

**Р. Хокни
Дж. Иствуд**

**Численное
моделирование
методом частиц**

Roger W. Hockney
Professor of Computer Science
Reading University
and
James W. Eastwood
Department of Computer Science
Reading University

COMPUTER SIMULATION
USING PARTICLES

McGraw-Hill
International Book Company

**Р. Хокни
Дж. Иствуд**

**Численное
моделирование
методом частиц**

Перевод с английского
А. С. Липатова и А. Н. Полюдова

под редакцией
Р. З. Сагдеева и В. И. Шевченко



Москва «Мир» 1987

ББК 22.193

X 46

УДК 519.6

Хокни Р., Иствуд Дж.

X 46 Численное моделирование методом частиц: Пер. с
англ. — М.: Мир, 1987. — 640 с.

Монография известных английских специалистов, охватывающая широкий круг проблем численного моделирования физических процессов методом частиц. Материал изложен подробно, представлены основные этапы моделирования, много внимания уделено тестированию программ на ЭВМ.

Для математиков-прикладников, специалистов по численным методам, физиков, инженеров, студентов и аспирантов университетов.

X $\frac{1702070000-391}{041(01)-87}$ 44-87, ч. 1

ББК 22.193

Редакция литературы по математическим наукам

© 1981 McGraw-Hill Inc.

© перевод на русский язык, с дополнениями,
«Мир», 1987

Предисловие редакторов перевода

Предлагаемая читателю монография известных английских специалистов Р. Хокни и Дж. Иствуда является первой в мировой литературе книгой, целиком посвященной методу частиц для моделирования плазмы и плазмоподобных сред. Бурное развитие вычислительной техники привело к появлению принципиально нового мощного средства исследования физических процессов — вычислительному эксперименту. В условиях, когда теоретические методы наталкиваются на серьезные трудности, а натурные эксперименты либо чрезмерно трудоемки, либо принципиально невозможны, численное моделирование часто оказывается единственным способом изучения рассматриваемых процессов. Важнейшей областью применения вычислительного эксперимента стала физика плазмы и плазмоподобных сред, в особенности исследование нелинейных процессов. Среди методов дискретного моделирования плазмы ведущую роль играет метод частиц. Высокая эффективность этого метода обусловлена его универсальностью, относительно невысокой стоимостью вычислительных экспериментов по сравнению с натурными и практически неограниченными возможностями диагностики моделируемых явлений.

В книге нашли отражение все этапы проведения вычислительного эксперимента: физическая постановка задачи, построение математической модели, разработка и реализация алгоритма и сам расчет. Особую ценность придает книге материал по многомерным моделям, что отражает во многом вклад авторов в развитие метода частиц. Большое внимание уделяется повышению эффективности алгоритмов. Кроме того, в книге изложены основные вопросы вычислительной математики, необходимые при разработке моделей частиц: описаны и исследованы соответствующие разностные схемы, методы решения уравнений для поля, изучены возможные нефизические численные эффекты. Применение метода частиц подробно проиллюстрировано на примере нескольких задач физики плазмы, астрофизики, физики полупроводников, твердого тела. Изложение ведется на «физическом» уровне строгости и, таким образом, книга обращена скорее к физикам, занимающимся вычислительным экс-

периментом, чем к математику-прикладнику. Цель книги, однако, заключается не в том, чтобы дать строгое математическое обоснование метода частиц и связанных с ним вопросов вычислительной математики. Более строгое изложение ряда вопросов читатель найдет в отечественной литературе по численным методам.

Нужно отметить, что авторы практически не отразили вклад советских исследователей в развитие метода частиц и смежных вопросов вычислительной математики. Поэтому мы сочли необходимым дополнить список литературы работами советских авторов.

Книга будет полезна математикам-прикладникам, физикам, использующим модели частиц в своих исследованиях, специалистам по вычислительной математике и программированию, разрабатывающим дискретные кинетические модели. Определенный интерес она представит также для разработчиков пакетов прикладных программ. В качестве учебного пособия она будет полезна студентам и аспирантам соответствующих специальностей.

Перевод книги выполнен А. С. Липатовым (гл. 1, 2, 6, 9—11) и А. Н. Полюдовым. Большую помощь в редактировании гл. 10 и 11 оказали Н. А. Баннов и А. А. Клыпин.

*Р. З. Сагдеев,
В. И. Шевченко*

Предисловие

Уже давно решающим моментом как в теоретических, так и экспериментальных исследованиях стало умение проследить движение большого числа частиц в полях, создаваемых ими самими и приложенных извне. Несмотря на конкуренцию со стороны гидродинамических методов расчета, вычислительные модели, основанные на прослеживании траекторий десятков тысяч частиц, будут, по-видимому, еще некоторое время занимать в подобных исследованиях центральное место. В настоящее время все более привлекательными становятся гибридные модели (частицы плюс гидродинамика), особенно в плазме, где характерные электронные и ионные времена очень сильно различаются вследствие малости отношения масс m_e/m_i . Долгая история численных исследований движения многих частиц, связанных и с теориями и с приложениями, предстает во множестве отдельных публикаций, которые заслуживают того, чтобы собрать их и издать вместе, как это сделали в данном случае авторы настоящей книги.

Во многих работах изучалось движение частиц в электровакуумных приборах, как, например, в первых работах Хартри и Бунемана по магнетронам и работах Нордсика 40-х гг. по лампе бегущей волны, на смену которым в 50-х гг. пришли усовершенствованные модели Полтера, Тьена и Рове. Эти авторы контролировали длину, или период, электронных волн, проходящих через изучаемые устройства, причем расчеты велись часами на ЭВМ, которые по сегодняшним меркам примитивны, однако были разработаны методы расчета движения частиц и соответствующих полей, что пригодилось позднее при работе на гораздо более мощных машинах. В их работах впервые получил объяснение нелинейный режим мощных усилителей и генераторов микроволнового диапазона. Кроме того, Тьен показал, что частота зависит от сглаживания пульсаций, возникающих в ограничивающем ток пространственном заряде (дебаевском облаке) — виртуальном катоде, обусловленном динамикой минимума потенциала. Теперь уже подобными исследованиями никого не удивить, и их подчас выполняют студенты в порядке освоения программы.

Примерно в 1960 г. Бунеман и Доусон показали, как нужно проводить моделирование одномерной плазмы с помощью большого числа частиц. В частности, Доусон выполнил различные вычислительные эксперименты по плазме, опирающиеся на классическую кинетическую теорию, и поразил воображение теоретиков-плазменщиков. Этот шаг вперед предопределил приложения моделирования к решению большого числа задач бесстолкновительной плазмы, важных для осуществления управляемого термоядерного синтеза, если не считать того, что одномерия было недостаточно, а в двумерном и трехмерном случаях прямые вычисления кулоновской силы были нереальны.

В 1963 г. Хокни, работая у Бунемана в Стэнфордском университете в Беркли, начал с того, что придумал метод быстрого решения уравнения Пуассона на двумерной пространственной сетке, и это сразу сделало двумерное электростатическое моделирование привлекательным и экономичным. В последующие несколько лет двумерное и трехмерное моделирование, включая теоретические исследования о влиянии сеток по пространству и времени и приложения к неустойчивостям и процессам переноса в термоядерной плазме, быстро и успешно развивалось в Стэнфорде, Беркли и Принстоне, а также в национальных лабораториях в Соединенных Штатах Америки, Великобритании, ФРГ и Японии.

Начальные знания выросли в целую науку, обогатившись теоретической базой и множеством самых разных изобретений (эффективные и точные алгоритмы, гибкая архитектура), которые улучшались в приложениях. Работы Хокни и Иствуда этого периода, а также 70-х гг. касаются многого и необходимы любому, кто занимается моделированием методом частиц.

Авторы собрали свои работы в своеобразном и полезном виде, начав с подробного рассмотрения сеток по пространству и времени, а также проектирования программы. Далее они приводят ряд интересных физических и отвечающих им численных моделей, причем вместе с вычислительными экспериментами, отражающими их собственную разностороннюю исследовательскую деятельность. В работе по плазме моделируются дальнедействующие кулоновские силы. Работа по модели Р³М добавляет короткодействующие силы и открывает возможности для других приложений. Моделирование полупроводникового прибора добавляет целый ряд столкновительных эффектов. Астрофизика распространяет модель с силой $1/r^2$ на гравитирующие тела, при этом шаг по времени составляет миллионы лет, в противоположность пикосекундным шагам в плазменных моделях. Наконец, они приводят работу по фазовым переходам в твердых телах и жидкостях.

Во многих современных приложениях, таких, как очень большие и дорогостоящие разрабатываемые термоядерные реакторы,

теория и конструирование находятся на очень высоком уровне, но и тому, и другому чрезвычайно помогают в получении надежных результатов вычислительный эксперимент и в значительной степени исследования систем многих частиц. Книг, которые могли бы служить одновременно учебниками и справочниками, очень сильно недостает. Авторы поступили совершенно правильно, приведя избранные результаты, которые они сочли важными. Содержащиеся в книге модели и методы, несомненно, будут полезны во многих задачах, которые еще возникнут как в существующих сегодня, так и в будущих важных областях.

Беркли, апрель 1980 г.

Чарльз К. Бэрдсол

От авторов

Исследование различных явлений с помощью моделирования на ЭВМ стало в настоящее время признанным и быстро развивающимся направлением в науке и технике. Наиболее очевидные примеры можно найти в области физики, где вычислительные модели используют для изучения эволюции спиральных структур в галактиках, устойчивости горячей газовой плазмы в термоядерных установках и турбулентности в жидкостях. Приложения к химии включают в себя изучение расплавов солей, водных растворов и фазовых переходов. Применение тех же методов моделирования к исследованию динамики больших биологических молекул, таких как ферменты, может, по-видимому, в самом недалеком будущем оказать важное влияние на биологию. В технике умение моделировать поток электронов в полупроводниковом приборе, скажем, полевом транзисторе, может стать важным средством проектирования для инженера-электронщика; при этом по сравнению с традиционными лабораторными методами экономятся средства и время разработок.

В исследованиях, основанных на таких вычислительных экспериментах, решающую роль играет представление искомой системы в виде соответствующей математической модели. Эта математическая модель после дискретизации представляет в вычислительном эксперименте как аппаратуру, так и исследуемый материал. Методы дискретизации можно, вообще говоря, подразделить на три класса: конечно-разностные аппроксимации, методы конечных элементов и моделирование с помощью частиц. Существует обширная учебная литература по методам конечных разностей и конечных элементов, но очень мало руководств по моделям частиц.

В данной книге нас интересуют модели частиц, хотя некоторые материалы можно было бы с равным успехом отнести и к остальным двум классам. Модели частиц обычно, хотя и не всегда, используют для нестационарных задач, т. е. изучения эволюции во времени решения задачи Коши с граничными условиями. Эти модели обладают соответствующим дискретным видом, одинаково пригодным как для систем, корпускулярных по природе, так и описываемых в лагранжевой формулировке.

Мы различаем три типа моделей: частица-частица (чисто лагранжева), частица-сетка (эйлерово-лагранжева модель) и гибридная модель (частица-частица—частица-сетка). Каждая из них имеет свою область приложения. При правильном использовании модели частиц в состоянии продемонстрировать явные преимущества над другими методами, хотя справедливо и обратное утверждение, если они используются неправильно.

В этой книге мы хотели показать, где уместно прибегать к различным методам с использованием частиц и как такие методы применять эффективно. Приводимые сведения отражают те меры, которые приходится принимать, чтобы получать осмысленные численные результаты; именно, рассматриваются все этапы моделирования: физическое явление — математическая модель — вычислительный алгоритм — программа моделирования — вычислительный эксперимент. Вместо того, чтобы просто перечислять различные варианты, мы постарались показать (и хотелось бы надеяться, что это удалось), как на каждом из этих этапов можно выявить определенные количественные критерии и применить их для достижения наилучшего компромисса между затратами и качеством. Примеры применения методов частиц, которые мы приводим, также подобраны так, чтобы проиллюстрировать разнообразие приложений таких моделей, а не чтобы дать исчерпывающий обзор.

Мы усматриваем для читателя два возможных варианта работы с книгой. При первом, т. е. вводном, чтении можно опустить более подробные математические выкладки, представленные в главах с 4 по 8. При этом происходит знакомство с целями вычислительных экспериментов, дается общее введение в модели частиц, разработку и тестирование больших программ моделирования методом сверху вниз, приводится простой пример, поясняющий процесс проведения перечисленных выше этапов, и иллюстрируется использование моделей частиц. Материал по математике, содержащийся в гл. 4—8, спланирован так, чтобы дать наиболее подготовленному студенту информацию, которая позволит ему изучать, разрабатывать и эффективно использовать многомерные модели частиц.

Представляется, что такой вводный курс (возможно, с добавлением более сложных материалов в зависимости от подготовки конкретного студента) подходит в качестве учебника для старшекурсников по использованию ЭВМ в научных исследованиях в дополнение к курсам по обработке экономической информации. Материал повышенной сложности, хотя он на самом деле и нетрудный, требует умения свободно владеть преобразованием Фурье, и, по-видимому, в большей степени будет к месту в спецкурсе для соискателей магистерской степени по численному анализу, информатике или вычислительной физике. Кроме того, курс примерно из сорока лекций можно было бы

назвать «Методы моделирования частицами на сетках», и он мог бы составить естественное дополнение к традиционным курсам, например, по численному решению уравнений в частных производных или численным методам в динамике жидкости.

Думается, что в этом руководстве содержится достаточно исчерпывающее описание современного состояния искусства моделирования методом частица-сетка. В нем имеется полное обсуждение методов NGP и CIC, которые уже интенсивно используются с самого момента зарождения данного направления в 60-х годах, и кроме того, рассмотрены существенные усовершенствования этих методов, сделанные примерно в 1973 г. (именно, P³M-алгоритм), которые до сих пор широко не применялись. В силу этого мы считаем, что это руководство займет свое место в библиотеке профессиональных вычислителей, работающих в научно-исследовательских организациях и университетах.

Было бы невозможно назвать всех тех, кто прямо или косвенно помог в подготовке этой книги. Тем не менее, особую благодарность мы должны выразить Джону Даусону и его группе: Дж. Эфстатью, Дж. Гамильтону, Д. Браунриггу, Б. Берману, М. Амини и Р. Уоринеру, которые великодушно разрешили нам использовать свои результаты, в ряде случаев до их опубликования. Кроме того, многие прочитали рукопись, внесли поправки и предположения по ее улучшению, и среди всех нам бы хотелось упомянуть А. Аарсет, С. Бёрда, Ч. К. Бёрдсола, О. Бунемана, К. Моглестью, К. Робертса, Р. Робсона, Дж. Вокса, Б. Уилсона и О. Уинtringама. Слова самой сердечной признательности мы обращаем к Джилле Дикинсон, без преданности которой делу, аккуратности и поистине блестящего владения машинкой рукопись не смогла бы быть завершена в установленный срок. Мы просим всех тех, чью работу мы, возможно, по нечаянности неверно процитировали, исказили или неаккуратно упомянули, не судить нас слишком строго и дать нам знать с тем, чтобы эти погрешности можно было исправить во втором издании, если нам с ним повезет. По этой же причине нам бы хотелось получить сведения о любых замеченных ошибках и опечатках. Мы благодарны нашим коллегам из Ридингского университета за то понимание, с которым они отнеслись к нашим непрерывным отлучкам в период долгой и тяжелой работы по подготовке рукописи.

Ваш интерес к моделированию частицами безусловно стимулировал Оскар Бунеман, которому мы с благодарностью посвящаем эту книгу. Все стороны рассматриваемого предмета пронизаны его идеями. Эта книга представляет собой результаты работы авторов в течение последних восемнадцати лет, в основном в Стэнфордском и Ридингском университетах, исследовательском центре НАСА в Лэнгли, исследовательском центре

фирмы IBM в Йорктауне и в лаборатории Резерфорд/Атлас под Оксфордом. И нам бы хотелось выразить признательность нашим коллегам из этих организаций, в том числе особо сотрудникам вычислительных центров за вдохновляющую поддержку и терпеливую помощь. Кроме содействия указанных организаций и щедрого выделения машинного времени на самых быстродействующих компьютерах, определенная работа финансировалась многими ведомствами, включая Комиссию по атомной энергии США, ВВС США и Национальное управление США по авионавигации и исследованию космического пространства. В Великобритании эта работа поддерживается Советом по научным исследованиям через его комитеты и Каламской лабораторией по термоядерному синтезу Управления по атомной энергии в течение уже более десяти лет, за что мы особенно благодарны

Тесная связь с лабораторным экспериментом является важным условием успешной работы в вычислительной науке, и нам хочется адресовать нашу признательность сотрудникам фирмы Ройал Сигналс Энд Радар Истэблишмент (Малверн и Балдок) и Исследовательского центра Плесси Аллен Кларк (Касвелл), которые обеспечили нас результатами лабораторных исследований и постановками задач для большей части работы по полупроводниковым приборам. Кроме того, преподаватели и студенты пользовались финансовой поддержкой Министерства обороны и компании Плесси Лимитед.

Ридингский университет
Август 1979

*Р. В. Хокни,
Дж. В. Иствуд*

Вычислительные эксперименты с использованием моделей частиц

1.1. ВВЕДЕНИЕ

В этой книге мы будем иметь дело с областью научного исследования, в которой компьютер играет центральную роль. Рассматриваемые нами темы обычно называют «вычислительной физикой», хотя такое название не вполне точно, поскольку представленный материал перекрывает многие известные предметные области: физику, химию, технику, численный анализ и вычислительную технику — все они играют роль в вычислительной физике. В связи с этим мы отдаем предпочтение более общему выражению «вычислительная наука».

В научном исследовании основой вычислительной науки служит математическая модель интересующего нас физического явления. Уравнения математической модели переводят в дискретную алгебраическую форму, поддающуюся численному решению. Дискретные алгебраические уравнения описывают *вычислительную модель*, которая, если ее перевести в последовательность машинных команд, дает *программу моделирования* для вычислительной машины. После этого компьютер и программа позволяют исследовать эволюцию модельной физической системы в *вычислительных экспериментах*.

Численное моделирование можно рассматривать как теоретическое упражнение по численному решению краевой задачи Коши. В момент времени $t = 0$ задается начальное состояние системы в некоторой ограниченной области пространства (*расчетная область*), на поверхности которой поддерживаются заданные граничные условия. Моделирование состоит в прослеживании временной эволюции этой конфигурации. Основной частью вычисления является *цикл по временному шагу*, в котором состояние физической системы продвигается вперед по времени на малый шаг DT . Экспериментальный аспект (а отсюда и название «вычислительный эксперимент») возникает, когда мы рассматриваем проблему измерений. Даже простейший модельный расчет порождает большое количество данных, и для получения результатов в удобной для анализа форме требуется экспериментальный подход.

Хотя объем информации, который можно обработать с помощью компьютеров, велик, он тем не менее ограничен. Чтобы

получить хорошие вычислительные модели физических систем в условиях существующих ограниченных машинных ресурсов, ученые-вычислители проявляют чудеса изобретательности. Методы дискретизации, используемые при создании вычислительных моделей, включают в себя конечно-разностные методы [332], методы конечных элементов [357] и методы, с которыми мы будем иметь дело в этой книге — *методы частиц*.

Термин «*модели частиц*» является общим для класса вычислительных моделей, в которых дискретное описание физических явлений включает использование взаимодействующих частиц. Само название возникло потому, что в большинстве приложений этим частицам можно прямо сопоставить физические объекты. Каждая частица имеет набор атрибутов, таких как масса, заряд, завихренность, положение, импульс. Состояние физической системы определяется атрибутами конечного ансамбля частиц, а эволюция системы определяется законами взаимодействия этих частиц. Особенность, которая делает модели частиц привлекательными с вычислительной точки зрения, состоит в том, что ряд атрибутов частицы сохраняется, и поэтому их не надо изменять, когда вычислительная модель развивается во времени.

Связь между частицами вычислительной модели и частицами физической системы в значительной степени определяется соотношениями ограниченных вычислительных ресурсов и характерных пространственных и временных масштабов физических систем. С этой точки зрения в разд. 1.3 рассматриваются примеры физических систем, которые успешно исследовались с помощью моделей частиц. При всем разнообразии перечисляемых явлений их объединяет сходство математических моделей физических систем (разд. 1.4), а также сходство используемых численных схем (разд. 1.5).

1.2. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

В научном исследовании метод моделирования становится возможным благодаря быстродействующим ЭВМ. Однако сам факт возможности исследования с помощью моделирования не всегда означает, что это наилучший или даже предпочтительный метод. Как и при теоретическом и экспериментальном исследовании, мы должны выяснить, каковы ограничения и преимущества вычислительного эксперимента. Только в тех случаях, когда выясняется, что вычислительный эксперимент будет оправдан, нам нужно подробнее рассмотреть вопросы, касающиеся разработки вычислительной модели и проведения вычислительного эксперимента.

1.2.1. Роль вычислительного эксперимента

Для того чтобы получить представление о значении моделирования, рассмотрим традиционную картину развития науки и возможные пути его ускорения с помощью вычислительных экспериментов.

Экспериментальная (и наблюдательная) работа связана в основном с накоплением опытных данных. Теоретическая работа направлена главным образом на приведение такой информации в логически последовательную систему для того, чтобы вывести законы, предсказывающие поведение материи. В своем развитии как теория, так и эксперимент проверяют и обогащают друг друга. В непосредственном воздействии теории на эксперимент можно выделить три момента. Теория ведет к экспериментам с целью:

1. определить те факты, которые, согласно математическим выкладкам, основанным на некоей системе принципов, существенны для описания конкретных явлений;
- 2, проверить теоретические предсказания и предположения, и, наконец,
3. определить эмпирические величины, которых требует теория.

Точно так же и экспериментальная работа оказывает стимулирующее воздействие на теорию: теорию модифицируют в соответствии с экспериментом, ее используют для предсказания наиболее существенных опытных данных, а затем обобщают с тем, чтобы включить в единую систему возросший объем информации. Обычная практика науки состоит в том, чтобы во все более сложных ситуациях сопоставлять теорию с опытом. Аномалии и неопределенности оказываются в центре внимания как теории, так и эксперимента, а их разрешение приводит либо к обобщению существующих теорий, либо к пересмотру принципов с целью построения новых и более совершенных теорий, адекватно описывающих новые явления.

Потребности новейшей технологии послужили дополнительным стимулом как для теории, так и для эксперимента перейти к исследованию ситуаций, с которыми математика справиться не может, а прямые эксперименты неосуществимы. Чтобы преодолеть возникающие трудности, широко используются компьютеры, например, для численного вычисления математических функций и автоматизированного сбора и обработки данных. Эти разработки, которые составляют часть традиционного научного исследования, не являются здесь предметом нашего внимания. Мы будем рассматривать компьютер как новое средство проведения экспериментов — вычислительных экспериментов, которые ликвидируют разрыв между теорией и экспериментом.

Для моделирования поведения сложных систем в вычислительных экспериментах используются математические модели.

Компьютер совместно с программой выступают и как прибор, и как исследуемая система. Свойства этой системы в рамках ограничений, налагаемых конечным и дискретным представлением на компьютере, целиком определяются самим «вычислителем-экспериментатором». Прибор для вычислительного эксперимента обеспечивает полностью контролируемую среду, в которой даже измерение физических величин не вызывает никаких возмущений. Можно легко исследовать результаты действия основных физических законов в сложных ситуациях, в частности благодаря тому, что можно легко изменять совокупность законов, относящихся к таким ситуациям. Нелинейность, большое (но конечное) число степеней свободы и отсутствие симметрии не являются препятствиями для исследования с помощью моделирования.

Значение вычислительного эксперимента особенно велико в тех областях, где имеется большой разрыв между возможностями теории и эксперимента. При разработке теорий обычно опираются на линейность или возможность разложения в степенные ряды, высокую степень симметрии, небольшое количество параметров и упрощающие приближения. Эксперименты в лаборатории сталкиваются со всей сложностью природы: условия могут с трудом поддаваться контролю, измерения не всегда легко осуществить, и, как следствие, результаты зачастую трудно интерпретировать однозначно. Вычислительный эксперимент благодаря его универсальности становится связующим звеном между теорией и экспериментом.

Вычислительные эксперименты по аналогии с лабораторными можно грубо подразделить на три категории. К первой относятся эксперименты, предназначенные для точного предсказания работы сложных устройств с тем, чтобы иметь возможность оценить большое количество вариантов, прежде чем строить оптимальную модель с применением дорогостоящей технологии. Вторая группа вычислительных экспериментов включает в себя эксперименты, имеющие целью получение информации, труднодостижимой в лабораторном эксперименте — так обстоит дело при моделировании галактик или субмикронных электронных приборов, когда рассматриваются очень большие или очень малые пространственные и временные масштабы. И наконец, имеется группа вычислительных экспериментов, «теоретические эксперименты», которые дают ценные рекомендации для развития теории. Сложные ситуации, которые, например, встречаются при изучении плазмы, можно разбить на более простые компоненты, что помогает разобраться в обилии явлений и получить ясную теоретическую картину. Более того, эта ясная картина сможет теперь существовать не только в теории, поскольку полученный из вычислительных экспериментов мультфильм дает точную и наглядную картину эволюции системы во времени.

Объединение вычислительного эксперимента, физического эксперимента и теории оказывается гораздо более эффективным при получении физически полезных результатов, чем любой из этих методов по отдельности или сочетание любых двух из них. Для получения результатов теория использует математический анализ и численную оценку, физический эксперимент использует приборы и обработку данных, а вычислительный эксперимент использует компьютер и программу моделирования. Недостатки каждого метода исследования компенсируются достоинствами других методов. Роль вычислительного эксперимента определяется его сильными сторонами: возможностью дополнить теоретические исследования в случаях, когда важную роль играют нелинейность, большое число степеней свободы или отсутствие симметрии, а также дополнить экспериментальные исследования, когда приборы дороги, данные недоступны для непосредственного измерения или явления очень сложны.

1.2.2. Постановка вычислительного эксперимента

Отправным пунктом всех вычислительных экспериментов является некоторое физическое явление. Их цель состоит в получении физически полезных результатов. Между этими двумя пунктами можно выделить ряд этапов. Эти этапы объединены на рис. 1.1.

Каждый этап вносит свои ограничения. Как всегда, математическая формулировка представляет собой лишь приближенное описание физического явления. Ученый-вычислитель не может прямо принять уравнения в готовом виде, хотя разработкой математического описания физических явлений обычно занимаются теоретики. Он должен знать, какие упрощающие предложения сделаны, чтобы определить те случаи, когда справедливы эти уравнения, а следовательно, и его вычислительная модель.

Более строгие ограничения возникают на этапе перехода к дискретной алгебраической аппроксимации, на котором непрерывные дифференциальные или интегральные уравнения математической модели заменяются алгебраическими аппроксимациями, для того чтобы сделать возможным численное решение на компьютерах. На этом этапе возникают вопросы, касающиеся влияния конечного временного шага, дискретных пространственных сеток и, в случае моделей частиц, конечного числа частиц. Эти вопросы обсуждаются в общих чертах в гл. 4—8, а для конкретных приложений в гл. 9—12. На экспериментаторе-вычислителе всегда лежит обязанность доказать, что результаты, полученные с помощью численной модели, физичны — в последующих главах мы увидим, что численные ошибки могут приводить к катастрофическим последствиям!

Выбор способа дискретизации зависит от численного алгоритма. В результате дискретизации непрерывные переменные математической модели заменяются дискретным множеством точек, а дифференциальные уравнения — алгебраическими уравнениями. Если возникающие десятки (а иногда и сотни) тысяч алгебраических уравнений не могут быть решены на компьютере быстро, то предлагаемое моделирование становится невозможным.

Этапы, приводящие к численному алгоритму, включают процесс «самонастройки», в котором ищется наилучший компромисс между качеством представления и вычислительными затратами. Только после того, как он установлен, можно строить «прибор» — программу моделирования. Как и с настоящим прибором, необходимо уделить много внимания построению программы для того, чтобы гарантировать, что она функционирует желательным образом. Качественно разработанная программа должна быть удобочитаемой, несложной в использовании и в случае необходимости легко модифицироваться. Она должна иметь четкую модульную структуру и встроенную диагностику. Аналогию с лабораторной практикой можно продолжить: заранее спроектированная структура должна собираться только из проверенных программных компонентов. Полная программа должна быть испытана на известных задачах для проверки как алгоритма, так и кода. Только когда испытания и «калибровка» завершены, установку можно считать готовой для проведения экспериментов. Глава 3 посвящена описанию методик программирования, предназначенных для достижения этих целей.

1.3. ПРОСТРАНСТВЕННЫЕ И ВРЕМЕННЫЕ МАСШТАБЫ

Физические явления, которые адекватно описываются классической и квазиклассической теорией, можно промоделировать с помощью моделей частиц. В принципе любую классическую систему можно описать, зная положения и скорости частиц, которые ее составляют, и закон их взаимодействия. К сожалению, из-за неограниченного числа частиц, вовлеченных в совершенно простые ситуации, детальное описание становится вообще

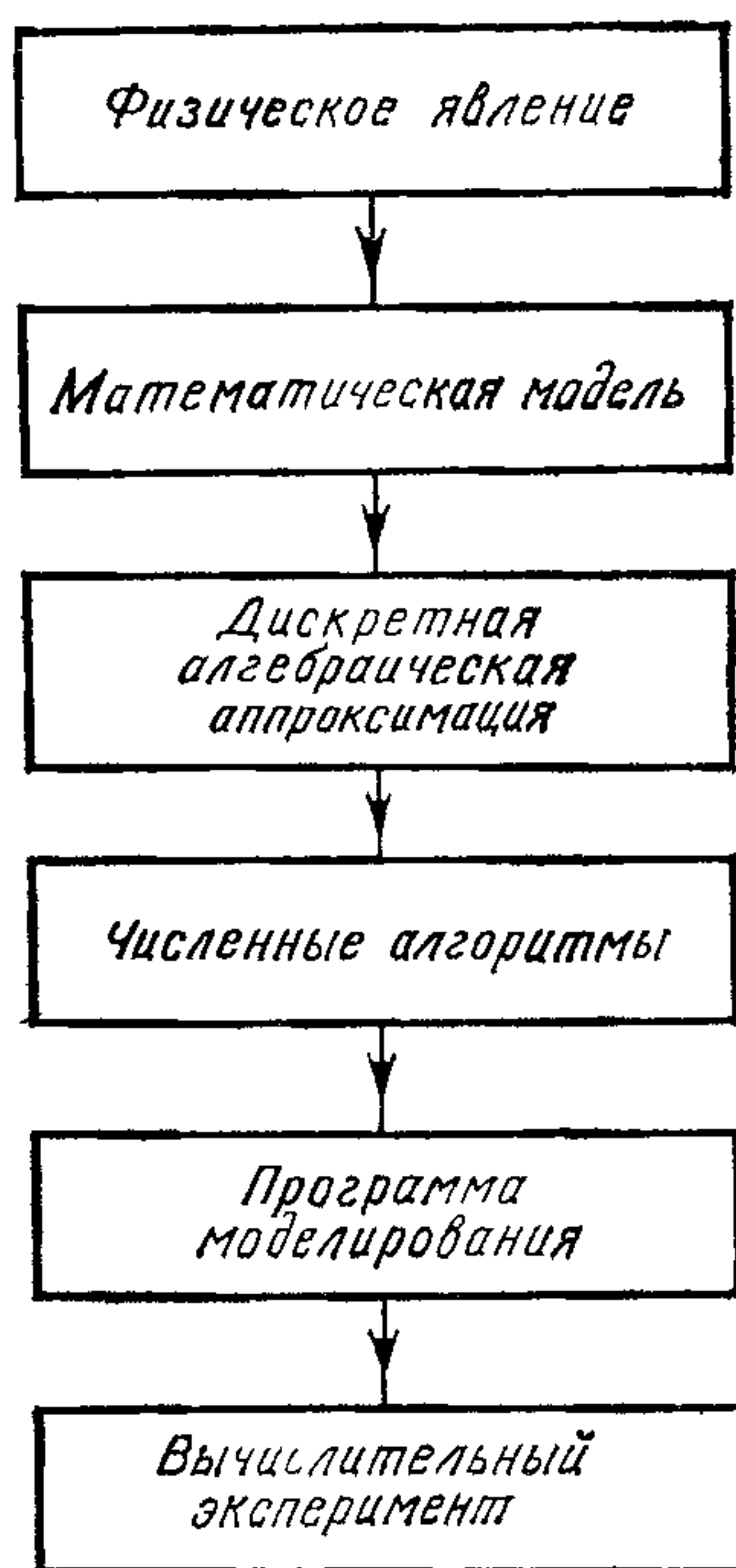


Рис 1.1 Основные этапы постановки вычислительного эксперимента

бесполезным как для понимания самой системы, так и для численного моделирования. Например, если бы мы исследовали течение воды через сопло, то детальная информация о положениях и скоростях каждого атома была бы малоинтересной. Действительно, если бы мы попытались выполнить такой расчет с помощью самого мощного из существующих компьютеров, то сумели бы закончить рассмотрение колебаний и вращений молекул только в микроскопическом элементе воды — течение имеет гораздо большие пространственные и временные масштабы, чем те, которые можно было надеяться промоделировать с помощью детальной модели.

Секрет успеха вычислительных экспериментов заключается в разработке модели, которая достаточно детализирована, чтобы точно воспроизвести важные физические эффекты, и все же не настолько подробна, чтобы сделать расчеты неосуществимыми. Наилучший выбор модели обусловлен характерными физическими масштабами длины и времени, примеры которых приведены в табл. 1.1 для самых разнообразных физических систем, которые были успешно промоделированы с помощью метода частиц. В случае течения воды через сопло модель, дающая среднемаховые скорости течения, может быть вполне подходящей (например, при моделировании жидкостей с помощью вихревых элементов). В другом предельном случае эксперименты по «молекулярной динамике» выполняются для изучения структуры жидкостей, в которых важны короткие масштабы длины и времени, соответствующие молекулярным движениям.

При моделировании молекулярной динамики имеется четкое взаимнооднозначное соответствие между физическими частицами и частицами вычислительной модели: каждый атом соответствует одной частице, и характеристики этой частицы такие же, как у атома (заряд, масса, положение, скорость). Аналогичные замечания справедливы для моделирования звездных скоплений и сгущения галактик, где звезды и галактики соответствуют частицам. Рассматриваемый масштаб длины есть среднее расстояние между частицами, а масштаб времени определяется периодом колебаний или орбитальным периодом соответственно. Расчеты сгущения звезд и галактик и молекулярной динамики подробно описываются в гл. 11 и 12 соответственно.

В другом предельном случае идентичность атомных строительных блоков (в вычислительной модели вихревой жидкости) полностью утрачивается. В математическом описании жидкость рассматривается как непрерывная, несжимаемая и вязкая среда. При дискретизации математической модели снова вводятся частицы (элементы жидкости), которые в данном случае характеризуются положением и завихренностью. Масштабы длины и времени определяются структурой и скоростью течения жидкости. Эти масштабы длины должны быть значительно больше,

чем атомные масштабы длины, для того чтобы такая модель имела смысл.

Третий тип модели частиц занимает промежуточное положение по отношению к приведенным выше примерам, т. е. моделью молекулярной динамики и вихревой моделью. В этих моделях «машинные» частицы в основном сохраняют идентичность атомным составным частям моделируемого вещества, однако атомы и частицы не соответствуют взаимнооднозначно друг другу, как в расчетах по молекулярной динамике. Моделирование разреженной плазмы, галактик и микроскопического полупроводникового прибора попадают в эту категорию.

При моделировании бесстолкновительной плазмы «электроны» и «ионы» соответствуют миллионам физических электронов и ионов. В математическом описании бесстолкновительной плазмы (уравнение Власова) ионы и электроны заменяются на непрерывную фазовую жидкость, так что вся информация о дискретности плазмы сглаживается. Если физические явления будут содержать длины волн, большие по сравнению со средними расстояниями между электронами и ионами, а временные масштабы окажутся малы по сравнению с временем, в течение которого дискретность проявится заметным образом (т. е. с временем столкновения), тогда это описание даст точное представление реальных процессов. Дискретизация гидродинамических уравнений Власова приводит к понятию счетных «макрочастиц». Дискретность, вносимая малым числом макрочастиц, минимизируется путем сглаживания сил на малых расстояниях. В одном из подходов макрочастицы рассматриваются как облака конечного размера, состоящие из электронов и ионов, причем положения и скорости макрочастиц соответствуют центрам масс и средним скоростям облаков. Другая интерпретация, которая проясняет связь с моделями жидких частиц, состоит в том, что макрочастицы считаются элементарными объемами несжимаемой фазовой жидкости, движущимися в фазовом пространстве (положение — скорость).

Пространственные и временные масштабы в бесстолкновительной модели плазмы определяются частотами и длинами волн собственных колебаний бесстолкновительной плазмы; таковыми являются плазменная частота и дебаевская длина соответственно. Данная модель справедлива до тех пор, пока дискретность системы макрочастиц не приведет к отклонению от поведения бесстолкновительной плазмы, — т. е. на временах, меньших времени столкновения. К счастью, сглаживание взаимодействия на малых расстояниях значительно увеличивает время столкновения в модели (см. разд. 9.2.1). В гл. 9 мы описываем свойства бесстолкновительной электростатической модели плазмы и ее использование при изучении диффузии плазмы поперек магнитного поля. Мы также представим модель,

Таблица 1.1. Примеры физических систем, представляемых моделями час
 Масштабы длины L даны в метрах времени T в секундах N_p — число час

Пример	Счетные частицы	Атрибуты частиц
1. Системы с корреляциями		
Ковалентные жидкости	Атомы или молекулы	Силовые константы, связанные с квантово-механическим дипольным и квадрупольным взаимодействиями, масса, сила, положение, скорость
Ионные жидкости	Положительные и отрицательные ионы	Заряд, масса, сила, положение, скорость, радиус
Звездные скопления	Звезды	Масса, положение, скорость, сила радиус
Скопления галактик	Галактики	Масса, положение, скорость, сила радиус
2. Бесстолкновительные системы		
Бесстолкновительная плазма	„Макрочастицы“ из 10^8 электронов или ионов	Заряд, масса, положение, скорость, радиус
Спиральные структуры галактик	„Макрочастицы“ из 10^6 звезд	Масса, положение, скорость, радиус
3. Столкновительные системы		
Субмикронные полупроводниковые приборы (микроскопическая модель Монте-Карло)	„Макрочастица“ = 10^4 волновых пакетов	Заряд, масса, положение, волновое число, радиус
4. Системы с сильными столкновениями		
Полупроводниковые приборы (диффузионная модель)	„Макрочастица“ = 10^4 электронов или дырок	Заряд, положение
Невязкая несжимаемая жидкость (вихри)	Вихревой элемент	Завихренность, положение

тиц.
тиц в кубе со стороной L .

Физические параметры			Параметры вычислительной модели		
N_p	L	T	N_p	L	T
10^5	10^{-8}	10^{-12}	10^3-10^4	$10^{-8}-10^{-9}$	10^{-12}
10^5	10^{-8}	10^{-12}	10^3-10^4	$10^{-8}-10^{-9}$	10^{-12}
10^2-10^3	10^{17}	10^{15}	10^2-10^3	10^{17}	10^{15}
$10^4 \div 10^5$	10^{23}	10^{17}	10^4-10^5	10^{23}	10^{17}
$10^9 \div 10^{12}$	$10^{-5} \div 10^{-2}$	$10^{-9} \div 10^{-12}$	$< 10^5$	$10^{-5} \div 10^{-2}$	$10^{-9} \div 10^{-12}$
$10^{10} \div 10^{11}$	10^{21}	10^{13}	$< 10^5$	10^{21}	10^{13}
10^8	10^{-7}	10^{-10}	$< 10^5$	10^{-7}	10^{-10}
10^9	10^{-6}	10^{-9}	$< 10^5$	10^{-6}	10^{-9}
Континуум	$10^{-3}-10^6$	$10^{-3}-10^5$	$< 10^5$	$10^{-3}-10^6$	$10^{-3}-10^5$

в которой частицы изображают малые элементы тока, и опишем ее использование для изучения взаимодействия плазмы солнечного ветра с геомагнитным полем с последующим образованием геомагнитного хвоста.

Моделирование галактик является еще одним примером моделирования бесстолкновительной системы. Макрочастицы в модели галактики представляют миллионы реальных звезд. Масштабы длины и времени определяются размерами и движением спиральных структур. Столкновительные эффекты опять уменьшены посредством сглаживания силы взаимодействия на малых расстояниях и включением в модель большого числа макрочастиц на интервале сглаживания. Важнейшее приложение модели бесстолкновительной галактики, а именно изучение спиральной структуры, описано в гл. 11.

Вычислительные модели полупроводников отличаются от моделей бесстолкновительной плазмы и галактики тем, что важные эффекты кристаллической решетки (рассеяние) с быстропротекающим временем учитываются путем изменения уравнений движения. Диффузионная модель полупроводникового прибора является наиболее подходящей, если размеры прибора велики по сравнению со средней длиной свободного пробега электрона. В этом случае частицы характеризуются зарядом и положением — при этом скорость определяется произведением электрического поля и подвижности. В субмикронных полупроводниковых приборах время, необходимое для достижения электроном стационарной скорости, соответствующей подвижности, становится сравнимым с временем жизни электрона в приборе. Следовательно, диффузионное приближение нужно заменить на микроскопическое описание, в котором используются статистические методы Монте-Карло для моделирования рассеяния электронов на кристаллической решетке, и должна учитываться инерция электронов. Тогда характеристиками частицы будут заряд, масса, положение и волновое число. Микроскопическая модель Монте-Карло полевого транзистора на арсениде галлия выбрана в качестве примера исследования и подробно описана в гл. 10.

1.4. ФИЗИЧЕСКИЕ СИСТЕМЫ

Иерархию физических систем и моделей частиц, перечисленных в табл. 1.1, полученную рассмотрением масштабов длины и времени, можно формально обосновать посредством применения математического аппарата классической статистической механики; эта статистическая связь между различными моделями описывается в общих чертах ниже в разд. 1.4.1—1.4.4. Хотя использованные там понятия статистической механики и полезны для понимания соотношения между различными моделями

частиц, знание деталей подобного анализа на протяжении этой книги не предполагается.

С точки зрения численного моделирования, единый подход, представляемый статистической механикой, проявляется как сходство математических моделей, и, следовательно, как сходство численных алгоритмов и вычислительных программ для большого разнообразия физических явлений. Все математические модели, к которым приложимы методы моделирования частицами, могут быть сформулированы как набор частиц, взаимодействующих через поля. Частицы имеют ряд сохраняющихся характеристик (например масса, заряд) и меняющихся характеристик (например положение, скорость). Переменные характеристики изменяются в соответствии с уравнением движения, а ответственные за движение члены этого уравнения определяются уравнениями для полей.

Уравнение Пуассона (1.7) является наиболее важным из уравнений для поля — и это относится ко всем примерам, перечисленным в табл. 1.1, за исключением первого. Эффективность моделирования частицами зависит от методов быстрого решения этого уравнения или сходного эллиптического дифференциального уравнения в частных производных, и именно поэтому мы посвящаем значительную часть гл. 6 методам его решения.

Уравнение Пуассона описывает скалярное (потенциальное) поле. Второе уравнение поля связывает векторное поле с этим скалярным полем. Для систем коррелированных, бесстолкновительных и со столкновениями векторное поле представляет собой поле силы (или ускорения), в то время как для систем с преобладанием столкновений векторное поле будет полем скоростей. Когда векторное поле является силовым полем, для определения скорости частиц интегрируется закон движения Ньютона $dv/dt \sim F$. Во всех случаях положение находится из скорости путем интегрирования уравнения $dx/dt = v$. Подходящие численные методы этого интегрирования по времени даны в гл. 4.

Естественным следствием сходства математических моделей становится сходство анализа погрешностей численных схем. Одно из преимуществ единого подхода к большому разнообразию физических явлений заключается в том, что достижения в численных методах в одной области приложений быстро проникают в другие области. Методы анализа погрешностей и конструирования алгоритмов, описанные в гл. 4—8, применимы к самым различным вычислительным моделям, кроме тех, которые мы рассматриваем в гл. 9—12.

Принципиальная польза от признания вычислительной науки вообще и моделей частиц в частности в качестве самостоятельных областей исследования появляется, когда мы рассматриваем машинную программу. Программы, разработанные для

изучения конкретных явлений, могут с небольшими модификациями быть использованы для изучения какого-либо совершенно другого явления. Например, одна и та же компьютерная программа может использоваться для изучения сгущения галактик (гл. 11) и ионных жидкостей (гл. 12). В числе других примеров представлено моделирование потока электронов в сильном магнитном поле и движение несжимаемой невязкой жидкости (разд. 14.4), а также бесстолкновительной плазмы (гл. 9) и спиральных структур в галактиках (гл. 11). Именно выявление сходства математических, численных и вычислительных задач и воплощение этих общих черт в стандартизованные структурированные программы позволяет достичь большой экономии человеческих и компьютерных ресурсов.

1.4.1. Системы с корреляциями

Первые три случая, перечисленные в табл. 1.1, в которых имеется взаимнооднозначное соответствие между физическими и счетными частицами, представляют примеры коррелированных систем. Вычислительные эксперименты с этими системами реализуют их статистикомеханическое описание грубым способом, проследив путь системы в sN -мерном фазовом пространстве Γ , где s — число степеней свободы каждой из N частиц, составляющих систему.

Вычислительные эксперименты как таковые являются детерминированными, а не статистическими. Начальное состояние определяется набором положений и скоростей $\{\mathbf{x}_i, \mathbf{v}_i; i = 1, N\}$ частиц в момент времени $t = 0$. Сила \mathbf{F}_i , действующая на частицу i , равна сумме сил, обусловленных остальными $N - 1$ частицами, и любых внешних сил, которые могут присутствовать:

$$\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij} + \mathbf{F}_i^{\text{ext}}. \quad (1.1)$$

Здесь \mathbf{F}_{ij} — сила, с которой частица j действует на частицу i . Хорошо знакомым примером служит закон Кулона, описывающий электростатическую силу взаимодействия двух заряженных частиц в виде

$$\mathbf{F}_{ij} = \frac{q_i q_j}{4\pi\epsilon_0} \frac{(\mathbf{x}_i - \mathbf{x}_j)}{|\mathbf{x}_i - \mathbf{x}_j|^3}, \quad (1.2)$$

где $\mathbf{x}_i, \mathbf{x}_j$ и q_i, q_j — положения и заряды частиц i и j соответственно. Потенциальная энергия Ψ_i частицы i может описываться аналогично в терминах попарно аддитивных потенциалов ψ_{ij} :

$$\Psi_i = \sum_{j \neq i} \psi_{ij} + \Psi_i^{\text{ext}}. \quad (1.3)$$

Для точечных зарядов ψ_{ij} имеют вид

$$\psi_{ij} = \frac{q_i q_j}{4\pi\epsilon_0} \frac{1}{|x_i - x_j|}. \quad (1.4)$$

Силы и потенциалы можно также рассматривать как поля, т. е. как величины, которые хотя и определяются характеристиками частиц, но сами частицами не являются. Поля силы и потенциальной энергии заполняют все пространство, в котором силовое поле равняется взятому с обратным знаком градиенту поля потенциальной энергии:

$$\mathbf{F} = -\nabla\psi. \quad (1.5)$$

Сила, действующая на частицу i , и ее потенциальная энергия даются соответственно выражениями

$$\mathbf{F}_i = \mathbf{F}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i}, \quad \Psi_i = \psi(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_i}. \quad (1.6)$$

Потенциальное поле связано с распределением источника посредством *уравнения поля*; например, уравнением поля, соответствующим (1.4), служит уравнение Пуассона

$$\nabla^2\phi = -\rho/\epsilon_0, \quad (1.7)$$

где $\Psi_i = q_i\phi|_{\mathbf{x}=\mathbf{x}_i}$, ϕ — электростатический потенциал, а ρ — плотность заряда.

Формулировки посредством *дальнодействия* (1.1)—(1.4) и *близкодействия* (1.5)—(1.7) эквивалентны. Для заданного распределения частиц они дают одни и те же силы. Заданные внешние силы и потенциалы в выражениях дальнодействия заменяются заданными граничными условиями в уравнении поля.

Описание физических систем завершается заданием граничных условий и уравнений движения. Граничные условия определяют внешние силы и объем пространства (*расчетную область*), в котором движутся частицы. Уравнения движения, например уравнение движения Ньютона,

$$\frac{d}{dt}\mathbf{x}_i = \mathbf{v}_i, \quad \frac{d}{dt}(m_i\mathbf{v}_i) = \mathbf{F}_i, \quad (1.8)$$

описывают динамику частиц.

Описание коррелированных систем на языке статистической механики, пригодное для проведения вычислительных экспериментов, основанных на моделях частиц, формально задается уравнением непрерывности Лиувилля для плотности вероятности ρ в пространстве Γ :

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^N \mathbf{v}_i \cdot \frac{\partial\rho}{\partial\mathbf{x}_i} + \sum_{i=1}^N \frac{\mathbf{F}_i}{m_i} \frac{\partial\rho}{\partial\mathbf{v}_i} = 0. \quad (1.9)$$

Здесь $\rho(\mathbf{x}_1 \dots \mathbf{x}_N, \mathbf{v}_1 \dots \mathbf{v}_N, t) d\mathbf{x}_1 \dots d\mathbf{x}_N d\mathbf{v}_1 \dots d\mathbf{v}_N$ есть вероятность того, что в момент времени t система находится в объеме $[(\mathbf{x}_1, \mathbf{x}_1 + d\mathbf{x}_1) \dots]$ пространства Γ .

Один вычислительный эксперимент соответствует ρ , имеющей вид δ -функции в пространстве Γ , причем точка сосредоточения этой δ -функции со временем перемещается по некоторой траектории в пространстве Γ . Теоретически, чтобы обеспечить один ансамбль для определения ρ , нужно выполнить серию экспериментов. На практике усреднение по времени обычно заменяет усреднение по ансамблю при вычислении ожидаемых значений для почти равновесной конфигурации, хотя в неравновесных расчетах важность начальных условий нельзя не учитывать.

С аналитической точки зрения с уравнением (1.9) справиться ничуть не легче, чем с полным набором траекторий частиц. Для достижения прогресса в этом направлении необходимо отбросить больше информации. Это достигается путем определения иерархии функций распределения и получения иерархии кинетических уравнений (цепочки БГКИ) посредством интегрирования (1.9) по последовательно меньшему числу координат частиц. Мы ограничимся кратким резюме и отсылаем читателей за дальнейшими подробностями к другим источникам¹⁾.

Состояние ансамбля из N -частиц в момент времени t можно определить с помощью точной одночастичной функции распределения:

$$F(\mathbf{x}, \mathbf{v}, t) = \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{x}_i(t)) \delta(\mathbf{v} - \mathbf{v}_i(t)), \quad (1.10)$$

где (\mathbf{x}, \mathbf{v}) теперь определяет точку в шестимерном фазовом пространстве (μ) для одной частицы. Редуцированное статистическое описание получается усреднением по ансамблю. Усредненная по ансамблю одночастичная функция распределения имеет вид

$$\begin{aligned} f_1(\mathbf{x}, \mathbf{v}, t) &= \langle F(\mathbf{x}, \mathbf{v}, t) \rangle = \int \rho F d\mathbf{x}_1 \dots d\mathbf{x}_N d\mathbf{v}_1 \dots d\mathbf{v}_N = \\ &= N \int \rho(\mathbf{x}, \mathbf{x}_2 \dots \mathbf{x}_N, \mathbf{v}, \mathbf{v}_2 \dots \mathbf{v}_N) d\mathbf{x}_2 \dots d\mathbf{x}_N d\mathbf{v}_2 \dots d\mathbf{v}_N. \end{aligned} \quad (1.11)$$

Второй интеграл в уравнении (1.11) получается из первого интегрированием δ -функций и переобозначением. Выражение $f_1(\mathbf{x}, \mathbf{v}, t) d\mathbf{x} d\mathbf{v}$ равно среднему числу частиц в объеме $d\mathbf{x} d\mathbf{v}$ в точке (\mathbf{x}, \mathbf{v}) пространства μ в момент времени t . Усредненная

¹⁾ Например, Монтгомери и Тидман [289, ч. I, гл. 6]; Исихара [216, гл. 12]; Клеммов и Доугерти [84, гл. 12]; Хансен и Макдональд [163].

Для ансамбля двухчастичная функция распределения определена аналогичным образом:

$$f_2(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t) = N(N-1) \int \rho(\mathbf{x}, \mathbf{x}', \mathbf{x}_3 \dots \mathbf{x}_N, \mathbf{v}, \mathbf{v}', \mathbf{v}_3 \dots \mathbf{v}_N) \times \\ \times d\mathbf{x}_3 \dots d\mathbf{x}_N d\mathbf{v}_3 \dots d\mathbf{v}_N; \quad (1.12)$$

при этом $f_2 d\mathbf{x} d\mathbf{v} d\mathbf{x}' d\mathbf{v}'$ интерпретируется как среднее произведение числа частиц в $d\mathbf{x} d\mathbf{v}$ и $d\mathbf{x}' d\mathbf{v}'$. Аналогично можно определить другие функции распределения $f_3 f_4$ и так далее, причем каждая последующая функция является более информативной.

Дифференциальное уравнение, описывающее эволюцию функции распределения f_s , получается из уравнения Лиувилля (1.9) посредством интегрирования по переменным, связанным с $N-1$ частицами. Уравнение для f_1 имеет вид

$$\frac{\partial}{\partial t} f_1 + \mathbf{v}_1 \cdot \frac{\partial f_1}{\partial \mathbf{x}_1} + \frac{\mathbf{F}_1^{\text{ext}}}{m_1} \cdot \frac{\partial f_1}{\partial \mathbf{v}_1} + \int \frac{\mathbf{F}_{12}}{m_1} \cdot \frac{\partial f_2}{\partial \mathbf{v}_1} d\mathbf{x}_2 d\mathbf{v}_2 = 0. \quad (1.13)$$

Уравнение для f_2 включает член, содержащий f_3 , а уравнение для f_s включает член с f_{s+1} . Для замыкания цепочки уравнений требуются некоторые приближения, и именно поэтому для исследования f_2 и f_3 было затрачено много усилий.

Если двухчастичную функцию распределения проинтегрировать по всем скоростям и выразить как функцию скалярного расстояния r между парой частиц, то получим парную или радиальную функцию распределения. Она широко используется для описания локальной структуры твердых тел и жидкостей и является обычным конечным результатом численного моделирования таких коррелированных систем. Изменение радиальной функции распределения в процессе плавления электронной пленки (разд. 12.2.5 и рис. 12.4) и при плавлении и стекловании в микрокристалле KCl (разд. 12.3.5 и 12.3.6 и рис. 12.9) дано в гл. 12. При исследовании сгущения галактик во Вселенной, однако, необходимо также измерять трехчастичную и четырехчастичную корреляционные функции (см. гл. 11, разд. 4).

1.4.2. Системы без корреляций (бесстолкновительные)

Замыкание наинизшего порядка иерархии кинетических уравнений приводит к уравнению Власова. В этом приближении частицы предполагаются нескоррелированными, что позволяет записать двухчастичную функцию распределения в виде произведения одночастичных функций распределения:

$$f_2(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t) = f(\mathbf{x}, \mathbf{v}, t) f_1(\mathbf{x}', \mathbf{v}', t). \quad (1.14)$$

Условия, при которых аппроксимация (1.14) является подходящей, можно наглядно получить из простых физических сообра-

жений. Выражение $f_2 d\tau d\tau'$ можно трактовать как вероятность одновременного обнаружения частицы в объеме $d\tau = dx dv$ около точки (x, v) и в объеме $d\tau' = dx' dv'$ около точки (x', v') в пространстве μ . Вероятность обнаружения частицы в объеме $d\tau'$ будет существенно зависеть от наличия частицы в $d\tau$, только если потенциальная энергия взаимодействий частиц велика по сравнению с их кинетической энергией. Приравнявая потенциальную энергию взаимодействия к типичной кинетической энергии, находим радиус взаимодействий. Для плазмы и галактик радиус a получается равным соответственно

$$a_p = \frac{e^2}{4\pi\epsilon_0 k_B T}, \quad a_g = \frac{Gm^2}{\frac{1}{2}mv^2}, \quad (1.15)$$

где e — заряд, k_B — постоянная Больцмана, ϵ_0 — диэлектрическая проницаемость вакуума, G — гравитационная постоянная, m — масса, v — скорость. Ясно, что если a мало по сравнению со средним расстоянием между частицами, то статистическая независимость нахождения частиц в различных объемах пространства μ , подразумеваемая в (1.14), будет вполне хорошим приближением. Если мы обозначим через n плотность частиц (электронов) в плазме, а через σ поверхностную плотность звезд в галактическом диске, тогда мерой качества бесстолкновительного приближения (1.14) служат величины

$$\epsilon_p = a_p n^{1/3}, \quad \epsilon_g = a_g \sigma^{1/2}. \quad (1.16)$$

Типичные скорости v во вращающихся гравитирующих дисках можно оценить, приравнявая центростремительное и гравитационное ускорения на краю однородного гравитирующего диска радиуса R , что дает $a_g \sim (\sigma R)^{-1}$. Подстановка выражений для a_p и a_g в (1.16) дает

$$\epsilon_p \sim (n\lambda_D^3)^{-2/3}, \quad \epsilon_g \sim (\sigma R^2)^{-1/2}, \quad (1.17)$$

где $\lambda_D = (\epsilon_0 k_B T / ne^2)^{1/2}$ — дебаевская длина. Следовательно, для справедливости формулы (1.14) требуется, чтобы число электронов в дебаевском кубе $n\lambda_D^3$ или число звезд в галактическом диске σR^2 было велико: в типичном случае $n\lambda_D^3 \sim 10^9$ и $\sigma R^2 \sim 10^{10}$.

Подставляя выражение (1.14) в уравнение (1.13) и опуская индекс «1», получаем уравнение Власова (дальнейшие подробности см., например, в [84]):

$$\frac{\partial f}{\partial t} \equiv \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \left\{ \frac{\mathbf{F}^{\text{ext}}}{m} + \int \frac{\mathbf{F}}{m} f(\mathbf{x}', \mathbf{v}', t) d\mathbf{x}' d\mathbf{v}' \right\} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \quad (1.18)$$

Интегральный член в (1.18) описывает самосогласованное ускорение электронов в плазме или звезд в галактическом диске.

Например, в случае плазмы, используя выражение (1.2) для \mathbf{F} , получим

$$\int \frac{q^2}{4\pi\epsilon_0 m} \frac{(\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3} f(\mathbf{x}', \mathbf{v}', t) d\mathbf{x}' d\mathbf{v}' = \frac{q}{m} \mathbf{E} = -\frac{q}{m} \nabla\phi, \quad (1.19)$$

или, что эквивалентно,

$$\nabla^2\phi = -\frac{q}{\epsilon_0} \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \quad (1.20)$$

где \mathbf{E} — электрическое поле, а ϕ — электростатический потенциал, возникающий из усредненного по ансамблю распределения плотности заряда $\rho = q \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}$ (здесь подразумевается суммирование по электронной и ионной компонентам). Уравнение (1.18) является бесстолкновительным в том смысле, что взаимодействия отдельных частиц уже нет; это выглядит так, как будто бы точки в пространстве μ , представляющие физические частицы, размазались с образованием непрерывной «фазовой» жидкости в пространстве μ .

В вычислительных моделях бесстолкновительной плазмы и галактики при дискретизации непрерывной фазовой жидкости вновь появляются частицы, хотя, как отмечалось в предыдущем разделе, в вычислительной модели во много раз меньше частиц, чем в физической системе. Приписывание этим счетным частицам конечного размера W сильно уменьшает силу взаимодействия частиц, расположенных на расстоянии, меньшем чем W . В этом случае, если радиус a существенно меньше W , вычислительная модель с хорошим приближением будет бесстолкновительной, поскольку число макрочастиц в макрочастице (nW^3 или σW^2) достаточно велико (см. разд. 9.2.1).

1.4.3. Системы со столкновениями

Явления, длины волн которых велики по сравнению со средним расстоянием между физическими частицами, но которые в то же время не описываются надлежащим образом в бесстолкновительном приближении, также поддаются моделированию с помощью моделей частиц. Требуемое математическое описание для таких явлений получается добавлением к приближению, отраженному в (1.14), небольшого корректирующего члена — корреляционной функции g :

$$f_2(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t) = f_1(\mathbf{x}, \mathbf{v}, t) f_1(\mathbf{x}', \mathbf{v}', t) + g(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t). \quad (1.21)$$

В этом случае уравнение (1.18) заменяется на

$$\frac{df}{dt} = - \int \frac{\mathbf{F}}{m} \cdot \frac{\partial}{\partial \mathbf{v}} g d\mathbf{x}' d\mathbf{v}' = \left(\frac{\partial f}{\partial t} \right)_c. \quad (1.22)$$

Уравнение (1.22) известно как уравнение Больцмана. Левая часть совпадает с левой частью уравнения Власова, в то время

как правая часть, формально представленная как скорость изменения f за счет столкновений, описывает влияние зернистости среды.

Было предложено несколько моделей плазмы со столкновениями, использующих для представления f счетные макрочастицы. Столкновительный член в (1.22) обычно представляют с помощью обмена импульсом и энергией [301] или с помощью обмена числом электронов, приходящимся на макрочастицу, и энергией между соседними макрочастицами [155]. Модель субмикронного полупроводникового прибора, указанная в табл. 1.1, служит еще одним примером модели столкновительной плазмы, в которой в данном случае плазма представляет собой газ электронов проводимости, движущихся в полупроводнике, а столкновения включают квантовомеханические взаимодействия.

Для рассматриваемых пространственных масштабов классическое уравнение Больцмана (1.22) хорошо описывает электронную проводимость в полупроводниках в том случае, когда влияние энергетических зон учитывается введением эффективной массы $m^* = \hbar^2 / (\partial^2 \mathcal{E} / \partial k^2)$, где $\mathcal{E}(\mathbf{k})$ — структура энергетической зоны, характерная для данного материала (см., например, [106]). Если энергетическая зона аппроксимируется параболой (разд. 10.2.2), то квантовомеханический волновой вектор \mathbf{k} и энергия \mathcal{E} , которые нужны для описания процессов рассеяния, связаны со скоростями частицы соотношениями

$$m^* \mathbf{v} = \hbar \mathbf{k}, \quad \mathcal{E} = \frac{\hbar^2 k^2}{2m^*}. \quad (1.23)$$

1.4.4. Системы с сильными столкновениями

Модели, основанные на уравнении Больцмана (1.22), требуют слишком больших вычислительных затрат, если расчет ведется для времени, превышающего несколько времен столкновений. При желании исследовать явления на больших временных (и пространственных) масштабах необходимо отказаться от значительной части информации. Это сокращение достигается путем перехода к импульсам в уравнении Больцмана. Формализм шестимерного пространства μ сводится к формализму трехмерного конфигурационного пространства (\mathbf{x}) . Детальное описание распределения по скоростям заменяется на среднюю скорость и ее разброс (температуру). Как и в иерархии кинетических уравнений, каждое моментное уравнение имеет член, включающий в себя более высокие моменты, поэтому для получения замкнутой системы уравнений необходимо дальнейшее приближение. Обычно для замыкания системы принимают определенное уравнение состояния для давления p : $p = p(\rho, T)$, где ρ — плотность и T — температура.

Уравнение для нулевого момента, описывающее сохранение массы, получается умножением (1.22) на массу частицы и интегрированием по пространству скоростей:

$$\int d\mathbf{v} m \left\{ \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} \right\} = \int d\mathbf{v} m \left(\frac{\partial f}{\partial t} \right)_c. \quad (1.24)$$

Из определения f следует, что

$$\int d\mathbf{v} m f = mn = \rho, \quad \int d\mathbf{v} m \mathbf{v} f = \rho \mathbf{u}, \quad (1.25)$$

где ρ — плотность и \mathbf{u} — средняя (гидродинамическая) скорость. Если число частиц сохраняется, правая часть уравнения (1.24) равна нулю. Используя (1.25) и необходимое условие, что $f \rightarrow 0$ при $\mathbf{v} \rightarrow \infty$, левую часть можно проинтегрировать; получим

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{u} = 0. \quad (1.26)$$

Умножение (1.22) на $m\mathbf{v}$ и интегрирование по пространству скоростей дает первое моментное уравнение (сохранение импульса):

$$\frac{\partial}{\partial t} \rho \mathbf{u} + \nabla [\rho \mathbf{u} \mathbf{u} + \mathbf{P}] - n\mathbf{F} = \int d\mathbf{v} m \mathbf{v} \left(\frac{\partial f}{\partial t} \right)_c, \quad (1.27)$$

где тензор давления \mathbf{P} определяется выражением

$$\mathbf{P} = \int d\mathbf{v} m f (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}). \quad (1.28)$$

Во многих приложениях тензор давления неплохо аппроксимируется скалярным давлением, $\mathbf{P} \approx p\mathbf{I}$, где p подчиняется уравнению идеального газа $p = \rho RT = nk_B T$. Использование этого приближения в сочетании с (1.26) и (1.27) приводит к уравнению движения, которое лежит в основе сильностолкновительных моделей частиц:

$$\rho \frac{d\mathbf{u}}{dt} = \rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla p + n\mathbf{F} + \int d\mathbf{v} m \mathbf{v} \left(\frac{\partial f}{\partial t} \right)_c. \quad (1.29)$$

Электроны (или соответственно дырки) в модели полупроводника с преобладанием столкновений рассматриваются как заряженная жидкость, при этом модельные частицы соответствуют элементам этой жидкости. Частицы характеризуются зарядом, эффективной массой и положением в пространстве. Темп изменения их положений определяется скоростью, а сама скорость, полученная из (1.29), рассматривается как полевая величина. На временах, больших времени электронно-решеточной релаксации τ , столкновительный член в уравнении (1.29) можно аппроксимировать простой моделью

$$\int d\mathbf{v} m^* \mathbf{v} \left(\frac{\partial f}{\partial t} \right)_c \simeq \int d\mathbf{v} m^* \mathbf{v} \frac{f_0 - f}{\tau} = -\frac{m^* n \mathbf{u}}{\tau}, \quad (1.30)$$

а инерционным членом $\rho du/dt$ можно пренебречь. Равновесное распределение (Ферми) f_0 симметрично по v и поэтому не дает никакого вклада в интеграл в (1.30). В результате при подстановке выражений для ρ и \mathbf{F} уравнение (1.29) принимает вид

$$0 = -k_B T \nabla n + nqE - \frac{m^* n u}{\tau}$$

где

$$\mathbf{u} = -\frac{D}{n} \nabla n + \mu E, \quad (1.31)$$

где D — коэффициент диффузии и μ — подвижность (см. [106, гл. 13]).

Математическая модель для моделирования вихревой гидродинамики с преобладанием столкновений также получается из (1.26) и (1.29). Если эти уравнения просуммировать по всем сортам частиц, образующих жидкость, то все столкновительные члены в (1.29) для каждого сорта, которые определяют скорость изменения импульса в единице объема за счет столкновений с другими сортами, дадут в результате суммирования нуль. При дополнительных предположениях о несжимаемости ($\rho = \text{const}$) и отсутствии электромагнитных и гравитационных массовых сил ($\mathbf{F} = 0$) получаем гидродинамическое описание

$$\nabla \cdot \mathbf{u} = 0, \quad \rho \frac{d\mathbf{u}}{dt} = -\nabla p. \quad (1.32)$$

Для двумерных течений (скажем, в плоскости $x-y$), введем функцию тока ψ и вектор завихренности ω , которые оба имеют компоненты только в направлении z . Эти переменные связаны с полем скорости \mathbf{u} соотношениями

$$\mathbf{u} = \text{rot } \psi, \quad (1.33)$$

$$\omega = \text{rot } \mathbf{u}. \quad (1.34)$$

Завихренность частицы — элемента жидкости — сохраняется, в чем можно убедиться при выполнении операции rot от уравнения движения (1.32).

Уравнения движения частиц в вихревой модели просто $dx/dt = \mathbf{u}$. Расписывая по компонентам (1.33), имеем

$$u_x = \frac{\partial \psi_z}{\partial y}, \quad u_y = -\frac{\partial \psi_z}{\partial x}, \quad (1.35)$$

и подставляя в (1.34), получаем уравнение поля

$$\nabla^2 \psi_z = -\omega_z. \quad (1.36)$$

Модель, математически подобная модели вихревой жидкости, получается для движения ведущего центра электронов перпендикулярно сильному магнитному полю. Уравнения ведущего центра получаются разделением масштаба времени быстрого

вращения по ларморовской орбите вокруг силовой линии магнитного поля и медленных дрейфовых движений. Математически это разделение выполняется с помощью разложения по малому параметру $\varepsilon = m_e/e$. В отсутствие столкновений уравнение (1.29) в наинизшем приближении ($\varepsilon = 0$) переходит в

$$\mathbf{F} = \mathbf{E} + \mathbf{u} \times \mathbf{B} = 0. \quad (1.37)$$

Предполагая однородность вдоль магнитного поля, выбирая поле в виде $\mathbf{B} = (0, 0, B_z)$ и записывая $\mathbf{E} = -\nabla\varphi$, получаем

$$u_x = -\frac{1}{B_z} \frac{\partial\varphi}{\partial y}, \quad u_y = \frac{1}{B_z} \frac{\partial\varphi}{\partial x}. \quad (1.38)$$

Потенциал и плотность заряда ρ связаны уравнением Пуассона

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0}, \quad (1.39)$$

причем

$$\rho = -e(n - n_0), \quad (1.40)$$

где e — абсолютная величина заряда электрона,
 n — плотность ведущих центров электронов,
 n_0 — плотность фона.

Сравнивая (1.35) и (1.36) с (1.38) и (1.39), мы видим, что электростатический потенциал φ играет ту же роль, что и функция тока ψ_z , а плотность заряда играет роль плотности завихренности ω_z . Следовательно, при соответствующем масштабировании одна и та же модель применяется к несжимаемой вязкой жидкой среде и электронным потокам в сильных магнитных полях!

Вероятно, самым первым моделированием завихренности частицами было моделирование, которое выполнили Абернаци и Кронауэр [9], используя метод РР. Успешно наблюдалось образование «вихревых дорожек». Двумерная модель РМ использовалась Леви и Хокни [251] в аналогичной задаче, связанной с электронным потоком. Сходная гидродинамическая модель была описана Христиансенем [77, 78] и широко использовалась при изучении взаимодействия между ограниченными в пространстве вихревыми структурами [81].

Другие случаи явлений, описываемых гидродинамическими уравнениями, также поддаются моделированию частицами. Сжимаемую жидкость можно рассматривать с помощью метода дробных шагов, чтобы решить уравнение движения: вызванные градиентом давления изменения в поле скорости вычисляются при фиксированном положении частиц, а затем элементы жидкости (частицы) перемещаются, перенося массу и импульс [164]. Жидкие среды в магнитном поле можно рассматривать

аналогично, причем в этом случае частицы имеют характеристики, связанные с магнитным потоком, в дополнение к положению, массе, импульсу и энтропии [278].

1.5. МОДЕЛИ ЧАСТИЦ

Мы выделяем три основных типа вычислительной модели частиц: модель частица-частица (РР), модель частица-сетка (РМ) и модель частица-частица — частица-сетка (РРМ или Р³М). В модели РР используется формулировка закона силы дальнего действия; в модели РМ сила рассматривается как полевая величина и аппроксимируется на сетке; модель Р³М является гибридом моделей РР и РМ. Выбор модели диктуется частично физикой изучаемого явления и частично соображениями вычислительных затрат.

1.5.1. Метод частица-частица

Метод РР является простейшим с понятийной и вычислительной точки зрения. В некоторый момент времени t состояние физической системы описывается набором положений и скоростей частиц $\{x_i(t), v_i(t); i = 1, N_p\}$. Цикл временного шага пересчитывает эти величины, используя силы взаимодействия и уравнения движения для получения состояния системы в более поздний момент времени $t + DT$ следующим образом:

Цикл шага по времени метода РР

1. Вычислить силы

Обнулить накопители сил

```
[ for i = 1 to Np do
  Fi := 0
```

Накопить силы

```
[ for i = 1 to Np - 1 do
  [ for j = i + 1 to Np do
    Найти силу Fij, с которой частица j действует на частицу i
    Fi := Fi + Fij
    Fj := Fj - Fij
```

2. Проинтегрировать уравнения движения

```
[ for i = 1 to Np do
  vinew := viold +  $\frac{F_i}{m_i} DT$ 
  xinew := xiold + vi DT
```

3. Изменить значение счетчика времени

$$t := t + DT$$

Для слежения за развитием системы во времени используется многократное повторение этого цикла.

Программирование цикла временного шага в методе РР является тривиальным упражнением. Если же рассмотреть вычислительные затраты, становится понятным, почему такая простая схема используется только в небольшом числе случаев. Возьмем, например, систему точечных заряженных частиц, в которой выражение для далекодействующих сил определено уравнением (1.2). Положим, что каждая операция $+$, $-$, \div и \times считается одной операцией с плавающей точкой и $()^{3/2}$ считается тремя операциями. Умножения на $m_i(q_i q_j / 4\pi\epsilon_0)$ и DT можно избежать тщательным подбором единиц измерения (см. гл. 2), и поэтому мы эти операции учитывать не будем. Расписывая цикл шага по времени, мы получаем следующее количество затраченных операций:

<i>Вычисления</i>	<i>Количество операций</i>
Обнуление накопителей силы выполняем для $N_p(N_p - 1)/2$ пар частиц	$3 N_p$
вычисление $x_i - x_j$	3
вычисление $ x_i - x_j ^3$	8
вычисление F_{ij}	3
изменение F_i и F_j	6
	<hr style="width: 50%; margin: 0 auto;"/>
	20
	<hr style="width: 50%; margin: 0 auto;"/>
изменение скорости и положения	$6 N_p$
	<hr style="width: 50%; margin: 0 auto;"/>
Итого	<u>$10 N_p^2 - N_p$</u>

В табл. 1.2 для различного числа частиц представлены оценки времени центрального процессора (ЦП), необходимого для расчета одного шага по времени в модели РР. Типичными для британских университетов являются компьютеры средней мощности с возможностью выхода на большие компьютеры научных расчетов в региональных вычислительных центрах. В некоторых научных учреждениях национального масштаба установлены высокопроизводительные векторные процессоры для параллельной обработки. Времена выполнения операции с плавающей точкой являются условными и достаточно умеренно учитывают другие необходимые вспомогательные операции. Их можно взять в качестве разумной основы для заниженной оцен-

ки жизнеспособности каждого предполагаемого расчета. Несомненно, очень тщательно запрограммированный код будет соответствовать меньшим условным временам, чем те, которые указаны в табл. 1.2. Эти времена относятся к компьютерам примерно 1978 г.

Таблица 1.2. Оценки затрат машинного времени на один временной шаг в модели РР. Типичное номинальное время выполнения операции с плавающей точкой включает затраты времени на служебные операции и относится к 1978 г.; N_p — число частиц.

Тип компьютера	Номинальное время на 1 операцию	Время работы ЦП ($= 10 N_p^2 \times 1 \text{ оп}$)		
		$N_p = 10^2$	$N_p = 10^3$	$N_p = 10^5$
Мини-ЭВМ	100 мкс	10 с	0.25 ч	~ 1 год
ЭВМ среднего класса	10 мкс	1 с	1.5 мин	~ 1 месяц
Большая ЭВМ для научных расчетов	1 мкс	0.1 с	10 с	~ 1 сутки
Большой векторный процессор	0.1 мкс	0.01 с	1 с	~ 1 ч

Как правило, для получения полезных результатов из вычислительного эксперимента требуется несколько тысяч шагов по времени, поэтому метод прямого суммирования РР целесообразен только для систем, содержащих приблизительно до тысячи частиц, если силы дальнодействующие, как, например, при изучении звездных скоплений. Однако если силы взаимодействия короткодействующие, то количество операций будет пропорционально не квадрату числа частиц, а числу частиц, умноженному на число соседних, достаточно близких к выделенной частице, чтобы значительно повлиять на действующую на нее силу. Модель атомных жидкостей Леннарда-Джонса служит примером системы с короткодействующими силами, для которой широко применялась модель РР.

1.5.2. Метод частица — сетка

В методе частица-сетка используется формализм близкоддействия и уравнение поля для потенциала [например, (1.5) и (1.7) соответственно]. В результате сила вычисляется гораздо быстрее, но обычно менее точно, чем при использовании метода РР.

Полевые величины, которые заполняют все пространство физической системы, приближенно представляются значениями в регулярно расположенных узлах сетки. Дифференциальные операторы, такие как лапсиан ∇^2 , заменяются конечно-разностными аппроксимациями на этой сетке. Потенциалы и силы в

месте положения частицы вычисляются посредством интерполяции по массиву сеточных значений. Сеточные плотности рассчитываются с помощью обратной процедуры раздачи характеристик частицы (например, заряда) в ближайшие узлы сетки для того, чтобы получить сеточные значения (например, плотность заряда).

Цикл шага по времени в методе РМ отличается от цикла метода РР только способом вычисления сил. Расчет силы в методе РМ, соответствующий приведенному выше примеру метода РР с заряженными частицами, состоит из трех этапов:

1. Раздача заряда на сетку.
2. Решение на этой сетке уравнения Пуассона (1.7).
3. Вычисление сил по сеточному потенциалу и интерполяция сил в точке расположения частиц.

Количество операций на этапах (1) и (3) пропорционально числу частиц N_p . Количество операций на этапе (2) зависит от числа узлов сетки N . Следовательно, полное количество операций определяется выражением

$$\text{число операций} = \alpha N_p + \beta(N),$$

где постоянная α и функция β зависят от конкретного вида используемой схемы РМ. Исследование этих различных видов схем будет продолжено в гл. 5. А пока, чтобы провести сравнение с методом РР, мы возьмем некоторые характерные значения:

$$\alpha = 20, \quad \beta = 5N^3 \log_2 N^3 \text{ для сетки } N \times N \times N.$$

Считая, что $N = 32$, $N_p = 10^5$ и что номинальное время центрального процессора (ЦП), затраченное на выполнение одной операции с плавающей точкой, равно 1 мкс, получаем

$$\text{время ЦП} = (20 \times 10^5 + 5 \times 32^3 \times 15) \times 10^{-6} \text{ с} \approx 4.5 \text{ с}$$

по сравнению ~ 1 сутки для метода РР!

Огромный выигрыш в скорости метода РМ по сравнению с методом РР достигается ценой потери разрешения в поле потенциала и силы. Только те поля, вариации которых имеют длину волны, большую, чем шаг пространственной сетки, можно точно представить с помощью сеточных значений. Поля потенциала и силы одиночного точечного заряда (или массы) на расстояниях, меньших шага сетки H , представляются неточно. Однако характер этих погрешностей таков, что неточное представление полей точечного заряда можно интерпретировать как точное представление полей от заряженного облака конечного размера, ширина которого порядка шага сетки H .

Ограниченное разрешение сетки оказывается неприемлемым для изучения коррелированных систем: сетка, достаточно мелкая для разрешения близких столкновений частиц, имела бы

настолько большое количество сеточных узлов, что время ЦП, необходимое для нахождения потенциала, превышало бы время исполнения цикла временного шага для метода РР. Для некоррелированных систем и систем со столкновениями (в том числе с сильными столкновениями) ситуация обратная; ограниченное сеточное разрешение приносит свойства, необходимые для подавления нефизических корреляций между конечным числом счетных макрочастиц. При условии что сеточный шаг меньше, чем характерная длина волны физической системы (например, дебаевская длина в плазме и ширина спирального рукава галактик) и что число счетных макрочастиц в каждой ячейке сетки достаточно велико (скажем, ~ 10) для поддержания низких уровней флуктуаций, модель РМ будет точно представлять математические модели некоррелированных физических систем и систем со столкновениями.

1.5.3. Метод частица-частица — частица-сетка

Мы выяснили, что метод РР может быть использован для небольших систем с далекодействующими силами или для больших систем, в которых силы взаимодействия отличны от нуля только на больших расстояниях между частицами. С другой стороны, метод РМ является быстрым в вычислительном отношении, но он пригоден только для гладко меняющихся сил. Метод Р³М сочетает достоинства методов РР и РМ и позволяет моделировать большие коррелированные системы с далекодействующими силами.

Прием, использованный в методе Р³М, состоит в расщеплении действующих между частицами сил на две части,

$$\mathbf{F}_{ij} = \mathbf{F}_{ij}^{sr} + \mathbf{F}_{ij}^m, \quad (1.41)$$

где быстроменяющаяся короткодействующая часть \mathbf{F}_{ij}^{sr} отлична от нуля только на нескольких межчастичных расстояниях, а медленно меняющаяся часть \mathbf{F}_{ij}^m достаточно гладкая для точного представления на сетке. Метод РР используется для нахождения суммарной короткодействующей силы, действующей на каждую частицу, метод РМ используется для нахождения суммарной медленно меняющейся составляющей силы. Сложение суммарной короткодействующей и суммарной медленно меняющейся составляющей силы дает результирующую силу, действующую на каждую частицу; эта результирующая используется для пересчета скоростей. Полученная в результате схема Р³М занимает промежуточное положение в том смысле, что можно представить близкие столкновения с той же точностью, как в методе РР, и вычислять далекодействующие силы с тем же быстродействием, как в методе РМ.

Расщепление сил (1.41) легче всего демонстрируется для сил, подчиняющихся закону обратных квадратов. Рассмотрим, например, однородно заряженные сферы с полным зарядом q и радиусом $a/2$. Из теоремы Гаусса ([330, гл. 2]) и третьего закона Ньютона (действие и противодействие равны по величине и противоположно направлены) следует, что сила взаимодействия двух таких сфер, расположенных на расстоянии $r \geq a$, равна

$$F(r) = \frac{q^2}{4\pi\epsilon_0} \frac{1}{r^2}, \quad r \geq 0, \quad (1.42)$$

и направлена вдоль линии, соединяющей центры сфер. Несколько более детальное вычисление показывает, что для $r \leq a$

$$F(r) = \frac{q^2}{4\pi\epsilon_0 a^2} \left(\frac{8r}{a} - \frac{9r^2}{a^2} + \frac{2r^4}{a^4} \right) r \leq a. \quad (1.43)$$

Таким образом, один из возможных способов расщепления силы — положить сеточную часть силы равной $F(r)$, а короткодействующую часть (не равную нулю только при $r < a$) — разности между силой, обратно пропорциональной квадрату расстояния, и $F(r)$. Более детально вопрос о расщеплении силы рассмотрен в гл. 8.

Количество операций, приходящихся на шаг по времени в методе P^3M , состоит из части $\alpha N_p + \beta(N)$, отвечающей PM и интегрированию уравнений движения, и части, пропорциональной $N_p N_n$ для этапа PP . Число соседей N_n приблизительно составляет $4\pi a^3 n_0/3$, где n_0 — средняя плотность частиц. Таким образом, полное число операций для модели P^3M можно записать в виде

$$\text{Число операций} = \alpha N_p + \beta(N) + \gamma N_n N_p. \quad (1.44)$$

Правильным выбором N , N_p и a добиваются линейной зависимости количества операций от числа макрочастиц. Время, затраченное на исполнение цикла шага по времени, приблизительно равно 1 с на тысячу частиц для типичных высокопроизводительных компьютеров, предназначенных для научных расчетов. Для десяти тысяч частиц это дает время цикла, равное 10 с, что можно сравнить с временем, приблизительно равным 20 мин, для той же ситуации при использовании метода PP .

Одномерная модель плазмы

В этой главе мы воспользуемся задачей моделирования электростатических волн в бесстолкновительной плазме, чтобы познакомиться с основными элементами схемы частица — сетка (РМ) и проиллюстрировать процесс построения вычислительных моделей. Этапы, которые нам предстоит разобрать, показаны на рис. 1.1, хотя сами программы моделирования мы рассмотрим в следующей главе. Схема РМ, которую мы должны получить, представляет собой одномерную схему NGP. Обобщение этой базисной схемы вместе с детальным анализом свойств вычислительной модели РМ дается в гл. 5 и 7.

2.1. ФИЗИЧЕСКАЯ СИСТЕМА

Плазма представляет собой горячий ионизированный газ, состоящий из ионов, электронов и нейтральных атомов. Великое множество физических процессов, протекающих с самыми разными пространственными и временными масштабами, обусловлено взаимодействием компонентов плазмы с электромагнитными полями. Математическая модель плазмы, которую мы будем здесь рассматривать, представляет собой идеализированную теоретическую модель в том смысле, что она учитывает только ограниченное подмножество процессов, происходящих в реальной плазме.

Идеализированная модель описывается уравнением Власова, характеризующим распределение электронов f и имеющим вид

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \quad (2.1)$$

и выражениями для потенциала ϕ , электрического поля \mathbf{E} и силы \mathbf{F} , отвечающими электростатическому локальному взаимодействию:

$$\nabla^2 \phi = -\rho/\epsilon_0, \quad \mathbf{F} = q\mathbf{E} = -q \nabla \phi. \quad (2.2)$$

Положительные ионы рассматриваются как неподвижный нейтрализующий фон с плотностью заряда ρ_0 , при этом полная

плотность заряда ρ определяется выражением

$$\rho(x) = q \int f dv + \rho_0, \quad (2.3)$$

где $q (= -e)$ — это заряд электрона. В уравнениях (2.1) — (2.3) неявно предполагается, что полный заряд электронов сохраняется, что электронный газ является бесстолкновительным (т. е. кинетическая энергия существенно превосходит потенциальную энергию), возмущения от положения равновесия параллельны электрическому полю (электростатическое приближение) и скорости намного меньше скорости света. Такая модель предназначена для изучения коллективных электронных колебаний в области с характерным масштабом, значительно большим, чем среднее расстояние между электронами, и на временах, значительно меньших времени столкновения электронов и характерного времени движения ионов (см. гл. 1).

Одномерную модель получают в предположении, что все величины (как, например, плотность заряда, потенциал и электрическое поле) зависят только от одной пространственной координаты, скажем x . По двум другим направлениям, y и z , плазма предполагается однородной и бесконечной. Из однородности сразу следует упрощение: положения и скорости вдоль y и z можно не учитывать, и (2.1) — (2.3) приводятся к виду

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \frac{F}{m} \frac{\partial f}{\partial v} = 0, \quad (2.4)$$

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho}{\epsilon_0}, \quad E = -\frac{d\varphi}{dx}, \quad F = qE, \quad (2.5)$$

$$\rho(x) = q \int f(x, v, t) dv + \rho_0. \quad (2.6)$$

Хотя уравнения (2.4) — (2.6) являются колоссальным упрощением полного описания плазмы, их исследование тем не менее представляет невероятно трудную математическую задачу. В линейной теории (см., например, [355]) переменные разлагают в ряд около равновесных решений (т. е. $f = f_0 + f_1$, $E = E_0 + E_1$ и т. д.) и удерживают только члены первого порядка. Применяя к полученным уравнениям первого порядка преобразование Фурье по пространству и преобразование Лапласа по времени, приходят к описанию возмущений в виде суперпозиции гармоник $\sim e^{i(kx - \omega t)}$. При исключении переменных первого порядка из линеаризованной версии уравнений (2.4) — (2.6) получается условие существования решения (дисперсионное соотношение), связующее ω и k :

$$1 = -\frac{\omega_p^2}{kn_0} \left\{ P \int_{-\infty}^{\infty} \frac{\partial f_0 / \partial v}{\omega - kv} dv - \frac{i\pi}{|k|} \frac{\partial f_0}{\partial v} \Big|_{\omega/k} \right\}, \quad (2.7)$$

где P обозначает интеграл в смысле главного значения.

Основные свойства уравнения (2.7) можно получить, подставляя вместо f_0 достаточно простую функцию, для которой интеграл можно было бы вычислить. Например, если мы возьмем функцию f_0 , первые три момента которой соответствуют распределению Максвелла с плотностью n_0 и тепловой скоростью v_T , т. е.

$$f_0 = \begin{cases} \frac{n_0}{2\sqrt{3}v_T} & \text{при } |v| < \sqrt{3}v_T, \\ 0 & \text{в остальных случаях,} \end{cases} \quad (2.8)$$

то (2.7) примет вид

$$1 \simeq \frac{\omega_p^2}{(\omega^2 - 3k^2v_T^2)} + \frac{i\pi\omega_p^2}{n_0k|k|} \frac{\partial f_0}{\partial v} \Big|_{\omega/k}. \quad (2.9)$$

Если мы положим $\omega = \omega_r + i\omega_i$, то для малых инкрементов $\omega_i \ll \omega_r$ уравнение (2.9) дает приближение тепловой плазмы для (2.7)

$$\omega_r^2 \simeq \omega_p^2 (1 + 3k^2\lambda_D^2), \quad \omega_i \simeq \frac{\pi}{2} \frac{\omega_p^2\omega_r}{n_0k|k|} \frac{\partial f_0}{\partial v} \Big|_{\omega/k}, \quad (2.10)$$

где $\lambda_D = v_T/\omega_p$ — дебаевская длина. Из вида вещественной части частоты ясно, что характерными частотами и длинами волн колебаний электростатической плазмы являются соответственно плазменная частота ω_p и дебаевская длина λ_D , где

$$\omega_p = \sqrt{\frac{nq^2}{\epsilon_0 m_e}}, \quad \lambda_D = \sqrt{\frac{\epsilon_0 k_B T}{nq^2}}. \quad (2.11)$$

Мнимая часть частоты описывает затухание Ландау. Если у f_0 имеется только один максимум, то все волны будут затухать; в противном случае амплитуды волн, знак фазовой скорости которых ω/k совпадает со знаком $\partial f_0/\partial v|_{\omega/k}$, будут экспоненциально нарастать.

Математическую теорию удается распространить на слабо-нелинейный режим с помощью оценки изменений f_0 , обусловленных f_1 [110, 31], но для изучения сильнонелинейных эффектов, связанных с неустойчивостями в пространстве скоростей, мы вынуждены обращаться к численному моделированию.

2.2. ПОСТРОЕНИЕ ДИСКРЕТНОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Первый шаг в переходе от математической к вычислительной модели состоит в аппроксимации математических уравнений (2.4) — (2.6) алгебраическими уравнениями, необходимыми для численных расчетов. Вычислительной моделью, которую мы по-

лучим, будет схема «ближайшего сеточного узла» (NGP) модели «частица-сетка» (PM). Чтобы понять, почему эта модель PM вполне оправдана, мы разберем процесс дискретизации достаточно подробно.

2.2.1. Уравнения движения макрочастиц

Первая стадия дискретизации состоит в замене уравнения (2.4) на уравнения его характеристик. Если мы представим, что фазовое пространство $x - v$ разбито регулярной сеткой на бесконечно малые ячейки объема $d\tau = dx dv$, где $d\tau$ достаточно мал, чтобы его мог занять только один электрон, тогда $f(x, v, t)$ дает вероятность того, что в момент времени t ячейка (x, v) занята. Если в момент времени t в этой ячейке находится электрон, то из этого следует, что в момент времени t' электрон будет находиться в ячейке (x', v') , где координаты (x', v') связаны с (x, v) посредством уравнений движения электронов:

$$x' = x + \int_t^{t'} v dt, \quad v' = v + \int_t^{t'} \frac{qE}{m} dt. \quad (2.12)$$

Следуя этому рассуждению, можно показать, что в общем случае

$$f(x', v', t) = f(x, v, t), \quad (2.13)$$

где (x', v') связаны с (x, v) посредством (2.12). Уравнения (2.12) и (2.13) представляют собой просто иную формулировку уравнения (2.4), что можно доказать, разложив уравнение (2.13) в ряд Тейлора и перейдя к пределу при $t' - t$, стремящемся к нулю. Уравнение движения (2.12) является характеристикой уравнения (2.4) или, что эквивалентно, f сохраняется вдоль траекторий электронов (равенство (2.13)). Таким образом, если бы мы знали значения f в каждой бесконечно малой ячейке фазового пространства в момент времени t , то можно было бы найти f для любого последующего момента времени, интегрируя уравнения движения.

Ясно, что задать f для каждой бесконечно малой ячейки с вычислительной точки зрения невозможно. Вместо этого мы берем выборку точек $\{x_i, v_i; i = 1, N_p\}$, где каждая точка представляет i -й элемент фазовой жидкости, соответствующий $N_s = \int_i f dx dv$ электронам плазмы, приходящимся на единицу площади в плоскости $y - z$. Траектории выборочных точек в фазовом пространстве определяются уравнениями

$$\frac{dx_i}{dt} = v_i, \quad M \frac{dv_i}{dt} = F(x_i), \quad (2.14)$$

где $M = N_s m_e$ и m_e — масса электрона. Физически привлекательная интерпретация выборочных точек как макрочастиц, введенная в предыдущей главе, следует естественным образом из (2.14).

Если плотность макрочастиц (выборочных точек) достаточно велика, тогда в дискретной модели не требуется явного задания функции распределения. Приближенные выражения для моментов функции распределения f строятся непосредственно по координатам макрочастиц. Например, моменты f по скорости можно аппроксимировать выражением

$$\int v^n f dv \simeq \frac{N_s}{\lambda} \int_{x-\lambda/2}^{x+\lambda/2} dx' \int v^n \tilde{f} dv, \quad (2.15)$$

$$\simeq \frac{N_s}{\lambda} \sum_i v_i^n, \quad (2.16)$$

где $\tilde{f} = \sum_{i=1}^{N_p} \delta(x - x_i) \delta(v - v_i)$ — распределение макрочастиц, а суммирование по индексу i в (2.16) производится по всем макрочастицам в окрестности $[x - (\lambda/2), x + (\lambda/2)]$ точки x , в которой оценивается момент. Такое локальное усреднение по окрестности необходимо, если мы хотим представить гладкую функцию распределения f : если бы усреднение не выполнялось, тогда макрочастицы вели бы себя как сильно коррелированные заряженные частицы, а не как элементы непрерывной гладко меняющейся фазовой жидкости (см. разд. 1.4).

Приближенная формула (2.15) дает некоторое представление о том, какую плотность макрочастиц можно считать достаточной. Характерной длиной волны электростатических колебаний является дебаевская длина λ_D (см. (2.10)). Для того чтобы адекватно представить эти колебания, необходимо, чтобы интервал усреднения был меньше λ_D или того же порядка. Количество макрочастиц на интервале усреднения $\lambda \sim \lambda_D$ должно быть достаточно большим (например, ~ 10), чтобы статистические флуктуации моментов были малы, из чего мы заключаем, что приемлемой плотностью выборки является плотность, при которой число макрочастиц на дебаевской длине велико.

Заключительная стадия дискретизации уравнения Власова (2.4) состоит в замене в (2.14) непрерывного времени дискретным набором временных слоев, разделенных во времени малым интервалом — временным шагом DT . Чаще всего применяется конечно-разностная схема с перешагиванием (leapfrog), названная так потому, что в ней положения и скорости «перешагивают» друг через друга по времени (см. рис. 2.1). Положения и поля определяются на целых временных слоях ($t = 0, DT, 2DT, 3DT \dots$), а скорости определяются на полуцелых временных слоях ($t = \frac{1}{2} DT, \frac{3}{2} DT \dots$). Конечно разностная ап-

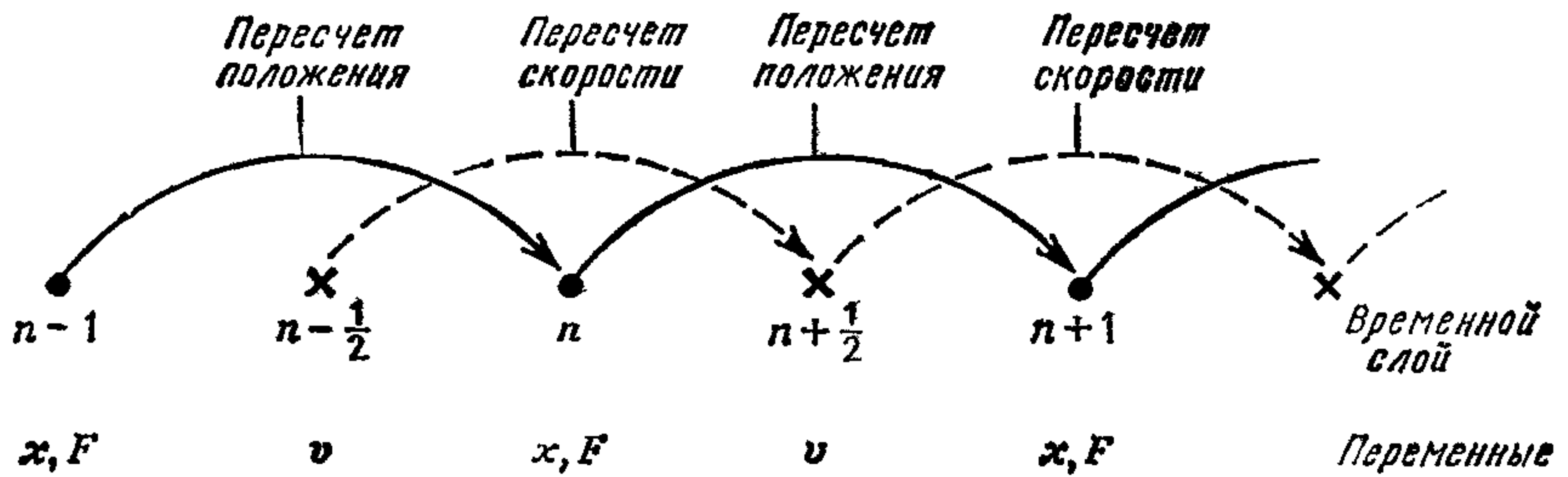


Рис. 2.1. Уравнения движения интегрируются по времени с использованием схемы с перешагиванием. Положения на временном слое $n-1$ пересчитываются с использованием скоростей с временного слоя $n-1/2$, причем для пересчета скоростей на слое $n-1/2$ берутся значения сил со слоя n и так далее.

проксимация с перешагиванием для уравнения (2.14) имеет вид

$$\begin{aligned} x_i^{n+1} - x_i^n &= v_i^{n+1/2} DT, \\ v_i^{n+1/2} - v_i^{n-1/2} &= \frac{F(x_i^n) DT}{m_i}. \end{aligned} \quad (2.17)$$

Свойства уравнения (2.17) подробно исследуются в гл. 4. Здесь мы лишь отметим, что временной шаг DT должен выбираться с учетом представления плазменных колебаний (с частотой $\sim \omega_p$), т. е. $\omega_p DT \ll 2$.

2.2.2. Уравнения для полей

Область пространства (т. е. область значений x), охватываемая вычислительной моделью (которая по необходимости ограничена), называется расчетной областью. Граничные условия задаются на концах области $x=0$ и $x=L$. Например, если бы концы были заземленными металлическими пластинами, то мы бы потребовали $\varphi=0$ на границах. Если заданы потенциалы на границах и распределение плотности заряда внутри расчетной области, то уравнение Пуассона (2.5) полностью определяет потенциал. Для настоящего примера больше подходят периодические граничные условия

$$\varphi(x) = \varphi(x + L). \quad (2.18)$$

Для периодических условий получается такое же дисперсионное соотношение (2.7), что и для бесконечной системы, за исключением того, что непрерывный спектр волновых чисел заменяется дискретным набором $k = 2\pi l/L$, длины волн которого λ удовлетворяют условию $L = l\lambda$, l целое.

Для того чтобы представить уравнения (2.5) для потенциала и электрического поля в удобной для численного решения фор-

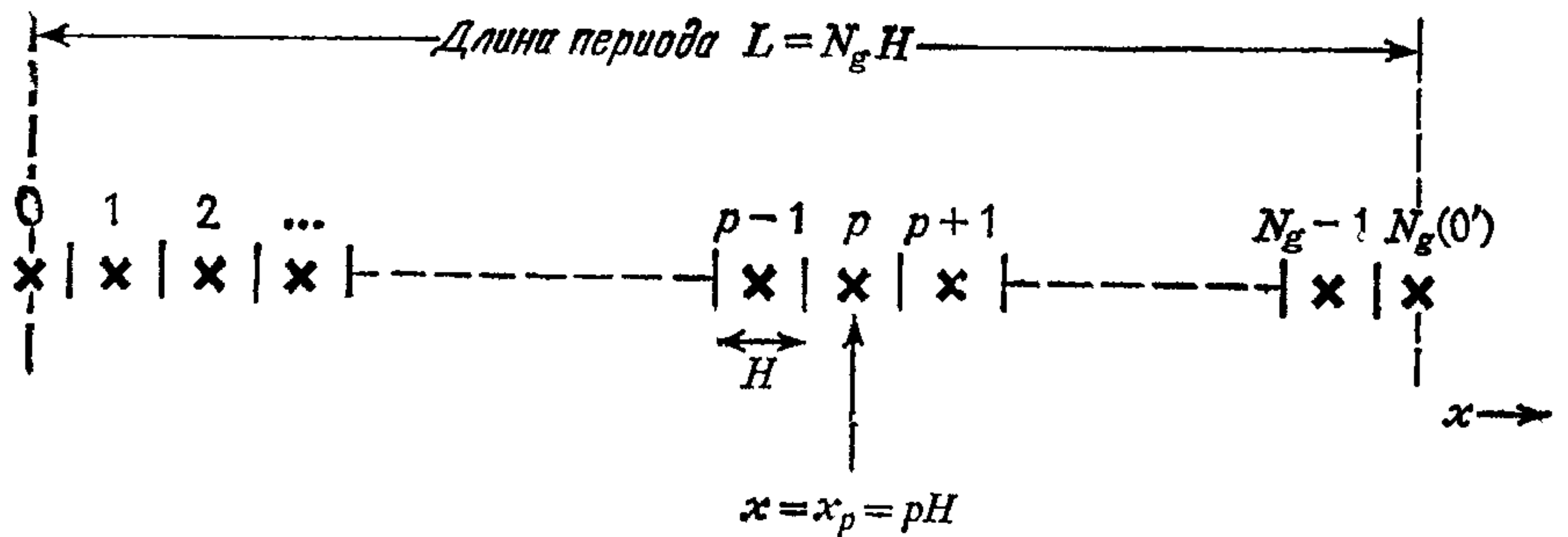


Рис. 2.2. Расчетная область $x \in [0, L]$ разбита на ячейки шириной H . В центре каждой ячейки расположен сеточный узел, в котором запоминаются значения плотности заряда, потенциала и электрического поля. Можно считать, что область периодически повторяется: сеточный узел N_g является периодическим отображением узла 0, узел $N_g + 1$ — это периодическое отображение узла 1 и т. д.

ме, вновь воспользуемся конечными разностями. Производная определяется выражением

$$\frac{d}{dx} f(x) = \lim_{h \rightarrow 0} \left\{ \frac{f(x + h/2) - f(x - h/2)}{h} \right\}. \quad (2.19)$$

В конечно-разностной аппроксимации h не устремляется к нулю, а принимает некоторое конечное значение H , которое выбирается из соображений достаточной точности и приемлемых вычислительных затрат. Для рассматриваемой модели плазмы вполне годится $H \leq \lambda_D$. Плотность заряда, потенциал и электрические поля представляются набором значений, расположенных с постоянным шагом H во всей расчетной области (сравните с представлением непрерывной логарифмической функции в виде таблицы логарифмов с постоянным шагом). Точки, в которых записываются величины, являются узлами сетки (или решетки) (см. рис. 2.2). Узлы сетки расположены в центрах ячеек шириной H . Если начало координат выбрано в узле 0, то узел с номером P имеет координату $x_p = pH$. Длина расчетной области L кратна ширине ячейки H ; для периодических граничных условий количество ячеек равно числу узлов сетки N_g :

$$L = N_g H. \quad (2.20)$$

Производные заменяются конечными разностями на сетке следующим образом:

$$\begin{aligned} \frac{d^2 \varphi}{dx^2} \Big|_{x_p} &= \frac{d}{dx} \left(\frac{d\varphi}{dx} \right) \Big|_{x_p} \\ &\approx \frac{1}{H} \left(\frac{d\varphi}{dx} \Big|_{x_{p+H/2}} - \frac{d\varphi}{dx} \Big|_{x_{p-H/2}} \right) \end{aligned}$$

$$\begin{aligned} &\simeq \frac{1}{H} \left(\frac{\varphi(x_{p+1}) - \varphi(x_p)}{H} - \frac{\varphi(x_p) - \varphi(x_{p-1}))}{H} \right) \\ &= \frac{\varphi(x_{p+1}) - 2\varphi(x_p) + \varphi(x_{p-1}))}{H^2} \end{aligned} \quad (2.21)$$

$$\text{и} \quad \left. \frac{d\varphi}{dx} \right|_{x_p} \simeq \frac{(\varphi(x_{p+1}) - \varphi(x_{p-1})))}{2H}. \quad (2.22)$$

В результате конечно-разностные уравнения для поля приобретают вид

$$\frac{\varphi_{p+1} - 2\varphi_p + \varphi_{p-1}}{H^2} = -\frac{\rho_p}{\epsilon_0}, \quad (2.23)$$

$$E_p = \frac{\varphi_{p-1} - \varphi_{p+1}}{2H}, \quad (2.24)$$

где использована более компактная форма записи величин с помощью целых индексов, а именно $\varphi_p \equiv \varphi(x_p)$ и т. д.

2.2.3. Раздача заряда и интерполяция силы

Чтобы завершить построение дискретной модели, требуется задать способ получения плотности заряда в сеточных узлах по распределению макрочастиц (раздача заряда) и получения сил в точках, где расположены макрочастицы, по сеточным электрическим полям (интерполяция силы).

Плотность заряда — это заряд в единице объема. В непрерывной среде она определяется так: берется полный заряд в некотором малом объеме, делится на этот объем и предполагается, что сам объем становится бесконечно малым. Тот же подход, за исключением предельного перехода, использован в схеме раздачи заряда NGP: плотность заряда в сеточном узле p равняется полному заряду в ячейке, окружающей узел p , деленному на объем этой ячейки:

$$\rho_p = \frac{1}{H} \sum_{\substack{\text{по частицам } i \\ \text{в ячейке } p}} N_s q + \rho_0. \quad (2.25)$$

Схема раздачи заряда NGP (2.25) сходна со схемой построения приближенных моментов функции распределения по координатам макрочастиц. Подставляя (2.15) в (2.6) и полагая $\lambda = H$, имеем

$$\rho(x) \simeq \frac{qN_s}{H} \int_{x-H/2}^{x+H/2} dx' \int_{-\infty}^{\infty} \tilde{f} dv + \rho_0, \quad (2.26)$$

$$= \frac{qN_s}{H} \int_{x-H/2}^{x+H/2} dx' n(x') + \rho_0, \quad (2.27)$$

$n(x') = \sum_{i=1}^{N_p} \delta(x' - x_i)$ является плотностью центров макрочастиц. Выполняя интегрирование по x' и полагая $x = x_p$, вновь получаем (2.25). Однако, когда мы будем анализировать раздачу заряда более подробно (гл. 5), полезной окажется другая форма записи, которая получается при введении в функции раздачи заряда W :

$$W(x) = \begin{cases} 1 & \text{при } |x| < H/2 \text{ или } x = H/2, \\ 0 & \text{в противном случае.} \end{cases} \quad (2.28)$$

Использование (2.28) позволяет нам переписать (2.27) в виде

$$\rho(x) = \frac{qN_s}{H} \int_0^L W(x' - x) n(x') dx' + \rho_0. \quad (2.29)$$

Значения ρ_p получаются путем выборки $\rho(x)$ в узлах сетки

$$\rho_p = \rho(x) |_{x=x_p}.$$

В схеме интерполяции силы NGP вычисление силы похоже на раздачу заряда. Силовое поле внутри ячейки приравнивается к его значению в сеточном узле в центре данной ячейки. Таким образом, сила, действующая на некоторую частицу i в точке x_i , равна

$$F(x_i) = N_s q E(x_p), \quad x_p - H/2 < x_i \leq x_p + H/2. \quad (2.30)$$

Это выражение также можно переписать в терминах функции раздачи заряда NGP:

$$F(x_i) = N_s q \sum_{p=0}^{N_g-1} W(x_i - x_p) E_p. \quad (2.31)$$

Заметим, что здесь суммирование проводится по сеточным узлам, тогда как в раздаче заряда суммирование проводится по частицам. Периодические граничные условия учтены в (2.31) тем, что сеточный узел с индексом N_g трактуется как сеточный узел с индексом 0.

2.2.4. Дискретная модель

Собирая результаты предыдущих трех подразделов вместе, получаем уравнения схемы NGP модели «частица — сетка» для одномерной электронной плазмы:

1. Раздача заряда

$$\rho_p^n = \frac{qN_s}{H} \sum_{i=1}^{N_p} W(x_i^n - x_p) + \rho_0. \quad (2.32)$$

2. Уравнения поля

$$\frac{\varphi_{p-1}^n - 2\varphi_p^n + \varphi_{p+1}^n}{H^2} = -\frac{\rho_p^n}{\epsilon_0}, \quad (2.33)$$

$$E_p^n = \frac{\varphi_{p-1}^n - \varphi_{p+1}^n}{2H}. \quad (2.34)$$

3. Интерполяция силы

$$F_i^n = F(x_i^n) = N_s q \sum_{p=0}^{N_g-1} W(x_i^n - x_p) E_p^n. \quad (2.35)$$

4. Уравнения движения

$$\frac{v_i^{n+1/2} - v_i^{n-1/2}}{DT} = \frac{F(x_i^n)}{N_s m_e}, \quad (2.36)$$

$$\frac{x_i^{n+1} - x_i^n}{DT} = v_i^{n+1/2}. \quad (2.37)$$

Периодические граничные условия накладываются на потенциал и положения частиц. Уравнения (2.32) — (2.37) вместе с граничными условиями полностью определяют эволюцию дискретной системы. Если состояние системы $\{x_i^n; v_i^{n-1/2}; i = 1, N_p\}$ известно, то последовательное решение уравнений (2.32) — (2.37) для $\{\rho_p^n\}$, $\{\varphi_p^n\}$, $\{E_p^n\}$, $\{F_i^n\}$, $\{v_i^{n+1/2}\}$, $\{x_i^{n+1}\}$ соответственно дает состояние системы $\{x_i^{n+1}, v_i^{n+1/2}; i = 1, N_p\}$ через время DT .

Для того чтобы можно было адекватно представить плазменные волны и чтобы модель была бесстолкновительной, параметры H , DT , L , N_p , фигурирующие в дискретной модели, должны удовлетворять следующим условиям:

1. $H \leq \lambda_D$.
2. $\omega_p DT \ll 2$.
3. $L \gg \lambda_D$.
4. $N_p \lambda_D \gg L$.

Дополнительное требование, состоящее в том, чтобы число макрочастиц со скоростями вблизи фазовой скорости неустойчивых волн было достаточно велико, как правило, выполняется, если обеспечено необходимое число частиц на дебаевской длине (условие 4).

Более точные количественные утверждения относительно указанных условий и качества представления физической системы можно получить, применяя к дискретной модели методы исследования, используемые для математической модели. Мы рассмотрим их позднее (гл. 5 и 7).

2.3. ЧИСЛЕННЫЕ АЛГОРИТМЫ

Для перевода дискретной модели в программу для вычислительной машины сначала разрабатывают последовательность операций (численный алгоритм), при исполнении которых определяются все требуемые неизвестные, входящие в уравнения дискретной модели, а затем эту последовательность представляют в виде ряда машинных команд. Мерой качества алгоритмов служит количество арифметических операций, требуемых для решения заданной системы алгебраических уравнений.

2.3.1. Безразмерные единицы

Нас в данном случае интересуют алгебраические уравнения (2.32) — (2.37). В этих уравнениях имеется ряд постоянных множителей, которые можно исключить тщательным выбором безразмерных единиц.

В качестве единиц длины и времени в системе дискретных уравнений естественно принять шаги дискретизации: ширину ячейки H и шаг по времени DT . Если обозначать безразмерные величины со штрихом, то

$$x' = x/H : \text{длина}, \quad (2.38)$$

$$t' = t/Dt : \text{время}. \quad (2.39)$$

Если, кроме того, измерять скорость в единицах ширины ячейки, деленных на шаг по времени, а ускорение в единицах ширины ячейки, деленных на квадрат шага по времени, т. е. взять

$$v' = v DT/H : \text{скорость}, \quad (2.40)$$

$$a' = a DT^2/H : \text{ускорение}, \quad (2.41)$$

то уравнения движения (2.36) — (2.37) принимают вид

$$v_i'^{n+1/2} - v_i'^{n-1/2} = a_i'^n, \quad (2.42)$$

$$x_i'^{n+1} - x_i'^n = v_i'^{n+1/2}, \quad (2.43)$$

где

$$a_i'^n = \frac{F_i^n DT^2}{N_s m_e H}. \quad (2.44)$$

Связь между безразмерным ускорением частицы $a_i'^n$ и определенным на сетке электрическим полем, вытекающая из (2.35),

$$a_i'^n = \sum_p W(x_i'^n - p) \frac{q DT^2}{m_e H} E_p^n, \quad (2.45)$$

подсказывает следующий выбор безразмерных единиц электрического поля:

$$E_p'^n = \frac{q \cdot DT^2}{m_e H} E_p^n, \quad (2.46)$$

что дает

$$a_i'^n = \sum_p W(x_i'^n - p) E_p'^n. \quad (2.47)$$

Единицы потенциала и плотности заряда получаются естественным образом при удалении постоянных коэффициентов из уравнений (2.34) и (2.33):

$$E_p'^n = \Phi_{p+1}'^n - \Phi_{p-1}'^n, \quad (2.48)$$

$$\Phi_{p-1}'^n - 2\Phi_p'^n + \Phi_{p+1}'^n = \rho_p'^n; \quad (2.49)$$

это дает

$$\Phi_p'^n = -\frac{q DT^2}{2m_e H^2} \Phi_p^n \quad (2.50)$$

и

$$\rho_p'^n = \frac{q DT^2}{2m_e \epsilon_0} \rho_p^n. \quad (2.51)$$

Уравнение (2.32) принимает вид

$$\rho_p'^n = \frac{\omega_p^2 DT^2}{2} \left[\sum_{i=1}^{N_p} \frac{W(x_i'^n - p)}{N_c} - 1 \right], \quad (2.52)$$

где $N_c = \rho_0 H / N_s |q|$ есть среднее число макрочастиц в ячейке (предполагается нейтральность полного заряда). Первый член в квадратных скобках в правой части (2.52) ясно показывает, каким образом флуктуация плотности в модели частиц связана с динамикой частиц, если N_c мало.

2.3.2. Раздача заряда

Выражение (2.52) дает простейшее представление о трудностях, возникающих при создании численных алгоритмов. Если бы оно использовалось для построения ρ_0' , ρ_1' , ρ_2' и т. д. по очереди, то вычисление $\{\rho_p'; p = 0, N_g - 1\}$ потребовало бы приблизительно $N_g N_p$ проверок, N_p сложений и N_g умножений.

Существенно более экономичная процедура проходит по частицам один раз, одновременно накапливая все плотности заряда:

Алгоритм

1. Инициализация накопителей плотности заряда:

for $p = 1$ *to* N_g *do*

$\rho_p := -N_c$

2. Накопление плотности заряда:

for $j=1$ to N_p do

$p := \text{nint}(x'_j)$ (определение ближайшего узла сетки)

$\rho_p := \rho_p + 1$ (приращение плотности заряда)

3. Масштабирование плотности заряда

for $p=1$ to N_g do

$$\rho_p'' := \frac{\omega_p^2 DT^2}{2N_c} \times \rho_p$$

Этот алгоритм требует N_p сложений и N_g умножений. В алгоритме имеются два момента, которые следует отметить особо: во-первых, на этапах 1 и 3, чтобы облегчить составление алгоритма на языке Фортран, используется периодически продолженный сеточный узел N_g , а не сеточный узел 0, и, во-вторых, на этапе 2 периодические граничные условия из координаты частиц налагаются таким образом, чтобы их координаты x лежали в интервале $0.5 \leq x \leq N_g + 0.5$. На практике этапы 2 и 3 объединяют (см. рис. 3.10) путем добавления на этапе 2 масштабированного приращения заряда, а не единицы.

2.3.3. Уравнение Пуассона

Решение уравнения (2.49) для значений потенциала представляет собой более сложную задачу, чем раздача заряда. Система уравнений, которые надо решить, если опустить штрих у безразмерных величин, имеет вид

$$\begin{aligned} \varphi_{N_g} - 2\varphi_1 + \varphi_2 &= \rho_1, \\ \varphi_1 - 2\varphi_2 + \varphi_3 &= \rho_2, \\ &\vdots \\ \varphi_{N_g-2} - 2\varphi_{N_g-1} + \varphi_{N_g} &= \rho_{N_g-1}, \\ \varphi_{N_g-1} - 2\varphi_{N_g} + \varphi_1 &= \rho_{N_g}, \end{aligned} \tag{2.53}$$

где периодичность использована для приведения индексов в первом и последнем уравнениях к диапазону $[1, N_g]$.

Чтобы эта система уравнений имела единственное решение, необходимо удовлетворить двум условиям. Первое состоит в том, что система является зарядово-нейтральной. Каждое значение потенциала фигурирует в трех разных уравнениях системы (2.53): один раз с множителем -2 и дважды с множителем $+1$. Складывая все эти уравнения друг с другом, получим слева

нуль, а справа полный заряд. Следовательно, для совместности должно выполняться

$$\sum_{p=1}^{N_g} \rho_p = 0 \quad (\text{зарядовая нейтральность}). \quad (2.54)$$

Второе условие состоит в том, что должен быть выбран базисный потенциал. Если построить набор величин $\{\psi_p = \varphi_p + c; p \in [1, N_g]\}$, где c — произвольная постоянная, то $\{\psi_p\}$ также является решением уравнения (2.53). Выбирая постоянную c так, чтобы $\varphi_{N_g} = 0$, приводим (2.53) к виду

$$\begin{aligned} -2\varphi_1 + \varphi_2 &= \rho_1 \\ \varphi_1 - 2\varphi_2 + \varphi_3 &= \rho_2 \\ \varphi_2 - 2\varphi_3 + \varphi_4 &= \rho_3 \\ \varphi_3 - 2\varphi_4 + \varphi_5 &= \rho_4 \\ &\vdots \\ \varphi_{N_g-2} - 2\varphi_{N_g-1} &= \rho_{N_g-1} \\ \varphi_1 + \varphi_{N_g-1} &= \rho_{N_g} \end{aligned} \quad (2.55)$$

Если мы умножим первое уравнение системы (2.55) на единицу, второе на двойку, третье на тройку и т. д. и сложим полученные уравнения друг с другом, то получим

$$N_g \varphi_1 = \sum_{p=1}^{N_g} p \rho_p. \quad (2.56)$$

Уравнение (2.56) дает величину φ_1 . Далее из первого уравнения системы (2.55) определяется φ_2 , из второго φ_3 и так далее, пока решение не будет найдено полностью.

Алгоритм

1. Вычисление потенциала в сеточном узле 1:

$$\varphi_1 := 0$$

for $p = 1$ to N_g do

$$\varphi_1 := \varphi_1 + p \rho_p$$

$$\varphi_1 := \varphi_1 / N_g$$

2. Вычисление потенциала в сеточном узле 2:

$$\varphi_2 := \rho_1 + 2\varphi_1$$

3. Вычисление остальных значений потенциала:

for $p = 3$ to N_g do

$$\varphi_p := \rho_{p-1} + 2\varphi_{p-1} - \varphi_{p-2}$$

Вычислением значения φ_{N_g} на этапе 3 алгоритма можно воспользоваться для проверки зарядовой нейтральности и влияния

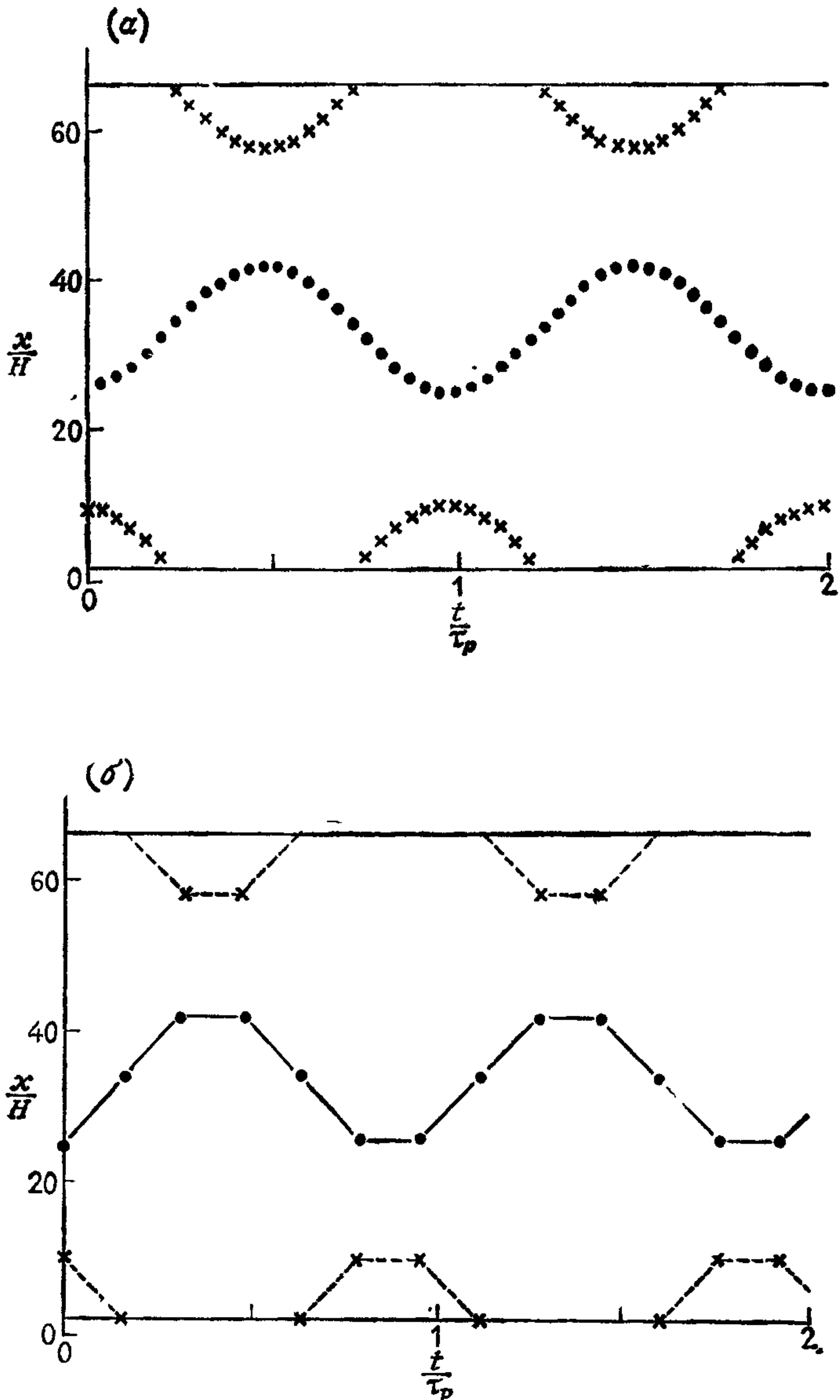
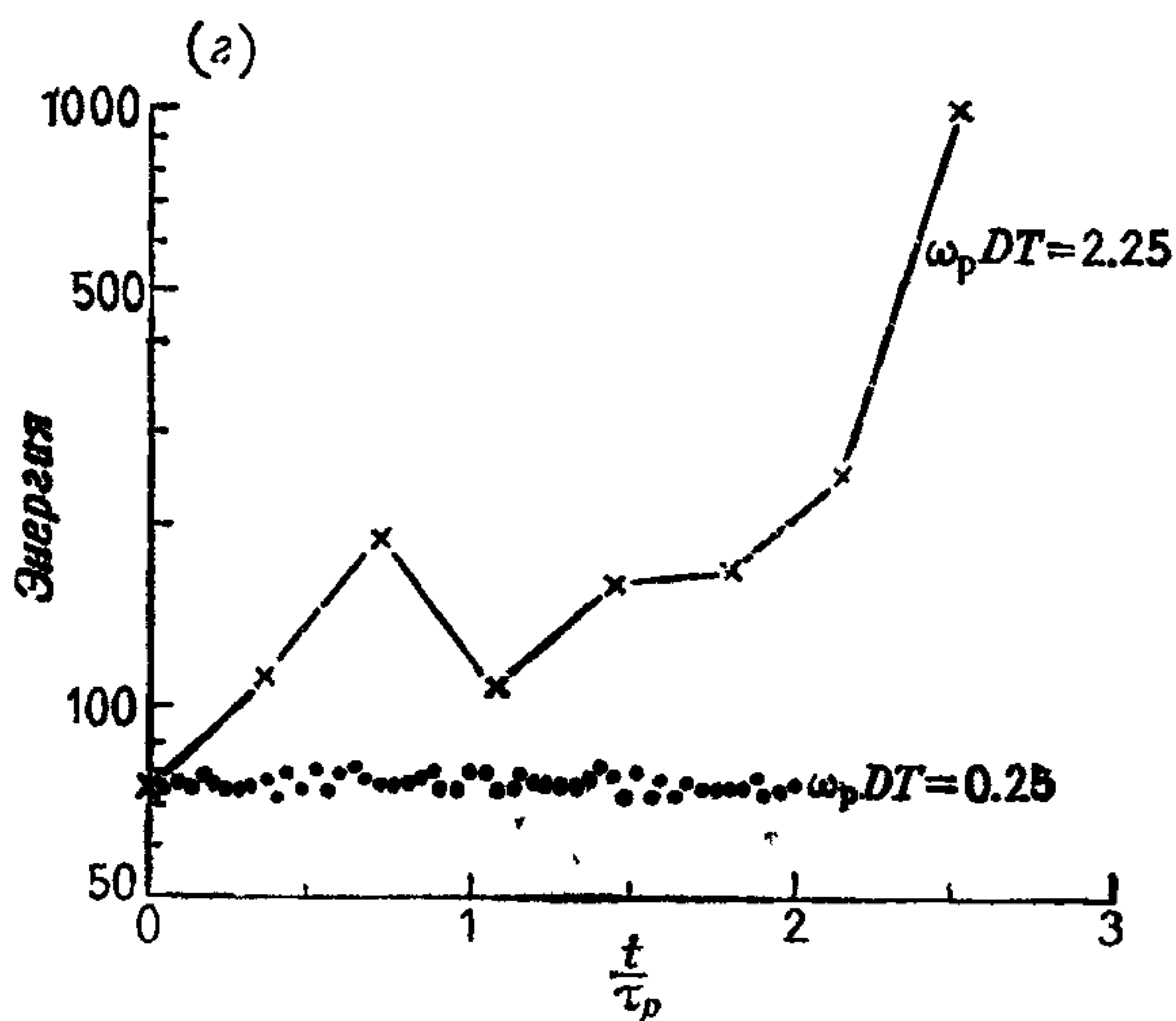
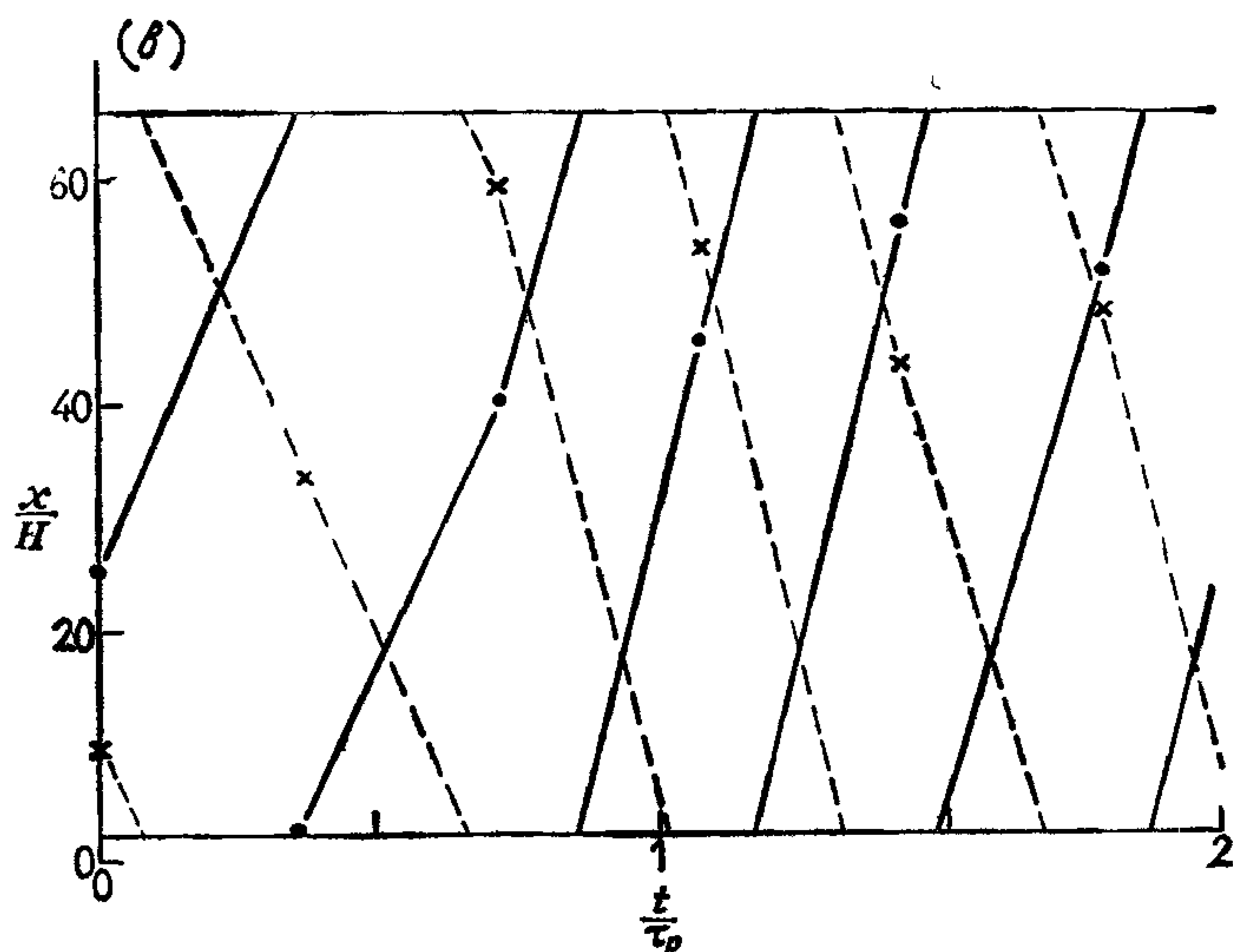


Рис. 23 Пространственно-временные графики колебаний пары электронов, разделенных расстоянием $L/4$, для различных значений временного шага. (a) $\omega_p DT = 0.25$ (устойчивое и точное); (б) $\omega_p DT = 1.0$ (устойчивое, но не точное); (в) $\omega_p DT = 2.25$ (неустойчивое); (г) полная энергия для случаев (a) и (б).



округления. Рассмотренный выше алгоритм относится к частному случаю трехдиагональных коэффициентов $[1, -2, 1]$. Алгоритм для случая произвольных коэффициентов представлен в разд. 6.4.1.

2.4. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Систематический метод перевода численного алгоритма в программу моделирования обсуждается в следующей главе. Результатом его использования является программа моделирования ES1D1V, показанная на рис. 3.10. Следующий этап состоит

в проведении контрольных экспериментов для того, чтобы продемонстрировать, что программа моделирования воспроизводит физическое поведение плазмы с удовлетворительной точностью. В результате таких экспериментов проверяется правильность программы и определяются нужные комбинации численных и физических входных параметров. Эксперименты, кратко описанные в следующих подразделах, были выбраны с таким расчетом, чтобы проиллюстрировать типы тестов, на которых обычно отрабатываются программы моделирования методом «частица — сетка», и чтобы продемонстрировать ловушки, в которые может попасть неосторожный вычислитель-экспериментатор.

2.4.1. Двухчастичный тест

Простую, но действенную проверку всего вычислительного цикла обеспечивает двухчастичный тест. Симметрию потенциала и т. д. можно легко проверить; действительно, в одномерном случае потенциалы, поля и силы можно получить аналитически (см. гл. 5). Колебание пары частиц дает возможность построить критерии для проверки величины энергии, граничных условий и подобия модели. Для малого шага по времени пара частиц, разведенных на любое расстояние, отличное от $L/2$ (подумайте об этом), колеблется с правильной физической плазменной частотой ω_p , определяемой формулой (2.11). Когда шаг по времени увеличивается, частота, наблюдаемая в моделировании, становится неверной и превышает физическую частоту. При значениях $\omega_p DT > 2$ частота модели становится комплексной, и интегрирование по времени делается неустойчивым (см гл. 4). Вычисляемое при этом решение не имеет никакого отношения к физическим колебаниям.

Те данные, которые приняты по умолчанию в программе одномерной плазмы, показанной на рис. 3.10, как раз отвечают задаче о двухчастичном колебании. Читателям, желающим провести эксперименты с помощью этой модели одномерной плазмы, будет полезно построить графики положения, скорости, кинетической энергии, потенциальной энергии и полной энергии как функций времени для 50 шагов по времени, используя данные по умолчанию. Влияние шага по времени на устойчивость и точность может быть наглядно показано при выполнении двухчастичного теста для следующих значений шага по времени:

1. $\omega_p DT = 0.25$ (значение по умолчанию). Шаг по времени, обеспечивающий устойчивое и точное решение. Плазменный период воспроизводится правильно, при этом период $\tau_p = 2\pi/\omega_p = 25.1$ шагов по времени (см. рис. 2.3а).

2. $\omega_p DT = 1.0$. Шаг по времени, дающий устойчивое, но неточное решение. На один период колебаний в модели приходится

6 шагов по времени вместо физически правильного значения $\tau_p = 2\pi/\omega_p = 6.28$ шагов по времени (см. рис. 2.3б).

3. $\omega_p DT = 2.25$. Шаг по времени, дающий неустойчивое решение, поскольку $\omega_p DT > 2$. Вычисленные результаты являются бессмысленными, и экспоненциальный рост численно неустойчивого решения уже через несколько шагов по времени может привести к арифметическому переполнению, если в программе это не будет предусмотрено. В программе, представленной на рис. 3.10, вычисление автоматически прекращается, если какая-нибудь частица за шаг по времени проходит расстояние, большее $2L$. В данном случае это происходит через 7 шагов по времени (см. рис. 2.3в). Неустойчивость отчетливо проявляется также в плохом сохранении полной энергии (см. рис. 2.3г), чего не наблюдается в рассмотренном выше случае 1.

2.4.2. Дисперсия волн

Проверку коллективных свойств модельной плазмы можно произвести путем измерения зависимости частоты от длины волны для волн малой амплитуды. Линейная теория предсказывает, что для непрерывной среды эта зависимость имеет вид (2.9), хотя в вычислительных моделях характер ее за счет численных эффектов может заметно меняться (см. гл. 7). Кроме того, вычислительная модель включает нелинейные эффекты, которые значительно усложняют простое соотношение $\omega = \omega(k)$, предсказываемое линейной теорией.

Измерение зависимости частоты от длины волны для отдельной моды можно выполнить следующим образом: в качестве начального состояния задать в программе тепловое распределение частиц, возбудить интересующую моду посредством модуляции скоростей частиц требуемой длиной волны, а потом измерить эволюцию во времени гармоник потенциала с той же длиной волны. Для слабо затухающих длинноволновых гармоник доминирующую частоту обычно удается определить из графика зависимости амплитуды от времени после затухания начального переходного процесса. Более систематический метод нахождения доминирующей частоты предполагает использование спектральных методов на определенном участке зарегистрированной зависимости амплитуды от времени. За дальнейшими подробностями измерения временных корреляций читатель отсылается к работе Брасвелла [46].

2.4.3. Холодная плазма

Предельный случай нулевой температуры, который часто используется для упрощения математического анализа плазменных явлений, недоступен в вычислительном эксперименте. Точно

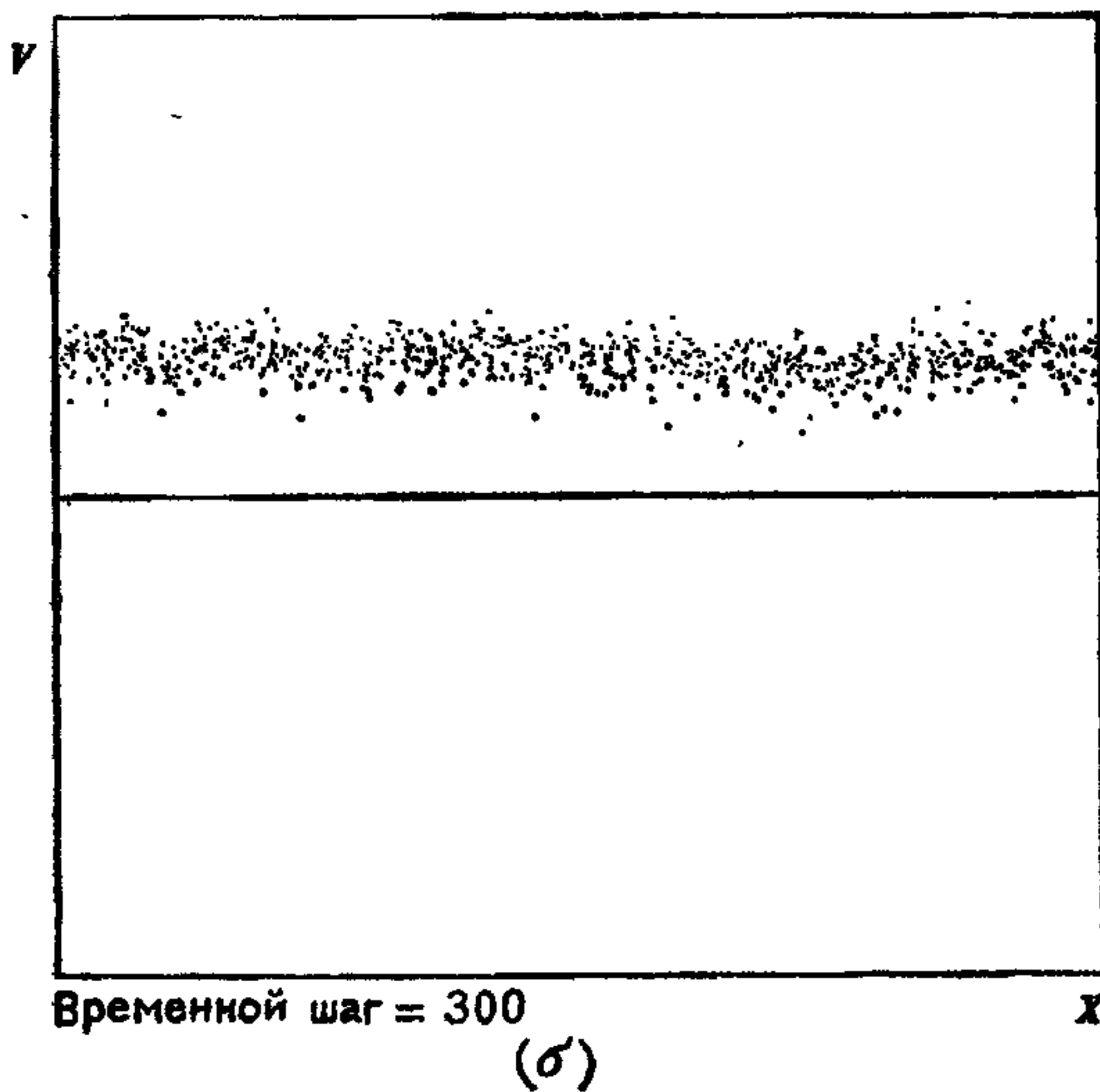
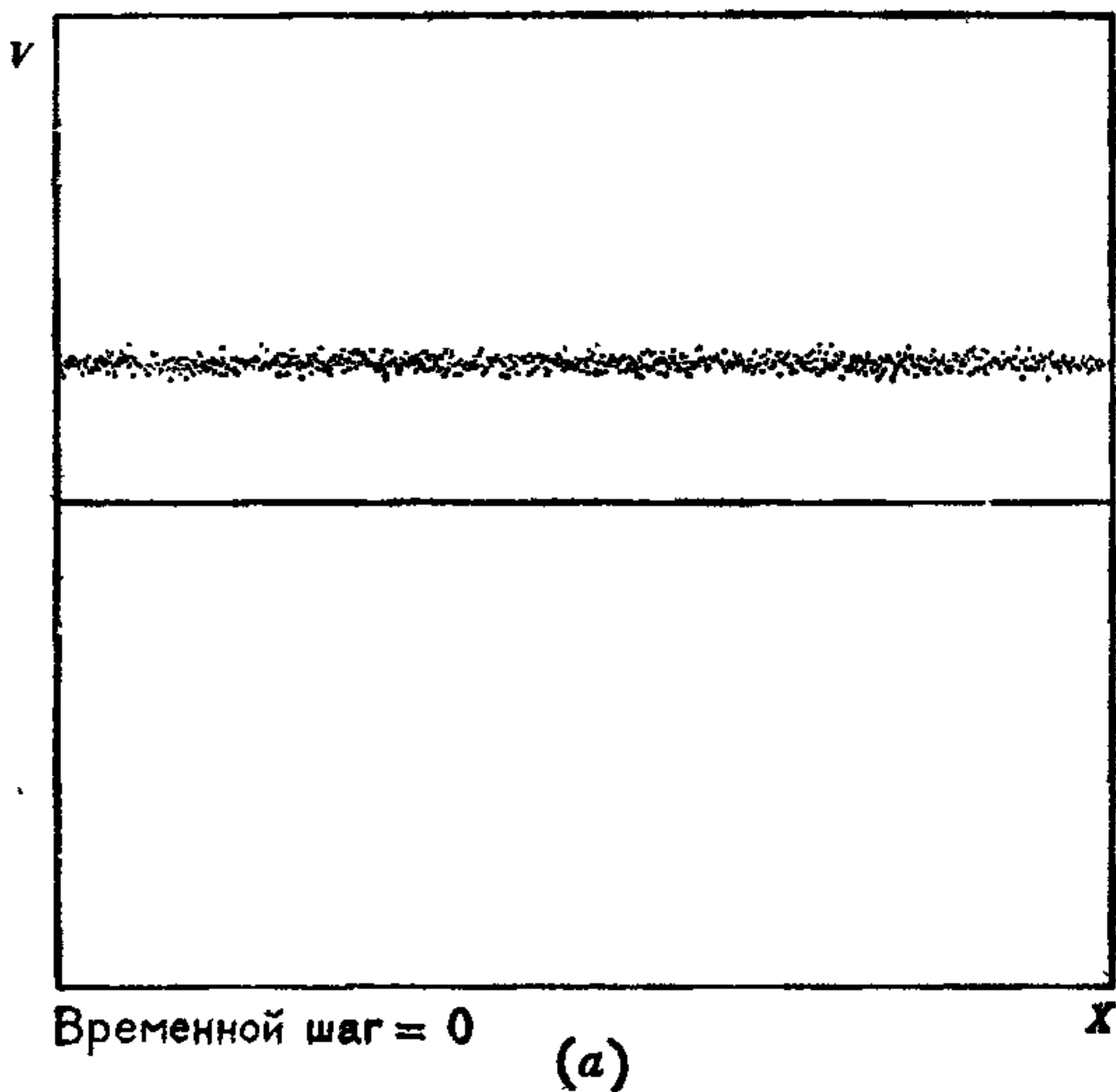


Рис. 2.4. Фазовый портрет одиночного пучка холодных электронов ($H = 22.5\lambda_D$), полученный по программе, показанной на рис. 3.10. Начальный разброс скорости (a) увеличивается из-за нефизических неустойчивостей, достигая устойчивого конечного состояния (б).

так же эта полезная математическая абстракция невозпроизводима в лаборатории. Если пытаться провести численный эксперимент с очень низкой температурой плазмы, то численное взаимодействие мод будет приводить к неустойчивости плазмы: в конечном счете эта неустойчивость будет проявляться в росте температуры и полной энергии до тех пор, пока плазма не станет устойчивой. В конце концов нефизическая неустойчивость прекращает расти и наступает насыщение при конечной амплитуде.

Дальнейшее обсуждение численных неустойчивостей в холодной плазме (т. е. плазме, в которой дебаевская длина λ_D существенно меньше, чем ширина ячейки H) и библиография по этому вопросу приведены в разд. 7.3. Эти неустойчивости можно обнаружить с помощью программы ES1D1V, подбирая тепловую скорость v_T и потоковую скорость v_0 таким образом, чтобы начальное состояние модели лежало в области, отмеченной на рис. 7.1 как «неустойчивая». Тогда графики кинетической, потенциальной и полной энергии покажут характерный рост, продолжающийся до тех пор, пока неустойчивость не стабилизируется. На рис. 2.4 показаны начальное и конечное распределение частиц в фазовом пространстве для вычислительного эксперимента с $H = 22.5 \lambda_D$, $N_g = 64$, $N_p = 1000$, $\omega_p DT = 0.25$. В этом примере неустойчивость насыщается быстрее, чем за сто шагов по времени.

2.4.4. Сохранение энергии

Даже в отсутствие нефизических неустойчивостей, упомянутых в разд. 2.4.3, моделирование будет тем не менее показывать устойчивый рост полной энергии. Причинами этого роста являются флуктуации численных погрешностей при определении силы и погрешности, обусловленные конечным размером шага по времени. В разд. 7.6.1 дается простая теория случайного блуждания, которая описывает это возрастание энергии. Используя программу ES1D1V (рис. 3.10), можно провести эксперименты для проверки области применимости зависимости скорости нагрева (см. (7.140)) от параметров $(\omega_p DT)$, (H/λ_D) и $(n_0 \lambda_D)$ по аналогии с тем, как это делалось в двумерном электростатическом случае, описываемом в разд. 9.2.2.

2.4.5. Двухпотоковая неустойчивость

Простейшей физической неустойчивостью электростатической плазмы является неустойчивость, наблюдаемая при прохождении двух одинаковых потоков электронов друг через друга, как это показано на рис. 2.5а. Взаимодействие пучков вызывает бунчировку электронов (макрочастиц), приводя к образованию потенциальных ям, которые, в свою очередь, далее усиливают бунчировку (рис. 2.5б). Свободная энергия течения быстро преобразуется в энергию поля, и частицы становятся захваченными в потенциальные ямы (рис. 2.5в). С течением времени энергия концентрируется в одном длинноволновом вихре в фазовом пространстве $x-v$ (рис. 2.5г).

Дополнительные сведения о линейной теории двухпотоковой неустойчивости можно найти в большинстве стандартных учебников по физике плазмы (напр., [84]). Для тех читателей, ко-

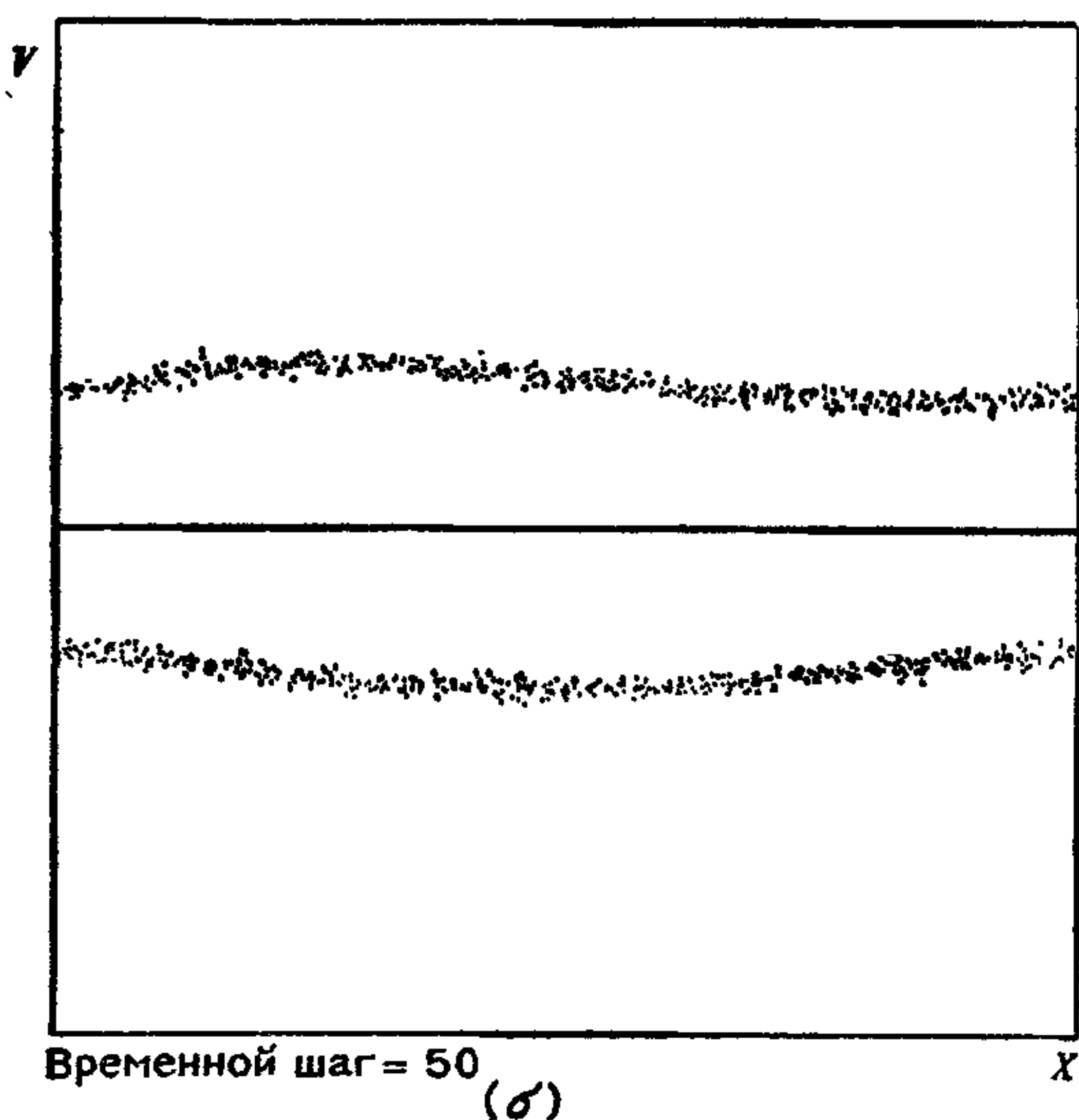
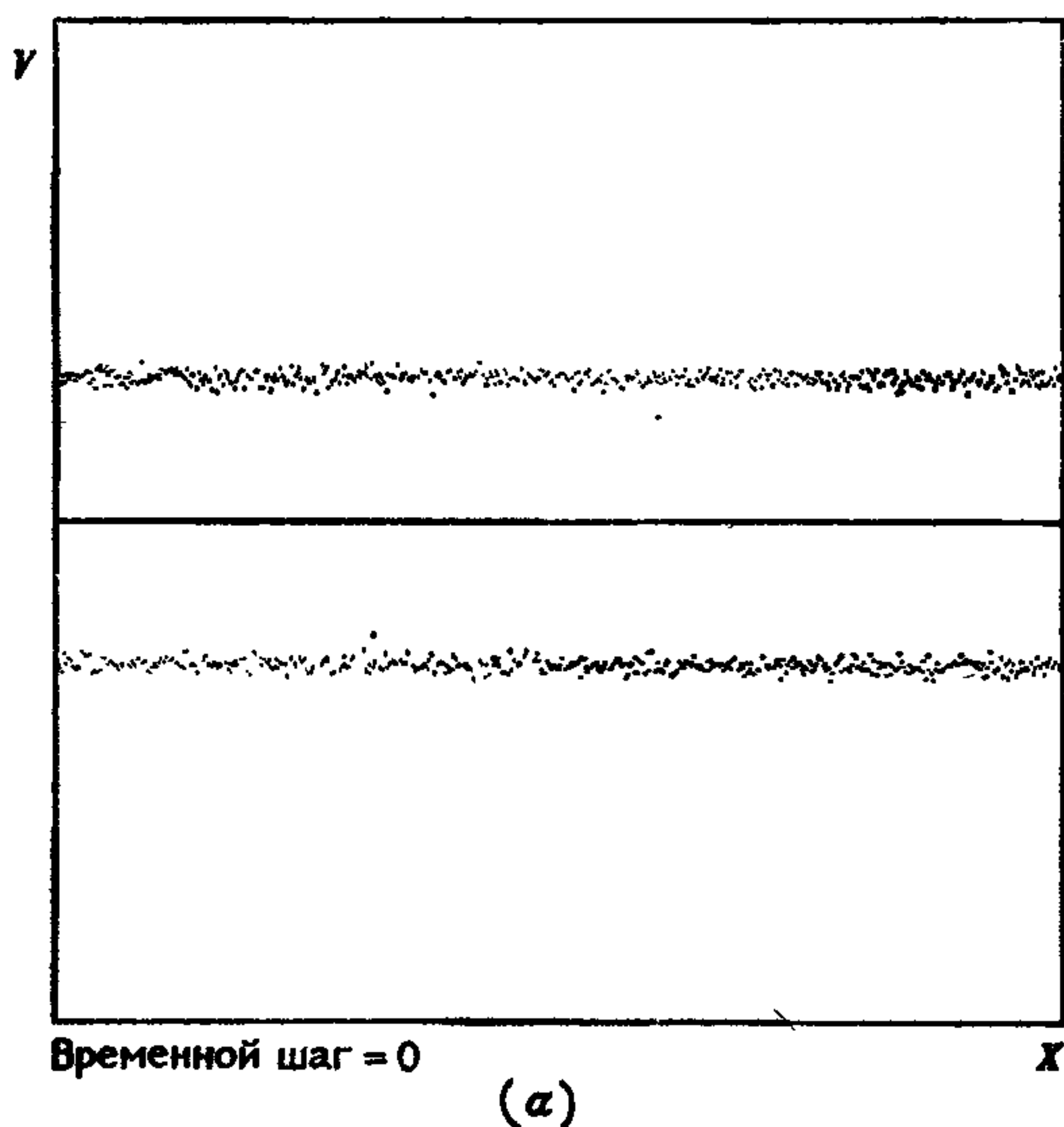
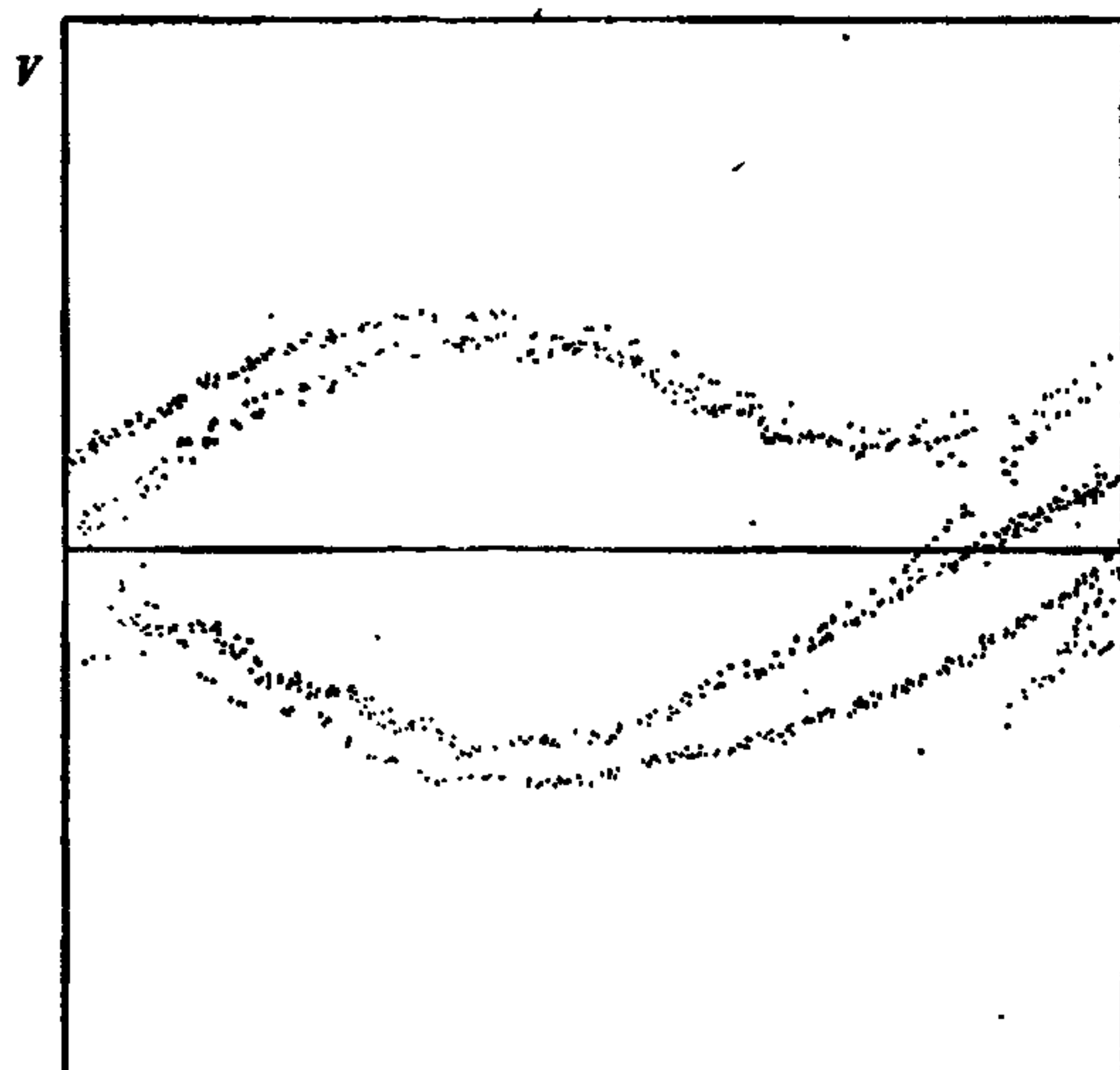


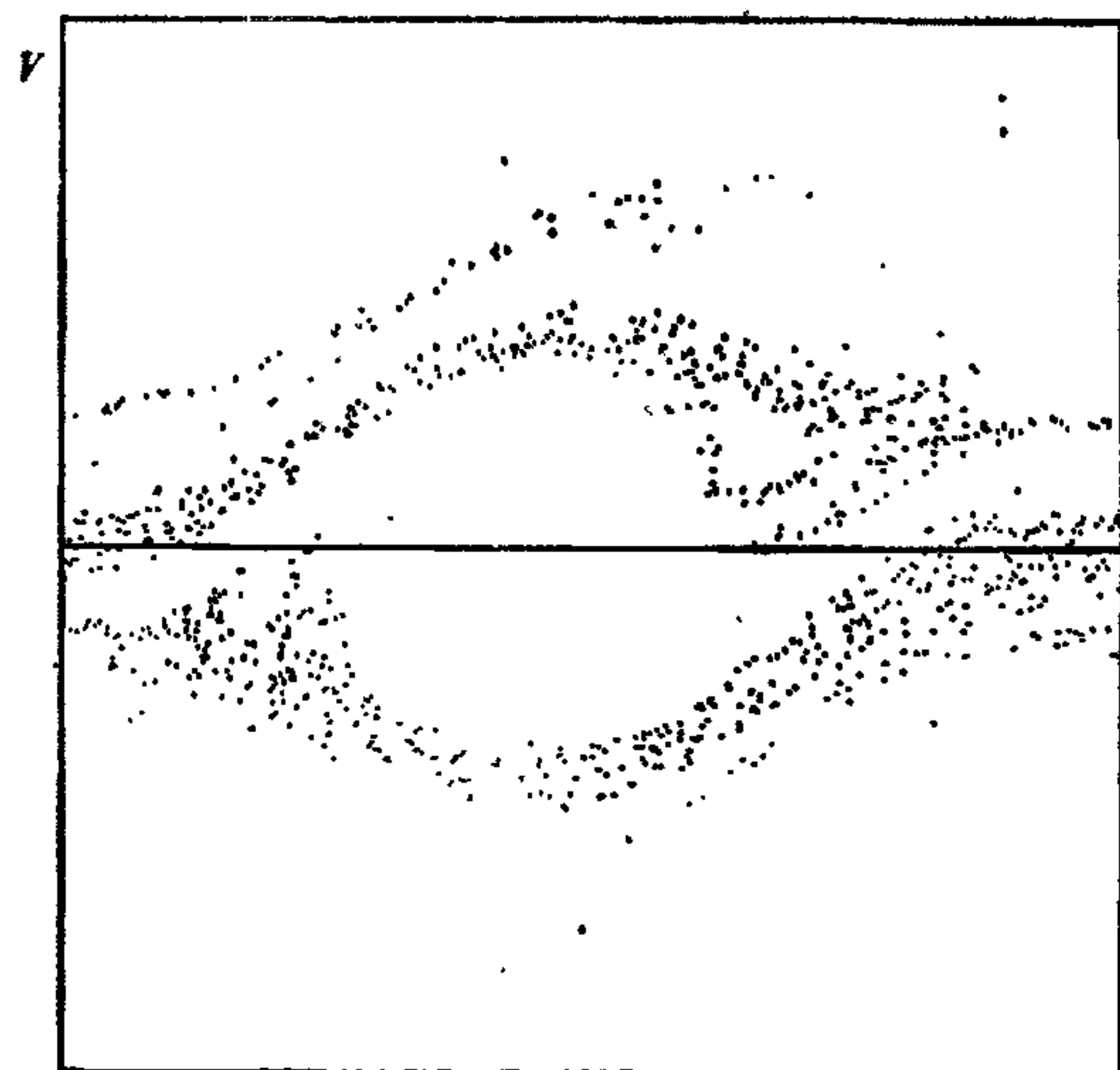
Рис. 2.5. Фазовый портрет нелинейной эволюции двухпоточковой неустойчивости. Использовались следующие параметры: $v_0/v_T = 25$; $H/\lambda_D = 1.25$; $\omega_p DT = 0,25$; $N_g = 64$ и $N_p = 2000$. Графикам (а), (б), (в) и (г) отвечают моменты времени 0, 50, 100 и 150 шагов соответственно.

торых интересуют эксперименты по двухпоточковой неустойчивости, мы приводим готовую программу одномерного моделирования. Кроме случая одинаковых пучков, данная одномерная



Временной шаг = 100

(a)



Временной шаг = 150

(b)

электростатическая модель может быть использована для исследования неодинаковых пучков, а также ион-электронных и ион-ионных неустойчивостей (с заниженным отношением массы иона к массе электрона, дабы ограничить необходимый объем машинного времени).

Программа моделирования

3.1. ВВЕДЕНИЕ

Характерной особенностью большинства программ, применяемых для проведения на ЭВМ вычислительных экспериментов, является их большой объем и сложность. Такие программы, как правило, насчитывают тысячи карт, а на их создание уходят не человеко-часы, а несколько человеко-лет. Совершенно очевидно, что без тщательного проектирования и программирования не обойтись, если мы хотим, чтобы затрачиваемые усилия привели к полезным физическим результатам.

Несмотря на то, что нас в этом смысле интересуют главным образом модели частиц, обсуждаемый в этой главе метод разработки программ находит гораздо более широкое применение. В общих чертах этот метод сводится к тому, что программа проектируется «сверху вниз», после чего также по принципу сверху вниз осуществляются программирование, тестирование и последовательное уточнение. Данный метод программирования обладает теми несомненными достоинствами, что развивающаяся программа обеспечивает себе собственный «испытательный стенд», который позволяет ее систематически тестировать в ходе всего времени разработки, и, кроме того, в итоге естественным образом получается хорошо структурированная программа.

Со стадией проектирования программы мы уже встречались: на рис. 1.1 кратко отражены все ее этапы от физического явления до численного алгоритма. Эти этапы довольно подробно разбирались в гл. 2 на одномерной плазменной модели NGP, которой мы еще воспользуемся в этой главе в качестве примера. Поэтому далее мы сосредоточимся на вопросах проектирования, связанных в основном с переходом от вычислительной модели к программе для ЭВМ.

Основная задача, стоящая при создании всякой программы моделирования, — обеспечить средство для получения физических результатов на основе данной вычислительной модели. Эта задача легко решается в случае простых расчетов, для которых требуется закодировать, скажем, порядка двух сотен карт; все, что нужно сделать, это просто оттранслировать такие алгоритмы в машинные коды. Однако относительно большие и сложные

программы, применяемые в многомерном моделировании частицами, требуют уже четкого понимания того, как взаимодействуют машина, программа и ее пользователь. По этой причине в следующем разделе мы слегка отклонимся от темы с тем, чтобы материал, представленный в разд. 3.3 и 3.4, предстал в нужном свете.

В разд. 3.3 содержится краткое описание системы программирования, а именно системы программирования OLYMPUS [333, 80, 212, 213, 214, 334, 210, 211, 121], воплотившей в себе нисходящий принцип структурного программирования, который мы всемерно поддерживаем. В разд. 3.4 эта методика практически продемонстрирована на одномерной программе моделирования плазмы ES1D1V.

3.2. ТРЕБОВАНИЯ К ПРОГРАММЕ И СПЕЦИФИКАЦИИ

3.2.1. Требования пользователя

Посмотрим сначала на программу моделирования с точки зрения пользователя. Это, как правило, физик, химик или инженер. В моделировании его интересуют в первую очередь физические результаты, а не численные или программистские аспекты. Нужно, чтобы ничто не мешало ему заниматься своим основным делом, а для этого в его распоряжении должен быть большой набор заранее сконструированных и проверенных программ, и он должен уметь ими воспользоваться. Он не должен тратить лишнее время и силы ни на детальное выяснение того, как работает каждая отдельная программа, ни на программирование своей собственной версии. Тем не менее в любой программе, которой он пользуется, он должен суметь легко определить все принятые в ней теоретические предположения и изменять или расширять их в соответствии со своими потребностями. Аналогично, ему должны быть известны все ограничения примененных численных схем с тем, чтобы он сам либо его коллеги могли модифицировать или заменять их в тех задачах, где они оказались непригодными.

Самое главное отличие больших программ, применяемых в научных расчетах, от большинства больших программ другого назначения заключается в мере их изменчивости. В редких случаях большую программу моделирования прогоняют в одном и том же виде несколько раз. Ее пользователь всегда пребывает в состоянии поиска в какой-нибудь ранее неизученной области исследований. Он обращается к многочисленным вариантам исходных теоретических посылок, начальных условий, конфигураций и численных схем, и, кроме того, изменяет форму выдачи результатов и управление программой. Как пользователю, ему нужно, чтобы программы моделирования обладали четкой мо-

дульной структурой в соответствии с некоторым стандартом и поддерживались обширной библиотекой заранее проверенных программ (или программных модулей).

Требования пользователя к программе моделирования сводятся, по-существу, к условиям, чтобы в такой программе было нетрудно разобраться, удобно ею пользоваться и легко адаптировать. О конкретной программе он в целом судит по тому, насколько хорошо она удовлетворяет этим условиям. На самом верхнем уровне он может считать программу «черным ящиком», на вход которого он поставляет исходные данные, а на выходе получает физические результаты. На следующем уровне программа видится уже состоящей из основной управляющей структуры, основных данных и набора модулей (пока «черных ящиков»). Основная управляющая структура определяет ход выполнения программы, основные данные характеризуют состояние физической системы, а модули (содержащие последовательности инструкций) выступают в качестве операторов высокого уровня, действующих на эти основные данные. Например, в модели, описанной в предыдущей главе, цикл по временному шагу можно рассматривать как оператор высокого уровня, действие которого состоит в том, чтобы текущее состояние системы сдвинуть вперед во времени на шаг DT . При более детальном рассмотрении в каждом модуле можно, в свою очередь, выделить управляющую структуру, данные и операторы, и так до тех пор, пока не дойдем до отдельных переменных и инструкций.

Противоречие между требованиями, которые предъявляет к программе пользователь (легкость чтения, использования и адаптации) и которые накладывает ЭВМ (минимизация объема памяти для данных и числа операций), в значительной степени устраняется с помощью систематического программирования сверху вниз. Требования пользователя отражаются, главным образом, на самых верхних уровнях программы, инструкции которых выполняются редко, в то время как машинные ограничения касаются в основном самого низкого уровня, т. е. отображения часто исполняемых частей алгоритмов в программные инструкции.

На практике указанными требованиями пользователя зачастую пренебрегают. Существует тенденция разрабатывать программы, начиная с некоего ядра, вокруг которого далее беспорядочным образом формируется архитектура программы, управляющая структура и организация данных. Двойное назначение языков программирования высокого уровня, а именно сделать программу понятной и человеку, и машине, обычно в полной мере не используется. Документация по программе зачастую скудна и не обновляется по мере внесения в нее изменений. Фактически многие программы становятся непригодными именно из-за того, что вносимые изменения не регистрируются и тесты

данной программной разработки, а также ее документация, оказываются неприменимыми. Эти проблемы в значительной степени снимаются при использовании системы программирования OLYMPUS, основанной на перечисленных в табл. 3.1 спецификациях.

Таблица 3.1. Спецификации для систематического программирования

Спецификация	Средства
Мобильность	Использование языка высокого уровня. Выделение системнозависимых и системнонезависимых частей.
Сохранность	Создание СПФ-инварианта. Опознавательный код для проверки.
Гибкость	Гибкость закладывается в управляющую структуру программы. Используются редактирующие файлы и системнонезависимый редактор.
Архитектура	Используется универсальная управляющая структура. Организация операторов и операндов ориентирована на пользователя.
Управление	Используется универсальная управляющая структура. Встраиваются переключатели и точки диагностики.
Диагностика Система обозначений	Используются стандартные утилиты выдачи. Соглашения по начальным буквам. Мнемоника.
Форма записи	Ступенчатая запись со стандартными отступами, разделительные строки и т. д.
Документация	Указатели, перекрестные ссылки, десятичная нумерация. Заголовки и подзаголовки.
Составление	Сверху вниз, исходя из скелета.
Тестирование	Сверху вниз, пользуясь диагностическим дампом.
Автоматическая генерация программ	Код и документация создаются на основе главных указателей. Автоматически приводятся к стандартному виду текст и заголовки.

3.2.2. Программные спецификации

Первые три пункта в табл. 3.1 касаются в основном самого верхнего уровня, т. е. программы в целом. Спецификации архитектуры, управления и диагностики обретают смысл на следующем уровне, когда рассматривается та схема, по которой должны собираться различные программные модули. Для ясности как

структуры в целом, так и отдельных модулей важны спецификации обозначений, формата записи и документации. Последние три спецификации касаются эффективного создания программ по правилам, которые определяются остальными девятью.

Мобильность является существенной характеристикой любой большой программы моделирования. Нередко программы разрабатывают на одной машине, а затем для проведения рабочего счета переносят на более мощную ЭВМ. Обмениваются программами и лаборатории, имеющие разные вычислительные установки. Даже те программы, которые разрабатывались и пропускались на вполне определенном компьютере, установленном в конкретном месте, нужно со временем переносить на новую машину, в противном случае при замене компьютера их придется выбрасывать.

Чтобы программа была мобильной, необходимо, как минимум, чтобы она была написана в основном или полностью на широко распространенном языке высокого уровня. В данном случае в качестве такого языка нами принят Фортран [292]. Хотя другие языки, как, например, Алгол-68 [386, 265] и Паскаль [220], и лучше отражают метод структурного программирования, Фортран реализован на всех современных компьютерах для научных расчетов. Если в программе необходимо учесть особенности конкретной вычислительной системы, то системно-зависимую часть нужно явным образом выделить, что сведет к минимуму трудности по переносу данной программы. Например, такие характерные особенности вычислительной системы, как соглашения о каналах и обращения к супервизору, можно обойти соответственно с помощью символических номеров каналов и обращений на языке высокого уровня к набору служебных подпрограмм.

Сохранность необходима, поскольку пользователь программы должен быть уверен в получаемых результатах, а гибкость нужна для того, чтобы он имел возможность приспособить программу к новым условиям. Сохранность и мобильность могут быть достигнуты посредством организации системнонезависимого стандартного программного файла (СПФ), неизменность которого может быть проконтролирована по условному коду, сформированному программой контроля [121]. Гибкость можно обеспечить с помощью отдельных редактирующих файлов (рис. 3 1), а также заранее предусматривая ее в проекте управляющей структуры программы.

Все программы моделирования, которые рассматриваются в этой книге, похожи друг на друга тем, что в них имеется стадия инициализации, за которой следуют цикл по времени и стадия завершения. В их состав входят управляющие программы, направляющие ход вычислений, а также программы вывода, осуществляющие передачу информации пользователю. Если сделать

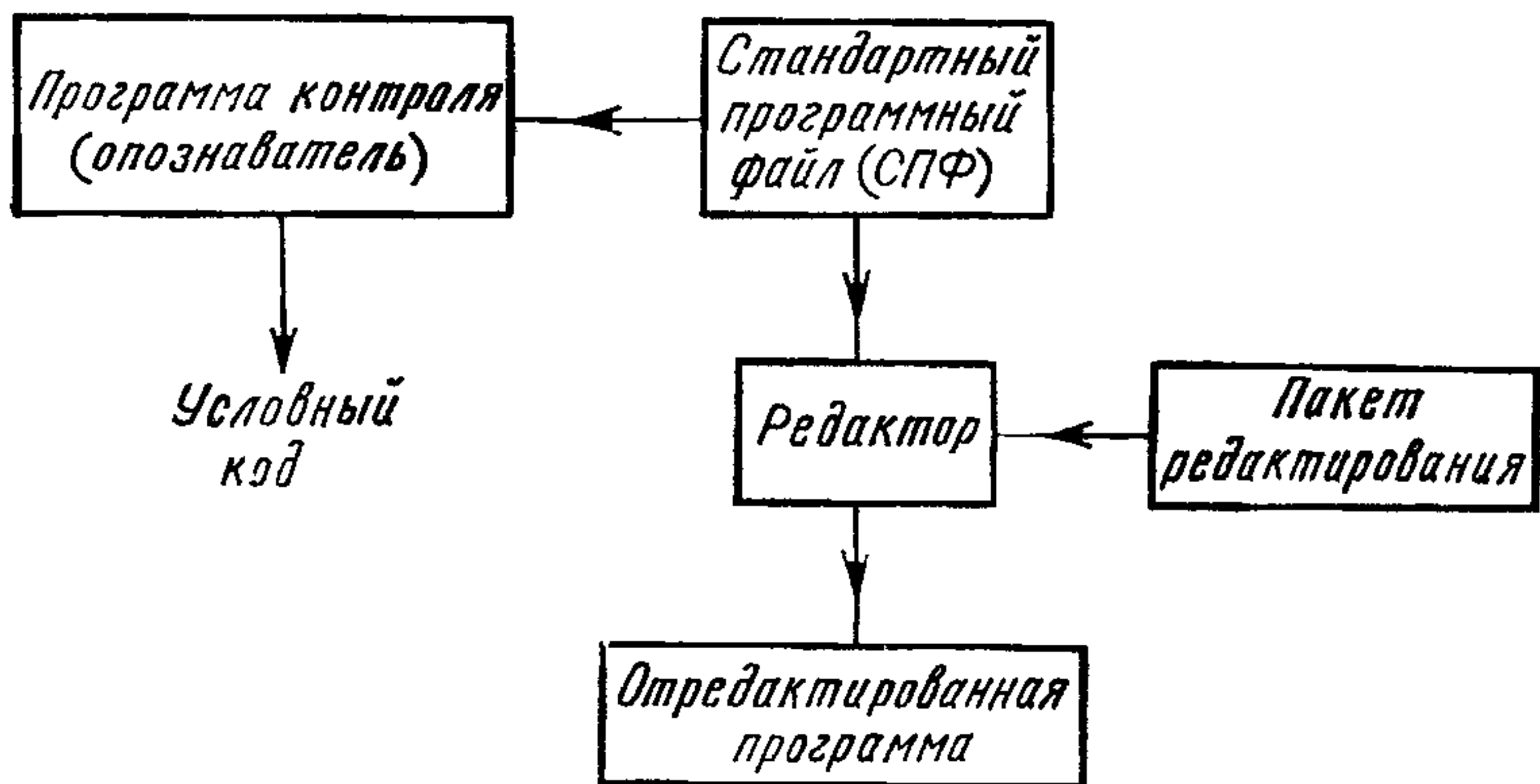


Рис. 3 1. Мобильность и сохранность программ можно обеспечить заданием системнонезависимого программного файла (СПФ), инвариантность которого проверяется программой контроля, генерирующей условный код. Гибкость обеспечивается с помощью пакета редактирования и редактора, формирующего предназначенную для исполнения отредактированную программу [334] (воспроизводится с разрешения *Atom*).

так, чтобы различные программы моделирования имели совершенно одинаковую архитектуру, то такие программы будет гораздо легче понять и использовать. Как будет показано в разд. 3.3, такую стандартизацию архитектуры можно распространить также на операторы (подпрограммы) и операнды (данные) программы.

Стандартизация архитектуры программы позволяет возложить большинство функций управления на универсальную, не зависящую от приложений скелетную программу, на основе которой уже можно разрабатывать программы для конкретных приложений. Благодаря логическим переключателям и точкам диагностики, содержащимся в скелетной программе, автоматически обеспечивается высокая степень гибкости разработанных таким способом прикладных программ. Если же, как это имеет место в системе OLYMPUS, в скелетную программу включены и простые диагностические служебные подпрограммы, то такая скелетная программа обеспечивает отличную среду для эффективной разработки программ.

Если, кроме того, установить стандарт на систему обозначений, форму записи и документацию, это еще больше усилит сходство архитектуры и управляющей структуры прикладных программ, порожденных одной универсальной скелетной программой. В результате повышается удобочитаемость программ, а значит, их легче понимать, использовать и адаптировать. Хотя польза обширной документации и соглашений относительно формы записи и обозначений может показаться сомнительной и на первый взгляд выглядит как напрасная трата времени,

в большинстве случаев справедливо как раз обратное. Благодаря соглашениям о системе обозначений и форме записи гораздо реже допускаются программистские ошибки, а документация способствует хорошему структурированию программ. Более того, если указанные соглашения включены в систему программирования, то можно разработать служебные программы, которые будут *автоматически гечерировать* многие из управляющих разделов программ и приводить к стандартному виду участки, набранные вручную.

3.3. СИСТЕМА ПРОГРАММИРОВАНИЯ OLYMPUS

OLYMPUS — это система программирования на Фортране, которая была разработана на основе спецификаций, приведенных в табл. 3.1. Мы используем в качестве примера именно эту систему по нескольким причинам. Во-первых, в ней воплощена хорошая программистская практика, и система побуждает следовать этой практике. Во-вторых, она широко используется, в том числе в коллективах, ведущих плазменные исследования. И в-третьих, ее можно легко получить из библиотеки журнала *Computer Physics Communications* в Белфасте.

Говоря о системе OLYMPUS, выделим четыре аспекта. Во-первых, и это самое главное, OLYMPUS представляет собой определенную методику эффективного создания хорошо структурированных программ. Остальные три аспекта призваны помочь программисту достигнуть целей этой методики: программное обеспечение предоставляет системнонезависимые средства разработки, система соглашений обеспечивает ясность программы и структуры данных, а дополнительное программное обеспечение частично автоматизирует процесс конструирования программы. Предлагаемая методика и система обозначений могут применяться (и не только к программам на Фортране) независимо от наличия программного обеспечения, хотя последнее и существенно упрощает дело. В этом разделе мы ограничимся кратким описанием соглашений и программного обеспечения, принятых в системе OLYMPUS. С методикой мы познакомимся в следующем разделе, рассматривая пример одномерной плазменной модели.

В системе OLYMPUS программы пишутся на стандартном Фортране, что отвечает требованиям мобильности, простоты использования, а также совместимости с большинством существующих научных программ и программных модулей. В каждом конкретном случае OLYMPUS-программа задается в системнонезависимом формате и называется стандартным программным файлом (СПФ). Сохраняемый СПФ содержит документацию, общие блоки, фортранные подпрограммы и тестовые данные. СПФ-процессор системы OLYMPUS преобразует СПФ

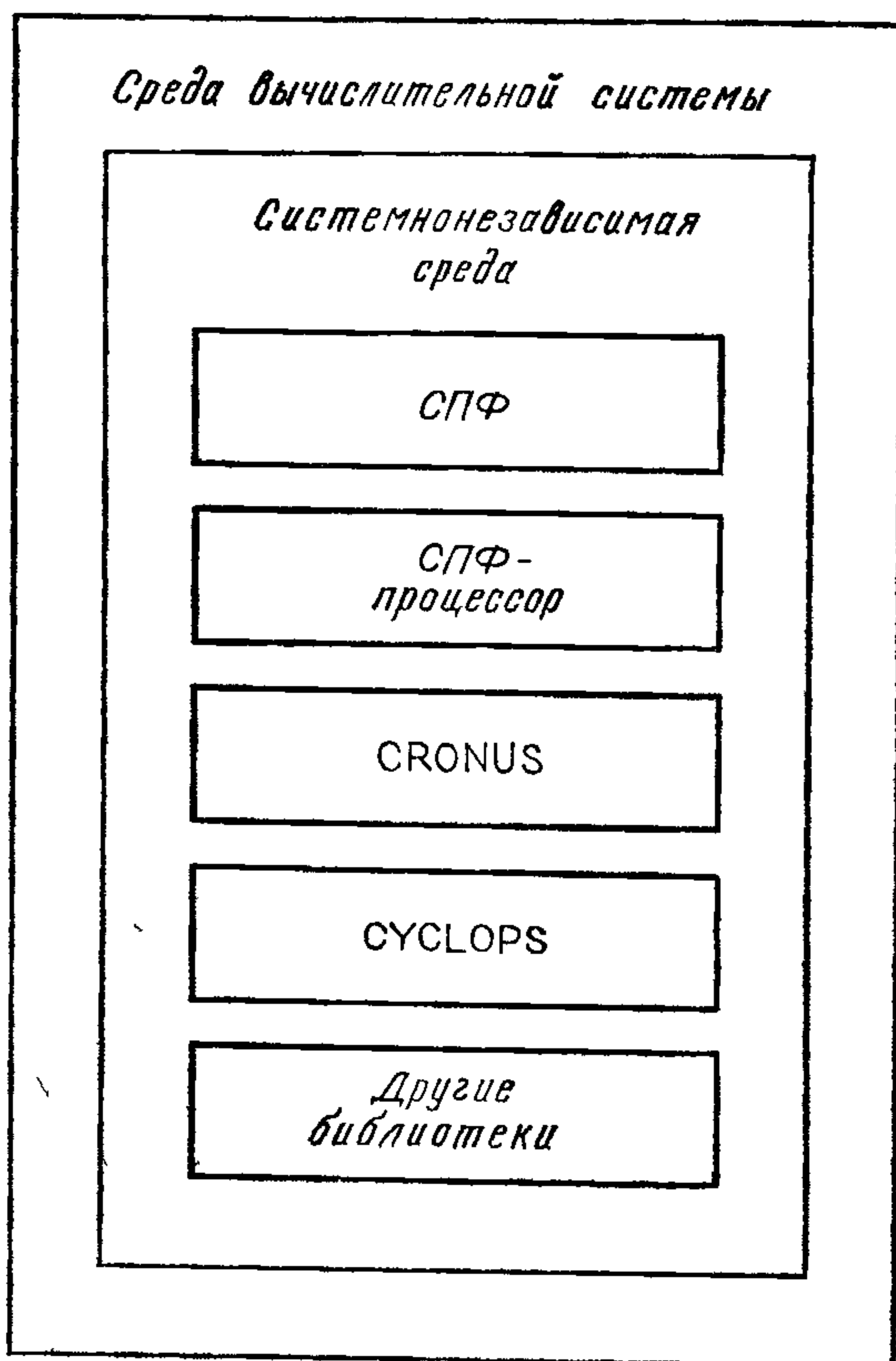


Рис. 3.2 Программное обеспечение системы OLYMPUS представляет собой системнонезависимую среду для программ пользователя. Прикладные программы в виде СПФ можно переносить на другие установки и прогонять на них без изменений.

