели программирования Shmem был обусловлен следующими факторами. Во-первых, первый гибридный кластер, появившийся у нас в стране - «МВС-Экспресс» [3], долгое время не имел другой модели параллельного программирования. Во-вторых, модель односторонних обменов, по мнению автора, существенно упрощает понимание программы и сокращает время разработки. В-третьих, существуют программные реализации shmem основанные на MPI [8, 16], которые хотя и несколько снижают производительность относительно реализаций shmem привязанных к оборудованию, позволяют упростить программирование, сохранив при этом переносимость кода на многие вычислительные системы.

Прежде чем обмениваться информацией между двумя графическими процессорами, находящимися на разных узлах вычислительной системы, необходимо сначала скопировать данные из памяти графического процессора в память универсального процессора, а по завершении обмена между универсальными процессорами, передать данные обратно графическому процессору. Это также вносит вклад в снижение параллельной эффективности приложений. Системы программирования которые позволили бы избежать такого двухуровневого обмена данными в настоящий момент только разрабатываются.

В таблице 3 приведены времена расчетов уравнений Эйлера (обтекание крылатого тела, см. Главу 4) для различного числа универсальных и графических процессоров. Количество блоков сетки – 132, общее число узлов $\sim 8 \cdot 10^6$.

	1x CPU core	11xCPU cores	3x GPU	6x GPU	9x GPU	12x GPU	24x GPU	30x GPU
Время, с	13510	1424	360	211	136	108	60	46
Ускорение	1,0	9,5	37,5	64,0	99,3	125,6	226,7	291,8
Эффективность CPU	100%	86%						
Эффективность GPU			100%	85%	88%	84%	77%	78%

Таблица 3 : Время расчета уравнений Эйлера на различном числе CPU и GPU.

Масштабирование данной задачи на более чем 30 вычислителей не целесообразно — количество блоков и их размеры не позволяют добиться качественной балансировки загрузки.

В таблице 4 приведены времена расчета системы квазигазодинамических уравнений (задача о каверне с движущейся крышкой, см главу 3). Количество блоков сетки – 216, размерность каждого блока $50 \times 50 \times 50$.

Число GPU	4	6	9	18	27	36	54	72	108
Время, с	11090	7580	5080	2670	1800	1360	940	730	530
Эффективность, %	100	97,5	97,0	92,3	91,3	90,6	87,4	84,4	77,5

Таблица 4 : Время расчета квазигазодинамических уравнений на различном числе GPU.

В **Третьей главе** описано построение численного метода для решения квазигазодинамической системы (1)-(2). Описание и физический смысл которой приведены в [5]

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} j_{mi} = 0, \quad (1)$$

$$\frac{\partial}{\partial t} \rho u_i + \frac{\partial}{\partial x_j} j_{mj} u_i + \frac{\partial}{\partial x_i} p = \frac{\partial}{\partial x_j} \Pi_{ji}, \quad (2)$$

$$\frac{\partial}{\partial t}E + \frac{\partial}{\partial x_j}j_{mj} \left(\frac{E+p}{\rho} \right) + \frac{\partial}{\partial x_j}q_j = \frac{\partial}{\partial x_j}\Pi_{ji}u_i. \quad (3)$$

Здесь и далее по повторяющимся индексам подразумевается суммирование, $E = \rho (\mathbf{u}^2/2 + \varepsilon)$ – полная энергия. Поток массы задается как

$$j_{mi} = \rho(u_i - w_i), \quad w_i = \frac{\tau}{\rho} \left(\frac{\partial}{\partial x_j}\rho u_j u_i + \frac{\partial}{\partial x_i}p \right), \quad (4)$$