к виду, необходимому для данной конфигурации ЭВМ, и, исходя из заказанных библиотек системы OLYMPUS, подключает независимые от конкретного приложения программные модули (которые, однако, могут быть машиннозависимыми), необходимые для создания выполняемой программы.

Простое, к тому же написанное большей частью на Фортране, программное обеспечение системы OLYMPUS обеспечивает интерфейс между СПФ и вычислительной системой. СПФ-программы могут свободно переноситься на любую машину, где создана среда системы OLYMPUS. Это схематически показано на рис. 3.2. На этом рисунке прямоугольник «CRONUS» обозначает универсальную скелетную программу, определяющую базовую архитектуру всех OLYMPUS-программ, а прямоугольник «CYCLOPS» обозначает библиотеку вспомогательных программ системы OLYMPUS, предназначенных для ускорения процесса разработки программ. «Другие библиотеки» обозначает как программное обеспечение системы OLYMPUS, применяемое для автоматизации конструирования программ, так и растущий запас имеющихся прикладных программ и программных модулей.

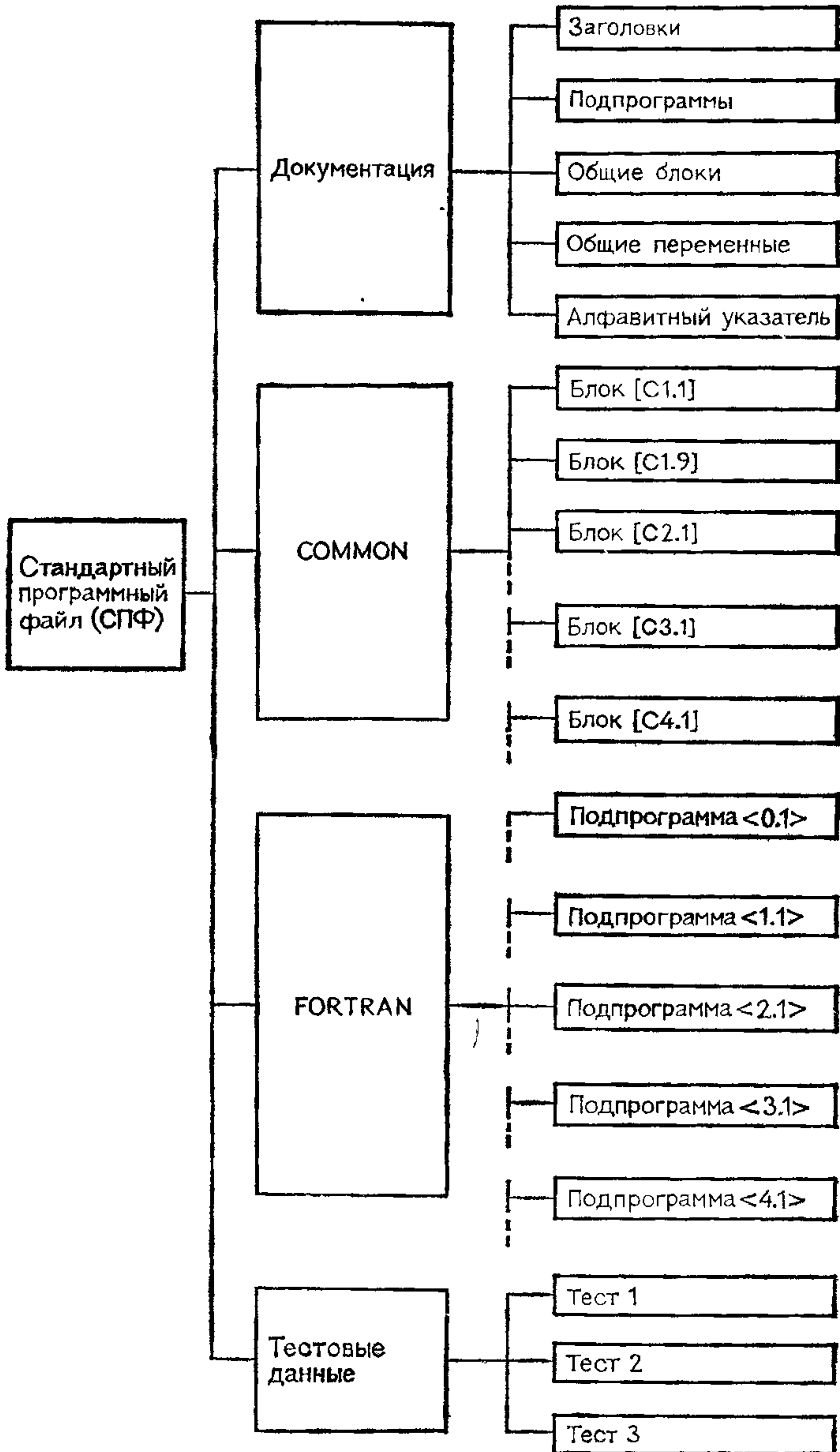


Рис. 33. СПФ делится на четыре основные секции, а каждая секция в свою очередь на модули. В программной колоде секциям и модулям предшествуют системнонезависимые управляющие карты. Общим блокам и фортран-ным подпрограммам присваиваются десятичные номера по системе, описанной в табл. 3.2.

На рис. 3.3 изображена структура стандартного программного файла (СПФ). СПФ делится на четыре основных секции: документация, общие блоки («COMMON»), фортранные подпрограммы («FORTRAN») и тестовые данные. Каждая секция делится на подсекции. Всем основным секциям и подсекциям предшествуют системнонезависимые управляющие инструкции, начинающиеся соответственно с «C/» и «C/MODULE». С помощью этих управляющих инструкций СПФ-процессор преобразует СПФ к той форме представления программы, которая требуется для исполнения на данной ЭВМ.

Секция «Документация» состоит из титульного модуля, идентифицирующего программы, и набора модулей, где перечислены номера подпрограмм, общих блоков, переменных каждого общего блока и всех общих переменных в алфавитном порядке. Каждый модуль секции «COMMON» содержит один документированный общий блок, а каждый модуль секции «FORTRAN» — одну фортранную подпрограмму. Подпрограммы рассматриваются как операторы, а данные из помеченных общих блоков считаются операндами. Каждая подпрограмма и каждый общий блок несут по возможности вполне определенную смысловую нагрузку, что отражается с помощью приписываемых им десятичных номеров. Подпрограммы делятся на классы и снабжаются десятичными номерами (класс. подпрограмма). Общие блоки делятся на группы и обозначаются [группа. блок]. Классы подпрограмм и группы общих блоков перечислены в табл. 3.2.

Таблица 3.2. Деление подпрограмм на классы и общих блоков на группы в системе OLYMPUS

Инструкции	0 Управление
	1 Пролог
	2 Расчет
Классы подпрограмм	3 Вывод
	4 Эпилог
	5 Диагностика
	U Утилиты
Данные	1 Общие данные OLYMPUS
	2 Физическая задача
Группы помеченных общих блоков	3 Численная схема
	4 Служебные
	5 Ввод-вывод и диагностика
	6 Обработка текста

Источник: [333]; с разрешения Computer Phys. Commn.

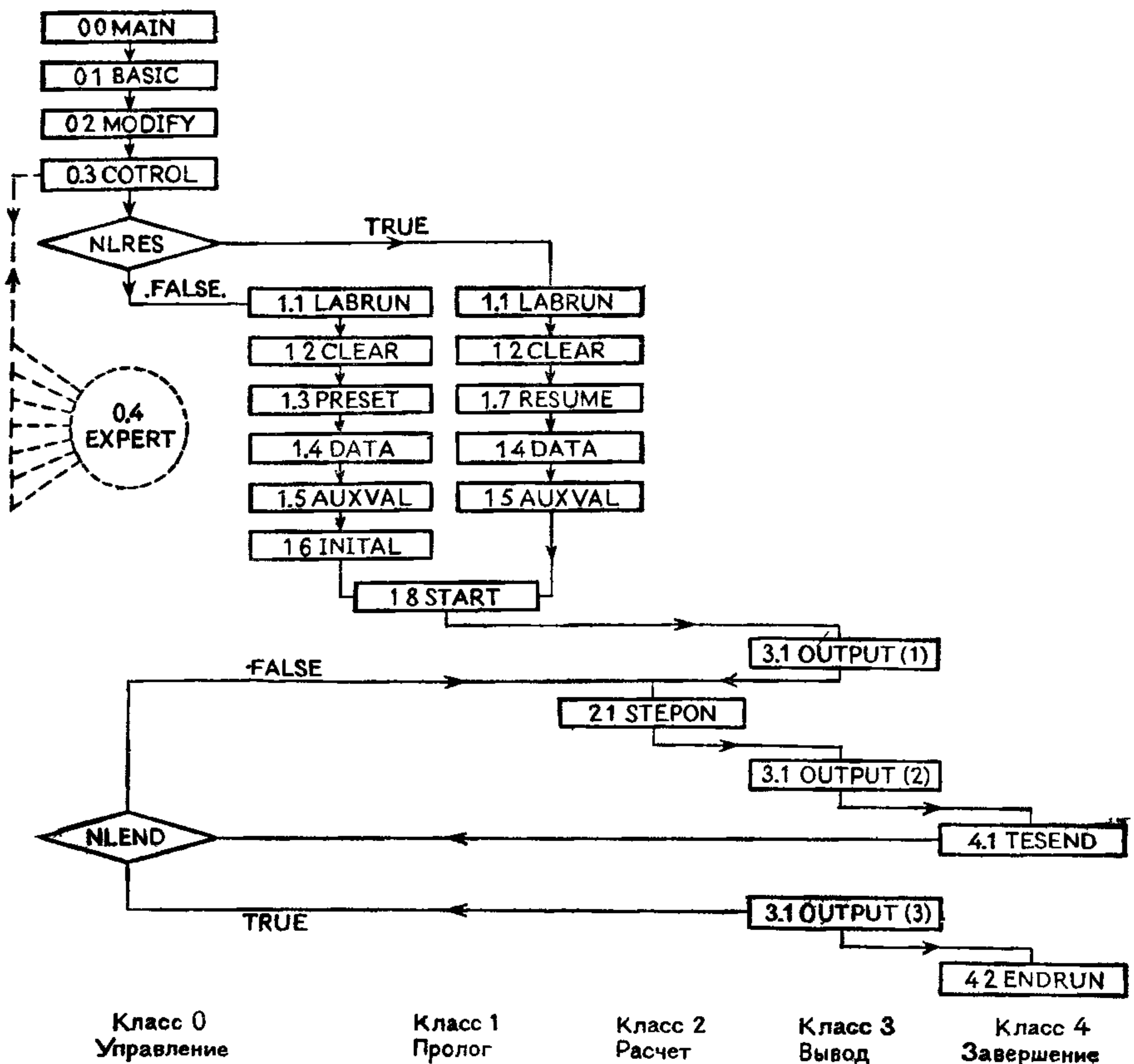


Рис. 3.4 Блок-схема скелетной программы CRONUS. Назначение подпрограмм указано в табл. 3.3 ([80], с любезного разрешения Computer Physics Communications.)

Стандартный скелетный СПФ системы OLYMPUS, называемый CRONUS, определяет базовую архитектуру всех OLYMPUS-программ. CRONUS (рис. 3.4 и табл. 3.3) содержит универсальные общие блоки управления ([C1.1] COMBAS) и диагностики ([C1.9] COMDDP), универсальные (класса 0) управляющие подпрограммы ($\langle 0.0 \rangle$ — $\langle 0.3 \rangle$), стандартную диагностическую структуру (рис. 3.6), а также набор фиктивных подпрограмм со стандартными именами и нумерацией. На рис. 3.5 приведен листинг стандартного варианта управляющей подпрограммы $\langle 0.3 \rangle$ COTROL. В любой «настоящей» OLYMPUS-программе фиктивные подпрограммы, входящие в CRONUS, заменяются по мере необходимости определенными версиями СПФ с теми же именами и десятичной нумерацией. Так, например, програм-

```

C/ MODULE COS3
C
      SUBROUTINE COTROL
C
C 0.3 Управление счетом
C
C      Version 2a      1/8/73      KVR/JPC      Culham
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C-----
C
      DATA ICLASS,ISUB/0,3/
      CALLMESSAGE(48H      0.3 НАЧАЛО КОНТРОЛЯ СЧЕТА
C-----
CL      1.      Пролог
C
      IF(NLRES) GO TO 170
C
C      A.      Новый вариант
C
CL      1.1      Маркировка расчета
110      CALL LABRUN
                        CALL EXPERT(ICLASS,ISUB,1 )
C
CL      1.2      Чистка переменных и массивов
120      CALL CLEAR
                        CALL EXPERT(ICLASS,ISUB,2 )
C
CL      1.3      Задание значений, принимаемых по умолчанию
130      CALL PRESET
                        CALL EXPERT(ICLASS,ISUB,3 )
C
CL      1.4      Задание данных конкретного варианта
140      CALL DATA
                        CALL EXPERT(ICLASS,ISUB,4 )
C
CL      1.5      Задание вспомогательных величин
150      CALL AUXVAL
                        CALL EXPERT(ICLASS,ISUB,5 )
C
CL      1.6      Задание физического начального состояния
160      CALL INITAL
                        CALL EXPERT(ICLASS,ISUB,6 )
      GO TO 180
C
C      B.      Продолжение сохраненного варианта
C
CL      1.7      Поиск записи, изменение нужных параметров
170      CONTINUE
                        CALL EXPERT(ICLASS,ISUB,7 )
C      Маркировка продолжения счета
      CALL LABRUN

```

Рис. 3.5 Исходный вид программы <0.3> COTROL из CRONUS, принимаемый по умолчанию ([80], с разрешения Computer Physics Communications).

```

CALL EXPERT(ICLASS,ISUB,8 )
C      Чистка переменных и массивов
      CALL CLEAR
CALL EXPERT(ICLASS,ISUB,9 )
C      Считывание записи и печать информации о текущем файле
      CALL RESUME
CALL EXPERT(ICLASS,ISUB,10)
C      Чтение всех необходимых новых данных
      CALL DATA
CALL EXPERT(ICLASS,ISUB,11)
C      Соответствующее изменение вспомогательных переменных
      CALL AUXVAL
CALL EXPERT(ICLASS,ISUB,12)
C
C      С.      Подготовительные действия
C
CL      .1.8      Запуск или продолжение расчета
180     CALL START
CALL EXPERT(ICLASS,ISUB,13)
C      Начальная выдача
      CALL OUTPUT(1)
CALL EXPERT(ICLASS,ISUB,14)
C
C
C-----
CL      2.      Расчет
C
CL      2.1      Шаг расчета
210     CALL STEPON
CALL EXPERT(ICLASS,ISUB,15)
C
C-----
CL      3.      Вывод
C
CL      3.1      Периодические выдачи
310     CALL OUTPUT(2)
CALL EXPERT(ICLASS,ISUB,16)
C
C-----
CL      4.      Эпилог
C
CL      4.1      Проверка на окончание расчета
410     CALL TESEND
CALL EXPERT(ICLASS,ISUB,17)
      IF(.NOT.NLEND) GO TO 210
C      Заключительная выдача
      CALL OUTPUT(3)
CALL EXPERT(ICLASS,ISUB,18)
C
CL      4.2      Завершение расчета
420     CALL ENDRUN
C
      RETURN
      END

```

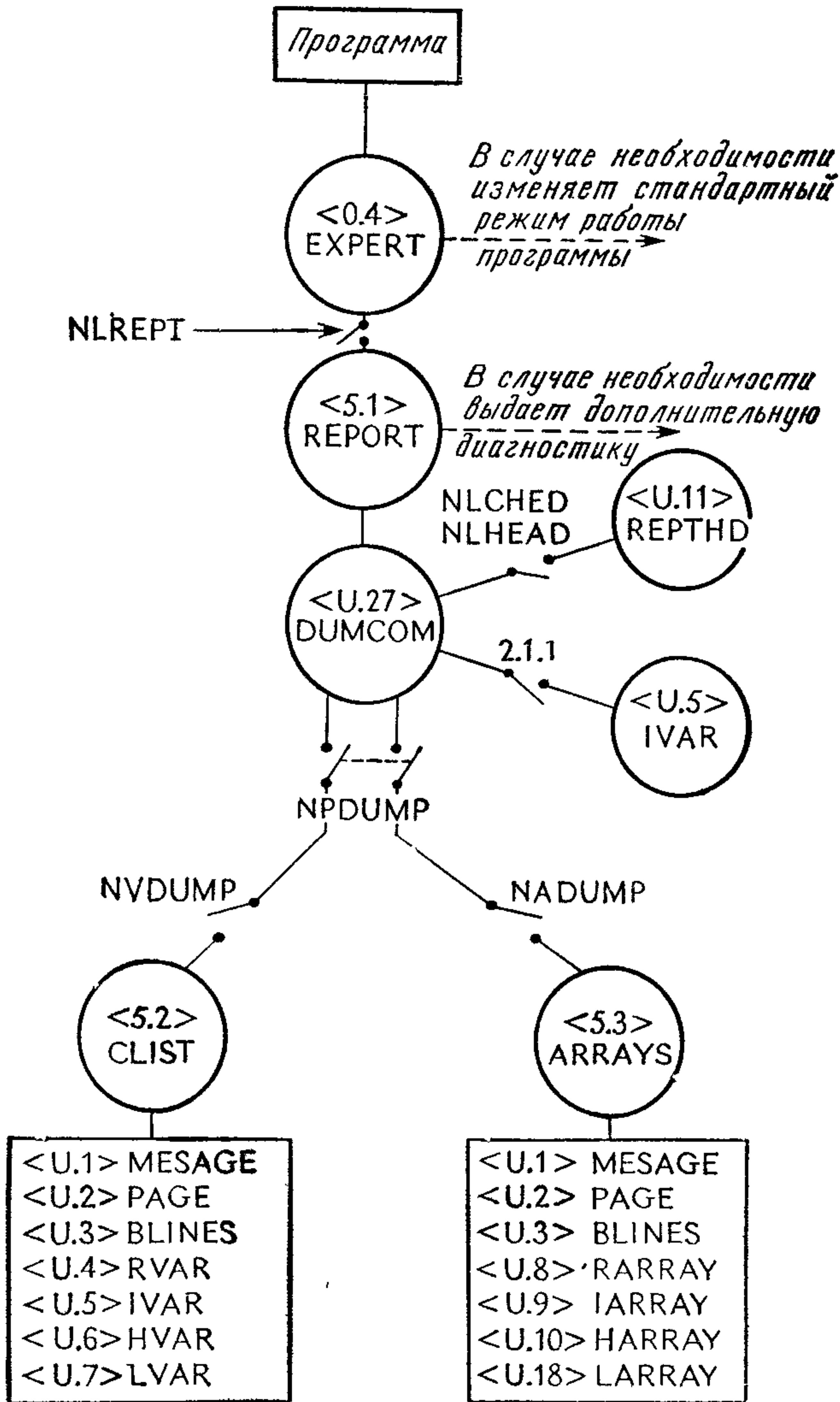


Рис. 3.6. Стандартные средства диагностики и управления в системе OLYMPUS ([80], с разрешения Computer Physics Communications.)

мист мог бы предусмотреть свою собственную версию подпрограммы <2.1> STEPON, которая управляет циклом по времени и вызывает другие подпрограммы, <2.2>, <2.3> и т. д., для фактического проведения вычислений. Подпрограмма <3.1> OUTPUT организует выдачу результатов и т. д.

Таблица 3.3. Программа CRONUS: имена подпрограмм и идентифицирующие номера

Имя	Номер	Формальные параметры	Назначение	Статус
Класс 0 Главное управление				
(MAIN)	0.0		Главная фортранная подпрограмма	P
BASIC	0.1		Инициализирует основные управляющие данные	P
MODIFY	0.2		Если необходимо, изменяет основные данные	D
COTROL	0.3		Управляет счетом	P
EXPERT	0.4	KCLASS, KSUB, KPOINT	Изменяет стандартный режим работы программы	M
Класс 1 Пролог				
LABRUN	1.1		Маркирует прогоняемый вариант	D
CLEAR	1.2		Очищает переменные и массивы	D
PRESET	1.3		Задаёт принятые по умолчанию значения	D
DATA	1.4		Определяет данные конкретного прогона	D
AUXVAL	1.5		Задаёт вспомогательные величины	D
INITAL	1.6		Определяет физические начальные условия	D
RESUME	1.7		Возобновляет счет с предыдущей записи	D, M
START	1.8		Начинает расчет	D
Класс 2 Расчет				
STEPON	2.1		Ведение расчета	D
Класс 3 Вывод				
OUTPUT	3.1	K	Управляет выдачей результатов	D

Таблица 3.3 (продолжение)

Имя	Номер	Формальные параметры	Назначение	Статус
Класс 4 Эпилог				
TESEND	4.1		Проверка на окончание расчета	M
END RUN	4.2		Завершение расчета	M
Класс 5 Диагностика				
REPORT	5.1	KCLASS, KSUB, KPOINT	Управляет диагности- кой	M
CLIST	5.2	KGROUP, KBLOCK	Печатает переменные общих блоков	D
ARRAYS	5.3	KGROUP, KBLOCK	Печатает массивы об- щих блоков	D

Кодировки статуса

- D Фиктивные программы; заменяются программистом на «настоящие»
M Может быть оставлена программистом в том же виде или модифицирована в зависимости от потребностей.
P Постоянная управляющая программа, изменяться не должна.

Источник [80], с разрешения Computer Phys. Commun

Таблица 3.4. Соглашения по начальным буквам, принятые в системе OLYMPUS

	Вещественные, комплексные	Целые (и коллеритовы)	Логические
Формальные параметры подпрограмм	P	K	KL
Имена переменных и массивов общих блоков	A—H, O, Q—Y	L, M, N	LL, ML, NL
Имена локальных переменных и массивов	Z	I	IL
Индексы цикла		J	

Источник [80]; с разрешения Computer Phys. Commun.

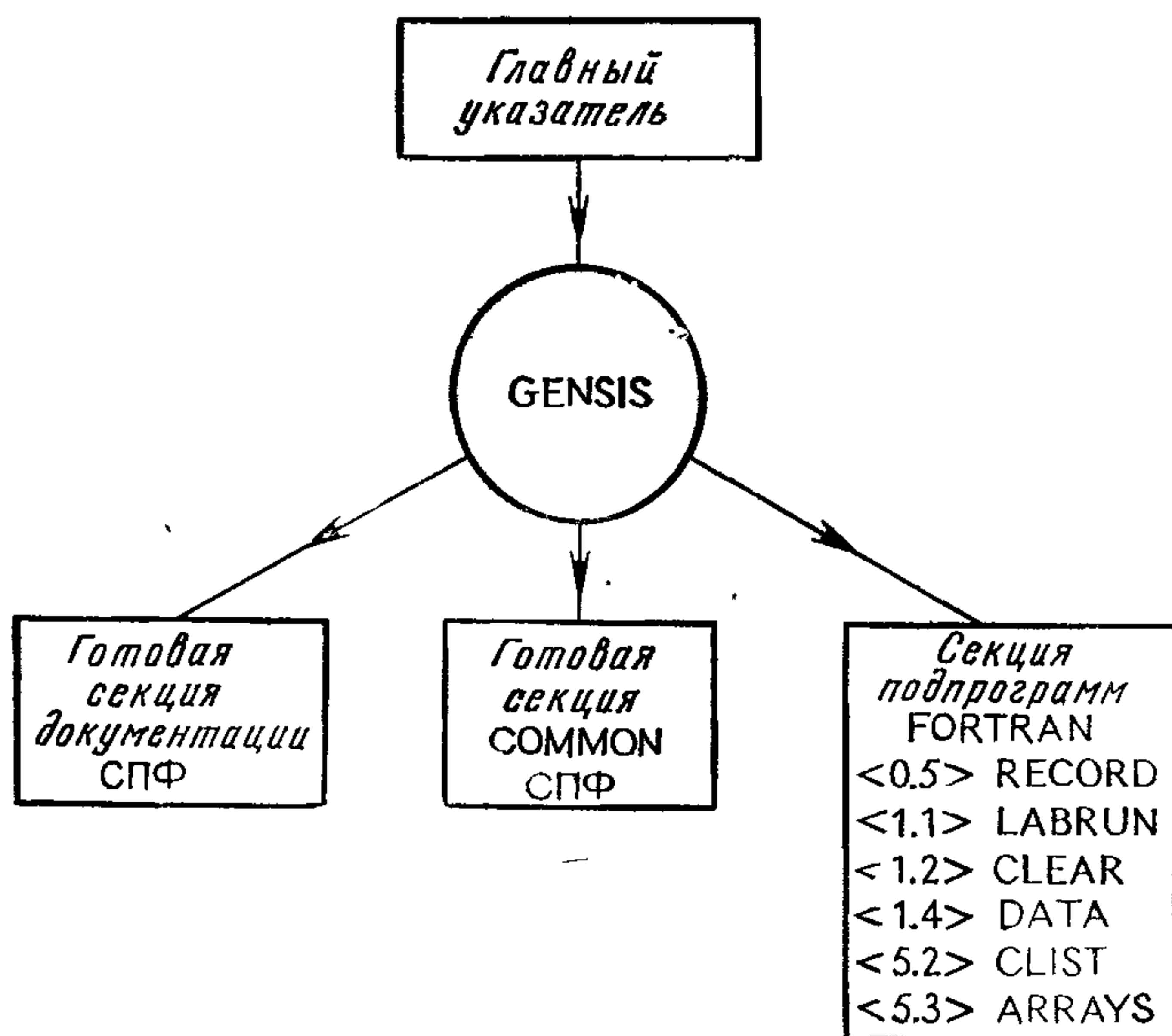


Рис. 3.7. Автоматический генератор программ GENSIS на основе компактного главного указателя создает секцию документации, секцию общих блоков, подпрограммы вывода диагностики (5.2) и (5.3), а также некоторые подпрограммы класса 1.

Строгие правила, принятые в системе OLYMPUS, распространяются также на систему обозначений, форму записи и документацию подпрограмм и общих блоков. Подпрограммы делятся на десятично занумерованные секции и подсекции. Чтобы смысл программы был понятнее, используются мнемонические имена, т. е. символьные обозначения вместо цифровых, и соглашения по первым буквам, которые приведены в табл. 3.4. Для повышения удобочитаемости программы применяются разделительные линии между секциями и ступенчатая форма записи с разными отступами в заголовках, подзаголовках, комментариях и инструкциях.

Четко определенная структура OLYMPUS-программ дает возможность почти полностью ликвидировать трудоемкую черновую часть работы по программированию. Автоматический генератор программ GENSIS [211] по компактно описанной структуре данных (главный указатель, см. рис. 3.8) генерирует секцию документации СПФ и секцию COMMON, а также ряд подпрограмм диагностики и инициализации. Сказанное поясняется на рис. 3.7. Кроме того, имеется также вспомогательная программа COMPOS [210], редактирующая подпрограммы; она преобразует текст подпрограмм к принятой в OLYMPUS стан-

Таблица 3.5. Служебные подпрограммы CYCLOPS

Они предназначены для выполнения четырех функций: (1) выдачи, (2) обработки массивов, (3) хронометрирования и (4) диагностики.

Номер	Имя	Параметры	Назначение
1	RVAR	KNAME, PVALUE	Печать имени вещественной переменной и ее значения.
2	IVAR	KNAME, KVALUE	Печать имени целой переменной и ее значения
3	HVAR	KNAME, KVALUE	Печать имени холлеритовой переменной и ее значения.
4	LVAR	KNAME, KLVAL	Печать имени логической переменной и ее значения.
5	RARRAY	KNAME, PA, KDIM	Печать вещественного массива.
6	IARRAY	KNAME, KA, KDIM	Печать целого массива.
7	HARRAY	KNAME, KA, KDIM	Печать холлеритового массива.
8	LARRAY	KNAME, KLA, KDIM	Печать логического массива.
9	RARRAY2	KNAME, PA, KDIMX, KX, KY	Печать двумерного вещественного массива.
10	RESETR	PA, KDIM, PVALUE	Засылка в вещественный массив заданного значения.
11	RESETI	KA, KDIM, KVALUE	Засылка в целый массив заданного значения.
12	RESETH	KA, KDIM, KVALUE	Засылка в холлеритовый массив заданного значения.
13	RESETL	KLA, KDIM, KLVAL	Засылка в логический массив заданного значения.
14	SCALER	PA, KDIM, PC	Масштабирование вещественного массива умножением на вещественную величину.
15	SCALEI	KA, KDIM, KC	Масштабирование целого массива умножением на целую величину.
16	COPYR	PA1, K1, PA2, K2, KDIM	Копирование одного вещественного массива в другой.
17	COPYI	KA1, K1, KA2, K2, KDIM	Копирование одного целого массива в другой.
18	SIGNR	PA, KDIM	Изменение знака вещественного массива.
19	SIGNI	KA, KDIM	Изменение знака целого массива.
20	PAGE		Выборка новой страницы в канал вывода.
21	BLINES	K	Печать K пустых строк
22	MESSAGE	KMESS	Печать 48-символьного сообщения.
23	REPTH0	KCLASS, KSUB, KPOINT	Печать заголовка для диагностического сообщения.
24	RUNTIM		Обновление времени ЦП и его печать.
25	DAYTIM		Печать даты и времени.
26	JOBTIM	PTIME	Запрос заказанного времени работы.
27	DUMCOM	KCLASS, KSUB, KPOINT	Распечатка выбранного общего блока.

дартной форме записи и десятичной нумерации. Пример использования COMPOS показан на рис. 3.9.

Неотъемлемой частью любой OLYMPUS-программы является обширная встроенная диагностика. Подпрограммы диагностики, отражающие специфику конкретной прикладной задачи, можно легко написать с помощью библиотеки обслуживающих программ CYCLOPS (табл. 3.5). Каждая программа, когда она пишется, оснащается средствами контроля, после чего стандартную диагностику, доступ к которой осуществляется путем обращений к подпрограмме <0.4> EXPERT, можно включать и отключать с помощью управляющих логических переменных из общего блока [C1.9] COMDDP.

3.4. ПРОГРАММА ES1D1V

Итоги всех этапов проектирования ES1D1V, проводимых последовательно от физического явления до численного алгоритма, можно представить в виде: (1) списка дифференциальных уравнений, (2) списка разностных уравнений и масштабных множителей и (3) перечня всех шагов численного алгоритма. Эти списки вместе со стандартной структурой и переменными, предусмотренными в OLYMPUS, служат отправным пунктом методики системы OLYMPUS-программирования сверху вниз с последовательным уточнением.

Приступая к разработке программы, первым делом необходимо выяснить, нет ли подходящей библиотечной программы или программных модулей, которыми можно воспользоваться для решения искомой задачи. В самом деле, если бы такие программы нашлись, они бы могли послужить отправной точкой для разработки данной программы или повлиять на нее. Например, если бы имелся подходящий модуль для решения уравнения Пуассона, то к указанным выше спискам (1)—(3) добавился бы список (4), содержащий таблицы имен подпрограмм, имен общих блоков и имен общих переменных. Таким способом удастся предотвратить излишние конфликтные ситуации с именами.

3.4.1. Управляющая структура программы

Как только списки составлены, а мы предлагаем читателю проделать это самостоятельно в качестве упражнения, сверяясь с материалом гл. 2, можно перейти к проектированию структуры всей программы. Поскольку ES1D1V есть не что иное, как программа решения краевой задачи Коши, то в качестве блок-схемы ES1D1V подойдут та управляющая структура и имена основных подпрограмм, которые использованы в CRONUS (рис. 3.5). По мере того как программа будет развиваться, фик-

тивные подпрограммы в CRONUS будут заменяться и дополняться рабочими подпрограммами ES1D1V. В более крупных программах ряд входящих в CRONUS подпрограмм, как, например, <2.1> STEPON, превратились бы просто в управляющие подпрограммы. STEPON обращалась бы к последовательности подпрограмм <2.2>, <2.3> и т. д., которые и осуществляли бы собственно численный счет. В данном же простом примере (см. рис. 3.10) основная часть вычислений производится в

```
PROGRAM ES1D2V
TITLE 1-D ELECTROSTATIC PLASMA MODEL
AUTHOR J.W.Eastwood and R.W.Hockney
VERSION 1   RWH/JWE                               Reading Univ.   1979
```

```
УКАЗАТЕЛЬ ПОДПРОГРАММ
2.2 POTSOL Нахождение потенциала
```

```
РАЗМЕРЫ МАССИВОВ
MAX1=256,NGMAX
MAX2=2048,NPMAX
```

```
УКАЗАТЕЛЬ ОБЩИХ ПЕРЕМЕННЫХ
```

```
§ 2.1 Основные физические величины
```

```
COMMON/COMPRQ/
R EPSO *Диэлектрическая проницаемость вакуума (СИ)
R ELMASS *Масса электрона (СИ)
R ELCHAG *Заряд электрона (СИ)
R API Pi
R BOLTZK *Постоянная Больцмана (СИ)
R DEBLNO Начальная дебаевская длина (СИ)
R VLMETR *Длина области (СИ)
R DENBAK *Плотность фона (СИ)
R FREQPO Начальная плазменная частота (СИ)
R TODEGK *Начальная температура (град К)
R VTHERO Начальная тепловая скорость (СИ)
RA ELPOT(MAX1) Электрический потенциал
RA CHARG(MAX1) Плотность заряда
RA XPOS(MAX2) Положения частиц
RA VEL(MAX2) Скорости частиц
R VDRIFT Направленная скорость электронов (СИ)
```

```
§ 2.2 Вспомогательные физические величины
```

```
COMMON/COMSPQ/
R BOXLEN Длина области
R DELCHG Вклад одной частицы в плотность заряда
R QBAKG Зарядовая плотность фона
R ELKIN Кинетическая энергия
R ENPOT Потенциальная энергия
```

Рис 3.8. Фрагмент входных данных GENESIS для программы ES1D1V. Результат работы представлен в листинге программы на рис. 3.10.

```

SUBROUTINE INITIAL
*
* 1.6 Задание физического начального состояния
*
* Примечание: XPOS и VEL - в безразмерных единицах
*INSERT COMBAS,COMDDP,COMPPQ,COMSPQ
*INSERT COMNUM,COMSCA,COMНОК,COMOUT
DATA ICLASS,ISUB/1,6/
CALL EXPERT
IF(NLOMT1(ISUB)) RETURN
C
C-----
CL          1.          Двухчастичный тест
C
          IF(NCASE.NE.1) GO TO 10
          XPOS(1) = X2PT1
          XPOS(2) = X2PT2
          VEL(1) = V2PT1
          VEL(2) = V2PT2
          RETURN
#Пучок с тепловым разбросом, плотность=const
10 CONTINUE
IF(NCASE.NE.2) GO TO 5
##Формирование положений
CALL RANPOS(XPOS,1,NP,XMIN,XMAX)
##Формирование скоростей
ZVDR = VDRIFF / SCVEL
ZVTH = VTHERO / SCVEL
CALL THERDI(VEL,1,NP,ZVDR,ZVTH)
RETURN
#Два "горячих" пучка
5
IF(NCASE.NE.3) GO TO 15
CALL RANPOS(XPOS,1,NP,XMIN,XMAX)
ZVDR1 = VDR1 / SCVEL
ZVDR2 = VDR2 / SCVEL
ZVTH1 = SQRT(BOLTZK * TST1 / ELMASS) / SCVEL
ZVTH2 = SQRT(BOLTZK * TST2 / ELMASS) / SCVEL
CALL THERDI(VEL,1,NST1,ZVDR1,ZVTH1)
CALL THERDI(VEL,NST2,NP,ZVDR2,ZVTH2)
RETURN
#Другие случаи
C
15
CALL EXPERT
*-----
RETURN
          END

```

Рис. 3.9. Пример, иллюстрирующий работу наборщика программ COMPOS. Показаны входные данные COMPOS (слева) и результат работы (справа).

<2.1> STEPON, которая обращается только к стандартной диагностической подпрограмме <0.4> EXPERT и подпрограмме решения уравнения Пуассона <2.2> POTSOL.

3.4.2. Главный указатель

Задача третьего этапа разработки программы состоит в получении главного указателя данных, для чего надо еще раз

```

C
      SUBROUTINE INITIAL
C
C 1.6 Задание физического начального состояния
C
C     Примечание: XPOS и VEL - в безразмерных единицах
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
      DATA          ICLASS,  ISUB/1,  6/
                  CALL EXPERT(ICLASS, ISUB, 1)
      IF(NLOMT1(ISUB)) RETURN
C
C-----
CL          1.          Двухчастичный тест
C
      IF(NCASE.NE.1) GO TO 200
      XPOS(1) = X2PT1
      XPOS(2) = X2PT2
      VEL(1) = V2PT1
      VEL(2) = V2PT2
      RETURN
C
C-----
CL          2.          Пучок с тепловым разбросом, плотность=const
C
200      CONTINUE
      IF(NCASE.NE.2) GO TO 300
C
      CL          2.1      Формирование положений
      CALL RANPOS(XPOS, 1, NP, XMIN, XMAX)
C
      CL          2.2      Формирование скоростей
      ZVDR = VDRIPT / SCVEL
      ZVTH = VTHERO / SCVEL
      CALL THERDI(VEL, 1, NP, ZVDR, ZVTH)
      RETURN
C
C-----
CL          3.          Два "горячих" пучка
C
300      CONTINUE
      IF(NCASE.NE.3) GO TO 400
      CALL RANPOS(XPOS, 1, NP, XMIN, XMAX)
      ZVDR1 = VDR1 / SCVEL
      ZVDR2 = VDR2 / SCVEL
      ZVTH1 = SQRT(BOLTZK * TST1 / ELMASS) / SCVEL
      ZVTH2 = SQRT(BOLTZK * TST2 / ELMASS) / SCVEL
      CALL THERDI(VEL, 1, NST1, ZVDR1, ZVTH1)
      CALL THERDI(VEL, NST2, NP, ZVDR2, ZVTH2)
      RETURN
C
C-----
CL          4.          Другие случаи
C
400      CONTINUE
                  CALL EXPERT(ICLASS, ISUB, 2)
C-----
      RETURN
      END

```


пробежаться по спискам (1)—(4). Первый уровень организации данных обеспечивается группами общих блоков системы OLYMPUS, перечень которых приведен в табл. 3.2. В каждой из этих групп нужно определить общие блоки, а затем и переменные, используемые внутри программы.

Универсальные управляющие блоки [C1.1] COMBAS и [C1.9] COMDDP уже описаны в CRONUS. COMBAS содержит основную информацию системы OLYMPUS, а COMDDP содержит параметры для диагностики и разработки. Переменные в группах 2—4 почти в точности повторяют переменные, фигурирующие в математической модели (список (1)), дискретной постановке задачи (список (2)) и численных алгоритмах (список (3)) соответственно.

В списке (1) (см. уравнения (3.5)—(3.17)) представлены физические константы и переменные. Эти величины соответствующим образом сгруппированы, что достигается введением, например, общего блока [C2.1] COMPRQ для основных физических величин и [C2.2] COMSPQ для вспомогательных. В дискретной формулировке задачи фигурируют разностные сетки, размеры ячеек, шаг по времени, макрочастицы и пр., а также масштабные множители. Отсюда — общие блоки [C3.1] COMNUM и [C3.2] COMSCA. Еще могут, вероятно, понадобиться дополнительные блоки для служебных переменных ([C4.1] COMНОК) и вывода результатов ([C5.1] COMOUT). В более крупных программах общие блоки удобно выбирать с учетом модульности. А именно все переменные, которые используются для связи между группой подпрограмм (т. е. вспомогательные физические величины, переменные, связанные с численным методом и служебные), выполняющих свою вполне определенную задачу (т. е. внутри модуля), логично выделить в самостоятельные общие блоки, что минимизирует вероятность случайной порчи информации.

Переменные и массивы группы 1 общих блоков, т. е. [C1.1] COMBAS и [C1.9] COMDDP, описываются и инициализируются в фиктивной программе CRONUS. Их имена, типы, размерность, назначение и значения, принимаемые по умолчанию, сведены в табл. 3.6. Величина ALTIME, первая в [C1.1] COMBAS, формируется служебной подпрограммой <U.17> JOBTIM, обращение к которой заранее встроено в подпрограмму нулевого класса <0 0> MAIN, входящую в CRONUS (рис. 3.4). Ее значение равно ресурсу времени ЦП, оставшемуся на момент начала шага «счет». В сочетании с CPTIME (значение которой обновляется и распечатывается программой <U.12> RUNTIM) эту величину можно использовать в <4.1> TESEND с тем, чтобы прекратить работу, если окажется, что время ЦП будет исчерпано до завершения запланированного числа счетных шагов. Третья переменная STIME, включенная в [C1.1], также связана с таймером

Таблица 3.6. Главные указатели стандартных общих блоков OLYMPUS
[C1.1] COMBAS и [C1.9] COMDDP

Имя	Тип	Размер	Назначение	Предварительно установленное значение
ALTIME	R		Время ЦП, выделенное для работы (с)	Вызов супервизора
CPTIME	R		Время ЦП, израсходованное к данному моменту (с)	0.0
NLEDGE	I		Канал для записей рестарта	30
NLEND	L		.TRUE если надо завершить счет	.FALSE.
NLRES	L		.TRUE если счет продолжается	.FALSE.
NONLIN	I		Канал для интерактивного ввода-вывода	1
NOUT	I		Текущий канал вывода	NPRINT
NPRINT	I		Канал для выдачи на печать	99
NREAD	I		Канал для ввода карт (или их эквивалентов)	99
NREC	I		Текущий номер записи	1
NRESUM	I		Возобновление счета с записи по этому каналу	NLEDGE
NSTEP	I		Текущий номер шага	0
STIME	R		Начальное время (с)	0.0
[C1.1] COMBAS				
LABEL1	IA	12	Метки, используемые для описания варианта, устанавливаются программой в LABRUN	
LABEL2	IA	12		
LABEL3	IA	12		
LABEL4	IA	12		
LABEL5	IA	12	Метки, доступные программисту	Пробел
LABEL6	IA	12		
LABEL7	IA	12	Метки, зарезервированные для системного использования	
LABEL8	IA	12		
NDIARY	I		Канал для протокола	NPUNCH
NIN	I		Текущий канал ввода	NREAD
NPUNCH	I		Канал для вывода отперфорированных карт (или их эквивалента)	98
NRUN	I		Максимальное число шагов	1
[C1.9] COMDDP				
MAXDUM	I		Максимальный размер массивов дампа	20
MXDUMP	I		Фактический размер массивов дампа	10
NADUMP	IA	20	Коды для дампов массивов	0
NCLASS	I		Класс в самом последнем сообщении	0
NPDUMP	I		Коды для точек дампа	0
NPOINT	IA	20	Точка в самом последнем сообщении	0
NSUB	I		Подпрограмма в самом последнем сообщении	1

Таблица 3.6 (продолжение)

Имя	тип	Размер	Назначение	Предвари- тельно установленное значение
NVDUMP	IA	20	Коды для переменных дампа	0
NLCHED	L		.TRUE. если нужны заголовки со- общений для центрального класса	.FALSE.
NLHEAD	LA	9	.TRUE. если нужны заголовки со- общений для классов 1—9	.FALSE.
NLOMT 1	LA	50	.TRUE. если нужно отключить под- программу класса 1	.FALSE.
NLOMT 2	LA	50	.TRUE. если нужно отключить под- программу класса 2	.FALSE.
NLOMT 3	LA	50	.TRUE. если нужно отключить под- программу класса 3	.FALSE.
NLREPT	L		.TRUE. если сообщения нужны	.FALSE.

Источник [80]; с разрешения Computer Phys Comm.

и предназначена для запоминания опорной точки отсчета времени, что необходимо на некоторых машинах. Назначение логических переключателей NLEND и NLRES должно быть ясно из рис. 3.4, а различных каналов ввода-вывода — из их названий, причем для нас в данном случае основное значение имеют каналы NREAD, NOUT и NPRINT¹⁾. Текущий номер шага NSTEP увеличивается на единицу в <2.1> STEPON, при этом максимальное число шагов содержится в NRUN. В массивах «LABELn» хранятся 48-литерные обозначения.

Все элементы [C1.9] связаны со встроенными в CRONUS обширными средствами разработки и диагностики. MAXDUM и MXDUMP содержат размерности, которые используются для очистки и вывода на печать массивов дампа NADUMP, NPDUMP и NVDUMP. Если ни в одном из запусков программы не планируется выдавать больше десяти различных дампов, то с этими размерностями нам иметь дело не придется. В переменных трассировки NCLASS, NSUB и NPOINT регистрируются сведения о самой последней подпрограмме, обращавшейся к <0.4> EXPERT, а именно класс этой подпрограммы, ее номер и номер точки, откуда вызывалась <0.4> EXPERT. Логические массивы NLOMT1, NLOMT2, NLOMT3 обеспечивают возможность отключать отдельные подпрограммы. Например, если бы нам понадобилось отключить <1.6> INITAL, то нужно было бы задать (вводя в <1.4> DATA) NLOMT1(6) = .TRUE..

¹⁾ Ввод, вывод и печать. — Прим. перев.

Назначение остальных элементов [C1.9] кратко поясняется на рис. 3.6. Переменная NLREPT является «главным выключателем» сообщений. Если хоть какие-нибудь сообщения нужны, она должна равняться .TRUE.. Программа <5.1> REPORT вызывает служебную программу вывода на печать <U.27> DUMCOM. Если NLCHED задана равной .TRUE., то при каждом обращении к <5.1> REPORT на печать выдается сообщение вида

КЛАСС = 0, ПОДПРОГРАММА = 3, ТОЧКА = 6

(данный пример соответствует вызову EXPERT из точки 6 программы <0.3> CONTROL). Для подпрограмм других классов аналогичную роль играет массив NLHEAD. Так, обращение к <0.4> EXPERT из точки 1 программы <2.1> STEPON вызовет (если NLREPT = .TRUE.) печать сообщения

ШАГ = (значение NSTEP)

* * * * *

Точки, в которых нужно выдать отдельные переменные и массивы, указываются в закодированном виде в массиве NPDUMP. Предположим, например, что нам требуется предусмотреть две выдачи: первую — в точке (класс. подпрограмма. точка) = 0.3.6, а вторую — в 3.1.4. В таком случае надо задать

$$\text{NPDUMP}(1) = 306 \quad (3.1)$$

$$\text{NPDUMP}(2) = 30104 \quad (3.2)$$

класс подпрограмма точка

Какую информацию необходимо выдавать в этих точках, указывают в соответствующих элементах массивов NVDUMP и NADUMP. Например, чтобы в точке первого дампа распечатать все переменные из блока [C2.1], все блоки группы 3 и блок [C4.1], нам следует задать

$$\text{NVDUMP}(1) = 213041 \quad (3.3)$$

блок [C2.1] блок [C4.1]

ноль обозначает вывод всех переменных группы 3

По такой же форме кодируется и NADUMP. Код 100 означает «выдавать все». Так, например, если бы нам захотелось распечатать все массивы в точке 2, мы бы задали

$$\text{NADUMP}(3) = 100 \quad (3.4)$$

Вообще говоря, в каждом элементе массивов NVDUMP и NADUMP можно упаковать до четырех заказов, по две цифры

на каждый. На основании (3.3) программы <U.27> DUMCOM породит обращения

```
CALL CLIST (2,1)
CALL CLIST (3,0)
CALL CLIST (4,1)
```

тогда как (3.4) приведет к

```
CALL ARRAYS (0,0)
```

Подпрограммы <5.2> CLIST и <5.3> ARRAYS автоматически генерируются на основании главного указателя переменных автоматическим генератором программ GENESIS, а все остальные стандартные диагностические средства предусмотрены в программе CRONUS. Таким образом, как только составлен главный указатель, программное обеспечение системы OLYMPUS сразу обеспечивает определенную среду для тестирования любой подпрограммы в том виде, как она написана. Даже если GENESIS отсутствует, благодаря служебным программам из библиотеки CYCLOPS (табл. 3.5) и заготовкам в CRONUS процесс написания <5.2> CLIST и <5.3> ARRAYS сводится к простой механической операции, как только сформирован главный указатель.

На следующем этапе производится уточнение главного указателя данных. Для этого текущий указатель опять сверяется со списками, в которых отражены итоги этапов проектирования. На этот раз детальное сопоставление физического, численного и алгоритмического описаний со всеми элементами общих блоков соответствующих групп имеет своей целью выявить и завести все необходимые константы, переменные и массивы.

Полученное нами в гл. 2 математическое описание одномерной электронной плазмы кратко выражается следующими уравнениями:

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = a, \quad (3.5), (3.6)$$

$$a = -\frac{q}{m} \frac{d\varphi}{dx}, \quad \frac{d^2\varphi}{dx^2} = -\frac{\rho}{\epsilon_0}, \quad (3.7), (3.8)$$

$$\rho = \rho_0 + qn_e = -qn_0 + qn_e. \quad (3.9)$$

Координаты частиц $\{x\}$ и потенциалы $\{\varphi\}$ подчиняются периодическим граничным условиям в расчетной области длиной L :

$$x_{\min} \leq x \leq x_{\max}, \quad (3.10)$$

$$\varphi(x) = \varphi(x + L), \quad (3.11)$$

где

$$x_{\max} = x_{\min} + L. \quad (3.12)$$

Начальное состояние обычно характеризуют величинами, относящимися ко всему объему: плотностью фона n_0 и распределением электронов по скоростям. В большинстве случаев распределение по скоростям описывают первыми двумя моментами скорости: скоростью направленного движения v_d и среднеквадратичной (тепловой) скоростью v_T , при этом v_T связана с температурой T соотношением

$$v_T = \sqrt{\frac{k_B T}{m_e}}, \quad (3.13)$$

где k_B — постоянная Больцмана. Пространственный и временной масштабы характеризуются соответственно дебаевской длиной λ_D и плазменной частотой ω_p :

$$\lambda_D = \sqrt{\frac{\epsilon_0 k_B T}{n q^2}}, \quad \omega_p = \sqrt{\frac{n q^2}{\epsilon_0 m_e}}. \quad (3.14), (3.15)$$

Полная кинетическая энергия \mathcal{E} , приходящаяся на один электрон, определяется по формуле

$$\mathcal{E} = \frac{1}{2} m_e v_T^2 = \frac{1}{2} k_B T, \quad (3.16)$$

а полная потенциальная энергия на электрон V равняется

$$V = \frac{1}{2n_0 L} \int dx \rho \phi. \quad (3.17)$$

Уравнения (3.5) — (3.17) образуют список (1). Для заполнения общих блоков [C2.1] COMPPQ и [C2.2] COMSPQ мы методично переносим переменные и константы из этих уравнений в секции главного указателя, относящиеся к обоим блокам, вместе с информацией, характеризующей смысл и тип этих переменных. Затем то же самое делается для общих блоков третьей и четвертой групп. Уравнения (2.32) — (2.37) дают дискретное алгебраическое описание плазмы. Блок численных параметров [C3.1] COMNUM содержит численные величины, приведенные в математические уравнения при дискретизации, блок [C3.2] COMSCA содержит масштабные множители и так далее.

В результате этой работы получается начальная версия главного указателя, содержащего входные данные для автоматического генератора программ GENESIS. Одна секция этих входных данных представлена на рис. 3.8. Условные размерности массивов MAX1 и MAX2 заменяются в общих блоках численными значениями, а при использовании в самой программе — символьными именами (NGMAX и NPMAX). Разумеется, NGMAX и NPMAX являются служебными величинами и поэтому включены в [C1.4] COMНОК. Звездочки, фигурирующие около описаний в ряде позиций, обозначают величины, ко-

торые задает пользователь. GENSIS порождает по этим данным указатели, общие блоки и ряд подпрограмм, показанных на рис. 3.10.

3.4.3. Класс 1: подпрограммы пролога

Как только GENSIS пропущена с первым вариантом главного указателя, скелетная программа CRONUS начинает обретать черты плазменной программы ES1DIV. Теперь уже разрабатываемая программа содержит операнды, средства диагностики и разработки (рис. 3.6), подпрограммы класса 0 (управляющие) и подпрограммы инициализации <1.1> LABRUN, <1.2> CLEAR и <1.4> DATA. Любые изменения, которые требуется вносить в эти программы по ходу разработки, выполняются посредством корректировки главного указателя и перезапуска GENSIS. Остальные подпрограммы класса 1 остаются пока в виде макета.

Подпрограммы <1.3>, <1.4> и <1.5> составляют стандартную структуру инициализации данных. В подпрограмме <1.3> PRESET основным переменным и константам, не равным нулю, присваиваются начальные значения, принимаемые по умолчанию. Набор основных данных должен быть полным в том смысле, что программа должна выполнять стандартный тестовый просчет без ввода каких бы то ни было дополнительных данных. Те значения основных параметров, которые принимаются по умолчанию (и задаются пользователем), для конкретных вариантов просчета можно модифицировать, включив имена величин, которые надо изменить, и их новые значения в списки NAMELIST вводимые подпрограммой <1.4> DATA. На основании основных величин, инициализированных в <1.3> PRESET и <1.4> DATA, <1.5> AUXVAL вычисляет вспомогательные величины.

Подпрограммы <1.3> и <1.5> дают, в частности, возможность выполнить простые упражнения по составлению программ методом сверху вниз. На первом этапе нужно обеспечить в подпрограммах доступ к переменным общих блоков. Это достигается с помощью включения управляющих карт СПФ вида

C/INSERT COMppp,

где COMppp обозначает имя блока (например, COMBAS). При исполнении программы СПФ-процессор заменяет эти карты копиями отвечающих им общих блоков из библиотек системы OLYMPUS (в случае COMBAS или COMDDP) либо из самой колоды СПФ (для всех остальных блоков), в зависимости от того, что требуется. На следующих двух этапах нужно разбить подпрограммы на секции (которые отражают группы данных) и подсекции (отвечающие характеру переменных в каждом общем блоке). Согласно принятому соглашению, секции и подсек-

```

C/ PACKAGE
C/ PROGRAM ES1D1V
C/ DOCUMENTATION
C/ MODULE TITLE
C
C
C
C
C          *****
C          *                *
C          *   E S I D I V   *
C          *                *
C          *****
C
C
C          I-D ELECTROSTATIC PLASMA MODEL
C
C          BY
C
C          J.W.Eastwood and R.W.Hockney
C
C
C
C
C          Computational Physics Group
C          Computer Science Department
C          University of Reading
C          Reading, Berkshire
C          England
C
C          / Перевод: М.:Мир,1987/
C
C          Базовый вариант одномерной плазменной модели
C          частица-сетка со схемой NGR. Его назначение -
C          проиллюстрировать нисходящий метод написания
C          программ моделирования методом частица-сетка
C          с помощью системы программирования OLYMPUS.
C
C          ПРИМЕЧАНИЕ(1): В указатели не включены общие переменные
C          системы OLYMPUS и неизменяемые подпрограммы из CRONUS.
C
C          ПРИМЕЧАНИЕ(2): В этой версии программы не приведены
C          программы вывода и соответствующие переменные
C          общих блоков.
C
C          VERSION 1   RWH/JWE                      Reading Univ.   1979
C
C/ MODULE INDSUB
C-----
CL                                УКАЗАТЕЛЬ ПОДПРОГРАММ
C
C VERSION 1   RWH/JWE                      Reading Univ.   1979
CL                                ПРОЛОГ                      КЛАСС 1
C RANPOS      Формирование случайных положений                      1.9
C THERDI      Формирование скоростей с тепловым распределением 1.10
CL                                РАСЧЕТ                      КЛАСС 2
C POTSOL      Нахождение потенциала                                2.2
C
C/ MODULE INDCOM
C-----
CL                                УКАЗАТЕЛЬ ОБЩИХ БЛОКОВ

```

Рис. 3.10. СПФ-листинг программы ES1D1V

C				
C	VERSION 1	RWH/JWE	Reading Univ.	1979
C				
CL			1. ОБЩИЕ ДАННЫЕ СИСТЕМЫ OLYMPUS	
C	COMBAS		БАЗОВЫЕ ПАРАМЕТРЫ СИСТЕМЫ	C1.1
C	COMDDP		ПАРАМЕТРЫ ДЛЯ ДИАГНОСТИКИ И РАЗВИТИЯ	C1.9
C				
CL			2. ФИЗИЧЕСКАЯ ЗАДАЧА	
C	COMPPQ		Основные физические величины	C2.1
C	COMSPQ		Вспомогательные физические величины	C2.2
C				
CL			3. ЧИСЛЕННАЯ СХЕМА	
C	COMNUM		Численные параметры	C3.1
C	COMSCA		Коэффициенты перевода	C3.2
C				
CL			4. СЛУЖЕБНЫЕ ВЕЛИЧИНЫ	
C	COMHOK		Служебные величины	C4.1
C				
CL			5. ВВОД-ВЫВОД И ДИАГНОСТИКА	
C	COMOUT		Выходные переменные	C5.1
C				
C	/ MODULE INDVAR			
C	-----			
CL	АЛФАВИТНЫЙ УКАЗАТЕЛЬ ОБЩИХ ПЕРЕМЕННЫХ			
C				
C	VERSION 1	RWH/JWE	Reading Univ.	1979
C				
C	API	Pi		R 2.1
C	BLMETR	*Длина области (СИ)		R 2.1
C	BOLTZK	*Постоянная Больцмана (СИ)		R 2.1
C	BOXLEN	Длина области		R 2.2
C	CHARG(MAX1)	Плотность заряда		RA 2.1
C	DEBLNO	Начальная дебаевская длина (СИ)		R 2.1
C	DELCHG	Вклад одной частицы в плотность заряда		R 2.2
C	DENBAK	*Плотность фона (СИ)		R 2.1
C	DT	*Шаг по времени		R 3.1
C	ELCHAG	*Заряд электрона (СИ)		R 2.1
C	ELMASS	*Масса электрона (СИ)		R 2.1
C	ELPOT(MAX1)	Электрический потенциал		RA 2.1
C	ELKIN	Кинетическая энергия		R 2.2
C	ENPOT	Потенциальная энергия		R 2.2
C	ENTOT	Полная энергия		R 2.2
C	EPSO	*Диэлектрическая проницаемость вакуума (СИ)		R 2.1
C	FACPE	Отношение безразм. единиц кин. и пот. энергии		R 3.2
C	FREQPO	Начальная плазменная частота (СИ)		R 2.1
C	FRPDT	*Плазменная частота * шаг по времени		R 3.1
C	H	Ширина ячейки		R 3.1
C	NCASE	*Номер варианта начального состояния		I 4.1
C	NG	*Количество узлов сетки		I 3.1
C	NGDIM	*Размер сеточных массивов		I 4.1
C	NGMAX	*Максимальный размер дампа сеточных массивов		I 4.1
C	NGP1	NG + 1		I 3.1
C	NGP2	NG + 2		I 3.1
C	NOPSEL	*Указатель выводимого набора данных		I 5.1
C	NP	*Число частиц		I 3.1
C	NPDIM	*Размер массивов для частиц		I 4.1
C	NPMAX	*Максимальный размер дампа массивов частиц		I 4.1
C	NS1	*Период выдачи результатов		I 5.1
C	NS2	* " " " "		I 5.1

C	NS3	* -- -- --	I	5.1
C	NS4	* -- -- --	I	5.1
C	NS5	* -- -- --	I	5.1
C	NST1	*Число частиц в пучке 1 (случай 3)	I	2.2
C	NST2	*Число частиц в пучке 2 (случай 3)	I	2.2
C	QBARG	Зарядовая плотность фона	R	2.2
C	RNC	Число частиц в ячейке	R	3.1
C	SCELF	Коэффициент перевода электростат. полей в СИ	R	3.2
C	SCENG	Коэффициент перевода энергий в СИ	R	3.2
C	SCPOS	Коэффициент перевода координат в СИ	R	3.2
C	SCPOT	Коэффициент перевода потенциалов в СИ	R	3.2
C	SCVEL	Коэффициент перевода скоростей в СИ	R	3.2
C	TODEGK	*Начальная температура (град К)	R	2.1
C	TDEGK	Электронная температура	R	2.2
C	TOLRNC	*Максимально допустимая ошибка в POTSOL	R	3.1
C	TOTMOM	Полный импульс	R	2.2
C	TST1	*Температура, пучок 1 (случай 3)	R	2.2
C	TST2	*Температура, пучок 2 (случай 3)	R	2.2
C	V2PT1	*VEL(1) для двухчастичного теста	R	2.2
C	V2PT2	*VEL(2) для двухчастичного теста	R	2.2
C	VDR1	*Направленная скорость, пучок 1 (случай 3)	R	2.2
C	VDR2	*Направленная скорость, пучок 2 (случай 3)	R	2.2
C	VDRIFT	Направленная скорость электронов (СИ)	R	2.1
C	VEL(MAX2)	Скорости частиц	RA	2.1
C	VTHERO	Начальная тепловая скорость (СИ)	R	2.1
C	X2PT1	*XPOS(1) для двухчастичного теста	R	2.2
C	X2PT2	*XPOS(2) для двухчастичного теста	R	2.2
C	XMAX	Максимальное положение частицы	R	2.2
C	XMIN	Минимальное положение частицы	R	2.2
C	XPOS(MAX2)	Положения частиц	RA	2.1
C				
C	C/ MODULE INDBLK			
C	-----			
CL	УКАЗАТЕЛЬ ОБЩИХ ПЕРЕМЕННЫХ			
C				
C	VERSION 1	RWH/JWE	Reading Univ.	1979
C				
CL	C2.1	Основные физические величины		
C	VERSION 1	RWH/JWE	Reading Univ.	1979
C	COMMON/COMPRQ/			
C				
C	API	Pi	R	2.1
C	BLMETR	*Длина области (СИ)	R	2.1
C	BOLTZK	*Постоянная Больцмана (СИ)	R	2.1
C	CHARG(MAX1)	Плотность заряда	RA	2.1
C	DEBLNO	Начальная дебаевская длина (СИ)	R	2.1
C	DENBAK	*Плотность фона (СИ)	R	2.1
C	ELCHAG	*Заряд электрона (СИ)	R	2.1
C	ELMASS	*Масса электрона (СИ)	R	2.1
C	ELPOT(MAX1)	Электрический потенциал	RA	2.1
C	EPSO	*Диэлектрическая проницаемость вакуума (СИ)	R	2.1
C	FREQPO	Начальная плазменная частота (СИ)	R	2.1
C	TODEGK	*Начальная температура (град К)	R	2.1
C	VDRIFT	Направленная скорость электронов (СИ)	R	2.1
C	VEL(MAX2)	Скорости частиц	RA	2.1
C	VTHERO	Начальная тепловая скорость (СИ)	R	2.1
C	XPOS(MAX2)	Положения частиц	RA	2.1
C				
CL	C2.2	Вспомогательные физические величины		

C	VERSION 1	RWH/JWE	Reading Univ.	1979
C	COMMON/COMSPQ/			
C				
C	BOXLEN	Длина области	R	2.2
C	DELCHG	Вклад одной частицы в плотность заряда	R	2.2
C	ELKIN	Кинетическая энергия	R	2.2
C	ENPOT	Потенциальная энергия	R	2.2
C	ENTOT	Полная энергия	R	2.2
C	QVAKG	Зарядовая плотность фона	R	2.2
C	TDEGK	Электронная температура	R	2.2
C	TOTMOM	Полный импульс	R	2.2
C	TST1	*Температура, пучок 1 (случай 3)	R	2.2
C	TST2	*Температура, пучок 2 (случай 3)	R	2.2
C	V2PT1	*VEL(1) для двухчастичного теста	R	2.2
C	V2PT2	*VEL(2) для двухчастичного теста	R	2.2
C	VDR1	*Направленная скорость, пучок 1 (случай 3)	R	2.2
C	VDR2	*Направленная скорость, пучок 2 (случай 3)	R	2.2
C	X2PT1	*XPOS(1) для двухчастичного теста	R	2.2
C	X2PT2	*XPOS(2) для двухчастичного теста	R	2.2
C	XMAX	Максимальное положение частицы	R	2.2
C	XMIN	Минимальное положение частицы	R	2.2
C	NST1	*Число частиц в пучке 1 (случай 3)	I	2.2
C	NST2	*Число частиц в пучке 2 (случай 3)	I	2.2

CL C3.1 Численные параметры
 C VERSION 1 RWH/JWE Reading Univ. 1979
 C COMMON/COMNUM/

C	DT	*Шаг по времени	R	3.1
C	FRPDT	*Плазменная частота * шаг по времени	R	3.1
C	H	Ширина ячейки	R	3.1
C	RNC	Число частиц в ячейке	R	3.1
C	TOLRNC	*Максимально допустимая ошибка в POTSOL	R	3.1
C	NG	*Количество узлов сетки	I	3.1
C	NGP1	NG + 1	I	3.1
C	NGP2	NG + 2	I	3.1
C	NP	*Число частиц	I	3.1

CL C3.2 Коэффициенты перевода
 C VERSION 1 RWH/JWE Reading Univ. 1979
 C COMMON/COMSCA/

C	FACPE	Отношение безразм. единиц кин. и пот. энергии	R	3.2
C	SCELF	Коэффициент перевода электростат. полей в СИ	R	3.2
C	SCENG	Коэффициент перевода энергий в СИ	R	3.2
C	SCPOS	Коэффициент перевода координат в СИ	R	3.2
C	SCPOT	Коэффициент перевода потенциалов в СИ	R	3.2
C	SCVEL	Коэффициент перевода скоростей в СИ	R	3.2

CL C4.1 Служебные величины
 C VERSION 1 RWH/JWE Reading Univ. 1979
 C COMMON/COMNOK/

C	NCASE	*Номер варианта начального состояния	I	4.1
C	NGDIM	*Размер сеточных массивов	I	4.1
C	NGMAX	*Максимальный размер дампа сеточных массивов	I	4.1
C	NPDIM	*Размер массивов для частиц	I	4.1
C	NPMAX	*Максимальный размер дампа массивов частиц	I	4.1


```

C
C/ MODULE CIS1
      SUBROUTINE LABRUN
C
C 1.1  МАРКИРОВКА РАСЧЕТА
C
C VERSION 1  RWH/JWE                      Reading Univ.    1979
C
C/ INSERT COMBAS
-----
C
C      ЧТЕНИЕ МЕТОК
      READ(NREAD,9900)LABEL1
      READ(NREAD,9900)LABEL2
      READ(NREAD,9900)LABEL3
      READ(NREAD,9900)LABEL4
C
C      ЗАПИСЬ ЗАГОЛОВКА
      CALL BLINES(8)
      WRITE(NOUT,9901)
      WRITE(NOUT,9902)
      CALL BLINES(4)
C
C      ЗАПИСЬ МЕТОК
      WRITE(NOUT,9903) LABEL1
      CALL BLINES(1)
      WRITE(NOUT,9903) LABEL2
      CALL BLINES(1)
      WRITE(NOUT,9903) LABEL3
      CALL BLINES(1)
      WRITE(NOUT,9903) LABEL4
      CALL BLINES(1)
C
      RETURN
9900  FORMAT(12A4)
9901  FORMAT(50X,27HP R O G R A M   E S I D I V )
9902  FORMAT(50X,27(1H*))
9903  FORMAT(12X,12A4)
      END
C
C/ MODULE CIS2
      SUBROUTINE CLEAR
C
C 1.2  ЧИСТКА ПЕРЕМЕННЫХ И МАССИВОВ
C
C VERSION 1  RWH/JWE                      Reading Univ.    1979
C
      COMMON/COMPPQ/ R21(4620)
      COMMON/COMSPQ/ R22(18),I22(2)
      COMMON/COMNUM/ R31(5),I31(4)
      COMMON/COMSCA/ R32(6)
      COMMON/COMHOK/ I41(5)
      COMMON/COMOUT/ I51(6)
-----
C
C
C      ZEROS=0.0
C
C      БЛОК COMPPQ

```

```

      CALL RESETR(R21,4620,ZEROS)
C   БЛОК COMSPQ
      CALL RESETR(R22,18,ZEROS)
      CALL RESETI(I22,2,0)
C   БЛОК COMNUM
      CALL RESETR(R31,5,ZEROS)
      CALL RESETI(I31,4,0)
C   БЛОК COMSCA
      CALL RESETR(R32,6,ZEROS)
C   БЛОК COMHOK
      CALL RESETI(I41,5,0)
      БЛОК COMOUT
      CALL RESETI(I51,6,0)
C
      RETURN
      END
C
C/ MODULE C1S3
      SUBROUTINE PRESET
C
C 1.3 Задание значений, принимаемых по умолчанию
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C
-----
CL          1.          Данные системы OLYMPUS
C
      NRUN = 50
C
-----
CL          2.          Физическая система
C
      Эти величины дают результаты в единицах плазм. частоты и деб. длины
C
CL          2.1         Физические константы
      API = 3.14159265
      BOLTZK = 1.0
      ELCHAG = -1.0
      ELMASS = 1.0
      EPSO = 1.0
C
CL          2.2         Физическое состояние
      BLMETR = 64.0
      DENBAK = 1.0
      TODEGK = 1.0
      VDRIFT = 0.0
C
CL          2.3         Вспомогательные физические величины
      XMIN = 2.0
C
CL          2.4         Информация о распределении частиц
      X2PT1 = 10.0

```

```

X2PT2 = 26.0
V2PT2 = 0.0
C
TST1 = 0.01
TST2 = 0.01
VDR1 = 1.0
VDR2 = -1.0
NST1 = 1000
NST2 = 1000
C
C-----
CL          3.          Численные параметры
C
NG = 64
NP = 2
FRPDT = 0.25
TOLRNC = 1.0E-5
C
C-----
CL          4.          Служебные величины
C
NGMAX = 70
NPMAX = 2
NGDIM = 250
NPDIM = 2048
NCASE = 1
C
C-----
CL          5.          Управление выводом
C
NOPSEL = 1
NS1 = 25
NS2 = 1
NS3 = 1
NS4 = 1
NS5 = 1
C-----
RETURN
END
C
C/ MODULE C1S4
SUBROUTINE DATA
C
C 1.4 ЗАДАНИЕ ДАННЫХ КОНКРЕТНОГО ВАРИАНТА
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C-----
NAMELIST/NEWRUN/
. NDIARY, NIN , NLEDGE, NONLIN, NOUT , NPRINT,
. NPUNCH, NRUN , NADUMP, NCLASS, NPDUMP, NPOINT,
. NSUB , NVDUMP, NLCHED, NLHEAD, NLOMT1, NLOMT2,
. NLOMT3, BLMETR, BOLTZK, DENBAK, ELCHAG, ELMASS,

```

```

      . EPSO ,   TODEGK,   TST1 ,   TST2 ,   V2PT1 ,   V2PT2 ,
      . VDR1 ,   VDR2 ,   X2PT1 ,   X2PT2 ,   NST1 ,   NST2 ,
      . DT ,     FRPDT ,   TOLRNC,   NG ,     NP ,     NCASE ,
      . NGDIM ,   NGMAX ,   NPDIM ,   NPMAX ,   NOPSEL,   NS1 ,
      . NS2 ,    NS3 ,    NS4 ,    NS5
C-----
      NAMELIST/RESET/
      . NDIARY,   NIN ,    NLEDGE,   NONLIN,   NOUT ,    NPRINT,
      . NPUNCH,  NRUN ,   NADUMP,   NCLASS,   NPDUMP,   NPOINT,
      . NSUB ,   NVDUMP,   NLCHED,   NLHEAD,   NLOMT1,   NLOMT2,
      . NLOMT3
C-----
CL          1.          Новый расчет
C
      IF(NLRES) GO TO 200
      READ(NREAD,NEWRUN)
      WRITE(NOUT,NEWRUN)
      RETURN
C
C-----
CL          2.          Продолжение расчета
C
      200  CONTINUE
           READ(NREAD,RESET)
           WRITE(NOUT,RESET)
C
           RETURN
           END
C
C/ MODULE CIS5
      SUBROUTINE AUXVAL
C
C 1.5 Задание вспомогательных величин
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C
C-----
CL          1.          Физическая система
C
C
CL          1.1         Основные физические величины
      DEBLNO = SQRT(EPSO*BOLTZK*TODEGK/(DENBAK*ELCHAG**2))
      FREQPO = SQRT(DENBAK*ELCHAG**2/(EPSO*ELMASS))
      VTHERO = FREQPO * DEBLNO
C
CL          1.2         Вспомогательные физические величины
      IF(NCASE.EQ.3) NP = NST1 + NST2
      BOXLEN = NG
      DELCHG = FRPDT**2 * BOXLEN / (2.0*FLOAT(NP))
      QBAKG = -FRPDT**2 / 2.0
      XMAX = XMIN + BOXLEN
C

```



```

C-----
CL          2.          Численные параметры
C
      NGP1 = NG + 1
      NGP2 = NG + 2
      DT   = FRPDT / FREQPO
      H    = BLMETR / BOXLEN
      RNC  = FLOAT(NP) / BOXLEN
C
C-----
CL          3.          Коэффициенты перевода
C
      SCPOS = H
      SCVEL = H / DT
      SCELF = ELCHAG * DT**2 / (ELMASS*H)
      SCPOT = -2.0 * ELMASS / ELCHAG * SCVEL**2
      SCENG = ELMASS * SCVEL**2 / (8.0*FLOAT(NP))
      FACPE = -16.0 * RNC / FRPDT**2
C
C-----
CL          4.          Проверки инициализации
C
      IF(NG.LE.NGDIM) GO TO 400
      CALLMESSAGE(48HNG.СЛИШКОМ ВЕЛИКО - ПРОГРАММНЫЙ ОСТАНОВ )
      STOP
C
400      IF(NP.LE.NPDIM) GO TO 401
      CALLMESSAGE(48HNP СЛИШКОМ ВЕЛИКО - ПРОГРАММНЫЙ ОСТАНОВ )
      STOP
C
401      IF(FRPDT.LT.2.0) GO TO 402
      CALLMESSAGE(48HPРЕДУПРЕЖДЕНИЕ - DT БОЛЬШЕ ГРАНИЦЫ УСТОЙЧИВОСТИ )
      RETURN
C
402      IF(FRPDT.LE.1.0) GO TO 403
      CALLMESSAGE(48HPРЕДУПРЕЖДЕНИЕ- DT НЕ ОБЕСПЕЧИТ ВЫСОКУЮ ТОЧНОСТЬ)
403      CONTINUE
C
C-----
      RETURN
      END
C
C/ MODULE C1S6
      SUBROUTINE INITAL
C
C 1.6 Задание физического начального состояния
C
C      Примечание: XOS и VEL - в безразмерных единицах
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
      DATA          ICLASS, ISUB/1, 6/
      CALL EXPERT(ICLASS,ISUB,1)
      IF(NLOMT1(ISUB)) RETURN

```

```

C
C-----
CL          1.          Двухчастичный тест
C
      IF(NCASE.NE.1) GO TO 200
      XPOS(1) = X2PT1
      XPOS(2) = X2PT2
      VEL(1) = V2PT1
      VEL(2) = V2PT2
      RETURN

C
C-----
CL          2.          Пучок с тепловым разбросом, плотность=const
C
200    CONTINUE
      IF(NCASE.NE.2) GO TO 300
      CALL RANPOS(XPOS,1,NP,XMIN,XMAX)
      ZVDR = VDRIFF / SCVEL
      ZVTH = VTHERO / SCVEL
      CALL THERDI(VEL,1,NP,ZVDR,ZVTH)
      RETURN

C
C-----
CL          3.          Два "горячих" пучка
C
300    CONTINUE
      IF(NCASE.NE.3) GO TO 400
      INST2 = NST1 + 1
      CALL RANPOS(XPOS,1,NST1,XMIN,XMAX)
      CALL RANPOS(XPOS,INST2,NP,XMIN,XMAX)
      ZVDR1 = VDR1 / SCVEL
      ZVDR2 = VDR2 / SCVEL
      ZVTH1 = SQRT(BOLTZK * TST1 / ELMASS) / SCVEL
      ZVTH2 = SQRT(BOLTZK * TST2 / ELMASS) / SCVEL
      CALL THERDI(VEL,1,NST1,ZVDR1,ZVTH1)
      CALL THERDI(VEL,INST2,NP,ZVDR2,ZVTH2)
      RETURN

C
C-----
CL          4.          Другие случаи
C
400    CONTINUE
                                           CALL EXPERT(ICLASS,ISUB,2)

C-----
      RETURN
      END

C
C/ MODULE C1S8
      SUBROUTINE START

C
C 1.8 Начало расчета
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK

```

```

C/ INSERT COMOUT
C-----
DATA      ICLASS,  ISUB/1,  8/
          CALL EXPERT(ICLASS, ISUB, 1)
          IF(NLOMT1(ISUB)) RETURN
C
C-----
CL          1.      Приведение положений к граничным условиям
C
          DO 104 JPART = 1, NP
C
100        IF(XPOS(JPART).GE.XMIN) GO TO 101
          XPOS(JPART) = XPOS(JPART) + BOXLEN
          GO TO 100
101        CONTINUE
C
102        IF(XPOS(JPART).LT.XMAX) GO TO 103
          XPOS(JPART) = XPOS(JPART) - BOXLEN
          GO TO 102
103        CONTINUE
C
104        CONTINUE
C-----
          RETURN
          END
C
C/ MODULE C1S9
          SUBROUTINE RANPOS(POS, KFRST, KLAST, PMIN, PMAX)
C
C 1.9 Формирование случайного распределения по ячейкам
C      Примечание: функция RND выдает случайное число в интервале (0,1)
C
C-----
C      POS      Массив координат
C      KFRST    Адрес первого элемента
C      KLAST    Адрес последнего элемента
C      PMIN     Значение минимального элемента POS
C      PMAX     Значение максимального элемента POS
C-----
C
          DIMENSION POS(1)
C
C-----
CL          1.      Инициализация
C
          ZINC = (PMAX-PMIN) / FLOAT(KLAST-KFRST+1)
          ZCEL = PMIN
          ZSEED = RND(127)
C
C-----
CL          2.      Формирование положений
C
          DO 200 JPAR = KFRST, KLAST
          POS(JPAR) = ZCEL + ZINC * RND(0)
          ZCEL = ZCEL + ZINC
200        CONTINUE
C
C-----
          RETURN

```

```

      END
C
C/ MODULE C1S10
      SUBROUTINE THRED1(PVEL,KFRST,KLAST,PDRIFT,PTHERM)
C
C 1.10 Формирование теплового распределения
C
C Скорости масштабируются на точное совпадение первых двух моментов
C Гауссоида обрезается на 3 * PTHERM
C-----
C PVEL      Массив скоростей
C KFRST     Адрес первого элемента
C KLAST     Адрес последнего элемента
C PDRIFT    Направленная скорость
C PTHERM    Тепловая скорость
C-----
      DIMENSION PVEL(1)
      DATA ZRMIN/1.1109E-2/, ZTUPI/6.283185/
C
C-----
CL      1.      Формирование немасштабированной гауссоиды
C
      ZSEED = RND(271)
C
      DO 101 JPART = KFRST,KLAST
100     ZRAND = RND(0)
         IF(ZRAND.LT.ZRMIN) GO TO 100
         PVEL(JPART) = SQRT(-2.0*ALOG(ZRAND))
101     CONTINUE
C
      DO 102 JPART = KFRST,KLAST
         ZARG = ZTUPI * RND(0)
         PVEL(JPART) = PVEL(JPART) * SIN(ZARG)
102     CONTINUE
C
C-----
CL      2.      Вычисление первых двух моментов
C
      ZMEANV = 0.0
      ZMVSQ = 0.0
C
      DO 200 JPART = KFRST,KLAST
200     ZMEANV = ZMEANV + PVEL(JPART)
         ZMVSQ = ZMVSQ + PVEL(JPART)**2
C
      ZNVALS = KLAST - KFRST + 1
      ZMEANV = ZMEANV / ZNVALS
      ZMVSQ = ZMVSQ / ZNVALS - ZMEANV**2
      ZRMSV = SQRT(ZMVSQ)
C
C-----
CL      3.      Масштабирование и сдвиг распределения
C
      DO 300 JPART = KFRST,KLAST
300     PVEL(JPART) = (PVEL(JPART) - ZMEANV) * PTHERM / ZRMSV + PDRIFT
C-----
      RETURN

```

```

      END
C
C/ MODULE C2S1
      SUBROUTINE STEPON
C
C 2.1 Продвижение на один шаг
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C      Version 1 JWE/RWH
C-----
      DATA          ICLASS,  ISUB/2,  1/
                    CALL EXPERT(ICLASS, ISUB, 1)
      IF(NLOMT2(ISUB)) RETURN
      NSTEP = NSTEP + 1
C
C-----
CL          1.          Раздача заряда
C
C
CL          1.1        Начальная установка сумматоров заряда
      DO 110 JMESH = 2, NGP1
110      CHARG(JMESH) = QBAKG
C
CL          1.2        Раздача заряда
      DO 120 JPART = 1, NP
      ICEL = XPOS(JPART)
120      CHARG(ICEL) = CHARG(ICEL) + DELCHG
C
C-----
CL          2.          Получение потенциалов и потенциальной энергии
C
C
CL          2.1        Решение уравнения Пуассона
      CALL POTSOL
C
CL          2.2        Вычисление потенциальной энергии
      ENPOT = 0.0
      DO 220 JMESH = 2, NGP1
220      ENPOT = ENPOT + CHARG(JMESH) * ELPOT(JMESH)
      ENPOT = ENPOT * FACPE
C
C-----
CL          3.          Ускорение частиц и вычисление энергий
C
      ENKIN = 0.0
      DO 300 JPART = 1, NP
      ICEL = XPOS(JPART)
      ICM1 = ICEL - 1
      ICP1 = ICEL + 1
      ZACCN = ELPOT(ICP1) - ELPOT(ICM1)
      ZVOLD = VEL(JPART)
      VEL(JPART) = ZVOLD + ZACCN

```



```

ENKIN = ENKIN + (ZVOLD+VEL(JPART))**2
300 CONTINUE
C
ENTOT = ENKIN + ENPOT
C
-----
CL          4.          Движение частиц
C
DO 421 JPART = 1,NP
C
CL          4.1        Движение
XPOS(JPART) = XPOS(JPART) + VE (JPART)
C
CL          4.2        Учет граничных условий
IF(XPOS(JPART).LT.XMAX) GO TO 420
XPOS(JPART) = XPOS(JPART) - BOXLEN
IF(XPOS(JPART).GE.XMAX) GO TO 500
GO TO 421
C
420 IF(XPOS(JPART).GE.XMIN) GO TO 421
XPOS(JPART) = XPOS(JPART) + BOXLEN
IF(XPOS(JPART).LT.XMIN) GO TO 500
421 CONTINUE
RETURN
C
-----
CL          5.          Аварийный выход при слишком большой скорости
C
500 CONTINUE
CALLMESSAGE(48НКОРОСТИ СЛИШКОМ ВЕЛИКИ - СЧЕТ ПРЕКРАЩАЕТСЯ )
NLEND = .TRUE.
RETURN
END
C
C/ MODULE C2S2
SUBROUTINE POTSOL
C
C 2.2 Нахождение потенциала в одномерной периодической системе
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C Version 1 JWE/RWH
C
-----
LOGICAL ILREDO
DATA ICLASS, ISUB/2, 2/
CALL EXPERT(ICLASS,ISUB,1)
IF(NLOMT2(ISUB)) RETURN
C
-----
CL          1.          Получение потенциалов в граничных узлах
C
ILREDO = .TRUE.
100 ZPOT1 = 0.0

```

```

C
DO 101 JMESH = 2,NGP1
101 ZPOT1 = ZPOT1 + FLOAT(JMESH-NGP2) * CHARG(JMESH)
ZPOT1 = ZPOT1 / BOXLEN
C
ELPOT(2) = ZPOT1
ELPOT(NGP1) = 0.0
C
CL          1.1      Добавление периодических изображений
ELPOT(1) = 0.0
ELPOT(NGP2) = ZPOT1
C
C-----
CL          2.      Расчет внутренних узлов
C
DO 200 JMESH = 3,NG
ELPOT(JMESH) = CHARG(JMESH-1)
+ +2.0 * ELPOT(JMESH-1) - ELPOT(JMESH-2)
200 CONTINUE
C
C-----
CL          3.      Проверка согласованности и попытка исправить
C
ZERROR = CHARG(NG)+2.0*ELPOT(NG)-ELPOT(NG-1)
IF(ABS(ZERROR).LT.TOLRNC) GO TO 323
C
CL          3.1      Предупреждающее сообщение
CALLMESSAGE(48НЕСОГЛАСОВАННОСТЬ В ПОТЕНЦИАЛАХ
CALL IVAR(8HNSTEP ,NSTEP)
CALL RVAR(8HZERROR ,ZERROR)
C
IF(ILREDO) GO TO 320
CALL EXPERT(ICLASS,ISUB,2)
CALLMESSAGE(48НЕСОБЕСТИМОСТЬ ВЕЛИЧИНЫ ЗАРЯДОВ В C2S2
STOP
C
CL          3.2      Нейтрализация заряда
320 ILREDO = .FALSE.
ZTOTCH = 0.0
C
DO 321 JMESH = 2,NGP1
321 ZTOTCH = ZTOTCH + CHARG(JMESH)
ZTOTCH = ZTOTCH / FLOAT(NG)
CALL RVAR(8HZTOTCH ,ZTOTCH)
C
DO 322 JMESH = 2,NGP1
322 CHARG(JMESH) = CHARG(JMESH) - ZTOTCH
GO TO 100
323 CONTINUE
C-----
RETURN
END
C
C/ MODULE C3S1
SUBROUTINE OUTPUT(K)
C
C 3.1 Управление выводом
C
C/ INSERT COMBAS

```

```

C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
      DATA          ICLASS,  ISUB/3,  1/
                        CALL EXPERT(ICLASS,ISUB,1)
      IF(NLOMT3(ISUB)) RETURN
C-----
      GO TO (100,200,300),K
C-----
CL          1.      Начальная выдача
C
100  CONTINUE
C
CL          1.1     Общая выдача
C
CL          1.2     Выдача набора данных 1
      IF(NOPSEL.NE.1) GO TO 130
      RETURN
C
CL          1.3     Выдача набора данных 2
130  CONTINUE
      IF(NOPSEL.NE.2) GO TO 140
      RETURN
C
CL          1.4     Выдача набора данных 3
140  CONTINUE
      IF(NOPSEL.NE.3) GO TO 150
      RETURN
C
CL          1.5     Выдача других наборов данных
150  CONTINUE
                        CALL EXPERT(ICLASS,ISUB,2)
      RETURN
C-----
CL          2.      Периодическая выдача
C
200  CONTINUE
C
CL          2.1     Общая выдача
C
CL          2.2     Выдача набора данных 1
C
CL          2.3     Выдача набора данных 2
C
CL          2.4     Выдача набора данных 3
C
CL          2.5     Выдача других наборов данных
                        CALL EXPERT(ICLASS,ISUB,3)
C-----
CL          3.      Заключительная выдача
C
300  CONTINUE

```

```

C
CL          3.1      Общая выдача
C
CL          3.2      Выдача набора данных 1
C
CL          3.3      Выдача набора данных 2
C
CL          3.4      Выдача набора данных 3
C
CL          3.5      Выдача других наборов данных
                      CALL EXPERT(ICLASS,ISUB,4)

```

```

C-----
      RETURN
      END

```

```

C
C/ MODULE C5S2
      SUBROUTINE CLIST(KGROUP,KBLOCK)

```

```

C
C 5.2 ПЕЧАТЬ ОБЩИХ ПЕРЕМЕННЫХ

```

```

C
C VERSION 1   RWH/JWE                      Reading Univ.    1979

```

```

C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT

```

```

C-----
C
      IF(KGROUP.EQ.0)GO TO 100
      IF((KGROUP.LT.0).OR.(KGROUP.GT.9))RETURN
      GO TO(100,200,300,400,500,999,999,999,999),KGROUP

```

```

C-----
C
CL          1.      ОБЩИЕ ДАННЫЕ СИСТЕМЫ OLYMPUS

```

```

C
      100  CONTINUE
           IF(KBLOCK.EQ.0)GO TO 101
           GO TO(110,999,999,999,999,999,999,999,190),KBLOCK
      101  CONTINUE

```

```

C
CL          1.1     БЛОК COMBAS

```

```

      110  CONTINUE
           CALL PAGE
           CALL MESSAGE(48H БЛОК COMBAS
           CALL BLINES(1)
           CALL IVAR(8HNDIARY ,NDIARY)
           CALL IVAR(8HNIN   ,NIN   )
           CALL IVAR(8HNLEDGE ,NLEDGE)
           CALL IVAR(8HNONLIN ,NONLIN)
           CALL IVAR(8HNOUT  ,NOUT  )
           CALL IVAR(8HNPRINT ,NPRINT)
           CALL IVAR(8HNPUNCH ,NPUNCH)
           CALL IVAR(8HNREAD  ,NREAD )
           CALL IVAR(8HNREC   ,NREC  )
           CALL IVAR(8HNRESUM ,NRESUM)

```

```

CALL IVAR(8HNRUN ,NRUN )
CALL IVAR(8HNSTEP ,NSTEP )
CALL LVAR(8HNLEND ,NLEND )
CALL LVAR(8HNLRES ,NLRES )
IF(KBLOCK.NE.0)RETURN

```

C
CL 1.9 БЛОК COMDDP

```

190 CONTINUE
CALL PAGE
CALLMESSAGE(48H БЛОК COMDDP )
CALL BLINES(1)
CALL IVAR(8HMAXDUM ,MAXDUM)
CALL IVAR(8HMXDUMP ,MXDUMP)
CALL IVAR(8HNCLASS ,NCLASS)
CALL IVAR(8HNPOINT ,NPOINT)
CALL IVAR(8HNSUB ,NSUB )
CALL LVAR(8HNLCHED ,NLCHED)
CALL LVAR(8HNLREPT ,NLREPT)
IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN

```

C

C

CL 2. ФИЗИЧЕСКАЯ ЗАДАЧА

C

```

200 CONTINUE
IF(KBLOCK.EQ.0)GO TO 201
GO TO(210,220,999,999,999,999,999,999,999),KBLOCK
201 CONTINUE

```

C

CL 2.1 БЛОК

```

210 CONTINUE
CALL PAGE
CALLMESSAGE(48H БЛОК COMPPQ)
CALL BLINES(1)
CALL RVAR(8HAPI ,API )
CALL RVAR(8HBLMETR ,BLMETR)
CALL RVAR(8HBOLTZK ,BOLTZK)
CALL RVAR(8HDEBLNO ,DEBLNO)
CALL RVAR(8HDENBAK ,DENBAK)
CALL RVAR(8HELCHAG ,ELCHAG)
CALL RVAR(8HELMASS ,ELMASS)
CALL RVAR(8HEPSO ,EPSO )
CALL RVAR(8HFREQPO ,FREQPO)
CALL RVAR(8HTODEGK ,TODEGK)
CALL RVAR(8HVDRIFT ,VDRIFT)
CALL RVAR(8HVTHERO ,VTHERO)
IF(KBLOCK.NE.0)RETURN

```

C

CL 2.2 БЛОК COMSPQ

```

220 CONTINUE
CALL PAGE
CALLMESSAGE(48H БЛОК COMSPQ )
CALL BLINES(1)
CALL RVAR(8HBOXLEN ,BOXLEN)
CALL RVAR(8HDELCHG ,DELCHG)
CALL RVAR(8HENKIN ,ENKIN )
CALL RVAR(8HENPOT ,ENPOT )
CALL RVAR(8HENTOT ,ENTOT )
CALL RVAR(8HQBAKG ,QBAKG )
CALL RVAR(8HTDEGK ,TDEGK )

```



```

CALL RVAR(8HTOTMOM ,TOTMOM)
CALL RVAR(8HTST1 ,TST1 )
CALL RVAR(8HTST2 ,TST2 )
CALL RVAR(8HV2PT1 ,V2PT1 )
CALL RVAR(8HV2PT2 ,V2PT2 )
CALL RVAR(8HVDR1 ,VDR1 )
CALL RVAR(8HVDR2 ,VDR2 )
CALL RVAR(8HX2PT1 ,X2PT1 )
CALL RVAR(8HX2PT2 ,X2PT2 )
CALL RVAR(8HXMAX ,XMAX )
CALL RVAR(8HXMIN ,XMIN )
CALL IVAR(8HNST1 ,NST1 )
CALL IVAR(8HNST2 ,NDT2 )
IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN
C
C-----
CL          3.          ЧИСЛЕННАЯ СХЕМА
C
300  CONTINUE
      IF(KBLOCK.EQ.0)GO TO 301
      GO TO(310,320,999,999,999,999,999,999,999),KBLOCK
301  CONTINUE
C
CL          3.1        БЛОК COMNUM
310  CONTINUE
      CALL PAGE
      CALLMESSAGE(48H БЛОК COMNUM
      CALL BLINES(1)
      CALL RVAR(8HDT ,DT )
      CALL RVAR(8HFRPDT ,FRPDT )
      CALL RVAR(8HH ,H )
      CALL RVAR(8HRNC ,RNC )
      CALL RVAR(8HTOLRNC ,TOLRNC)
      CALL IVAR(8HNG ,NG )
      CALL IVAR(8HNGP1 ,NGP1 )
      CALL IVAR(8HNGP2 ,NGP2 )
      CALL IVAR(8HNP ,NP )
      IF(KBLOCK.NE.0)RETURN
C
CL          3.2        БЛОК COMSCA
320  CONTINUE
      CALL PAGE
      CALLMESSAGE(48H БЛОК COMSCA
      CALL BLINES(1)
      CALL RVAR(8HFACPE ,FACPE )
      CALL RVAR(8HSCELF ,SCELF )
      CALL RVAR(8HSCENG ,SCENG )
      CALL RVAR(8HSCPOS ,SCPOS )
      CALL RVAR(8HSCPOT ,SCPOT )
      CALL RVAR(8HSCVEL ,SCVEL )
      IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN
C
C-----
CL          4.          СЛУЖЕБНЫЕ ВЕЛИЧИНЫ
C
400  CONTINUE
      IF(KBLOCK.EQ.0)GO TO 401
      GO TO(410,999,999,999,999,999,999,999,999),KBLOCK
401  CONTINUE

```

```

C
CL          4.1      БЛОК СОМНОК
  410      CONTINUE
          CALL PAGE
          CALLMESSAGE(48H БЛОК СОМНОК
          CALL BLINES(1)
          CALL IVAR(8HNCASE      ,NCASE )
          CALL IVAR(8HNGDIM     ,NGDIM )
          CALL IVAR(8HNGMAX     ,NGMAX )
          CALL IVAR(8HNPDIM     ,NPDIM )
          CALL IVAR(8HNPMAX     ,NPMAX )
          IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN
C
C-----
CL          5.      ВВОД-ВЫВОД И ДИАГНОСТИКА
C
  500      CONTINUE
          IF(KBLOCK.EQ.0)GO TO 501
          GO TO(510,999,999,999,999,999,999,999),KBLOCK
  501      CONTINUE
C
CL          5.1      БЛОК СОМОУТ
  510      CONTINUE .
          CALL PAGE
          CALLMESSAGE(48H БЛОК СОМОУТ
          CALL BLINES(1)
          CALL IVAR(8HNOSEL     ,NOSEL)
          CALL IVAR(8HNS1      ,NS1  )
          CALL IVAR(8HNS2      ,NS2  )
          CALL IVAR(8HNS3      ,NS3  )
          CALL IVAR(8HNS4      ,NS4  )
          CALL IVAR(8HNS5      ,NS5  )
          IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN
C
C-----
  999      CONTINUE
          RETURN
          END
C
C/ MODULE C5S3
          SUBROUTINE ARRAYS(KGROUP,KBLOCK)
C
C 5.3 ПЕЧАТЬ ОБЩИХ МАССИВОВ
C
C VERSION 1   RWH/JWE                      Reading Univ.    1979
C
C/ INSERT COMBAS
C/ INSERT COMDDP
C/ INSERT COMPPQ
C/ INSERT COMSPQ
C/ INSERT COMNUM
C/ INSERT COMSCA
C/ INSERT COMHOK
C/ INSERT COMOUT
C-----
C
          IF(KGROUP.EQ.0)GO TO 100
          IF((KGROUP.LT.0).OR.(KGROUP.GT.9))RETURN

```

```

GO TO(100,200,300,400,500,999,999,999,999),KGROUP
C
C-----
CL          1.          ОБЩИЕ ДАННЫЕ СИСТЕМЫ OLYMPUS
C
100  CONTINUE
      IF(KBLOCK.EQ.0)GO TO 101
      GO TO(110,999,999,999,999,999,999,999,190),KBLOCK
101  CONTINUE
C
CL          1.1        БЛОК COMBAS
110  CONTINUE
      CALL PAGE
      CALL BLINES(1)
      CALLMESSAGE(48H БЛОК COMBAS
      CALL BLINES(1)
      CALL HARRAY(8HLABEL1 ,LABEL1,12)
      CALL HARRAY(8HLABEL2 ,LABEL2,12)
      CALL HARRAY(8HLABEL3 ,LABEL3,12)
      CALL HARRAY(8HLABEL4 ,LABEL4,12)
      CALL HARRAY(8HLABEL5 ,LABEL5,12)
      CALL HARRAY(8HLABEL6 ,LABEL6,12)
      CALL HARRAY(8HLABEL7 ,LABEL7,12)
      CALL HARRAY(8HLABEL8 ,LABEL8,12)
      IF(KBLOCK.NE.0)RETURN
C
CL          1.9        БЛОК COMDDP
190  CONTINUE
      CALL PAGE
      CALL BLINES(1)
      CALLMESSAGE(48H БЛОК COMDDP
      CALL BLINES(1)
      CALL IARRAY(8HNADUMP ,NADUMP,20)
      CALL IARRAY(8HNPDUMP ,NPDUMP,20)
      CALL IARRAY(8HNVDUMP ,NVDUMP,20)
      CALL LARRAY(8HNLNEAD ,NLNEAD,9)
      CALL LARRAY(8HNLOMT1 ,NLOMT1,50)
      CALL LARRAY(8HNLOMT2 ,NLOMT2,50)
      CALL LARRAY(8HNLOMT3 ,NLOMT3,50)
      IF((KGROUP.NE.0).OR.(KBLOCK.NE.0))RETURN
C
C-----
CL          2.          ФИЗИЧЕСКАЯ ЗАДАЧА
C
200  CONTINUE
      IF(KBLOCK.EQ.0)GO TO 201
      GO TO(210,999,999,999,999,999,999,999,999),KBLOCK
201  CONTINUE
C
CL          2.1        БЛОК COMPPQ
210  CONTINUE
      CALL PAGE
      CALL BLINES(1)
      CALLMESSAGE(48H БЛОК COMPPQ
      CALL BLINES(1)
      CALL RARRAY(8HCHARG ,CHARG ,NGMAX)
      CALL RARRAY(8HELPOТ ,ELPOT ,NGMAX)
      CALL RARRAY(8HVEL ,VEL ,NPMAX)
      CALL RARRAY(8HXPOS ,XPOS ,NPMAX)

```

```

      IF(KBLOCK.NE.0)RETURN
C
C-----
CL          3.          ЧИСЛЕННАЯ СХЕМА
C
  300      CONTINUE
           IF(KBLOCK.EQ.0)GO TO 301
           GO TO(999,999,999,999,999,999,999,999,999),KBLOCK
  301      CONTINUE
C
C-----
C          4.          СЛУЖЕБНЫЕ ВЕЛИЧИНЫ
C
  400      CONTINUE
           IF(KBLOCK.EQ.0)GO TO 401
           GO TO(999,999,999,999,999,999,999,999,999),KBLOCK
  401      CONTINUE
C
C-----
CL          5.          ВВОД-ВЫВОД И ДИАГНОСТИКА
C
  500      CONTINUE
           IF(KBLOCK.EQ.0)GO TO 501
           GO TO(999,999,999,999,999,999,999,999,999),KBLOCK
  501      CONTINUE
C
C-----
C          999      CONTINUE
                   RETURN
                   END
C
C/ TEST DATA
C/ MODULE TEST1
1-D Plasma Model
Two Particle Test
R.W.Hockney and J.W.Eastwood
Reading University 1979
C/ NAMELIST NEWRUN
  NLREPT = T,
  NPDUMP(1)=313,NADUMP(1)=100,NVDUMP(1)=100,
  NPDUMP(2)=30101,NADUMP(2)=21,NVDUMP(2)=22,
C/ END NAMELIST
C/ END PROGRAM
C/ END PACKAGE

```

ции начинаются с заголовка, снабженного десятичным номером, при этом секции отделяются друг от друга сплошной линией (см. рис. 3.10). Как только такое разбиение произведено, в соответствующие места можно вносить выполняемые инструкции. Польза такого подхода заключается в том, что в процессе своего развития программа сохраняет аккуратный вид. Если вдруг потребуются какие-нибудь новые переменные, то всегда имеется вполне определенное место, куда их можно вставить.

Подпрограмма <1.6> INITIAL задает физические начальные условия. Ее разработка ведется таким же образом, как и описанных выше <1.3> PRESET и <1.5> AUXVAL. Однако в этом случае имеется еще один уровень последовательного уточнения. Все выполняемые инструкции можно подразделить на две группы: инструкции, касающиеся управления, и инструкции, связанные с вычислением начальных координат и скоростей. Сначала пишутся первые, чтобы заготовить каркас, на котором потом можно монтировать инструкции второй группы, точно так же, как CRONUS обеспечивает определенную управляющую структуру, в которую могут вставляться OLYMPUS-подпрограммы со стандартными именами. Стоит еще обратить внимание на управляющую структуру <1.6> (см. рис. 3.10). Она построена таким образом, чтобы можно было легко подключать другие варианты, либо в каждом конкретном случае предусматривая специальную версию <0.4> EXPERT, задающую координаты и скорости при обращении к ней с аргументами (1, 6, 2), либо постоянно — добавив в конец дополнительные секции. В обоих случаях можно воспроизвести все предшествующие запуски программы, ничего в ней не трогая.

Подпрограмма <1.7> RESUME, фигурирующая в CRONUS, оставлена в ESIDIV в виде макета. Последняя в CRONUS подпрограмма пролога, а именно <1.8> START, обычно используется для подготовки к интегрированию по времени, когда применяются разностные схемы, не обладающие свойством самозапуска (например, когда для интегрирования на один шаг в основном счетном цикле требуются данные не только с текущего временного слоя). В ESIDIV <1.8> START используется только для проверки на согласованность. Остальные подпрограммы класса 1 (см. рис. 3.10) служат для формирования координат и скоростей частиц.

3.4.4. Подпрограммы счета и вывода результатов

После того как начальные версии программ класса 1 проверены с помощью стандартных средств дампа системы OLYMPUS, можно параллельно разрабатывать основные счетные подпрограммы (класс 2) и подпрограммы вывода результатов (класс 3). Главными управляющими подпрограммами этих двух

классов являются, соответственно, <2.1> STEPON и <3.1> OUTPUT.

Метод разработки подпрограммы <2.1> STEPON не отличается от описанного выше для подпрограмм класса 1. Сначала с помощью управляющих карт C/INSERT, вставляемых в макетный вариант, организуется доступ к переменным общих блоков. Затем добавляют стандартную структуру переключателей и диагностики

```
DATA ICLASS, ISUB/2,1/
```

```
CALL EXPERT(ICLASS, ISUB, 1)
```

```
IF(NLOMT2 (ISUB)) RETURN
```

вместе с инструкцией присваивания

```
NSTEP = NSTEP + 1
```

которая обновляет значение счетчика шагов по времени. После этого описывают заголовки секций и подсекций и добавляют инструкции, которые управляют логикой выполнения программы в этих секциях. И наконец, программируют инструкции, реализующие требуемые действия, а также обращение к подпрограмме более низкого уровня, а именно <2.2> POTSOL.

Что касается программ большего объема, то обычно все инструкции, которые фактически реализуют вычисления, почти полностью выносят из <2.1> STEPON в подпрограммы более низкого уровня. В таких случаях заключительная стадия разработки <2.1> STEPON состоит в том, чтобы вставить обращения к соответствующим подпрограммам и завести сами эти подпрограммы — вначале в виде макета, чтобы тем самым сохранить разрабатываемую программу в выполняемом виде. Более того, на этом структурирование программы может и не заканчиваться. В больших программах часто требуется, чтобы подпрограммы, вызываемые из STEPON, сами были управляющими программами по отношению к расчетным модулям. Если в этих модулях требуются, в свою очередь, дополнительные общие блоки или переменные либо и то, и другое, то их можно добавить в главный указатель и включить в структуру стандартной инициализации класса 1.

Поскольку разработка <3.1> OUTPUT и связанных с ней программ класса 3 ведется по той же схеме, что и программ класса 1 и класса 2, разбирать их здесь подробно мы не будем. В приведенной на рис. 3.10 распечатке программа <3.1> OUTPUT оставлена в недописанном виде и иллюстрирует управляющую структуру, в которую по мере необходимости можно вводить обращения к конкретным программам.

3.5. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Программы моделирования составляют инструментарий учебного-вычислителя. Как и в случае настоящих лабораторных установок, которые они заменяют, с ростом объема и сложности таких программ возрастает необходимость в систематическом управлении их разработкой. Чтобы формализовать разработку ESIDIV, мы воспользовались системой OLYMPUS, хотя выбор этой системы и не является принципиальным с точки зрения тех понятий, которые мы постарались донести до читателя. Не будь системы OLYMPUS, разработка все равно прошла бы те же основные этапы, но при этом пришлось бы трудоемкую техническую работу выполнять вручную.

Основные этапы разработки программ можно коротко сформулировать в виде следующего плана проведения работ.

1. Прежде чем начинать программировать, тщательно проведите этапы проектирования (рис. 1.1). По каждому этапу проектирования составьте итоговые списки, что обеспечит готовые перекрестные ссылки. Если для задачи, которую надо решить, подойдет какая-то хорошо проверенная программа, используйте ее.
2. Взяв за основу CRONUS, спланируйте архитектуру своей программы. Выделите модульные части, причем если есть готовые модули, включите их в программу.
3. Ориентируясь на группы данных, принятые в системе OLYMPUS, и модульный принцип организации программы, составьте план общей организации данных. Приступайте к созданию главного указателя переменных.
4. Напишите и отладьте программы самого верхнего уровня, используя вначале на нижележащих уровнях программного дерева фиктивные подпрограммы. На этой стадии введите переключатели, управляющие ходом выполнения программы, диагностические контрольные точки и вывод диагностики. Величины, входящие в [C1.1] COMBAS и [C1.9] COMDDP (табл. 3.6), наглядно иллюстрируют характер переменных, нужных на этой стадии.
5. Разработайте подпрограммы инициализации первого уровня (класс 1), добавляя в подходящие места организации данных основные физические, численные и служебные переменные и записывая в главный указатель их имена, тип и краткое описание. Заполните фиктивные подпрограммы, добавляя по очереди невыполняемые части, точки переключения и диагностики, секции и подсекции, управляющую структуру и, наконец, либо выполняемые инструкции, которые фактически и производят вычисления, либо обращения к подпрограммам более низкого уровня или и то и другое.

6. Повторите п. 5 для программ первого уровня: основных счетных подпрограмм (класс 2), подпрограмм вывода (класс 3) и подпрограмм завершения расчета, расширяя при необходимости главный указатель переменных и программы класса 1.
7. Переходите к подпрограммам следующего уровня, пользуясь для отладки новых программ на каждой стадии разработки имеющимися диагностическими средствами. Для автономных расчетных модулей вводите дополнительные общие блоки, инициализирующие подпрограммы и т. д.
8. Повторите п. 7, уточняя и расширяя, если потребуется, ранее разработанные части программы, до тех пор, пока не получите полностью протестированную работающую программу. На всех стадиях следите за тем, чтобы исправления, вносимые в главные указатели, средства диагностики, общие блоки и комментарии были согласованы между собой и с выполняемыми частями программы.

В приведенном рабочем плане неявно фигурируют тестовые прогоны частично разработанной программы на каждой стадии ее развития. Простейший и зачастую самый полезный тест заключается в ручной проверке полного дампа, выданного в тестовых запусках с малым числом частиц и грубыми сетками. Однако такие проверки очень трудоемки и не гарантируют обнаружения всех ошибок. Вот почему всегда имеет смысл обратиться к известным свойствам моделей частиц, которые могут помочь в проверке правильности программы: из простейших статистических соображений (теорема Байеса) ясно, что чем больше будет независимых проверок, тем более вероятно, что в программе не останется ошибок.

Применительно к моделям частиц полезными оказываются три важных вида проверок правильности программы — тесты на согласованность, симметрию и сходимость. В качестве примеров тестов на согласованность можно назвать проверку сохранения полного импульса и условия зарядовой нейтральности в периодических системах. Погрешности значений этих величин не должны превышать погрешностей округления. В тестах на симметрию проверяется правильность формул, а не значений входящих в них величин. Такие тесты можно, например, применить к конфигурации, состоящей из одной или двух частиц, и проверить симметрию распределения зарядовой плотности, потенциала и сил. Для периодических и изолированных систем можно предусмотреть дополнительные проверки на симметрию, основанные на свойстве инвариантности указанных характеристик по отношению к смещению. Более того, метод зеркальных изображений [330, гл. 3] позволяет распространить проверки инвариантности к сдвигу и на модели с граничными условиями Неймана и Дирихле. Тесты на сходимость используются для

проверки того, стремятся ли численные решения к известным аналитическим решениям с измельчением размера ячейки и шага по времени. В основе двух таких тестов лежит знание потенциала отдельной частицы и колебательного движения пары частиц. При этом, однако, не следует забывать, что в случае бесстолкновительных систем интерес представляет коллективное поведение (поскольку $N_p \gg 1$), а «самая хорошая» парная сила может и не дать в сумме наилучшего приближения (см. гл. 7).

Схемы интегрирования по времени

4.1. ВВЕДЕНИЕ

В большинстве моделей частиц их движение происходит в соответствии с законами Ньютона (см. табл. 1.1)

$$\frac{dx}{dt} = v, \quad (4.1)$$

$$m \frac{dv}{dt} = F. \quad (4.2)$$

Вид силы F зависит от задачи. Например, для частицы с зарядом q , движущейся в электрическом поле E и магнитном поле B , сила F равна сумме электрической силы qE и силы Лоренца $qv \times B$.

В процессе перехода от математического описания физических явлений к программе моделирования непрерывные дифференциальные уравнения (4.1) и (4.2) заменяются линейными алгебраическими соотношениями. Непрерывные функции x и v заменяются их значениями в дискретные моменты времени. Одним из примеров дискретизации, с которым мы уже встречались (см. гл. 2), является схема с перешагиванием:

$$\frac{x^{n+1} - x^n}{DT} = v^{n+1/2}, \quad (4.3)$$

$$m \frac{(v^{n+1/2} - v^{n-1/2})}{DT} = F(x^n), \quad (4.4)$$

где DT — шаг по времени, а верхний индекс обозначает номер временного слоя ($t = nDT$ и т. д.).

Уравнения (4.3) и (4.4) — это только один из возможных вариантов. Существует множество способов аппроксимации уравнений (4.1) и (4.2) дискретными аналогами, хотя, как правило, дискретные аппроксимации имеют вид линейных многошаговых уравнений, связывающих величины на дискретных временных слоях:

$$\sum_{i=0}^k a_{k-i} x^{n+k-i} = \frac{DT^2}{m} \sum_{i=0}^k b_{k-i} F^{n+k-i}. \quad (4.5)$$

В зависимости от значения b_k дискретные уравнения, записанные в такой форме, относятся к одному из двух типов. Если b_k равно нулю, схема называется *явной* и в этом случае x^n может быть выражено непосредственно через известные величины. Если b_k не равно нулю, схема является *неявной* и, за исключением случаев, когда F имеет простой вид (например, линейная функция от x), x^n нужно находить с помощью итерационных методов.

Из уравнения (4.5) очевидно, что необходимы какие-то критерии, помогающие оценивать и выбирать схемы из почти неограниченного числа вариантов. Такие критерии мы рассмотрим ниже. Их можно озаглавить так:

согласованность,
точность,
устойчивость,
эффективность.

Проводимое ниже обсуждение этих критериев связано с практическими рекомендациями по интегрированию конкретных уравнений движения, чаще всего встречающихся при моделировании методом частиц. Тех читателей, которых интересует строгое математическое рассмотрение численных методов интегрирования обыкновенных дифференциальных уравнений, частным случаем которых являются уравнения движения, мы отсылаем к книгам [171], [100] или большинству учебников по численным методам.

4.2. СОГЛАСОВАННОСТЬ

Первое необходимое условие любой алгебраической аппроксимации дифференциального уравнения состоит в том, что она должна переходить в это дифференциальное уравнение в пределе бесконечно малого временного шага. Так, например, можно сказать, что дискретное уравнение

$$\frac{x^{n+1} - x^n}{DT} = v^n \quad (4.6)$$

согласовано с дифференциальным уравнением

$$\frac{dx}{dt} = v, \quad (4.7)$$

поскольку

$$\begin{aligned} x^{n+1} - x^n &= x(t^n + DT) - x(t^n) = \\ &= \left[x(t^n) + \frac{dx}{dt} \Big|_{t=t^n} DT + \dots \right] - x(t^n) \end{aligned} \quad (4.8)$$

дает

$$\lim_{DT \rightarrow 0} \left[\frac{x(t^n + DT) - x(t^n)}{DT} \right] = \frac{dx}{dt} = v. \quad (4.9)$$

Второе (желательное) свойство согласованности дискретной аппроксимации с дифференциальным уравнением касается совпадения их свойств симметрии относительно времени. Дифференциальные уравнения (4.1) и (4.2) обратимы во времени, т. е. если движение частицы проинтегрировать вперед по времени в заданном силовом поле, а затем обратить время (и скорость), частица повторит пройденный ею путь и вернется в исходную точку.

Обратимые во времени разностные аппроксимации получают, вводя центрированные по времени производные. Примером центрированных по времени разностных аппроксимаций служат уравнения (4.3) и (4.4): разность $(x^{n+1} - x^n)$ центрирована относительно времени $t^{n+1/2}$, а разность $(v^{n+1/2} - v^{n-1/2})$ центрирована относительно t^n . Уравнение (4.6) содержит разностную аппроксимацию, которая должным образом не центрирована по времени: $(x^{n+1} - x^n)$ центрирована относительно $t^{n+1/2}$, в то время как правая часть (v^n) определена в момент времени t^n .

К сожалению, не всегда удастся воспользоваться должным образом центрированными по времени схемами, поскольку обычно они приводят к неявным уравнениям для переменных на новом временном слое. В явных разностных уравнениях значение функции на новом временном слое входит только в член с производной, что дает возможность получать новые значения простым вычислением выражения, зависящего лишь от старых значений. В неявных же схемах новые значения присутствуют как в члене с производной, так и в правой части. А это значит, что неявные схемы требуют решения системы уравнений, которая может быть и нелинейной. Например, для уравнений движения

$$\frac{dx}{dt} = v, \quad (4.10)$$

$$\frac{dv}{dt} = v \times \Omega(x, t) \quad (4.11)$$

можно написать неявную схему

$$(x^{n+1} - x^n) = \frac{1}{2} (v^{n+1} + v^n) DT, \quad (4.12)$$

$$(v^{n+1} - v^n) = \frac{1}{2} (v^{n+1} \times \Omega^{n+1} + v^n \times \Omega^n) DT. \quad (4.13)$$

Если Ω не является простой алгебраической функцией x , уравнения (4.12) и (4.13) не удастся переписать таким образом, чтобы новые значения (на временном слое $n+1$) были выражены через старые (на временном слое n). Затраты на итера-

ционное решение неявных уравнений приемлемы для небольшого числа частиц, но не для больших ансамблей, используемых во многих моделях частица — сетка. Вот почему иногда пользуются явными схемами, которые должным образом не центрированы по времени.

4.3. ТОЧНОСТЬ

Понятия точности и устойчивости разностных схем связаны с требованием малости отклонения вычисленных значений координат и скоростей от точного решения дифференциального уравнения.

Точность связана с локальными погрешностями. Есть два источника их возникновения. Первым являются погрешности округления, проистекающие из-за конечной длины представления чисел в слове компьютера, а второй обусловлен погрешностями аппроксимации, вызванными представлением непрерывных переменных дискретным набором значений. Третьим возможным источником погрешностей в вычислениях является программистская ошибка — думается, что вы этого избежите, применяя систематический подход к программированию, описанный в гл. 3, а ежели нет, тогда она отразится на согласованности!

Как правило, погрешности округления намного меньше погрешностей аппроксимации, и при условии, что схема устойчива (см. следующий раздел), ими обычно можно пренебречь. Погрешности аппроксимации, которые возникают из-за представления непрерывных величин дискретными наборами значений, можно описать любым из двух способов: на языке выборки (см. гл. 5), где эти погрешности проявляются в явном виде в частотной области как ложные спектральные компоненты, или исходя из разности между дифференциальным и алгебраическим уравнениями. В последнем случае мерой малости погрешностей аппроксимации служит порядок разностной схемы. Порядок p означает, что при малых DT погрешности аппроксимации пропорциональны $(DT)^p$.

Схема с перешагиванием (4.3) — (4.4) аппроксимирует уравнения движения (4.1) и (4.2) со вторым порядком точности. Исключая v из уравнений (4.3) и (4.4), мы получим (в одномерном случае)

$$\frac{x^{n+1} - 2x^n + x^{n-1}}{DT^2} = \frac{F(x^n)}{m}. \quad (4.14)$$

Пусть X является решением дифференциальных уравнений (4.1) и (4.2)

$$\frac{d^2X}{dt^2} = F/m. \quad (4.15)$$

Тогда погрешность аппроксимации на временном слое n , δ^n , мы определим, подставляя X в уравнение (4.14):

$$\frac{X^{n+1} - 2X^n + X^{n-1}}{DT^2} = \frac{F(X^n)}{m} - \delta^n. \quad (4.16)$$

Разлагая X^{n+1} и X^{n-1} в ряд Тейлора около $X^n = X(t^n)$, можно записать уравнение (4.16) в виде

$$\frac{d^2X}{dt^2} + \frac{DT^2}{12} \frac{d^4X}{dt^4} + \text{члены более высокого порядка} = \frac{F(X^n)}{m} - \delta^n, \quad (4.17)$$

где производные нужно вычислять в точке $t = t^n$. После чего, вычитая уравнение (4.17) из уравнения (4.15), получаем требуемый результат:

$$\delta^n = -\frac{DT^2}{12} \frac{d^4X}{dt^4} + \text{члены более высокого порядка}. \quad (4.18)$$

Это значит, что схема с перешагиванием имеет второй порядок точности ($\delta^n \sim DT^2$).

4.4. УСТОЙЧИВОСТЬ

Устойчивость связана с распространением погрешностей. Даже в том случае, когда погрешности аппроксимации и округления очень малы, от схемы будет мало пользы, если влияние этих малых погрешностей быстро нарастает со временем. Возникновение неустойчивости связано с наличием нефизического решения дискретных уравнений. Если данные дискретные уравнения имеют решения, которые растут во много раз быстрее искомого решения дифференциального уравнения, можно быть уверенным в том, что даже совсем небольшая погрешность округления в конце концов забьет это решение и сделает все численные результаты бессмысленными.

Говорят, что численный метод устойчив, если малая погрешность, внесенная на любой стадии расчета, не накапливается. Чтобы получить количественную меру устойчивости, мы должны найти уравнение, описывающее эволюцию погрешностей, а затем исследовать, как ведут себя решения этого уравнения. В качестве примера вновь возьмем схему с перешагиванием.

Записывая схему с перешагиванием как функцию одних координат, имеем

$$x^{n+1} - 2x^n + x^{n-1} = \frac{F(x^n)}{m} DT^2. \quad (4.19)$$

Если задать в качестве начальных условий $x^0 = X^0$, $x^1 = X^1$, то, не будь погрешностей округления, уравнение (4.19) дало бы

значения решения X^2, X^3, \dots , где

$$X^2 - 2X^1 + X^0 = \frac{F(X^1)}{m} DT^2, \quad (4.20)$$

$$X^3 - 2X^2 + X^1 = \frac{F(X^2)}{m} DT^2 \quad (4.21)$$

и так далее. Однако арифметические операции с плавающей точкой выполняются в компьютерах неточно, в результате чего значения точного решения $X^2, X^3, X^4 \dots$ разностных уравнений заменяются приближенными x^2, x^3, x^4, \dots , где

$$x^2 - 2X^1 + X^0 = \frac{F(X^1)}{m} DT^2, \quad (4.22)$$

$$x^3 - 2x^2 + X^1 = \frac{F(x^2)}{m} DT^2 \quad (4.23)$$

и т. д. Погрешность в координате на втором временном слое за счет округления равна

$$\varepsilon^2 = x^2 - X^2. \quad (4.24)$$

Каждая арифметическая операция всегда сопровождается погрешностями округления. Нас в данном случае интересует, как распространяется эта погрешность, внесенная на определенном временном слое, поэтому давайте посмотрим, каким образом ε^2 сказывается на решении на более поздних временных слоях при отсутствии каких бы то ни было других округлений.

Пусть погрешность на втором временном слое приводит к погрешности в решении уравнения (4.19), равной

$$\varepsilon^n = x^n - X^n. \quad (4.25)$$

Тогда из уравнений (4.23) и (4.25) мы имеем

$$(X^3 + \varepsilon^3) - 2(X^2 + \varepsilon^2) + X^1 = \frac{F(X^2 + \varepsilon^2)}{m} DT^2. \quad (4.26)$$

Вычитая уравнение (4.21), получаем

$$\varepsilon^3 - 2\varepsilon^2 = \frac{F(X^2 + \varepsilon^2)}{m} DT^2 - \frac{F(X^2)}{m} DT^2. \quad (4.27)$$

Предположим, что погрешность округления мала, так что правую часть уравнения (4.27) можно разложить в окрестности X^2 в ряд Тейлора, что дает

$$\varepsilon^3 - 2\varepsilon^2 \simeq \varepsilon^2 \left(\frac{\partial F}{\partial x} \Big|_{x=X^2} \frac{DT^2}{m} \right). \quad (4.28)$$

Аналогично,

$$\varepsilon^4 - 2\varepsilon^3 + \varepsilon^2 = \varepsilon^3 \left(\frac{\partial F}{\partial x} \Big|_{x=X^3} \frac{DT^2}{m} \right) \quad (4.29)$$

и на n -м временном слое

$$\varepsilon^{n+1} - 2\varepsilon^n + \varepsilon^{n-1} = \varepsilon^n \left(\frac{\partial F}{\partial x} \Big|_{x=X^n} \frac{\partial T^2}{m} \right). \quad (4.30)$$

Решить уравнение (4.30) ничуть не легче, чем уравнение (4.19), если, конечно, не рассматривать случай $\partial F/\partial x = \text{const}$. Поскольку мы сейчас ищем наихудший случай (нас интересует только сам факт — неустойчива схема или нет, а не насколько она неустойчива), для получения уравнения эволюции погрешности заменим производную $(\partial F/\partial x)|_{x=X^n}$ ее максимальным отрицательным значением $-|\partial F/\partial x|_{\text{max}}$:

$$\varepsilon^{n+1} - 2\varepsilon^n + \varepsilon^{n-1} = -\frac{1}{m} \left| \frac{\partial F}{\partial x} \right|_{\text{max}} DT^2 \varepsilon^n \quad (4.31)$$

$$= -\Omega^2 DT^2 \varepsilon^n. \quad (4.32)$$

Модуль и отрицательный знак, т. е. $\partial F/\partial x \rightarrow -|\partial F/\partial x|_{\text{max}}$, возникают из предположения, что решение уравнения (4.19) имеет осцилляторный вид. В общем случае это не обязательно так. Для неограниченного (неосцилляторного) решения уравнения (4.19) нам пришлось бы задаться вопросом: «Растет ли накопленная погрешность ε^n медленнее, чем X^n , с увеличением n ?» Для анализа устойчивости ограниченного (осцилляторного) решения нужно ответить лишь на существенно более простой вопрос: «Растет ли ε^n с увеличением n ?» Мы заранее знаем, что X^n никогда не превзойдет амплитуды колебания.

Уравнение (4.32) линейно по ε и поэтому всегда имеет решения в форме $\varepsilon^n = \lambda^n = \exp(i\omega n DT)$, где верхний индекс n у λ обозначает показатель степени (квадрат, куб и т. д.). Подставляя это пробное решение в уравнение (4.32) и деля все члены на λ^{n-1} , получим квадратное уравнение относительно λ

$$\lambda^2 - 2\lambda + 1 = -(\Omega DT)^2 \lambda, \quad (4.33)$$

решениями которого служат две характеристические функции λ_+ и λ_-

$$\lambda_{\pm} = 1 - \frac{(\Omega DT)^2}{2} \pm \left[\frac{(\Omega DT)^2}{2} \right] \left[1 - \frac{4}{(\Omega DT)^2} \right]^{1/2}. \quad (4.34)$$

Общее решение уравнения (4.32) может быть записано в виде

$$\varepsilon^n = a\lambda_+^n + b\lambda_-^n, \quad (4.35)$$

где постоянные a и b определяются начальными погрешностями. В общем случае постоянные a и b , а также характеристические функции λ_+ и λ_- являются комплексными величинами.

Характеристические функции λ_+ и λ_- полностью определяют поведение решения уравнения эволюции погрешности (4.32).

Если $|\lambda_+|$ и $|\lambda_-|$ меньше или равны единице, то любая погрешность, привнесенная на каком-нибудь шаге, будет оставаться малой, и такая схема будет устойчивой. Однако если $|\lambda_+|$ или $|\lambda_-|$ превышают единицу, накопление погрешности происходит экспоненциально, и в конечном счете она забывает решение. Таким образом, критерий устойчивости принимает следующий вид:

Чтобы схема интегрирования по времени была численно устойчивой, характеристические решения λ ее уравнения эволюции погрешности должны лежать внутри или на границе единичного круга, т. е. $|\lambda| \leq 1$.

Вопрос о зависимости устойчивости от шага по времени сводится к задаче нахождения траектории характеристических решений (иначе говоря, корней) λ в комплексной λ -плоскости при изменении временного шага от нуля до бесконечности. Если какой-нибудь корень пересечет единичную окружность, то такая численная схема становится уже неустойчивой. Величина временного шага, при котором это происходит, является границей устойчивости данной разностной схемы.

4.4.1. Метод корневого годографа

При исследовании устойчивости численных схем методом корневого годографа используется представление λ в комплексной плоскости. В этом методе расположение корней рассматривается как функция коэффициента «усиления» Γ . Основное достоинство этого подхода состоит в том, что зачастую общее поведение данной схемы удается установить без нахождения корней λ в явном виде.

Возвращаясь к примеру схемы с перешагиванием, из уравнения (4.33) имеем

$$\frac{(\lambda - 1)(\lambda - 1)}{\lambda} = -\Gamma, \quad (4.36)$$

где

$$\Gamma = (\Omega DT)^2, \quad (4.37)$$

при этом Γ может принимать значения $0 \leq \Gamma \leq \infty$. Для многошаговой схемы общего вида (4.5) коэффициент усиления можно представить в виде

$$\frac{\rho(\lambda)}{\sigma(\lambda)} = \frac{\prod_{i=0}^{l-k} (\lambda - \alpha_i)}{\prod_{i=0}^{l-k} (\lambda - \beta_i)} = -\Gamma, \quad (4.38)$$

где

$$\rho(\lambda) = \sum_{i=0}^k a_{k-i} \lambda^{k-i}; \quad \sigma(\lambda) = \sum_{i=0}^k b_{k-i} \lambda^{k-i},$$

$$\Gamma = (\Omega DT)^2, \quad \Omega = \left(\frac{1}{m} \left| \frac{\partial F}{\partial x} \right|_{\max} \right)^{1/2}.$$

Из вида уравнения (4.38) следует, что при $\Gamma = 0$ (малые DT) его корни лежат около нулей α_i функции $\rho(\lambda)$, а при $\Gamma = \infty$ (большие DT) располагаются около нулей β_i функции $\sigma(\lambda)$ или в бесконечности. Общая многошаговая схема согласована, если $\rho(\lambda)$ имеет двойной корень в точке $\lambda = +1$, и является точной на решении уравнения Ньютона, если по мере увеличения Γ от нуля эти корни движутся по единичной окружности.

Для схемы с перешагиванием (уравнение (4.36)) указанные два предельных случая Γ дают

$$\begin{aligned} \lambda_+ = \lambda_- = 1 & \quad \text{при } \Gamma = 0, \\ \lambda_+ = 0, \quad \lambda_- = -\infty & \quad \text{при } \Gamma = \infty. \end{aligned} \quad (4.39)$$

На комплексной λ -плоскости, изображенной на рис. 4.1, эти точки отмечены символами 0 и ∞ . Следующая задача — проследить траекторию между этими крайними положениями. Логарифмируя (4.38), получаем для вещественной части φ логарифма и мнимой части ψ выражения:

$$\varphi = \sum_i \log |\lambda - \alpha_i| - \sum_i \log |\lambda - \beta_i| = \log \Gamma, \quad (4.40)$$

$$\psi = \sum_i \arg(\lambda - \alpha_i) - \sum_i \arg(\lambda - \beta_i) = \pi. \quad (4.41)$$

Оказывается, что второе соотношение (4.41) более полезно. Оно утверждает, что корень λ перемещается по комплексной плоскости таким образом, что сумма углов, стягиваемых линиями, проведенными в нули α_i , минус сумма углов, стягиваемых линиями, проведенными в полюса β_i , равняется π . Теперь можно отметить участки вещественной оси, где могут располагаться корни. Заметим, что парные корни (или двойные вещественные, или комплексно сопряженные), как и любые корни слева от λ , при вычислении на вещественной оси дают нулевой вклад в суммы (4.41).

В случае схемы с перешагиванием ни один корень не может располагаться на вещественной оси в областях

$+1 < \lambda_r$, поскольку все корни и полюса лежат левее и стягивают нулевой угол, и

$0 < \lambda_r \leq +1$, поскольку каждый из нулей в точке $\alpha = +1$ дает вклад π , что в сумме составляет 2π или нуль радиан,

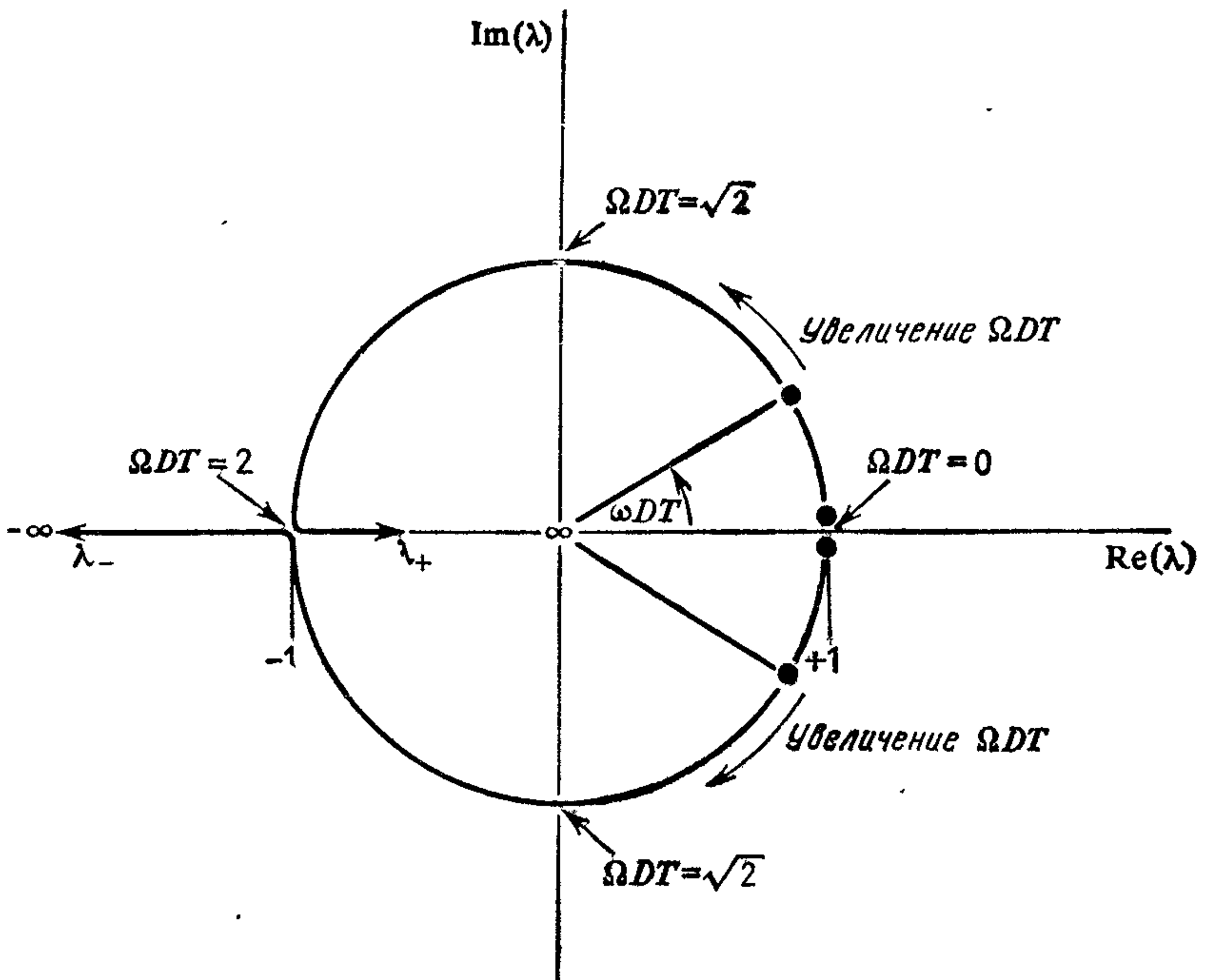


Рис 4.1 Изменения корней λ_+ и λ_- в разностной схеме с перешагиванием при изменении ΩDT от нуля до бесконечности. При $\Omega DT = 0$ и λ_+ и λ_- равны единице. По мере увеличения ΩDT λ_+ и λ_- перемещаются по единичной окружности соответственно против и по часовой стрелке до тех пор, пока снова не становятся вырожденными в точке $\lambda_+ = \lambda_- = -1$ при $\Omega DT = 2$. При $\Omega DT > 2$ корень λ_- движется к $-\infty$, а λ_+ к нулю, причем оба корня остаются вещественными. Поведение λ_+ и λ_- можно связать с частотой разностного аналога гармонического осциллятора, полученного по схеме с перешагиванием, для чего полагают $\lambda = e^{i\omega DT}$ (см. разд. 4.6 и рис. 4.3).

однако корни существуют в области

$-\infty < \lambda_r < 0$, поскольку простой полюс в начале координат дает вклад π радиан.

Корни передвигаются от нулей в точке $\lambda = \pm 1$ к полюсам в начале координат и бесконечности, двигаясь по единичной окружности. В этом нетрудно убедиться, поскольку свободный член уравнения (4.34), который есть произведение корней, равен единице. Раз между $+1$ и -1 корни не могут быть вещественными, они должны составлять пару комплексно сопряженных чисел с модулем единица, а следовательно, должны лежать на единичной окружности.

Из нашей картинки (рис. 4.1) понятно, что корень может покинуть единичную окружность только в том месте, где вещественная ось пересекает единичную окружность, т. е. в точке

$\lambda = -1$. Подставляя это значение в исходное уравнение (4.36), мы сразу получаем границу устойчивости:

$$-\Gamma = -(\Omega DT)^2 = -4 \quad \text{или} \quad \Omega DT = 2. \quad (4.42)$$

Надо сказать, что в случае схемы с перешагиванием мы аппроксимируем дифференциальное уравнение второго порядка разностным уравнением второго порядка. Обычно это желательно, потому что общее решение как дифференциального, так и разностного уравнений может быть представлено в виде линейной комбинации одинакового числа (в данном случае двух) характеристических функций (см. (4.35)). В этом случае каждое характеристическое решение разностного уравнения аппроксимирует одно характеристическое решение дифференциального уравнения.

Более точные разностные аппроксимации неизбежно требуют привлечения большего числа слоев по времени. Это приводит к тому, что число характеристических решений разностного уравнения превышает число характеристических решений физической задачи, и именно эти нефизические паразитные решения часто становятся неустойчивыми в первую очередь, что приводит к более жестким ограничениям на временной шаг, чем в аппроксимациях более низкого порядка.

В качестве примера подобной схемы высокого порядка рассмотрим пятислойную схему, определяемую формулой

$$\frac{x^{n+2} - 2x^n + x^{n-2}}{4DT^2} = \frac{1}{3} [F^{n+1} + F^n + F^{n-1}] + O(DT)^4. \quad (4.43)$$

Эта схема явная, поскольку F^{n+2} не встречается в правой части, и имеет точность порядка DT^4 в отличие от схемы с перешагиванием, точность которой порядка DT^2 .

Уравнение эволюции погрешности для схемы (4.43) имеет вид

$$\varepsilon^{n+2} - 2\varepsilon^n + \varepsilon^{n-2} = -\frac{4(\Omega DT)^2}{3} [\varepsilon^{n+1} + \varepsilon^n + \varepsilon^{n-1}], \quad (4.44)$$

что приводит к следующему характеристическому уравнению:

$$\lambda^4 - 2\lambda^2 + 1 = -\Gamma [\lambda^3 + \lambda^2 + \lambda], \quad (4.45)$$

где

$$\Gamma = 4(\Omega DT)^2/3. \quad (4.46)$$

Теперь для нахождения корней нам надо решить алгебраическое уравнение более высокого порядка. Несмотря на то что в данном случае полное расположение корней можно найти довольно легко (см. рис. 4.2), будет полезно обсудить далее другой метод определения границы устойчивости, который можно всегда применить для уравнений более высокого порядка.

Если посмотреть на разностные схемы, исследуемые в этой

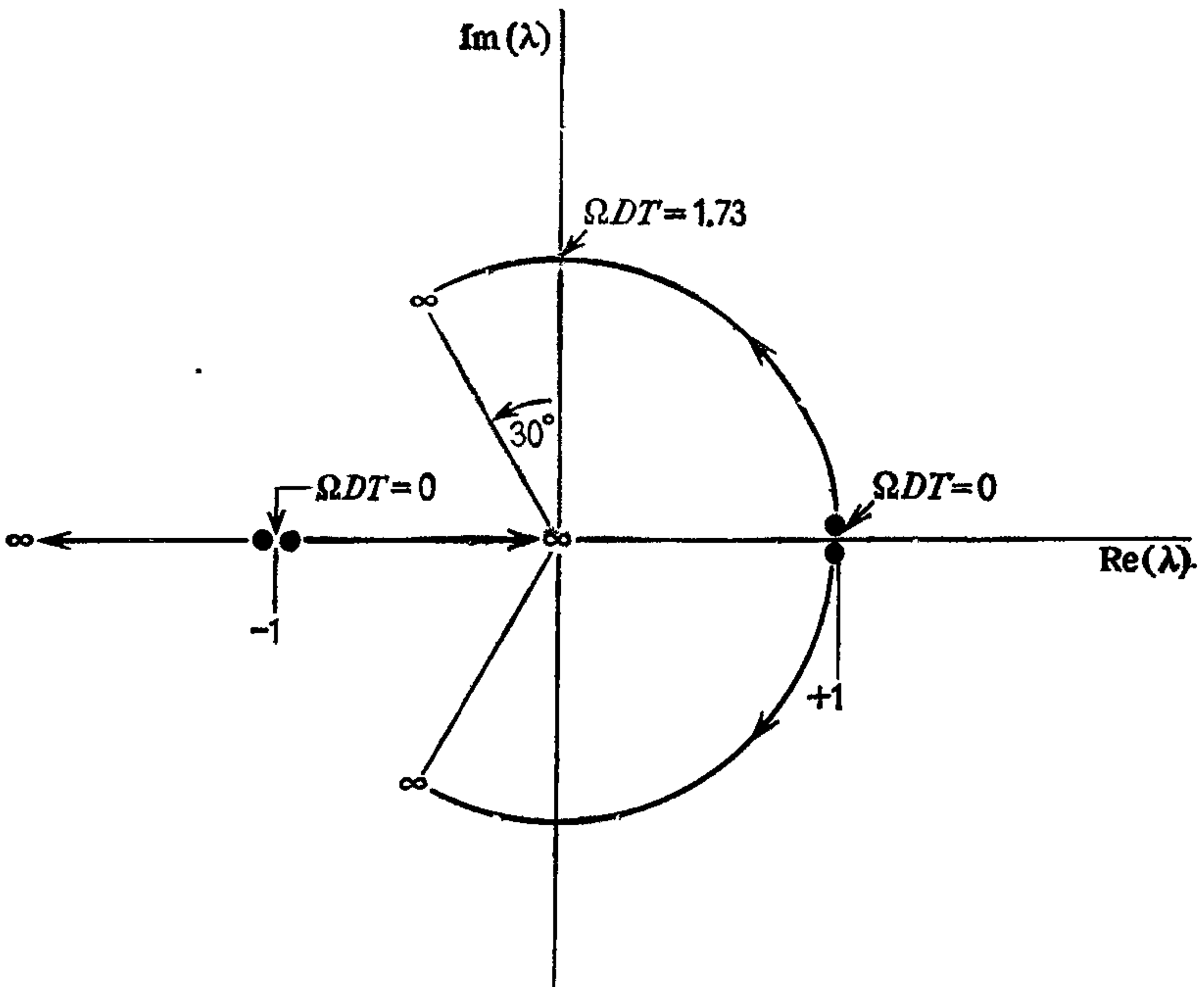


Рис 4.2. Корневой годограф для пятислойной явной схемы (4.43) Символами 0 и ∞ на траекториях корней обозначены значения Γ .

главе, то мы обнаружим, что границе устойчивости всегда отвечает $\lambda = -1$. Поэтому представляется разумным принять это утверждение в качестве начальной гипотезы и попытаться доказать, что любой корень действительно покидает единичную окружность в этой точке. Это можно сделать, проводя разложение $\lambda = -(1 + \beta)$ и доказывая, что при $\beta > 0$ существует решение для Γ в области $0 \leq \Gamma \leq \infty$.

Подставляя $\lambda = -(1 + \beta)$ в уравнение (4.45), получаем с точностью до низшего порядка по β нужное решение для Γ :

$$\Gamma = 4\beta^2. \quad (4.47)$$

Граница устойчивости имеет место при $\beta = 0$, а этому в силу (4.46) отвечает $DT = 0$. Следовательно, мы можем сделать вывод, что пятислойная схема, хотя она и точнее схемы с перешагиванием, неустойчива при любых положительных DT .

Схема расположения корней для пятислойной схемы (4.43) показана на рис. 4.2. Корень, который сразу уходит в область неустойчивости для ненулевого DT , выходит из нулей в точке $\lambda = -1$, которые появились при дифференцировании по удвоенному временному шагу в левой части уравнения (4.43). Эти корни не возникают в более простой схеме с перешагиванием

и являются паразитными. Анализ этой схемы расположения корней мы оставляем читателю в качестве упражнения.

Заканчивая обсуждение метода корневого годографа, важно сказать и о том, какую точность он практически обеспечивает. Для осцилляторных задач непаразитные корни разностных уравнений аппроксимируют корни $\lambda = \exp(\pm \Omega DT)$ дифференциального уравнения. Выбирая шаг по времени равным четверти периода ($\Omega DT = \pi/2 \simeq 1.57$), имеем $\lambda = i$. Подставляя $\lambda = i$ в уравнения (4.36) и (4.45), получим, что для схемы с перешагиванием $\Omega DT = \sqrt{2} \simeq 1.41$, а для пятислойной схемы $\Omega DT = \sqrt{3} \simeq 1.73$. В обоих случаях отличие от точного значения примерно одинаковое. Таким образом, хотя схема (4.43) и точнее схемы с перешагиванием при малых DT , это преимущество быстро теряется по мере увеличения DT (не говоря уже, разумеется, о том, что схема (4.43) безусловно неустойчивая!).

4.4.2. Матрица перехода

Полезный метод получения характеристического уравнения для схем интегрирования по времени, являющихся дискретными аппроксимациями системы дифференциальных уравнений, состоит в том, чтобы сформулировать задачу на собственные значения на основе матрицы перехода G . Чтобы проиллюстрировать использование матрицы перехода, давайте вновь исследуем на устойчивость схему с перешагиванием:

$$x^{n+1} - x^n = v^{n+1/2} DT, \quad (4.48)$$

$$v^{n+1/2} - v^{n-1/2} = \frac{F^n}{m} DT. \quad (4.49)$$

Для получения уравнений эволюции погрешности положим $x = X + \varepsilon_x$ и $v = V + \varepsilon_v$, считая ε достаточно малым, чтобы $F(X^n + \varepsilon_x^n)$ можно было разложить в ряд Тейлора, а затем вычтем из приближенных решений решения, полученные в отсутствие погрешностей округления. Это дает

$$\varepsilon_x^{n+1} - \varepsilon_x^n = \varepsilon_v^{n+1/2} DT, \quad (4.50)$$

$$\varepsilon_v^{n+1/2} - \varepsilon_v^{n-1/2} = \frac{dF^n/dx}{m} DT \varepsilon_x^n = -\Omega^2 DT \varepsilon_x^n. \quad (4.51)$$

Если определить вектор погрешностей

$$\varepsilon^n = \begin{bmatrix} \varepsilon_x^n \\ \varepsilon_v^{n-1/2} \end{bmatrix}, \quad (4.52)$$

то уравнения (4.50) и (4.51) можно записать как одно матричное уравнение

$$\varepsilon^{n+1} - \varepsilon^n = \begin{bmatrix} 0 & DT \\ 0 & 0 \end{bmatrix} \varepsilon^{n+1} + \begin{bmatrix} 0 & 0 \\ -\Omega^2 DT & 0 \end{bmatrix} \varepsilon^n. \quad (4.53)$$

Приводя подобные члены, имеем

$$\begin{bmatrix} 1 & -DT \\ 0 & 1 \end{bmatrix} \mathbf{e}^{n+1} = \begin{bmatrix} 1 & 0 \\ -\Omega^2 DT & 1 \end{bmatrix} \mathbf{e}^n. \quad (4.54)$$

и, умножая на матрицу $\begin{bmatrix} 1 & DT \\ 0 & 1 \end{bmatrix}$, обратную к матрице в левой части, получаем

$$\mathbf{e}^{n+1} = \begin{bmatrix} 1 - \Omega^2 DT^2 & DT \\ -\Omega^2 DT & 1 \end{bmatrix} \mathbf{e}^n, \quad (4.55)$$

т. е.

$$\mathbf{e}^{n+1} = \mathbf{G} \mathbf{e}^n, \quad (4.56)$$

где \mathbf{G} и есть матрица перехода.

Для любой схемы интегрирования по времени можно получить матричное уравнение вида (4.56). Если провести преобразование подобия, которое диагонализует \mathbf{G} , то уравнение (4.56) примет вид

$$(S^{-1} \mathbf{e}^{n+1}) = (S^{-1} \mathbf{G} S) (S^{-1} \mathbf{e}^n), \quad (4.57)$$

или

$$\mathbf{e}'^{n+1} = \mathbf{G}' \mathbf{e}'^n, \quad (4.58)$$

где $\mathbf{e}' = S^{-1} \mathbf{e}$.

Матрица \mathbf{G}' диагональна, и ее диагональные элементы суть собственные значения \mathbf{G} :

$$\mathbf{G}' = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \lambda_3 & \\ & & & \ddots \\ 0 & & & & \ddots \end{bmatrix}. \quad (4.59)$$

Погрешность на временном слое $n+1$ можно связать с внесенной начальной погрешностью \mathbf{e}'^1 с помощью многократного применения формулы (4.58):

$$\mathbf{e}'^{n+1} = \mathbf{G}' \mathbf{e}'^n = (\mathbf{G}')^2 \mathbf{e}'^{n-1} = \dots = (\mathbf{G}')^n \mathbf{e}'^1 \dots, \quad (4.60)$$

где

$$(\mathbf{G}')^n = \begin{bmatrix} \lambda_1^n & & & \\ & \lambda_2^n & & 0 \\ & & \ddots & \\ 0 & & & \ddots \end{bmatrix}. \quad (4.61)$$

Из (4.60) и (4.61) следует, что если какое-нибудь из собственных значений \mathbf{G} лежит вне единичного круга (т. е. $|\lambda| > 1$), то

начальная погрешность ϵ^1 будет расти. В результате эта погрешность растет экспоненциально, приблизительно как $|\lambda|_{\max}^n$, где $|\lambda|_{\max}$ — наибольшее по модулю собственное значение матрицы G .

Количественный критерий устойчивости на языке матрицы перехода G можно сформулировать следующим образом:

Схема интегрирования по времени устойчива, если все собственные значения ее матрицы перехода лежат на единичной окружности или внутри нее, т. е.

$$|\lambda|_{\max} \leq 1. \quad (4.62)$$

Собственные значения G являются корнями уравнения

$$\det(G - \lambda I) = 0. \quad (4.63)$$

Так, например, для схемы с перешагиванием

$$\det(G - \lambda I) = \begin{vmatrix} 1 - \Omega^2 DT^2 - \lambda & DT \\ -\Omega^2 DT & 1 - \lambda \end{vmatrix} = 0 \quad (4.64)$$

дает

$$\lambda^2 - (2 - \Omega^2 DT^2)\lambda + 1 = 0, \quad (4.65)$$

что совпадает с полученным ранее уравнением (см. (4.33)).

В общем случае при выводе уравнений эволюции погрешности получают матричные уравнения вида

$$A\epsilon^{n+1} = B\epsilon^n, \quad (4.66)$$

дающие матрицу перехода $G = A^{-1}B$. Тем не менее можно избежать обращения матрицы A , если воспользоваться тем свойством, что

$$\begin{aligned} \det(G - \lambda I) &= \det(A^{-1}(B - \lambda A)) \\ &= \frac{1}{\det A} \det(B - \lambda A) = 0, \end{aligned} \quad (4.67)$$

$$\Rightarrow \det(B - \lambda A) = 0. \quad (4.68)$$

Вычислять (4.68) для нахождения $\{\lambda\}$ обычно гораздо легче, чем обращать A , и в этом можно будет убедиться на примерах, обсуждаемых в разд. 4.7.

4.5. ЭФФЕКТИВНОСТЬ

Эффективность особенно важна при моделировании методом частица — сетка вследствие очень большого числа (до миллиона!) используемых частиц. Два фактора, которые необходимо принимать во внимание при оценке стоимости той или иной схемы, это память и время. Ограничения по памяти заставляют отдавать предпочтение схемам, которые в процессе продвиже-

ния положений и скоростей на один шаг по времени затрагивают как можно меньше временных слоев. Ограничения по времени требуют схем с небольшим числом операций, приходящихся на частицу за шаг по времени.

Требование устойчивости обычно приводит к ограничению, состоящему в том, чтобы шаг по времени DT был мал по сравнению с характерными периодами рассматриваемой системы. Чтобы достигнуть компромисса между точностью и эффективностью, можно действовать двумя способами — либо использовать схему более высокого порядка с крупным шагом по времени, либо использовать схему низкого порядка с более мелким шагом. Первый подход имеет три недостатка: (1) шаг по времени ограничивается собственными частотами системы, (2) схемы высших порядков в большинстве случаев обладают более жесткими условиями устойчивости по временному шагу и (3) в схемах высокого порядка нужны значения силы на нескольких временных слоях. Как правило, в расчетах систем многих тел наилучший компромисс между точностью, устойчивостью и эффективностью достигается при использовании простых схем второго порядка точности (типа схемы с перешагиванием) с соответствующим выбором шага по времени.

Примером того, к каким бессмысленным затратам может привести недостаточно тщательный анализ ограничений, налагаемых эффективностью, устойчивостью и точностью, служит следующая схема интегрирования:

$$x^{n+1} = x^n + v^n DT + \frac{DT^2}{6} \left[4 \frac{F^n}{m} - \frac{F^{n-1}}{m} \right], \quad (4.69)$$

$$v^{n+1} = v^n + \frac{DT}{6} \left[2 \frac{F^{n+1}}{m} + 5 \frac{F^n}{m} - \frac{F^{n-1}}{m} \right]. \quad (4.70)$$

Предполагалось, что уравнения (4.69) и (4.70) обладают определенными преимуществами по сравнению со схемой с перешагиванием. Они предусматривают примерно в четыре раза больше операций на частицу за шаг по времени, чем схема с перешагиванием, и требуют сохранения ускорений с двух предыдущих временных слоев. Если, однако, вычесть уравнения (4.69), относящиеся к двум последовательным временным слоям,

$$x^n - x^{n-1} = v^{n-1} DT + \frac{DT^2}{6} \left[4 \frac{F^{n-1}}{m} - \frac{F^{n-2}}{m} \right], \quad (4.71)$$

$$x^{n+1} - x^n = v^n DT + \frac{DT^2}{6} \left[\frac{4F^n}{m} - \frac{F^{n-1}}{m} \right], \quad (4.72)$$

то мы найдем, что

$$x^{n+1} - 2x^n + x^{n-1} = (v^n - v^{n-1}) DT + \frac{DT^2}{6} \left[4 \frac{F^n}{m} - 5 \frac{F^{n-1}}{m} + \frac{F^{n-2}}{m} \right]. \quad (4.73)$$

Исключая из (4.70) и (4.73) разность $(v^n - v^{n-1})$, получим

$$x^{n+1} - 2x^n + x^{n-1} = \frac{F^n DT^2}{m}. \quad (4.74)$$

Уравнение (4.74) в точности совпадает с тем, которое было получено из схемы с перешагиванием при исключении из нее скорости! Следовательно, уравнения (4.69) и (4.70) эквивалентны схеме с перешагиванием, только, давая тот же самый результат, они требуют почти в четыре раза больше действий!

4.6. СХЕМА С ПЕРЕШАГИВАНИЕМ ДЛЯ ГАРМОНИЧЕСКОГО ОСЦИЛЛЯТОРА

Уравнение движения частицы, колеблющейся в параболической потенциальной яме, имеет вид

$$\frac{d^2X}{dt^2} = -\Omega^2 X. \quad (4.75)$$

Решение этого уравнения — положение частицы X — синусоида с частотой Ω . В общем случае

$$X = a \cos \Omega t + b \sin \Omega t, \quad (4.76)$$

где постоянные a и b определяются начальными условиями.

Аппроксимация уравнения (4.75) схемой с перешагиванием дает

$$x^{n+1} - 2x^n + x^{n-1} = -\Omega^2 DT^2 x^n. \quad (4.77)$$

Если мы будем искать решения (4.77) в виде $x = x_0 \exp(i\omega t)$, то получим

$$e^{i\omega DT} - 2 + e^{-i\omega DT} = -\Omega^2 DT^2, \quad (4.78)$$

что даст

$$\sin \frac{\omega DT}{2} = \pm \frac{\Omega DT}{2}. \quad (4.79)$$

Следовательно, можно выписать общее решение уравнения (4.77)

$$x^n = a \cos \omega t^n + b \sin \omega t^n, \quad (4.80)$$

где $t^n = nDT$, т. е. схема с перешагиванием (4.77) также дает синусоидальное изменение положений частицы, но с другой частотой ω , определяемой уравнением (4.79). При увеличении шага по времени эта частота ω растет до тех пор, пока не будет достигнута граница устойчивости $\Omega DT = 2$. В точке, отвечающей границе устойчивости, на одном периоде укладывается два временных шага — последовательные значения x принимают поочередно положительные и отрицательные значения. Из (4.79) следует, что за границей устойчивости ω становится комплекс-

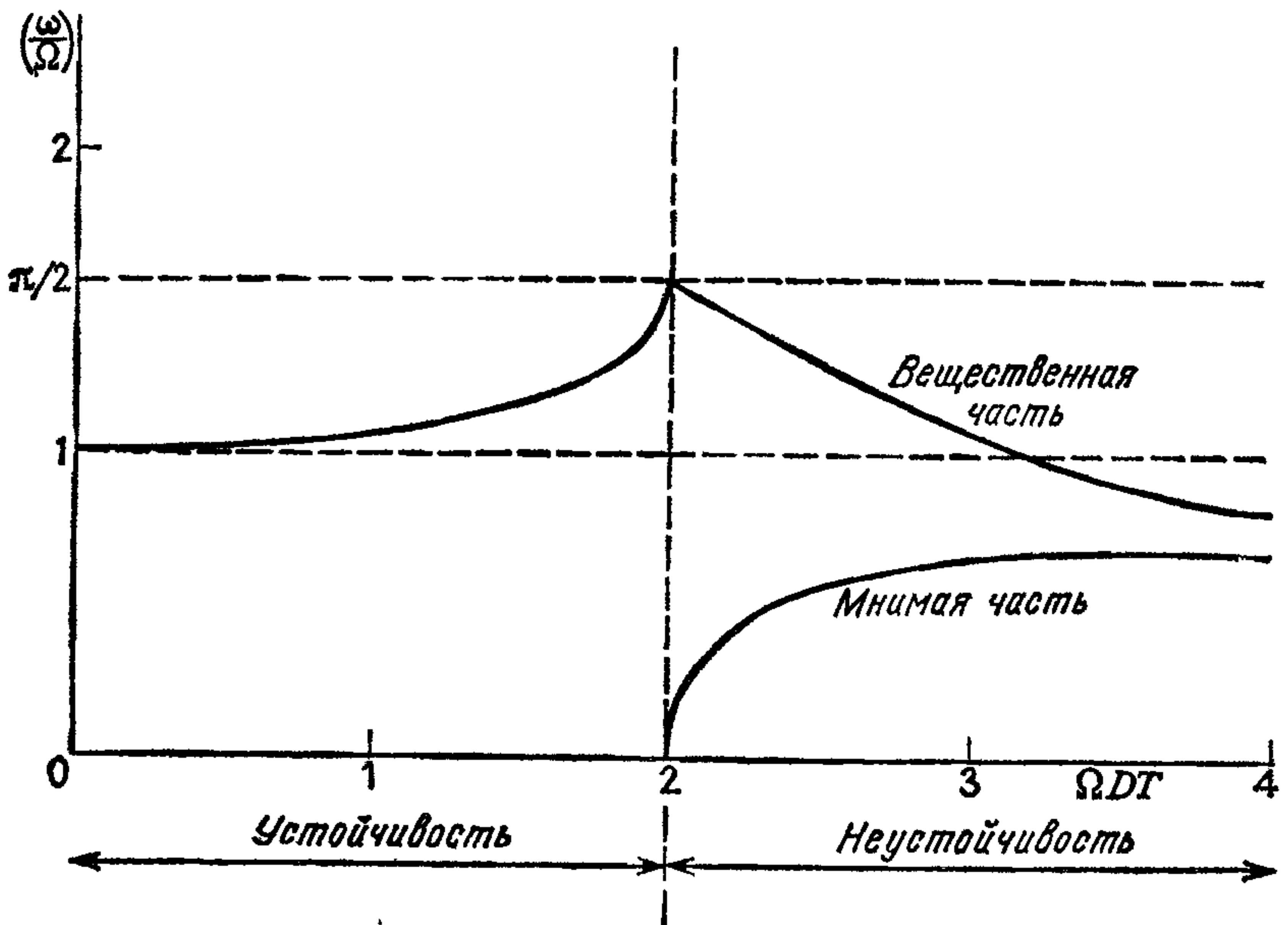


Рис. 4.3. Зависимость частоты ω разностного аналога гармонического осциллятора (схема с перешагиванием) от временного шага. За границей устойчивости частота становится комплексной (что соответствует экспоненциальному росту амплитуды колебаний). Ω — истинная частота гармонического осциллятора.

ной. Подстановка $\omega = \omega_r + i\omega_i$ в (4.79) показывает, что вещественная часть частоты по-прежнему дает два временных шага на период (т. е. $\omega_r DT = \pi$), а мнимая часть определяется уравнением

$$\operatorname{ch} \frac{\omega_i DT}{2} = \pm \frac{\Omega DT}{2}. \quad (4.81)$$

На рис. 4.3 показана зависимость от временного шага частоты гармонического осциллятора в приближении схемы с перешагиванием. Примеры траекторий частицы, рассчитанные по схеме с перешагиванием для нескольких шагов по времени, представлены на рис. 4.4. При $\Omega DT > 2$ амплитуда колебаний экспоненциально растет.

Энергия гармонического осциллятора, описываемого дифференциальным уравнением (4.75), постоянна:

$$E = (V)^2 + \Omega^2 (X)^2 = \text{const}, \quad (4.82)$$

где $V = dX/dt$. У разностного аналога осциллятора, отвечающего схеме с перешагиванием, также существует аналогичная энергетическая постоянная. Из уравнения (4.77) имеем

$$v^{n+1/2} - v^{n-1/2} = -\Omega^2 DT x^n, \quad (4.83)$$

где

$$v^{n+1/2} = (x^{n+1} - x^n)/DT. \quad (4.84)$$

Умножая (4.83) на $(v^{n+1/2} + v^{n-1/2})$, после преобразований получаем

$$E'^n = (v^{n+1/2})^2 + \Omega^2 x^n x^{n+1} = \text{const}. \quad (4.85)$$

Сохранением E' можно объяснить рост амплитуды колебаний при $\Omega DT > 2$. Кроме того, удастся понять рост полной энергии в расчетах методом частица — сетка. Если к правой части уравнения (4.77) добавить случайную силу ADT^2 (представляющую погрешность вычисления силы в методе частица — сетка), то с помощью несложных рассуждений о случайном блуждании можно показать, что

$$E'^n = E'^0 + \langle A^2 \rangle n DT^2. \quad (4.86)$$

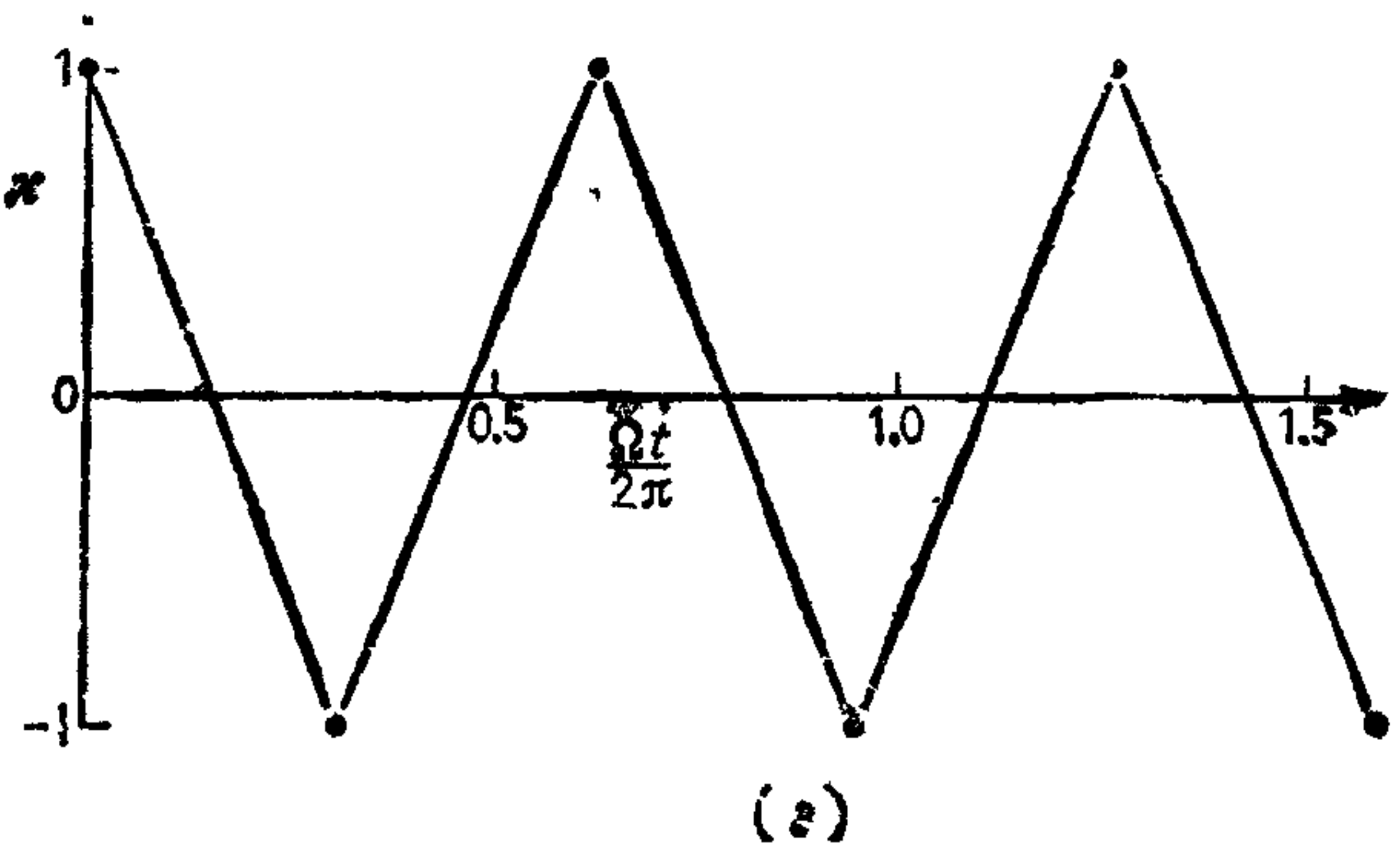
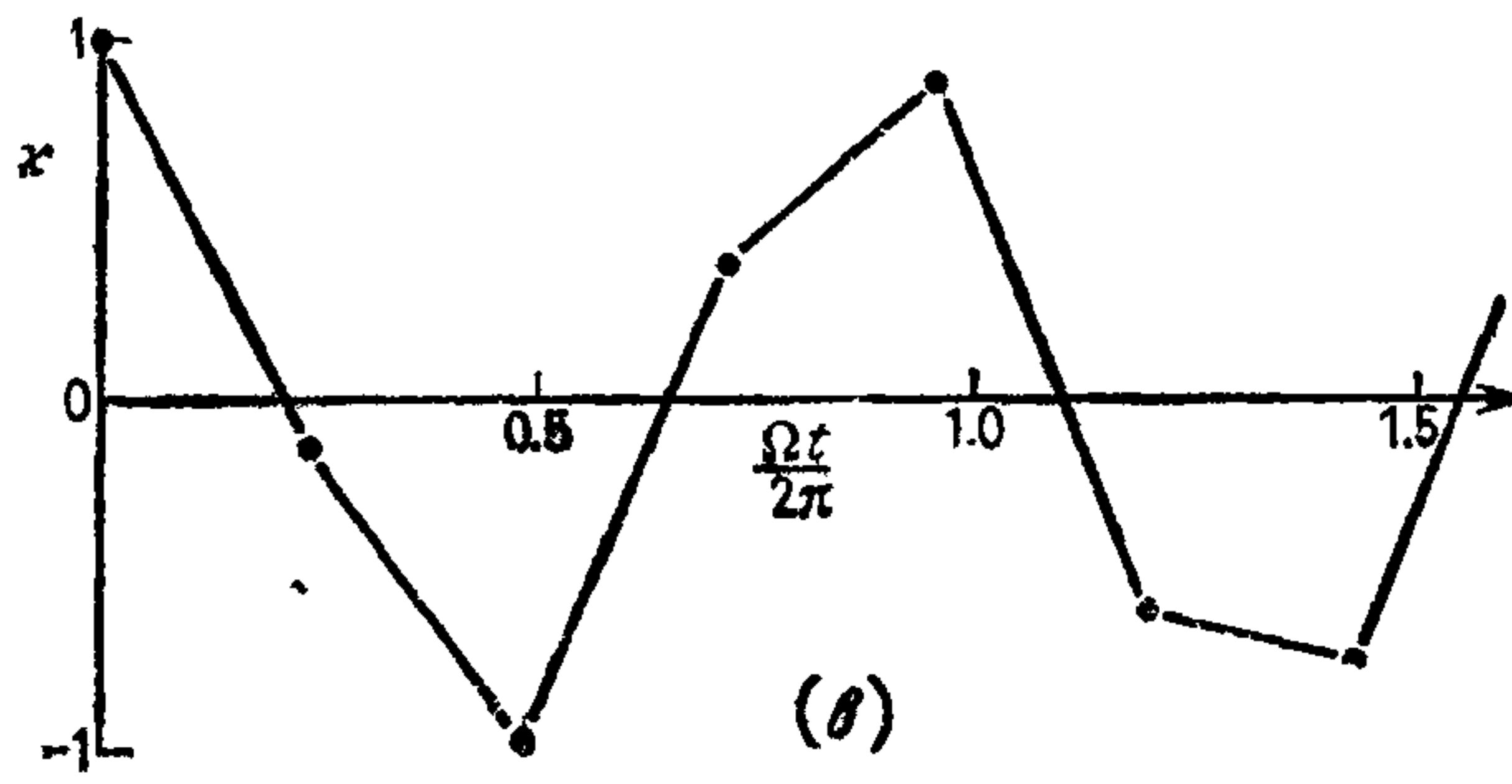
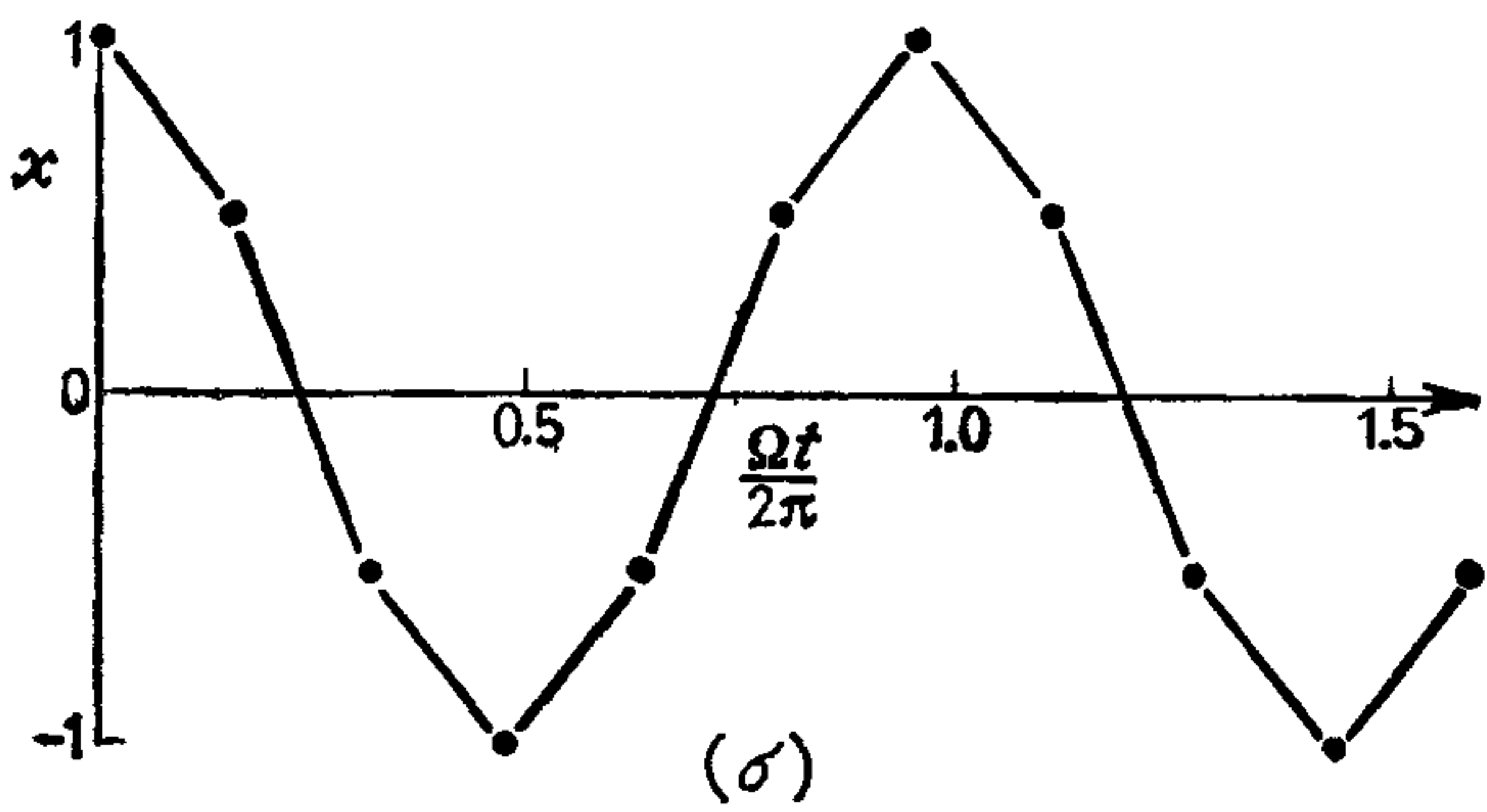
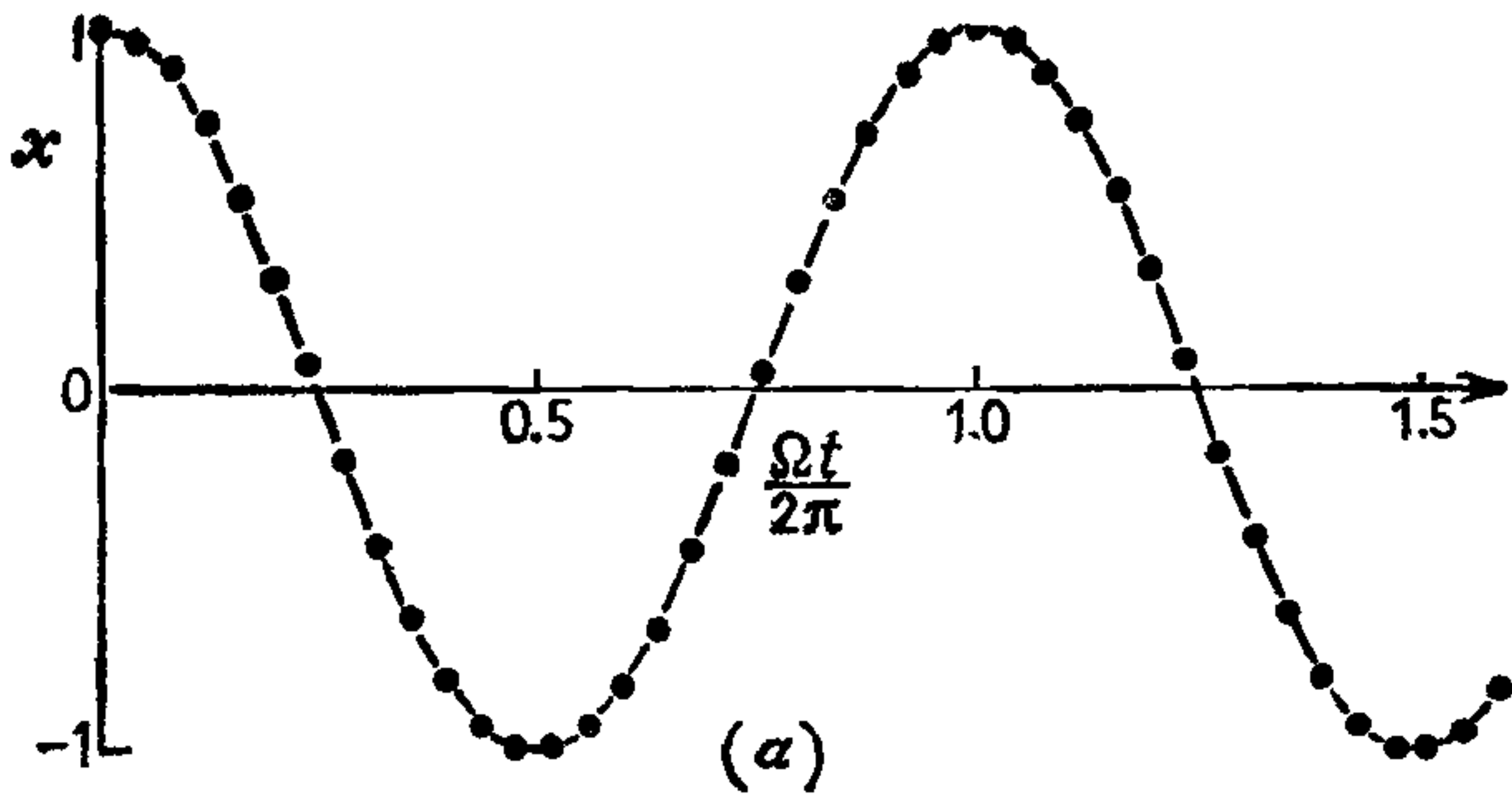
Можно добиться большего по сравнению с конечно-разностной аппроксимацией (4.77). Сравнивая (4.76) с (4.80), мы видим, что уравнение (4.77) дает правильную форму решения, но с неверной частотой. Поэтому, заменяя в правой части (4.77) частоту Ω некоторой другой частотой Ω' , можно добиться того, чтобы решения разностного уравнения в точности совпали с истинным решением. Подставляя в разностное уравнение точное решение (формула (4.76)) уравнения (4.75), имеем

$$X^{n+1} - 2X^n + X^{n-1} = - \left(\Omega \frac{\sin \frac{\Omega DT}{2}}{\frac{\Omega DT}{2}} \right)^2 DT^2 X^n. \quad (4.87)$$

Отсюда, заменяя в уравнении (4.77) Ω на

$$\Omega' = \Omega \left(\frac{\sin \frac{\Omega DT}{2}}{\frac{\Omega DT}{2}} \right), \quad (4.88)$$

мы получаем разностное уравнение, которое воспроизводит значения точного решения. Кроме того, такая модифицированная схема сохраняет устойчивость при всех DT . К сожалению, в случае сил, которые нелинейны по x , коррекция частоты так просто не проходит, хотя похожие схемы с коррекцией и удастся построить для моделирования молекулярной динамики и сгущения галактик, чтобы справиться с трудностями, связанными с наличием быстро меняющихся короткодействующих сил. Заметим в заключение, что в другой интерпретации уравнения (4.87) оно рассматривается как некое алгебраическое соотношение между выборками из решения уравнения (4.75), взятыми с интервалом DT по времени. В гл. 5 мы покажем, что такая



трактовка на языке выборки оказывается чрезвычайно полезной при анализе погрешностей, вносимых дискретизацией по пространству.

4.7.1. ПРИМЕРЫ СХЕМ ИНТЕГРИРОВАНИЯ

4.7.1. Интегрирование силы Лоренца

При наличии магнитного поля \mathbf{B} на частицу с зарядом q , движущуюся со скоростью \mathbf{v} , действует сила Лоренца

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}. \quad (4.89)$$

Положения и скорости находятся путем интегрирования уравнений

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}, \quad \frac{d\mathbf{v}}{dt} = \mathbf{v} \times \boldsymbol{\Omega}, \quad (4.90)$$

где $\boldsymbol{\Omega} = q\mathbf{B}/m$ — циклотронная (или иначе гиромагнитная) частота.

Рассмотрим сначала случай $\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}$, где $\Omega = \text{const}$. Уравнения (4.90) описывают круговые орбиты радиуса $|\mathbf{v}|$ в плоскости (v_x, v_y) и радиуса $|\mathbf{v}|/\Omega$ в плоскости (x, y) , где в обоих случаях угловая частота равняется Ω . Согласованная центрированная по времени конечно-разностная аппроксимация уравнений (4.90) имеет вид

$$\mathbf{x}^{n+1} - \mathbf{x}^n = \mathbf{v}^{n+1/2} DT, \quad (4.91)$$

$$\mathbf{v}^{n+1/2} - \mathbf{v}^{n-1/2} = (\mathbf{v}^{n+1/2} + \mathbf{v}^{n-1/2}) \times \frac{\boldsymbol{\Omega} DT}{2}. \quad (4.92)$$

Эта схема была предложена Бунеманом [57] и применялась Хокни [176] при исследовании аномальной диффузии плазмы (см. гл. 9). Эти уравнения, хотя они и неявные по скорости, путем соответствующей перегруппировки в (4.92) можно легко разрешить относительно новых значений, выразив их через старые (см. ниже). Они обратимы во времени и имеют второй порядок точности. Разностные уравнения дают такую же полную энергию ($=|\mathbf{v}|^2$), что и дифференциальные уравнения. Кроме того, схема (4.91) — (4.92) безусловно устойчива: это можно легко установить методом матрицы перехода, дающим четыре

Рис. 4.4. Временная зависимость положения частицы, движущейся в параболической потенциальной яме в соответствии с разностными уравнениями движения (схема с перешагиванием). Использовались начальные условия: $x^{(0)} = 1$, $v^{(0)} = 0$ ($\Rightarrow v^{-1/2} = \Omega^2 x_0 DT/2$). Кружками отмечены положения частиц на каждом временном шаге. Шаг по времени составлял: (а) $\Omega DT = 0.25$, (б) $\Omega DT = 1$, (в) $\Omega DT = 1.5$ и (г) $\Omega DT = 2$.