Также рассматривается модифицированная квазигазодинамическая система (5)-(7):

$$\frac{\tau}{2} \frac{\partial^2 \rho}{\partial t^2} + \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}j_{mi} = 0, \quad (5)$$

$$\frac{\tau}{2} \frac{\partial^2 \rho u_i}{\partial t^2} + \frac{\partial}{\partial t}\rho u_i + \frac{\partial}{\partial x_j}j_{mj}u_i + \frac{\partial}{\partial x_i}p = \frac{\partial}{\partial x_j}\Pi_{ji}, \quad (6)$$

$$\frac{\tau}{2} \frac{\partial^2 E}{\partial t^2} + \frac{\partial}{\partial t}E + \frac{\partial}{\partial x_j}j_{mj} \left(\frac{E+p}{\rho} \right) + \frac{\partial}{\partial x_j}q_j = \frac{\partial}{\partial x_j}\Pi_{ji}u_i. \quad (7)$$

Тензор вязких напряжений $\mathbf{\Pi}$, тепловой поток \mathbf{q} и соотношения, замыкающие систему, определяются следующими соотношениями:

$$\Pi_{ij} = \Pi_{ij}^{NS} + \tau \rho u_i \left(u_k \frac{\partial}{\partial x_k} u_j + \frac{1}{\rho} \frac{\partial}{\partial x_j} p \right) + \tau \delta_{ij} \left(u_k \frac{\partial}{\partial x_k} p + \gamma p \frac{\partial}{\partial x_k} u_k \right), \quad (8)$$

$$\Pi_{ij}^{NS} = +\mu \left(\frac{\partial}{\partial x_i} u_j + \frac{\partial}{\partial x_j} u_i - \frac{2}{3} \delta_{ij} \frac{\partial}{\partial x_k} u_k \right), \quad (9)$$

$$q_i = q_i^{NS} - \tau \rho u_i u_j \left(\frac{\partial}{\partial x_j} \varepsilon + p \frac{\partial}{\partial x_j} \frac{1}{\rho} \right), \quad q_i^{NS} = -\kappa \frac{\partial}{\partial x_i} T, \quad (10)$$

$$p = \rho \varepsilon (\gamma - 1), \quad T = \frac{p}{\rho R}. \quad (11)$$

Динамическая вязкость μ , коэффициент теплопроводности κ и релаксационный параметр τ , имеющий размерность времени, имеют следующий вид

$$\mu = \mu_\infty \left(\frac{T}{T_\infty} \right)^\omega, \quad \kappa = \mu \frac{\gamma R}{(\gamma - 1) \text{Pr}}, \quad \tau = \frac{\mu}{p \text{Sc}}, \quad (12)$$

КГД система (1) – (3) получена из балансного уравнения в предположении малости времени τ по сравнению с характерным газодинамическим временем $t_{\text{газ}}$. Учитывая следующий член разложения функции распределения в ряд Тейлора, приходим к варианту системы (5) – (7).

Эта система имеет гиперболический тип [7]. Важным свойством КГД системы уравнений (5) – (7) является ее близость к классическим уравнениям Навье – Стокса (НСУ). Связь между этими двумя системами может быть представлена в виде

$$\text{КГД} = \text{НСУ} + O(\text{Kn}^2), \quad (13)$$

где Kn – число Кнудсена. Следует отметить, что сами уравнения Навье – Стокса получены из уравнения Больцмана с точностью до членов второго порядка малости по числу Кнудсена. Дополнительные диссипативные члены в КГД системе, содержащие множитель τ , обращаются в нуль в тех областях течения, где решение удовлетворяет стационарным уравнениям Эйлера. Заметим также, что наличие вторых производных по времени в уравнениях (5) – (7) окажется важным фактором при формировании вычислительного алгоритма.

Конечно разностная аппроксимация систем (1) – (3) и (5) – (7), строится на основе метода контрольных объемов. Все газодинамические параметры относятся к центрам ячеек разностной сетки, а в качестве контрольного объема берется сама ячейка. Интегрируя уравнения (1) – (3) по объему ячейки, мы получаем законы сохранения в интегральной форме. При этом изменение газодинамических параметров в ячейке определяется суммой потоков консервативных переменных (плотности ρ , импульса $\rho \mathbf{u}$ и полной энергии E) через все ее грани. Для аппроксимации пространственных производных, входящих в выражения для потоков, используются центральные разности, а значения газодинамических переменных в центрах граней вычисляются с помощью линейной интерполяции. Для обеспечения устойчивости счета системы (1) – (3) по явной схеме к релаксационному параметру τ в (12) добавляется слагаемое, пропорциональное шагу пространственной сетки.