корня характеристического уравнения:

$$\begin{aligned}\lambda_1 &= \lambda_2 = 1, \\ \lambda_3 &= \lambda_4^* = \frac{1 + i\Omega DT/2}{1 - i\Omega DT/2}.\end{aligned}\quad (4.93)$$

Первая пара корней (в точке $\lambda = 1$) отвечает уравнению (4.91), тогда как вторая пара корней (которые проходят единичную окружность в виде пары комплексно сопряженных чисел, движущихся из $\lambda = +1$ в $\lambda = -1$ по мере того, как DT увеличивается от 0 до бесконечности) соответствует уравнению (4.92).

Дискретная аппроксимация (4.92) дает скорости, лежащие на окружности радиуса $|\mathbf{v}|$ в плоскости (v_x, v_y) , и положения, лежащие на окружности радиуса R' в плоскости (x, y) . Вследствие конечности шага по времени частота будет превышать истинную частоту Ω (см. разд. 4.6), а радиус R' будет отличаться от циклотронного радиуса $R = |\mathbf{v}|/\Omega$. Частоту ω орбит, описываемых уравнением (4.92), можно найти, положив λ_3 или λ_4 в выражении (4.93) равными $\exp(i\omega DT)$, что дает

$$\operatorname{tg} \frac{\omega DT}{2} = \pm \frac{\Omega DT}{2}.\quad (4.94)$$

Радиус R' можно определить из (4.91) и (4.92), используя (4.94) и тот факт, что $|\mathbf{v}| = \text{const}$:

$$R' = \frac{|\mathbf{v}|}{\Omega} \sec \frac{\omega DT}{2}.\quad (4.95)$$

Как и в случае схемы с перешагиванием для гармонического осциллятора, погрешность частоты можно устранить корректировкой частоты, фигурирующей в разностных уравнениях (см. [176], стр. 93—98). В данном случае, исходя из (4.94), нам нужно заменить в уравнении (4.92) $\Omega DT/2$ на $\operatorname{tg}(\Omega DT/2)$.

Результаты для $\Omega = \text{const}$ непосредственно обобщаются на случай переменной Ω заменой в (4.92) Ω на $\Omega^n = \Omega(x^n)$. Если наряду с этим имеется электрическое поле, то можно объединить схему с перешагиванием для электрической силы со схемой интегрирования силы Лоренца, что приводит к схеме

$$\mathbf{x}^{n+1} - \mathbf{x}^n = \mathbf{v}^{n+1/2} DT,\quad (4.96)$$

$$\mathbf{v}^{n+1/2} - \mathbf{v}^{n-1/2} = \frac{q\mathbf{E}^n}{m} DT + (\mathbf{v}^{n+1/2} + \mathbf{v}^{n-1/2}) \times \frac{\Omega^n DT}{2}.\quad (4.97)$$

Для компонент уравнений (4.96) и (4.97), параллельных магнитному полю, получается просто схема с перешагиванием, и поэтому для них должно выполняться условие устойчивости схемы с перешагиванием. Компоненты, перпендикулярные магнитному полю, безусловно устойчивы и при этом обладают тем

важным свойством, что при больших шагах по времени переходят в уравнение для адиабатического дрейфа:

$$\bar{\mathbf{v}}^n = \frac{(\mathbf{v}^{n+1/2} + \mathbf{v}^{n-1/2})}{2} = \frac{\mathbf{E}^n \times \mathbf{B}^n}{|\mathbf{B}^n|^2} + O(DT^{-1}). \quad (4.98)$$

Этим свойством воспользовались Леви и Хокни [251] при изучении электронных пучков в скрещенных полях, и оно также широко использовалось Бэрдсомом и его сотрудниками ([36]).

Вводя две промежуточные скорости \mathbf{v}_1^* и \mathbf{v}_2^* , можно разделить в уравнении (4.97) члены, описывающие ускорение в электрическом поле, и члены, связанные с вращением (силу Лоренца):

$$\mathbf{v}_1^* = \mathbf{v}^{n-1/2} + \frac{q}{m} \mathbf{E}^n \frac{DT}{2}, \quad (4.99)$$

$$\mathbf{v}_2^* = \mathbf{v}_1^* + (\mathbf{v}_2^* + \mathbf{v}_1^*) \times \frac{\Omega^n DT}{2}, \quad (4.100)$$

$$\mathbf{v}^{n+1/2} = \mathbf{v}_2^* + \frac{q}{m} \mathbf{E}^n \frac{DT}{2}. \quad (4.101)$$

Уравнение (4.100) можно еще упростить, для чего умножим его векторно на $\Omega DT/2$ и исключим двойное векторное произведение, содержащее \mathbf{v}_2^* . Это дает

$$\mathbf{v}_2^* = \mathbf{v}_1^* + \frac{2}{1 + \left(\frac{\Omega DT}{2}\right)^2} \mathbf{v}_3^* \times \Omega, \quad (4.102)$$

где

$$\mathbf{v}_3^* = \mathbf{v}_1^* + \mathbf{v}_1^* \times \frac{\Omega DT}{2}. \quad (4.103)$$

Полученные факторизованные уравнения ((4.99), (4.103), (4.102) и (4.101)) лежат в основе алгоритма CYLRAD, принадлежащего Бóрису [41].

Введенный в случае постоянной Ω поправочный множитель для частоты можно также включить и в более общую формулу (4.97). Если считать, что \mathbf{E} и Ω в пределах одного шага по времени приблизительно постоянны, то уравнение движения

$$\dot{\mathbf{v}} = \frac{q\mathbf{E}}{m} + \mathbf{v} \times \Omega \quad (4.104)$$

можно проинтегрировать по этому малому временному интервалу. Разбивая скорость на параллельную (\mathbf{v}_{\parallel}) и перпендикулярную (\mathbf{v}_{\perp}) магнитному полю компоненты и представляя \mathbf{v}_{\perp} в виде суммы скорости дрейфа \mathbf{u} и циклотронной скорости \mathbf{w} ,

получим

$$\mathbf{v}_{\parallel} = \mathbf{v}_{\parallel}^0 + \frac{q}{m} \mathbf{E}_{\parallel} t, \quad (4.105)$$

$$\mathbf{v}_{\perp} = \mathbf{u} + \mathbf{w}, \quad (4.106)$$

$$\mathbf{u} = \frac{q}{m} \frac{\mathbf{E} \times \boldsymbol{\Omega}}{|\boldsymbol{\Omega}|^2}, \quad (4.107)$$

$$\dot{\mathbf{w}} = \mathbf{w} \times \boldsymbol{\Omega}. \quad (4.108)$$

Интегрируя (4.108) на отрезке $[(n - 1/2)DT, (n + 1/2)DT]$, исключая с помощью (4.106) и (4.107) \mathbf{w} и \mathbf{u} , находим:

$$\begin{aligned} \left(\mathbf{v}_{\perp}^{n+1/2} - \frac{q}{m} \frac{\mathbf{E}^n \times \boldsymbol{\Omega}^n}{|\boldsymbol{\Omega}^n|^2} \right) &= \left(\mathbf{v}_{\perp}^{n-1/2} - \frac{q}{m} \frac{\mathbf{E}^n \times \boldsymbol{\Omega}^n}{|\boldsymbol{\Omega}^n|^2} \right) \cos \Omega^n DT \\ &+ \left(\mathbf{v}_{\perp}^{n-1/2} - \frac{q}{m} \frac{\mathbf{E}^n \times \boldsymbol{\Omega}^n}{|\boldsymbol{\Omega}^n|^2} \right) \times \boldsymbol{\Omega}^n \frac{\sin \Omega^n DT}{\Omega^n}. \end{aligned} \quad (4.109)$$

После упрощения получим

$$\mathbf{v}_{\perp}^{n+1/2} = \mathbf{v}_{\perp}^{n-1/2} + \alpha \left[\frac{q \mathbf{E}^n}{m} DT + (\mathbf{v}_{\perp}^{n+1/2} + \mathbf{v}_{\perp}^{n-1/2}) \times \frac{\boldsymbol{\Omega}^n DT}{2} \right], \quad (4.110)$$

что совпадает с уравнением (4.97) с точностью до поправочного множителя

$$\alpha = \frac{\operatorname{tg} \frac{\Omega DT}{2}}{\frac{\Omega DT}{2}}. \quad (4.111)$$

4.7.2. Схемы интегрирования уравнений с вязкими силами

В моделях частиц, не учитывающих инерцию, уравнения движения Ньютона заменяются уравнениями первого порядка, описывающими релаксацию к стационарному состоянию. К таким моделям относятся, например, модель полупроводника с сильными столкновениями, гидродинамическая модель вихрей и плазменная модель ведущего центра (см. разд. 1.4 и табл. 1.1). В таких случаях скорость определяется действующей силой:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}). \quad (4.112)$$

Для линейной \mathbf{f} мы имеем

$$\frac{d\mathbf{x}}{dt} = -\alpha \mathbf{x}, \quad \alpha > 0. \quad (4.113)$$

Одна из возможных конечно-разностных аппроксимаций уравнения (4.113) имеет вид

$$\mathbf{x}^{n+1} - \mathbf{x}^{n-1} = -2\alpha \mathbf{x}^n DT. \quad (4.114)$$

Схема (4.114) явная, согласованная, центрированная по времени и обратимая, и к тому же имеет второй порядок точности. Несмотря на это, она приводит к уравнению эволюции погрешности

$$\varepsilon^{n+1} - \varepsilon^{n-1} = -2\alpha DT \varepsilon^n, \quad (4.115)$$

характеристическое уравнение которого имеет корни

$$\lambda_{\pm} = -\alpha DT \pm (1 + \alpha^2 DT^2)^{1/2}. \quad (4.116)$$

Поскольку $\lambda_- < -1$, такая схема безусловно неустойчива. Это еще один пример, когда нефизическое паразитное решение быстро забывает спадающее решение, имеющее физический смысл. Появление корня, отвечающего паразитной неустойчивости, связано с тем, что дифференциальное уравнение первого порядка (4.112) было аппроксимировано разностным уравнением (4.114) более высокого (второго) порядка.

Для аппроксимации уравнения (4.112) больше подходит следующая неявная схема второго порядка точности:

$$\mathbf{x}^{n+1} - \mathbf{x}^n = -\alpha \frac{(\mathbf{x}^{n+1} + \mathbf{x}^n)}{2} DT. \quad (4.117)$$

Поскольку в этой схеме фигурируют только два временных слоя, она представляет собой разностную аппроксимацию первого порядка дифференциального уравнения первого порядка. Уравнение (4.117) дает характеристическое уравнение

$$\lambda = \frac{1 - \frac{\alpha DT}{2}}{1 + \frac{\alpha DT}{2}}, \quad (4.118)$$

означающее, что $|\lambda| \leq 1$ для всех DT , т. е. схема безусловно устойчива. Недостаток этой схемы состоит в том, что в случае, когда действующие силы нелинейны, на каждом слое по времени нужно решать неявное разностное уравнение

$$\mathbf{x}^{n+1} - \mathbf{x}^n = [f(\mathbf{x}^{n+1}) + f(\mathbf{x}^n)] \frac{DT}{2}. \quad (4.119)$$

Поэтому такая схема, как правило, неэффективна для расчета систем многих тел.

На рис. 4.5 для сравнения приведены результаты расчетов по двум центрированным по времени схемам второго порядка точности: явной схеме (4.114) и неявной схеме (4.117). Видно, что решение, полученное по неявной схеме, экспоненциально убывает к нулю, как и точное решение. Решение, полученное по явной схеме, также сначала убывает, но через всего семь шагов по времени оно поглощается экспоненциально растущим осцилляторным паразитным решением, и к концу расчета ре-

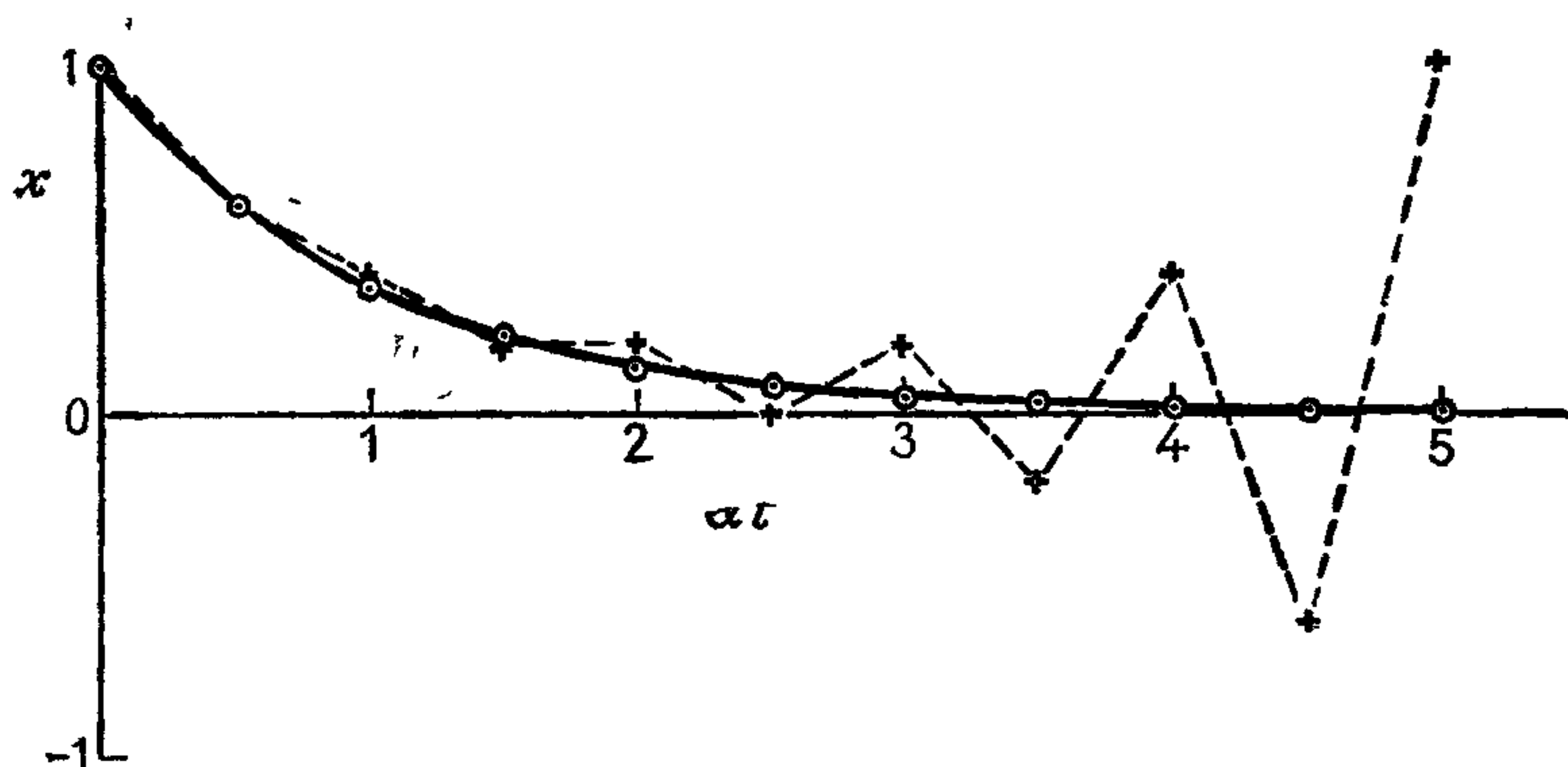


Рис. 4.5. Интегрирование уравнения $\dot{x} = -\alpha x$. Непрерывная кривая отвечает точному решению $e^{-\alpha t}$, крестики, соединенные штриховой линией, — решению по схеме (4.114), а кружки — решению по схеме (4.117). Использовалось начальное условие $x^{(0)} = 1$, а шаг по времени был таким, что $\alpha DT = 1/2$. Начальные значения для схемы (4.114) находились с помощью (4.117).

шение по явной схеме не имеет ничего общего с решением дифференциального уравнения. Результаты, полученные по явной схеме, бессмысленны, поскольку она численно неустойчива, хотя эта схема и центрирована по времени, и имеет второй порядок точности.

Для того чтобы избежать паразитных решений, характеристическое уравнение должно быть того же порядка, что и дифференциальное уравнение. Единственной согласованной явной разностной аппроксимацией уравнений (4.112) и (4.113) этого вида является схема Эйлера

$$x^{n+1} - x^n = -\alpha x^n DT, \quad (4.120)$$

характеристическое уравнение которой имеет вид

$$\lambda = 1 - \alpha DT, \quad (4.121)$$

что приводит к условию устойчивости

$$\alpha DT < 2. \quad (4.122)$$

Заметим, что, как и для частоты в осцилляторных случаях, можно подобрать коэффициент затухания в линейном случае так, чтобы улучшить конечно-разностные аппроксимации. Что касается описанных здесь случаев, то для нахождения поправочного множителя можно положить физический корень характеристического уравнения равным $\lambda = \exp(-\alpha t)$, а затем найти такое α' для использования в разностном уравнении, чтобы в характеристическом уравнении $\alpha = \alpha'$.

Еще один метод решения диссипативного уравнения состоит в том, что инерциальный член сохраняют, получают какую-нибудь строго устойчивую схему, а затем используют крупный шаг

по времени. Мы уже встречались с таким подходом в разд. 4.7.1 при рассмотрении уравнения движения ведущего центра. Для задачи о полупроводнике мы можем аппроксимировать дифференциальные уравнения

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = \frac{qE}{m^*} - \frac{v}{\tau} \quad (4.123)$$

разностными уравнениями

$$x^{n+1} = x^n + v^{n+1/2} DT, \quad (4.124)$$

$$v^{n+1/2} - v^{n-1/2} = \frac{qE^n}{m^*} DT - \frac{(v^{n+1/2} + v^{n-1/2})}{2} \frac{DT}{\tau}, \quad (4.125)$$

где при большом шаге по времени ($DT \gg \tau$) уравнение (4.125) переходит в

$$\frac{(v^{n+1/2} + v^{n-1/2})}{2} = \bar{v}^n = \frac{q\tau E^n}{m^*} = \mu E^n, \quad (4.126)$$

которое представляет собой просто искомое дрейфовое приближение.

С математической точки зрения уравнению (4.123) эквивалентно уравнение движения галактик в сопутствующих координатах:

$$\dot{V}' + 2NV' = F', \quad (4.127)$$

где N — постоянная Хаббла, а F' — приведенная гравитационная сила. В соответствующих модельных расчетах, о которых пойдет речь в разд. 11.4.3, для аппроксимации уравнения (4.127) была использована схема, аналогичная схеме (4.124)—(4.125).

4.7.3. Экономичные по памяти схемы Рунге — Кутты

Явный метод Рунге — Кутты ([171], гл. 2) принадлежит к одному из самых популярных методов интегрирования обыкновенных дифференциальных уравнений. С этим названием связывают целое семейство неитерационных одношаговых методов, которые могут строиться с высоким порядком аппроксимации и в которых не требуется знания ни одного из предыдущих значений решения. Неявные методы Рунге — Кутты, разработанные Бутчером [70], хотя они и точнее, обычно меньше подходят для больших систем, возникающих в методе частиц, из-за того, что для нахождения значений решения на новом шаге нужно решить итерационным образом нелинейные уравнения, если только силы не зависят линейно от расстояния (что вряд ли типично).

В схемах Рунге — Кутты скорость и положение определяются в один и тот же момент времени, шаг по времени можно с легкостью менять, и обычно погрешность аппроксимации меньше, чем у конечно-разностных схем, которые обсуждались выше в

этой главе. Их главным практическим недостатком является необходимость предусматривать память для нескольких вспомогательных векторов, каждый длиной, равной количеству переменных интегрирования. По сравнению с этим в более простой конечно-разностной схеме с перешагиванием новые значения можно записывать на место старых и никакой вспомогательной памяти не требуется. Объем дополнительной памяти, требуемой в методе Рунге — Кутты, можно свести к минимуму тщательным выбором констант самой схемы либо тщательным планированием арифметических операций. Способы минимизации памяти обсуждаются в работах [387] и [346]. Далее мы опишем схемы четвертого и третьего порядка, которые построены таким образом, чтобы память была минимальна.

Все задачи интегрирования по времени, рассмотренные в этой главе, представляют собой частные случаи общей задачи

$$\frac{dx}{dt} = f(x), \quad (4.128a)$$

где x — вектор длины N^1), содержащий все переменные, по которым ведется интегрирование, а $f(x)$ — функция этих переменных, как правило, нелинейная. Если она явным образом зависит от времени, это учитывается добавлением t в список переменных, представленных в векторе x , и введением дополнительного уравнения

$$\frac{dt}{dt} = 1. \quad (4.128b)$$

В нашем случае $f(x)$, как правило, не бывает явной функцией времени, за исключением тех случаев, когда, например, имеется зависящее от времени внешнее магнитное поле, поэтому при оценке требуемой памяти мы будем считать это условие выполненным. В случае уравнений Ньютона (4.1) и (4.2) вектор x имеет вид

$$x = \{x_i, y_i, z_i, v_{xi}, v_{yi}, v_{zi}; i = 1, \dots, N_p\}, \quad (4.129a)$$

где $N = 6N_p$, i — номер частицы, а N_p — полное число частиц. С другой стороны, в случае движения под действием силы внутреннего трения (4.112) скорости хранить не надо, и тогда

$$x = \{x_i, y_i, z_i; i = 1, \dots, N_p\}, \quad (4.129b)$$

$$N = 3N_p.$$

В обоих случаях x зависит от времени неявно через изменение значений координат частиц.

¹⁾ Имеется в виду N -мерный вектор, представляемый в программе массивом из N элементов. — Прим. перев.

В классической схеме Рунге — Кутты четвертого порядка [338, 231] вычисляется градиент в четырех точках, находящихся в окрестности решения, по формулам

$$\begin{aligned} k_1 &= DT f(x^n), \\ k_2 &= DT f\left(x^n + \frac{1}{2} k_1\right), \\ k_3 &= DT f\left(x^n + \frac{1}{2} k_2\right), \\ k_4 &= DT f(x^n + k_3) \end{aligned} \quad (4.130a)$$

и значение x на новом временном слое получается в итоге из взвешенного среднего этих значений:

$$x^{n+1} = x^n + \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4). \quad (4.130b)$$

Погрешность аппроксимации в данном случае пропорциональна DT^4 . Казалось бы, требуемый в этом методе объем памяти составляет $4N^1$). Один вектор длиной N можно использовать для хранения как векторов x^n , так и x^{n+1} , поскольку при вычислении выражения (4.130b) последний можно записывать на место первого. Кроме этого, необходимы, по-видимому, три вспомогательных вектора: один — для накопления суммы векторов k_i в соответствии с (4.130b), а второй и третий — для хранения аргумента функции $f(x)$ и ее значения в процессе последовательных вычислений k_i по формулам (4.130a).

Тем не менее Блам показал [40], что можно уменьшить расход памяти до $3N$ путем организации вычислений в следующем порядке:

$$\begin{aligned} z_0 &= x^n, \\ p_0 &= DT f(z_0), \\ z_1 &= z_0 + p_0/2, \\ q_1 &= p_0, \\ p_1 &= DT f(z_1), \\ z_2 &= z_1 + p_1/2 - q_1/2, \\ q_2 &= q_1/6, \\ p_2 &= DT f(z_2) - p_1/2, \\ z_3 &= z_2 + p_2, \\ q_3 &= q_2 - p_2, \\ p_3 &= DT f(z_3) + 2p_2, \end{aligned} \quad (4.131)$$

¹⁾ Память измеряется в словах — единицах хранения чисел. — Прим. перев.

и, наконец, $\mathbf{x}^{n+1} = \mathbf{z}_4 = \mathbf{z}_3 + \mathbf{q}_3 + \mathbf{p}_3/6$. Выражения (4.131) и (4.130) алгебраически эквивалентны, но теперь совершенно ясно, что требуются только три вектора — один для всех \mathbf{z}_i , один для всех \mathbf{q}_i и один для всех \mathbf{p}_i . Вычисления ведутся в следующей последовательности: \mathbf{z}_{i+1} записываются на место \mathbf{z}_i , \mathbf{q}_{i+1} записываются на место \mathbf{q}_i и \mathbf{p}_{i+1} записываются на место \mathbf{p}_i для $i = 0, 1, 2, 3$. В результате общий объем памяти, который требуется для интегрирования трехмерных уравнений Ньютона, составляет $18N_p$, по сравнению с $6N_p$ для схемы с перешагиванием. Ранее Гилл [143] показал, как с другим набором констант в схеме Рунге — Кутты получается метод четвертого порядка, который требует только $3N$ слов памяти. Позднее Файф [139] доказал, что всякую схему Рунге — Кутты четвертого порядка можно организовать так, чтобы требуемый объем памяти составлял $3N$.

Если число частиц велико, скажем, больше 10 000, то, по видимому, не стоит использовать схему, которая требует как минимум в три раза больше памяти, чем N . Для уравнений Ньютона в трехмерном случае это составляет $> 180\,000$ чисел, или 0.72 Мбайт при 32-битовой точности. Несмотря на это, Уильямсон [387] рассмотрел класс схем второго, третьего и четвертого порядка, которые требуют полную память $2N$. Он рекомендует следующую схему третьего порядка:

$$\begin{aligned}
 \mathbf{z}_0 &= \mathbf{x}^n, \\
 \mathbf{q}_1 &= DT f(\mathbf{z}_0), \\
 \mathbf{z}_1 &= \mathbf{z}_0 + \frac{1}{3} \mathbf{q}_1, \\
 \mathbf{q}_2 &= DT f(\mathbf{z}_1) + \frac{5}{9} \mathbf{q}_1, \\
 \mathbf{z}_2 &= \mathbf{z}_1 + \frac{15}{16} \mathbf{q}_2, \\
 \mathbf{q}_3 &= DT f(\mathbf{z}_2) - \frac{153}{128} \mathbf{q}_2, \\
 \mathbf{x}^{n+1} &= \mathbf{z}_3 = \mathbf{z}_2 + \frac{8}{15} \mathbf{q}_3.
 \end{aligned} \tag{4.132}$$

Ясно, что в процессе вычислений по формулам (4.132) нужно хранить только два вектора (один для всех \mathbf{q}_i и один для всех \mathbf{z}_i), поскольку \mathbf{q}_{i+1} можно записывать на место \mathbf{q}_i , а \mathbf{z}_{i+1} — на место \mathbf{z}_i . Если эту схему применить к уравнениям с вязкой силой (4.112), то потребность в памяти составит $6N_p$, а это совпадает с тем, что требуется при учете инерциального члена в соответствии с уравнениями (4.123) — (4.125). Однако эта схема Рунге — Кутты имеет погрешность аппроксимации, пропорциональную DT^3 , а не DT^2 , как в схеме с сохранением инерциального члена.

Схема Уильямсона (4.132) была использована для решения двумерных уравнений ведущего центра (или вихрей) для пары частиц, движущихся в одной плоскости. Удобно представить эти уравнения в комплексной записи:

$$\frac{dr}{dt} = i \frac{r}{|r|^2}. \quad (4.133)$$

Здесь комплексная переменная r представляет собой расстояние между частицами. Точной орбитой является окружность, причем частота обратно пропорциональна квадрату радиуса. При использовании шага по времени в $1/8$ периода относительная погрешность радиуса после интегрирования в течение одного периода составила 0.001% . Указанная схема Рунге — Кутты третьего порядка, примененная для моделирования $10\,000$ частиц, сохраняет потенциальную энергию (которая должна сохраняться точно) с погрешностью 0.25% через 350 шагов по времени.

Вычисление силы в методе частица — сетка

5.1. ВВЕДЕНИЕ

В гл. 2 и 3 для иллюстрации проектирования и разработки программ использовался простейший вариант метода частица — сетка, а именно схема NGR (ближайшего сеточного узла). В данной главе мы узнаем, как на основе этой схемы можно перейти к схемам более высокого порядка точности и больших размерностей. Для определенности мы будем рассматривать заряженные системы, хотя основные выводы с равным успехом относятся и к другим системам, для которых применим метод частица — сетка (см. гл. 1).

Напомним, что, как говорилось в гл. 1, расчет методом частица — сетка состоит из четырех основных этапов:

1. Распределение заряда в узлы сетки.
2. Решение на этой сетке уравнения для поля.
3. Вычисление сеточного поля сил.
4. Нахождение сил, действующих на частицы, с помощью интерполяции.

При этом силы, получаемые на этапе 4, используются для интегрирования уравнений движения (см. гл. 4). На каждом этапе вносятся погрешности. На этапе 1 распределение заряда частиц, положения которых изменяются непрерывным образом во всей расчетной области, заменяется конечным набором значений плотности заряда. Очевидно, что такой конечный набор значений в состоянии передать только распределение плотности, которое является в некотором смысле «гладким»; в дальнейшем мы поясним, что имеется в виду (разд. 5.3.2 и 5.6.2). На этапах 2 и 3 за счет замены дифференциальных соотношений разностными (или интегралов суммами) вносятся погрешности аппроксимации. Погрешности последнего этапа, т. е. этапа 4, обусловлены интерполяцией. В результате получается приближенная сила, и ее погрешности искажают далее интегрирование по времени из-за неточного определения ускорения (для схем интегрирования сильностолкновительных систем из-за неточности в скорости) в уравнениях движения.

Качество конкретной схемы расчета силы определяется суммарной погрешностью силы, а не составляющими ее погреш-

ностями, ибо только полная сила и определяет динамику системы. На каждом из указанных этапов вычисления силы допустимы достаточно большие погрешности, если в конечном счете они дадут точную силу. Чтобы понять, как добиться такой взаимной компенсации погрешностей, необходимо представлять основные свойства погрешностей, вносимых на каждом этапе, и изменения этих погрешностей при простых модификациях численных схем.

Для исследования алгоритмов частица — сетка мы используем два дополняющих друг друга (и до некоторой степени эквивалентных) метода. В первом случае рассмотрение ведется в x -пространстве, и схемы упорядочены по степеням малого параметра, в качестве которого используется размер ячейки H . Для определения свойств погрешностей привлекаются такие понятия, как гладкость, порядок аппроксимации и угловая анизотропия. Во втором методе операции, выполняемые на разных этапах вычисления, формулируются на языке преобразований Фурье; здесь, чтобы пролить свет на свойства соответствующих приближений, используются понятия, более привычные в обработке сигналов — выборка, наложение, свертки. Первый подход применяется главным образом для выяснения влияния изменений отдельных элементов схемы на качество и вычислительную трудоемкость данной аппроксимации, тогда как второй, более мощный метод, можно использовать с целью оптимального сочетания этих элементов.

В следующих трех разделах этой главы используется анализ только в x -пространстве. В последнем разделе (5.6) вводится спектральный формализм. В дальнейшем (гл. 7 и 8) с помощью спектрального описания и методов математической физики мы покажем, каким образом можно получить всестороннее описание «физики» численной модели и как такую «физику» можно оптимизировать.

5.2. СИЛЫ В ОДНОМЕРНОМ СЛУЧАЕ

Рассмотрим вначале простую периодическую систему, а именно одномерную (т. е. плоскосимметричную, см. гл. 2) систему двух одинаковых по величине зарядов разных знаков. Несмотря на простоту такой системы, результаты, которые мы получим, справедливы и для системы многих частиц, поскольку силы аддитивны, а более общие граничные условия эквивалентны фиктивным зарядам (изображениям).

5.2.1. Непрерывная система

Пусть имеется непрерывная периодическая система с длиной периода L , в которой заряд q находится в точке $x = x_1 = x'$, а заряд $-q$ в точке $x = x_2 = -x'$. Тогда потенциал определяется

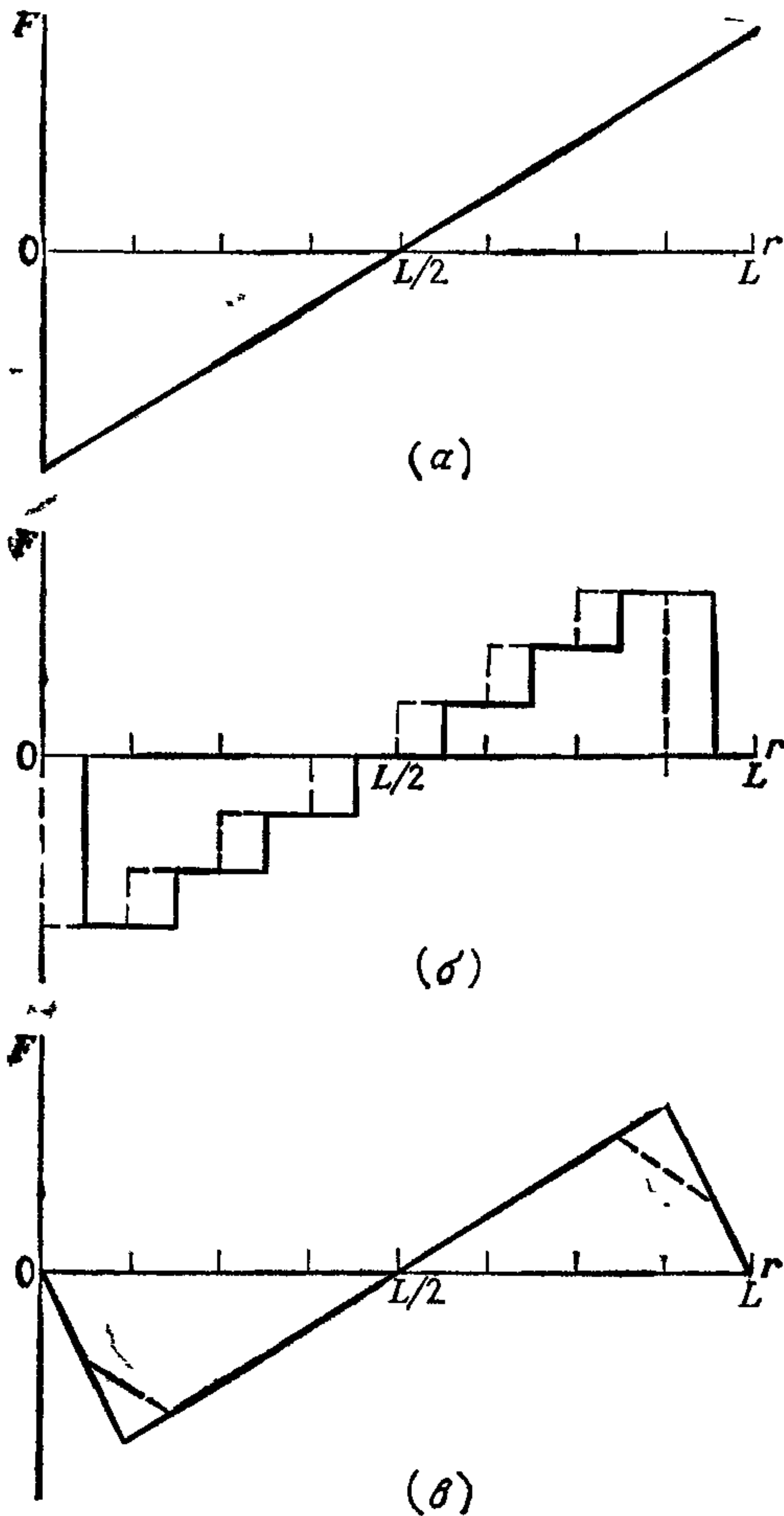


Рис. 5.1. Зависимость силы F , действующей на заряд q со стороны заряда $-q$, от расстояния r между частицами в одномерной периодической системе с длиной периода L (8 ячеек) для (а) непрерывного случая, (б) схемы NGP, (в) схемы CIC. Сплошные линии на (б) и (в) отвечают случаю заряда $-q$ в сеточном узле, а штриховые — заряду $-q$ на границе ячейки.

решением уравнения

$$\frac{d^2\varphi}{dx^2} = -\frac{q}{\epsilon_0} (\delta(x-x') - \delta(x+x')), \quad (5.1)$$

где

$$\varphi(x) = \varphi(x+L), \quad (5.2)$$

символ δ обозначает δ -функцию Дирака (см. приложение). Легко показать, что решение уравнения (5.1) с условием (5.2) имеет вид

$$\varphi(x) = \begin{cases} \frac{q}{\epsilon_0 L} (L - 2x')x, & |x| \leq x', \\ \frac{q}{\epsilon_0 L} (L - 2x)x', & x' \leq |x| \leq \frac{L}{2}, \end{cases} \quad (5.3)$$

где произвольная постоянная в потенциале выбрана так, чтобы средний потенциал равнялся нулю. Принимая, что при $|x| = x'$ электрическое поле

$$E = - \frac{d\varphi}{dx} \quad (5.4)$$

равняется среднему арифметическому его значений на краях разрыва,

$$E = - \frac{q}{2\epsilon_0 L} (L - 4x'), \quad |x| = x', \quad (5.5)$$

найдем силы $F_e(x_1)$ и $F_e(x_2)$, действующие на заряженные частицы в точках $x_1 = x'$ и $x_2 = -x'$ соответственно:

$$F_e(x_1) = -F_e(x_2) = \begin{cases} -\frac{q^2}{2\epsilon_0 L} (L - 4x'), & |x'| \neq 0. \\ 0, & |x'| = 0 \text{ и } L/2. \end{cases} \quad (5.6)$$

На рис. 5.1(a) показана зависимость силы, действующей на частицу 1, от расстояния между частицами $r = 2x'$.

5.2.2. Схема NGP

Непрерывному случаю наиболее полно отвечает расчет силы по схеме NGP (см. гл. 2), когда обе частицы расположены в узлах сетки. Если частицы 1 и 2 имеют заряды q и $-q$ и находятся соответственно в узлах сетки с $x_1 = x_{p'}$ и $x_2 = x_{-p'} = -x_{p'}$ то сеточная плотность заряда равняется

$$\rho_p \equiv \rho(x_p) = \frac{q}{H} [\delta_{p, p'} - \delta_{p, -p'}], \quad (5.7)$$

где $\delta_{p, p'}$ — символ Кронекера (см. разд. П.4.4), а $x_p = pH$ — это положение сеточного узла p , p целое.

Сеточный потенциал определяется решением системы алгебраических уравнений

$$\varphi_{p-1} - 2\varphi_p + \varphi_{p+1} = -\frac{\rho_p H^2}{\epsilon_0} \quad (5.8)$$

с граничными условиями

$$\varphi_p = \varphi_{p+N_g}, \quad (5.9)$$

где N_g — количество узлов сетки на длине периода L .

Решение уравнения (5.8) с условием (5.9) имеет вид

$$\varphi_p = \begin{cases} \frac{q}{\epsilon_0 L} (L - 2x_{p'}) x_p, & |x_p| \leq x_{p'}, \\ \frac{q}{\epsilon_0 L} (L - 2x_p) x_{p'}, & x_{p'} \leq |x_p| < L/2. \end{cases} \quad (5.10)$$

Вычисляя электрическое поле по формуле

$$E_p = - \left[\frac{\varphi_{p+1} - \varphi_{p-1}}{2H} \right], \quad (5.11)$$

получаем, что силы, действующие на частицы 1 и 2, а именно F_1 и F_2 , равны

$$F_1 = - F_2 = \begin{cases} \frac{q^2}{2\epsilon_0 L} (L - 4x_{p'}), & x_{p'} \neq 0, \\ 0, & x_{p'} = 0. \end{cases} \quad (5.12)$$

Из сравнения выражений (5.3) с (5.10) и (5.6) с (5.12) обнаруживается на первый взгляд удивительный факт: трехточечная аппроксимация лапласиана и конечно-разностная аппроксимация производной приводят в точности к тому же результату, что и дифференциальное уравнение, если частицы расположены в узлах сетки! Немного подумав, находим, что этого и следовало ожидать. Конечно-разностные аппроксимации первой и второй производных потенциала имеют второй порядок точности, причем старшие члены погрешностей пропорциональны соответственно третьей и четвертой производным потенциала. Согласно разложению Тейлора,

$$\begin{aligned} E_p &= \frac{\varphi(x_p - H) - \varphi(x_p + H)}{2H} = - \frac{d\varphi}{dx} \Big|_{x_p} - \frac{H^2}{6} \frac{d^3\varphi}{dx^3} \Big|_{x_p} + O(H^3), \\ \frac{\varphi(x_p - H) - 2\varphi(x_p) + \varphi(x_p + H)}{H^2} &= \frac{d^2\varphi}{dx^2} \Big|_{x_p} + \frac{H^2}{12} \frac{d^4\varphi}{dx^4} \Big|_{x_p} + O(H^4). \end{aligned} \quad (5.13)$$

В одномерном непрерывном случае потенциал единичного изолированного заряда (функция Грина дифференциального уравнения) кусочно-линейен. Поэтому третья и более старшие производные равны нулю, и, следовательно, уравнения (5.8) и (5.11) удовлетворяются точно. Поскольку любое решение этих уравнений может быть записано в виде линейной суперпозиции взвешенных функций Грина со сдвинутым аргументом для задачи об изолированном заряде, из этого следует, что решения дифференциального и разностного уравнений должны тождественно совпадать.

Нулевые погрешности простой конечно-разностной аппроксимации лапласиана характерны для одномерного случая. В случае двух и трех измерений функции Грина соответственно логарифмически и обратно пропорционально зависят от расстояния, и поэтому все их производные будут ненулевые. Тем не менее, используя для потенциала формализм функции Грина, а не уравнения Пуассона, и в многомерном случае вновь удается

получить точные значения потенциалов в узлах сетки, хотя в определенных ситуациях вычислительные затраты такого метода могут оказаться значительными.

Схема NGP приписывает весь заряд данной частицы ближайшему к ней сеточному узлу, а в качестве значения силы, действующей на частицу, берется сеточная сила из ближайшего узла. Вследствие этого сила взаимодействия между частицами может принимать только значения, определяемые выражением (5.12). Поэтому при изменении расстояния между частицами сила их взаимодействия меняется скачком, когда частицы пересекают границы ячеек. Это показано на рис. 5.1(б). Более того, сила зависит не только от расстояния между частицами, но и от их расположения относительно сетки. Если расстояние между частицами зафиксировать, а сами частицы перемещать относительно сетки, то сила их взаимодействия будет флуктуировать с периодом, определяемым шириной ячейки H . Именно это нарушение инвариантности по отношению к сдвигу самым пагубным образом сказывается на физической реальности результатов моделирования. Что касается схемы NGP, то самые большие флуктуации силы испытывают частицы, находящиеся на расстоянии, меньшем H , поскольку когда одна частица пересекает границу ячейки, сила меняется скачком от нуля до своего максимального значения.

5.2.3. Схема CIC

Большие погрешности в величине межчастичных сил в одномерной модели NGP обусловлены только грубостью схемы распределения заряда и интерполяции силы. Более хорошую аппроксимацию силы, хотя и более дорогую (в смысле числа арифметических операций, приходящихся на частицу за шаг по времени), можно получить, заменяя распределение заряда в ближайший узел сетки схемой, использующей два ближайших соседа. Эта более точная схема называется схемой «облаков в ячейке» (cloud-in-cell), или CIC [35], и ее название связано с физической интерпретацией процедуры распределения заряда (см. разд. 5.4).

Для того чтобы продемонстрировать повышение точности, которое дает использование при распределении заряда и интерполяции силы двух узлов вместо одного, вновь рассмотрим силу взаимодействия частиц. Возьмем сначала случай, когда частица 2 находится в узле сетки (скажем, узле с номером 0 при $x = 0$), а частица 1 расположена в точке x между сеточными узлами p и $p + 1$, где $p > 1$. Мы не учитываем пока никаких сил, с которыми частица 1 может действовать сама на себя. Точное значение силы F_e , действующей на частицу 1 со стороны частицы 2,

получается из (5.6) при расстоянии между частицами $2x' = x$:

$$F_e(x) = \frac{-q^2}{2\epsilon_0 L} (L - 2x). \quad (5.14)$$

Сеточные силы в узлах p и $p + 1$ (согласно схеме NGP) равны

$$F(x_p) = -\frac{q^2}{2\epsilon_0 L} (L - 2x_p), \quad (5.15)$$

$$F(x_{p+1}) = -\frac{q^2}{2\epsilon_0 L} (L - 2x_{p+1}). \quad (5.16)$$

Если вместо значения силы, принятой в схеме NGP, взять линейную комбинацию значений из обоих ближайших узлов, то получим

$$F(x) = \alpha F(x_p) + (1 - \alpha) F(x_{p+1}), \quad (5.17)$$

где α является некоторой функцией от $x - x_p$. Погрешность ϵ_1 в силе, действующей на частицу 1, составляет теперь

$$\begin{aligned} \epsilon_1 &= F_e - F \\ &= \frac{q^2}{\epsilon_0 L} [(x - x_{p+1}) - \alpha(x_p - x_{p+1})]. \end{aligned} \quad (5.18)$$

Следовательно, если выбрать

$$\alpha = \frac{x_{p+1} - x}{x_{p+1} - x_p} = 1 - \frac{x - x_p}{H}, \quad (5.19)$$

то мы получим $\epsilon_1 = 0$. Подставляя выражение для α (5.19) в (5.17), после небольшой перегруппировки получаем

$$F = \left(1 - \frac{x - x_p}{H}\right) F(x_p) + \left(1 + \frac{x - x_{p+1}}{H}\right) F(x_{p+1}). \quad (5.20)$$

Определяя в схеме SIC интерполяционную функцию для силы в форме

$$W(x - x_p) = \begin{cases} 1 - \frac{|x - x_p|}{H}, & |x - x_p| \leq H, \\ 0 & \text{в других случаях,} \end{cases} \quad (5.21)$$

можно записать (5.20) в более общем виде:

$$F(x) = \sum_p W(x - x_p) F(x_p), \quad (5.22)$$

где сумма по p берется по всем узлам сетки. Выражение (5.22) по форме в точности совпадает с полученным в гл. 2 для схемы NGP (формула (2.31)).

Похожие выкладки можно проделать и для схемы SIC распределения заряда. Рассмотрим ситуацию, где частица 1 расположена между сеточными узлами p и $p \pm 1$, а частица 2 нахо-

дится теперь в точке x , лежащей между узлами 0 и 1. Если бы заряд распределялся по схеме NGP, то сила, действующая на частицу 1 со стороны частицы 2, была бы нечувствительна к смещениям частицы 2 в пределах $x \in [0, H/2]$ и $x \in [H/2, H]$, а при $x = H/2$ менялась бы скачком. Однако если заряд $-q$ частицы 2 распределить между обоими ближайшими к ней узлами сетки, а именно узлами 0 и 1 с координатами x_0 и x_1 ,

$$\rho(x_0) = \frac{(-q)}{H} \beta, \quad (5.23)$$

$$\rho(x_1) = \frac{(-q)}{H} (1 - \beta), \quad (5.24)$$

то, выбирая $\beta = 1 - x/H$, можно и в этом случае погрешность в величине силы свести к нулю (в предположении, что сеточные узлы, участвующие в распределении заряда и интерполяции, различны). Указанный выбор β равносителен использованию функции распределения заряда, тождественно равной интерполяционной функции для силы, определяемой формулой (5.21):

$$\rho(x_0) = \frac{(-q)}{H} W(x - x_0), \quad (5.25)$$

$$\rho(x_1) = \frac{(-q)}{H} W(x - x_1). \quad (5.26)$$

Что касается ансамбля из N_p заряженных частиц, то формулы (5.25) и (5.26) обобщаются и дают

$$\rho(x_p) = \frac{1}{H} \sum_{i=1}^{N_p} q_i W(x_i - x_p), \quad (5.27)$$

где x_i и q_i — положение и заряд i -й частицы соответственно. С другой стороны, (5.27) можно выразить через плотность частиц, и в этом случае снова получается выражение для распределения заряда, найденное для схемы NGP (см. (2.29)).

На рис. 5.1(в) показана сила взаимодействия двух частиц, полученная по схеме СИС, в которой распределение заряда и интерполяция силы выполняются по формулам (5.27) и (5.22) с одинаковой функцией распределения-интерполяции. При сравнении рис. 5.1(а) — (в) сразу видно, что:

1. СИС дает значительно более гладкую силу, чем схема NGP (кусочно-линейную вместо кусочно-постоянной).
2. СИС снижает амплитуду флуктуаций силы взаимодействия частиц при их смещении относительно сетки.
3. В случае схемы СИС остаточная погрешность силы более локализована.

Полное подавление флуктуаций силы при расстояниях между частицами, превышающих две ячейки, характерно лишь для

одномерного случая, хотя все три указанные выше моменты остаются по-прежнему справедливы и в случае двух или трех измерений.

5.2.4. Смешанные схемы

В каждой из схем, NGP и CIC, для распределения заряда и для интерполяции силы используются одинаковые функции. В результате такой симметрии эти схемы сохраняют импульс, т. е. две частицы действуют друг на друга с одинаковыми по величине и противоположными по направлению силами, и к тому же сила действия на частицу ее собственного поля (самодействие) равна нулю (см. разд. 5.3.3).

Одна из возможных модификаций заключается в том, что для распределения заряда и для интерполяции силы используются разные функции. В таких случаях в силу симметрии самих функций силы, действующие на каждую из двух частиц, по-прежнему будут одинаковы по величине и направлены в противоположные стороны, но нулевое самодействие больше не обеспечивается. Наличие самодействия в лучшем случае приводит к нефизическому ограничению на шаг по времени, а в худшем ведет к катастрофическим последствиям.

Чтобы наглядно показать те трудности, к которым приводит самодействие, оценим самодействие изолированного заряда q для случаев, когда заряд распределяется по схеме CIC, а сила интерполируется по NGP, и наоборот. В обоих случаях мы будем предполагать, что сеточные потенциалы и электрическое поле определяются соотношениями (5.8) и (5.11) соответственно.

Решая уравнение (5.8) для изолированного единичного заряда, расположенного в сеточном узле 0, и принимая $\varphi_0 = 0$, получаем

$$\varphi_p = -\frac{q}{2\epsilon_0} |x_p|. \quad (5.28)$$

Если заряд q находится в точке x , где $0 \leq x \leq H/2$, и для распределения заряда используется схема CIC, то потенциал является суперпозицией потенциалов вида (5.28) с зарядом $q(1 - x/H)$ в нулевом узле сетки и с зарядом qx/H в узле 1. Прodelывая выкладки до конца, получим для заряда, расположенного в ячейке 0 (т. е. $|x| \leq H/2$),

$$\varphi_p = -\frac{q}{2\epsilon_0} |x - x_p|. \quad (5.29)$$

Отсюда, используя (5.11), (2.28) и (5.22), находим, что в случае, когда заряд распределяется по схеме CIC, а сила интерполируется по NGP, сила самодействия частицы с зарядом q , распо-

ложенной в ячейке 0, равняется

$$F_{\text{сам}} = -\frac{q^2}{2\epsilon_0 H} x. \quad (5.30)$$

Уравнение (5.30) описывает простой гармонический осциллятор с частотой

$$\omega_{\text{сам}} = \left(\frac{q^2}{2\epsilon_0 m H} \right)^{1/2}. \quad (5.31)$$

при условии, что частица не покидает области $|x| \leq H/2$. Для обеспечения устойчивости интегрирования по времени по схеме с перешагиванием необходимо, чтобы кроме всех ограничений на временной шаг, налагаемых физикой системы многих тел, выполнялось условие $\omega_{\text{сам}} DT < 2$. Для типичных наборов параметров моделирования бесстолкновительных систем ограничение шага по времени, связанное с $\omega_{\text{сам}}$, как правило, слабее, чем ограничения, налагаемые собственными физическими частотами (см. гл. 7). Тем не менее ограничение временного шага, обусловленное $\omega_{\text{сам}}$, нельзя полностью игнорировать, поскольку оно устанавливает верхний предел отношения DT^2/H .

Если повторить вычисление силы самодействия с распределением заряда по схеме NGP и интерполяцией силы по CIC, то сила самодействия получится равной

$$F_{\text{сам}} = \frac{q^2}{2\epsilon_0 H} x. \quad (5.32)$$

Таким образом, тогда как расчет силы по схемам CIC/NGP вносит лишь сравнительно неопасные колебания силы самодействия, в случае схем NGP/CIC самодействие имеет противоположный знак, что приводит к экспоненциально неустойчивому движению даже в пределе нулевого шага по времени. Ясно, что, поскольку в любой численной схеме подобная ситуация неприемлема, мы должны ограничить порядок схемы интерполяции силы величиной, равной или меньшей порядку схемы распределения заряда. Кроме того, если мы хотим обеспечить точное сохранение импульса, то необходимо, чтобы функция распределения заряда и интерполяционная функция для силы тождественно совпадали (см. разд. 5.3.3).

5.3. ИЕРАРХИЯ СХЕМ РАСПРЕДЕЛЕНИЯ ЗАРЯДА

Распределение заряда и интерполяция силы сопровождаются численными погрешностями, и это приводит к тому, что силы взаимодействия частиц получаются неточными по величине и флуктуируют при смещениях. Увеличение количества узлов сетки, участвующих в распределении заряда и интерполяции силы, позволяет ослабить нежелательные численные эффекты,

но за счет повышения вычислительных затрат. С практической точки зрения важно добиться того, чтобы при заданном числе используемых сеточных узлов погрешность была минимальна, иными словами, достигнуть наилучшего компромисса качество—затраты.

Иерархию функций распределения заряда и интерполяции силы, упорядоченную по затратам и качеству, можно построить, требуя выполнения следующих критериев:

1. На больших по сравнению с шагом сетки расстояниях между частицами флуктуации должны становиться пренебрежимо малыми (т. е. погрешности должны быть локализованы в пространстве).
2. При перемещении частицы по сетке заряд, распределяемый в сеточные узлы, и действующая на нее сила, определяемая интерполяцией по сеточным значениям, должны изменяться гладко (т. е. флуктуации пространственно локализованных погрешностей должны быть малы).
3. Импульс должен сохраняться (т. е. необходимо избегать проблем, связанных с самодействием).

Первый критерий позволяет, исходя из разложений по малым параметрам, связанным с шагом сетки, вывести систему условий, ограничивающих вид функций распределения заряда. Вторым критерий сводит множество функций распределения заряда к иерархии схем, два первых члена которой оказываются схемами NGP и CIC. Третий критерий делает функцию интерполяции силы совпадающей с функцией распределения заряда, хотя другим критерий, которым его можно заменить, а именно сохранение энергии, приводит к другим схемам (см. разд. 5).

5.3.1. Условия на больших расстояниях

Критерий 1 рассматривается применительно к полю или потенциалу, порожденным одиночным единичным зарядом. Используя линейность и принцип суперпозиции, результаты можно обобщить на случай других значений заряда и систем многих тел. Рассмотрим сначала одномерную схему, в которой заряд частицы приписывается m узлам сетки, ближайшим к ней.

Пусть $W_p(x)$ — доля заряда, распределенная в узел p , находящийся в точке x_p , от частицы, находящейся в точке x , и пусть $G(x' - x_p)$ (функция Грина) — это потенциал в точке x' , порожденный единичным зарядом, находящимся в узле p . Тогда распределение заряда приведет к потенциалу в точке x' , описываемому выражением

$$\varphi(x') = \sum_{p=1}^m W_p(x) G(x' - x_p). \quad (5.33)$$

$W_p(x) [= W(x - x_p)]$ — это просто функция распределения заряда, частные случаи которой были рассмотрены в предыдущем разделе. Точный потенциал φ_c , порожденный единичным зарядом, помещенным в точке x , зависит только от расстояния $|x' - x|$, тогда как приближенный потенциал зависит и от $|x' - x|$, и от $(x - x_p)$, т. е. от расположения точки x относительно узлов сетки. Согласно критерию 2 (и физическим соображениям!), именно эту зависимость от $(x - x_p)$ и надлежит подавлять.

Разлагая в (5.33) функцию $G(x' - x_p)$ в ряд Тейлора в точке $(x' - x)$, имеем

$$\varphi(x') = \sum_{p=1}^m W_p(x) G(x' - x) + \sum_{p=1}^m W_p(x) \sum_{n=1}^{\infty} \frac{\Delta_p^n d^n G(x - x')}{dx^n}, \quad (5.34)$$

$$\Delta_p = x - x_p. \quad (5.35)$$

Условие сохранения заряда дает первое ограничение на вид W :

$$\sum_{p=1}^m W_p(x) = 1. \quad (5.36)$$

Если заряд приписывается только одному узлу сетки ($m = 1$), то соотношение (5.36) полностью определяет схему распределения заряда, а именно схему NGP. Если, однако, $m > 1$, то появляющиеся дополнительные степени свободы можно использовать для ослабления сеточной зависимости функции $\varphi(x')$: для этого достаточно потребовать, чтобы члены высоких порядков в выражении (5.34) (упорядоченные по степеням малого параметра $\Delta_p/|x - x'|$) последовательно не зависели от сетки, т. е. мы требуем, чтобы

$$\sum_{p=1}^m W_p \Delta_p^n = \text{const} \quad (5.37)$$

для все больших значений n по мере того, как растет количество сеточных узлов m , участвующих в распределении заряда. Поскольку как φ_c , так и G — четные функции своих аргументов, в уравнения (5.37) нужно, очевидно, ввести дополнительные ограничения, что дает следующие условия:

$$\sum_{p=1}^m W_p \Delta_p^n = \begin{cases} 0, & n \text{ нечетные,} \\ \text{const,} & n \text{ четные.} \end{cases} \quad (5.38)$$

Для $m = 2$ можно удовлетворить только условиям нулевого (5.36) и первого порядка ($n = 1$ в (5.38)), при этом старший

член остатка в (5.34) порядка $(\Delta_p)^2$. С учетом (5.35) условия принимают вид

$$W_1 + W_2 = 1, \quad (5.39)$$

$$W_1 x_1 + W_2 x_2 = x, \quad (5.40)$$

где все W_p для $p \neq 1$ и 2 равны нулю и $x_1 \leq x \leq x_2$. Решая (5.39) и (5.40), получаем

$$W_1(x) = 1 - \frac{(x - x_1)}{H}, \quad (5.41)$$

$$W_2(x) = 1 + \frac{(x - x_2)}{H}, \quad (5.42)$$

где H (ширина ячейки) равно расстоянию между узлами сетки. Используя свойство инвариантности функции распределения заряда по отношению к смещениям

$$W_p(x) = W(x - x_p), \quad (5.43)$$

можно объединить (5.41) с (5.42), что дает

$$W(x) = \begin{cases} 1 - \frac{|x|}{H}, & |x| < H, \\ 0 & \text{в других случаях,} \end{cases} \quad (5.44)$$

а это, очевидно, есть функция распределения заряда схемы СИС.

Увеличивая m до трех, можно удовлетворить уравнениям (5.38) вплоть до $n = 2$:

$$W_1 + W_2 + W_3 = 1, \quad (5.45)$$

$$W_1 x_1 + W_2 x_2 + W_3 x_3 = x, \quad (5.46)$$

$$W_1 x_1^2 + W_2 x_2^2 + W_3 x_3^2 = C + x^2, \quad (5.47)$$

где

$$W_p = 0 \quad \text{для } p \neq 1, 2, \text{ или } 3 \quad (5.48)$$

и $-H/2 \leq x - x_2 \leq H/2$. В этом случае уравнение (5.34) принимает вид

$$\varphi(x') = G(x' - x) + \frac{C}{2} \frac{d^2 G(x' - x)}{dx^2} + O(\Delta^3). \quad (5.49)$$

Решение (5.45) — (5.47) дает

$$W_1(x) = (X^2 - HX + C)/2H^2, \quad (5.50)$$

$$W_2(x) = 1 - (X^2 + C)/H^2, \quad (5.51)$$

$$W_3(x) = (X^2 + HX + C)/2H^2, \quad (5.52)$$

где $X = x - x_2$ и $-H/2 \leq X \leq H/2$. Соотношения (5.50) — (5.52) описывают распределение заряда в такой форме, которая позволяет построить алгоритм распределения заряда. Используя

свойство инвариантности относительно смещений (см. (5.43)) и соотношение (5.48), можно переписать выражения (5.50) — (5.52), как это было сделано для схемы СИС (см. (5.44)), в виде одной кусочно-полиномиальной функции распределения заряда:

$$W(x) = \begin{cases} (x^2 + 3Nx + 2N^2 + C)/2N^2, & -3N/2 \leq x < -N/2, \\ 1 - (x^2 + C)/N^2, & -N/2 \leq x < N/2, \\ (x^2 - 3Nx + 2N^2 + C)/2N^2, & N/2 \leq x < 3N/2, \\ 0 & \text{в других случаях.} \end{cases} \quad (5.53)$$

Если C положить равным нулю, то сеточный потенциал (5.49) будет воспроизводить точный потенциал φ_c с точностью до $O(\Delta^3)$ при условии, что G задано равным φ_c . Если же C выбрано ненулевым, то нужно, чтобы G удовлетворяло уравнению

$$\frac{C}{2} \frac{d^2 G}{dx^2} + G = \varphi_c \quad (5.54)$$

с тем, чтобы компенсировать изменение формы потенциала, обусловленное процедурой распределения заряда, и сохранить второй порядок точности. Оказывается, что на практике выгоднее подправлять сами силы взаимодействия между частицами, а потенциал оставлять с погрешностями (см. гл. 7 и 8).

В двумерном и трехмерном случаях посредством обобщения выражения (5.33) и его разложения получаются аналогичные условия на больших расстояниях, упорядоченные по степеням малого параметра вида $\Delta/|x - x'|$. Замена x на $x [= (x, y)$ в двумерном и (x, y, z) в трехмерном случае] и т. д. дает

$$\varphi(x') = \sum_p W_p(x) G(x' - x_p). \quad (5.55)$$

Вектору p отвечает теперь пара (p_1, p_2) или тройка (p_1, p_2, p_3) целых чисел, обозначающих p -й узел сетки, расположенный в точке x_p , а сумма берется по m узлам сетки, ближайшим к источнику в точке x .

Разложение (5.55) в двумерном случае имеет вид

$$\varphi(x') = \sum_p W_p(x) \sum_{r,s=0}^{\infty} \frac{\Delta_1^r \Delta_2^s}{r!s!} \frac{\partial^{r+s} G}{\partial x^r \partial y^s}, \quad (5.56)$$

а в трехмерном случае

$$\varphi(x') = \sum_p W_p(x) \sum_{r,s,t=0}^{\infty} \frac{\Delta_1^r \Delta_2^s \Delta_3^t}{r!s!t!} \frac{\partial^{r+s+t} G}{\partial x^r \partial y^s \partial z^t}, \quad (5.57)$$

где

$$\Delta = (\Delta_1, \Delta_2, \Delta_3) = x - x_p. \quad (5.58)$$

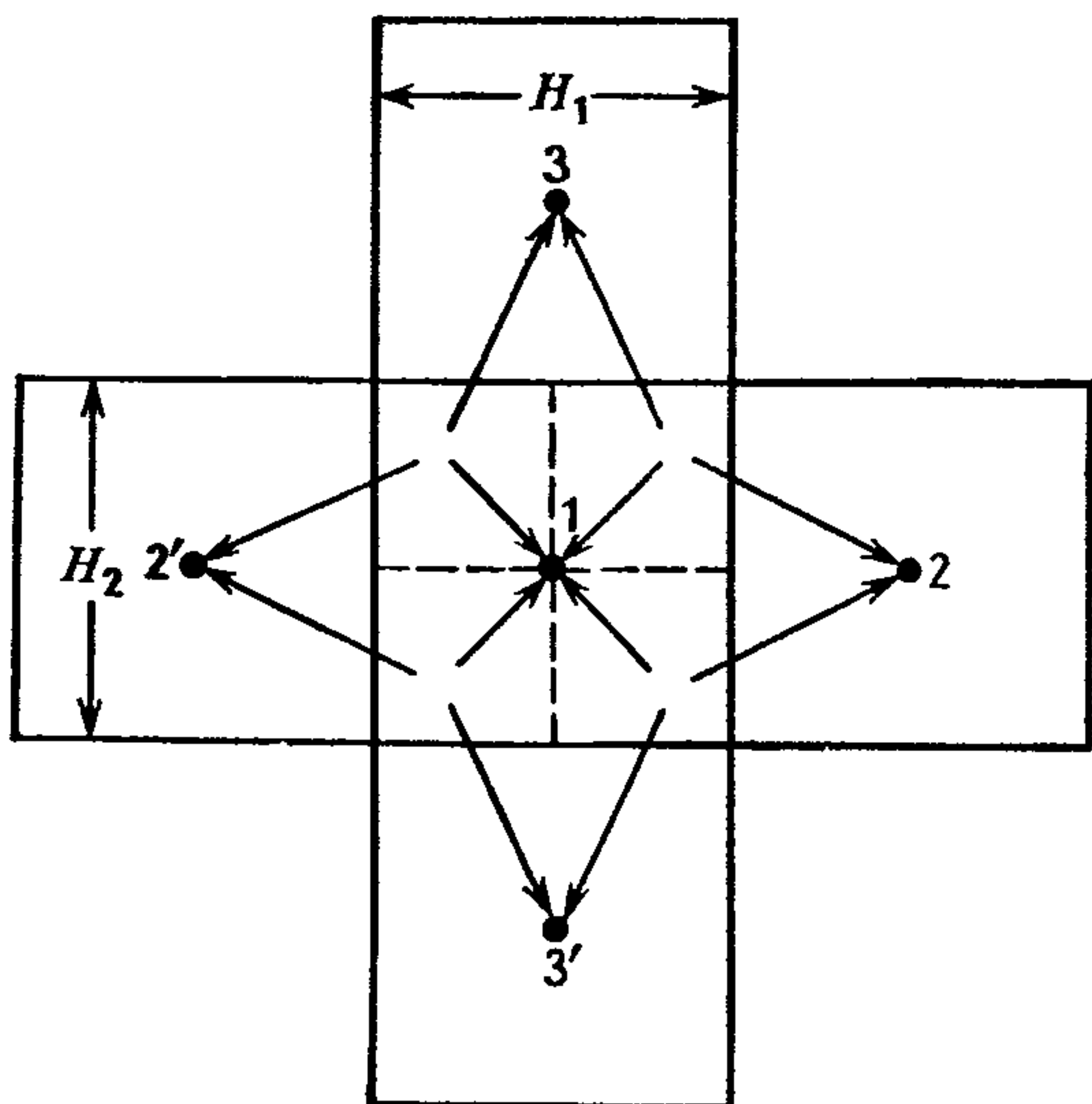


Рис. 5.2. Двумерная трехточечная схема размазывания заряда приписывает заряд ближайшему узлу сетки (обозначен цифрой 1), а также следующим ближайшим узлам в направлении на восток или запад (обозначены 2 или 2') и север или юг (отмечены цифрами 3 или 3'). Выбор очередной пары ближайших узлов определяется квадрантом ячейки, в котором лежит частица.

Если, как и в одномерном случае, наложить на G только минимальное ограничение, состоящее в том, чтобы $G(\mathbf{x}) = G(|\mathbf{x}|)$, то из (5.56) или (5.57) мы получим иерархию условий на функцию распределения заряда:

$$\sum_p W_p(\mathbf{x}) = 1, \quad (\text{сохранение заряда}), \quad (5.59)$$

$$\sum_p W_p(\mathbf{x}) \Delta_i = 0, \quad (\text{первый порядок}), \quad (5.60)$$

$$\sum_p W_p(\mathbf{x}) \Delta_i \Delta_j = C \delta_{ij} \quad (\text{второй порядок}), \quad (5.61)$$

$$\sum_p W_p(\mathbf{x}) \Delta_i \Delta_j \Delta_k = 0 \quad (\text{третий порядок}) \quad (5.62)$$

и так далее.

Уравнения (5.59) — (5.62) применяются таким же образом, как их одномерные аналоги. Например, двумерная функция распределения заряда самого низкого порядка, удовлетворяющая только уравнению (5.59), получается путем приписывания всего заряда ближайшему сеточному узлу:

$$W(\mathbf{x}) = \begin{cases} 1, & -H/2 \leq x < H/2, \\ 0 & \text{в других случаях.} \end{cases} \quad (5.63)$$

$\mathbf{H} = (H_1, H_2)$ — это вектор, характеризующий размеры ячейки используемой сетки. Распространение процедуры распределения заряда на три ближайших к исходному заряду узла сетки позволяет удовлетворить как условию сохранения заряда, так и условию первого порядка. В обозначениях рис. 5.2 эти условия

принимают вид

$$W_1 + W_2 + W_3 = 1, \quad (5.64)$$

$$W_1 x_1 + W_2 x_2 + W_3 x_3 = x, \quad (5.65)$$

$$W_1 y_1 + W_2 y_2 + W_3 y_3 = y, \quad (5.66)$$

что дает

$$W_1 = 1 - \frac{|x|}{H_1} - \frac{|y|}{H_2}, \quad (5.67)$$

$$W_2 = \frac{|x|}{H_1}, \quad (5.68)$$

$$W_3 = \frac{|y|}{H_2} \quad (5.69)$$

для $-H/2 \leq x \leq H/2$.

С помощью свойства инвариантности по отношению к смещениям можно объединить (5.67) — (5.69) и записать выражение для функции распределения заряда:

$$W(\mathbf{x}) = \begin{cases} 1 - \frac{|x|}{H_1} - \frac{|y|}{H_2}, & -\frac{H}{2} \leq x < \frac{H}{2}, \\ 1 - \frac{|x|}{H_1}, & x = \frac{H_1}{2} \text{ или } \frac{H_1}{2} < |x| < H_1, \quad -\frac{H_2}{2} \leq y < \frac{H_2}{2}, \\ 1 - \frac{|y|}{H_2}, & -\frac{H_1}{2} \leq x < \frac{H_1}{2}, \quad y = \frac{H_2}{2} \text{ или } \frac{H_2}{2} < |y| < H_2, \\ 0 & \text{в других случаях.} \end{cases} \quad (5.70)$$

Увеличение количества используемых узлов сетки до четырех по-прежнему позволяет удовлетворить условиям только до первого порядка, однако дает дополнительную степень свободы при выборе W . Например, если мы зададимся схемой с четырьмя узлами в виде произведения

$$W(x, y) = P(x)P(y), \quad (5.71)$$

то, удовлетворяя условию первого порядка, получим следующую функцию распределения заряда:

$$W(\mathbf{x}) = \begin{cases} \left(1 - \frac{|x|}{H_1}\right) \left(1 - \frac{|y|}{H_2}\right), & -H < x < H, \\ 0 & \text{в других случаях.} \end{cases} \quad (5.72)$$

Функция распределения заряда (5.72) отвечает двумерному аналогу схемы СИС.

Список примеров двумерных схем распределения заряда можно продолжать до бесконечности: можно построить схемы, в которых задействовано пять, шесть, семь и более узлов сетки, что позволяет удовлетворять условиям все более высокого порядка или вводить дополнительные степени свободы. Основная трудность, связанная со многими такими схемами, состоит в том, что при отсутствии регулярного постоянного шаблона узлов

(как это имеет место в четырехточечной, симметричной пятиточечной или девятиточечной схемах распределения) для определения местонахождения нерегулярных узлов нужна значительная вычислительная работа. Среди всех девятиточечных схем, по-видимому, самой полезной является схема второго порядка точности, которая представляет собой произведение функций вида (5.53).

Трехмерные схемы строятся путем повторения двумерных выкладок с добавлением третьей компоненты. В трехмерном аналоге двумерной трехточечной схемы используются четыре узла и т. д. Поскольку эта задача не добавит ничего, кроме дополнительных алгебраических выкладок, мы на этом больше останавливаться не будем.

С физической точки зрения схемы разложения (5.56) и (5.57) можно трактовать как разложение по мультиполям. Члены $O(\Delta)$ соответствуют дипольному моменту, члены $O(\Delta^2)$ — квадрупольному и т. д. Эта связь легко усматривается при подстановке $G = \log r/2\pi\epsilon_0$ в (5.56) либо $G = (4\pi\epsilon_0 r)^{-1}$ в (5.57). Если распределение заряда строится по схеме нулевого порядка (NGP), то на больших расстояниях старший член остатка потенциала и поля пропорционален дипольному моменту. В схеме первого порядка (например, SIC) старший член остатка имеет характер квадрупольного и т. д. Мультипольная интерпретация выявляется при внимательном рассмотрении примеров схем распределения заряда. Например, в трехточечной схеме (5.67) — (5.69) можно выделить единичный заряд в сеточном узле 1 и наряду с ним диполь в направлении С—Ю с моментом H_2y и диполь в направлении В—З с моментом $H_1(x)$. Подобным образом интерпретируются и другие примеры.

В двумерном случае дипольный характер погрешности приводит к погрешности потенциала, ведущей себя как $\cos \theta$. Для квадрупольной погрешности (т. е. для схемы первого порядка) погрешность меняется как $\cos 2\theta$, и так далее для схем более высокого порядка, где угол θ меняется от 0 до 2π при обходе исходного заряда против часовой стрелки. Как указывалось выше, это можно показать подстановкой в (5.56) $G = \log r/2\pi\epsilon_0$ или, проще, полагая $G = \log z/2\pi\epsilon_0$ и используя равносильное (5.56) представление в комплексных переменных (см. [119]):

$$\begin{aligned} \psi &= \frac{1}{2\pi\epsilon_0} \sum_p W_p(z) \log(z - z_p) \\ &= \frac{1}{2\pi\epsilon_0} \left[\log z - \sum_p \sum_{n=1}^{\infty} \frac{W_p}{n} \left(\frac{z_p}{z}\right)^n \right]. \end{aligned} \quad (5.73)$$

Вещественная часть выражения (5.73) дает потенциал ϕ . Здесь z — обычная комплексная переменная $z = x + iy$, p — комплекс-

ный индекс $p = p_1 + ip_2$, обозначающий узлы сетки, а исходный заряд помещен в начало координат.

Погрешность потенциала определяется выражением

$$\operatorname{Re}[\varepsilon] = \operatorname{Re}\left[\sum_{n=1}^{\infty} \varepsilon_n\right] = -\operatorname{Re}\left[\frac{1}{2\pi\varepsilon_0} \sum_{n=1}^{\infty} \left(\frac{1}{nz^n} \sum_p W_p z_p^n\right)\right]. \quad (5.74)$$

Для схемы NGP старшим членом остатка является ε_1 , что дает

$$\varepsilon_1 = -\frac{1}{2\pi\varepsilon_0} \frac{z_1}{z} = -\frac{1}{2\pi\varepsilon_0} \frac{r_1 e^{i\theta_1}}{r e^{i\theta}}, \quad (5.75)$$

$$\operatorname{Re}[\varepsilon_1] = -\frac{r_1}{2\pi\varepsilon_0 r} \cos(\theta_1 - \theta), \quad (5.76)$$

т. е. схема NGP дает погрешность, имеющую дипольный характер, причем она максимальна вдоль линии, соединяющей исходный заряд с ближайшим сеточным узлом. Для схем первого порядка старший член остатка имеет вид

$$\varepsilon_2 = -\frac{1}{4\pi\varepsilon_0} \sum_p \frac{W_p z_p^2}{z^2}, \quad (5.77)$$

что дает, например для погрешности трехточечной схемы и схемы SIC в случае заряда, расположенного посередине между соседними сеточными узлами, величину

$$\operatorname{Re}[\varepsilon_2] = -\frac{1}{4\pi\varepsilon_0 r^2} r_1^2 \cos 2\theta. \quad (5.78)$$

В конечном счете эти погрешности обуславливают зависимость потенциала от направления, иначе говоря угловую анизотропию. На эту нефизическую угловую зависимость потенциала, а также полей, накладываются далее погрешности, возникающие из-за аппроксимации уравнений для поля (см. разд. 5.4). Другой метод рассмотрения мультипольного разложения описан в работах [336], [228].

5.3.2. Условия гладкости

Уравнения, налагающие условия на вид функции распределения заряда, порождают иерархию схем, обладающих тем свойством, что чем больше узлов сетки участвует в распределении заряда, тем быстрее спадают погрешности с увеличением расстояния от исходного заряда. Однако, несмотря на то, что у схем высших порядков локализация ошибок выражена сильнее, они могут оказаться и не лучше схем более низкого порядка, если флуктуации силы на малых расстояниях между частицами велики. Второй критерий — условие гладкости — и связан с уменьшением этих флуктуаций.

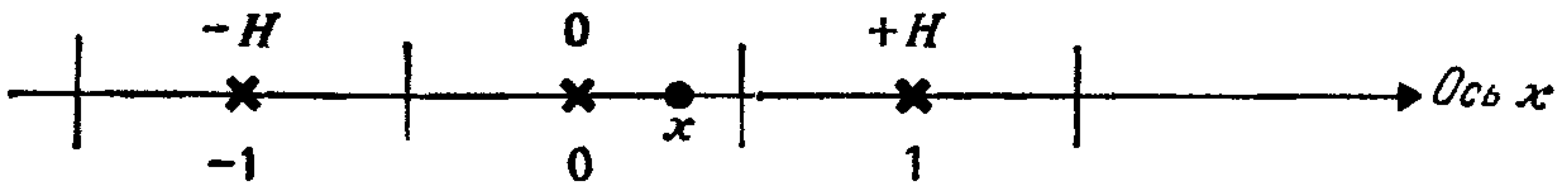


Рис 5.3. В одномерной трехточечной схеме заряд частицы, находящейся в точке x , где $|x| \leq H/2$, распределяется в три ближайших сеточных узла $(-1, 0, 1)$, имеющих соответственно координаты $(-H, 0, H)$. Вертикальными черточками показаны границы ячеек.

В x -пространстве гладкость количественно выражается непрерывностью производных приближенной функции. Примерами могут служить силы, изображенные на рис. 5.1: схема самого низкого порядка (NGP) дает разрывную силу, тогда как CIC дает непрерывную силу, но с разрывами в первой производной. Дальнейшее повышение гладкости достигается сначала требованием непрерывности как самой силы, так и ее первой производной, а затем (дополнительно) непрерывности второй производной, и так далее. Сказанное относится и к сеточной плотности заряда, и к потенциалу. В k -пространстве гладкость отражается на скорости убывания амплитуды фурье-гармоник функции с ростом волнового числа k (см. разд. 5.6).

Достижимая степень гладкости определяется количеством используемых сеточных узлов. В одномерном случае через n точек можно провести единственным образом полином степени не выше $n - 1$, и поэтому наивысший порядок производной, которую можно сделать непрерывной, равен $n - 1$ (производная порядка $n = 0$ — это сама функция, а производная порядка $n = -1$ — это интеграл от функции). Следовательно, согласно критерию гладкости, самыми хорошими функциями распределения заряда и интерполяции силы в случае одного и двух узлов являются соответственно функции NGP и CIC (см. (2.28) и (5.21)).

Непрерывность значения функции и ее производной имеет место в схеме TSC («треугольное облако»), которая в одномерном случае использует три сеточных узла и обладает кусочно-квадратичными функциями распределения заряда и интерполяции силы. Чтобы проиллюстрировать, каким образом можно применить критерий гладкости, опишем в общих чертах его использование при выводе функции распределения заряда (обратная интерполяция) и интерполяционной функции для силы в схеме TSC.

Рассмотрим ситуацию, изображенную на рис. 5.3. Если частица, находящаяся в точке x , обладает единичным зарядом, то доли (W_{-1}, W_0, W_1) заряда, приписанные узлам сетки $(-1, 0, 1)$, удовлетворяют соотношению

$$W_{-1} + W_0 + W_1 = 1, \quad (5.79)$$

где каждая из долей является квадратичной функцией координаты x . Из соображений симметрии следует, что

$$W_0(x) = ax^2 + b, \quad (5.80)$$

$$W_{-1}(x) = W_1(x) = cx^2 + dx + e. \quad (5.81)$$

Применительно к распределению заряда критерий гладкости требует, чтобы в каждом узле величина приписанного ему заряда и ее производная при движении частицы изменялись непрерывно. Полиномиальная форма (5.80) и (5.81) автоматически обеспечивает это для $|x| < H/2$, и поэтому нам необходимо рассмотреть только переходы в точках $x = H/2$ или $x = -H/2$. Когда частица пересекает границу ячейки при $x = H/2$, тремя ближайшими узлами сетки становятся (0, 1, 2), и поэтому для непрерывности величины заряда мы потребуем

$$W_{-1}(H/2) = 0, \quad (5.82)$$

$$W_0(H/2) = W_1(H/2), \quad (5.83)$$

а для непрерывности производной

$$\frac{d}{dx} W_{-1}(H/2) = 0, \quad (5.84)$$

$$\frac{d}{dx} W_0(H/2) = -\frac{d}{dx} W_1(H/2). \quad (5.85)$$

Решая уравнения (5.79) — (5.85), получаем схему TSC:

$$W_0(x) = \frac{3}{4} - \left(\frac{x}{H}\right)^2, \quad (5.86)$$

$$W_{-1}(-x) = W_1(x) = \frac{1}{2} \left(\frac{1}{2} + \frac{x}{H}\right)^2. \quad (5.87)$$

Используя свойство инвариантности по отношению к смещениям, объединяем (5.86) с (5.87) и получаем функцию распределения заряда схемы TSC:

$$W(x) = \begin{cases} \frac{3}{4} - \left(\frac{x}{H}\right)^2, & |x| \leq \frac{H}{2}, \\ \frac{1}{2} \left(\frac{3}{2} - \frac{|x|}{H}\right)^2, & \frac{H}{2} \leq |x| \leq \frac{3H}{2}, \\ 0 & \text{в других случаях.} \end{cases} \quad (5.88)$$

Из сравнения (5.88) с (5.53) мы видим, что схема TSC является частным случаем (отвечающим $C = H^2/4$) трехточечной схемы второго порядка точности, полученной с помощью условий на больших расстояниях.

Одним из последствий применения схемы TSC распределения заряда является расплывание самого заряда. Как следует из выражений (5.86) и (5.87), единичный заряд, помещенный в

точке $x = 0$, порождает в сеточных узлах $(-1, 0, +1)$ дробные заряды, равные соответственно $1/8, 3/4$ и $1/8$. В некоторых случаях, но не во всех, это может расцениваться как полезное увеличение ширины частицы конечного размера, приобретаемое вместе с более высокой степенью гладкости. Гораздо чаще бывает желательно добиться гладкости не за счет расплывания, причем достижению этой цели служит следующий за распределением заряда этап восстановления исходной формы.

Чтобы провести параллель с тем, что обсуждалось в разд. 5.3.1, рассмотрим опять задачу о потенциале единичного заряда. Пусть единичный заряд, находящийся в сеточном узле p' , создает в узле p потенциал $G_{p-p'}$. Тогда вследствие распределения заряда потенциал в p -м узле сетки, обусловленный единичным зарядом в точке $x = 0$, получим равным

$$\Phi_p = \frac{1}{8} G_{p+1} + \frac{3}{4} G_p + \frac{1}{8} G_{p-1}, \quad (5.89)$$

т. е.

$$\Phi_p = G_p + \frac{H^2}{8} \frac{(G_{p+1} - 2G_p + G_{p-1}))}{H^2}. \quad (5.90)$$

Выражение (5.90) представляет собой просто конечно-разностную аппроксимацию уравнения (5.54) с $C = (H/2)^2$, т. е. C принимает значение, которое приводит (5.53) к функции распределения заряда схемы TSC (5.88). Если точный потенциал в p -м сеточном узле от единичного заряда в узле p' равен $G_{p-p'}^+$, то с помощью выражения (5.90) можно определить значения G , которые компенсируют последствия расплывания, порожденные распределением заряда:

$$G_p^+ = \left(1 + \frac{H^2}{8} D^2\right) G_p. \quad (5.91)$$

В уравнении (5.91) D^2 является разностным оператором второго порядка, отвечающим лапласиану для непрерывного случая.

Для системы зарядов выражение (5.89) принимает вид

$$\Phi_p = \sum_{p'} G_{p-p'} \rho_{p'} H, \quad (5.92)$$

где $(\rho_{p'} H)$ представляет собой заряд, приписываемый сеточному узлу p' , или, другими словами, множество значений $\{\rho_{p'}\}$ имеет смысл сеточной плотности заряда. В выражении (5.92) значения G скорректированы в соответствии с (5.91) с тем, чтобы компенсировать расплывание заряда. С другой стороны, можно было бы исходить из точного потенциала взаимодействия $\Phi_{p-p'}$ и получить

$$\Phi_c = \sum_{p'} G_{p-p'}^+ \rho_{p'}^* H, \quad (5.93)$$

где скорректированная плотность заряда $\{\rho_p^*\}$ определяется соотношениями (5.91)—(5.93). Подставляя (5.91) в (5.93), производя перегруппировку слагаемых в сумме и вычитая (5.92), получаем связь между ρ и ρ^* :

$$\left(1 + \frac{H^2}{8} D^2\right) \rho_p^* = \rho_p. \quad (5.94)$$

Уравнения (5.91)—(5.94) подсказывают два возможных способа компенсации эффектов распыления, вызванного процедурой распределения заряда. Либо мы сначала решаем уравнение (5.94) и получаем $\{\rho_p^*\}$, а затем находим потенциал из уравнения

$$\nabla^2 \varphi_p = - \frac{\rho_p^*}{\epsilon_0}, \quad (5.95)$$

либо сначала находим нескорректированный потенциал

$$\nabla^2 \varphi_p^* = - \frac{\rho_p}{\epsilon_0}, \quad (5.96)$$

а потом его исправляем, решая относительно $\{\varphi_p\}$ уравнение

$$\left(1 + \frac{H^2}{8} D^2\right) \varphi_p = \varphi_p^*. \quad (5.97)$$

В обоих этих подходах никак не учитываются конечно-разностные погрешности аппроксимации лапласианов в уравнениях (5.95) и (5.96).

Критерий гладкости, когда его применяют для интерполяции, обычно называют аппроксимацией сплайном. Обращаясь вновь к рис. 5.3, рассмотрим некоторую величину A (например, потенциал или силовое поле), значения которой в сеточных узлах $(-1, 0, 1)$ равны соответственно (A_{-1}^+, A_0^+, A_1^+) . Тогда интерполированное значение A в точке x равно

$$A(x) = W_{-1} A_{-1}^+ + W_0 A_0^+ + W_1 A_1^+. \quad (5.98)$$

Учет симметрии, а также непрерывности $A(x)$ и ее производной в точке $x = H/2$ даст соотношения (5.80)—(5.85). Требуя совпадения $A(x)$ с сеточными значениями для функции $A(x) = \text{const}$, вновь придем к уравнению (5.79) и тем самым к (5.88) для интерполяционной функции.

Интерполяция приводит к расплыванию, подобному обнаруженному для распределения заряда. Как и раньше, это расплывание можно компенсировать, вводя дополнительный этап коррекции. Если мы потребуем, чтобы интерполированная функция $A(x)$ принимала точные значения $A(x_p) = A_p$ в узлах сетки, то соотношение (5.98) даст уравнение относительно сеточных величин $\{A_p^+\}$:

$$A_p = \left(1 + \frac{H^2}{8} D^2\right) A_p^+. \quad (5.99)$$

В случае потенциала уравнение (5.99) просто повторяет способ введения коррекции в распределение заряда.

Для расчета силы методом частица — сетка с коррекциями как распределения заряда, так и интерполяции силы, последовательность этапов принимает следующий вид:

1. Распределение заряда:

$$\rho_p = \frac{q}{H} \sum_{i=1}^{N_p} W(x_i - x_p). \quad (5.100)$$

2. Коррекция плотности заряда:

$$L_1 \rho_p^* = \rho_p. \quad (5.101)$$

3. Нахождение потенциала и поля:

$$\nabla^2 \varphi_p = -\rho_p^* / \epsilon_0, \quad (5.102)$$

$$F_p = -q \frac{d\varphi_p}{dx}. \quad (5.103)$$

4. Коррекция сил:

$$L_2 F_p^+ = F_p. \quad (5.104)$$

5. Интерполяция:

$$F(x_i) = \sum_p W(x_i - x_p) F_p^+ \quad (5.105)$$

Операторы L_1 и L_2 , фигурирующие на этапах 2 и 4, представляют собой операторы «обострения», которые компенсируют расплывание, обусловленное соответственно распределением заряда и интерполяцией. В проведенном рассмотрении сеточные величины подгонялись под взаимодействие между точечными частицами. Практически гораздо реалистичнее ориентироваться на частицы конечного размера, и в этом случае операторы обострения L_1 и L_2 будут отличаться от описанных. Кроме того, нигде пока не учитывалось влияние погрешностей на этапе 3. Несмотря на то что такие погрешности можно включить в рассмотрение описанным выше способом, выкладки становятся чрезвычайно громоздкими, в особенности в двумерном и трехмерном случаях. Как упоминалось во введении, полный анализ лучше всего проводить с помощью спектральных методов.

Чтобы свойства гладкости одномерных функций распределения заряда и интерполяции силы сохранились, функции распределения заряда для двумерных схем нужно брать в виде произведения

$$W(\mathbf{x}) = W(x, y) = W(x)W(y), \quad (5.106)$$

где обе функции, составляющие произведение в правой части (5.106), являются одномерными функциями распределения за-

ряда с требуемой степенью гладкости. Точно так же, в трехмерном случае функции распределения заряда должны иметь вид

$$W(\mathbf{x}) = W(x, y, z) = W(x)W(y)W(z) \quad (5.107)$$

с тем, чтобы получить всюду непрерывность самой величины (первый порядок), непрерывность самой величины и ее первой производной (второй порядок) и т. д. Такие факторизованные выражения обладают дополнительным преимуществом, состоящим в том, что факторизация позволяет вычислить их за относительно небольшое число арифметических операций. Например, для определения весов заряда в 27 узлах по трехмерной схеме TSC количество арифметических операций с плавающей точкой составит не более 33 умножений и 15 сложений на все 27 весов заряда, т. е. приблизительно одно умножение и полтора сложения на сеточный узел для каждой частицы.

Факторизованные формы, которые последовательно удовлетворяют условиям на больших расстояниях и условиям гладкости все более высокого порядка, образуют последовательность схем распределения заряда NGP, CIC, TSC, ..., требующих $1^d, 2^d, 3^d, \dots$, узлов сетки, где d — размерность пространства. Вообще говоря, этим схемам следует отдать предпочтение, хотя в трехмерном случае из-за кубической зависимости числа используемых узлов от порядка схемы более рентабельными могут оказаться промежуточные схемы. Промежуточные схемы можно построить таким образом, чтобы условия гладкости вплоть до некоторого порядка выполнялись всюду, а в определенных направлениях выполнялись условия более высокого порядка. Например, трехточечная схема (см. (5.70)), занимающая по экономичности промежуточное положение между схемами NGP и CIC, условиям на больших расстояниях удовлетворяет в первом порядке (давая квадрупольный остаточный член), а непрерывность изменения заряда и силы при пересечении частицами границ ячеек обеспечивает только вдоль прямых, проходящих через узлы сетки параллельно осям x, y и $|x| = |y|$. Для этой трехточечной схемы максимальный скачок величины заряда, распределяемого в данный сеточный узел от частицы с единичным зарядом, равняется $1/2$, что нужно сравнить с 1 для NGP и 0 для CIC.

5.3.3. Условие сохранения импульса

Любой алгоритм метода частица — сетка тождественно сохраняет импульс (в отсутствие погрешностей округления), если:

1. функции для распределения заряда и интерполяции силы одинаковы и

2. разностные аппроксимации производных правильно сцентрированы по пространству (или, на языке функции Грина, дискретная функция Грина обладает правильной симметрией).

Условие 1 утверждает, что если от частицы с зарядом q , находящейся в точке \mathbf{x} , в сеточный узел \mathbf{p} распределяется заряд $\delta\rho(\mathbf{x}_p)$, равный

$$\delta\rho(\mathbf{x}_p) = \frac{q}{V_c} W(\mathbf{x} - \mathbf{x}_p), \quad (5.108)$$

то сила $\mathbf{F}(\mathbf{x})$, действующая на эту частицу, определяется интерполяционной формулой

$$\mathbf{F}(\mathbf{x}) = \sum_p q W(\mathbf{x} - \mathbf{x}_p) \mathbf{E}(\mathbf{x}_p), \quad (5.109)$$

где $\mathbf{E}(\mathbf{x}_p)$ — сеточное электрическое поле в p -м узле сетки, а сумма берется по всем узлам. Условие 2 требует, чтобы приближенные уравнения, используемые для нахождения сеточного потенциала и полей, можно было формально представить в виде

$$\mathbf{E}(\mathbf{x}_p) = V_c \sum_{p'} \mathbf{d}(\mathbf{x}_p; \mathbf{x}_{p'}) \rho(\mathbf{x}_{p'}) \quad (5.110)$$

где

$$\mathbf{d}(\mathbf{x}_p; \mathbf{x}_{p'}) = -\mathbf{d}(\mathbf{x}_{p'}; \mathbf{x}_p). \quad (5.111)$$

В выражениях (5.108) и (5.110) V_c обозначает объем сеточной ячейки: для одномерных схем $V_c = H$, для двумерных $V_c = H_1 H_2$ и для трехмерных $V_c = H_1 H_2 H_3$. Соотношение (5.111) выражает требование, чтобы компоненты сеточного поля, обусловленные зарядом в точке $\mathbf{x} = 0$, были нечетными функциями x , y и z .

Для того чтобы показать, что условия 1 и 2 гарантируют сохранение импульса, мы должны доказать, что (а) самодействие равно нулю и (б) частицы действуют друг на друга с силами, равными по величине и противоположными по направлению.

(а) Самодействие, $\mathbf{F}_{\text{сам}}$:

Объединяя (5.108) — (5.110) для одной частицы с зарядом q , расположенной в точке \mathbf{x} , получаем

$$\mathbf{F}_{\text{сам}}(\mathbf{x}) = q^2 \sum_{p, p'} \mathbf{d}(\mathbf{x}_p; \mathbf{x}_{p'}) W(\mathbf{x} - \mathbf{x}_p) W(\mathbf{x} - \mathbf{x}_{p'}), \quad (5.112)$$

или, меняя местами p и p' ,

$$= q^2 \sum_{p, p'} \mathbf{d}(\mathbf{x}_{p'}; \mathbf{x}_p) W(\mathbf{x} - \mathbf{x}_{p'}) W(\mathbf{x} - \mathbf{x}_p). \quad (5.113)$$

Складывая (5.112) с (5.113) с учетом (5.111), получаем требуемый результат: $\mathbf{F}_{\text{сам}} = 0$.

(б) *Силы, действующие на частицы:*

Сила, с которой на частицу 1, имеющую заряд q_1 и расположенную в точке x_1 , действует частица 2 с зарядом q_2 в точке x_2 , согласно (5.108) — (5.110) равна

$$F_{12} = q_2 q_1 \sum_{p, p'} W(x_1 - x_p) d(x_p; x_{p'}) W(x_2 - x_{p'}). \quad (5.114)$$

Аналогично, сила, действующая на частицу 2 со стороны частицы 1, есть

$$F_{21} = q_1 q_2 \sum_{p, p'} W(x_2 - x_{p'}) d(x_{p'}; x_p) W(x_1 - x_p). \quad (5.115)$$

Складывая (5.114) с (5.115), приходим к результату

$$F_{21} + F_{12} \equiv 0, \quad (5.116)$$

т. е. силы равны по величине и направлены в противоположные стороны.

5.3.4. Форма облака и форма функции распределения заряда

Иерархии схем распределения заряда, построенной на основе условий на больших расстояниях и условий гладкости, можно придать физическую интерпретацию, если считать, что заряд каждой частицы распределен в области конечных размеров. Преимущество этой интерпретации перед формализмом мультипольного разложения состоит в том, что она позволяет с некоторыми незначительными изменениями применить для исследования вычислительных моделей существующую кинетическую теорию волновых процессов, описывающую сплошную среду дифференциальными уравнениями. Однако частицами конечного размера нужно пользоваться с осторожностью, поскольку, в зависимости от того, какая физическая величина используется для определения их формы, последняя получается различной.

Понятия формы облака и формы функции распределения заряда происходят от несколько разных трактовок сеточной плотности заряда. Если мы пользуемся определением сеточной плотности заряда, которое применялось в разд. 2.2.3, а именно считаем, что плотность заряда в узле сетки есть полный заряд в ячейке, окружающей этот узел, поделенный на объем ячейки, то соответствующая форма — это форма облака. С другой стороны, мы можем считать, что распределение перекрывающихся частиц конечного размера порождает гладко меняющуюся непрерывную плотность заряда, выборочные значения из которой записываются в узлы сетки. В этом случае соответствующая форма — это форма функции распределения заряда.

В качестве формы облака $S(x')$ частицы с единичным зарядом выступает плотность ее заряда, где расстояние x' отсчи-

тывается от центра частицы. Доля заряда, попавшая в p -й сеточный узел с координатой x_p от частицы с формой S , находящейся в точке x , определяется перекрыванием данного облака с ячейкой p , т. е.

$$W(x - x_p) = W_p(x) = \int_{x_p - H/2}^{x_p + H/2} S(x' - x) dx'. \quad (5.117)$$

Используя функцию прямоугольного импульса Π (см. приложение) и формулу (5.117), получаем связь между формой облака S и функцией W :

$$W(x) = \int \Pi\left(\frac{x'}{H}\right) S(x' - x) dx'. \quad (5.118)$$

Если мы ограничимся для S четными функциями, то (5.118) можно представить в компактной записи:

$$W(x) = \Pi\left(\frac{x}{H}\right) * S(x). \quad (5.119)$$

С помощью метода, аналогичного использованному нами в разд 5.3.1 и 5.3.2 для функции распределения заряда, можно получить форму облаков S для иерархии схем. Поскольку, однако, функции W уже получены, мы можем для нахождения формы воспользоваться соотношением (5.119). Используя определения функций Π и Λ , приведенные в приложении, можно переписать известные функции распределения заряда для схем NGP, CIC и TSC (формулы (2.28), (5.44) и (5.88) соответственно) в следующем виде:

$$\text{NGP: } W(x) = \Pi\left(\frac{x}{H}\right) \equiv \frac{1}{H} \Pi\left(\frac{x}{H}\right) * \delta\left(\frac{x}{H}\right), \quad (5.120)$$

$$\text{CIC: } W(x) = \Lambda\left(\frac{x}{H}\right) = \frac{1}{H} \Pi\left(\frac{x}{H}\right) * \Pi\left(\frac{x}{H}\right), \quad (5.121)$$

$$\text{TSC: } W(x) = \frac{1}{H} \Lambda\left(\frac{x}{H}\right) * \Pi\left(\frac{x}{H}\right) = \frac{1}{H^2} \Pi\left(\frac{x}{H}\right) * \Pi\left(\frac{x}{H}\right) * \Pi\left(\frac{x}{H}\right). \quad (5.122)$$

Следовательно, облака имеют следующую форму:

$$\text{NGP: } S(x) = \frac{1}{H} \delta\left(\frac{x}{H}\right) = \delta(x), \quad (5.123)$$

$$\text{CIC: } S(x) = \frac{1}{H} \Pi\left(\frac{x}{H}\right), \quad (5.124)$$

$$\text{TSC: } S(x) = \frac{1}{H} \Lambda\left(\frac{x}{H}\right). \quad (5.125)$$

Каждая последующая функция распределения заряда более высокого порядка получается сверткой предыдущей функции с $(1/H)\Pi(x/H)$, так что приведенную в табл. 5.1 последовательность схем можно продолжать неограниченно.

Таблица 5.1. Иерархия факторизованных схем распределения заряда ($H = 1$)

Схема	Порядок	Число узлов	Форма облака	Форма функции распределения заряда	Сила
NGP	0	1^d	δ	Π	Кусочно-постоянная
CIC	1	2^d	Π	$\Lambda = \Pi * \Pi$	Непрерывная кусочно-линейная
TSC	2	3^d	Λ	$\Pi * \Pi * \Pi$	Непрерывная вместе с первой производной
PQS	3	4^d	$\Lambda * \Pi$	$\Pi * \Pi * \Pi * \Pi$	Непрерывная вместе с первой и второй производными

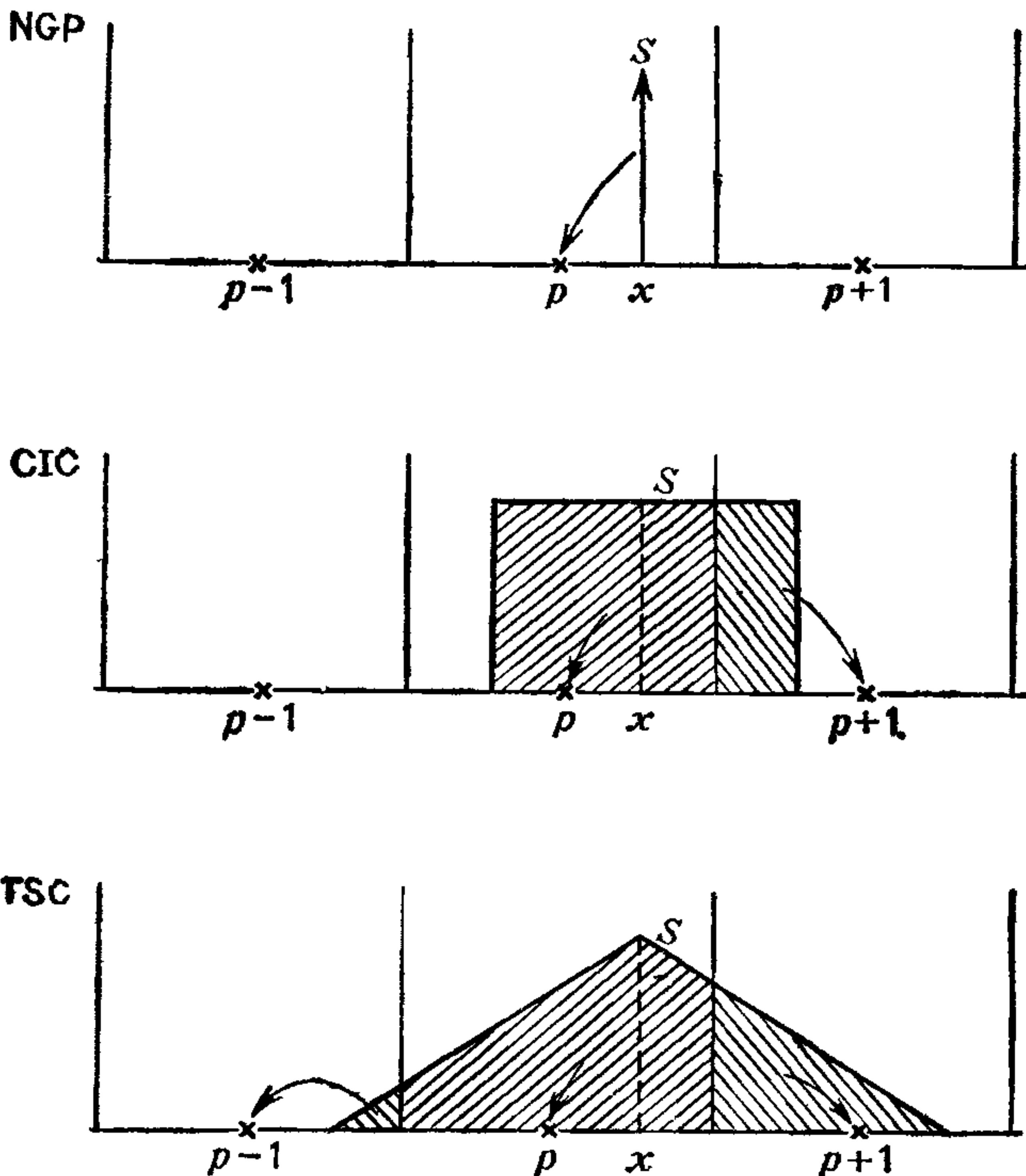


Рис 5.4 Интерпретация распределения заряда на основе формы облака. Доля заряда, приписываемого данному сеточному узлу от частицы с координатой x , равна площади перекрывания облака формы S с ячейкой, содержащей этот узел

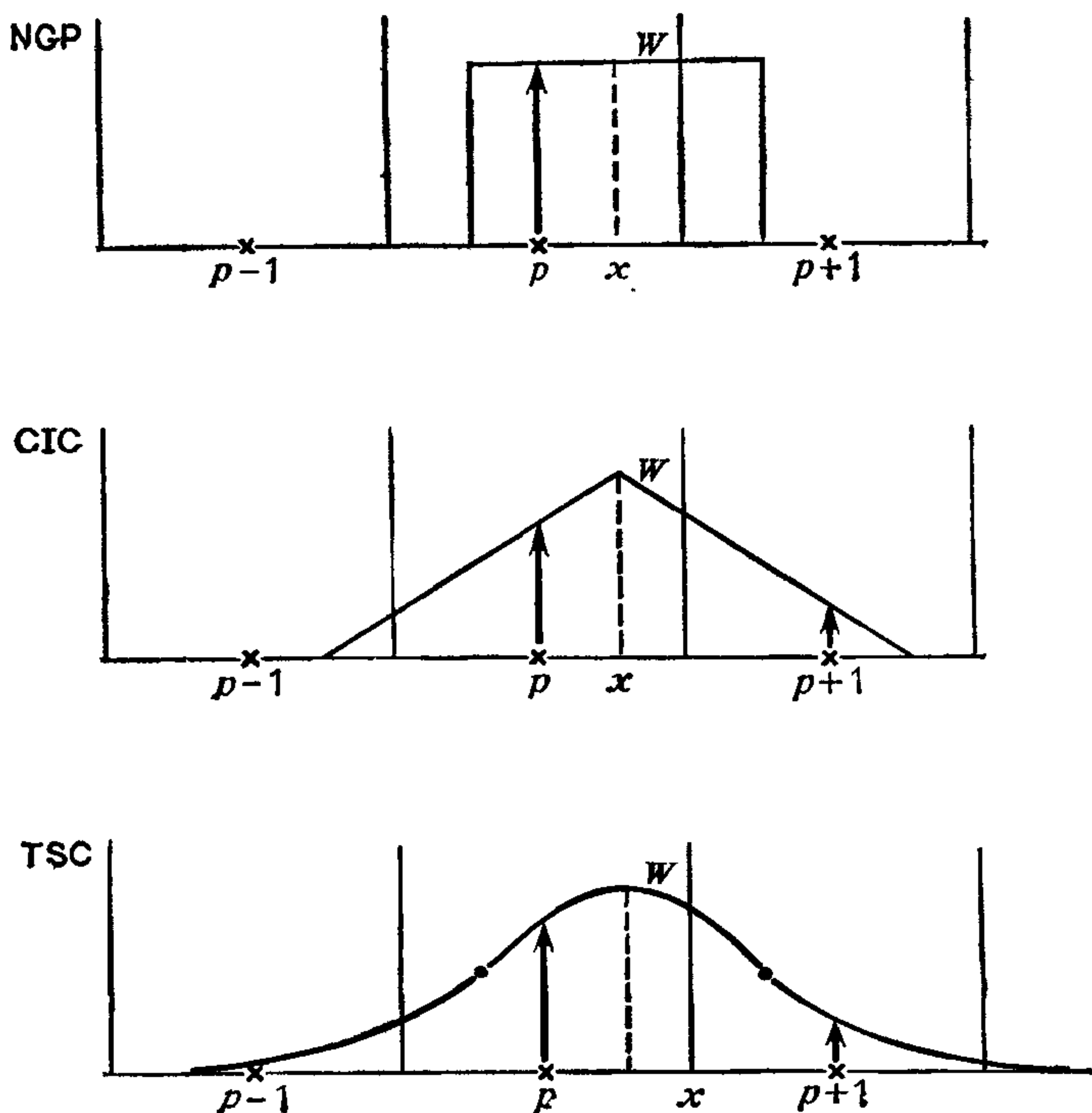


Рис. 5.5. Интерпретация распределения заряда на основе формы весовой функции. Доля заряда, приписываемого данному сеточному узлу от частицы, находящейся в точке x , равна значению весовой функции W в этом узле.

Рис. 5.4 дает наглядное представление об этой трактовке формы облака. С частицей связывают облако определенной формы, отвечающей данной схеме распределения заряда. Площадь перекрывания облака с ячейкой определяет долю заряда, приписываемую сеточному узлу в центре этой ячейки. Соответствующая ситуация для интерпретации процедуры распределения заряда на основе формы функции распределения заряда показана на рис. 5.5. В этом случае с частицами связывают функцию распределения, и значение этой функции в каждом узле сетки дает долю заряда, приписанного к этому узлу.

В многомерных версиях схем, перечисленных в табл. 5.1, форма облаков S и форма функций распределения W выражаются произведением соответствующих одномерных функций для каждой компоненты. На рис. 5.6 показана графическая интерпретация двумерной схемы CIC: произведение функций Π придает частице форму, проекция которой совпадает с ячейкой сетки. Площадь перекрывания облака дает долю заряда, рас-

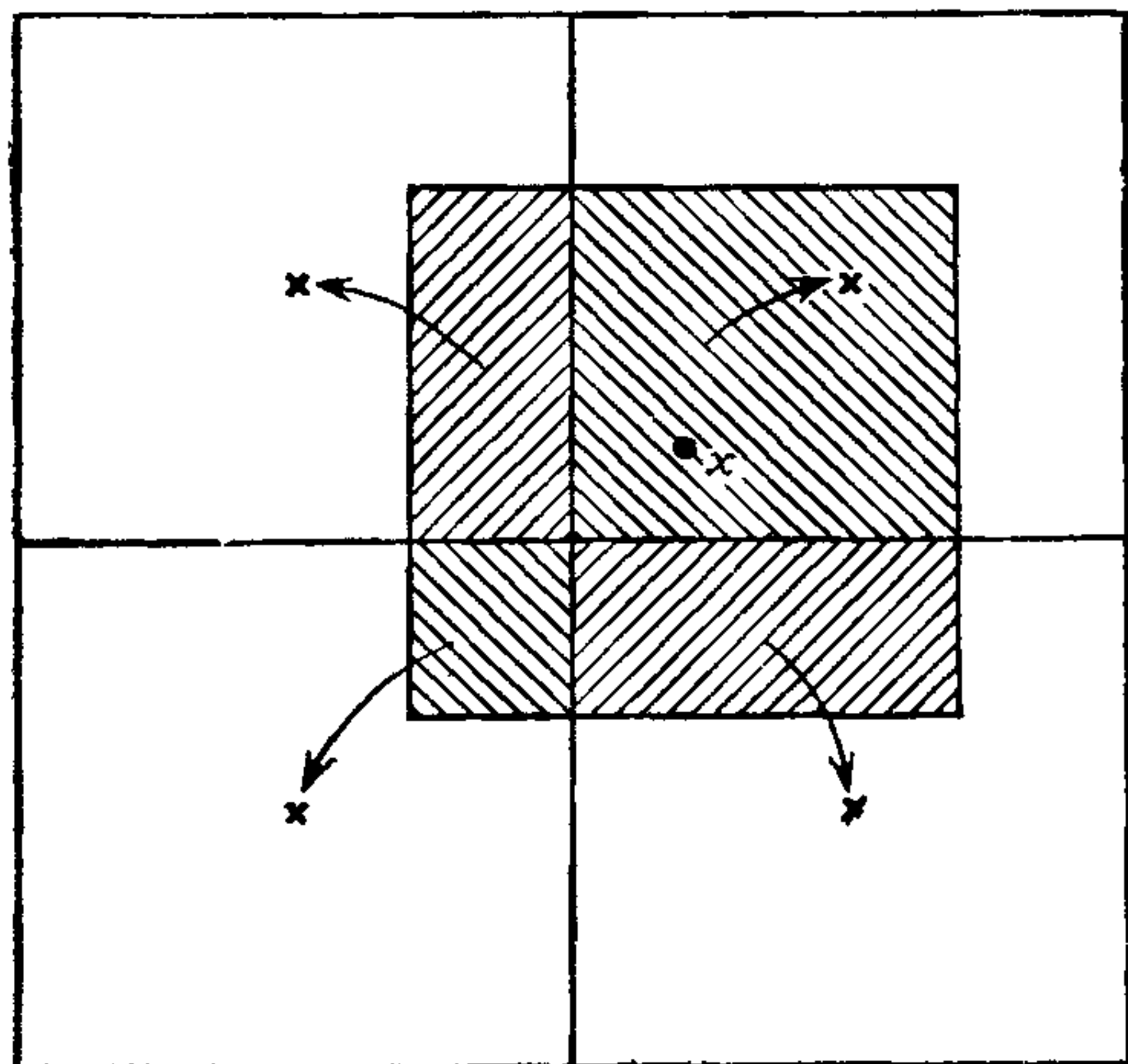


Рис 5.6. Двумерная схема СІС или пространственного взвешивания. Доля заряда, приписываемого четырем ближайшим сеточным узлам от частицы с координатой x , определяется площадью перекрывания ее облака с ячейками, содержащими эти узлы.

пределяемого в каждый из четырех узлов сетки вокруг данной частицы: именно поэтому схему СІС иногда называют «пространственным взвешиванием» [164, 290].

Форму облаков можно определить также и для тех двумерных и трехмерных схем, которые не записываются в виде произведения. Например, для трехточечной схемы распределения (5.70) облако имеет форму

$$S(x, y) = \frac{1}{H_1 H_2} \left[\Pi \left(\frac{x}{H_1} \right) \delta \left(\frac{y}{H_2} \right) + \delta \left(\frac{x}{H_1} \right) \Pi \left(\frac{y}{H_2} \right) - \delta \left(\frac{x}{H_1} \right) \delta \left(\frac{y}{H_2} \right) \right], \quad (5.126)$$

а функция распределения заряда

$$W(x, y) = \Pi \left(\frac{x}{H_1} \right) \Pi \left(\frac{y}{H_2} \right) * S(x, y) \quad (5.127)$$

$$= \Lambda \left(\frac{x}{H_1} \right) \Pi \left(\frac{y}{H_2} \right) + \Pi \left(\frac{x}{H_1} \right) \Lambda \left(\frac{y}{H_2} \right) - \Pi \left(\frac{x}{H_1} \right) \Pi \left(\frac{y}{H_2} \right). \quad (5.128)$$

Облако этой трехточечной схемы, если его изобразить тем же способом, что и на рис. 5.6 для схемы СІС, имеет вид креста, образованного из отрезков, параллельных сторонам ячейки и имеющих ее размеры.

При описании вычисления силы методом частица-сетка пользуются еще одной формой частицы конечного размера — эффективной формой частицы. Если все этапы вычисления силы свести воедино, то получится некоторая приближенная сила взаимодействия частиц, которая будет отличаться от закона силы точечных частиц. Если для удаления флуктуирующей компоненты

эту приближенную силу усреднить по всем смещениям относительно сетки, то, считая взаимодействие точечным и соответствующим физическому закону, полученную силу можно использовать для определения эффективной формы частицы. Если для получения количественных результатов по модели частица-сетка мы захотим обратиться к математическим методам, применяемым для исследования физических систем, то нужно использовать именно эффективную форму частицы.

5.4. ПОГРЕШНОСТИ АППРОКСИМАЦИИ

В одномерном примере, который обсуждался в разд. 5.2, погрешности аппроксимации отсутствовали всюду, кроме самой точки источника, из-за характера решения одномерного уравнения Пуассона. В двумерном и трехмерном случаях такое отсутствие погрешностей для простых конечно-разностных аппроксимаций не имеет места.

В простейшей схеме решения уравнения Пуассона используется пятиточечная аппроксимация лапласиана, которую для квадратной сетки ($H_1 = H_2 = H$) можно записать в виде

$$\nabla^2 \varphi \simeq D^2 \varphi = \frac{\varphi_{p_1+1, p_2} - 2\varphi_{p_1, p_2} + \varphi_{p_1-1, p_2}}{H^2} + \frac{\varphi_{p_1, p_2+1} - 2\varphi_{p_1, p_2} + \varphi_{p_1, p_2-1}}{H^2}, \quad (5.129)$$

где сеточный узел $\mathbf{p} = (p_1, p_2)$ расположен в точке $\mathbf{x}_p = \mathbf{p}H$. Разлагая $D^2 \varphi$ в ряд Тейлора в окрестности точки $\mathbf{x} = \mathbf{x}_p$, получаем

$$\nabla^2 \varphi = D^2 \varphi - \frac{H^2}{12} \left(\frac{\partial^4 \varphi}{\partial x^4} + \frac{\partial^4 \varphi}{\partial y^4} \right) + \text{члены высших порядков}. \quad (5.130)$$

Таким образом, пятиточечная аппроксимация имеет второй порядок точности, причем второй член в правой части (5.130) дает погрешность аппроксимации δ .

Для единичного заряда (стержня), расположенного в начале координат, оценку погрешности аппроксимации можно получить подстановкой известного решения $\varphi = \log r / 2\pi\epsilon_0$ соответствующей задачи в выражение для погрешности аппроксимации, что дает

$$\delta = -H^2 \frac{\cos 4\theta}{2\pi\epsilon_0 r^4}, \quad (5.131)$$

где θ — азимутальный угол. Важно отметить, что погрешность аппроксимации в (5.131) будет приводить к погрешности величины потенциала, зависящей как от расстояния от исходного заряда, так и от направления на него. Точное обращение пяти-

точечного оператора (5.129), выполненное Бунеманом [39] для единичного заряда, показывает, что величина погрешности потенциала ведет себя как $\cos 4\theta$, причем наибольшие отрицательные погрешности имеют место вдоль осей $\pm x$ и $\pm y$, а наибольшие положительные погрешности — вдоль осей $|x| = |y|$.

Чтобы уменьшить погрешность аппроксимации и связанную с ней угловую анизотропию потенциала, можно прибегнуть к разностным аппроксимациям более высокого порядка, например, девятиточечной аппроксимации [60]. Однако если граничные условия таковы, что запись потенциала через функцию Грина для уравнения Пуассона принимает инвариантную по отношению к смещениям форму

$$\varphi(\mathbf{x}_p) = V_c \sum_{p'} G(\mathbf{x}_p - \mathbf{x}_{p'}) \rho(\mathbf{x}_{p'}), \quad (5.132)$$

то вычислительные затраты на нахождение решения через функцию Грина будут по порядку величины такими же, что и для разностных уравнений (см. гл. 6). Если рассматривать значения G в качестве подгоночных параметров, то погрешность аппроксимации потенциала можно сделать любой, какую мы выберем. В расчетах методом частица-сетка самый разумный выбор тот, который дает самую хорошую аппроксимацию сил.

Еще один источник погрешностей аппроксимации связан с заменой градиента потенциала на конечно-разностную аппроксимацию. Точное электрическое поле

$$\mathbf{E} = -\nabla\varphi \quad (5.133)$$

заменяется симметричной двухточечной конечно-разностной схемой второго порядка точности

$$\mathbf{E}(\mathbf{x}_p) = -\left(\frac{\varphi_{p_1+1, p_2} - \varphi_{p_1-1, p_2}}{2H}, \frac{\varphi_{p_1, p_2+1} - \varphi_{p_1, p_2-1}}{2H} \right). \quad (5.134)$$

Из разложения (5.134) в ряд Тейлора получим погрешность аппроксимации \mathbf{e} электрического поля:

$$\mathbf{e} = \frac{H^2}{6} \left(\frac{\partial^3 \varphi}{\partial x^3}, \frac{\partial^3 \varphi}{\partial y^3} \right). \quad (5.135)$$

В случае единичного заряда в начале координат выражение (5.135) принимает вид

$$\mathbf{e} = \frac{H^2}{6\pi\epsilon_0 r^3} (-\cos 3\theta, \sin 3\theta). \quad (5.136)$$

Таким образом, в самом первом приближении абсолютная величина погрешности $|\mathbf{e}|$ не зависит от направления. Направление погрешности совпадает с направлением точного поля, когда $\operatorname{tg} 3\theta = \pm \operatorname{tg} \theta$, т. е. при $\theta = n\pi/4$, и перпендикулярно точному полю при $\theta = (2n + 1)\pi/8$, где n целое.

Аппроксимация градиента в следующем порядке требует четырех узлов сетки по каждой компоненте. Эта четырехточечная схема принимает вид линейной комбинации центрированных разностных аппроксимаций по двум ближайшим к данному узлам и двум следующим ближайшим соседям:

$$E_x(\mathbf{x}_p) = -\alpha \frac{(\varphi_{p_1+1, p_2} - \varphi_{p_1-1, p_2})}{2H} - (1-\alpha) \frac{(\varphi_{p_1+2, p_2} - \varphi_{p_1-2, p_2})}{4H}. \quad (5.137)$$

Если в (5.137) положить $\alpha = 4/3$, то члены $O(H^2)$ в погрешности аппроксимации взаимно уничтожаются и остается

$$\varepsilon = -\frac{H^4}{30} \left(\frac{\partial^5 \varphi}{\partial x^5} \cdot \frac{\partial^5 \varphi}{\partial y^5} \right). \quad (5.138)$$

В случае единичного заряда в начале координат абсолютная величина погрешности вновь не зависит от направления и ориентирована вдоль направления точного поля при $\theta = n\pi/4$ или $n\pi/6$, где n целое.

В принципе, используя интегральное представление для поля, погрешности аппроксимации поля можно исключить. Практически же такой метод используется редко, поскольку подобный расчет обычно предъявляет высокие требования к памяти и быстродействию ЭВМ. Принятый в алгоритме Р³М (гл. 8) компромисс состоит в том, что для компонент поля используется схема (5.137), а значения функции Грина в (5.132) подбираются таким образом, чтобы компенсировать погрешность в величине силы.

5.5. СХЕМЫ, СОХРАНЯЮЩИЕ ЭНЕРГИЮ

До сих пор при анализе методов расчета РМ-силы использовался подход, в котором все сеточные приближения рассматривались как конечно-разностные. Другой метод, введенный Х. Р. Льюисом [252, 253], основан на использовании метода конечных элементов. В результате применения метода Льюиса получаются так называемые «схемы, сохраняющие энергию». Преимущества формализма конечных элементов состоят в том, что он приводит к классу РМ-алгоритмов, которые являются оптимальными для выбранных базисных функций, и что он может быть применен к полностью электромагнитным системам, сложным геометриям и сеткам произвольной структуры. Тем не менее для рассматриваемых в этой книге моделей РМ и Р³М обнаруживается, что более рентабельным вариантом являются оптимизированные конечно-разностные схемы («сохраняющие импульс»).

Метод конечных элементов исходит из интегрального, а не дифференциального описания физической системы. Для дискре-

тизации системы фигурирующие в интегралах функции заменяются функциями из ограниченного класса, причем этот класс содержит функции, непрерывные вместе с достаточным числом производных, что позволяет избежать сингулярностей в подынтегральных выражениях (т. е. функции «хорошие»). Для квадратичных интегралов, возникающих, например, в схемах, сохраняющих энергию, дискретные линейные уравнения получаются применением к приближенному выражению для интеграла вариационных методов.

Соответствующий интеграл для системы заряженных частиц, взаимодействующих по закону Кулону, имеет вид

$$I[\Phi, \mathbf{x}] = \int_{t_0}^{t_1} L dt, \quad (5.139)$$

где L — общий лагранжиан частиц и поля [145]:

$$L = \int d\mathbf{x}_0 d\mathbf{v}_0 f(\mathbf{x}_0, \mathbf{v}_0, 0) \left[\frac{1}{2} m v^2 - q\Phi(\mathbf{x}) \right] + \int d\mathbf{x}' \left(\epsilon_0 \frac{(\nabla\Phi)^2}{2} - \rho_0 \Phi \right). \quad (5.140)$$

Интеграл I является функционалом потенциала Φ и координат частиц \mathbf{x} . При минимизации I вновь получают дифференциальные уравнения, описывающие систему. Первый интеграл в (5.140) берется по всей расчетной области и всем скоростям, $f(\mathbf{x}_0, \mathbf{v}_0, 0)$ обозначает функцию распределения частиц в момент времени $t = 0$, а положение \mathbf{x} и скорость $\mathbf{v} = \dot{\mathbf{x}}$ являются лагранжевыми координатами (соответствующими орбитам частиц):

$$\mathbf{x} = \mathbf{x}(\mathbf{x}, \mathbf{v}, t), \quad \mathbf{v} = \dot{\mathbf{x}}(\mathbf{x}_0, \mathbf{v}_0, t). \quad (5.141)$$

Второй интеграл берется по расчетной области. Величина ρ_0 представляет собой фоновую плотность заряда.

Вариация интеграла (5.139) по \mathbf{x} дает

$$\begin{aligned} \delta_{\mathbf{x}} I &= \int dt \int d\mathbf{x}_0 d\mathbf{v}_0 f(\mathbf{x}_0, \mathbf{v}_0, 0) [m\mathbf{v} \cdot \delta\dot{\mathbf{x}} - q\delta\mathbf{x} \cdot \nabla\Phi] = \\ &= \int dt \int d\mathbf{x}_0 d\mathbf{v}_0 f(\mathbf{x}_0, \mathbf{v}_0, 0) [-m\dot{\mathbf{v}} - q\nabla\Phi] \cdot \delta\mathbf{x} + \\ &+ \int d\mathbf{x}_0 d\mathbf{v}_0 f(\mathbf{x}_0, \mathbf{v}_0, 0) m\mathbf{v} \cdot \delta\mathbf{x} \Big|_{t_0}^{t_1}. \end{aligned} \quad (5.142)$$

Ограничиваясь вариациями, которые при (произвольных) временах t_0 и t_1 обращаются в нуль¹⁾, получим, что второй член

¹⁾ Это означает, что \mathbf{x} варьируется в классе функций, совпадающих при $t = t_0$ и $t = t_1$ с искомым решением. — Прим. перев.

в (5.142) равен нулю. Далее, поскольку для всех остальных вариаций δx первый член также должен обращаться в нуль, получим

$$m\dot{\mathbf{v}} = -q\nabla\Phi, \quad (5.143)$$

т. е. уравнение движения. Вариация по Φ дает

$$\delta_{\Phi}I = \int dt \left\{ \int d\mathbf{x}_0 d\mathbf{v}_0 f(\mathbf{x}_0, \mathbf{v}_0, 0) [-q\delta\Phi(x)] + \right. \\ \left. + \int d\mathbf{x}' [\epsilon_0 \nabla\varphi \cdot \nabla\delta\Phi - \rho_0 \delta\Phi(\mathbf{x}')] \right\}, \quad (5.144)$$

Используя граничные условия для Φ на границе расчетной области и тот факт, что $f(\mathbf{x}, \mathbf{v}, t) = f(\mathbf{x}_0, \mathbf{v}_0, 0)$ (см. разд. 2.2.1), можно записать (5.144) в виде

$$\delta_{\Phi}I = \int dt \int d\mathbf{x} \left[-q \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} - \epsilon_0 \nabla^2\Phi - \rho_0 \right] \delta\Phi = 0. \quad (5.145)$$

Поскольку выражение (5.145) справедливо для любых $\delta\Phi$, удовлетворяющих заданным граничным условиям, из этого следует, что

$$\nabla^2\Phi = -\frac{1}{\epsilon_0} \left[\rho_0 + q \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} \right]. \quad (5.146)$$

Аналогично можно определить лагранжианы для нескольких сортов частиц и для полей более общего вида [252].

Для построения дискретного аналога лагранжиана (5.140) заменим Φ приближенным выражением:

$$\Phi \simeq \varphi(\mathbf{x}, t) = \sum_p \varphi_p(t) W_p(\mathbf{x}), \quad (5.147)$$

где сумма берется по всем узлам сетки, $\{\varphi_p\}$ — сеточные значения потенциала и $W_p(\mathbf{x})$ — базисные функции. Для согласованности необходимо, чтобы φ была по крайней мере непрерывной. Следовательно, если выражение (5.147) представляет собой кусочно-полиномиальную аппроксимацию на прямоугольной сетке, $W_p(\mathbf{x})$ должна быть весовой функцией схемы СИС или более высокого порядка (см. разд. 5.3.2 и 5.3.4). Записывая функцию распределения частиц в виде

$$f(\mathbf{x}_0, \mathbf{v}_0) = \sum_{i=1}^{N_p} \delta(\mathbf{x}_i - \mathbf{x}_{0i}) \delta(\mathbf{v}_i - \mathbf{v}_{0i}), \quad (5.148)$$

получим явную связь с орбитами N_p частиц.

Подстановка (5.147) и (5.148) в (5.139) и (5.140) дает аппроксимацию интеграла

$$I = \int dt L = \int dt \left\{ \sum_{i=1}^{N_p} \left[\frac{1}{2} m v_i^2 - q \sum_p \varphi_p W_p(\mathbf{x}_i) \right] + \right. \\ \left. + \int d\mathbf{x}' \left[\frac{\epsilon_0}{2} \left(\sum_p \varphi_p \nabla W_p(\mathbf{x}') \right)^2 - \rho_0 \sum_p \varphi_p W_p(\mathbf{x}') \right] \right\}. \quad (5.149)$$

Вновь применяя к интегралу (5.149) вариационный принцип, приходим к уравнениям Эйлера — Лагранжа:

$$\frac{\partial L}{\partial \mathbf{x}_i} - \frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}_i} = 0, \quad i \in [1, N_p], \quad (5.150)$$

т. е.

$$m \dot{\mathbf{v}}_i = -q \sum_p \varphi_p \nabla W_p(\mathbf{x}_i) \quad (5.151)$$

и

$$\frac{\partial L}{\partial \varphi_p} = 0, \quad (5.152)$$

т. е.

$$\epsilon_0 \sum_{p'} \varphi_{p'} \int d\mathbf{x}' \nabla W_{p'}(\mathbf{x}') \cdot \nabla W_p(\mathbf{x}') = \rho_0 \int d\mathbf{x}' W_p(\mathbf{x}') + q \sum_{i=1}^{N_p} W_p(\mathbf{x}_i). \quad (5.153)$$

Происхождение названия «схемы, сохраняющие энергию» можно понять, если записать скалярное произведение уравнения (5.151) с \mathbf{v} и воспользоваться тем, что φ — неявная функция времени:

$$\frac{d}{dt} \frac{m v^2}{2} = -\mathbf{v} \cdot \nabla \left(q \sum_p \varphi_p W_p \right) = -\frac{d}{dt} \left(q \sum_p \varphi_p W_p \right),$$

$$\text{т. е.} \quad \frac{m v_i^2}{2} + q \sum_p \varphi_p W_p = \text{const.} \quad (5.154)$$

К сожалению, точное сохранение энергии, предполагаемое равенством (5.154), утрачивается при замене производных по времени конечно-разностными аппроксимациями [254].

После выбора функции W уравнения (5.151) и (5.153) полностью определяют оптимальную (в смысле минимизации I) схему, сохраняющую энергию. Например, в одномерном случае выбор W , соответствующий самому низкому порядку, а именно функции распределения заряда схемы СИС (см. (5.121)), дает

для уравнений (5.151) и (5.153)

$$m\dot{v} = q \frac{\Phi_p - \Phi_{p+1}}{H} \Pi \left(\frac{x - x_{p+1/2}}{H} \right), \quad (5.155)$$

$$\frac{\Phi_{p-1} - 2\Phi_p + \Phi_{p+1}}{H^2} = -\frac{1}{\epsilon_0} \left[\rho_0 + \frac{q}{H} \sum_{i=1}^{N_p} \Lambda \left(\frac{x_i - x_p}{H} \right) \right]. \quad (5.156)$$

Правая часть уравнения (5.156) — это обычное распределение заряда по одномерной схеме CIC. Однако выражение для интерполяции силы в правой части (5.155) отвечает схеме NGP, но на сетке, смещенной на пол-ячейки относительно сетки для заряда. Электрическое поле на этой смещенной сетке определяется по центрированной разностной схеме с половинным шагом по сравнению по схемой NGP, сохраняющей импульс (разд. 5.2.2). Из приведенного в разд. 5.2.3 рассмотрения следует, что вследствие самодействия такая смешанная схема CIC/NGP не сохраняет импульс. Изолированная частица, помещенная в произвольном месте сетки, будет колебаться около границ ячейки с частотой [235]

$$\omega_{\text{сам}} = \left(\frac{q^2}{\epsilon_0 m H} \right)^{1/2}, \quad (5.157)$$

причем отличие этой частоты в $\sqrt{2}$ раз от частоты, определяемой формулой (5.31), связано с разным определением электрических полей в этих двух случаях.

5.6. СПЕКТРАЛЬНЫЙ МЕТОД

Преобразование Фурье или разложение в ряды Фурье обладает четырьмя характерными свойствами, которые делают его особенно ценным при исследовании методов вычисления РМ-силы: теорема о свертке, теорема об аналитическом представлении сигнала, компактность представления спектров кусочно-полиномиальных функций и связь прямого и обратного преобразований позволяют провести волновое и кинетическое рассмотрение физической системы. Инициатором приложения спектральных методов к РМ-моделям выступил Бэрдсол с сотрудниками в Беркли [38, 238, 299]. В частности, именно работы Лэнгдона [233—237] создали современную прочную теоретическую основу для моделирования бесстолкновительных систем. Осознание того, что дает спектральный метод, привело к усовершенствованиям к разработке бесстолкновительных моделей [75, 115, 116, 118] и позволило создать вполне жизнеспособный алгоритм Р³М [[117, 120].

Этапы вычисления РМ-силы суммарно выражаются следующими уравнениями:

Распределение заряда

$$\rho(\mathbf{x}_p) = \frac{q}{V_c} \sum_{i=1}^{N_p} W(\mathbf{x}_i - \mathbf{x}_p). \quad (5.158)$$

Нахождение потенциала

$$\varphi(\mathbf{x}_p) = V_c \sum_{p'} G(\mathbf{x}_p - \mathbf{x}_{p'}) \rho(\mathbf{x}_{p'}). \quad (5.159)$$

Вычисление сил

$$\mathbf{E}(\mathbf{x}_p) = -\mathbf{D}\varphi(\mathbf{x}_p), \quad (5.160)$$

$$\mathbf{F}(\mathbf{x}_i) = q \sum_p W(\mathbf{x}_i - \mathbf{x}_p) \mathbf{E}(\mathbf{x}_p) \quad (5.161)$$

или
$$\mathbf{F}(\mathbf{x}_i) = -q \sum_p \nabla W(\mathbf{x}_i - \mathbf{x}_p) \varphi(\mathbf{x}_p), \quad (5.162)$$

где система обозначений совпадает с использованной ранее, но с добавлением в (5.160) условного обозначения \mathbf{D} для оператора дифференцирования потенциала. В уравнении (5.159) функцию Грина G можно определить так, чтобы она вобрала в себя операторы «обострения» (см. разд. 5.3.2), которые можно включить в расчетный цикл. При проведении преобразования Фурье уравнений (5.158) — (5.162) мы будем для ясности брать в качестве иллюстрации случай одномерной бесконечной системы с шагом сетки H ; обобщение на более высокие размерности, которое сопряжено с заменой скаляров векторами и переопределением, где это необходимо, нормировочных коэффициентов будет подробнее рассмотрено в гл. 7 и 8.

5.6.1. Распределение заряда

Если плотность центров частиц в одномерной бесконечной системе определить в виде

$$n(x) = \sum_{i=1}^{N_p} \delta(x - x_i), \quad (5.163)$$

где x_i — положение i -й частицы, и если (что всегда имеет место для рассматриваемых случаев) W — четная функция, то распределение заряда (5.158) описывается формулой

$$\rho(x_p) = \frac{q}{H} \int n(x') W(x_p - x') dx'. \quad (5.164)$$

Сравнивая (5.164) с формулами из табл. П.3, мы видим, что сеточные значения плотности заряда представляют собой отдельные значения (выбираемые в точках $x_p = pN$) свертки плотности частиц с функцией распределения заряда, т. е.

РАСПРЕДЕЛЕНИЕ ЗАРЯДА \equiv **СВЕРТКА + ДИСКРЕТИЗАЦИЯ**

Для непрерывной плотности заряда

$$\rho'(x) = \frac{q}{H} \int n(x') W(x - x') dx', \quad (5.165)$$

значения которой при $x = x_p$ дают сеточную плотность, можно легко найти фурье-образ, используя теорему о свертке (табл. П. 3):

$$\rho'(x) \supset \hat{\rho}'(k) = \frac{q}{H} \hat{n}(k) \hat{W}(k). \quad (5.166)$$

Символ \supset обозначает «преобразуется в» или «чей образ равен», а «крышка» над буквой указывает на то, что данная величина является фурье-образом (или гармоникой) соответствующей величины в x -пространстве, обозначаемой тем же символом. Мы употребляем один и тот же символ \supset для преобразования Фурье, разложения в ряд и конечного преобразования Фурье, поскольку в каждом конкретном случае тип преобразования обычно очевиден из контекста. Например, в случае ρ' из соотношений (П.1) и (П.2) мы имеем

$$\rho'(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{\rho}'(k) e^{ikx}, \quad (5.167)$$

$$\hat{\rho}'(k) = \int_{-\infty}^{\infty} dx \rho'(x) e^{-ikx}. \quad (5.168)$$

Соответствующим преобразованием для набора величин, определяемых выражением (5.164), является ряд Фурье, т. е. (П.7) и (П.8). Результат преобразования $\{\rho(x_p)\}$ имеет вид

$$\hat{\rho}(k) = H \sum_{p=-\infty}^{\infty} \rho(x_p) e^{-ikx_p}, \quad (5.169)$$

а обратного преобразования

$$\rho(x_p) = \int_{k_g} \frac{dk}{2\pi} \hat{\rho}(k) e^{ikx_p}, \quad (5.170)$$

где $k_g = 2\pi/H$ — длина периода в k -пространстве. Однако поскольку $\rho'(x_p) = \rho(x_p)$, мы можем воспользоваться (5.167) и получить

$$\rho(x_p) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{\rho}'(k) e^{ikx_p}. \quad (5.171)$$

Разбивая область интегрирования в (5.171) на участки длиной k_g и производя замену переменных ($k \rightarrow k + nk_g$, n целое) для

приведения всех интегралов к одному интервалу интегрирования длиной k_g , получим

$$\rho(x_p) = \int_{k_g} \frac{dk'}{2\pi} \sum_{n=-\infty}^{\infty} \hat{\rho}'(k') e^{ik'x_p}, \quad (5.172)$$

что при сравнении с (5.170) дает

$$\hat{\rho}(k) = \sum_{n=-\infty}^{\infty} \hat{\rho}'(k - nk_g) = \quad (5.173)$$

$$= \frac{q}{H} \sum_{n=-\infty}^{\infty} \hat{n}(k - nk_g) \hat{W}(k - nk_g). \quad (5.174)$$

Тот же результат, а именно (5.173), можно получить строго, вычисляя пределы соответствующих определенных интегралов с конечными пределами и сумм, однако, как будет показано ниже, операции над фурье-образами удобнее выполнять с помощью обобщенных функций.

На рис. 5.7(а) показана графическая интерпретация вещественной части выражения (5.173). Фурье-образ $\hat{\rho}$ набора сеточных значений получается суммированием вкладов от копий образа $\hat{\rho}'$ непрерывной функции плотности, при этом копии смещены друг относительно друга на длину, кратную $k_g (= 2\pi/H)$. Вклад в $\hat{\rho}$ члена $\hat{\rho}'(k - nk_g)$ называется «наложением n -ой сеточной гармоники», а сумма по n в (5.173) — «суммой по сеточным гармоникам». Аналогичная графическая интерпретация справедлива и для мнимой части выражения (5.173) с тем единственным отличием, что мнимые части $\hat{\rho}$ и $\hat{\rho}'$ являются нечетными функциями от k , поскольку (вещественное) \supset (эрмитово) (см. табл. П. 2).

Происхождение названия «наложение» становится понятным, если рассмотреть отдельную гармонику в основной зоне ($|k| \leq \pi/H$). На рис. 5.7(б) сплошной кривой показано поведение в k -пространстве одной гармоники $\hat{\rho}'$ с волновым числом $k = \pi/2H$, или, что эквивалентно, с длиной волны $\lambda = 4H$. Амплитуда этой синусоиды равна $|\hat{\rho}'(\pi/2H)|$, а ее фаза определяется соотношением вещественной и мнимой частей $\hat{\rho}'(\pi/2H)$. Две штриховые кривые — это синусоиды, отвечающие сеточным гармоникам с n , равным $+1$ ($\lambda_{+1} = 4H/3$) и -1 ($\lambda_{-1} = 4H/5$). Амплитуды этих двух синусоид определяются соответственно величинами $|\hat{\rho}'(k - k_g)|$ и $|\hat{\rho}'(k + k_g)|$, где $k = \pi/2H$. На рис. 5.7(б) крестики на оси x соответствуют узлам сетки. При выборе дискретных значений $\rho'(x)$ в точках x_p узлы сетки попадают последовательно в нуль, максимум, нуль, минимум основной гармоники и сеточных гармоник $+1$ и -1 , при условии, что относительные фазы такие, как показаны на рис. 5.7(б).

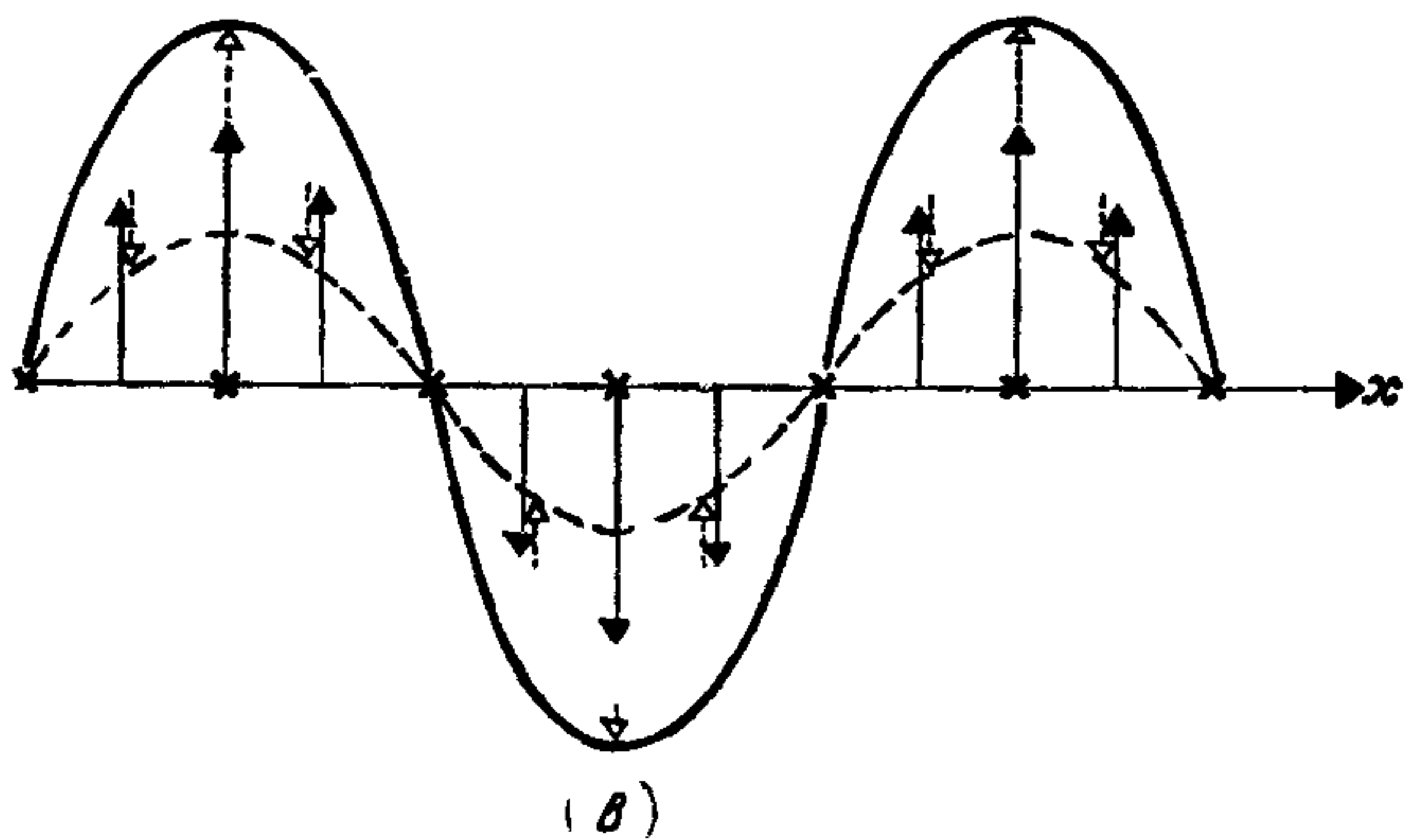
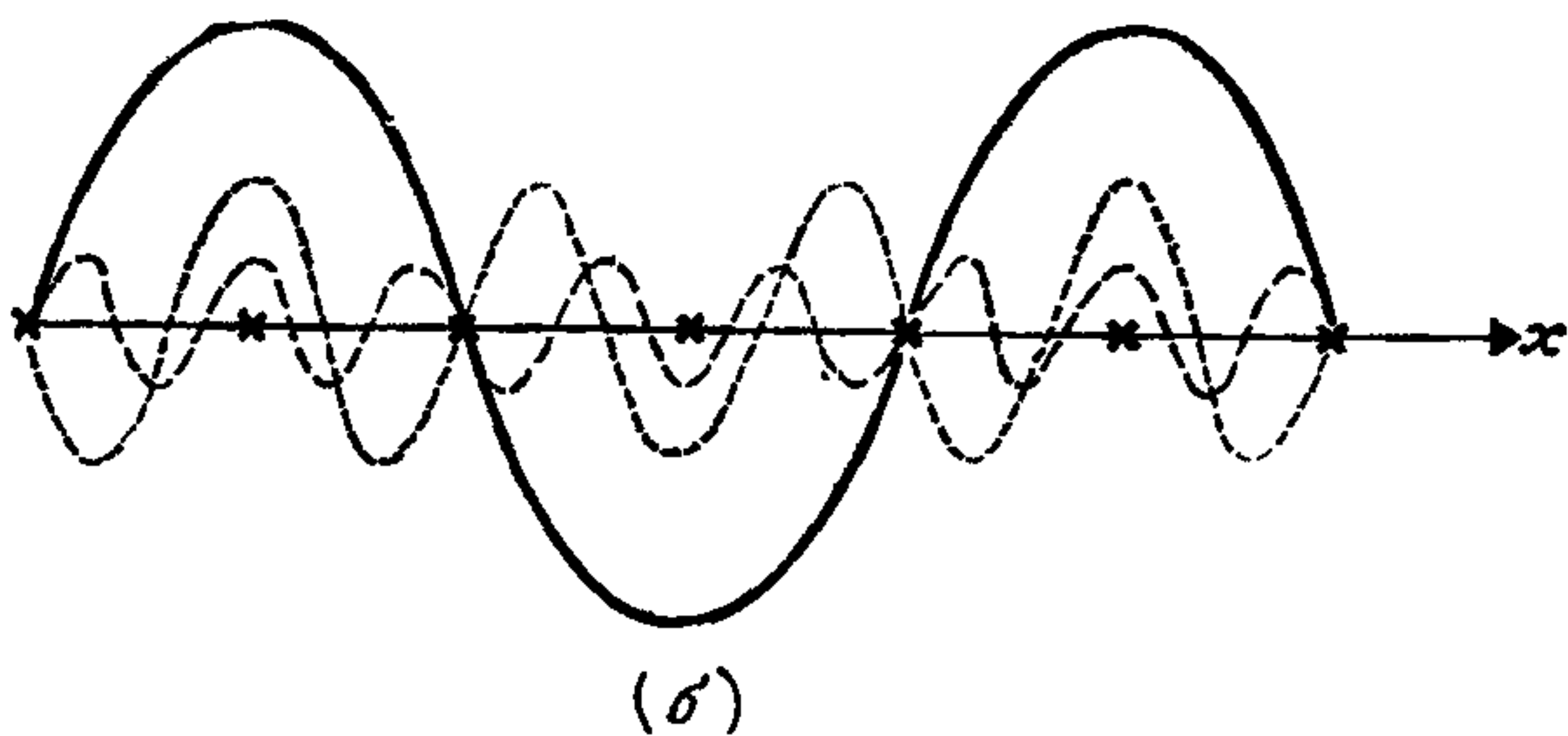
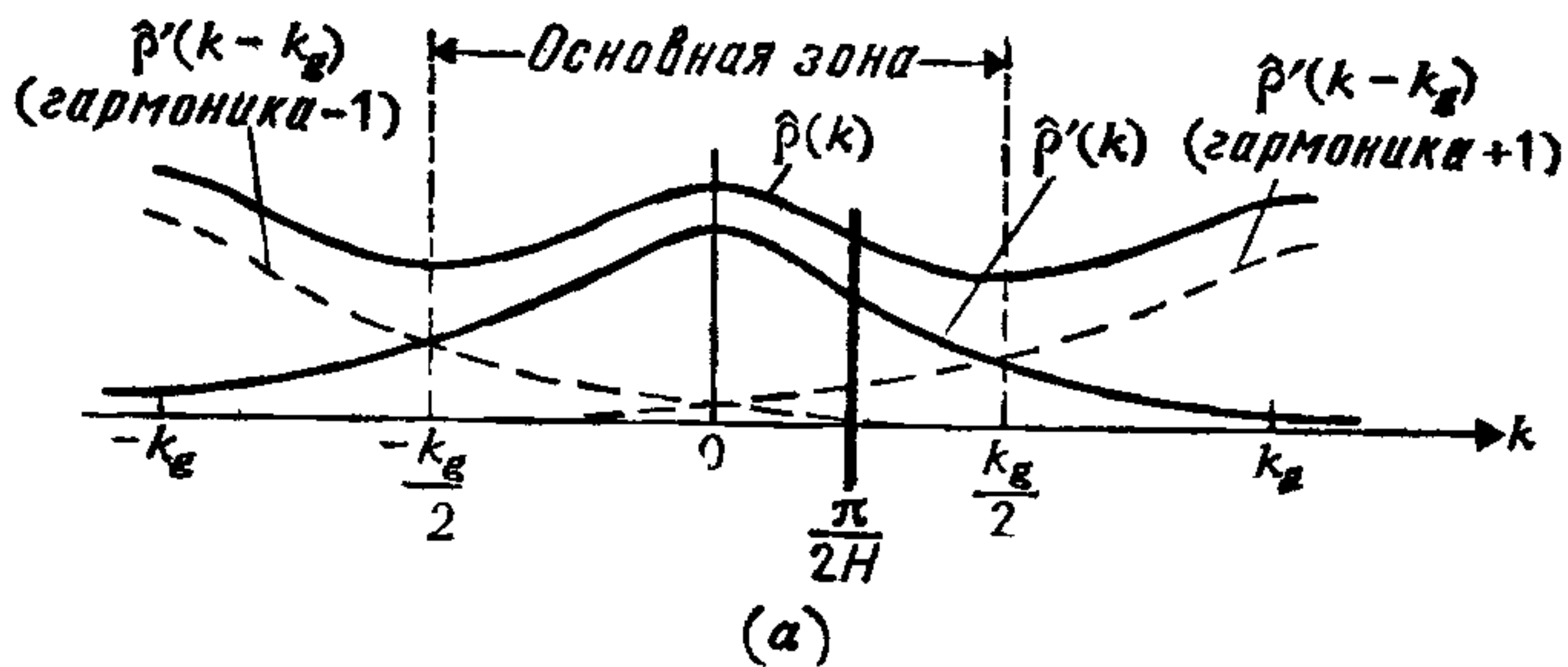


Рис 57 (а) Фурье-образ $\hat{\rho}$ дискретного набора значений плотности заряда на сетке с шагом H равен сумме вкладов от копий образа $\hat{\rho}'$ непрерывного распределения плотности, сдвинутых на интервал $k_g = 2\pi/H$ (б) При определенной длине волны (показанной как $k = \pi/2H$) в гармонику $\hat{\rho}$ дают вклад основная величина $\hat{\rho}'$ и связанные с ней сеточные гармоники (в) В зависимости от соотношения фаз основной и сеточных гармоник амплитуда гармоники может восприниматься сеткой как сильно завышенная или сильно заниженная.

Очевидно, что сетка никак не может разрешить основную гармонику от сеточных, иными словами, основная гармоника и связанные с ней сеточные гармоники дают вклад в гармонику $\hat{\rho}$ с волновым числом, равным волновому числу основной гармоники $\hat{\rho}'$, т. е. при вычислении на сетке те гармоники плотности

заряда, чьи волновые числа лежат вне основной зоны $|k| \leq \pi/H$, маскируются под гармоники с волновыми числами из основной зоны, иначе говоря, накладываются на них. В примере, показанном на рис. 5.7(б), в результате наложения сеточных гармоник амплитуда гармоники $\hat{\rho}(\pi/2H)$ получается слишком большой.

Если бы в результате наложения гармоники всегда получались с сильно завышенной амплитудой, то это было бы не страшно, поскольку можно было ввести в расчет корректирующие множители и компенсировать погрешность. Однако дело обстоит не так. На рис. 5.7(в) сплошной кривой показана гармоника $\hat{\rho}(\pi/2H)$ из примера рис. 5.7(б), когда крестики совпадали с узлами сетки. Если бы вместо этого узлы сетки располагались посередине между крестиками, то вклад сеточных гармоник $+1$ и -1 отличался бы по знаку от основной гармоники. Сложение всех трех гармоник в данном случае даст гармонику, амплитуда которой будет сильно занижена. Это показано штриховыми кривыми на рис. 5.7(в).

Выводы, которые можно сделать на примере одной гармоники, показанной на рис. 5.7, а именно, что распределение заряда ведет к потере информации, носят общий характер. В k -пространстве эта потеря информации проявляется в виде ложных спектральных компонент — нефизическом взаимодействии гармоник, а в x -пространстве — как потеря инвариантности по отношению к смещениям. На рис. 5.7(в) потеря инвариантности относительно смещений выглядит как флуктуация значений сеточной плотности заряда, флуктуация, которая в конечном счете приводит к флуктуациям силы взаимодействия частиц, с которыми мы встречались в первых разделах этой главы. Сеточные эффекты, как их ни описывать — через взаимодействие гармоник или как флуктуации — могут только пагубно повлиять на качество физики модели частиц, и поэтому их надо по возможности уменьшать, не выходя за пределы установленных вычислительных затрат.

Преобразование Фурье на этапе распределения заряда, где фигурируют как непрерывные функции, так и дискретные наборы величин, можно выразить в компактном виде с помощью обобщенной функции Π . Функция выборки Π представляет собой бесконечную гребенку δ -функций, расположенных с единичным интервалом (см. разд. П.4.5). Если непрерывную функцию f умножить на Π , то в результате получится функция f^+ , имеющая вид бесконечной гребенки импульсных функций, идущих с единичным интервалом, при этом амплитуда импульса при $x = n$ равна значению f при $x = n$, т. е.

$$f(n) = \int_{n-1/2}^{n+1/2} f^+(x) dx = \int_{n-1/2}^{n+1/2} \Pi(x) f(x) dx. \quad (5.175)$$

Таким образом, вместо того, чтобы описывать сеточную плотность заряда набором значений $\{\rho(x_p); x_p = pN, p — \text{целое}\}$, можно определить обобщенную функцию

$$\rho^+(x) = \Pi\left(\frac{x}{H}\right) \rho'(x), \quad (5.176)$$

которая содержит бесконечную гребенку импульсных функций с интервалом H , амплитуды которых дают значения сеточной плотности заряда. Преобразование Фурье функции $\rho^+(x)$ согласно теореме о свертке имеет вид

$$\hat{\rho}^+(k) = H \Pi\left(\frac{kH}{2\pi}\right) * \hat{\rho}'(k), \quad (5.177)$$

где $\Pi(x/H) \supseteq H \Pi(kH/2\pi)$ в силу (П.29) и теоремы подобия (табл. П.4). Вычисляя в (5.177) свертку в явном виде, получим

$$\hat{\rho}^+(k) = \int_{-\infty}^{\infty} dk' \sum_{n=-\infty}^{\infty} \delta(k - nk_g) \hat{\rho}'(k - k') = \quad (5.178)$$

$$= \sum_{n=-\infty}^{\infty} \hat{\rho}'(k - nk_g) = \hat{\rho}(k). \quad (5.179)$$

Равенство между $\hat{\rho}$ и $\hat{\rho}^+$ можно установить также непосредственно, поскольку по определению

$$\hat{\rho}^+(k) = \int_{-\infty}^{\infty} dx \rho^+(x) e^{-ikx} = \int_{-\infty}^{\infty} dx \Pi\left(\frac{x}{H}\right) \rho'(x) e^{-ikx}, \quad (5.180)$$

что при вычислении дает

$$\hat{\rho}^+(k) = H \sum_p \rho'(x_p) e^{-ikx_p} = H \sum_p \rho(x_p) e^{-ikx_p} = \hat{\rho}(k). \quad (5.181)$$

Потеря информации, обусловленная дискретным представлением непрерывной функции, зависит от гладкости функции, из которой производится выборка, и шага сетки. Определение гладкости, использованное в разд. 5.3.2, можно перевести на язык асимптотического убывания амплитуды гармоник с ростом k : функция f , непрерывная вместе со своими n производными, имеет образ \hat{f} , убывающий асимптотически как $k^{-(n+1)}$. Дополнительное представление о важности гладкости при выборе функции распределения заряда дает другая точка зрения на гладкость в k -пространстве, а именно с позиций энергосодержания спектра функции в области волновых чисел, больших $|k| = \pi/H$.

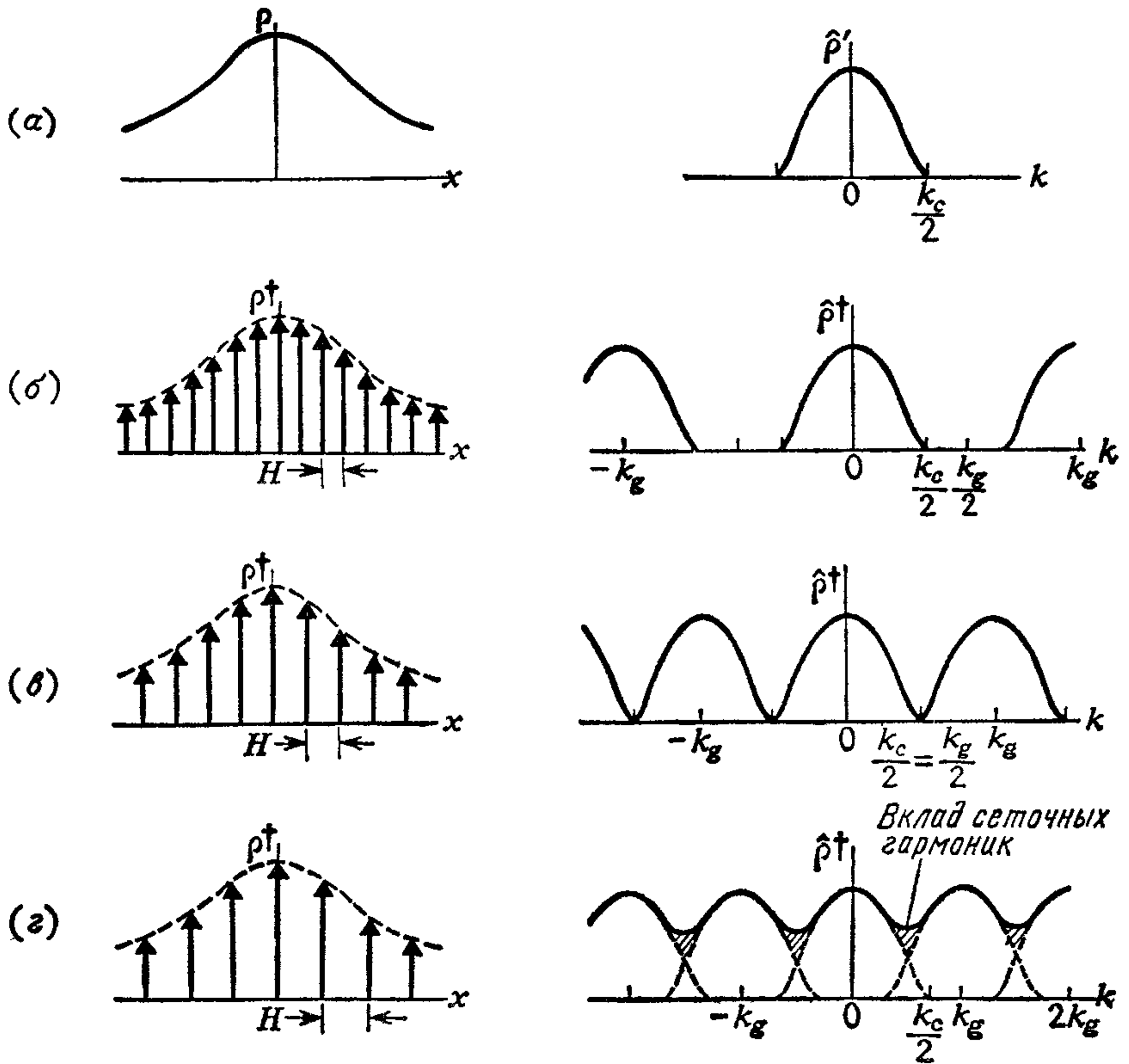


Рис. 5.8. (а) В предельном случае гладкого распределения заряда ρ' его спектр имеет конечную ширину. (б) Если шаг сетки H таков, что $k_g (= 2\pi/H) > k_c$, то ρ' имеет избыточную частоту дискретизации (в) Если H таково, что $k_g = k_c$, то ρ' имеет критическую частоту дискретизации. (г) Если $k_g < k_c$, ρ' имеет недостаточную частоту дискретизации. Перекрытие смещенных копий $\hat{\rho}'$, составляющих $\hat{\rho}'$ вызывает эффект наложения. Вертикальные стрелки на рис. (б) — (г) в x -пространстве, которые отслеживают огибающую, передающую функцию ρ' , представляют собой импульсные функции, амплитуды которых пропорциональны значениям ρ' в точках расположения стрелок.

На рис. 5.8(а) показан предельный случай гладкого распределения плотности $\rho'(x)$. Его спектр ограничен конечным интервалом волновых чисел шириной k_c , т. е. $\hat{\rho}'(k) = 0$ для всех $|k| \geq k_c/2$. На рис. 5.8(б—г) показаны изменения в x - и k -пространствах по мере того, как меняется шаг сетки H . В x -пространстве стрелки указывают амплитуды и координаты импульсов обобщенной плотности заряда (см. (5.176)). Умножению $\rho'(x)$ на $\Pi(x/H)$ в x -пространстве соответствует свертка $\hat{\rho}'(k)$ с $H\Pi(k/k_g)$ в k -пространстве. Свертки с $H\Pi(k/k_g)$ приводят к тому, что в k -пространстве $\rho'(k)$ повторяется с периодом $k_g =$

$= 2\pi/H$. На рис. 5.8(б) показан случай избыточной частоты дискретизации ($k_g > k_c$). Поскольку k_g больше k_c , в k -пространстве нет никакого перекрывания и, следовательно, $\rho'(x)$ можно восстановить по выборочной функции $\rho^+(x)$ умножением $\hat{\rho}^+(k)$ на $\Pi(k/k_g)$ и выполнением обратного преобразования. Точно так же и $\rho'(x)$ можно построить по $\rho^+(x)$ посредством свертки с $1/H \operatorname{sinc}(x/H) [\supset \Pi(k/k_g)]$:

$$\rho'(x) = \frac{1}{H} \operatorname{sinc}\left(\frac{x}{H}\right) * \rho^+(x) = \quad (5.182)$$

$$= \sum_p \operatorname{sinc}\left(\frac{x - x_p}{H}\right) \rho(x_p). \quad (5.183)$$

По мере увеличения сеточного шага H до $H = 2\pi/k_c$ период повторения k_g в k -пространстве уменьшается. На рис. 5.8(в), показан критический период дискретизации $k_g = k_c$, т. е. максимальный шаг сетки, при котором функция $\rho'(x)$ может быть еще восстановлена по ее выборке. За пределами критической точки $\rho'(x)$, как говорят, недостаточно дискретизирована. Из рис. 5.8(г) ясно, что $\rho'(x)$ нельзя однозначно восстановить по недостаточно дискретизированному набору значений, поскольку на интервале $|k| \leq k_g/2$ уже нет взаимно однозначного соответствия между $\hat{\rho}^+$ и $\hat{\rho}'$. Разница между $\hat{\rho}^+$ и $\hat{\rho}'$ в случае недостаточной дискретизации равна вкладу сеточных гармоник.

Если бы удалось построить схему распределения заряда, которая бы действовала как фильтр, пропускающий лишь низкие частоты, удаляя все гармоники \hat{n} для $|k| \geq \pi/H$, то $\rho'(x)$ была бы представлена на сетке с критическим периодом дискретизации и сеточные гармоники были бы полностью подавлены. Такие схемы, действительно, существуют, например, функция $W(x) = (1/H) \operatorname{sinc}(x/H) \supset \hat{W}(k) = \Pi(k/k_g)$ действует как идеальный низкочастотный фильтр: она дает гармоники $\hat{\rho}$ распределенной на сетке плотности, в точности совпадающие с $q\hat{n}$ для $|k| < \pi/H$, и, следовательно, значения $\{\rho(x_p)\}$ совпадают с выборочными значениями функции с конечной шириной спектра, являющейся наилучшим среднеквадратичным приближением к точной плотности заряда $qn(x)$. На самом деле любые функции $W(x) \supset \hat{W}(k)$, спектр которых ограничен интервалом $|k| < k_g/2$, будут полностью подавлять ложные частоты, а связанное с ними расплывание формы частицы можно компенсировать (при условии, что $\hat{W}(k) \neq 0$ для $|k| < k_g/2$) для гармоник основной зоны, для чего к соответствующим сеточным значениям плотности необходимо применить оператор обострения $L \supset [\hat{W}(k)]^{-1}$. К сожалению, функции распределения заряда, спектр которых имеет конечную ширину, не могут одновременно иметь конечную ширину в x -пространстве (т. е. быть отличными от нуля только для $|x| \leq x_c/2$), и поэтому они должны предусматривать

распределение заряда во все сеточные узлы. Следовательно, полное исключение сеточных гармоник в вычислительном отношении невозможно.

Среди функций распределения заряда, которые используют только небольшое число сеточных узлов, функции, удовлетворяющие как условиям гладкости, так и условиям на больших расстояниях (см. табл. 5.1), хорошо аппроксимируют и фильтры низких частот в k -пространстве. Из них схема самого низкого порядка — схема NGP. Исходя из того, что распределение заряда — это свертка с последующей дискретизацией, можно интерпретировать схему NGP как вычисление сначала скользящего среднего (по интервалу H) от точной плотности заряда $qn(x)$ с получением сглаженной плотности $\rho'(x)$:

$$\rho'(x) = \frac{1}{H} \int_{x-H/2}^{x+H/2} qn(x') dx', \quad (5.184)$$

а затем дискретную выборку $\rho'(x)$ в точках $\{x = x_p\}$. Аналогично, в схеме первого порядка — CIC — берется скользящее среднее от скользящего среднего

$$\rho'(x) = \frac{1}{H} \int_{x-H/2}^{x+H/2} dx' \frac{1}{H} \int_{x'-H/2}^{x'+H/2} qn(x'') dx'', \quad (5.185)$$

а затем выбирают значения этой сглаженной плотности и получают сеточную плотность заряда. В каждой последующей схеме более высокого порядка из иерархии, приведенной в табл. 5.1, перед дискретизацией берется еще одно скользящее среднее.

Фурье-образы семейства весовых функций NGP, CIC, TSC легче всего найти, применяя к соотношениям (5.120) — (5.122) теорему о свертке. Из (П.18) и теоремы подобия (табл. П.4) имеем

$$\Pi\left(\frac{x}{H}\right) \supset H \operatorname{sinc}\left(\frac{k}{k_g}\right) = H \frac{\sin\left(\frac{kH}{2}\right)}{\left(\frac{kH}{2}\right)}. \quad (5.186)$$

Отсюда, используя теорему о свертке, находим

$$\text{NGP: } \hat{W}(k) = H \operatorname{sinc}\left(\frac{k}{k_g}\right), \quad (5.187)$$

$$\text{CIC: } \hat{W}(k) = H \operatorname{sinc}^2\left(\frac{k}{k_g}\right), \quad (5.188)$$

$$\text{TSC: } \hat{W}(k) = H \operatorname{sinc}^3\left(\frac{k}{k_g}\right) \quad (5.189)$$

и т. д. Образ весовой функции в схеме m -го порядка равен $H \operatorname{sinc}^{m+1}(k/k_g)$.

Мерой эффективности функции распределения заряда W в смысле отфильтровывания n -й сеточной гармоники волнового числа k (длины волны $\lambda = 2\pi/k$) из основной зоны ($|k| < k_g/2$, $\lambda > 2H$), служит относительная величина этой сеточной гармоники и основной гармоники. Для схемы m -го порядка в иерархии NGP-CIC-TSC и т. д. это отношение равно

$$\left| \frac{\widehat{W}(k - nk_g)}{\widehat{W}(k)} \right| = \left(1 - n \frac{k_g}{k} \right)^{-(m+1)} = \left(1 - \frac{n\lambda}{H} \right)^{-(m+1)}, \quad (5.190)$$

т. е. ослабление увеличивается с номером сеточной гармоники, длиной волны и порядком схемы.

5.6.2. Нахождение потенциала

В случае одномерной бесконечной системы уравнение (5.159) принимает вид

$$\varphi(x_p) = H \sum_{p'=-\infty}^{\infty} G(x_p - x_{p'}) \rho(x_{p'}). \quad (5.191)$$

Снова обращаясь к табл. П.3, мы видим, что это выражение представляет собой свертку для системы, которая дискретна в x -пространстве и периодична в k -пространстве. Следовательно,

$$\varphi(x_p) \supset \hat{\varphi}(k) = \hat{G}(k) \hat{\rho}(k), \quad (5.192)$$

где периодичность означает, что для целых n

$$\hat{\varphi}(k + nk_g) = \hat{\varphi}(k). \quad (5.193)$$

Свертку двух последовательностей, в данном случае (5.191), можно представить в виде интеграла типа свертки, построив непрерывные функции φ' и G' для потенциала и функции Грина. Ясно, что существует бесконечное число всевозможных непрерывных функций φ' и G' , которые можно было бы выбрать, и в то же время есть только один вариант, который не вносит паразитной информации. Этим частным вариантом является как раз тот, где наборы величин $\{\varphi(x_p)\}$ и $\{G(x_p)\}$ получаются посредством дискретизации φ' и G' с критическим шагом, т. е.

$$G'(x) \supset \hat{G}'(k) = \Pi\left(\frac{k}{k_g}\right) \hat{G}(k) \Rightarrow \quad (5.194)$$

$$\Rightarrow G'(x) = \sum \operatorname{sinc}\left(\frac{x - x_p}{H}\right) G(x_p), \quad (5.195)$$

и аналогично для φ' . Переписывая уравнение (5.191) через φ' и G' и используя соотношение (5.176), чтобы представить плот-

ность заряда как функцию непрерывного аргумента x' , получим'

$$\varphi'(x) = \int_{-\infty}^{\infty} dx' G'(x - x') \rho^+(x'), \quad (5.196)$$

что при переходе к фурье-образу дает

$$\varphi'(x) \supset \hat{\varphi}'(k) = \hat{G}'(k) \hat{\rho}^+(k) = \Pi\left(\frac{k}{k_g}\right) \hat{\varphi}(k). \quad (5.197)$$

Выражение (5.196) для потенциала имеет в точности тот же вид, что и интегральная форма записи уравнения Пуассона через функцию Грина для физической системы. В непрерывной системе функция Грина g удовлетворяла уравнению

$$\nabla^2 g = -\frac{\delta(x)}{\epsilon_0}, \quad (5.198)$$

т. е.

$$\hat{g}(k) = \frac{1}{\epsilon_0 k^2}, \quad (5.199)$$

где в двумерном и трехмерном случаях $k^2 = k \cdot k$. Если в качестве $\varphi'(x)$ взять наилучшее (в смысле метода наименьших квадратов) приближение к точному потенциалу

$$\Phi(x) = \int_{-\infty}^{\infty} dx' g(x - x') \rho^+(x'), \quad (5.200)$$

то мы найдем, что

$$\hat{G}'(k) = \Pi\left(\frac{k}{k_g}\right) \hat{g}(k), \quad (5.201)$$

что соответствует функции влияния $\hat{G}(k) = (1/\epsilon_0 k^2)$, применяемой в «дешевом методе» решения уравнения Пуассона [42]. Более того, если приближенную функцию влияния положить равной нулю для волновых чисел, превышающих некоторое предельное значение $\alpha k_g/2$; $0 < \alpha < 1$, то, задавая ненулевые элементы \hat{G} равными \hat{g} , т. е.

$$\hat{G}'(k) = \Pi\left(\frac{k}{k_g}\right) \Pi\left(\frac{k}{\alpha k_g}\right) \hat{g}(k), \quad (5.202)$$

получим, что $\varphi'(x)$ по-прежнему является наилучшей средне-квадратичной аппроксимацией функции $\Phi(x)$. В качестве альтернативного критерия, который нашел применение, например, в моделировании изолированных гравитирующих систем (см. гл. 11), можно потребовать, чтобы значения потенциала в узлах сетки были точными. Это предопределяет вид функции влияния:

$$\hat{G}'(k) = \Pi\left(\frac{k}{k_g}\right) \sum_{n=-\infty}^{\infty} \hat{g}(k - nk_g). \quad (5.203)$$

При использовании конечно-разностных аппроксимаций лапласиана в уравнении Пуассона функции влияния получают посредством непосредственного преобразования разностных уравнений. Например, в одномерном случае трехточечная аппроксимация (уравнение (5.8)) дает

$$\hat{G}'(k) = \Pi\left(\frac{k}{k_g}\right) \frac{(H/2)^2}{\sin^2\left(\frac{kH}{2}\right)}, \quad (5.204)$$

а в соответствующем двумерном случае (уравнение (5.129))

$$\hat{G}'(k_1, k_2) = \Pi\left(\frac{k_1}{k_g}\right) \Pi\left(\frac{k_2}{k_g}\right) \frac{(H/2)^2}{\sin^2\left(\frac{k_1 H}{2}\right) + \sin^2\left(\frac{k_2 H}{2}\right)}. \quad (5.205)$$

Сравнивая выражения (5.204) и (5.205), мы видим, что для конечно-разностной аппроксимации коротковолновые (т. е. при $|k|$, близких к $k_g/2$) гармоники потенциала получаются намного больше, чем для наилучшей среднеквадратичной аппроксимации, т. е. погрешности дискретизации и погрешности распределения заряда действуют в противоположном направлении: погрешности, связанные с распределением заряда, ослабляют коротковолновые компоненты, тогда как погрешности дискретизации их усиливают. Следовательно, если к распределенному на сетку заряду применяется оператор обострения (см. разд. 5.3.2) и при этом не учитывается влияние погрешностей аппроксимации в уравнении для поля на восстановление исходной формы частицы, то в результате можно прийти к тому, что аппроксимация получится хуже, чем если бы не делалось никаких попыток коррекции!

5.6.3. Интерполяция силы

Как с очевидностью следует из проведенного обсуждения этапов распределения заряда и нахождения потенциала, выражения (5.160) и (5.161) для вычисления силы в схемах, сохраняющих импульс, и соответствующее выражение (5.162) для схем, сохраняющих энергию, также можно рассматривать как свертки.

Для схем, сохраняющих импульс, первым шагом при вычислении силы является построение конечно-разностной схемы, связывающей значения потенциала и электрического поля в узлах сетки. Простейшую одномерную схему (5.11) можно записать в виде

$$E(x_p) = -H \sum_{p'} \frac{1}{H} \frac{(\delta_{p+1-p'} - \delta_{p-1-p'})}{2H} \Phi(x_{p'}), \quad (5.206)$$

где δ — символ Кронекера (см. (П.27)). Обращаясь к теореме о свертке для дискретных величин (FS (ii) в табл. П.3), можно из (5.206) найти фурье-образ E . Для того чтобы получить интегральную форму уравнения (5.206), по значениям $\{E(x_p)\}$ строится непрерывная функция $E(x)$, имеющая конечную ширину спектра, точно так же, как это делалось для G' и φ' (см. (5.194) и (5.196)), и правая часть записывается через φ' :

$$E'(x) = - \int dx' \left[\frac{\delta(x+H-x') - \delta(x-H-x')}{2H} \right] \varphi'(x'). \quad (5.207)$$

Член в скобках в (5.207) является двухточечным разностным оператором, аппроксимирующим ∇ . Для четырехточечного оператора разностный оператор D принимает вид

$$D(x) = \alpha \frac{\delta(x+H) - \delta(x-H)}{2H} + (1-\alpha) \frac{\delta(x+2H) - \delta(x-2H)}{4H} \quad (5.208)$$

(ср. (5.137)). В общем случае для одномерных схем мы имеем

$$E'(x) = - \int dx' D(x-x') \varphi'(x'), \quad (5.209)$$

что дает

$$\hat{E}'(k) = - \hat{D}(k) \hat{\varphi}'(k). \quad (5.210)$$

Для схем, сохраняющих импульс, интерполяционную формулу для силы (5.161) можно записать в интегральной форме

$$F(x) = \frac{q}{H} \int W(x-x') E^+(x') dx', \quad (5.211)$$

где для представления сеточного поля введена обобщенная функция E^+ :

$$E^+(x) = \Pi\left(\frac{x}{H}\right) E'(x). \quad (5.212)$$

Отсюда

$$\hat{F}(k) = \frac{q}{H} \hat{W}(k) \hat{E}^+(k), \quad (5.213)$$

где

$$\hat{E}^+(k) = H \Pi\left(\frac{k}{k_g}\right) * \hat{E}'(k) = \quad (5.214)$$

$$= \sum_{n=-\infty}^{\infty} \hat{E}'(k - nk_g) = \quad (5.215)$$

$$= - \hat{D}(k) \hat{\varphi}(k). \quad (5.216)$$

Сумма по сеточным гармоникам (5.125) нужна только для того, чтобы периодически повторить E' вне интервала $|k| < k_g/2$, поскольку $E'(x)$, в силу способа ее построения, является функцией

с конечной шириной спектра. Это отражено в (5.216), где сумма по сеточным гармоникам заменяется периодической величиной $\hat{\phi}(k) \supset \{\phi(x_p)\}$, т. е. $\phi^+(x)$.

В многомерном случае выражение (5.216) превращается в

$$\hat{E}^+(k) = -\hat{D}(k)\hat{\phi}(k). \quad (5.217)$$

Точный градиент потенциала равен

$$\hat{E}(k) = -ik\phi(k). \quad (5.218)$$

Из сравнения (5.217) и (5.218) ясно, как влияют погрешности аппроксимации на гармоники поля \hat{E}^+ : погрешности в абсолютной величине вызывают отличие $|\hat{D}|$ от $|k|$, а погрешности в направлении приводят к тому, что $\hat{D} \times k$ не равно нулю.

Определяя обобщенную функцию потенциала

$$\phi^+(x) = \Pi\left(\frac{x}{H}\right)\phi'(x), \quad (5.219)$$

можно записать выражение для силы в схемах, сохраняющих энергию, а именно (5.162), в виде интеграла типа свертки, что дает (в одномерном случае)

$$F(x) = -\frac{q}{H} \int \frac{d}{dx} W(x-x')\phi^+(x') dx'. \quad (5.220)$$

Применяя теорему о свертке (табл. П.3) и теорему о производной (табл. П.4), сразу находим фурье-образ:

$$\hat{F}(k) = -\frac{q}{H} ik\hat{W}(k)\hat{\phi}(k), \quad (5.221)$$

а для многомерного случая

$$\hat{F}(k) = -\frac{q}{V_c} ik\hat{W}(k)\hat{\phi}(k). \quad (5.222)$$

Таким образом, хотя во всех схемах, сохраняющих энергию, импульс и не сохраняется (см. разд. 5.5), применяемые в них схемы интерполяции силы действительно не порождают угловых погрешностей силы, которые при прочих равных условиях в схемах, сохраняющих импульс, приносятся конечно-разностным дифференцированием потенциала.

5.6.4. Сила парного взаимодействия частиц

Фурье-образ силы, действующей на заряд q в точке x_2 со стороны заряда $-q$ в точке x_1 , получается сведением воедино результатов разд. 5.6.1—5.6.3,

Сила, действующая на заряд в точке x_2 , равна

$$F(x_2) = \int \frac{dk}{2\pi} \hat{F}(k) e^{ikx_2}. \quad (5.223)$$

Исходя из (5.213), выполняем преобразование силы для схем, сохраняющих импульс:

$$\begin{aligned} \hat{F}(k) &= \frac{q}{H} \hat{W} \hat{E}^+ = \\ &= -\frac{q}{H} \hat{W} \hat{D} \hat{\phi} \quad (\text{в силу (5.216)}) \\ &= -\frac{q}{H} \hat{W} \hat{D} \hat{G} \hat{\rho} \quad (\text{в силу (5.192)}) \\ &= \frac{q^2}{H^2} \hat{W} \hat{D} \hat{G} \sum_n \hat{W}(k - nk_g) e^{-i(k - nk_g)x_1}, \quad (5.224) \end{aligned}$$

где в последней строке (5.224) для исключения $\hat{\rho}$ были использованы соотношения (5.179) и (5.166) с $n(x) = \delta(x - x_1) \Rightarrow \hat{n}(k) = e^{-ikx_1}$. Повторяя такую же процедуру исключения для схем, сохраняющих энергию (формула (5.221)), получаем

$$\hat{F}(k) = \frac{q^2}{H^2} \hat{W} ik \hat{G} \sum_n \hat{W}(k - nk_g) e^{-i(k - nk_g)x_1}. \quad (5.225)$$

Объединяя (5.223) и (5.224), находим выражение для силы взаимодействия частиц, отвечающее методу частица — сетка:

$$F(x; x_1) = \int \frac{dk}{2\pi} \left(\frac{q^2}{H^2} \hat{W} \hat{D} \hat{G} \sum_n \hat{W}(k - nk_g) e^{-ink_g x_1} \right) e^{ikx}, \quad (5.226)$$

где $x = x_2 - x_1$ — расстояние между частицами. Наличие в (5.226) суммы по сеточным гармоникам, возникающей от распределения заряда, является единственной причиной потери инвариантности относительно смещений. Члены с $n \neq 0$ в этой сумме вносят в k -пространстве периодичность с длиной периода k_g . То же самое справедливо и для силы взаимодействия частиц, получаемой подстановкой (5.225) в (5.223).