Для построения численного алгоритма для гиперболического варианта КГД–системы уравнений (5) – (7) перепишем ее в компактной форме:

$$\tau_f \frac{\partial^2}{\partial t^2} \mathbf{U} + \frac{\partial}{\partial t} \mathbf{U} = \nabla \mathbf{F}_{QGD}. \quad (14)$$

Здесь $\mathbf{U} = (\rho, \rho u_1, \rho u_2, \rho u_3, E)^T$ – вектор консервативных переменных, \mathbf{F}_{QGD} – потоки, соответствующие системе (1) – (3), τ_f – новый релаксационный параметр. Физический смысл дополнительных слагаемых со второй производной по времени, а с ними и параметра τ_f становится понятным, если переписать (14) в форме релаксации потоков:

$$\frac{\partial}{\partial t} \mathbf{U} = \nabla \mathbf{F}, \quad \tau_f \frac{\partial}{\partial t} \mathbf{F} = \mathbf{F}_{QGD} - \mathbf{F}. \quad (15)$$

В соответствии с (15) потоки консервативных переменных не могут в отличие от системы (1) – (3) мгновенно достичь значений \mathbf{F}_{QGD} , они стремятся к ним, стартуя со значений в предыдущий момент времени, и τ_f – характерное время релаксации потоков. При малых значениях τ_f можно пренебречь левой частью второго из уравнений (15), и мы получим систему (1) – (3), поскольку в этом случае $\mathbf{F} = \mathbf{F}_{QGD}$.

Обратим внимание, что идея релаксации потоков не нова. Еще в середине прошлого века физики обратили внимание на нефизичное свойство уравнения теплопроводности, основанного на законе Фурье: бесконечная скорость распространения возмущений, свойственная параболическим уравнениям. Этот факт приводит к явному расхождению с экспериментальными данными при исследовании быстро протекающих процессов с большими градиентами температур, таких, как течения разреженных газов, низкотемпературный теплоперенос в твердых телах, электронная теплопроводность в плазме и др. Чтобы избежать этого и обеспечить конечную скорость распространения возмущений, было предложено гиперболическое уравнение теплопроводности [16].

При построении разностной схемы для гиперболической системы (5) – (7) используем ту же пространственную аппроксимацию потоков \mathbf{F}_{QGD} , как и для системы (1) – (3). Второе уравнение в (15) является линейным ОДУ первого порядка, и может быть проинтегрировано аналитически на интервале времени (t^j, t^{j+1}) :

$$\mathbf{F}^{j+1} = \mathbf{F}^j D + \mathbf{F}_{QGD}(1 - D), \quad D = \exp(-\Delta t/\tau_f), \quad \Delta t = t^{j+1} - t^j. \quad (16)$$

Относительно выбора параметра τ_f , можно сказать, что с одной стороны, он должен быть достаточно большим, чтобы обеспечить наилучшую устойчивость схемы, а с другой – достаточно малым, чтобы получаемое решение не слишком сильно отличалось от решения исходной системы.

Переход на новый временной слой осуществляется в следующем порядке. Сначала по значениям газодинамических величин в момент времени t^j вычисляются потоки \mathbf{F}_{QGD} для уравнений (1) и (3), описывающих изменение скалярных переменных. Затем в соответствии с (16) находим новые значения потоков этих переменных \mathbf{F}^{j+1} . Используя простейшую аппроксимацию по времени первого уравнения (15), $\mathbf{U}^{j+1} = \mathbf{U}^j + \Delta t \cdot \nabla \mathbf{F}^{j+1}$, вычисляем новые значения плотности и энергии. После этого та же процедура производится с уравнением импульса (2), причем потоки \mathbf{F}_{QGD} для этого уравнения вычисляются по вновь полученным значениям ρ^{j+1} и ε^{j+1} . Переход на слой t^{j+1} завершается вычислением новых значений скоростей \mathbf{u}^{j+1} .

$$\tau = \frac{\mu}{pSc} + \alpha \frac{h}{c}, \quad (17)$$

где c – локальная скорость звука, $h = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2}$, $\Delta x_1, \Delta x_2, \Delta x_3$ – шаги прямоугольной пространственной сетки, α – числовой параметр порядка единицы, подбираемый экспериментально.