Усреднение F в (5.226) по всем положениям исходного заряда дает

$$\langle F(x; x_1) \rangle = \frac{1}{H} \int_H F(x; x_1) dx_1 = \int \frac{dk}{2\pi} \frac{q^2}{H} (\hat{W} \hat{D} \hat{G} \hat{W}) e^{ikx}. \quad (5.227)$$

Члены в скобках в правой части (5.227) отвечают описанию в k -пространстве соответственно интерполяции силы, конечно-раз-

ностного дифференцирования потенциала, нахождения самого потенциала и усредняющего влияния распределения заряда. Каждое действие определенным образом влияет на форму гармоник силы взаимодействия частиц, при этом их совокупное влияние можно оценить с помощью выражений, выводимых из (5.223) — (5.227). Более того, эти выражения являются точным описанием, поэтому выводы, делаемые на их основе, не связаны малостью параметров, как чаще всего бывает при рассмотрении в x -пространстве.

Вопросы обобщения полученных результатов на бóльшие размерности и системы с границами мы откладываем до глав 7 и 8, где покажем, каким образом используются спектральные методы при объяснении свойств моделей частиц и оптимизации этих свойств с точки зрения экономичности вычислений.

Решение уравнений поля

6.1. ВВЕДЕНИЕ

При практической реализации описанных алгоритмов моделирования методом частиц необходим эффективный метод решения уравнений поля. Если мы рассмотрим типичный вычислительный эксперимент, включающий 1000 шагов по времени, причем доступное машинное время равно 1 ч, то станет ясно, что время решения уравнений поля (которое при хорошо сбалансированном алгоритме вычислений будет составлять около половины времени исполнения цикла) обычно не должно превышать нескольких секунд. Для машины со средним временем исполнения арифметической операции, равным 1 мкс, и для пространственной сетки, содержащей приблизительно $N_g = 10\,000$ узлов (скажем, 20^3 или 100^2)¹⁾, нам потребуются методы, позволяющие получить решение менее чем за 100 арифметических операций (или 100 мкс) на один узел. На самом деле лучшие имеющиеся спектральные методы (там, где они применимы) могут затрачивать на решение около 20 операций на один узел.

⁸ Важно осознать, что такое малое число операций является чрезвычайно строгим требованием, вынуждающим использовать специальные методы. Действительно, в ряде случаев необходимо переделать математическую модель на такую, к которой можно применить быстрые специальные методы для того, чтобы довести ее исследование до конца. Для сравнения отметим, что прямое решение уравнений для 10 000 сеточных узлов методом исключения Гаусса без использования свойств разреженной матрицы ($\simeq N_g^3$ операций) потребовало бы $\sim 10^{12}$ операций, или около одного месяца машинного времени для непрерывного счета одного шага по времени!

Какой из имеющихся методов решения уравнений в частных производных, описывающих поля, следует применять, очень сильно зависит от конкретного вида уравнений, и нет единого метода, универсального и наилучшего для всех случаев. Выбор метода определяется следующими характеристиками задачи:

¹⁾ N_g равно полному числу сеточных узлов, а n в этой главе равно числу сеточных узлов по каждому направлению. Следовательно, для однородной сетки с размерностью d получаем $N_g = n^d$.

видом уравнений, т. е.

- a)* линейностью,
- b)* размерностью,
- c)* разделяемостью переменных,
- d)* наличием переменных или постоянных коэффициентов,

а при наличии границы свойствам области и краевых условий, а именно:

- a)* прямоугольная или произвольная область,
- b)* смешанные или простые граничные условия,
- c)* односвязная или неодносвязная область.

Имеющиеся методы можно разбить на следующие большие классы.

1. Методы сеточной релаксации. В этих методах первоначально взятые наугад сеточные значения релаксируются к решению с помощью систематического прохода по сетке и уточнения значений. Например:

- a)* метод Гаусса — Зайделя (GS),
- b)* метод последовательной верхней релаксации (SOR),
- c)* метод последовательной линейной верхней релаксации (SLOR),
- d)* метод переменных направлений (ADI).

2. Матричные (прямые) методы. В этих методах конечно-разностные уравнения рассматриваются как линейные матричные уравнения и решаются с помощью различных вариантов прямых и итерационных матричных методов. Например:

- a)* алгоритм Томаса¹⁾ для трехдиагональной системы линейных алгебраических уравнений,
- b)* методы исключения для разреженной матрицы (SM),
- c)* алгоритм сопряженных градиентов (CGA),
- d)* строго неявная процедура Стоуна (SIP),
- e)* метод неполного разложения Холецкого — сопряженных градиентов (ICCG).

3. Методы быстрого решения эллиптических уравнений. В них специальные методики, основанные на быстром преобразовании Фурье (FFT) и циклической редукции (CR), используются для решения некоторых классов задач не более чем за $O(N_g \log_2 N_g)$ операций. Например:

- a)* циклическая редукция, или алгоритм Бунемана (CR),
- b)* кратное преобразование Фурье (MFT),
- c)* анализ Фурье/циклическая редукция (FACR),

¹⁾ В отечественной литературе известен как метод прогонки. — Прим. перев.

- d) методы свертки,
- e) итерационный метод Конкуса и Голуба.

Далее мы обсудим более подробно, как влияют различные аспекты постановки проблемы на выбор метода решения.

6.1.1. Выбор метода

В выборе наиболее подходящего метода для решения уравнений поля помогут рис. 6.1 и табл. 6.1. На рисунке мы даем дерево решений для облегчения выбора метода, и в таблице указаны требуемые количество операций и память, приходящаяся на один сеточный узел.

Первым возникает вопрос о линейности рассматриваемого уравнения. Если уравнение нелинейно, то обычно оно решается с помощью локальной линеаризации, сопровождаемой решением последовательности изменяющихся линейных задач. Этот метод последовательного приближения представляет обобщение итераций по Ньютону для нахождения корней скалярных уравнений. Он обсуждается в разд. 6.2.1. Коэффициенты линейных уравнений, появляющиеся в итерациях по Ньютону, зависят от значения текущего решения, и поэтому нельзя предполагать, что они имеют какую-либо особую симметрию или значения. Мы называем такую задачу задачей с произвольными коэффициентами и, за исключением одномерного случая, хороший общепринятый метод для ее решения отсутствует.

В одномерном случае задача сводится к решению трехдиагональной системы уравнений, и ее можно эффективно решить для произвольных коэффициентов, используя вариант метода исключения Гаусса, часто называемый алгоритмом Томаса (см. разд. 6.4.1).

В двумерном и трехмерном случаях целесообразно убедиться в том, можно ли представить уравнения в данном ниже виде с разделяющимися переменными:

$$a(x) \frac{\partial^2 \varphi}{\partial x^2} + b(x) \frac{\partial \varphi}{\partial x} + c(x) \varphi + \text{аналогичные члены}$$

$$\text{по } y \text{ и } z = g(x, y, z), \quad (6.1)$$

где оператор в левой части представляет сумму трех членов, каждый из которых зависит только от одной координаты. Эта форма включает оператор ∇^2 , выраженный в любой из распространенных координатных систем (цилиндрической, сферической, полярной и т. д.). Если уравнение допускает разделение переменных, то можно использовать специальные быстрые методы решения эллиптических уравнений (RES); они описываются в разд. 6.5.1—6.5.7,

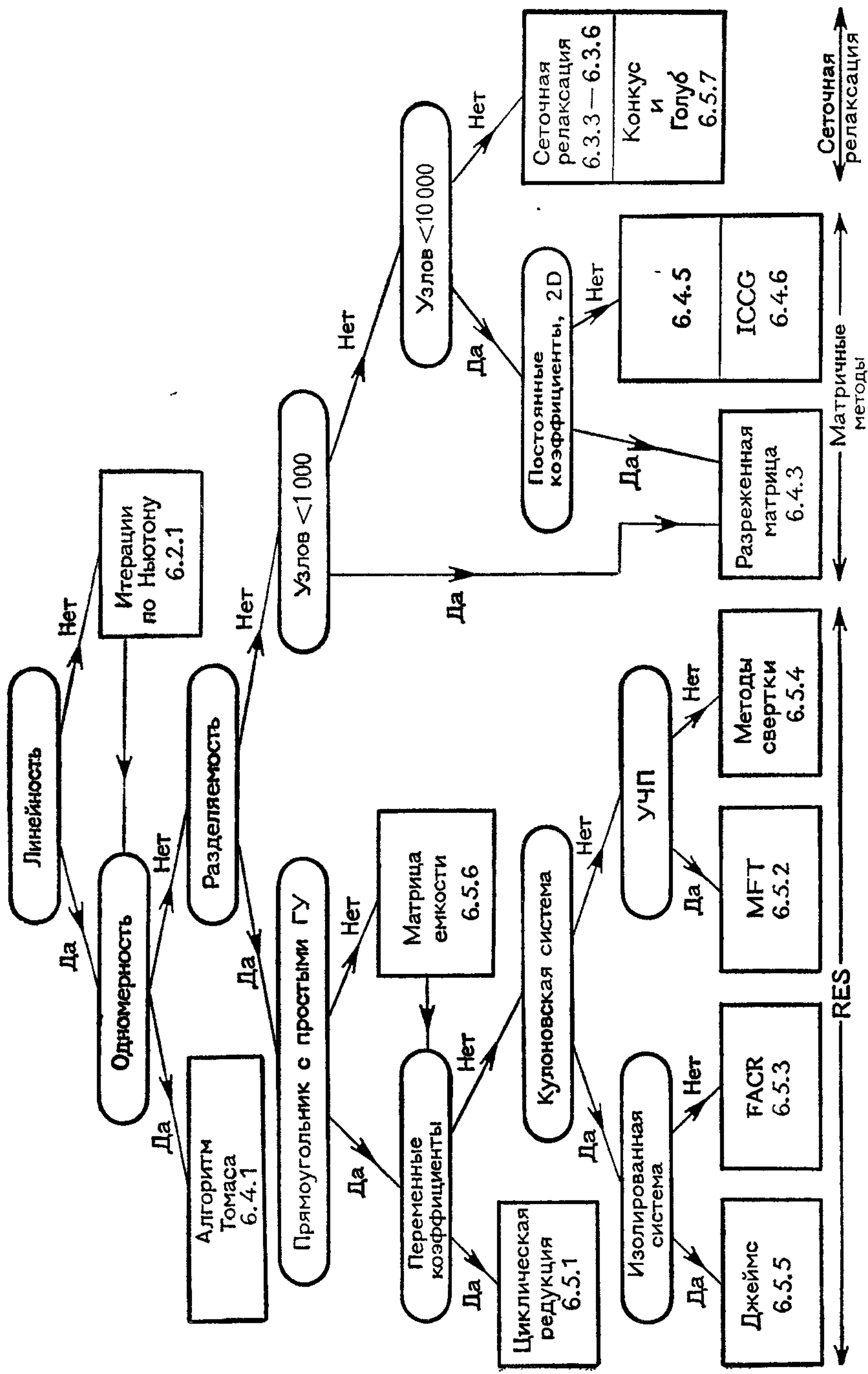


Рис. 6.1. Дерево решений, принимаемых при выборе методов решения уравнений поля. Нумерация означает разделы текста, в которых описан соответствующий метод.

Таблица 6.1. Число операций и объем памяти, требуемые для различных методов

Вид коэффициентов	Метод	Операций на узел		Память на узел		Операций на узел [§]		
		разложение	решение	основная	дополнительная	разложение	решение	
Методы RES:								
Специальные	FACR (*)	$4.5 \log_2 (\log_2 N_g)$		1	0	17		
	FACR (1)	$1.25 [\log_2 N_g + 5]$		1	0	25		
	MFT общий (3D)	$10 \log_2 N_g$		1	1	140		
	CR (Бунеман)	$3 \log_2 N_g$		1	0	42		
	Алгоритм Джеймса	$5 \log_2 N_g$		1	0	70		
	CR общий с разделением переменных	$16 \log_2 N_g + 88$		1	$\frac{\log_2 N_g}{N_g^{1/2}}$	317	136	
Матричные методы:								
Произвольные	Алгоритм Томаса (1D)	3	5	4	1	8	5	
	Метод для разреженных матриц	2D	$10N_g^{1/2}$	$12 \log_2 N_g$	6	$3 \log_2 N_g$	1280	168
		3D	$4N_g$	$112N_g^{1/3}$			6×10^4	10^4
	ICCG [†]	$27 \times 1.3N^{1/d}$		6	3	270 [#]		
	Процедура Стоуна [†] SIP	27×10		6	7	270 [#]		
Релаксационные методы:								
	Процедура Конкуса и Голуба [†]	$4.5 \log_2 (\log_2 N_g) \times 10$		6	2	170 [#]		
	SOR [†]	$11 \times 2N_g^{1/d}$		6	1	2816		
	SLOR [†]	$11 \times \sqrt{2} N_g^{1/d}$		6	1	2000		
	ADI [†]	$18 \times 12 \frac{\log_2 N_g}{d}$		6	1	1500		

[#] Надежных данных о различиях в сходимости этих методов не существует. Мы принимаем типичное значение 10 итераций для всех методов.

[§] На двумерной сетке 128×128 .

[†] Форма записи: (число операций на одну итерацию) \times (число итераций, необходимое для снижения погрешности в 10^6 раз); d — размерность.

Если разделение переменных неосуществимо, то можно использовать или матричные методы, или методы сеточной релаксации. Выбор легко определяется количеством используемых сеточных узлов. Для малых систем с $N_g < 1000$ задачу с произвольными коэффициентами можно решить непосредственно с помощью методов исключения для разреженных матриц (разд. 6.4.3). Поскольку решение при этом заведомо получается за известное время, то рекомендуется этот метод. Если система большая, $N_g > 1000$, то время, затраченное на разложение матрицы с помощью методов исключения для разреженных матриц, вероятно, становится чрезмерно большим для многократного использования. Однако если коэффициенты уравнений остаются неизменными для большого количества правых частей уравнения, то требуется только одно разложение, и время, затраченное на решение с каждой правой частью с помощью методов SM, будет приемлемо вплоть до нескольких тысяч уравнений. Количество операций на один сеточный узел на этапе разложения и этапе решения даны по-отдельности в табл. 6.1.

Если произвольные коэффициенты уравнения меняются для каждой правой части или количество сеточных узлов превосходит 1000, то нужно использовать итерационные методы решения. Если число узлов меньше 10 000, то пригодны строго неявная процедура Стоуна (SIP) (см. разд. 6.4.5) или метод неполного разложения Холецкого — сопряженных градиентов (см. разд. 6.4.6). В последних итерационных методах требуется от трех до семи дополнительных сеток для промежуточных векторов, и для сеток, больших чем в 10 000 узлов, эта дополнительная память, вероятно, заставляет от таких методов отказаться. Методы сеточной релаксации (см. разд. 6.3.3—6.3.6), такие как SLOR или ADI, требуют только одного дополнительного массива и могут быть использованы для таких больших задач. В других случаях рекомендуется повторное использование алгоритма RES, которое предложено Конкусом и Голубом (см. разд. 6.5.7) и, вероятно, должно дать более быструю сходимость. Однако этот метод требует двух дополнительных сеток.

Теперь вернемся к выбору методов, когда уравнения допускают разделение переменных. Можно непосредственно использовать метод RES для таких задач, если расчетная область является квадратом или прямоугольником, а граничные условия простые (например, задано значение или градиент искомого решения). В противном случае можно использовать метод матрицы емкости (см. 6.5.6) для удерживания требуемых граничных условий в области произвольной формы, вложенной в прямоугольник. Это требует вычисления матрицы емкости, связанного с большими затратами времени, а затем двух применений метода RES для получения решения. Если геометрия неизменна, то матрицу емкости нужно рассчитывать только один раз, и

необходимое время решения становится сравнимым с временем решения методом RES.

Если коэффициенты задачи, допускающей разделение переменных, переменные, то можно использовать циклическую редукцию и решить уравнения за $3 \log_2 N_g$ операций на один узел (см. разд. 6.5.1). Если коэффициенты постоянные, то пригодны более быстрые методы. При подходящих граничных условиях обычное дифференциальное уравнение в частных производных с постоянными коэффициентами можно решить с помощью метода преобразования Фурье (MFT) (см. разд. 6.5.2). Если потенциал выражается в виде суммы вкладов от отдельных источников, то пригодны методы свертки, использующие MFT. В них рассматриваются взаимодействие источников с функцией Грина произвольной формы, либо в периодической геометрии, либо как изолированный ансамбль в пространстве (см. разд. 6.5.4). Если взаимодействие между частицами кулоновское, то применимы более быстрые методы, как для изолированных систем (см. алгоритм Джеймса, разд. 6.5.5), так и для уравнения Пуассона в прямоугольнике (см. оптимальный алгоритм FACR, который имеет наименьшее количество операций $4.5 \log_2(\log_2 N_g)$, разд. 6.5.3).

Теперь мы дадим краткую оценку различных методов решения, начиная с методов решения наиболее общих нелинейных задач и кончая специальными методами решения уравнения Пуассона в прямоугольной области.

6.2. НЕЛИНЕЙНЫЕ ЗАДАЧИ

6.2.1. Итерации по Ньютону

Если уравнение поля нелинейно, то решение обычно получают итерационным применением методов решения соответствующих линейных задач.

Например, пусть

$$N(\varphi) = 0 \quad (6.2)$$

является нелинейным дифференциальным уравнением для функции $\varphi(x, y, z)$ и пусть $\varphi^{(t)}$ является приближенным решением уравнения (6.2) на итерации t . Если $\Delta\varphi = \varphi^{(t+1)} - \varphi^{(t)}$ — требуемое приращение (предполагаемое малым) за итерацию, то (6.2) можно разложить относительно имеющегося приближения и получить

$$N(\varphi^{(t)} + \Delta\varphi) = N(\varphi^{(t)}) + \frac{\partial N}{\partial \varphi} \Big|^{(t)} \Delta\varphi + \text{члены более высокого порядка.} \quad (6.3)$$

Оба выражения (для N и ее производной) в правой части вычисляются с использованием $\varphi^{(t)}$ и являются известными операторами.

Если мы пренебрегаем членами более высокого порядка в (6.3) и требуем, чтобы $\varphi^{(t+1)}$ было решением уравнения (6.2), то получаем итерационную схему первого порядка:

$$\left. \frac{\partial N}{\partial \varphi} \right|^{(t)} \Delta \varphi = -N(\varphi^{(t)}) \quad (6.4a)$$

или

$$\left. \frac{\partial N}{\partial \varphi} \right|^{(t)} \varphi^{(t+1)} = -N(\varphi^{(t)}) + \left. \frac{\partial N}{\partial \varphi} \right|^{(t)} \varphi^{(t)}. \quad (6.4b)$$

В уравнениях (6.4) левые части суть известные линейные операторы, действующие на искомое решение, а правые части — известные функции аргумента $\varphi^{(t)}$, которые найдены на предыдущей итерации. Как показано в примере ниже, N будет обычно зависеть от других заданных функций, которые мы обозначим здесь через $n(x)$.

В уравнении (6.4a) невязка $N(\varphi)$ обуславливает изменение решения $\Delta \varphi$ за время одной итерации. Такой подход наиболее приемлем в компьютерах с малой длиной слова, в которых погрешность округления должна тщательно контролироваться. Не говоря о матрице коэффициентов, для которых необходима память объемом в $4N_g$, требуются еще четыре массива для n , $\varphi^{(t)}$, $N(\varphi)$ и $\Delta \varphi$. Однако новое значение $\varphi^{(t+1)}$ находится сразу из уравнения (6.4b), и требуются только три массива дополнительно к матрице коэффициентов: для n , $\varphi^{(t)}$ и правой части уравнения (6.4b). Полная память, требуемая для решения нелинейной задачи этим методом, равна $7N_g$ или $8N_g$ в дополнение к временной памяти, необходимой для решения линейного уравнения. Эту память можно сопоставить с полной памятью размером N_g , требуемой алгоритмом RES для частного случая задач с постоянными коэффициентами (см. разд. 6.5.1—6.5.7).

Для того чтобы понять, как этот алгоритм работает на практике, рассмотрим нелинейное дифференциальное уравнение

$$N(\varphi) = \nabla^2 \varphi - (e^\varphi - 1)n(x) = 0, \quad (6.5)$$

у которого имеется нелинейная зависимость от решения во втором члене. При нормированном подходящим образом переменных это есть электростатический потенциал электронной плазмы с фиксированным фоном из положительных ионов $n(x)$, который является функцией вектора координат x . Экспоненциальная плотность находится в тепловом равновесии в соответствии с законом Больцмана. Уравнение можно было бы использовать для расчета потенциала в гибридной модели, в которой медленно движущиеся ионы продвигаются с помощью уже описанных методов частиц и предполагается, что легкие подвижные

электроны сами мгновенно подстраиваются в соответствии с законом Больцмана $n_e \sim \exp(e\varphi/k_B T)$. В этом случае ионный фон (представляемый функцией координат $n(\mathbf{x})$ в (6.5)) должен быть известной функцией.

Дифференцируя уравнение (6.5) по φ , получаем

$$\frac{\partial N}{\partial \varphi} = \nabla^2 - e^\varphi n(\mathbf{x}). \quad (6.6)$$

Заметим, что $n(\mathbf{x})$ зависит от координат, но не от φ , и, следовательно, производная от нее равна нулю. Поэтому итерационная схема в (6.4b) принимает вид

$$[\nabla^2 - e^{\varphi^{(t)}} n(\mathbf{x})] \varphi^{(t+1)} = [(1 - \varphi^{(t)}) e^{\varphi^{(t)}} - 1] n(\mathbf{x}). \quad (6.7)$$

Если (6.7) дифференцируется обычным способом (см. гл. 2, (2.21)), то видно, что оператор в левой части является трех-, пяти- или семиточечным разностным оператором в зависимости от числа измерений. Однако коэффициенты оператора переменные, поскольку они зависят от последнего решения $\varphi^{(t)}$, которое является функцией координат. Это приводит к важному следствию, состоящему в том, что уравнения, представленные формулой (6.7), имеют различные коэффициенты на каждой итерации, и всякий раз уравнения должны снова обращаться.

Итерационная схема в уравнении (6.4) служит обобщением метода Ньютона — Рафсона для решения нелинейных алгебраических уравнений, и можно показать, что сходимость очень быстрая (фактически она является квадратичной), если начальное приближение достаточно близко к решению. Это имеет место при моделировании с продвижением по времени, когда в качестве приближенной оценки для решения на шаге n выбирается решение на шаге $(n-1)$.

Для того чтобы найти сходимость уравнения (6.4), разложим его вблизи точного решения φ^* , и пусть $\varepsilon^{(t)}$ будет погрешностью на итерации t . Тогда

$$\left. \begin{aligned} \varphi^{(t)} &= \varphi^* + \varepsilon^{(t)}, \\ N(\varphi^{(t)}) &= N(\varphi^*) + \frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} \varepsilon^{(t)} = \frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} \varepsilon^{(t)}, \\ \frac{\partial N}{\partial \varphi} \Big|_{\varphi^{(t)}} &= \frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} + \frac{\partial^2 N}{\partial \varphi^2} \Big|_{\varphi^*} \varepsilon^{(t)}. \end{aligned} \right\} \quad (6.8)$$

Вычитая $\partial N / \partial \varphi \Big|_{\varphi^*} \varepsilon^{(t)}$ из обеих частей уравнения (6.4b) и подставляя (6.8), получим

$$\left\{ \frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} + \frac{\partial^2 N}{\partial \varphi^2} \Big|_{\varphi^*} \varepsilon^{(t)} \right\} \varepsilon^{(t+1)} = -\frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} \varepsilon^{(t)} + \left\{ \frac{\partial N}{\partial \varphi} \Big|_{\varphi^*} + \frac{\partial^2 N}{\partial \varphi^2} \Big|_{\varphi^*} \varepsilon^{(t)} \right\} \varepsilon^{(t)} \quad (6.9)$$

Первый и второй члены (первого порядка по ε) в правой части уравнения (6.9) взаимно уничтожаются, и, удерживая

только наиболее важные члены самого низкого порядка в обеих частях, получаем

$$\frac{\partial N}{\partial \varphi} \Big|_{\varepsilon^{(t+1)}}^* = \frac{\partial^2 N}{\partial \varphi^2} \Big|_{\varepsilon^{(t)}}^* (\varepsilon^{(t)})^2. \quad (6.10)$$

Уравнение (6.10) является линейным дифференциальным уравнением для определения погрешности на итерации $(t+1)$ через погрешность на итерации (t) . Заметим, что член, представляющий источник в правой части, пропорционален квадрату погрешности на последней итерации. Поскольку уравнение линейно по $\varepsilon^{(t+1)}$, новая погрешность также пропорциональна квадрату предыдущей погрешности. Эта сходимость называется квадратичной и является очень быстрой. Типичные величины погрешностей на последовательных итерациях равны

$$\left. \begin{aligned} \varepsilon^{(0)} &\sim 10^{-1} \\ \varepsilon^{(1)} &\sim 10^{-2} \\ \varepsilon^{(2)} &\sim 10^{-4} \\ \varepsilon^{(3)} &\sim 10^{-8} \end{aligned} \right\}, \quad (6.11)$$

и решение в пределах погрешности округления обычно получается за четыре-пять итераций.

Сравним эту скорость сходимости с экспоненциальной скоростью, которая наблюдается в большинстве схем сеточной релаксации (см. разд. 6.3.1—6.3.6). При экспоненциальной сходимости

$$\varepsilon^{(t+1)} = \lambda \varepsilon^{(t)}, \quad (6.12)$$

где λ — постоянная, модуль которой меньше единицы. В этом случае

$$\varepsilon^{(t+1)} \sim \lambda^t \varepsilon^{(0)} \quad (6.13)$$

и количество итераций, необходимое для снижения погрешности до 10^{-8} , равно $-8/\log_{10}\lambda$. Для типичного значения $\lambda \sim 0.9$ требуется около 100 итераций (вместо 4 или 5, необходимых при квадратичной сходимости).

Неудобство метода Ньютона состоит в том, что уравнения имеют переменные коэффициенты, и прямое использование быстрого метода решения эллиптических уравнений невозможно (см. разд. 6.5.1—6.5.7). Если сетка содержит лишь несколько сотен сеточных узлов, то подходят методы исключения для разреженной матрицы (см. разд. 6.4.3). Для большего числа сеточных узлов необходимо применять какую-нибудь итерационную схему. Для нескольких тысяч узлов подойдут методы неполного разложения (разд. 6.4.4—6.4.6), а если число узлов превышает 10 000, то используется метод сеточной релаксации или итерации по методу RES.

6.3. СЕТОЧНАЯ РЕЛАКСАЦИЯ

Если дифференциальные уравнения для всех сеточных узлов записать в виде одного матричного уравнения, то имеем

$$A\varphi = q, \quad (6.14)$$

где A — матрица размера $N_g \times N_g$, а φ и q — соответственно векторы из N_g неизвестных величин и заданных источников.

Методы сеточной релаксации определяются разбиением матрицы A на две части,

$$A = B + R, \quad (6.15)$$

где B легко обратима, а R — остаток. Тогда (6.14) можно записать как

$$B\varphi = -R\varphi + q \quad (6.16a)$$

или

$$B\varphi = -(A - B)\varphi + q. \quad (6.16b)$$

Сеточная релаксация проводится, отправляясь от начального приближения $\varphi^{(0)}$, образованием серии итераций согласно выражению

$$B\varphi^{(t+1)} = -(A - B)\varphi^{(t)} + q, \quad t = 0, 1, \dots \quad (6.17)$$

Поскольку матрица B выбрана легко обратимой, то уравнение (6.17) для нового приближения легко решить:

$$\varphi^{(t+1)} = -B^{-1}(A - B)\varphi^{(t)} + B^{-1}q, \quad (6.18a)$$

или

$$\varphi^{(t+1)} = M\varphi^{(t)} + B^{-1}q, \quad (6.18b)$$

где $M = -B^{-1}(A - B)$ называется матрицей перехода.

Если теперь определить вектор погрешности на итерации t равным

$$e^{(t)} = \varphi^{(t)} - \varphi, \quad (6.19)$$

где φ — точное решение, то подстановкой в (6.18a) получаем уравнение для погрешности

$$e^{(t+1)} = Me^{(t)} \quad (6.20a)$$

или

$$e^{(t)} = M^{(t)}e^{(0)}, \quad (6.20b)$$

где $e^{(0)}$ — начальный вектор погрешностей. Различные итерационные процессы определяются разбиением матрицы A разными способами, т. е. посредством определения различных видов матрицы B .

Выше мы предполагали, что матрица перехода M остается неизменной на всех итерациях. Это не обязательно так, поэтому

уравнение (6.20) записывают в более общем виде

$$\mathbf{e}^{(t)} = \mathbf{M}^{(t)}\mathbf{e}^{(0)}, \quad (6.21)$$

где $\mathbf{M}^{(t)}$ может быть полиномом от $\mathbf{M}^{(1)}$.

Сходимость данного процесса можно оценить с помощью определения одной скалярной меры, называемой нормой вектора погрешностей. Мы будем использовать норму L_2 , которая определяется как

$$\|\mathbf{e}\| = \sum_{i=1}^{N_g} \varepsilon_i^2, \quad (6.22)$$

где ε_i является погрешностью в сеточном узле i .

Вычисление нормы от обеих частей уравнения (6.21) дает

$$\frac{\|\mathbf{e}^{(t)}\|}{\|\mathbf{e}^{(0)}\|} \leq \|\mathbf{M}^{(t)}\|, \quad (6.23)$$

т. е. гарантируется, что относительное уменьшение величины погрешности после t итераций меньше или равно норме матрицы $\|\mathbf{M}^{(t)}\|$. Поскольку норма L_2 матрицы равняется модулю ее наибольшего собственного значения (называемого спектральным радиусом), в конечном итоге сходимость будет иметь место при условии $\|\mathbf{M}^{(t)}\| < 1$ для $t > t_0$. Ясно, что сходимость для меньших значений $\|\mathbf{M}^{(t)}\|$ более быстрая. Следовательно, матрица \mathbf{B} , использованная при разложении \mathbf{A} , должна выбираться не только из соображений удобства ее обращения, но и чтобы норма матрицы перехода была минимальна. Проблема, возникающая в связи с релаксационными методами, обычно заключается в том, что некоторые собственные векторы матрицы $\mathbf{M}^{(t)}$ имеют собственные значения, по абсолютной величине меньшие единицы, но очень близкие к ней, и сходимость векторов погрешностей такого вида итераций действительно может быть очень медленной. Это показывается ниже в методе Гаусса — Зайделя (разд. 6.3 2).

На рис. 6.2 изображена величина $\|\mathbf{M}^{(t)}\|$ для различных методов сеточной релаксации. Важно понимать, что на рис. 6.2 изображен не график убывания погрешности для какого-то конкретного случая, а огибающая, ниже которой должны лежать все кривые убывания погрешности. Равенство в (6.23) гарантирует, что для любой точки линий на рис. 6.2 существует начальный вектор погрешности, кривая убывания которого будет касаться этой точки огибающей. В любом частном случае гарантируется лишь то, что фактическая кривая убывания лежит ниже линии для соответствующего процесса на рис. 6.2, и вполне возможно, что конкретные кривые не будут касаться этих линий.

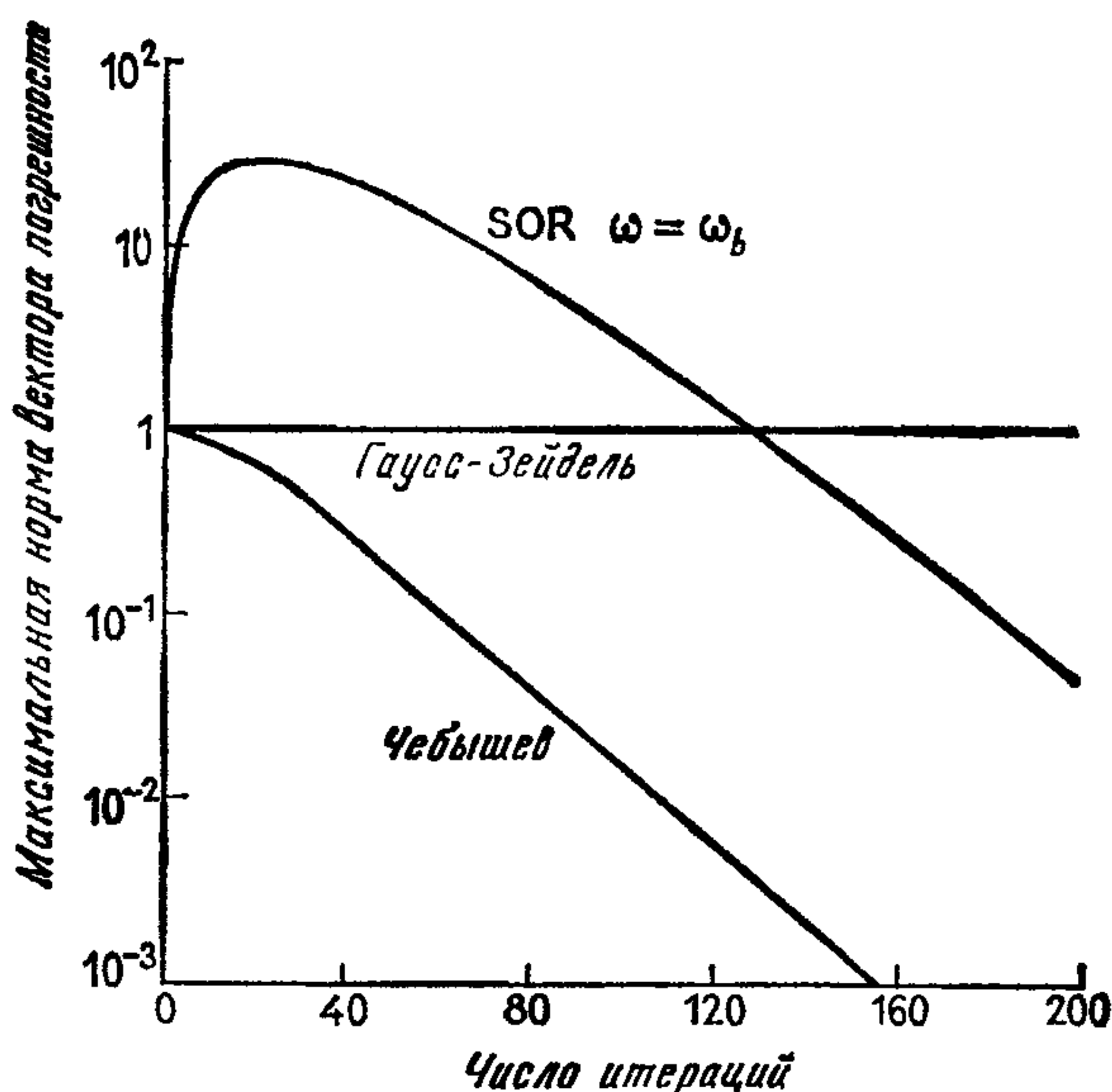


Рис. 6.2. Зависимость теоретических границ изменения максимально возможной нормы вектора погрешности (по отношению к ее начальному значению) от числа итераций для методов последовательной верхней релаксации (SOR), Гаусса — Зейделя и Чебышёва на сетке 128×128 при использовании нечетно-четного обхода узлов [180].

Если требуется получить решение с точностью p десятичных знаков (погрешность $< 10^{-p}$), то необходимое количество итераций t^* находится из соотношения

$$\|M^{(t^*)}\| = 10^{-p}. \quad (6.24a)$$

Для простоты методы обычно сравниваются на основе их асимптотической сходимости для больших t . Если λ — коэффициент асимптотической сходимости, то

$$\lambda^{t^*} = 10^{-p}$$

и

$$t^* = -p \frac{\ln 10}{\ln \lambda}. \quad (6.24b)$$

Релаксационные методы для решения дифференциальных уравнений в частных производных хорошо описаны Янгом [390] и Форсайтом и Вазовым [135]. Более полное обсуждение теории матриц и нахождения скоростей сходимости дано Варгой [376].

Матричное представление данной выше релаксационной процедуры удобно для разработки теории сходимости. Матрицы A и B , которые появляются в выражениях, никогда, конечно, явно не запоминаются и не обращаются. Однако с точки зрения

программиста релаксационные схемы лучше всего рассматривать на языке уравнений для отдельных сеточных узлов и операций, совершаемых над ними.

В общем случае пятиточечное разностное уравнение, соответствующее одной строке матрицы \mathbf{A} , можно записать в виде

$$a_{p,q}\varphi_{p,q-1} + b_{p,q}\varphi_{p,q+1} + c_{p,q}\varphi_{p-1,q} + d_{p,q}\varphi_{p+1,q} + e_{p,q}\varphi_{p,q} = f_{p,q}, \quad p, q = 1, \dots, n, \quad (6.25)$$

где коэффициенты a, b, c, d, e , вообще говоря, меняются от сеточного узла к сеточному узлу, и индексы p, q указывают номера сеточных узлов по x и y , соответственно. Для уравнения Пуассона на равномерной сетке $a = b = c = d = 1$ и $e = -4$, но это очень частный случай, и было бы совершенно неправильно использовать для этой задачи общий итерационный метод. Существуют гораздо более быстрые и более надежные методы для таких задач, и они описаны в разд. 6.5.1—6.5.7. Эти методы RES обычно в пять-десять раз быстрее, чем описываемые сейчас итерационные методы, но они применимы лишь в этих важных специальных случаях.

Уравнение (6.25) превращают в явную итерационную схему путем выражения значения φ в сеточном узле (p, q) через ее значения в соседних сеточных узлах следующим образом:

$$\varphi_{p,q}^* = \frac{1}{e_{p,q}} (f_{p,q} - a_{p,q}\varphi_{p,q-1} - b_{p,q}\varphi_{p,q+1} - c_{p,q}\varphi_{p-1,q} - d_{p,q}\varphi_{p+1,q}), \quad (6.26a)$$

$$\varphi_{p,q}^{\text{new}} = \omega\varphi_{p,q}^* + (1 - \omega)\varphi_{p,q}^{\text{old}}, \quad (6.26b)$$

где ω — постоянный или переменный релаксационный параметр, выбираемый из соображений оптимизации сходимости итераций.

В реальном расчете полезно выразить (6.26) через невязку

$$r_{p,q} = a_{p,q}\varphi_{p,q-1} + b_{p,q}\varphi_{p,q+1} + c_{p,q}\varphi_{p-1,q} + d_{p,q}\varphi_{p+1,q} + e_{p,q}\varphi_{p,q} - f_{p,q}. \quad (6.27a)$$

Тогда уравнение (6.26) принимает вид

$$\varphi_{p,q}^{\text{new}} = \varphi_{p,q}^{\text{old}} - \frac{r_{p,q}}{e_{p,q}} \omega. \quad (6.27b)$$

В этой формулировке норму вектора невязки можно удобно накопить и использовать в качестве критерия окончания итерации (см. разд. 6.3.4).

Итерационные процедуры отличаются: 1) выбором ω , 2) порядком, в котором обрабатываются сеточные узлы, и 3) моментом, когда в ходе процедуры новые значения заменяют старые значения на сетке.

На рис. 6.3 показаны два важных способа обхода узлов. В первом способе сеточные узлы обрабатываются последова-

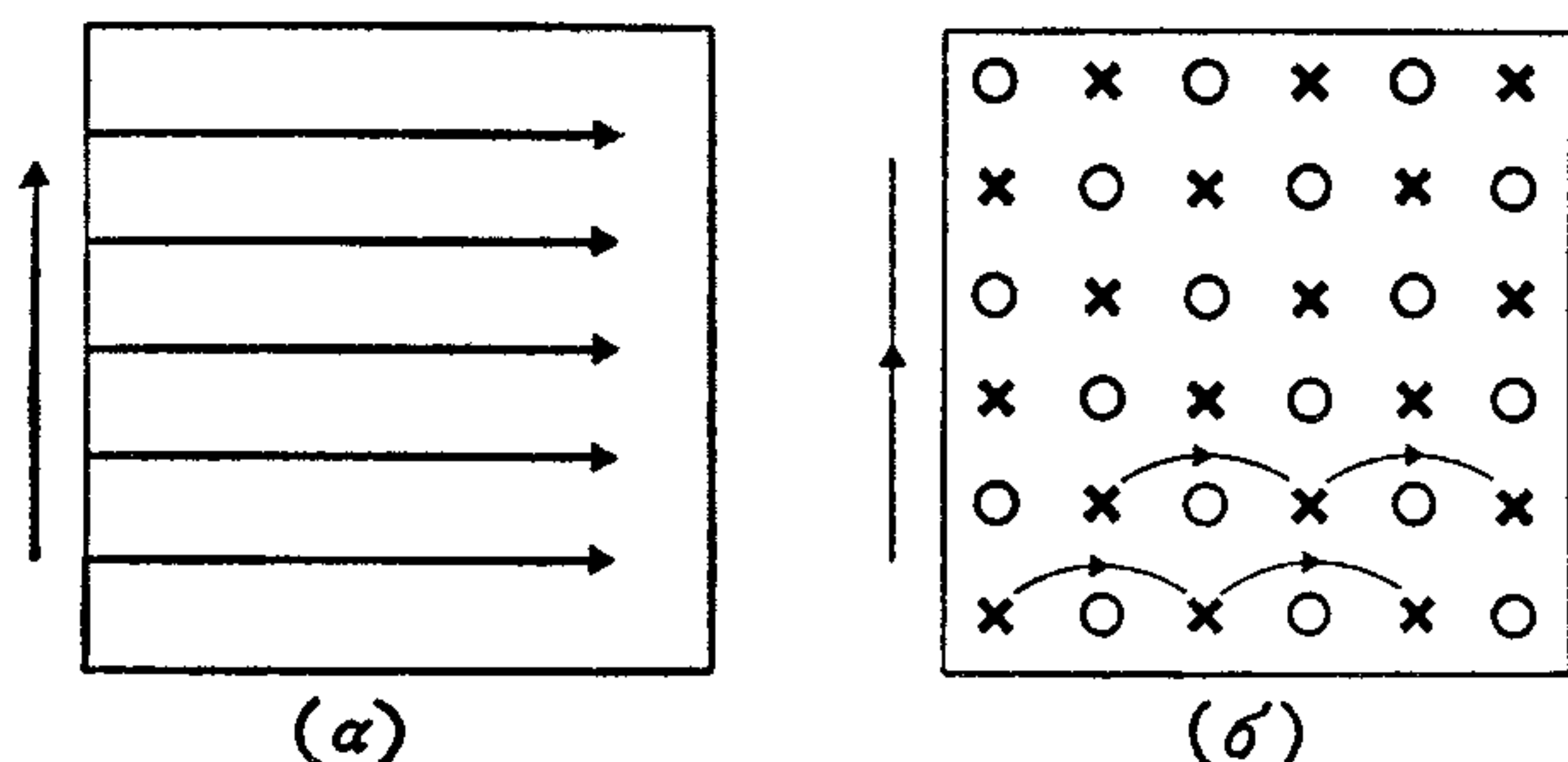


Рис. 6.3. Обходы узлов, применяемые в схемах сеточной релаксации: (а) обход по принципу пишущей машинки; (б) нечетно-четный (в шахматном порядке) обход.

тельно, строка за строкой, как при чтении слов в книге. В нечетно-четном обходе сеточные узлы разбиваются на две группы, как черные и белые квадраты шахматной доски. Черные, или нечетные, сеточные узлы проходят в первой половине обхода, а белые, или четные, — во второй его половине. Поскольку в уравнение (6.26а) для нечетных узлов входят только значения в четных узлах, все переменные в правой части — это старые значения, взятые из последней половины обхода. Отсюда следует, что порядок, в котором проходят нечетные (или четные) узлы, несуществен и не меняет алгоритма. Однако для удобства программирования узлы внутри каждой группы обычно проходят последовательно.

Теперь мы опишем наиболее важные релаксационные методы с точки зрения только что изложенных идей.

6.3.1. Метод Якоби (J)

В методе Якоби матрица B является диагональной, образованной из диагональных элементов матрицы A . Таким образом, решение для следующей итерации получается путем решения уравнения для каждого сеточного узла относительно значения в этом узле в предположении, что все окружающие значения правильны. Полученные таким образом новые значения не замещают старые значения до тех пор, пока не будут пересчитаны все сеточные величины. Таким образом, на языке сеточных узлов, все значения ϕ в правой части уравнения (6.26а) суть старые значения, взятые из последней итерации и $\omega \equiv 1$. Этот метод представляет только теоретический интерес из-за его медленной сходимости. Мы упоминаем его, поскольку скорость сходимости всех последующих методов полностью выражается через спектральный радиус ρ матрицы перехода Якоби. Для сравнения возьмем в качестве модельной задачи уравнение Пуассона с пятиточечной разностной аппроксимацией лапла-

сиана на регулярной сетке $n \times n$. Область считаем квадратной, $\varphi \equiv 0$ границе. Для этой модельной задачи имеем

$$\rho = \cos(\pi/n) \simeq 1 - \frac{1}{2} \frac{\pi^2}{n^2}, \quad (6.28a)$$

и коэффициент асимптотической сходимости равен

$$\lambda_J = \rho. \quad (6.28b)$$

6.3.2. Метод Гаусса — Зайделя (GS)

Если матрица \mathbf{B} является нижней треугольной матрицей, образованной из элементов нижнего треугольника матрицы \mathbf{A} , то мы имеем итерацию Гаусса — Зайделя. Обращение матрицы \mathbf{B} проводится прямой подстановкой. С вычислительной точки зрения этот метод сходен с методом Якоби, за исключением того, что вычисляемые вновь значения для следующей итерации замещают старые значения из предыдущей итерации сразу после того, как они вычисляются. С точки зрения сеточных уравнений (6.26), в правой части уравнения (6.26a) берутся самые последние имеющиеся значения и $\omega = 1$.

При использовании нечетно-четного обхода коэффициент асимптотической сходимости метода Гаусса — Зайделя равен

$$\lambda_{GS} = \rho^2, \quad (6.29)$$

где ρ — спектральный радиус метода Якоби, и

$$\lambda_{GS} = \cos^2(\pi/n) \simeq 1 - \frac{\pi^2}{n^2}. \quad (6.30a)$$

Следовательно,

$$t^* = \frac{\ln 10}{\pi^2} (pn^2) \simeq \frac{pn^2}{4} \simeq \frac{p(N_g)^{2/d}}{4}. \quad (6.30b)$$

Из (6.26) находим, что количество операций, приходящихся на один узел, равно 9; следовательно, количество операций, требуемое для решения уравнений методом Гаусса — Зайделя, равно

$$T_{GS} \simeq 2pn^4 = 2pN_g^2. \quad (6.31)$$

Поскольку число сеточных узлов N_g , вероятно, будет составлять по крайней мере 1000, то для приемлемой сходимости потребуется несколько тысяч итераций. По этой причине метод Гаусса — Зайделя совершенно бесполезен для вычислительных целей. Мы рассмотрели его здесь в качестве предупреждения, а также потому, что он пригодится при описании последующих методов. Это заключение подкрепляется рис. 6.2, на котором показано, что за сто итераций в модельной задаче с сеткой (128×128) снижение погрешности практически отсутствует.

6.3.3. Последовательная верхняя релаксация (SOR)

В методе последовательной точечной верхней релаксации учитывается, что сходимость метода Гаусса — Зайделя можно существенно улучшить использованием линейной комбинации старого значения φ_p^{old} и значения, полученного в методе Гаусса — Зайделя, которое выражено в (6.26b). Параметр релаксации ω лежит в диапазоне $1 \leq \omega \leq 2$ и соответствует взятию избыточной коррекции в точке для того, чтобы упредить будущие коррекции. В некоторых случаях сходимость можно получить только при $\omega < 1$; в этом случае процесс называется нижней релаксацией.

Если спектральный радиус ρ матрицы Якоби известен, то можно показать, что наилучшее значение ω равно

$$\omega_b = \frac{2}{1 + \sqrt{1 - \rho^2}}. \quad (6.32)$$

В этом случае коэффициент асимптотической сходимости равен

$$\lambda_{\text{SOR}} = \omega_b - 1. \quad (6.33)$$

В случае двумерного уравнения Пуассона в квадрате с сеткой $n \times n$ для больших n получаем

$$\rho = \cos(\pi/n) \simeq 1 - \frac{1}{2} \frac{\pi^2}{n^2}, \quad (6.34a)$$

$$\omega_b \simeq 2(1 - \pi/n), \quad (6.34b)$$

$$\lambda_{\text{SOR}} = \omega_b - 1 = 1 - 2\pi/n. \quad (6.34c)$$

Количество итераций t^* , асимптотически необходимое для снижения погрешности в 10^p раз, дается выражением

$$\lambda_{\text{SOR}}^t = (1 - 2\pi/n)^t = 10^{-p} \quad (6.35a)$$

или

$$t^* \approx \frac{\ln 10}{2\pi} (np) \simeq \frac{pn}{3}. \quad (6.35b)$$

Если требуется точность 10^{-3} ($p = 3$), то число итераций приблизительно равно количеству сеточных узлов n в каждом направлении. Для достижения точности 10^{-6} необходимо удвоить это число итераций. Таким образом, в задаче с пространственным разрешением около 1% ($n \simeq 100$) выражение (6.35b) показывает, что любая приемлемая точность решения потребует несколько сотен итераций. Количество операций, приходящихся на один узел в схеме (6.26), равно 11; следовательно, полное число операций, необходимое для решения уравнения (6.25)

с точностью 10^{-p} , оценивается следующим образом:

$$\begin{aligned} T_{\text{SOR}} &= 11N_g \frac{p}{3} (N_g)^{1/d} \\ &\cong 22N_g (N_g)^{1/d} \text{ для } p=6, \end{aligned} \quad (6.36)$$

где d — размерность задачи.

Из этого обсуждения вытекает важное заключение: количество итераций, требуемое для решения задачи методом SOR, пропорционально не только необходимому числу верных знаков, а именно p , но также числу сеточных узлов в одном направлении сетки — время решения пропорционально $N_g^{(1+1/d)}$, а не N_g . Это крупный недостаток, когда число сеточных узлов становится большим.

6.3.4. Ускорение по Чебышёву

На рис. 6.2 показано, что для SOR асимптотическая скорость сходимости может не появиться до тех пор, пока не будет сделано приблизительно n итераций. На самом деле за первые n итераций вектор погрешности может вырасти примерно в 20 раз. Этот рост начальной погрешности наблюдается на практике и делает метод SOR ненадежным, пока число итераций не превысит нескольких n . Это случается редко.

Тривиальная модернизация метода SOR, известная как чебышевское ускорение, позволяет преодолеть эту трудность. В методе Чебышёва используется нечетно-четный способ обхода, причем ω изменяется на каждой половине прохода следующим образом:

$$\begin{aligned} \omega^{(0)} &= 1, \\ \omega^{(1/2)} &= 1 / \left(1 - \frac{1}{2} \rho^2 \right), \\ \omega^{(t+1/2)} &= 1 / \left(1 - \frac{1}{4} \rho^2 \omega^{(t)} \right), \quad t = \frac{1}{2}, 1, \dots, \infty, \\ \omega^{(\infty)} &\rightarrow \omega_b. \end{aligned} \quad (6.37)$$

Можно показать, что максимально возможная норма вектора погрешности всегда уменьшается с ростом числа итераций. Поэтому метод называется «уменьшением нормы». Коэффициент асимптотической сходимости оказывается тем же, что и в методе SOR, но гарантированное улучшенное поведение для малого числа итераций делает метод Чебышёва предпочтительнее SOR. Действительно, поскольку количество операций на итерацию одно и то же для обоих случаев, то, по-видимому, нет причин использовать метод SOR без ускорения по Чебышёву.

В очевидных фортраноподобных обозначениях алгоритм ускорения по Чебышёву представляется следующим образом:

```

      ω = 1.0
      → DO 2 t = 1, MAXIT
          Норма = 0.0
          → DO 1 p = 1, n
              DO 1 q = 1, n
                  IF (MOD(p + q, 2).NE.MOD(t, 2)) go to 1
                  Невязка = ap,qφp,q-1 + bp,qφp,q+1 + cp,qφp-1,q +
                      + dp,qφp+1,q + ep,qφp,q - fp,q
                  Норма = Норма + (Невязка)2
                  φp,q = φp,q - ω × Невязка/ep,q
              1 CONTINUE
          * ω = 1.0/(1 - 0.25 × ρ2 × ω)
          * IF (t.EQ.1) ω = 1.0/(1 - 0.5 × ρ2)
          IF (Норма .LT.EPS × ||f||) решение найдено
      2 CONTINUE
  
```

(6.38)

Здесь $\|f\|$ — это заранее вычисленная норма правой части уравнения, а EPS — заранее определенный предел точности (например, 10^{-6}). MAXIT — заранее заданное максимальное число итераций, введенное для предотвращения заикливания в случаях медленной сходимости. Обычно MAXIT = 100 или 1000.

Этот алгоритм можно преобразовать в алгоритм классического метода Гаусса — Зайделя, опуская из него инструкции, помеченные звездочкой. Если, кроме того, правая часть первой инструкции заменяется на ω_b , то получается итерация SOR.

6.3.5. Блочные методы

До сих пор рассматривались методы, предусматривающие поточечное уточнение приближенного решения. Однако обсуждение решения одномерной системы с трехдиагональной матрицей в разд. 6.4.1 показывает, что можно столь же просто найти решение непосредственно для строки неизвестных величин, если значения решения на двух соседних строках являются точными. Метод последовательной построчной верхней релаксации (SLOR) можно определить следующим образом. Строки можно разбить на нечетные и четные группы и ускорить процесс с использованием чебышевской методики в соответствии с (6.37). Для таких однострочных методов в описанной модельной задаче получаем

$$\rho = \cos(\pi/n)/[2 - \cos(\pi/n)] \simeq 1 - \pi^2/n^2, \quad (6.39a)$$

следовательно,

$$\omega_b = 2/2(1 + \sqrt{1 - \rho^2}) \simeq 2 - 2\sqrt{2} \frac{\pi}{n} \quad (6.39b)$$

и

$$\lambda_{\text{SLOR}} = \omega_b - 1 \simeq 1 - 2\sqrt{2} \frac{\pi}{n}, \quad (6.39c)$$

тогда

$$t_{\text{SLOR}}^* = \frac{\ln 10}{2\sqrt{2}\pi} (np) \simeq \frac{pn}{4}, \quad (6.39d)$$

$$T_{\text{SLOR}} = 11N_g \frac{pn}{4} \simeq 3pn^3. \quad (6.39e)$$

Можно также определить двухстрочный блочный метод SOR, в котором используется эффективная процедура нахождения значений в двух соседних строках, если заданы значения в следующих смежных строках. Сходимость двухстрочного метода в $\sqrt{2}$ раз лучше, чем однострочного метода SOR.

Общая особенность этих методов состоит в том, что скорость сходимости улучшается с увеличением размера блоков, решаемых неявно. Другие варианты метода можно получить, поочередно находя значения для строк, идущих в направлениях x , y и z при последовательных обходах узлов сетки. Наглядно можно представлять себе, что погрешность свободно перемещается к границе по строке в направлении обхода, но очень медленно в направлениях, перпендикулярных к нему. Цель заключается в перемещении погрешности к границе, где она поглощается. Это достигается быстрее, если направление обхода меняется циклически.

Последнее усовершенствование блочных методов описывается Конкусом и Голубом [86], которые используют программу, решающую уравнение Пуассона на всей сетке, в качестве основы итерационного решения задачи, имеющей сходство с уравнением Пуассона. Этот подход можно рассматривать как блочную релаксацию, и он описывается вместе с методами быстрого решения эллиптических уравнений в разд. 6.5.7.

Блочные методы применяются реже, чем точечный метод SOR, в основном потому, что они сложнее для программирования. Однако поскольку они отличаются улучшенными скоростями сходимости, ими следует пользоваться, когда доступна библиотечная программа или можно затратить дополнительные усилия на программирование.

6.3.6. Неявный метод переменных направлений (ADI)

Метод переменных направлений (ADI) является блочно-итерационным методом, первоначально разработанным для решения нестационарного уравнения теплопроводности

$$\frac{\partial \varphi}{\partial t} = \nabla^2 \varphi + q. \quad (6.40)$$

Он применим для решения эллиптических задач, поскольку, если найдено стационарное решение уравнения (6.40) с $\partial\varphi/\partial t = 0$, то φ является решением уравнения

$$-\nabla^2\varphi = q. \quad (6.41)$$

В большей общности можно применить этот метод для решения линейной задачи

$$L\varphi = q \quad (6.42)$$

посредством решения нестационарной задачи

$$\frac{\partial\varphi}{\partial t} = -L\varphi + q. \quad (6.43)$$

Оператор L , в качестве которого можно сразу брать требуемый конечно-разностный оператор, является положительно определенным (подобно оператору $-\nabla^2$). Расщепим теперь оператор L на две части

$$L = L_x + L_y, \quad (6.44)$$

где L_x представляет дифференцирование по x , а L_y — дифференцирование по y .

В случае пятиточечной аппроксимации на равномерной сетке оператора $-\nabla^2\varphi$ мы бы использовали

$$-\frac{\partial^2\varphi}{\partial x^2} \simeq L_x\varphi = 2\varphi_{p,q} - (\varphi_{p-1,q} + \varphi_{p+1,q}), \quad (6.45a)$$

$$-\frac{\partial^2\varphi}{\partial y^2} \simeq L_y\varphi = 2\varphi_{p,q} - (\varphi_{p,q-1} + \varphi_{p,q+1}) \quad (6.45b)$$

$$\text{и} \quad -\nabla^2\varphi \simeq L\varphi = [L_x + L_y]\varphi. \quad (6.45c)$$

Более сложные операторы тоже можно расщепить аналогично, но полезнее основывать расщепление на естественном расщеплении исходных физических дифференциальных уравнений, как это и было сделано выше. Мы знаем, что физическая задача имеет стационарное решение, поскольку известно, что температура φ должна приближаться к термодинамическому равновесию после начального переходного процесса. Если расщепление и конечно-разностное дифференцирование отражают это поведение, то можно иметь некоторую уверенность относительно сходимости метода.

Уравнение с зависимостью от времени

$$\frac{\partial\varphi}{\partial t} = -(L_x + L_y)\varphi + q \quad (6.46)$$

далее дифференцируется неявно на двух полушагах:

$$\frac{\varphi^{n+1/2} - \varphi^n}{\Delta t/2} = -L_x\varphi^{n+1/2} - L_y\varphi^n + q, \quad (6.47a)$$

$$\frac{\varphi^{n+1} - \varphi^{n+1/2}}{\Delta t/2} = -L_x\varphi^{n+1/2} - L_y\varphi^{n+1} + q. \quad (6.47b)$$

В матрично-векторных обозначениях, положив $r = \Delta t/2$, получим

$$[\mathbf{I} + r\mathbf{L}_x] \Phi^{n+1/2} = [\mathbf{I} - r\mathbf{L}_y] \Phi^n + r\mathbf{q}, \quad (6.48a)$$

$$[\mathbf{I} + r\mathbf{L}_y] \Phi^{n+1} = [\mathbf{I} - r\mathbf{L}_x] \Phi^{n+1/2} + r\mathbf{q}. \quad (6.48b)$$

Операторы в левой части уравнения (6.48) являются трехдиагональными матрицами, и соответствующие системы можно быстро решить построчным применением алгоритма Томаса (разд. 6.4.1). Заметим, что правую часть уравнения (6.48) нужно сначала вычислять с использованием всюду «старых» значений Φ из последнего обхода узлов, прежде чем решать какие-либо трехдиагональные уравнения.

Если исключить $\Phi^{n+1/2}$, то получится матрица перехода \mathbf{M} для одного двойного обхода:

$$\Phi^{n+1} = \mathbf{M}\Phi^n + \mathbf{N}, \quad (6.49a)$$

где

$$\mathbf{M} = (\mathbf{I} + r\mathbf{L}_y)^{-1} (\mathbf{I} + r\mathbf{L}_x)^{-1} (\mathbf{I} - r\mathbf{L}_x) (\mathbf{I} - r\mathbf{L}_y), \quad (6.49b)$$

$$\mathbf{N} = 2(\mathbf{I} + r\mathbf{L}_y)^{-1} (\mathbf{I} + r\mathbf{L}_x)^{-1} r\mathbf{q}. \quad (6.49c)$$

Если \mathbf{L}_x и \mathbf{L}_y коммутируют и симметричны, то их собственные векторы совпадают. Если собственные значения оператора \mathbf{L}_x суть λ_x , а оператора \mathbf{L}_y суть λ_y , то собственные числа матрицы перехода \mathbf{M} равны

$$\lambda(\mathbf{M}) = \frac{(1 - r\lambda_x)(1 - r\lambda_y)}{(1 + r\lambda_x)(1 + r\lambda_y)}. \quad (6.50)$$

Из выражения (6.50) ясно, что собственные значения матрицы \mathbf{M} по абсолютной величине меньше единицы для всех положительных r . Этого и следовало ожидать, поскольку положительное r соответствует интегрированию уравнения диффузии вперед по времени, т. е. физически устойчивому процессу. Погрешность можно считать диффундирующей к границе, где она поглощается. Использование отрицательного r , или отрицательного шага по времени, соответствовало бы интегрированию назад по времени. Эта отрицательная диффузия приводит к образованию выбросов или накоплению погрешностей, и нельзя ожидать, что такой процесс будет сходящимся.

Параметр r можно варьировать на каждом двойном обходе узлов с целью улучшения сходимости. Если

$$0 < \lambda_{\min} \leq \lambda_x, \lambda_y \leq \lambda_{\max}, \quad (6.51)$$

то последовательность r_k выбирается с помощью соотношения

$$r_k = 1/(\lambda_{\max} x^{k-1}), \quad k = 1, 2, \dots, t, \quad (6.52a)$$

где

$$x = \left(\frac{\lambda_{\min}}{\lambda_{\max}} \right)^{1/(t-1)}. \quad (6.52b)$$

Длина цикла t должна быть выбрана так, чтобы минимизировать

$$S_t = \left(\frac{1 - \sqrt{x}}{1 + \sqrt{x}} e^{-x^{3/2}/(1+x)} \right)^{4/t}. \quad (6.53)$$

Параметры r_k меняются в диапазоне от $r_1 = 1/\lambda_{\max}$ до $r_t = 1/\lambda_{\min}$, и для модельной задачи, заданной в квадратной области, имеем

$$\lambda_{\min} = 4 \sin^2(\pi/2n), \quad \lambda_{\max} = 4 \cos^2(\pi/2n), \quad t = 9.$$

Для практических целей подходящие значения для λ_{\min} и λ_{\max} можно оценить, погружая реальную задачу в квадратную область и используя указанные выше значения.

Количество итераций, требуемое для получения заданной точности с помощью метода ADI, пропорционально $\log N_g$, а не $N_g^{1/d}$, как в методе SOR или SLOR. Однако условия, при которых известно, что метод ADI сходится, являются более ограниченными, чем для метода верхней релаксации, и нередко обнаруживается расходимость метода ADI в задачах, в которых сходится метод SOR, хотя бывает и наоборот. Поэтому можно сказать, что метод ADI следует испытывать на задачах в первую очередь, потому что в случае успеха его скорость сходимости скорее всего окажется выше, чем скорость сходимости метода SOR. В случае неуспеха, как правило, можно перейти к последовательной верхней релаксации.

6.4. МАТРИЧНЫЕ МЕТОДЫ

В изложенных ниже матричных методах уравнения, возникающие при конечно-разностном дифференцировании на сетке, рассматриваются просто как большая система линейных уравнений, и используются различные стандартные методы, разработанные для линейных систем общего вида. К ним относятся методы исключения Гаусса, разложения Холецкого и методы сопряженных градиентов. Такие методы применимы к линейным задачам самого общего вида, но их недостаток — неиспользование известной структуры уравнений.

6.4.1. Трехдиагональный алгоритм Томаса

Приложение конечных разностей с переменным сеточным шагом к обыкновенным дифференциальным уравнениям второго порядка (ОДУ) приводит к трехдиагональной системе уравнений с произвольными коэффициентами. Аналогичные уравнения получаются из метода Ньютона в одномерном случае (разд. 6.2). Такие уравнения можно эффективно решать за $5N_g$ операций

при любых коэффициентах стандартным методом исключения Гаусса.

Если разностные уравнения имеют вид

$$a_i \varphi_{i-1} + b_i \varphi_i + c_i \varphi_{i+1} = d_i, \quad i = 1, \dots, N_g, \quad (6.54)$$

то решение можно получить с помощью двух циклов на языке Фортран, отражающих рекуррентные соотношения для прямой и обратной прогонки.

Прямая прогонка:

$$\omega_1 = \frac{c_1}{b_1}, \quad \omega_i = \frac{c_i}{b_i - a_i \omega_{i-1}}, \quad i = 2, 3, \dots, N_g - 1, \quad (6.55)$$

$$g_1 = \frac{d_1}{b_1}, \quad g_i = \frac{d_i - a_i g_{i-1}}{b_i - a_i \omega_{i-1}}, \quad i = 2, 3, \dots, N_g. \quad (6.56)$$

Обратная прогонка:

$$\varphi_{N_g} = g_{N_g}, \quad \varphi_i = g_i - \omega_i \varphi_{i+1}, \quad i = N_g - 1, N_g - 2, \dots, 1. \quad (6.57)$$

Для частного периодического случая с $a = c = 1$ и $b = -2$ в гл. 2 дается другой алгоритм (уравнения (2.54) и (2.55)).

Если коэффициенты a_i, b_i, c_i изменяются на каждой итерации, как, например, в схеме Ньютона, то требуемое число операций равно $8N_g$. Однако если коэффициенты остаются постоянными и меняется только член d_i , описывающий источник, то промежуточные векторы ω_i и $1/(b_i - a_i \omega_{i-1})$ можно заранее рассчитать и запомнить. В этом случае решение получается за $5N_g$ операций. Если, кроме того, $a_i = c_i = 1$, то число операций можно еще больше понизить до $4N_g$.

Во время прямой прогонки можно g_i записывать на место правой части d_i , а при обратной прогонке φ_i можно записывать на место g_i . Однако для переменных ω_i необходима дополнительная память, на что требуется в дополнение к памяти для коэффициентов полная память объемом $2N_g$.

6.4.2. Алгоритм сопряженных градиентов (CGA)

Алгоритм сопряженных градиентов представляет собой общий метод нахождения минимума функции. Он может использоваться для решения системы разностных уравнений

$$\mathbf{A}\varphi = \mathbf{q} \quad (6.58)$$

посредством определения квадратичной функции

$$V(\varphi) = \frac{1}{2} \varphi^T \mathbf{A}\varphi - \mathbf{q}^T \varphi, \quad (6.59)$$

где \mathbf{q} обозначает вектор, элементы которого равны N_g значениям заряда на сетке, а $\boldsymbol{\varphi}$ — вектор неизвестных значений потенциалов. Матрица \mathbf{A} размерности $(N_g \times N_g)$ появляется в результате применения разностных уравнений (6.25) для каждого из N_g узлов сетки.

Метод сопряженных градиентов применим, если матрица \mathbf{A} симметричная и положительно определенная. Ее можно сделать такой, если \mathbf{A} является конечно-разностным представлением оператора $-\nabla^2$. В этом случае, дифференцируя уравнение (6.59), получаем вектор градиента \mathbf{r} , который оказывается невязкой исходных уравнений

$$\frac{\partial V}{\partial \boldsymbol{\varphi}} = \mathbf{A}\boldsymbol{\varphi} - \mathbf{q} = \mathbf{r}, \quad (6.60)$$

и, снова дифференцируя, мы показываем, что матрица \mathbf{A} является матрицей вторых производных:

$$A_{ij} = \frac{\partial^2 V}{\partial \varphi_i \partial \varphi_j}. \quad (6.61)$$

Уравнения (6.60) и (6.61) показывают, что все первые производные на решении исходных уравнений (6.58) равны нулю и если \mathbf{A} положительно определенная, то этот экстремум соответствует минимуму функции V .

В методе CGA минимум квадратичной функции от N_g переменных, например определенной формулой (6.59), находится самое большее за N_g итераций. В реальных задачах часто сообщается о значительно более быстрой сходимости, однако можно показать, что в отсутствие погрешностей округления точное решение уравнения (6.58) должно находиться за N_g итераций. В методе CGA последовательно находятся минимумы в подпространствах N_g переменных, начиная с одномерного подпространства. Найдя минимум в одном подпространстве, алгоритм увеличивает размерность подпространства на единицу и определяет точку нахождения нового минимума. Решение найдено, когда размерность подпространства достигает N_g — числа исходных уравнений.

В двумерном случае, проиллюстрированном на рис. 6.4, процесс протекает следующим образом. Градиент \mathbf{r}_0 вычисляется в точке начального приближения $\boldsymbol{\varphi}_0$. Начальное одномерное подпространство задается направлением поиска $\mathbf{s}_0 = -\mathbf{r}_0$, которое является направлением наибыстрейшего спуска из точки $\boldsymbol{\varphi}_0$. Затем вычисляется минимум $V(\boldsymbol{\varphi})$ в направлении \mathbf{s}_0 . Если ω — расстояние вдоль \mathbf{s}_0 , то $V(\boldsymbol{\varphi}_0 + \omega \mathbf{s}_0)$ является скалярной квадратичной функцией от ω и можно непосредственно найти минимум при $\omega = \omega_0$. Это положение $\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_0 + \omega_0 \mathbf{s}_0$ определяет исходную точку для следующей итерации. Если продолжить этот процесс поиска вдоль линии наибыстрейшего спуска, то

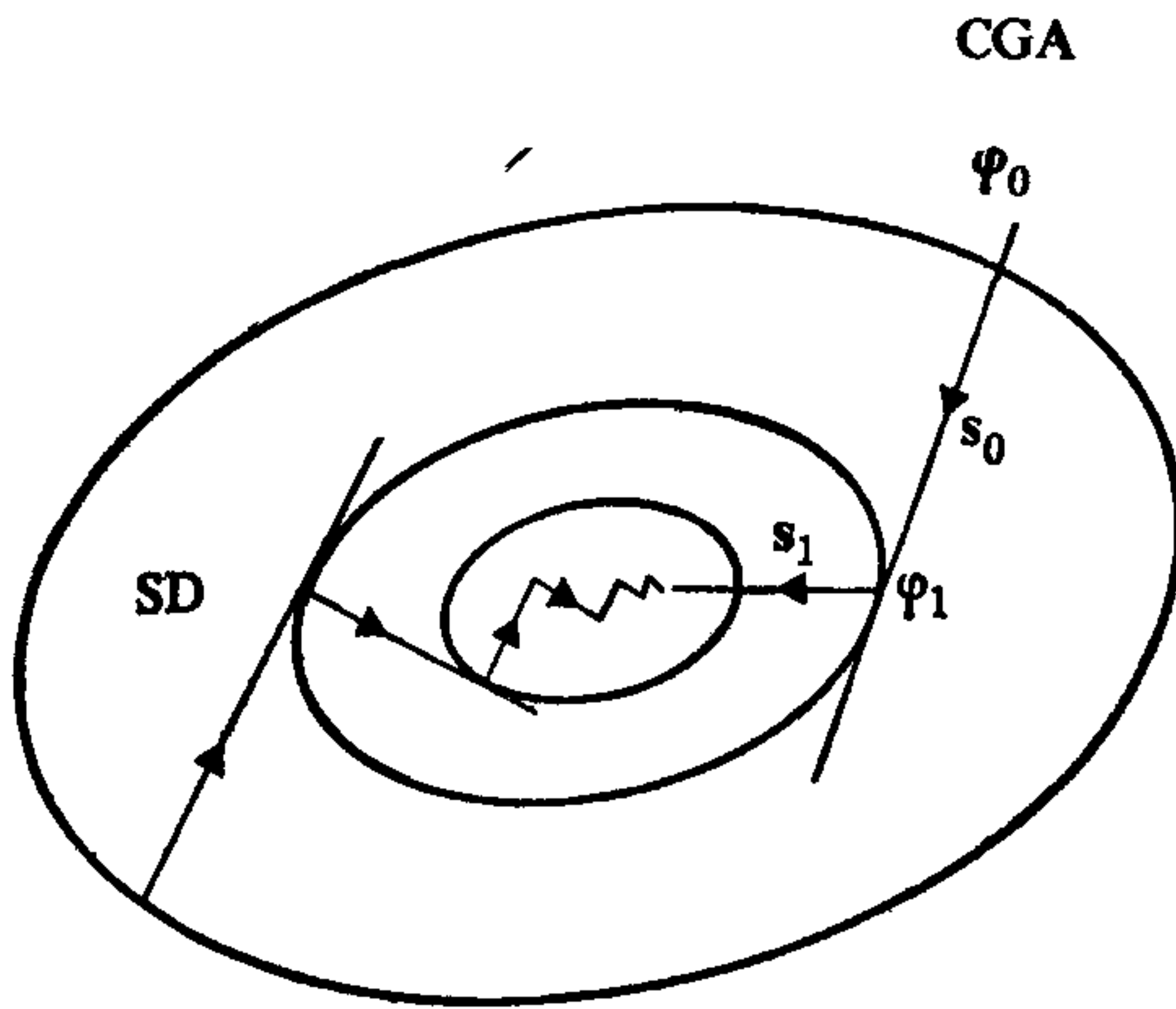


Рис. 6.4. CGA: алгоритм сопряженных градиентов, сходящийся за два шага SD: алгоритм наискорейшего спуска медленно сходится в узкой долине с $\lambda_{\max}/\lambda_{\min} \gg 1$.

мы получим метод «наискорейшего спуска». Когда отношение максимального и минимального собственных чисел матрицы A (называемое числом обусловленности) велико, т. е. долина на рис. 6.4 очень длинная и узкая, этот метод может крайне медленно сходиться. В этом случае последовательные итерации приводят к перескакиванию с одной стороны долины на другую и лишь очень медленному продвижению к минимуму, как это показано слева на рис. 6.4. В методе CGA эта трудность преодолевается выбором в качестве очередного направления поиска s_1 линейной комбинации s_0 и $-r_1$ — нового направления наискорейшего спуска в точке φ_1 , а именно

$$s_1 = -r_1 + \beta s_0, \quad (6.62a)$$

где

$$\beta = \frac{r_1 \cdot r_1}{r_0 \cdot r_0}. \quad (6.62b)$$

Можно показать, что s_1 есть направление минимума в двумерном подпространстве, содержащем r_1 и s_0 . В двумерном примере на рис. 6.4 минимум вдоль s_1 оказывается решением задачи. Он достигнут за два шага.

Двумерный случай можно рассматривать как срез трехмерной задачи, и метод естественным образом обобщается на более высокие размерности. Векторы последовательных невязок обладают важным свойством ортогональности друг к другу. Поскольку в N_g -мерном пространстве может быть не более N_g различных ортогональных векторов, то $(N_g + 1)$ -й вектор невязки должен быть нулевым вектором, и решение найдено.

Свойство ортогональности никогда точно не удовлетворяется из-за погрешностей округления, и считается, что сам по себе метод CGA непригоден для решения конечно-разностных уравнений. Однако найдено, что он особенно эффективен в качестве

дополнения к методам, в которых матрица A близка к единичной, как, например, в методе ICSSG (разд. 6.4.6).

Общий алгоритм в понятном фортраноподобном представлении имеет вид

<i>Начальная приближенная оценка:</i>	<i>Число операций</i>
Φ_0	
EPS = требуемая точность	
DO 1 $j = 0, N_g - 1$	
<i>Невязка:</i>	
$r_j = A\Phi_j - q$	$10 N_g$
<i>Погрешность:</i>	
$\ e_j\ = r_j^T r_j$	$2 N_g$
<i>Досрочное окончание:</i>	
если ($\ e_j\ \leq EPS \times \ q\ $) задача решена	
<i>Направление поиска:</i>	
$s_0 = -r_0$	
$\beta_j = \ e_j\ / \ e_{j-1}\ $	$2 N_g$
$s_j = -r_j + \beta_j s_{j-1}$ } $j > 0$	
<i>Расстояние до минимума:</i>	
$\omega_j = \frac{\ e_j\ }{s_j^T A s_j}$	$11 N_g$
<i>Следующая итерация:</i>	
$\Phi_{j+1} = \Phi_j + \omega_j s_j$	$2 N_g$
1 CONTINUE	
Число операций на итерацию $27 N_g$	

Для решения задачи методом CGA нужен объем памяти $3N_g$ для хранения Φ_j , r_j и s_j . К этому нужно добавить память объемом $5N_g$ для матрицы A и N_g для правой части q . Количество операций при использовании трех сеток равно $27N_g$. Его можно уменьшить до $17N_g$, если мы введем четвертую сетку для хранения вектора $t_j = As_j$ во время расчета ω_j . Тогда следующую невязку можно найти из выражения

$$r_{j+1} = r_j + \omega_j t_j. \quad (6.63)$$

и отпадет необходимость вычислять $A\Phi_j$. Однако для большинства задач с уравнениями в частных производных и достаточно большими сетками эта дополнительная память будет неприемлема. Алгоритм сопряженных градиентов широко обсуждается Фаддеевым и Фаддеевой [129] и Рейдом [324]. Конкус, Голуб и О'Лири [87] полагают, что метод наиболее эффективен, когда его используют для ускорения существующих итераций.

6.4.3. Методы исключения для разреженных матриц (SM)

Есть много вариантов метода исключения Гаусса прямого решения систем уравнений, разработанных для матриц с большим числом нулей. Такие разреженные матрицы возникают при дискретизации дифференциальных уравнений в частных производных с помощью конечных разностей или конечных элементов, а также в теории управления, при анализе сетей энергоснабжения, в геологической съемке, в теории цепей и в структурном анализе. Отличный обзор дан Рейдом [325]. Применение пятиточечной конечно-разностной аппроксимации к уравнению Пуассона на сетке 128×128 приводит к $\sim 16\,000$ уравнениям, каждое из которых содержит пять переменных. Полная матрица содержит 256 миллионов элементов, однако только $\sim 80\,000$ (или $\sim 0.03\%$) из них отличны от нуля. Это является типичной ситуацией, для которой развиты методы разреженных матриц. В таких методах хранятся только ненулевые элементы и минимизируется образование дальнейших ненулевых элементов во время процесса исключения. Методика хранения обычно включает некоторую форму связного списка, в котором с каждым ненулевым элементом связан адресный указатель, указывающий на следующий ненулевой элемент в строке или столбце. Тогда вводить или исключать ненулевые элементы можно, изменяя значения этих связей, а не перемещая данные в памяти.

Общая задача заключается в нахождении решения уравнения

$$Ax = b, \quad (6.64)$$

где A есть $(N \times N)$ -матрица с элементами $a_{i,j}$. Разложение матрицы A методом исключения Гаусса проводится по этапам k , $k = 1, 2, \dots, N$. На каждом этапе из строк и столбцов матрицы, которые еще не разложены (т. е. тех, для которых $i, j \geq k$), выбирается ведущий элемент. Эти строки и столбцы переставляются, хотя необязательно перемещаются в памяти, для того чтобы перевести ведущий элемент на позицию $a_{k,k}$. Затем векторы, пропорциональные ведущей строке k , добавляются ко всем ниже расположенным строкам для того, чтобы исключить элементы столбца ниже ведущего, для чего используются соотношения

$$a_{i,j}^{(k+1)} = a_{i,j}^{(k)} - l_{i,k} a_{k,j}^{(k)}, \quad i, j > k, \quad (6.65)$$

$$b_i^{(k+1)} = b_i^{(k)} - l_{i,k} b_k^{(k)}, \quad i > k, \quad (6.66)$$

где

$$l_{i,k} = [a_{k,k}^{(k)}]^{-1} a_{i,k}^{(k)}$$

для $k = 1, 2, \dots, N$, начиная с $A^{(1)} = A$ и $b^{(1)} = b$. Этот процесс исключает элемент $a_{i,k}$ в каждой строке, но ясно, что он может

приводить к появлению в новой строке других ненулевых элементов, где их раньше не было. Это называется заполнением. Новые значения $a_{i,j}^{(k+1)}$ можно записывать на место $a_{i,j}^{(k)}$ и $l_{i,k}$ можно записывать на место $a_{i,k}^{(k)}$. Наконец, на этапе N треугольное разложение завершается. Множители $l_{i,k} = a_{i,j}^{(N)}$ ($i > j$) являются элементами нижней треугольной матрицы L с единичными диагональными элементами, которые не запоминаются, а $a_{i,j}^{(N)}$ ($i \leq j$) являются элементами верхней треугольной матрицы U , такой, что

$$\mathbf{A} = \mathbf{LU}. \quad (6.67)$$

Далее находим решение уравнения $\mathbf{Ax} = \mathbf{b}$, для чего проводим прямую подстановку с получением вспомогательного вектора \mathbf{y}

$$\mathbf{Ly} = \mathbf{b}, \quad (6.68)$$

за которой следует обратная подстановка

$$\mathbf{Ux} = \mathbf{y}, \quad (6.69)$$

которая дает решение \mathbf{x} .

В принципе любой ненулевой элемент редуцированной матрицы можно выбрать в качестве следующего ведущего элемента. Этот выбор определяют три соображения:

1. численная устойчивость,
2. минимизация заполненности,
3. минимизация числа арифметических операций.

Численная устойчивость гарантируется для симметричной положительно определенной матрицы, если мы проводим симметричные перестановки строк и столбцов в ходе исключения ведущих элементов. Для несимметричного случая важно не выбирать ведущий элемент, который слишком мал, поскольку тогда коэффициенты $l_{i,k}$ станут большими. В этом случае точность теряется, когда вектор, пропорциональный ведущей строке (теперь численно большой), добавляется к неведущей строке. Для предотвращения нежелательного роста величины элементов матрицы в процессе исключения ведущий элемент обычно выбирают так, чтобы он был больше некоторой доли (как правило, $1/4$) от максимального элемента в той же строке.

Чтобы свести к минимуму появление новых ненулевых элементов, или заполненность, требуется провести вычисления, которые трудно реализовать эффективно. Тем не менее критерий Марковица, который минимизирует число арифметических операций, существенно проще осуществить, и на практике он оказался ограничивающим заполненность почти так же успешно, как и процедура минимального заполнения (см. табл. 6.2). На этой таблице показано, что процедура минимального заполнения в шесть раз дороже по затратам машинного времени и

Таблица 6.2. Степень заполнения в процессе разложения.

(Воспроизводится из работы [325] с любезного разрешения издательства Academic Press, © Institute of Mathematics and its Applications)

Порядок матрицы			54	57	199	
Число ненулевых элементов в A			291	281	701	
Число ненулевых элементов в L/U	эле-	{	по критерию Марковица	381	315	1387
			минимальное заполнение	361	291	1372
Время работы программы ANALYZE (в мс на IBM 370/165)	эле-	{	по критерию Марковица	140	110	900
			минимальное заполнение	800	570	6400

лишь чуть-чуть лучше по уменьшению заполненности, чем критерий Марковица. В критерии Марковица в качестве ведущего выбирается тот элемент, для которого произведение числа других ненулевых элементов в его строке и числа других ненулевых элементов в его столбце минимально. Это значит, что в качестве ведущего берут элемент с минимальным значением выражения

$$\rho_{i,j} = (r_i - 1)(c_j - 1), \quad (6.70)$$

где r_i — число ненулевых элементов в строке i и c_j — число ненулевых элементов в столбце j .

Организация пакета программ для разреженных матриц, например программ MA18 [98] и MA28 из библиотеки Harwell, состоит из трех частей:

1. ANALYZE: анализируется структура матрицы для определения порядка выбора ведущих элементов, который будет численно устойчивым и уменьшит образование ненулевых элементов. Метод исключения Гаусса применяется к переставленной таким образом матрице с целью получения LU-разложения.

2. FACTORIZE: новая матрица с той же самой структурой, что и у предварительно проанализированной в (1), разлагается в произведение L и U .

3. OPERATE; по заданному разложению, полученному в (2), находится решение для одной правой части уравнения.

Для всякой полностью новой задачи необходимо сначала применить программу ANALYZE. Если далее задать другую матрицу с той же самой структурой (т. е. с теми же самыми положениями ненулевых элементов), что и у проанализированной ранее матрицы, то ее можно разложить с помощью программы FACTORIZE без повторного обращения к программе ANALYZE. Аналогично, если необходимо решить систему для

ряда различных правых частей с матрицей, которая уже была разложена, то необходимо вызвать только программу OPERATE. Важность такой разбивки пакета программ ясно видна из табл. 6.3, в которой показано, что использование программы ANALYSE в десять раз дороже, чем программы FACTORIZE, и что использование программы FACTORIZE в десять раз дороже, чем программы OPERATE.

Таблица 6.3. Относительная стоимость этапов анализа, разложения и решения для пятиточечной разностной аппроксимации уравнения Лапласа на сетке 31×31 . (Воспроизводится из работы [325] с любезного разрешения издательства Academic Press, © Institute of Mathematics and its Applications)

Порядок матрицы		930
Число запоминаемых ненулевых элементов	{ переменная полоса критерий Марковица	20 245
		11 328
Время на IBM 370/165 в мс	{ ANALYZE FACTORIZE	критерий Марковица 12 200
		переменная полоса 1300
	{ OPERATE	критерий Марковица 600
		переменная полоса 160
		критерий Марковица 120

Достоинство методов для разреженных матриц состоит в том, что с их помощью в принципе можно приступать к решению общего неразделяющегося дифференциального уравнения в частных производных с переменными коэффициентами прямыми методами и таким образом получать решение задачи вполне общего вида за фиксированное и известное время. Их практическая применимость существенно зависит от роста ненулевых элементов и соответствующих требований к объему памяти. Хотя мы знаем, что для уравнения Пуассона в квадратной области всегда следует использовать метод RES, его решение является полезной тестовой задачей с той же самой структурой, что и у дифференциального уравнения в частных производных с переменными коэффициентами наиболее общего вида. Поскольку время, затрачиваемое на решение методом для разреженных матриц, зависит от структуры матрицы, но не от численных значений ее элементов, то найденные при этом затраты машинного времени будут теми же, что для решения задачи с переменными коэффициентами, к которой нельзя применить методы RES (кроме итерации Конкуса и Голуба, разд 6 5.7).

Лучшим первоначальным упорядочением конечно-разностных уравнений является обход узлов, предложенный в [142]; оно называется «вложенным разбиением» (см. рис. 6.5). Если уравнения для всех сеточных узлов, образующих центральный крест (называемый первым разрезом), записываются в конце, остающиеся уравнения образуют четыре несвязанных системы для

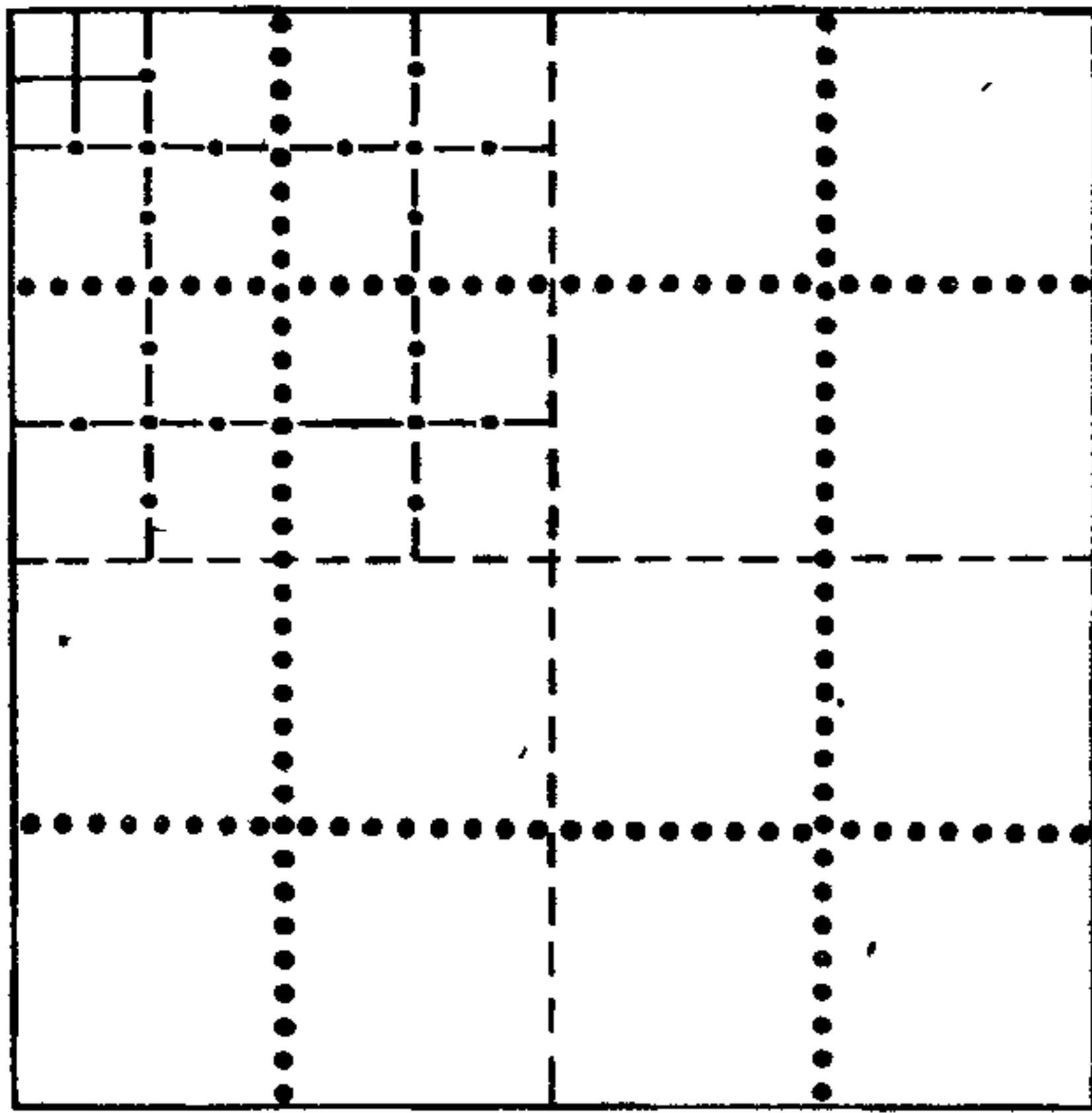


Рис. 6.5. Обход узлов при решении конечно-разностных уравнений методом вложенного разбиения Георга [142]. Первый разрез — — —, второй разрез . . ., третий разрез — . —, четвертый разрез ———.

внутренних узлов каждой четверти. Если вначале записаны все уравнения для верхней левой четверти, за ними — уравнения для верхней правой четверти, то для остальных уравнений получится блочная диагональная матрица. Уравнения в каждом из этих блоков мож-

но теперь переупорядочить в точности таким же образом, как были упорядочены исходные уравнения, введением второго разреза на рисунке. Процесс повторяется до тех пор, пока слева не окажется только набор всех нечетных узлов, и это будут первые $N_g/2$ уравнений. Процедура была обобщена Даффом и др. [111] без потери эффективности на сетки с любым количеством узлов. Таким образом, использованное в нашем примере неявное допущение, что число сеточных узлов является степенью двойки, необязательно. Кроме того, можно доказать, что обход узлов «вложенным разбиением» дает минимальное заполнение для об-суждаемой выше задачи.

Дафф и др. [111] сообщают следующие результаты для такого исходного упорядочивания, за которым следует применение критерия Марковица для последующего исключения. Для сравнения в квадратных скобках даются соответствующие показатели оптимального алгоритма RES [оптимальный алгоритм FACR (*), разд. 6.5.3].

Ненулевые элементы

$$\begin{array}{ll}
 \text{в } \mathbf{A} : 5N_g & [5], \\
 \text{в } \mathbf{L} = \mathbf{U}^T : 3N_g \log_2 N_g & [0], \\
 \text{в } \mathbf{b} : N_g & [N_g].
 \end{array} \quad (6.71)$$

Арифметические операции

$$\left. \begin{array}{l}
 \text{FACTORIZE: } 10N_g^{3/2} \\
 \text{OPERATE: } 12N_g \log_2 N_g
 \end{array} \right\} [4.5N_g \log_2(\log_2 N_g)]$$

Из этих результатов видно, что для заранее разложенной матрицы получение решения методами для разреженных матриц теоретически потребует для сеток в диапазоне от 32^2 до 128^2 в восемь—десять раз больше времени, чем при использовании

оптимального алгоритма FACR. В случае неоптимального алгоритма FACR(1) (например, программы POT1 [180]), для реализации которого требуется число операций, равное $1.25N_g(\log_2 N_g + 4.8)$, ожидаемый коэффициент равен 6.6. Хокни и Дафф выполнили фактическое сравнение программы POT1 для алгоритма FACR(1) и программы HARWELL MA17 для метода разреженной матрицы; его результаты приведены в табл. 6.4. Таблица показывает, что для сетки 32×32 различия между этими методами невелики, но для сеток, больших чем 64×64 , метод разреженной матрицы значительно медленнее, чем метод RES. Скорости счета приближаются к ожидаемому отношению, когда число сеточных узлов возрастает.

Таблица 6.4. Затраты машинного времени и объем памяти, требуемые для решения 5-точечного разностного уравнения Пуассона. Время в миллисекундах для программы на Фортране, пересчитано на параметры IBM 360/168

Сетка	Метод для разреженных матриц [†]		RES, POT1=FACR (1)		
	Этап OPERATE	Байты*	Решение	Байты [§]	
32×32	87	120К	71 [§] 66 ^{**}	71 [#]	4К
64×64	(438)	576К	192 140	223	16К
128×128	(2028)	3М	618 441	748	64К
256×256	(9270)	12М	[2472] [1764]	2491	256К

[†] И. С. Дафф, частное сообщение.

* $K = 1024$, $M = K^2$.

[§] В 2.7 раз больше времени, измеренного на CDC 7600.

[#] В 2.3 раз больше времени, измеренного на IBM 360/195.

** На IBM 360/168, язык Ассемблера.

() Экстраполировано по формуле $8.84 N_g \log_2 N_g$ мс.

[] Экстраполировано.

Основное различие между этими методами заключается в требуемой памяти, которая равна N_g для метода RES и $3N_g \log_2 N_g$ для методов, использующих разреженность матрицы. Если мы находим набор решений для различных правых частей уравнения и хотим производить вычисления по сути дела как при физическом нестационарном моделировании, то это обстоятельство само по себе ограничит использование методов

для разреженных матриц относительно малыми сетками, например (32×32) , тогда как метод RES стандартно применяется к сеткам (256×256) с $N_g \cong 65\,000$.

Поскольку для решения задачи с переменными коэффициентами с помощью процедуры Конкуса и Голуба потребовалось бы по крайней мере пять итераций для алгоритма RES, то для задачи с переменными коэффициентами предпочтительнее метод разреженной матрицы с обходом, определяемым вложенным разбиением, при условии, что левая часть уравнения остается той же самой и памяти хватает. Если левая часть уравнения меняется, то этап разложения также должен выполняться заново каждый раз. Поскольку этот этап в десять — двадцать раз медленнее, чем итерация RES, то в этом случае предпочтительнее итерация по Конкусу и Голубу или методы неполного разложения.

Для больших сеток ($N_g \gg 1000$) методы разреженных матриц, вероятно, должны быть непрактичными из-за огромной памяти, требуемой для L и U . Однако решение посредством метода RES с использованием итерации по Конкусу и Голубу является вполне осуществимым по памяти ($2N_g$) и требуемому времени (в 5 раз больше времени решения методом RES).

В трехмерном случае оцененные количества операций и память для метода вложенного разбиения таковы:

операции в FACTORIZE: $4N_g^2$,
 ненулевые элементы в L : $28N_g^{4/3}$,
 операции в OPERATE: $112N_g^{4/3}$.

Эти цифры в 100 раз больше, чем в итерации RES или ICCG, и в десять раз больше, чем в решении методом сеточной релаксации. Поэтому представляется, что методы разреженных матриц неконкурентоспособны в трехмерном случае даже при тщательном выборе порядка обхода узлов.

6.4.4. Неполное разложение

Из уравнения (6.71) можно видеть, что если требуется точное (или полное) разложение разреженной матрицы A в произведение LU , то количество ненулевых элементов возрастает приблизительно, в $(3/5) \log_2 N_g$ раз. Альтернативный подход состоит в построении приближенного (неполного) LU -разложения, столь же разреженного, как и исходная матрица A , и затем в нахождении точного разложения с помощью итерации.

Приближенное разложение матрицы A будет точным разложением матрицы P , близкой к A :

$$P = A + N = LU. \quad (6.72)$$

Здесь матрица \mathbf{N} должна быть малой, чтобы метод достиг цели. Итерационное решение уравнения

$$\mathbf{A}\varphi = \rho \quad (6.73)$$

определяется соотношением

$$(\mathbf{A} + \mathbf{N})\varphi^{(t+1)} = \mathbf{N}\varphi^{(t)} + \rho$$

или

$$\mathbf{LU}\varphi^{(t+1)} = \mathbf{N}\varphi^{(t)} + \rho. \quad (6.74a)$$

Поскольку \mathbf{LU} является такой же разреженной, как и \mathbf{A} , то решение для следующей итерации $\varphi^{(t+1)}$ будет простым и быстрым.

Имеет смысл выразить (6.74a) через вектор невязки

$$\mathbf{r} = \mathbf{A}\varphi - \rho \quad (6.74b)$$

и решать уравнение для изменения в решении $\Delta\varphi$. Тогда итерация принимает вид

$$\mathbf{L}\mathbf{y} = -\mathbf{r}, \quad (6.74c)$$

$$\mathbf{U}\Delta\varphi = \mathbf{y} \quad (6.74d)$$

и

$$\varphi^{(t+1)} = \varphi^{(t)} + \Delta\varphi. \quad (6.74e)$$

Поскольку $(\mathbf{LU})^{-1}$ является приближенным обращением для решения на всей области, то можно ожидать существенно более быструю сходимость, чем в методе SOR, в котором используется только точечное или построчное обращение.

Мы рассматриваем два метода, основанных на этой идее: строго неявную процедуру, принадлежащую Стоуну [356], в которой используется только что описанная итерация, и метод неполного разложения по Холецкому, принадлежащий Мейеринку и Ворсту [280], в котором применяется итерация методом сопряженных градиентов. В другом родственном алгоритме, примененном Петравиком и Куо — Петравиком [311], неполное LU-разложение объединяется с итерацией, использующей сопряженные градиенты.

6.4.5. Строго неявная процедура Стоуна (SIP)

На рис. 6.6 показан вид матриц \mathbf{A} , \mathbf{P} , \mathbf{L} и \mathbf{U} в методе SIP для случая пятиточечной разностной схемы. \mathbf{A} является пятидиагональной матрицей с произвольными элементами, возникающими из дифференциального уравнения в частных производных общего вида, неразделяемого, с переменными коэффициентами, переменными сеточными шагами и границами произвольной формы. На сетке $n \times n$ ненулевые элементы расположены на главной диагонали и непосредственно над и под ней. Кроме того, имеются еще две диагонали на расстоянии n от

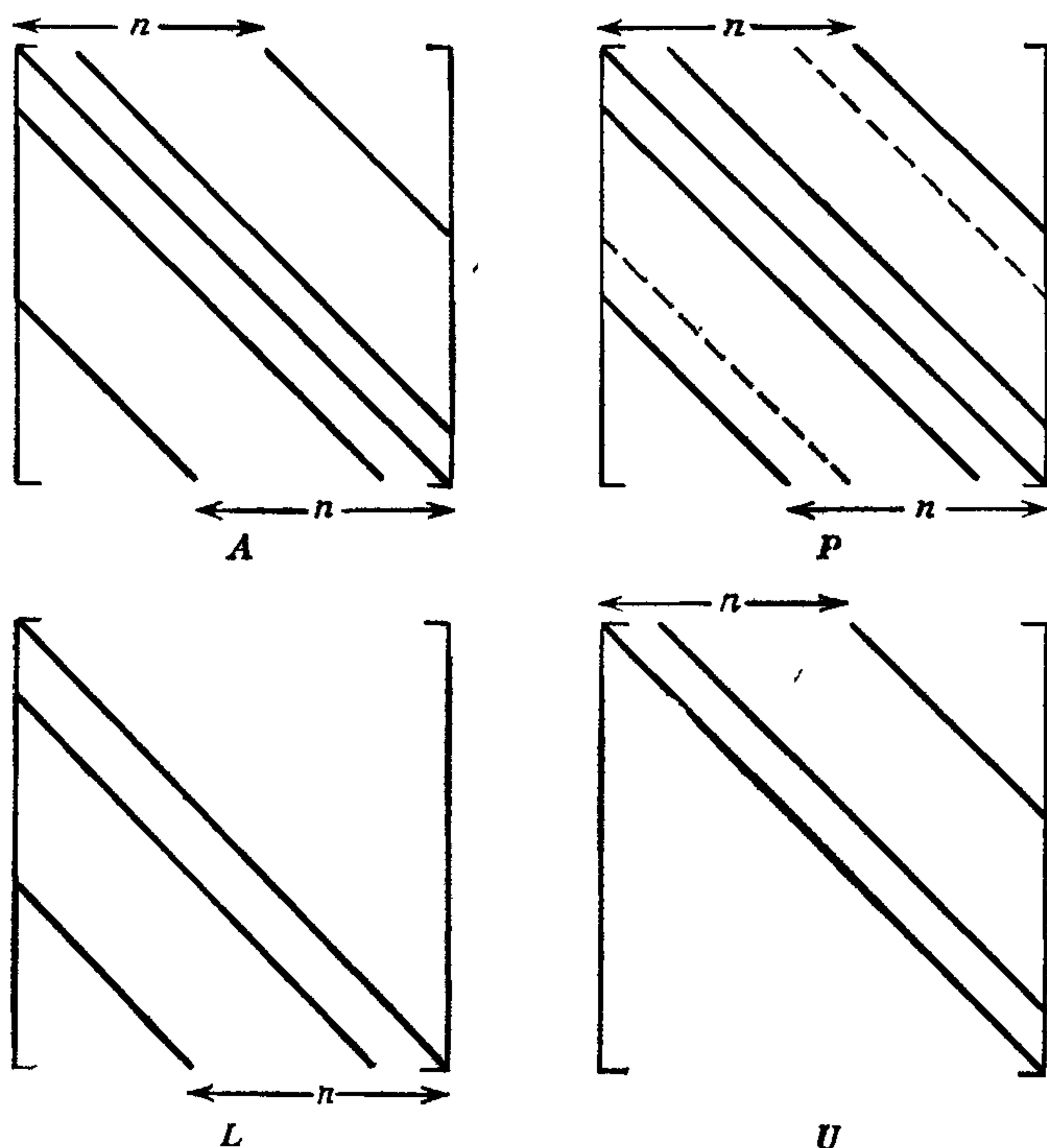


Рис 66 Структура матрицы A и ее приближенных разложений L , U P является произведением матриц LU ; штрихами показаны нежелательные элементы, которые не соответствуют элементам матрицы A . Матрица A выводится из пятиточечного разностного уравнения Пуассона на сетке $(n \times n)$. Все элементы, не помеченные линиями, равны нулю

главной. Верхняя и нижняя треугольные матрицы L и U определяются так, чтобы элементы имелись только на тех позициях, на которых имеются элементы в матрице A , так что и L , и U имеют только $3N_g$ элементов. С другой стороны, точное разложение матрицы A имело бы $\sim N_g^{3/2}$ элементов, поскольку матрица должна быть заполненной в направлении от внешней диагонали, показанной на рис 6.6, к главной диагонали. Произведение матриц L и U , имеющих эту структуру, — это матрица P , которая в дополнение к элементам, соответствующим элементам матрицы A , имеет еще две лишние диагонали, показанные штриховыми линиями. Они лежат непосредственно рядом с внешними диагоналями. Эти нежелательные элементы матрицы можно интерпретировать как соответствующие несуществующим членам вида $\partial^2/\partial x \partial y$ в исходных уравнениях.

Если коэффициенты в пяти диагоналях матрицы A (см. рис 6.6), соответствующие сеточному узлу (p, q) , равны $(B_{p,q}, D_{p,q}, E_{p,q}, F_{p,q}, H_{p,q})$, то элементы $(b_{p,q}, c_{p,q}, d_{p,q})$ матрицы

L и $(1, e_{p,q}, f_{p,q})$ матрицы U вычисляются по следующим формулам:

$$\begin{aligned} b_{p,q} &= B_{p,q}/(1 + \alpha e_{p,q-1}^*), \\ c_{p,q} &= D_{p,q}/(1 + \alpha f_{p-1,q}^*), \\ d_{p,q} &= E_{p,q} + b_{p,q}(\alpha e_{p,q-1}^* - f_{p,q-1}^*) + c_{p,q}(\alpha f_{p-1,q}^* - e_{p-1,q}^*), \\ e_{p,q} &= (F_{p,q} - \alpha b_{p,q} e_{p,q-1}^*)/d_{p,q}, \\ f_{p,q} &= (H_{p,q} - \alpha c_{p,q} f_{p-1,q}^*)/d_{p,q}. \end{aligned} \quad (6.75)$$

Здесь величины, отмеченные звездочкой, уже известны из вычислений для более ранних значений p или q , а прописные буквы обозначают известные элементы матрицы A .

Параметр α служит итерационным параметром, который должен изменяться циклически в диапазоне от 0 до 1. Значение α_{\max} можно определить из соотношения

$$(1 - \alpha_{\max}) = \min \left[\frac{2 \Delta x^2}{1 + \frac{DY \Delta x^2}{DX \Delta y^2}}, \frac{2 \Delta y^2}{1 + \frac{DX \Delta y^2}{DY \Delta x^2}} \right], \quad (6.76)$$

где $\Delta x, \Delta y$ — сеточные шаги по x и y , а DX и DY — коэффициенты диффузии в исходном уравнении

$$\frac{\partial}{\partial x} DX \frac{\partial}{\partial x} \varphi + \frac{\partial}{\partial y} DY \frac{\partial}{\partial y} \varphi = q. \quad (6.77)$$

Если α_{\max} вычисляется для каждого узла сетки из (6.76), то среднее значение по всей сетке α_{\max}^* можно использовать для определения цикла величин α , каждая из которых используется в последовательности дважды. Если длина цикла равна M , то значения α образуют геометрическую прогрессию и лежат в диапазоне $0 \leq \alpha \leq \alpha_{\max}^*$; они вычисляются из выражения

$$1 - \alpha_m = (1 - \alpha_{\max}^*)^{m/(M-1)}, \quad m = 0, 1, \dots, M-1. \quad (6.78)$$

Обычно используются значения m , изменяющиеся от 4 до 9.

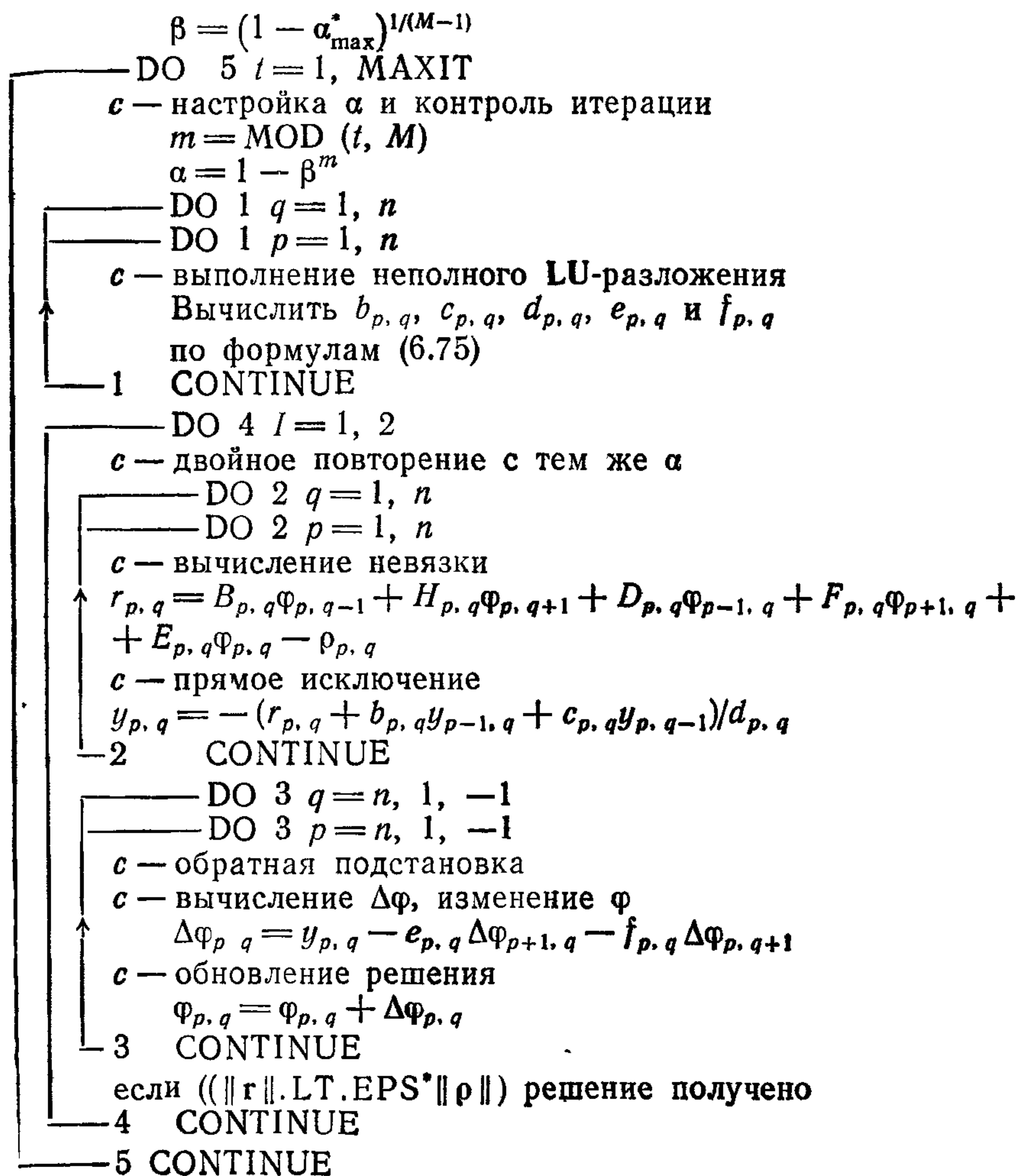
Стоун сообщает об удовлетворительной сходимости для широкого класса задач, включая некоторые задачи с большой асимметрией в коэффициентах диффузии, например $DX/DY = 100$.

Чтобы получить снижение погрешности до $10^{-5} - 10^{-6}$, достаточно провести от 20 до 30 итераций по процедуре SIP. Стоун [356] утверждает, что число итераций, требуемых для решения, нечувствительно к количеству сеточных узлов, сложности граничных условий и к изменению коэффициентов исходных уравнений. Для сравнения заметим, что число итераций, требуемых для решения с заданной точностью методом SOR, пропорционально числу сеточных узлов в каждом направлении (см. разд. 6.3.3).

Коэффициенты LU-разложения можно получить из (6.75) за $17N_g$ арифметических операций, и прямой и обратной подстановкой можно получить новую итерацию из уравнения (6.74) еще за $19N_g$ операций. Поскольку используется то же самое значение α и, следовательно, то же самое LU для двух итераций, то на итерацию требуется в среднем полное число операций, равное $27N_g$.

В дополнение к памяти, необходимой для формулировки задачи (хранение A и p), требуется память объемом $5N_g$ для LU-разложения и $2N_g$ для $\varphi^{(t)}$ и $\Delta\varphi$.

Программа для метода Стоуна включает в себя три цикла DO, как можно видеть на следующем наброске:



Действующая программа для строго неявной процедуры была опубликована Джесхопом [222]. Программа SIPSOL имеется в библиотеке Computer Physics Communications Белфастского университета и содержит набор подпрограмм для решения линейных уравнений, возникающих из дифференциальных уравнений в частных производных.

6.4.6. Метод неполного разложения Холецкого — сопряженных градиентов (ICCG)

Для симметричной и положительно определенной матрицы A Мейеринк и Ворст [280] предложили другое неполное разложение. Оно основано на разложении Холецкого для симметричной матрицы с последующей итерацией методом сопряженных градиентов, которая естественна для решения положительно определенной системы

Схемы разрежения, определенные для матриц L и U , те же, которые использовались в методе SIP. Однако элементы матрицы вычисляются иначе. Поскольку A симметрична, $U = L^T$, и нужно вычислять одну только матрицу L .

Разложение Холецкого для симметричной матрицы A можно записать в форме, которая не требует вычисления квадратного корня, введением диагональной матрицы D , такой, что

$$A = L D L^T. \quad (6.79)$$

Полное разложение приводит к ленточной нижнетреугольной матрице L с полушириной полосы n и $\sim N_g^{3/2}$ элементами. Для ее расчета требуется $\sim N_g^2$ операций. При неполном разложении мы добиваемся той же самой разреженности матрицы L , которая существует в матрице A , и рассчитываем только $3N_g$ элементов, соответствующих трем диагоналям в нижней части матрицы A . В дальнейшем L обозначает такое неполное разложение

Мы обозначаем элементы матрицы A через a_i , b_i и c_i , где a_i — главная диагональ, b_i представляет ближайшую верхнюю диагональ, и c_i является n -й верхней диагональю. Элементы транспонированной матрицы L^T в позициях, соответствующих a_i , b_i и c_i , обозначаются через \tilde{a}_i , \tilde{b}_i и \tilde{c}_i , а элементы матрицы D обозначаются через \tilde{d}_i . Тогда неполное разложение проводится рекуррентно по формулам

$$\begin{aligned} \tilde{b}_i &= b_i, & \tilde{c}_i &= c_i, \\ \tilde{a}_i &= a_i - \tilde{b}_{i-1}^2 \tilde{d}_{i-1} - \tilde{c}_{i-n}^2 \tilde{d}_{i-n}, & i &= 1, 2, \dots, N_g, \\ \tilde{d}_i &= \tilde{a}_i^{-1}, \end{aligned} \quad (6.80)$$

причем все остальные элементы заменяются нулями.

Формулы (6.80) показывают, что в отличие от метода Стоуна, недиагональные элементы матрицы L (и, следовательно, матрицы L^T) те же самые, что недиагональные элементы матрицы A . Дополнительная память требуется лишь для новых диагональных элементов \tilde{d}_i .

Для того чтобы найти точное решение, мы перепишем уравнение

$$A\varphi = \rho \quad (6.81a)$$

в виде

$$B y = D^{-1} L^{-1} \rho, \quad (6.81b)$$

$$L^T \varphi = y, \quad (6.81c)$$

где

$$B = [D^{-1} L^{-1} A (L^T)^{-1}]. \quad (6.81d)$$

Алгоритм сопряженных градиентов используется теперь для решения уравнения (6.81b) (см. разд. 6.4.2). Поскольку $A \approx \approx LDL^T$, то B приближенно является единичной матрицей. Следовательно, характеристические числа матрицы B группируются вокруг единицы, и квадратичная поверхность, на которой нужно найти минимум, почти сферическая. При этом все направления наискорейшего спуска устремлены к минимуму, и метод CGA сходится очень быстро.

Количество операций, необходимых для вычисления приближения, равно $6N_g$. За этим следует $27N_g$ операций на каждую итерацию, требуемую для алгоритма CGA. Во время итераций CGA матрица B остается неизменной, так что число затраченных операций в основном определяется процедурой CGA. Результаты, представленные Мейеринком и Ворстом [280], показывают, что количество операций, требуемых для снижения погрешности до 10^{-6} , приблизительно равно $1,3n$. Это та же самая зависимость от числа узлов сетки, которая найдена в оптимальном методе SOR.

Во время решения нужна память, равная $4N_g$, из них $3N_g$ для хранения векторов алгоритма CGA и N_g для диагональных элементов при неполном разложении. Эта память на $4N_g$ меньше, чем в процедуре Стоуна, что связано со специальным видом матрицы (ее симметричностью) и с тем фактом, что недиагональные элементы матрицы L расположены так же, как и у матрицы A .

Теперь мы видим, что важное отличие любых матричных методов (разрежения или неполного разложения) от сеточно-итерационных методов, таких как SOR, ADI и метод Конкуса и Голуба, состоит в требованиях к объему памяти. Первые методы требуют от трех до семи сеток для временного хранения в процессе счета по сравнению только с одной дополнительной сеткой, необходимой для последних методов (см. табл. 6.1).

Это означает, что при заданном объеме памяти с помощью матричных методов можно получить меньшее количество сеточных узлов (и, следовательно, меньшее пространственное разрешение), чем с помощью сеточных методов. По этой причине матричные методы с ростом N_g становятся все менее практичными при $N_g > 1000$.

6.5. БЫСТРОЕ РЕШЕНИЕ ЭЛЛИПТИЧЕСКИХ УРАВНЕНИЙ (RES)

Мы используем термин «быстрое решение эллиптических уравнений» для описания класса прямых методов, которые получают точное решение системы из N_g дифференциальных уравнений за $N_g \log_2 N_g$ (или меньше) операций и требуют память объемом в N_g . Эти методы настолько специальные, что они не позволяют непосредственно решить общее дифференциальное уравнение в частных производных второго порядка с переменными коэффициентами и с границей, имеющей произвольную форму. Однако с их помощью достаточно широкий круг важных частных случаев таких задач решается существенно быстрее, чем более общими методами, и, следовательно, о них следует сказать особо. Поскольку они вне конкуренции для задач, к которым они приложимы, то использовать в таких случаях более общие методы нет смысла. Методы RES основаны на циклической редукции (CR) и быстром преобразовании Фурье (FFT) или на их комбинации (алгоритм FACR). Сравнительные теоретические обзоры алгоритмов RES опубликовали Дорр [108], Шварцтраубер [359] и Хокни [186], а сравнительные характеристики машинных программ представлены Хокни [183], Темпертоном [365] и Шуманом [344]. В последней работе сообщается о состязании, проведенном между двадцатью программами, которые использовались для решения уравнения Пуассона на сетке 32×32 с граничными условиями Дирихле и Неймана. Программы прогонялись в одинаковых условиях на IBM 370/168 с использованием одного и того же транслятора. Состязание выиграла оптимизированная программа FACR(1) пакета PSOLVE Темпертона [366].

Первым появившимся алгоритмом RES был метод FACR, предложенный Хокни [175], который предшествовал даже быстрому преобразованию Фурье (FFT) [88]. Метод был обобщен на произвольное число уровней циклической редукции и связан с FFT Хокни [180]. Существование оптимального числа уровней редукции эмпирически отметил Хокни [180] и затем аналитически показал Шварцтраубер [359]. Этот оптимизированный алгоритм FACR также широко изучался на практике Темпертоном [366]. Реализация алгоритма FACR на очень больших сет-

ках 2D и 3D, например $128 \times 128 \times 128$, также рассмотрена Темпертоном [364]. Алгоритм FACR описывается в разд. 6.5.3 и применяется начиная с 1965 г. при моделировании плазмы (гл. 9), электронных устройств (гл. 10) и взаимодействия вихрей [77—78].

Вторым алгоритмом RES, по-видимому, был стабилизированный метод циклической редукции Бунемана [58], описанный в разд. 6.5.1. Первоначально этот метод был предложен для уравнения Пуассона с количеством сеточных узлов, равным степени числа два. Он был обобщен на общее разделяющееся уравнение Шварцтраубером [358] и на произвольное число сеточных узлов Свитом [361—362]. Теорию и анализ устойчивости предложили Базби, Голуб и Нильсон [72].

После опубликования быстрого преобразования Фурье в 1965 году многие авторы (см., например, [42], [244]) пользуются методом преобразования Фурье (разд. 6.5.2) для быстрого решения уравнения Пуассона и ему подобных уравнений в прямоугольных областях, в особенности когда амплитуды гармоник интересны сами по себе. Метод преобразования Фурье также (через теорему о свертке) приложим к произвольному потенциалу взаимодействия и к изолированному распределению частиц (разд. 6.5.4). В случае кулоновского взаимодействия более экономичный метод для изолированных систем предложил Джеймс [217]; он описан в разд. 6.5.5.

В другом классе алгоритмов RES используются процедуры, которые являются численно неустойчивыми, но тем не менее могут быть успешно реализованы на компьютерах с очень большой длиной слова (например, CDC 7600 и CRAY-1). Примерами служат многопроходные процедуры [269, 22, 21, 126] и метод точечной циклической редукции [107]. В этих методах за счет излишней арифметической точности машинного представления числа достигается повышенная скорость получения решения, и в случае их применимости они могут оказаться самыми быстрыми из всех известных методов [186]. Они привлекли внимание еще и потому, что их часто можно применять непосредственно к сложным неразделяющимся задачам. Тем не менее они не описываются далее в настоящей книге из-за того, что их можно реализовать лишь на определенных компьютерах.

Алгоритмы RES первоначально были развиты для простой геометрии области, такой как квадрат или прямоугольник. Однако для задач, включающих границы произвольной формы, Хокни [179, 184] использовал методику матрицы емкости (разд. 6.5.6), которая была усовершенствована для большего числа граничных узлов Видландом [385] и Проскуровским и Видландом [317]. Мартин [275] распространил метод на более общие граничные условия и использовал его при решении аэродинамических задач. Для области решения, представляющей

собой объединение прямоугольников, известны методы сращения решений, полученных алгоритмами RES в каждом из составляющих ее прямоугольников [71].

6.5.1. Циклическая редукция (CR)

Циклическую редукцию, устойчивый вариант которой предложил Бунеман [58], можно использовать для решения произвольного разделяющего уравнения вида

$$a(x) \frac{\partial^2 \varphi}{\partial x^2} + b(x) \frac{\partial \varphi}{\partial x} + c(x) \varphi + d(y) \frac{\partial^2 \varphi}{\partial y^2} + e(y) \frac{\partial \varphi}{\partial y} + f(y) \varphi = g(x, y) \quad (6.82)$$

при условии, что форма области и граничные условия являются «простыми». Шварцтраубер и Свит [360] опубликовали набор программ на языке Фортран для решения уравнений этого типа. Решение дифференциального уравнения в частных производных можно представить в виде линейной комбинации собственных значений левой части оператора. В случае разделяющегося уравнения общего вида эти собственные векторы являются произведениями вида $U_n(x, y) = X_n(x) Y_n(y)$, что можно продемонстрировать подстановкой, причем функции $X_n(x)$, $Y_n(y)$ удовлетворяют для заданного собственного значения n^2 уравнениям

$$a(x) \frac{d^2 X_n}{dx^2} + b(x) \frac{dX_n}{dx} + c(x) X_n = n^2 X_n, \quad (6.83a)$$

$$d(y) \frac{d^2 Y_n}{dy^2} + e(y) \frac{dY_n}{dy} + f(y) Y_n = -n^2 Y_n. \quad (6.83b)$$

После несложных преобразований общая разделяющаяся форма уравнений включает уравнение Гельмгольца в стандартных системах координат. Например:

декартова (x, y)

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \lambda \varphi = q, \quad (6.84a)$$

полярная (r, θ)

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \varphi}{\partial \theta^2} + \lambda \varphi = q, \quad (6.84b)$$

цилиндрическая осесимметрическая (r, z)

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \varphi}{\partial r} \right) + \frac{\partial^2 \varphi}{\partial z^2} + \frac{\lambda}{r^2} \varphi = q, \quad (6.84c)$$

сферическая осесимметрическая (r, θ)

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \varphi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \varphi}{\partial \theta} \right) + \frac{\lambda}{r^2 \sin^2 \theta} \varphi = q, \quad (6.84d)$$

сферическая (на сфере) (θ, φ)

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \cdot \frac{\partial u}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \varphi^2} + \lambda u = q. \quad (6.84e)$$

Общая форма уравнений с разделяющимися переменными включает задачу с переменным коэффициентом диффузии вида

$$\frac{\partial}{\partial x} D(x) \frac{\partial \varphi}{\partial x} + \frac{\partial}{\partial y} E(x) \frac{\partial \varphi}{\partial y} + \lambda \varphi = g(x, y). \quad (6.85a)$$

но, к сожалению, не включает часто встречающуюся задачу более общего вида

$$\frac{\partial}{\partial x} D(x, y) \frac{\partial \varphi}{\partial x} + \frac{\partial}{\partial y} E(x, y) \frac{\partial \varphi}{\partial y} + \lambda \varphi = q. \quad (6.85b)$$

Теперь мы объясним, что мы имеем в виду под «простой» областью и «простыми» граничными условиями. Решение ищется во внутренней части пространства, определенной двумя парами значений для независимых переменных. Например, в декартовых координатах прямоугольник определен граничными значениями координат (x_1, y_1) , (x_2, y_2) ; в полярных координатах сектор кольца определен значениями (r_1, θ_1) и (r_2, θ_2) . Граничные условия, допускаемые на этих поверхностях постоянных координат, представляют какую-либо комбинацию следующих условий:

1. Задано значение искомого решения φ . (6.86a)
2. Смешанное условие, например

$$\alpha \varphi(x, y) + \frac{\partial}{\partial x} \varphi(x, y) = p(y) \quad (6.86b)$$

вдоль $x = a$, где $p(y)$ задано.

Если $\alpha = 0$, мы имеем условие с заданным градиентом искомого решения.

3. Периодичность. (6.86c)

Если общее дифференциальное уравнение в частных производных с разделяющимися переменными переводится в конечно-разностное, то получается набор алгебраических уравнений вида

$$l_j \varphi_{j-1} + \Lambda_j \varphi_j + m_j \varphi_{j+1} = h_j, \quad (6.87a)$$

где φ_j и h_j — вектор решения и вектор правой части на сеточной строке j . Матрица Λ_j является трехдиагональной с переменными коэффициентами и имеет вид

$$\Lambda_j = \mathbf{B} + p_j \mathbf{I}. \quad (6.87b)$$

Разделение переменных в уравнении гарантирует, что трехдиагональная матрица \mathbf{B} , представляющая дифференцирование по x вдоль строки, одна и та же для всех строк, и что коэффициенты l_j , p_j , m_j , представляющие дифференцирование по y

между строками, одни и те же для всех переменных внутри строки. Следовательно, разностные уравнения можно записать в виде построчных уравнений (6.87а).

Шварцтраубер [358] предложил метод циклической редукции для общего разделяющегося уравнения (6.87а). Подробности слишком обширны, чтобы излагать их здесь, поэтому мы проиллюстрируем принцип метода на примере простого случая, когда коэффициенты уравнений не зависят от y , чтобы можно было опустить индекс j при l , m , p и, следовательно, при Λ . Заметим, что случай $l = m = 1$, $p = -2$ охватывает любое дифференциальное уравнение в частных производных, дифференциальный оператор которого в одном координатном направлении состоит лишь из одной второй производной. Это включает уравнения Пуассона и Гельмгольца в декартовых, полярных, цилиндрических и сферических поверхностных координатах (ур. 6.84а, б, с и е).

Выбирая три последовательных уравнения вроде уравнения (6.87а), имеем

$$l\varphi_{j-2} + \Lambda\varphi_{j-1} + m\varphi_j = h_{j-1}, \quad (6.88a)$$

$$l\varphi_{j-1} + \Lambda\varphi_j + m\varphi_{j+1} = h_j, \quad (6.88b)$$

$$l\varphi_j + \Lambda\varphi_{j+1} + m\varphi_{j+2} = h_{j+1}. \quad (6.88c)$$

Умножая эти три уравнения на l , $-\Lambda$ и m соответственно, а затем складывая, получаем уравнение, включающее значения φ только с каждой второй строки сетки,

$$l^{(1)}\varphi_{j-2} + \Lambda^{(1)}\varphi_j + m^{(1)}\varphi_{j+2} = h_j^{(1)}, \quad (6.89a)$$

где

$$\Lambda^{(1)} = 2ml\Lambda - \Lambda^2, \quad (6.89b)$$

$$h_j^{(1)} = lh_{j-1} - \Lambda h_j + mh_{j+1}, \quad (6.89c)$$

$$l^{(1)} = l^2, \quad m^{(1)} = m^2. \quad (6.89d)$$

Итак, мы завершили один уровень циклической редукции, вдвое снизив число уравнений. К тому же три уравнения (6.89) совпадают по виду с уравнением (6.87), и весь процесс редукции можно повторить до тех пор, пока не останется только одно уравнение для переменных центральной строки:

$$\Lambda^{(t)}\varphi_{n/2} = h_{n/2}^{(t)} - l^{(t)}\varphi_0 - m^{(t)}\varphi_n. \quad (6.90)$$

В случае когда число узлов по y является степенью числа два, φ_0 и φ_n будут известными граничными значениями, так что уравнение (6.90) можно решить для переменных, расположенных в центральной строке $\varphi_{n/2}$. Связь между известными и вычисленными переменными на каждом уровне редукции показана на рис. 6.7.

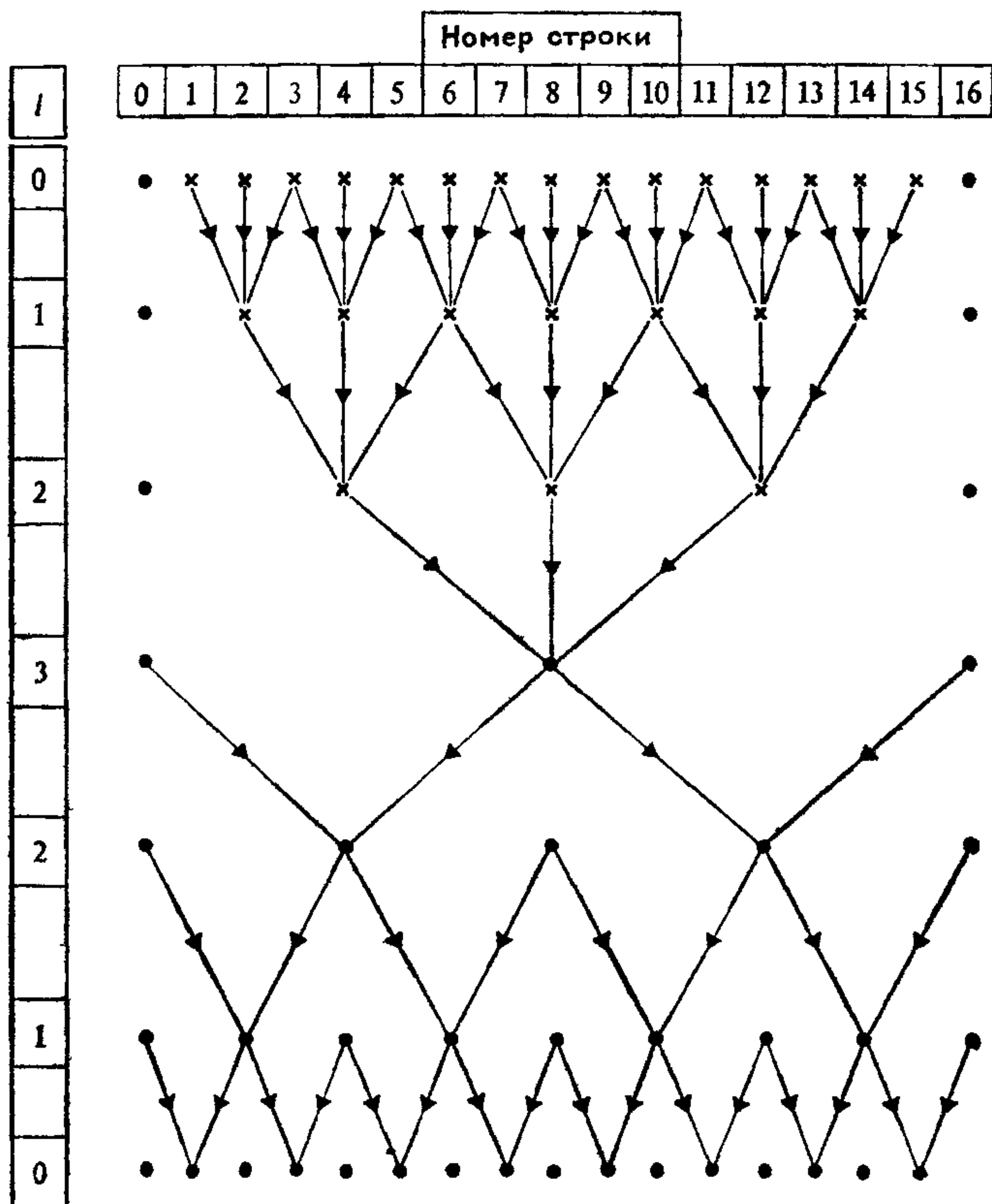


Рис. 67 Диаграмма, показывающая строки величин, вычисляемых на каждом уровне циклической редукции. Стрелки указывают данные, которые вносят вклад в новые значения. Известные граничные значения и значения решения помечены сплошными кружками.

Матрицу $\Lambda^{(t)}$ можно представить как произведение множителей, линейно зависящих только от исходной трехдиагональной матрицы $\Lambda^{(0)} = \Lambda$:

$$\Lambda^{(t)} = - \prod_{k=1}^{2t} (\Lambda - \beta_k I). \quad (6.91a)$$

Для случая $l = m = 1$ корни β_k находятся аналитически и равны

$$\beta_k = 2 \cos \left[\frac{2(k-1)\pi}{2^{t+1}} \right]. \quad (6.91b)$$

В общем случае разделяющихся переменных корни не находятся аналитически, но могут быть заранее найдены численно. В любом случае уравнение (6.90) можно решить путем после-

довательного решения 2^t трехдиагональных систем, используя алгоритм Томаса (см разд. 6.4.1). Количество требуемых уровней редукции будет равно $t = \log_2 n$ в случае периодических граничных условий или $t = \log_2 n - 1$ в случае заданных граничных значений

Для того чтобы получить неизвестные значения $\varphi_{n/4}$ и $\varphi_{3n/4}$, мы используем уравнения, полученные на уровне $(t-1)$, где только что вычисленное значение $\varphi_{n/2}$ можно перенести в правую часть:

$$\Lambda^{(t-1)}\varphi_{n/4} = h_{n/4}^{(t-1)} - l\varphi_0^{(t-1)} - m\varphi_{n/2}^{(t-1)}, \quad (6.92)$$

и аналогично для $\varphi_{3n/4}$. Таким способом можно последовательно отыскать промежуточные значения решения

Вычисление правых частей из равенства (6.89с) численно неустойчиво, и вместо этого необходимо использовать приведенную ниже схему, предложенную Бунеманом [58]. Правая часть на уровне r находится из выражения

$$\mathbf{h}^{(r)} = \Lambda^{(r)}\mathbf{p}_j^{(r)} + \mathbf{q}_j^{(r)}, \quad (6.93)$$

где для случая $l = m = 1$ векторы \mathbf{p}_j и \mathbf{q}_j подчиняются рекуррентным соотношениям

$$\begin{aligned} \mathbf{p}_j^{(r)} &= \mathbf{p}_j^{(r+1)} + (\Lambda^{(r-1)})^{-1} (\mathbf{q}_j^{(r-1)} - \mathbf{p}_{j-2^{r-1}}^{(r-1)} - \mathbf{p}_{j+2^{r-1}}^{(r-1)}), \\ \mathbf{q}_j^{(r)} &= \mathbf{q}_{j-2^{r-1}}^{(r-1)} + \mathbf{q}_{j+2^{r-1}}^{(r-1)} - 2\mathbf{p}_j^{(r)}, \\ \mathbf{p}_j^{(0)} &= 0 \quad \text{и} \quad \mathbf{q}_j^{(0)} = \mathbf{h}_j. \end{aligned} \quad (6.94)$$

Асимптотически число операций для циклической редукции с $l = m = 1$ равно $\sim 6n^2 \log_2 n$ для сетки $(n \times n)$ или $3N_g \log_2 N_g$. Требуемая память равна N_g . Это число операций увеличивается до $16N_g \log_2 N_g$ для общего случая разделения переменных для уравнения с переменными коэффициентами. Для решения задачи с переменными коэффициентами на сетке 127×127 требуется 13 с на CDC 7600, согласно Шварцтрауберу и Свиту [360]. Это время следует сравнить с временем, равным 0.23 с для решения уравнения Пуассона с постоянными коэффициентами посредством алгоритма FACR (см разд. 6.5.3 и [183]). Обе программы написаны на языке Фортран.

6.5.2. Метод преобразования Фурье (MFT)

Общее дифференциальное уравнение в частных производных второго порядка с постоянными коэффициентами и периодическими граничными условиями

$$a \frac{\partial^2 \varphi}{\partial x^2} + b \frac{\partial \varphi}{\partial x} + c\varphi + d \frac{\partial^2 \varphi}{\partial y^2} + e \frac{\partial \varphi}{\partial y} + f\varphi = g(x, y) \quad (6.95)$$

можно решить, используя периодическое быстрое преобразование Фурье (FFT). Если сеточные значения неизвестных $\varphi_{p,q}$ и правые части $g_{p,q}$ разложить в конечные двойные ряды Фурье

$$\begin{aligned}\varphi_{p,q} &= \sum_{k,l} \hat{\varphi}_{k,l} e^{2\pi i (pk+ql)/n}, \\ g_{p,q} &= \sum_{k,l} \hat{g}_{k,l} e^{2\pi i (pk+ql)/n}\end{aligned}\quad (6.96)$$

и подставить в обычную форму уравнения (6.95) с центральными конечными разностями, на равномерной сетке получаются уравнения для n^2 амплитуд гармоник $\hat{\varphi}_{k,l}$:

$$\hat{\varphi}_{k,l} = \hat{G}_{k,l} \hat{g}_{k,l}, \quad 0 < k, l < n-1, \quad (6.97a)$$

где для единичного сеточного шага имеем

$$\begin{aligned}\hat{G}_{k,l} = \left[2a \left(\cos \frac{2\pi k}{n} - 1 \right) + ib \sin \frac{2\pi k}{n} + c + 2d \left(\cos \frac{2\pi l}{n} - 1 \right) + \right. \\ \left. + ie \sin \frac{2\pi l}{n} + d \right]^{-1}.\end{aligned}\quad (6.97b)$$

Процедура решения состоит в следующем.

1. Проводят комплексное преобразование Фурье правой части $\hat{g}_{k,l}$, используя алгоритм FFT. Требуется $5N_g \log_2 N_g$ операций.
2. Умножают каждую гармонику на известный комплексный коэффициент $\hat{G}_{k,l}$. Это дает преобразование Фурье решения $\hat{\varphi}_{k,l}$. На это требуется $2N_g$ операций, что незначительно по сравнению с пп. (1) и (3).
3. Проводят обратное преобразование Фурье гармоник $\hat{\varphi}_{k,l}$ для получения решения $\varphi_{p,q}$ с помощью FFT. Требуется $5N_g \log_2 N_g$ операций.

Полное число операций составляет $10N_g \log_2 N_g$, и решение можно записать на место правой части. Если член с первой производной отсутствует, то все преобразования являются вещественными и количество операций сокращается наполовину. Метод можно использовать для задачи с постоянными коэффициентами в трехмерном случае с помощью выполнения тройного преобразования Фурье.

Метод преобразования Фурье наиболее прост для программирования из всех методов RES, если имеется подходящая испытанная библиотечная программа для FFT. Однако этот метод дорог по сравнению с другими методами RES для решения конечно-разностной формы дифференциальных уравнений в частных производных и, следовательно, не рекомендуется для этой цели. Однако он привлекателен, если многомерное преобразование Фурье все равно необходимо, как, например, в алгоритме P³M или при произвольных законах изменения силы с

расстоянием (см. разд. 6.5.4), или если ищется решение для изолированной системы (например, для звездной галактики см. разд. 6.5.5).

В методе MFT гармоники в разложении должны быть собственными функциями разностного оператора левой части уравнения (6.95) и удовлетворять граничным условиям задачи. Это гарантирует, что уравнения для гармоник (6.97а) являются независимыми друг от друга и могут быть решены отдельно умножением каждой гармоники на численный множитель $\hat{G}_{k,l}$. Ясно, что комплексные функции $\exp(2\pi i r k/n)$ в (6.96) удовлетворяют этим условиям для общей задачи с постоянными коэффициентами (уравнение (6.95)) с периодическими граничными условиями. Для граничных условий с нулевым значением, которые удовлетворяются разложением в ряд Фурье по синусам, синусоидальные функции являются собственными функциями оператора, только когда члены с первой производной отсутствуют ($b = e = 0$). Аналогичное утверждение справедливо для разложения в ряд Фурье по косинусам в случае граничных условий, налагаемых на градиент решения.

К счастью, члены с первой производной можно исключить из (6.95) с помощью подстановки $\varphi(x, y) = e^{-\alpha x} e^{-\beta y} \psi(x, y)$ с $\alpha = b/2a$ и $\beta = e/2d$. После умножения на $e^{(\alpha x + \beta y)}$ получается

$$a \frac{\partial^2 \psi}{\partial x^2} + (c - b^2/4a) \psi + d \frac{\partial^2 \psi}{\partial y^2} + (f - e^2/4d) \psi = e^{(\alpha x + \beta y)} g(x, y). \quad (6.97c)$$

Решение $\varphi(x, y)$ находится следующим образом:

1. Образум модифицированную правую часть $e^{(\alpha x + \beta y)} g(x, y)$.
2. Решаем уравнение (6.97с) для ψ , используя рассмотренный ниже метод.
3. Находим решение $\varphi(x, y) = e^{-(\alpha x + \beta y)} \psi(x, y)$.

В последующем обсуждении мы предполагаем, что члены с первой производной уже этим способом исключены. С другой стороны, если уравнение (6.95) записать сначала в разностной форме, то можно использовать подстановку $\psi_{i,j} = \alpha^i \beta^j \varphi_{i,j}$ для исключения влияния первой производной [244, с. 463], где теперь $\alpha = [(1 + bH/2a)/(1 - bH/2a)]^{1/2}$ и $\beta = [(1 + eH/2d)/(1 - eH/2d)]^{1/2}$. В этом случае шаг сетки H должен выбираться так, чтобы $H < 2a/b$ и $< 2d/e$.

Разделение переменных позволяет рассматривать граничные условия по x отдельно от условий по y , и далее мы обсудим граничные условия по x . Совершенно аналогичное рассмотрение применимо в направлении y . Чтобы задача допускала разделение переменных, граничное условие по x не должно зависеть от координаты y . В следующих трех случаях это условие очевидно выполняется:

1. *Периодичность* $\varphi(x+L) = \varphi(x)$. В этом случае пригодно разложение (6.96) с $L = nH$.

2. *Нулевые значения* (т. е. условия Дирихле). $\varphi(L) = \varphi(0)$. В этом случае пригодно дискретное синус-преобразование Фурье

$$\varphi_p = \sum_{k=1}^{n-1} \hat{\varphi}_k \sin(\pi kp/n), \quad (6.98a)$$

и без учета постоянного множителя, который можно ввести в рассмотрение в любом другом месте, получаем

$$\hat{\varphi}_k = \sum_{p=1}^{n-1} \varphi_p \sin(\pi kp/n). \quad (6.98b)$$

3. *Нулевой градиент* (т. е. условия Неймана): $\frac{\partial \varphi}{\partial x} \Big|_{x=0} = \frac{\partial \varphi}{\partial x} \Big|_{x=L} = 0$. В этом случае подходит дискретное косинус-преобразование Фурье

$$\varphi_p = \sum_{k=0}^n E(k) \hat{\varphi}_k \cos(\pi kp/n), \quad (6.99a)$$

и, опять отбрасывая постоянный множитель, получаем

$$\hat{\varphi}_k = \sum_{p=0}^n E(p) \varphi_p \cos(\pi kp/n), \quad (6.99b)$$

где

$$E(i) = \begin{cases} 1/2, & \text{если } i = 0 \text{ или } n, \\ 1 & \text{в противном случае.} \end{cases} \quad (6.99c)$$

Случаи (1) и (2) можно обобщить для включения заданных на границе значений (отличных от нуля) методом эквивалентных зарядов:

4. *Заданные значения искомого решения*. Уравнение (6.95), выраженное в конечных разностях для сеточного узла (p, q) , имеет вид

$$a_{p,q} \varphi_{p,q-1} + b_{p,q} \varphi_{p,q+1} + c_{p,q} \varphi_{p-1,q} + d_{p,q} \varphi_{p+1,q} + e_{p,q} \varphi_{p,q} = f_{p,q}, \quad (6.100)$$

где коэффициенты в сеточных узлах $a_{p,q}, \dots, f_{p,q}$ являются функциями постоянных a, \dots, f и $g(x, y)$ исходных уравнений. Поскольку первая производная по x отсутствует, $c_{p,q} = a_{p,q}$. Предположим, что левая граница по x расположена при $p = 0$; тогда уравнение для сеточного узла, расположенного непосредственно справа от левой границы ($p = 1$), имеет вид

$$a_{1,q} \varphi_{1,q-1} + b_{1,q} \varphi_{1,q+1} + d_{1,q} \varphi_{2,q} + e_{1,q} \varphi_{1,q} = f_{1,q} - c_{1,q} \varphi_{0,q}, \quad (6.101)$$

где мы перенесли известное граничное значение $\varphi_{0,q}$ в правую часть уравнения. Заметим, что (6.101) можно рассматривать в качестве уравнения для задачи с нулевыми граничными условиями в том случае, когда слой эквивалентных источников, равный $-c_{1,q}\varphi_{0,q}$, помещен в столбец сеточных узлов непосредственно с внутренней стороны границы. Тогда для решения задачи в случае 3 можно использовать разложение по синусам. Аналогичным образом, заданные значения на правой границе можно заменить на слой эквивалентного заряда $-d_{n-1,q}\varphi_{n,q}$ в столбце сеточных узлов, примыкающих непосредственно слева к правой границе ($p = n - 1$).

5. *Заданный градиент искомого решения.* Рассмотрим, как прежде, случай левой границы по x . Заданный градиент в точке $(0, q)$ будет выражаться в конечно-разностной форме в виде

$$\frac{\varphi_{1,q} - \varphi_{-1,q}}{2H} \simeq \frac{\partial \varphi}{\partial x} = h_q. \quad (6.102a)$$

Это уравнение должно удовлетворяться вместе с уравнением на границе в точке $(0, q)$

$$a_{0,q}\varphi_{0,q-1} + b_{0,q}\varphi_{0,q+1} + c_{0,q}\varphi_{-1,q} + d_{0,q}\varphi_{1,q} + e_{0,q}\varphi_{0,q} = f_{0,q}, \quad (6.102b)$$

где опять $c_{0,q} = a_{0,q}$ из-за отсутствия первой производной по x . Если мы введем новую переменную

$$\varphi_{-1,q}^* = \varphi_{-1,q} + 2Hh_q, \quad (6.103)$$

то уравнение (6.102) можно переписать в виде

$$\varphi_{1,q} - \varphi_{-1,q}^* = 0, \quad (6.104a)$$

$$a_{0,q}\varphi_{0,q-1} + b_{0,q}\varphi_{0,q+1} + c_{0,q}\varphi_{-1,q}^* + d_{0,q}\varphi_{1,q} + e_{0,q}\varphi_{0,q} = f_{0,q} + 2Hh_q c_{0,q}. \quad (6.104b)$$

Уравнения (6.104) являются уравнениями для нулевого градиента, если слой эквивалентного заряда, равный $2Hh_q c_{0,q}$, помещен в граничные точки $(0, q)$. Уравнения можно решить с помощью дискретного косинус-преобразования Фурье, см. рассмотренный выше случай 3.

6.5.3. Метод FACR

В задаче с постоянными коэффициентами минимальное число операций достигается разумным сочетанием анализа Фурье и циклической редукции в алгоритме FACR [175, 180]. На уровне r циклической редукции получаются уравнения

$$l^{(r)}\varphi_{j-2^r} + \Lambda^{(r)}\varphi_j + m^{(r)}\varphi_{j+2^r} = h^{(r)}. \quad (6.105)$$

В алгоритме FCSR используется то обстоятельство, что редукцию можно остановить на любом уровне r и решить полученные уравнения с помощью преобразования Фурье уравнения (6.105). Уровень r , на котором редукция останавливается, служит теперь параметром, который можно варьировать для минимизации числа арифметических операций. Преобразование Фурье проводится в одном измерении (по координате x), следуя методам и условиям, данным в разд. 6.5.2.

После преобразования Фурье уравнение (6.105) принимает вид

$$l^{(r)}\hat{\phi}_{j-2^r}^k + \lambda_k^{(r)}\hat{\phi}_j^k + m^{(r)}\hat{\phi}_{j+2^r}^k = \hat{h}_j^k, \quad (6.106)$$

где $\lambda_k^{(r)}$ являются собственными значениями оператора $\Lambda^{(r)}$, соответствующими k -ой Фурье-гармонике. Оно представляет одну независимую трехдиагональную систему в направлении y для каждой из амплитуд гармоник. Трехдиагональные системы решаются с помощью алгоритма Томаса, а обратное преобразование Фурье дает решение на $n/2^r$ строках на уровне r . Заполнение промежуточных строк происходит так же, как в методе CR.

В простом случае уравнения Пуассона в квадратной области, количество операций на один сеточный узел, требуемых в методе FCSR (если редукция остановлена на уровне r), равно [180]

$$S(r) = 3 + 4.5r + (5 \log_2 n - 4)2^{-r}, \quad (6.107)$$

где первые два члена соответствуют циклической редукции и последний член — прямому и обратному преобразованию Фурье по каждой 2^r -ой строке.

Обычно мы называем этот метод алгоритмом FCSR(l), где максимальная глубина редукции задается в скобках. Однако здесь мы используем символ r вместо l для того, чтобы избежать путаницы с коэффициентом в уравнении (6.105). Кривая $S(r)$ нанесена на рис. 6.8 для случая $n = 128$; виден минимум при $r = 2$. В большей общности, следуя Шварцтрауберу [359], можно продифференцировать S и найти минимум аналитически:

$$\frac{dS}{dr} = 4.5 - \ln 2 (5 \log_2 n - 4)2^{-r} + 0; \quad (6.108a)$$

следовательно, в минимуме

$$2^r = \frac{\ln 2}{4.5} (5 \log_2 n - 4), \quad (6.108b)$$

$$\text{и} \quad S_{\min} = 3 + \frac{4.5}{\ln 2} + 4.5 \log_2 \left[\frac{\ln 2 (5 \log_2 n - 4)}{4.5} \right]. \quad (6.108c)$$

Таким образом, в пределе при больших n оптимальный алгоритм FCSR, который мы обозначаем как FCSR(*), содержит

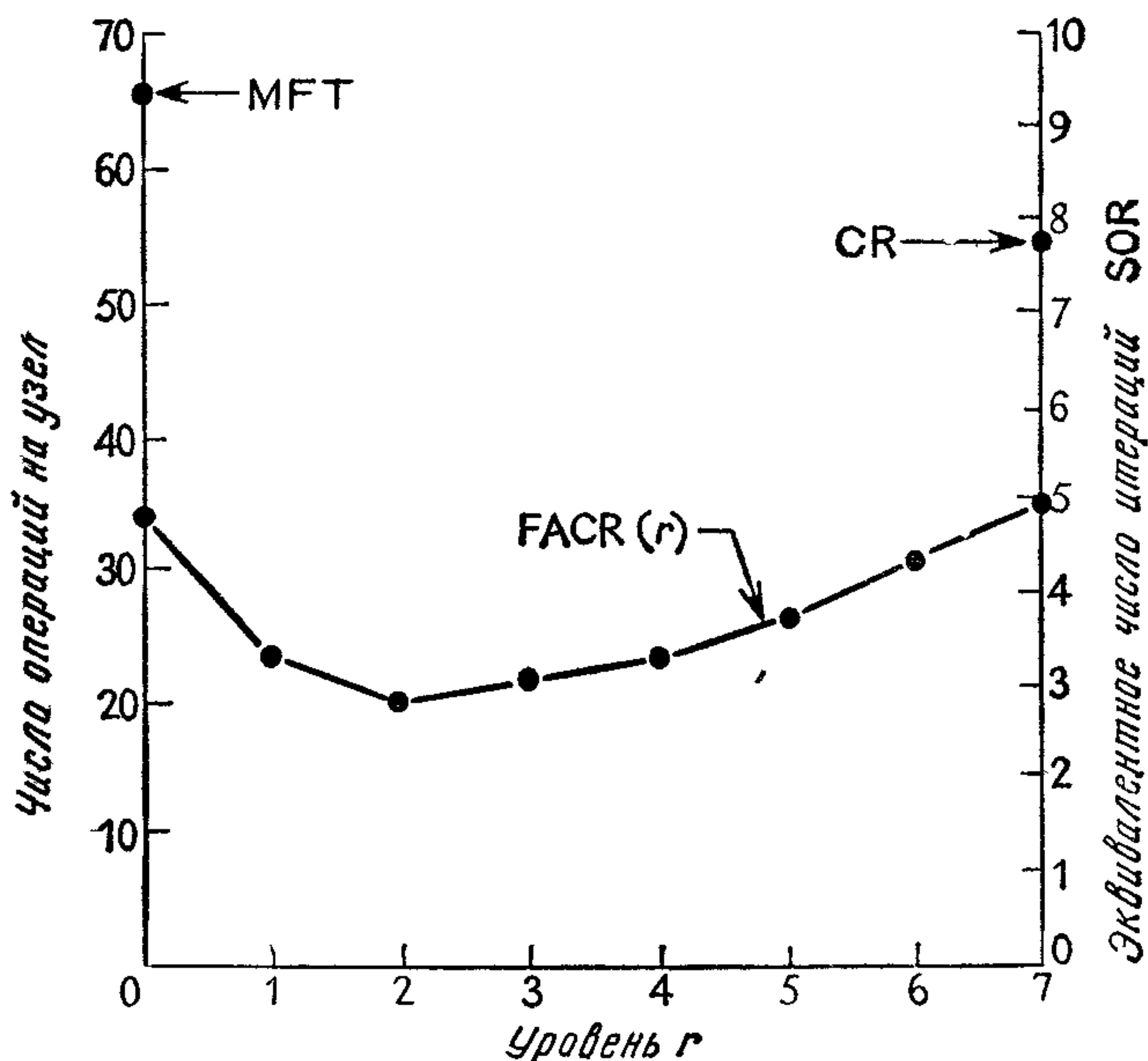


Рис. 6.8. Количество операций на один узел сетки (128×128) для алгоритмов FACR(r), MFT и CR. Дано также эквивалентное число итераций для метода SOR в предположении, что число операций в одном узле на одну итерацию составляет наименьшее значение, равное 7. Во всех случаях граничные условия выбраны в форме, наиболее благоприятной для данного алгоритма. Отметим, что минимальное число операций достигается при $r = 2$ [180].

следующее количество операций:

$$S_{\min}(n \rightarrow \infty) = 4.5 \log_2(\log_2 n), \quad (6.109a)$$

и

$$r(n \rightarrow \infty) = \log_2(\log_2 n), \quad (6.109b)$$

а предельное полное число операций равно

$$T_{\text{FACR}} = 4.5 N_g \log_2(\log_2 N_g). \quad (6.109c)$$

Для практически применяемых значений n асимптотические формулы (6.109) достаточно точны. Для $n = 128$ получаются значения $r \sim 3$ и $S_{\min} \sim 14$, сравнимые со значениями $r \sim 2$ и $S_{\min} \sim 20$, полученными из полных формул (6.108). Обширное исследование минимума было выполнено Темпертоном [366] с использованием программы с переменным r .

В методе CR (разд. 6.5.1) используется значение $r = \log_2 n$ ($= 7$ в верхнем примере), и редукция проводится до конца. Количество операций в этом случае равно

$$T_{\text{CR}} = 4.5 N_g \log_2 N_g. \quad (6.109d)$$

Таким образом, мы убеждаемся, что метод CR неоптимален и всегда медленнее, чем оптимальный метод FACR. Асимптотические времена для этих двух алгоритмов относятся как

$$\frac{T_{\text{FACR}}}{T_{\text{CR}}} = \frac{\log_2(\log_2 N_g)}{\log_2 N_g}. \quad (6.109e)$$

Для $n \sim 128$ это отношение приблизительно равно 2 и медленно увеличивается с ростом n .

Приведенные выше данные о числе операций относятся к редукции по «неустойчивому» алгоритму (6.89). На такой машине, как CDC-7600, точность выполнения арифметических операций которой составляет $\sim 10^{-14}$, вполне допустимо потерять несколько десятичных знаков, применяя «неустойчивый» алгоритм. Подпрограмма ROT1 [180] отвечает алгоритму FACR(1) и, следовательно, является не совсем оптимальной. Здесь число операций равно $1.25(\log_2 N_g + 4.8)$. Из табл. 6.5 видно, что для

Таблица 6.5. Затраты машинного времени и величина погрешности при решении уравнения Пуассона (двумерного) в квадратной области с использованием подпрограммы ROT1. Применялся алгоритм FACR(1), тщательно запрограммированный на языке Ассемблера.

Сетка	CDC 7600		IBM 360/195	
	t , мс	Погрешность ¹⁾	t , мс	Погрешность ¹⁾
32×32	6	1×10^{-12}	12	8×10^{-5}
64×64	36	5×10^{-12}	39	2×10^{-4}
128×128	145	2×10^{-11}	137	4×10^{-4}
256×256	—	—	559	1×10^{-3}

¹⁾ Это погрешность конечно-разностной аппроксимации, а не погрешность округления.

сетки 128×128 на этой машине получается точность $\sim 10^{-12}$. На машинах IBM с разрядностью 32 бита (погрешность округления $\sim 10^{-6}$) точность решения понижается до $\sim 10^{-4}$. Следовательно, на одном уровне «неустойчивой» редукции теряется около двух десятичных знаков. Для того, чтобы на машинах IBM проводить вычисления на больших сетках, можно или перейти на двойную точность (64 бита), или выполнять шаги циклической редукции с помощью «устойчивого» алгоритма Бунемана (6.94). Недостаток использования алгоритма Бунемана состоит в том, что он вводит дополнительные арифметические операции, и множитель 4.5 в формулах (6.108) и (6.109) следует заменить на 6. Однако асимптотическое отношение времен

исполнения оптимального FACR и CR алгоритмов в (6.109e) остается неизменным. В табл. 6.6 сравниваются времена для алгоритмов CR, FACR(1) и MFT на различных сетках.

Таблица 6.6. Затраты машинного времени при решении двумерного уравнения Пуассона на квадратной области различными методами. Сетка 128×128 , время в миллисекундах на CDC-7600, программирование на Фортране

Сетка	PWSCRT [†] CR	POT1* FACR (1)	POT 3* MFT
32×32	49	26	34
64×64	211	71	142
128×128	926	229	604
256×256	—	630	—

[†] По данным работы [360].

* Транслятор с Фортрана FTN 4.6.

Оценка для CDC-7600, полученная умножением времени работы фортрановской программы на IBM 360/195 на коэффициент 0.6.

В алгоритме FACR разложение в ряд Фурье проводится только по координате x . Поэтому алгоритм можно применить к более общим задачам вида

$$\nabla D(y) \nabla \varphi(x, y) + k^2(y) \varphi(x, y) = s(x, y), \quad (6.110)$$

в которых коэффициенты уравнения и сеточные шаги могут зависеть от y , но не от x . Примером является решение уравнения Пуассона в цилиндрических координатах, для которого Хьюс [209] опубликовал программу FACR(0). Программы FACR(1) для той же самой задачи в декартовых координатах опубликовали Хокни [180] и Христиансен и Хокни [79].

6.5.4. Методы свертки

Часто необходимо вычислять потенциал, обусловленный распределением источников, когда потенциал взаимодействия нельзя представить уравнением в частных производных. Примером служит взаимодействие Леннард-Джонса между атомами, которое имеет вид $G(r) \sim \{r^{-12} - r^{-6}\}$, где $G(r)$ — потенциал двух атомов, разделенных расстоянием r . Если взаимодействие попарно аддитивно, то потенциал в сеточном узле (p, q) можно всегда записать как сумму вкладов от всех дру-

гих точечных источников (p', q') :

$$\Phi_{p, q} = \sum G_{p-p', q-q'} \rho_{p', q'}. \quad (6.111)$$

Формула (6.111) выражает потенциал через свертку распределения источника ρ с функцией Грина взаимодействия (или с потенциалом взаимодействия) G . Теорема о свертке (см. приложение А.3) утверждает, что преобразование Фурье от свертки равно произведению преобразований Фурье от свертываемых величин. Поэтому формулу (6.111) можно записать в виде

$$\hat{\Phi}_{k, l} = \hat{G}_{k, l} \hat{\rho}_{k, l}, \quad (6.112)$$

где $\hat{}$ означает преобразование Фурье и (k, l) — волновые числа гармоник.

Можно найти потенциал, обусловленный любым периодическим множеством источников с произвольным потенциалом взаимодействия. Сначала рассматриваемый периодический потенциал взаимодействия G записывается на сетке для единичного источника в начале координат. Преобразование Фурье \hat{G} этого потенциала взаимодействия находится и запоминается. Такой расчет необходимо выполнить лишь один раз. На каждом временном шаге моделирования распределение источника ρ получается на второй сетке. Затем находим преобразование Фурье $\hat{\rho}$ от этого распределения; его можно записать на место исходного распределения источников ρ . Для получения преобразования Фурье от распределения потенциала преобразование Фурье от источника умножается на преобразование Фурье от потенциала взаимодействия. Затем с помощью обратного преобразования Фурье получаем распределение потенциала, обусловленное исходным распределением. Все операции на второй сетке можно записать друг на друга, и требуемая полная память приблизительно равна $2N_g$.

Многомерное вещественное преобразование Фурье на N_g узлах требует $2.5N_g \log_2 N_g$ операций (независимо от числа измерений), поэтому полное количество операций составляет $5N_g \log_2 N_g$. Память, требуемая для произвольного взаимодействия, равна $2N_g$ — одна сетка для источника и решения и вторая сетка — для функции Грина и ее преобразования. Однако если G зависит только от расстояния между частицами и мы работаем в области, имеющей форму куба, то можно использовать симметрию и уменьшить память для \hat{G} приблизительно до $N_g/48$. В этом случае полная требуемая память по существу равна N_g .

Описанный метод свертки будет находить решение периодической системы источников с произвольным видом взаимодействия. В системе не допускаются проводники или границы и, если они присутствуют, необходимо использовать метод матрицы

емкости (см. разд. 6.5.6). При подходящем выборе функции Грина метод свертки позволяет найти потенциал для изолированного распределения источников. Под изолированными системами мы понимаем системы, подобные галактике, в которых единственное граничное условие состоит в том, что потенциал правильным образом убывает до нуля на бесконечности (т. е. $\Phi \sim r^{-1}$).

Метод, который мы описывали в двумерном случае, можно приспособить к изолированным системам, если пойти на то, чтобы для описания распределения зарядов использовалась только четверть имеющегося количества сеточных узлов, например нижний левый квадрант, определяемый соотношениями $0 \leq p, q \leq n/2$. Распределение источников в оставшихся трех квадрантах системы задается тождественно равным нулю. Если в качестве примера взять взаимодействие точечных зарядов, то потенциал взаимодействия будет строиться следующим образом:

$$\left. \begin{aligned} G_{p, q} &= (p^2 + q^2)^{-1/2} \\ G_{n-p, q} &= G_{p, n-q} = G_{n-p, n-q} = G_{p, q} \\ G_{0, 0} &= 1 \end{aligned} \right\} \begin{aligned} &0 \leq p, q \leq n/2, \\ &p + q \neq 0. \end{aligned} \quad (6.113)$$

Когда этот потенциал продолжен периодически, видно, что правильная зависимость вида r^{-1} для потенциала точечного заряда, помещенного в начале координат, будет иметь место в области $-n/2 \leq p, q \leq n/2$. На границе этой области потенциал имеет точку возврата, и вне ее он принимает неверные значения. Однако если использовать только нижний левый квадрант для описания распределения зарядов и рассматривать потенциал только в этой области, то будет получено верное значение потенциала изолированной системы. Вне нижнего левого квадранта потенциал будет неверным; в самом деле, он будет содержать все нефизические точки возврата потенциала взаимодействия. Последнее, однако, не имеет значения, поскольку этот потенциал никогда не используется. В случае изолированной системы применение сетки иллюстрируется на рис. 6.9.

Если N_g равно полному числу узлов в физической расчетной области (активная сетка), то расчет должен проводиться на $2^d N_g$ узлах, где d — размерность. Это вычисление асимптотически требует $5(2^d N_g) \log_2 N_g$ операций. Период преобразований равен $n = 2N_g^{1/d}$.

Память, требуемая в двумерном случае, очевидно, равна $4N_g$, однако ее можно понизить до $2N_g$ следующим образом. Мы запоминаем только нижнюю половину рис. 6.9 и сначала находим преобразование Фурье всех данных в направлении x . Это заполняет числами нижнюю половину рис. 6.9. Далее мы берем данные на одной строке по y и выполняем следующие три операции: 1) находим преобразование Фурье строки по y и полу-

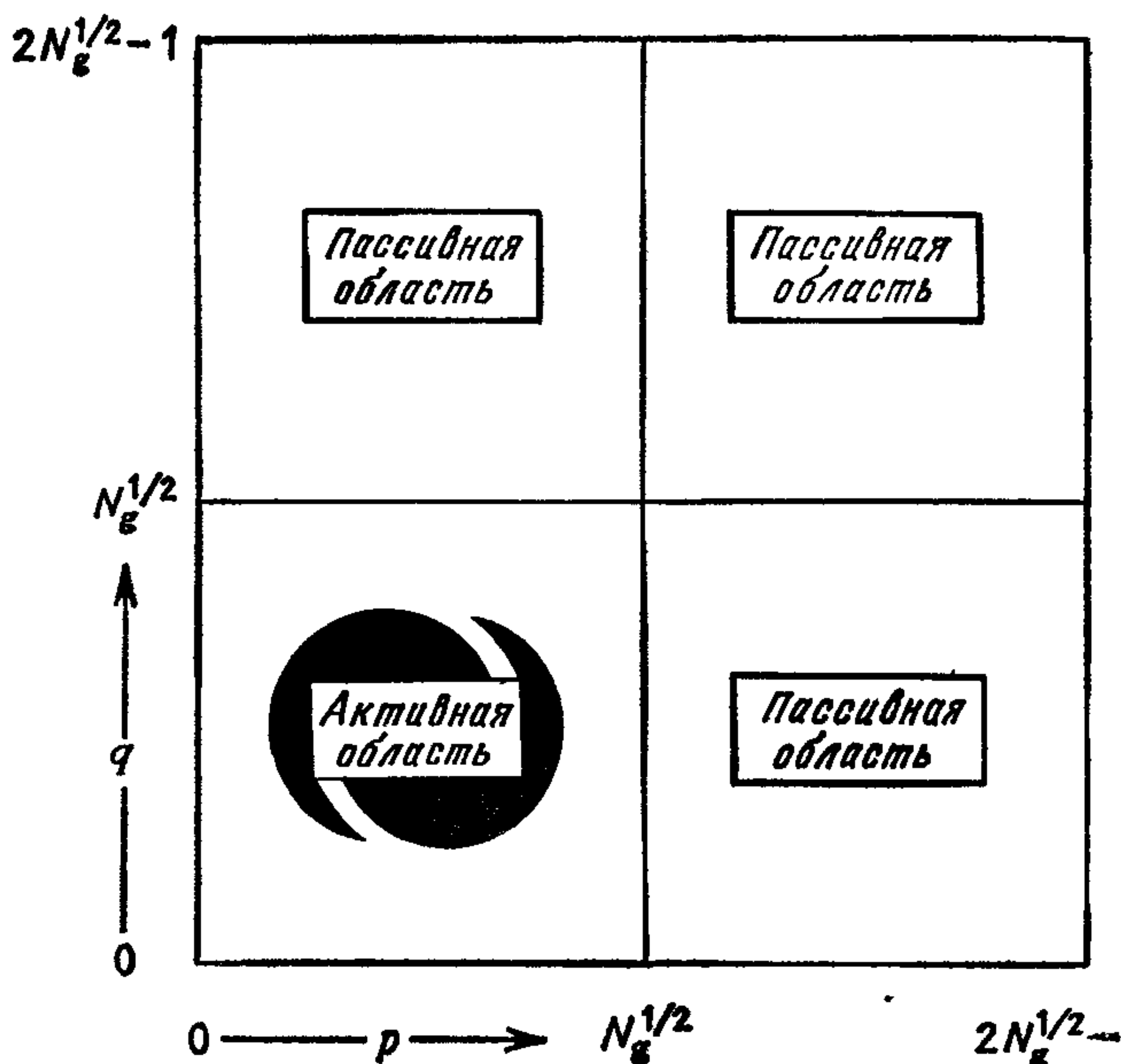


Рис 69 Использование сетки для расчетов потенциала в изолированной системе. Для описания зарядов и масс можно использовать только активную четверть сетки. Во время вычисления требуются остальные пассивные области [180].

чаем временный вектор длиной $2N_g^{1/2}$; 2) умножаем его на \hat{G} ; 3) производим обратное преобразование по y и запоминаем результат в нижней левой четверти на рис. 69. Данные, которые можно было бы запомнить в верхнем левом квадранте, не будут влиять на решение в активной области, и они отбрасываются. Полная требуемая память равна $2N_g + 2N_g^{1/2} \approx 2N_g$. Количество требуемых машинных операций равно $15N_g \log_2 N_g$ в сравнении с $20N_g \log_2 N_g$, которое потребовалось бы без упомянутой выше реорганизации процесса.

Для выполнения подобной процедуры в трехмерном случае требуется память для запоминания двух активных сеток $2N_g$, временной плоскости объемом в $2N_g^{2/3}$ узлов и временного вектора в $2N_g^{1/3}$ узлов. Сначала проводится преобразование Фурье всех начальных данных в направлении x . Затем каждый раз выбирается одна плоскость, и на ней производится преобразование Фурье в направлении y . Из этой плоскости каждый раз выбирается по одному вектору и находится преобразование Фурье этого вектора в направлении z . Затем можно провести умножение на функцию Грина, после чего следует обратное преобразование по z . Это повторяется для каждого вектора в направ-

лении z , пока не пройдена вся плоскость. Тогда проводится обратное преобразование по y внутри этой плоскости, и когда такие обратные преобразования по y для всех плоскостей завершены, то для получения решения проводится обратное преобразование по x .

Обобщенное таким образом на случай расчета в d измерениях количество арифметических операций составляет

$$\frac{5 \times 2^d}{d} N_g \log_2 N_g,$$

и требуемая память равна

$$2(N_g^{1/d} + N_g^{2/d} + \dots + N_g) \approx 2N_g.$$

Дополнительную память, требуемую для запоминания функции Грина, можно сократить до N_g/d , используя симметрии функции G . Программа ROT5A была написана Браунриггом на основе описанного выше метода для трехмерного изолированного распределения масс [54, 118]. Время выполнения программы на IBM 360/195 составляет 46 с для активной сетки $32 \times 32 \times 16$, а требуемая память равна $\sim 2.5N_g$.

Марун, Уэлтон и Уонг [276] утверждали, что рассмотренный выше метод для изолированных систем не устраняет изображения источников полностью, а только делает их более удаленными. Однако Иствуд и Браунригг [118] доказали, что это утверждение неверно и что рассмотренный выше метод правильно решает задачу для изолированной системы.

6.5.5. Алгоритм Джеймса

Метод MFT для свертки, описанный в разд. 6.5.4, приложим к произвольному потенциалу взаимодействия. Метод дорогостоящ как по потребляемой памяти, так и по затратам машинного времени, поскольку для исключения изображений источников он требует дублирования системы в каждом координатном направлении. Джеймс [217] предложил усовершенствованный алгоритм, обходящийся без удвоения при решении уравнения Пуассона для изолированной системы.

Когда распределение источников задается в кубической области, содержащей $N_g = (n \times n \times n)$ узлов, процедура решения выполняется следующим образом:

1. Вычисляем потенциал ϕ внутри куба методом MFT. На границе задаем нулевое значение потенциала.

2. Полагая потенциал $\phi = 0$ вне куба, применяем оператор Пуассона к узлам границы, чтобы определить наведенный на границе «экранирующий» заряд из соотношения

$$q = -\nabla^2 \phi. \quad (6.114)$$

3. Вычисляем поправку ψ , обусловленную этими отрицательными экранирующими зарядами. Процедура удвоения (разд. 6.5.4) используется для расчета потенциала, обусловленного изолированной системой экранирующих зарядов, но удваиваются только граничные величины, а не вся сетка.

Точное решение является суммой исходного потенциала ϕ и поправки ψ , обусловленной фиктивными граничными зарядами. Далее поправку на граничные заряды можно ввести в преобразованном пространстве решения методом MFT (на этапе (1)), подобно тому, как вводятся граничные электроды в методах матрицы емкости (см. разд. 6.5.6).

Алгоритм Джеймса примерно в три раза быстрее, чем метод свертки, и требует примерно вдвое меньшей памяти. Джеймс описывает следующие характеристики программы на языке Фортран (FTN4.4) при использовании CDC-7600 и сетки $33 \times 33 \times 33$:

1. Проведение FFT на полной сетке ($15n^3 \log_2 n$ операций)	1.0 с
2. Расчет экранирующего заряда ($6n^3$ операций)	0.3 с
3. Расчет граничного потенциала ($18n^3 + 108n^2 \log_2 n$ операций)	0.6 с
4. Организационные затраты времени	0.1 с
<hr/>	
Полное время	2.0 с

Поскольку вычисление проводится только на активной сетке, состоящей из N_g узлов, то требуемая память лишь незначительно превышает N_g . Асимптотически требуется $\sim 5N_g \log N_g$ операций.

6.5.6. Метод матрицы емкости

Могло показаться, будто только что описанные прямые методы для решения уравнения Пуассона имеют весьма ограниченное применение, в особенности при моделировании электронных приборов (см. гл. 10, в частности разд. 10.4.7), поскольку электроды внутри области не допускаются, хотя, конечно, они могут служить ее границами. Один из способов обойти это ограничение заключается в предварительном расчете матрицы емкости (назовем ее C), которая связывает потенциал и заряд в целом ряде узлов внутри области. Эти узлы могут быть электродами, поддерживаемыми при заданном потенциале, и цель состоит в вычислении поверхностного заряда, индуцированного на электродах окружающим пространственным зарядом. Решение получаем, дважды решая уравнение Пуассона, сле-

дующим образом. Сначала решаем уравнение Пуассона с нулевым зарядом на электродных узлах и записываем отклонение потенциала на электродных узлах от требуемых значений. Это отклонение после умножения на матрицу емкости и замены дает искомый поверхностный заряд в каждом электродном узле. Уравнение Пуассона решается снова с подсчитанным поверхностным зарядом, и теперь решение справедливо всюду в области, включая электродные узлы.

Способ непосредственного отыскания матрицы емкости состоит в помещении единичного заряда в каждый электродный узел по очереди (с нулевым зарядом на других электродных узлах) и решении относительно потенциала. Значения потенциала на электродных узлах образуют элементы одного столбца обратной матрицы емкости $\mathbf{A} = \mathbf{C}^{-1}$. Повторение этого процесса для каждого из l электродных узлов заполняет l столбцов матрицы \mathbf{A} . Затем матрица емкости находится обращением матрицы \mathbf{A} .

Если граничные условия области являются периодическими, то потенциал, обусловленный единичным зарядом, зависит только от вектора смещения между узлами потенциала и заряда. В этом случае обратную матрицу емкости можно построить, исходя из единственного решения для потенциала от единичного заряда, помещенного в любую точку, подходящим образом выбирая исходные данные [385]. Случаи заданных значений искомого решения на противоположных поверхностях или заданного градиента искомого решения можно свести к периодическому случаю в подходящим образом удвоенной области и рассматривать их аналогично. Эти методы снижают число решений для потенциала, необходимых для нахождения элементов матрицы \mathbf{A} , до четырех или еще меньшего числа, но они не устраняют требующую большего времени задачу обращения матрицы \mathbf{A} для получения матрицы емкости \mathbf{C} .

Необходимость запоминать и обращать матрицу \mathbf{A} ограничивает применение описанного выше метода матрицы емкости до 100—200 узлов. Матрица емкости \mathbf{C} , которую можно записать на место матрицы \mathbf{A} в памяти, является симметричной, так что требуемая память равна $l(l+1)/2$ для l электродных узлов. Поскольку матрица емкости зависит только от геометрии задачи, но не от распределения пространственного заряда, то ее можно в начале вычислить раз и навсегда и запомнить. Время, затраченное на умножение отклонения решения на матрицу емкости, незначительно, поэтому если влияние электродов учитывают таким образом, то машинное время, требуемое для определения потенциала, удваивается. Мы будем называть этот метод прямым методом матрицы емкости (DCM).

Мартин [275] показал, что метод можно обобщить на более широкий класс условий, чем условие заданного потенциала

*Верхняя и нижняя
границы*

1. $\varphi(x)$ задан
2. $\frac{\partial \varphi}{\partial y} = 0$
3. φ периодический
4. $a\varphi(x) + b \frac{\partial \varphi(x)}{\partial n} = c(x)$

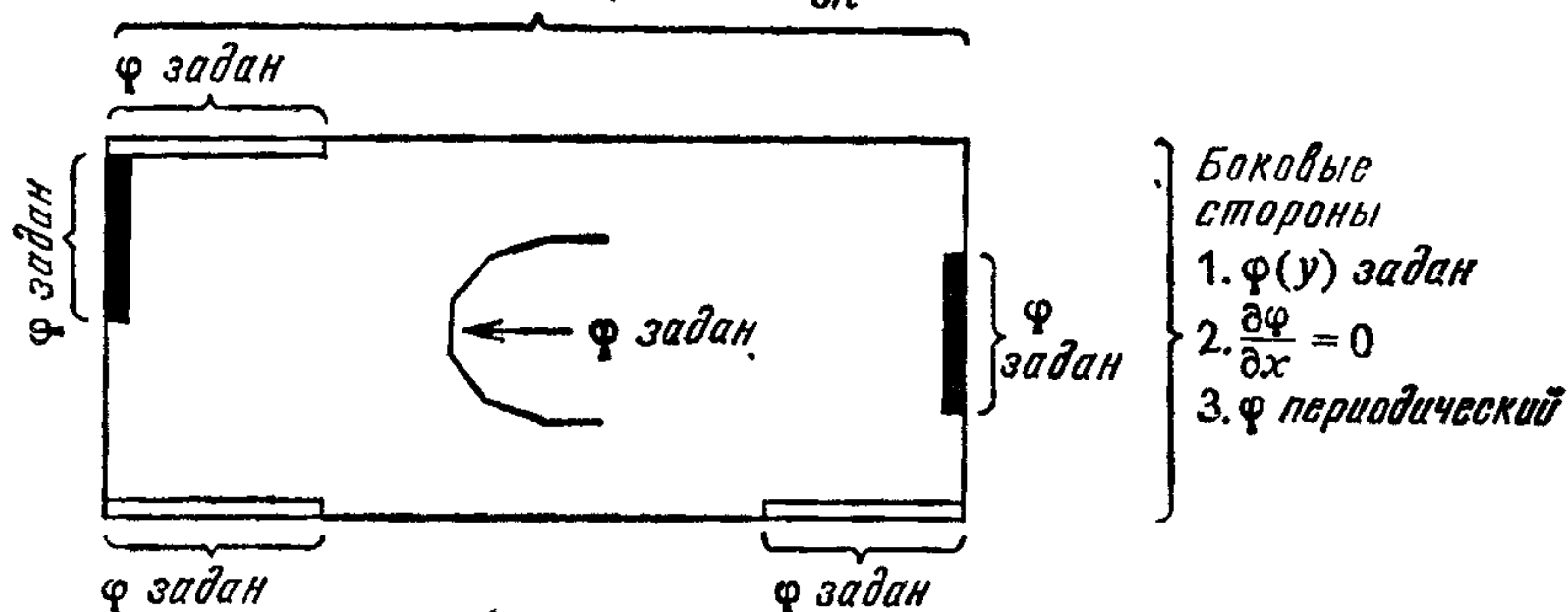


Рис. 6.10 Примеры типов задач Пуассона, решаемых программой РОТ4. Тонкими линиями изображены граничные электроды, а жирными линиями — внутренние электроды [180].

электродов. Отклонение на «электродном» узле можно рассчитать в более общем виде как

$$\varepsilon = a\varphi + b \frac{\partial \varphi}{\partial n} - c. \quad (6.115)$$

Поскольку в методе полагается $\varepsilon = 0$, то первый и последний члены дают фиксированный потенциал, а второй член может учитывать разрывы коэффициента диффузии, диэлектрической постоянной или проводимости. Матрица емкости вычисляется так же, как прежде, с тем отличием, что элементами обратной матрицы емкости служат значения ε из уравнения (6.115), соответствующие помещению единичных зарядов в электродные узлы.

Программа РОТ4 [184] реализует применение алгоритма FASR(1) и прямого метода матрицы емкости для решения широкого класса задач, возникающих при проектировании полупроводниковых устройств. На рис. 6.10 показаны условия, при которых можно найти решение с помощью этой программы. Прямоугольная геометрия является хорошим представлением эпитаксиального полупроводника. Если электроды задаются на верхней и нижней сторонах прямоугольника, то можно применить поправку с помощью матрицы емкости на половине цикла по алгоритму FASR, и полное решение получается однократным применением указанного алгоритма. Однако если какой-нибудь электрод находится внутри области или на ее боковых сторонах,

то требуются два обращения к алгоритму FACR, а поправка с помощью матрицы емкости вводится между этими двумя обращениями. На табл. 6.7 и 6.8 приведены времена работы такого алгоритма для различных сеток и наборов электродов. Подготовка матрицы емкости (колонки, озаглавленные SEPT4) — это

Таблица 6.7. Затраты машинного времени в секундах на IBM 360/91 для решения тестовой задачи со 100 граничными и 100 внутренними электродами и с условиями Неймана на всех других границах при использовании POT4 на сетке $NX \times NY$

NX	NY	Погрешность	SEPT4 [†]	POT4*	Без электродов [§]
128	32	8.4×10^{-4}	38.8	0.73	0.23
256	64	5.1×10^{-4}	99.2	1.98	0.79
512	128	1.3×10^{-3}	360	7.0	3.20
1024	256	— [#]	1440	28.0	13.4

[†] Время расчета матриц емкости.