Построенная таким образом разностная схема была применена к решению упоминавшейся выше задачи о течении в каверне.

Для различных значений релаксационного параметра τ_f были численно определены максимальные значения β , обеспечивающие вычислительную устойчивость. Результаты вычислений для случая $Ma = 0.05$ изображены на рис. 2. Точки на оси ординат соответствуют $\tau_f = 0$, т.е. решению исходной системы (3) – (5). Видно, что значения β_{max} оказываются значительно больше, чем при решении исходной системы и практически не зависят от шага пространственной сетки. Таким образом, мы имеем практически Курантовское условие устойчивости для существенно дозвукового течения вязкого га-

за. При этом оптимальное значение τ_f , при котором достигается β_{max} , почти пропорционально шагу h .

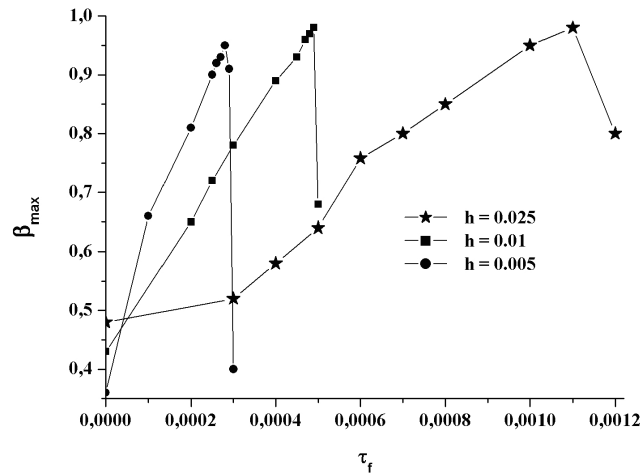


Рис. 2. Зависимость максимально допустимого числа Куранта β_{max} от релаксационного параметра τ_f для различных пространственных сеток.

При решении той же задачи для случая $Ma = 0.1$ были получены аналогичные результаты. Максимально допустимые значения β_{max} близки к единице, а оптимальные значения τ_f , при которых эти значения достигаются удвоились по сравнению с вариантом $Ma = 0.05$. В Таблице 5 представлено сравнение оптимальных значений параметра τ_f и соответствующих им чисел Куранта β_{max} для различных вариантов расчета.

h	Параметры	$Ma = 0.05$	$Ma = 0.1$
0.025	τ_f	1.1×10^{-3}	2.2×10^{-3}
	β_{max}	0.98	0.99
0.01	τ_f	4.9×10^{-4}	9.8×10^{-4}
	β_{max}	0.98	0.95
0.005	τ_f	2.8×10^{-4}	5.9×10^{-4}
	β_{max}	0.95	0.94
0.0025	τ_f	1.62×10^{-4}	
	β_{max}	0.88	
0.00125	τ_f	1.09×10^{-4}	
	β_{max}	0.82	

Таблица 5 : Оптимальные значения параметра τ_f и соответствующие значения β_{max}

Анализируя эти результаты, можно заметить, что для всех вариантов расчета оптимальное значение τ_f близко к величине $Ma \cdot h$, т.е. релаксационный параметр должен быть одного порядка с шагом по времени. С физической точки зрения этот результат представляется вполне естественным. Следует отметить, что безразмерное время установления стационарного течения практически не зависит ни от сетки, ни от релаксационного параметра. Таким образом, достигнутое увеличение шага по времени приводит к реальной экономии времени расчета задачи.

Отметим, что установившиеся при β_{max} распределения газодинамических величин и структура течения, полученные по модифицированной системе, отличаются от полученных по системе (1) – (3) на доли процента. Это означает, что модификация схемы не вносит сколько-нибудь существенных деформаций в решение задачи. Заметим также, что можно еще больше уменьшить разницу между решениями исходной и модифицированной задачи, выбирая τ_f пропорциональным h , но с меньшим коэффициентом пропорциональности. При этом также получится близкое к курантовскому условие устойчивости, но с несколько меньшим коэффициентом пропорциональности, чем при выборе оптимального τ_f .

В **Четвертой главе** с помощью разработанного программного комплекса проводится расчет аэродинамических характеристик летательного аппарата (крылатого тела) рис. 3 обтекаемого сверхзвуковым потоком газа под различными углами атаки при различных числах Маха набегающего потока.

Расчетная сетка состояла из 132 блоков. Общее число ячеек сетки $1.8 \cdot 10^6$.

Для каждого из углов атаки $\alpha = 5^\circ, 10^\circ, 15^\circ$ моделировались течения с различными числами Маха $M = 1.5, 2.0, 2.5$ набегающего потока. Получены зависимости сопротивления подъемной силы в зависимости от углов атаки и чисел Маха набегающего потока.

Общая картина течения для $M = 2.0, \alpha = 10^\circ$ изображена на рисунке 4

На рисунке 5 представлены зависимости подъемной силы от угла атаки для различных чисел Маха. Для всех чисел Маха зависимость подъемной силы от угла атаки практически линейно, что хорошо согласуется с теорией [14].

Давление торможения определяется параметрами набегающего потока,

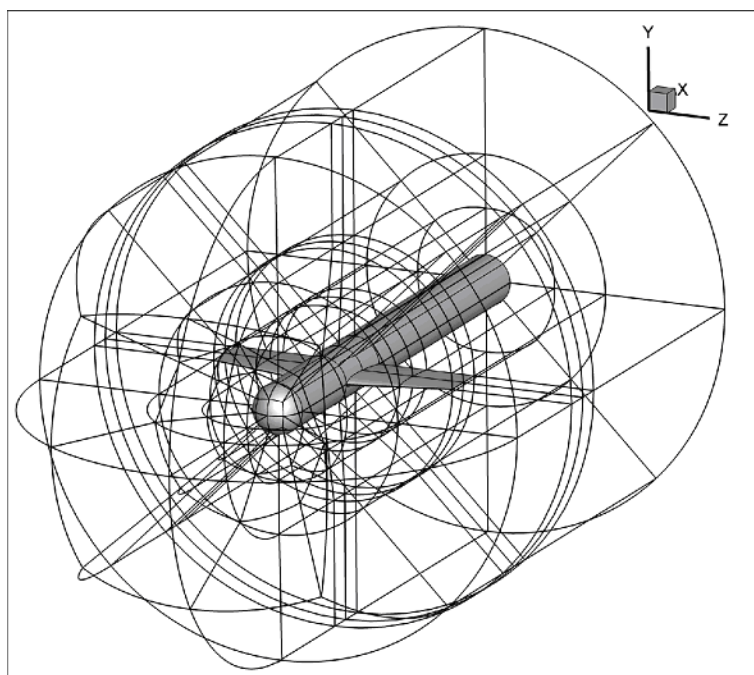


Рис. 3. Модель крылатого тела и расположение блоков индексной сетки.