* Время решения задачи о потенциале при известных матрицах емкости.

[§] Время решения задачи о потенциале без учета электродов.

[#] Памяти недостаточно для хранения точного решения, и поэтому погрешность подсчитать нельзя.

Источник: [184], с разрешения Advance Publications.

Таблица 6.8. Затраты машинного времени в секундах на IBM 360/91 для решения задачи Пуассона на сетке 128×32 для различного числа электродов с заданными на всех других границах условиями Неймана

N	Только граничные		Только внутренние		Оба типа электродов [†]	
	SEPT4*	POT4 [§]	SEPT4*	POT4 [§]	SFTPTO*	POT4 [§]
0	0.006	0.21	0.006	0.21	0.006	0.21
25	0.94	0.25	5.5	0.43	7.5	0.51
50	2.16	0.29	11.2	0.46	15.6	0.56
75	3.94	0.30	17.5	0.51	26.4	0.65
100	6.52	0.33	24.9	0.58	38.8	0.73

[†] N внутренних и N граничных.

* Затраты времени для выполнения подпрограммы SEPT4. В основном это время вычисления матрицы емкости.

[§] Время решения задачи Пуассона с заданной матрицей емкости. При наличии внутренних электродов требуется два обращения к подпрограмме POT4.

Источник: [184], с разрешения Advance Publications.

дополнительные расходы, которые требуются только в случае изменения геометрии электродов. Матрица емкости не меняется при изменении распределения заряда или приложенных потенциалов. Поэтому при нестационарном моделировании для постоянной геометрии матрица емкости вычисляется только один раз. Времена в колонках, озаглавленных ROT4, служат мерой стоимости шага по времени в этом случае.

Если набор электродных узлов слишком велик, чтобы допустить вычисление и хранение матрицы емкости, то поверхностный заряд можно итеративно установить по локальному отклонению в электронном узле решения уравнения Пуассона, по-прежнему получаемого прямым методом. Другой способ — полностью отбросить прямые методы и использовать какой-нибудь из описанных выше методов решения общего неразделяющегося уравнения. В зависимости от числа сеточных узлов могут быть подходящими методы разреженных матриц (см. разд. 6.4.3), неполного разложения (разд. 6.4.4—6.4.6) или сеточной релаксации (разд. 6.3.3—6.3.6 и 6.5.7).

В итерационном методе матрицы емкости (ICM), разработанном Проскуровским и Уидландом [317], набор поправочных диполей в граничных узлах определяется решением уравнения $A_{\mu} = -\varepsilon$, где ε представляет отклонение в граничных условиях. Матрица A (путанно и неправильно, если иметь в виду электростатическую терминологию, названная Проскуровским матрицей емкости), является дипольным аналогом обратной матрицы емкости. Решение находится итеративно с помощью метода сопряженных градиентов, а использование поправочных диполей вместо зарядов обеспечивает близость собственных чисел матрицы A к единице и быструю сходимость (см. разд. 6.4.2). Поскольку в методе CG (сопряженных градиентов) требуется только вычисление вектора невязки $r = A_{\mu} + \varepsilon$, который можно найти однократным применением RES, то нет необходимости находить или хранить элементы матрицы A . Таким образом полная память понижается приблизительно до 30%. На самом деле требуются два применения алгоритма RES на одну итерацию CG, поскольку матрица A является несимметричной и нужно работать с симметричной матрицей AA^T .

Метод Проскуровского и Уидланда подходит для областей, расположенных полностью внутри или полностью снаружи от замкнутой границы. Поскольку строка диполей вызывает разрыв потенциала поперек границы, то поправка не является аналогичной действию физического электрода, на котором индуцируются заряды, а не диполи. Это означает, что метод не подходит для границ, которые не образуют замкнутой кривой, например для внутренних электродов на рис. 6.10. Используя свою программу HLMHLZ, Проскуровский [316] получил решение уравнения Пуассона с граничными условиями Дирихле

внутри круга с 528 граничными узлами на сетке 256×256 . Для снижения погрешности до 10^{-6} требуется 16 итераций CG (32 решения методом RES) и затрачено 24 с на CDC-7600. Метод был обобщен на трехмерный случай О'Лири и Уидландом [300].

Выбор между прямым и итерационным методами матрицы емкости в основном определяется количеством электродных узлов l . При заданной матрице емкости можно получить решение методом DCM за время выполнения двух решений методом RES или одной итерации метода ICM. Поэтому прямой метод предпочтительней, если необходимо найти множество решений при фиксированной геометрии, когда стоимостью предварительного вычисления матрицы емкости можно пренебречь. С другой стороны, если рассматривается одноразовое решение, то необходимо сравнить время обращения матрицы $(l \times l)$ (приблизительно равное $\sim 2l^3$ операций для алгоритма DCM, использующем периодический метод Уидланда для получения матрицы A) с временем выполнения 32 решений методом RES (равное $\sim 80n^2 \log_2 n$ для алгоритма ICM на сетке $(n \times n)$). Если выбрать $l = 2n$ и $n = 128$, то найдем, что метод DCM быстрее метода ICM, при $l \leq 140$. Это значение близко к практическому пределу для обращения матриц, так что, вероятно, метод DCM следует применять всегда, когда это осуществимо практически. Метод ICM наиболее подходит, если число электродных узлов составляет несколько сотен или тысяч, когда, очевидно, нельзя использовать метод DCM. Именно так обстоит дело при точном описании любой достаточно сложной геометрии, например электронной пушки.

Наиболее общее уравнение поля с произвольными коэффициентами будет состоять из областей с относительно гладким изменением коэффициентов, разделенных границами, на которых сами коэффициенты или их производные терпят разрывы. В такой задаче любой итерационный метод (например, итерация Конкуса и Голуба, разд. 6.5.7), предполагающий гладкое изменение коэффициентов по всей области решения, будет сходиться очень медленно или не сходиться вовсе из-за разрывов. В этих случаях рекомендуется, чтобы разрывы были включены в корректировку матрицы емкости, оставляя таким образом задачу с гладким изменением коэффициентов для решения итерационным методом.

6.5.7. Итерация по Конкусу и Голубу

Уравнение с разделяющимися переменными вида (6.82) являются наиболее общим уравнением, которое можно решить непосредственно методами RES. Конкус и Голуб [86] показали, что итеративным использованием алгоритма RES можно решить более общие уравнения. Они добиваются решения в прямоуголь-

ной области общей задачи с переменными коэффициентами

$$-\nabla [D(x, y) \nabla \varphi] = f(x, y), \quad (6.116)$$

не допускающей разделения переменных. Это уравнение представляет диффузию с переменным коэффициентом диффузии или задачу об электростатическом потенциале с переменной диэлектрической проницаемостью. В случае когда D постоянно, имеем модельное уравнение Пуассона, к которому можно применить все методы RES. Если $D(x, y) = X(x)Y(y)$, т. е. представимо в виде произведения, то уравнение (6.116) допускает разделение переменных, и можно применить непосредственно метод CR (разд. 6.5.1). Такой вид коэффициента D является весьма искусственным и не соответствует какой-либо очевидной физической ситуации. Поэтому нам нужно искать решение уравнения (6.116) с произвольным изменением величины $D(x, y)$.

Если $D(x, y)$ имеет вторую производную, то можно ввести замену переменной

$$W(x, y) = [D(x, y)]^{1/2} \varphi(x, y). \quad (6.117)$$

Тогда уравнение (6.116) принимает вид

$$-\nabla^2 W + p(x, y) W = q(x, y), \quad (6.118a)$$

где
$$p(x, y) = D^{-1/2} \nabla^2 (D^{1/2}) \quad (6.118b)$$

и
$$q(x, y) = D^{-1/2} f. \quad (6.118c)$$

Таким образом, замена переменной преобразует дифференциальную часть оператора к виду, который наиболее подходит для решения методами RES.

Итерационная процедура определяется следующим образом:

$$(-\nabla^2 + K) W^{(n+1)} = [K - p(x, y)] W^{(n)} + q(x, y), \quad (6.119a)$$

где n — номер итерации и K — параметр, который можно менять для улучшения сходимости. Величину K выбирают так, чтобы минимизировать наибольшее собственное значение матрицы перехода: она равна

$$K = \frac{1}{2} (\min(D) + \max(D)). \quad (6.119b)$$

Решение дискретизированной формы уравнения (6.119a) проводится алгоритмом RES, поскольку левая часть уравнения представляет задачу с постоянными коэффициентами. Если p на самом деле постоянно, то решение получается за одну итерацию. Это происходит, например, когда $D(x, y) = \sin^2(x + y)$.

Дискретизированная форма уравнения (6.119a) имеет вид

$$(-\nabla_h^2 + KI) W^{(n+1)} = (KI - P) W^{(n)} + Q, \quad (6.120)$$

где ∇_h^2 и \mathbf{P} являются матрицами, полученными из конечно-разностного приближения операторов ∇^2 и $p(x, y)$, \mathbf{Q} является вектором значений правой части $q(x, y)$.

Погрешность удовлетворяет уравнению

$$\varepsilon^{(n+1)} = \mathbf{M}\varepsilon^{(n)}, \quad (6.121a)$$

где матрица перехода равна

$$\mathbf{M} = (-\nabla_h^2 + \mathbf{KI})^{-1}(\mathbf{KI} - \mathbf{P}). \quad (6.121b)$$

Если собственные значения матрицы \mathbf{M} лежат в диапазоне $[-\rho, \rho]$, то для последовательных итераций можно применить ускорение по Чебышёву:

$$\mathbf{W}^{*(n+1)} = \omega_{n+1} (\mathbf{W}^{(n+1)} - \mathbf{W}^{*(n-1)}) + \mathbf{W}^{*(n-1)}, \quad (6.122)$$

где

$$\omega_0 = 1,$$

$$\omega_1 = 2/(2 - \rho^2), \quad (6.123)$$

и

$$\omega_{n+1} = 1/(1 - \rho^2\omega_n/4), \quad n = 1, 2, \dots$$

$\mathbf{W}^{*(n+1)}$ является уточненным значением $\mathbf{W}^{(n+1)}$, и $\mathbf{W}^{(n+1)}$ в уравнении (6.122) представляет решение уравнения (6.120), в котором $\mathbf{W}^{(n)}$ в правой части заменено на \mathbf{W}^{*n} . Спектральный радиус ρ можно оценить из отношения

$$\rho \simeq \|\mathbf{W}^{(n)} - \mathbf{W}^{(n-1)}\| / \|\mathbf{W}^{(n-1)} - \mathbf{W}^{(n-2)}\|. \quad (6.124)$$

Конкус и Голуб отмечают сходимость до погрешности $\sim 10^{-6}$ за 4 итерации, когда изменение $D(x, y)$ относительно гладкое. Для примеров с разрывом $D(x, y)$ после 10 итераций остается погрешность $\sim 10^{-2}$. Для таких задач рекомендуются матричные методы (см. разд. 6.4.1—6.4.6). Важное заключение состоит в том, что в отличие от случая итерационных методов SOR число итераций, требуемое для решения, по существу не зависит от количества сеточных узлов.

Кроме памяти для $p(x, y)$ и $q(x, y)$, необходимой для постановки задачи, неускоренная схема Голуба и Конкуса требует сетки в N_g узлов для вектора решения $\mathbf{W}^{(n)}$. Если используется ускорение по Чебышеву, то, кроме того, требуется сетка в N_g узлов для хранения $\mathbf{W}^{(n+1)}$.

Возможны дальнейшие усовершенствования итерации Конкуса и Голуба. Поскольку с помощью оптимального алгоритма FASR можно решить задачу, коэффициенты которой зависят только от одной из двух зависимых переменных (см. разд. 6.5.3), то можно предложить следующую итерацию:

$$[-\nabla^2 + X(x)] \mathbf{W}^{(n+1/2)} = [X(x) - p(x, y)] \mathbf{W}^{(n)} + q, \quad (6.125a)$$

$$[-\nabla^2 + Y(y)] \mathbf{W}^{(n+1)} = [Y(y) - p(x, y)] \mathbf{W}^{(n+1/2)} + q, \quad (6.125b)$$

где

$$\left. \begin{aligned} X(x) &= \frac{1}{L} \int p(x, y) dy, \quad 0 \leq y \leq L \\ \text{и} \quad Y(y) &= \frac{1}{L} \int p(x, y) dx, \quad 0 \leq x \leq L \end{aligned} \right\} \quad (6.125c)$$

Уравнение (6.125a) будет решаться с помощью алгоритма FASR с проведением преобразования Фурье в направлении y , а уравнение (6.125b) — с проведением преобразования Фурье в направлении x .

6.6. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

В этой главе мы пытались сделать обзор вопросов, связанных с выбором метода для решения уравнений поля, и краткий обзор главных особенностей наиболее важных методов. Мы надеемся, что этого будет достаточно, чтобы подсказать тип метода, скорее всего наиболее пригодный в каждом конкретном случае, и избежать выбора чрезвычайно неэффективных методов. Затем следует обратиться к оригинальным работам или к описаниям подходящих библиотечных программ за подробностями, которые мы здесь опустили.

Бесстолкновительные модели частиц

7.1. ВВЕДЕНИЕ

Физические системы, моделируемые методом частиц, математически можно описать с помощью иерархии функций распределения (см. гл. 1). При математическом исследовании таких систем иерархию обрывают с помощью некоторого замыкающего допущения, после чего находят приближенное решение получающейся системы нелинейных уравнений.

Замыкание самого низкого порядка устанавливается для разреженных систем, где разложение по малому параметру отношения потенциальной энергии к кинетической приводит к бесстолкновительному уравнению (уравнению Власова). Аналитическое рассмотрение уравнений Власова (см. гл. 2) проводится для малых времен посредством линеаризации и преобразования Фурье — Лапласа линеаризованных уравнений. Применительно к плазме этот подход дает в результате теорию плазменных волн [355], а для гравитирующих систем — теорию волн плотности в приближении туго закрученной спирали [257]. В последнем случае (а также в первом при наличии магнитных полей) анализ становится довольно сложным из-за необходимости интегрирования по невозмущенным орбитам для получения линеаризованного отклика. Теория линейного отклика распространяется на медленно меняющиеся системы посредством допущения, что в нулевом порядке распределение медленно меняется во времени (квазилинейная теория, см., например, [31]).

Следующее по порядку замыкание, также для разреженных систем, оставляет и уравнения для одночастичной функции распределения, и уравнения для парной корреляционной функции. Методы исследования этой системы уравнений разнообразны: от простого рассмотрения дискретного строения (зернистости) в виде последовательности мгновенных бинарных столкновений до всестороннего анализа, приводящего к уравнению Ленарда — Балеску — Гарнси¹⁾ [250, 20, 154]. В конечном счете все эти методы анализа приводят к уравнениям диффузии типа Фоккера — Планка для одночастичной функции распреде-

¹⁾ В отечественной литературе принято называть соответствующее кинетическое уравнение уравнением Ленарда — Балеску. — *Прим. перев.*

ления:

$$\left(\frac{\partial f}{\partial t}\right)_c = \frac{\partial}{\partial v_i} D_{ij} \frac{\partial f}{\partial v_j} + \frac{\partial}{\partial v_i} A_{ij} f. \quad (7.1)$$

В этих уравнениях вид коэффициентов диффузии D_{ij} и коэффициентов трения A_{ij} определяется использованными приближениями. В простейшей формулировке диффузия характеризуется временем столкновения τ_c , а трение — временем затухания τ_s (см. гл. 9).

Теория для плотных систем строится по той же схеме, что и для разреженных систем, но для получения требуемых замыкающих соотношений используются более сложные методы разложения [163, 331]. В результате таких разложений получается целый ряд замыканий, область успешного применения которых зависит от конкретных условий.

Полный охват статистической теории систем многих тел выходит далеко за рамки этой книги. Однако такая теория тесно связана с подходом к моделированию с помощью частиц и служила побудительной причиной многих предпринятых расчетов. В этой главе на примере бесстолкновительных систем, используя в качестве модели электростатическую плазму, мы хотим проиллюстрировать такую связь в обратном направлении: именно, как физическая теория может быть использована для анализа моделей частиц. Использование этой связи в обоих направлениях порождает процесс последовательного взаимного уточнения, в котором теоретическое рассмотрение приводит к более совершенным вычислительным моделям, которые в свою очередь обеспечивают более точную информацию, необходимую для улучшения теоретических приближений.

В следующем разделе мы получим кинетические уравнения для моделей частиц. Далее в разд. 7.3 будет кратко проиллюстрирован линейный анализ поведения мод для бесстолкновительных вычислительных моделей плазмы и в разд. 7.5 — некоторые наиболее яркие результаты, касающиеся столкновительных членов в кинетических уравнениях. В разд. 7.6—7.8 показывается, как результаты такого анализа можно применить для улучшения конструкций вычислительных моделей. Хотя получаемые нами результаты относятся к простому случаю электростатической бесстолкновительной плазмы, рассмотренные методы исследования можно распространить на другие вычислительные модели частица — сетка.

7.2. КИНЕТИЧЕСКИЕ УРАВНЕНИЯ

При выводе кинетических уравнений для физической системы, состоящей из N_p частиц, исходят либо из уравнения Лиувилля в $6N_p$ -мерном Γ -пространстве, т. е. уравнения (1.9), либо

из эквивалентного уравнения в 6-мерном μ -пространстве

$$\frac{\partial F}{\partial t} + \mathbf{v} \cdot \frac{\partial F}{\partial \mathbf{x}} + \mathbf{a} \cdot \frac{\partial F}{\partial \mathbf{v}} = 0, \quad (7.2)$$

где F обозначает точную одночастичную функцию распределения, определяемую выражением (1.10).

7.2.1. Предел малого шага по времени

Ограниченный объем памяти компьютера заставляет использовать в расчетах методом частица — сетка сравнительно грубые сетки по пространству. Шаг по времени ограничивается в меньшей степени, и в практических расчетах он берется гораздо меньше предельных значений, определяемых точностью и устойчивостью разностных аппроксимаций временных производных, входящих в уравнения движения. Поэтому имеет смысл сначала исследовать изменения в непрерывном математическом описании, вносимые введением сетки по пространству, не учитывая пока влияния конечного шага по времени DT .

В пределе $DT \rightarrow 0$ точные уравнения движения отличаются от приближенных только силой. Поэтому можно привлечь законы сохранения, используемые в непрерывном случае, изменив только определение силы (или, что то же самое, ускорения), фигурирующей в различных уравнениях. Уравнение Лиувилля для плотности вероятности ρ в пространстве Γ принимает вид

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{N_p} \left\{ \mathbf{v}_i \cdot \frac{\partial \rho}{\partial \mathbf{x}_i} + \frac{\mathbf{F}_i}{m} \cdot \frac{\partial \rho}{\partial \mathbf{v}_i} \right\} = 0, \quad (7.3)$$

где силы $\{\mathbf{F}_i\}$ отвечают методу РМ. Как отмечалось в гл. 1, иерархия уравнений для s -частичной функции распределения получается интегрированием (7.3) по всем остальным $N_p - s$ координатам. В случае $s = 1$ это дает уравнение для усредненной по ансамблю одночастичной функции распределения:

$$\frac{\partial f_1}{\partial t} + \mathbf{v} \cdot \frac{\partial f_1}{\partial \mathbf{x}} + \frac{\mathbf{F}_{\text{ext}}}{m} \cdot \frac{\partial f_1}{\partial \mathbf{v}} + \int \frac{\mathbf{F}(\mathbf{x}; \mathbf{x}')}{m} \cdot \frac{\partial f_2}{\partial \mathbf{v}} d\mathbf{x}' d\mathbf{v}'. \quad (7.4)$$

Здесь \mathbf{F}_{ext} обозначает внешние силы, действующие на данную систему; $\mathbf{F}(\mathbf{x}; \mathbf{x}')$ — сила, действующая на частицу в точке \mathbf{x} со стороны другой частицы, находящейся в точке \mathbf{x}' , а $f_2(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t)$ — усредненная по ансамблю двухчастичная функция распределения.

Полагая в (7.4) $f_2 = f_1 f_1$, получим уравнение Власова (т. е. бесстолкновительное уравнение) для модели частица — сетка. Уравнение Ленарда [250], относящееся к уравнениям типа Фоккера — Планка, выводится из уравнений, которые получаются подстановкой в (7.4) $f_2 = f_1 f_1 + g$ и в соответствующее уравне-

ние для f_2 : $f_2 = f_1 f_1 + g$, $f_3 = f_1 f_1 f_1 + \Sigma f_1 g$, после чего используется приближение об упорядочении временных масштабов, состоящее в том, что характерное время релаксации f_2 к своему асимптотическому значению во много раз меньше характерного времени изменения f_1 . В соответствии с этим и, кроме того, в предположении пространственной однородности, уравнение для g сводится к выражению, в котором g является функционалом только от f_1 , и уравнение для f_1 принимает вид

$$\frac{\partial f_1}{\partial t} = - \int \frac{F(\mathbf{x}; \mathbf{x}')}{m} \cdot \frac{\partial}{\partial \mathbf{v}} g(\mathbf{x}, \mathbf{x}', \mathbf{v}, \mathbf{v}') d\mathbf{x}' d\mathbf{v}'. \quad (7.5)$$

Следовательно, заменяя точную силу взаимодействия частиц сеточным приближением и повторяя известные этапы рассмотрения физической системы (с использованием описанных в гл. 5 спектральных методов), можно получить кинетическое уравнение для моделей частиц в пределе $DT \rightarrow 0$.

7.2.2. Конечный шаг по времени

Если рассматривать конечный шаг по времени, то прямое копирование анализа дифференциальных уравнений для непрерывной системы более невозможно. Если раньше уравнения движения давали характеристики дифференциального уравнения для плотности вероятности, то теперь вместо этого они дают правила перехода для последовательных состояний системы. Поэтому если, например, эволюция системы определяется схемой с перешагиванием и ее состояние характеризуется на n -м временном слое координатами $\{\mathbf{x}_i = \mathbf{x}_i^n; \mathbf{v}_i = (\mathbf{v}_i^{n+1/2} + \mathbf{v}_i^{n-1/2})/2; i \in \in[1, N_p]\}$, то закон сохранения в Γ -пространстве можно выразить в виде

$$\rho(\{\mathbf{x}_i\}, \{\bar{\mathbf{v}}_i\}, t^n) = \rho(\{\mathbf{x}'_i\}, \{\bar{\mathbf{v}}'_i\}, t^{n+1}). \quad (7.6)$$

В бесстолкновительном (власовском) приближении частицы некоррелированы, поэтому плотность вероятности ρ может быть записана как произведение одночастичных функций распределения

$$\rho(\{\mathbf{x}_i\}, \{\bar{\mathbf{v}}_i\}, t^n) = \prod_{i=1}^{N_p} f_1(\mathbf{x}_i, \bar{\mathbf{v}}_i, t^n), \quad (7.7)$$

и в этом случае уравнение (7.6) может быть приведено к уравнению Власова для моделей частиц с конечным шагом по времени, т. е.

$$f_1(\mathbf{x}, \bar{\mathbf{v}}, t^n) = f_1(\mathbf{x}', \bar{\mathbf{v}}', t^{n+1}), \quad (7.8)$$

где $\mathbf{x} = \mathbf{x}^n$, $\bar{\mathbf{v}} = (\mathbf{v}^{n+1/2} + \mathbf{v}^{n-1/2})/2$, $\mathbf{x}' = \mathbf{x}^{n+1}$, $\bar{\mathbf{v}}' = (\mathbf{v}^{n+3/2} + \mathbf{v}^{n+1/2})/2$

связаны между собой разностными уравнениями движения (схема с перешагиванием). Следующее по порядку замыкание получается в предположении, что

$$f_2(\mathbf{x}, \bar{\mathbf{v}}, \mathbf{x}', \bar{\mathbf{v}}', t^n) = f_1(\mathbf{x}, \bar{\mathbf{v}}, t^n) f_1(\mathbf{x}', \bar{\mathbf{v}}', t^n) + g(\mathbf{x}, \mathbf{v}, \mathbf{x}', \mathbf{v}', t^n), \quad (7.9)$$

$$\rho = \prod_{i=1}^{N_p} f_i + \prod_{i=1}^{l-1} g \prod_{k \neq i, j}^{N_p} f_i. \quad (7.10)$$

Далее можно проинтегрировать уравнение (7.6) по всем координатам, кроме первой и второй, и получить кинетическое уравнение для слабо коррелированных систем.

7.3. ДИСПЕРСИОННОЕ СООТНОШЕНИЕ

Дисперсионное соотношение $\omega = \omega(\mathbf{k})$ представляет собой условие согласованности для волн малой амплитуды. Оно дает информацию о характерных временных и пространственных масштабах, фазовой и групповой скоростях волн, а также об устойчивости данной конфигурации. В непрерывном случае диапазон волновых чисел неограничен. С введением дискретной сетки по пространству с шагом H область возможных волновых чисел сокращается до основной зоны (см. разд. 5.6), т. е. $|k| < \pi/H$. Любая гармоника с волновым числом k , большим $|k| = \pi/H$ (т. е. с длиной волны меньше $2H$), проявляется на сетке в виде вклада в гармонику основной зоны с волновым числом $k' = k \bmod k_g$. К аналогичным последствиям приводит и дискретизация по времени. Введение конечного шага по времени DT ограничивает область частот ω диапазоном $|\omega| < \pi/DT$. Поскольку, однако, устойчивость схемы интегрирования по времени требует $|\omega|DT < 2$ (для схемы с перешагиванием см. гл. 4), то результаты вычислений со слишком большим шагом по времени, который не отслеживает частоты системы, будут в любом случае явно неверными.

7.3.1. Предел малого шага по времени

Полагая в уравнении (7.4) f_2 равной произведению одночастичных функций распределения, получаем уравнение Власова для моделей частиц в пределе нулевого шага по времени:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \left\{ \frac{\mathbf{F}^{\text{ext}}(\mathbf{x})}{m} + \int \frac{\mathbf{F}(\mathbf{x}; \mathbf{x}')}{m} f(\mathbf{x}', \mathbf{v}', t) d\mathbf{x}' d\mathbf{v}' \right\} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \quad (7.11)$$

В этом уравнении и в дальнейшем изложении индекс 1 у одночастичной функции распределения будем опускать.

Рассмотрим для простоты одномерный вариант уравнения (7.11) и положим шаг пространственной сетки в бесконечной системе равным H . Уравнение (7.11) принимает вид

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \left\{ \frac{F^{\text{ext}}}{m} + \int \frac{F(x, x')}{m} n(x') dx' \right\} \frac{df}{dv} = 0, \quad (7.12)$$

где $n(x)$ обозначает плотность частиц,

$$n(x) = \int f(x, v') dv', \quad (7.13)$$

а $F(x, x')$ — сила, действующая в точке x со стороны заряда в точке x' , которая вычисляется по схеме частица — сетка. Используя введенную в разд. 5.6 компактную форму записи, можно записать этапы цикла вычисления силы следующим образом:

Распределение заряда

$$\rho^+(x) = \Pi\left(\frac{x}{H}\right) \frac{q}{H} W(x) * n(x). \quad (7.14)$$

Нахождение потенциала

$$\varphi'(x) = G'(x) * \rho^+(x). \quad (7.15)$$

Интерполяция силы (схема, сохраняющая импульс)

$$E'(x) = -D(x) * \varphi'(x), \quad (7.16)$$

$$E^+(x) = \Pi\left(\frac{x}{H}\right) E'(x), \quad (7.17)$$

$$F(x) = \frac{q}{H} W(x) * E^+(x), \quad (7.18)$$

где штрихом отмечены непрерывные функции с конечной шириной спектра, построенные по дискретным наборам значений, а крестиками — функции, состоящие из гребенки импульсов, амплитуда которых равна сеточным значениям в точках расположения этих импульсов. В частности, для уединенного заряда, помещенного в точке $x = x'$, выражение (7.14) принимает вид

$$\rho^+(x) = \Pi\left(\frac{x}{H}\right) \frac{q}{H} W(x) * \delta(x - x'). \quad (7.19)$$

Объединяя (7.15) — (7.19), имеем

$$F(x, x') = -\frac{q^2}{H^2} W(x) * \Pi\left(\frac{x}{H}\right) \times \\ \times \left[D(x) * G'(x) * \Pi\left(\frac{x}{H}\right) (W(x) * \delta(x - x')) \right]. \quad (7.20)$$

Зависимость $F(x, x')$ от x' обусловлена лишь членом $\delta(x - x')$, связанным с точечным зарядом, поэтому мы можем символи-

чески представить (7.20) в виде

$$F(x, x') = h(x) * \delta(x - x'). \quad (7.21)$$

Отсюда

$$\begin{aligned} \int F(x, x') n(x') dx' &= h(x) * \int \delta(x - x') n(x') dx' = h(x) * n(x) \\ &= -\frac{q^2}{H^2} W(x) * \Pi\left(\frac{x}{H}\right) \left[D(x) * G'(x) * \Pi\left(\frac{x}{H}\right) (W(x) * n(x)) \right]. \end{aligned} \quad (7.22)$$

Соотношение (7.22) выражает силу в точке x , вычисленную по схеме частица — сетка и обусловленную плотностью $n(x)$. Полная сила в точке x получается добавлением к (7.22) внешней силы, в которой отражен вклад неподвижного нейтрализующего заряженного фона. Таким образом, по аналогии с непрерывным случаем, уравнение (7.12) можно переписать в виде

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + \frac{F}{m} \frac{\partial f}{\partial v} = 0, \quad (7.23)$$

где F описывает теперь полную силу в точке x , обусловленную распределением f и заряженным фоном. То же самое справедливо и для варианта модели частица — сетка, сохраняющего энергию. В обоих случаях результаты обобщаются на системы, состоящие из частиц многих сортов, для чего для новых сортов добавляются дополнительные переменные i , где необходимо, производят суммирование по сортам.

Уравнение (7.23) по форме совпадает с точным описанием системы N_p тел в μ -пространстве, т. е. уравнением (7.2). Оно отличается от точного в двух отношениях: точная одночастичная функция распределения заменена на гладкую усредненную по ансамблю функцию распределения f , и ускорение получается из усредненного по ансамблю, а не точного пространственного распределения частиц.

Простейшее усредненное по ансамблю (бесстолкновительное) состояние, которое поддается анализу, — однородная зарядово-нейтральная система. Такая система находится в равновесии:

$$f = f(v), \quad F = 0. \quad (7.24)$$

Для описания малых отклонений от равновесного состояния в уравнении (7.23) полагают $f = f^0(v) + f^1(x, v, t)$, $F = F^1$ и т. д. и удерживают члены только первого порядка:

$$\frac{\partial f^1}{\partial t} + v \frac{\partial f^1}{\partial x} = -\frac{F^1}{m} \frac{\partial f^0}{\partial v}. \quad (7.25)$$

Делая преобразование Фурье по пространству и по времени, получаем уравнение стационарного отклика

$$-i(\omega - kv) \hat{f}^1(k, v, \omega) = -\frac{\hat{F}(k, \omega)}{m} \frac{\partial f^0}{\partial v}. \quad (7.26)$$

(Строго говоря, следовало бы использовать по времени преобразование Лапласа и вывести уравнение (7.26) из асимптотического поведения решения задачи Коши; см., например, [355].) Перегруппировка членов в (7.26) и интегрирование по v дают в первом порядке гармоника плотности

$$\hat{n}(k) = \int \hat{f}^1 dv = -i \frac{\hat{F}(k)}{m} \int \frac{1}{\omega - kv} \frac{\partial f^0}{\partial v} dv. \quad (7.27)$$

Для уравнений (7.14)–(7.18) преобразование Фурье дает

$$\begin{aligned} \hat{\rho}(k) &= \frac{q}{H} H \Pi \left(\frac{k}{k_g} \right) * \hat{W}(k) \hat{n}(k) = \\ &= \frac{q}{H} \sum_{n=-\infty}^{\infty} \hat{W}(k - nk_g) \hat{n}(k - nk_g), \end{aligned} \quad (7.28)$$

$$\hat{\phi}'(k) = \hat{G}'(k) \hat{\rho}(k), \quad (7.29)$$

$$\hat{E}'(k) = -\hat{D}(k) \hat{\phi}'(k), \quad (7.30)$$

$$\hat{E}(k) = H \Pi \left(\frac{k}{k_g} \right) * \hat{E}'(k) = -\hat{D}(k) \hat{\phi}(k), \quad (7.31)$$

$$\hat{F}(k) = \frac{q}{H} \hat{W}(k) \hat{E}(k), \quad (7.32)$$

где крестики в обозначениях опущены, поскольку $\hat{\rho}^\dagger = \hat{\rho}$ и $\hat{E}^\dagger = \hat{E}$. В (7.31) использованы свойства периодичности \hat{D} и конечности ширины спектра штрихованных величин, что позволяет записать $H \Pi(k/k_g) * \hat{\phi}'(k) = \hat{\phi}(k)$ (более подробно см. разд. 5.6). Умножая (7.27) на $q\hat{W}/H$ и образуя сумму, фигурирующую в правой части (7.28), имеем

$$\hat{\rho}(k) = -i \sum_n \frac{q\hat{W}(k_n)}{H} \frac{\hat{F}(k_n)}{m} \int \frac{1}{(\omega - k_n v)} \frac{\partial f^0}{\partial v} dv. \quad (7.33)$$

Записывая с помощью (7.29)–(7.33) \hat{F} как функцию $\hat{\rho}$, вынося $\hat{\rho}$, \hat{G} и \hat{D} в силу их периодичности за сумму в правой части (7.33) и, наконец, деля обе части на $\hat{\rho}$, получаем дисперсионное соотношение для (бесстолкновительной) плазменной модели частица — сетка, сохраняющей импульс:

$$1 = -(-i\hat{D})(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 \int \frac{1}{\omega - k_n v} \frac{\partial f^0}{\partial v} dv. \quad (7.34)$$

В (7.33) и (7.34) выражение $k - nk_g$ сокращенно обозначено через k_n , а интегралы по v берутся по контуру Ландау в комплексной v -плоскости. Эти интегралы можно переписать через их главные значения и вклады от особенностей:

$$1 = -(-i\hat{D})(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 \times \\ \times \left(P \int_{-\infty}^{\infty} \frac{\frac{\partial f^0}{\partial v} dv}{\omega - k_n v} - \frac{i\pi}{|k_n|} \frac{\partial f^0}{\partial v} \Big|_{\omega/k_n} \right). \quad (7.35)$$

Для схем, сохраняющих энергию, получим похожее дисперсионное соотношение. Следуя вышеприведенной процедуре с заменой (7.16) — (7.18) на соответствующие выражения для схем, сохраняющих энергию, а именно

$$\varphi^\dagger(x) = \Pi\left(\frac{x}{H}\right) \varphi'(x), \quad (7.36)$$

$$F(x) = -\frac{q}{H} \left(\frac{d}{dx} W(x) \right) * \varphi^\dagger(x), \quad (7.37)$$

где

$$F(x) \supset \hat{F}(k) = -\frac{q}{H} ik \hat{W}(k) \hat{\varphi}(k), \quad (7.38)$$

получим

$$1 = -(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} k_n \left(\frac{W(k_n)}{H} \right)^2 \times \\ \times \left(P \int_{-\infty}^{\infty} \frac{\frac{\partial f^0}{\partial v} dv}{\omega - k_n v} - \frac{i\pi}{|k_n|} \frac{\partial f^0}{\partial v} \Big|_{\omega/k_n} \right). \quad (7.39)$$

7.3.2. Конечный шаг по времени

Если в дисперсионном анализе сохранять конечный DT , то в этом случае надо пользоваться уравнением Власова в форме (7.8). Используя схему с перешагиванием (уравнения (4.3) и (4.4)), можно представить (7.8) для одномерной системы в виде

$$f(x, \bar{v}, t) = f\left(x + \bar{v} DT + a(x, t^n) \frac{DT^2}{2}, \right.$$

$$\left. \bar{v} + [a(x, t^n) + a(x', t^{n+1})] \frac{DT}{2}, t + DT \right), \quad (7.40)$$

где $a(x, t^n)$ — ускорение в точке x в момент времени t^n . Соответствующее выражение для многомерных систем получается заменой x , \bar{v} и a их векторными аналогами.

Разлагая (7.40) в ряд Тейлора, имеем

$$f(x, \bar{v}, t) = f(x + \bar{v} DT, \bar{v}, t + DT) + a(x, t^n) \frac{DT^2}{2} \frac{\partial f}{\partial x} + \\ + [a(x, t^n) + a(x', t^{n+1})] \frac{DT}{2} \frac{\partial f^0}{\partial v} + \dots \quad (7.41)$$

Уравнение (7.41) линеаризуется подстановкой

$$f = f^0 + f^1, \quad a = a^0 + a^1 \quad (7.42)$$

с удержанием членов только первого порядка. В частности, в случае однородной плазмы

$$f^0 = f^0(v), \quad a^0 = 0 \quad (7.43)$$

получается

$$f^1(x, v, t) = f^1(x + v DT, v, t + DT) + \\ + [a^1(x, t^n) + a^1(x + v DT, t + DT)] \frac{DT}{2} \frac{\partial f^0}{\partial v}. \quad (7.44)$$

Выполняя преобразование (7.44) по пространству и времени (что дает гармоники $\sim \exp i(kx - \omega t)$) и применяя теорему о сдвиге (табл. П 4) к аргументам $x + vDT$ и $t + DT$, имеем

$$\hat{f}^1(k, v, \omega) = \hat{f}^1(k, v, \omega) e^{ikv DT} e^{-i\omega DT} + \\ + [\hat{a}^1(k, \omega) + \hat{a}^1(k, \omega) e^{ikv DT} e^{-i\omega DT}] \frac{DT}{2} \frac{\partial f^0}{\partial v}, \quad (7.45)$$

что после приведения подобных членов дает

$$\hat{f}^1(k, v, \omega) = -i \frac{DT}{2} \frac{\hat{F}}{m} \operatorname{ctg} \left((\omega - kv) \frac{DT}{2} \right) \frac{\partial f^0}{\partial v}, \quad (7.46)$$

где вместо \hat{a} подставлено \hat{F}/m . Теперь можно повторить выкладки, отвечающие (7.27) — (7.34), используя вместо (7.26) выражение (7.46), и получить дисперсионное соотношение для плазменных моделей частица — сетка, сохраняющих импульс:

$$1 = -(-i\hat{D})(\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 \int \frac{DT}{2} \times \\ \times \operatorname{ctg} \left((\omega - k_n v) \frac{DT}{2} \right) \frac{\partial f^0}{\partial v} dv. \quad (7.47)$$

Для моделей, сохраняющих энергию, аналогично находим:

$$1 = -(\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} k_n \left(\frac{\hat{W}(k_n)}{H} \right)^2 \int \frac{DT}{2} \times \\ \times \operatorname{ctg} \left((\omega - k_n v) \frac{DT}{2} \right) \frac{\partial f^0}{\partial v} dv. \quad (7.48)$$

Используя тождество

$$\operatorname{ctg} x = \sum_{n=-\infty}^{\infty} \frac{1}{(x - \pi n)}, \quad (7.49)$$

можно записать интегралы в (7.47) и (7.48) в той же форме, которая получена в непрерывном случае, т. е.

$$\int \frac{DT}{2} \operatorname{ctg} \left((\omega - k_n v) \frac{DT}{2} \right) \frac{\partial f^0}{\partial v} dv = \sum_{m=-\infty}^{\infty} \int \frac{\frac{\partial f^0}{\partial v}}{(\omega_m - k_n v)} dv, \quad (7.50)$$

где

$$\omega_m = \omega - \frac{2\pi}{DT} m. \quad (7.51)$$

В результате учета конечного шага по времени в область частот вносится периодичность с периодом $2\pi/DT$. Частоты, которые могут быть представлены в моделях частица — сетка, ограничены основной зоной $|\omega| \leq \pi/DT$. В расчетах с конечным шагом по времени все частоты ω_m неотличимы от основной частоты ω .

7.3.3. Приближение тепловой плазмы

Сравнивая дисперсионные соотношения, которые получаются из (7.35) или (7.39), с найденным для непрерывного случая (2.7), можно исследовать вопрос, как сказывается дискретная пространственная сетка, а также вид функции влияния \hat{G} и интерполяции силы на дисперсионных свойствах модельной плазмы. Для учета влияния конечного шага по времени можно воспользоваться уравнениями (7.47) и (7.48). Если имеются магнитные поля, то, следуя методу, применяемому для непрерывной системы [355], для получения дисперсионных соотношений дискретных моделей f^1 выражают через интеграл вдоль невозмущенных орбит. Однако даже в простейших приближениях наличие в уравнениях системы частица — сетка сумм по сеточным гармоникам обычно не позволяет получить явное аналитическое решение $\omega = \omega(k)$, хотя в некоторых случаях и удастся получить замкнутые выражения, содержащие тригонометрические функции.

В приближении тепловой плазмы при вычислении интегралов по скорости, входящих в дисперсионные соотношения, максвелловская функция распределения

$$f^0(v) = \frac{n_0}{v_T \sqrt{2\pi}} \exp\left(-\frac{v^2}{2v_T^2}\right) \quad (7.52)$$

аппроксимируется ее первыми двумя моментами. Формально этой процедуре отвечает замена в интегралах по скорости функ-

ции (7.52) на

$$f^0(v) = \frac{n_0}{2\sqrt{3}v_T} \Pi\left(\frac{v}{2\sqrt{3}v_T}\right). \quad (7.53)$$

Это дает

$$P \int_{-\infty}^{\infty} \frac{\partial f^0/\partial v}{\omega - k_n v} dv \simeq \frac{-n_0 k_n}{\omega^2 - 3k_n^2 v_T^2}. \quad (7.54)$$

Аналогично, для сдвинутого максвелловского распределения с направленной скоростью v_0

$$P \int_{-\infty}^{\infty} \frac{\partial f^0/\partial v}{\omega - k_n v} dv \simeq \frac{-n_0 k_n}{(\omega - k_n v_0)^2 - 3k_n^2 v_T^2}. \quad (7.55)$$

Подставляя (7.54) в (7.34), получаем дисперсионное соотношение для модельной плазмы, сохраняющей импульс, в приближении тепловой плазмы:

$$1 = (-i\hat{D})(\epsilon_0 \hat{G}) \omega_p^2 \sum_n \left(\frac{\hat{W}}{H}\right)^2 \frac{k_n}{\omega^2 - 3k_n^2 v_T^2}. \quad (7.56)$$

В частности, для одномерной плазменной модели, описанной в гл. 2, выбор двухточечной конечно-разностной аппроксимации (КРА) для производной, трехточечной КРА для лапласиана и весовой функции NGR дают соответственно

$$(-i\hat{D}) = \frac{\sin kH}{H} = \frac{\sin \frac{kH}{2} \cos \frac{kH}{2}}{(H/2)}, \quad (7.57)$$

$$\epsilon_0 \hat{G} = \frac{(H/2)^2}{\sin^2 \frac{kH}{2}}, \quad (7.58)$$

$$\frac{\hat{W}}{H} = \frac{\sin \frac{kH}{2}}{\frac{kH}{2}}. \quad (7.59)$$

В результате уравнение (7.56) принимает вид

$$1 = \omega_p^2 \operatorname{ctg} \frac{k_n H}{2} \sum_n \frac{\sin^2 \frac{k_n H}{2}}{\frac{k_n H}{2}} \frac{1}{\omega^2 - 3k_n^2 v_T^2}. \quad (7.60)$$

Используя соотношение

$$\sin^2 \frac{k_n H}{2} = \sin^2 \left(\frac{kH}{2} - n\pi \right) = \sin^2 \frac{kH}{2}, \quad (7.61)$$

можно вынести синус за знак суммы, что дает

$$1 = \omega_p^2 \cos \frac{kH}{2} \sin \frac{kH}{2} \sum_{n=-\infty}^{\infty} \frac{1}{\frac{k_n H}{2} (\omega^2 - 3k_n^2 v_T^2)}. \quad (7.62)$$

В пределе «холодной» плазмы $v_T \rightarrow 0$ можно применить к (7.62) тождество (7.49) и получить

$$\omega^2 = \omega_p^2 \cos^2 \frac{kH}{2}. \quad (7.63)$$

Для иерархии весовых функций схем NGP ($p=1$), CIC ($p=2$), TSC ($p=3$) и т. д.,

$$\frac{\hat{W}(k)}{H} = \left(\frac{\sin \frac{kH}{2}}{\frac{kH}{2}} \right)^p, \quad (7.64)$$

дисперсионное соотношение для тепловой плазмы принимает вид

$$1 = \omega_p^2 (-i\hat{D})(\epsilon_0 \hat{G}) \left(\sin \frac{kH}{2} \right)^{2p} \frac{2}{H} \sum_n \frac{1}{\left(\frac{k_n H}{2} \right)^{2p-1}} \frac{1}{\omega^2 - 3k_n^2 v_T^2}. \quad (7.65)$$

Тождество

$$\frac{(-1)^s}{s!} \frac{d^s}{dx^s} \operatorname{ctg} x = \sum_{n=-\infty}^{\infty} \frac{1}{(x - \pi n)^{s+1}} \quad (7.66)$$

позволяет записать (7.65) в пределе холодной плазмы в виде

$$\omega^2 = \omega_p^2 (-i\hat{D})(\epsilon_0 \hat{G}) \frac{2}{H} \times \begin{cases} \cos \frac{kH}{2} \sin \frac{kH}{2}, & \text{NGP} \\ \cos \frac{kH}{2} \sin \frac{kH}{2}, & \text{CIC} \\ \cos \frac{kH}{2} \frac{\left(2 + \cos^2 \frac{kH}{2} \right) \sin \frac{kH}{2}}{3}, & \text{TSC.} \end{cases} \quad (7.67)$$

Если в проведенном рассмотрении учитывать конечный шаг по времени DT , дисперсионное соотношение для тепловой плазмы (7.65) заменится на

$$1 = \omega_p^2 (-i\hat{D})(\epsilon_0 \hat{G}) \left[\sin \frac{kH}{2} \right]^{2p} \left(\frac{2}{H} \right) \sum_n \sum_m \frac{1}{\left(\frac{k_n H}{2} \right)^{2p-1}} \frac{1}{(\omega_m^2 - 3k_n^2 v_T^2)}. \quad (7.68)$$

В пределе холодной плазмы (7.68) с учетом (7.66) приводит к дисперсионным соотношениям, отличающимся от (7.67) лишь заменой ω^2 на $\sin^2(\omega DT/2) / (DT/2)^2$.

Дисперсионное соотношение для тепловой плазмы можно получить в замкнутой тригонометрической форме, применяя в дисперсионном соотношении к выражениям, входящим в суммы по сеточным гармоникам, метод элементарных дробей и вычисляя эти суммы с помощью разложения (7.66). Например, в одномерной модели NGR, описанной в гл. 2, дисперсионное соотношение для плазмы с тепловой скоростью v_T равняется (для $DT = 0$):

$$1 = \omega_p^2 \cos \frac{kH}{2} \sin \frac{kH}{2} \sum_n \frac{1}{\left(\frac{k_n H}{2} \omega^2 - 3k_n^2 v_T^2\right)}. \quad (7.69)$$

Слагаемое в сумме по сеточным гармоникам можно представить в виде

$$\frac{y^2}{\omega^2} \frac{1}{x_n (y - x_n) (y + x_n)} = \frac{1}{\omega^2} \left[\frac{1}{x_n} - \frac{1}{2} \left(\frac{1}{x_n - y} + \frac{1}{x_n + y} \right) \right], \quad (7.70)$$

где

$$y = \frac{\omega H}{2 \sqrt{3} v_T}, \quad x_n = \frac{k_n H}{2}, \quad (7.71)$$

что после некоторых преобразований дает

$$\omega^2 = \omega_p^2 \left[\cos^2 \frac{kH}{2} + \frac{1}{2} \left(\operatorname{ctg} \frac{\omega H}{\sqrt{3} v_T} - \operatorname{ctg} kH \right)^{-1} \right]. \quad (7.72)$$

В табл. 7.1 и 7.2 приведены дисперсионные соотношения электростатической плазмы для схем, сохраняющих соответственно импульс и энергию. Дисперсионное соотношение с усредненной силой — это соотношение, которое получается при использовании силы, усредненной по смещениям относительно сетки, причем это усреднение приводит к тому, что в сумме по сеточным гармоникам обращаются в нуль все члены с $n \neq 0$ (см разд. 5.6.4).

Дисперсионные соотношения с усредненной силой можно записать в виде

$$1 = - \frac{\omega_p^2 k}{n_0 k^2} \hat{S}_e^2(k) \int \frac{\partial f^0}{\partial v} \frac{dv}{\omega - kv}, \quad (7.73)$$

где для схем, сохраняющих импульс,

$$\hat{S}_e^2(k) = \left(- \frac{i\hat{D}}{k} \right) (\epsilon_0 \hat{G} k^2) \left(\frac{\hat{W}}{H} \right)^2. \quad (7.74)$$

а для схем, сохраняющих энергию,

$$\hat{S}_e^2(k) = (\epsilon_0 \hat{G} k^2) \left(\frac{\hat{W}}{H} \right)^2. \quad (7.75)$$

Таблица 7.1. Дисперсионные соотношения для плазменных моделей частица-сетка, сохраняющих импульс

 $DT \neq 0$

$$1 = -(-i\hat{D})(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega_m - k_n v} dv$$

 $DT = 0$

$$1 = -(-i\hat{D})(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega - k_n v} dv$$

 $DT = 0$, усредненная сила

$$1 = -(-i\hat{D})(\epsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \left(\frac{\hat{W}(k)}{H} \right)^2 \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega - kv} dv$$

Тепловая плазма, $DT \neq 0$

$$1 = (-i\hat{D})(\epsilon_0\hat{G}) \omega_p^2 \sum_n \sum_m \left(\frac{\hat{W}(k_n)}{H} \right)^2 \frac{k_n}{(\omega_m - k_n v_0)^2 - 3k_n^2 v_T^2}$$

$$\int_C \frac{\frac{\partial f^0}{\partial v}}{\omega_m - k_n v} dv = P \int_{-\infty}^{\infty} \frac{\frac{\partial f^0}{\partial v}}{\omega_m - k_n v} dv - \frac{i\pi}{|k_n|} \frac{\partial f^0}{\partial v} \Big|_{\omega_m/k_n}$$

$$k_n = k - n \frac{2\pi}{H},$$

$$\omega_m = \omega - m \frac{2\pi}{DT},$$

 v_0 — направленная скорость v_T — тепловая скорость

В вычислительных моделях частиц $S(x) \Rightarrow \hat{S}(k)$ можно интерпретировать как эффективную форму частицы конечного размера. Как показано в работах [38, 238, 299], если отвечающая $S(x)$ ширина порядка дебаевской длины, свойства модели частиц конечного размера близки к свойствам точной бесстолкновительной системы.

Дисперсионное уравнение (7.56) было решено численно для потока плазмы с тепловым разбросом по скоростям (с заменой ω на $\omega - k_n v_0$, где v_0 — скорость потока) и подтверждено результатами численного эксперимента [37]. Было установлено, что развивающиеся в холодном плазменном потоке нефизические неустойчивости [233] выходят на насыщение, как только отношение дебаевской длины к сеточному шагу λ_D/H делается

Таблица 7.2. Дисперсионные соотношения для плазменных моделей частица-сетка, сохраняющих энергию

$DT \neq 0$

$$1 = - (\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 k_n \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega_m - k_n v} dv$$

$DT = 0$

$$1 = - (\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \sum_{n=-\infty}^{\infty} \left(\frac{\hat{W}(k_n)}{H} \right)^2 k_n \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega - k_n v} dv$$

$DT = 0$, усредненная сила

$$1 = - (\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \left(\frac{\hat{W}(k_n)}{H} \right)^2 k \int_C \frac{\frac{\partial f^0}{\partial v}}{\omega - kv} dv$$

Тепловая плазма, $DT \neq 0$

$$1 = (\epsilon_0 \hat{G}) \omega_p^2 \sum_n \sum_m \left(\frac{\hat{W}(k_n)}{H} \right)^2 \frac{k_n^2}{(\omega_m - k_n v_0)^2 - 3k_n^2 v_T^2}$$

достаточно большим, и что если в начальный момент тепловой разброс скоростей достаточно велик, то неустойчивости не возникают вовсе. Эти результаты отражены на рис. 7.1. Если мо-

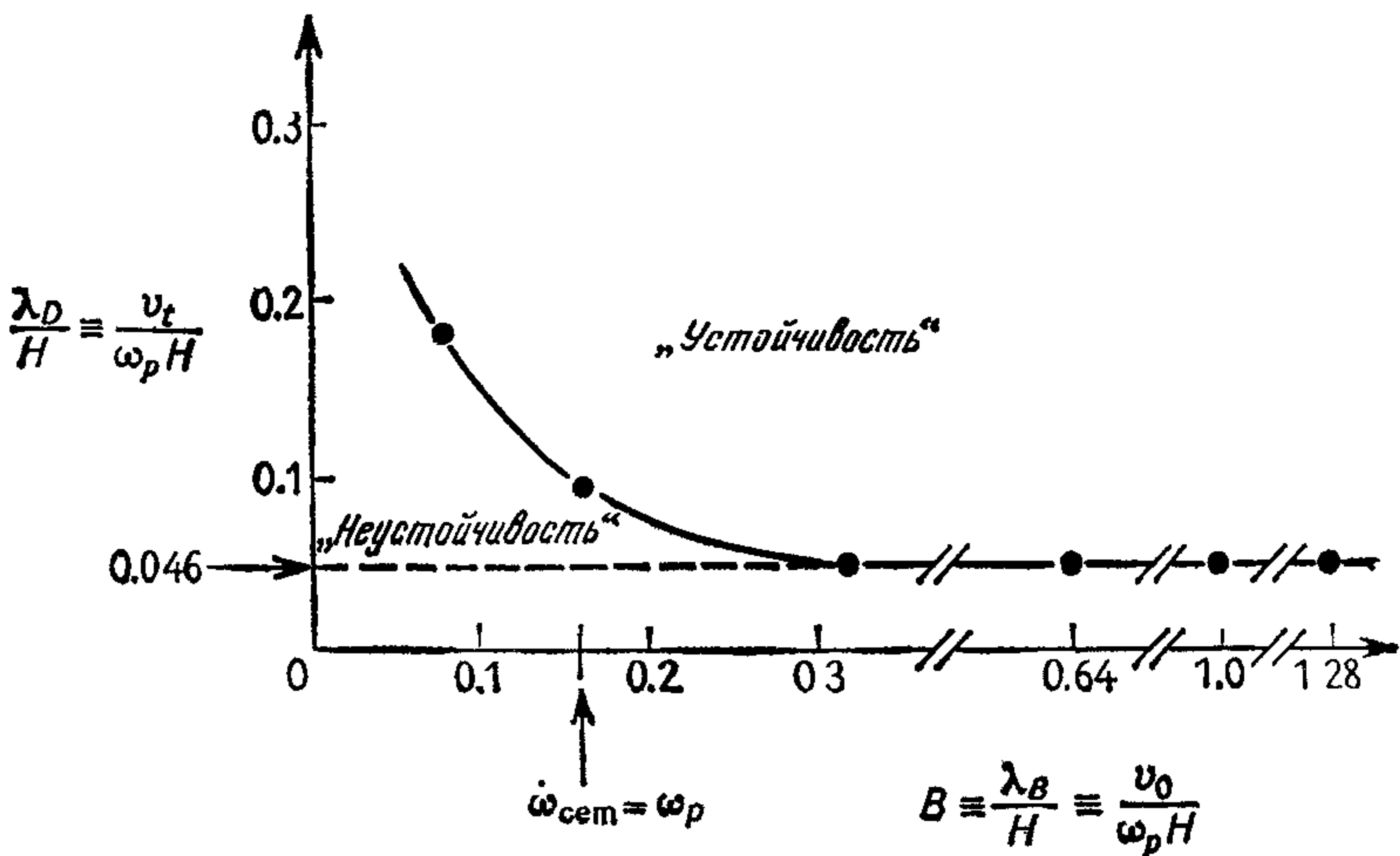


Рис. 7.1 Экспериментальное определение теплового разброса, необходимого для устойчивости пучка в программе моделирования плазмы, сохраняющей импульс (из работы [37], рис. 10, с любезного разрешения Journal of Computational Physics).

делирование начинается в области, помеченной как «неустойчивость», то нефизический рост энергии сопровождается повышением температуры до тех пор, пока система не перейдет в область, обозначенную как «устойчивость». Если же начальные условия моделирования заданы выше пороговой кривой, то неустойчивость отсутствует. В отличие от неустойчивостей, возникающих из-за слишком большого шага по времени ($DT > 2/\omega_p$, см. гл. 4), данная неустойчивость, являющаяся результатом выбора $N \gg \lambda_D$, не опасна в том отношении, что ее действие направлено на перевод системы из начального состояния в состояние, которое может быть адекватно представлено дискретной сеткой.

7.3.4. Взаимодействие мод

Лэнгдон [234] численно решил уравнение (7.35) и соответствующее дисперсионное уравнение с усредненной силой (табл. 7.1) для максвелловской f^0 и приведенными в (7.57) — (7.59) функциями D , G и W . Результаты его расчетов представлены на рис. 7.2 и 7.3. На рис. 7.2 изображены зависимости вещественной и мнимой частей частоты от волнового числа k в основной зоне для $N = \lambda_D$. В этом случае влияние членов с $n \neq 0$ в сумме по сеточным гармоникам ничтожно мало, за исключением очень коротких волн. На рис. 7.3, где размер ячейки в десять раз превышает дебаевскую длину, ситуация коренным образом меняется. Обе приведенные зависимости вещественной части частоты, и соответствующая точному дисперсионному уравнению, и соответствующая уравнению с усредненной силой, качественно ведут себя одинаково и дают групповую скорость $d\omega/dk$, по знаку противоположную скорости, полученной для власовской плазмы ($d\omega/dk \simeq 6k\lambda_D^2$). Более необычно поведение мнимой части частоты, где слабое затухание Ландау в относительно длинноволновой области сменяется неустойчивостью за счет вклада сеточных мод. Таким образом, если $N \gg \lambda_D$, то поведение модельной плазмы в качественном (как, впрочем, и в количественном) отношении является неверным, поскольку физически максвелловская плазма устойчива. Существование этой неустойчивости и ее инкремент были подтверждены в численных экспериментах Окуды [296, 297].

Нелинейный характер такой неустойчивости похож на неустойчивость, найденную для холодного пучка. Данная неустойчивость вызывает нефизический нагрев плазмы, увеличивая λ_D до состояния, в котором сеточные моды относительно неважны. Оба примера следует воспринимать как предостережение: хотя указанные нефизические неустойчивости не катастрофичны, они могут вводить в заблуждение, поскольку форма их проявления

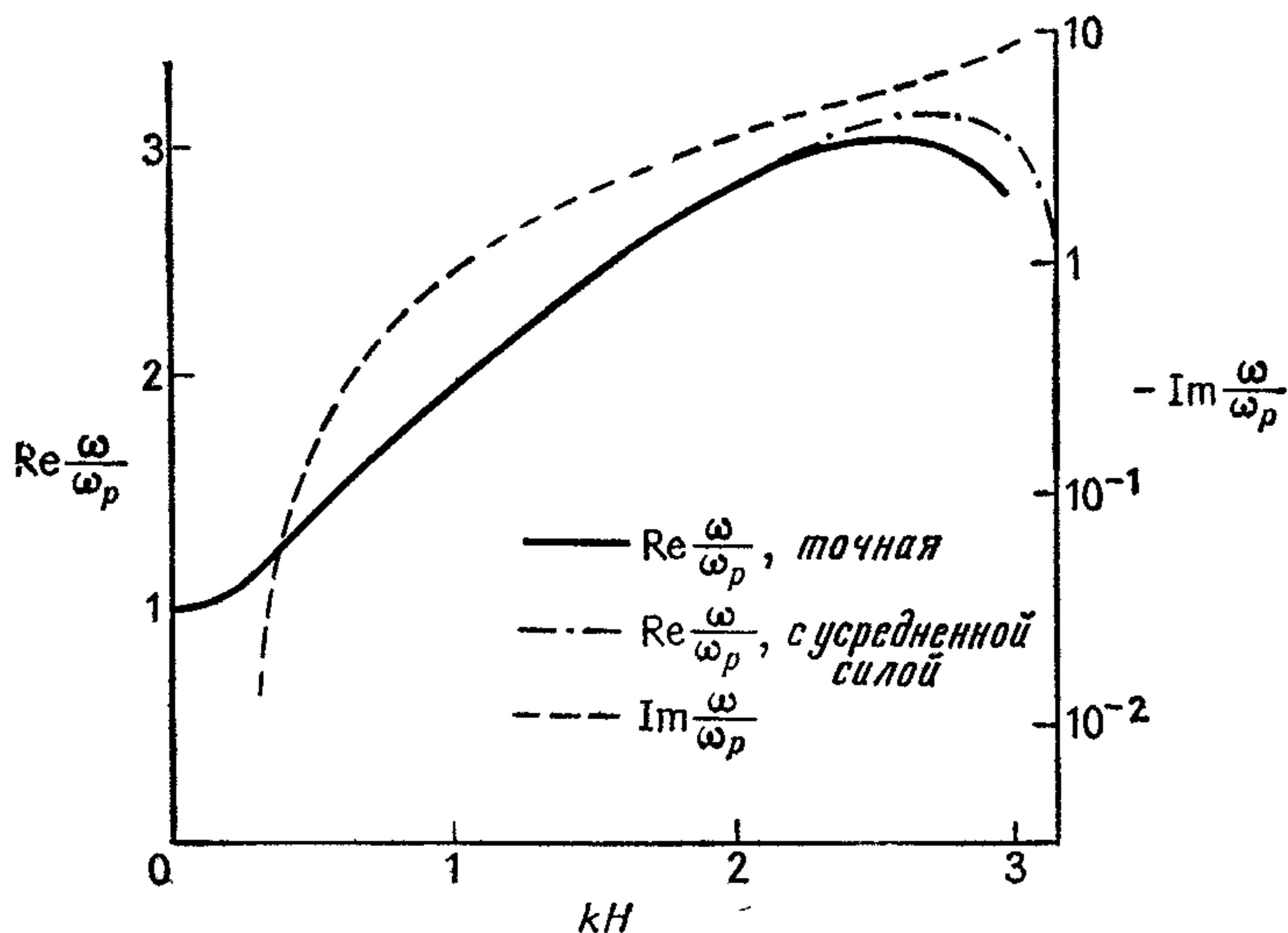


Рис 7.2. Решение дисперсионных уравнений с точной и усредненной силой для максвелловского распределения по скоростям и $H = \lambda_D$ в одномерной численной модели NGP (из работы [234], рис. 5, с любезного разрешения Journal of Computational Physics).

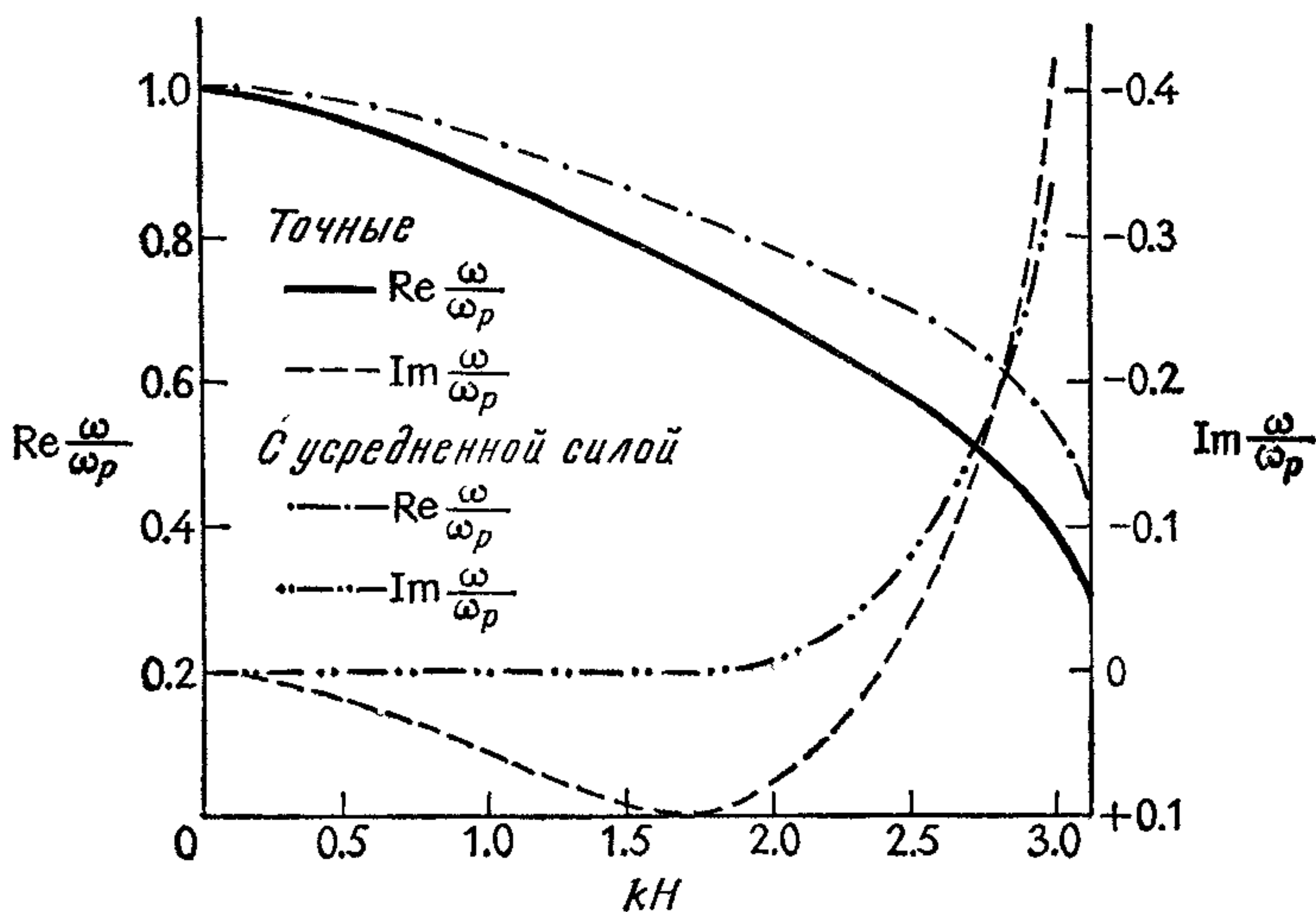


Рис. 7.3. Решение дисперсионных уравнений с точной и усредненной силой при $H = 10\lambda_D$ в одномерной модели NGP (из работы [234], рис. 6, с любезного разрешения Journal of Computational Physics).

весьма похожа на физические неустойчивости, которые, возможно, и являются предметом исследования.

Причиной нефизических неустойчивостей, связанных с пространственной дискретизацией, служит, как отмечалось ранее, нефизическое взаимодействие мод, которое возникает из-за недостаточной частоты дискретизации распределения плотности, обеспечиваемого сеткой. Вычисляя в (7.35) интеграл в смысле главного значения в приближении холодной плазмы, имеем

$$1 = (-i\hat{D})(\varepsilon_0\hat{G})\omega_p^2 \sum_n \left(\frac{\hat{W}}{H}\right)^2 \frac{k_n}{\omega^2} + \\ + (-i\hat{D})(\varepsilon_0\hat{G}) \frac{\omega_p^2}{n_0} \sum_n \left(\frac{\hat{W}}{H}\right)^2 \frac{i\pi}{|k_n|} \frac{\partial f^0}{\partial v} \Big|_{\omega/k_n}. \quad (7.76)$$

Умножая (7.76) на ω^2 , записывая $\omega = \omega_r + i\omega_i$ и считая $|\omega_i| \ll \ll |\omega_r|$, получаем выражение для мнимой части частоты:

$$\omega_i = (-i\hat{D})(\varepsilon_0\hat{G})\omega_p^2 \sum_n \left(\frac{\hat{W}(k_n)}{H}\right)^2 \frac{\pi}{2} \frac{\omega_r}{|k_n|} \frac{1}{n_0} \frac{\partial f^0}{\partial v} \Big|_{\omega_p/k_n}. \quad (7.77)$$

Для стационарного максвелловского распределения (7.52) имеем

$$\omega_i = -\frac{\varepsilon_0\hat{G}\omega_p^2\pi}{\sqrt{2\pi}v_T^3} \sum_n \left(\frac{\hat{W}}{H}\right)^2 \frac{-i\hat{D}(k)}{k_n} \frac{\omega_r}{|k_n|} \exp\frac{-\omega_r^2}{2k_n^2v_T^2}. \quad (7.78)$$

Поскольку знаки у $(-i\hat{D})$ и k совпадают, а все остальные множители в (7.78) положительны, основной член с $n=0$ дает вклад в затухание Ландау, т. е. $\delta_0\omega_i < 0$, где $\delta_n\omega_i$ обозначает вклад от n -й сеточной моды в инкремент затухания или роста (пропорциональных $e^{\omega_i t}$). Аналогично, поскольку

$$\text{sgn}(-i\hat{D}) = \text{sgn}(k + |n|k_g), \quad (7.79)$$

в затухание дают вклад и все сеточные моды с $n < 0$. Однако для всех членов с $n > 0$ знак меняется на обратный,

$$\text{sgn}(-i\hat{D}) = -\text{sgn}(k - nk_g), \quad (7.80)$$

поскольку $|k| < k_g/2$. Следовательно, все сеточные моды с положительными n дают в ω_i вклад со знаком плюс.

Если сумма положительных ($n > 0$) членов превышает по абсолютной величине сумму вкладов от основного члена и отрицательных ($n < 0$) сеточных мод, тогда ω_i будет положительна, что означает нефизическую неустойчивость. Сопоставим, например, величину основной моды и сеточных мод $+1$ и -1 . Из

(7.78) и (7.64) находим

$$\frac{\delta_0 \omega_i}{\delta_1 \omega_i} = - \left(\frac{k - k_g}{k} \right)^{2p+2} \exp \left[\frac{\omega_r^2}{2v_T^2} \left(\frac{1}{(k - k_g)^2} - \frac{1}{k^2} \right) \right], \quad (7.81)$$

$$\frac{\delta_1 \omega_i}{\delta_{-1} \omega_i} = - \left(\frac{k + k_g}{k - k_g} \right)^{2p+2} \exp \left[\frac{\omega_r^2}{2v_T^2} \left(\frac{1}{(k + k_g)^2} - \frac{1}{(k - k_g)^2} \right) \right]. \quad (7.82)$$

Множители перед экспонентами в правых частях (7.81) и (7.82) в обоих случаях больше единицы. Следовательно, увеличение порядка p весовой функции для распределения заряда и интерполяции силы уменьшает величину $\delta_1 \omega_i$ относительно $\delta_0 \omega_i$, $\delta_{-1} \omega_i$ относительно $\delta_1 \omega_i$ и т. д. Оба аргумента в экспонентах отрицательны; если эти экспоненты малы, основная мода будет доминировать, что даст физически правильное поведение. Если, однако, эти экспоненты велики, будут преобладать сеточные моды, что приведет к нефизическому поведению.

Влияние экспоненциального множителя наглядно видно из численного примера. Если мы примем схему NGR ($p = 1$), положим приближенно ω_r равной плазменной частоте ω_p , а волновое число возьмем в середине основной зоны $k = k_g/4 = \pi/2H$, то выражения (7.81), (7.82) примут вид

$$\frac{\delta_0 \omega_i}{\delta_1 \omega_i} = - (3)^4 \exp \left[- \frac{16}{9\pi^2} \left(\frac{H}{\lambda_D} \right)^2 \right] \quad (7.83)$$

и

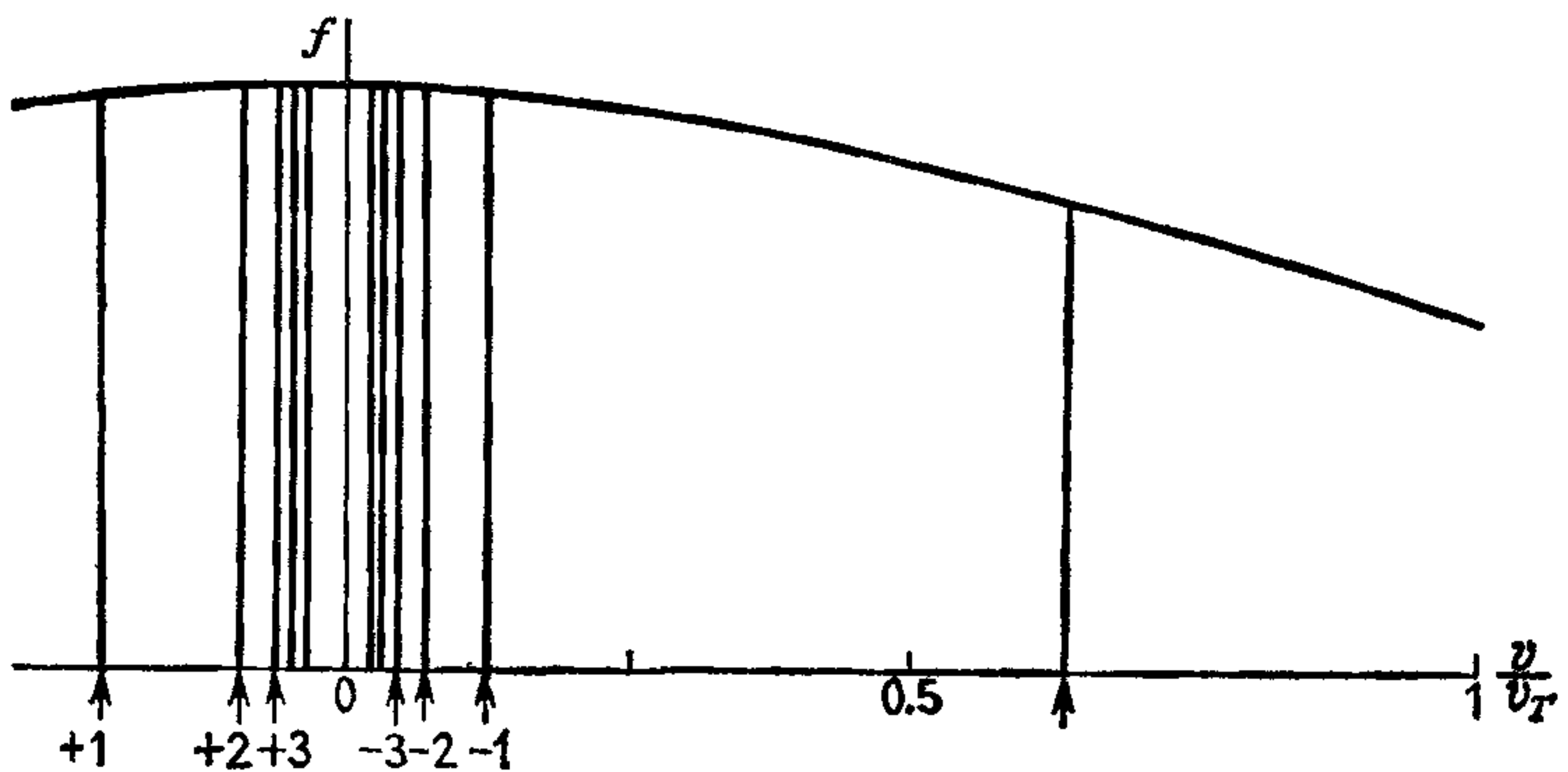
$$\frac{\delta_1 \omega_i}{\delta_{-1} \omega_i} = - \left(\frac{5}{3} \right)^4 \exp \left[- \frac{32}{225\pi^2} \left(\frac{H}{\lambda_D} \right)^2 \right]. \quad (7.84)$$

Если $H \lesssim \lambda_D$, экспоненты в (7.83) и (7.84) приближенно равны единице, и поэтому относительные амплитуды основной и сеточных мод почти целиком определяются весовой функцией распределения заряда: чем ближе \mathcal{W} к функции с конечной шириной спектра, тем слабее делаются сеточные моды.

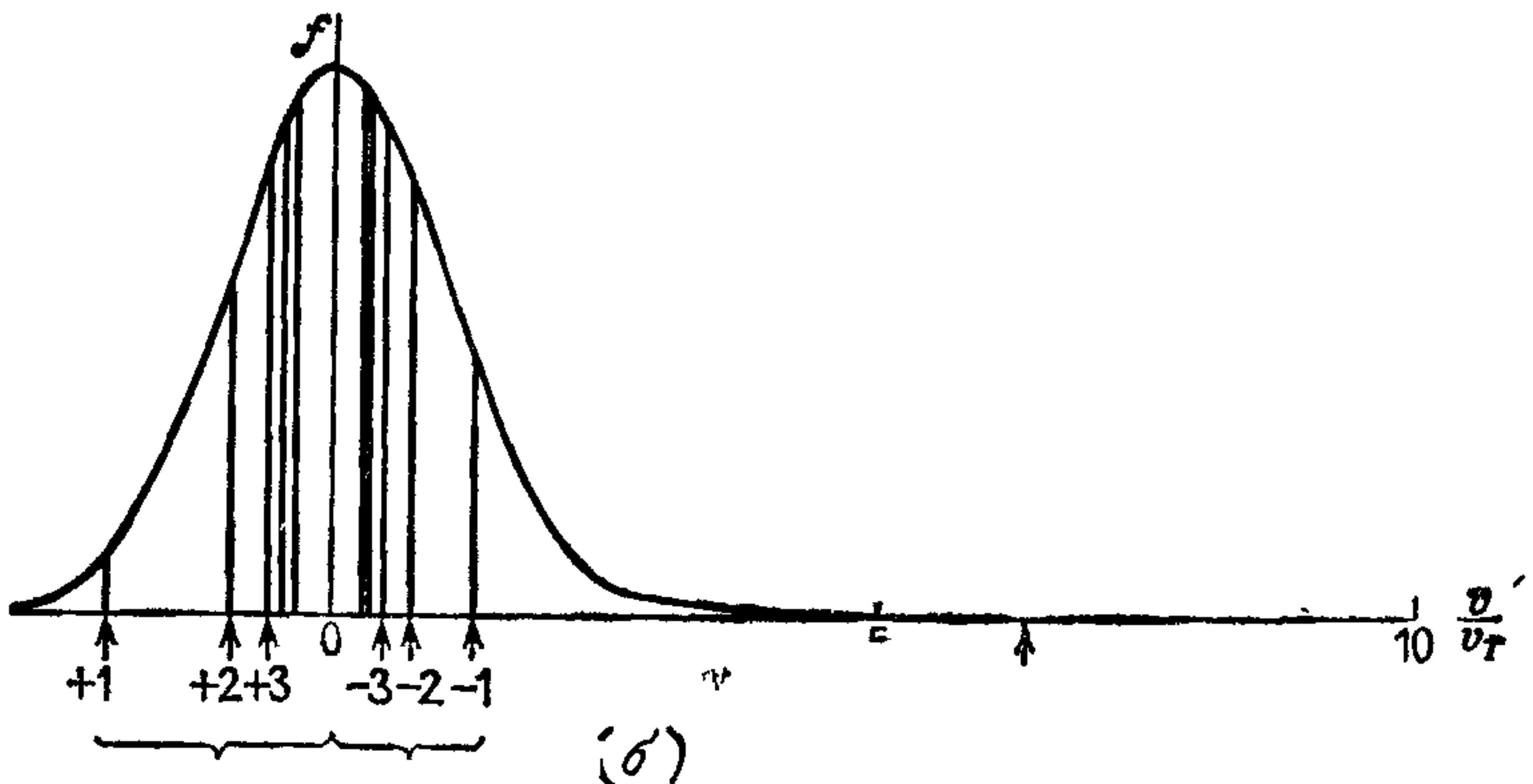
Для $H = \lambda_D$, что отвечает случаю, показанному на рис. 7.2,

$$\delta_0 \omega_i = -68 \delta_1 \omega_i = 515 \delta_{-1} \omega_i. \quad (7.85)$$

Рис. 7.4(a) позволяет дать физическую интерпретацию взаимодействия мод для этого случая. Основная мода, отвечающая фазовой скорости ω/k , в полном соответствии с физикой затухает по Ландау. Моды с волновыми числами k и $k_n = k - nk_g$ на сетке неразличимы, поэтому частицы с фазовыми скоростями $\omega/(k - nk_g)$ также находятся в резонансе с волной, имеющей фазовую скорость ω/k , и если $(\partial f / \partial v)|_{\omega/k_n}$ и ω/k имеют разные знаки, то для частиц со скоростями, близкими к ω/k_n , результирующий поток энергии будет направлен от частиц к волне, соответствующей этой сеточной моде. Поскольку в данном случае



(a)



(б)

Рис. 7.4. Максвелловское распределение f^0 как функция v . Вертикальными линиями показаны фазовые скорости основной волны (в правой части рисунка) и некоторых из волн, отвечающих сеточным модам (положительные и отрицательные сеточные моды выделены фигурными скобками): (а) для случая $H = \lambda_D$ и (б) для $H = 10\lambda_D$. Все положительные сеточные моды (фазовые скорости < 0) приводят к росту, все отрицательные моды и основная мода ($n = 0$) — к затуханию.

шаг пространственной сетки H равен дебаевской длине λ_D , фазовые скорости сеточных мод ω/k_n лежат вблизи $v = 0$, где df/dv мало и поэтому сеточные моды почти совсем не влияют.

На рис. 7.4(б) показана ситуация, отвечающая рис. 7.3, где $H = 10\lambda_D$. Вычисление по формулам (7.81) и (7.82) для этого случая дает

$$\delta_1 \omega_i = -1.8 \delta_{-1} \omega_i \simeq -8.2 \cdot 10^5 \delta_0 \omega_i. \quad (7.86)$$

Теперь пренебрежимо мал вклад в инкремент основной моды. Энергия, которая перекачивается в волну от частиц со скоростями, близкими к $\omega/(k - nk_g)$ для положительных ($n > 0$) сеточных мод, превышает энергию, отбираемую частицами со ско-

ростями, близкими к фазовым скоростям отрицательных ($n < 0$) сеточных мод, что приводит к неустойчивости.

Можно получить простую оценку того, насколько большим можно брать шаг H , чтобы взаимодействие мод не сильно сказывалось на физике системы. Для этого потребуем $|\delta_0 \omega_i| \gg \gg |\delta_1 \omega_i|$. Полагая в рассмотренном выше примере $|\delta_0 \omega_i| > > 10|\delta_1 \omega_i|$, из (7.83) имеем $H < 3.4\lambda_D$.

В вариантах моделей частица — сетка, сохраняющих энергию, обнаруживается интересное качественное отличие во взаимодействии мод. Повторяя преобразования (7.76) — (7.78) для дисперсионного соотношения (7.39), получаем инкремент для схем, сохраняющих энергию:

$$\omega_i = - \frac{\varepsilon_0 \hat{G} \omega_p^2 \pi}{\sqrt{2\pi v_T^3} 2} \sum_n \left(\frac{\hat{W}(k_n)}{H} \right)^2 \frac{\omega_r}{|k_n|} \exp\left(- \frac{\omega_r^2}{2k_n^2 v_T^2} \right). \quad (7.87)$$

Внимательное рассмотрение (7.87) показывает, что основная мода и все сеточные моды вносят вклад в затухание, и, таким образом, нефизическая неустойчивость подавлена. Однако если имеется сдвинутое, а не стационарное максвелловское распределение, нефизические неустойчивости при $H \gg \lambda_D$ вновь оказываются возможными, хотя они будут и значительно слабее, чем в схемах, сохраняющих импульс [234, 235]

Еще одно ограничение на параметры схем, сохраняющих энергию, обусловлено колебаниями самодействия (см. разд. 5.5). Мету значимости самодействия дает сравнение потенциальной энергии самодействия с тепловой энергией. Глубина энергетической потенциальной ямы самодействия $W_{\text{сам}}$ равняется

$$W_{\text{сам}} = \frac{m}{2} \omega_{\text{сам}}^2 \left(\frac{H}{2} \right) = \frac{m}{2} \frac{\omega_p^2}{n_0 H} \left(\frac{H}{2} \right)^2. \quad (7.88)$$

Следовательно,

$$\frac{W_{\text{сам}}}{\frac{1}{2} m v_T^2} = \frac{1}{4} \left(\frac{H}{\lambda_D} \right)^2 \frac{1}{(n_0 H)}. \quad (7.89)$$

Таким образом, если в схемах, сохраняющих энергию, использовать $H \gg \lambda_D$, то число частиц в ячейке $n_0 H$ также должно быть большим, если мы хотим, чтобы потенциальная энергия самодействия была несущественна. Отсюда мы делаем вывод, что как для схем, сохраняющих энергию, так и для схем, сохраняющих импульс, ширина ячейки H не должна превосходить характерной длины λ_D , чтобы дисперсионные свойства бесстолкновительных численных моделей адекватно передавали дисперсионные свойства физической системы.

7.4. МНОГОМЕРНЫЕ ОГРАНИЧЕННЫЕ СИСТЕМЫ

Обсуждение в предыдущем разделе дисперсионных свойств и ранее в разд. 5.6 силы межчастичного взаимодействия касалось одномерной бесконечной системы. Полученные результаты, дополненные некоторым ограничением на допустимые значения волновых чисел, применимы для ограниченных периодических систем. Указанное ограничение состоит в том, что значения волновых чисел k ограничиваются дискретным набором равноотстоящих значений $k = lk_0$, где l — целое, $k_0 = 2\pi/L$, а L — длина периодической области. Более того, если взять периодические системы длиной $2L$ и рассматривать на этом периоде только четные (нечетные) функции, то полученные результаты приложимы и к системам длиной L с однородными граничными условиями Дирихле (Неймана).

Результаты для одномерных ограниченных систем непосредственно обобщаются на случай двух и трех измерений. Единственное существенное изменение, вносимое переходом к большим размерностям, — это угловая анизотропия сил, дисперсионных свойств и т. д.

7.4.1. Периодичность

Один из способов доказать, что результаты, полученные для периодических систем, отличаются от результатов для бесконечных систем только ограничением на допустимые значения волновых чисел — это повторить все выкладки с самого начала, используя $FS(i)$ вместо FT и FFT вместо $FS(ii)$ (см. табл. П.1). Символически эту замену можно представить в виде

$$\int \frac{dk}{2\pi} \rightarrow \frac{1}{L} \sum_l = \sum_l \frac{k_0}{2\pi}, \quad (7.90)$$

где l — целое, обозначающее определенное значение волнового числа и принимающее в основной зоне значения от $-N/2$ до $N/2 - 1$. Здесь N равняется количеству узлов сетки, используемому на длине периода L в x -пространстве.

Указанную аналогию результатов для периодических и бесконечных систем можно наглядно продемонстрировать с помощью функции выборки Π . Всякую периодическую систему можно рассматривать как суперпозицию множества бесконечных систем, каждая из которых сдвинута относительно предыдущей на длину периода L : пример такой системы, на который мы уже ссылались, встречается в методе изображений в электростатике. Процесс сдвига с повторениями, как мы установили в гл. 5, можно описать операцией свертки с Π .

Пусть в бесконечной системе задана функция $f(x) \supset \hat{f}(k)$. Тогда в периодической системе ей отвечает функция $g(x)$, равная

$$g(x) = \frac{1}{L} \Pi\left(\frac{x}{L}\right) * f(x) = \quad (7.91)$$

$$\begin{aligned} &= \frac{1}{L} \int \sum_n \delta\left(\frac{x' - nL}{L}\right) f(x - x') dx' = \\ &= \sum_{n=-\infty}^{\infty} f(x - nL), \end{aligned} \quad (7.92)$$

которая, разумеется, периодична с периодом L и получается суммированием сдвинутых «изображений» функции $f(x)$. Далее, в силу (П.29) и теоремы подобия (табл. П.4), имеем

$$\frac{1}{L} \Pi\left(\frac{x}{L}\right) \supset \Pi\left(\frac{k}{k_0}\right). \quad (7.93)$$

Отсюда

$$\hat{g}(k) = \Pi\left(\frac{k}{k_0}\right) \hat{f}(k). \quad (7.94)$$

Для приведения $g(x)$ к длине одного периода умножим ее на $\Pi(x/L)$, т. е.

$$g'(x) = \Pi\left(\frac{x}{L}\right) g(x), \quad (7.95)$$

где

$$\begin{aligned} g'(x) \supset \hat{g}'(k) &= L \operatorname{sinc}\left(\frac{k}{k_0}\right) * \Pi\left(\frac{k}{k_0}\right) \hat{f}(k) = \\ &= \sum_l \operatorname{sinc}\left(\frac{k - lk_0}{k_0}\right) \hat{f}(lk_0). \end{aligned} \quad (7.96)$$

Сравнивая (7.96) с (5.183), мы видим, что (7.96) имеет тот же вид, что и непрерывная функция $\hat{g}(k)$, которая эквивалентна набору значений $\{\hat{f}(lk_0)\}$, т. е. если задана некоторая функция $f \supset \hat{f}$ для бесконечной системы, фурье-образ ее периодического аналога равняется $\hat{f}(lk_0)$, l целое. Соотношения между FT , FS и FFT перечислены на рис. П.2 и П.3.

7.4.2. Двумерные и трехмерные системы

Преобразования Фурье в направлениях x и y в двумерном случае и по x , y и z в трехмерном производятся независимо друг от друга. Следовательно, результаты для двумерных и трехмерных систем можно получить повторением известных этапов одномерного рассмотрения с заменой скалярных величин векторными, а сумм — двойными и тройными суммами соответственно. Правила этих замен приведены в приложении в разд. П.6.

Чтобы проиллюстрировать, как делаются многомерные преобразования, давайте воспроизведем сейчас вывод дисперсионного соотношения для электростатической плазмы, а именно (7.34), для двумерных и трехмерных систем. Как и ранее, будем пользоваться компактной формой записи, введенной в разд. 5.6; штрихом помечены непрерывные функции, соответствующие сеточным величинам, а крестики обозначают обобщенные функции, представляющие собой наборы сеточных значений. За исключением плотности заряда ρ' , все штрихованные величины строятся по сеточным значениям таким образом, чтобы их спектр имел конечную ширину (см. разд. 5.6.1). Получающиеся при данном выводе уравнения с равным успехом относятся к периодическим и бесконечным системам, причем эти два случая отличаются лишь определениями (1) прямого и обратного преобразований (см. табл. П.1), (2) свертки (см. табл. П.3) и (3) функции выборки (см. разд. П.4.5).

Плотность заряда

$$\rho'(\mathbf{x}) = \frac{q}{V_c} W(\mathbf{x}) * n(\mathbf{x}) \supset \hat{\rho}'(\mathbf{k}) = \frac{q}{V_c} \hat{W}(\mathbf{k}) \hat{n}(\mathbf{k}), \quad (7.97)$$

$$\begin{aligned} \rho^+(\mathbf{x}) = \Pi(\mathbf{x}; \mathbf{H}) \rho'(\mathbf{x}) &\supset \hat{\rho}(\mathbf{k}) = V_c \Pi(\mathbf{k}; \mathbf{k}_g) * \hat{\rho}'(\mathbf{k}) \\ &= \frac{q}{V_c} \sum_n \hat{W}(\mathbf{k}_n) \hat{n}(\mathbf{k}_n). \end{aligned} \quad (7.98)$$

Потенциал

$$\varphi'(\mathbf{x}) = G'(\mathbf{x}) * \rho^+(\mathbf{x}) \supset \hat{\varphi}'(\mathbf{k}) = \hat{G}'(\mathbf{k}) \hat{\rho}(\mathbf{k}), \quad (7.99)$$

$$\varphi^+(\mathbf{x}) = \Pi(\mathbf{x}; \mathbf{H}) \varphi'(\mathbf{x}) \supset \hat{\varphi}(\mathbf{k}) = \hat{G}(\mathbf{k}) \hat{\rho}(\mathbf{k}). \quad (7.100)$$

Электрическое поле

$$\mathbf{E}'(\mathbf{x}) = -\mathbf{D}(\mathbf{x}) * \varphi'(\mathbf{x}) \supset \hat{\mathbf{E}}'(\mathbf{k}) = -\hat{\mathbf{D}}(\mathbf{k}) \hat{\varphi}'(\mathbf{k}), \quad (7.101)$$

$$\mathbf{E}^+(\mathbf{x}) = \Pi(\mathbf{x}; \mathbf{H}) \mathbf{E}'(\mathbf{x}) \supset \hat{\mathbf{E}}(\mathbf{k}) = -\hat{\mathbf{D}}(\mathbf{k}) \hat{\varphi}(\mathbf{k}). \quad (7.102)$$

Сила

$$\mathbf{F}(\mathbf{x}) = \frac{q}{V_c} W(\mathbf{x}) * \mathbf{E}^+(\mathbf{x}) \supset \hat{\mathbf{F}}(\mathbf{k}) = \frac{q}{V_c} \hat{W}(\mathbf{k}) \hat{\mathbf{E}}^+(\mathbf{k}). \quad (7.103)$$

В трехмерном случае

$$V_c = H_1 H_2 H_3 = \text{объем ячейки}, \quad (7.104)$$

$$\mathbf{H} = (H_1, H_2, H_3) = \text{вектор ячейки}, \quad (7.105)$$

$$\Pi(\mathbf{x}; \mathbf{H}) = \Pi\left(\frac{x}{H_1}\right) \Pi\left(\frac{y}{H_2}\right) \Pi\left(\frac{z}{H_3}\right), \quad (7.106)$$

$$\mathbf{n} = (n_1, n_2, n_3) = \text{номер сеточной моды, } n_i \text{ целое}, \quad (7.107)$$

$$\mathbf{k}_g = 2\pi \left(\frac{1}{H_1}, \frac{1}{H_2}, \frac{1}{H_3} \right) = \text{сеточное волновое число}, \quad (7.108)$$

$$\mathbf{k}_n = \mathbf{k} - 2\pi \left(\frac{n_1}{H_1}, \frac{n_2}{H_2}, \frac{n_3}{H_3} \right). \quad (7.109)$$

Одномерные функции выборки в правой части равенства (7.106) определяются соотношениями (П.28) для бесконечных систем либо (П.31) для периодических систем. Если, как это почти всегда бывает, $\hat{W}(\mathbf{k})$ является произведением функций, то

$$\hat{W}(\mathbf{k}) = \hat{w}(k_1) \hat{w}(k_2) \hat{w}(k_3), \quad (7.110)$$

где $\hat{w}(k_i)$ — фурье-образ i -й компоненты одномерной функции распределения заряда. Соответствующие (7.104) — (7.110) двумерные выражения получаются отбрасыванием компонент по z (с индексом 3), т. е. $V_c = H_1 H_2$, и т. д.

Векторный аналог уравнения (7.27) имеет вид

$$\hat{n}(\mathbf{k}) = -\frac{i}{m} \int \frac{\hat{\mathbf{F}}(\mathbf{k}) \cdot \frac{\partial f^0}{\partial \mathbf{v}}}{(\omega - \mathbf{k} \cdot \mathbf{v})} d\mathbf{v}, \quad (7.111)$$

что в итоге, после исключения \hat{n} , $\hat{\mathbf{F}}$, $\hat{\rho}$, $\hat{\phi}$ и $\hat{\mathbf{E}}$ с помощью (7.97) — (7.103), дает

$$1 = -(\epsilon_0 \hat{G}) \frac{\omega_p^2}{n_0} \sum_{\mathbf{n}} \left(\frac{\hat{W}(\mathbf{k}_{\mathbf{n}})}{V_c} \right)^2 \int \frac{-i \hat{\mathbf{D}}(\mathbf{k}) \cdot \frac{\partial f^0}{\partial \mathbf{v}}}{(\omega - \mathbf{k}_{\mathbf{n}} \cdot \mathbf{v})} d\mathbf{v}. \quad (7.112)$$

7.5. СТОЛКНОВЕНИЯ

В разд. 7.3 мы выяснили, что если размер ячейки порядка дебаевской длины, то модельная плазма по своим дисперсионным свойствам близка к плазме частиц конечного размера, при этом фурье-образ эффективной формы этих частиц определяется выражениями (7.74) или (7.75). Следовательно, можно предположить, что если работать в области параметров, где приближение усредненной силы в дисперсионном соотношении удовлетворительно, то и описание столкновительных процессов с усредненной силой также будет применимо; это предположение действительно подтверждается практикой.

Столкновительные процессы в приближении частиц конечного размера (с усредненной силой) подробно исследованы [299, 238, 38]. Было показано, что при переходе к частицам конечного размера в кинетическом уравнении для плазмы точечных частиц [250].

$$\left(\frac{\partial f}{\partial t} \right)_c = \frac{\partial}{\partial \mathbf{v}} \int d^3 v' \frac{1}{8\pi^2} \int d^3 k \delta(\mathbf{k} \cdot \mathbf{v} - \mathbf{k} \cdot \mathbf{v}') \frac{\mathbf{k} \mathbf{k} \hat{\Phi}(\mathbf{k})}{|\epsilon(\mathbf{k} \cdot \mathbf{v}, \mathbf{k})|^2} \times \\ \times \left(\frac{\partial f(\mathbf{v})}{\partial \mathbf{v}} f(\mathbf{v}') - \frac{\partial f(\mathbf{v}')}{\partial \mathbf{v}'} f(\mathbf{v}) \right), \quad (7.113)$$

меняется только фурье-образ потенциальной энергии парного взаимодействия $\hat{\Phi}(\mathbf{k})$ и диэлектрическая проницаемость плазмы $\epsilon(\omega = \mathbf{k} \cdot \mathbf{v}, \mathbf{k})$ (нули которой дают дисперсионное соотношение);

они заменяются соответствующими выражениями для плазмы частиц конечного размера:

$$\hat{\Phi}(\mathbf{k}) = \frac{q^2 \hat{S}_e^2(\mathbf{k})}{\epsilon_0 k^2}, \quad (7.114)$$

$$\epsilon(\omega, \mathbf{k}) = 1 + \frac{\omega_p^2}{n_0 k^2} \hat{S}_e^2(\mathbf{k}) \int \mathbf{k} \cdot \frac{\partial f}{\partial \mathbf{v}} \frac{d\mathbf{v}}{\omega - \mathbf{k} \cdot \mathbf{v}}. \quad (7.115)$$

Уравнение (7.113) можно записать в форме уравнения (7.1) посредством выделения членов с первой и второй производными. Получающиеся в результате коэффициенты диффузии и трения были рассчитаны численно Окудой и Бэрдсолом [299] для максвелловского распределения $f(\mathbf{v}')$ и облаков гауссовой формы ($\hat{S}_e(\mathbf{k}) = \exp(-2k^2 R^2)$) как для двумерных, так и для трехмерных систем. Ими показано, что облака с размером R от H до $2H$ приводят к уменьшению величины этих коэффициентов на один-два порядка по сравнению с соответствующими значениями для точечных частиц, и далее с ростом R они продолжают уменьшаться.

Можно показать, что с математической точки зрения указанное уменьшение коэффициентов диффузии и трения обусловлено множителем $\hat{S}_e^4(\mathbf{k})$, появляющимся в уравнении (7.113) за счет конечного размера частицы. На самом деле этот множитель устраняет проблемы расходимости при больших k , с которыми приходится сталкиваться при рассмотрении точечных частиц. Из этого следует, что если заставить $\hat{S}_e(\mathbf{k})$ спадать более быстро (что отвечает большим значениям R), то величина коэффициентов диффузии уменьшится.

Рассматривая движение пробной частицы через плазму, имеющую относительно небольшое число частиц в дебаевской сфере, можно получить простую физическую картину влияния конечного размера частицы. Орбитой пробной частицы в «размазанной» (т. е. усредненной по ансамблю) бесстолкновительной плазме является прямая линия. В плазме точечных частиц пробная частица чувствует сильно флуктуирующее поле от небольшого числа частиц в дебаевской сфере, ее окружающей, и поэтому будет испытывать множество сильных возмущений, заставляющих ее быстро отклониться (соответственно большой диффузии скорости) от бесстолкновительной прямой траектории. Придание частице конечного размера устраняет сильное рассеяние от встреч на близких расстояниях. Если размер и число частиц достаточно велики, чтобы пробная частица перекрывалась многими из них, то пробная частица будет «чувствовать» фон частиц конечного размера как однородную среду, т. е. бесстолкновительную плазму.

Максимально допустимый радиус R частиц определяется коллективными свойствами. Когда этот радиус значительно превышает дебаевскую длину, он сам выступает в роли дебаевской длины, и это приводит к тому, что коллективные свойства становятся существенно отличными от свойств физической системы: это изменение наглядно иллюстрируют дисперсионные кривые на рис. 7.2 и 7.3.

Для целей моделирования бесстолкновительных систем, будь то плазма или гравитирующие системы, достаточно знать, будет ли вычислительная модель вести себя как бесстолкновительная система за расчетный промежуток времени или нет. Даже в тех системах, где столкновительные процессы накладываются на бесстолкновительное движение, как это, например, делается при моделировании полупроводников (см. гл. 10), важно, чтобы процессы физических столкновений доминировали над численными. Простую оценку времени, в течение которого можно проводить вычислительный эксперимент, прежде чем столкновительные эффекты не станут значительными, дает время столкновения τ_c . При экспериментальном определении τ_c (см. разд. 9.2.3) установлено, что для тех размеров частиц, которые обычно используются в вычислительных моделях плазмы, время столкновения в случае двумерных моделей хорошо описывается эмпирической зависимостью

$$\frac{\tau_c}{\tau_p} = N_D \left[1 + \left(\frac{W}{\lambda_D} \right)^2 \right] = \frac{\omega_p / 2\pi}{\nu_c}, \quad (7.116)$$

где $\omega_p = 2\pi/\tau_p$ — плазменная частота,

ν_c — частота столкновений,

N_D — число частиц в дебаевской сфере ($= n\lambda_D^2$),

$W (=2R)$ — ширина частицы.

Полученное в [181, 190] эмпирическое соотношение (7.116) показывает хорошее согласие с теоретическими временами столкновений, найденными в [299] для модели плазмы с частицами конечного размера. На рис. 7.5 показаны рассчитанные значения эффективных сечений, найденные для частиц конечного размера с однородной плотностью. Поскольку частота столкновений приближенно равна $\nu \sim n\nu\sigma$, рис. 7.5 можно использовать для того, чтобы оценить, как изменится время столкновения при изменении размера частицы. Аналогичные оценки частоты столкновений, полученные в [238], для частиц с эффективной формой $S_e(x) \supset \hat{S}_e(k) = \exp[-k^2 R^2/2]$ и $S_e \supset \hat{S}_e(k) = \Pi(2k/R)$ показывают, что частота столкновений не зависит от формы частицы в практически важной области значений радиуса частицы R порядка $R = \lambda_D$.

Дальнейшие экспериментальные результаты о влиянии столкновений в плазме частиц конечного размера [303, 298, 277]

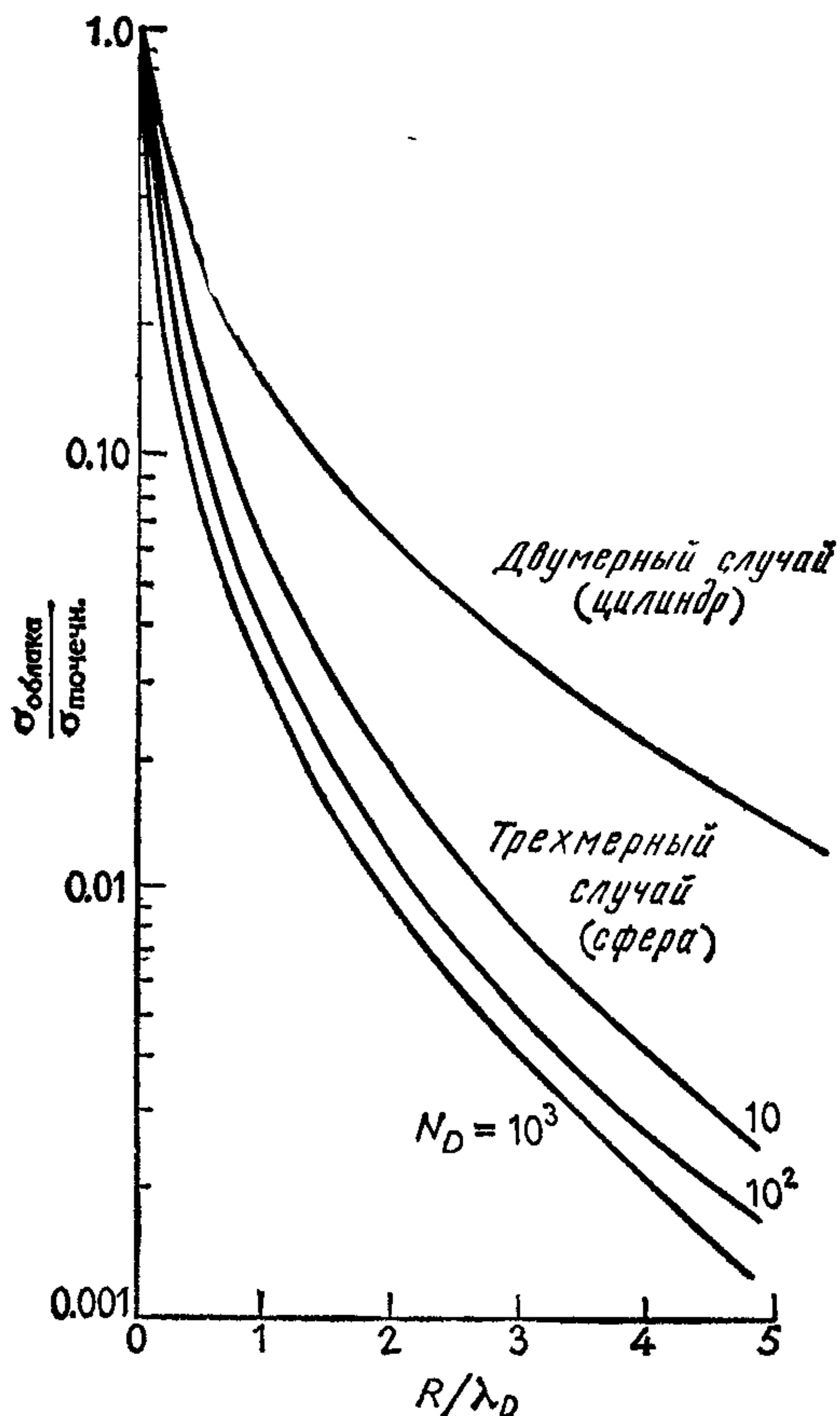


Рис. 7.5 Эффективные сечения однородно заряженных частиц конечного размера, нормированные на сечение точечной частицы, для двумерного и трехмерного случаев (из работы [299], рис. 7, с разрешения журнала *Physics of Fluids*).

подтверждают выводы кинетической теории частиц конечного размера с усредненной силой. Более общая кинетическая теория вычислительных моделей частица-сетка, в которой учитывается влияние сеточных гармоник, обусловленных конечными шагами по пространству и по времени, была разработана Лэнгдоном [233, 236, 237]. Дополнительные детали общей кинетической теории вычислительных моделей бесстолкновительной плазмы читатель может почерпнуть из указанных работ.

7.6. ЗАКОНЫ СОХРАНЕНИЯ

Подавляющее большинство вычислительных моделей частиц отвечает консервативным физическим системам, т. е. физическим системам, в которых сохраняются энергия, импульс и угловой момент. Наличие в этих вычислительных моделях погрешностей дискретизации приводит к тому, что некоторые или все

законы сохранения становятся приближенными. При проведении практических расчетов все физически сохраняющиеся величины обычно непрерывно контролируются, что обеспечивает проверку в целом качества моделирования. Аналогичные тесты удается нередко разработать и для неконсервативных систем, поскольку экспериментатор-вычислитель, как правило, обладает достаточной информацией об источниках и стоках в рассматриваемой системе, что позволяет ему построить уравнения баланса полной энергии и импульса.

К нарушению законов сохранения приводят многие факторы: эффект наложения, конечно-разностные погрешности при вычислении полей, погрешности интегрирования по времени с конечным шагом, погрешности округления, а также погрешности измерения физических характеристик. Последний фактор может рассматриваться отдельно, поскольку его влияние состоит в том, что в наблюдаемые величины вносятся погрешности эксперимента, поддающиеся количественному определению, и они никак не затрагивают динамику. Погрешности округления можно не учитывать вследствие их малости по сравнению с погрешностями аппроксимации и недостаточной дискретизации, при условии, что сама схема численно устойчива (см. гл. 4). Но всестороннее рассмотрение влияния оставшихся погрешностей по-прежнему представляет собой трудную аналитическую проблему. Тем не менее, как будет показано ниже, несколько упрощающих предположений позволят нам количественно оценить доминирующие факторы.

7.6.1. Энергия

В консервативной системе полная энергия E равняется сумме кинетической и потенциальной энергий. Кинетическая энергия \mathcal{T} определяется суммой кинетических энергий всех частиц,

$$\mathcal{T} = \sum_{i=1}^{N_p} \frac{1}{2} m v_i^2, \quad (7.117)$$

а потенциальная энергия V равняется сумме потенциальных энергий парного взаимодействия $\Phi_{ij} = \Phi(x_i; x_j)$:

$$V = \frac{1}{2} \sum_{i=1}^{N_p} \sum_{\substack{j=1 \\ i \neq j}}^{N_p} \Phi_{ij}. \quad (7.118)$$

Измерение \mathcal{T} в вычислительных экспериментах, проводимых методом частица — сетка, не представляет никаких трудностей. Чтобы измерить V , требуется знать потенциал в точках, где расположены частицы; его можно вычислить с помощью

интерполяции по известным сеточным значениям потенциала, аналогично тому, как находится сила по сеточному полю (см. (5.161)). Если частица с зарядом q , находящаяся в точке x_i , создает потенциал, изображаемый набором сеточных значений $\{\varphi(x_p)\}$, или, что равносильно, обобщенным потенциалом φ^\dagger , то потенциальная энергия парного взаимодействия этой частицы с зарядом q в точке x_j определяется равенством

$$\begin{aligned}\Phi(x_i; x_j) &= q \sum_p W(x_j - x_p) \varphi(x_p) = \\ &= \frac{q}{H} \int W(x_j - x') \varphi^\dagger(x') dx'.\end{aligned}\quad (7.119)$$

Исключая с помощью (5.158) и (5.159) $\varphi(x_p)$, имеем

$$\Phi(x_i; x_j) = q^2 \sum_p \sum_{p'} W(x_j - x_p) G(x_p - x_{p'}) W(x_i - x_{p'}). \quad (7.120)$$

Подставляя (7.120) в (7.118) (пренебрегая ограничением $i \neq j$) и проводя перегруппировку членов, получаем приближенное выражение для потенциальной энергии:

$$\begin{aligned}V &= \frac{H}{2} \sum_p \left(\frac{q}{H} \sum_i W(x_j - x_p) \right) \left[H \sum_{p'} G(x_p - x_{p'}) \times \right. \\ &\quad \left. \times \left(\frac{q}{H} \sum_i W(x_i - x_{p'}) \right) \right]\end{aligned}\quad (7.121)$$

$$= \frac{H}{2} \sum_p \rho(x_p) \varphi(x_p) = \frac{1}{2} \int \rho' \varphi^\dagger dx = \frac{1}{2} \int \rho^\dagger \varphi' dx. \quad (7.122)$$

Пренебрежение ограничением $i = j$ приводит к тому, что в (7.121) входит член потенциальной энергии самодействия. Поскольку этот член в моделях частица — сетка конечен, его присутствие не вызывает никаких существенных изменений в определении энергии (заметим, однако, что из-за погрешностей дискретизации потенциальная энергия самодействия не является строго постоянной). Отождествляя в (7.121) выражения в круглых скобках с сеточной плотностью заряда, а член в квадратных скобках с потенциалом, можно записать потенциальную энергию через известные сеточные величины или эквивалентные им обобщенные функции (см. (7.122)).

В любых практических расчетах, по-видимому, предпочтительнее при вычислении потенциальной энергии использовать выражение (7.122), а не (7.118), из-за значительно меньшего числа операций. Однако в двумерных и трехмерных задачах обычно принято на стадии нахождения потенциала с целью экономии памяти записывать значения потенциала на место плотности заряда, а в этом случае $\{\rho(x_p)\}$ и $\{\varphi(x_p)\}$ одновременно недоступны. Возможный выход из положения заключается в том,

чтобы накапливать V на этапе интерполяции силы, поскольку вклад i -й частицы в полную потенциальную энергию равен, исходя из (7.121),

$$\delta V_i = q \sum_p W(x_i - x_p) \varphi(x_p). \quad (7.123)$$

Если в интерполяции силы, действующей на i -ю частицу участвует s узлов сетки, то дополнительные накладные расходы на вычисление V при использовании (7.123) составляют sN_p сложений и sN_p умножений по сравнению с N_g сложениями и N_g умножениями при использовании (7.122). Так как число частиц N_p обычно значительно больше количества узлов сетки N_g , последний метод в вычислительном отношении дешевле.

Возможен еще один метод вычисления V , когда уравнения для полей решаются с помощью быстрого преобразования Фурье (см. гл. 6). Используя теорему Парсеваля (табл. П. 4), для периодической системы с длиной периода L можно переписать выражение (7.122) в виде

$$V = \frac{1}{2} \sum_l^{N_g-1} \left(\frac{k_0}{2\pi} \right) \hat{\rho}^*(k) \hat{\phi}(k). \quad (7.124)$$

Здесь k_0 — ширина «ячейки» в k -пространстве, $k_0 = 2\pi/L$, $\hat{\rho}^*$ — комплексно сопряженное $\hat{\rho}$ ($\hat{\rho}^*(k) = \hat{\rho}(-k)$, поскольку ρ вещественное $\supset \hat{\rho}$ эрмитово). Если V накапливается в счетном цикле, где вычисляется $\hat{\phi} = \hat{G}\hat{\rho}$, то число операций опять составляет $\sim 2N_g$. Аналогичные результаты получаются и для алгоритмов, использующих разложения по синусам или косинусам, а также для двумерных и трехмерных задач.

Потенциальная энергия парного взаимодействия, определяемая соотношением (7.119), проявляет такую же сеточную зависимость (возникающую из-за погрешностей дискретизации при распределении заряда), как и сила взаимодействия частиц. Следуя выкладкам, которые проводились для силы взаимодействия частиц (разд. 5.6.4), можно записать потенциальную энергию взаимодействия двух частиц с зарядом q , находящихся на расстоянии x , в виде

$$\Phi(x; x_1) = \int \frac{dk}{2\pi} e^{ikx} \left(\frac{q^2}{H^2} \hat{W}(k) \hat{G}(k) \sum_n \hat{W}(k_n) e^{-ink_g x_1} \right), \quad (7.125)$$

и после усреднения по смещениям получим

$$\langle \Phi(x; x_1) \rangle = \int \frac{dk}{2\pi} \left(q^2 \frac{\hat{W}^2}{H^2} \hat{G} \right) e^{ikx}. \quad (7.126)$$

Записывая гармонику среднего потенциала парного взаимодействия в форме (7.114), извлекаем эффективную форму ча-

стицы:

$$\hat{S}_e^2(k) = \left(\frac{\hat{W}}{H}\right)^2 (\epsilon_0 \hat{G} k^2). \quad (7.127)$$

Сравнивая (7.127) с (7.74) и (7.75), мы видим, что эффективные формы частицы, отвечающие средней силе и средней энергии парного взаимодействия, совпадают для схем, сохраняющих энергию, но для схем, сохраняющих импульс, они различаются. В случае схем, сохраняющих импульс, более хорошая аппроксимация потенциальной энергии получается при замене (7.124) на

$$V = \frac{1}{2} \sum_l \frac{k_0}{2\pi} \hat{\rho}^*(k) \hat{\phi}(k) \left(\frac{-i\hat{D}(k)}{k}\right). \quad (7.128)$$

Множитель $(-i\hat{D}/k)$ приводит к тому, что для расчета силы и потенциальной энергии будет использоваться одна и та же (усредненная по смещениям) эффективная форма частицы. В результате этого если функции \hat{W} и \hat{n} (обе вместе или по отдельности) достаточно быстро убывают с ростом k , чтобы $\hat{\rho} = q\hat{W}\hat{n}/H$ хорошо аппроксимировалось (т. е. сеточные гармоники были пренебрежимо малы), то и силы, и энергии описывают одну и ту же «бессеточную» систему частиц конечного размера.

Обсуждавшиеся до сих пор погрешности в энергии суть погрешности измерения. Гораздо важнее погрешности, вносимые неточностью динамики. В пределе нулевого шага по времени варианты схем частица — сетка, сохраняющие энергию, действительно заставляют энергию сохраняться за счет того, что сила, хотя она и приближенная, может быть записана в виде градиента некоторого скаляра (см. разд. 5.5). В схемах же, сохраняющих импульс, требования симметрии, обуславливающие сохранение импульса (разд. 5.3.3), не позволяют записать силу как градиент скаляра, и именно неконсервативная часть этой силы вызывает рост полной энергии. Если в схемах, сохраняющих энергию, силу взаимодействия частиц представить в виде

$$\mathbf{F}(x; x_1) = -\nabla\Phi + \mathbf{f}, \quad (7.129)$$

мы можем рассматривать $-\nabla\Phi$ как усредненную по смещениям часть силы, а \mathbf{f} — как (флуктуирующий) неконсервативный остаток.

Рассмотрим однородную бесстолкновительную модельную (частица — сетка) плазму, в которой сила определяется по схеме, сохраняющей импульс. Если плотность частиц достаточно велика, чтобы систему можно было считать бесстолкновительной, то энергия пробной частицы равняется ее кинетической энергии:

$$E(t) = \frac{mv^2}{2}. \quad (7.130)$$

На следующем временном шаге энергия равна

$$E(t + DT) = \frac{m}{2} (\mathbf{v} + \Delta\mathbf{v})^2, \quad (7.131)$$

где

$$\Delta\mathbf{v} = \frac{DT}{m} \sum_{i=1}^{N_p} \mathbf{f}_i. \quad (7.132)$$

В (7.132) сделано упрощающее предположение, что консервативная часть силы в сумме дает нуль. Сила \mathbf{f}_i является неконсервативной частью силы, действующей на пробную частицу со стороны i -й частицы, где i -я частица с координатой \mathbf{x}_i расположена в расчетной области объема V_b случайным образом. Если рассматривать $E(t + DT)$ как функционал от $E(t)$ и совокупности координат частиц $\{\mathbf{x}_i; i = 1, N_p\}$, то математическое ожидание энергии $E(t + DT)$ можно записать в виде

$$\langle E(t + DT) \rangle = \int EQ(E) dE = \int E \prod_{i=1}^{N_p} P(\mathbf{x}_i) d\mathbf{x}_i, \quad (7.133)$$

где $Q(E)$ — плотность распределения по энергиям, а $P(\mathbf{x}_i) (= V_b^{-1})$ — плотность распределения координат i -й частицы. При получении (7.133) предполагалось, что координаты частиц некоррелированы. Подстановка (7.131) в (7.133) дает

$$\begin{aligned} \langle E(t + DT) \rangle = \langle E(t) \rangle + m\mathbf{v} \cdot \int \frac{DT}{m} \sum_j \mathbf{f}_j \prod_i P(\mathbf{x}_i) d\mathbf{x}_i + \\ + \frac{DT^2}{2m} \int \sum_j \mathbf{f}_j \cdot \sum_{j'} \mathbf{f}_{j'} \prod_i P(\mathbf{x}_i) d\mathbf{x}_i. \end{aligned} \quad (7.134)$$

Полагая $P = V_b^{-1}$ и используя тот факт, что $\int \mathbf{f}_j d\mathbf{x}_j = 0$, удается проинтегрировать (7.134) и получить

$$R = \frac{\langle E(t + DT) \rangle - \langle E(t) \rangle}{DT} = \frac{DTn_0}{2m} \int |\mathbf{f}(\mathbf{x}; \mathbf{x}')|^2 d\mathbf{x}. \quad (7.135)$$

Рассматривая по очереди каждую частицу в качестве пробной, находим среднюю по ансамблю удельную скорость нагрева $\langle r \rangle$:

$$\langle r \rangle = \frac{\langle R \rangle}{\frac{1}{2} m v_T^2} = \frac{DTn_0}{m^2 v_T^2} \frac{1}{V_c} \int d\mathbf{x}' \int d\mathbf{x} |\mathbf{f}(\mathbf{x}; \mathbf{x}')|^2. \quad (7.136)$$

При выводе (7.136) использовалась периодичность \mathbf{f} для приведения усредняющего интеграла к одной ячейке. Дополнительное предположение об отсутствии временных корреляций позволяет использовать (7.136) на последующих шагах по времени, что дает линейный рост энергии, характерный для стохастических процессов (см. разд. 9.2.2).

Применяя к интегралу по x (7.136) теорему Парсевала (табл. П.4) и используя для фурье-образа \hat{f} функции f соотношения (5.223) и (5.224), получаем (для одномерной системы)

$$\frac{1}{H} \int dx' \int dx |f(x; x')|^2 = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{q^4}{H^4} \hat{W}^2 |\hat{D}|^2 \hat{G}^2 \left[\sum_n \hat{W}^2(k_n) - \hat{W}^2(k) \right] \quad (7.137)$$

$$= \int_{k_g} \frac{dk}{2\pi} q^4 |\hat{D}|^2 \hat{G}^2 \left[\left(\sum_n \frac{\hat{W}^2(k_n)}{H^2} \right)^2 - \sum_n \frac{\hat{W}^4(k_n)}{H^4} \right]. \quad (7.138)$$

В случае ограниченной системы с длиной периода L , $\int_{k_g} dk/2\pi$ заменяется на $\sum_{l=0}^N g^{-1} k_0/2\pi = \sum_l 1/L$, а для многомерных систем k заменяется на \mathbf{k} , D на \mathbf{D} и т. д., как описывалось в разд. 7.4.2. Из выражения (7.138) следует, что если бы спектр \hat{W} имел конечную ширину, то флуктуации исчезли бы и (в пределе $DT \rightarrow 0$) энергия сохранялась. Для схем, применяемых на практике (NGP, SIC, TSC и т. д.), выражение в квадратных скобках в (7.138) принимает наибольшее значение вблизи границ основной зоны $|k| = \pi/H$, поскольку именно там сеточные гармоники дают самый большой вклад, хотя может оказаться, что наибольший вклад во флуктуации полной силы придется на более длинноволновую часть благодаря множителю $|\hat{D}|^2 \hat{G}^2 \sim 1/k^2$. Например, используя (7.59) и (7.66), находим, что для схемы NGP

$$\left(\sum_n \frac{\hat{W}^2(k_n)}{H^2} \right)^2 - \sum_n \frac{\hat{W}^4(k_n)}{H^4} = \frac{2}{3} \sin^2 \frac{kH}{2}. \quad (7.139)$$

Это дает

$$\langle f^2 \rangle = \frac{1}{H} \int dx' \int dx |f(x; x')|^2 = \int_{k_g} \frac{dk}{2\pi} \frac{q^4}{\varepsilon_0^2} |\hat{D}|^2 (\varepsilon_0 \hat{G})^2 \left(\frac{2}{3} \sin^2 \frac{kH}{2} \right). \quad (7.140)$$

Из равенства (7.140) ясно, что использование более широкой частицы конечного размера (приводящее к более быстрому спаду \hat{G} с ростом k) сильно уменьшает $\langle f^2 \rangle$ и тем самым улучшает сохранение энергии: см. гл. 9 и работу [303].

Для модели NGP, описанной в гл. 2, интеграл (7.140) мож-

фициент, определяющий ширину частицы), приведем (7.140) к виду

$$\langle f^2 \rangle = \frac{q^4}{\epsilon_0^2} \frac{H^2}{3} \int_0^{(\alpha k_g)/2} \frac{dk}{2\pi} \cos^2 \frac{kH}{2} = \frac{q^4}{\epsilon_0^2} H \frac{\alpha}{12} [1 + \text{sinc } \alpha] \quad (7.141)$$

и получим для удельной скорости нагрева (7.136) выражение

$$\langle r \rangle_{\text{NGP}} = \omega_p (\omega_p DT) \frac{H}{\lambda_D} \frac{1}{n_0 \lambda_D} \left(\frac{\alpha}{12} (1 + \text{sinc } \alpha) \right). \quad (7.142)$$

Аналогично, для схемы СІС

$$\langle f^2 \rangle = \int_{k_g} \frac{dk}{2\pi} \frac{q^4}{\epsilon_0^2} |\hat{D}|^2 (\epsilon_0 \hat{G})^2 \left[\frac{2}{45} \sin^4 \frac{kH}{2} \left(1 + \frac{2}{7} \sin^2 \frac{kH}{2} \right) \right]. \quad (7.143)$$

Подстановка в (7.143) выражений для \hat{D} и \hat{G} из (7.57) и (7.58) дает скорость нагрева для схемы СІС:

$$\langle r \rangle_{\text{СІС}} = \frac{\omega_p^2 HDT}{n_0 \lambda_D^2} \left[\frac{\alpha}{315} \left(1 - \frac{\sin c2\alpha}{2} - \text{sinc } \alpha \frac{\sin^2 \pi\alpha}{24} \right) \right]. \quad (7.144)$$

При оставлении всех гармоник ($\alpha = 1$) уменьшение скорости нагрева, вызванное заменой весовой функции NGP на функцию СІС, равно

$$\frac{\langle r \rangle_{\text{NGP}}}{\langle r \rangle_{\text{СІС}}} = \frac{\frac{1}{12}}{\frac{1}{315}} = 26 \frac{1}{4},$$

т. е. время, за которое стохастический нагрев заметно изменит полную энергию, в одномерной схеме СІС примерно в 25 раз больше, чем в NGP. Если, кроме того, ввести обрезание по волновому числу ($\alpha < 1$), то скорость нагрева будет уменьшаться в обеих схемах, хотя в СІС это уменьшение будет проявляться сильнее.

Представленные здесь результаты простого анализа роста энергии в модельной плазме довольно хорошо согласуются с результатами измерений [309], несмотря на то, что его применимость ограничена малыми $\omega_p DT$, $H \gtrsim \lambda_D$ и $(n_0 \lambda_D) \gg 1$. Если не налагать этих ограничений, то картина усложнится за счет погрешностей интегрирования по времени, нефизической связи гармоник, а также корреляционных (т. е. кинетических) эффектов. Несмотря на эти оговорки, такие результаты, как (7.142) и (7.143), служат полезным ориентиром в оценке изменений в сохранении энергии, вызываемых изменениями характерных параметров типа $\omega_p DT$, H/λ_D или $n\lambda_D$, а также сменой алгоритма вычисления силы (что предполагает изменения в \hat{D} , \hat{G} и \hat{W}).

7.6.2. Импульс

Основной причиной несохранения энергии в схемах, сохраняющих импульс, является наличие неконсервативной части силы, обусловленной недостаточной степенью дискретизации. Несохранение импульса в схемах, сохраняющих энергию, — другое проявление тех же самых погрешностей. Переход к форме сеточной аппроксимации силы, позволяющей записать силу в виде градиента скалярного потенциала, противоречит той симметрии, которая требуется для сохранения импульса.

Частью силы, ответственной за нарушение симметрии в схемах, сохраняющих энергию, является член самодействия (см. гл. 5). Отсюда изменение полного импульса Δp за один временной шаг равно

$$\Delta p = \sum_{i=1}^{N_p} DT F^{\text{сам}}. \quad (7.145)$$

Из тех же соображений о случайном блуждании, которые использовались в предыдущем разделе для энергии, найдем, что математическое ожидание скорости изменения полного импульса и скорости изменения полного квадрата импульса равны соответственно

$$\frac{\langle \Delta p \rangle}{DT} = 0, \quad \frac{\langle \Delta p^2 \rangle}{DT} = DT \frac{N_p}{H} \int_{-H/2}^{H/2} (F^{\text{сам}})^2 dx. \quad (7.146)$$

Равенства (7.146) означают, что в среднем, как и следовало ожидать, схемы, сохраняющие энергию, импульс сохраняют, но при этом отклонения от сохранения могут нарастать во времени (в положительную или отрицательную сторону) по линейному закону.

Чтобы обеспечить малость дрейфа полного импульса, потребуем, чтобы характерное увеличение неопределенности импульса за время одного шага было мало, т. е.

$$\frac{\langle \Delta p^2 \rangle}{N_p m^2 v_T^2} = \frac{DT^2}{m^2 v_T^2 H} \int (F^{\text{сам}})^2 dx \ll 1. \quad (7.147)$$

Для одномерной схемы, сохраняющей энергию, в которой использованы кусочно-линейные базисные функции (разд. 5.5), сила самодействия определяется выражением (см. разд. 5.2.4 и 5.5)

$$F^{\text{сам}} = -\frac{q^2}{\epsilon_0 H} x, \quad -\frac{H}{2} < x \leq \frac{H}{2}, \quad (7.148)$$

что дает

$$\frac{\langle \Delta p^2 \rangle}{N_p m^2 v_T^2} = \frac{\omega_p^2 DT^2}{12 (n_0 \lambda_D)^2} = \frac{\omega_{\text{сам}}^2 DT^2}{12 (n_0 \lambda_D)} \left(\frac{H}{\lambda_D} \right). \quad (7.149)$$

Следовательно, если $\omega_p DT \ll 1$ и на дебаевскую длину приходится большое число частиц ($n_0 \lambda_D \gg 1$), медленный дрейф полного импульса не доставит сколь-либо серьезных неприятностей в схемах, сохраняющих энергию. Экспериментальные измерения дрейфа полного импульса приведены в работе [254].

7.6.3. Момент количества движения

В задачах о галактиках (см. гл. 12) важно, чтобы сохранялся момент количества движения. Скорость изменения полного момента количества движения системы N_p частиц равна вращающему моменту:

$$\frac{d\mathbf{L}}{dt} = \sum_{i=1}^{N_p} \mathbf{x}_i \times \mathbf{F}_i = \sum_{i=1}^{N_p-1} \sum_{j=i+1}^{N_p} (\mathbf{x}_i - \mathbf{x}_j) \times \mathbf{f}_{ij}, \quad (7.150)$$

где \mathbf{F}_i — полная сила, действующая на i -ю частицу, а \mathbf{f}_{ij} — сила, с которой j -я частица действует на i -ю частицу. Для частиц с симметричными распределениями заряда (либо массы для гравитирующих систем), т. е. такими распределениями, которые являются функциями только от расстояния до центра частиц, \mathbf{f}_{ij} параллельны $\mathbf{x}_i - \mathbf{x}_j$. В этом случае вращающий момент равен нулю и \mathbf{L} сохраняется. Однако в периодических системах за счет зарядов-изображений сохранение \mathbf{L} нарушается, как, скажем, в задачах гл. 12.

Вращающий момент системы частиц, движение которых подчиняется уравнениям схемы с перешагиванием, равен $\sum_i \mathbf{x}_i^n \times \mathbf{F}_i^n$, где n обозначает номер временного шага. Если момент количества движения на временном шаге $n + 1/2$ определить формулой

$$\mathbf{L}^{n+1/2} = m \sum_i \mathbf{x}_i^n \times \mathbf{v}_i^{n+1/2} = \frac{m}{DT} \sum_i \mathbf{x}_i^n \times \mathbf{x}_i^{n+1}, \quad (7.151)$$

то в соответствии с уравнениями движения схемы с перешагиванием получим

$$\frac{\mathbf{L}^{n+1/2} - \mathbf{L}^{n-1/2}}{DT} = \sum_i \mathbf{x}_i^n \times \mathbf{F}_i^n. \quad (7.152)$$

Следовательно, если \mathbf{f}_{ij}^n параллельна $\mathbf{x}_i^n - \mathbf{x}_j^n$, то в системе, подчиняющейся разностным уравнениям движения (схема с перешагиванием), момент количества движения также сохраняется. Однако, как мы видели в гл. 5, сила взаимодействия частиц, полученная с помощью сеточной аппроксимации, вообще говоря, не параллельна линии, соединяющей обе частицы, причем это отклонение обусловлено погрешностями, сопровождающими процедуры распределения заряда и интерполяции силы, анизотропией потенциала, а также погрешностями представле-

ния направления, вносимыми разностной аппроксимацией градиента (для схем, сохраняющих импульс). Тем не менее с помощью метода Q-оптимизации, о котором говорится ниже и в гл. 8, эти погрешности, приводящие к зависимости силы от направления, удается сделать весьма незначительными.

7.7. ОПТИМИЗАЦИЯ

При создании всякой программы моделирования бесстолкновительных систем методом частица — сетка мы сталкиваемся с большим разнообразием возможных схем распределения заряда, нахождения потенциала, дифференцирования и интерполяции силы. Что касается двумерных и трехмерных моделей, то ограниченность объема памяти и быстродействия компьютеров делают неизбежным компромисс между точностью представления рассматриваемой физической системы и стоимостью расчета.

Ограниченность вычислительных ресурсов сказывается прежде всего в том, что экспериментатор-вычислитель должен обойтись относительно малым числом частиц и грубой сеткой. Если число частиц на дебаевской длине $N_D = n_0 \lambda_D$ (или $n_0 \lambda_D^2$ и $n_0 \lambda_D^3$ в двумерном и трехмерном случаях) становится малым, корреляции частиц могут заметным образом изменить коллективные свойства модельной плазмы. Характерное время, за которое такие корреляции становятся доминирующими, определяется временем столкновения τ_c . Как показано в разд. 7.5 и 9.2.3, увеличивая эффективную ширину частицы конечного размера, удается сделать τ_c приемлемо большим для относительно малых N_D (5—10). Влияние грубости сетки двойко. Она ограничивает столкновительные эффекты за счет того, что эффективная ширина частицы не может быть меньше размера ячейки H , и в то же время она ограничивает область коллективных мод гармониками с волновыми числами из основной зоны $|k| < \pi/H$.

Влияние размеров частицы на коллективные свойства можно оценить с помощью дисперсионного соотношения с усредненной силой. Из приближения тепловой плазмы (разд. 7.3.3) ясно, что наклон дисперсионной кривой $\omega = \omega(k)$ падает с увеличением эффективного размера частицы, и для некоторых профилей частиц, размер которых намного превышает дебаевскую длину, групповая скорость $d\omega/dk$ может отличаться по знаку от групповой скорости в физической системе. Кроме того, взаимодействие мод, которое привносится недостаточной частотой дискретизации плотности заряда (разд. 7.3.4), влияет на инкременты, и если ширина ячейки H больше одной-двух дебаевских длин, оно может приводить к нефизическим неустойчивостям. Однако при условии, что $H \leq \lambda_D$, сеточные эффекты сводятся главным образом к медленному нефизическому росту полной энергии системы (см. разд. 7.6.1 и 9.2.3).

Цель, которая ставится при оптимизации бесстолкновительных вычислительных моделей — добиться наилучшего представления коллективных свойств при заданных вычислительных затратах. Эффект наложения ограничивает максимально возможную ширину ячейки приблизительно дебаевской длиной. Длина расчетной области L должна быть намного больше рассматриваемых длин волн, чтобы плотность допустимых волновых чисел (расположенных с шагом $k_0 = 2\pi/L$) была достаточно велика и описывала физическое нелинейное взаимодействие мод, т. е. $H \leq \lambda_D \ll L$. Выбор значений H и L определяет количество узлов сетки $N_g = L/H$. После этого число частиц диктуется в основном необходимостью обеспечить бесстолкновительность, т. е. нужно, чтобы время столкновения было больше времени моделирования — $n_{\max}DT \ll \tau_c$.

7.7.1. Сила взаимодействия частиц

При определении свойств дискретной вычислительной модели важной величиной является сила взаимодействия частиц. Различия между свойствами вычислительной модели и физической системы непосредственно связаны с конечным шагом по времени, ограниченным числом частиц и отличием сил межчастичного взаимодействия в этих двух случаях. При заданных ограничениях на размер сетки и число частиц, связанных с ограниченными машинными ресурсами, вопрос об оптимизации сводится к поиску оптимальной комбинации известных элементов из цикла вычисления силы. В лучшем случае вычисление силы методом частица-сетка может передать только некоторую эталонную силу с ограниченной шириной спектра $\mathbf{R}(\mathbf{x}) \supset \hat{\mathbf{R}}(\mathbf{k})$, где $\mathbf{R}(\mathbf{x})$ выбирается (с помощью теории частиц конечного размера) так, чтобы получить требуемые коллективные свойства.

В качестве количественного критерия Q качества конкретной схемы принимается среднеквадратичное отклонение силы $\mathbf{F}(\mathbf{x}, \bar{\mathbf{x}})$, рассчитанной методом частица-сетка, от эталонной силы $\mathbf{R}(\mathbf{x})$:

$$Q = \frac{1}{V_c} \int_{V_c} d\bar{\mathbf{x}} \int d\mathbf{x} (\mathbf{F}(\mathbf{x}, \bar{\mathbf{x}}) - \mathbf{R}(\mathbf{x}))^2, \quad (7.153)$$

где усреднение по относительным положениям сетки $\bar{\mathbf{x}}$ приведено в одной ячейке с учетом периодичности \mathbf{F} по этому аргументу. Теорема Парсеваля (табл. П.4) позволяет записать Q через гармоники $\hat{\mathbf{R}}(\mathbf{k}) \supset \mathbf{R}(\mathbf{x})$ и $\hat{\mathbf{F}}(\mathbf{k}, \bar{\mathbf{x}}) \supset \mathbf{F}(\mathbf{x}, \bar{\mathbf{x}})$:

$$Q = \frac{1}{V_c} \int_{V_c} d\bar{\mathbf{x}} \int \frac{d\mathbf{k}}{(2\pi)^d} \{ |\hat{\mathbf{F}}(\mathbf{k}, \bar{\mathbf{x}})|^2 + |\hat{\mathbf{R}}(\mathbf{k})|^2 - 2\hat{\mathbf{F}}(\mathbf{k}, \bar{\mathbf{x}}) \cdot \hat{\mathbf{R}}^*(\mathbf{k}) \}, \quad (7.154)$$

где d ($= 1, 2$, или 3) — размерность пространства, а $\hat{\mathbf{R}}^*$ — функция, комплексно сопряженная $\hat{\mathbf{R}}$. Выражая $\hat{\mathbf{F}}$ через фурье-образы величин из цикла вычисления силы (см. разд. 5.6.4 и 7.4) и приводя интегралы по \mathbf{k} к основной зоне, для единичных зарядов в схеме, сохраняющей импульс, получаем

$$Q = \int_{k_g} \frac{d\mathbf{k}}{(2\pi)^d} \left\{ \hat{G}^2 |\hat{\mathbf{D}}|^2 \left[\sum_{\mathbf{n}} \hat{U}^2 \right]^2 - 2\hat{G}\hat{U}^2\hat{\mathbf{D}} \cdot \hat{\mathbf{R}}^* + |\hat{\mathbf{R}}|^2 \right\}, \quad (7.155)$$

где $O = \mathcal{W}/H$. В случае периодических систем интеграл $\int_{k_g} d\mathbf{k}/(2\pi)^d$ заменяется суммой по гармоникам $\sum_{\mathbf{n}} V_b^{-1}$ (см. разд. 7.4).

Если для нахождения потенциала используется конечное преобразование Фурье (см. гл. 6), мы вправе в качестве \hat{G} выбрать любую вещественную четную в основной зоне функцию. Очевидно, что оптимальным вариантом будет такой, который минимизирует Q , что, исходя из (7.155), дает оптимальную функцию влияния

$$\hat{G}(\mathbf{k}) = \frac{\hat{\mathbf{D}}(\mathbf{k}) \cdot \hat{\mathbf{R}}^*(\mathbf{k}) \hat{U}^2(\mathbf{k})}{|\hat{\mathbf{D}}(\mathbf{k})|^2 \left[\sum_{\mathbf{n}} \hat{U}^2(\mathbf{k}_{\mathbf{n}}) \right]^2} \quad (7.156)$$

с минимальным Q :

$$Q = \int_{k_g} \frac{d\mathbf{k}}{(2\pi)^d} |\hat{\mathbf{R}}|^2 \left\{ 1 - \frac{|\hat{\mathbf{D}} \cdot \mathbf{k}|^2 U^4}{|\hat{\mathbf{D}}|^2 |\mathbf{k}|^2 \left[\sum_{\mathbf{n}} \hat{U}^2 \right]^2} \right\}. \quad (7.157)$$

Выражение (7.156) указывает на то, что точность вычисления силы получится выше, если в алгоритм решения уравнения для потенциала ввести поправки, которые компенсируют средние погрешности, связанные с распределением заряда и интерполяцией силы, а также сглаживающие эффекты от дифференцирования потенциала. Остаточные ошибки (см. (7.157)) обусловлены тем, что разностный градиент не параллелен истинному градиенту (т. е. $|\hat{\mathbf{D}} \cdot \mathbf{k}| \neq |\hat{\mathbf{D}}| |\hat{\mathbf{k}}|$) и весовая функция не является функцией с ограниченной шириной спектра (т. е. $\sum_{\mathbf{n}} \hat{W}(\mathbf{k}_{\mathbf{n}}) \neq \hat{W}(\mathbf{k})$).

Улучшения, достигаемые изменением вида функции \hat{G} , не влекут за собой никаких дополнительных расходов на нахождение потенциала, если используются методы быстрого преобразования Фурье. Более точные, но и более дорогие в вычислительном отношении схемы получают применением операторов градиента \mathbf{D} высокого порядка (для ослабления угловой зависимости погрешностей градиента) или интерполяционных функций более высокого порядка, чтобы получить быстрое убывание \mathcal{W} по \mathbf{k} , или же обоих способов одновременно. В качестве альтернативы увеличению порядка интерполяционных функций в

разд. 7.8 описан метод чередующихся сеток. Кроме того, в следующей главе будет рассмотрен порядок согласованного выбора операторов градиента и схем интерполяции.

7.7.2. Одномерные схемы

В одномерных схемах последствия дифференцирования потенциала можно полностью исключить при помощи компенсирующего множителя в \hat{G} , так что для одномерной периодической системы длиной L выражение для Q принимает вид

$$Q = \frac{1}{L} \sum_{l=0}^{N_g-1} |\hat{R}|^2 \left\{ 1 - \frac{\hat{W}^4}{\left[\sum_n \hat{W}^2 \right]^2} \right\} = \frac{1}{L} \sum_l Q(k). \quad (7.158)$$

Относительная среднеквадратичная погрешность в l -й гармонике ($k = 2\pi l/H$) определяется (в процентах) выражением:

$$\sqrt{Q^+} = 100 \frac{Q(k)}{|\hat{R}(k)|^2}. \quad (7.159)$$

На рис. 7.6 изображены графики $\sqrt{Q^+}$ для схем, сохраняющих импульс (т. е. NGP, CIC и TSC), и для схем, сохраняющих энергию, в которых использованы линейная интерполяция (ELI) и квадратичные сплайны (EQS). Схема ELI представляет собой комбинацию схем CIC и NGP, а EQS соответствует схемам TSC/CIC. Величина $\sqrt{Q^+}$ складывается из двух факторов: флуктуирующей части, появление которой связано с недостаточной частотой дискретизации распределения заряда, и нефлуктуирующей части, возникающей из-за того, что спектр сеточной силы не ограничен конечным интервалом волновых чисел [115]. Флуктуирующая часть характеризует степень интенсивности взаимодействия мод (см. разд. 7.3.4) и вклад в стохастический нагрев (см. разд. 7.6.1) в зависимости от волнового числа, а нефлуктуирующая часть показывает меру отклонения вещественной части дисперсионного соотношения от соответствующей зависимости для эталонной силы. Таким образом, $\sqrt{Q^+}/k$ дает простейшую меру качества представления коллективных свойств.

Показанные на рис. 7.6 кривые позволяют для любой конкретной задачи выбрать наиболее рентабельную схему. Например, предположим, что мы хотим промоделировать систему длиной $100\lambda_D$ и нам надо, чтобы для всех длин волн, больших $2\lambda_D$, амплитуды гармоник силы отличались от гармоник для физической системы точечных частиц меньше чем на 10%. Какая же из пяти схем, показанных на рис. 7.6, самая дешевая по вычислительным затратам на одну частицу за шаг по времени?

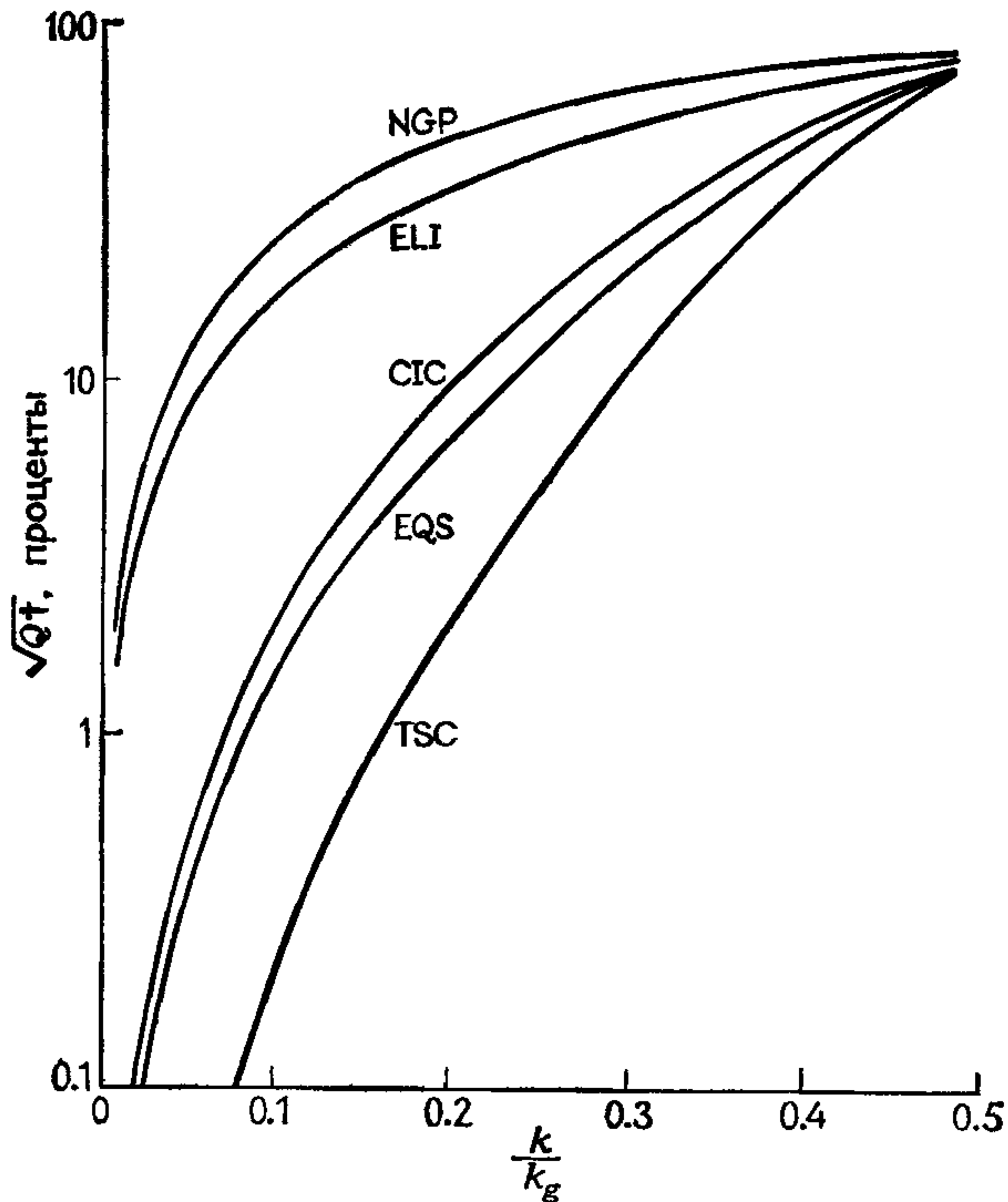


Рис 76. Среднеквадратичное относительное отклонение амплитуд гармоник сеточной силы от гармоник эталонной силы как функция волнового числа для произвольной силы со спектром конечной ширины. Надписи у кривых соответствуют типу интерполяционных функций (см в тексте) (из работы [115], с любезного разрешения *Journal of Computational Physics*).

Находим из рис. 7.6 максимальные значения μ_i отношения k/k_g , при которых прямая $\sqrt{Q^+} = 10\%$ пересекает кривые для каждой из пяти схем. Приравнивая отвечающую μ_i длину волны значению $2\lambda_D$ и используя соотношения $L = N_g H = 100\lambda_D$ и $k_g = 2\pi/H$, получаем, что минимальное число узлов сетки, при котором i -я схема может отвечать предъявленному условию, равно $N_i = 50\mu_i$. Если считать, что уравнения для полей решаются с помощью быстрого преобразования Фурье, то число операций T_i для i -й схемы приблизительно составит

$$T_i = \alpha_i N_p - \beta N_i \log N_i. \quad (7.160)$$

Здесь α_i определяет полное количество операций, приходящихся в расчетном цикле на одну частицу за шаг по времени, на этапах распределения заряда, интерполяции силы и интегрирования по времени, N_p — число частиц и N_i — число сеточных узлов. Число частиц m_{ij} , при котором i -я и j -я схемы одинаково

выгодны, находится из равенства $T_i = T_j$:

$$m_{ij} = \frac{-\beta (N_i \log N_i - N_j \log N_j)}{(\alpha_i - \alpha_j)}.$$

Используя для оценки α_i число арифметических операций с вещественными числами, имеем для схем TSC/EQS, EQS/CIC, CIC/ELI и ELI/NGP соответственно $\Delta\alpha = \alpha_i - \alpha_j = 5, 6, 3$ и 4 . Используя $\beta = 5$ (см. разд. 6.5.2) и соответствующие значения N_i , находим области значений N_p , в пределах которых каждая схема является самой дешевой по затратам на частицу за шаг по времени:

$$\text{NGP: } N_p > 8000,$$

$$\text{CIC: } 325 < N_p < 8000,$$

$$\text{TSC: } N_p < 325.$$

Нигде схемы, сохраняющие энергию, самыми рентабельными не оказываются.

Для того чтобы получить гармоники (а значит, и дисперсионное соотношение), соответствующие физической системе точечных частиц, $\hat{R}(k)$ полагается равной известному значению для точечных частиц для волновых чисел, меньших $k_{\max} = 2\pi/(2\lambda_D)$. Волновые числа, бóльшие k_{\max} , подавляются посредством задания $\hat{G} = 0$ для $k > k_{\max}$ с тем, чтобы улучшить сохранение энергии и уменьшить число частиц, необходимое для получения достаточно большого времени столкновения.

Условие $N_D \gg 1$, определяющее время столкновения, исключает режим, где схема TSC наиболее экономична. Выясняется, что в той области значений N_D ($\sim 10-20$), для которой получаются удовлетворительные времена столкновений, наилучшей для рассматриваемого примера схемой является CIC (с $N_g = 256$). После того как эта схема выбрана, чтобы убедиться, что результаты вычислительных экспериментов будут отражать физические, а не численные эффекты, можно воспользоваться заранее вычисленными дисперсионными характеристиками (см. разд. 7.3), временами столкновений (разд. 7.5) и временами нагрева (разд. 7.6.1), а также провести контрольные эксперименты по программе.

Преимущество оптимизированных схем по сравнению с другими возможными схемами иллюстрируется рис. 7.7. Если нам необходимо, чтобы для удерживаемых длин волн дисперсионные свойства моделей частица—сетка отвечали результату для точечных частиц, то различные алгоритмы можно сравнивать по относительному отклонению амплитуд гармоник от эталонной силы

$$\hat{R}(k) = -\frac{ik}{\epsilon_0 k^2}.$$

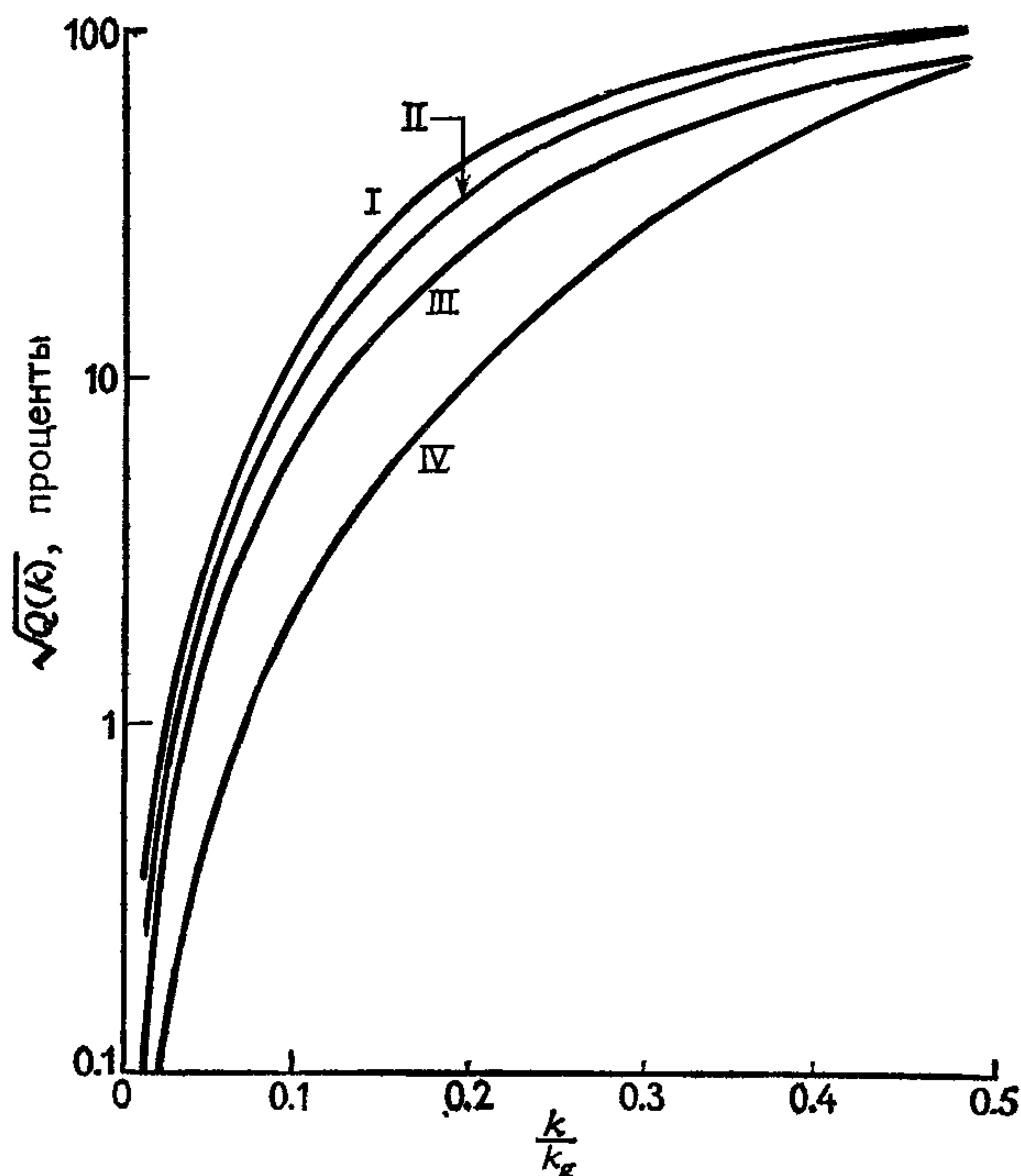


Рис. 7.7. Среднеквадратичное относительное отклонение амплитуд гармоник сеточной силы от гармоник кулоновской силы для четырех вариантов схемы СИС (из работы [115], с разрешения журнала *Journal of Computational Physics*).

На рис. 7.7 такое сравнение проведено для четырех вариантов схемы СИС, отличающихся методом определения поля по плотности заряда. Обозначения у кривых следующие:

I: «Дешевый метод» решения уравнения Пуассона [42]:

$$E_p = \frac{\varphi_{p-1} - \varphi_{p+1}}{2H}, \quad \hat{\varphi}(k) = \frac{\hat{\rho}(k)}{\varepsilon_0 k^2}.$$

II: Конечно-разностная аппроксимация

$$E_p = \frac{\varphi_{p-1} - \varphi_{p+1}}{2H}, \quad \frac{\varphi_{p-1} - 2\varphi_p + \varphi_{p+1}}{2H} = \frac{\rho_p}{\varepsilon_0}.$$

III: Прямое решение в k -пространстве

$$\hat{E}(k) = -\frac{i\hat{\rho}(k)}{\varepsilon_0 k}.$$

IV: Оптимальная (с минимальным Q) схема.

Погрешности, встроенные в функцию влияния, компенсируют погрешности распределения заряда и интерполяции силы до та-

кой степени, что суммарные погрешности для оптимизированной схемы обычно раза в три меньше, чем для других схем. Более заметные улучшения получаются для интерполяционных схем более высокого порядка или оптимизированных схем с чередованием, поскольку ограничивающим фактором является отношение величин основных и сеточных гармоник.

7.8. ЧЕРЕДОВАНИЕ

В расчетах методом частица—сетка основные трудности доставляет эффект наложения. Наличие в спектре ложных компонент приводит к взаимодействию мод и несохранению энергии или импульса, а также, в более явном виде, к флуктуациям сил взаимодействия частиц при их смещении относительно сетки. Флуктуации возникают из-за того, что при смещении выборочных точек фаза сеточных гармоник относительно основной гармоники меняется (см. разд. 5.6.1 и рис. 5.7). Это изменение фазы и используется в методе чередования для подавления некоторых из сеточных гармоник.

Самым простым применением идеи чередования является операция «баттерфляй», на которой базируется алгоритм быстрого преобразования Фурье. Дискретное преобразование Фурье $\{\hat{D}_l; l \in [0, 2N - 1]\}$ на множестве $2N$ значений $\{D_p\}$, по определению, есть

$$\hat{D}_l = \sum_{p=0}^{2N-1} D_p e^{-i2\pi lp/2N}. \quad (7.161)$$

Расщепляя сумму в правой части (7.161) на члены с четными и нечетными p и вводя обозначения $D_{2p} = A_q$, $D_{2p+1} = B_q$, получаем:

$$\hat{D}_l = \sum_{q=0}^{N-1} A_q e^{-i2\pi lq/N} + e^{-i\pi l/N} \sum_{q=0}^{N-1} B_q e^{-i2\pi lq/N}. \quad (7.162)$$

Сравнивая суммы в (7.162) с определением преобразования (7.161), мы видим, что эти суммы определяют конечные преобразования по половинному числу точек, т. е.

$$\begin{aligned} \hat{D}_l &= \hat{A}_l + e^{-i\pi l/N} \hat{B}_l, \\ \hat{D}_{l+N} &= \hat{A}_l - e^{-i\pi l/N} \hat{B}_l, \end{aligned} \quad 0 \leq l \leq N - 1. \quad (7.163)$$

Соотношения (7.163) определяют операцию «баттерфляй», названную так из-за внешнего вида схемы потоков данных, изображающей вычисление \hat{D} по \hat{A} и \hat{B} . Важный момент, который необходимо отметить по поводу соотношений (7.163), состоит в

том, что если мы знаем по отдельности результат преобразования наборов величин с четными и нечетными p , то с помощью операций «баттерфляй» — сложения и вычитания — можно найти преобразование полного набора

Чтобы приложить указанные идеи четных и нечетных выборок к непрерывным величинам, воспользуемся функцией выборки Π . Если мы имеем некоторую функцию $A(x) \supset \hat{A}(k)$, то функция $C_1^+ = \Pi(x/H) A(x)$ задает «четную» выборку, а

$$C_2^+ = \Pi\left(\frac{x}{H} + \frac{1}{2}\right) A(x)$$

— «нечетную» выборку. Используя теорему о свертке, теорему о сдвиге и определение функции Π , находим:

$$C_1^+ = \Pi\left(\frac{x}{H}\right) A(x) \supset \hat{C}_1 = H \Pi\left(\frac{k}{k_g}\right) * \hat{A}(k) = \sum_n \hat{A}(k - nk_g),$$

$$C_2^+ = \Pi\left(\frac{x}{H} + \frac{1}{2}\right) A(x) \supset \hat{C}_2 = H \Pi\left(\frac{k}{k_g}\right) e^{+ikH/2} * \hat{A}(k) \\ = \sum_n (-1)^n \hat{A}(k - nk_g).$$

(7.164)

Отсюда

$$\hat{C} = \frac{1}{2} (\hat{C}_1 + \hat{C}_2) = \sum_{\substack{n=-\infty \\ n-\text{четные}}}^{\infty} \hat{A}(k - nk_g). \quad (7.165)$$

Графическая интерпретация этих результатов представлена на рис 7.8. Сеточные гармоники второй (нечетной) выборки меняют поочередно знак за счет множителя $e^{-ink_g H/2} = (-1)^n$.

Аналогичные результаты, но чуть более сложным образом, можно получить и для двух чередующихся выборок, несцентрированных относительно средней точки. Для исключения всех сеточных гармоник, кроме гармоник с n , кратным 3, можно использовать три чередующиеся выборки; комбинация четырех выборок позволяет сохранить лишь каждую четвертую сеточную гармонику и т. д. Три чередующиеся выборки соответствуют алгоритму быстрого преобразования Фурье по основанию 3, четыре — алгоритму по основанию 4 и т. д.

Фазовый множитель (или «вертушка») $e^{i\pi n/N}$, фигурирующий в (7.163), присутствует также и в (7.165), но неявно, поскольку он входит в само определение преобразования:

$$\hat{C}_2 = \int dx \Pi\left(\frac{x}{H} - \frac{1}{2}\right) A(x) e^{-ikx}.$$

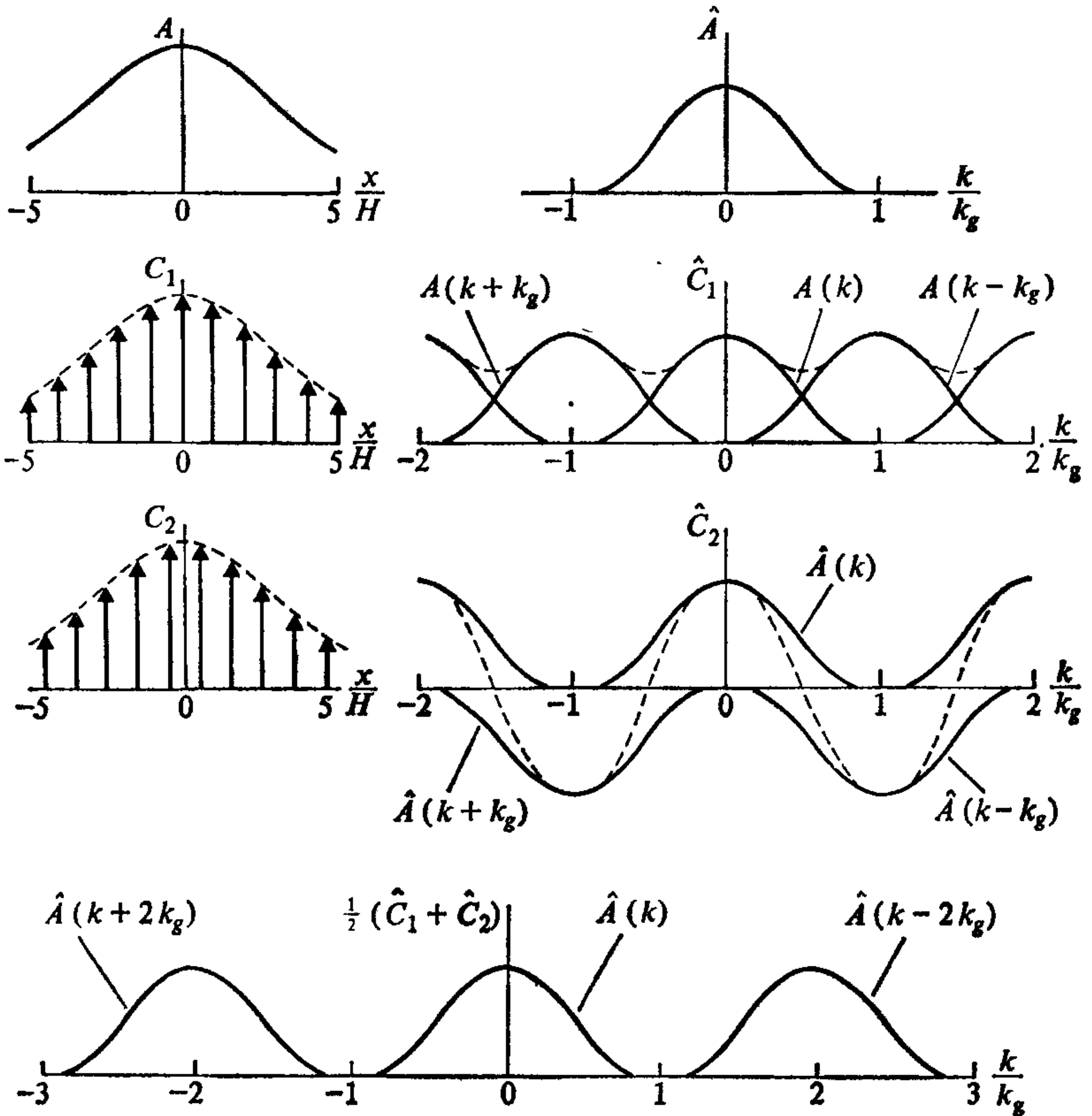


Рис. 7.8 Графическое представление выражений (7.164) и (7.165). Усреднение фурье-образов «четной» выборки $C_1^+ \supset \hat{C}_1$ и «нечетной» выборки $C_2^+ \supset \hat{C}_2$ приводит к взаимному уничтожению всех нечетных сеточных гармоник функции A в $\hat{C} = (\hat{C}_1 + \hat{C}_2)/2$.

Этот множитель можно выделить в явном виде, если операцию сдвига во второй выборке проводить не для $\Pi(x/H)$, а для функции $A(x)$ (что равносильно). В этом случае находим:

$$\begin{aligned}
 C_2^+ &= \Pi\left(\frac{x}{H}\right) A(x - H/2) \supset \hat{C}_2 = H \Pi\left(\frac{k}{k_g}\right) * e^{-ikH/2} \hat{A}(k) \\
 &= e^{-ikH/2} \sum_n (-1)^n \hat{A}(k - nk_g) \\
 \Rightarrow \hat{C} &= \frac{1}{2} (\hat{C}_1 + e^{-ikH/2} \hat{C}_2).
 \end{aligned}$$

(7.166)

7.8.1. Усреднение силы

Впервые применить чередующуюся выборку для уменьшения сеточных эффектов в расчетах методом частица—сетка предложили Чен и др. [75]. В их алгоритме вычисление силы на каждом шаге по времени производится в два этапа. На первом этапе сила вычисляется обычным образом, что мы символически изобразим (пользуясь введенными в разд. 5.6 обозначениями) в виде

$$n \rightarrow \rho_1^\dagger \rightarrow \varphi_1' \rightarrow E_1^\dagger \rightarrow F_1.$$

Последовательные стрелки символически изображают распределение заряда, нахождение потенциала, дифференцирование потенциала и получение полей и, наконец, интерполяцию силы, а индексом 1 обозначен этап 1. После этого половина силы $F_1/2$ добавляется к скоростям частиц. Второй этап является повторением первого, но используется сетка, сдвинутая по пол-ячейки относительно первой. Таким образом, цикл продвижения на шаг по времени можно символически изобразить в следующем виде:

$$\begin{array}{c} \rho_1^\dagger \rightarrow \varphi_1' \rightarrow E_1^\dagger \rightarrow F_1 \\ \swarrow \quad \searrow \\ n \\ \searrow \quad \swarrow \\ \rho_2^\dagger \rightarrow \varphi_2' \rightarrow E_2^\dagger \rightarrow F_2 \end{array} \quad F = \frac{(F_1 + F_2)}{2}$$

Фурье-образ силы F_1 , вычисляемый на первом этапе цикла по временному шагу, не отличается от получаемого по схеме без чередования:

$$\hat{F}_1 = q^2 \frac{\hat{W}}{H} \hat{D} \hat{G} \sum_{n=-\infty}^{\infty} \frac{\hat{W}(k_n)}{H} \hat{n}(k_n). \quad (7.167)$$

Для вывода формулы преобразования Фурье на всех шагах второго этапа требуется лишь дополнительно воспользоваться теоремой о сдвиге (табл. П.4):

$$\begin{aligned} \rho_2^\dagger = \Pi \left(\frac{x}{H} + \frac{1}{2} \right) \rho(x) &\supset \hat{\rho}_2(k) = H \Pi \left(\frac{k}{k_g} \right) e^{ikH/2} * \hat{\rho}' = \\ &= \frac{q}{H} \sum_n (-1)^n \hat{W}(k_n) \hat{n}(k_n), \end{aligned} \quad (7.168)$$

$$\varphi_2' = G' * \rho_2^\dagger \supset \hat{\varphi}_2' = \hat{G}' \hat{\rho}_2, \quad (7.169)$$

$$\begin{aligned} E_2^\dagger = \Pi \left(\frac{x}{H} + \frac{1}{2} \right) D * \varphi_2' &\supset \hat{E}_2 = H \Pi \left(\frac{k}{k_g} \right) e^{ikH/2} * \hat{D} \hat{\varphi}_2' = \\ &= D \sum_{n'} (-1)^{n'} \hat{G}'(k - n'k_g) \hat{\rho}_2(k - n'k_g), \end{aligned} \quad (7.170)$$

$$F_2 = \frac{q}{H} W * E_2^\dagger \supset \hat{F}_2 = \frac{q}{H} \hat{W} \hat{E}_2. \quad (7.171)$$

Объединяя (7.168) — (7.171), после несложных преобразований получаем

$$\hat{F}_2(k) = q^2 \frac{\hat{W}}{H} \hat{D} \hat{G} \sum_{n=-\infty}^{\infty} (-1)^n \frac{\hat{W}(k_n)}{H} \hat{A}(k_n), \quad (7.172)$$

где по-прежнему $\hat{G}(k) = \sum_n G'(k_n)$. Следовательно,

$$F = \frac{(F_1 + F_2)}{2} \Rightarrow \hat{F} = q^2 \frac{\hat{W}}{H} \hat{D} \hat{G} \sum_{\substack{n=-\infty \\ n-\text{четное}}}^{\infty} \frac{\hat{W}(k_n)}{H} \hat{A}(k_n). \quad (7.173)$$

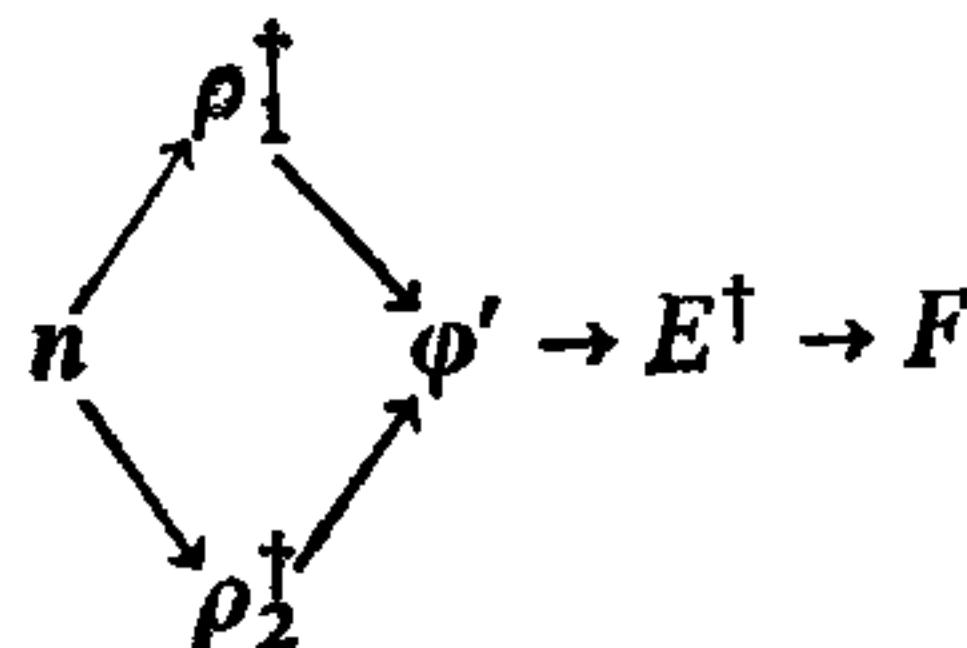
Аналогично на схемы с чередованием переносится анализ дисперсионных и других свойств. Везде, где встречается сумма по сеточным гармоникам распределения заряда, она заменяется суммой по четным n . Аналогичные результаты справедливы и для схем, сохраняющих энергию. В обоих вариантах расчета искомой силы данный метод обобщается на i чередующихся сеток с уменьшением n до кратных i .

Двухпроходная схема с чередованием почти эквивалентна обычной однопроходной схеме вычисления силы с половинным сеточным шагом $H/2$. Их отличие состоит в том, что \hat{G} и \hat{D} периодичны с длиной периода $k_g = 2\pi/H$, тогда как в обычных схемах с сеточным шагом $H/2$ они имеют период $4\pi/H$. Тем не менее, ограничивая возможные варианты \hat{G} , \hat{D} и \hat{W} в обычной схеме, можно вновь получить нужный результат, т. е. (7.173).

Чисто с точки зрения количества операций, схемы усреднения силы с чередованием проигрывают с самого начала. Они удваивают время цикла только ради того, чтобы сократить наполовину один шаг операций «баттерфляй» при решении уравнения Пуассона с помощью быстрого преобразования Фурье. Однако в двумерных и трехмерных схемах, когда ограниченность быстрой памяти с произвольным доступом становится важным фактором, метод чередующихся сеток заслуживает уже большего внимания.

7.8.2. Усреднение гармоник

В вычислительном отношении можно реализовать метод чередования с меньшими затратами, если объединить гармоники, относящиеся к чередующимся выборкам плотности заряда, а уже потом вычислять потенциалы и т. д. [117]. Символически такой алгоритм изображается в виде



Данный алгоритм дороже, чем алгоритм Чена и др. [75], с точки зрения памяти, поскольку он требует двух сеток, а не одной, для хранения значений плотности заряда. Его можно обобщить на i чередующихся выборок плотности заряда, причем схема с i чередованиями по-прежнему требует только двух сеток для плотности заряда.

Процедура усреднения гармоник в варианте с двумя чередующимися сетками включает в себя следующие шаги:

1. Распределение заряда на сетку 1.
2. Сдвиг сетки (или, что равносильно, координат частиц) на $H/2$ и распределение заряда на сетку 2.
3. Преобразование Фурье плотности заряда как на сетке 1, так и на сетке 2.
4. Объединение полученных гармоник

$$\hat{\rho}(k) = (\hat{\rho}_1(k) + e^{+ikH/2} \hat{\rho}_2(k))/2. \quad (7.174)$$

5. Нахождение потенциала, поля и силы на сетке 1 с помощью обычного метода частица—сетка.

На практике шаги 1 и 2 объединяются в одном проходе по частицам с тем, чтобы использовать простое соотношение, связывающее доли заряда (т. е. значения функции распределения заряда), приписываемые сеткам 1 и 2 каждой частицей. Соответственно фигурирующий на шаге 4 множитель 2 упрятывается в запоминаемые значения функции влияния.

Фурье-анализ схемы усреднения гармоник с чередованием повторяет анализ схемы усреднения силы, описанный выше. Однако в этом случае полезнее считать, что сдвигаются координаты частиц, а не сетка, чтобы выделить в явном виде множители «вертушки», необходимые в практической реализации алгоритма:

1. Распределение заряда на сетку 1.

$$\rho'_1 = \frac{q}{H} W * n \Rightarrow \hat{\rho}'_1 = \frac{q}{H} \hat{W} \hat{n}, \quad (7.175)$$

$$\rho_1^\dagger = \Pi \left(\frac{x}{H} \right) \rho'_1 \Rightarrow \hat{\rho}_1 = H \Pi \left(\frac{k}{k_g} \right) \hat{\rho}'_1 = \frac{q}{H} \sum_n \hat{W}(k_n) \hat{n}(k_n). \quad (7.176)$$

2. Сдвиг координат в положительном направлении оси x на $H/2$ и распределение заряда на сетку 2.

$$\rho'_2 = \frac{q}{H} W(x) * n(x - H/2) \Rightarrow \hat{\rho}'_2 = \frac{q}{H} \hat{W} \hat{n} e^{-ikH/2}, \quad (7.177)$$

$$\rho_2^\dagger = \Pi \left(\frac{x}{H} \right) \rho'_2 \Rightarrow \hat{\rho}_2 = e^{-ikH/2} \frac{q}{H} \sum_n (-1)^n \hat{W}(k_n) \hat{n}(k_n). \quad (7.178)$$

3. Усреднение гармоник.

$$\hat{\rho} = \frac{1}{2} (\hat{\rho}_1 + e^{+ikH/2} \hat{\rho}_2) = \frac{q}{H} \sum_{\substack{n \\ n \text{ четные}}} \hat{W} \hat{\rho}_n. \quad (7.179)$$

4. Нахождение потенциала, поля и силы.

$$\varphi' = G' * \rho \supset \hat{\varphi}' = \hat{G}' \hat{\rho}', \quad (7.180)$$

$$E^+ = \Pi \left(\frac{x}{H} \right) D * \varphi' \supset \hat{E} = H \Pi \left(\frac{k}{k_g} \right) * \hat{D} \hat{\varphi}' = \hat{D} \hat{\varphi}, \quad (7.181)$$

$$F = \frac{q}{H} W * E^+ \supset \hat{F} = \frac{q}{H} \hat{W} \hat{E}. \quad (7.182)$$

Объединяя (7.179) — (7.182), получаем тот же результат (7.173).

Стоимость схемы усреднения гармоник с чередованием только в худшем случае выше, чем у обычной схемы с удвоенным количеством сеточных узлов в каждом пространственном измерении. Например, использование для распределения заряда в процедуре с чередованием схемы CIC влечет за собой две дополнительные арифметические операции на частицу, что надо сравнить с 25 %-ным уменьшением времени на нахождение потенциала. Если отбрасываемые гармоники ($\pi/H \leq |k| \leq 2\pi/H$), неважны, самой рентабельной будет обычная схема с большим числом частиц в ячейке. Однако, как и для схемы усреднения силы, этот вывод требует повторной проверки, когда в двумерных и трехмерных задачах приходится сталкиваться с ограничениями на машинную память.

7.8.3. Многомерные схемы

В двумерном и трехмерном аналогах одномерных схем с двумя чередующимися сетками силы или гармоники усредняются, соответственно, по четырем или восьми чередующимся сеткам. Повторяя фурье-анализ поля сил, получаем векторный аналог выражения (7.173):

$$\mathbf{F} \supset \hat{\mathbf{F}} = q^2 \frac{\hat{W}}{V_c} \hat{D} \hat{G} \sum_{\substack{n \\ n \text{ четные}}} \frac{\hat{W}(k_n)}{V_c} \hat{\rho}(k_n), \quad (7.183)$$

где условие четности \mathbf{n} в сумме по сеточным гармоникам означает, что компоненты n_1 , n_2 и (в трехмерном случае) n_3 принимают целые четные значения.

В сумму по сеточным гармоникам основной вклад вносят гармоники, ближайšie к основной зоне, т. е. гармоники с $\mathbf{n} = (10)$ или (01) , или же, в трехмерном случае, с $\mathbf{n} = (100)$, (010) , (001) . Для подавления этих доминирующих членов можно построить промежуточные схемы с чередованием, которые ис-

пользуют только две чередующиеся сетки (сравните с промежуточным распределением заряда, которое бы осуществлялось более чем в одном измерении). В таких многомерных схемах с двумя чередующимися сетками удобнее всего располагать сетки относительно друг друга таким образом, чтобы узлы второй сетки находились в углах ячеек первой сетки. Анализ этих схем дает для поля силы выражение, совпадающее с (7.183), но с тем ограничением, что сумма компонент \mathbf{n} ($n_1 + n_2$ или $n_1 + n_2 + n_3$) должна быть четной. Если представить себе k -пространство в виде шахматной доски, причем основную зону связать с белым полем, то указанная схема с двумя чередующимися сетками удаляет все сеточные гармоники, расположенные в центре черных полей.

Вдоль направления главных осей ($x = \text{const}$ и т. д. в x -пространстве или $k_1 = \text{const}$ и т. д. в k -пространстве) сеточные гармоники дают такой же вклад в силу, дисперсионное соотношение и другие характеристики, какой бы получился при использовании обычной схемы с шагом сетки $H/2$, а по сравнению с обычной схемой с шагом H этот вклад меньше везде. Следовательно, если (как и должно быть для реалистичного моделирования бесстолкновительных систем) волны с $|k| > \pi/H$ несущественны, то оба алгоритма с чередующимися сетками при шаге H дают модель, по качеству сравнимую с моделью, в которой используется одна сетка с шагом $H/2$. Если применяется схема с усреднением гармоник, то в трехмерном случае нахождение потенциала при двух чередующихся сетках требуется приблизительно в четыре раза меньше времени ЦП и памяти, чем в обычной схеме с более мелкой сеткой, и вследствие этого такая схема может стать конкурентоспособной с точки зрения затрат машинного времени и памяти.

Идеи чередования находят еще одно применение в задачах об изолированных гравитирующих системах (см. гл. 11) для удаления периодических изображений в x -пространстве при нахождении потенциала. Дополнительные подробности по этому поводу читатель может найти в работе [118].

Алгоритмы частица-частица — частица-сетка (P^3M)

8.1. ВВЕДЕНИЕ

Алгоритмы частица-частица — частица-сетка (P^3M)¹⁾ представляют собой класс гибридных алгоритмов, разработанный авторами [189, 119, 115, 117, 120]. Эти алгоритмы позволяют моделировать коррелированные системы с далекодействующими (например, кулоновскими) силами для гораздо больших ансамблей частиц, чем было возможно ранее. Сущность метода состоит в том, что сила взаимодействия частиц представляется в виде суммы двух частей — короткодействующей силы f^{sr} , которая отлична от нуля только при расстояниях между частицами, меньших некоторого радиуса усечения r_e , и гладко меняющейся силы R , спектр которой приближенно имеет конечную ширину (т. е. заметно отличен от нуля только в ограниченной области значений k). Полная короткодействующая сила F^{sr} , действующая на каждую частицу, вычисляется прямым суммированием сил парного взаимодействия данной частицы с другими (метод частица — частица (PP)), а гладко меняющаяся часть силы приближенно вычисляется методом частица-сетка (PM). Хотя в этой главе мы ограничиваемся рассмотрением далекодействующих сил, которые являются чисто кулоновскими, получаемые результаты можно обобщить на произвольные зависимости силы от расстояния заменой заряда другой подходящей мерой интенсивности силовых центров, а кулоновской эталонной силы R — соответствующим аналогом.

В P^3M -алгоритмах используются две сетки: одна сетка для заряда и потенциала и одна более грубая сетка, называемая цепочечной сеткой. Сетка заряда-потенциала используется на различных стадиях PM-расчета для запоминания поочередно значений плотности заряда, гармоник заряда, гармоник потенциала и значений потенциала. Цепочечная сетка представляет собой регулярный массив ячеек, линейные размеры которых превышают или равны радиусу усечения r_e короткодействующей силы. С каждой ячейкой этой сетки связан один элемент в массиве заголовков цепочек, этот массив адресов используется вместе с дополнительной координатой частицы, а именно координатой

¹⁾ От Particle-Particle — Particle-Mesch (англ). — Прим. перев.

связного списка, для определения пар соседних частиц при вычислении короткодействующей компоненты. Практически массив заголовков цепочек и координаты связного списка располагаются в том же месте памяти, что и сетка заряда-потенциала (см. разд. 8.7), и поэтому никакой дополнительной памяти для них не требуется.

Для интегрирования орбит частиц по времени используется схема с перешагиванием

$$\mathbf{x}_i^{n+1} = \mathbf{x}_i^n + \frac{\mathbf{p}_i^{n+1/2}}{m_i} DT, \quad (8.1)$$

$$\mathbf{p}_i^{n+1/2} = \mathbf{p}_i^{n-1/2} + (\mathbf{F}_i + \mathbf{F}_i^{sr}) DT. \quad (8.2)$$

Положения частиц $\{\mathbf{x}_i\}$ определяются на целых временных шагах, а импульсы $\{\mathbf{p}_i\}$ — на полуцелых. В формуле (8.2) \mathbf{F}_i представляет собой вычисляемую по сеточным значениям аппроксимацию полной гладко меняющейся части силы, действующей на i -ю частицу, а \mathbf{F}_i^{sr} является полной короткодействующей силой. Использование в Р³М-алгоритмах импульсов вместо скоростей обусловлено соображениями экономии числа операций.

На рис. 8.1 в виде блок-схемы показан цикл по времени Р³М-алгоритма. Набор координат частиц в начале цикла равен $\{\mathbf{x}^n, \mathbf{p}^{n-1/2}\}$, а в конце цикла — $\{\mathbf{x}^{n+1}, \mathbf{p}^{n+1/2}\}$. РМ-часть расчета силы является вариантом РМ-алгоритма расчета силы, применяемого для бесстолкновительных систем, который обладает повышенной точностью. Метод достижения этой точности обсуждается далее в разд. 8.3 и 8.6. РР-часть (см. разд. 8.4) полностью отделена от вычисления РМ-силы, и ее можно было бы использовать самостоятельно для расчета сил межчастичного взаимодействия в отсутствие далекодействующих сил.

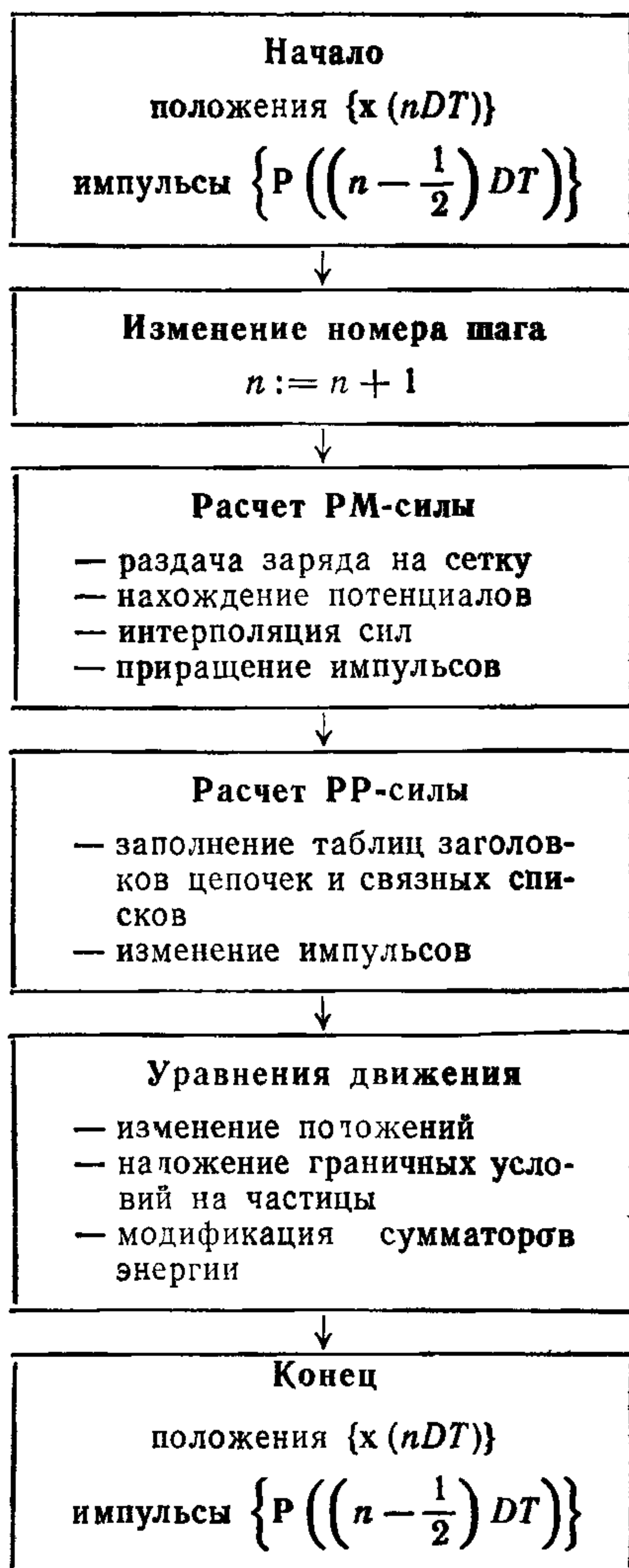


Рис. 8.1 Цикл по временному шагу в Р³М алгоритме

Целью столь сложной процедуры расчета силы является достижение высокой скорости счета. Как показано в разд. 8.5, в P^3M -алгоритмах время цикла почти линейно зависит от числа частиц, тогда как в обычных схемах с далекодействующими силами время цикла пропорционально квадрату числа частиц. Рассматриваемые алгоритмы таковы, что накладные расходы на вычисление на каждом шаге по времени термодинамических характеристик, функции радиального распределения для малых r и преобразования парной корреляционной функции минимальны. Кроме того, при использовании P^3M -алгоритма снимается проблема формирования периодических изображений, которая присутствует в методах прямого суммирования и в вычислительном отношении обходится дорого. Если для нахождения потенциала используется быстрое преобразование Фурье, то периодичность вносится в сеточную часть силы автоматически. Периодические изображения в короткодействующей части силы нужны только для тех частиц, которые попадают в слой толщиной r_e у границ расчетной области.

Несмотря на свое очевидное преимущество в быстродействии, алгоритм P^3M может показаться непривлекательным из-за трудностей расчета силы. На самом деле, как будет показано в разд. 8.3 и 8.6, эти трудности кажущиеся. Реализация метода Q -минимизации (см. разд. 7.7) в сочетании с соотношениями для времени счета программы дает возможность по заданной точности силы и числу частиц автоматически находить оптимальные комбинации размера сетки, расщепления силы, радиуса усечения, схемы распределения заряда, функции Грина и оператора разностного дифференцирования. Все что остается сделать потенциальному пользователю — это получить P^3M -программу [120], выбрать нужную зависимость силы от расстояния и решить, какой компромисс между точностью силы, числом частиц и временем цикла его устраивает для той конфигурации, которую он хочет исследовать.

8.2. РАСЩЕПЛЕНИЕ СИЛЫ

Изменение импульса i -й частицы на каждом шаге по времени определяется полной силой, действующей на эту частицу. Поэтому мы вправе решать, как лучше поделить полную силу между короткодействующей частью и гладко меняющейся, или, иными словами, как разбить силу межчастичного взаимодействия на f^{sr} и R .

Под эталонной силой R мы будем понимать силу межчастичного взаимодействия, которую мы желаем воспроизвести вычислением с сеткой. Из соображений экономии нам нужно, чтобы радиус усечения r_e был как можно меньше и тем самым R равнялась полной силе межчастичного взаимодействия вплоть до

расстояний, возможно более малых. К сожалению, мы не можем сделать эталонную и полную силы межчастичного взаимодействия одинаковыми вплоть до сколь угодно малых расстояний, ибо сетка, применяемая в РМ-части расчета, не в состоянии сколько-нибудь правильно передать спектр $\hat{\mathbf{R}} \supset \mathbf{R}$ за пределами основной зоны (т. е. для волновых чисел \mathbf{k} , компоненты которых k_i таковы, что $|k_i| > \pi/H_i$). Например, если бы нам захотелось представить силу кулоновского взаимодействия для точечных частиц, то не имело бы никакого смысла задавать $\mathbf{R} \sim \mathbf{x}/|\mathbf{x}|^3$, поскольку это дало бы $|\hat{\mathbf{R}}| \sim k^{-1}$.

Спектр эталонной силы сужается путем сглаживания. Применительно к кулоновской дальнедействующей силе подходящей формой эталонной силы является зависимость, повторяющая кулоновский закон силы для точечных частиц на расстояниях, больших радиуса усечения r_e , и гладко спадающая до нуля в пределах этого радиуса (см., например, рис. 12.7). Чем глаже спадает $\mathbf{R}(\mathbf{x})$ при $\mathbf{x} \rightarrow 0$ и чем больше становится r_e , тем быстрее затухают гармоники $\hat{\mathbf{R}}(\mathbf{k})$ с ростом \mathbf{k} . Практически, делая $\mathbf{R}(\mathbf{x})$ непрерывной вместе с одной или более производными и используя радиус усечения, больший двух шагов сетки потенциала, можно сделать погрешности, обусловленные конечной шириной спектра эталонной силы, пренебрежимо малыми по сравнению с остальными погрешностями.

Сглаживание кулоновской эталонной силы межчастичного взаимодействия эквивалентно приписыванию заряженной частице конечного размера (см разд. 1.5.3). Следовательно, прямой метод включения сглаживания состоит в том, чтобы связать с эталонной силой межчастичного взаимодействия некоторый простой профиль плотности $S(\mathbf{x})$. Примерами распределений, которые, как было практически установлено, описывают с соизмеримой точностью полную силу, являются однородно заряженная сфера (форма S1), сфера с линейно спадающей плотностью (форма S2)

$$S2: S(r) = \begin{cases} \frac{48}{\pi a^4} \left(\frac{a}{2} - r \right), & r < a/2, \\ 0 & , \quad r \geq a/2, \end{cases} \quad (8.3)$$

и плотность с гауссовым распределением (форма S3). Впредь мы ограничимся рассмотрением формы S2, поскольку в трехмерных схемах при заданном эффективном радиусе усечения r_e она показала чуть-чуть лучшую точность. Гармоники эталонной силы (для единичных зарядов) связаны с формой профиля $S \supset \hat{S}$ соотношением

$$\hat{\mathbf{R}}(\mathbf{k}) = - \frac{i\mathbf{k}\hat{S}^2}{k^2}. \quad (8.4)$$

В формуле (8.3) радиус усечения короткодействующей силы равняется a , а не r_e . На практике r_e можно брать значительно меньше a , поскольку за счет непрерывности производных при $r = a$ эталонная сила близка к силе взаимодействия точечных частиц на расстояниях, несколько меньших a . Опытным путем было установлено, что в качестве оценки r_e снизу хорошо брать радиус корреляции форм-фактора заряда, т. е.

$$r_e^3 \geq \frac{\int d\mathbf{x} \int d\mathbf{x}' S(\mathbf{x}) S(\mathbf{x} + \mathbf{x}')}{\int S^2(\mathbf{x}) d\mathbf{x}}, \quad (8.5)$$

что для формы S_2 дает

$$r_e \geq \left(\frac{5\pi}{48}\right)^{1/3} a \simeq 0.7a. \quad (8.6)$$

Если эталонная сила межчастичного взаимодействия R для РМ-части расчета уже выбрана, короткодействующую часть f^{sr} можно найти путем вычитания R из полной силы взаимодействия частиц f^{tot} :

$$f^{sr} = f^{tot} - R. \quad (8.7)$$

Типичное расщепление силы межчастичного взаимодействия показано на рис. 12.7 для модели ионной жидкости: сеточное приближение R является чисто кулоновской силой для зарядов формы S_2 . Короткодействующая сила f^{sr} равняется разности между силами кулоновского взаимодействия для зарядов формы S_2 и для точечных зарядов плюс быстро меняющаяся короткодействующая сила, которая воспроизводит отталкивание ионизированных атомов. В случае, изображенном на рис. 12.7, сила f^{sr} полагается равной нулю на расстояниях, больших $r_e = 0.735a$.

8.3. СЕТОЧНАЯ СИЛА

Полная «сеточная» часть силы, действующей на i -ю частицу,

$$\mathbf{F}_i \left(\simeq \sum_{j=1}^{N_p} \mathbf{R}(\mathbf{x}_i - \mathbf{x}_j) \right)$$

вычисляется РМ-методом, т. е.:

1. Раздаем заряд на сетку заряда-потенциала.
2. Находим потенциал.
3. Дифференцируем потенциалы и находим поля в узлах сетки.
4. Находим с помощью интерполяции силы \mathbf{F}_i , действующие на частицы, $i \in [1, N_p]$.

Результаты, получаемые на каждом из этих этапов, полезно рассматривать как всего лишь артефакты, необходимые для дости-

жения наилучшего приближения F_i к $\sum_j R(x_i - x_j)$. Именно точность сил определяет точность моделирования динамики. Погрешности, вносимые на промежуточных шагах, почти не имеют значения, если они компенсируют друг друга. Оптимизация вычисления РМ-силы направлена на то, чтобы умышленным введением погрешностей потенциалов компенсировать неизбежные погрешности, вносимые при распределении заряда, интерполяции силы и дифференцировании потенциала (см. аналогичные результаты для бесстолкновительных систем в разд. 7.7).

8.3.1. Распределение заряда

В гл. 5 мы показывали, какими свойствами должна обладать функция распределения заряда. А именно нужно, чтобы распределение заряда захватывало небольшое число узлов сетки, чтобы вдали от исходного заряда поле приближалось к точному полю (условие на больших расстояниях) и чтобы зависимость сеточного заряда от положения частицы была гладкой (условие гладкости). Выбор схемы распределения заряда, обеспечивающей необходимую в Р³М-алгоритме точность РМ-части расчета, есть следствие компромисса между числом узлов сетки и гладкостью. Количество сеточных узлов определяет вычислительные затраты. Гладкость определяет спектр гармоник сеточной силы за пределами основной зоны. Если энергия этой части спектра велика, то эффект наложения (разд. 5.6) будет вызывать флуктуации вычисляемых сил, давать ошибочные корреляции и приводить к плохому сохранению энергии.

Влияние гладкости легче всего понять с помощью спектрального описания распределения заряда. Если функция распределения W гладкая, то фурье-образ $\hat{\rho}$ ее свертки ρ' с плотностью n быстро убывает с ростом k . Фурье-образ функции W , непрерывной вместе со своими n производными, убывает как $k^{-(n+1)}$, поэтому $\hat{\rho}' \sim W\hat{n}$ убывает намного быстрее, чем \hat{n} . Дискретизация ρ' приводит к периодичности в k -пространстве, причем перекрывание периодических копий с основной зоной вызывает появление ложных спектральных компонент.

Идеальная функция распределения заряда — это фильтр низких частот. Она имеет фурье-образ W , который отфильтровывает все гармоники функции n с $|k_i| > \pi/N$. На практике такая функция не годится, поскольку тогда пришлось бы распределять заряд каждой частицы по всем узлам сетки в расчетной области. Последовательность все более хороших (и все более дорогих в смысле числа арифметических операций, приходящихся на частицу за шаг по времени) приближений к фильтрам низких частот образует иерархию функций распределения заряда, удовлетворяющих условиям на больших расстояниях и условиям

гладкости все более высокого порядка, т. е. NGP (0-й порядок), CIC (1-й порядок), TSC (2-й порядок) и т. д. В трехмерном случае

$$W(\mathbf{x}) = \frac{1}{V_c} w\left(\frac{x}{H_1}\right) w\left(\frac{y}{H_2}\right) w\left(\frac{z}{H_3}\right), \quad (8.8)$$

где

$$w(x) \supset \hat{w}(k) = \left(\frac{\sin k/2}{k/2}\right)^{p+1}. \quad (8.9)$$

Эти полиномиальные функции распределения не только понижают амплитуду сеточных гармоник относительно основных гармоник, но меняют также амплитуду самих гармоник, лежащих в основной зоне. Одна из компенсирующих погрешностей, вводимая в потенциал, и призвана погасить изменение амплитуды основных гармоник. Более подробное обсуждение иерархии схем распределения заряда, спектральной трактовки, дискретизации и эффекта наложения было проведено выше в главах 5 и 7.

8.3.2. Вычисление силы

Для того чтобы получить силу с наивысшей возможной при заданном порядке функции распределения заряда точностью, в P³M-алгоритмах используется вариант PM-силы, сохраняющий импульс. Поэтому на основании условия сохранения импульса (разд. 5.3.3) интерполяционная функция для силы задается равной функции распределения заряда.

Если нигде не используется чередование, то этапы вычисления силы можно формально представить в следующем виде:

Распределение заряда:

$$\rho^+(\mathbf{x}) = \frac{q}{V_c} \Pi(\mathbf{x}; \mathbf{H}) W(\mathbf{x}) * n(\mathbf{x}). \quad (8.10)$$

Нахождение потенциала:

$$\phi'(\mathbf{x}) = G'(\mathbf{x}) * \rho^+(\mathbf{x}). \quad (8.11)$$

Дифференцирование потенциала:

$$\mathbf{E}^+(\mathbf{x}) = -\Pi(\mathbf{x}; \mathbf{H}) \mathbf{D}(\mathbf{x}) * \phi'(\mathbf{x}). \quad (8.12)$$

Интерполяция силы:

$$\mathbf{F}(\mathbf{x}) = \frac{q}{V_c} W(\mathbf{x}) * \mathbf{E}^+(\mathbf{x}). \quad (8.13)$$

Используя результаты разд. 5.6.4 и 7.4, можно применить к соотношениям (8.10) — (8.13) преобразование Фурье и после ряда выкладок получить силу (вычисляемую через сеточную силу), действующую на единичный заряд в точке \mathbf{x}_2 со стороны отри-

цательного единичного заряда в точке \mathbf{x}_1 :

$$\mathbf{F}(\mathbf{x}_2; \mathbf{x}_1) = \frac{1}{V_b} \sum_I \frac{\hat{W}}{V_c} \hat{\mathbf{D}} \hat{G} \sum_n \frac{\hat{W}(k_n)}{V_c} \exp(-ik_n \cdot \mathbf{x}_1) \cdot \exp(ik \cdot \mathbf{x}_2). \quad (8.14)$$

При выводе (8.14) предполагается периодичность системы, поэтому интеграл по волновым числам переходит в сумму $V_b^{-1} \sum_I$, где I — тройка целых чисел (в трехмерном случае), отвечающих номерам гармоник (см. разд. 7.4).

Метод моделирования с помощью Р³М-алгоритма таков, что в оптимальной рабочей точке на одну ячейку приходится меньше одной частицы. В этих условиях алгоритм с усреднением гармоник на чередующихся сетках (разд. 7.8.2) становится вполне конкурентоспособен обычному РМ-алгоритму (для бесстолкновительных вычислительных моделей, где при $N \sim \lambda_D$ требуется довольно большое число частиц в ячейке, обычно справедливо обратное). Например, если дискретизация распределения заряда производится на двух сетках, причем вторая сдвинута на $\mathbf{H}/2$ относительно первой, то сумма по сеточным гармоникам, т. е. по \mathbf{n} , в формуле (8.14) сводится к сумме по таким значениям $\mathbf{n} = (n_1, n_2, n_3)$, для которых $(n_1 + n_2 + n_3)$ четное.

8.3.3. Погрешности в силе

Мерой качества вычисляемой с помощью сетки аппроксимации $\mathbf{F}(\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1; \mathbf{x}_1)$ эталонной силы межчастичного взаимодействия $\mathbf{R}(\mathbf{x})$ служит усредненный по смещениям относительно узлов сетки полный квадрат отклонения \mathbf{F} от \mathbf{R} :

$$Q = \frac{1}{V_c} \int_{V_c} d\mathbf{x}_1 \int_{V_b} d\mathbf{x} |\mathbf{F}(\mathbf{x}; \mathbf{x}_1) - \mathbf{R}(\mathbf{x})|^2. \quad (8.15)$$

Для конкретной функции распределения заряда и интерполяции силы W и конкретного разностного оператора \mathbf{D} можно в силу соотношений (8.10) — (8.13) рассматривать Q как квадратичный функционал от набора дискретных значений функции Грина $\{G_p = G(\mathbf{x}_p); p_i \in [0, N_i/2]\}$. Величины компонент p_i вектора $\mathbf{p} = (p_1, p_2, p_3)$ ограничены половинным числом узлов сетки $N_i/2$ в направлении оси i периодической расчетной области, поскольку G должна быть четной функцией. В принципе можно применить к Q вариационные методы и получить дискретные системы уравнений для оптимального набора величин

$\{G_q^{\text{opt}}; q_i \in [0, N_i/2]\}$, т. е.

$$\frac{\partial Q}{\partial G_q} = \frac{1}{V_c} \int d\mathbf{x}_1 \int_{V_b} d\mathbf{x}_2 (\mathbf{F} - \mathbf{R}) \cdot \frac{\partial \mathbf{F}}{\partial G_q} = 0. \quad (8.16)$$

Уравнение (8.16) дает алгебраическое уравнение для G_q и, следовательно, определяет оптимальную (минимизирующую Q) форму алгебраических уравнений, связывающих сеточные значения потенциала со значениями плотности заряда (сравните с уравнением (5.153) для схем, сохраняющих энергию).

Оптимизацию можно распространить на ограниченный набор разностных операторов градиента, для чего \mathbf{D} представляется в виде линейной комбинации операторов градиента \mathbf{D}_m , i -е компоненты которых представляют собой двухточечные центральные разностные производные, взятые по узлам, отстоящим на $2mH_i$, $m = 1, 2$; и т. д., а затем функционал (8.15) минимизируется по коэффициентам этой линейной комбинации. Например, в случае четырехточечной разностной производной (формула (5.137))

$$\mathbf{D} = \alpha \mathbf{D}_1 + (1 - \alpha) \mathbf{D}_2, \quad (8.17)$$

и Q_{opt} получается минимизацией Q по α и $\{G_p\}$.

Практически интегралы по \mathbf{x} в формуле (8.16) вычислить трудно. Один из возможных выходов из этого положения заключается в том, что функцию Грина записывают в виде суммы функции Грина, отвечающей конечно-разностной аппроксимации уравнения Пуассона, и корректирующих членов с неопределенными коэффициентами, после чего численно минимизируют Q относительно этих неизвестных коэффициентов [119]. Более мощный метод состоит в том, что выражение (8.15) представляется через фурье-компоненты и решается задача для значений функции влияния $\{G(\mathbf{k})\}$ [115, 117].

Применение теоремы Парсеваля к выражению (8.15) дает

$$Q = \frac{1}{V_c} \int_{V_c} d\mathbf{x}_1 \frac{1}{V_b} \sum_{\mathbf{l}} \left[|\hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1)|^2 - 2\hat{\mathbf{R}}^*(\mathbf{k}) \cdot \hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1) + |\hat{\mathbf{R}}(\mathbf{k})|^2 \right]. \quad (8.18)$$

Из (8.14) имеем

$$\hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1) = \frac{\hat{W}}{V_c} \hat{\mathbf{D}} \hat{G} \sum_{\mathbf{n}} \frac{\hat{W}(\mathbf{k}_n)}{V_c} \exp(i(\mathbf{k} - \mathbf{k}_n) \cdot \mathbf{x}_1). \quad (8.19)$$

Следовательно,

$$Q = \frac{1}{V_b} \sum_{\mathbf{l}} \left\{ \hat{U}^2 |\hat{\mathbf{D}}|^2 \hat{G}^2 \sum_{\mathbf{n}} \hat{U}^2(\mathbf{k}_n) - 2\hat{\mathbf{R}}^* \cdot \hat{\mathbf{D}} \hat{G} \hat{U}^2 + |\hat{\mathbf{R}}|^2 \right\}, \quad (8.20)$$

где $O = \hat{W}/V_c$. Можно воспользоваться периодичностью $\hat{\mathbf{D}}$ и \hat{G} , т. е. $\hat{\mathbf{D}}(\mathbf{k}_n) = \hat{\mathbf{D}}(\mathbf{k})$, $\hat{G}(\mathbf{k}_n) = \hat{G}(\mathbf{k})$, и переписать (8.20) через гармоники одного периода в k -пространстве:

$$Q = \frac{1}{V_b} \sum_{\mathbf{l}=0}^{N-1} \left\{ |\hat{\mathbf{D}}|^2 \hat{G} \left[\sum_{\mathbf{n}} \hat{U}^2 \right]^2 - 2\hat{G} \hat{\mathbf{D}} \cdot \sum_{\mathbf{n}} \hat{\mathbf{R}}^* \hat{U}^2 + \sum_{\mathbf{n}} |\hat{\mathbf{R}}|^2 \right\}. \quad (8.21)$$

Минимизация Q относительно набора значений функции влияния $\{\hat{G}(\mathbf{k})\}$ дает оптимальную функцию влияния

$$\hat{G}(\mathbf{k}) = \frac{\hat{\mathbf{D}}(\mathbf{k}) \cdot \sum_n \hat{U}^2(\mathbf{k}_n) \hat{\mathbf{R}}(\mathbf{k}_n)}{|\hat{\mathbf{D}}(\mathbf{k})|^2 \left[\sum_n \hat{U}^2(\mathbf{k}_n) \right]^2} \quad (8.22)$$

с минимальным $Q = Q_{\text{opt}}$:

$$Q_{\text{opt}} = \frac{1}{V_b} \sum_{l=0}^{N-1} \left\{ \sum_n |\hat{\mathbf{R}}|^2 - \frac{|\hat{\mathbf{D}} \cdot \sum_n \hat{U}^2 \hat{\mathbf{R}}^*|^2}{|\hat{\mathbf{D}}|^2 \left[\sum_n \hat{U}^2 \right]^2} \right\}. \quad (8.23)$$

Оптимизация распространяется на ограниченное семейство операторов \mathbf{D} путем представления $\hat{\mathbf{D}}$ в виде линейной комбинации фурье-образов двухточечных операторов (см. (8.17)) и минимизации (8.23) относительно коэффициентов этой линейной комбинации.

Возникновение остаточных погрешностей, составляющих $Q_{\text{opt}} = \min_{\{\alpha, \hat{G}\}} [Q]$, связано с тем, что сеточные гармоники плотности заряда, обусловленные дискретизацией, подавлены не полностью, разностный градиент не всегда параллелен истинному градиенту и эталонная сила \mathbf{R} и вычисляемая с помощью сетки сила \mathbf{F} не являются функциями с конечной шириной спектра. Задача определения относительной значимости различных факторов, дающих вклад в полную погрешность сеточной силы, решается посредством вычисления трех вспомогательных величин E_{ref} , P и Z .

Величина Z представляет собой полное квадратичное отклонение от эталонной силы сеточной силы, усредненной по смещениям, и таким образом дает меру отклонения нефлуктуирующей части \mathbf{F} от \mathbf{R} . Исходя из (8.19), имеем

$$\frac{1}{V_c} \int_{V_c} d\mathbf{x}_1 \hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1) = \langle \hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1) \rangle = \hat{U} \hat{\mathbf{D}} \hat{G} \hat{U}.$$

Отсюда

$$\begin{aligned} Z &= \int d\mathbf{x} |\langle \mathbf{F}(\mathbf{x}; \mathbf{x}_1) \rangle - \mathbf{R}(\mathbf{x})|^2 \\ &= \frac{1}{V_b} \sum_{\mathbf{l}} |\langle \hat{\mathbf{F}}(\mathbf{k}; \mathbf{x}_1) \rangle - \hat{\mathbf{R}}(\mathbf{k})|^2 \\ &= \frac{1}{V_b} \sum_{\mathbf{l}} \{ |\hat{\mathbf{D}}|^2 \hat{G}^2 \hat{U}^4 - 2 \hat{G} \hat{\mathbf{D}} \cdot \hat{\mathbf{R}}^* \hat{U}^2 + |\hat{\mathbf{R}}|^2 \} \\ &= \frac{1}{V_b} \sum_{l=0}^{N-1} \left\{ |\hat{\mathbf{D}}|^2 \left[\sum_n \hat{U}^4 \right] - 2 \hat{G} \hat{\mathbf{D}} \cdot \left[\sum_n \hat{\mathbf{R}}^* \hat{U}^2 \right] + \sum_n |\hat{\mathbf{R}}|^2 \right\}. \end{aligned} \quad (8.24)$$

Величина P выражает полное квадратичное отклонение сеточной силы от ее среднего по смещениям:

$$\begin{aligned}
 P &= \frac{1}{V_c} \int_{V_c} dx_1 \int_{V_b} dx |F(x; x_1) - \langle F(x; x_1) \rangle|^2 \\
 &= \frac{1}{V_b} \sum_l |\hat{D}|^2 \hat{G}^2 \hat{U}^2 \left\{ \left[\sum_n \hat{U}^2 \right] - \hat{U}^2 \right\} \\
 &= \frac{1}{V_b} \sum_{l=0}^{N-1} |\hat{D}|^2 \hat{G}^2 \left\{ \left[\sum_n \hat{U}^2 \right]^2 - \sum_n \hat{U}^4 \right\}.
 \end{aligned} \tag{8.25}$$

P оценивает вклад в Q сеточных гармоник. Если бы спектр функции распределения заряда точно, а не лишь приближенно имел конечную ширину, то P равнялась бы нулю. Величина E_{ref} — это наименьшее остаточное отклонение сеточной силы от эталонной. Если бы функция распределения заряда обладала спектром конечной ширины и одновременно отсутствовали угловые погрешности в операторе градиента, то оптимальность \hat{G} привела бы к полной компенсации погрешностей в основной зоне ($l \in [-N/2, N/2 - 1]$, или, что эквивалентно, $l \in [0, N - 1]$) и к остаточной погрешности, равной

$$E_{\text{ref}} = \frac{1}{V_b} \sum_{l=0}^{N-1} \left\{ \left[\sum_n |\hat{R}|^2 \right] - |\hat{R}|^2 \right\}. \tag{8.26}$$

Из определения Q , P и Z следует, что

$$Q = P + Z. \tag{8.27}$$

Если преобладает P , это указывает на то, что основной причиной погрешностей является недостаточная частота дискретизации плотности заряда. В этом случае можно добиться заметных улучшений, прибегнув к чередованию или интерполяционной функции более высокого порядка. Если преобладает Z , значит, определяющим фактором выступают угловые погрешности в градиенте и слишком большая энергия части спектра эталонной силы, лежащей за пределами основной зоны. Сравнение величины Z с E_{ref} помогает уточнить, определяется Z в основном угловыми погрешностями или же шириной спектра. Если Z гораздо больше E_{ref} , то положение можно существенно улучшить, воспользовавшись оператором градиента более высокого порядка; в противном случае, чтобы добиться уменьшения Z и тем самым Q , нужно использовать другую эталонную силу или увеличить радиус усечения. Практическое применение этих величин обсуждается далее в разд. 8.6.

Выражения (8.20) — (8.25) относятся к обычному РМ-алгоритму. Если применяется чередование, выражение (8.21) прини-

мает вид

$$Q = \frac{1}{V_b} \sum_{l=0}^{N-1} \left\{ |\hat{\mathbf{D}}|^2 G^2 \left[\sum_{\mathbf{n}} \hat{U}^2 \right] \left[\sum'_{\mathbf{n}'} \hat{U}^2 \right] - 2\hat{G}\hat{\mathbf{D}} \cdot \sum_{\mathbf{n}} \hat{\mathbf{R}}^* \hat{U}^2 + \sum_{\mathbf{n}} |\hat{\mathbf{R}}|^2 \right\}, \quad (8.28)$$

где штрих у суммы по \mathbf{n}' указывает, что здесь \mathbf{n}' пробегают те значения, которые остаются после чередования. При двух чередующихся сетках этот штрих означает, что в сумме по \mathbf{n}' значения $n'_1 + n'_2 + n'_3$ четные. Все остальные величины и суммы в (8.28) такие же, как для обычных схем. Оптимальная функция влияния принимает вид

$$\hat{G}(\mathbf{k}) = \frac{\hat{\mathbf{D}} \cdot \sum_{\mathbf{n}} \hat{\mathbf{R}}^* \hat{U}^2}{|\hat{\mathbf{D}}|^2 \left[\sum_{\mathbf{n}} \hat{U}^2 \right] \left[\sum'_{\mathbf{n}'} \hat{U}^2 \right]}. \quad (8.29)$$

Выражение для Z (т. е. (8.24)) остается неизменным, а (8.25) переходит в

$$P = \frac{1}{V_b} \sum_{l=0}^{N-1} |\hat{\mathbf{D}}|^2 \hat{G}^2 \left\{ \left[\sum_{\mathbf{n}} \hat{U}^2 \right] \left[\sum'_{\mathbf{n}'} \hat{U}^2 \right] - \sum_{\mathbf{n}} \hat{U}^4 \right\}. \quad (8.30)$$

8.4. КОРОТКОДЕЙСТВУЮЩАЯ СИЛА

Полная короткодействующая часть силы, действующая на i -ю частицу, находящуюся в точке \mathbf{x}_i , равняется сумме короткодействующих сил межчастичного взаимодействия:

$$\mathbf{F}_i^{\text{sr}} = \sum_{j=1}^{N_p} \mathbf{f}_{ij}^{\text{sr}}. \quad (8.31)$$

Простейший метод вычисления \mathbf{F}_i^{sr} состоит в том, чтобы просмотреть все частицы $j = 1, \dots, N_p$, проверить, не меньше ли расстояние $r_{ij} = |\mathbf{x}_i - \mathbf{x}_j|$, чем r_e , и если это окажется так, подсчитать $\mathbf{f}_{ij}^{\text{sr}}$ и добавить ее к \mathbf{F}_i^{sr} . Совершенно очевидно, что такой подход непрактичен, поскольку для каждого из N_p значений i пришлось бы проверять $N_p - 1$ расстояний r_{ij} , что дает число операций порядка N_p^2 .

8.4.1. Цепочечная сетка

Вычислительные затраты, связанные с выявлением тех частиц j , которые дают вклад в короткодействующую силу, действующую на частицу i , значительно снижаются, если частицы упорядочены по координатам таким образом, что проверки координат частиц j на выполнение условия $r_{ij} \leq r_e$ нужно произ-

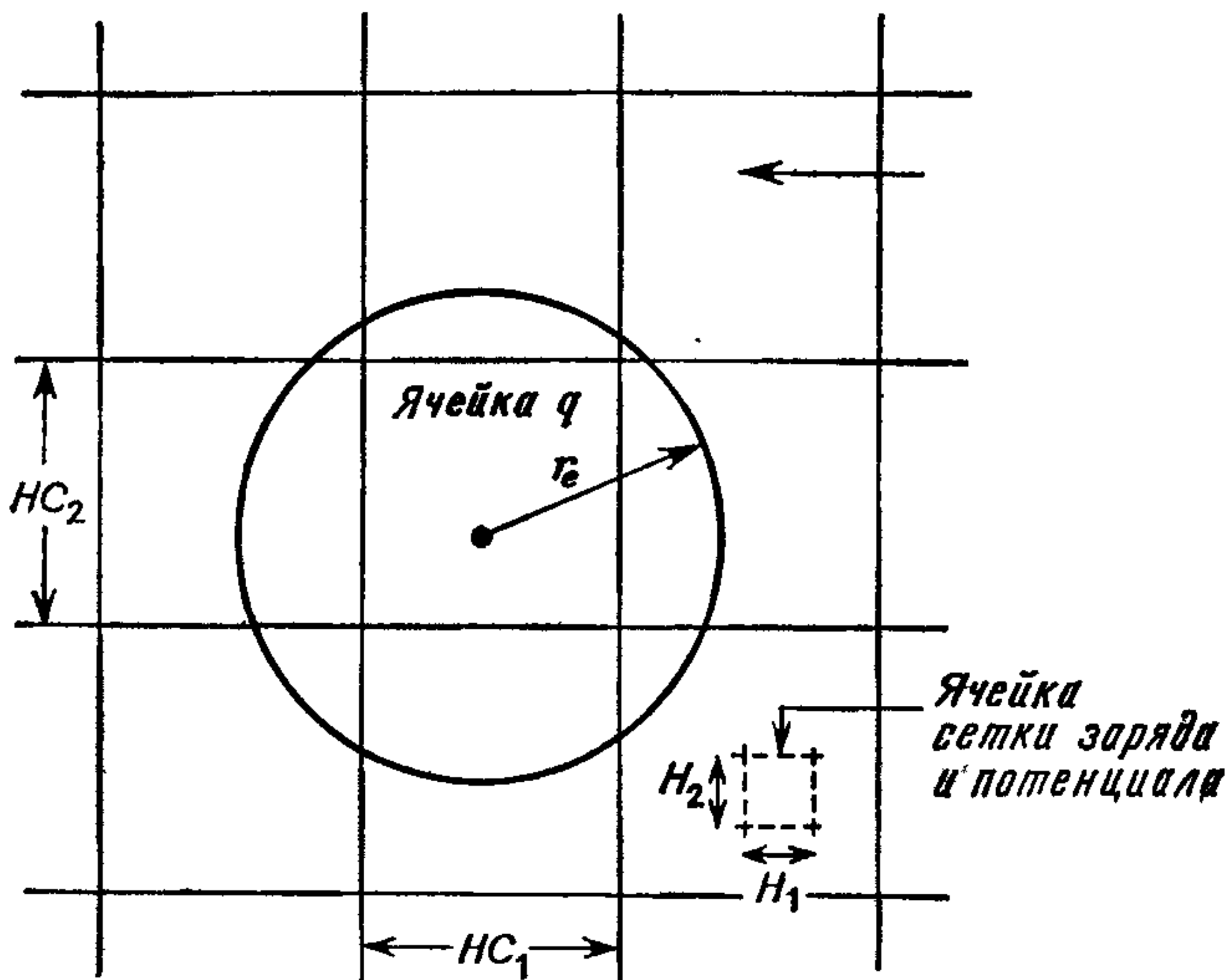


Рис. 8.2. При вычислении короткодействующей силы расчетная область разбивается на цепочечные ячейки (сплошные линии). Ненулевой вклад в силу F_i^{sr} , которая действует на частицу i , находящуюся в ячейке q , вносят только частицы j из ячейки q и соседних ячеек.

водить только для небольшого подмножества N_n полного числа частиц N_p . Именно по этой причине и вводится цепочечная сетка. Цепочечная сетка (в трехмерном случае) представляет собой регулярную решетку из $(M_1 \times M_2 \times M_3)$ ячеек, покрывающих расчетную область (со сторонами $L_1 \times L_2 \times L_3$) примерно таким же образом, как $(N_1 \times N_2 \times N_3)$ ячеек гораздо более мелкой сетки заряда-потенциала. Количество ячеек M_s в направлении s равняется наибольшему целому, не превосходящему L_s/r_e . Тем самым, длины сторон ячеек цепочечной сетки всегда больше или равны радиусу усечения r_e .

На рис. 8.2 изображена цепочечная сетка в двумерном случае. Типичные длины сторон ячеек цепочечной сетки H_{Cs} в 3—4 раза превышают соответствующие размеры H_s ячеек сетки заряда-потенциала. Окружность радиуса r_e с центром в i -й частице, находящейся в цепочечной ячейке q , очерчивает область, в которую должны попасть частицы j , чтобы их вклад в F_i^{sr} был ненулевым. Поскольку $H_{Cs} \geq r_e$ для всех s , из этого следует, что те частицы j , которые дают ненулевой вклад в F_i^{sr} , должны находиться либо в той же самой ячейке q , что и частица i , либо в одной из восьми соседних ячеек. Если для каждой цепочечной ячейки координаты частиц упорядочены в списки, то для нахождения силы F_i^{sr} , действующей на частицу i , требуется приблизи-

тельно $9N_c$ проверок, где $N_c (= N_p/M_1M_2)$ — среднее число частиц в цепочечной ячейке. Таким образом, если учесть третий закон Ньютона, то полное число проверок на нахождение всех короткодействующих сил приблизительно равно $N_n N_p \simeq 4.5 N_c N_p$ по сравнению с N_p^2 в простейшем методе. Аналогично, в трехмерном случае упорядочение координат в цепочечные ячейки дает число проверок $N_n N_p \sim 13 N_c N_p$.

8.4.2. Связные списки

На сериальных компьютерах (но не всегда на машинах с векторными или матричными процессорами) гораздо эффективнее в вычислительном отношении сортировать адреса координат, а не сами координаты. Адресная сортировка реализуется с помощью связного списка LL.

Пусть НОС(q) обозначает элемент массива заголовков цепочек, отвечающий цепочечной ячейке q , и пусть LL(i) является связующей координатой частицы i . Тогда процедуру сортировки координат в списки для каждой цепочечной ячейки можно в итоге организовать с помощью адресной сортировки следующим образом:

1. Задать НОС(q) = 0 для всех q .
2. Выполнить для всех частиц:
 - (а) определить ячейку, в которой находится частица:

$$q := \text{int}(x_1/NC_1, x_2/NC_2, x_3/NC_3);$$

- (б) добавить частицу i в заголовок списка для ячейки q :

$$\begin{aligned} \text{LL}(i) &= \text{НОС}(q), \\ \text{НОС}(q) &= i. \end{aligned}$$

В двумерном случае третьи компоненты у q и x опускаются.

Принцип работы этой процедуры сортировки поясним на примере. Рассмотрим случай, когда три частицы i_1 , i_2 и i_3 находятся в цепочечной ячейке q , и пусть $i_1 < i_2 < i_3$. Будем изображать координаты прямоугольником из трех полей

i	X_i	LL_i
-----	-------	--------

здесь i — адрес (или элемент массива в Фортране), X_i — физические координаты частицы (x_i, p_i) и LL_i — координата связного списка. Если просматривать координаты частиц в порядке возрастания значений i , то формирование связного списка для ячейки q происходит следующим образом:

изначально (после шага 1)

НОС(q) → — 0

$i_1 < i < i_2$

НОС(q) → —

i_1	X_1	LL_1
-------	-------	--------

 → — 0

$i_2 < i < i_3$

НОС(q) → —

i_2	X_2	LL_2
-------	-------	--------

 → —

i_1	X_1	LL_1
-------	-------	--------

 → — 0

$i_3 < i$

НОС(q) → —

i_3	X_3	LL_3
-------	-------	--------

 → —

i_2	X_2	LL_2
-------	-------	--------

 → —

i_1	X_1	LL_1
-------	-------	--------

 → — 0

Скорость и простота создания связанных списков заново делают бессмысленными их сохранение и модификацию на каждом шаге по времени. Весь процесс сортировки требует в трехмерном случае только три арифметические операции с вещественными числами в расчете на одну частицу, или две — в двумерном случае.

Когда массивы НОС и LL заполнены, нулевой элемент в НОС(q) указывает на то, что в цепочечной ячейке q нет ни одной частицы. Ненулевой элемент содержит адрес координат первой частицы этого списка. Связующая координата частицы дает либо адрес координат следующей частицы, входящей в этот список, либо равна нулю, что указывает на исчерпание списка. Следовательно, имея НОС и LL, можно просмотреть координаты в каждой ячейке без какого бы то ни было поиска.

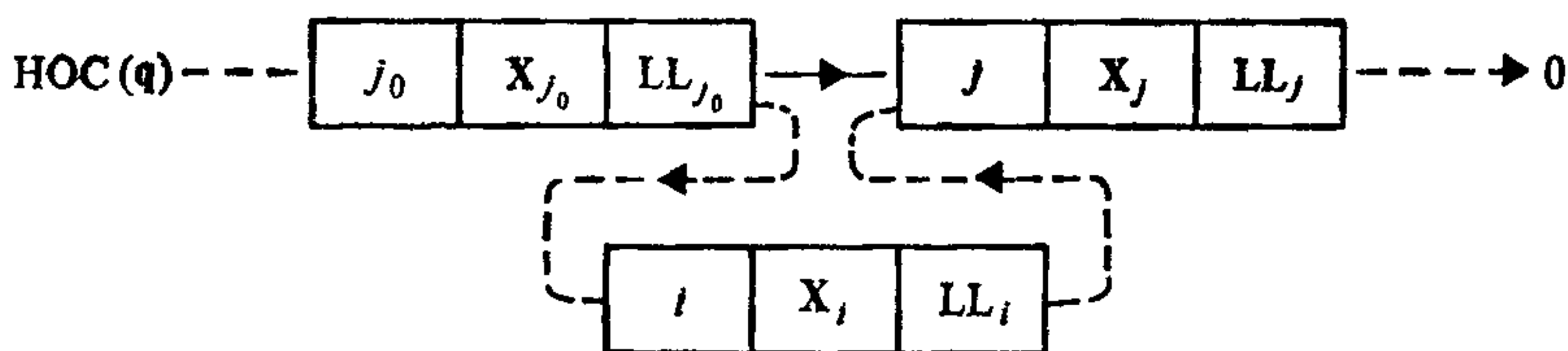
8.4.3 Изменение импульса

Процесс вычисления вклада короткодействующей силы в приращение импульса частицы на каждом шаге по времени осуществляется в два приема. На первой стадии заполняются, как было описано выше, массивы НОС и LL, а на второй — накапливаются приращения в импульсах.

В трехмерном случае частицы, дающие ненулевой вклад в F_i^{sr} , должны находиться либо в той же самой цепочечной ячейке, что и частица i , либо в одной из двадцати шести соседних цепочечных ячеек. Для того, чтобы исключить ненужное дублирование при вычислении значений f_{ij}^{sr} , используемых в приращениях импульсов как частицы i , так и частицы j , желательно обрабатывать каждую пару частиц (i, j) , $i \neq j$, только один раз. Один из приемов достижения этого состоит в том, чтобы просматривать до конца каждую цепочечную ячейку, которую мы

будем называть текущей ячейкой. Для текущей цепочечной ячейки получают f_{ij}^{sr} для всех пар (i, j) частиц, принадлежащих текущей ячейке, а также для всех пар с одним элементом i из текущей ячейки, а другим — из одной из тринадцати соседних ячеек. Можно, например, выбрать в качестве соседей цепочечной ячейки $q = (q_1, q_2, q_3)$ ячейки $\{(q_1 + s, q_2 - 1, q_3 + t), (q_1 + s, q_2, q_3 - 1), (q_1 - 1, q_2, q_3)\}; s, t \in [-1, 1]\}$. Пусть $\mathcal{P}(q)$ обозначает множество пар частиц, первая из которых, i , находится в цепочечной ячейке q а другая j — в ячейке q или в од-

каждой пары «текущая ячейка — соседняя ячейка», как только встретится $\Delta y = y_i - y_j > r_e$. Для девяти ячеек $\{(q_1 + s, q_2, q_3 + t); (s, t) \in [-1, 1]\}$, соседних с цепочечной ячейкой q , это приведет к тому, что в среднем количество расстояний между частицами, которое надо будет проверять, уменьшится на 50%. Как правило, экономия времени на стадии приращения импульсов с лихвой возмещает дополнительные затраты на формирование связных списков. Изменения в алгоритме построения связного списка состоят в том, что значение y нового i -го элемента y_i сравнивается последовательно с имеющимися в списке значениями y , и, как только первый раз встретится элемент j с $y_j > y_i$, перед ним вставляется элемент i :



В этом примере замена старой связи новой связью (штриховая линия) осуществляется операциями

$$\begin{aligned} LL(i) &:= j, \\ LL(j_0) &:= i. \end{aligned}$$

8.5. ВРЕМЕННОЕ УРАВНЕНИЕ

Метод Q -минимизации, описанный в общих чертах в разд. 8.3.3, дает оптимальные комбинации интерполяционной функции, функции влияния и разностного оператора дифференцирования. Выбирая различные интерполяционные функции, используя чередование, меняя ограничения на тип разностного оператора или просто изменяя эталонную силу, можно найти много оптимальных комбинаций. Каждая из них предъявляет свои требования к машинному времени и памяти. Величина Q дает возможность оценить качество таких оптимальных комбинаций. А соотношение для времени счета (временное уравнение) дает дополнительную информацию, необходимую для определения того, какой из многих способов достижения данного значения Q наилучший, причем самой оптимальной схемой является та, которая обеспечивает требуемую точность вычисления силы с наименьшими вычислительными затратами.

Еще одним фактором, входящим в определение вычислительных затрат, является ограниченный объем памяти. Несмотря на то что разбиение на порции частиц или сеточных данных либо и того и другого, а также использование дополнительной памяти с последовательным доступом не представляет особых трудно-

стей, практический опыт показывает, что там, где возможно, этого лучше избегать. Тем самым исключается дорогостоящий этап физической сортировки координат частиц и предотвращается ситуация, когда время счета определяется в основном скоростью обмена с диском. В данной работе предполагается, что ограничения по памяти учитываются простейшим образом. Если для какой-нибудь конкретной схемы требуется больший объем памяти с произвольным доступом, чем может быть предоставлено, то она исключается из списка возможных вариантов. После чего оставшиеся варианты сравниваются по временам цикла и находится оптимальный вариант.

В табл. 8.1 перечислены этапы, составляющие цикл по времени, и для каждого из них приведены зависимости количества

Таблица 8.1. Типичные значения числа операций для этапов цикла по времени в Р³М-алгоритме

Расчет РМ-силы

1. Раздача заряда на сетку	$C_1 N_p$
2. Нахождение потенциала	$C_2 N^3 \log N^3$
3. Интерполяция сил и ускорение частиц	$C_3 N_p$

Расчет РР-силы

1. Заполнение массивов НОС и LL	$C_4 N_p$
2. Интерполяция сил и ускорение частиц	$C_5 \left(\frac{r_e^3 N_p}{H^3 N^3} \right) N_p$

Уравнения движения

Перемещение частиц	$C_6 N_p$
--------------------	-----------

Затраты на диагностику

Внутренняя энергия	$C_7 N^3 + C_8 N_t$
Кинетическое давление	$C_9 N_p$
Давление, обусловленное взаимодействием	$C_9 N_t$

Функция радиального распределения (для $r < r_e$)	$C_{10} \left(\frac{r_e^3 N_p}{H^3 N^3} \right) N_p$
--	---

Структурный фактор (для $kH < \pi$)	$C_{11} N^3$
--------------------------------------	--------------

операций от числа частиц N_p , количества узлов сетки заряда-потенциала N^3 и радиуса усечения r_e короткодействующей силы. Для ясности мы ограничиваемся рассмотрением кубической сетки заряда-потенциала со стороной, равной N узлов. Величины C_i ; $i = 1, \dots, 11$ являются константами, значение которых определяется особенностями алгоритма, используемого на том или ином этапе. В первом приближении $\{C_i\}$ пропорциональны числу арифметических операций с вещественными числами. Функциональная зависимость от N_p , N и т. д. установлена путем исследования роста числа операций. Зависимость $N^3 \log N^3$, связанная

с нахождением потенциала с помощью быстрого преобразования Фурье, рассматривалась в гл. 6. Величина $(r_e^3 N_p / H^3 N^3)$, относящаяся к расчету короткодействующей силы парного взаимодействия, представляет собой с точностью до константы среднее число соседей, окружающих каждую частицу в сфере радиуса r_e . Для некоторых этапов вычисления силы количество операций и определение масштабных коэффициентов рассматриваются далее в разд. 8.7. Для остальных этапов определение числа операций и масштабов мы оставляем читателю в качестве упражнения.

В конце табл. 8.1 приведен раздел, озаглавленный «Затраты на диагностику», призванный особо подчеркнуть тот факт, что в программе P³M накладные расходы на измерение физических величин невелики. Измерение функции радиального распределения g для $r \leq r_e$ требует одну дополнительную операцию на пару частиц в цикле вычисления короткодействующей силы, если эта функция набирается по схеме NGR, или три, если применяется обратная линейная интерполяция. Если g определена, то ценой двух операций на каждую табличную точку g можно получить давление и внутреннюю энергию, отвечающие короткодействующей части взаимодействия. N_t обозначает полное число табличных точек, используемых при накоплении g . Члены, пропорциональные N_t , составляют ничтожную долю времени полного цикла. Кинетическое давление получается из кинетической энергии частиц. Часть внутренней энергии, отвечающая дальнему действию, и структурный фактор (т. е. соответствующее преобразование корреляционной функции плотности) для малых k получаются непосредственно из гармоник плотности. Дополнительные подробности об измерении физических характеристик приведены в разд. 12.3.4.

Полное время T цикла одного шага по времени в алгоритме P³M получается сложением вкладов всех этапов. Это дает

$$T = \alpha N_p + \beta N^3 \log N^3 + \gamma \left(\frac{r_e}{H} \right)^3 \frac{N_p^2}{N^3}, \quad (8.32)$$

где коэффициенты α , β и γ зависят от конкретного вида различных этапов счетного цикла. Отметим, что в табл. 8.1 и формуле (8.32) оставлены только доминирующие факторы — такие факторы, как, например, поверхностные члены в определении потенциала (пропорциональные N^2) не учитываются.

Вычислительные затраты Γ в расчете на частицу за шаг по времени определяются выражением

$$\Gamma = \frac{T}{N_p} = \alpha + \frac{\beta N^3 \log N^3}{N_p} + \gamma \left(\frac{r_e}{H} \right)^3 \frac{N_p}{N^3}. \quad (8.33)$$

Если N_p мало, то главным фактором является время, затрачиваемое на нахождение потенциала (член $N^3 \log N^3$ в (8.33)).

По мере увеличения N_p доля времени нахождения потенциала, приходящаяся на частицу, снижается и доминирующим членом становится время вычисления сил парного взаимодействия на этапе расчета РР-силы. При заданном числе сеточных узлов N^3 оптимальная рабочая точка достигается, когда

$$\frac{\partial \Gamma}{\partial N_p} = - \frac{\beta N^3 \log N^3}{N_p^2} + \gamma \left(\frac{r_e}{H} \right)^3 \frac{1}{N^3} = 0, \quad (8.34)$$

т. е. когда

$$\beta N^3 \log N^3 = \gamma \left(\frac{r_e}{H} \right)^3 \frac{N_p^2}{N^3}. \quad (8.35)$$

Из сравнения (8.35) с (8.32) видно, что самая рентабельная рабочая точка для Р³М программы достигается в том случае, когда на нахождение потенциала и на вычисление короткодействующей части сил затрачивается одинаковое машинное время.

Преобразуя (8.35), получаем оптимальное число частиц в ячейке:

$$\left(\frac{N_p}{N^3} \right)_{\text{opt}} = \left(\frac{\beta \log N^3}{\gamma (r_e/H)^3} \right)^{1/2}. \quad (8.36)$$

Если работать в оптимальной точке, то временное уравнение (8.32) упрощается и принимает вид

$$T = \left[\alpha + 2 \left(3\gamma\beta \left(\frac{r_e}{H} \right)^3 \log N \right)^{1/2} \right] N_p. \quad (8.37)$$

Таким образом, если не говорить о слабой нелинейной зависимости от N в виде $\sqrt{\log N}$, время цикла Р³М-алгоритмов пропорционально числу частиц. Как следствие этого закона Р³М-алгоритмы сделали реальным моделирование коррелированных систем с размерами, ранее недостижимыми.

8.6. ОПТИМИЗАЦИЯ

Результаты разд. 8.3.3 и 8.5 можно свести воедино и получить количественный метод конструирования Р³М-программ и выбора численных параметров для использования в таких программах. Чтобы проиллюстрировать этот метод, мы рассмотрим его применительно к конструкции программы РЗМЗДР [117, 120]. Примеры расчетов, выполненных с помощью РЗМЗФР, описаны в гл. 11 и 12.

8.6.1. Вычисление точности силы

Первая стадия оптимизации конструкции состоит в том, чтобы выделить комбинацию элементов расчета силы, которые дают силы с приемлемой точностью. Полученные в разд. 8.3.3

выражения для погрешностей сил межчастичного взаимодействия дают возможность осуществить этот шаг для большого разнообразия вариантов, не взваливая на себя тягостной задачи по кодированию (и отладке!) каждого варианта. Все варианты параметризуются посредством фурье-образов трех величин \hat{D} , \hat{U} и \hat{R} и, кроме того, ограничениями на значения n в суммах по сеточным гармоникам для схем с чередованием. Очевидно, что нет никакого смысла использовать какие бы то ни было варианты \hat{G} , кроме оптимальных (выражения (8.22) или (8.29)), которые полностью определяются \hat{D} , \hat{U} , \hat{R} и ограничениями на n . Зная эти величины, можно оценить точность сил межчастичного взаимодействия, вычисляя величины Q , P , Z и E_{ref} по формулам (8.23)—(8.26) либо аналогичных им для схем с чередованием.

Хотя выражения для Q и т. д. содержат тройные бесконечные суммы (по n), их, как оказывается, несложно вычислить. Эти суммы таковы, что они либо очень быстро сходятся, либо могут быть записаны в замкнутой функциональной форме. Рассмотрим, например, выражение (8.23). Используя (8.4), первую бесконечную сумму можно записать в виде

$$\sum_n |\hat{R}(\mathbf{k}_n)|^2 = \sum_n \frac{\hat{S}^4(k_n)}{|\mathbf{k}_n|^2}. \quad (8.38)$$

Ограничимся рассмотрением кубической сетки ($H = H_1 = H_2 = H_3$), так что $\mathbf{k}_n = \mathbf{k} + n\mathbf{k}_g$, где $k_g = 2\pi/H$. На практике возможные варианты $S \supset \hat{S}$ ограничивают сферически симметричными функциями $S(\mathbf{x}) = S(r)$, $r = |\mathbf{x}|$. Иначе при вычислении короткодействующей силы приходилось бы использовать дорогие в вычислительном отношении многомерные таблицы коррекции силы. Для сферически симметричных формфакторов S трехмерное преобразование Фурье приводит к

$$\hat{S}(k) = 4\pi \int_0^\infty r^2 dr S(r) \frac{\sin kr}{kr}, \quad (8.39)$$

где

$$k = |\mathbf{k}| = (k_1^2 + k_2^2 + k_3^2)^{1/2}. \quad (8.40)$$

Отметим, что в (8.39) мы использовали преобразование Фурье по бесконечной области. Можно использовать это преобразование также и для силы, поскольку дискретизация гармоник бесконечной системы дает гармоники периодической системы. Такая простота соотношения между гармониками разительно отличается от тех сложностей, которые вносит периодичность в x -пространстве. Если бы для вычисления Q мы захотели воспользоваться формулой (8.15), то возникающие ин-

тегралы приводили бы к эллиптическим функциям (в двумерном случае) или разложениям по сферическим гармоникам (в трехмерном случае). Вычисление же интеграла (8.39) для частицы с формой $S2$ (формула (8.3)) дает

$$\hat{S}(k) = \frac{12}{(ka/2)^4} \left(2 - 2 \cos \frac{ka}{2} - \frac{ka}{2} \sin \frac{ka}{2} \right). \quad (8.41)$$

Следовательно, для частицы с формой $S2$ величина \hat{S}^4 спадает как $(ka/2)^{-16}$. При условии что $a/H \gtrsim 1$, те члены, у которых компоненты \mathbf{n} по величине больше единицы или двойки, пренебрежимо малы.

Члены второй суммы в (8.23), куда входят гармоники эталонной силы, также быстро убывают по величине с уходом компонент \mathbf{n} от нуля. \hat{R}^* спадает как k^{-9} , а \hat{U}^2 — как $k^{-2(p+1)}$, где p — порядок интерполяционной функции. В частности, для иерархии схем $NGP - CIC - TSC - \dots$, $\hat{U}(k) = \hat{U}(k_1, k_2, k_3)$ имеет вид

$$\hat{U}(\mathbf{k}) = \left(\frac{\sin \frac{k_1 H}{2}}{\frac{k_1 H}{2}} \frac{\sin \frac{k_2 H}{2}}{\frac{k_2 H}{2}} \frac{\sin \frac{k_3 H}{2}}{\frac{k_3 H}{2}} \right)^{p+1}. \quad (8.42)$$

Оставшуюся в (8.23) сумму по сеточным гармоникам можно записать в замкнутом виде, воспользовавшись введенными в предыдущей главе тождествами для тригонометрических функций. Например, схема TSC ($p = 2$) дает сумму

$$\sum_{\mathbf{n}} \hat{U}^2 = \prod_{i=1}^3 \left[\sin^6 \frac{k_i H}{2} \sum_{n_i=-\infty}^{\infty} \frac{1}{\left(\frac{k_i H}{2} - \pi n_i \right)^6} \right]. \quad (8.43)$$

Тождество

$$\sin^{-6} x \left(1 - \sin^2 x + \frac{2}{15} \sin^4 x \right) = \sum_{n=-\infty}^{\infty} \frac{1}{(x - \pi n)^6} \quad (8.44)$$

приводит (8.43) к виду

$$\sum_{\mathbf{n}} \hat{U}^2 \prod_{i=1}^3 \left(1 - \sin^2 \frac{k_i H}{2} + \frac{2}{15} \sin^4 \frac{k_i H}{2} \right). \quad (8.45)$$

Преобразование Фурье разностного оператора градиента \hat{D} получается путем преобразования уравнения, связывающего сеточные потенциалы с сеточными полями. В случае двухточечных центральных разностей, взятых по узлам, находящимся на расстояниях $2mH$, $m = 1, 2, \dots$, это дает для j -й компоненты ($j = 1, 2, 3$)

$$\hat{D}_j = \frac{i \sin mk_j H}{mH}. \quad (8.46)$$

Следовательно, для четырехточечных конечно-разностных приближений (8.17) имеем

$$\hat{D}_j = i\alpha \frac{\sin k_j H}{H} + i(1 - \alpha) \frac{\sin 2k_j H}{2H}. \quad (8.47)$$

Объединяя результаты, выраженные в (8.38) — (8.47), с (8.23), получаем сравнительно простое выражение, которое можно вычислить для нахождения меры качества силы Q . С вычислением Q связан один практический вопрос, состоящий в том, что выражение (8.23) представляет собой сумму величин $\hat{Q}(k) = \hat{Q}(l_1, l_2, l_3)$, которые являются четными и периодическими относительно своих аргументов:

$$Q = \frac{1}{V_b} \sum_{l_1=0}^{N_1-1} \sum_{l_2=0}^{N_2-1} \sum_{l_3=0}^{N_3-1} \hat{Q}(l_1, l_2, l_3). \quad (8.48)$$

Вследствие этого за счет приведения подобных членов достигается восьмикратное уменьшение объема вычислений:

$$Q = \frac{1}{V_b} \sum_{l_1=0}^{N_1/2} \sum_{l_2=0}^{N_2/2} \sum_{l_3=0}^{N_3/2} \alpha(l_1) \alpha(l_2) \alpha(l_3) \hat{Q}(l_1, l_2, l_3), \quad (8.49)$$

где

$$\alpha(l_i) = \begin{cases} 1, & \text{если } l_i = 0 \text{ или } N_i/2, \\ 2 & \text{в других случаях.} \end{cases} \quad (8.50)$$

Более того, если расчетная область является кубом ($N = N_1 = N_2 = N_3$), то можно использовать полную 48-кратную симметрию кубической решетки и получить дополнительное уменьшение числа слагаемых в шесть раз, т. е.

$$Q = \frac{1}{V_b} \sum_{l_1=0}^{N/2} \sum_{l_2=0}^{l_1} \sum_{l_3=0}^{l_2} \alpha(l_1) \alpha(l_2) \alpha(l_3) \gamma(l_1, l_2, l_3) \hat{Q}(l_1, l_2, l_3), \quad (8.51)$$

где

$$\gamma(l_1, l_2, l_3) = \begin{cases} 1, & \text{если } l_1 = l_2 = l_3, \\ 3, & \text{если } l_1 = l_2 \neq l_3, l_3 = l_1 \neq l_2 \text{ или } l_2 = l_3 \neq l_1, \\ 6 & \text{в других случаях.} \end{cases} \quad (8.52)$$

Величины G , P , Z и E_{ref} находятся аналогичным образом.

8.6.2. Сравнение схем

На рис. 8.3—8.5 показаны примеры численного расчета Q , P , Z и E_{ref} для нескольких схем, в которых для распределения заряда и интерполяции силы использована схема TSC. Схемы бо-

лее низкого порядка, NGP или CIC, обладающие большими погрешностями при распределении заряда и интерполяции силы, не годятся для точного моделирования силы. Схемы более высокого порядка, чем TSC, можно вообще не принимать во внимание по причине больших вычислительных затрат. Так, в следующей по порядку за TSC схеме число узлов распределения заряда возрастает с 27 до 64, что приводит примерно к трехкратному увеличению величины α во временном уравнении. Если одновременно с этим не перейти в весьма дорогой в вычислительном отношении режим с большим r_e/H , то указанными дополнительными вычислениями почти не достигается улучшение точности силы, поскольку точность ограничивается количеством гармоник, которые может передать сетка.

Функциональная зависимость величины погрешностей показывает лишь слабую зависимость от формы облака (см. разд. 8 2). Поэтому, хотя результаты приведены только для облаков формы S2, аналогичные вычисления для однородного распределения заряда (форма S1) и гауссова распределения (форма S3) не меняют общего характера результатов. С хорошей точностью показанные на рис. 8 3—8 5 кривые можно отнести к любой гладко меняющейся сферически симметричной форме частиц, отвечающих эталонной силе взаимодействия с характерной шириной распределения r_e (см. (8.5)), т. е. можно считать Q и т. д. функцией r_e , а не функционалом от формы S1, S2, S3 и т. д.

Вторым фактором, который очень мало влияет на Q , P , Z и E_{ref} , является размер расчетной области. Если зафиксировать размер ячейки H и увеличивать число сеточных узлов, то изменения в Q и т. д. будут пренебрежимо малы за исключением случаев, где $r_e \simeq L/2$ (L — длина стороны расчетной области). Можно показать, что в x -пространстве это свойство обусловлено пространственной локализацией погрешностей в силе межчастичного взаимодействия. При расстояниях между частицами, превышающих несколько шагов сетки, отличие РМ-силы от эталонной силы пренебрежимо мало. В k -пространстве последовательность выражений для Q , отвечающая увеличению числа узлов сетки, может интерпретироваться как последовательные приближения к интегралам, получаемым в пределе $N \rightarrow \infty$, которые вычисляются методом трапеций на все более мелких сетках в k -пространстве, т. е.

$$Q_{\infty} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \hat{Q}(\mathbf{k}) = \lim_{\substack{N \rightarrow \infty \\ H = \text{const}}} \left(\sum_{l=-N/2}^{N/2-1} \frac{k_0^3}{(2\pi)^3} Q(\mathbf{k}) \right), \quad (8.53)$$

где $k_0 = 2\pi/L = 2\pi/NH$. Интеграл в (8.53) берется по основной зоне гармоник (компоненты k_j вектора \mathbf{k} лежат в области $|k_j| \leq \pi/H$).

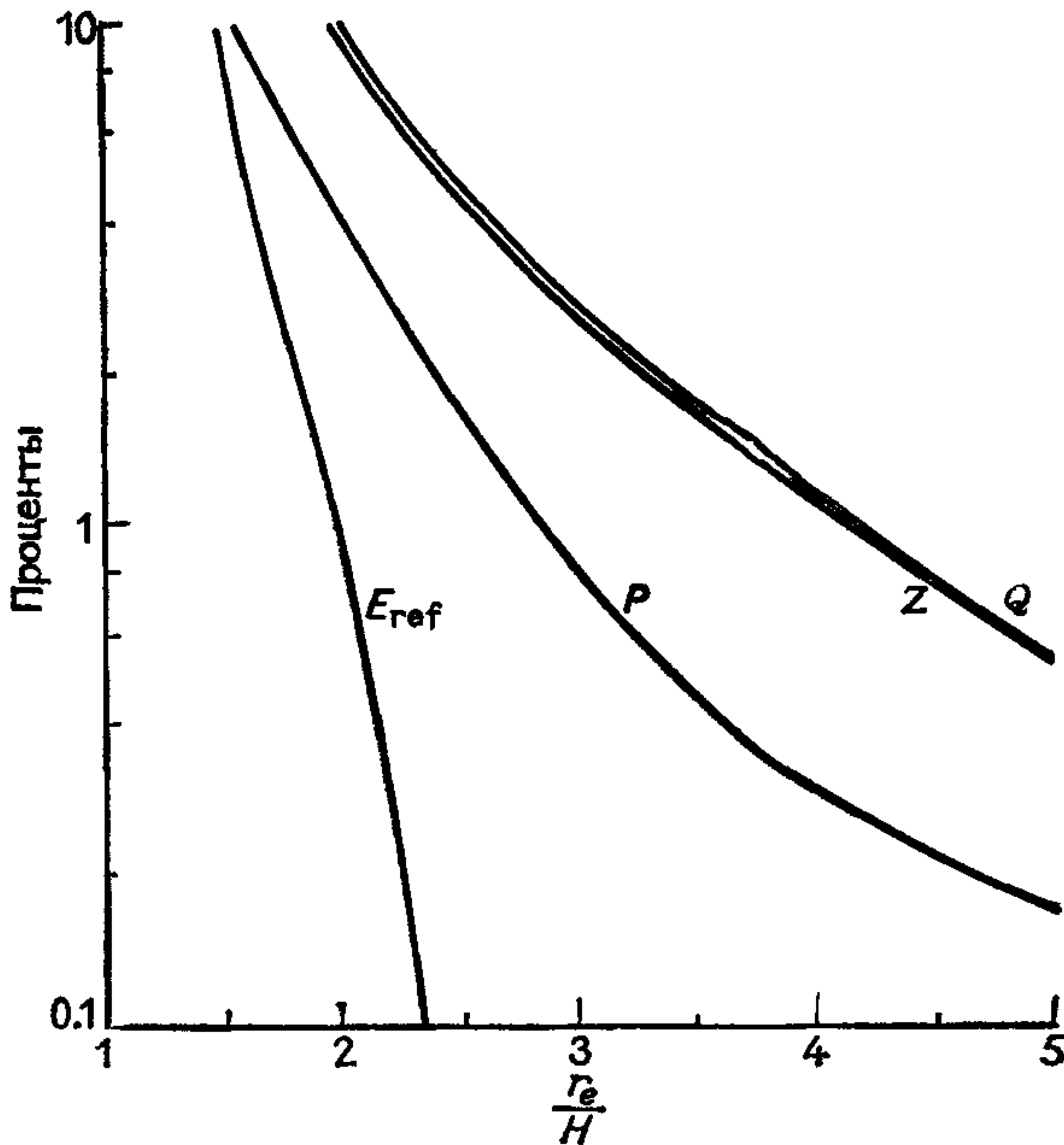


Рис. 8.3. Зависимость $\sqrt{Q^+}$, $\sqrt{P^+}$, $\sqrt{Z^+}$ и $\sqrt{E_{\text{ref}}^+}$ от эффективного радиуса усечения r_e короткодействующей силы для эталонной силы с формой S2 и обычного алгоритма РМ-силы, в котором использованы интерполяционная функция TSC и двухточечные конечно-разностные аппроксимации для компонент градиента потенциала. Подробности относительно нормировки см. в тексте. (Из работы [117], с разрешения Advance Publications, Ltd.)

Величины Q и др., представленные на рис. 8.3—8.5, нормированы таким образом, чтобы отразить пространственную локализацию погрешностей. Q , P , Z и E_{ref} имеют размерность (сила)² × объем. Переход к безразмерным величинам осуществляется путем деления на квадрат кулоновской силы взаимодействия единичных зарядов, находящихся на расстоянии H , умноженный на объем шара радиуса H . Таким образом,

$$Q^+ = \frac{Q}{\left(\frac{1}{4\pi\epsilon_0 H^2}\right)^2 \left(\frac{4}{3}\pi H^3\right)}, \quad (8.54)$$

и аналогично для P^+ , Z^+ и E_{ref}^+ . На рис. 8.3—8.5 приведены выраженные в процентах зависимости величин $100\sqrt{Q^+}$ и др. от r_e/H , там же указано, какой величине — Q , P , Z или E_{ref} — отвечает соответствующая кривая. Поскольку на расстояниях, больших H , встречаются погрешности не исчезающе малые (см. рис. 8.8), нормировка, использованная на рис. 8.3—8.5, дает (слабую) оценку сверху отклонения РМ-силы межчастичного взаи-

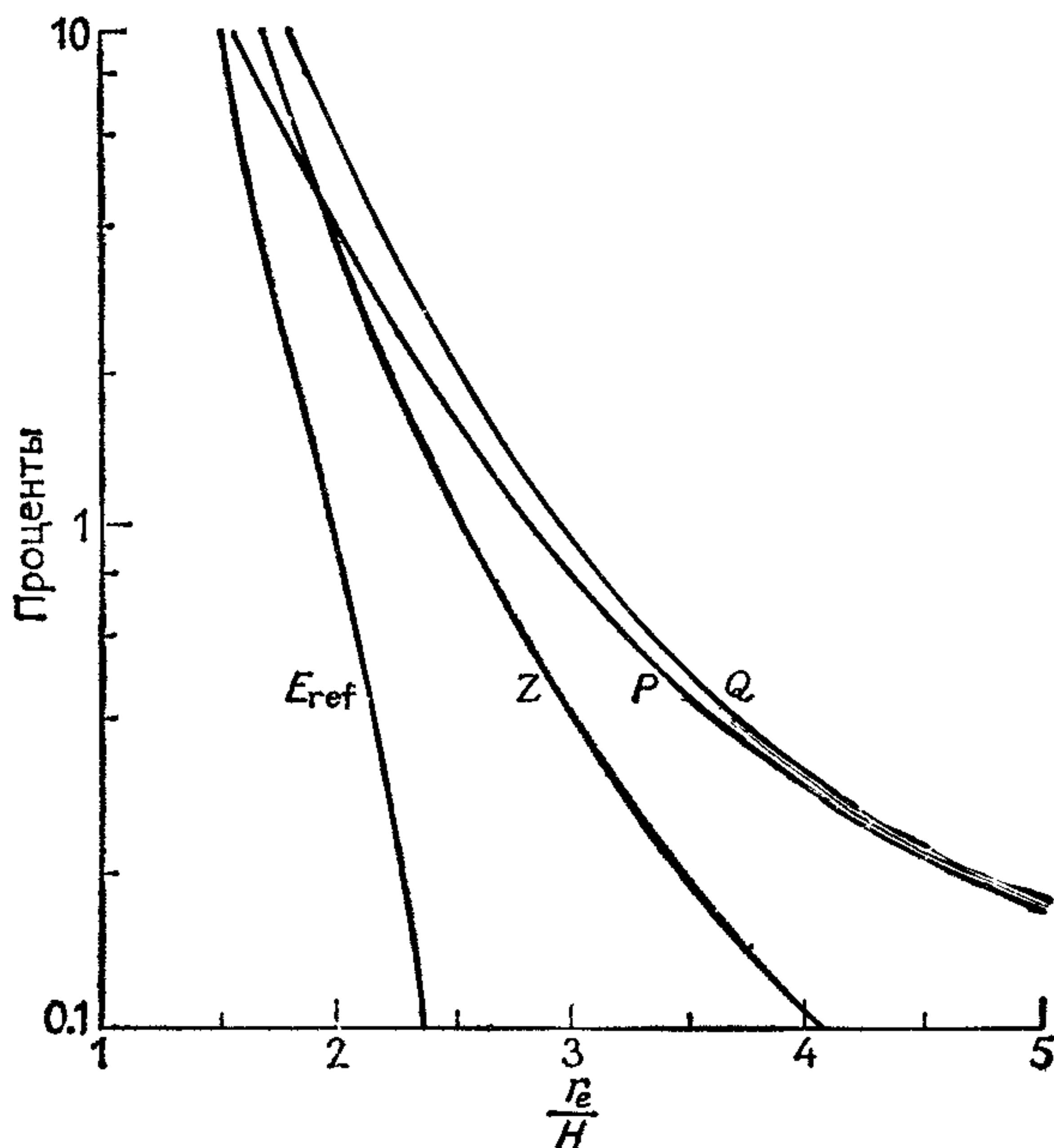


Рис. 8.4. То же, что на рис. 8.3, но с заменой двухточечных конечно-разностных аппроксимаций на оптимизированные четырехточечные. (Из работы [117], с разрешения Advance Publications Ltd.)

модействия от эталонной силы для любого расстояния между частицами и их расположения относительно сетки.

На рис. 8.3 представлены кривые для обычного расчета РМ-силы, в котором использованы интерполяционная функция TSC и двухточечная конечно-разностная аппроксимация компонент градиента потенциала. В этом примере основной вклад в Q вносит Z , и при этом Z намного больше E_{ref} . Следовательно, погрешности дифференцирования потенциала преобладают над погрешностями силы межчастичного взаимодействия для всех r_e . Чтобы добиться существенного улучшения, нужна схема дифференцирования потенциала более высокого порядка. Эффект от повышения порядка интерполяционной функции был бы весьма незначительным.

Рис. 8.4 дает наглядное представление об очень заметном улучшении, являющимся результатом замены двухточечной разностной аппроксимации на четырехточечную аппроксимацию, в которой коэффициент линейной комбинации α ($\approx 4/3$) выбирался так, чтобы Q было минимально. По сравнению с рис. 8.3 определяющий вклад в Q дает теперь флуктуирующая часть P , за исключением малых r_e . При малых r_e ограниченность диапазона волновых чисел, которые могут передаваться сеткой, при-

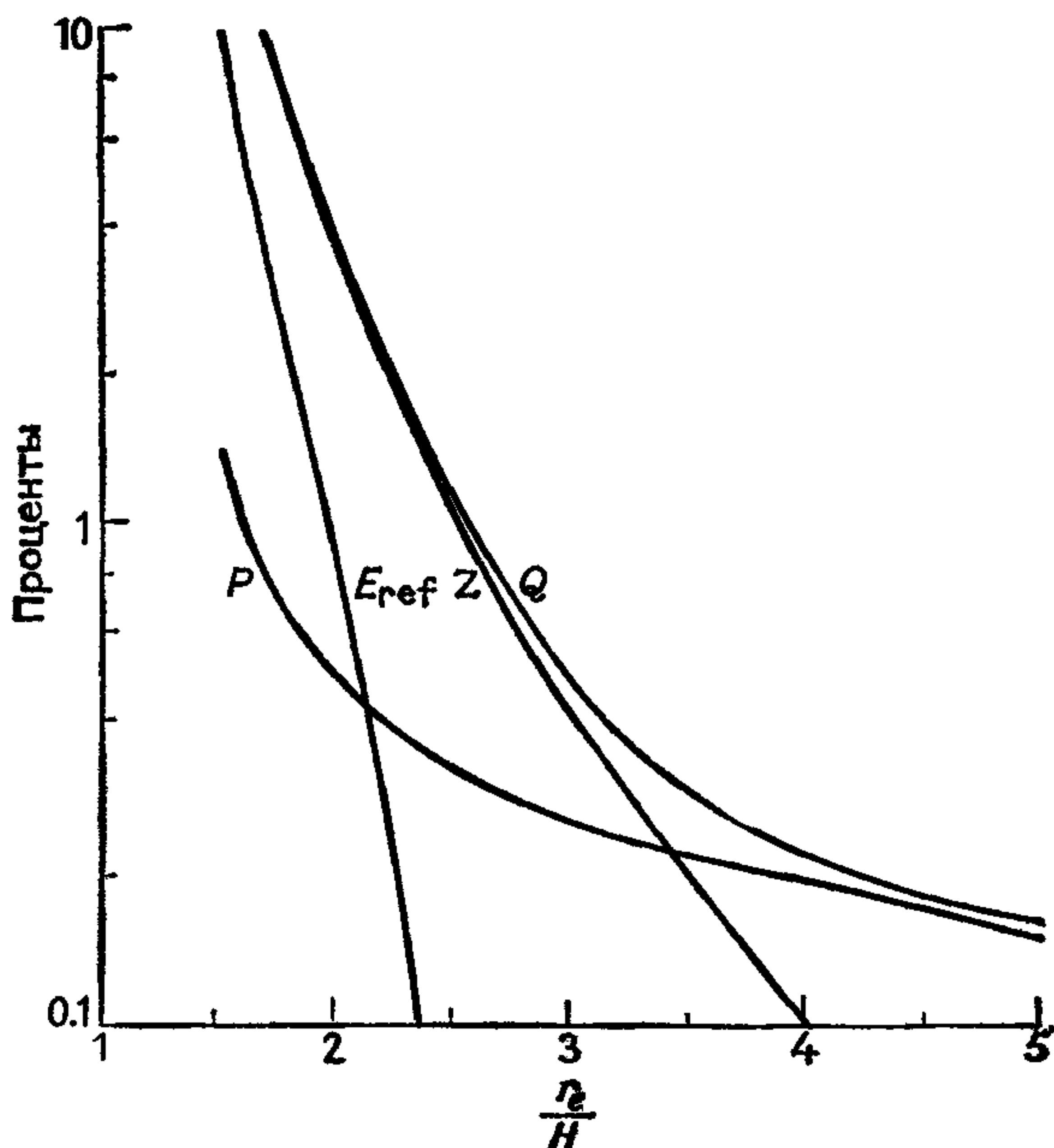


Рис. 8.5. То же, что на рис. 8.4, но для раздачи заряда по схеме TSC на две чередующиеся сетки. (Из работы [117], с разрешения Advance Publications Ltd.)

водит к тому, что Z начинает играть основную роль. Теперь самый эффективный способ уменьшить погрешность при данном r_e состоит в том, чтобы использовать функцию распределения заряда более высокого порядка либо ввести чередование. На рис. 8.5 показан результат введения одного чередования (т. е. распределения заряда на две сетки). Для малых r_e опять преобладает Z , что указывает на то, что для уменьшения Q необходима схема дифференцирования более высокого порядка, тогда как при больших r_e преобладают флуктуации, а это показывает, что наиболее эффективным способом уменьшить Q явилось бы распределение заряда по схеме более высокого порядка или дополнительное чередование.

8.6.3. Соотношение цена — качество

Вопрос о том, как лучше всего добиться требуемого Q — увеличивать r_e или вводить в расчет РМ-силы элементы более высокого порядка — решается рассмотрением зависимости между вычислительными затратами и качеством для различных вариантов. Затраты определяются с помощью временного уравнения (8.32), в котором коэффициенты α , β и γ можно оценить

по количеству операций вещественной арифметики или проводя хронометрирование в программных модулях. Качество определяется величиной Q .

Для получения количественных результатов воспользуемся здесь значениями α , β и γ , полученными путем усреднения ряда замеров времени, сделанных на машине IBM 360/195, в программных модулях, использующих компилятор FORTRAN H. По данным, использованным для построения рис. 8.3—8.5, для каждого варианта строится обратная зависимость $r_e = r_e(Q)$. Тогда время счета в формуле (8.32) получается как функция числа частиц, размера сетки и качества: $T = T(N_p, N, Q)$. Если работать в оптимальной рабочей точке, временное уравнение для схемы s превращается в функцию двух переменных: $T_s = T_s(N_p, Q)$. Практически значения N берутся равными степени двойки, чтобы воспользоваться эффективным алгоритмом быстрого преобразования Фурье по основанию 2. В этом случае временное уравнение превращается в функцию от N_p и Q , если положить

$$T_s(N_p, Q) = \min_r \{T_s(N_p, N = 2^r, Q)\}. \quad (8.55)$$

Формула (8.55) дает время счета одного варианта s алгоритма Р³М. Оптимальной из множества вариантов Р³М-алгоритмов является та конкретная схема s , которая минимизирует время счета:

$$T_{\text{opt}} = \min_s \{T_s(N_p, Q)\}. \quad (8.56)$$

На рис. 8.6 представлены линии уровня оптимального времени счета в $(\sqrt{Q^+}, N_p)$ -плоскости. Промежутки между ними на рис. 8.6 неодинаковы (их постоянность означала бы, что $T_{\text{opt}} \sim \sim N_p$), поскольку выбор $N = 2^r$ препятствует оптимальному согласованию (равенство (8.35)) N_p и N во всем диапазоне изменения $\sqrt{Q^+}$ и N_p . Времена счета, которые указаны у кривых, отвечают алгоритму, запрограммированному полностью на Фор-тране, за исключением программ преобразования Фурье, которые были закодированы вручную на языке ассемблера. Измерения времени относятся к однородному распределению частиц, заполняющих всю расчетную область. В некоторых задачах, как, например, о сгущивании галактик, могут встречаться несколько бóльшие (примерно на 50 %) времена счета, поскольку время вычисления короткодействующей силы чувствительно к степени неоднородности распределения частиц.

На рис. 8.7 в плоскости параметров $\sqrt{Q^+}$ и N_p для каждого варианта показана область, в которой этот вариант оказался наиболее рентабельным. Эта диаграмма вместе с результатами, отраженными на рис. 8.3—8.6, сводит проблему подготовки чис-

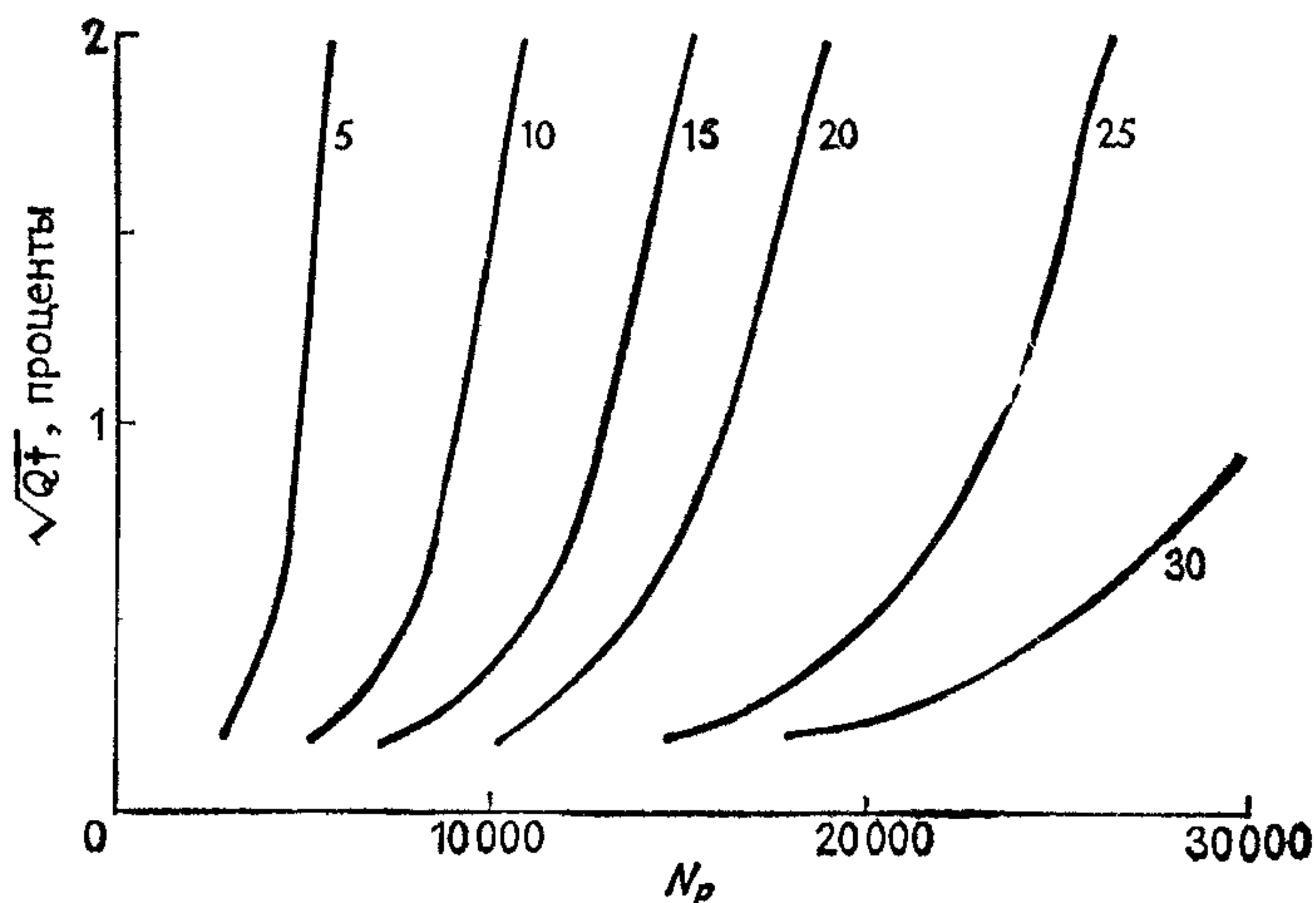


Рис 8.6 Изолинии счетного времени одного прохождения основного цикла по временному шагу P^3M -алгоритма в плоскости качество — число частиц. Числа около кривых обозначают время в секундах для IBM 360/195 (Из работы [117], с разрешения Advance Publications Ltd)

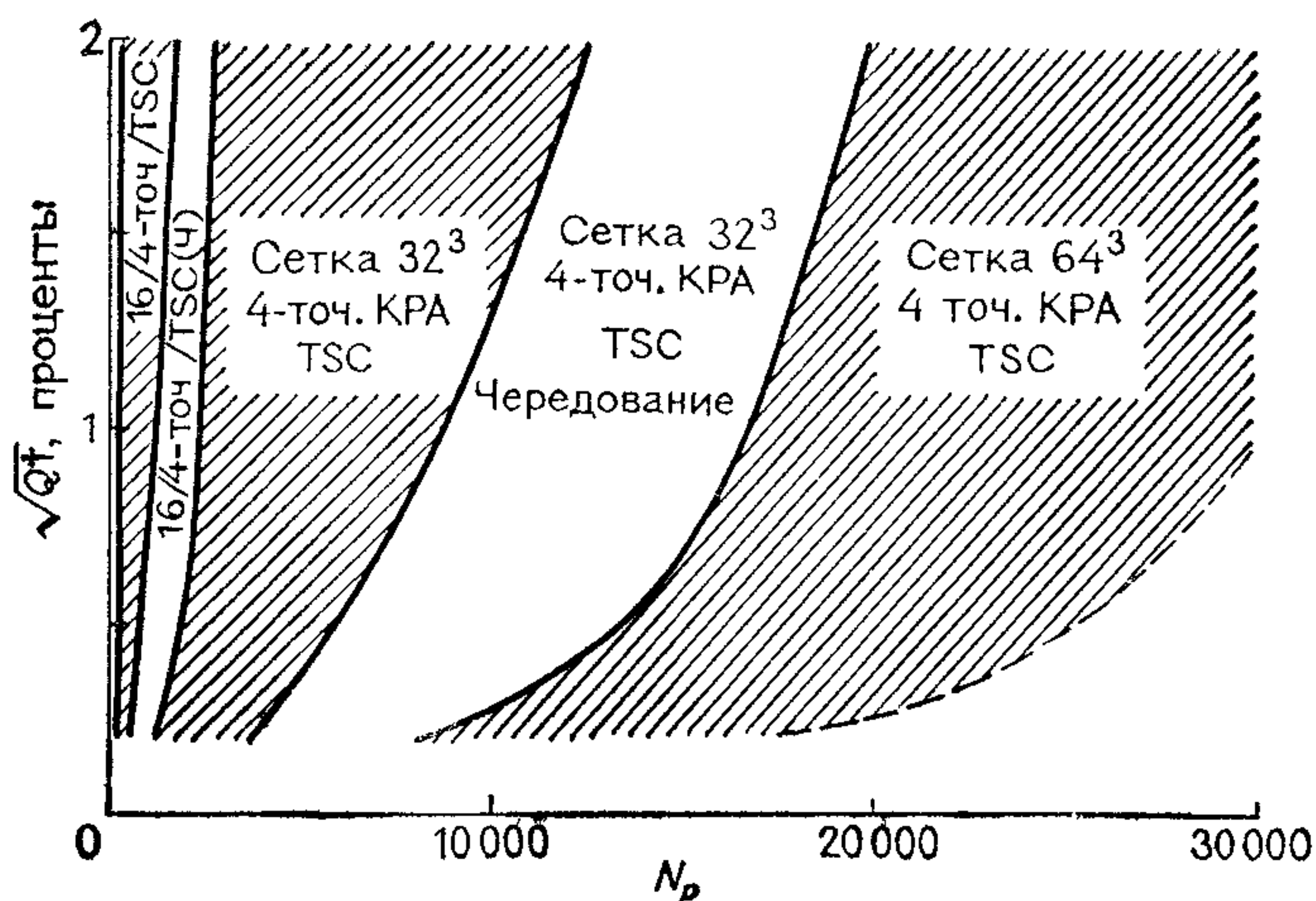


Рис. 8.7 В плоскости качество — число частиц выделены области, где указанные варианты P^3M -алгоритма являются в вычислительном отношении самыми дешевыми. В узкой области, показанной в крайней левой части рисунка, наиболее экономичен метод прямого суммирования (PP). Штриховой кривой в правой части показана изолиния счетного времени, равного 30 с. (Из работы [117], с разрешения Advance Publications Ltd.)

ленного эксперимента по модели P^3M к автоматической процедуре. Все что нужно сделать потенциальному пользователю P^3M -программы, это выбрать определенный компромисс между точностью силы, временем счета и числом частиц (рис. 8.6).

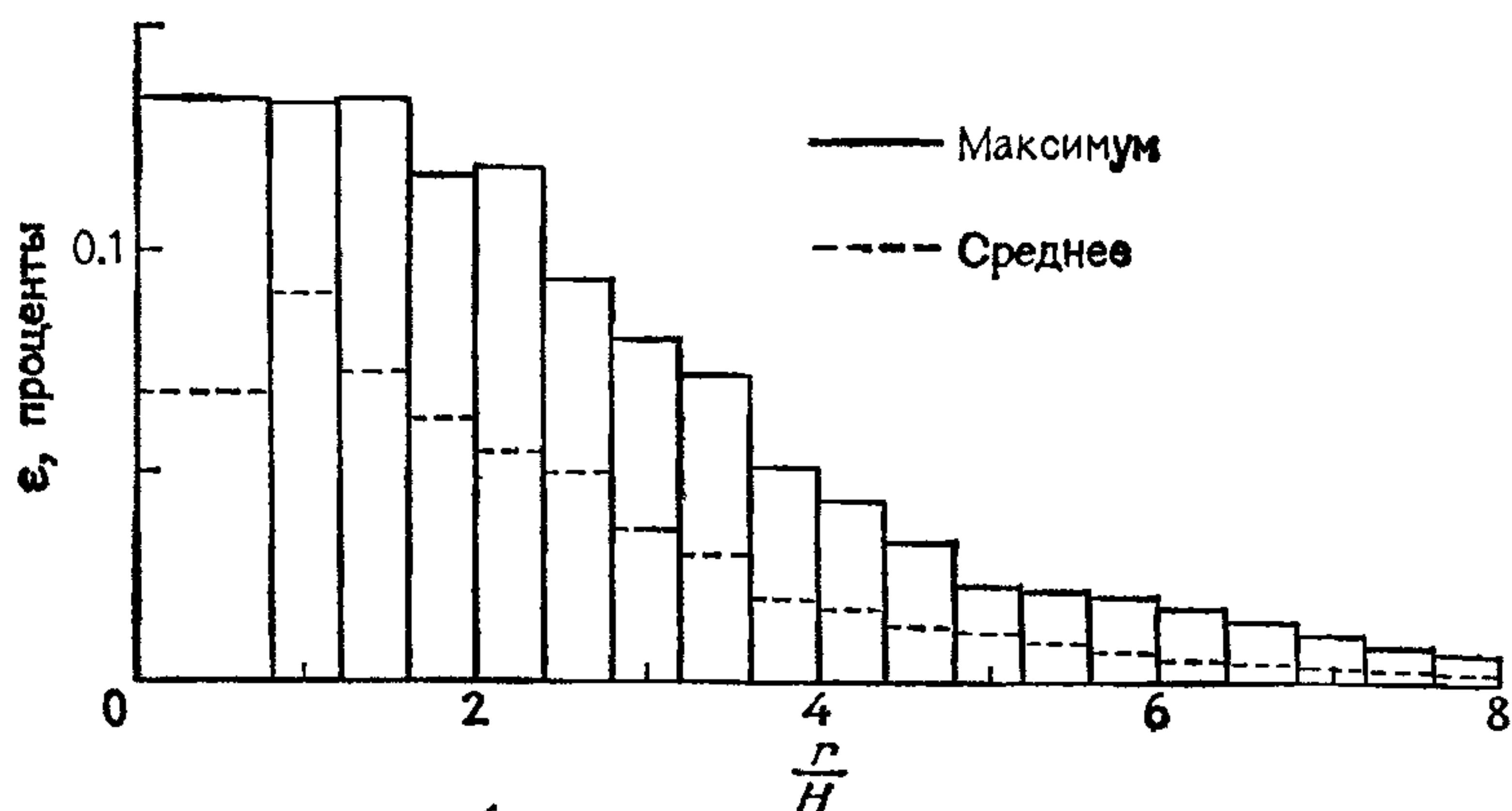


Рис 8.8 Радиальное распределение максимального и среднего значений отклонения ε силы межчастичного взаимодействия от точной силы в варианте $\sqrt{Q^+} = 1\%$, сетка 32^3 , конечно-разностная аппроксимация производной по 4 узлам, $r_e = 3H$, $\alpha = 1.435$, интерполяция TSC (Из работы [117], с разрешения Advance Publications Ltd.)

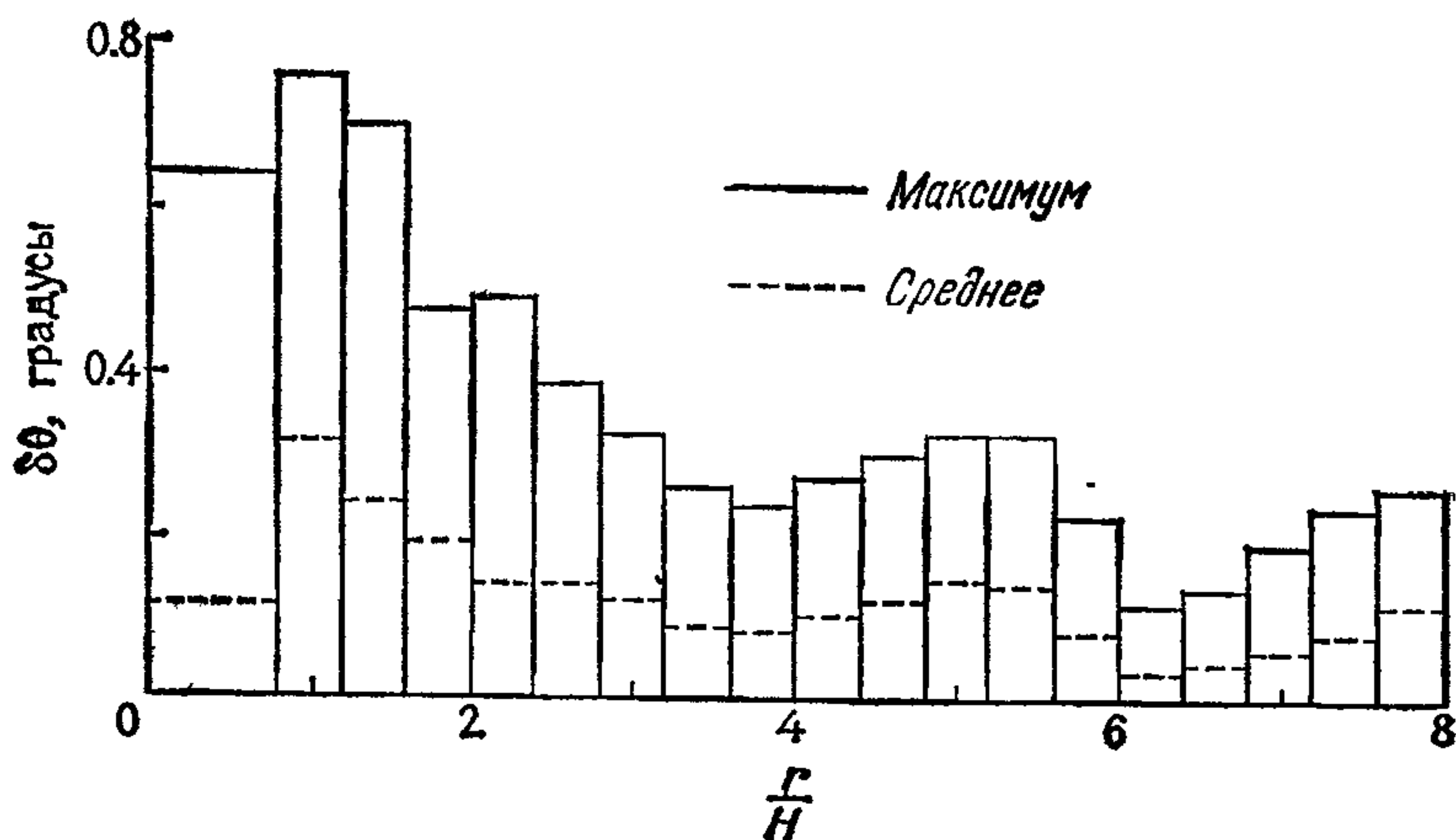


Рис. 8.9. Распределение угловых погрешностей в силе межчастичного взаимодействия для случая, показанного на рис. 8.8 (Из работы [117], с разрешения Advance Publications Ltd)

Тогда соответствующая точка на рис. 8.7 укажет вид РМ-части вычисления силы и число узлов. Соответственно рис. 8.3, 8.4 или 8.5 дает радиус усечения r_e короткодействующей силы, отвечающий выбранному значению $\sqrt{Q^+}$. Выражение (8.22) (или, для схем с чередованием, (8.29)) определяет значения функции влияния.

Практика показала, что $\sqrt{Q^+}$ порядка одного процента в

большинстве случаев обеспечивает достаточную точность. Выбирая $\sqrt{Q^+} = 1\%$ и $N_p = 5000$, из рис. 8.6 определяем, что $T \sim 5$ с. Рис. 8.7 показывает, что при таком выборе необходимы сетка 32^3 , четырехточечный конечно-разностный оператор для производной, интерполяционная функция TSC и не нужно ни одного чередования. Рис. 8.4 дает соответствующее значение $r_e/H (\approx 3)$. При этих параметрах программа P³M3DP обеспечила получение сил межчастичного взаимодействия с погрешностями по величине и углу, показанными на рис. 8.8 и 8.9. Приведенные на рис. 8.8 погрешности в величине силы отнормированы на кулоновскую силу взаимодействия частиц, находящихся на расстоянии $r = H$.

8.7. ПРАКТИЧЕСКИЕ СООБРАЖЕНИЯ

Ценность любого численного алгоритма полностью определяется его пригодностью к практической реализации. Зачастую переход от численного алгоритма к программе моделирования занимает большую часть времени подготовки вычислительного эксперимента. Несмотря на это, практическим вопросам разработки программ обычно уделяется мало внимания. В данном разделе мы рассмотрим некоторые из практических вопросов, с которыми пришлось столкнуться при разработке P³M-алгоритма, воплощенного в программе P³M3DP. Что касается более подробного описания, мы отсылаем читателя к работе [120] и самой программе P³M3DP, которая имеется в библиотеке программ журнала Computer Physics Communication.

8.7.1. Программа

Методика программирования в системе OLYMPUS (гл. 3) включает в себя определенные соглашения на систему обозначений, форму записи и документацию. Скелетная программа CRONUS (рис. 3.4) предопределяет программную архитектуру и обеспечивает «испытательный стенд», на котором можно отрабатывать программные модули. За счет включения в стратегических точках программы обращений к <0.4> EXPERT и логических переключателей вроде NLOMT1 и др. управляющая структура программы отличается гибкостью. Все данные разбиваются на группы в соответствии с правилами системы OLYMPUS (табл. 3.2).

Основными модулями, составляющими цикл по времени (рис. 8.1), являются: (1) расчет PМ-силы, (2) расчет PР-силы и (3) интегрирование уравнений движения. Эти три модуля в значительной степени независимы. Они связаны друг с другом только через основные величины, где основными величинами называется минимальный набор параметров, определяющих ре-

жим работы программы, а именно параметры физической системы, численной схемы и служебные. Вспомогательные величины определяются в отдельных общих блоках для организации внутренних связей между подпрограммами конкретного модуля.

Ход основных вычислений, выполняемых каждым из указанных модулей, определяет управляющая программа временного цикла $\langle 2.1 \rangle$ STEPON. С каждым из модулей, вызываемых в цикле по времени, связаны программы инициализации, управляемые $\langle 1.5 \rangle$ AUXVAL или $\langle 1.6 \rangle$ INITAL, смотря по тому, что необходимо. На стадии инициализации модуля расчета РМ-силы, исходя из введенных N и r_e , формируются значения \hat{G} , α и различные адресные указатели. При инициализации расчета РР-силы формируются таблицы короткодействующей силы, если заданы значения r_e , N_t и вид функциональной зависимости короткодействующей силы. Инициализация уравнений движения заключается в задании начальных положений и импульсов частиц.

Каждый из основных модулей можно в свою очередь подразделить на модули. Например, модуль расчета РМ-силы состоит из частей, связанных с распределением заряда, нахождением потенциала и интерполяцией силы. Модуль нахождения потенциала подразделяется дальше: он вызывает модуль быстрого преобразования Фурье для основной части вычислений. Аналогичная иерархическая разбивка заданий производится по всей программе.

8.7.2. Организация данных

Требуемый для Р³М-программы объем памяти определяется в основном местом для хранения значений функции влияния, плотности заряда и ее гармоник, потенциала и его гармоник, координат частиц, таблиц коррекции короткодействующей силы (и энергии), а также для заголовков цепочек и координат связанных списков. Значения функции влияния и таблицы коррекции короткодействующей силы вычисляются только один раз в программе инициализации и сохраняются в течение всего прогона программы моделирования. Новые координаты частиц $\{x_i, P_i; i \in [1, N_p]\}$ по мере их вычисления записываются на место старых. Остальные же величины, которые нужны только на определенных стадиях расчета, можно располагать так, чтобы они занимали один и тот же участок физической памяти машины, что снижает требования к общему объему памяти программы.

Одним из способов использования одного и того же участка физической памяти под различные переменные может быть употребление фортранного оператора EQUIVALENCE. Однако практика показывает, что легче поддерживать простую, гибкую

и эффективную организацию данных, если использование оператора EQUIVALENCE свести к минимуму. Вместо этого необходимые под сеточные данные большие многомерные массивы явным образом преобразуются в одномерные. Полное описание этого преобразования для программы P3M3DP содержится в работе [120]. Здесь же мы ограничимся иллюстрацией преобразования для сетки заряда-потенциала и для заголовков цепочек и координат связанных списков.

Пусть рабочий массив для сеточных значений заряда-потенциала называется CORE. Если сетка заряда-потенциала образует кубическую решетку из $(N \times N \times N)$ узлов, помеченных вектором $\mathbf{p} = (p_1, p_2, p_3)$, где $0 \leq p_i < N - 1$, значения плотности заряда $\rho(\mathbf{p})$ и потенциала $\phi(\mathbf{p})$ хранятся в элементе массива CORE с номером

$$L_0 + p_1 + p_2 L_2 + p_3 L_3.$$

Здесь L_0 отвечает адресу элемента $(0, 0, 0)$, L_2 и L_3 — приращение адресов между соседними элементами данных в узлах сетки, расположенных вдоль осей y и z соответственно. L_2 и L_3 связаны с размером сетки N :

$$L_2 = N + 2L_B, \quad L_3 = (L_2)^2, \quad (8.57)$$

где $L_B (\geq 0)$ — толщина окаймляющего слоя. Если хватает памяти, то окаймляющий слой можно использовать для того, чтобы исключить проверки на границах при распределении заряда и интерполяции сил. Если говорить о распределении заряда, то заряд каждой частицы раздается в соседние узлы сетки независимо от положения частицы относительно границ. Периодические граничные условия накладываются далее путем прибавления плотностей заряда из окаймляющего слоя к соответствующим плотностям заряда внутри периодической области, отвечающей интервалам $0 \leq p_i \leq N - 1$. Аналогично осуществляется экономия при интерполяции сил, для чего значения потенциала периодически продолжаются в окаймляющий слой перед тем, как интерполировать силы в точки нахождения частиц. Если L_B не равно нулю, то компоненты p_i вектора \mathbf{p} принимают значения в диапазоне $-L_B \leq p_i \leq N - 1 + L_B$, а L_0 должно быть больше, чем $L_B(L_2 + L_3)$.

Фурье-образы функций ρ и ϕ эрмитовы и периодичны и вследствие этого могут храниться в том же месте, что и значения самих функций. Схема расположения в памяти значений фурье-образов и используемый порядок перезаписи обсуждаются далее в разд. 8.7.4.

Как только расчет PМ-сил закончен, значения потенциалов в массиве CORE уже не нужны. Поэтому рабочий массив целого типа ICORE для таблиц заголовков цепочек и связанных списков можно с помощью EQUIVALENCE разместить в памяти,

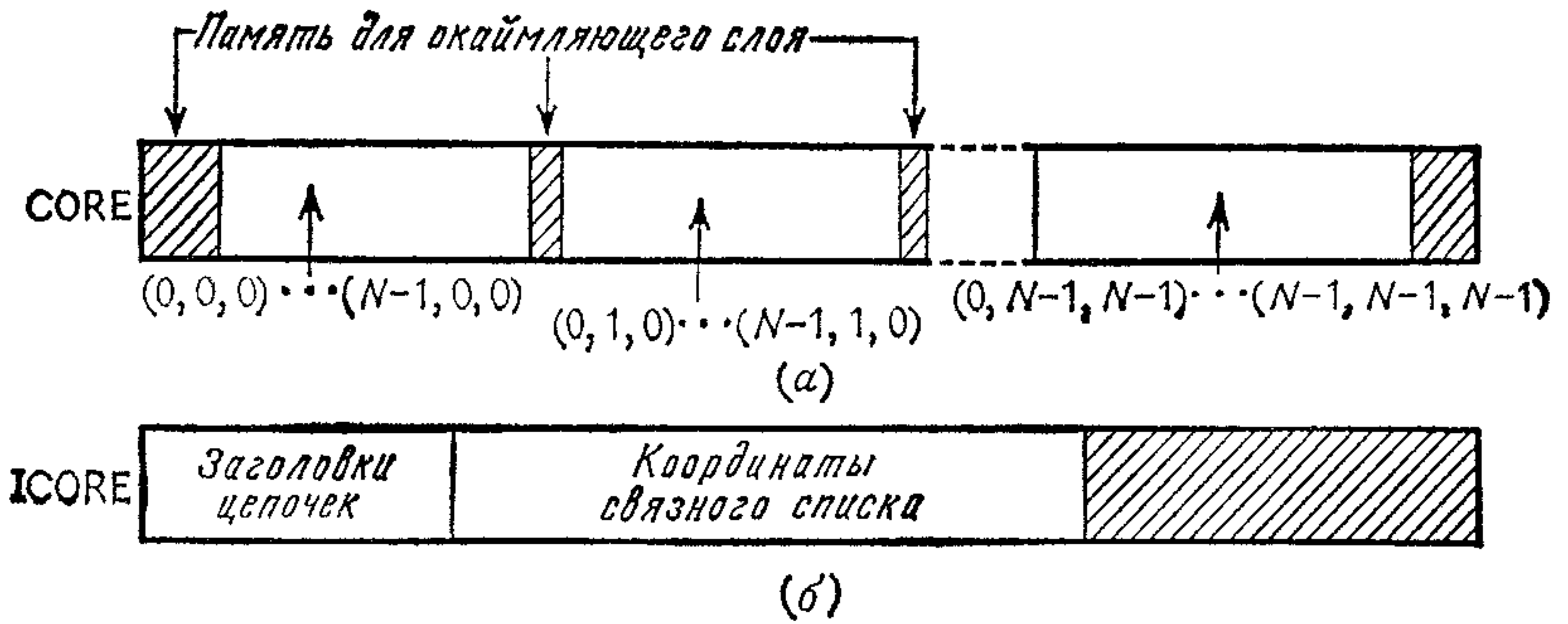


Рис 810 Распределение памяти рабочего массива CORE для хранения (а) значений заряда и потенциала и (б) заголовков цепочек и координат связанных списков

отведенной массиву CORE, что даст дополнительную экономию памяти. На рис. 810(б) схематически показано расположение в памяти таблицы заголовков цепочек и связанных списков. Таблицы заголовков цепочек для трехмерной $(M \times M \times M)$ цепочечной сетки размещаются в памяти точно так же, как значения заряда и потенциала. В отсутствие окаймления величина $НОС(q)$, где $q = (q_1, q_2, q_3)$ и $0 \leq q_i \leq M - 1$, хранится в элементе массива ICORE с номером

$$L_c = M_0 + q_1 + q_2 M + q_3 M^2,$$

где $L_c > 0$, а M_0 — адрес элемента $(0, 0, 0)$.

Координаты связанного списка $\{LL(i); i \in [1, N_p]\}$ хранятся в ICORE сразу за M^3 значениями заголовков цепочек. Координаты связанного списка расположены в том же порядке, что и основные координаты (положения и импульсы), так что i -я частица имеет положение x_i , импульс P_i и связующую координату $LL(i)$.

8.7.3. Распределение заряда и интерполяция силы

Во всех программах моделирования методом частиц хорошо пользоваться в главном счетном цикле удобными в вычислительном отношении безразмерными единицами (см. гл. 2). В системе единиц, где $H = 1$, интерполяционная функция (см. (8.8)) принимает вид

$$W(x) = w(x) w(y) w(z), \tag{8.58}$$

где для схемы TSC

$$w(x) = \Pi(x) * \Lambda(x). \tag{8.59}$$

Для практических целей полезней выразить (8.58) и (8.59) через долю заряда, раздаваемого в каждый узел сетки. Если ча-

стица расположена в ячейке p и ее положение относительно узла сетки в центре этой ячейки есть $x' = x - x_p$, то доля заряда, приписываемая сеточному узлу $p + t$, равняется

$$W_{p+t} = w_{t_1}(x') w_{t_2}(y') w_{t_3}(z'), \quad (8.60)$$

где на основании (5.86) и (5.87)

$$w_t(x') = \begin{cases} \frac{1}{2} \left(\frac{1}{2} + x' \right)^2, & t = 1, \\ \frac{3}{4} - x'^2, & t = 0, \\ \frac{1}{2} \left(\frac{1}{2} - x' \right)^2, & t = -1, \\ 0 & \text{в других случаях.} \end{cases} \quad (8.61)$$

Пусть $v_t = 2w_t$ и пусть q_s составляет одну восьмую часть заряда частицы сорта s . Тогда алгоритм распределения заряда по схеме TSC для каждой частицы сорта s имеет следующий вид:

вычисление v_{t_i} ; $i \in [1, 3]$, $t \in [-1, 1]$ (15 операций)

```

do for  $t_1 \in [-1, 1]$ 
   $T_1 = q_s v_{t_1}(x')$  (3 операции)
  do for  $t_2 \in [-1, 1]$ 
     $T_2 = T_1 v_{t_2}(y')$  (9 операций)
    do for  $t_3 \in [-1, 1]$ 
       $T_3 = T_2 v_{t_3}(z')$  (27 операций)
   $\rho(p + t) = \rho(p + t) + T_3$  (27 операций)

```

Числа в круглых скобках указывают количество арифметических операций с вещественными числами, приходящихся на одну частицу. Складывая эти величины, получаем, что для этапа распределения заряда количество арифметических операций с вещественными числами, приходящихся на частицу за шаг по времени, составляет $3 \times 27 = 81$.

Алгоритм интерполяции силы отличается от алгоритма распределения заряда тем, что в самом внутреннем цикле распределение заряда заменяется интерполяцией, т. е. строка

$$\rho(p + t) = \rho(p + t) + T_3 \quad (8.62)$$

заменяется на

$$P_{i,s} = P_{i,s} + T_3 E(p + t), \quad (8.63)$$

где $P_{i,s}$ обозначает импульс i -й частицы сорта s , а $E(p + t)$ — электрическое поле (построенное посредством дифференцирования значений потенциала) в сеточном узле $p + t$.

8.7.4. Нахождение потенциала

Решение уравнения для потенциала методом преобразования Фурье (разд. 6.5.2) состоит из трех этапов:

1. Преобразование $\rho \supset \hat{\rho}$.
2. Вычисление $\hat{\phi} = \hat{G}\hat{\rho}$.
3. Преобразование $\hat{\phi} \supset \phi$.

Трехмерное преобразование Фурье плотности заряда в случае $N = 1$ имеет вид

$$\hat{\rho}(\mathbf{k}) = \sum_{\mathbf{p}=0}^{N-1} \rho(\mathbf{p}) e^{-i\mathbf{k}\cdot\mathbf{p}}. \quad (8.64)$$

Выражение (8.64) можно записать в явном виде через целые l ($\mathbf{k} = 2\pi\mathbf{l}/N$), отвечающие номерам гармоник:

$$\hat{\rho}(l_1, l_2, l_3) = \sum_{p_3=0}^{N-1} e^{-i2\pi l_3 p_3/N} \times \\ \times \left[\sum_{p_2=0}^{N-1} e^{-i2\pi l_2 p_2/N} \left(\sum_{p_1=0}^{N-1} e^{-i2\pi l_1 p_1/N} \rho(p_1, p_2, p_3) \right) \right]. \quad (8.65)$$

Член в круглых скобках в (8.65) соответствует одномерному дискретному преобразованию Фурье (FFT) плотности заряда в направлении оси x , т. е.

$$\tilde{\rho}(l_1, p_2, p_3) = \sum_{p_1=0}^{N-1} e^{-i2\pi l_1 p_1/N} \rho(p_1, p_2, p_3). \quad (8.66)$$

Член в квадратных скобках описывает набор одномерных преобразований в направлении оси y :

$$\tilde{\rho}(l_1, l_2, p_3) = \sum_{p_2=0}^{N-1} e^{-i2\pi l_2 p_2/N} \tilde{\rho}(l_1, p_2, p_3). \quad (8.67)$$

Самая внешняя сумма в (8.65) дает преобразования в направлении оси z :

$$\hat{\rho}(l_1, l_2, l_3) = \sum_{p_3=0}^{N-1} e^{-i2\pi l_3 p_3/N} \tilde{\rho}(l_1, l_2, p_3). \quad (8.68)$$

Алгоритмическое описание формулы (8.66) имеет вид:

```

do for  $p_3 = 1, N - 1$ 
do for  $p_2 = 1, N - 1$ 
пересылка  $\{\rho(p_1, p_2, p_3); p_1 = 1, N - 1\}$  на входной буфер FFT
получение значений фурье-образов
пересылка выходного буфера FFT на то место, где были
 $\{\rho(p_1, p_2, p_3); p_1 = 1, N - 1\}$ 

```

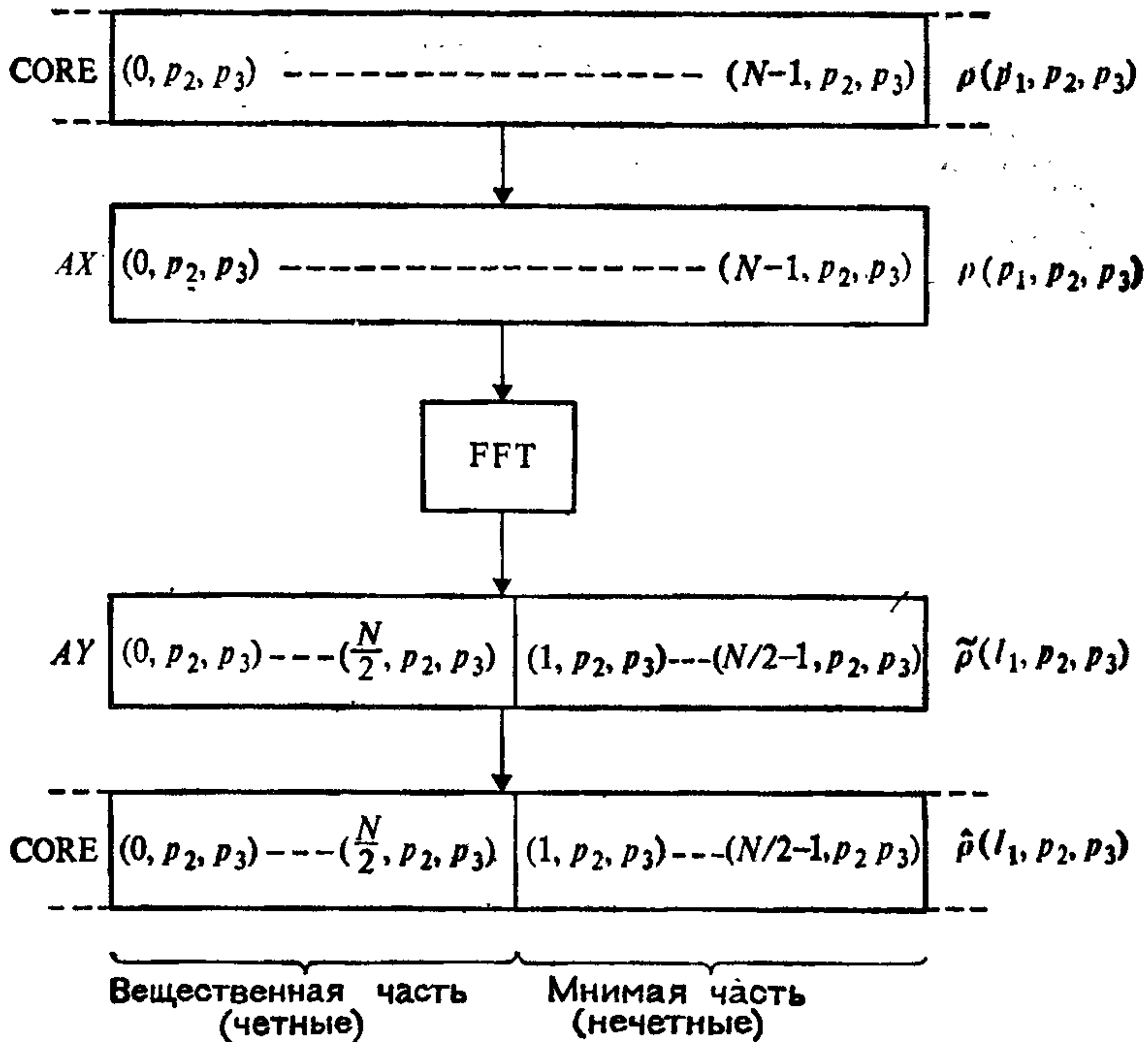



Рис. 8.11. Строка значений данных плотности заряда пересылается на входной буфер AX программы FFT. Результат преобразования, находящийся в буфере AY, пересылается обратно в рабочий массив CORE, причем значения преобразования занимают те же самые адреса памяти, что и исходные величины

На рис. 8.11 схематично показано это преобразование для одной строки. Из рабочего массива CORE величины пересылаются в AX — входной буфер FFT. По содержимому AX модуль FFT формирует в выходном буфере AY результат преобразования. Первые $N/2 + 1$ элементов AY содержат вещественные (четные) части гармоник для $l_1 \in [0, N/2]$, а следующие $N/2 - 1$ элементов — мнимые (нечетные) части гармоник для $l_1 \in [1, N/2 - 1]$. Поскольку мнимая часть является нечетной, все мнимые части гармоник $\tilde{\rho}(0, p_2, p_3)$ и $\tilde{\rho}(N/2, p_2, p_3)$ равны нулю и поэтому не хранятся.

Процедуры вычисления (8.67) и (8.68) имеют тот же вид, что и для (8.66). Для преобразования (8.67) на буфер FFT пересылаются данные из узлов сетки заряда-потенциала, расположенных в направлении оси y , выполняется преобразование, а затем выходной буфер FFT переписывается на место исходных данных. Те же действия повторяются для линий узлов в направлении оси z . В итоге на том месте памяти, где были

значения плотности, получаются гармоники $\{\hat{\rho}\}$. Каждая гармоника хранится в виде восьми частей:

$$\hat{\rho}(l_1, l_2, l_3) = (\hat{\rho}_{\text{ччч}} + \hat{\rho}_{\text{чнн}} + \hat{\rho}_{\text{нчн}} + \hat{\rho}_{\text{ннч}}) + \\ + i(\hat{\rho}_{\text{ччн}} + \hat{\rho}_{\text{чнч}} + \hat{\rho}_{\text{нчч}} + \hat{\rho}_{\text{ннн}}). \quad (8.69)$$

Индексы „ч“ и „н“ обозначают четные и нечетные по индексам l_1, l_2, l_3 части $\hat{\rho}$. Например, $\hat{\rho}_{\text{нчн}}$ нечетна по индексам l_1 и l_3 и четна по l_2 . Гармоники, находящиеся вне интервала $0 \leq l_i \leq N/2$, можно получить из (8.69) с учетом свойств симметрии и периодичности каждой компоненты.

Расположение в памяти указанных частей гармоник $\hat{\rho}(l_1, l_2, l_3)$ определяется соотношением

$$L_0 + l_1^* + l_2^* L_2 + l_3^* L_3,$$

где

$$l_i^* = \begin{cases} l_i & \text{для четных по } l_i \text{ частей,} \\ l_i + N/2 & \text{для нечетных по } l_i \text{ частей} \end{cases}$$

и $0 \leq l_i \leq N/2$ для четных частей и $0 < l_i < N/2$ для нечетных.

Те же восемь частей гармоник потенциала (нечетные и четные по каждому из их индексов) находятся умножением соответствующих частей $\hat{\rho}$ на значения функции влияния G и запоминаются в рабочем массиве CORE на месте элементов $\hat{\rho}$. Именно на этой стадии накапливается сеточная часть потенциальной энергии (см. (12.58) — (12.60)). Значения потенциала получаются с помощью обратного преобразования, проводимого (включая запись результатов) в порядке, обратном использованному при нахождении гармоник плотности заряда по значениям плотности заряда.

Распишем несколько подробнее три этапа алгоритма нахождения потенциала, сформулированные в начале данного подраздела.

1. Преобразование всех N^2 строк данных, параллельных оси x :

$$\rho(p_1, p_2, p_3) \Rightarrow \begin{cases} \bar{\rho}_{\text{ч}}(l_1, p_2, p_3), & 0 \leq l_1 \leq N/2, \\ \bar{\rho}_{\text{н}}(l_1, p_2, p_3), & 0 < l_1 < N/2. \end{cases}$$

2. Преобразование всех N^2 строк данных, параллельных оси y :

$$\bar{\rho}_{\text{ч}}(l_1, p_2, p_3) \Rightarrow \begin{cases} \bar{\rho}_{\text{чч}}(l_1, l_2, p_3), & 0 \leq l_2 \leq N/2, \\ \bar{\rho}_{\text{чн}}(l_1, l_2, p_3), & 0 < l_2 < N/2, \end{cases}$$

$$\bar{\rho}_{\text{н}}(l_1, p_2, p_3) \Rightarrow \begin{cases} \bar{\rho}_{\text{нч}}(l_1, l_2, p_3), & 0 \leq l_2 \leq N/2, \\ \bar{\rho}_{\text{нн}}(l_1, l_2, p_3), & 0 < l_2 < N/2. \end{cases}$$

3. Преобразование всех N^2 строк данных, параллельных оси z :

$$\tilde{\rho}_{\text{чч}}(l_1, l_2, p_3) \supset \begin{cases} \hat{\rho}_{\text{ччч}}(l_1, l_2, l_3), & 0 \leq l_3 \leq N/2, \\ \hat{\rho}_{\text{ччн}}(l_1, l_2, l_3), & 0 < l_3 < N/2, \end{cases}$$

и т. д.

4. Вычисление гармоник потенциала:

$$\begin{aligned} \hat{\Phi}_{\text{ччч}}(l_1, l_2, l_3) &= \hat{G}(l_1, l_2, l_3) \hat{\rho}_{\text{ччч}}(l_1, l_2, l_3), & 0 \leq (l_1, l_2, l_3) \leq N/2, \\ \hat{\Phi}_{\text{ччн}}(l_1, l_2, l_3) &= \hat{G}(l_1, l_2, l_3) \hat{\rho}_{\text{ччн}}(l_1, l_2, l_3), & 0 \leq l_1, l_2 \leq N/2, \\ & & 0 < l_3 < N/2 \end{aligned}$$

и т. д.

5. Преобразование всех N^2 строк данных, параллельных оси z :

$$\left. \begin{aligned} \hat{\Phi}_{\text{ччч}}(l_1, l_2, l_3) \\ \hat{\Phi}_{\text{ччн}}(l_1, l_2, l_3) \end{aligned} \right\} \supset \tilde{\Phi}_{\text{чч}}(l_1, l_2, p_3), \quad 0 \leq p_3 \leq N - 1$$

и т. д.

6. Преобразование всех N^2 строк данных, параллельных оси y :

$$\left. \begin{aligned} \tilde{\Phi}_{\text{чч}}(l_1, l_2, p_3) \\ \tilde{\Phi}_{\text{чн}}(l_1, l_2, p_3) \end{aligned} \right\} \supset \tilde{\Phi}_{\text{ч}}(l_1, p_2, p_3), \quad 0 \leq p_2 \leq N - 1,$$

$$\left. \begin{aligned} \tilde{\Phi}_{\text{нч}}(l_1, l_2, p_3) \\ \tilde{\Phi}_{\text{нн}}(l_1, l_2, p_3) \end{aligned} \right\} \supset \tilde{\Phi}_{\text{н}}(l_1, p_2, p_3), \quad 0 \leq p_2 \leq N - 1.$$

7. Преобразование всех N^2 строк данных, параллельных оси x :

$$\left. \begin{aligned} \tilde{\Phi}_{\text{ч}}(l_1, p_2, p_3) \\ \tilde{\Phi}_{\text{н}}(l_1, p_2, p_3) \end{aligned} \right\} \supset \Phi(p_1, p_2, p_3), \quad 0 \leq p_1 \leq N - 1.$$

8.7.5. Короткодействующая сила

Основные практические трудности, связанные с вычислением короткодействующей силы, вызывает определение местоположения соседних частиц и организация такого порядка действий, при котором силы парного взаимодействия не приходится вычислять дважды. Поскольку эти два вопроса были уже детально рассмотрены в разд. 8.4, мы на них здесь больше останавливаться не будем.

Значения короткодействующих сил, которые закладываются в силы межчастичного взаимодействия, получаются на основании формулы (8.7). Отметим, что нужно учитывать только силы взаимодействия изолированной пары частиц. Силы, обусловленные периодическими изображениями пары частиц, включены в РМ-часть расчета силы. В таблицах хранятся значения $f^{\text{sr}}(r)/r$ (см. разд. 8.4), а не значения $f^{\text{sr}}(r)$, чем исключаются деления на r , и, кроме того, эта функция табулируется с постоянным шагом по r^2 , чтобы избежать вычисления квадратного корня $\sqrt{(r^2)}$ в самом внутреннем цикле суммирования сил пар-

ного взаимодействия. Табличные значения \mathcal{F} задаются по формуле

$$\mathcal{F}(r) = \frac{f^{sr}(r)}{r} = \frac{f_{(r)}^{\text{tot}} - R(r)}{r}, \quad (8.70)$$

где

$$R(r) = \mathbf{R}(\mathbf{x}) \cdot \hat{\mathbf{r}} = \frac{\hat{\mathbf{r}}}{4\pi\epsilon_0} \cdot \int d\mathbf{x}' \int d\mathbf{x}'' S(\mathbf{x}'') S(\mathbf{x}' - \mathbf{x}'') \frac{\mathbf{x}' - \mathbf{x}''}{|\mathbf{x}' - \mathbf{x}''|^3}. \quad (8.71)$$

Для частицы с формой $S2$ S задается выражением (8.3). Подстановка (8.3) в (8.71) после некоторых простых, но длинных выкладок дает

$$R(r) = \frac{1}{4\pi\epsilon_0} \times \begin{cases} \frac{1}{35a^2} (224\xi - 224\xi^3 + 70\xi^4 + 48\xi^5 - 21\xi^6), & 0 \leq \xi \leq 1, \\ \frac{1}{35a^2} \left(\frac{12}{\xi^2} - 224 + 896\xi - 840\xi^2 + 224\xi^3 + 70\xi^4 - 48\xi^5 + 7\xi^6 \right), \\ \frac{1}{r^2}, & \xi \geq 2, \end{cases} \quad (8.72)$$

где

$$\xi = 2r/a. \quad (8.73)$$

Табличные значения короткодействующей части потенциальной энергии парного взаимодействия $\varphi^{sr}(r)$ вычисляются аналогичным образом, исходя из разности между полной потенциальной энергией парного взаимодействия φ и потенциальной энергией парного взаимодействия φ^m , отвечающей сеточной эталонной силе межчастичного взаимодействия \mathbf{R} (см. разд. 12.3.4), по формуле

$$\varphi^{sr}(r) = \varphi(r) - \varphi^m(r). \quad (8.74)$$

Здесь для сеточной эталонной силы взаимодействия частиц с формой $S2$

$$\varphi^m = \frac{1}{4\pi\epsilon_0} \times \begin{cases} \frac{1}{70a} [208 - 112\xi^2 + 56\xi^4 - 14\xi^5 - 8\xi^6 + 3\xi^7], & 0 \leq \xi \leq 1, \\ \frac{1}{70a} \left[\frac{12}{\xi} + 128 + 224\xi - 448\xi^2 + 280\xi^3 - 56\xi^4 - 14\xi^5 + 8\xi^6 - \xi^7 \right], & 1 \leq \xi \leq 2, \\ \frac{1}{r}, & \xi \geq 2. \end{cases} \quad (8.75)$$

В задачах моделирования ионных систем для предотвращения неустойчивостей при интегрировании по времени (возникающих при $\omega DT > 2$; см. гл. 4), силы межчастичного взаимодействия ограничиваются сверху таким образом, чтобы $\omega_{\max} DT < 2$ (см. также разд. 12.2.2 и 12.3.2). Если шаг по времени DT выбран надлежащим образом, то вероятность того, что два иона сблизятся на расстояние, меньшее радиуса b , в пределах которого сила установлена равной ее предельно допустимому значению, будет ничтожна. В качестве простого средства контроля за тем, не подходят ли частицы друг к другу ближе, чем на b , используется задание табличных значений φ^{st} до некоторого физически недостижимого значения. Тогда, если меньшие b расстояния все же встретятся, они приведут к «выбросам» в энергии, что будет указывать на то, что не закончились начальные переходные процессы, или на необходимость уменьшить шаг по времени.

8.7.6. Выбор параметров

Последнее, что необходимо сделать прежде, чем от программы можно будет получать физически полезные результаты, это задать исходные данные для моделирования. Неудачный выбор параметров может стать причиной неэффективности программы или, что еще хуже, привести к физически бессмысленным результатам. В настоящем разделе мы опишем процедуру выбора и оценки параметров, которая, как мы убедились, полезна при подготовке P³MЗDP к моделированию ионных жидкостей.

Задание входных данных для программы включает в себя пять основных стадий: необходимо описать характеристики физической системы, численную схему, начальное состояние, выходные наборы данных, управление и диагностику. Исследуемое вещество задается количеством сортов частиц и определенными атрибутами этих сортов: зарядом, массой, видом (короткодействующей) силы отталкивания и т. д. Количество вещества описывается числом ионов каждого сорта, а объем пространства, которое оно занимает, определяется размером расчетной области: в совокупности число ионов и объем области задают среднюю плотность. Численная схема задается размером сетки N , радиусом усечения a , эффективным радиусом усечения r_e , шагом по времени DT , а также порогом силы для таблиц короткодействующей силы.

Подходящая комбинация физических и численных параметров получается следующим образом:

1. Задаем вещество:

— количество сортов

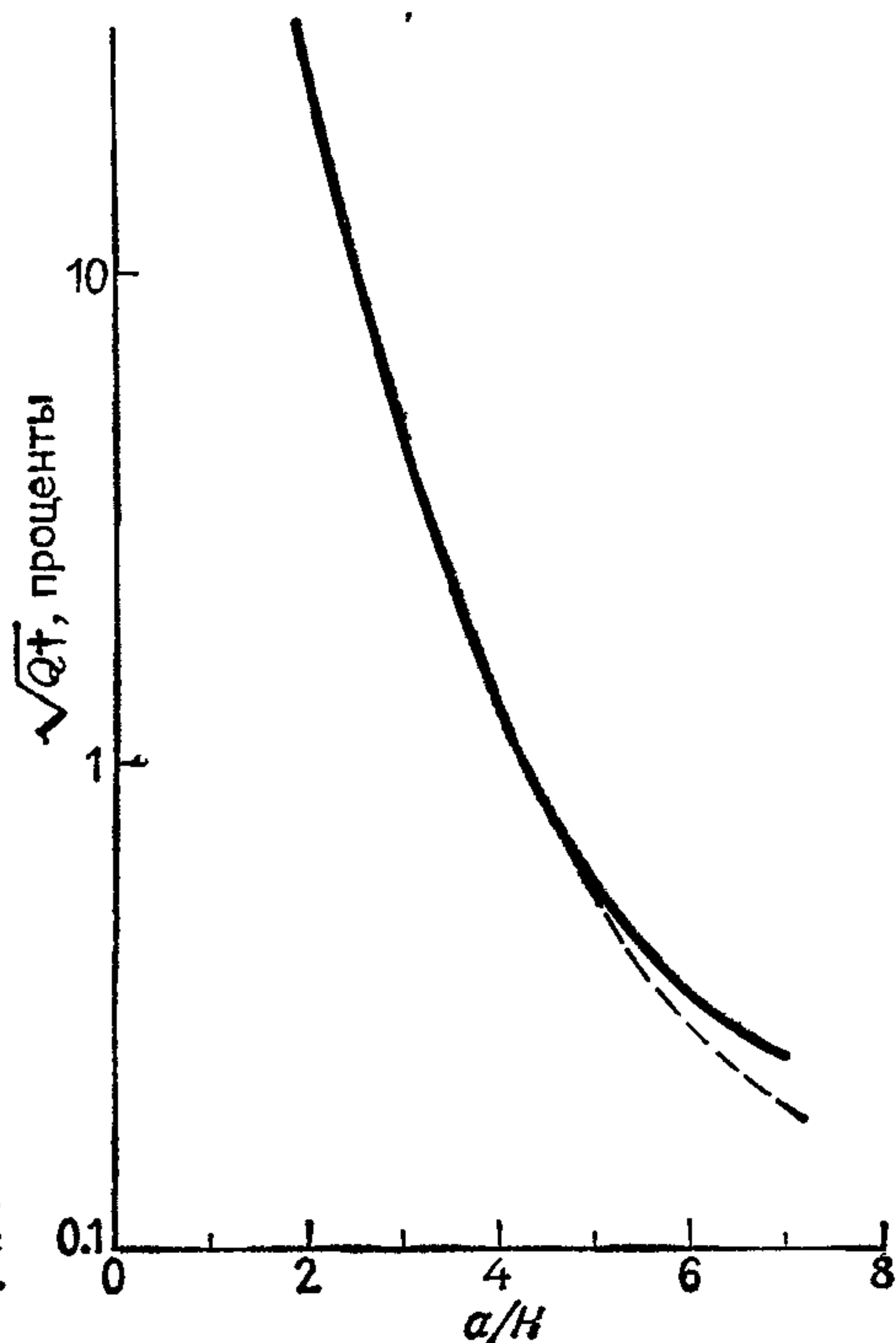


Рис. 8.12. Мера качества $\sqrt{Q^+}$ как функция радиуса усечения a . Штриховая кривая получена с сеткой 16^3 , а сплошная кривая — с сеткой 32^3 .

— заряд и массу ионов

— закон короткодействующей силы: область действия некулоновской части обуславливает минимальный радиус усечения r_e^0 .

2. Выбираем среднюю плотность n_0 .

3. Выбираем значение $\sqrt{Q^+}$ и по нему, исходя из рис. 8.12, (a/H). Задаем отношение r_e/a для обрывания таблиц коррекции короткодействующей силы на радиусе (r_e/H), где табличные значения (выраженные в единицах кулоновской силы при $r = H$) малы по сравнению с $\sqrt{Q^+}$, лучше на том N , которое немного превышает целую часть от (r_e/H). Типичные значения $0.7 \leq r_e/a \leq 1.0$.

4. Шаг 3 дал (r_e/H). Из рис. 8.13 определяем полное число частиц N_p и размер сетки N . Не исключено, что для достижения приемлемого компромисса «цена — качество» шаги 3 и 4 придется повторить несколько раз.

5. Предыдущие шаги конкретизируют значения r_e^0 , (r_e/H), n_0 , N_p и N , и поэтому мы можем теперь вычислить

$$r_e = \left(\frac{r_e}{H}\right) \frac{1}{N} \left(\frac{N_p}{n_0}\right)^{1/3}.$$

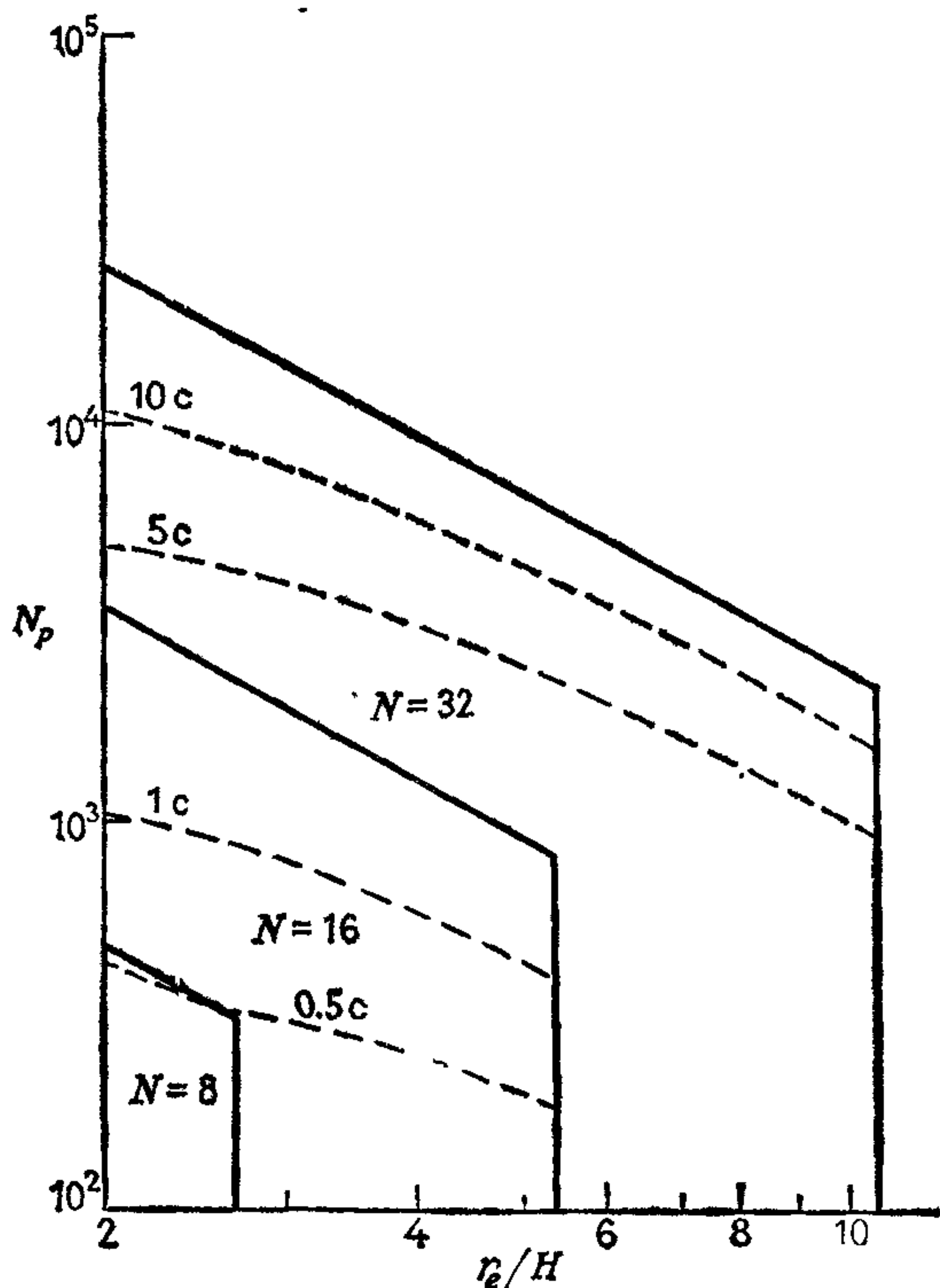


Рис. 8.13. Изолинии счетного времени (штриховые кривые) и области в плоскости число частиц — эффективный радиус усечения, где указанные размеры сетки являются наиболее подходящими (сплошные линии). Вертикальные участки сплошных линий обусловлены ограничением $3(r_e/H) < N$. Времена счета даны в секундах работы машины IBM 360/195 (показаны на кривых как 10 с, 5 с и т. д.) и включают в себя время на вычисление давления, температуры, энергии и функции радиального распределения.

- Если оказалось, что $r_e < r_e^0$, увеличиваем по отдельности или вместе (r_e/a) , a и N_p до тех пор, пока не получим $r_e \geq r_e^0$.
6. Вычисляем $L = (N_p/n_0)^{1/3}$ и изменяем пропорционально число частиц каждого сорта так, чтобы полное число частиц равнялось выбранному N_p и сохранилось правильное соотношение между сортами.
 7. Выбираем шаг DT , который обеспечивает устойчивость и точность схемы с перешагиванием интегрирования по времени. Критерий устойчивости $\omega DT < 2$, где ω — локальная частота колебаний

$$\omega = \left| 2 \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \frac{\partial F}{\partial r} \right|^{1/2},$$

вводит наиболее жесткие ограничения для легких ионов и малых расстояний. Большие значения ω , которые получаются при очень малых расстояниях между ионами, предотвращаются заданием постоянных сил межчастичного взаимодействия при малых r (см. разд. 12.2.2 и 12.3.2).

Выбор начальных условий и выходного набора данных определяется конкретными условиями проводимого эксперимента. Эти вопросы обсуждаются далее в последних главах. Одной из

проблем, связанных с инициализацией плотных ионных систем, с которой, как правило, не приходится встречаться в бесстолкновительных расчетах РМ-методом, является проблема начальных переходных процессов. В плотных системах потенциальная энергия и кинетическая энергия сравнимы. Способ формирования начального состояния системы, задающий положения ионов и их импульсы (с заданными первым и вторым моментами функции распределения), обладает тем свойством, что система обычно получается гораздо горячее, чем требуется. Эта трудность преодолевается перемасштабированием импульсов на протяжении первых одной-двух сотен шагов по времени таким образом, чтобы получались требуемые моменты скорости (т. е. несдвинутое максвелловское распределение с температурой T).

Простые проверки пригодности комбинации входных параметров, отобранных вышеуказанным методом, проводятся в тестовых прогонах, выполняемых с системой, состоящей только из двух ионов, — это дает результаты, которые можно проверить аналитически. Весьма ценным оказывается графический вывод результатов, поскольку он позволяет быстро проконтролировать отсутствие ошибок в виде зависимости силы от расстояния, в расщеплении силы, в начальном распределении, а также в траектории. Полезную перекрестную проверку на пригодность эффективного радиуса усечения дают погрешности аппроксимации энергии и давления (см. разд. 12.3.4(4)). Текущий контроль хода вычислительных экспериментов ведется по полному импульсу, моменту количества движения и энергии. Эти характеристики отражают погрешности и округления и аппроксимации, и, кроме того, их можно использовать для регистрации переходных процессов, если таблицы коррекции энергии короткодействия сформированы указанным выше способом (разд. 8.7.5).

Экспериментально установлено, что на IBM 360/195 время счета программы P3M3DP определяется (с точностью до 10 %) формулой

$$T = (50 + 2N_n) 10^{-5} N_p + \beta(N) \text{ секунд}, \quad (8.76)$$

где N_n — число частиц в сфере радиуса r_e , N_p — полное число частиц, $\beta(N)$ — время нахождения потенциала ($\beta = 0.3, 2$ и 12 с для $N = 16, 32$ и 64 соответственно). T включает в себя время на расчет динамики и время на вычисление давления, температуры, энергий и функции радиального распределения. Для типичных параметров, применяемых в моделировании ионных жидкостей, получено время счета порядка 1.5 с на тысячу частиц. Например, замеры времени счета [186] дали около 6 с для 4096 частиц и 33 с для 21952 частиц. Это стоит сравнить с получаемым в методах прямого суммирования временем счета порядка 1 с на 256 частиц, что в пересчете на 21952 частицы составляет приблизительно 2 часа!

Моделирование плазмы

9.1. ВВЕДЕНИЕ

Плазма представляет собой высокотемпературный полностью ионизированный газ, который может рассматриваться как совокупность положительных ионов и отрицательных электронов, взаимодействующих посредством самосогласованных электрического и магнитного полей (\mathbf{E} и \mathbf{B}). Поля связаны с зарядом ρ и плотностью тока \mathbf{j} уравнениями Максвелла. Для вакуума имеем

$$\operatorname{div} \mathbf{B} = 0, \quad (9.1a)$$

$$\operatorname{rot} \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \quad (9.1b)$$

$$\operatorname{div} \mathbf{E} = \rho / \epsilon_0, \quad (9.1c)$$

$$\operatorname{rot} \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}, \quad (9.1d)$$

где $c = (\epsilon_0 \mu_0)^{-1/2}$ — скорость света в вакууме.

Сила, действующая на заряд q , движущийся в этих полях со скоростью \mathbf{v} , равна

$$\mathbf{F} = q [\mathbf{E} + \mathbf{v} \times \mathbf{B}], \quad (9.2a)$$

а движение заряда определяется законом Ньютона

$$m \frac{d\mathbf{v}}{dt} = \mathbf{F}. \quad (9.2b)$$

Если в малом объеме пространства V (например, в ячейке сетки) находятся N частиц с зарядами q_i и скоростями \mathbf{v}_i ($i = 1, \dots, N$), то плотности заряда и тока могут быть вычислены путем суммирования:

$$\rho = \sum_{i=1}^N q_i / V, \quad (9.3a)$$

$$\mathbf{j} = \sum_{i=1}^N q_i \mathbf{v}_i / V. \quad (9.3b)$$

Полное трехмерное моделирование описанных систем [уравнения (9.1)—(9.3)] находится на грани возможностей таких компьютеров, как IBM 360/195 или CDC-7600 [61, 239]. Однако при продолжительности расчета одного временного шага, равной нескольким минутам, машинное время, требуемое для приемлемого по затратам моделирования, слишком велико для большинства проблем. Для таких расчетов необходимы компьютеры с параллельными процессорами, такие, как CRAY-1 с производительностью, по крайней мере в 10 раз превосходящей скорость IBM 360/195 (см. [185, 192]).

Поэтому мы вынуждены исследовать способы упрощения упомянутых выше уравнений. Хотя возможны различные приближения, однако здесь мы изучим два предельных случая — магнитогидродинамику (МГД) и электростатическую модель (ЭС). В обоих случаях рассматриваются временные масштабы, существенно превосходящие время прохождения света через систему, и не учитывается второй член в правой части (9.1b), пропорциональный c^{-2} . В этом подходе отсутствуют электромагнитные волны и подразумевается, что поля E и B мгновенно подстраиваются под заданное распределение зарядов и токов. Теперь рассмотрим эти приближения подробно.

9.1.1. Магнитная гидродинамика

Рассмотрим плазму с малой плотностью в низкочастотном приближении, в котором пространственный масштаб значительно превышает дебаевскую длину. Дебаевская длина λ_D определяет максимальное расстояние, на котором еще проявляется разделение зарядов в горячей плазме. Следовательно, если нас интересуют характеристики, осредненные по масштабу, существенно превосходящему λ_D , то можно считать плазму нейтральной и не учитывать правую часть (9.1c). Иначе, можно рассматривать все токи текущими по замкнутым цепям и поэтому не приводящими к образованию зарядов. В связи с этим МГД-плазма может рассматриваться как нейтральная жидкость, по которой текут токи. Приближение зарядовой квазинейтральности исключает механизм, ответственный за возбуждение плазменных колебаний, и связанные с ним высокочастотные движения.

Обычно магнитные поля в МГД-плазме имеют большое значение и являются заданными (например, в термоядерных установках), а поведение системы определяется токами и действующими на них магнитными силами. Токи сами возбуждают значительные магнитные поля согласно (9.1b). Действие электрического поля является менее важным и проявляется через индукционный эффект при изменении во времени B согласно уравнению (9.1d).

В этом приближении уравнения Максвелла принимают вид

$$\operatorname{rot} \mathbf{B} = \mu_0 \mathbf{j}, \quad (9.4a)$$

$$\operatorname{rot} \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}, \quad (9.4b)$$

$$\operatorname{div} \mathbf{B} = \operatorname{div} \mathbf{E} = 0. \quad (9.4c)$$

Если проводимость обозначить через σ , то уравнение для плотности тока преобразуется к виду

$$\mathbf{j} = \sigma [\mathbf{E} + \mathbf{v} \times \mathbf{B}], \quad (9.5)$$

где \mathbf{v} — массовая скорость жидкости.

Если проводимость высокая, а величина \mathbf{j} ограничена, то из уравнения (9.5) следует

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0, \quad (9.6)$$

которое в комбинации с (9.4b) приводит к уравнению

$$\frac{\partial \mathbf{B}}{\partial t} = \operatorname{rot} (\mathbf{v} \times \mathbf{B}). \quad (9.7)$$

Уравнение (9.7) является условием сохранения магнитного потока $\int \mathbf{B} \cdot d\mathbf{s}$ через любую поверхность, движущуюся со скоростью жидкости \mathbf{v} . Следовательно, это уравнение может интерпретироваться как «вмороженность» силовых линий магнитного поля, вынужденных двигаться с жидкостью.

Уравнения МГД дополняются уравнениями неразрывности и баланса сил для жидкости

$$\frac{\partial}{\partial t} \rho_m = - \operatorname{div} (\rho_m \mathbf{v}), \quad (9.8a)$$

$$\rho_m \frac{d}{dt} \mathbf{v} = - \operatorname{grad} p + \mathbf{j} \times \mathbf{B}, \quad (9.8b)$$

где субстанционная производная, или скорость изменения, «отслеживаемая» частицами, движущимися с жидкостью, имеет вид

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \operatorname{grad} \quad (9.8c)$$

и ρ_m — массовая плотность.

Давление p обычно связывается с плотностью при помощи адиабатического закона $p/\rho_m^\gamma = \operatorname{const}$, где γ — отношение удельной теплоемкости при постоянном давлении к удельной теплоемкости при постоянном объеме.

Уравнения МГД, как правило, решаются при помощи конечно-разностных методов, известных в динамике жидкости

(например, метод Лакса — Вендорфа). Эти методы хорошо описаны в книгах «Введение в вычислительную физику» [314] и «Разностные методы решения краевых задач» [332]. Такие методы выходят за рамки настоящей книги и не будут рассматриваться.

Поскольку вывод уравнений МГД содержит масштаб осреднения, существенно превышающий орбиты ионов и электронов, могло показаться, что это приближение не годится для моделирования методом частиц. Однако в отдельных частных случаях методы частиц могут успешно применяться. Например, при исследовании ленгмюровских волн и вистлеров Ратманн, Вомворидис и Денавит [320] разделяют плазму на холодную компоненту с большой плотностью, которая составляет основную среду, в которой распространяются колебания, и энергичную электронную компоненту с малой плотностью, ответственную за взаимодействие резонансных частиц. Затем холодная компонента представляется гидродинамической моделью, по которой рассчитывается изменение во времени амплитуд и фаз ряда волн. После этого орбиты электронов с высокой энергией, движущихся в этих волнах, вычисляются аналитически в первом приближении по электрическим полям. В этом подходе временной шаг связан с частотой захвата частиц на волнах и может представлять интеграл по многим плазменным периодам. Кроме того, предполагается, что амплитуда и фаза волн медленно изменяются за период осцилляции. Этот метод расчета известен как «гибридное моделирование», поскольку в нем объединяются кинетическое и гидродинамическое описания. В этом подходе имеются дополнительные возможности экономичного применения частиц, так как они используются для представления компоненты с малой плотностью. Интегрирование по времени относится к моделированию по так называемому «медленному» времени (LTS). В разд. 9.4 описываются два других интересных примера использования методов частиц в МГД-пределе. Они применяются для изучения взаимодействия солнечного ветра с магнитным диполем Земли, которое приводит к образованию магнитосферы и геомагнитного хвоста.

9.1.2. Плазма в электростатическом поле

В противоположность магнитной гидродинамике, плазма в электростатическом приближении (ЭС) содержит высокие частоты (например, плазменные частоты) и малые пространственные масштабы порядка дебаевской длины. В этом случае может появиться разделение зарядов, и мы сохраняем правую часть (9.1c). Однако предполагается, что токи малы, так что самоогласованное поле незначительно, и можно пренебречь правыми частями (9.1b), (9.1d). В этом приближении уравнения Мак-

свелла приобретают вид

$$\operatorname{div} \mathbf{E} = \rho/\epsilon_0, \quad (9.9a)$$

$$\operatorname{div} \mathbf{B} = 0, \quad (9.9b)$$

$$\operatorname{rot} \mathbf{B} = \operatorname{rot} \mathbf{E} = 0. \quad (9.9c)$$

В этих случаях удобно ввести электростатический потенциал φ , так что

$$\mathbf{E} = -\nabla\varphi, \quad (9.10a)$$

тогда

$$\operatorname{rot} \mathbf{E} = -\operatorname{rot} \operatorname{grad} \varphi = 0, \quad (9.10b)$$

как требовалось, и, кроме того,

$$\operatorname{div} \mathbf{E} = -\operatorname{div} \operatorname{grad} \varphi = -\nabla^2\varphi = \rho/\epsilon_0. \quad (9.10c)$$

Таким образом, уравнения Максвелла приведены к (9.10c), известному как уравнение Пуассона.

Поскольку интересующий нас пространственный масштаб по порядку величины теперь равен дебаевской длине и важны орбиты отдельных ионов или электронов, то для моделирования плазмы в электростатическом приближении вполне подходят методы «частиц в ячейке».

Полная система уравнений в этом случае имеет вид для полей

$$\nabla^2\varphi = -\rho/\epsilon_0, \quad (9.11a)$$

$$\mathbf{E} = -\operatorname{grad} \varphi, \quad (9.11b)$$

$$\rho = \sum_i q_i/V, \quad (9.11c)$$

для движения каждой частицы с индексом i

$$m_i \frac{d}{dt} \mathbf{v}_i = q_i [\mathbf{E} + \mathbf{v}_i \times \mathbf{B}_0], \quad (9.11d)$$

$$\frac{d}{dt} \mathbf{x}_i = \mathbf{v}_i. \quad (9.11e)$$

Здесь \mathbf{B}_0 описывает влияние заданного внешнего магнитного поля.

В разд. 9.2 подробно обсуждаются свойства двумерной электростатической модели «частица—сетка» и в разд. 9.3 описывается ее применение при изучении диффузии плазмы поперек сильного магнитного поля. Сильное магнитное поле используется для удержания высокотемпературной плазмы в установках, предназначенных для управляемого термоядерного синтеза. Неожиданно высокая аномальная диффузия плазмы поперек удерживающего поля явилась главным фактором, затормозившим при-

менение этих установок. Численное моделирование сыграло важную роль в расширении наших знаний об основных механизмах, вызывающих аномальную диффузию.

9.1.3. Исторический обзор

Гораздо большие усилия в численном моделировании затрачены на изучение высокотемпературных плазм в связи с большим капиталовложением на изучение управляемого термоядерного синтеза. Большинство обсуждаемых работ выполнено в исследовательских группах, организованных в Соединенных Штатах [в Научной лаборатории Лос-Аламоса (LASL), Вычислительном центре по термоядерной энергии с магнитным удержанием (MFECSS) в Ливерморе, Лоуренсовской ливерморской лаборатории (LLL), Военно-морской исследовательской вашингтонской лаборатории (NRL), Калифорнийском университете в Беркли (UCB) и Лос-Анджелесе (UCLA) и университетах Стэнфорда, Принстона и Колумбии] и в Британии (в Каламской лаборатории УКАЕА, Королевском колледже и Ридингском университете), хотя важный вклад был сделан и в других странах (например, в Институте Макса Планка, Гарчинге, ФРГ).

Начиная с 1960 г. литература стала настолько обширной, что здесь можно затронуть лишь наиболее значительные достижения в разработке проблемы. В центре внимания будут статьи, касающиеся развития методов моделирования, использующих частицы. Первые статьи вышли из печати в основном в журнале «Physics of Fluids», а сейчас подобные работы появляются в «Journal of Computational Physics», который специализируется на вычислительных методах с подробным изложением структуры моделей, зачастую лишь поверхностно упоминающейся в других журналах. Более подробные заказные статьи появляются в серии томов «Methods in Computational Physics» (ред. Б. Олдер, С. Фернбах и М. Ротенберг, «Академик Пресс», Нью-Йорк и Лондон), среди которых том 9 (1970) и том 16 (1976) посвящены физике плазмы и управляемому термоядерному синтезу соответственно. Эти тома вместе с трудами конференций специалистов по численному моделированию плазмы (LASL, 1968, Вашингтон, NRL, 1970, LLL, 1973, и Университет Нью-Йорка, 1975) представляют собой основные материалы по методам моделирования плазмы. Специальные обзоры о роли вычислительного моделирования в физике плазмы выполнены Бунеманом и Дунном [65] и Бердсолом и Доусоном [34].

В 1975 г. в Калифорнийском университете вышли отпечатанные записи курса под названием «Вычислительный эксперимент в физике плазмы» (Бердсол и Лэнгдон, четвертое издание вышло в 1978 г.). В этом курсе детально, включая образцы программы, рассматриваются методы и теория электростатиче-

ских кодов при наличии внешнего магнитного поля и без него. Коммерческое издание этих лекций осуществлено в 1981 г. Бердсолом и Лэнгдоном. Отлаженные стандартные машинные программы для некоторых моделей опубликованы в журнале «Computer Physics Communications» (North-Holland Publishing Co., Amsterdam). Помещенные в этом журнале программы могут быть приобретены за минимальную стоимость в библиотеке программ СРС в Королевском университете в Белфасте.

Первые модели частиц для электростатической плазмы были построены в пионерской работе Бунемана [56] в Стэнфорде и Бердсола и Бриджса [32] в Беркли. Эти модели были одномерными и не использовали сетку для расчета полей. Такие модели развиты Доусоном (103) в Принстоне и Элдриджем и Фейксом [127]. Важный шаг, состоящий в введении сетки для расчета полей, приводящий к первым алгоритмам РМ, был сделан в Стэнфорде независимо Бюргером [67] в одномерном приближении и Хокни [177] в двумерном. Последнее потребовало параллельного развития быстрых прямых методов для решения уравнения Пуассона, о которых надо говорить как о первых алгоритмах RES [175, гл. 6]. В упомянутом выше алгоритме РМ использовались распределение зарядов согласно NGR и интерполяция поля для экономии машинного времени. Свыше 3000 частиц могло промоделироваться в одномерном приближении и до 2000 частиц в двумерном на IBM 7090. Схемы интерполяции высокого порядка (SIC) впервые были введены группой в Беркли [35] для понижения шума при моделировании и Морсом и Нилсоном [291], использующими метод PIC Харлоу [164] для моделирования динамики жидкости. Метод PIC, примененный к физике плазмы, также обсуждался Морсом [290]. Работы по одномерной электростатической модели были рассмотрены в обзоре Доусона [104], а методы для эффективно работающих кодов — Борисом и Робертсом [42] и Бунеманом [61].

В Принстоне и UCLA ([228], [76]) была развита другая методика для расчета полей — метод мультипольного разложения. Этот метод основан на разложении распределения заряда в каждой ячейке сетки по монопольному, дипольному и квадрупольному моментам относительно центра каждой ячейки. Эти отдельные распределения анализируются методом Фурье и суммируются для получения преобразования Фурье монопольного, дипольного и квадрупольного выражений для силы поля. Анализ Фурье этих выражений дает мультипольный момент силы поля в узлах сетки. Разложение этих сил в ряд Тейлора определяет поле в точке нахождения каждой частицы. Даже если в мультипольном разложении берутся только дипольные члены, в основном требуется больше вычислений и памяти по сравнению с алгоритмами РМ, описанными в этой книге. Обобщение на трехмерную ситуацию представляет особый случай, в котором

имеются четыре распределения для анализа, синтеза и запоминания (монополю и три компоненты диполя) по сравнению с распределением одного эквивалентного монополя, использованного в алгоритме РМ. Тем не менее успешно применяются одномерные и двумерные мультипольные коды.

Теория алгоритма РМ была впервые разработана последовательным образом Бердсолом, Лэнгдоном и их коллегами в Беркли в серии важных работ. Тот факт, что алгоритм РМ представляет собой модель плазмы с частицами конечного размера или облаками, был раскрыт Хокни [177] в описании самой ранней модели РМ, а теория моделирования, использующая такие частицы конечного размера, была развита Лэнгдоном и Бердсолом [238]. Были последовательно рассмотрены эффекты пространственной сетки [233, 234] и конечного временного шага [75, 233, 237] соответственно на дисперсионное соотношение модели. Эти работы были обобщены Лэнгдоном в 1973 г. на консервативные алгоритмы Льюиса [252, 253]. Упомянутые выше статьи лежат в основе гл. 7. Кинетическая теория флуктуаций и шума в таких вычислительных моделях была представлена Лэнгдоном [236].

Концепция частиц конечного размера была использована Окудой и Бердсолом [299] для расчета сильного понижения темпа столкновений в результате использования метода РМ, который по существу моделирует частицы, сравнимые по размеру с ячейкой сетки. Этот эффект был подтвержден обширным эмпирическим исследованием характерных времен столкновения и нагрева, выполненным Хокни [181] и представленным в разд. 9.2.1—9.2.3, и Окудой [298]. Свойства сохранения энергии и импульса различными кодами РМ были также сопоставлены Льюисом, Сайксом и Вессоном [254], которые показали, что из-за эффекта конечного временного шага так называемые «сохраняющие энергию» схемы (см. разд. 5.5) не имеют заметного преимущества в сохранении энергии по сравнению с кодами, сохраняющими импульс.

Теории Лэнгдона для сеточных эффектов в алгоритмах РМ были подтверждены экспериментальной проверкой при изучении нефизических неустойчивостей в холодной плазме, обусловленных движением через вычислительную сетку, и неустойчивостей горячей плазмы, когда $H/\lambda_D > \pi$ [8, 37, 75, 235, 297]. К счастью, ни те ни другие неустойчивости не являются разрушительными для используемого метода. Неустойчивость холодной плазмы насыщается при малой амплитуде, а неустойчивость горячей плазмы перестает нарастать, после того как увеличивается температура, а следовательно, и λ_D , так что выполняется условие $H/\lambda_D < \pi$, и, таким образом, плазма стабилизируется. Тем не менее хорошие модели должны минимизировать эти нежелательные эффекты.

В результате усилий, затраченных на понижение эффекта пространственной сетки, обе группы из Беркли и Ридинга пришли к схеме следующего более высокого порядка распределения заряда по ячейкам и интерполяции поля. В этой схеме распределение осуществляется с помощью облака с треугольнообразной формой плотности (TSC, см. гл. 5), а интерполяция — с использованием квадратичного сплайна (QS) согласно системе обозначений, принятой в Беркли. Группа из Ридинга также показала, что функции Грина, использованной при вычислении потенциала, может быть придана форма, компенсирующая квадратную форму сетки. В результате появился алгоритм, в котором сеточный эффект был исключен для всех практических целей. Впервые это было продемонстрировано в двумерном случае эмпирическим построением функции Грина [119, 190]. Иствуд [115—117] (см. также гл. 7—8) развил теорию, которая дает выражение для наилучшего выбора функции Грина с помощью процедуры *Q-минимизации*. Эти алгоритмы обладают, по крайней мере, в сто раз меньшим уровнем шума по сравнению с исходным методом СИС и известны как «бесшумные» коды «частица — сетка» (QPM). Добавление прямых взаимодействий типа частица — частица с соседними частицами (образующее код P^3M) дает возможность включать любое желаемое короткодействующее взаимодействие и может быть достигнуто внутрисеточное разрешение. Трехмерный вычислительный код P^3M^3DP был опубликован с учетом этих идей и является доступным из библиотеки программ SRC [120].

Появление лазеров и возможность лазерного термоядерного синтеза привели к разработке полностью электромагнитных (в отличие от электростатических) кодов, в которых все уравнения Максвелла (9.1) для электрического и магнитного полей решаются совместно с движением заряженных частиц. Такие модели обсуждаются Лэнгдоном и Лазински [239], а в безизлучательном пределе Нилсоном и Льюисом [295]. Переход от двумерных электростатических кодов к трехмерным электромагнитным обсуждается Бунеманом [61], описавшим электромагнитный код, разработанный в Стэнфорде. Этот код включает этап движения частиц, использующий релятивистское обобщение центрированной во времени конечной разности второго порядка для ускорения Лоренца (разд. 4.7.1), и этап обновления значений полей E и B с применением тройного преобразования Фурье и решения в k -пространстве. Имеется десять переменных для каждого узла сетки (ρ и по три компоненты векторов E , B и j) и шесть переменных на частицу (три компоненты векторов x и v). Из-за их большого количества все эти переменные запоминаются на дисках и передаются в быструю память компьютера и возвращаются обратно с целью обновления их значений. Проявлена большая изобретательность для того, чтобы операции обмена

между двумя уровнями памяти производились одновременно с операциями арифметической обработки [23]. На IBM 360/91 (производительность составляет от 1/2 до 1/3 от скорости CDC-7600, см. [183]) коды используют сетку $64 \times 64 \times 64$ с интерполяцией с помощью квадратичного сплайна (например, 3D TSC). Этап движения частиц занимает 250 мкс на частицу, а пересчет полей — 70 с. На один временной шаг для 500 000 частиц затрачивается приблизительно 4 мин. Версия этого кода, хорошо оптимизированная на CDC-7600 Нилсеном [293], Грином и Мирабеллой, получившая название SPLASH и имеющаяся в библиотеке MFESS в Ливерморе [64], была использована для изучения динамики эволюции Z-пинча Нилсеном и др. [294]. Затраты машинного времени на этапе движения частиц в системе, включающей 300 000 частиц, составляют 130 мкс на частицу и 4 с на пересчет поля на сетке $32 \times 32 \times 32$. Векторизованная версия, описанная Бунеманом [62] для ЭВМ CRAY-1, использовала линейную тетраэдрическую интерполяцию и сетку $128 \times 128 \times 256$ (4 миллиона узлов сетки). Тетраэдрическая интерполяция [63] является формой трехмерной линейной интерполяции, для которой требуются данные только из четырех ближайших узлов сетки, по сравнению с трилинейной интерполяцией, использующей ближайшие восемь узлов в трехмерном коде SIC. Шесть координат каждой частицы запакованы в два слова длиной 64 бита на GRAY-1, с фиксированной запятой с точностью в 21 бит. Код оформлен на языке ассемблера CAL, и расчет движения частиц требует всего 7 мкс на частицу вследствие комбинированного эффекта снижения порядка интерполяции (ускорение в 4 раза) и повышения производительности за счет тщательного кодирования на GRAY-1 (ускорение в 5 раз). Благодаря этому коду моделирование с применением нескольких миллионов частиц затрачивает приблизительно одну минуту на один временной шаг.

Описанные к настоящему моменту модели, расчет которых проходит на микроскопической шкале с прослеживанием за орбитами индивидуальных частиц, ограничены при изучении фундаментальных свойств плазмы временными и пространственными масштабами, существенно меньшими, чем это необходимо в реальных термоядерных установках. Для изучения последних необходимо перейти к макроскопическим уравнениям магнито-гидродинамики (9.4) — (9.5) или гибридной модели, включающей как макроскопическое, так и микроскопическое описание. Роль чисто микроскопического моделирования состоит в обеспечении понимания основного поведения плазмы и задания, например, коэффициентов переноса для использования в макроскопической модели. Исследование последней различными общепринятыми методами, использующими совместные эйлерову

и лагранжеву методике, дано в обзорах Робертса и Поттера [335] и Брэкбилла [47].

Методы частиц нашли свое место в макроскопическом моделировании жидкостей, в частности при рассмотрении движущихся поверхностей раздела. Метод, основанный на известном методе «частица в ячейке» (PIC), развитом Харлоу [164] для гидродинамических расчетов, был успешно распространен на МГД Мардером [272] и использован, например, Тоддом [370] для изучения нагрева плазмы с помощью ударных волн. Другой подход представляет модель «водяного мешка», в которой система определяется с помощью хранения и продвижения координат частиц, задающих контуры некоторой величины. Этот метод впервые был использован в одномерной постановке Берком и Робертсом [28] для слежения за плотностью контуров в фазовом пространстве (x, v) , но с тех пор был распространен Поттером [315] на изучение контуров магнитного поля в реалистических тороидальных термоядерных установках в двумерной постановке.

Довольно часто параметры плазмы находятся между значениями, которые могут удовлетворительно моделироваться соответственно с помощью микроскопической и макроскопической моделей. Таким является случай сильного магнитного поля, когда ларморовский радиус электронов мал и им можно пренебречь, но необходимо учесть влияние ларморовского радиуса ионов. В этом случае удобно рассматривать электроны как жидкость и рассчитывать орбиты ионов с помощью бесстолкновительного метода частиц. Такие гибридные методы обсуждаются в работах [73, 172].

9.2. ДВУМЕРНАЯ ЭЛЕКТРОСТАТИЧЕСКАЯ МОДЕЛЬ

Мы уже использовали одномерную модель электростатической плазмы в качестве примера, иллюстрирующего метод частица — сетка, изложенный в гл. 2 и 3. При более высокой размерности основной цикл остается тем же, но для решения уравнения Пуассона необходимо использовать другие методы. Выбор подходящих методов уже полностью обсужден в гл. 6. Ограничения на затраты машинного времени обычно заставляют использовать методы быстрого преобразования (RES-алгоритмы, разд. 6.5.1—6.5.7), если геометрия задачи достаточно простая. Поведение таких моделей характеризуется двумя основными особенностями — временем столкновения и временем нагрева, которые были изучены в работах [181, 190]. Результаты представлены ниже.

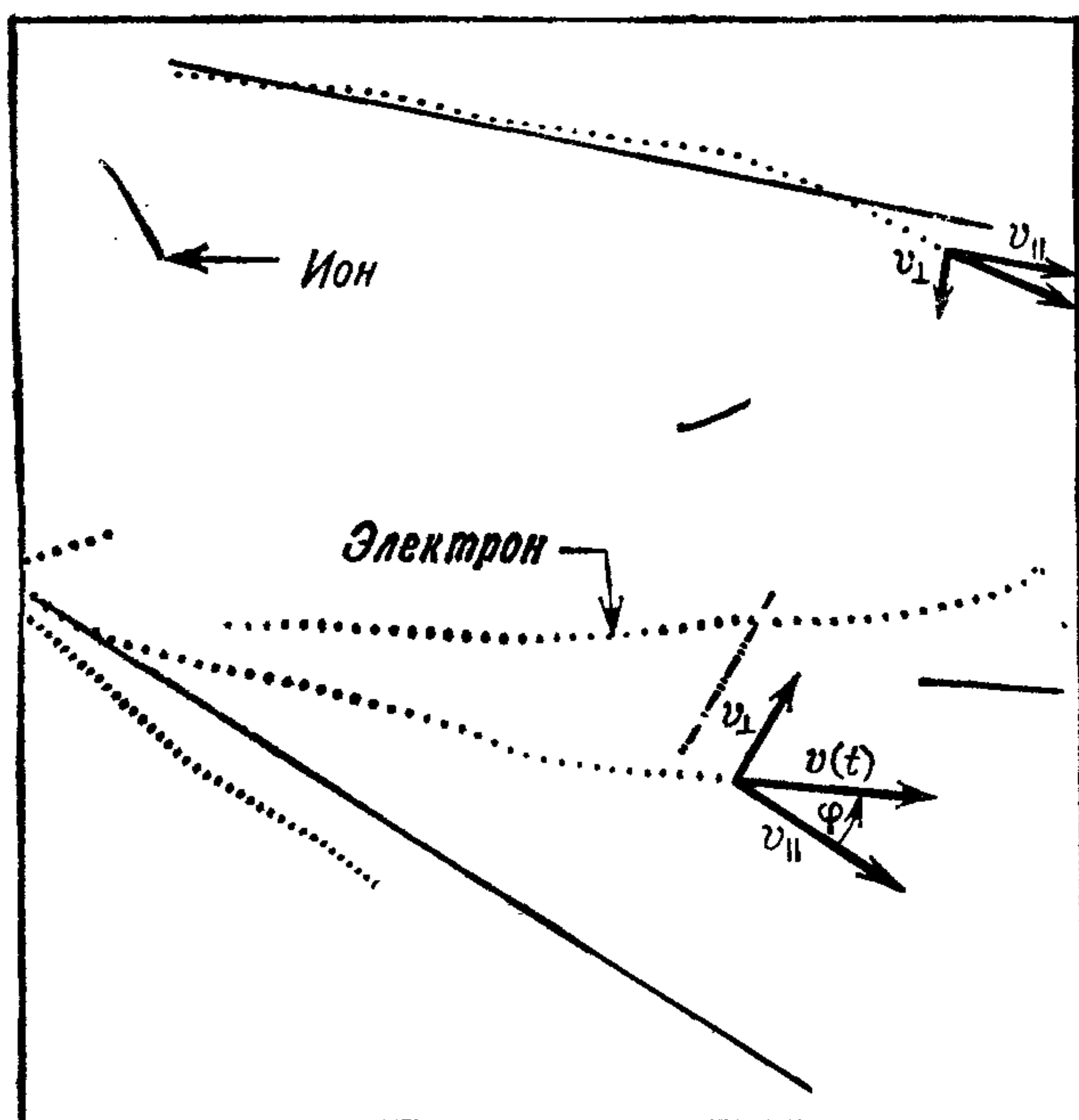


Рис. 9.1. Типичные орбиты в двумерной модели частица — сетка электростатической плазмы. ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc)

9.2.1. Время столкновения

Для изучения фундаментальных свойств электростатической плазмы рассмотрим квадратную область с периодическими граничными условиями по обоим направлениям. На рис. 9.1 показаны некоторые типичные орбиты для такой модели. В идеальной однородной бесстолкновительной плазме орбиты всех ионов и электронов являются прямыми линиями. То, что наблюдаемые в модели орбиты не являются прямыми, связано с влиянием столкновительных эффектов в модели. Количественно этот эффект можно выразить через определение времени столкновения следующим образом.