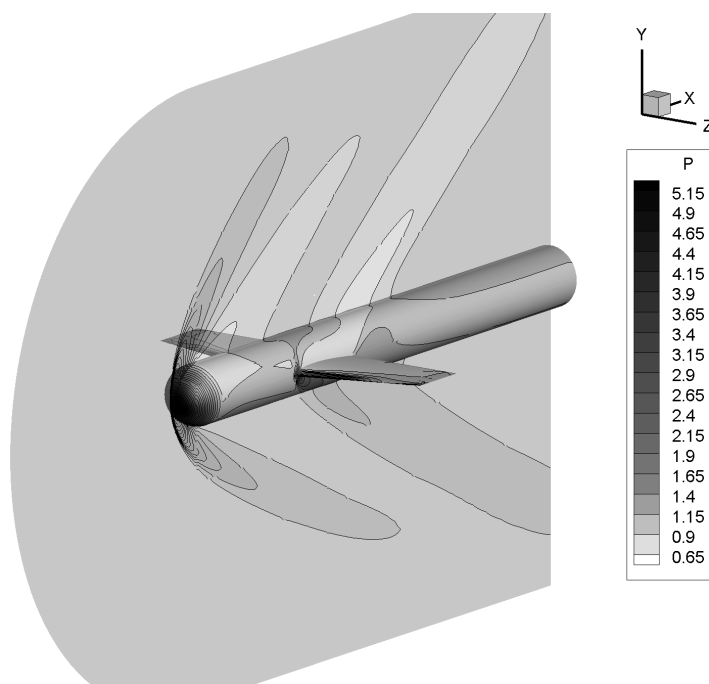


Рис. 4. Общая картина течения, уровни давления.

и не должно зависеть от угла атаки, при этом положение точки торможения, конечно, меняется. В таблице 6 приведено давление торможения для числа Маха набегающего потока $M = 2.5$ и погрешность относительно аналитического решения $P = 8.52613589$ [14].

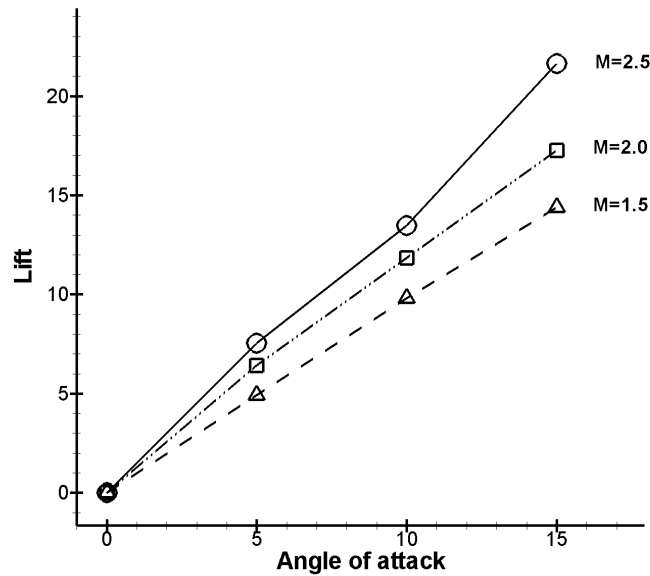


Рис. 5. Зависимость подъемной силы от угла атаки.

Угол атаки	Давление торможения	Погрешность относительно точного решения
5	8.6378	1.31 %
10	8.5646	0.45 %
15	8.4986	0.32 %

Таблица 6 : Давление торможения для $M = 2.5$

В **Пятой главе** в рамках модели модифицированных квазигазодинамических уравнений проводится моделирование истечения тяжелой жидкости из резервуара.

Модель резервуара представляет из себя куб с равномерной пространственной сеткой. В начальный момент жидкость покоится рис. (6). В одной из боковых стенок имеется отверстие прямоугольной формы совпадающее с ячейками сетки. Для отверстия задаются условия внешнего давления, плотность и компоненты скорости “сносятся” из ячейки резервуара. Считая что объем вытекающей жидкости много меньше объема резервуара положим на свободной верхней границе входящий поток, равный по массе вытекающему