Предположим, что орбита частицы с индексом i отклонена от первоначального направления на угол $\varphi_i(t)$ за время t ; тогда можно определить среднеквадратичное отклонение как

$$\langle \varphi^2(t) \rangle^{1/2} = \left(\frac{1}{N} \sum_{i=1}^N \varphi_i^2(t) \right)^{1/2}. \quad (9.12a)$$

Время столкновения τ_c равно времени, за которое среднее отклонение достигает 90° , т. е.

$$\langle \varphi^2(\tau_c) \rangle^{1/2} = \pi/2. \quad (9.12b)$$

Везде в этом разделе мы используем ломаные скобки для обозначения величины, усредненной по ансамблю рассматриваемых частиц. Например, если a_i — значение некоторой величины для частицы i , то

$$\langle a \rangle = \frac{1}{N} \sum_{i=1}^N a_i. \quad (9.12c)$$

Нам также будут нужны некоторые результаты теории случайного блуждания частицы за n шагов с длиной вектора шага δa . Предположим, что длина шага $|\delta a|$ имеет постоянную величину, но случайное направление. После n шагов частица i проходит расстояние Δa_i , и можно получить следующие усреднения:

$$\langle \Delta a \rangle = 0, \quad (9.13a)$$

$$\langle |\Delta a|^2 \rangle = n |\delta a|^2. \quad (9.13b)$$

Первый результат следует из симметрии задачи. Второй результат более тонкий и утверждает, что квадрат полного изменения можно получить путем сложения квадратов отдельных шагов. Таким образом, среднеквадратичное значение величины растет

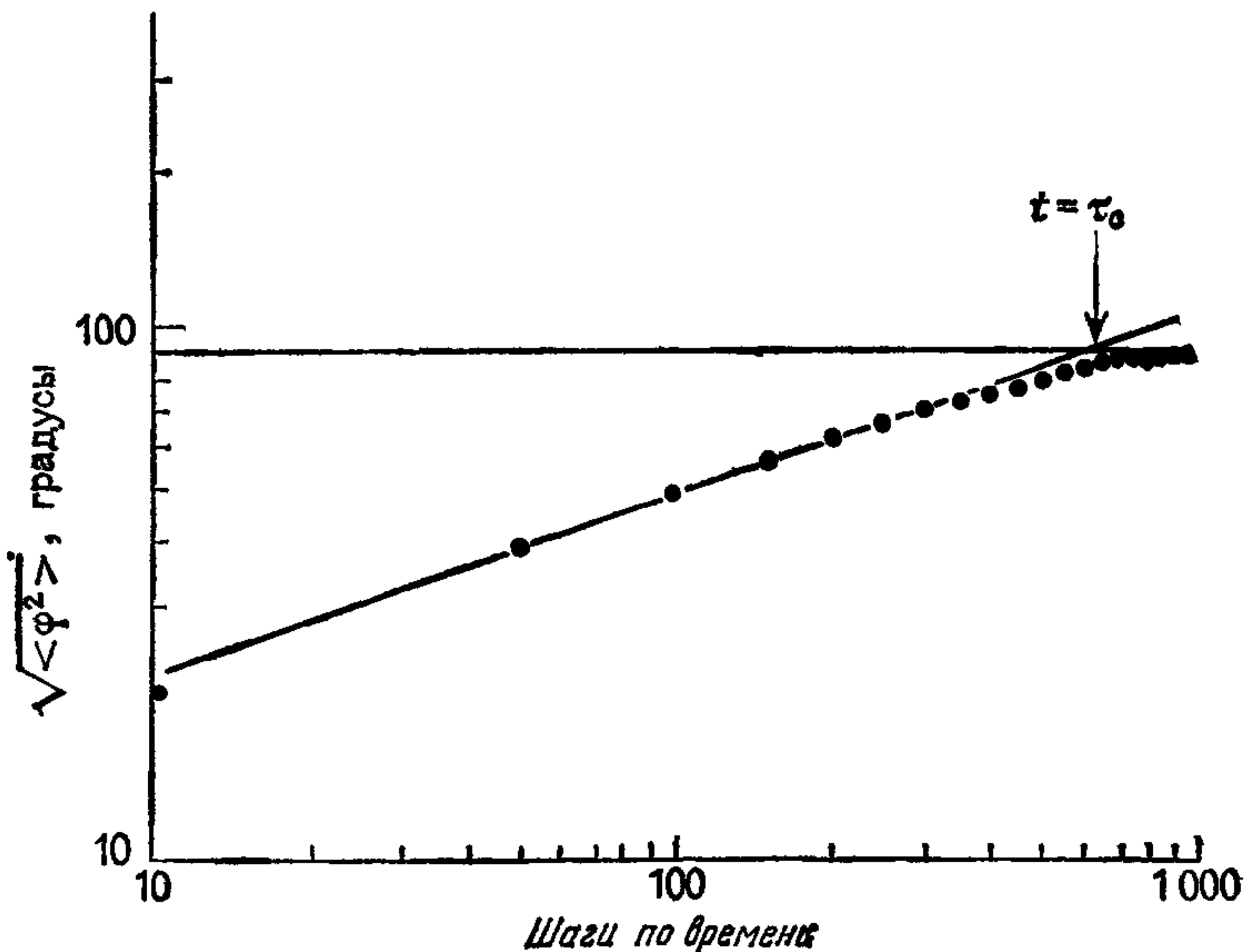


Рис. 92 Измерение времени отклонения ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc.)

линейно с числом шагов. Если δa равно смещению, то (9.13b) является уравнением диффузии облака частиц, а $\langle |\Delta a|^2 \rangle^{1/2}$ служит мерой ширины этого облака.

Если полное отклонение φ является результатом большого числа малых отклонений $\delta\varphi$, которые имеют стохастическую природу, можно рассматривать процесс рассеяния как случайное блуждание по углу и применить (9.13b) для получения полного квадратичного отклонения за n шагов

$$\langle \varphi^2 \rangle = n (\delta\varphi)^2, \quad (9.14a)$$

или

$$\langle \varphi^2 \rangle^{1/2} = \delta\varphi n^{1/2}. \quad (9.14b)$$

Таким образом, среднеквадратичное отклонение пропорционально квадратному корню из числа шагов. На рис. 9.2 показано измерение величины $\langle \varphi^2 \rangle^{1/2}$ в электростатической модели плазмы. Видно, что среднеквадратичное отклонение пропорционально $n^{1/2}$ в соответствии с данной выше стохастической теорией. Время столкновения определяется довольно точно

9.2.2. Время нагрева

Другой величиной, важной для определения качества плазменной модели, является время нагрева. Если $h_i(t)$ есть отклонение кинетической энергии частицы i от ее первоначального значения, то можно определить среднее по ансамблю изменение кинетической энергии на частицу

$$\langle h(t) \rangle = \frac{1}{N} \sum_{i=1}^N h_i(t). \quad (9.15a)$$

Время нагрева τ_H определяется как время увеличения средней кинетической энергии на $(1/2)k_B T$ на частицу, т. е.

$$\langle h(\tau_H) \rangle = (1/2) k_B T, \quad (9.15b)$$

где k_B — константа Больцмана и T — абсолютная температура.

В изолированной неизлучающей физической плазме полная энергия сохраняется, и в тепловом равновесии средняя кинетическая энергия (т. е. температура) постоянна, следовательно, $\langle h(t) \rangle \equiv 0$. Энергия, теряемая одной частицей, всегда приобретает другой. Следовательно, в физической системе $\tau_H = \infty$, а измерение конечного значения τ_H в модели является выражением неполного сохранения энергии.

Для того чтобы понять, почему в реальной модели $\tau_H < \infty$, поучительно рассмотреть все погрешности вычислительной модели, вызывающие стохастическое поле погрешностей. Мы собираем вместе погрешности, обусловленные арифметическим

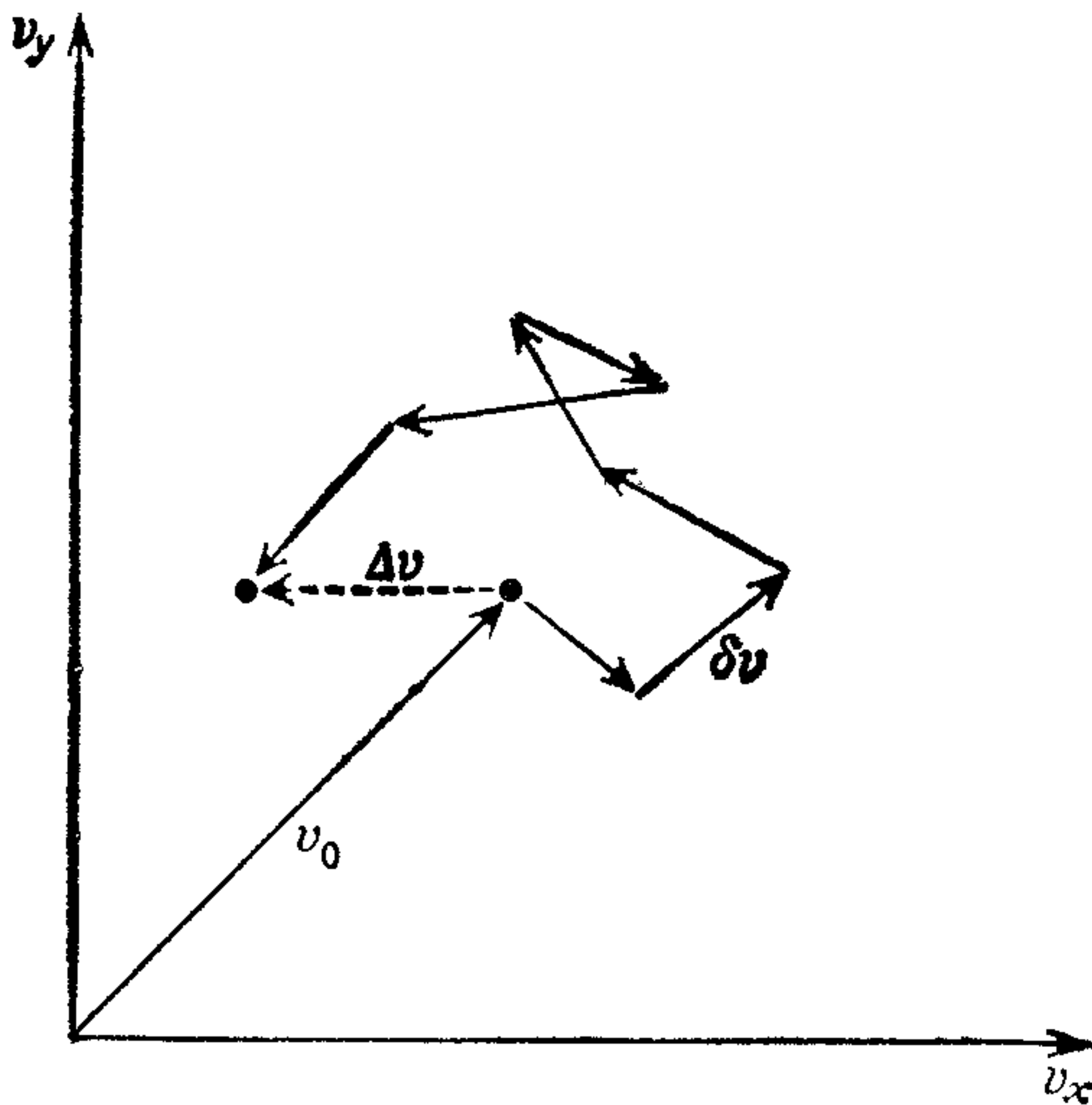


Рис 9.3 Случайное блуждание в пространстве скоростей под действием случайных флуктуаций поля

округлением, конечным размером временного шага, нахождением конечных разностей в уравнениях поля и использованием макрочастиц, в единое поле погрешностей δE , которое для простоты будем рассматривать постоянным по величине, но случайным по направлению. Это поле приложено к каждой частице и вызывает изменение количества движения, равное импульсу. Следовательно,

$$m \delta \mathbf{v} = q \delta E DT, \quad (9.16a)$$

$$\delta \mathbf{v} = (q/m)DT \delta E. \quad (9.16b)$$

Поле погрешностей отличается для каждой частицы и изменяется за каждый временной шаг случайным образом. Влияние большого числа временных шагов сказывается в том, что каждая частица случайно блуждает в пространстве скоростей относительно ее начальной скорости \mathbf{v}_0 , как показано на рис. 9.3.

Если мы имеем ансамбль частиц, обладающих в начальный момент скоростью \mathbf{v}_0 , то после n шагов, используя (9.13), получим

$$\langle \Delta \mathbf{v} \rangle = 0, \quad (9.17a)$$

$$\langle |\Delta \mathbf{v}|^2 \rangle = n \frac{q^2}{m^2} DT^2 |\delta E|^2 \quad (9.17b)$$

и среднее изменение кинетической энергии на частицу равно выражению

$$\begin{aligned} \langle h \rangle &= \frac{1}{2} m \langle |\mathbf{v}_0 + \Delta \mathbf{v}|^2 \rangle - \frac{1}{2} m \langle |\mathbf{v}_0|^2 \rangle = \\ &= \frac{1}{2} m (2\mathbf{v}_0 \cdot \langle \Delta \mathbf{v} \rangle) + \frac{1}{2} m \langle |\Delta \mathbf{v}|^2 \rangle, \end{aligned}$$

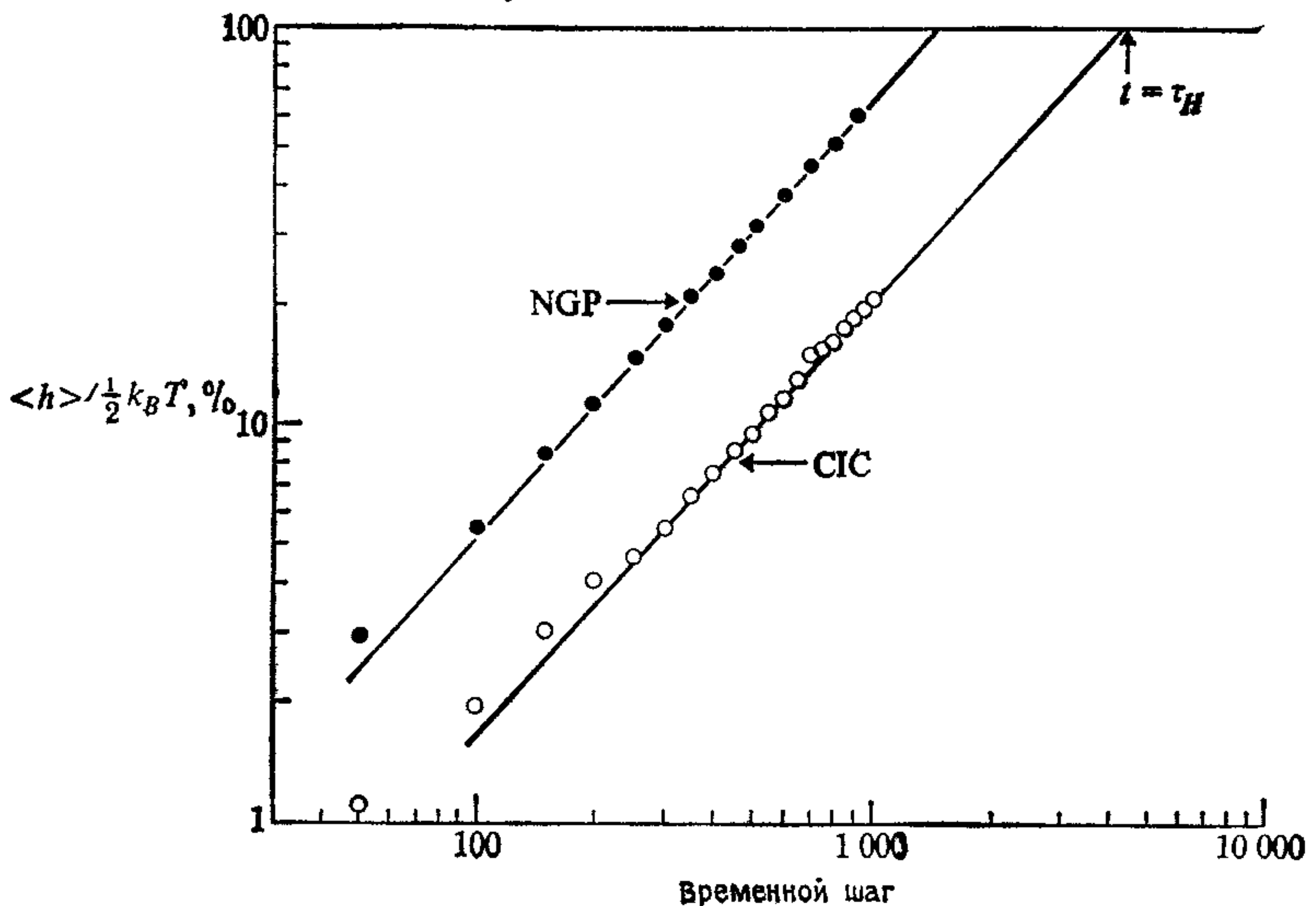


Рис 9.4 Измерение времени нагрева τ_H для моделей NGP и CIC. ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc.)

которое при использовании (9.17a) принимает вид

$$\langle h \rangle = \frac{1}{2} \frac{q^2}{m} DT^2 \langle \delta E \rangle^2 n. \quad (9.18)$$

Уравнение (9.18) применяется для любой группы частиц независимо от их начальной скорости v_0 и, следовательно, применимо к системам с распределением по скоростям, таким, как максвелловское. Важный результат состоит в том, что средняя кинетическая энергия возрастает линейно с количеством шагов, если только поля погрешностей являются случайными. Это явление носит название *стохастический нагрев* и до некоторой степени имеется во всех вычислительных моделях. Поскольку этот нефизический эффект нельзя исключить, важно знать его величину и как его контролировать.

На рис. 9.4 показано измерение $\langle h \rangle$ как функции времени в логарифмическом масштабе. График отчетливо демонстрирует линейное возрастание кинетической энергии со временем в соответствии с данной выше простой стохастической теорией. Измерения для моделей NGP и CIC также показывают, что стохастический нагрев значительно слабее в модели CIC, чем в модели NGP.

В плазме потенциальная энергия мала, и ею можно пренебречь в сравнении с кинетической энергией. Далее (9.18)

показывает, что темп стохастического нагрева обратно пропорционален массе рассматриваемых частиц. Поэтому стохастический нагрев проявляется в основном как увеличение температуры легких электронов. Скорость роста температуры ионов гораздо ниже, и обычно ею можно пренебречь. По этой причине величина $\langle h \rangle$ измеряется только для электронов. Поскольку в двумерном случае средняя энергия электронов равна $k_B T$ ($1/2 k_B T$ на каждую степень свободы), и первоначальная кинетическая энергия ионов равна кинетической энергии электронов, время нагрева определяется временем увеличения полной энергии системы на 25 %.

9.2.3. Эмпирические корреляции

Вычислительная плазма отличается от реальной плотностью частиц n и конечной дискретностью по пространству и времени (с пространственной ячейкой H и временным шагом DT). Для корреляции экспериментальных измерений времени столкновения и нагрева использованы следующие безразмерные формы этих переменных.

1. $N_C = N_D \{1 + (W/\lambda_D)^2\}$ — число, характеризующее время столкновений, где W — ширина частицы. Оно изменяется в пределах от числа частиц на дебаевском квадрате $N_D (= n\lambda_D^2)$ для частиц с малым размером $W \ll \lambda_D$ до числа частиц в квадрате Дебая с большим размером $W \gg \lambda_D$.

2. H/λ_D — количество дебаевских длин на ячейке пространственной сетки.

3. $\omega_{pe}DT$ — произведение электронной плазменной частоты на временной шаг, т. е. часть дебаевской длины, пройденная тепловым электроном за временной шаг.

На рис. 9.5 показана корреляция времени столкновения для широкого диапазона вычислительных моделей с единственной переменной N_C . Все измерения могут с точностью до 20 % аппроксимироваться соотношением

$$\left(\frac{\tau_c}{\tau_{pe}}\right) = N_C = N_D \left\{1 + \left(\frac{W}{\omega_D}\right)^2\right\} = n(\lambda_D^2 + W^2). \quad (9.19a)$$

Соответственно для частоты столкновений $\nu_c (= \tau_c^{-1})$ имеем

$$\left(\frac{\nu_c}{\omega_{pe}}\right) = \frac{1}{2\pi N_C}. \quad (9.19b)$$

Уравнение (9.19a) показывает, что время столкновения в основном определяется плотностью частиц (переменная n) и только незначительно величиной W , связанной с типом используемой модели (например, NGR или CIC). Например, если модель со-

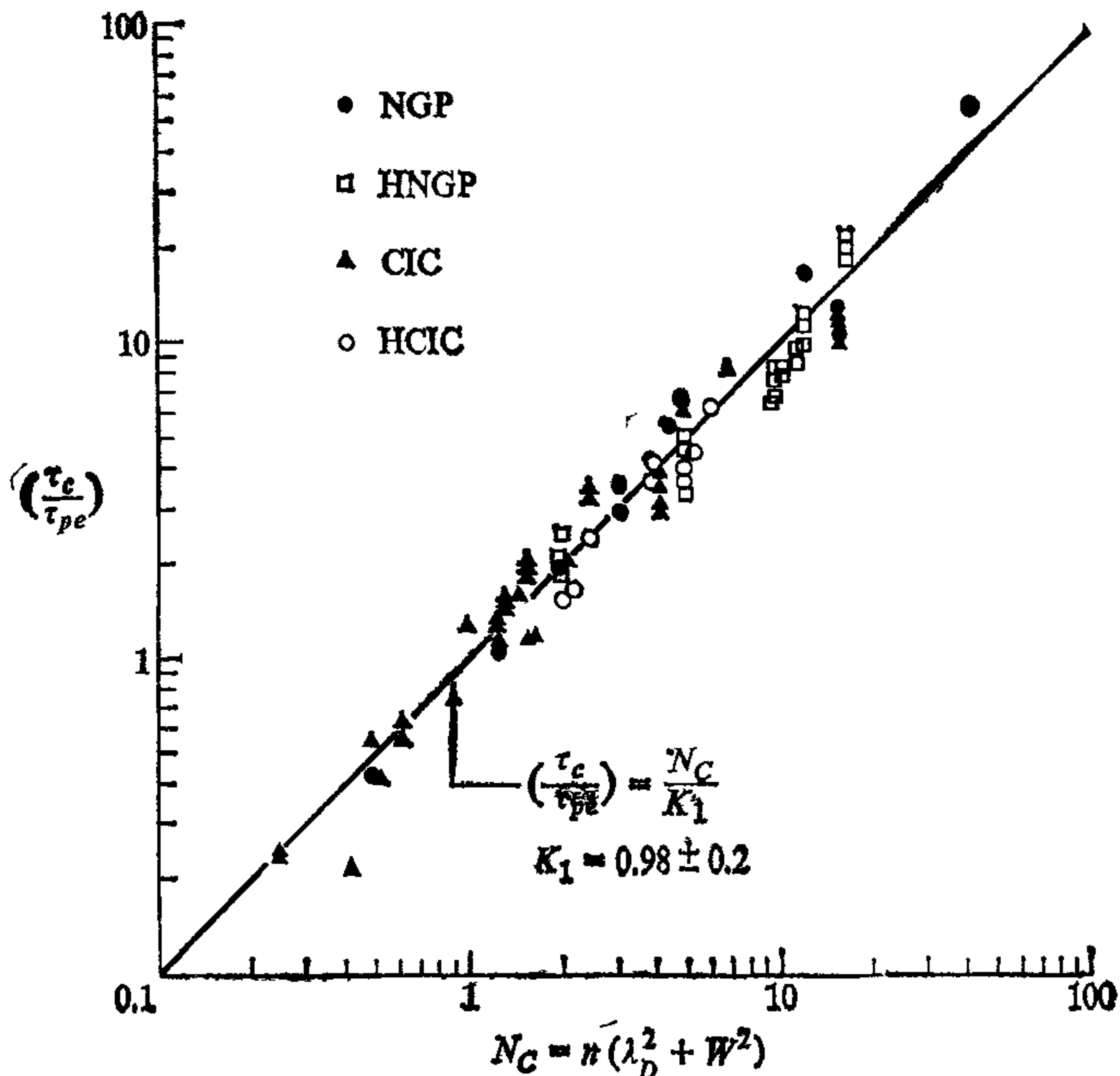


Рис. 9.5. Корреляция времени столкновений с числом частиц. ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc.)

держит только одну миллионную от числа частиц в реальной системе, то столкновительные эффекты будут усилены и время столкновения уменьшится на тот же коэффициент.

Размер частицы W зависит от выбранного типа распределения заряда по ячейкам и формы сглаживания поля, и типичные значения для обсуждаемых в этой книге моделей равны

$$\begin{aligned}
 \text{NGP с 5-точечным уравнением Пуассона} & \quad W = H, \\
 \text{CIC с 5-точечным уравнением Пуассона} & \quad W = H, \\
 \text{TSC с 9-точечным уравнением Пуассона} & \quad W = 2H, \\
 \text{TSC с } Q\text{-минимизацией} & \quad W = 3H.
 \end{aligned} \tag{9.20}$$

Заметим, что процесс сглаживания поля и понижения сеточных эффектов, который появляется при процедуре Q -минимизации (разд. 8.6), увеличивает размер частицы и снижает столкновительные эффекты примерно в 3 раза. Однако необходимо понимать, что это сделано за счет пространственного разрешения. Подавление коротковолновых гармоник с помощью увеличения размера частиц имеет двойной эффект понижения интенсивности столкновений и изменения дисперсионных свойств (см. гл. 7). Если только эти коротковолновые гармоники не

имеют большого значения для коллективных свойств, можно применять увеличение размера частицы для понижения эффектов столкновения. Иначе необходимо использовать большее число частиц.

В отличие от времени столкновения время нагрева сильно зависит от H и DT . Зависимости времени нагрева и времени столкновения от N_c одинаковы, и мы можем исключить эту зависимость, рассматривая отношение этих времен τ_H/τ_c . В любом случае именно отношение этих величин является важным, поскольку оно определяет количество столкновительных времен, которое можно вычислить до того, как нарушение сохранения энергии станет серьезным. Например, если мы выбираем параметры модели, для которой $\tau_H/\tau_c = 10$, то можно вести расчет в течение одного времени столкновения с 2.5 % -ным нарушением сохранения энергии.

Отношение времен нагрева и столкновения является сложной функцией от H/λ_D и $\omega_{pe}DT$, и ее вариации измерены в плоскости параметров этих двух переменных. Интересующая нас область показана на рис. 9.6. При $\omega_{pe}DT > 2$ явная конечно-разностная аппроксимация законов движения Ньютона неустойчива (разд. 4.4.1) и эту область плоскости нельзя использовать. Ошибки конечно-разностного обрезания станут серьезными, если тепловой электрон пересекает больше одной ячейки за временной шаг, т. е. если $v_T DT/H > 1$, поскольку в этом случае многие частицы «перепрыгивают» через вариации поля. Так как разрешение в описании поля равно H , мы ничего не выиграем от выбора временного шага короче, чем, например, одна десятая времени пересечения ячейки тепловой частицей. Следовательно, можно не обращать внимания на левую верхнюю часть плоскости, для которой $v_T DT/H < 0.1$.

Внутри интересующей нас области определим оптимальный путь, показанный пунктирной линией на рис. 9.6 и заданный соотношением

$$(\omega_{pe} DT)_{opt} = \min [1/2 H/\lambda_D, 1]. \quad (9.21)$$

Эта линия представляет наиболее разумный выбор временного шага для заданного H/λ_D в том смысле, что понижение DT вызывает относительно небольшое увеличение времени нагрева, тогда как его увеличение вызывает резкое ухудшение свойств модели (например, уменьшение τ_H).

Если предполагается, что временной шаг выбран вдоль оптимального пути, то можно установить зависимость отношения времен нагрева и столкновения от единственной переменной H/λ_D и получить (рис. 9.7)

$$\frac{\tau_H}{\tau_c} = \frac{K_H}{(H/\lambda_D)^2}, \quad (9.22)$$

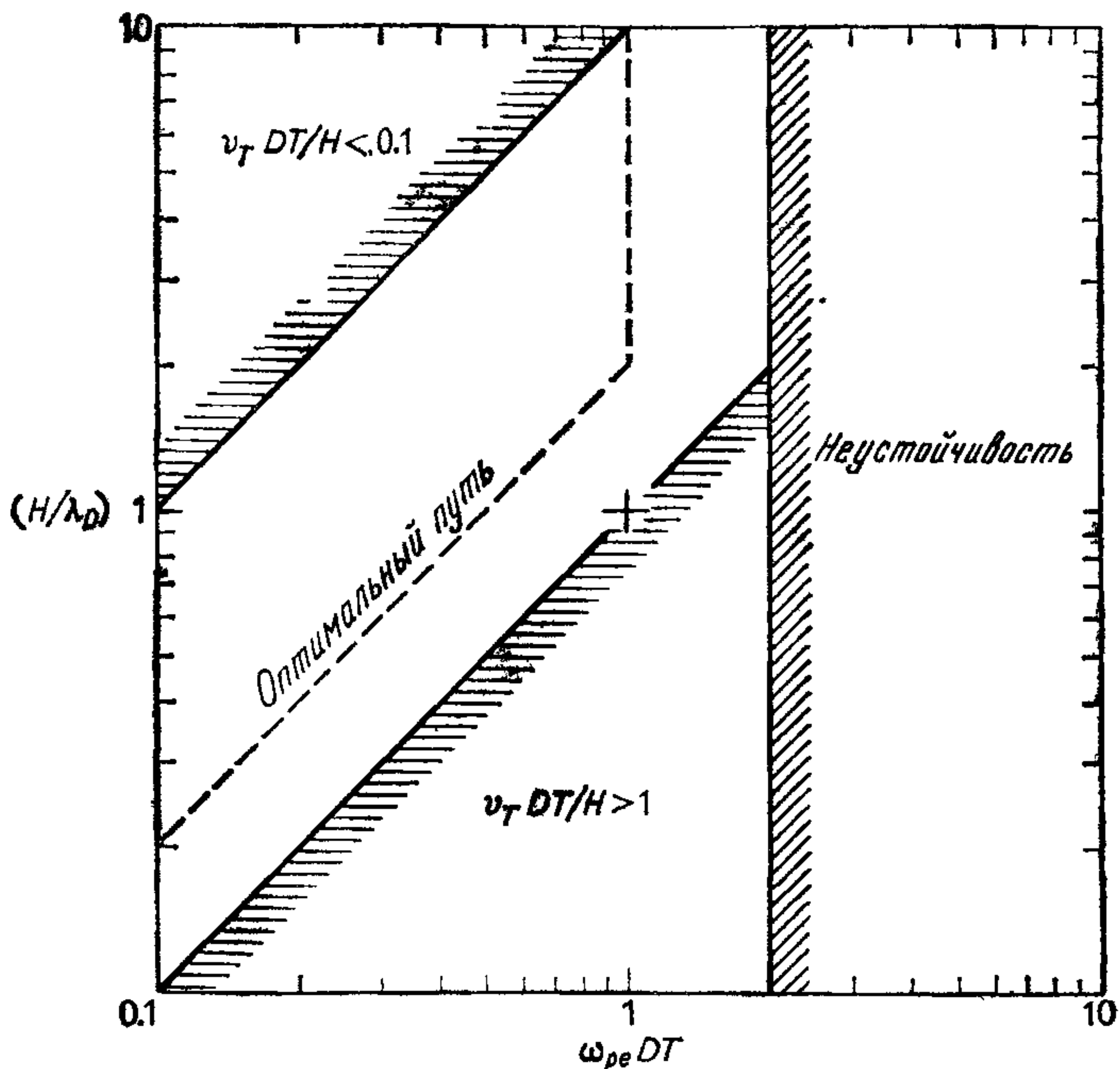


Рис. 9.6 Оптимальный путь для выбора шага по времени в электростатической модели плазмы. ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc)

где постоянная нагрева K_H сильно зависит от типа используемой модели:

$$\begin{aligned}
 K_H &= 2 \text{ для NGP с 5-точечным уравнением Пуассона,} \\
 &= 40 \text{ для CIC с 5-точечным уравнением Пуассона,} \\
 &= 150 \text{ для TSC с 9-точечным уравнением Пуассона,} \quad (9.23) \\
 &= 600^1) \text{ для CIC с } Q\text{-минимизацией,} \\
 &= 3000^1) \text{ для TSC с } Q\text{-минимизацией.}
 \end{aligned}$$

Эти результаты показывают поразительное улучшение качества модели, достигнутое путем распределения заряда высокого порядка и процедурами Q -минимизации сглаживания силы. Видно также, что существенные улучшения можно получить с помощью добавления Q -минимизации к модели CIC. Однако необходимо помнить, что использование схем высокого порядка

¹⁾ Результаты взяты из [190] и используют ранние эмпирические методы минимизации. Аналитические методы гл. 8 должны дать более высокие значения K_H

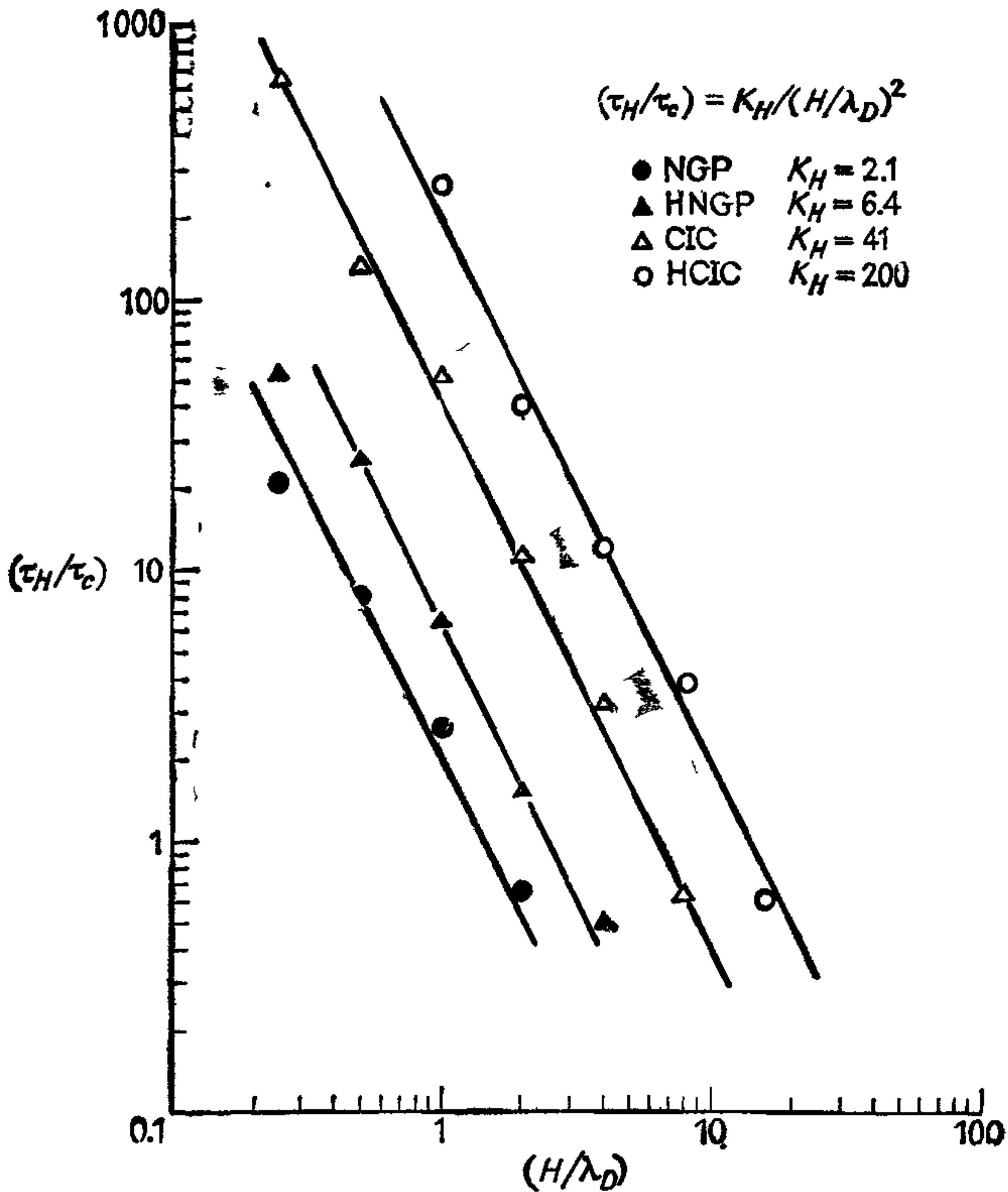


Рис. 9.7 Корреляция времени нагрева с шириной ячейки сетки вдоль оптимального пути. ([181], с любезного разрешения Journal of Computational Physics, © Academic Press, Inc)

приводит к большому расходу машинного времени. Ниже приводятся некоторые типичные затраты времени центрального процессора на IBM 360/195 для фазы ускорения частиц в двумерной электростатической модели плазмы:

NGP 1-точечная схема	23 мкс/частица,	
CIC 4-точечная схема	44 мкс/частица,	(9.24)
TSC 9-точечная схема	104 мкс/частица.	

Использование Q -минимизации не влияет на эти времена, поскольку оно только изменяет величину функции Грина, примененной при решении уравнения Пуассона. Числа представлены для кода на Фортране без оптимизации, и существенные улучшения, вероятно, должны быть получены тщательным ручным кодированием на языке ассемблера.

9.3. АНОМАЛЬНАЯ ДИФФУЗИЯ

Двумерная электростатическая модель частица — сетка была впервые применена для изучения аномально высокой диффузии плазмы поперек магнитного поля [176, 177]. Большинство установок, запланированных для производства энергии из дейтерия морской воды, предназначены для удержания очень горячей ($\sim 10^8$ К) дейтериево-третиевой плазмы внутри тороидального вакуумного резервуара с сильным магнитным полем. Чтобы избежать радиоактивности, экспериментальные лабораторные установки для изучения магнитного удержания обычно используют чисто дейтериевую или водородную плазму. Для моделирования любых этих установок необходимо моделировать движение заряженных частиц в сильном магнитном поле. Орбита такой частицы представляет спираль с осью, параллельной силовым линиям магнитного поля. Движение вдоль поля не подвергается воздействию магнитного поля, но движение, спроектированное на плоскость, перпендикулярную к полю, представляет окружность с радиусом, обратно пропорциональным силе магнитного поля. Кроме вращения частиц плазмы вокруг силовых линий магнитного поля, частицы плазмы дополнительно испытывают дрейфы, вызванные градиентом и кривизной магнитного поля, а также наличием электрических полей (см., например, [268, гл. VI]).

Основная проблема магнитного удержания заключается в нахождении конфигурации поля, внутри которой плазма может находиться достаточно долго, чтобы успевал произойти термоядерный синтез. В дополнение к температурным требованиям, $T \sim 10^8$ К, условие жизнеспособности термоядерного реактора выражено критерием Лоусона, согласно которому произведение плотности плазмы на время удержания должно превышать $\sim 10^{20}$ м⁻³ с. Если силовые линии магнитного поля движутся вместе с плазмой из внутренней области тороидальной установки, движение плазмы вдоль силовых линий поля не приводит к потере плазмы из этой установки. Потери возникают из-за движения частиц плазмы перпендикулярно к силовым линиям поля, время удержания уменьшается, и поэтому сейчас мы ограничимся рассмотрением этих двумерных движений в плоскости, перпендикулярной к силовым линиям магнитного поля. Кроме того, будем рассматривать плазму, содержащую электроны с зарядом $-e$ и массой m_e и одним сортом положительных ионов с зарядом $+e$ и массой m_i . Модуль электронного заряда e равен 1.60210×10^{-19} Кл.

Орбита частицы с зарядом $\pm e$, массой m и скоростью v , перпендикулярной к магнитному полю напряженностью B , представляет собой стационарное вращение по окружности с гирора-

диусом (или ларморовским радиусом)

$$a = mv/eB \quad (9.25a)$$

и гирочастотой

$$\Omega = eB/m. \quad (9.25b)$$

Следовательно, при отсутствии взаимодействия с другими частицами плазмы (т. е. нет столкновений и коллективных неустойчивостей) система должна идеально удерживать плазму. Каждое столкновение должно перемещать центр вращения в случайном направлении по вектору смещения δa , так что $0 \leq |\delta a| \leq 2a$. Поэтому берем среднее значение $|\delta a| = a$. В системе происходит диффузия плазмы поперек магнитного поля, подчиняющаяся уравнению для случайного блуждания (9.13). Среднеквадратичное смещение за счет диффузии после n шагов будет равно

$$\langle |\Delta a|^2 \rangle = n |\delta a|^2 = n \left(\frac{m^2 v^2}{e^2 B^2} \right) = \quad (9.26a)$$

$$= \frac{v_c m^2 v^2}{e^2 B^2} t, \quad (9.26b)$$

где $|\cdot|$ означает модуль вектора, а $v_c = n/t$ — физическую частоту столкновений. Коэффициент диффузии D_c обычно связан со среднеквадратичным смещением частицы [см. разд. 12.3.4 (6), уравнение (12.82)]

$$\langle |\Delta a|^2 \rangle = 2 d D_c t, \quad (9.27a)$$

где d — размерность задачи. В нашем случае рассматривается двумерная диффузия в плоскости, перпендикулярной магнитному полю, поэтому $d = 2$. Следовательно,

$$D_c = \frac{v_c}{2d} \frac{m^2 v^2}{e^2 B^2} = \frac{v_c}{4} \frac{v^2}{\Omega^2}. \quad (9.27b)$$

Если в выражении для интенсивности физических столкновений используются реальные параметры, то теория предсказывает очень слабую диффузию и отличное удержание плазмы. Первые термоядерные установки были построены на основе таких вселяющих надежду оценок. Именно это обстоятельство и явилось главным препятствием в развитии исследований по термоядерному синтезу, поскольку диффузия оказалась на много порядков величины большей, чем это давали оценки чисто столкновительной диффузии. Это явление стало широко известным как «аномальная диффузия». В результате этой диффузии установки не могли удерживать плазму достаточно долго, чтобы произошел термоядерный синтез.

Удобный масштаб при обсуждении явлений аномальной диффузии был введен Бомом:

$$D_B = \frac{1}{16} \frac{k_B T}{eV} = \frac{1}{16} \frac{v_T^2}{\Omega}. \quad (9.28)$$

Он получил название бомовского коэффициента и действительно соответствует аномальной диффузии, наблюдаемой в некоторых диапазонах параметров (см. разд. 9.3.5). Заметим, что бомовский коэффициент меняется как B^{-1} , тогда как коэффициент столкновительной диффузии пропорционален B^{-2} (см. уравнение (9.27b)). Поскольку аномальная диффузия наблюдалась и в случаях, когда все магнитогидродинамические (МГД) неустойчивости были подавлены, то естественно пытаться объяснить ее за счет эффектов, таких, как разделение зарядов, которыми пренебрегали в МГД-приближении. Для изучения этой проблемы была поэтому разработана двумерная электростатическая модель [176, 177] с использованием модели NGR и уравнений Ньютона, проинтегрированных с помощью методов разд. 4.7.1. Уравнение Пуассона решалось на сетке 48×48 с использованием алгоритма FACR(1) (см. разд. 6.5.3 и [175]).

9.3.1. Диффузионный эксперимент

Для исключения торцевых эффектов большинство термоядерных установок имеет тороидальную форму, и на рис. 9.8 (слева) дано поперечное сечение вакуумной камеры такой установки. Плазма образуется вблизи центра и удерживается сильным магнитным полем, направленным перпендикулярно к ри-

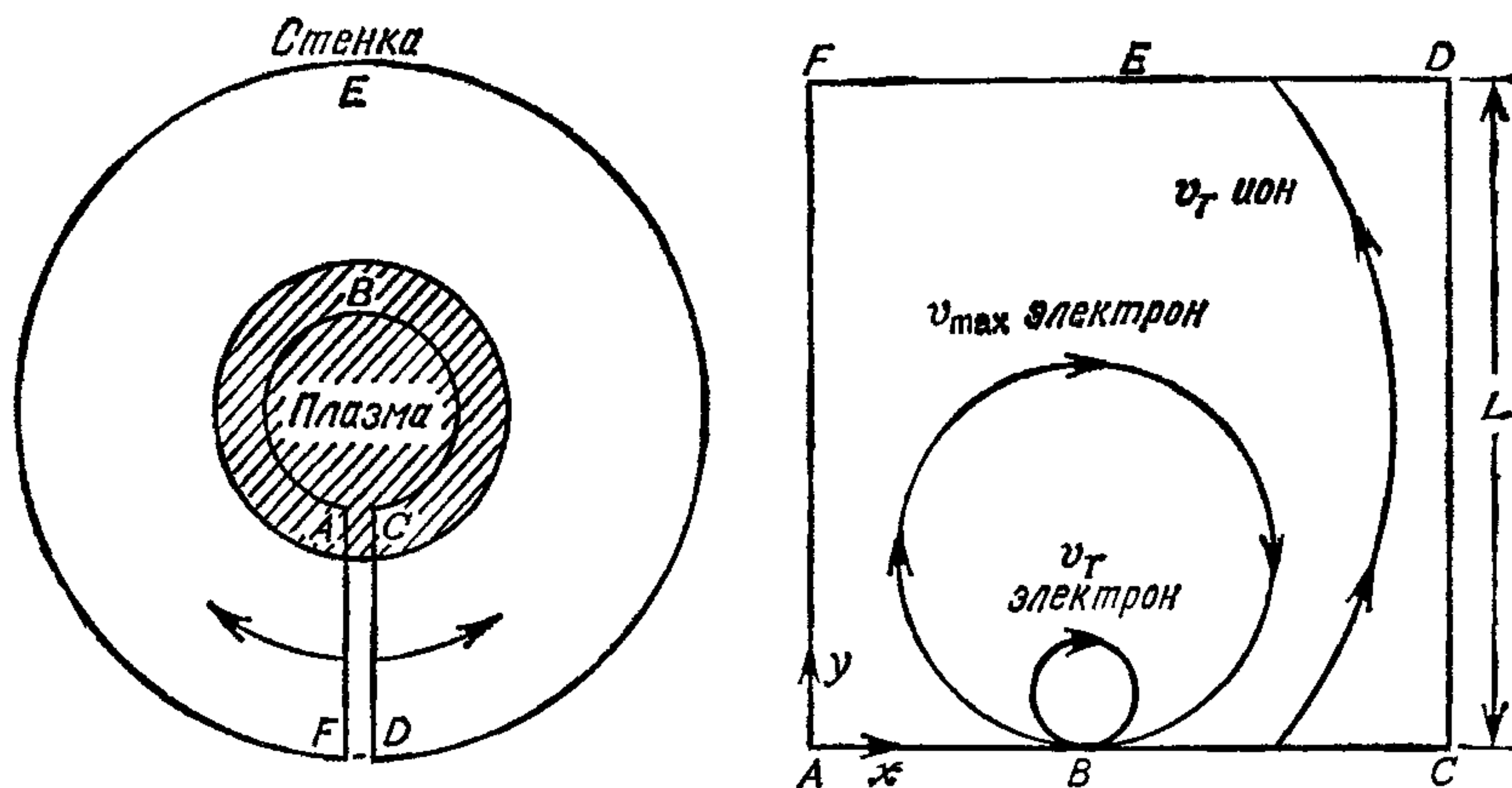


Рис 9.8 Геометрия типичной термоядерной установки (слева); идеализация, используемая в вычислительных целях (справа) [176].

сунку. Мы хотим моделировать диффузию плазмы поперек этого магнитного поля от центра к стенке. Описываемый эксперимент выполнен примерно в 1964—1965 г. на компьютере IBM 7090, когда для того, чтобы моделирование стало реальным, пришлось сделать значительные упрощения в геометрии.

Геометрия задачи показана на рис. 9.8 (справа). Она представляет собой квадратную область с плазменными частицами, инжектированными из поверхности основания ABC по направлению к стенке, представленной верхней поверхностью FED . Предполагается, что модель периодична в направлении x , так что поверхности AF и CD как бы примыкают друг к другу. Расчетную геометрию (справа) можно получить из фактической геометрии (слева) с помощью разреза вдоль контура $FABCD$ и развертывания цилиндра. Тогда периодичность по азимуту в реальной задаче переходит в периодичность по x в модели.

Определение масштаба также показано на рис. 9.8 (справа). Электрон с типичной тепловой скоростью $v_T = (k_B T/m)^{1/2}$ имеет гирорадиус меньше расстояния L между краем плазмы и стенкой. Даже максимальная скорость, заданная первоначально электронам, $v_{\max} \simeq 3v_T$, приводит к гирорадиусу, меньшему L . Следовательно, при отсутствии других эффектов электроны не могут достичь стенки. С другой стороны, ионы являются более тяжелыми, чем электроны, и, согласно (9.25а), имеют существенно больший гирорадиус. Следовательно, при наличии только внешнего поля B ионы вполне могут достичь стенки. Условие, заданное на стенке, состоит в том, что любые прибывающие заряды прилипают к ней в точке их столкновения со стенкой. Этот пристеночный заряд исчезает, только если он разряжается прибытием противоположно заряженных электронов.

Если в начальный момент область является пустой и инжекция плазмы начинается с нижней поверхности, ионы движутся к стенке, пока потенциал на стенке ϕ_W не станет достаточно положительным, чтобы начать отталкивать вновь прибывающие ионы. Этот эффект начнет сказываться, когда барьер потенциальной энергии, который должен преодолеваться тепловыми ионами, сравнивается с их характерной энергией, т. е.

$$e\phi_W \simeq k_B T, \quad (9.29a)$$

$$\phi_W \simeq \frac{k_B T}{e}. \quad (9.29b)$$

Электрическое поле, связанное с этим потенциальным барьером, перпендикулярно стенке и имеет значение

$$E_{\perp} \simeq \frac{k_B T}{eL}. \quad (9.30)$$

Действие электрического поля на вращение заряженной частицы должно вызывать дрейф центра вращения со скоростью, перпен-

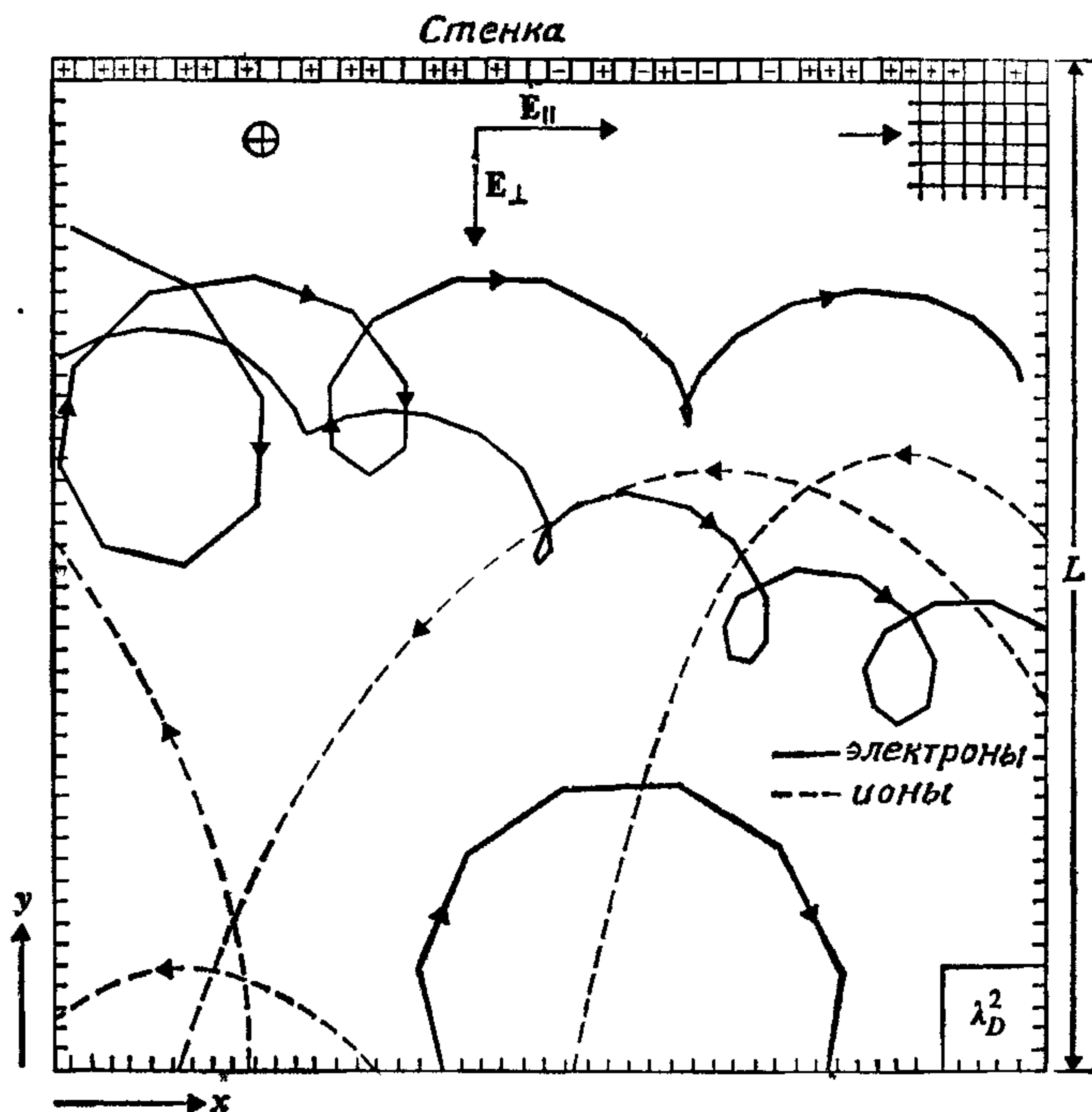


Рис. 9.9. Типичные орбиты частиц после того, как стенка заряжена положительно. Отмечен дрейф $\mathbf{E} \times \mathbf{B}$ электронов. Вектор \mathbf{B} направлен от читателя. ([177], с любезного разрешения Physics of Fluids, © American Institute of Physics.)

дикулярной полям \mathbf{E} и \mathbf{B} . Эта скорость называется «дрейфом ведущего центра» (или «дрейфом $\mathbf{E} \times \mathbf{B}$ ») и имеет величину

$$\mathbf{v}_g = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (9.31)$$

Важно заметить, что этот дрейф имеет равное значение для всех заряженных частиц, независимо от их знака, заряда или массы. Последние три величины определяют направление вращения, гиропериод и гирорадиус, но не скорость дрейфа центра вращения. На рис. 9.9 показаны орбиты типичных электронов и ионов в моделировании, после того как заряд на стенке устанавливается. Тяжелые ионы описывают примерно параболические орбиты, когда они «карабкаются» на потенциальный барьер и падают с него. Электроны вращаются и дрейфуют слева направо согласно (9.25), (9.31). В этой ситуации плазма все еще удерживается, поскольку хотя частицы дрейфуют, направление этого дрейфа всегда оказывается параллельным стенке и никогда не может быть направлено в ее сторону.

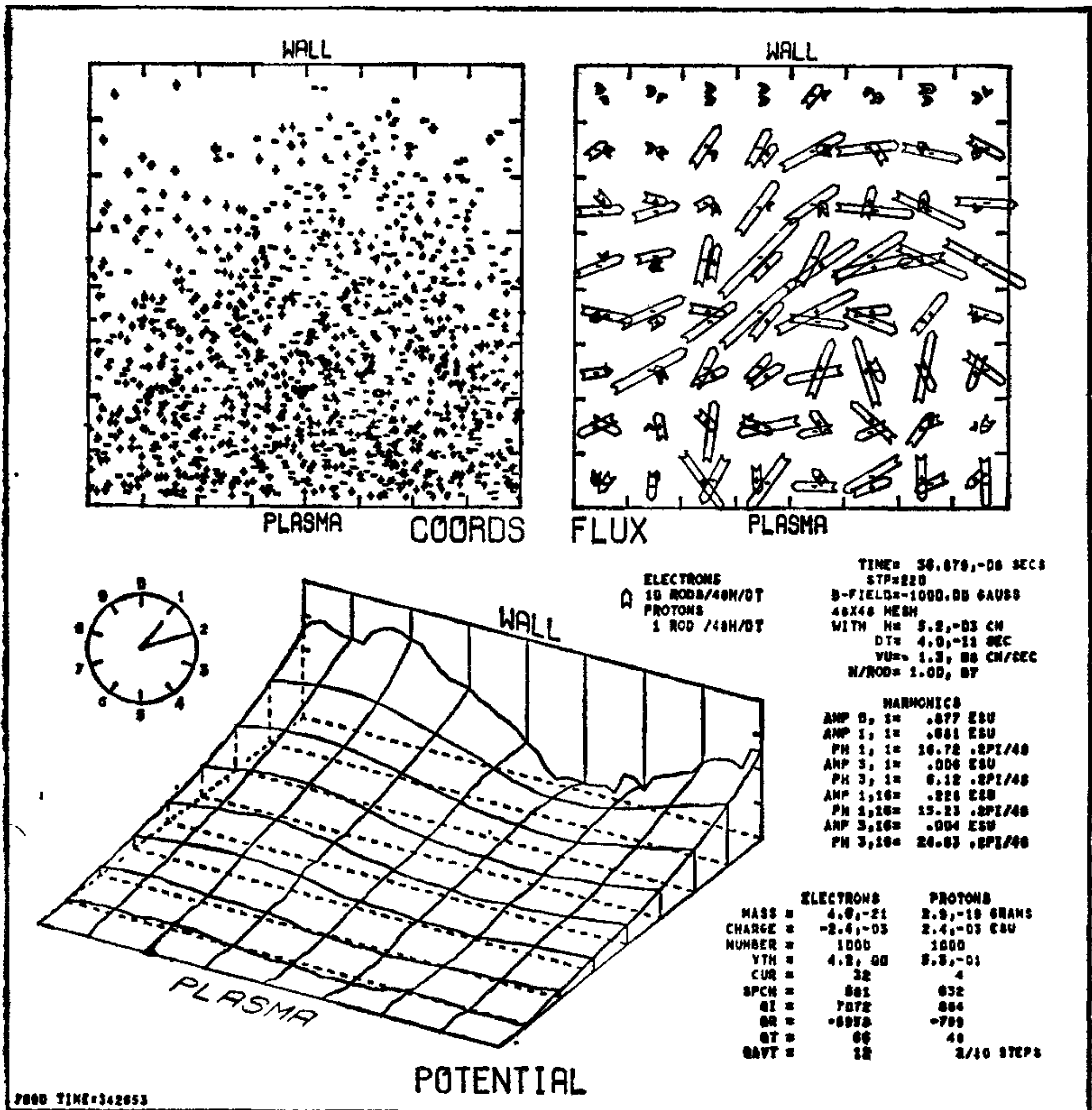


Рис. 9.10. Ситуация после развития сильной неустойчивости плазмы с $m = 1$. Стрелки (вверху справа) показывают течение плазмы к стенке, обусловленное электрическим полем волны E_{\parallel} . ([177], с любезного разрешения Physics of Fluids, © American Institute of Physics.)

Диффузия к стенке появится только тогда, когда будет существовать компонента электрического поля, параллельная стенке, E_{\parallel} . Такая компонента может возникнуть в связи с развитием неустойчивости в плазме. На рис. 9.10 показана ситуация в моделировании, после того как неустойчивость уже выросла до амплитуды насыщения. Возмущение потенциала, обусловленное неустойчивостью, имеет форму

$$\varphi' = A \frac{k_B T}{e} \cos\left(\frac{2\pi m x}{L} + \delta\right), \quad (9.32)$$

где A — амплитуда (в единицах $k_B T/e$), m — волновое число и δ — фаза. Из рис. 9.10 ясно, что преобладающая неустойчивость соответствует моде волны с $m = 1$. Амплитуда и фаза волны

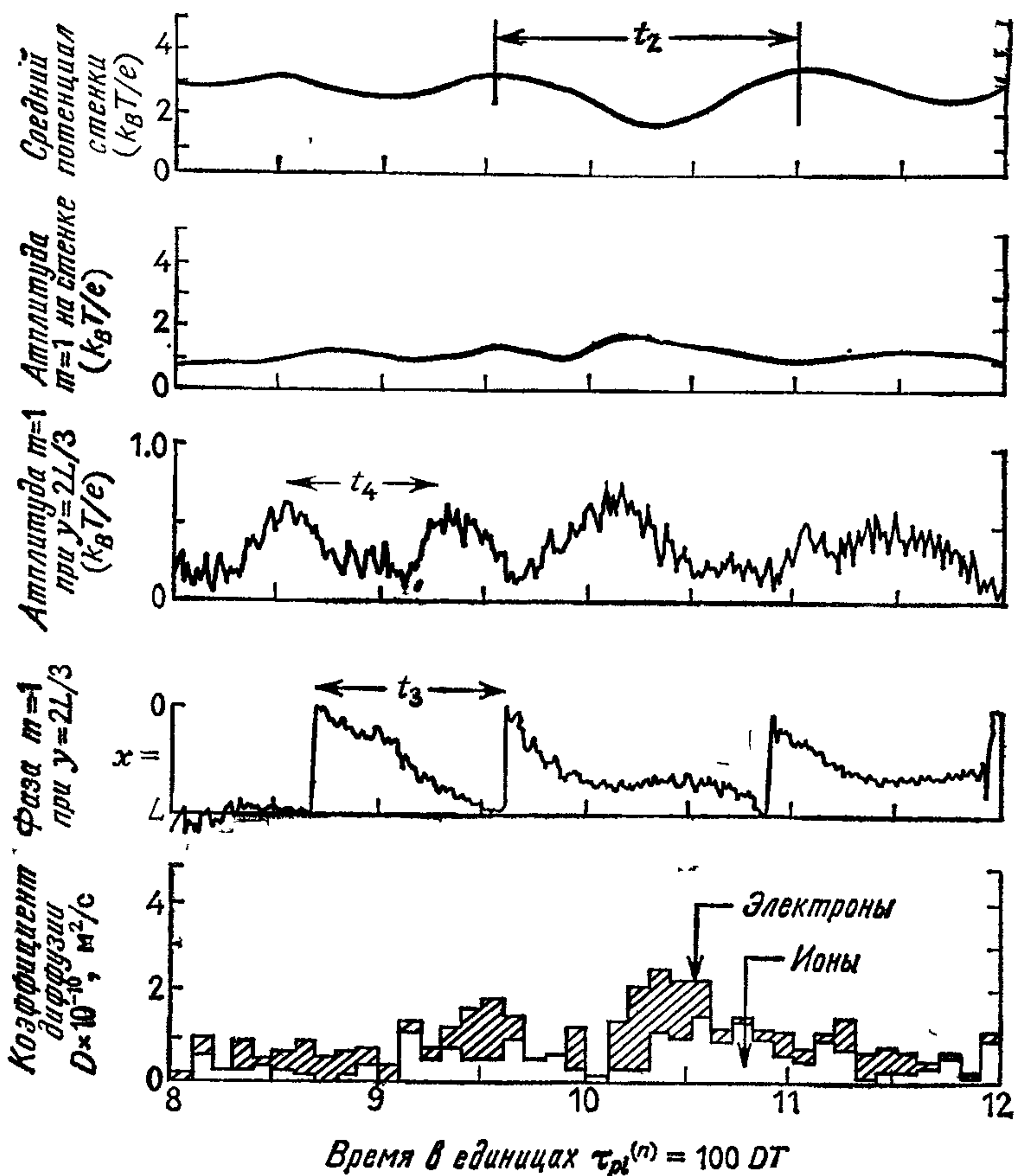


Рис. 9.11. Амплитуда и фаза неустойчивости с $m = 1$ и выброс плазмы на стенку во время практически стационарной диффузии. Единица времени $\tau_{pi}^{(n)} = 1.4\tau_{pi}^{(0)} = 100DT$. ([177], с любезного разрешения Physics of Fluids, © American Institute of Physics.)

с $m = 1$, на стенке и при $y = \frac{2}{3}L$ являются результатом нахождения решения для потенциала с помощью техники быстрого преобразования Фурье (FFT). Они показаны на рис. 9.11 вместе с прибывающими на стенку электронами и ионами.

Суммарную диффузию на стенку, вызванную волной, можно рассчитать следующим образом. Скорость дрейфа частиц к стенке равна (выбираем для удобства $\delta = 0$)

$$v_{\parallel} = \frac{E_{\parallel}}{B} = -\frac{1}{B} \frac{\partial \phi'}{\partial x} = \frac{A}{B} \frac{k_B T}{e} \frac{2\pi}{L} \sin \frac{2\pi x}{L}, \quad (9.33)$$

средний поток на стенку соответственно равен

$$\Gamma = \frac{n_W}{L} \int_{x=0}^{L/2} v_y dx = \quad (9.34a)$$

$$= 2 \frac{A}{B} \frac{k_B T}{e} \frac{n_W}{L}, \quad (9.34b)$$

где n_W — плотность плазмы на стенке, и мы проинтегрировали по той половине длины волны, для которой скорость дрейфа v_y положительна и направлена к стенке. Чтобы интерпретировать этот поток как диффузию, заметим, что по определению

$$\Gamma = -D \nabla n \simeq D \frac{n_p}{L}, \quad (9.35)$$

где $n_p (\gg n_W)$ — плотность плазмы (нижняя поверхность области). Следовательно,

$$D = 2A \frac{n_W}{n_p} \frac{k_B T}{eB}. \quad (9.36)$$

Сравнивая (9.36) с физической формулой (9.28), видим, что зависимость от температуры и поля B та же самая и поэтому

$$D = 32A \frac{n_W}{n_p} D_B = \quad (9.37a)$$

$$= 1.2D_B \quad (9.37b)$$

после введения значений A , n_W , n_p , полученных при моделировании. Измеренная скорость диффузии, рассчитанная по прибытию ионов и электронов на стенку, равна $0.9D_B$, и (9.37b) показывает, что в пределах погрешностей эксперимента вся эта диффузия может быть связана с воздействием волны с $m = 1$. Для подтверждения этого был проведен второй численный эксперимент, в котором амплитуда волны с $m = 1$ была искусственно установлена на нуль после расчета потенциала. Наблюдавшаяся при этом диффузия составляла меньше 4 % от значения в первоначальных экспериментах, подтверждая прямо сделанный выше вывод. Эта процедура иллюстрирует большую гибкость вычислительного эксперимента в определении механизма физического явления. В отличие от численного моделирования в лабораторном эксперименте было бы очень трудно подавить эту волну.

В описанной выше двумерной модели не учитывалось влияние ширины магнитного поля, и диффузия окажется больше, чем в физических экспериментах по достаточно долгому удержанию. Вычислительный код, включающий этот эффект, описан в разд. 9.3.5.

9.3.2. Простая количественная теория

Одна из возможностей численного эксперимента состоит в том, что он может указать направления построения простой аналитической теории. Сам эксперимент продемонстрировал, что неустойчивости действительно растут в ограниченной электростатической плазме и что они вызывают диффузию поперек поля, по величине подобную аномально высокой диффузии, наблюдавшейся в лаборатории. Цель теоретического анализа теперь состоит в попытке понять механизм неустойчивости с помощью простого рассмотрения. Наблюдение поведения ионов и электронов в численной модели указывает, что следующие упрощенные аналитические модели могли бы хорошо описать явление:

1. *Электроны.* Гирорадиусы относительно малы, и их движение в основном определяется дрейфом ведущего центра. Поэтому будем рассматривать электроны как жидкость с массой, равной нулю, движение которой подчиняется уравнению

$$\mathbf{v} = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (9.38)$$

2. *Ионы.* Ионные гирорадиусы велики, так что их гироорбиты незамкнуты. Движение ведущего центра отсутствует, и орбиты ионов в основном определяются их инерцией и электростатическим полем. Другими словами, магнитное поле не влияет на движение ионов. Уравнение их движения имеет вид

$$m_i \frac{d\mathbf{v}_i}{dt} = - \frac{1}{n_i} \frac{\partial p}{\partial \mathbf{x}} + e\mathbf{E}, \quad (9.39a)$$

где m_i , v_i и n_i — соответственно масса скорости и плотности ионов, а давление выбираем в форме

$$p = n_i k_B T. \quad (9.39b)$$

3. *Непрерывность.* Потери электронов и ионов нет, и уравнение непрерывности применяется к обеим жидкостям

$$\frac{\partial n_i}{\partial t} = - \frac{\partial}{\partial \mathbf{x}} (n_i \mathbf{v}_i), \quad (9.40a)$$

$$\frac{\partial n_e}{\partial t} = - \frac{\partial}{\partial \mathbf{x}} (n_e \mathbf{v}_e), \quad (9.40b)$$

где индекс e относится к электронам.

4. *Поля.* Магнитное поле постоянно в направлении z , и в модели пренебрегают самосогласованным полем от движущихся зарядов. Следовательно,

$$\mathbf{B} = (0, 0, B_z). \quad (9.41)$$

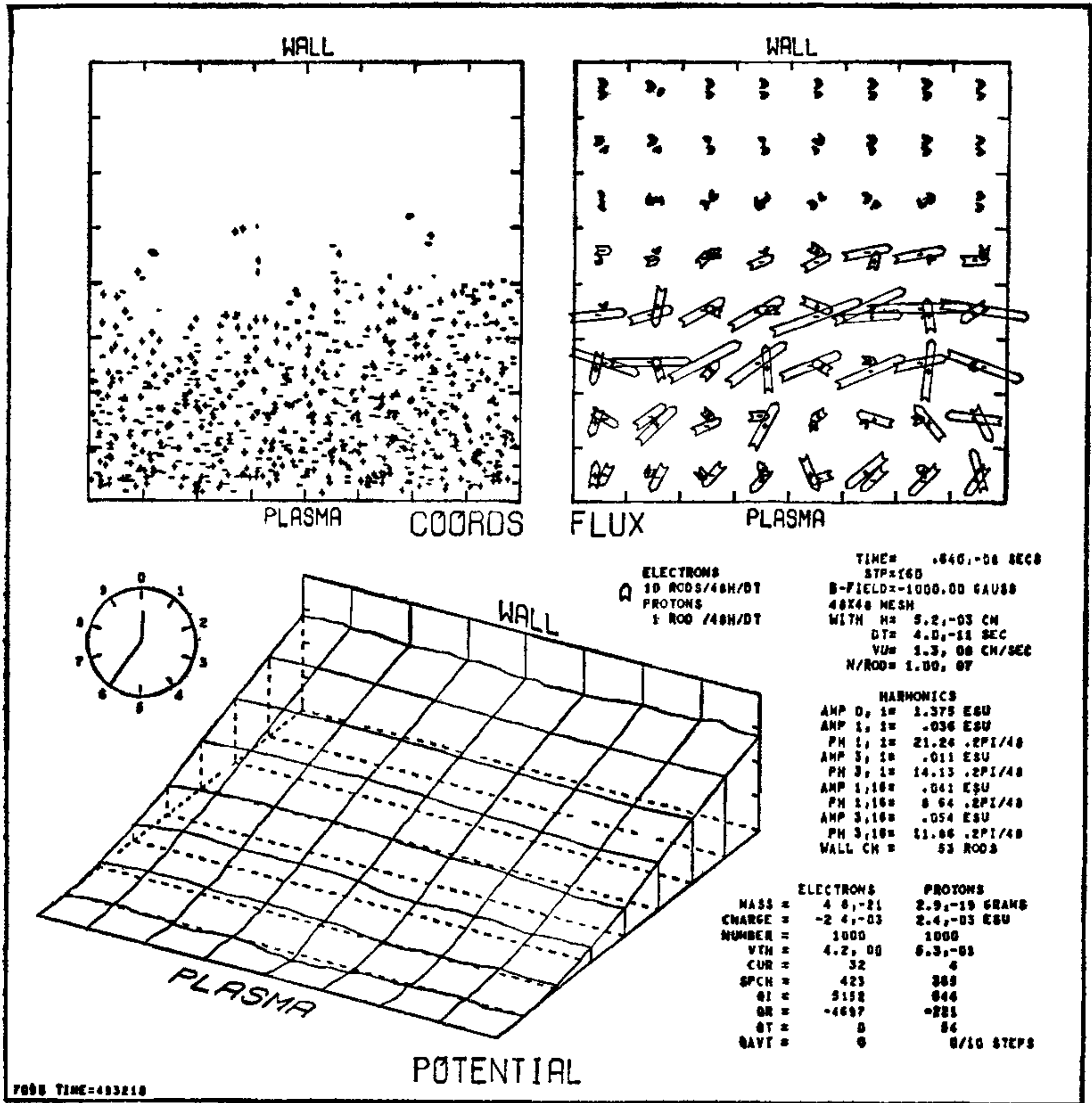


Рис 9.12 Состояние равновесия, предшествующее росту неустойчивости с $m = 1$. Поверхности потенциала показывают постоянное поле E , перпендикулярное к стенке, а стрелки — результирующий дрейф $E \times B$ электронов параллельно стенке. Это состояние является начальным для анализа неустойчивости. ([177], с любезного разрешения Physics of Fluids © American Institute of Physics.)

В случае если разделение зарядов существует, электрическое поле определяется с помощью уравнения Пуассона

$$\nabla^2 \phi = -\rho/\epsilon_0, \tag{9.42a}$$

$$E = -\nabla \phi. \tag{9.42b}$$

Данное выше описание соответствует двухжидкостной модели и остается только определить начальное равновесие (обозначенное величинами со звездочкой) и выполнить для него анализ устойчивости. Анализ результатов моделирования опять дает ключ к решению задачи. На начальной стадии эксперимента после того, как стенка зарядилась, и до того, как произошло развитие неустойчивости моды с $m = 1$, распределение потен-

циала является таким, как показано на рис. 9.12. Это состояние характеризуется следующим образом:

1. Постоянное поле E_y^* , перпендикулярное стенке,

$$\varphi^* = -E_y^* y. \quad (9.43a)$$

2. Дрейф ведущего центра для электронной жидкости задается выражением

$$v_x^* = \frac{E_y^*}{B_z}. \quad (9.43b)$$

3. Градиент ионной плотности определен распределением Больцмана

$$n_i^*(y) \sim \exp(-e\varphi^*/k_B T), \quad (9.44a)$$

$$n_i^*(y) \sim \exp(eE_y^* y/k_B T). \quad (9.44b)$$

4. Нейтральность заряда внутри области

$$n^*(y) = n_e^*(y) = n_i^*(y) \quad (9.45)$$

и впредь мы опускаем индекс переменной плотности.

Можно показать прямой подстановкой, что определенные выше равновесные величины удовлетворяют двухжидкостным уравнениям. Чтобы определить, является ли это равновесие устойчивым, мы налагаем малые возмущения на состояние равновесия, отмеченное штрихованными переменными

$$n(x, y, t) = n^*(y) + n'(x, y, t), \quad (9.46)$$

и разлагаем возмущенные величины

$$n'(x, y, t) = \hat{n}(y) e^{i(kx - \omega t)}, \quad (9.47)$$

где $k = 2\pi/L$ для моды $m = 1$.

Подстановка (9.47) в исходные уравнения и пренебрежение произведениями возмущенных величин приводят к набору линейных дифференциальных уравнений для $\hat{n}(y)$. Эти уравнения имеют решение только для некоторых значений частоты ω , которые являются собственными значениями уравнения. Эти уравнения можно решить с помощью разложения Фурье для $\hat{n}(y)$ и нахождения собственных значений результирующих уравнений для амплитуд гармоник.

Волна неустойчива, если собственное число ω комплексное,

$$\omega = \omega_r + i\gamma, \quad (9.48)$$

потому что в подстановке в (9.47) зависимость моды от времени имеет вид

$$e^{-i\omega t} = e^{\gamma t} e^{-i\omega_r t} \quad (9.49)$$

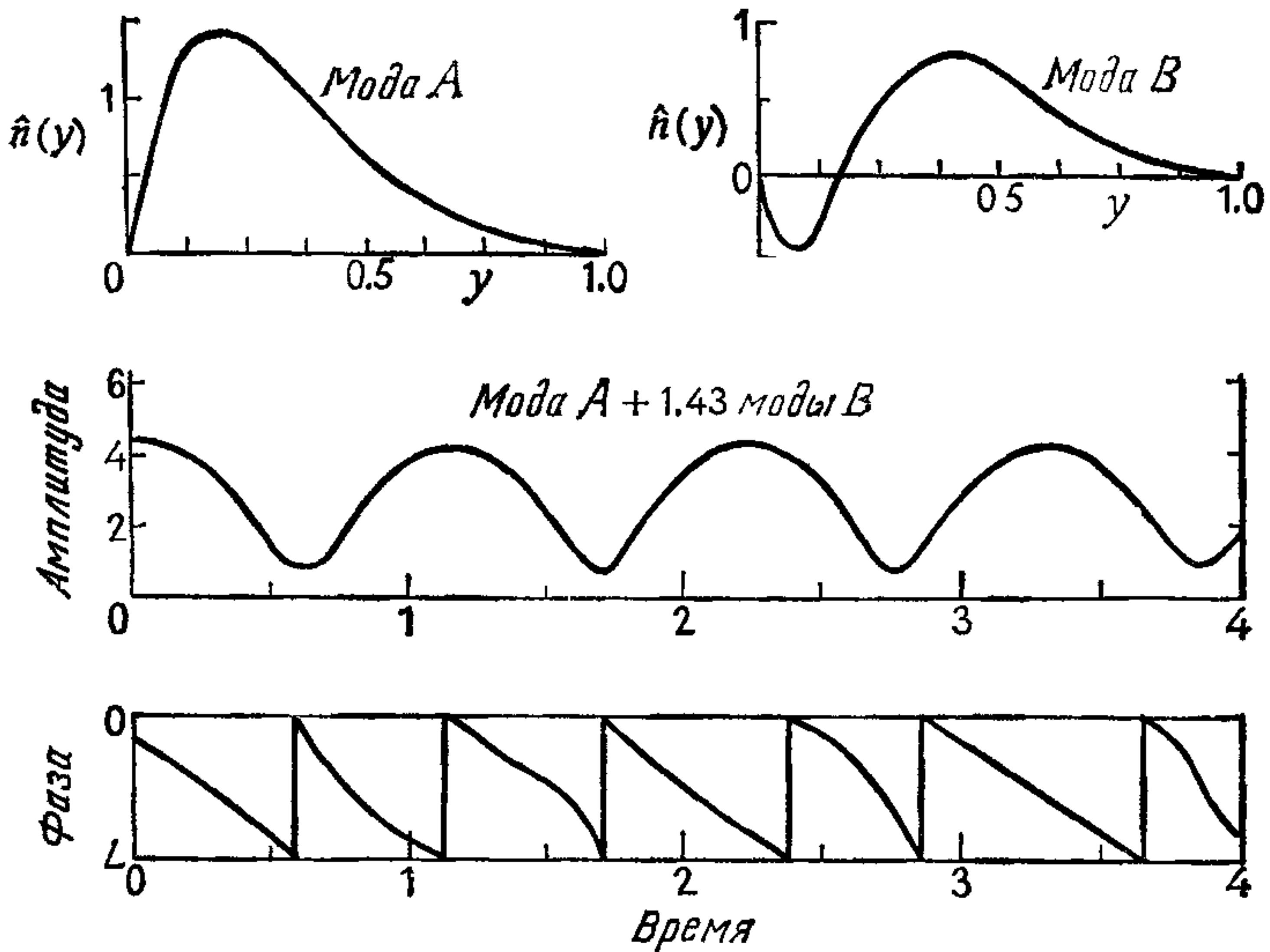


Рис. 9.13 Зависимость по y моды А (наверху слева) и моды В (наверху справа). Изменение соответственно амплитуды и фазы, происходящее в результате интерференции мод А и В, взятых в отношении 1 к 1.43 (в середине и внизу). Единица времени $\tau_{pi}^{(n)} = 1.4$; $\tau_{pi}^{(0)} = 100 DT$ [176].

и амплитуда возмущения растет экспоненциально со скоростью роста γ .

Результат этого анализа показывает, что две наиболее неустойчивые моды с $m = 1$ имеют вариации по y , как показано на рис. 9.13 (вверху). После учета стабилизирующего эффекта затухания Ландау частота ω , время прохождения t_3 и скорость роста γ для этих волн имеют значения

$$\begin{aligned} \text{мода А: } \quad \omega &= 0.52\omega_{pi}^{(0)}, \\ t_3 &= 1.9\tau_{pi}^{(0)}, \\ \gamma &= 0.34\omega_{pi}^{(0)}; \end{aligned} \tag{9.50a}$$

$$\begin{aligned} \text{мода В: } \quad \omega &= 1.15\omega_{pi}^{(0)}, \\ t_3 &= 0.87\tau_{pi}^{(0)}, \\ \gamma &= 0.14\omega_{pi}^{(0)}, \end{aligned} \tag{9.50b}$$

где $\omega_{pi}^{(0)}$ и $\tau_{pi}^{(0)}$ — ионная плазменная частота и период на плоскости инъекции.

В численном эксперименте (см. рис. 9.11) мы наблюдаем, что фаза неустойчивой волны с $m = 1$ перемещается поперек области с временем прохождения, варьирующимся между $0.7 \tau_{pi}^{(0)}$ и $\tau_{pi}^{(0)}$. В то же самое время амплитуда волны осциллирует

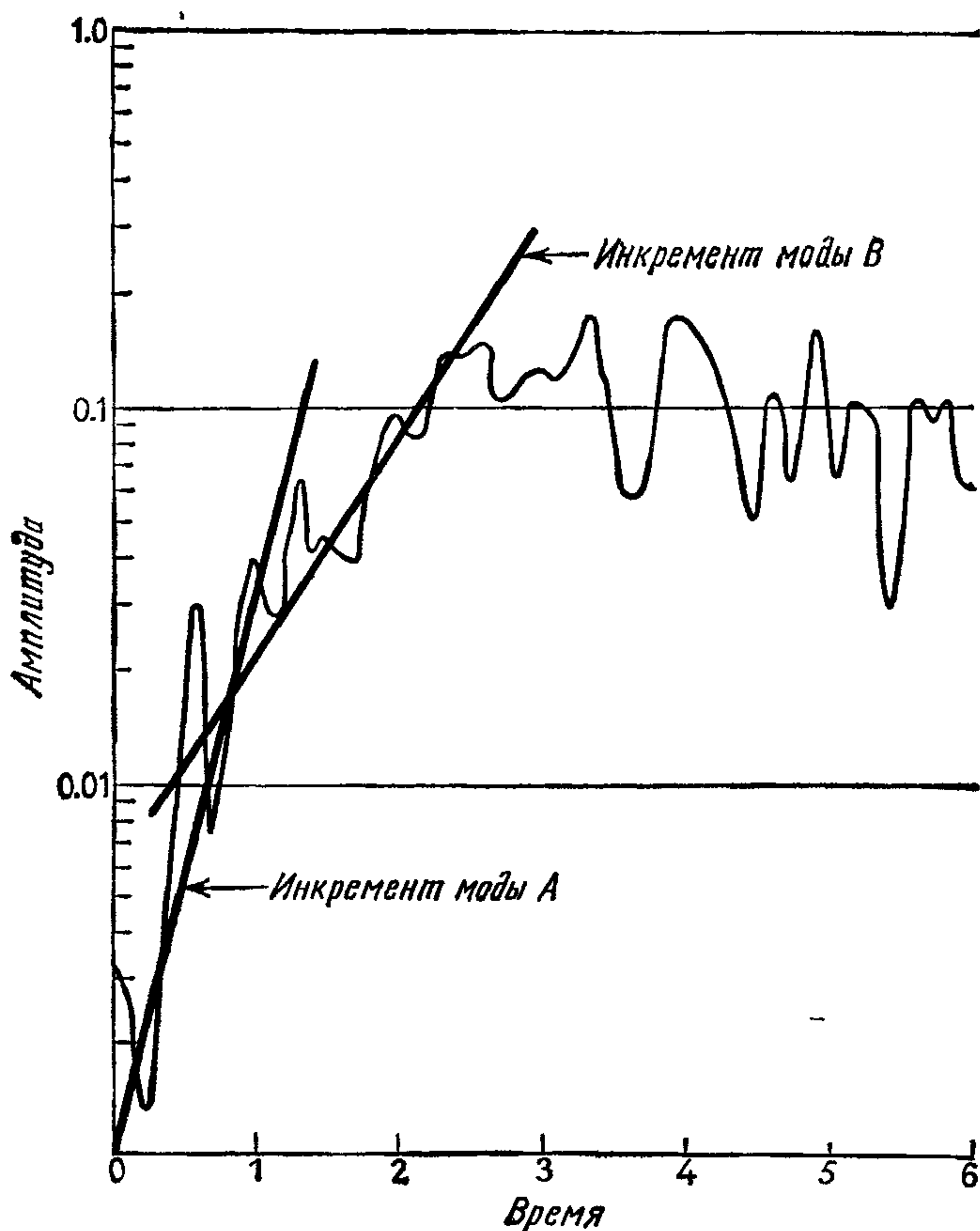


Рис. 9.14. Первоначальный рост неустойчивости, проанализированный с точки зрения мод A и B . Единица времени $\tau_{pi}^{(n)} = 100 DT$ [176].

с периодом $t_4 \approx 0.7\tau_{pi}^{(0)}$. Одиночная неустойчивая мода будет перемещаться поперек системы с постоянными амплитудой и временем прохождения. Изменение в амплитуде, наблюдавшееся в эксперименте, и изменение времени прохождения могут быть объяснены только интерференцией двух волн, движущихся со слегка отличающимися скоростями. На рис. 9.13 (в середине, внизу) показано изменение полной амплитуды и фазы, полученное при интерференции мод A и B в отношении 1:143. Этот результат находится в отличном согласии с численным экспериментом, представленным на рис. 9.11.

Дополнительно подтверждает теорию тот факт, что начальный рост неустойчивости определяется двумя инкрементами, что можно объяснить тем, что растущие моды соответствуют модам A и B . Это показано на рис. 9.14.

9.3.3. Выбор временного шага и размера ячейки

Выбор связан с характерными физическими частотами и пространственными масштабами задачи. Перечислим их:

электронная плазменная частота

$$\omega_{pe} = \left(\frac{n_e e^2}{\epsilon_0 m_e} \right)^{1/2}, \quad (9.51a)$$

электронная гирочастота

$$\Omega_e = \frac{eB}{m_e}, \quad (9.51b)$$

ионная плазменная частота

$$\omega_{pi} = \left(\frac{n_i e^2}{\epsilon_0 m_i} \right)^{1/2}, \quad (9.51c)$$

ионная гирочастота

$$\Omega_i = \frac{eB}{m_i}, \quad (9.51d)$$

дебаевская длина

$$\lambda_D = \left(\frac{k_B T \epsilon_0}{n e^2} \right)^{1/2}, \quad (9.51e)$$

электронный гирорадиус

$$a_e = \frac{m v_{Te}}{eB} = \frac{v_{Te}}{\Omega_e}, \quad (9.51f)$$

где характерная тепловая скорость определена как

$$v_{Te} = (k_B T / m_e)^{1/2}, \quad (9.51g)$$

ионный гирорадиус

$$a_i = \frac{m v_{Ti}}{eB} = \frac{v_{Ti}}{\Omega_i}, \quad (9.51h)$$

$$v_{Ti} = (k_B T / m_i)^{1/2}. \quad (9.51i)$$

Эти формулы показывают, что отношение электронной частоты к плазменной равно квадратному корню из отношения масс иона и электрона. Поскольку отношение физических масс иона к электрону для дейтериевой плазмы приблизительно равно 4000, отношение частот в физической задаче очень большое. Временной шаг должен быть выбран таким, чтобы $\omega DT < 2$ для наибольшей частоты в задаче. Это означает, что если бы мы использовали реальное отношение масс, то это потребовало бы несколько тысяч временных шагов, чтобы описать хотя бы одно колебание наименьшей частоты. Этот факт представляет собой основную трудность микроскопического моделирования методом частица — сетка и препятствует использованию физического отношения масс в практических расчетах. Однако если для раз-

деления частот использовано достаточно большое отношение масс и результаты выражены через безразмерные соотношения, то результаты могут вполне обоснованно применяться для реальных физических систем. Типичный выбор, использованный в этом эксперименте, состоит в следующем:

$$m_i/m_e = 64, \quad (9.52a)$$

следовательно,

$$\omega_{pe}/\omega_{pi} = 8, \quad (9.52b)$$

$$\Omega_e/\Omega_i = 64. \quad (9.52c)$$

Наибольшими частотами в задаче являются электронная плазменная и гирочастоты. Временной шаг должен быть выбран так, чтобы удовлетворить критерию устойчивости для схемы с перешагиванием $\omega DT < 2$. Для интегрирования с разумной точностью электронных орбит требуется значительно меньший временной шаг. Выбранный в моделировании временной шаг удовлетворяет условиям

$$\omega_{pe}DT = 0.72, \quad (9.53a)$$

$$\Omega_e DT = 0.7, \quad (9.53b)$$

$$\omega_{pi} DT = 0.09, \quad (9.53c)$$

$$\Omega_i DT = 0.01. \quad (9.53d)$$

При таком выборе временного шага ионные орбиты будут вычислены точно. С другой стороны, электронные орбиты имеют только ~ 9 шагов на гиропериод. Хотя это интегрирование ги-роорбиты и является устойчивым, оно имеет низкую точность. Используемые конечно-разностные схемы являются такими же, как в гл. 4 (уравнения (4.96) и (4.97)), и там показано, что дрейф ведущего центра определен точно даже для больших временных шагов (уравнение (4.98)). Развитая теория неустойчивости показала, что неустойчивость вызывается скоростью дрейфа, а не деталями гироорбиты. Следовательно, неточность в расчете электронной гироорбиты не является серьезной ошибкой.

Пространственный размер ячейки H должен быть выбран меньше, чем любой физический размер, который требуется разрешить. Выбранная величина удовлетворяет соотношениям

$$H/\lambda_D = 0.23, \quad (9.54a)$$

$$H/a_e = 0.12, \quad (9.54b)$$

$$H/a_i = 0.002, \quad (9.54c)$$

$$H/L = 1/48 = 0.02, \quad (9.54d)$$

$$H/\lambda_1 = 0.02, \quad (9.54e)$$

где λ_1 — длина неустойчивой волны в $m = 1$.

Можно проверить определенный выше выбор временно́го шага и пространственной ячейки в связи с результатами исследования столкновений и «дробового» шума, данных в разд. 9.2.1—9.2.3. Известно, что столкновительные эффекты вызывают диффузию плазмы поперек магнитного поля, и наши вычислительные модели значительно повышают столкновительные эффекты по сравнению с физической плазмой. Поэтому важно с приемлемой точностью оценить столкновительные эффекты, возникающие в модели, и показать, что они не являются причиной какой-либо наблюдающейся диффузии. Если это так, то следовало бы изучить неточную модель и измеренную в ней диффузию, обусловленную численными эффектами, которые не имеют ничего общего с процессами в физической плазме. В нашей модели именно перенос электронов поперек магнитного поля контролирует течение всей плазмы к стенке. На движение ионов магнитное поле почти не действует, и нужное число ионов всегда будет достигать стенки, чтобы нейтрализовать заряд любых прибывающих электронов. Следовательно, в последующем обсуждении необходимо рассматривать только диффузию электронов.

Вычисляя отношение диффузии, обусловленной столкновительными эффектами D_c [уравнение (9.27b)] и наблюдаемой аномальной диффузией D_B [уравнение (9.28)], получим

$$\frac{D_c}{D_B} = 16 \frac{\Omega_e}{v_{Te}^2} \frac{v_c}{4} \frac{v_{Te}^2}{\Omega_e^2} = 4 \left(\frac{v_c}{\Omega_e} \right). \quad (9.55)$$

Частота столкновений v_c в модели может быть получена из (9.19b), тогда

$$\frac{D_c}{D_B} = \frac{4}{2\pi N_C} \left(\frac{\omega_{pe}}{\Omega_e} \right). \quad (9.56a)$$

В плоскости, где происходит инжекция плазмы, имеем

$$N_D = n\lambda_D^2 \simeq 10, \quad (9.56b)$$

$$\left(\frac{W}{\lambda_D} \right) = 0.2, \quad (9.56c)$$

$$\frac{\omega_{pe}}{\Omega_e} \simeq 1, \quad (9.56d)$$

поэтому

$$N_C \simeq N_D (1 + (W/\lambda_D)^2) \simeq 10, \quad (9.56e)$$

$$(D_c/D_B) |_{\text{плазма}} = 0.06. \quad (9.56f)$$

Следовательно, можно сделать вывод, что на плоскости, где осуществляется инжекция, около 6% наблюдаемой диффузии связано с нефизическими эффектами (столкновениями за счет

дискретности модели). Для оценки столкновительной диффузии в каком-либо месте плазмы необходимо учесть изменение ее плотности.

В двумерной плазме с постоянной температурой число частиц на квадрате со стороной, равной дебаевской длине, не зависит от плотности из-за того, что

$$N_D = n\lambda_D^2 = \frac{k_B T e_0}{q^2}, \quad (9.57)$$

где q — заряд модельных макрочастиц на единицу длины. К тому же, поскольку Ω_e не зависит от плотности (см. уравнение (9.51b)) и ω_{pe} пропорционально $n^{1/2}$ (см. (9.51a)), заключаем

$$D_c/D_B \sim n^{1/2}. \quad (9.58)$$

Плотность на стенке наблюдалась равной 1/100 плотности плазмы, следовательно,

$$(D_c/D_B)|_{\text{стенка}} = 0.6\%. \quad (9.59)$$

Поэтому ясно, что вблизи стенки лишь небольшую часть наблюдаемой диффузии можно отнести за счет ошибок моделирования. Оценки, сделанные выше, подтверждаются тем фактом, что измеренная диффузия в случае искусственного подавления волны с $m = 1$ (4% D_B) лежит как раз между двумя рассчитанными значениями.

9.3.4. Обсуждение эксперимента

Описанный нами численный эксперимент был одним из самых первых по моделированию аномальной диффузии, и теперь отчетливо видны все его недостатки:

1. *Геометрия.* Идеализированная декартова геометрия не учитывает некоторых важных особенностей реальных установок, например кривизну силовых линий поля и изменение силы поля с радиусом. Кроме того, не учитывается изменение направления силовых линий с радиусом (или шир), которое является важным стабилизирующим механизмом в термоядерных устройствах.

2. *Размер.* Линейный размер вычисляемой системы ($\sim 10\lambda_D \approx 0.2 \times 10^{-2}$ м) по крайней мере в сто раз меньше, чем физическая система. К тому же ионный гирорадиус оказывается больше размеров установки, тогда как в большинстве реальных устройств ситуация оказывается противоположной.

3. *Граничные условия.* Условие на стенке выбрано так, чтобы соответствовать изолирующему резервуару, сделанному, например, из стекла. Эксперимент показывает, что распределение

потенциала на стенке имеет важное, даже определяющее влияние на движение плазмы. Поскольку большинство термоядерных установок имеет металлическую стенку, более реалистичное условие должно было бы описывать проводящую стенку с постоянным потенциалом. Плотность плазмы быстро уменьшается по мере удаления от нижней границы; отсюда следует, что профиль плотности определяется скорее положением инжектирующей поверхности, чем физическим состоянием плазмы. Может быть, эти довольно искусственные граничные условия в сильной степени влияют на полученные результаты.

4. *Динамика.* В физической плазме движение в направлении z , параллельное линиям поля B , является важным механизмом для нейтрализации флуктуаций заряда. Поскольку диффузия происходит вследствие таких флуктуаций заряда, результаты могли бы нарушиться из-за пренебрежения в модели движением в направлении z .

5. *Временной шаг.* Исследования двумерной электростатической плазмы в разд. 9.2.3 показали, что оптимальный выбор временного шага определяется в нашем масштабе (уравнение 9.54а) с помощью выражения

$$\omega_{pe} DT |_{\text{opt}} = \min \left[\frac{1}{2} (H/\lambda_D), 1 \right] = 0.12. \quad (9.60)$$

Действительное значение используемого временного шага в шесть раз больше, а именно $\omega_{pe} DT = 0.72$ (см. 9.53а). В этом случае среднее число ячеек сетки, пересекаемых тепловым электроном за временной шаг, равно

$$\frac{v_{Te} DT}{H} = \frac{\omega_{pe} DT}{H/\lambda_D} = 3. \quad (9.61)$$

Использование такого большого значения $v_{Te} DT/H$, значительно превышающего оптимальное значение $1/2$, является неудовлетворительным и приводит к значительным нефизическим эффектам нагрева.

9.3.5. Модели с размерностью $2^{1/2}$ и 3

Чтобы преодолеть некоторые из этих возражений, Доусон, Окуда и Розен [105] разработали представленную ниже электростатическую модель и провели большую серию экспериментов. В их так называемой модели с размерностью $2^{1/2}$ допускается движение в направлении z и для каждой частицы рассматриваются пять координат, а именно (x, y, v_x, v_y, v_z) . Включение компоненты скорости v_z позволяет грубо учесть магнитное отражение частиц в реальной установке. Верхняя граница квадратной расчетной области представляет магнитное зеркало.

Когда частицы приближаются к этой границе, вычисляется отношение компонент скоростей, параллельной и перпендикулярной \mathbf{B} . Если это отношение меньше, чем $[B_{\max}/B_{\min} - 1]^{1/2}$, где B_{\max}/B_{\min} — пробочное отношение, то частица отражается и компонента скорости, параллельная \mathbf{B} , меняет знак. В противном случае частица переходит через границу и входит в систему с другой стороны.

Эта модель допускает также, что поле \mathbf{B} располагается под углом к оси z , следовательно, существуют все компоненты поля $\mathbf{B} = (B_x, B_y, B_z)$. Вращение вокруг таких силовых линий поля связано со всеми тремя компонентами скорости $\mathbf{v} = (v_x, v_y, v_z)$, и учет v_z необходим для описания этого более общего случая. Поскольку модель находится на полпути (она не содержит переменной z) к трехмерной модели, она называется «модель с размерностью $2^{1/2}$ ». Магнитный шир может быть представлен в этой модели с помощью линейной вариации направления \mathbf{B} либо по x , либо по y .

Чтобы избежать искусственных эффектов, обусловленных нереальными граничными условиями, диффузия исследуется в однородной плазме, где отсутствует дрейфовое движение. Ведется наблюдение за движением пробной частицы из некоторого начального положения. Если Δr есть смещение за время t , то, поскольку рассматривается двумерный случай ($d = 2$), уравнение (9.27a) показывает, что

$$D = \frac{1}{4} \lim_{t \rightarrow \infty} \langle |\Delta r|^2 \rangle / t, \quad (9.62)$$

где ломаные скобки означают усреднение по частицам системы. Используя этот метод измерений, авторы нашли аномально высокую диффузию даже в однородной устойчивой тепловой плазме. Эта диффузия обусловлена конвективными движениями $\mathbf{E} \times \mathbf{B}$, связанными с тепловыми флуктуациями в поле \mathbf{E} . Этот эксперимент принципиально отличается от описанного ранее в этой главе, в котором аномальная диффузия возникала из-за электростатической неустойчивости плазмы. Ясно, что аномальная диффузия может возникать из-за различных причин.

Численные эксперименты Доусона, Окуды и Розена [105] проведены для различных значений поля \mathbf{B} . В результате были обнаружены три области, соответствующие различным режимам. В слабых магнитных полях $(\omega_{pe}/\Omega_e) > 3$ диффузия менялась как B^{-2} , что находилось в согласии с классической столкновительной теорией (уравнение (9.27b)). Для промежуточного диапазона $0.3 < (\omega_{pe}/\Omega_e) < 3$ диффузия приблизительно не зависит от B , а в сильных полях $(\omega_{pe}/\Omega_e) < 0.3$ диффузия аномальна и согласуется с бомовской диффузией (aB^{-1} , уравнение (9.28)).

Дальнейшие эксперименты были выполнены для неоднородной плазмы с градиентом плотности. В этом случае плазма не-

устойчива относительно желобковой дрейфовой волны и далее диффузия усиливается (в 3 раза) дрейфами $\mathbf{E} \times \mathbf{B}$, связанными с электрическим полем этой волны. Как и в экспериментах, описанных в данной главе, этот пример аномальной диффузии обусловлен неустойчивостью плазмы. Имеется весьма ограниченное число трехмерных экспериментов, в которых рассматривалось шесть координат (x, y, z, v_x, v_y, v_z) для каждой частицы. Эти эксперименты демонстрируют три области диффузии, но переход к бесстолкновительной диффузии появляется в более сильных полях при $\omega_{pe}/\Omega_e \simeq 1$.

Обобщение, в котором \mathbf{B} может быть наклонено к оси z , позволяет рассмотреть три разных случая, которые могут быть важны для исследований в тороидальных термоядерных установках. Если \mathbf{B} параллельно z , то силовые линии магнитного поля замыкаются после одного прохождения системы. При других наклонах поле \mathbf{B} может либо замыкаться после конечного числа прохождений, либо продолжать эргодически наполнять систему без замыкания.

Отсюда ясно, что относительно простой учет дополнительной координаты v_z (и соответствующего уравнения движения) значительно расширяет применение рассматриваемой модели. Однако необходимо указать, что модель размерности $2^{1/2}D$ вызывает сомнения в ее адекватности из-за представления тороидальной установки с помощью декартовых координат и из-за малого размера модельной системы.

9.3.6. Диагностика и визуализация

Основной особенностью численных экспериментов является то, что затраты времени на программирование диагностических программ и подпрограмм визуализации часто превышают затраты на основной цикл временного шага. Это имеет место даже в случае программирования с использованием системы OLYMPUS (см. гл. 3), которая снабжает некоторыми стандартными пакетами ввода и вывода. С другой стороны, полезность кода в основном зависит от способа, которым должна задаваться входная информация, и от способа анализа и представления результатов. Эти аспекты программирования радикально отличаются для различных приложений, и достигнута лишь очень ограниченная стандартизация. Имеются некоторые полезные программы, за исключением, возможно, печатания таблиц с числами.

Интерпретация результатов значительно облегчается с помощью их представления в графической форме, и в этой главе показаны несколько примеров на рисунках. К сожалению, до сих пор нет стандартного языка для описания графического вывода, хотя имеется несколько кандидатур (например, система Culham GHOST, GINO и некоторые коммерческие изделия). Вы-

вод состоит в том, что графические подпрограммы должны часто писаться с использованием языка очень низкого уровня и обычно не переносятся без труда с одного компьютера на другой. Рассмотрим теперь некоторые методы диагностики и визуализации, которые удобны для моделирования методом «частица — сетка».

Неразумно вычислять вслепую, не имея возможности подробно увидеть, как происходит моделирование. На самом низком уровне является существенной возможностью рассматривать орбиты нескольких (допустим, 10) частиц в моделировании. Для этой цели можно запомнить положения выбранных частиц, скажем, для 100 шагов, а потом их начертить на перьевом самописце «Калькомп». Такой график показан на рис. 9.9. В такой демонстрации непосредственно выделяются грубые ошибки в выборе временного шага и другие неправильные представления в масштабировании. Если используется графический дисплей с заполнением (например, Tektronix 4010) или аппарат для записи микрофильмов (например, IIFR80), необходимо только начертить положения выбранных частиц на каждом временном шаге и орбиты будут постепенно строиться в виде серии отдельных точек. Такая кривая показана на рис. 9.1. Если орбиту трудно разглядеть, то, вероятно, временной шаг слишком велик и его следует уменьшить.

На рис. 9.10 демонстрируется крайний случай, в котором нанесены все качественные особенности эксперимента вместе с ключом входных данных. Такая демонстрация содержит около 10 000 точек и линий и может быть выполнена только на высокоскоростном устройстве для записи микрофильма. На таком устройстве одна картина может наноситься за несколько секунд. Для этого рисунка использовалось устройство для записи микрофильма Stromberg — Carlson 4020 (рис. 9.15). Этот прибор воспроизводит символы на любом (1024×1024) растере возможных для адресации точек. Линии также могут быть отрисованы между парами таких точек. Задания на рисование, сформированные во время выполнения программы моделирования, обычно запоминаются на магнитной ленте, по которой впоследствии проводится рисование в автоматическом режиме на SC 4020. Однако в некоторых системах записывающее устройство можно более непосредственно подключить к основной системе компьютеров и рисование может проводиться по заданиям, хранящимся во временных файлах на дисках.

Команды на рисование читаются с ленты магнитофоном (1) и при помощи электроники (2) они управляют электронным пучком высокопрецизионной катодно-лучевой трубки (3). Изображение, получаемое на экране, раздваивается полупрозрачным зеркалом и направляется на кинокамеру (5) и камеру для снятия копии. Первая воспроизводит изображение для 16- или 32-мм кинофильма, а вторая воспроизводит изображение на

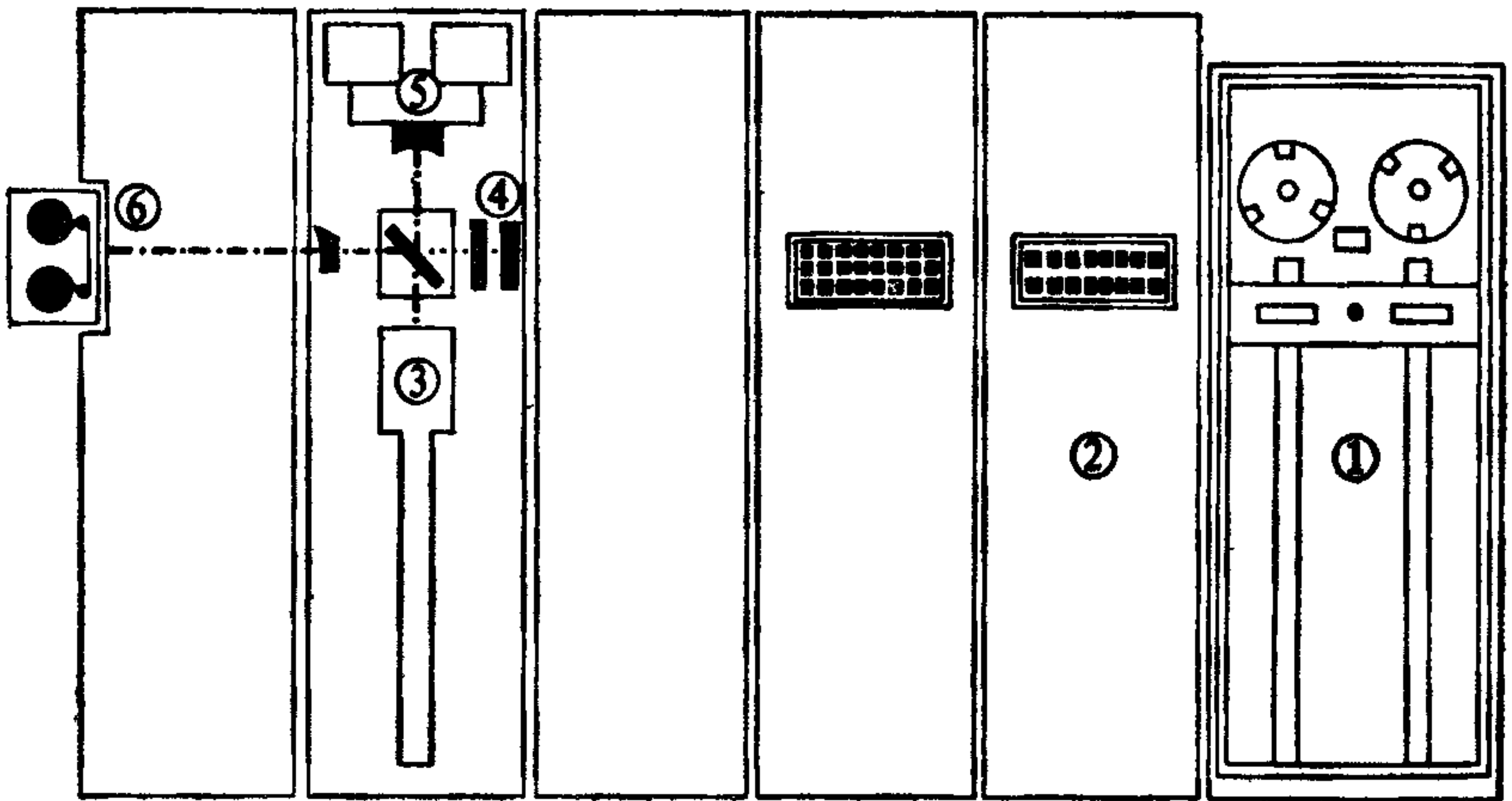


Рис 9.15. Аппарат для записи микрофильмов Стромберг — Карлсон 4020: магнитофон (1), электроника (2), высокопрецизионная электронно-лучевая трубка (3), устройство для изготовления слайдов (4), кинокамера (5), камера для твердой копии на бумаге (6). (С любезного разрешения Стромберг — Карлсона, Division of General Dynamics.)

10-дюймовую фоточувствительную бумагу. Если требуется фиксированное изображение фона, может быть введен формирующий слайд (4), содержащий это изображение. Изображение в обеих камерах накапливается (или рисуется) фоточувствительной средой до получения команды на продвижение фильма. В ЭЛТ встроена маска для придания электронному пучку формы одного из 64 возможных аппаратно реализуемых символов. Могут быть построены и другие наборы символов и форм, которые строятся из прямолинейных сегментов и самостоятельно программируются пользователем.

Слева вверху на рис. 9.10 каждая частица в плазменном моделировании записана с помощью аппаратных символов плюс и минус, соответственно для ионов и электронов. Надписи WALL и PLASMA вне расчетной области являются примерами изготовления надписей программными средствами с помощью прямолинейных отрезков. Вверху справа средняя скорость каждого типа частиц рассчитана на массиве квадратов (8×8). Затем нарисована стрелка, длина и направление которой пропорциональны среднему значению вектора скорости. Эта визуализация непосредственно показывает основные характеристики течения, которое представляет собой сильный поток электронов к стенке вблизи центра расчетной области. Внизу имеются часовой циферблат, указывающий течение времени, и изометрическая проекция поверхности электростатического потенциала. Визуализация дополнена записью маленькими аппаратными символами основных численных постоянных моделирования (времен-

ной шаг, пространственная ячейка, поле V , заряд, масса частиц и амплитуды волн).

Если запись, как, например, рис. 9.10, отображается в виде 16-мм мультфильма с одним кадром, соответствующим каждому временному шагу, и проецируется со скоростью 16 кадров в секунду, то обычно получается удовлетворительная динамическая картина временной эволюции эксперимента. Такая визуализация весьма полезна при идентификации качественных особенностей любой неустойчивости, которая недостаточно четко выявляется анализом статических картин, выбранных, например, через каждые 100 временных шагов. Изготовление мультфильма с помощью универсальных графических программных средств, вероятно, должно быть дорого при использовании основного компьютера и аппарата для микрофильмирования. По этой причине в редких случаях возможно изготовить фильмы для всех рассчитанных моментов, и стандартный вывод должен ограничиться статической визуализацией. Однако благоразумно изготавливать мультфильм всякий раз, когда происходит большое изменение в масштабе или геометрии. Неудача такой попытки может означать, что важное физическое явление или ошибки программирования проходят невыявленными. Если, с другой стороны, написана простая подпрограмма в кодах ассемблера с оптимизацией (например, только для рисования точек), то, как обнаружили некоторые авторы, изготовление кадра может составить малую долю временного цикла и оно становится обычной формой вывода [251, 42].

Целью статической визуализации и показа мультфильмов является выдвижение на первый план качественного явления, возникающего в моделях, как, например, волна с $m = 1$ на рис. 9.10. Мы также узнаем, является волна стоячей или бегущей. В редких случаях из таких визуализаций можно выполнить количественные измерения, и, как только природа явления становится понятной, важно написать специальную программу диагностики для численного исследования ключевых особенностей данного явления. В нашем случае ясно, что на поверхности потенциала имеется почти синусоидальная волна с $m = 1$. Волна определена количественно с помощью вычисления ее амплитуды и фазы, которые выведены на устройстве визуализации (см., например, рис. 9.11, выполненный на перьевом самописце «Калькомп»). Данные о волне воспроизведены как для стенки, так и для внутренней области плазмы. Также отражены прибывающие на стенку электроны и ионы. Для количественного исследования наиболее удобны двумерные графики на бумаге, с которых могут быть выполнены измерения. В этом примере время прохождения t_3 может быть измерено из графика фазы, а период t_4 — из графика амплитуды.

Stromberg — Carlson 4020, вероятно, был наиболее широко используемым и успешным устройством для микрофильмирования и доступным для ученых-вычислителей в период между началом 1960 г. и серединой 1970 г. Очевидно, такое устройство является существенной частью любой компьютерной установки, привлеченной в крупномасштабную вычислительную науку. Более совершенные установки (приблизительно 1975 г.), такие, как IIFR80, работают на том же принципе, но дают большую точность (растр с адресацией $[16\,000 \times 16\,000]$) и другие возможности (например, создание цветных фильмов). Введение цвета имеет очевидные реальные приложения для отображения результатов моделирования. Они включают выделение типов частиц и линий уровня с помощью цвета.

Иногда представляется возможным использовать устройство для микрофильмирования для записи изображения прямо на звуковую дорожку 16- или 35-мм мультфильма и, следовательно, генерировать звуки различной высоты. Это могло найти применение для фиксации попадания моделируемых частиц на стенку, причем различные высоты звука использовались для различных типов частиц. Однако для такого применения не созданы устройства для микрофильмирования и возможности искусственно сгенерированного звука ни в какой мере пока что не были использованы.

9.4. МАГНИТОСФЕРА

Видимое Солнце окружено протонной плазмой, называемой солнечной короной и имеющей температуру около 10^6 К. Поток квазинейтральной плазмы, называемой солнечным ветром, выбрасывается из короны и непрерывно бомбардирует Землю и другие планеты. Типичная скорость потока плазмы около Земли составляет примерно $7.5 \cdot 10^5$ м/с. Приближение в замороженности применимо очень хорошо, и вместе с солнечным ветром от Солнца переносится и замороженное в него магнитное поле. Межпланетное магнитное поле разделяется на два или четыре сектора, в которых поле имеет либо северное, либо южное направление. Проблема магнитосферы может быть решена с помощью полных трехмерных расчетов взаимодействия солнечного ветра с магнитным полем Земли. Важно также иметь нестационарное моделирование, потому что полярные сияния и связанные с ними магнитные бури, наблюдавшиеся на Земле, связаны с изменениями в распределении плазмы и магнитного поля в магнитосфере, которые в свою очередь инициируются изменениями в солнечном ветре и его магнитном поле.

На рис. 9.16 представлена картина магнитного поля и течения плазмы, которая была рассмотрена Данжи [112] для случая, когда солнечный ветер переносит магнитное поле южного

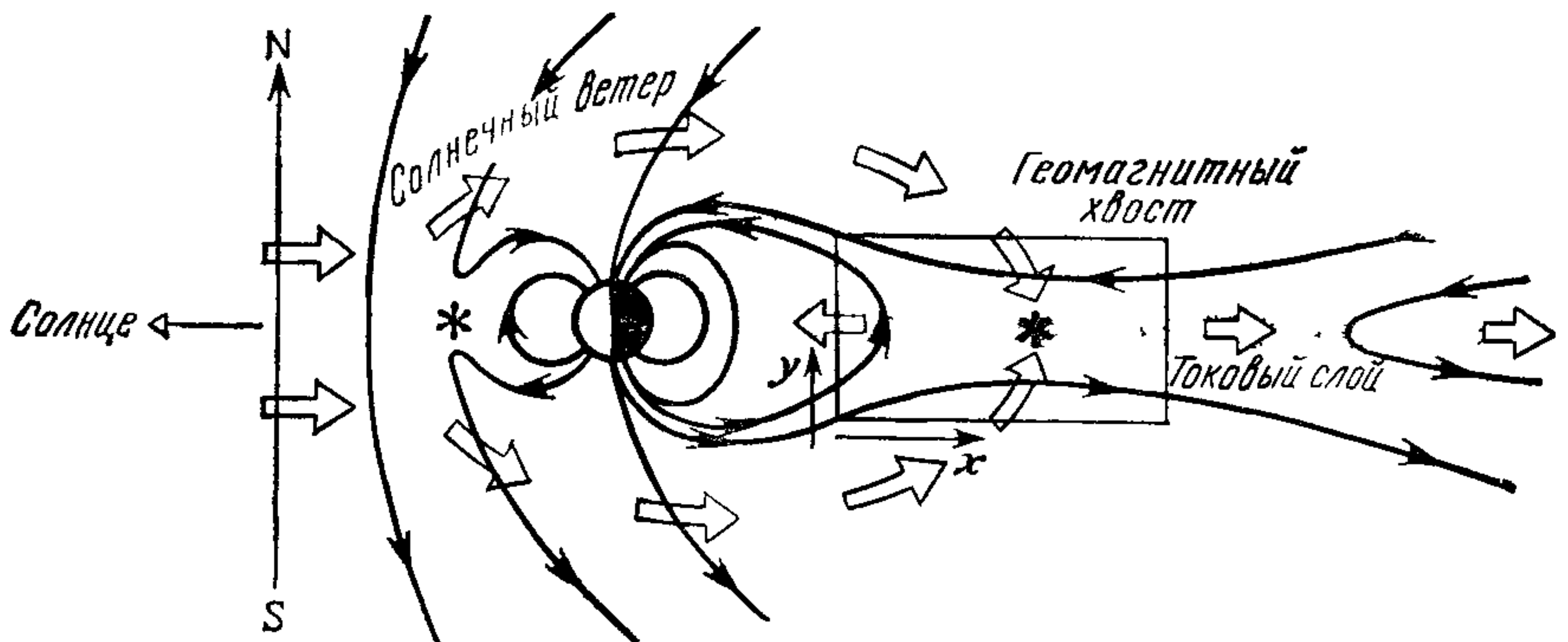


Рис 9.16 Модель магнитосферы Данжи с магнитным полем в солнечном ветре, направленном к югу. Плазма солнечного ветра (большие стрелки) течет через полюсы к нейтральной точке, находящейся вниз по потоку солнечного ветра от Земли, откуда она либо возвращается к Земле на низких широтах, либо уходит на бесконечность. Позади Земли плазма образует относительно тонкий токовый слой. Выделенная на рисунке область в виде прямоугольника, исследуется на рис 9.19. ([112], с любезного разрешения Physical Review Letters, © American Institute of Physics.)

направления. Линия Земля — Солнце лежит в плоскости страницы с изображением Солнца слева. Рисунок располагается в плоскости, перпендикулярной орбите Земли, а геомагнитный диполь изображен для случая, в котором он перпендикулярен линии Земля — Солнце, хотя обычно ось диполя может иметь наклон приблизительно до 45° от этого направления, что из-за устранения всех симметрий еще больше усложняет проблему.

В случае южного (т. е. направленного на юг) поля солнечного ветра в магнитосфере имеются две нейтральные точки, в которых магнитные поля Солнца и Земли взаимно уничтожаются. Они изображены в виде двух звездочек перед Землей и за ней вниз по потоку солнечного ветра. В этих точках силовые линии магнитного поля, переносимые движущейся плазмой, соприкасаются в виде символа X, а затем расходятся и пересоединяются в противоположном направлении. Это важное явление пересоединения изучается с помощью моделирования в разд. 9.4.4. Физические аспекты проблемы обсуждались Р. З. Сагдеевым [339].

Рисунок 9.16, таким образом, должен рассматриваться динамически с движением плазмы солнечного ветра (большие стрелки) и вмороженными в нее силовыми линиями магнитного поля, которые непрерывно сталкиваются с полем земного диполя. Падающие на магнитосферу силовые линии межпланетного магнитного поля встречаются и пересоединяются в нейтральной точке на подсолнечной стороне Земли с силовыми линиями поля, движущимися от Земли. Одна пара линий показана сразу после пересоединения, правее нейтральной точки. Затем

пересоединенные линии вместе с плазмой солнечного ветра текут через полюсы к нейтральной точке на ночной стороне. На рисунке показаны две линии непосредственно перед пересоединением, когда они движутся навстречу друг к другу и к плоскости земной орбиты. После пересоединения линия, показанная слева, представляет теперь уже часть геомагнитного поля и вместе с некоторым захваченным на нее количеством плазмы солнечного ветра движется по направлению к Земле. Остальные пересоединившиеся силовые линии вместе с потоком солнечного ветра движутся в направлении от Земли.

Приближение потоков плазмы к плоскости земной орбиты, возникающее в результате рассмотренного выше процесса пересоединения, приводит к появлению относительно тонкого и плоского слоя плазмы позади Земли, который известен как плазменный слой геомагнитного хвоста. Он имеет ширину приблизительно 40 земных радиусов (т. е. перпендикулярно к странице в плоскости земной орбиты) и толщину приблизительно 12 земных радиусов (т. е. в направлении север — юг в плоскости страницы). Внутри плазменного слоя магнитосферного хвоста должны течь сильные токи в плоскости земной орбиты (т. е. поперек страницы), которые создают магнитное поле, меняющее свое направление (к Солнцу вверх и против Солнца вниз). Благодаря тому, что параллельные элементы тока притягиваются друг к другу, эти токи взаимодействуют и образуют очень тонкий токовый слой толщины около 0.1 земного радиуса. Образование этого токового слоя в прямоугольной области, показанное на рис. 9.16, будет моделироваться методом частиц в разд. 9.4.3 и 9.4.4. Формирование картины общего течения плазмы и поля, включая моделирование магнитных суббурь, является предметом следующих двух разделов (разд. 9.4.1 и 9.4.2).

9.4.1. Магнитогидродинамическая модель частиц

Уравнения идеальной магнитной гидродинамики с идеальной проводимостью получены в разд. 9.1 (уравнения (9.4) — (9.8)). Их можно записать в следующем виде:

$$\frac{\partial \rho_m}{\partial t} = - \operatorname{div}(\rho_m \mathbf{v}), \quad (9.63a)$$

$$\rho_m \frac{d\mathbf{v}}{dt} = - \nabla p - \frac{1}{2\mu_0} \nabla B^2 + \frac{1}{\mu_0} \nabla \cdot \mathbf{B}\mathbf{B}, \quad (9.63b)$$

$$\frac{\partial}{\partial t} \mathbf{B} = \operatorname{rot}(\mathbf{v} \times \mathbf{B}), \quad (9.63c)$$

$$\rho \sim \rho_m', \quad (9.63d)$$

где член $\mathbf{j} \times \mathbf{B}$ в правой части уравнения (9.8b) выражен через тензор магнитного давления.

В традиционных методах эти уравнения целиком заменяются конечно-разностными на фиксированной эйлеровой сетке и решаются для определения ρ_m , \mathbf{v} и \mathbf{B} на следующем временном слое. Эти методы сталкиваются с трудностями при приложении их к расчетам магнитосферы, поскольку встречающиеся в вычислениях низкие плотности ведут к неустойчивостям, а высокие скорости жидкости могут давать нефизические отрицательные плотности при решении уравнения неразрывности (9.63а). Один способ обхода этих трудностей предложен в работе [245], где МГД-плазма рассмотрена как собрание частиц. При таком рассмотрении частицы не могут потеряться, так что уравнение неразрывности автоматически выполняется и отрицательные плотности возникнуть не могут. Уравнение (9.63b) становится уравнением движения частиц плазмы. Следовательно, для частицы i имеем

$$\frac{d\mathbf{v}_i}{dt} = m^{-1} \left(-\nabla p - \frac{1}{2\mu_0} \nabla B^2 + \frac{1}{\mu_0} \nabla \cdot \mathbf{B}\mathbf{B} \right), \quad (9.64a)$$

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i, \quad (9.64b)$$

где m — масса частицы. Вывод полевых величин, таких, как плотность частиц n , массовая плотность ρ_m и скорость плазмы \mathbf{v} , из усреднений параметров частиц, проводится с помощью их распределения на фиксированную эйлерову сетку с использованием весовой функции w таким же способом, как осуществлялось распределение заряда в электростатическом коде. Таким образом, имеем

$$n(\mathbf{r}) = \sum_i w(\mathbf{r} - \mathbf{r}_i), \quad (9.65a)$$

$$\rho_m(\mathbf{r}) = mn(\mathbf{r}),$$

$$\mathbf{v}(\mathbf{r}) = \sum_i \mathbf{v}_i w(\mathbf{r} - \mathbf{r}_i) / n(\mathbf{r}). \quad (9.65b)$$

Весовая функция, используемая в работе [245], соответствует рассмотрению каждой частицы как массового облака с плотностью, имеющей форму гауссовой кривой и распределенной симметрично относительно координаты частицы. Давление в каждой точке сетки пересчитывается из ρ_m при помощи (9.63d), а сила давления $-\nabla p$ получается локальным дифференцированием или при помощи преобразования Фурье [228]. Члены с тензором магнитного давления в правой части (9.64а) вычисляются обычными конечно-разностными методами.

Магнитное поле пересчитывается во времени по (9.63с) при помощи консервативного конечно-разностного метода Лакса [314]. Применение этого метода вводит численную диффузию,

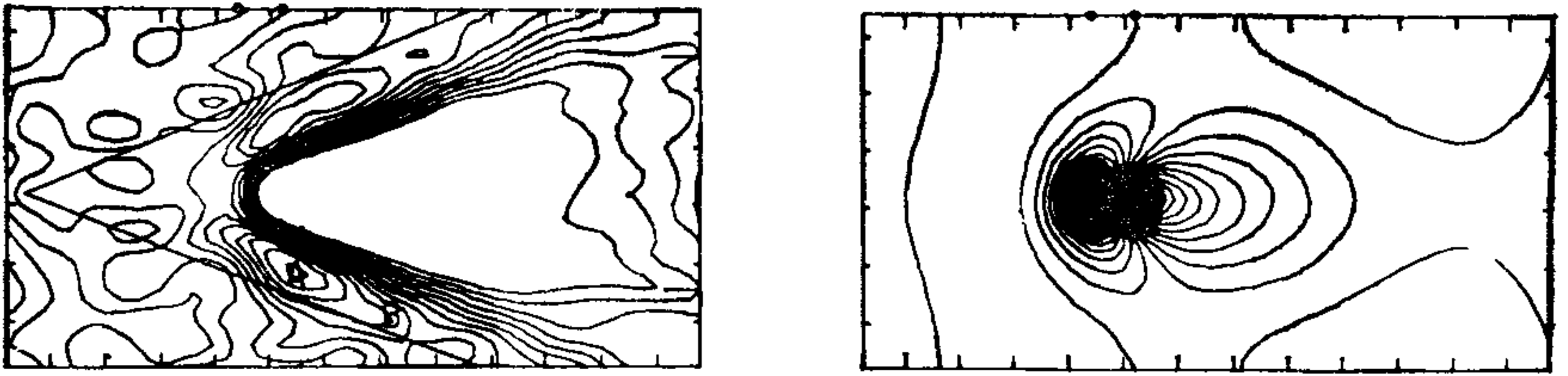


Рис 9.17. Моделирование взаимодействия солнечного ветра без магнитного поля с магнитным диполем, имитирующим магнитное поле Земли. Слева даны контуры постоянной плотности, показывающие конус Маха и магнитопаузу; справа — магнитные силовые линии, демонстрирующие магнитосферу, вытянутую в направлении солнечного ветра. ([247], с любезного разрешения Geophysical Monograph Series, © American Geophysical Union.)

поскольку фактически мы решаем уравнение

$$\frac{\partial \mathbf{B}}{\partial t} - \text{rot}(\mathbf{v} \times \mathbf{B}) = (H^2/4DT - DTc^2/2) \nabla^2 \mathbf{B}, \quad (9.66)$$

где c — наибольшая скорость распространения на сетке. Члены в правой части возникают из-за конечного размера интервала пространственной сетки H и временного шага DT . Эти члены, отвечающие погрешностям аппроксимации, не могут всюду взаимно уничтожаться из-за изменения скорости распространения в зависимости от положения на сетке. Однако необходимо, чтобы их сумма была положительной для того, чтобы избежать введения отрицательной диффузии, вызывающей численную неустойчивость. Это приводит к условию Куранта — Фридрихса — Леви для максимального размера временного шага

$$DT < 2^{-1/2} H/c. \quad (9.67)$$

Наличие положительной численной диффузии означает, что, хотя мы начали решение, используя уравнения с идеальной проводимостью (9.63с), моделирование в результате оказалось проведенным с конечной проводимостью. Силовые линии магнитного поля в итоге начинают диффундировать через плазму, и свойства замороженности выполняются лишь приближенно.

9.4.2. Глобальная модель магнитосферы

При изучении магнитосферы использован вариант рассмотренного выше кода с размерностью $2^{1/2}$ с двумя пространственными измерениями и тремя измерениями для скорости и поля [246, 247]. В эксперименте, показанном на рис. 9.17, солнечный ветер, лишенный собственного магнитного поля, входит слева в область моделирования и налетает на магнитный диполь, который соответствует Земле. Ветер имеет сверхзвуковую скорость, конус Маха для него и контуры плотности показаны

слева. Контуры, соответствующие плотностям, меньшим, чем плотность набегающего солнечного ветра, убраны с картины, чтобы сделать ее более наглядной. Справа показаны силовые линии магнитного поля, образующие по существу закрытую каплевидную магнитосферу, подобную предсказанной Джонсоном [223]. Форма наиболее удаленных силовых линий поля подвергается сильному влиянию граничных условий с двойной периодичностью, наложенных на прямоугольник.

Для моделирования условий, рассмотренных Данжи [112] и показанных на рис. 9.16, было выполнено моделирование, в котором магнитное поле солнечного ветра переключено с ориентации восток — запад на южное направление, как это могло бы произойти во время магнитной бури, связанной с вхождением Земли в сегмент солнечного ветра с обращенным к югу магнитным полем. На рис. 9.18 слева поле с южным направлением еще не достигло Земли и на подсолнечной стороне Земли образовалась нейтральная точка с силовыми линиями, разрывающимися и пересоединяющимися в этой точке. Справа показано распределение, когда развернутое поле прошло полностью через систему и в ней возникает профиль поля, подобный предсказанному Данжи, включая геомагнитный хвост, хотя он оказался намного короче, чем это должно было бы быть, за счет конечной проводимости модели. Тем не менее моделирование показало все качественные особенности модели Данжи. Для того чтобы моделирование можно было бы рассматривать как удовлетворительное, требуется рассмотреть полную трехмерную модель и значительно понизить численную диффузию. Такое моделирование было бы возможно на компьютерах, подобных CRAY-1, со скоростью, в несколько десятков раз превосходящей скорость IBM 360/195.

Очевидно, что рассмотренные методики для моделирования МГД-жидкости могут применяться для моделирования и других жидкостей методами частиц. В гидродинамике члены с тензо-

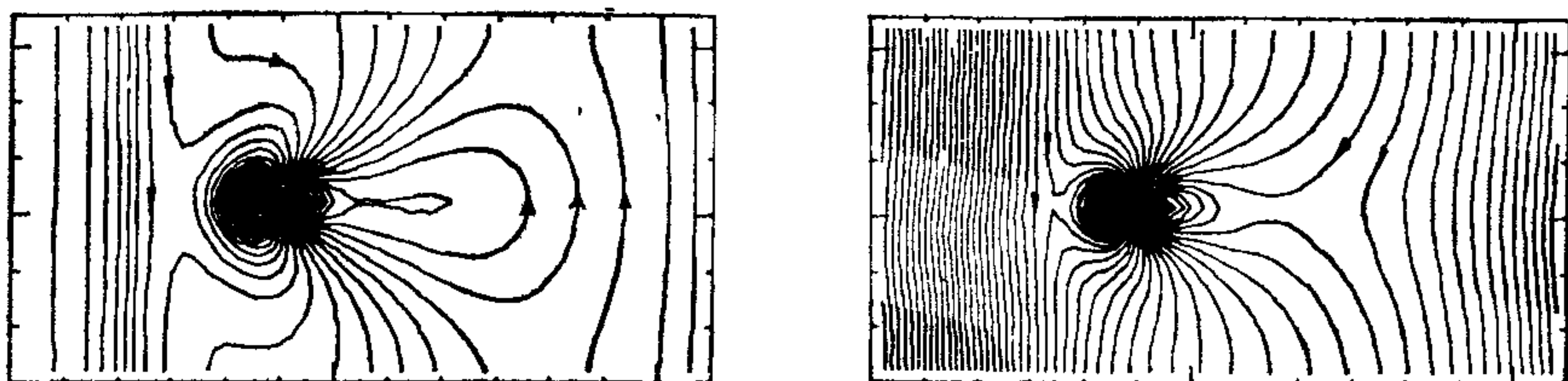


Рис 9.18. Моделирование магнитной суббури. Магнитное поле, переносимое солнечным ветром, меняет направление с восточно-западного на южное. Слева: область с направленным на юг полем еще не достигла Земли; справа: она полностью проходит через систему, приводя к распределению поля, очень напоминающему картину Данжи [112] на рис. 9.16. ([247], с любезного разрешения Geophysical Monograph Series, © American Geophysical Union.)

ром магнитного давления в (9.63b) заменены вязкими членами, связанными с производными от поля скоростей, а уравнение для магнитного поля (9.63c) отсутствует. Однако рассмотрение движения жидкости как движения отдельных частиц для того, чтобы обеспечить положительную плотность и сохранение массы, и вычисление силы, действующей на жидкие частицы, из полевых величин, полученных на фиксированной сетке, является надежным вычислительным методом и должно найти широкое приложение в динамике жидкости, хотя до сих пор они еще не привлекли того внимания исследователей, которое несомненно заслуживают.

9.4.3. Модель частиц Ампера

Для детального изучения пересоединения и образования токовых слоев Иствуд и Гамильтон развили другой вариант модели частиц. Эта модель [159] уже является полностью двумерной и может быть применена для изучения геомагнитного хвоста в прямоугольной области, показанной на рис. 9.16. Как мы уже отмечали, размер хвоста в направлении, перпендикулярном плоскости рисунка (z), во много раз больше его толщины в плоскости рисунка (x, y).

В силу этого можно предположить, что магнитное поле ограничено плоскостью (x, y), следовательно,

$$\mathbf{B} = (B_x, B_y, 0), \quad (9.68)$$

а токи, электрическое поле и векторный потенциал направлены вдоль оси z :

$$\mathbf{j} = (0, 0, j_z), \quad (9.69a)$$

$$\mathbf{E} = (0, 0, E_z), \quad (9.69b)$$

$$\mathbf{A} = (0, 0, A_z). \quad (9.69c)$$

Магнитное поле связано с векторным потенциалом \mathbf{A} соотношением

$$\mathbf{B} = \text{rot } \mathbf{A}, \quad (9.70a)$$

следовательно,

$$B_x = \frac{\partial}{\partial y} A_z, \quad B_y = -\frac{\partial}{\partial x} A_z. \quad (9.70b)$$

Уравнение Ампера (9.4a)

$$\text{rot } \mathbf{B} = \mu_0 \mathbf{j} \quad (9.71a)$$

дает

$$j_x = j_y = 0, \quad \frac{\partial}{\partial x} B_y - \frac{\partial}{\partial y} B_x = \mu_0 j_z, \quad (9.71b)$$

или, подставляя (9.70b), получаем

$$\nabla^2 A_z = -\mu_0 j_z. \quad (9.71c)$$

Электрическое поле можно найти из уравнения (9.4b):

$$\operatorname{rot} \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}; \quad (9.72a)$$

последнее при подстановке (9.70a) и (9.69b, c) можно выразить через векторный потенциал:

$$E_z = - \frac{\partial A_z}{\partial t}. \quad (9.72b)$$

Для того чтобы получить выражение для тока j_z , необходимо рассмотреть орбиты ионов, которые из-за их большого ларморовского радиуса должны, как легко показать, переносить большую часть тока [113, 114]. Уравнение движения для иона с массой m и зарядом q в электрическом и магнитном полях имеет вид

$$\frac{d\mathbf{v}}{dt} = \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (9.73)$$

В координатной форме с использованием допущений (9.68) и (9.69) это уравнение принимает вид

$$m \frac{d}{dt} v_x = -q B_y v_z, \quad (9.74a)$$

$$m \frac{d}{dt} v_y = q B_x v_z, \quad (9.74b)$$

$$m \frac{d}{dt} v_z = q (E_z + v_x B_y - v_y B_x). \quad (9.74c)$$

Последнее уравнение (9.74c) можно проинтегрировать с помощью подстановки (9.70b) и (9.72b) и в результате найти

$$m \frac{dv_z}{dt} = -q \left(\frac{\partial A_z}{\partial t} + v_x \frac{\partial A_z}{\partial x} + v_y \frac{\partial A_z}{\partial y} \right). \quad (9.75a)$$

Выражение в скобках в правой части равно производной A_z по времени dA_z/dt , которую «видит» движущаяся частица. Следовательно,

$$\frac{d}{dt} (mv_z + qA_z) = 0, \quad (9.75b)$$

или, интегрируя, получаем

$$p_z = mv_z + qA_z = \text{const}. \quad (9.75c)$$

Поэтому величина p_z , называемая обобщенным импульсом, является интегралом движения. Значение p_z определяется начальными условиями и обычно различно для каждого иона. Скорость иона находится теперь из значения для векторного потенциала

$$v_z = m^{-1} (p_z - qA_z), \quad (9.76a)$$

а ток задается выражением

$$j_z = nq\bar{v}_z = nqm^{-1}(\bar{p}_z - qA_z), \quad (9.76b)$$

где n — плотность частиц, а \bar{v}_z и \bar{p}_z — локальные средние скорость и импульс частиц.

Подставляя (9.76b) в уравнение Ампера (9.71c), получаем уравнение поля для рассматриваемой задачи

$$\left(-\nabla^2 + \frac{\mu_0 q^2 n}{m}\right) A_z = \frac{\mu_0 q}{m} (n\bar{p}_z). \quad (9.77)$$

Уравнение (9.77) дискретизируется, а величины A_z рассчитываются на регулярном наборе узлов сетки s с координатами \mathbf{r}_s . Затем значения плотности и импульса в этих узлах сетки получаются с помощью распределения параметров частицы по узлам сетки с использованием весовой функции ω способом, аналогичным распределению заряда в электростатической плазме. Таким образом,

$$n_s = n_z H^{-2} \sum_i \omega(\mathbf{r}_s - \mathbf{r}_i), \quad (9.78a)$$

$$(n\bar{p}_z)_s = n_z H^{-2} \sum_i \omega(\mathbf{r}_s - \mathbf{r}_i) p_{zi}, \quad (9.78b)$$

где суммирование проводится по всем частицам i с радиус-вектором \mathbf{r}_i . Сетка предполагается квадратной с шагом сетки H , а n_z — число протонов на метр в направлении z , которое представлено одной частицей в моделировании. Если область моделирования равна $L_x \times L_y$ и имеется N частиц, представляющих среднюю трехмерную плотность протонов n_0 , то

$$n_z = L_x L_y n_0 / N. \quad (9.78c)$$

Поскольку плотность n является произвольной функцией положения и распределения частиц, определяемой выражением (9.78a), уравнение поля (9.77) представляет несепарабельную задачу типа Пуассона. Так как нас удовлетворяет решение внутри прямоугольника и требуется большое число узлов сетки, рекомендуется решение с помощью итераций Конкуса и Голуба (см гл. 6, рис 6.1).

В методе Конкуса и Голуба вводится постоянное смещение k в обе части (9.77) и определяется схема

$$(-\nabla^2 + k) A_z^{(i+1)} = (k - \mu_0 q^2 m^{-1} n) A_z^{(i)} + \mu_0 q m^{-1} (n\bar{p}_z), \quad (9.79)$$

где верхний индекс i означает номер итерации. Очевидно, что итерация сходится за один шаг, если плотность постоянна и $k = \mu_0 q^2 m^{-1} n$. Если плотность меняется, как в обычном случае, то лучший выбор представляется значением

$$k = \frac{1}{2} \mu_0 q^2 m^{-1} (n_{\max} - n_{\min}), \quad (9.80)$$

где n_{\max} и n_{\min} — максимальное и минимальное значения плотности n , соответствующей узлам сетки согласно (9.78а). Сходимость является наиболее быстрой для почти однородной плазмы с низким уровнем флуктуаций плотности на сетке. При заданной физической вариации флуктуации меньше и сходимость лучше, чем больше число частиц, чем крупнее сетка и выше порядок схемы распределения заряда (т. е. СІС лучше, чем NGR).

Уравнение (9.79) является сепарабельным, и $A_z^{(t+1)}$ можно найти при помощи любого метода решения эллиптических уравнений (см. разд. 6.5.1—6.5.4). В программе COIRCE [160] используются методы свертки (разд. 6.5.4), основанные на двойном быстром преобразовании Фурье. Если используется схема приписывания СІС на сетке 32×32 , то одна итерация уравнения (9.79) занимает 0.046 с на IBM 360/195. Решение уравнения поля (9.77) для 8192 частиц, представляющих плазму, находится с точностью 1/1000 за 10 итераций, или 0.5 с.

Уравнения движения (9.74) интегрируются с помощью разностной схемы с перешагиванием (разд. 2.2.1).

$$v_{xi}^{(t+1/2)} = v_{xi}^{(t-1/2)} - B_{yi}^{(t)} v_{zi}^{(t)} \frac{q}{m} DT, \quad (9.81a)$$

$$x_i^{(t+1)} = x_i^{(t)} + v_{xi}^{(t+1/2)} DT \quad (9.81b)$$

и аналогично для v_y и y из уравнения (9.74b), кроме того

$$p_{zi}^{(t+1)} = p_{zi}^{(t)} \quad (9.81c)$$

из сохранения канонического импульса. Значения B_y и v_z в месте положения частицы получаются интерполяцией сеточных значений с той же весовой функцией, которая использована в распределении заряда. Таким образом, поле B_{yi} для i -й частицы с радиусом-вектором \mathbf{r}_i равно

$$B_{yi}^{(t)} = \sum_s \omega(\mathbf{r}_i - \mathbf{r}_s) B_y^{(t)}(\mathbf{r}_s), \quad (9.82a)$$

где суммирование проводится по ячейкам сетки s и

$$B_y^{(t)}(\mathbf{r}_s) = -\frac{\partial}{\partial x} A_z \simeq -[A_z^{(t)}(\mathbf{r}_s + H) - A_z^{(t)}(\mathbf{r}_s - H)]/2H \quad (9.82b)$$

и аналогично для B_{xi} . Скорость вдоль z в уравнении (9.81a) получается из уравнения (9.76а)

$$v_{zi}^{(t)} = m^{-1} (p_{zi}^{(t)} - qA_{zi}^{(t)}), \quad (9.83a)$$

где

$$A_{zi}^{(t)} = \sum_s \omega(\mathbf{r}_i - \mathbf{r}_s) A_z^{(t)}(\mathbf{r}_s). \quad (9.83b)$$

В программе COIRCE используется схема взвешивания СІС для интерполяции между сеточными значениями в (9.82а) и (9.83b).

Цикл по временному шагу для модели Ампера может быть описан следующим образом.

Заполнить координаты каждой i -й частицы

$$(x, y, p_z)_i^{(t)} (v_x, v_y)_i^{(t-1/2)}. \quad (9.84)$$

1. Распределить плотности и компоненты импульса в направлении z на сетке по (9.78) согласно $(x, y, p_z)^{(t)}$.

2. Решить уравнения поля (9.77) для $A_z^{(t)}$.

3. Проинтегрировать уравнения движения для нахождения $(v_x, v_y)^{t+1/2}$ и $(x, y, p_z)^{(t+1)}$ согласно (9.81).

В описанной выше модели в явной форме рассчитывается трехмерное движение тяжелых протонов с учетом их инерции. При этом учитываются также эффекты конечности гирорадиуса и гиропериода, которые отсутствуют в МГД-модели, рассматривавшейся в разд. 9.4.1. Это позволяет использовать модель в окрестности области обращения поля в геомагнитном хвосте, а также рассматривать звуковые и альфвеновские волны. Роль электронов, которые в модели явно не представлены, состоит в поддержании нейтральности заряда и в устранении любых протонных токов в плоскости (x, y) . Это объясняет, почему мы считаем, что $j_x = j_y = 0$, хотя учитывается, что скорости заряженных частиц v_x и $v_y \neq 0$. Малый гирорадиус электронов не позволяет им компенсировать протонные токи в направлении z . Движение частиц представляет собой проекцию протонных орбит на плоскость (x, y) . Они могут также рассматриваться как «стержни» нейтральной плазмы, движущиеся поперек поля (B_x, B_y) и переносящие индуцированные токи вдоль оси z .

9.4.4. Геомагнитный хвост

Модель Ампера была использована для изучения пересоединения силовых линий магнитного поля в нейтральной точке в области, расположенной вниз по потоку солнечного ветра, и исследования устойчивости токового слоя в геомагнитном хвосте.

На рис. 9.19 показана модель пересоединения. Нейтральная точка в центре создается двумя стационарными «стержнями» на левом и правом краях расчетной области, несущими ток в одном и том же направлении. Плазма инжектируется из вершины и основания в первоначально пустую систему. Силовые линии поля приносятся в систему вместе с плазмой, как это можно видеть на рисунке для последовательных моментов времени. Немного позже в нижней левой части рисунка можно увидеть, что силовые линии магнитного поля должны разорваться и пересоединиться в противоположном направлении. Внизу справа можно видеть, что пересоединившиеся линии «вдвигаются» в фиксированные токовые стержни, и новая пара линий

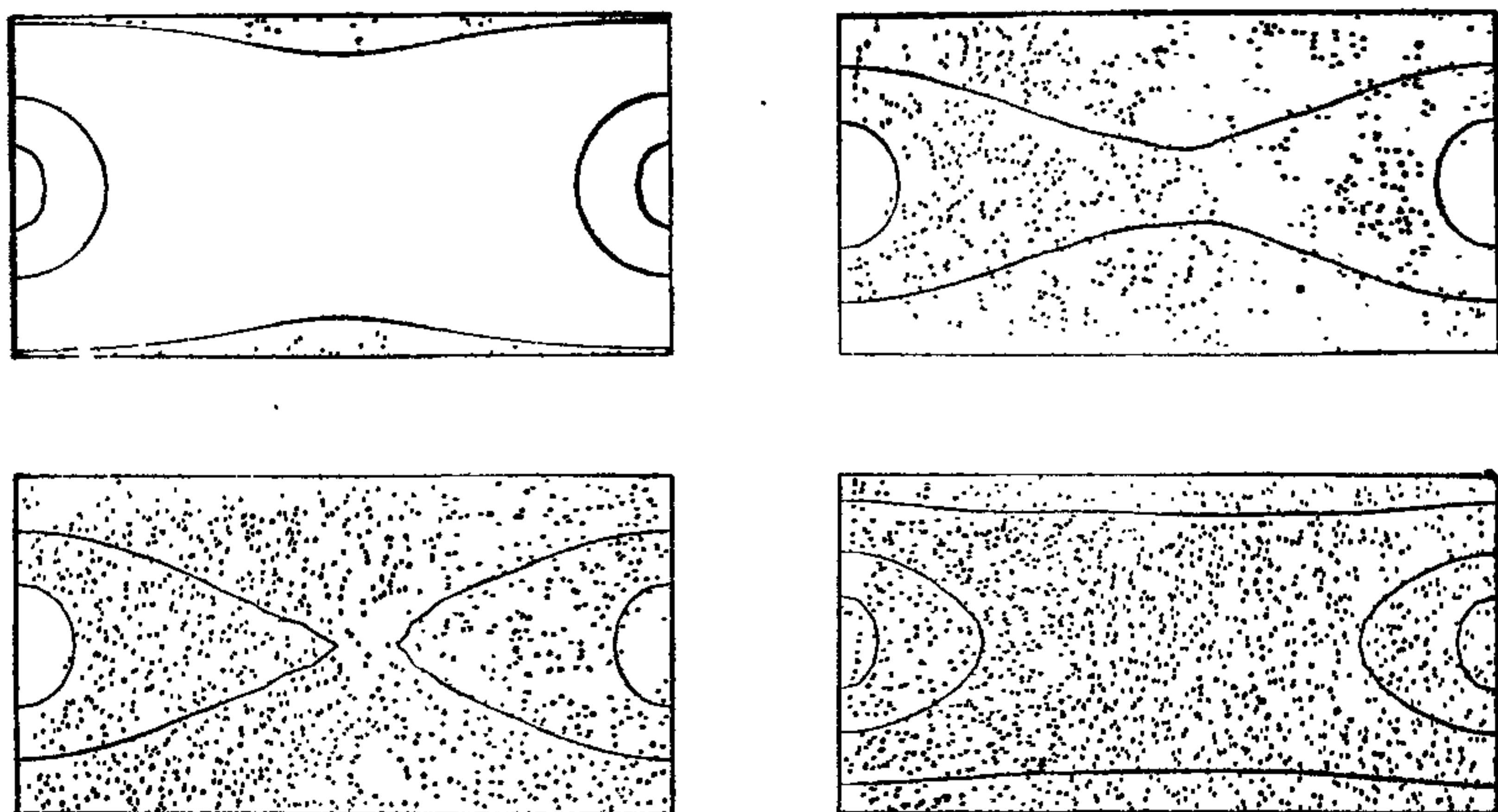


Рис. 9.19. Моделирование разрыва и пересоединения силовых линий магнитного поля в нейтральной точке (любезно передано Дж. Гамильтоном и Дж. Иствудом, не опубликовано).

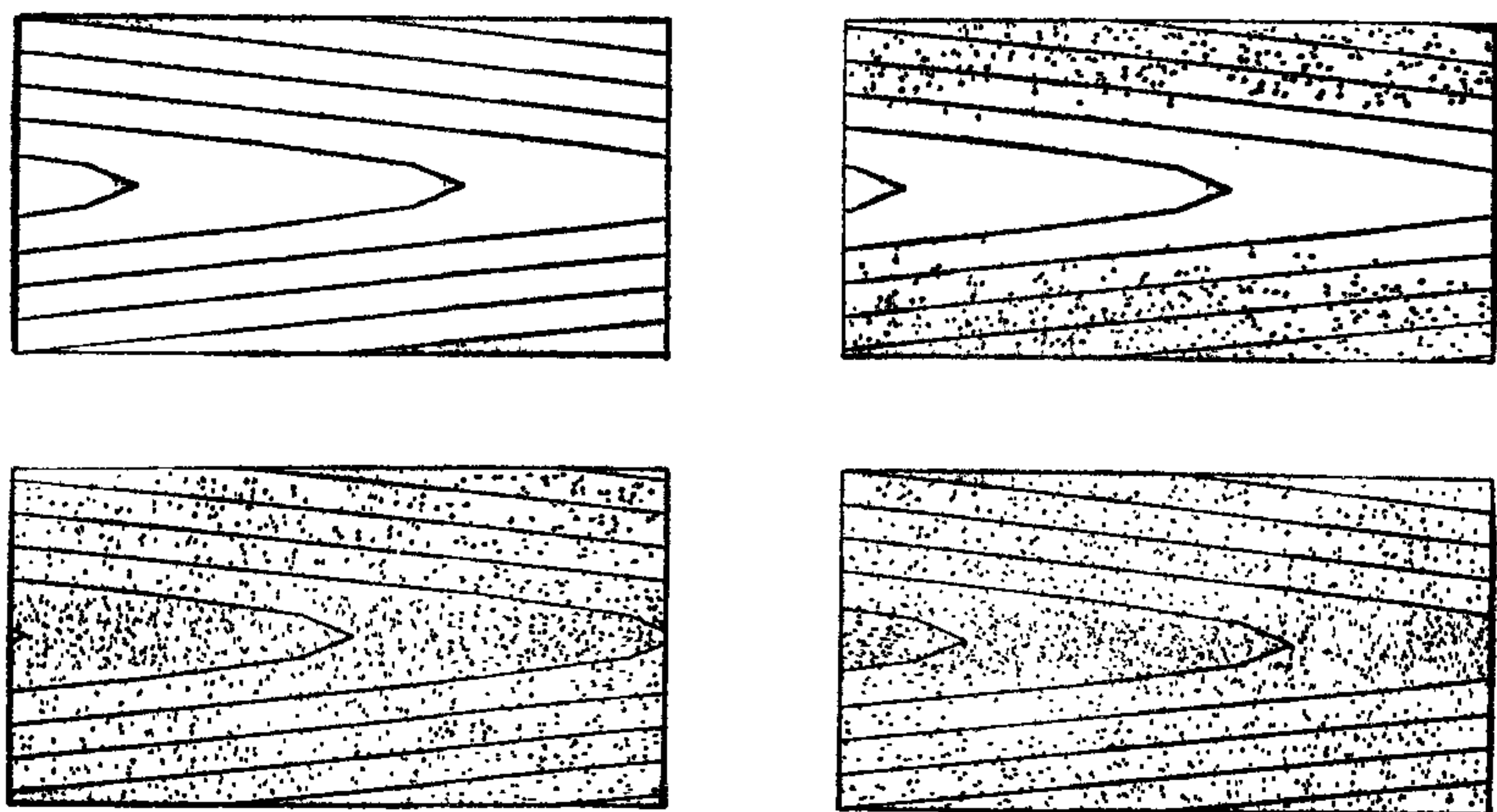


Рис. 9.20. Образование токового слоя в области слева от нейтральной точки (любезно передано Дж. Гамильтоном и Дж. Иствудом, не опубликовано).

входит в систему сверху и снизу. В течение этого процесса плазма течет из вершины и основания области моделирования к левой и правой сторонам.

На рис. 9.20 показано моделирование сегмента токового слоя, погруженного в плазменный слой геомагнитного хвоста. Имеются источники и стоки плазмы вдоль верхней и нижней границ и периодичность в направлении x . Конечность ширины токового слоя учитывается тем, что частицы, которые смещаются от плоскости листа больше, чем на некоторое заданное расстояние,

устраняются из системы. Земля располагается слева, а магнитное поле, обусловленное тонким токовым слоем, первоначально налагается на пустую область. Затем плазма инжектируется вдоль верхней и нижней границ до тех пор, пока она не заполнит всю область. Это показано на временной последовательности вверху (слева, справа), внизу (слева, справа). Когда постепенно устанавливается самосогласованное магнитное поле, наложенное поле понижается до тех пор, пока в конечном счете полное магнитное поле токового слоя не окажется созданным движущейся плазмой. Притяжение токовых стержней ведет к концентрации плазмы вдоль центральной оси в относительно тонкий токовый слой (внизу, справа). Для больших компонент поля B_y , соединяющихся через токовый слой, конфигурация остается устойчивой, как показано на рис. 9.20. Однако для малых значений B_y токовый слой становится неустойчивым и движется к новой устойчивости конфигурации с более сильным током и с точками O- и X-типа, в которых поле обращается в нуль. Этот динамический процесс может объяснить механизм накопления и высвобождения энергии магнитного поля в магнитосферных суббурях [161].

Моделирование полупроводниковых приборов

10.1. ВВЕДЕНИЕ

10.1.1. Цель моделирования

Конструирование полупроводниковых приборов является одним из наиболее перспективных приложений численного моделирования. Если надежная вычислительная модель полупроводникового прибора может быть создана, то ее можно использовать в качестве инструмента для конструирования. При определенных условиях характеристики нового прибора могут быть найдены посредством численного моделирования более экономично, чем традиционными лабораторными методами. Во всяком случае исчерпывающая программа численного моделирования является разумным сопровождением любой программы лабораторных экспериментов. В этой главе мы обсуждаем применение численного моделирования для конструирования одного класса полупроводниковых устройств, а именно полевых транзисторов (ПТ) на арсениде галлия (GaAs). Однако эти методы могут быть использованы в большинстве случаев для конструирования других приборов.

Преимущества численного моделирования могут быть изложены по пунктам следующим образом:

1. *Параметры.* Используемые параметры в любом численном моделировании точно известны, поскольку они определяются используемыми входными данными. В противоположность этому такие параметры, как уровень и профиль легирования, имеющиеся в любом лабораторном приборе, известны только приблизительно из-за трудности измерений.

2. *Диапазон параметров.* Диапазон параметров (например, расстояние между электродами), доступный для лабораторного измерения, ограничен существующей промышленной технологией. Однако в численном моделировании могут использоваться существенно более широкие диапазоны параметров. Исследование диапазонов параметров на компьютере, которые невозможно достичь в настоящее время в лаборатории, может указать области, в которых следует проводить инвестирование для развития лучших лабораторных методов. С другой стороны, вычислительные эксперименты могут указать области параметров, для которых *нецелесообразно* развитие новых лабораторных методов. В том и другом случае большие капитальные затраты по-

казывают, что промышленность электронных приборов игнорирует численное моделирование на свой страх и риск.

3. *Скорость.* Если лабораторное измерение на приборе нуждается в каком-либо техническом процессе, как, например, изготовление маски, то почти обязательно характеристики прибора могут быть получены существенно быстрее с помощью численного моделирования, чем в лаборатории.

Перечисляя преимущества численного моделирования, благоразумно помнить, что моделирование исключительно так же надежно, как и используемая вычислительная модель. В конце концов численное моделирование является лишь приближенной моделью реального мира и в конечном счете его результаты должны уступать результатам, полученным в лаборатории.

10.1.2. Постановка задачи

Первая стадия моделирования заключается в разработке упрощенной модели реального прибора. На рис. 10.1 показаны два ПТ на GaAs, изготовленные фирмой Plessey Ltd. Они используются как высокочастотные усилители в аэрокосмических приложениях и могут применяться в диапазоне частот от

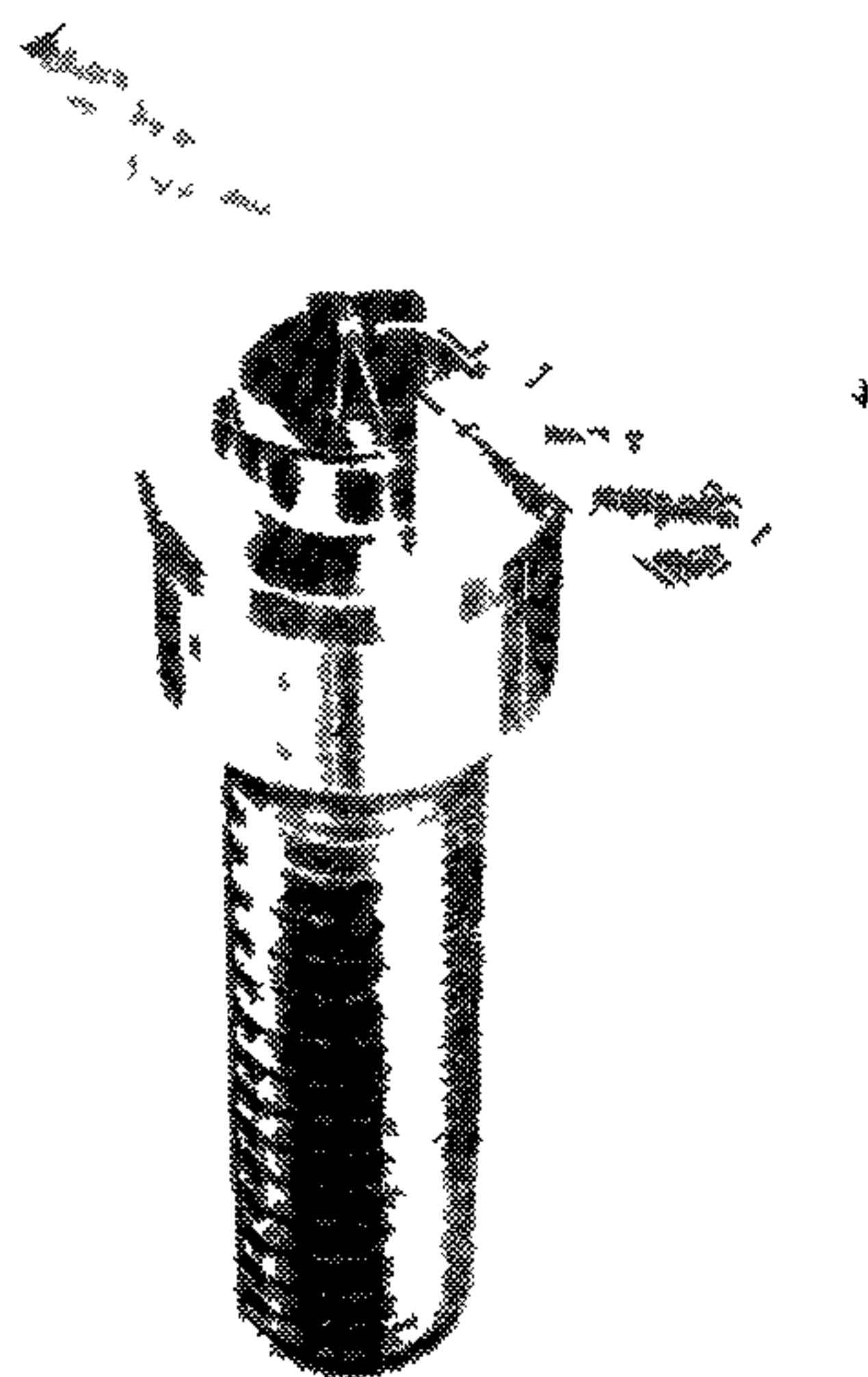
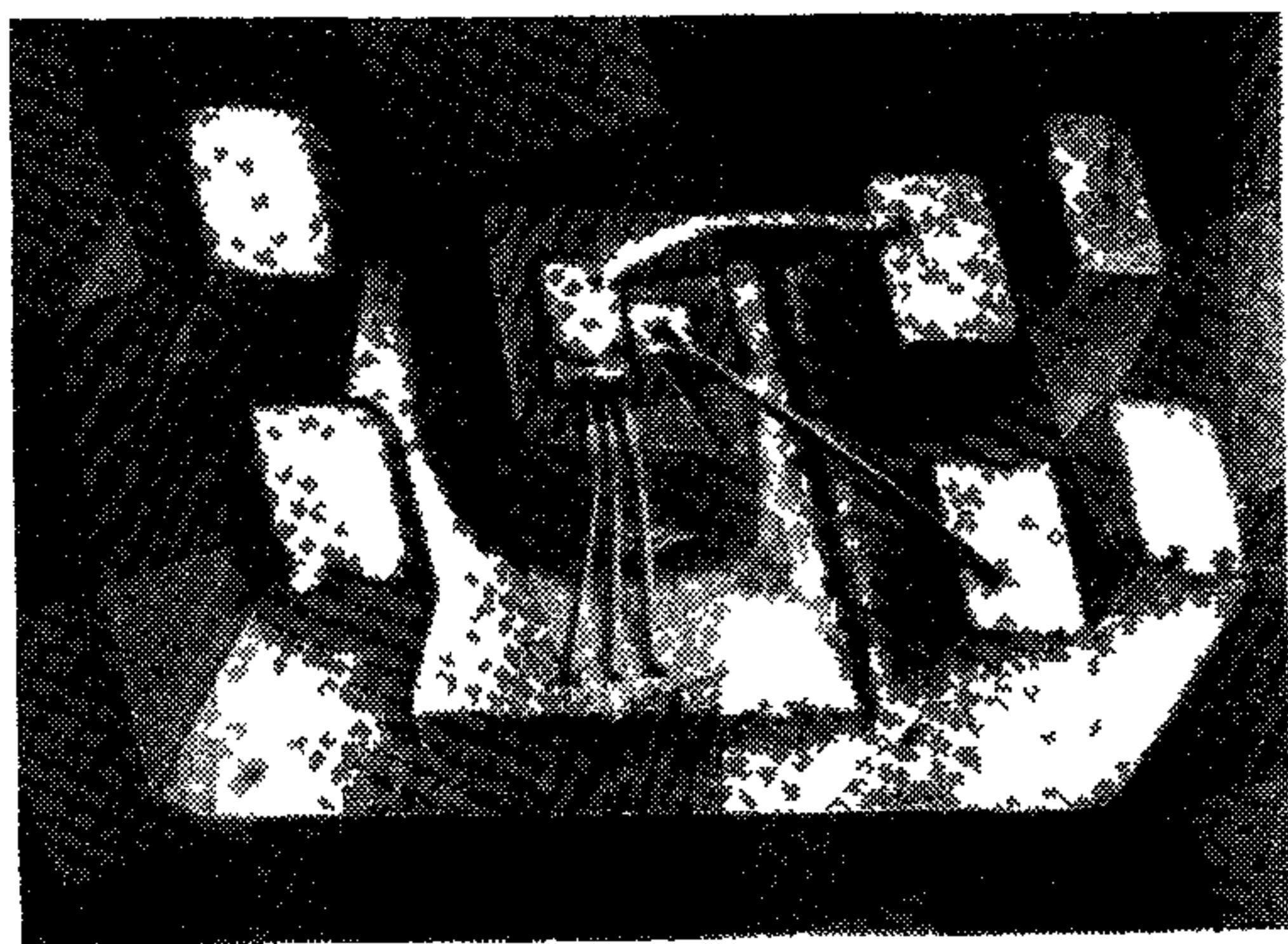


Рис 10.1. Два полевых транзистора на арсениде галлия, изготовленные фирмой Plessey Ltd. Эти приборы сочетают свойства низкого шума, высокого усиления и большой динамический диапазон. Они особенно пригодны в качестве высокочастотных усилителей в диапазоне частот до 200 МГц до 12 ГГц (Фотографии любезно представлены Р. Батлином и Plessey Research Ltd., Allen Clark Research Centre, Caswell.)

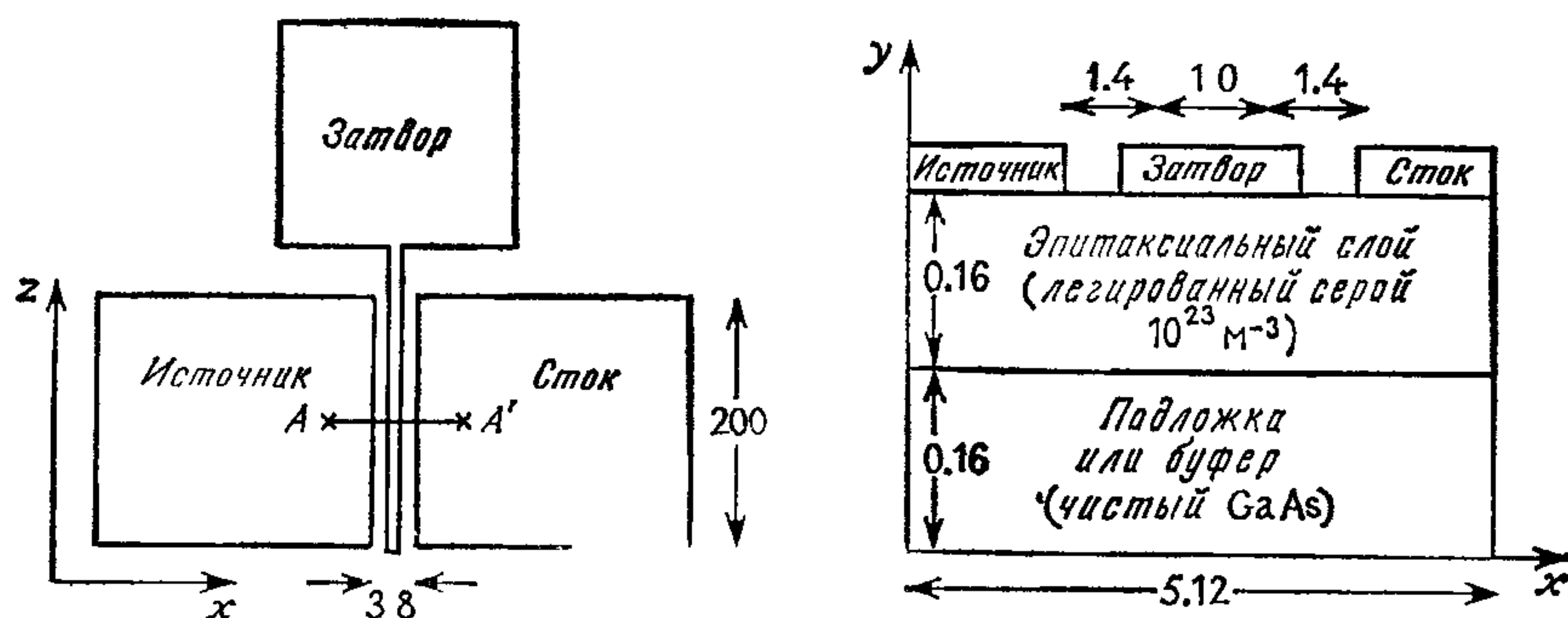


Рис. 10.2 Слева — вид в плане прибора, изображенного на рис 10.1, и справа — идеализированное вертикальное сечение по линии $A-A'$, показывающее активную область, моделируемую на компьютере. Размеры даны в микрометрах, не в масштабе

200 МГц до 12 ГГц. Их преимущества состоят в низком уровне шума, высоком коэффициенте усиления и широком динамическом диапазоне. Сами приборы имеют размеры порядка 0.3 мм на 0.3 мм и смонтированы на держателях. Слева показан бескорпусной перевернутый прибор размером 2 мм на 1 мм, используемый до 8 ГГц. Справа находится микрополосный корпус длиной 7 мм и максимальным диаметром 2 мм. Его верхняя часть удалена для показа транзисторного чипа. Имеются три электрода, источник S , затвор G и сток D . Протекание тока через прибор от источника к стоку управляется напряжением на затворе. Для тех, кто хорошо знаком с вакуумной электроникой, отметим, что этот прибор действует подобно триодному венгилю, причем источник, сток и затвор соответствуют в указанном порядке катоду, аноду и сетке.

На рис. 10.2 (слева) показан вид ПТ в плане. Три большие площадки размером (0.2×0.2) мм² образуют контактные выводы. Активная область прибора располагается в маленьком зазоре (~ 4 мкм) между площадками источника и стока. Сам затвор представляет собой длинную узкую полоску шириной в 1 мкм, лежащую в середине зазора между истоком и стоком. Геометрия прибора такова, что, за исключением нескольких микрометров от концов зазора (только 1% его длины), будут отсутствовать изменения величин в направлении z , параллельном сторонам зазора. Поэтому мы будем предполагать, что необходимо рассматривать только изменения поперек зазора (направление x) и по нормали к поверхности (направление y). Известно, что большая часть тока выходит из источника и поступает в сток внутри площадок источника и стока шириной около 1 мкм. Следовательно, хотя прибор занимает площадь около (400×400) мкм², только небольшая область, представ-

ленная сечением $A-A'$, играет активную роль в определении характеристик прибора.

На рис. 10.2 (справа) показано сечение ниже $A-A'$, образующее область, которую моделируют на компьютере. Полевые транзисторы изготавливаются путем наращивания тонкого эпитаксиального слоя (~ 0.16 мкм) легированного GaAs на пластинке кристаллического изолирующего GaAs ($\sim 10^5$ Ом·м), называемого подложкой. Эпитаксиальный слой выращивается в лаборатории в тщательно контролируемых условиях из пара, содержащего свободный мышьяк, и соответствующего соединения галлия с добавкой серы. Атомы серы после потери одного из внешних электронов в качестве свободного электрона проводимости становятся изоэлектронными с мышьяком и могут заменять мышьяк в решетке GaAs. В результате фиксированные положительные ионы серы жестко связываются как часть кристаллической решетки и, если они полностью ионизованы, равное число свободных электронов высвобождается в качестве проводящих электронов. Поскольку эти свободные носители заряжены отрицательно, говорят, что слой легирован как n -тип. Типичный уровень легирования составляет 10^{23} ионов серы и 10^{23} электронов проводимости на кубический метр.

В дополнение к отрицательным электронам все полупроводники содержат положительные дырки в качестве носителей тока. При комнатной температуре произведение электронной и дырочной плотностей постоянно ($\approx 10^{26}$ м⁻⁶), следовательно, в эпитаксиальном слое дырочная плотность составляет 10^3 м⁻³ [227, с. 369; 363, с. 58]. Эта плотность так мала по сравнению с электронной плотностью, что дырками можно пренебречь и рассматривать ПТ как униполярный прибор. Мы отличаем эту ситуацию от моделирования биполярных транзисторов, в которых необходимо представлять оба типа носителей тока.

Подложка вырезается из монокристалла арсенида галлия, выращенного методом медленного вытягивания из расплава. Такая подложка может содержать много примесей, в частности кремний, попадающий из кристаллизатора, который действует как донор электронов. Поэтому для снижения концентрации электронов проводимости до низкого уровня и для получения тем самым полуизолирующего материала добавляют хром, который захватывает электроны. Говорят, что хром компенсирует первичные примеси и в результате создается материал с низким уровнем активных примесей. В этой работе мы будем предполагать, что компенсация является точной, и рассматривать полуизолирующую подложку как чистый, беспримесный арсенид галлия.

Типичная подложка имеет толщину 250 мкм; однако только ее приповерхностная область, непосредственно примыкающая к эпитаксиальному слою, влияет на проводимость между источ-

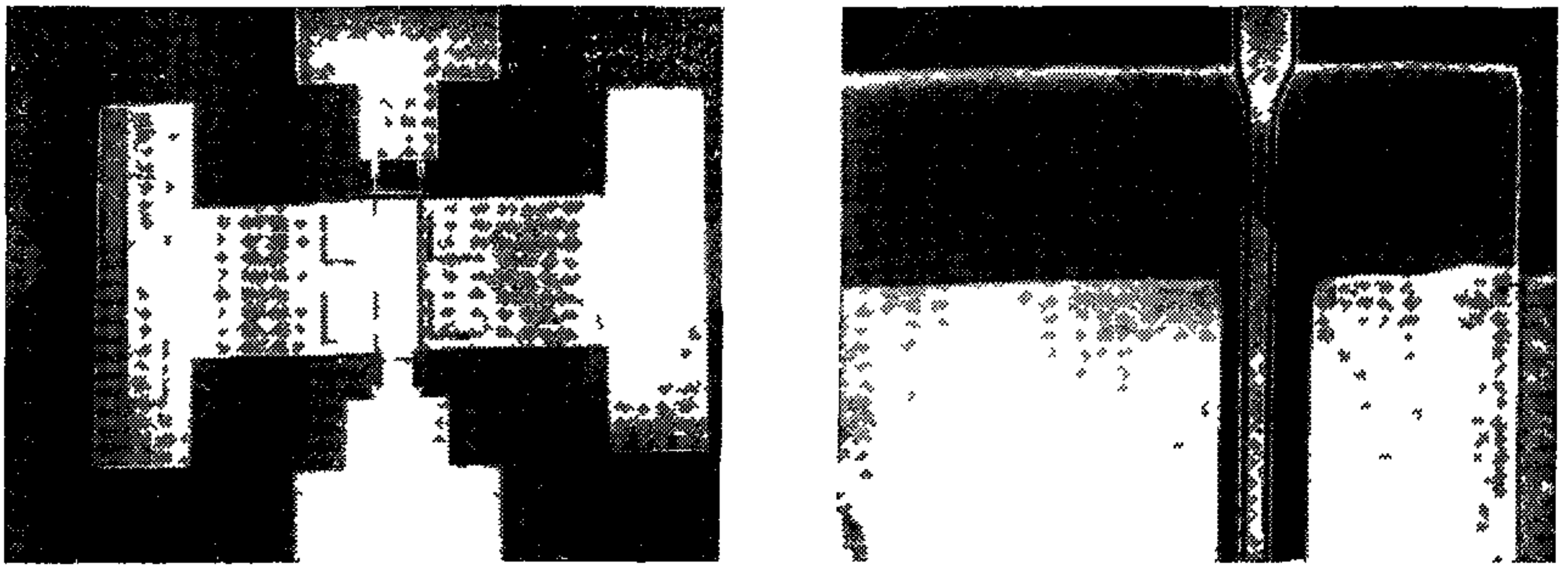


Рис 10.3 Слева — фотография, выполненная на сканирующем электронном микроскопе (1:120) пары полевых транзисторов с общим стоком, полный размер составляет приблизительно 1 мм. Справа — фотография, выполненная на сканирующем электронном микроскопе (1:4500), области источника и стока размером 3 мкм, которая моделировалась на компьютере. (Photographs Crown Copyright, 1980)

ником и стоком. Следовательно, обычно надо моделировать только часть подложки с толщиной, приблизительно равной толщине эпитаксиального слоя, как это показано на рис. 10.2 (справа). В некоторых приборах нелегированный слой чистого GaAs наращивается на полуизолирующую подложку до выращивания легированного слоя. Этот буферный слой действует как барьер, препятствующий диффузии примесей от подложки к активному легированному слою в процессе технологических операций. Подложка в описанной нами вычислительной модели является в действительности моделью буферного слоя, и наше рассмотрение его как чистого GaAs особенно подходит. Однако на практике в буферный слой может диффундировать довольно много хрома.

Электроды источника и стока изготовлены из сплава золота, индия и германия, который агломерирован на активный слой для получения контактов с малым омическим сопротивлением. Поэтому напряжение, приложенное к этим электродам, является также напряжением, приложенным к активному слою GaAs. Электрод затвора представляет собой слой алюминия, который образует хороший барьер Шоттки с GaAs. Это означает, что напряжение в GaAs, примыкающем к затвору, на 0.8 В меньше напряжения, приложенного к электроду затвора.

Геометрию типичного ПТ можно увидеть в увеличенном виде на фотографиях реальных приборов. На рис. 10.3 (слева) представлена выполненная сканирующим электронным микроскопом фотография ($\sim \times 120$) пары ПТ с общим стоком и соединенными затворами, соответствующая схеме на рис. 10.2 (слева) и ее зеркальному изображению относительно центра стока. Два зазора между источником и стоком видны как две черные

вертикальные линии, затворы же при этом увеличении неразличимы. Возрастание увеличения до $\sim \times 4500$ у правого зазора позволяет увидеть затвор (рис. 10.3, справа). За исключением того обстоятельства, что источник и сток переставлены, это точно соответствует диаграмме на рис. 10.2 (слева). Затвор выглядит как белая вертикальная линия шириной, приблизительно равной 1 мкм, расположенная примерно по центру черного зазора шириной ~ 3 мкм между прямоугольными площадками источника и стока.

Выделив геометрическую область для моделирования, мы должны теперь задать граничные условия. Электроны текут из источника к стоку под воздействием их собственного пространственного заряда и внешнего напряжения, приложенного к электродам. Что касается электростатического потенциала, ясно, что на границе, примыкающей к трем электродам, потенциал должен принимать заданные значения. На других участках границы требуется обращение в нуль нормальной компоненты электрического поля. Иначе говоря, нормальная компонента градиента потенциала равна нулю. Такие условия вокруг поверхности прямоугольника идеально подходят для решения с помощью метода быстрого решения эллиптических уравнений (RES), и вычислительные программы, как, например, POT4 (см. разд. 6.5.6), были написаны как раз для таких условий. Оправданием нулевых условий для поля на левой, правой и нижней границах является естественная тенденция свободного заряда компенсировать любое изменение поля в масштабе дебаевской длины (см. разд. 9.3.3) от любого заданного изменения (например, край электрода).

На верхней поверхности нулевые условия для поля являются хорошей аппроксимацией из-за больших изменений диэлектрической постоянной от эпитаксиального слоя ($\epsilon_e \approx 12$) до окружающего воздуха ($\epsilon_e \approx 1$). На таком разрыве нормальное значение электрического смещения $D = \epsilon E$ непрерывно. Следовательно,

$$\epsilon_e E_e = \epsilon_a E_a, \quad (10.1a)$$

$$E_e = \frac{\epsilon_a}{\epsilon_e} E_a \approx 0, \quad (10.1b)$$

где индексы e и a относятся соответственно к эпитаксиальной области и воздуху. Поскольку $\epsilon_e \gg \epsilon_a$ и E_a имеет конечное значение, электрическое поле внутри эпитаксиального слоя должно быть малым. В действительности более тонкий расчет включил бы вычисление электрического поля в воздушной области. Однако ясно, что это является уточнением, которое нет нужды рассматривать в первую очередь.

Рассмотрим теперь условия, которые необходимо наложить на движение электронов у границ. Требование отсутствия тока

через границы, за исключением электродов, лучше всего удовлетворяется отражением любых электронов, которые попадают на границу. Иными словами, нормальную к границе компоненту скорости электрона обращают. На электродах источника и стока омический контакт, как предполагается, поглощает все попадающие на электрод электроны и инжектирует на каждом временном шаге достаточное количество электронов для поддержания плотности 10^{23} м^{-3} в строке ячеек сетки потенциала, примыкающих к электродам. Инжекция происходит с полумаксвелловским распределением по скоростям тем же способом, как для примера по физике плазмы (см. разд. 9.3.1). На электроде затвора контакт с барьером Шоттки представляется с помощью напряжения, равного разности приложенного потенциала и высоты барьера 0.8 В. Здесь имеется поглощение частиц, а инжекция отсутствует. Эти последние условия на затворе на практике неважны, поскольку при обычных условиях смещения очень мало электронов будет достигать затвора.

10.1.3. Типы моделей

Остальная часть этой главы посвящена использованию методов частиц при моделировании полевых транзисторов. Чтобы показать их место, мы сначала приведем обзор различных типов моделей, которые использованы при моделировании полупроводников. Сейчас имеется обширная литература и в библиографии [12] даны 486 ссылок вплоть до 1974 г. Читатель отсылается также к книге по моделированию полупроводниковых приборов [158].

Простейшая модель представляет такой прибор с помощью эквивалентной цепи с сосредоточенными параметрами, которая описана, например, в разд. 10.4.4. Свойства прибора получают затем стандартными методами анализа цепей [132]. В [109] изучены динамические характеристики ПТ этим методом. Такие расчеты требуют небольших затрат машинного времени, и возможен широкий диапазон параметров; однако это достигается игнорированием распределенного характера параметров (например, сопротивления и емкости) и сложной геометрии прибора. Поэтому метод имеет лишь ограниченную применимость, но в том случае, если эти ограничения хорошо известны и понятны при сравнении с более точными моделями, он является привлекательным из-за минимального потребления вычислительных ресурсов. Метод также популярен, поскольку выражает поведение прибора через элементы цепи, свойства которых хорошо известны инженерам по электронным цепям.

Следующая наиболее сложная модель описывает течение электрического тока при помощи гидродинамических уравнений;

1. Уравнение Пуассона для электрического поля E

$$\nabla^2 \varphi = \frac{e}{\epsilon \epsilon_0} (n - N), \quad E = -\nabla \varphi. \quad (10.2a)$$

2. Уравнение для потока

$$\Gamma = n\mathbf{v} - D\nabla n, \quad \mathbf{v} = -\mu E. \quad (10.2b)$$

3. Уравнение неразрывности

$$\frac{\partial n}{\partial t} = -\nabla \cdot \Gamma. \quad (10.2c)$$

В написанных выше уравнениях ток переносится зарядами $-e$ ¹⁾ с концентрацией n и средней скоростью \mathbf{v} ; Γ — поток носителей и N — концентрация примесных зарядов $+e$. Мы выписали уравнения для одного типа носителей (электронов), что соответствует ПТ. Если важны более одного типа носителей, уравнения (10.2b) и (10.2c) следует выписать для каждого типа и в правую часть уравнения Пуассона добавить дополнительные члены. В упомянутые выше уравнения физические свойства материала входят только через подвижность μ , коэффициент диффузии D и диэлектрическую постоянную ϵ . В [327—329] обсуждаются двумерные конечно-разностные аппроксимации уравнений (10.2) и неустойчивости, обусловленные временным шагом, которые могут возникнуть. Приложения к изучению ПТ даны в [326, 328, 389]. Процедура начинается с данных о распределении плотности в заданный момент, вычисляется электростатический потенциал и поле при помощи (10.2a) и потом находится плотность на следующем временном слое при помощи (10.2b) и (10.2c). Затем цикл по временному шагу повторяется. Уравнение Пуассона является линейным по φ и в большинстве случаев может быть решено с помощью метода быстрого решения эллиптических уравнений (см. разд. 6.5.1—6.5.7), хотя многие авторы использовали итерационные схемы (см. разд. 6.3.1—6.3.6), методы конечных элементов (см. [24, 25] и разд. 6.4.3) или строго неявные процедуры (см. [221] и разд. 6.4.5). Неявные конечно-разностные схемы по времени для (10.2c) оказались необходимыми, чтобы избежать чрезмерно малого временного шага. Обычно используется обобщенная схема Кранка — Никольсона [97] с равными весами для новых и старых временных слоев.