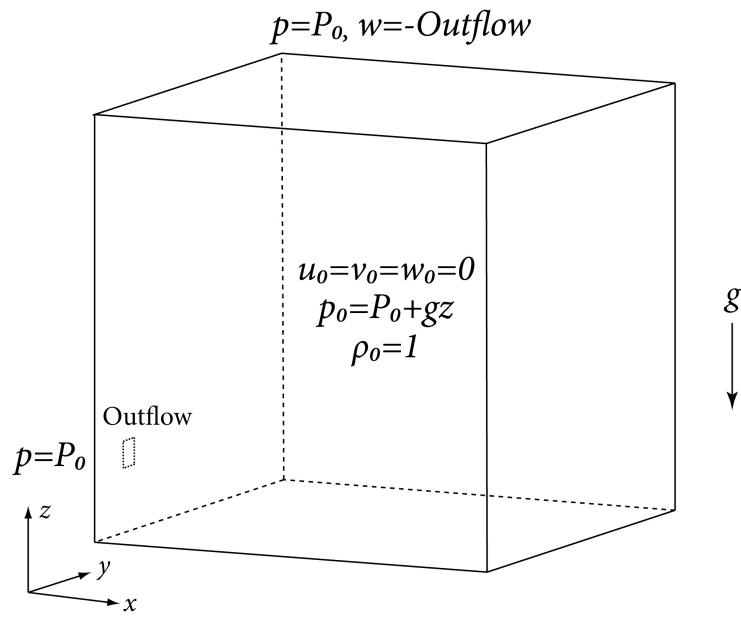


Рис. 6. Модель резервуара.

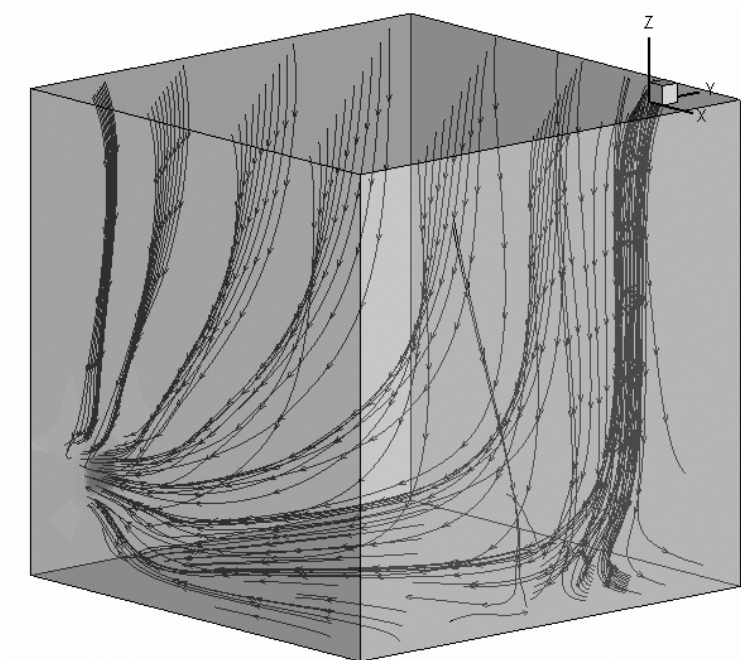


Рис. 7. Линии тока установившегося течения.

через отверстие.

На рисунке 7 показаны линии тока установившегося течения.

В работе были исследованы течения с различными положениями отверстия при различных формах резервуаров.

В **заклучении** приводятся основные выводы и результаты работы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

1. Разработан и реализован комплекс программ «Express-3D» для решения задач газовой динамики на многоблочных индексных сетках. Комплекс позволяет решать задачи, как на универсальных кластерах, так и на гибридных вычислительных системах. Модульная структура комплекса позволяет в короткие сроки добавлять к нему новые математические модели и алгоритмы. Показана высокая параллельная эффективность комплекса при расчетах на большом числе графических процессоров.
2. В рамках разработанного комплекса реализованы алгоритмы для решения уравнений Эйлера, алгоритмы для расчета квазигазодинамических и модифицированных (гиперболизированных) квазигазодинамических уравнений для архитектуры CUDA.

3. При помощи программного комплекса «Express-3D»

Проведено экспериментальное исследование устойчивости численного алгоритма для модифицированных (гиперболизированных) квазигазодинамических уравнений. Определены зависимости максимально допустимых (при которых сохраняется устойчивость) чисел Куранта от значений релаксационного параметра.

Проведено численное исследование аэродинамических характеристик летательного аппарата (крылатого тела) при различных углах атаки и числах Маха набегающего потока. Получены зависимости подъемной силы от угла атаки и чисел Маха.

Построена модель и исследованы параметры истечения тяжелой жидкости из резервуара через отверстие. Определены особенности течения в зависимости от положения отверстия и формы резервуара.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ РАБОТЫ

Статьи в рецензируемых журналах, рекомендованных ВАК

- [1] *Давыдов А. А.* Применение графических сопроцессоров на суперкомпьютере «МВС Экспресс» для расчета задач аэро-газодинамики// Научно-технический вестник СПГУ ИТМО, № 54, Технологии высокопроизводительных вычислений и компьютерного моделирования, сс. 178-180, 2008 г.
- [2] *Давыдов А. А.* Численное моделирование задач аэро-газодинамики на гибридном суперкомпьютере «МВС-Экспресс»// Журнал Математическое моделирование, том 22, № 4, 2010 г., с. 90-98.
- [3] *Давыдов А. А., Лацис А. О., Луцкий А. Е., Смольянов Ю. П., Четверушкин Б. Н., Шильников Е. В.* Многопроцессорная вычислительная система гибридной архитектуры «МВС-Экспресс»// ДОКЛАДЫ АКАДЕМИИ НАУК, 2010, том 434, № 4, с. 459-463
- [4] *Давыдов А. А., Четверушкин Б. Н., Шильников Е. В.* Моделирование течений несжимаемой жидкости и слабосжимаемого газа на многоядерных гибридных вычислительных системах// Журнал вычислительной математики и математической физики, 2010, том 50, № 12, с. 2275-2284

Список литературы

- [5] *Елизарова Т. Г.* Квазигазодинамические уравнения и методы расчета вязких течений. – М: Научный мир, 2007. – 352 с.
- [6] *Воеводин В. В., Воеводин Вл. В.* Параллельные вычисления. - СПб.: БХВ-Петербург, 2002. - 600 с.
- [7] *Злотник А.А., Четверушкин Б.Н.* Параболичность квазигазодинамической системы уравнений, гиперболичность одной ее модификации и

- устойчивость малых возмущений для них // Ж. вычисл. мат. и матем. физ., 2008, т. 49, №3, сс. 445 - 472.*
- [8] *Корж А. А. Масштабирование Data-Intensive приложений с помощью библиотеки DISLIB на суперкомпьютерах Blue Gene/P и “Ломоносов” // Труды конференции “Научный сервис в сети Интернет-2011”*
- [9] *Романовский И. В. Алгоритмы решения экстремальных задач, “Наука”, М., 1977, с.248.*
- [10] *Четверушкин Б. Н. Кинетические схемы и квазигазодинамическая система уравнений. М.: Макс Пресс, 2004.*
- [11] *http://delphiworld.narod.ru/base/stones_in_boxes.html*
- [12] *<http://www.kiam.ru/MVS/resources/k100.html>*
- [13] *Климов Ю. А., Лацис А. О. Руководство по использованию сети МВС-Экспресс на К-100.
<http://www.kiam.ru/MVS/documents/k100/mvseuserguide.html>*
- [14] *Ландау Л. Д., Лифшиц Е. М. Теоретическая физика: Учебное пособие. В 10 т. Т. VI. Гидродинамика. – 3-е изд. перераб. – М: Наука. Гл. ред физ-мат. лит., 1986. – 736 с.*
- [15] *<http://www.shmem.org/>*
- [16] *<http://www.kiam.ru/MVS/research/>*
- [17] *<http://www.nvidia.com>*

Подписано к печати 23.04.2012. Формат 60x84/16. Усл. печ. л. 1,5. Тираж 100 экз. Заказ.
ИПМ им.М.В.Келдыша. 127047, Москва, Миусская пл., 4