Решение двумерных установившихся уравнений ((10.2) с $\partial/\partial t = 0$) получено Кеннеди и О'Брайном [225] при помощи итерации по Пикару. Конечно-разностные аппроксимации уравнений (10.2) были получены путем сеточной релаксации, и выполнен анализ работы ПТ с $(p-n)$ -переходом на кремнии. Анало-

¹⁾ $e \equiv +1.60210 \times 10^{-19}$ Кл — абсолютная величина заряда электрона.

гичная модель использовалась в [174] при изучении ПТ с $(p-n)$ -переходом на GaAs. ПТ типа металл — полупроводник из кремния, GaAs и InP промоделированы в [379] при помощи программы быстрого решения эллиптических уравнений POTH (см. разд. 6.5.6) для решения уравнения Пуассона. Уравнение неразрывности решено с использованием нижней релаксации при помощи метода [342] для предотвращения численных неустойчивостей. Были промоделированы МОП (металл — окись — полупроводник) ПТ на кремнии (MOST) [266] и биполярные транзисторы [226] с помощью конечно-разностных аппроксимаций и итераций по Ньютону [349].

Плотность электронов в полупроводниковых приборах изменяется на много порядков величины, потому что при термодинамическом равновесии плотность пропорциональна $\exp(e\phi/k_B T)$, и малое изменение в потенциале ϕ приводит к большим изменениям в плотности. В [348] предложено заменить переменную плотности на

$$\Phi = n \exp(-e\phi/k_B T). \quad (10.3)$$

Из-за введения экспоненциального множителя в ее определение новая переменная не имеет больших количественных изменений. Следовательно, погрешности дискретизации понижаются. В терминах новой переменной Слотбума уравнение (10.2) принимает вид

$$\nabla^2 \Phi = \frac{e}{\epsilon \epsilon_0} [\Phi \exp(e\phi/k_B T) - N], \quad (10.4a)$$

$$\frac{\partial}{\partial t} [\Phi \exp(e\phi/k_B T)] = -\nabla \cdot [D \exp(e\phi/k_B T) \nabla \Phi], \quad (10.4b)$$

где использовано соотношение Эйнштейна $\mu = eD/k_B T$ для связи подвижности и коэффициента диффузии. Процедура нахождения решения состоит в решении (10.4a) для потенциала ϕ при заданном значении Φ и вычислении затем Φ на новом временном слое из (10.4b) при заданной величине ϕ . Заметим, что в новых переменных уравнение Пуассона (10.4a) нелинейно и может быть решено только итерациями. Решение обычно находится с помощью метода Ньютона (см. п. 6.2.1). Уравнение (10.4b) является, как прежде, линейным уравнением по Φ с переменными коэффициентами и может быть решено с помощью метода сеточной релаксации или матричных методов (см. разд. 6.3.1—6.4.6). Поэтому расплатой за введение переменной с более хорошим свойством является возросшее усилие, требуемое для решения уравнения Пуассона.

При условии установившегося течения уравнение (10.4b) переходит в уравнение

$$\nabla \cdot [D \exp(e\phi/k_B T) \nabla \Phi] = 0, \quad (10.5)$$

которое может рассматриваться как уравнение Лапласа с переменными коэффициентами для переменной Слотбума Φ и решается с помощью сеточной релаксации, матричных методов или метода Конкуса и Голуба (см. разд. 6.5.7). Полное решение задачи с установившимся течением требует одновременного решения уравнений (10.4а) и (10.5). Многие авторы используют алгоритм, описанный в [156] для одномерного случая. Метод основан на итерациях по Ньютону, улучшен и переформулирован на двумерный случай в [19] с помощью вариации переменных Слотбума.

В проведенном выше обсуждении предполагалось, что температура прибора остается постоянной. При моделировании теплового пробоя мощных приборов также важно включать эффекты генерации тепла в приборе и изменения подвижности и коэффициента диффузии с температурой. Это требует решения уравнения теплопроводности

$$\nabla \cdot (K \nabla T) = \rho c \frac{\partial T}{\partial t} + H \quad (10.6)$$

для температуры T . В этом уравнении K — теплопроводность, ρ — массовая плотность, c — теплоемкость, H — источник тепла. Решение уравнения (10.6) для стационарного течения в сочетании с уравнениями Пуассона и неразрывности (10.2) или (10.4) широко изучалось в [141]. В качестве переменных используются плотности носителей, и нелинейные уравнения решаются с помощью итераций по Ньютону. Для решения задачи Пуассона с переменными коэффициентами используется метод Стоуна (см. разд. 6.4.6). Зависящее от времени решение этих уравнений найдено в одномерном [82] и двумерном [15] случаях с помощью следующего цикла:

1. Нахождение потенциала из (10.2а).
2. Вычисление E , μ , D как функций поля и температуры.
3. Решение уравнения неразрывности (10.2с) для плотности.
4. Вычисление генерации тепла и решение (10.6) для нового значения температуры.

Заметим, что в данном выше цикле решение уравнения типа Пуассона с переменными коэффициентами требуется в трех случаях: во время выполнения (1), (3) и (4).

Описанные выше модели предполагают, что скорость носителей зарядов мгновенно подстраивается к электрическому полю через соотношение для подвижности $v = \mu E$. Здесь не принято во внимание время, требуемое для ускорения электронов до их предельной скорости, или временные запаздывания, связанные с рассеянием электронов в состоянии с более высокой энергией. Для полного описания этих эффектов требуется модель, учитывающая рассеяние частиц с помощью процедуры Монте-Карло.

Это является предметом оставшейся части данной главы. Такие временные запаздывания не имеют значения для материалов с одной энергетической зоной, подобных кремнию, если при этом период осцилляции много больше времени рассеяния, что обычно бывает выполнено. Для других материалов, подобных арсениду галлия, электроны могут находиться в двух энергетических зонах (см. п. 10.2.1) и необходимо принимать во внимание временные запаздывания, связанные с рассеянием между этими зонами.

В работах [48, 347], а также Робсоном из Шеффилдского университета показано, как эти эффекты могут частично учитываться введением дифференциальных уравнений первого порядка для средней дрейфовой скорости \mathbf{v} и хаотической энергии (т. е. температуры) электронов

$$\frac{d\mathbf{v}}{dt} = -\frac{e\mathbf{E}}{m^*} - \frac{\mathbf{v}}{\tau_m}, \quad (10.7a)$$

$$\frac{d\xi}{dt} = -e\mathbf{v} \cdot \mathbf{E} - \frac{\xi - \xi_L}{\tau_e}, \quad (10.7b)$$

где τ_m и τ_e — соответственно постоянные времени релаксации импульса и энергии и ξ_L — хаотическая тепловая энергия решетки (т. е. $3/2 k_B T$). Эффективная масса m^* должна учитывать пропорцию электронов в каждой энергетической зоне:

$$m^* = m_1(1 - p_2) + m_2 p_2, \quad (10.8)$$

где m_1 и m_2 — массы электронов в зоне I и зоне II соответственно и p_2 — пропорция электронов в зоне II.

Временные константы найдены из рассмотрения решения уравнения (10.7) в установившемся состоянии ($d/dt = 0$); следовательно,

$$\tau_m = \frac{m^* v}{e E} = \frac{m^*}{e} \mu, \quad (10.9a)^1)$$

$$\tau_e = -\frac{\xi - \xi_L}{e\mathbf{v} \cdot \mathbf{E}} = \frac{\xi - \xi_L}{e\mu E^2}, \quad (10.9b)$$

где $\mu = v/E$ — подвижность. Величины v , p_2 и m^* найдены из моделирования методом Монте-Карло (см. [133, 173]) как функции энергии теплового движения ξ , а временные константы оценены с помощью уравнения (10.9) также как функции ξ . Найдено, что упомянутые выше константы при использовании в уравнении (10.7) дают необыкновенно хорошее согласие с нестационарным моделированием методом Монте-Карло всплеска скорости в однородном электрическом поле. Таким образом

¹⁾ Векторные величины (например, \mathbf{E}) помечены жирным шрифтом, а их модули (например, E) — курсивом.

успешно предсказано зависящее от времени поведение используемых параметров, полученное из стационарных данных. Это приводит к уверенности относительно допущений в этом приближении. В более сложных ситуациях, обусловленных геометрией реального прибора, согласие менее удовлетворительное.

Преимущество модифицированной гидродинамической модели состоит в том, что она требует меньше вычислительных затрат, чем полная модель частиц, основанная на методе Монте-Карло. Однако нельзя ожидать, что все сложности электронного рассеяния можно представить двумя временными константами, и такая модель должна всегда сравниваться с последним типом модели, к описанию исторического развития которой мы переходим.

Моделирование полупроводников с использованием частиц стало возможным с помощью методов, первоначально развитых для вакуумных ламп и плазмы, в которой отсутствует рассеяние на фоновой среде. Эти методы были разработаны в период второй мировой войны проф Д. Р. Хартри, когда он руководил группой Манчестерского университета при Комитете по развитию электронных ламп (CVD) Британского адмиралтейства. В период с 1941 по 1944 гг. проф. Хартри и Филлис Никольсон [167] подсчитали с помощью арифмометра орбиты около 30 взаимодействующих электронов в магнетроне, используя одномерную аппроксимацию электрического поля (см. [65, с. 36]). В их модель включены эффект пространственного заряда, электрические и магнитные поля, но из-за ограниченности одномерностью она не в состоянии представить полное двумерное течение зарядов в магнетроне. Позднее Хартри [166] выполнил аналогичным образом исследование течения электронов в одномерном диоде и наблюдал начальный переход, появляющийся, когда пустой диод наполняется электронами. Примерно в это время стали доступными первые цифровые компьютеры и, возможно, первое моделирование на таких компьютерах представляло изучение шума в высокочастотном диоде, выполненное в Белл-лаборатории Тьеном и Мошменом [368]. Они применили Univac 1 для моделирования одномерного диода с использованием примерно 360 электронных листов и физического шага по времени, равного 2 пс. Было просчитано около 3000 временных шагов, на каждый из которых затрачено от 25 до 40 с машинного времени. Полный расчет задачи должен причисляться к одному из наиболее впечатляющих достижений раннего моделирования прибора, требующего для выполнения около 25 ч надежного вычисления на первом поколении компьютеров на вакуумных электронных лампах с памятью всего 1000 чисел на ртутной задерживающей линии.

Ранние расчеты Хартри были приспособлены для автоматических вычислений на ЭВМ EDSAC 2 в Кембриджском университете и продолжены на существенно бóльшие времена его учеником Ломаксом [267], обнаружившим неожиданные осцилляции пространственного заряда около ожидавшегося статического решения уравнения. Эти плазменные колебания наблюдались также на аналогичной модели примерно в то же самое время Бэрдсолом и Бриджесом [32, 33] в Калифорнийском университете в Беркли. Другая ранняя одномерная модель была развита Бунеманом [56] в Стэнфордском университете, который сам являлся учеником Хартри и работал в группе CVD в Манчестере. Эта модель с 512 листами была запрограммирована на ЭВМ Univac 1103AF Д. Тоу (D. Thoe) из Локхида и использована для демонстрации неожиданно быстрой диссипации токов при вхождении в ионизованную среду. Позже Бюргер, работая в качестве ученика Бунемана, усовершенствовал одномерную модель путем использования арифметики целых чисел и сетки для расчета полей. Эта модель допускала моделирование 10 000 электронных листов на IBM 7090 и была использована для изучения действия цезиевого термоэлектронного преобразователя и плазменных диодов [66—67, 69].

Распространение моделирования на двумерный случай стало возможным в 1964 г. с появлением нового метода решения уравнения Пуассона, предложенного Хокни ([175], алгоритм FACR, гл. 6), также ученика Бунемана, но занимавшегося моделированием плазмы. Этот метод позволил осуществить моделирование 2000 зарядов, представленных в форме стержней, и был успешно применен для демонстрации аномальной диффузии в плазме (см. [177] и гл. 9); для моделирования двумерного течения зарядов в магнетронах [391], завершая, таким образом, 20 годами позднее работы Хартри во время войны; для моделирования нейтрализации ионных пучков, предназначенных для реактивных двигателей космических аппаратов [380].

Для включения столкновений электронов с нейтральными атомами при моделировании одномерного плазменного диода Бюргер [68] впервые ввел в упомянутую выше бесстолкновительную модель рассеяние на фоновой среде. Для выбора средней длины свободного пробега при упругих столкновениях с жесткими сферами использовались случайные числа. Тот же общий метод был использован Лебвалом и Прайсом [248] для моделирования течения электронов в арсениде галлия и образования доменов Ганна. Двумерные модели с рассеянием были введены в работе [193] и развиты далее в [381—384]. Эти модели являются содержанием остальной части этой главы.

10.2. ПЕРЕНОС ЭЛЕКТРОНОВ В ПОЛУПРОВОДНИКАХ

10.2.1. Уравнения движения

Перед тем как перейти к моделированию движения электронов в полупроводнике, необходимо рассмотреть основные физические законы. Движение электронов в полупроводнике описывается в терминах зонной структуры (рис. 10.4), которая представляет соотношение $\mathcal{E}(\mathbf{k})$ между энергией электрона \mathcal{E} и волновым вектором \mathbf{k} (иногда его называют \mathbf{k} -вектором). Зонная структура находится из квантово-механических расчетов, которые здесь нас не интересуют (см., например, [227, гл. 9 и 10]). С точки зрения моделирования прибора можно взять зонную структуру как заданную и необходимо только узнать, как ее использовать для определения движения наших моделируемых электронов.

Электрон с волновым вектором \mathbf{k} имеет импульс \mathbf{p} и энергию \mathcal{E} , заданные выражениями

$$\mathbf{p} = \hbar \mathbf{k}, \quad \mathcal{E} = \mathcal{E}(\mathbf{k}), \quad (10.10a)$$

где \hbar — постоянная Планка, деленная на 2π . Электрон движется с групповой скоростью волнового пакета

$$\mathbf{v} = \frac{d\mathbf{x}}{dt} = \frac{1}{\hbar} \frac{d\mathcal{E}}{d\mathbf{k}} = \frac{1}{\hbar} \left(\frac{\partial \mathcal{E}}{\partial k_x}, \frac{\partial \mathcal{E}}{\partial k_y}, \frac{\partial \mathcal{E}}{\partial k_z} \right). \quad (10.10b)$$

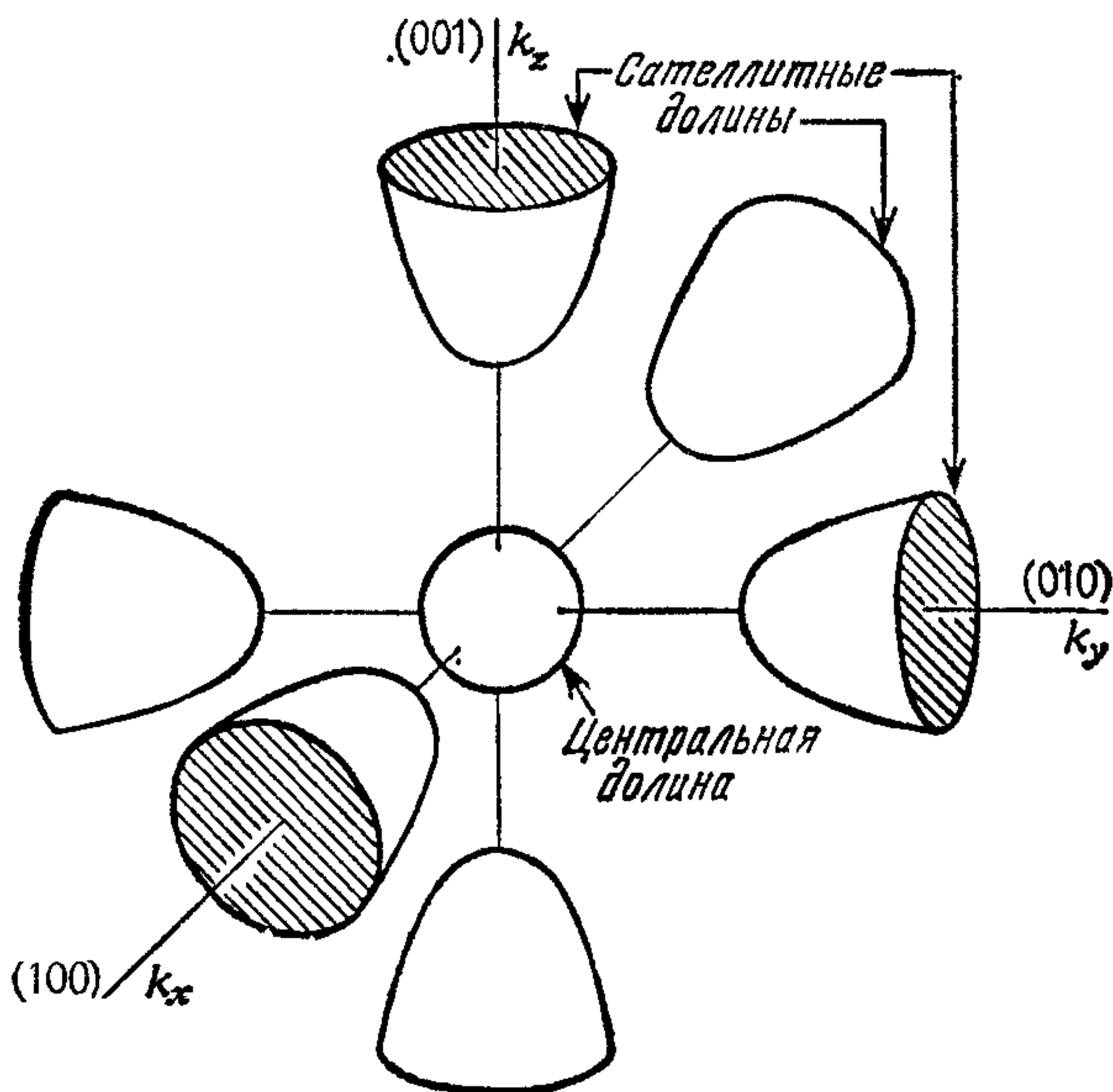


Рис. 10.4. Трехмерная зонная структура арсенида галлия. Показана одна поверхность постоянной энергии в центральной долине и в каждой из шести сателлитных долин. (Любезно предоставлено Уорринером, [381].)

Скорость изменения импульса всегда равна приложенной силе F , следовательно,

$$\frac{d\mathbf{k}}{dt} = \frac{1}{\hbar} \mathbf{F}. \quad (10.10c)$$

Уравнения (10.10b) и (10.10c) полностью определяют движение электрона, если известна зонная структура $\mathcal{E}(\mathbf{k})$. Мы запоминаем координаты каждого электрона (x, y, k_x, k_z) и продвигаем \mathbf{k} -вектор во времени согласно (10.10c), а положения частицы в соответствии с (10.10b). Заметим также, что скорость направлена по градиенту энергетической поверхности в \mathbf{k} -пространстве. Следовательно, скорость направлена так же, как и вектор \mathbf{k} , если поверхности энергии являются сферическими относительно начала вектора \mathbf{k} .

Чтобы связать движение электрона с движением классической частицы и известным понятием массы, уравнение (10.10c) может быть выражено эквивалентно в терминах ускорения (см. [39, с. 208])

$$\frac{dv_i}{dt} = F_i / m_{ij}^*, \quad (10.11a)$$

$$\frac{dx_i}{dt} = v_i, \quad (10.11b)$$

где

$$m_{ij}^* = \hbar^2 \left(\frac{\partial^2 \mathcal{E}}{\partial k_i \partial k_j} \right)^{-1}, \quad (10.11c)$$

i и j указывают направление координат x, y, z . Мы запоминаем обычные координаты (x, y, v_x, v_y, v_z) каждой частицы и продвигаем скорости согласно (10.11a), а положения частиц согласно (10.11b).

По аналогии с законом Ньютона для движения классической частицы с массой m , т. е.

$$\frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m}, \quad \frac{d\mathbf{x}}{dt} = \mathbf{v}, \quad (10.12)$$

мы видим, что величина m_{ij}^* , которая обратно пропорциональна кривизне энергетической поверхности, играет роль массы и называется *эффективной массой* электрона. Обычно она зависит от энергии электрона и существенно отличается от массы покоя m_e неподвижного электрона в свободном пространстве. Будучи зависимой от кривизны $\partial^2 \mathcal{E} / \partial k_i \partial k_j$ энергетической зоны, масса может стать даже отрицательной. Эффективная масса имеет 9 компонент (т. е. она является тензором)

$$m^* = \begin{bmatrix} m_{xx}^* & m_{xy}^* & m_{xz}^* \\ m_{yx}^* & m_{yy}^* & m_{yz}^* \\ m_{zx}^* & m_{zy}^* & m_{zz}^* \end{bmatrix}, \quad (10.13)$$

и уравнение (10.11a) показывает, что, если все компоненты m^* отличны от нуля, сила, направленная вдоль одной координаты, производит ускорение по всем трем координатным направлениям. Другими словами, ускорение и вызывающая его сила не параллельны.

Данная выше формулировка в терминах эффективной массы является такой же общей, как уравнения (10.10b) и (10.10c), и показывает, что движение электрона в полупроводнике можно рассматривать как движение классической частицы с зависящей от энергии массой в тензорной форме. Необычный характер тензора эффективной массы m_{ij}^* просто отражает квантово-механическое воздействие полупроводящей среды на движение электрона. Эта формулировка особенно полезна при качественном анализе движения электрона посредством введения понятия эффективной массы. Однако вектор \mathbf{k} не вычисляется в (10.11), и трудности при выражении эффективной массы в терминах скоростей делают формулировку неудобной для вычислительных целей, за исключением простого случая, когда эффективная масса не зависит от энергии. Это имеет место в так называемых параболических энергетических зонах, в которых энергия пропорциональна квадрату вектора \mathbf{k} .

К счастью, при комнатной температуре большая часть электронов будет располагаться около дна (т. е. в пределах 0.025 эВ) минимумов в зонной структуре, которых может быть несколько. В этом случае зонная структура вблизи каждого минимума может приближенно представляться как параболическая. Если, кроме того, предположим, что зона является сферически симметричной относительно минимума, то, измеряя \mathbf{k} от минимума, получаем

$$\mathcal{E}(\mathbf{k}) = ck^2 = c(k_x^2 + k_y^2 + k_z^2), \quad (10.14a)$$

где c — константа, и, следовательно,

$$\mathbf{v} = \frac{2c}{\hbar} \mathbf{k} = \frac{2c}{\hbar} (k_x, k_y, k_z) = \frac{\hbar}{m^*} \mathbf{k}, \quad (10.14b)$$

$$m_{ij}^* = \frac{\hbar^2}{2c} \delta_{ij} = m^*. \quad (10.14c)$$

Таким образом, находим, что масса является скалярной величиной m^* , не зависящей от направления и энергии. К тому же недиагональные элементы m_{ij}^* равны нулю, что означает, что ускорение направлено вдоль силы. Поэтому движение идентично движению классической частицы m' с массой m^* , подчиняющемуся законам Ньютона

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}, \quad (10.15a)$$

$$\frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m^*}, \quad (10.15b)$$

где $m^* = \hbar/2c$ и энергия рассматривается как обычная кинетическая энергия

$$\mathcal{E}(k) = \frac{\hbar^2}{2m^*} k^2 = \frac{1}{2} m^* v^2. \quad (10.15c)$$

Если сферически-параболическая аппроксимация является удовлетворительной, а это определенно имеет место в качестве первого приближения, то программы для ускорения и движения, написанные для моделирования плазмы (см. гл. 9), могут использоваться при моделировании полупроводников для описания свободного пролета при выборе правильной эффективной массы. Формулировка эффективной массы может также применяться, если энергетические поверхности являются эллипсоидами и изменяются параболически вдоль каждой оси. В этом случае масса по-прежнему не зависит от энергии, но уже не является скалярной величиной. Имеются различные эффективные массы вдоль трех осей эллипсоидов, и должна использоваться тензорная форма ускорения (10.11a).

Вдали от минимума зоны становятся важными эффекты непараболичности, и для зонной структуры обычно используются следующая аппроксимация:

$$\frac{\hbar^2 k^2}{2m^*} = \mathcal{E} (1 + \alpha \mathcal{E}), \quad (10.16a)$$

где

$$\alpha = \frac{1}{E_g} \left(1 + \frac{m^*}{m_e} \right)^2 \quad (10.16b)$$

и E_g — энергетическая щель между проводящей и валентной зонами. Соотношение между скоростью и волновым вектором получается из (10.10b):

$$\mathbf{v} = \frac{\hbar \mathbf{k}}{m^* (1 + 2\alpha \mathcal{E})}. \quad (10.16c)$$

В слабых электрических полях, когда энергия электронов мала, эффект непараболичности тоже мал. Однако в сильных электрических полях, которые появляются между затвором и стоком ПТ, энергия электрона достигает около $0.24 E_g$ и непараболический член $2\alpha \mathcal{E}$ в знаменателе уравнения (10.16c) приблизительно равен 0.5. В этих случаях, очевидно, необходимо использовать (10.16c) вместо (10.14b). Поскольку эффективная масса теперь зависит от энергии, должна использоваться формулировка уравнений движения в терминах вектора \mathbf{k} (10.10 b, c).

10.2.2. Зонная структура арсенида галлия

Трехмерная зонная структура арсенида галлия показана на рис. 10.4. Поверхности постоянной энергии $\mathcal{E}(\mathbf{k})$ изображены в трехмерном пространстве векторов \mathbf{k} . При обычных температу-

рах и в слабых электрических полях все электроны проводимости находятся в центральной (000) долине вблизи ее минимума при $k = 0$. Поверхности постоянной энергии являются сферами, центрированными в начале координат, и на рисунке показана одна такая сфера. Электроны в центральной долине ведут себя подобно классическим частицам с эффективной массой $m_1^* = 0,067 m_e$, хотя при больших энергиях должна быть сделана поправка на непараболичность долины с помощью (10.16с) с $\alpha = 0.576 \text{ эВ}^{-1}$. Электроны в этой долине являются, таким образом, легкими и подвижными. Если электрическое поле достаточно велико, чтобы сообщить электронам энергию $\sim 0.36 \text{ эВ}$, то имеется большая вероятность того, что они будут рассеиваться в одну из шести спутных (или верхних) долин, которые лежат в направлениях (100), (010) и (001) в обратной решетке (см. (227, гл. 2]). Хотя спутные долины являются эллип-

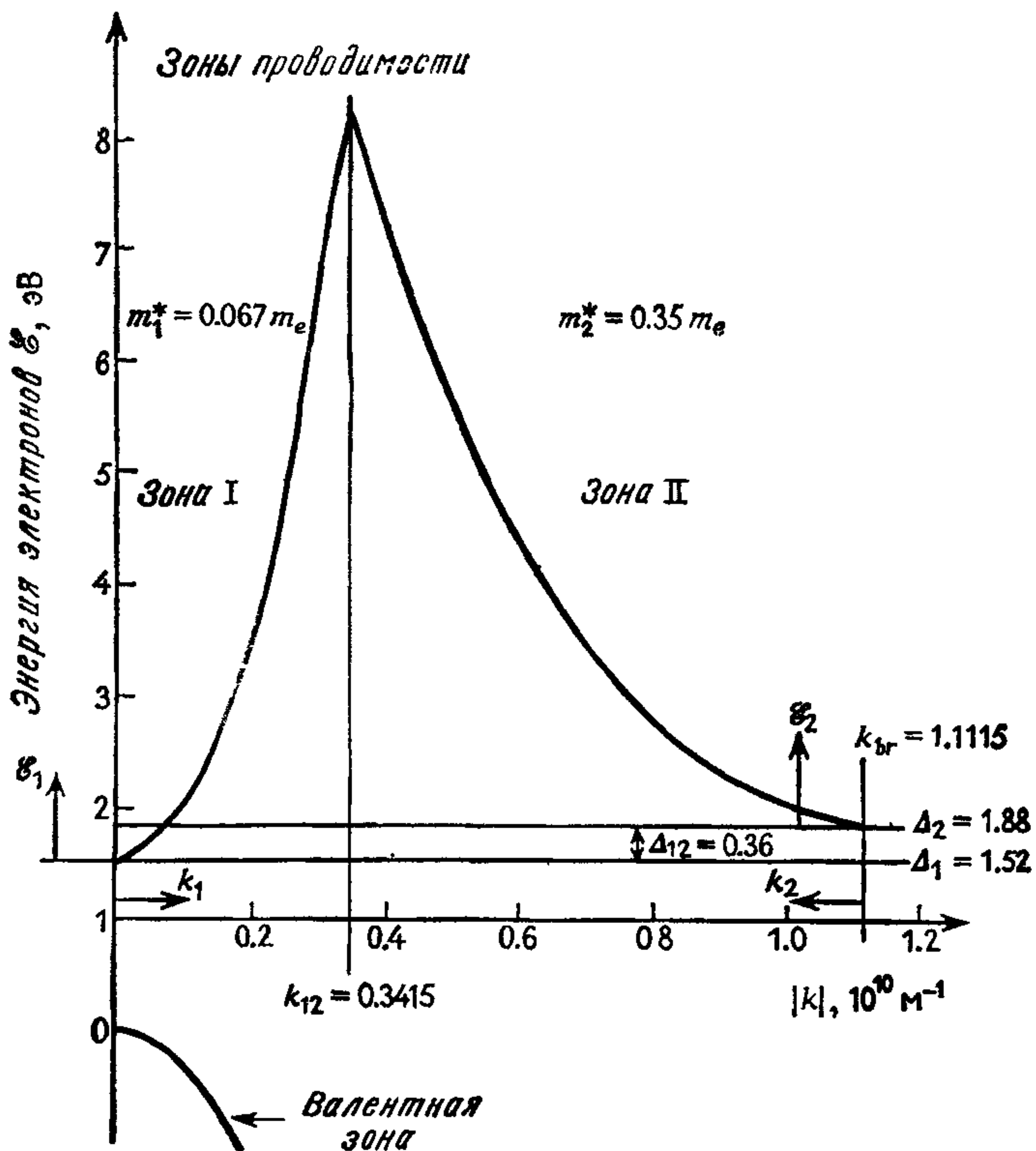


Рис 105 Упрощенная одномерная зонная структура арсенида галлия, использованная в вычислительной модели. Шесть спутных долин из рис. 10.4 объединены в одну составную зону II. Энергии отсчитываются от вершины валентной зоны в электронвольтах ($1 \text{ эВ} = 1.6 \times 10^{-19} \text{ Дж}$). (Любезно предоставлено Уорринером, [381])

соидальными и имеют тензорную эффективную массу, они будут рассматриваться как сферические зоны со скалярной эффективной массой $m_2^* = 0.35m_e$. В таком состоянии электрон в пять раз массивней, чем если бы он находился в центральной долине. Следовательно, электроны в сателлитных долинах инертны и имеют низкую подвижность.

В вычислительной модели мы не делаем попытки учесть все сложности трехмерной зонной структуры, хотя в принципе это можно сделать. Вычислительные расходы таких программ были бы, вероятно, неудовлетворительными. Вместо этого используем упрощенную одномерную зонную структуру, включающую две долины. Эта структура показана на рис. 10.5. Зона I соответствует центральному минимуму, в то время как шесть сателлитных минимумов объединены в составную зону II. Эта упрощенная зонная структура легче для расчета и содержит главные физические особенности материала. Рисунок начерчен в масштабе и показывает валентную зону, энергетическую щель $E_g = 1.52$ эВ между вершиной валентной зоны и дном центральной долины зоны проводимости и энергетический зазор в 0.36 эВ между дном сателлитной и центральной долин.

10.2.3. Процессы рассеяния

Уравнение движения электрона в совершенной решетке, вытекающее из (10.10с), при температуре, равной нулю по Кельвину, имеет вид

$$\frac{d\mathbf{k}}{dt} = \frac{-e}{\hbar} \mathbf{E}, \quad (10.17a)$$

где \mathbf{E} — электрическое поле. Для случая сферическо-параболической зоны (уравнение (10.15)) последнее переходит в закон движения Ньютона в его обычной форме

$$m^* \frac{d\mathbf{v}}{dt} = -e\mathbf{E}. \quad (10.17b)$$

В дефектной решетке при конечной температуре электроны не двигаются свободно через решетку в соответствии с (10.17). Отклонения от этого движения обусловлены механизмами рассеяния, которые могут изменять как величину, так и направление волнового вектора \mathbf{k} с известными вероятностями. В арсениде галлия важны следующие семь процессов рассеяния:

<i>Зона I</i>	<i>Физический источник</i>
1. Полярное оптическое	Тепловые колебания решетки
2. Акустическое	Тепловые колебания решетки
3. Междолинное	Неэквивалентное междолинное рассеяние из зоны I в зону II

Зона II

4. Полярное оптическое	Тепловые колебания решетки
5. Акустическое	Тепловые колебания решетки
6. Междолинное	Неэквивалентное междолинное рассеяние из зоны II в зону I
7. Внутримолинное	Эквивалентное междолинное рассеяние между шестью сателлитными долинами зоны II

Каждый процесс рассеяния описывается правилом для расчета нового волнового вектора k' из старого волнового вектора k и формулой для вероятности процесса рассеяния. Пусть $\lambda_i(\mathcal{E}) =$ вероятность рассеяния на i -ом процессе, $p_i(\theta) d\theta =$ вероятность нахождения угла между векторами k и k' в интервале θ и $\theta + d\theta$. Новая энергия после рассеяния будет равна

$$\mathcal{E}' = \mathcal{E} \pm \hbar\omega_i, \quad (10.18)$$

где ω_i — частота фонона. Используется положительный знак, если решетка передает энергию электрону (т. е. фонон поглощается), и отрицательный знак, если электрон передает энергию (т. е. фонон излучается). В этом случае процесс рассматривается как «столкновение» между «частицей» решеточных колебаний, получившей название «фонон», и электроном. Имея эту информацию, можно рассчитать новый волновой вектор следующим образом:

1. Вычисляем старую энергию через старый волновой вектор

$$\mathcal{E} = \frac{\hbar^2 k^2}{2m^*}. \quad (10.19)$$

2. Вычисляем вероятность рассеяния и выбираем процесс рассеяния (см. разд. 10.3.2)

$$\lambda_i(\mathcal{E}). \quad (10.20)$$

3. Вычисляем новую энергию для выбранного процесса

$$\mathcal{E}' = \mathcal{E} \pm \hbar\omega_i. \quad (10.21)$$

4. Вычисляем модуль нового волнового вектора, используя (10.16а)

$$k' = [2m^*\mathcal{E}'(1 + \alpha\mathcal{E}')^{1/2}/\hbar]. \quad (10.22)$$

5. Выбираем направление нового волнового вектора, исходя из зависимости $p_i(\theta)$.

Было бы неуместным приводить формулы для λ_i и p_i для всех рассеивающих механизмов. Формулы, использованные в численном моделировании, которое будет описано здесь, взяты из [133].

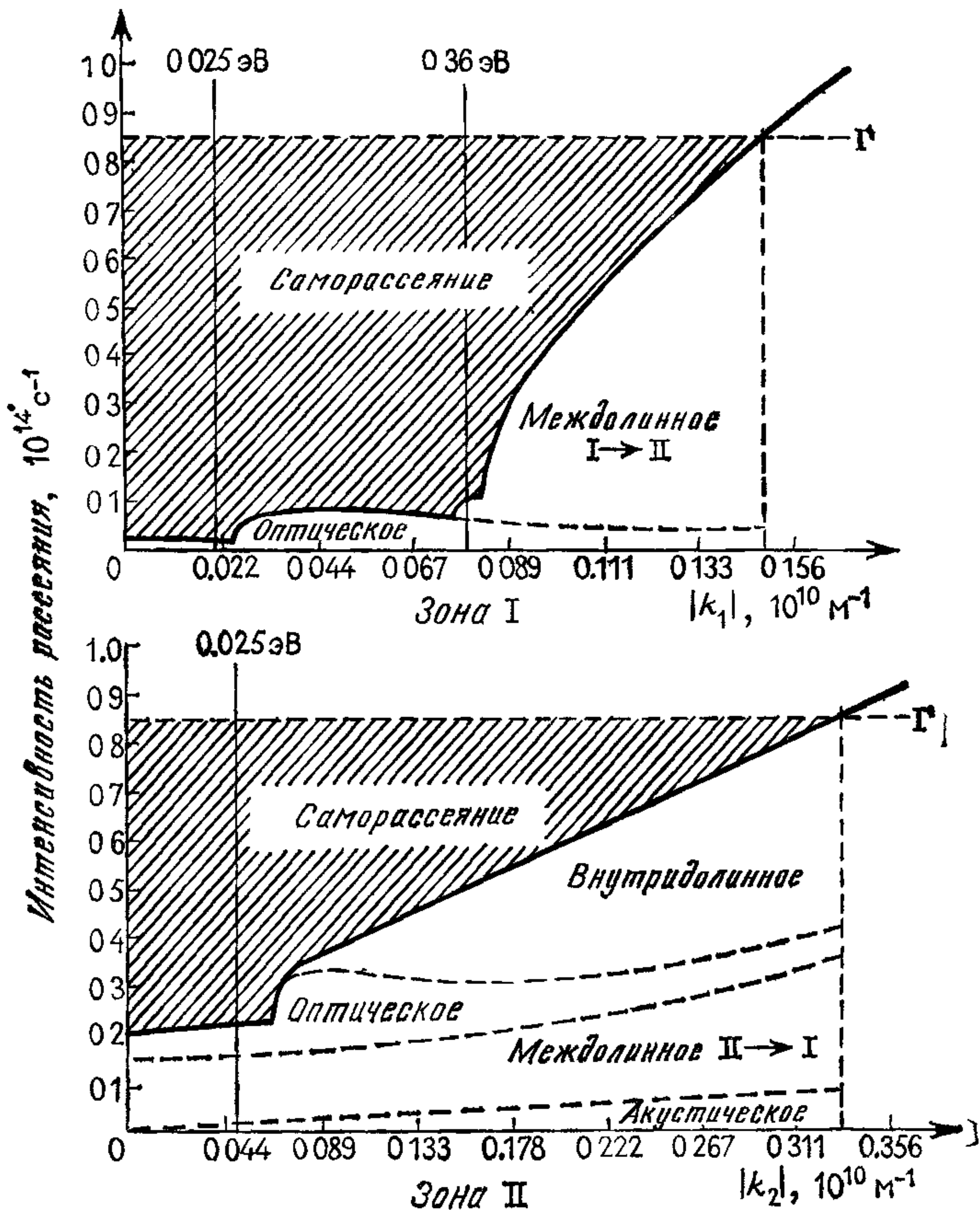


Рис 106 Полная интенсивность рассеяния (сплошная кривая) для электронов в долине I и долине II как функция волнового вектора. Акустическое рассеяние незначительно в долине I. Волновые векторы k_1 и k_2 отсчитываются от центра каждой долины, как на рис 105. Показана энергия электрона при комнатной температуре (0.025 эВ) и разности энергии между долинами (0.36 эВ) (Любезно предоставлено Уорринером, [381])

На рис 106 даны величины различных вероятностей рассеяния и полная вероятность рассеяния для электронов в зоне I и зоне II. В зоне I акустическое рассеяние несущественно и при низких энергиях преобладающим процессом является полярное оптическое рассеяние. При энергиях, превышающих приблизительно 0.36 эВ, или $k = 0.08 \times 10^{10} \text{ м}^{-1}$, междолинное рассеяние электронов в зону II доминирует в рассеянии. Для электронов в зоне II всегда имеется вероятность обратного рассеяния в зону I. Акустическое рассеяние должно учитываться во всем энергетическом интервале, так как, хотя его вероятность относительно мала, рассеяние может производить большие изменения в век-

торе k . При $k \simeq 0.05 \times 10^{10} \text{ м}^{-1}$ становится важным полярное рассеяние. При более высоких значениях k начинает доминировать внутривалинное рассеяние между шестью составными минимумами зоны II.

10.2.4. Подвижность

Уравнение (10.17а) показывает, что при наличии постоянного электрического поля и в отсутствие какого-либо рассеяния вектор k (и скорость) электрона должен увеличиваться линейно со временем. Эффект рассеяния должен ограничивать дрейфовую скорость, которую получают усреднением по электронам:

$$\bar{v} = -\mu(E) E, \quad (10.23a)$$

где μ — подвижность, которая связана с интенсивностью столкновений формулой

$$\mu = \frac{1}{2} \frac{e}{m^*} \tau = \frac{1}{2} \frac{e}{m^* \sigma}, \quad (10.23b)$$

где τ — среднее время рассеяния и σ — полная вероятность рассеяния. Вышеприведенная формула является приближенной, но годится для демонстрации влияния вероятности столкновений и эффективной массы на подвижность. Как ожидалось, чем больше масса и вероятность рассеяния, тем меньше подвижность. Формула получена в предположении, что скорость электрона понижается в среднем до нуля при каждом столкновении. Полученная средняя скорость поэтому равна половине произведения ускорения в электрическом поле и времени между столкновениями (или обратной величине вероятности столкновений).

На рис. 10.7 дана зависимость стационарной дрейфовой скорости электронов от приложенного электрического поля. Для слабых приложенных полей, меньших пороговой величины, равной приблизительно $4 \times 10^5 \text{ В} \cdot \text{м}^{-1}$, дрейфовая скорость пропорциональна полю и GaAs представляет собой обычное сопротивление с приблизительно постоянной подвижностью порядка $0.6 \text{ м}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$. Реальные образцы GaAs, легированные до уровня 10^{23} м^{-3} , имеют подвижность в слабом поле в интервале от 0.4 до $0.5 \text{ м}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$ и пороговые поля от 3×10^5 до $4 \times 10^5 \text{ В} \cdot \text{м}^{-1}$. Это противоречие объясняется неучетом рассеяния на ионизированной примеси в описываемой сейчас вычислительной модели. Физически уменьшение подвижности возникает из-за рассеяния электронов проводимости на фиксированных ионизированных атомах серы в решетке. При плотностях легирования около 10^{21} м^{-3} примесное рассеяние становится важным механизмом. Рассеяние на ионизированной примеси было учтено в последующей работе, и

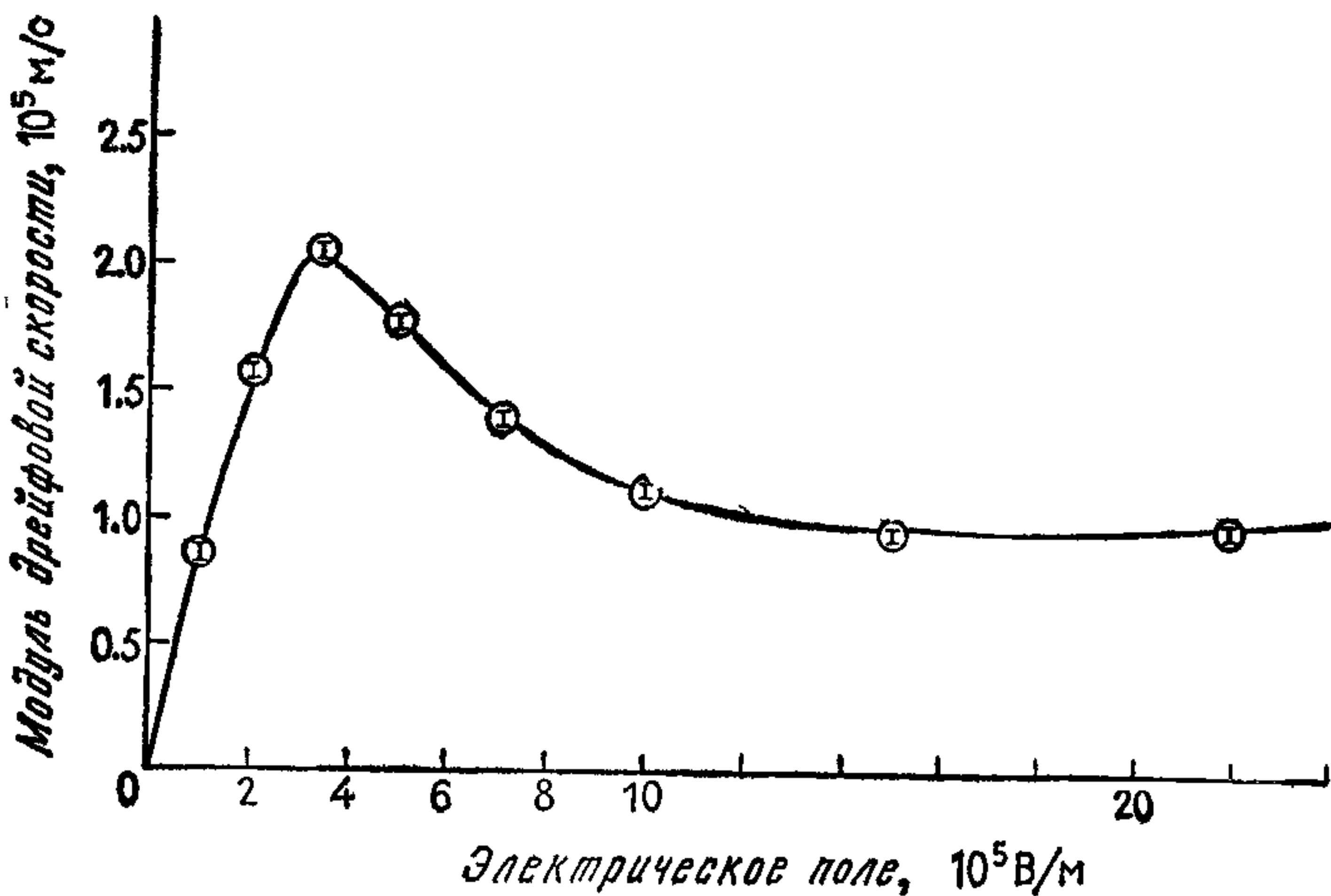


Рис. 10.7. Зависимость установившейся дрейфовой скорости от приложенного электрического поля. Сплошная кривая соответствует результатам [133], кружки [382]. ([382], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers.)

тогда результирующие подвижности и сопротивления находятся в хорошем согласии с экспериментом (см. табл. 10.2).

Для приложенных полей порядка $4 \times 10^5 \text{ В} \cdot \text{м}^{-1}$ и более сильных все большее число электронов достигает энергии $\sim 0.36 \text{ эВ}$ — энергетической щели между сателлитной и центральной долинами. На рис. 10.6 показано, что при этих энергиях вероятность электронного рассеяния в долину II становится высокой и, когда это рассеяние происходит, эффективная масса электронов возрастает в пять раз. В то же самое время большая часть их энергии превращается в потенциальную энергию долины II и скорость электрона уменьшается до малого значения. Общий результат состоит в том, что дрейфовая скорость падает с увеличением поля. В области полей между 4×10^5 и $12 \times 10^5 \text{ В} \cdot \text{м}^{-1}$, хотя полная подвижность $\mu = \bar{v}/E$ все еще положительна, дифференциальная подвижность отрицательна. Это означает, что материал ведет себя как отрицательное сопротивление для любых возмущений состояния стационарного потока. Уменьшение электрического поля вызывает увеличение, а не уменьшение скорости и тока. Ясно, что такое состояние неустойчиво и вызывает возникновение обогащенных и обедненных зарядом областей.

Это необычное свойство может быть использовано в определенных электронных приборах. Ганн [157] обнаружил, что обычный диод из GaAs будет многократно генерировать дипольные области заряда, названные доменами Ганна, которые перемещаются от катода к аноду. Таким образом, эти приборы преоб-

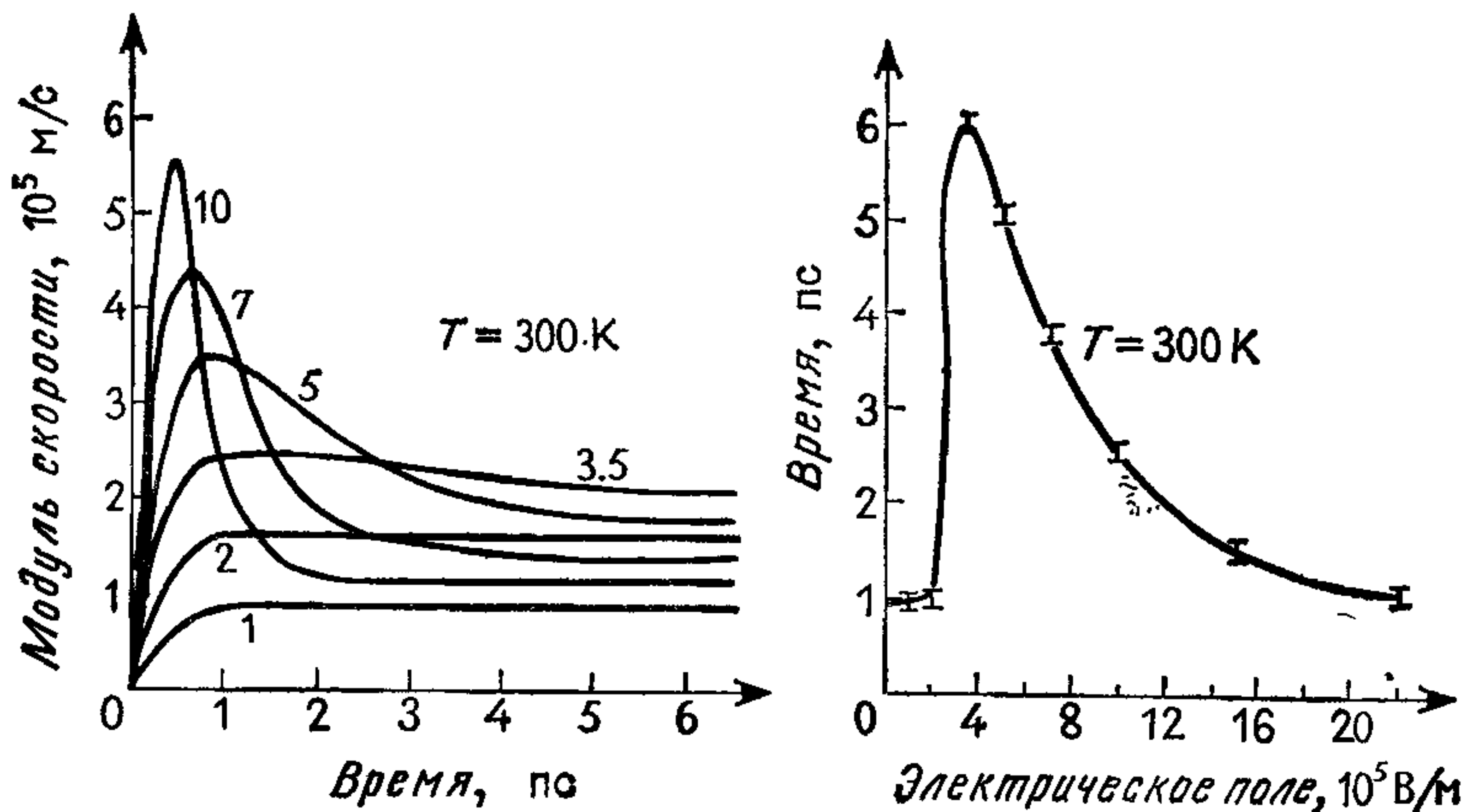


Рис. 10.8. Зависимость дрейфовой скорости электронов от времени как реакция на скачкообразное изменение электрического поля (слева). Исходное поле равно нулю, и кривые помечены изменением поля в единицах 10^5 В/м. Время достижения конечного установившегося значения (точность 5 %) как функция изменения поля. ([382], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers.)

разуют напряжение постоянного смещения в микроволновой ток, частота которого зависит от времени прохождения доменов Ганна через диод. Свойства таких доменов изучены Уорринером [383] с помощью численной модели, описанной в этой главе. Превосходный обзор работ по эффекту Ганна в арсениде галлия выполнен Боттом и Фаусеттом [45].

В полях, превышающих 20×10^5 В·м⁻¹, большинство электронов находится в долине II, и материал ведет себя как обычное положительное сопротивление. Вследствие пятикратного увеличения массы и приблизительно четырехкратного увеличения интенсивности рассеяния в долине II соотношения (10.23) показывают, что подвижность в долине II составляет только одну двадцатую часть от подвижности в слабом поле в долине I.

10.2.5. Релаксационные эффекты в переходных процессах

Важно понимать, что кривая дрейфовой скорости, описанная выше в терминах механизмов рассеяния, справедлива только в стационарном состоянии после того, как все переходные процессы, обусловленные конечными временами рассеяния, закончились. Эти эффекты подробно изучены Уорринером [382] и проиллюстрированы на рис. 10.8. Кривая слева показывает временную эволюцию средней дрейфовой скорости, когда электрическое поле увеличено от нуля до установившегося значения. Установившиеся значения на правом краю графика те же, что и при-

веденные на рис. 10.7. Кривые показывают большие скорости переходного процесса в сильных полях порядка 10^6 В·м⁻¹ и медленный отклик вблизи порога. На рис. 10.8 (справа) представлено время релаксации как функция поля. Оно определяется как время установления дрейфовой скорости в пределах 5 %. Наибольшие времена релаксации продолжительностью 5 пс (5×10^{-12} с) имеют место при пороговых значениях полей от 3×10^5 до 4×10^5 В·м⁻¹. Временные запаздывания появляются из-за конечности времени ускорения электронов и из-за того, что требуется несколько времен рассеяния, до того как установится состояние стационарного течения.

Существование переходных эффектов имеет несколько следствий. Наиболее важное из них касается размеров и рабочей частоты прибора. Типичный ПТ на GaAs (приблизительно 1978 г.) работает при напряжении примерно 3 В между затвором и стоком при расстоянии между ними в 1 мкм. Это соответствует среднему электрическому полю, равному 30×10^5 В·м⁻¹. В этих полях средняя скорость порядка 10^6 м·с⁻¹ поддерживается около 1 пс перед тем, как установится стационарное состояние. В течение этого времени типичный электрон пройдет расстояние около 1 мкм. Другими словами, расстояние, требуемое для установления стационарного течения, равно зазору между затвором и стоком. Следовательно, отсутствует область, где можно использовать стационарную зависимость \bar{v} от E , приведенную на рис. 10.7.

Чтобы правильно представить движение электронов через зазор между затвором и стоком, существенно корректно учесть запаздывания рассеяния. Максимальная рабочая частота может быть оценена исходя из 5 пс на каждую четверть периода осцилляции. При этой частоте переходные процессы закончены перед следующим изменением направления поля. Эта оценка дает 20 пс для периода осцилляции или частоту в 50 ГГц. Поскольку рабочие частоты 30 ГГц достигнуты для ПТ на GaAs, запаздывания рассеяния должны быть включены в моделирование таких приборов. Из предшествующего рассмотрения ясно, что учет запаздывания рассеяния является важным даже для микронных приборов. В связи с усовершенствованиями в технологии будут сделаны низко- и высокочастотные приборы, для которых сделанное выше утверждение будет даже более обоснованным.

В этой главе описывается метод моделирования, учитывающий отдельные рассеяния и возникающие из-за этого запаздывания. Такая микроскопическая модель рассеяния может корректно описать движение электронов в GaAs. Мы сопоставляем эту модель с традиционной диффузионной моделью для электронного переноса в транзисторах, которая использована многими авторами, включая Рейсера [328]. В последней модели электроны дрейфуют и диффундируют через решетку с подвиж-

ностью и коэффициентом диффузии, описанными кривой на рис. 10.7. Предполагается, что дрейфовая скорость должна мгновенно подстраиваться к изменению электрического поля, и, следовательно, все запаздывания рассеяния игнорируются. Такая модель, хотя и удовлетворительна для больших и медленных приборов, является несовершенной для моделирования малых ПТ на GaAs, рассматриваемых в этой главе.

10.3. РАЗРАБОТКА ВЫЧИСЛИТЕЛЬНОЙ МОДЕЛИ

Сделав обзор физических свойств электронов в полевом транзисторе на GaAs, можно построить вычислительную модель. Модель, которая будет описана, впервые опубликована Хокни, Варринером и Рейсером [193] и впоследствии улучшена и обобщена Варринером [381].

Если бы не было рассеяния электронов на решетке, электроны должны были бы вести себя подобно бесстолкновительной плазме, для которой подходят методы моделирования частица — сетка, описанные в гл. 9. Помимо уже обсужденных граничных условий (разд. 10.1.2), включение рассеяния на решетке является тем, что еще необходимо рассмотреть для распространения плазменной модели на задачу о ПТ. Процессы рассеяния будут выбраны на статистической основе с помощью случайных чисел, и по этой причине процедуру часто называют выборкой по Монте-Карло. Рассмотрим сейчас отдельно этапы моделирования частица — сетка и выборку по Монте-Карло и в заключение, как они объединяются вместе в модифицированном цикле шага по времени.

10.3.1. Расчеты по методу частица — сетка

В разд. 9.2.3 мы изучили влияние временного шага и пространственной ячейки на свойства двумерной модели частица — сетка в плазме. В частности, мы заметили, что временной шаг и пространственная ячейка сетки должны быть связаны с физической плазменной частотой и дебаевской длиной плазмы. Отсутствие между ними такой связи могло бы привести к неустойчивой и сильностолкновительной плазме. При моделировании полевого транзистора мы хотим включить физические эффекты столкновений в модель частица — сетка, которая для практических целей является бесстолкновительной. Поэтому важно выбрать параметры, которые гарантируют, что интенсивность столкновений, присущая модели частица — сетка, меньше, чем физические механизмы, которые мы хотим включить.

Рассмотрим сначала вопрос об устойчивости и требование, что $\omega_p DT \leq 2$, где ω_p — плазменная частота. В наихудшем слу-

чае легких электронов долины I имеем

$$\omega_{p1} = \left(\frac{ne^2}{\epsilon\epsilon_0 m} \right)^{1/2} = 2.12 \times 10^{13} \text{ с}^{-1}, \quad (10.24a)$$

где

$$\begin{aligned} n &= 10^{23} \text{ м}^{-3}, \\ \epsilon &= 10 \text{ (диэлектрическая константа)}, \\ m &= 0.067m_e. \end{aligned} \quad (10.24b)$$

Временной шаг, выбранный для расчета частица — сетка, будет называться «настроенным по полю» временным шагом, чтобы отличить его от времен свободного пробега, использованных далее при выборе процессов рассеяния. Выбранное для моделирования значение было равно

$$DT = 5 \times 10^{-14} \text{ с}, \quad (10.25a)$$

что приводит к

$$\omega_{p1} DT = 1.06, \quad (10.25b)$$

и, следовательно, плазменный период равен

$$\tau_{p1} = \frac{2\pi}{\omega_{p1}} = 6DT = 30 \times 10^{-14} \text{ с}. \quad (10.25c)$$

Аналогично, для тяжелых электронов ($m = 0,35m_e$) в долине II имеем

$$\omega_{p2} DT = 0.464, \quad (10.26a)$$

$$\tau_{p2} = 13.5 DT. \quad (10.26b)$$

Поэтому выбор DT удовлетворяет критерию устойчивости для обоих сортов электронов.

Чтобы рассчитать интенсивность столкновений в модели, мы должны вычислить дебаевскую длину λ_D и число модельных частиц в квадрате со стороной, равной дебаевской длине N_{DM} . Дебаевская длина не зависит от массы и поэтому одинакова для обоих сортов электронов

$$\lambda_D = \left(\frac{\epsilon\epsilon_0 k_B T}{ne^2} \right)^{1/2} = \frac{v_T}{\omega_p} = 1.2 \times 10^{-8} \text{ м}, \quad (10.27a)$$

где тепловая скорость определяется как

$$\begin{aligned} v_T &= \left(\frac{k_B T}{m} \right)^{1/2} = 2.6 \times 10^5 \text{ м} \cdot \text{с}^{-1} \text{ (долина I)}, \\ &= 1.1 \times 10^5 \text{ м} \cdot \text{с}^{-1} \text{ (долина II)}. \end{aligned} \quad (10.27b)$$

Эпитаксиальный слой модели (см. рис. 10.2) представляет область $5.12 \times 0.16 \text{ (мкм)}^2 = 430 \times 13.3\lambda_D^2$. Она первоначально за-

полнена 8000 модельных электронов и, следовательно,

$$N_{DM} = \frac{8000}{430 \times 13.3} = 1.4. \quad (10.27c)$$

Характерное время столкновения для модели τ_{CM} определяется согласно (9.19a) и равно

$$\frac{\tau_{CM}}{\tau_{pl}} = N_{DM} \left\{ 1 + \left(\frac{W}{\lambda_D} \right)^2 \right\}, \quad (10.28)$$

где W — ширина частицы. Модель ПТ использует раздачу параметров частицы до ближайшего узла сетки, для которой $W = 2H = 4 \times 10^{-8}$ м = 0.04 мкм — двойной пространственный шаг ячейки. Подставляя эти значения, получим

$$\frac{W}{\lambda_D} = 3.34, \quad (10.29a)$$

$$\tau_{CM} = 17.0 \tau_{pl} = 510 \times 10^{-14} \text{ с}, \quad (10.29b)$$

или выраженное как интенсивность столкновений

$$\nu_{CM} = \tau_{CM}^{-1} = 0.0020 \times 10^{14} \text{ с}^{-1}. \quad (10.29c)$$

Наименьшая полная интенсивность физического рассеяния соответствует низкоэнергетическим электронам долины I и равна $0.02 \times 10^{14} \text{ с}^{-1}$ (см. рис. 10.6). Поэтому интенсивность искусственного модельного рассеяния равна самое большее одной десятой доли от физических процессов. Она появляется в области источник — затвор прибора, которая считается относительно мало-важной для характеристики прибора. В имеющей большое значение области затвор — сток интенсивность физического расстояния в десять раз выше и интенсивность модельного рассеяния пропорционально является менее важной.

Для описания потенциала в расчетной области 5.12×0.32 (мкм)² на рис. 10.2 (справа) использована сетка (256×16) квадратных ячеек, имеющих размер $H = 0.02$ мкм. Расчеты проведены по программе POT4 при помощи алгоритма FCSR(1) (см. разд. 6.5.3). На электродах поддерживались заданные потенциалы с помощью метода матрицы емкости (см. разд. 6.5.6). Электростатический потенциал определяется уравнением Пуассона, которое в диэлектрической среде имеет вид

$$\nabla^2 \varphi = - \frac{\rho}{\epsilon \epsilon_0}, \quad (10.30a)$$

где ϵ — диэлектрическая постоянная. Эффект диэлектрической константы можно учесть наиболее просто путем подходящего уменьшения заряда частицы при моделировании. Для расчета поля моделируемая частица представляет собой заряженный

стержень с эффективным зарядом на единицу длины

$$q = \frac{-enL_xL_y}{eN}, \quad (10.30b)$$

где n — плотность легирования (10^{23} м^{-3}); L_x, L_y — размеры области легирования (5.12×0.16) (мкм)²; ϵ — диэлектрическая константа (~ 10); N — число моделируемых частиц, использованных в модели (~ 8000).

При высоких плотностях легирования может оказаться необходимым сглаживание рассчитанного из (10.30a) потенциала, перед тем как продвинуть частицы. Удобным методом является расчет сглаженного потенциала φ^* как среднего по окружающим восьми значениям:

$$\varphi_{i,j}^* = \frac{1}{8} (\varphi_{i-1,j-1} + \varphi_{i-1,j} + \varphi_{i-1,j+1} + \varphi_{i,j+1} + \varphi_{i,j-1} + \varphi_{i+1,j-1} + \varphi_{i+1,j} + \varphi_{i+1,j+1}). \quad (10.30c)$$

Эта модель известна как NGP (HNGP)-модель.

Мы уже упоминали, что взвешивание заряда и интерполяция силы по ближайшему узлу сетки (HNGP) используются в модели ПТ без затруднений. Преимущество этой методики заключается в простоте и малых вычислительных затратах. Ясно, что могли бы использоваться (CIC) или «облако с треугольной формой» (TSC). Все эти схемы имеют близкие времена столкновения (W изменяется от H до $3H$), но радикально отличаются в свойствах сохранения энергии. Поэтому выбор модели в первую очередь будет зависеть от требуемого уровня сохранения энергии. Соответствующие формулы уже даны в связи с бесстолкновительным моделированием плазмы (9.22). Потери в сохранении энергии измеряются временем нагрева τ_H , которое задается выражением

$$\frac{\tau_H}{\tau_{CM}} = \frac{K_H}{(H/\lambda_D)^2}, \quad (10.31a)$$

где $K_H=2$ (NGP), 6 (HNGP), 40 (CIC), 150 (TSC), 3000 (QPM).

Для модели HNGP имеем

$$\left(\frac{H}{\lambda_D}\right) = 1.67 \quad \text{и} \quad \frac{\tau_H}{\tau_{CM}} = 2.15, \quad (10.31b)$$

следовательно,

$$\tau_H = 1.097 \times 10^{-14} \text{ с} = 220 DT. \quad (10.31c)$$

При каждом физическом столкновении с решеткой температура электронного распределения частично релаксирует к температуре решетки, следовательно, важной величиной является относительная потеря в сохранении энергии за время столкновения. Если мы говорим, что время физического столкновения меняется

приблизительно от 5×10^{-14} с в зазоре между затвором и стоком до 50×10^{-14} с в области источника, данные выше времена нагрева соответствуют относительному сохранению энергии от 0.2 до 2 % ¹⁾. Если используется раздача заряда TSC, то относительное сохранение энергии может быть улучшено от 0.004 до 0.04 %, но эта методика значительно увеличит затраты компьютерного времени на расчет этапа свободного пробега частицы. Физический эффект относительно короткого времени нагрева эквивалентен использованию слегка повышенной температуры решетки. Это может быть, конечно, скомпенсировано введением пониженной температуры решетки при вычислении процессов рассеяния. Поскольку температура электрона определяется в основном заданными интенсивностями физического рассеяния, ясно, что можно проводить счет с временами нагрева, существенно меньшими тех значений, которые были бы приемлемы при моделировании бесстолкновительной плазмы с помощью чистого кода частица — сетка.

10.3.2. Выбор рассеяния с помощью процедуры Монте-Карло

Выбор времени свободного пробега между различными рассеяниями значительно упрощается, если полная интенсивность рассеяния для всех процессов является постоянной, скажем Γ с⁻¹. В этом случае вероятность нахождения свободного пробега в интервале между t и $t + dt$ равна

$$P(t) dt = \Gamma e^{-\Gamma t} dt. \quad (10.32)$$

Задача о выборе набора чисел с заданной неоднородной вероятностью $P(t)$, когда имеется в наличии только генератор случайных чисел, дающий однородное распределение, неоднократно встречается в моделировании по Монте-Карло. Сначала необходимо найти аналитически или численно решение для функции распределения

$$c(t) = \int_0^t P(t') dt'. \quad (10.33a)$$

Поскольку $P(t)$ есть вероятность, то $c(\infty) = 1$ и, следовательно $0 \leq c < 1$ и $dc = P(t) dt$.

$$0 \leq c < 1 \text{ и } dc = P(t) dt. \quad (10.33b)$$

Если значение c выбирается как случайное число r из распределения, однородного в диапазоне $0 \leq r < 1$, то вероятность выбора значения в интервале dc равна dc . Следовательно, соглас-

¹⁾ За время τ_H кинетическая энергия электронов увеличивается на 50 % от средней тепловой энергии электрона (см. разд. 9 2 2).

но (10.33), вероятность выбора значения в интервале dt равна $P(t)dt$, как требовалось. Поэтому положим

$$r = c(t), \quad (10.34a)$$

$$t = c^{-1}(r), \quad (10.34b)$$

где c^{-1} — функция, обратная $c(r)$, т.е. результат решения (10.34a) для различных t .

Применяя этот метод к (10.32), получим

$$r = c(t) = 1 - e^{-\Gamma t}, \quad (10.35a)$$

$$\delta t = -\Gamma^{-1} \ln(r_1), \quad (10.35b)$$

где введен символ δt вместо t для времени свободного пробега. Кроме того, для упрощения этой формулы определили новое случайное число $r_1 = 1 - r$, которое распределено однородно на интервале $0 < r_1 \leq 1$.

На рис. 10.6 показано, что полная интенсивность рассеяния для реальных процессов не является постоянной и может показаться, что предыдущую простую процедуру выборки нельзя использовать. Однако Рис [321, 322] предложил включить фиктивный процесс «саморассеяния», который не дает изменений k -вектора частицы. Интенсивность саморассеяния при любой энергии выбирается такой, чтобы привести реальные процессы к выбранной постоянной Γ ; она показана заштрихованной областью на рис. 10.6. Если теперь учесть процесс саморассеяния, то полная вероятность рассеяния будет постоянной, и свободные пробеги могут быть выбраны в соответствии с (10.35b).

После выбора времени свободного пробега электрон движется в электрическом поле, определяемом с помощью сетки в течение времени δt . Для сеточной ячейки, окружающей сеточную точку (i, j) , имеем

$$E_x = (\varphi_{i-1, j} - \varphi_{i+1, j})/2H, \quad (10.36a)$$

$$E_y = (\varphi_{i, j-1} - \varphi_{i, j+1})/2H. \quad (10.36b)$$

Новый k -вектор получается из соотношения

$$k_x^{(t+\delta t)} = k_x^{(t)} - eE_x \delta t / \hbar, \quad (10.37)$$

и аналогично для k_y . Тогда новое положение частицы получается из выражения

$$x^{(t+\delta t)} = x^{(t)} + \frac{\hbar}{2m} (k_x^{(t)} + k_x^{(t+\delta t)}) \delta t, \quad (10.38)$$

и аналогично для $y^{(t+\delta t)}$.

После этапа свободного пробега выбор процесса рассеяния проводится с помощью второго равномерно распределенного случайного числа r_2 в интервале

$$0 \leq r_2 < \Gamma. \quad (10.39)$$

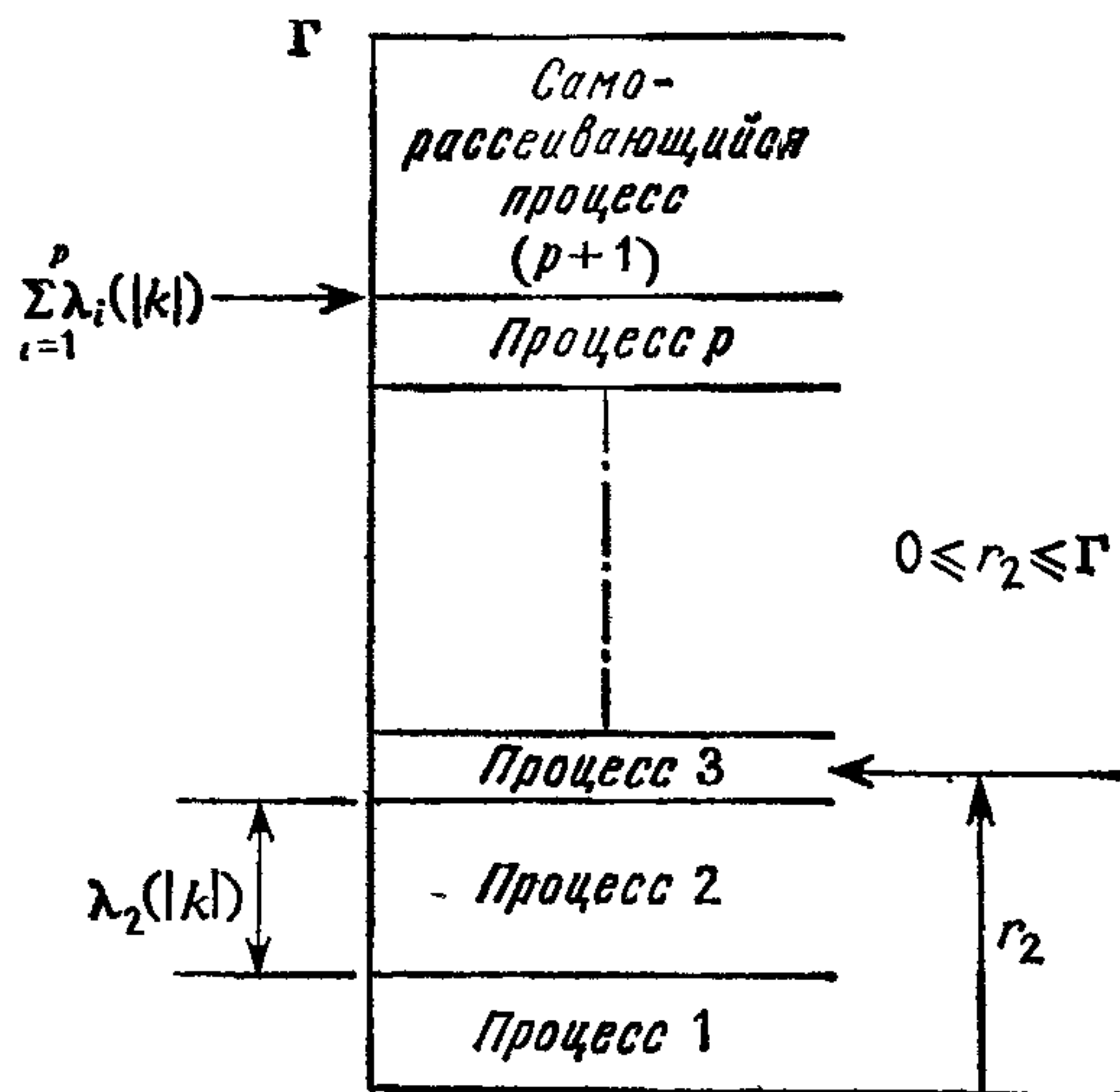


Рис 10.9 Диаграмма, показывающая выбор процесса рассеяния. Имеется p реальных процессов. Фиктивный саморассеивающийся процесс $(p+1)$ обеспечивает равенство полной интенсивности рассеяния постоянной Γ . Значение r_2 выбирает процесс номер 3. (Любезно предоставлено Уорринером, [381])

Процесс рассеяния m выбирается, если

$$\sum_{i=1}^{m-1} \lambda_i(k) \leq r_2 < \sum_{i=1}^m \lambda_i(k), \quad (10.40)$$

где $\lambda_i(k)$ — вероятность рассеяния для процесса i . Этот метод выбора иллюстрируется на рис. 10.9, где размер по вертикали, занятый каждым процессом, равен вероятности рассеяния на данном процессе и поэтому пропорционален вероятности появления этого процесса.

Представленный выше метод выбора может быть очень неэкономичным по затратам компьютерного времени, если вероятность рассеяния сильно меняется с энергией, как это происходит в некоторых материалах, таких, как GaAs. Вероятность выбора искусственного процесса саморассеяния пропорциональна заштрихованной области на рис. 10.6 и, очевидно, большая для большинства электронов, которые находятся в зоне I с низкими энергиями. Кроме того, формула для $\lambda(k)$ сложна, и компьютерное время было бы сэкономлено, если бы не нужен был пересчет в (10.40) для каждого испытания и каждого электрона.

Чтобы увеличить эффективность вычислительной программы, желательно табулировать суммы в (10.40) в виде двумерной таблицы, как это показано на рис. 10.10. Тогда случайное число r_2 по очереди сравнивается с табличными суммами:

$$\text{Если } r_2 \geq \sum_{i=1}^q \lambda_i, \text{ то выбирается процесс } (q+1). \quad (10.41a)$$

Процесс k	1	2	3	4	—	—	—	p
1	λ_1	$\sum_{i=1}^2 \lambda_i$	$\sum_{i=1}^3 \lambda_i$	$\sum_{i=1}^4 \lambda_i$	-----			$\sum_{i=1}^p \lambda_i$
2	λ_1	$\sum_{i=1}^2 \lambda_i$	$\sum_{i=1}^3 \lambda_i$					
3								
4								

Рис 10.10. Табулирование сумм, требуемых для выбора процесса рассеяния, для экономии машинного времени. Сравнение проводится справа налево. Если r_2 больше, чем табулированное значение для процесса q , то выбирается процесс $q + 1$. Если первое испытание успешно, то выбирается процесс саморассеяния. В этом примере r_2 выбирает процесс номер 4. (Любезно предоставлено Уорринером, [381].)

Поскольку саморассеяние, вероятно, является наиболее вероятным процессом для большинства электронов, сравнение должно быть сделано, начиная с саморассеяния в следующем порядке:

$$q = p, (p - 1), \dots, 1. \quad (10.41b)$$

Тогда, если первое испытание имеет успех, то выбирается саморассеяние, и если последнее испытание безуспешно, то выбирается процесс 1.

Процедура табулирования, хотя уменьшает повторение арифметических операций, не препятствует высокой вероятности непроизводительному выбору саморассеяния. При некоторых затратах памяти этому выбору можно препятствовать путем введения для каждой частицы индивидуального Γ_j , которое запоминается как дополнительная координата вместе с положением частицы и k -вектором. Единственное требование, налагаемое на Γ_j , состоит в том, чтобы оно было больше, чем полная интенсивность рассеяния для частицы j . Следовательно, низкоэнергетические частицы со слабыми интенсивностями рассеяния могут иметь значительно меньшее Γ_j , чем высокоэнергетические. Конечно, может случиться, что на некоторой стадии движения электрона полная интенсивность рассеяния превзойдет Γ_j . В этом случае Γ_j должно быть увеличено и используется следующая эвристика:

Полагаем $\Gamma_j = \sum_{i=1}^p \lambda_i(k^{(t)})$, суммируем по реальным процессам, выбираем r_1 и определяем свободный пробег с помощью Γ_j

$$\delta t = -\frac{1}{\Gamma_j} \ln(r_1)$$

Вычисляем новый k -вектор (10.42)

$$k^{(t+\delta t)} = k^{(t)} - \frac{e}{\hbar} E \delta t$$

Если $\sum_{i=1}^p \lambda_i(k^{(t+\delta t)}) \leq \Gamma_j$, то Γ_j удовлетворительно и переходим к следующей частице.

Если нет, то увеличиваем Γ_j до $\Gamma_j + \Delta\Gamma$ и делаем попытку опять с тем же самым r_1 .

Коэффициент экономии компьютерного времени при замене постоянной Γ без табулирования переменной Γ с табулированием равен приблизительно 10 на этапе движения частиц (табл. 10.1).

Таблица 10.1. Затраты машинного времени в секундах на IBM 360/195 для различных этапов вычислительного цикла.

Часть цикла	Постоянная Γ (без табулирования)	Переменная Γ с табулированием
QSHARE	0.07	0.07
POT4	0.35	0.35
SCATTER	9.80	1.20
Цикл	10.22	1.62

Используется сетка, имеющая (256×32) узлов с 12 000 частицами. Требуется размер области в 700 кбайт

Выбор нового k -вектора включает расчет его величины k' из (10.22) и вычисление угла рассеяния θ согласно вероятности $p_i(\theta)$ для выбранного процесса рассеяния. Зная $p_i(\theta)$, можно выбрать θ с помощью общего метода, данного соотношениями (10.33) и (10.34). Все процессы рассеяния, кроме полярного оптического рассеяния, таковы, что все направления равновероятны для окончательного k -вектора. Поэтому нет необходимости при определении окончательного направления учитывать направление k -вектора до рассеяния, как это обычно делается при определении угла рассеяния. В качестве базы можно взять

любое направление, и если выбирается окончательное направление в полярных координатах относительно оси z , то вероятность нахождения окончательного k -вектора $p(\theta)q(\varphi)d\theta d\varphi$ между θ и $\theta + d\theta$, φ и $d\varphi + \varphi$ задается относительной площадью поверхности единичной сферы с этими координатами. Другими словами,

$$p(\theta) d\theta = \sin \theta d\theta/2, \quad 0 \leq \theta \leq \pi, \quad (10.43a)$$

$$q(\varphi) d\varphi = d\varphi/2\pi, \quad 0 \leq \varphi \leq 2\pi. \quad (10.43b)$$

Вычисляя кумулятивное распределение, получаем

$$c_1(\theta) = \int_0^\theta p(\theta') d\theta' = (1 - \cos \theta)/2, \quad (10.44a)$$

$$c_2(\varphi) = \int_0^\varphi q(\varphi') d\varphi' = \varphi/2\pi, \quad (10.44b)$$

и, следовательно, θ и φ могут выбираться с помощью двух случайных чисел r_3, r_4 , равномерно распределенных на интервале $0 \leq r_3, r_4 \leq 1$:

$$\theta = c_1^{-1}(r_3) = \arccos(1 - 2r_3), \quad (10.45a)$$

$$\varphi = c_2^{-1}(r_4) = 2\pi r_4. \quad (10.45b)$$

Зная окончательное направление, компоненты вектора можно получить прямо из r_3 и второго из следующих равенств:

$$k'_z = k' \cos \theta = k' a, \quad (10.46a)$$

$$k'_x = k' \sin \theta \cos \varphi = k' b \cos \varphi, \quad (10.46b)$$

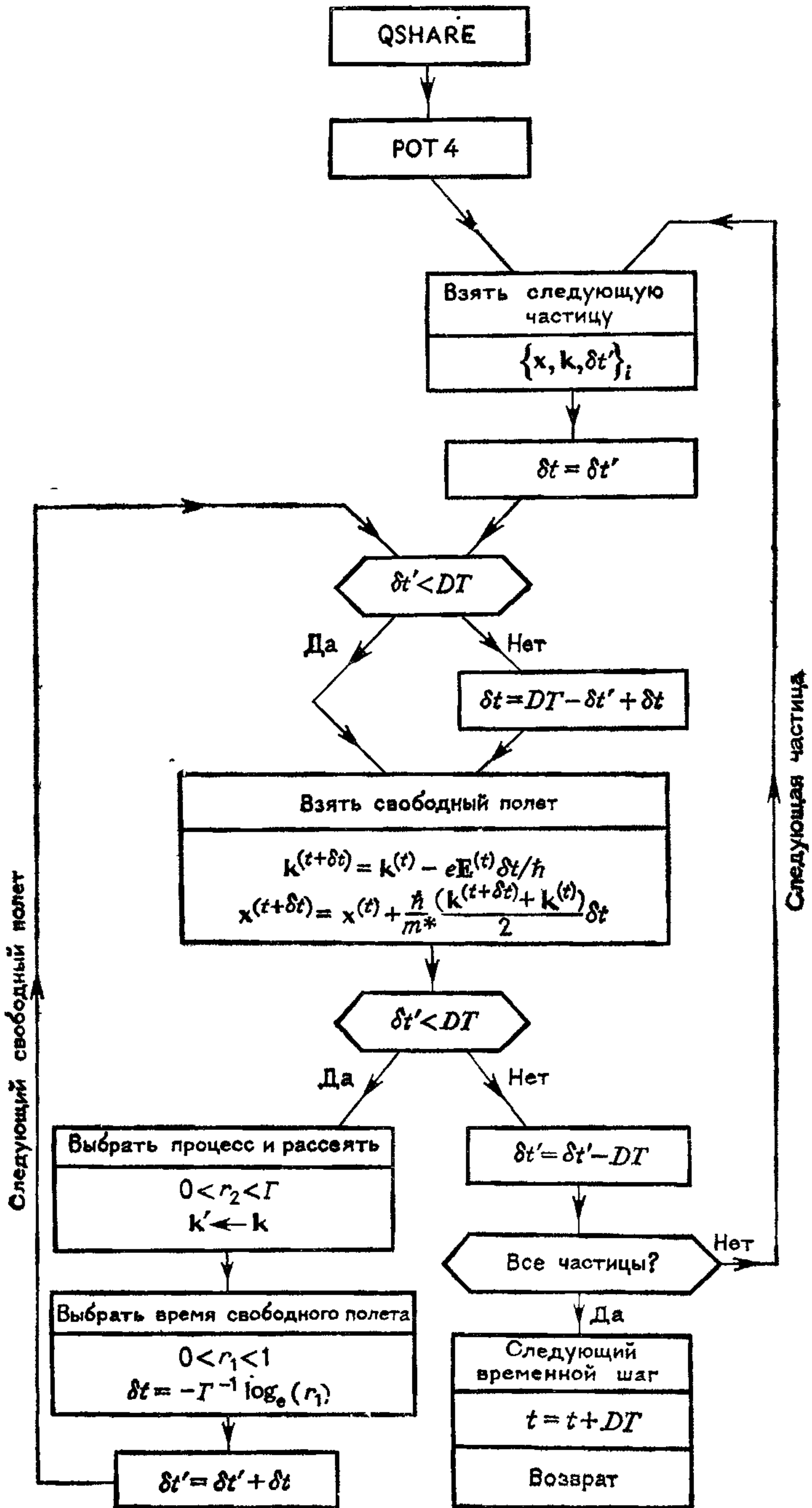
$$k'_y = k' \sin \theta \sin \varphi = k' b \sin \varphi, \quad (10.46c)$$

где $a = (1 - 2r_3)$ и $b = (1 - a^2)^{1/2}$.

Для процессов, которые не являются случайными, направление должно быть измерено относительно падающего k -вектора. Вероятность рассеяния не зависит от угла φ относительно этой оси, и (10.45b) все еще может использоваться для ее выбора. Метод выбора угла рассеяния θ в более сложных случаях, таких, как полярное оптическое рассеяние, в котором имеется преимущество для прямого рассеяния (θ мало), читатель может найти в книге [133, с. 197].

10.3.3. Модифицированный цикл временного шага

Сейчас мы можем рассмотреть модификации временного показано на рис. 10.11, который дает текущую диаграмму для цикла обычного кода частица — сетка, которые требуются для включения рассеяния, реализуемого методом Монте-Карло. Это показано на рис. 10.11, на котором представлена блок-схема



подпрограммы STEPON. Эта подпрограмма вызывается при постоянных интервалах DT (в секундах) управляющей программой CRONUS (см. рис. 3.4). Первое действие подпрограммы STEPON состоит в пересчете электрического поля, следовательно, DT называется временным шагом коррекции поля. Это делается обращением к подпрограммам QSHARE (определение заряда на сетке) и POT4 (нахождение потенциала) тем же самым способом, как в обычном алгоритме частица — сетка. В процессе этой части вычислений моделируемая частица представляет заряд в форме двумерного стержня, поскольку с высокой степенью точности можно считать, что вариация плотности заряда в направлении z отсутствует, и, следовательно, не нужно следить за координатой z .

После этой корректировки поля вводится внутренний цикл для расчета рассеяния методом Монте-Карло, в течение которого электрическое поле остается неизменным. В этом внутреннем цикле каждая частица просматривается по очереди и подвергается рассеянию по методу Монте-Карло до тех пор, пока сумма их времен свободного пробега не превысит время, при котором должна иметь место следующая корректировка поля. Величина этого превышения времени запоминается как координата $\delta t'$, так что свободный пробег может быть корректно продолжен с новым значением поля после следующего уточнения поля. В процессе внутреннего цикла каждая из моделируемых частиц рассеивается в трехмерном пространстве k , как если бы это был одиночный электрон с эффективной массой и зарядом, соответствующими его состоянию. Рассеяние изменяет только k -вектор, и, следовательно, мы запоминаем три величины (k_x , k_y , k_z). В процессе рассеяния 8000 моделируемых частиц рассматриваются в качестве выборки из распределения в реальном приборе, тогда как при расчете поля частицы рассматриваются как макрочастицы (см. (10.30b)), так что они правильно распределяют заряды на сетке. Важно тщательно различать различные роли, которые играют модельные частицы на различных этапах цикла.

Для высокоэнергетических частиц в зазоре затвор — сток с полной интенсивностью рассеяния $\sim 0.5 \times 10^{14} \text{ с}^{-1}$ будет в среднем $\sim 2^{1/2}$ событий рассеяния в течение шага корреляции поля $DT = 5 \times 10^{-14} \text{ с}^{-1}$. Эта ситуация иллюстрируется на рис. 10.12. Для низкоэнергетических частиц зоны I в области источника, где интенсивность рассеяния равна $\sim 0.02 \times 10^{14} \text{ с}^{-1}$, будет иметь

Рис. 10.11 Временной цикл модифицированного метода частица — сетка для программы моделирования ПТ с рассеянием по методу Монте-Карло (т. е. подпрограмма STEPON). Для каждой частицы запоминаются координаты (x , y , k_x , k_y , k_z , $\delta t'$, Γ), где $\delta t'$ — время свободного пробега, отложенное от последнего временного шага пересчета поля; $\delta t'$ также собирает сумму времен свободного пробега от начала текущего временного шага

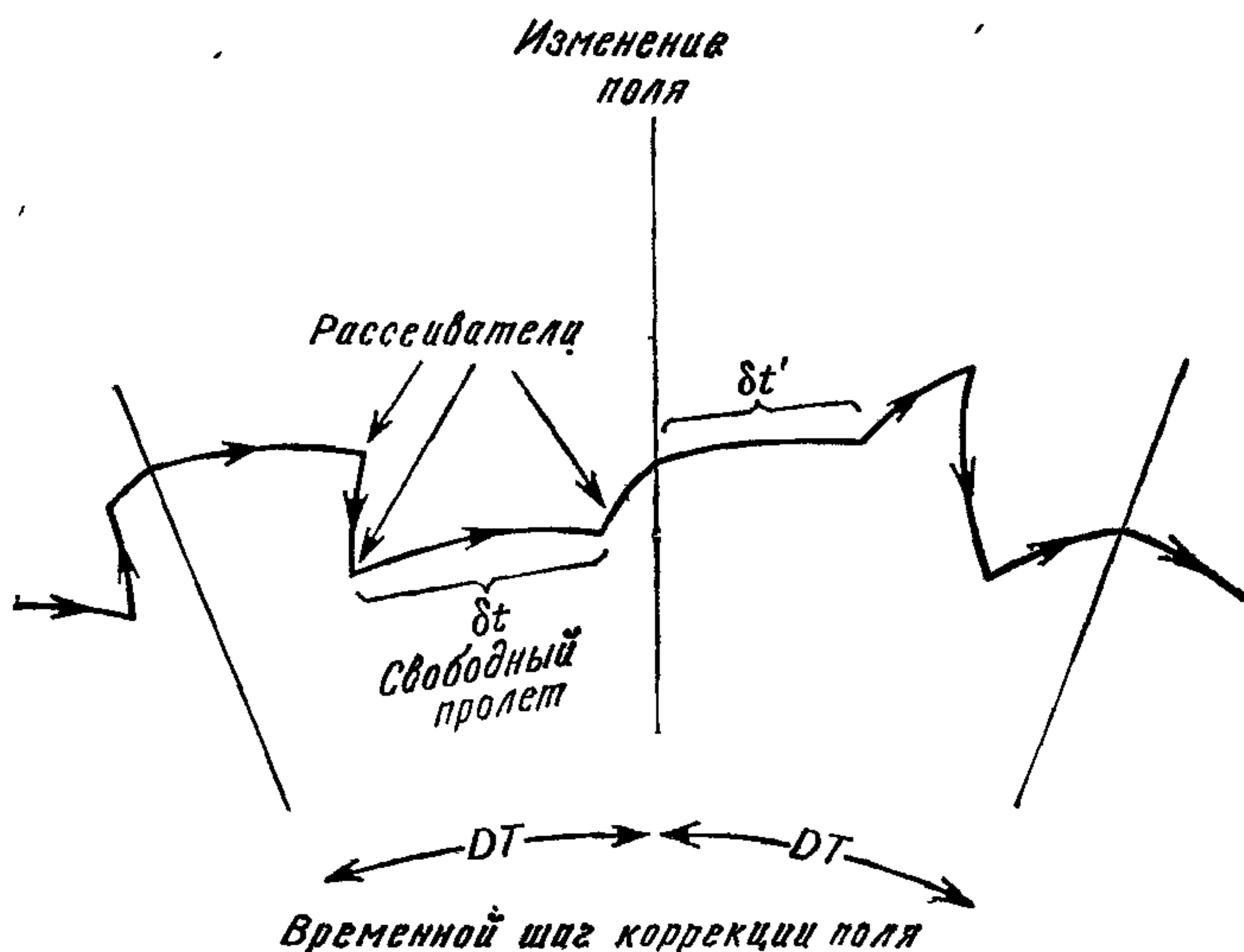


Рис. 10.12 Типичная траектория электрона для высокой скорости рассеяния, демонстрирующая несколько свободных пролетов за временной шаг корректировки поля. Рассеяния отмечены резким изменением траектории. Существует слабое изменение траектории за каждые DT секунд из-за пересчета электрического поля.

место одно событие рассеяния в среднем приблизительно на каждые десять временных шагов корректировки поля (или $1\frac{1}{2}$ плазменного периода). В этом последнем случае эффекты столкновения довольно слабы, и частица совершает плазменные колебания аналогично бесстолкновительной плазме.

На табл. 10.1 представлены вычислительные характеристики типичной программы моделирования ПТ методом рассеяния по Монте-Карло, используемой в Ридингском университете. Важность запоминания переменного Γ для каждой частицы и табулирования интенсивностей рассеяния является очевидной.

10.4. МОДЕЛИРОВАНИЕ ПОЛЕВЫХ ТРАНЗИСТОРОВ (ПТ)

Создав программы моделирования ПТ, необходимо измерять статические и динамические характеристики моделируемого прибора. Потом они могут сопоставляться с типичными лабораторными измерениями для того, чтобы обосновать вычислительную программу. Только после успешной проверки модель может использоваться для изучения геометрических форм и диапазона параметров ПТ, которые недоступны в лаборатории. Далее рассмотрим, как проводятся такие измерения, и способы заглянуть внутрь работающего прибора, то невозможно в лаборатории. Будет дан также метод получения параметров эквивалент-

ной цепи и картины шума. Наконец, модель будет применена для изучения потенциально очень быстродействующего прибора, работающего при температурах жидкого азота, который еще не создан, — охлаждаемый полевой транзистор (охлаждаемый ПТ). В этом случае численное моделирование является ведущим методом в разработке приборов, выявляя такие типы устройств, которые заслуживают построения в лаборатории.

10.4.1. Статические характеристики

Статические и динамические характеристики моделируемого прибора измеряются посредством приложения фиксированного напряжения к электродам и наблюдения возникающих при этом токов. Полный ток на электроде состоит из тока частиц и тока смещения. Они получаются из первого и второго членов соответственно в выражении (10.47). С момента старта программы моделирования хранится запись суммарного числа моделируемых электронов, поглощенных каждым электродом, и на каждом временном шаге корректировки поля вычисляется интеграл от нормальной компоненты электрического поля вдоль электрода. Затем рассчитывается эквивалентный полный заряд для каждого электрода из выражения

$$Q(t) = q(N_a - N_i) + \epsilon\epsilon_0 \int E_y(x, t) dx, \quad (10.47)$$

где интеграл берется по поверхности электрода и q — заряд, приходящийся на 1 м макрочастицы согласно (10.30b); N_a — полное число частиц, поглощенных к моменту времени t ; N_i — полное число частиц, инжектированных к моменту времени t ; $E_y(x, t)$ — электрическое поле, направленное в прибор в момент времени t в точке x .

Заметим, что в процессе этой фазы расчета, когда мы рассматриваем полный поток заряда без рассеяния, частицы должны рассматриваться как двумерные заряженные стержни или макрочастицы (см. разд. 10.3.1).

Начальным условием для большинства экспериментов будет состояние с нейтральзованным зарядом, показанное на рис. 10.13. Частицы распределены однородно в легированной эпитаксиальной области для нейтрализации примесного заряда. Их координаты (x, y) получаются с помощью датчика случайных чисел с однородным распределением. Условие нейтрализации заряда выбирается для того, чтобы избежать очень сильных полей и нестационарностей, которые могут появиться, если разделение зарядов не может происходить естественным образом из динамики частиц.

Первоначально все электроны располагаются в долине Γ и задаются k -векторами, соответствующими максвелловскому распределению скоростей при температуре решетки T (300 К). Рас-

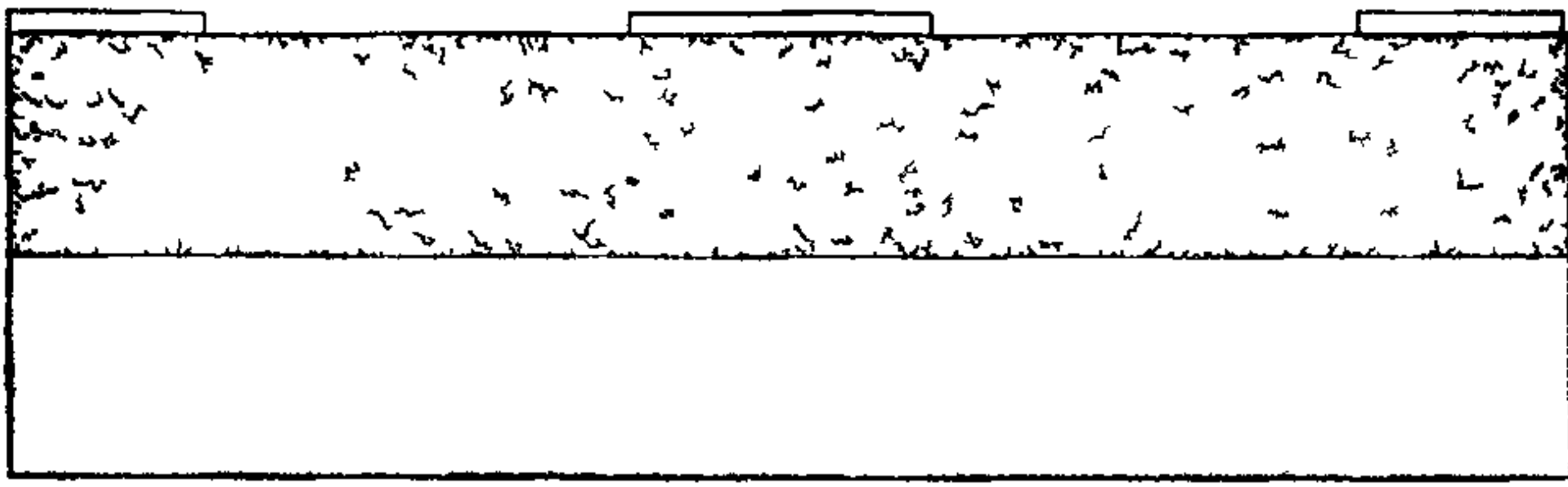


Рис 10 13 Начальное условие нейтральности заряда Показанные моделируемые электроны располагаются в долине I при 300 К Напряжение на всех электродах равно нулю (Любезно предоставлено С Дж Бердом)

пределение скоростей одно и то же по всем трем координатным направлениям Выбирая направление x , получаем, что вероятность нахождения скорости в интервале $v_x, v_x + dv_x$ равна

$$p(v_x) dv_x = (\sqrt{2\pi} v_T)^{-1} \exp(-v_x^2/2v_T^2), \quad (10.48a)$$

где $v_T = (k_B T/m_1^*)^{1/2}$ — среднеквадратичная скорость вдоль направления x Сейчас мы опишем два метода для определения этого распределения Сначала можно следовать общему методу, описанному в разд. 10 3 2, находим кумулятивное распределение

$$c(v_x) = 1/2 \operatorname{erf}(v_x/(2v_T^2)^{1/2}), \quad (10.48b)$$

где erf — интеграл ошибок, определяемый следующим образом:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx.$$

Поскольку $c(v_x)$ находится в интервале $-1/2 \leq c(v_x) < 1/2$, положим $c(v_x) = (r - 1/2)$, где r — случайное число из нормального интервала $0 \leq r < 1$. Тогда для требуемой скорости v_x получаем

$$v_x = \sqrt{2} v_T \operatorname{erf}^{-1}(2r - 1). \quad (10.48c)$$

Поэтому этот метод нуждается в подпрограмме для оценки обратной функции ошибок erf^{-1} или заранее вычисленной таблицы для этой функции После этого величина волнового числа получается из соотношения

$$k_x = \frac{m_1^*}{\hbar} v_x = \frac{\sqrt{2} v_T m_1^*}{\hbar} \operatorname{erf}^{-1}(2r - 1). \quad (10.49)$$

Для выбора каждой компоненты \mathbf{k} -вектора из (10 49) на каждую частицу требуется три случайных числа.

Другой метод выбора начальных скоростей использует большие возможности центральной предельной теоремы статистики (см., например, [337, с 252] или множество книг по статистике).

Эта теорема дает фундаментальные свойства случайного блуждания и всех явлений, основанных на случайном блуждании. Она утверждает:

«Если X_1, X_2, \dots , образуют последовательность независимых и идентично распределенных случайных величин, каждая из которых имеет среднее значение μ и стандартное отклонение σ , то сумма n таких случайных чисел, $\sum_{i=1}^n X_i$, приводит к нормальному (гауссовскому или максвелловскому) распределению со средним значением μ и стандартным отклонением $\sigma \sqrt{n}$, когда n становится большим».

Приведем примеры явлений, управляемых этой теоремой:

1. Случайное блуждание в физическом пространстве (диффузия)

2 Случайное блуждание в пространстве скоростей (стохастический нагрев)

3 Повторная свертка функции с собой приводит к максвелловскому распределению

Последовательность случайных чисел r_i на интервале от 0 до 1 характеризуется значениями $\mu = 1/2$ и $\sigma = (12)^{-1/2}$, следовательно, максвелловское распределение со средним значением 0 и средней квадратичной скоростью v_T получается из соотношения

$$v_x = v_T \left(\sum_{i=1}^n r_i - n\mu \right) / \left(\frac{n}{12} \right)^{1/2} \quad (10.50a)$$

(см. [10, с 952]) или для случая $n = 3$ из соотношения

$$v_x = 2v_T (r_1 + r_2 + r_3 - 1.5). \quad (10.50b)$$

При выборе n , не превышающем 3, мы умышленно минимизируем требуемое количество случайных чисел и объем машинного времени. Этим способом мы получаем приближенное максвелловское распределение в качестве начального условия и полагаемся на то, что динамика моделирования должна превратить его в точное максвелловское распределение посредством процессов столкновения с решеткой, существующих в моделировании. Этот метод, хотя и более простой, требует генерации девяти случайных чисел на частицу и, вероятно, является более медленным, чем использование (10 49), при условии что значение egf^{-1} находится интерполяцией из таблицы для этой функции.

Если к электродам теперь приложено типичное рабочее напряжение, то заряды входят и выходят из электродов, пока не установится условие стационарного течения. Этот процесс иллюстрируется на рис. 10.14, который показывает эквивалентный

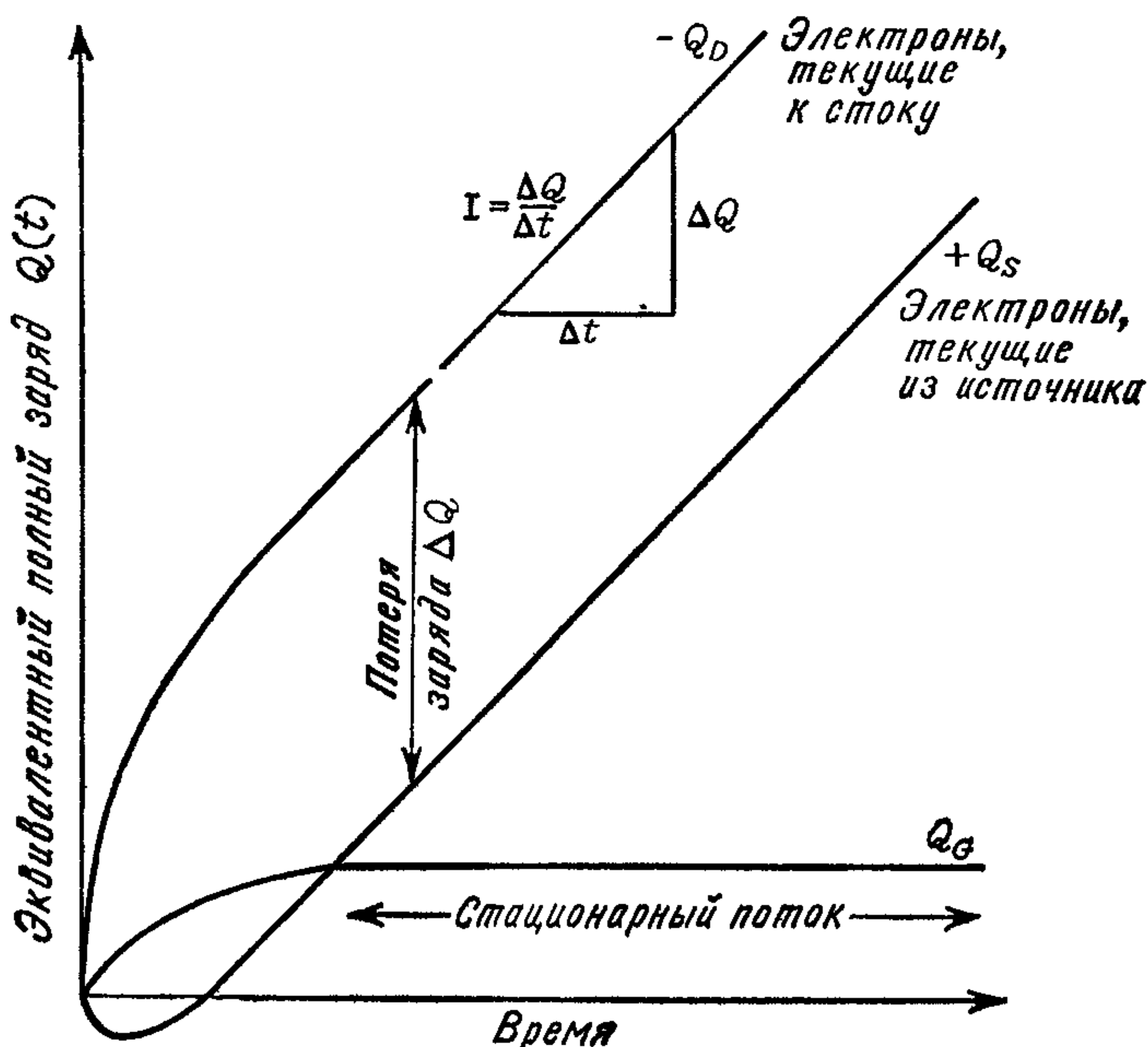


Рис. 10.14. Эквивалентный полный заряд как функция времени для электродов источника, стока и затвора после скачка, заданного в вольтах на стоке. В качестве начального условия выбран нейтральный прибор на рис. 10.13. Разница между потерей электронов на стоке и электронами, инжектированными в источник, является потерей электронов в течение образования обеднения под затвором.

заряд $Q(t)$, протекший через каждый электрод к моменту времени t . Заметим, что кривые отвечают суммарному числу электронов, поглощенных стоком и инжектированных источником. Первоначально электроны отражаются отрицательным потенциалом на затворе и пока в области ниже затвора число носителей убывает, покидают прибор через источник и сток. В течение этого начального процесса установления поле нарастает у затвора, что приводит к наведению заряда смещения на затворе. Разумеется, на затворе отсутствует заряд или поток частиц, поскольку затвор отражает электроны. Примерно после 10 пс устанавливается состояние стационарного потока со стационарным распределением заряда, и среднее количество электронов, выходящих из источника, равно среднему количеству электронов, входящих в сток. Ток, проходящий через каждый электрод, задается выражением

$$I(t) = \frac{dQ(t)}{dt}, \quad (10.51a)$$

где $Q(t)$ определяется из (10.47).

Из-за больших флуктуаций в $Q(t)$ трудно получить надежную оценку для I_{ss} (установившийся ток) подгонкой прямой линии через график зависимости заряда от времени. Однако I_{ss} можно оценить аккуратно с помощью автокорреляции тока $I(t)$. В установившемся условии ток может выражаться как сумма установившейся величины I_{ss} плюс ток $i(t)$, обусловленный случайным шумом, который имеет среднее значение, равное нулю; таким образом,

$$I(t) = I_{ss} + i(t). \quad (10.51b)$$

Величина $I(t)$ получается из $Q(t)$ дифференцированием. Если использовать верхние индексы для обозначения временного слоя, то

$$I^{(n)} = (Q^{(n+1)} - Q^{(n)})/DT, \quad (10.51c)$$

где $I^{(n)} = I(nDT)$ и $Q^{(n)} = Q(nDT)$. Автокорреляционная функция $A(\tau)$ определяется выражением

$$A(\tau) = \frac{1}{N-n'} \sum_{n=1}^{N-n'} I^{(n)} I^{(n+n')}, \quad (10.52a)$$

где корреляционное время $\tau = n'DT$ и всего имеется N шагов. Подставляя (10.51b) в (10.52a), получаем

$$\begin{aligned} A(\tau) = & \frac{1}{N-n'} \sum_{n=1}^{N-n'} I_{ss}^2 + \frac{I_{ss}}{N-n'} \sum_{n=1}^{N-n'} i^{(n)} + \\ & + \frac{I_{ss}}{N-n'} \sum_{n=1}^{N-n'} i^{(n+n')} + \frac{1}{N-n'} \sum_{n=1}^N i^{(n)} i^{(n+n')}, \end{aligned} \quad (10.52b)$$

Первый член в (10.52b) имеет постоянное значение I_{ss}^2 . Второй и третий члены представляют усреднение по $i(t)$ и поэтому стремятся к нулю при увеличении длины записи. Четвертый член является автокорреляцией случайных флуктуаций и, следовательно, стремится к нулю с увеличением корреляционного времени. Таким образом, установившийся ток получается из выражения

$$I_{ss} = [\lim A(\tau)]^{1/2}, \quad (10.52c)$$

где предел берется для больших τ , что на практике означает $\tau > 1$ пс. Ошибки, присущие измерению во втором, третьем и четвертом членах, могут понижаться далее усреднением $I(t)$ перед расчетом автокорреляции. Сглаженная величина $I_s(t)$ получается взятием скользящего среднего по m временным шагам:

$$I_s^{(n)} = \frac{1}{m} \sum_{n'=0}^{m-1} I^{(n+n')}. \quad (10.52d)$$

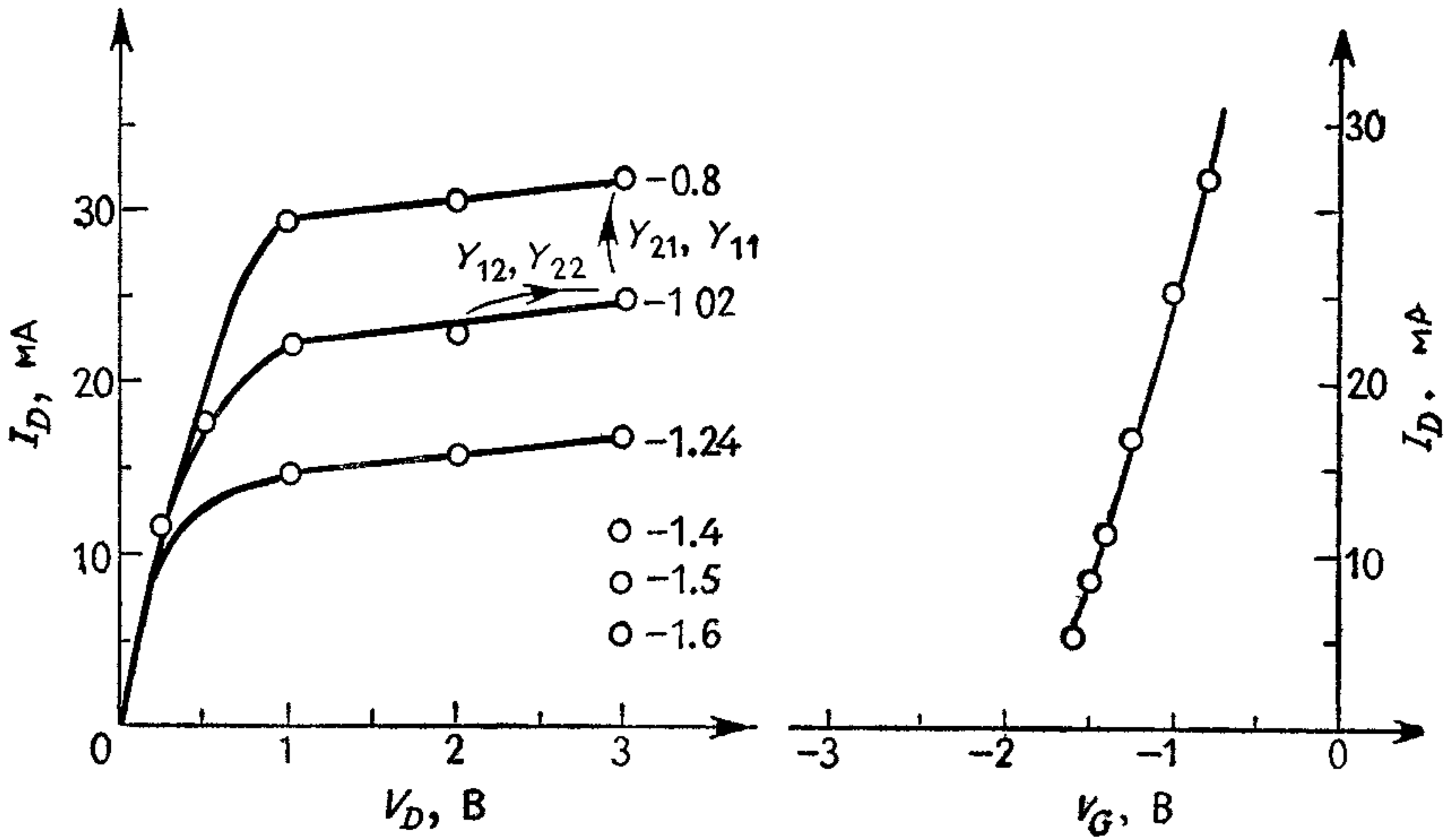


Рис 10 15 Измеренные статические характеристики для моделируемого прибора. Точка нормальной работы располагается при $V_D = 3,0$ В и $V_G = -1,02$ В. Показаны скачкообразные изменения, которые обеспечивают параметры Y в рабочей точке. Наклон характеристики I_D/V_D (слева) — выходное сопротивление, наклон характеристики I_D/V_G (справа) представляет крутизну. ([384], с любезного разрешения Solid State and Electron Devices, © Institution of Electrical Engineers)

Полная емкость затвора C_G может быть найдена из заряда обедненной области ΔQ_D и напряжения на затворе V_G :

$$C_G = \frac{\Delta Q_D}{V_G}. \quad (10.53)$$

Эта емкость будет суммой емкостей систем источник — затвор C_{sg} и сток — затвор C_{gd} , которые обсуждаются ниже в разд. 10.4.4.

Результатом проведения серии таких экспериментов со стационарными потоками является статическая характеристика I_D/V_D , как показано на рис. 10 15 (слева). Здесь изменение тока стока в зависимости от напряжения на стоке получается для ряда смещений затвора. Наиболее важными являются три характеристики: выходное сопротивление R_{out} , крутизна характеристики g_m и коэффициент усиления напряжения незамкнутого контура G_v . Эти величины определяются следующим образом:

$$R_{out} = \frac{\Delta V_D}{\Delta I_D} \text{ при постоянном } V_G, \quad (10.54a)$$

$$g_m = \frac{\Delta I_D}{\Delta V_G} \text{ при постоянном } V_D, \quad (10.54b)$$

$$G_o = \frac{\Delta V_D}{\Delta V_G} \text{ при постоянном } I_D, \\ = g_m R_{out}. \quad (10.54c)$$

Участок нормальной работы ПТ соответствует пологой области характеристики, т. е. насыщению. Для рассматриваемого прибора это положение находится при $V_D = 3,0$ В и $V_G = -1,02$ В. Обычно положение начала насыщения («колени» на кривой I_D/V_D) и выходное сопротивление должны находиться из трех прогонов программы, начинающихся с условия нейтральности заряда:

$$\begin{aligned} \text{Режим 1 } V_G &= -0,8 \text{ В, } V_D = 0,5, 1,0, 2,0, 3,0 \text{ В,} \\ \text{Режим 2 } V_G &= -1,02 \text{ В, } V_D = 0,5, 1,0, 2,0, 3,0 \text{ В,} \\ \text{Режим 3 } V_G &= -1,24 \text{ В, } V_D = 0,5, 1,0, 2,0, 3,0 \text{ В.} \end{aligned} \quad (10.55)$$

Вычислительную программу можно организовать так, чтобы увеличивать напряжение на стоке по этапам через каждые 300—400 временных шагов, поскольку это равно времени, необходимому для достижения условия стационарности потока. В этом случае условия течения, полученные из последней пары напряжений, становятся начальным условием для следующей пары, и переходные процессы сокращены до минимума. Расчет каждой точки на характеристике требует от 10 до 15 мин машинного времени на ЭВМ IBM 360/195.

Вышеприведенные величины V_G являются напряжениями, используемыми в качестве граничных условий в численном моделировании, и напряжениями в GaAs, как раз примыкающим к затвору. Поскольку численное моделирование не включает потенциал барьера Шоттки, равный 0,8 В, между алюминиевым затвором и материалом из арсенида галлия, для получения эквивалентного лабораторного напряжения к этим величинам нужно добавить 0,8 В. Источник и сток являются омическими контактами без потенциального барьера, и поэтому значения V_D и V_S в лабораторных экспериментах те же самые, что в моделировании.

Качественный ПТ будет поддерживать высокое значение g_m при низких выходных токах, следовательно, важно измерять вариации g_m при дальнейшем росте отрицательного напряжения на затворе. Эту информацию можно получить дальнейшим вычислительным процессом, начиная с точки нормальной работы, координаты частиц для которой хранятся в памяти:

$$\text{Режим 4 } V_D = 3,0 \text{ В, } V_G = -1,24, -1,40, -1,55, -1,70, -1,80 \text{ В.} \quad (10.56)$$

Результат такого эксперимента показан на рис. 10.15 (справа). Крутизна характеристики соответствует наклону кривой I_D/V_G . Целью конструирования приборов является получение высокого значения g_m при низком токе I_D , поскольку это производит как большое усиление (см. (10.54с)), так и низкий шум (см. разд. 10.4.5). Однако большинство приборов показывает снижение g_m , когда понижается ток, и в разд. 10.4.5 даны результаты изучения оптимизации параметров для снижения этого эффекта. Это типично для роли, которую может играть численное моделирование в конструировании прибора.

Вычисление характеристик по результатам режимов 1—4 для одного прибора заняло бы около 3 ч на ЭВМ IBM 360/195. Более ограниченный набор точек, например не включающий режимы 1 и 3, получил бы R_{out} и g_m за 1½ ч. Характеристики, полученные на рис. 10.15, являются типичными для измеренных в лаборатории.

10.4.2. Наблюдение внутренних процессов ПТ

Измерение статических характеристик является стандартной лабораторной процедурой. Нашей целью в проведении измерений на моделируемых приборах является обоснование и конструирование. Численное моделирование, однако, дает нечто исключительно новое, а именно возможность наблюдать внутренние процессы в ПТ способом, совершенно невозможным в лаборатории. На каждом временном шаге координаты всех частиц доступны для наблюдения вместе с распределением элек-

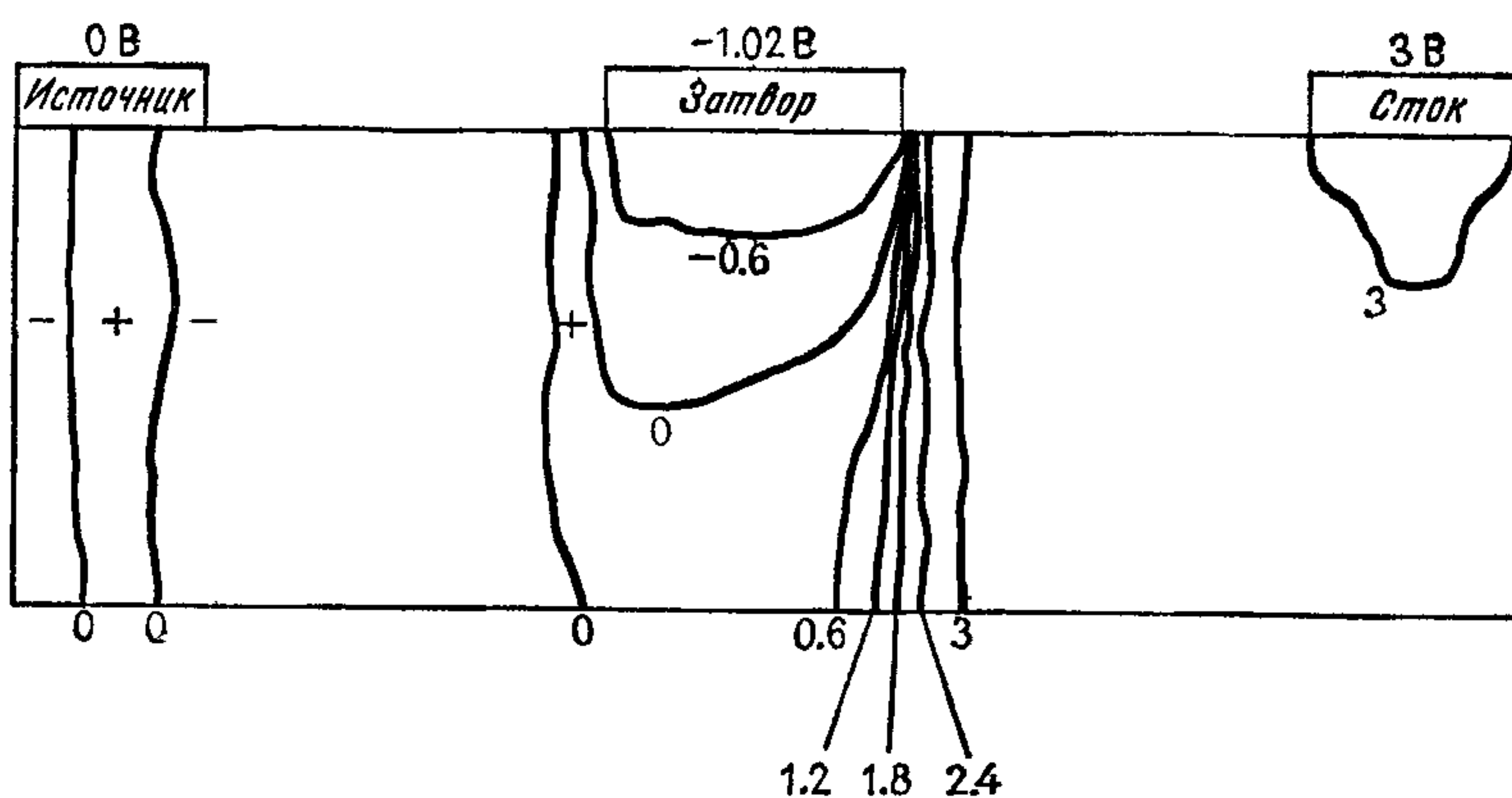


Рис. 10.16. Контуры электростатического потенциала в моделируемом ПТ. Отметим область сильного поля как раз справа от затвора, которая движет электроны долины I в долину II; в результате получается распределение, показанное на рис. 10.17. ([384], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers)

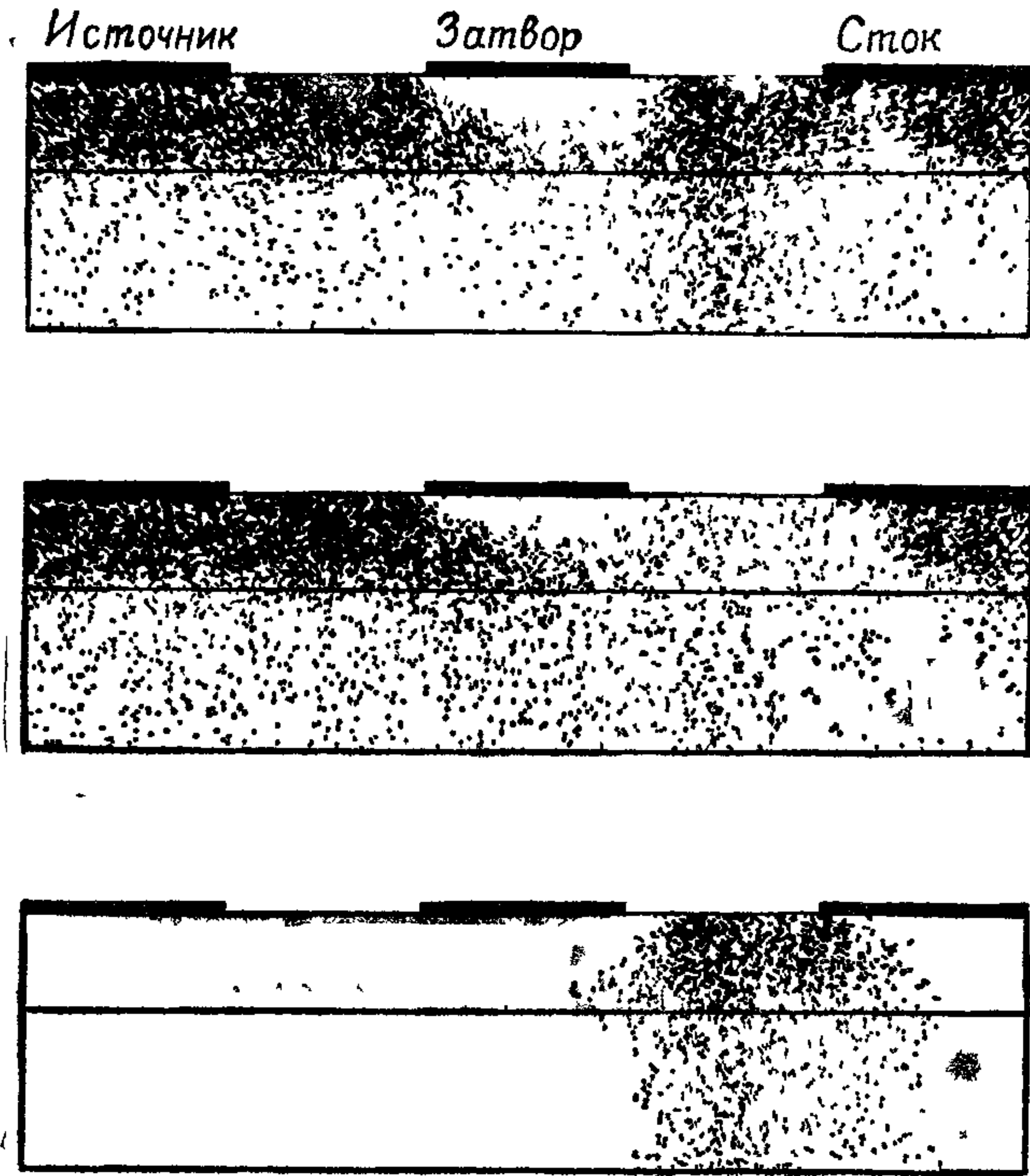


Рис. 10.17. Распределение электронов в точке нормальной работы ($V_D = 30$ В; $V_G = -102$ В). Показаны все электроны (вверху); электроны долины I (в центре); электроны долины II (внизу). Отметим электроны долины II в области сильного поля между затвором и стоком и значительный ток, текущий в подложке. Граница между эпитаксиальным слоем и подложкой показана горизонтальной линией. ([193], с любезного разрешения Electronics Letters, © Institution of Electrical Engineers.)

тростатического потенциала по всему прибору. Эта информация может анализироваться, чтобы дать полное описание функции распределения и электрического поля. Это детальное знание того, что происходит внутри прибора, дает информацию для улучшенного физического понимания происходящих в приборе процессов.

В качестве примера мы показали на рис. 10.16 контуры потенциала в приборе, соответствующие точке нормальной работы. Непосредственно можно видеть, что поля слабые, за исключением области $0,3$ мкм вблизи затвора со стороны стока. На этом коротком участке вся разность потенциалов между затвором и стоком падает на 4 В. Это поле имеет среднее значение 130×10^5 В/м, которого достаточно для заброса почти всех электронов проводимости в долину II. Это явление подтверждается

проверкой пространственного распределения всех электронов, что показано на рис. 10.17. Верхний, средний и нижний рисунки дают распределение всех электронов, электронов долины I и электронов долины II соответственно. Видно, что электроны долины II существуют только как раз справа в области сильного поля, лежащей примерно на середине между затвором и стоком. Этот результат противоречит выводам, вытекающим из диффузионной модели, согласно которой скорость мгновенно подстраивается под поле в соответствии с кривой \bar{v}/E (см. рис. 10.7). В этих моделированиях в области сильного поля непосредственно справа от затвора наблюдается аккумуляция электронов. Пренебрежение эффектами релаксации переходных процессов (разд. 10.2.5) в этих моделях приводит к ошибке в определении места аккумуляции около 0.5 мкм.

Другое заключение, вытекающее из рис. 10.17, состоит в неожиданно большом токе, текущем через слой подложки. Это приводит к мягкой отсечке для I_D/V_G и низкой величине g_m для низких I_D , как замечено ранее (разд. 10.4.1). Численное моделирование показало, что причина ухудшения характеристик заключается в проводимости слоя подложки. Подготовленные по традиции электронные инженеры не ожидали этого результата, потому что беспримесный подложечный слой имеет высокое сопротивление электрическому току при измерениях в лаборатории. Однако вычислительный эксперимент показал, что, если распределение потенциала является подходящим, в подложке будут течь существенные токи.

На основе вычислительного результата предложено включение буферного слоя p -типа между подложкой и эпитаксиальным слоем для того, чтобы образовать потенциальный барьер для потока электронов в подложку. Была проведена серия вычислительных экспериментов (см. разд. 10.4.5) для выбора наилучшей толщины и уровня легирования буферного слоя до изготовления прибора в лаборатории. Снова мы видим важное взаимодействие между численным моделированием и программой лабораторного исследования.

Наблюдаемое на рис. 10.16 очень быстрое изменение потенциала приводит к вопросу: достаточное ли разрешение для описания заряда имеет пространственная сетка? Имеется 15 сеточных ячеек с сеточным шагом $H = 0.02$ мкм на участке 0.3 мкм, на котором изменяется потенциал, и 50 сеточных ячеек поперек зазора между затвором и стоком. Эта сетка обеспечит адекватное представление изменения потенциала. Однако в моделях со значительно меньшим количеством сеточных ячеек между затвором и стоком, скажем с десятью ячейками или меньше, сеточная ячейка была бы сравнима с областью, на которой происходит изменение. Поскольку изменение потенциала не может иметь место на участке в несколько сеточных ячеек, крупная

ячейка этого типа будет неправильно размазывать заряд. Это приведет к недооценке электрического поля и даст неправильное распределение электронов в этой важной области прибора.

10.4.3. Динамические характеристики

Обсуждавшиеся в разд. 10.4.1 статические характеристики описывают стационарные токи, текущие под действием постоянных (т. е. статических) напряжений, приложенных к электродам. Они описывают поведение прибора при нулевой частоте или в условиях постоянного тока. Если мы хотим знать поведение прибора в любых условиях, необходимо исследовать ток в зависимости от приложенного напряжения на любой частоте. Тогда можно выразить реакцию прибора на любое изменение напряжения и иметь полное динамическое описание прибора. Параметры Y дают удобное динамическое описание, и сейчас покажем, как можно их измерить.

Полевой транзистор представляет собой двухпортовый прибор. Вход (порт 1) располагается между клеммами затвора и источника, а выход (порт 2) — между клеммами стока и источника. Зависящие от частоты параметры Y определяются матричным соотношением

$$\begin{bmatrix} \hat{i}_1(\omega) \\ \hat{i}_2(\omega) \end{bmatrix} = \begin{bmatrix} Y_{11}(\omega) & Y_{12}(\omega) \\ Y_{21}(\omega) & Y_{22}(\omega) \end{bmatrix} \begin{bmatrix} \hat{v}_1(\omega) \\ \hat{v}_2(\omega) \end{bmatrix}, \quad (10.57)$$

где \hat{v}_1 и \hat{v}_2 — амплитуды малых синусоидальных колебаний с частотой ω относительно точки нормальной работы, и \hat{i}_1 и \hat{i}_2 — амплитуды малых синусоидальных откликов тока. Некоторые из этих параметров (при нулевой частоте) уже знакомы как статические характеристики:

$$Y_{21}(0) = g_m, \quad (10.58a)$$

$$Y_{22}(0) = (R_{out})^{-1}, \quad (10.58b)$$

$$C_{\sigma} = - \left(\frac{Y_{21}(0)}{Y_{22}(0)} \right). \quad (10.58c)$$

Если мы всегда сохраняем напряжение источника равным нулю, тогда v_1 — вариация напряжения затвора и v_2 — вариация напряжения стока. Чтобы измерить $Y_{p,q}(\omega)$, мы должны были сохранять все напряжения постоянными, за исключением v_q , и измерять ток реакции i_p ; тогда

$$Y_{p,q}(\omega) = \frac{\hat{i}_p(\omega)}{\hat{v}_q(\omega)}. \quad (10.59)$$

Мы могли приложить по очереди различные частоты к порту q , но это было бы слишком дорого в отношении машинного времени. Взамен мы прикладываем ступенчатое изменение напряжения, которое содержит компоненты всех частот, к порту q , и получаем параметр Y для всех частот с помощью преобразования Фурье тока реакции \hat{i}_p .

Разложение Фурье зависящей от времени функции имеет вид (см. приложение А с $j = \sqrt{-1}$)

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} t(\omega) e^{j\omega t} d\omega, \quad (10.60a)$$

где фурье-преобразования даются выражением

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} t(t) e^{-j\omega t} dt. \quad (10.60b)$$

Поскольку $v_q(t)$ — ступенчатая функция:

$$\begin{aligned} v_q(t) &= \Delta v_q \quad \text{при } t \geq 0, \\ &= 0 \quad \text{при } t < 0, \end{aligned} \quad (10.61)$$

преобразования могут находиться непосредственно из соотношения

$$\hat{v}_q(\omega) = \int_0^{\infty} \Delta v_q e^{-j\omega t} dt = -\frac{\Delta v_q}{j\omega}, \quad (10.62)$$

где условно ω выбирается с малой отрицательной мнимой частью для того, чтобы значение интеграла при верхнем пределе было равно нулю. Ток реакции в канале p измеряется и численно преобразуется для получения соотношения

$$\hat{i}_p(\omega) = \int_0^{+\infty} i_p(t) e^{-j\omega t} dt. \quad (10.63)$$

Тогда находится параметр

$$Y_{p,q}(\omega) = \frac{\hat{i}_p(\omega)}{\hat{v}_q(\omega)} = j\omega \frac{\hat{i}_p(\omega)}{\Delta v_q}. \quad (10.64)$$

Вычисление $\hat{i}_p(\omega)$ требует некоторого дальнейшего объяснения. Значения $I(t)$, полного тока, проходящего через электрод, получаются из $Q(t)$ дифференцированием с помощью (10.51с). Если сделано изменение от стационарного тока I_{ss1} к стационарному току I_{ss2} в ответ на ступенчатое изменение приложенного

напряжения, то изменение тока около начального состояния I_{ss1} дается выражением

$$i_p^{(n)} = \begin{cases} \frac{1}{2} (I^{(n)} - I_{ss1}) \left(1 + \cos \frac{(n+8)\pi}{9} \right), & 1 \leq n \leq 10, & (10.65a) \\ (I^{(n)} - I_{ss1}), & 10 < n < 0.8N, & (10.65b) \\ (I_{ss2} - I_{ss1}) + \frac{1}{2} (I^{(n)} - I_{ss2}) \times \\ \times \left(\cos \frac{(n-0.8N)\pi}{0.2N} + 1 \right), & 0.8N \leq n \leq N, & (10.65c) \end{cases}$$

где $I^{(n)} = I(nDT)$ и $i_p^{(n)} = i_p(nDT)$ — полный ток и вариация тока на временном шаге n , всего имеется N шагов. Назначение членов с косинусом в (10.65a) и (10.65c) состоит в сглаживании измеренных величин в начальном и конечном состояниях и необходимости подавлять ложные гармоники в анализе Фурье величин $i_p(t)$. После N временных шагов (обычно около ~ 400) достигается окончательное установившееся значение; тогда получаем

$$i_p(t) = i_p(T) = I_{ss2} - I_{ss1}, \quad t > T = NDT. \quad (10.66)$$

Тогда для вычисления интеграла в (10.63) следует разбить его на две части:

$$\begin{aligned} \hat{i}_p(\omega) &= \int_0^T i_p(t) e^{-j\omega t} dt + \int_T^\infty i_p(T) e^{-j\omega t} dt = \\ &= \int_0^T i_p(t) e^{-j\omega t} dt + i_p(T) \frac{e^{-j\omega T}}{j\omega}. \end{aligned} \quad (10.67)$$

Первый член вычисляется с помощью численной квадратуры следующим образом:

$$\int_0^T i_p(t) e^{-j\omega t} dt = DT \sum_{n=1}^N i_p^{(n)} e^{-j\omega_l nDT}, \quad (10.68a)$$

где

$$\omega_l = 2\pi l/T, \quad l = 0, 1, \dots, N/2. \quad (10.68b)$$

В качестве примера этого метода рассмотрим определение выходного импеданса $[Y_{22}(\omega)]^{-1}$. Вычислительный процесс с 300 временными шагами являлся достаточным для достижения условия равновесного стационарного тока. Измеренный импеданс, полученный с помощью преобразования Фурье $i_2(t)$ и уравнения (10.64), показан на рис. 10 18. Этот отклик может интерпретироваться как сопротивление R_1 , параллельное последовательно соединенным сопротивлению R_2 и емкости C .

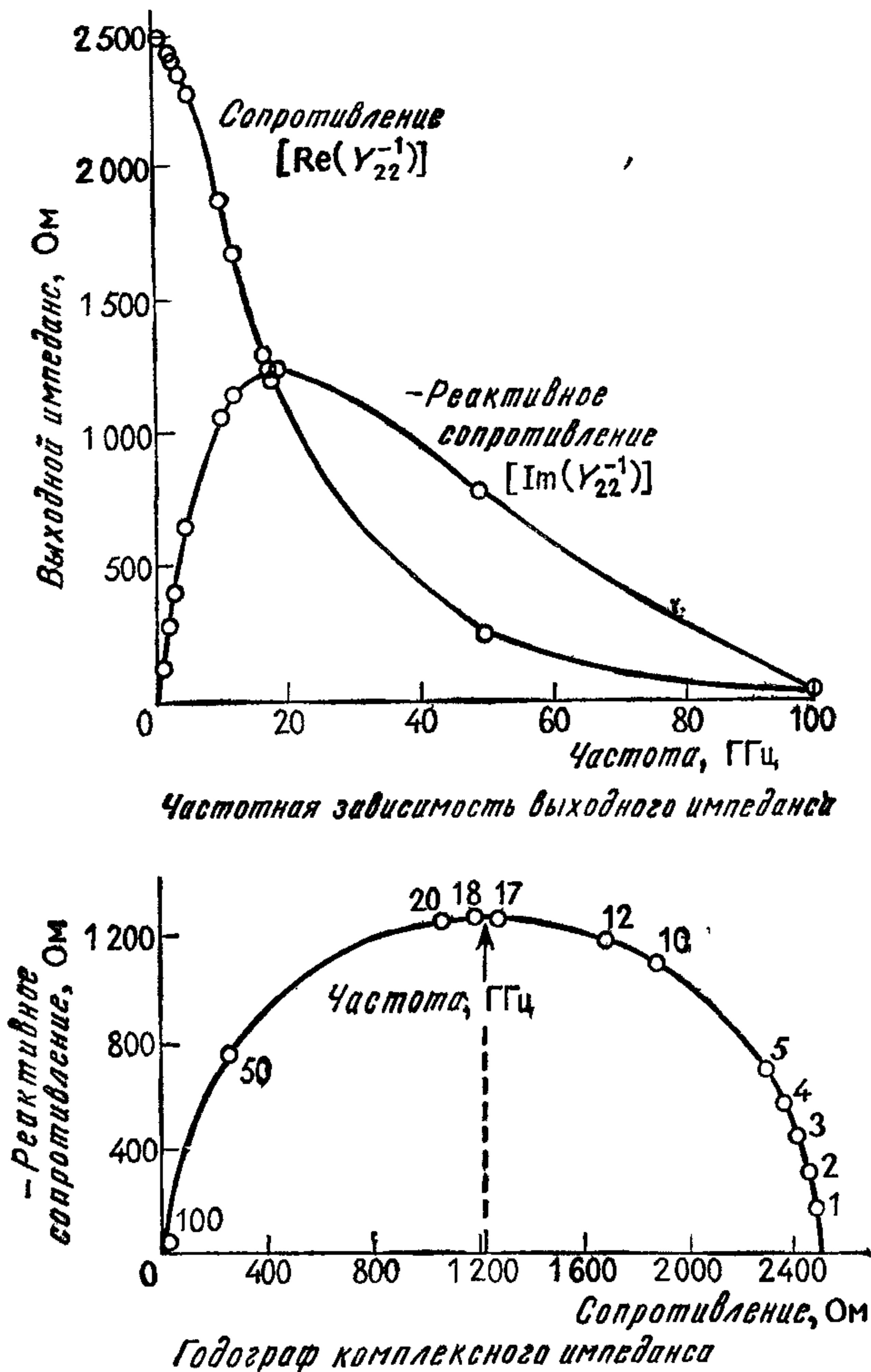


Рис. 10.18. Частотная зависимость выходного импеданса $[Y_{22}(\omega)]^{-1}$, полученная с помощью анализа Фурье отклика на ступенчатую функцию. ([384], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers.)

Полный импеданс этой цепи равен

$$Z_{\text{out}} = [Y_{22}(\omega)]^{-1} = \{R_1^{-1} + [R_2 + (j\omega C)^{-1}]^{-1}\}^{-1} = \text{Re}(Y_{22}^{-1}) + j\text{Im}(Y_{22}^{-1}), \quad (10.69a)$$

где

$$\frac{\text{Re}(Y_{22}^{-1})}{R_1} = \frac{1 + \omega^2 C^2 R_2 (R_1 + R_2)}{1 + \omega^2 C^2 (R_1 + R_2)^2}, \quad (10.69b)$$

$$\frac{\text{Im}(Y_{22}^{-1})}{R_1} = \frac{-\omega C R_1}{1 + \omega^2 C^2 (R_1 + R_2)^2}. \quad (10.69c)$$

При низких и высоких частотах импеданс действительный (т. е. омический). Если измеренные величины равны R_0 и R_∞ соответственно, то

$$R_1 = R_0 = 2500 \text{ Ом}, \quad (10.70a)$$

$$R_2 = R_0 R_\infty / (R_0 + R_\infty) = 25 \text{ Ом}. \quad (10.70b)$$

Емкость может оцениваться из мнимой (т. е. реактивной) части импеданса (10.69с), которая имеет максимальное отрицательное значение при $\omega C (R_1 + R_2) = 1$. Если максимуму соответствует частота f_m , то

$$C = [\omega_m (R_1 + R_2)]^{-1} = 0.0034 \text{ пФ}, \quad (10.70c)$$

так как $\omega_m = 2\pi f_m$ и $f_m = 18 \text{ ГГц}$.

Характерное время релаксации τ для цепи дается постоянной времени RC разряда конденсатора C через сопротивление $(R_1 + R_2)$,

$$\tau = C (R_1 + R_2) = \omega_m^{-1} = 8.6 \text{ пс}. \quad (10.70d)$$

Время релаксации является мерой минимального времени переключения прибора. Усиление пропорционально Y_{22}^{-1} (см. (10.58с)), и на рис. 10.18 показано, что f_m является частотой, выше которой усиление быстро падает. Поэтому она может рассматриваться как максимальная рабочая частота прибора.

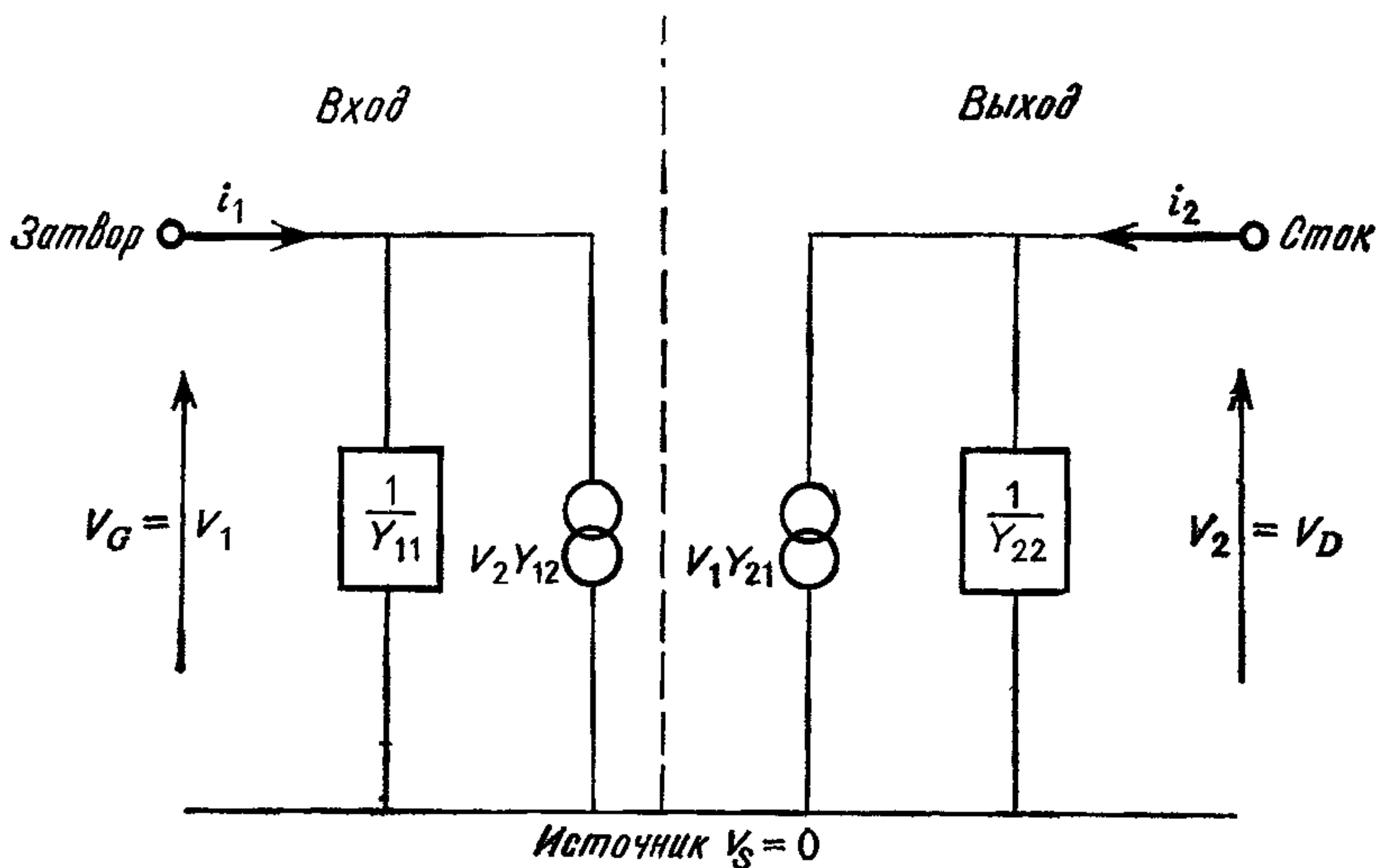


Рис. 10.19. Эквивалентная цепь для ПТ, использующая динамические параметры Y . Если заданы четыре величины $Y_{11}(\omega)$, $Y_{12}(\omega)$, $Y_{21}(\omega)$ и $Y_{22}(\omega)$ как функции частоты ω , то отклик ПТ на малый импульс в любой цепи может получаться с помощью стандартных методов анализа цепи. (Любезно предоставлено Уорринером, [381].)

Четыре параметра Y , $Y_{11}(\omega)$, $Y_{12}(\omega)$, $Y_{21}(\omega)$ и $Y_{22}(\omega)$, могут быть найдены как функция частоты в двух вычислительных режимах. В первом режиме создается скачок в напряжении стока, а во втором — скачок в напряжении затвора. В обоих случаях реакция тока измеряется на затворе и стоке и анализируется, как описано выше. Заметим, что такие изменения уже были вычислены в процессе измерения статических характеристик (режимы 2 и 4), и поэтому параметры Y могут быть получены анализом переходных откликов, полученных в этих режимах. На рис. 10.15 показаны два скачкообразных изменения, которые дают четыре параметра Y вблизи точки нормальной работы. Зная четыре параметра Y , можно вычислить отклик на малый импульс ПТ в любой цепи с помощью стандартной техники анализа цепей [132]. Параметры Y эквивалентной цепи ПТ для использования в этом способе показаны на рис. 10.19. Отклик на входе можно вычислить при рассмотрении ПТ в качестве входного импеданса Y_{11}^{-1} , параллельного источнику тока величиной $Y_{12}v_2$, а отклик на выходе — при рассмотрении ПТ в качестве выходного импеданса Y_{22}^{-1} , параллельного источнику тока величиной $Y_{21}v_1$. Параметры Y для ПТ получены в [288] с помощью рассмотренных выше методик.

10.4.4. Эквивалентная цепь с сосредоточенными параметрами

Хотя четыре параметра Y полностью описывают отклик ПТ на малый импульс, они не могут быть связаны физически с деталями прибора. Например, мнимая часть Y_{22}^{-1} , несомненно, представляет емкостной эффект, но нет информации о том, как надо изменить геометрию прибора, чтобы его ослабить. Конструктор прибора заинтересован в эквивалентной цепи, которая собирает пространственно распределенные сопротивления и емкости в малое число сосредоточенных параметров, которые, таким образом, связаны с различными областями прибора. Одна такая широко используемая эквивалентная цепь показана на рис. 10.20. На верхней части диаграммы показаны сопротивления и емкости, изображенные в тех участках прибора, с которыми они связаны. Величины R_s и R_d — сопротивления, связанные с материалом между источником и затвором и затвором и стоком соответственно. Поскольку они являются в сущности паразитными явлениями, предпочтительно держать путь тока в этих районах широким, а связанные с этим сопротивления низкими. Область под затвором рассматривается как канал, и здесь токопроводящая область сужается отрицательным потенциалом на затворе. Поэтому сопротивление канала R_c велико. Контролирующее действие затвора представлено с помощью g_m , которое означает, что проходящий в канале ток равен $g_m V_G$. Пол-

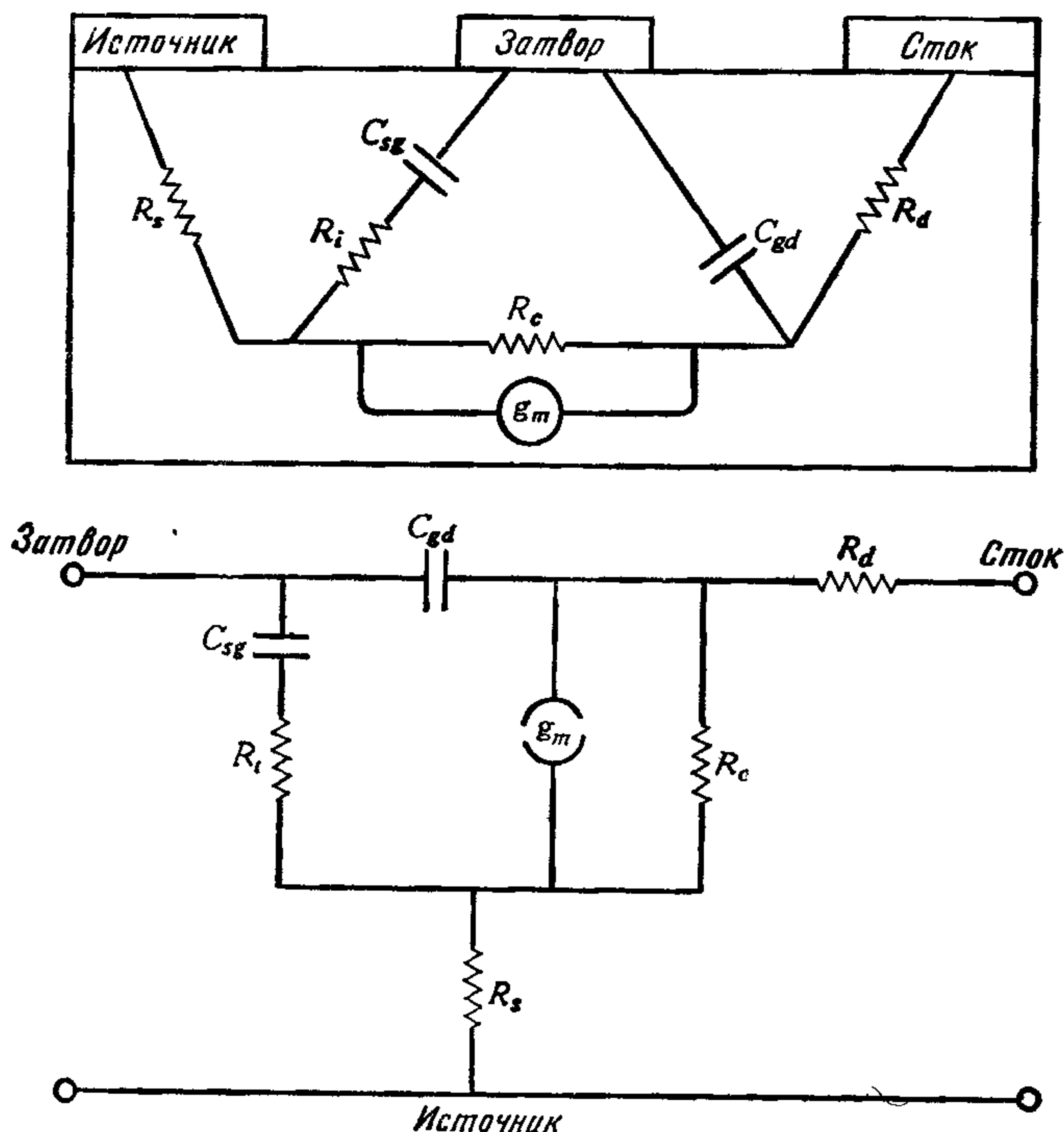


Рис 10.20 Эквивалентная схема с сосредоточенными параметрами для ПТ. Каждое сопротивление или емкость представляет распределенные свойства в области прибора ([384], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers)

ная емкость затвора, упомянутая в связи с рис. 10.14, распределена между источником C_{sg} и стоком C_{gd} .

Уорринером [384] предложена следующая процедура для измерения значений общих параметров:

1. $(R_s + R_i)$. Начиная с условия нейтральности заряда и нулевого напряжения на всех электродах (см. рис. 10.13), к затвору прикладывается скачок напряжения, равный -0.8 В. Это задает прибор с двумя путями от источника к затвору. Эквивалентный заряд, текущий из затвора, будет иметь форму

$$Q_G(t) = \Delta Q_G (1 - e^{-t/\tau}), \quad (10.71)$$

где константа времени $\tau = 1/2 (R_s + R_i) C_{sg}$. Множитель $1/2$ появляется из-за удвоенных путей источника. Константу времени τ можно получить методом наименьших квадратов из (10.71) или альтернативно очень быстро оценить по точке пересечения начального наклона $Q(t)$ с ее конечной величиной ΔQ_G , как показано на рис. 10.21. Емкость получается из ее определения как

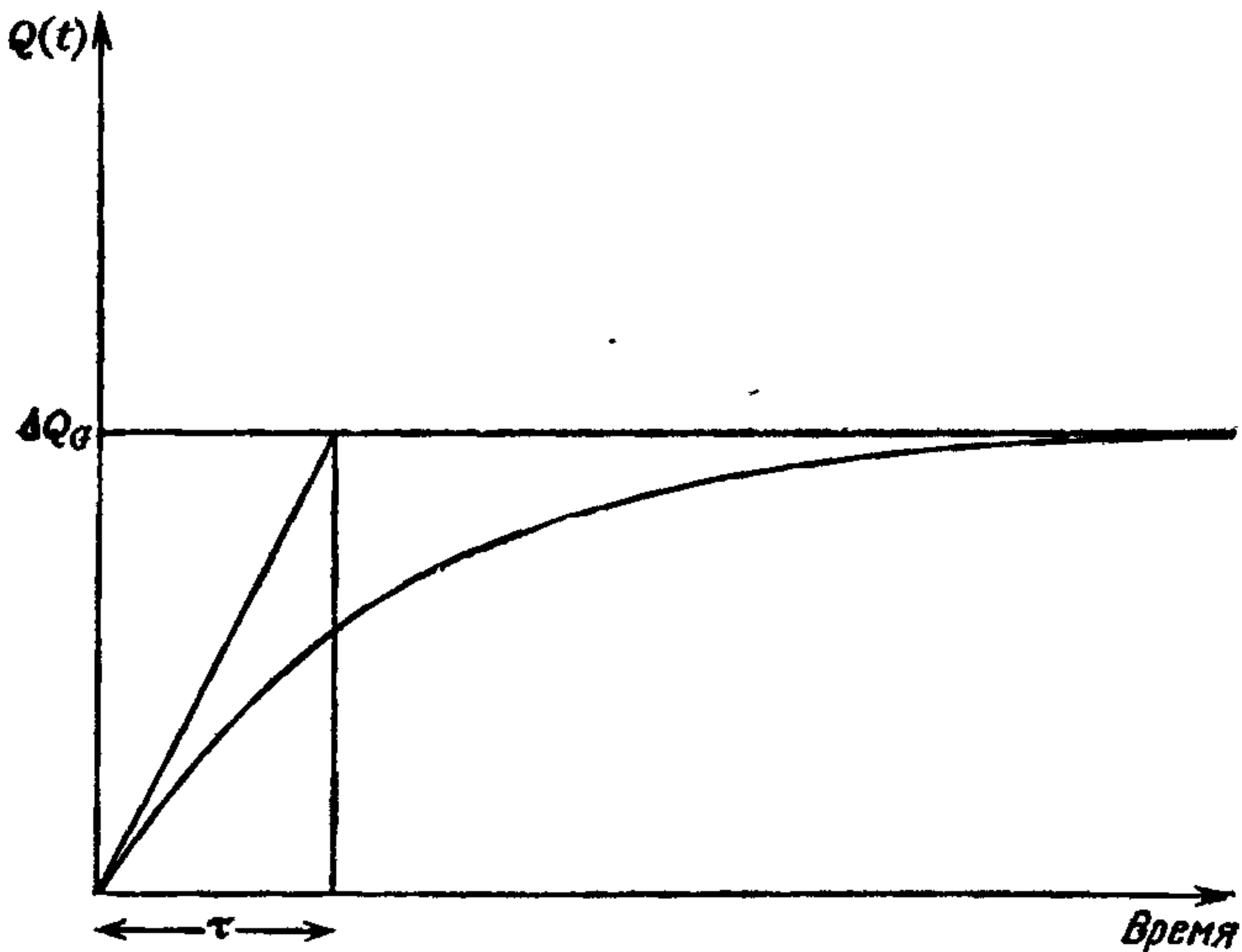


Рис. 10.21. Вычисление постоянной времени τ из кривой, имеющей зависимость $Q(t) + \Delta Q_G (1 - e^{-t/\tau})$.

перенесение заряда на единицу напряжения, следовательно,

$$C_{sg} = \frac{\Delta Q_G}{\Delta V_G}, \quad (10.72a)$$

$$(R_s + R_t) = \frac{2\tau}{C_{sg}}. \quad (10.72b)$$

Этим методом невозможно разделить R_s и R_t . Действительно, когда мы смотрим на распределение заряда в приборе (см. 10.17), то очевидно, что в этом разделении довольно мало смысла.

2. R_c . Поскольку $R_c \gg R_s + R_d$, сопротивление канала R_c можно получить из наклона статического отношения I_D/V_D (рис. 10.15) вблизи рабочей точки:

$$R_c = \frac{\Delta V_D}{\Delta I_D}. \quad (10.73)$$

3. R_d и C_{gd} . Поскольку $R_c \gg R_d$, действие скачкообразного изменения напряжения стока ΔV_D должно будет состоять в зарядке емкости C_{gd} , главным образом через сопротивление R_d . Эквивалентный заряд, текущий к затвору, будет изменяться согласно (10.71) с $\tau = C_{gd}R_d$. Константу времени можно получить,

как в шаге 1; тогда

$$C_{gd} = \frac{\Delta Q_G}{\Delta V_D}, \quad (10.74a)$$

$$R_d = \frac{\tau}{C_{gd}}. \quad (10.74b)$$

В табл. 10.2 сравниваются общие параметры, полученные в трех смоделированных приборах, с рассмотренными типичными величинами для реального прибора. Моделированные приборы без

Таблица 10.2. Значения сосредоточенных параметров для трех смоделированных и одного реального ПТ

Параметры цепи	Без РИП		С РИП *	Реальный прибор
	**	*		
$R_s + R_t$	2.8 Ом	2.8 Ом	9.8 Ом	~10 Ом
R_d	10 Ом	6 Ом	10.2 Ом	~10 Ом
C_{sg}	0.38 пФ	0.36 пФ	0.14 пФ	0.1 ÷ 1.0 пФ
C_{gd}	0.02 пФ	0.03 пФ	—	~0.01 пФ
R_c	750 Ом	857 Ом	1000 Ом	~1000 Ом
g_m	37.5 мСм	43.2 мСм	—	30 мСм

РИП — рассеяние на ионизированной примеси.

* подложка из чистого GaAs.

** граница раздела между полностью отражающей подложкой и эпитаксиальным слоем.

рассеяния на примеси и, естественно, предсказывают слишком малое значение сопротивлений R_s , R_t , R_d и R_c из-за пренебрежения этим важным резистивным эффектом. Проведено сравнение подложки из чистого GaAs и прибора, отражающего все электроны, которые попадают на границу подложки. Величины, полученные после учета рассеяния ионизированной примесью, находятся в хорошем согласии со значениями для аналогичных лабораторных приборов.

10.4.5. Шумы

Одна из принципиальных целей конструирования ПТ заключается в создании приборов с низким уровнем шума. Описанное в этой главе моделирование рассеяния по Монте-Карло особенно подходит для наблюдения шума, потому что все физические процессы, вызывающие шум в физическом приборе, также имеются и при моделировании. Шум связан с флуктуациями

выходного тока относительно среднего значения, которые возникают в результате движения и рассеяния отдельных электронов. Заметим, что все эти эффекты отсутствуют при моделировании обычного уравнения диффузии, которое описывает движение электрона только посредством заданной характеристики \bar{v}/E (см. разд. 10.2.5). Ниже мы дадим метод измерения, предложенный Хокни.

Полный ток $I(t)$ около электрода можно разбить на три части:

$$I(t) = \bar{I} + i(\omega) e^{i\omega t} + \delta i(t), \quad (10.75)$$

где первый член представляет средний ток, обусловленный стационарным смещающим напряжением, второй член — синусоидальный импульс тока частоты ω и амплитуды $\hat{i}(\omega)$ и третий член — случайные флуктуации тока, обусловленные дискретной природой заряда электронов, которые мы называем шумом. Средние по времени от сигнала и флуктуаций шума равны нулю.

Полную мощность, обусловленную этим током, можно точно так же разделить на три части, пропорциональные величинам

$$\langle \bar{I}^2 \rangle — \text{диссипация мощности, обусловленная токами смещения,} \quad (10.76a)$$

$$|\hat{i}(\omega)|^2 — \text{мощность сигнала,} \quad (10.76b)$$

$$\langle (\delta i)^2 \rangle — \text{мощность шума,} \quad (10.76c)$$

где ломаные скобки представляют усреднение по времени, а вертикальные черточки — модуль возможной комплексной величины.

Важной величиной является отношение мощности сигнала, несущего информацию, к мощности шума, уничтожающего информацию; отсюда определяем коэффициент шума

$$NF = \frac{(\text{Сигнал/шум}) \cdot \text{Мощность на входе}}{(\text{Сигнал/шум}) \cdot \text{Мощность на выходе}} \quad (10.77)$$

как коэффициент качества прибора и попытаемся сделать его как можно меньшим.

Подставляя величины для мощности из уравнений (10.76), получаем

$$NF = \frac{|\hat{i}_G|^2}{\langle (\delta i_G)^2 \rangle} \div \frac{|\hat{i}_D|^2}{\langle (\delta i_D)^2 \rangle} = \quad (10.78a)$$

$$= \frac{\langle (\delta i_D)^2 \rangle}{\langle (\delta i_G)^2 \rangle} \times \frac{|\hat{i}_G|^2}{|\hat{i}_D|^2}, \quad (10.78b)$$

где входной ток (порт 1) отождествляется с затвором, а выходной ток (порт 2) — со стоком.

Первый множитель в (10.78b) является отношением среднеквадратичных флуктуаций тока на стоке к той же величине на затворе. Обе эти величины являются непосредственно измеряемыми в течение вычислительного процесса со стационарным током при нормальных нагрузочных напряжениях и могут быть получены в качестве побочного продукта измерения характеристик прибора (см. разд. 10.4.1—10.4.3). Второй множитель устанавливает отношение между амплитудами сигнала на затворе и стоке и может получаться из измеренных параметров Y . Выходной сигнальный ток имеет максимальное значение, когда выход коротко замкнут ($v_D = v_2 = 0$). Тогда из (10.57) получаем

$$\hat{i}_1 = Y_{11} \hat{v}_1, \quad (10.79a)$$

$$\hat{i}_2 = Y_{21} \hat{v}_1. \quad (10.79b)$$

Следовательно,

$$\frac{|\hat{i}_G|}{|\hat{i}_D|} = \frac{|\hat{i}_1|}{|\hat{i}_2|} = \frac{|Y_{11}|}{|Y_{21}|}, \quad (10.79c)$$

и окончательно минимальный коэффициент шума равен

$$NF_{\min} = \left| \frac{Y_{11}}{Y_{21}} \right|^2 \frac{\langle (\delta i_D)^2 \rangle}{\langle (\delta i_G)^2 \rangle}. \quad (10.80)$$

Уравнение (10.80) связывает коэффициент шума с измеренными величинами в моделированном приборе в нормальной рабочей точке. Наблюдавшиеся в модели флуктуации тока будут усилены на много порядков величины в связи с использованием макрочастиц при вычислении поля. Однако этот множитель перехода между величиной физических флуктуаций и величиной модельных флуктуаций в (10.80) сокращается. Следовательно, коэффициент шума, полученный из (10.80), для модели будет тем же самым, что и для физического прибора.

Задача прямого измерения шума, использующая (10.80), представляет интервал частот, по которым может измеряться шум. Если временной шаг равен DT , и шум измерен из численного моделирования по N временным шагам по интервалу времени $T = NDT$, то максимальные и минимальные значения частот, которые можно измерить, равны

$$f_{\max} = \frac{1}{2DT} = 10000 \text{ ГГц}, \quad (10.81a)$$

$$f_{\min} = \frac{1}{T} = \frac{2}{N} f_{\max} = 20 \text{ ГГц}, \quad (10.81b)$$

где подставлены типичные значения $DT = 5 \times 10^{-14}$ с и $N = 1000$ временных шагов. Максимальная частота является наивысшей измеримой частотой, при которой нет сеточных гармоник,

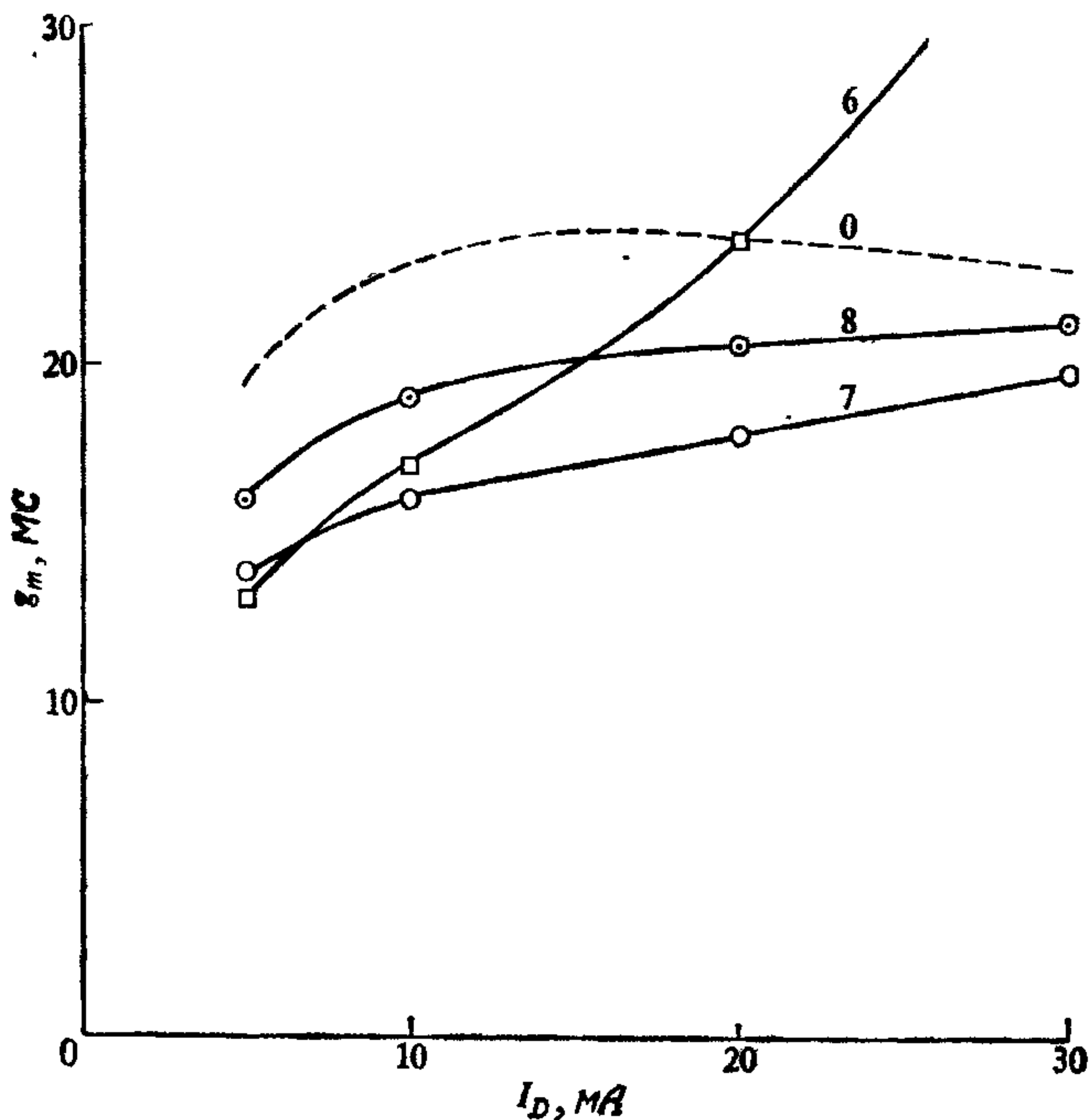
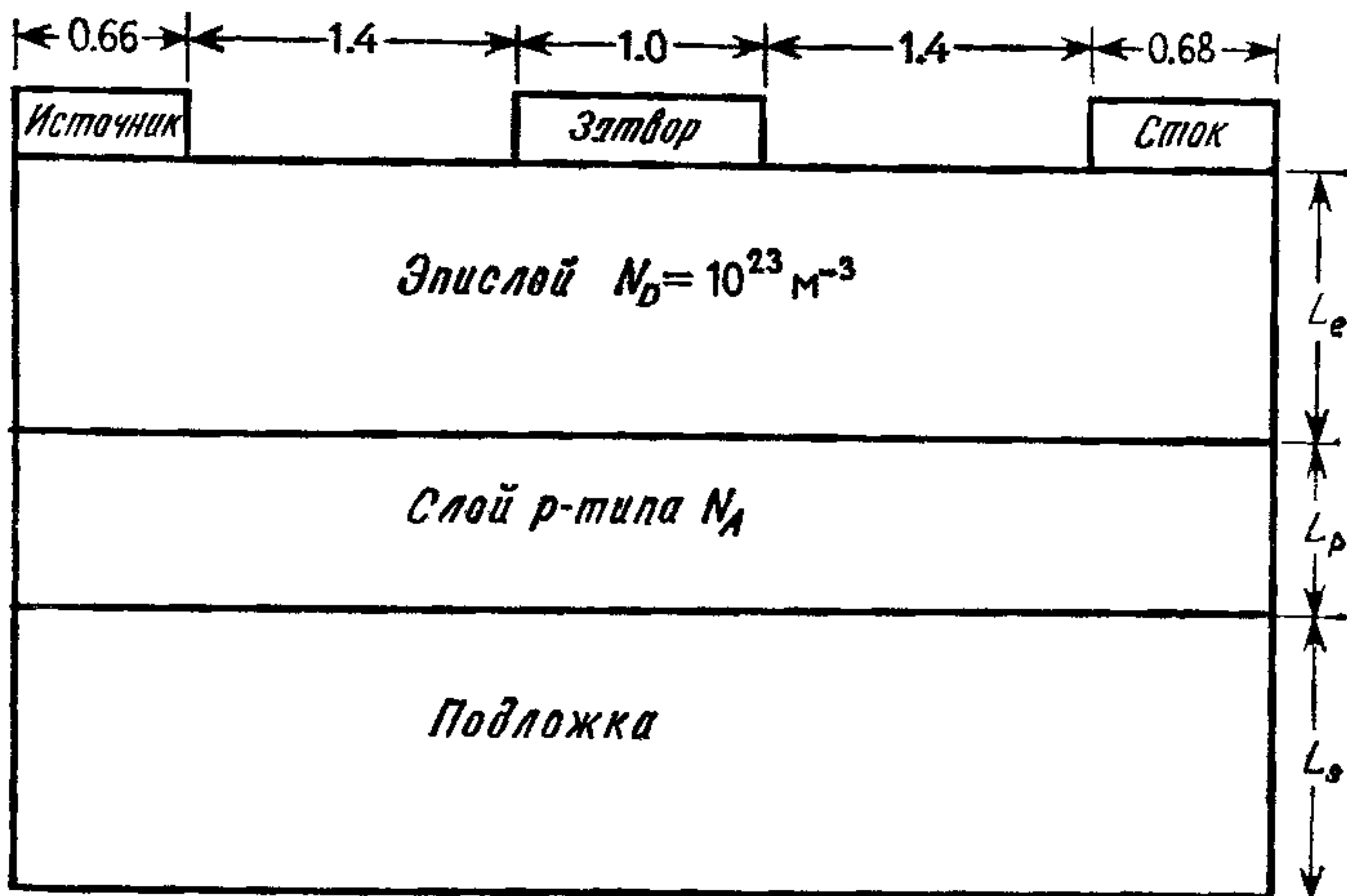


Рис. 10.22. Оптимизация плотности и толщины буферного слоя p -типа для получения высокого g_m при слабом I_D и, следовательно, низкого шума. Геометрия прибора (справа), кривые g_m/I_D (слева). Размеры даны в микрометрах. (Любезно предоставлено Сангхерой и Криссафисом.)

когда выборка производится на интервалах DT (см. разд. 7.3.2) и минимальная частота как раз допускает один период волны на интервале T и поэтому является наименьшей измеренной частотой при проведении анализа Фурье по этому интервалу. Минимальная частота приближается к рабочим частотам высокочастотных ПТ, но ясно, что потребовался бы чрезмерно длинный прогон ($> 10\,000$ временных шагов или 5 ч на ЭВМ IBM 360/195) для измерения шума на низких частотах.

Другой, но менее прямой подход должен связывать шум с воздействием на токовые диполи, распределенные в канале прибора. Простая аналитическая теория, развитая в работах [49, 50], дает вывод, что шум понижается, если g_m велико при малом I_D . Соотношение такого типа можно также усмотреть из (10.80), если мы заметим, что флуктуации тока вдоль канала подобны «дробовому» шуму, тогда как флуктуации тока поперек канала подобны тепловому шуму Джонсона. Первый приводит к флуктуациям в токе стока (следовательно, $\langle (\delta i_D)^2 \rangle \sim I_D$; [132, с. 108]), а последний — к флуктуациям тока затвора (сле-



Тип прибора	0	6	7	8
L_e	0.16	0.16	0.22	0.18
L_p	0.00	0.20	0.20	0.20
L_s	0.16	0.28	0.22	0.26
$N_A, \text{ м}^{-3}$	-	10^{22}	10^{22}	10^{22}

довательно, $\langle (\delta i_G)^2 \rangle \sim k_B T$; [132, с. 107]). Вспоминая также, что $Y_{21} = g_m$, имеем

$$NF_{\min} \sim I_D / g_m^2. \quad (10.82)$$

Экспериментально также найдено [378], что температура шума в приборе примерно пропорциональна величине I_D / g_m .

Изучение оптимизации прибора проведено Сангхерой и Хрисафисом с помощью численного моделирования для наилучшего выбора плотности легирования и геометрии ПТ с целью получить наилучшее g_m при слабом I_D и, следовательно, наименьший шум. Результат представлен на рис. 10.22. Сравниваются четыре прибора. Прибор 0 состоит только из эпитаксиального слоя с отражающей нижней границей. Этот прибор является идеальным и дает отличные характеристики из-за запрещения тока в подстилающем слое прибора. Однако он не может изготавливаться с помощью традиционных методов, описанных в разд. 10.1.2. Для приближения к идеальному прибору между эпитаксиальным слоем и подложкой вводится буферный слой р-типа. Этот буферный слой обеднен носителями, и отрицательные заряды легирующей примеси отталкивают падающие электроны и создают частично отражающую границу. Геометрия и

плотности легирования для приборов даны на рис. 10.22 (справа) и измеренные кривые g_m/I_D на рис. 10.22 (слева). Идеальный прибор 0 поддерживает примерно постоянное g_m при токе стока от 5 до 30 мА. Прибор 6, несмотря на то что имеет наивысшее g_m при больших токах, дает малое g_m при малых токах. Это быстрое изменение крутизны характеристики вследствие изменения рабочих условий является нежелательным. В приборе 7 мы имеем лучшее поведение и гладкую кривую g_m/I_D , однако значение g_m при слабом I_D (5 мА) является в сущности тем же, что в приборе 6. В приборе 8 мы имеем совершенно близкое приближение к идеальному поведению прибора 0 и значительно большую величину g_m при 5 мА, чем найдено в приборах 6 или 7. Описанное выше оптимизационное исследование типично для метода численного моделирования и может направлять выбор геометрии прибора.

Поддержанные упомянутым выше численным предсказанием, что прибор без подложки, подобный прибору 0, имел бы более высокое g_m и поэтому более низкий шум, исследователи в Royal Signals and Radar Establishment (Белдок) развили новую технологию производства таких совершенных приборов [378]. Процедура начинается с традиционного планарного ПТ, такого, который показан на рис. 10.22 (справа), содержащего легированный эпитаксиальный слой арсенида галлия толщиной ~ 0.5 мкм, выращенный на беспримесном буферном слое $\text{Ga}_{0.4}\text{Al}_{0.6}\text{As}$ толщиной ~ 0.7 мкм и подложке из компенсированного хромом арсенида галлия. Верхний слой этого прибора с его электродами затем устанавливался на стеклянную опору и переворачивался. Окончательно, и подложка, и буферный слой из GaAlAs избирательно вытравливались, оставляя перевернутый ПТ без подложки лежащим на стеклянной основе). Прибор может использоваться в этой перевернутой форме (тип В) или оголенная поверхность из GaAs может быть связана со стеклом, а исходная опора устранена, что приводит к обычному неперевернутому прибору с электродами на верхней свободной поверхности (тип А). С другой стороны, перевернутый тип В может иметь второй затвор, напыленный на GaAs напротив исходного затвора (тип С). В этом случае g_m примерно удваивается, потому что электронный ток сдвигается между двумя затворами. Специальная техника травления и соединения, требуемая в изготовлении этих совершенных приборов, описана в [150].

Эти приборы демонстрируют улучшенные значения g_m , что подтверждает вычислительные предсказания, и в принципе технология допускает любые комбинации электродов источника, затвора и стока на обеих сторонах тонкого активного слоя, приводящие к многочисленным возможным новым типам приборов. Развитие этих новых приборов как раз подтверждает важность распространения программы численного моделирова-

ния на любую лабораторию, которая занимается разработкой приборов. Для этого случая моделирование показывает, почему величина g_m для обычных приборов падает при низких I_D из-за того, что течение тока в изолирующей подложке — факт, которому вначале не верили многие разработчики прибора, поскольку подложки умышленно сделаны с высоким сопротивлением. Предложено также направление, в котором стоило бы усовершенствовать технологию производства, т. е. изготовить прибор без подложки.

Упомянутый выше опыт подтверждает сделанное в начале этой главы утверждение, что производство прибора игнорирует численное моделирование на свой страх и риск. Для такой программы моделирования требуется использование небольшой группы специалистов-вычислителей, имеющих квалификацию для развития, модификации и использования больших программ моделирования и обеспечения доступа к самым производительным из имеющихся в распоряжении компьютеров, на которых необходимо проводить моделирование. К сожалению, по состоянию на 1979 г. лишь немногие изготовители обеспечивали этот масштаб вычислительной поддержки своих разработок, хотя необходимые компьютеры, программы, знания и обучение были доступны в ряде университетов в течение многих лет.

10.4.6. Охлаждаемый ПТ

Отвод тепла, выделяемого триггерными элементами большой интегральной схемы (LSI), является весьма серьезной проблемой, особенно когда плотность упаковки логических элементов велика. Поэтому основной целью разработки прибора является построение триггерных элементов на ПТ с существенно меньшей диссипацией мощности, чем описанные до сих пор приборы. Рассмотренный в разд. 10.4.1 прибор имеет выходную диссипацию мощности 60 мВ (произведение напряжения и тока стока). Теперь опишем прибор с диссипацией, приблизительно в тысячу раз меньшей, возможности которого впервые найдены с помощью численного моделирования. Этот особый прибор, охлаждаемый ПТ, работает при температурах жидкого азота и, следовательно, создает специальные проблемы в лаборатории; однако исследование его с помощью численного моделирования представляет собой обычную процедуру применения методов анализа ПТ, описанных в разд. 10.4.1—10.4.5. Для исследования требуется около 3 ч времени на IBM 360/195 и 3—4 человеко-месяца для завершения работы. Мы излагаем здесь результаты, опубликованные в [323].

Для хорошего функционирования ПТ должен бы иметь высокую частоту отсечки (т. е. короткое время отклика) и высокий

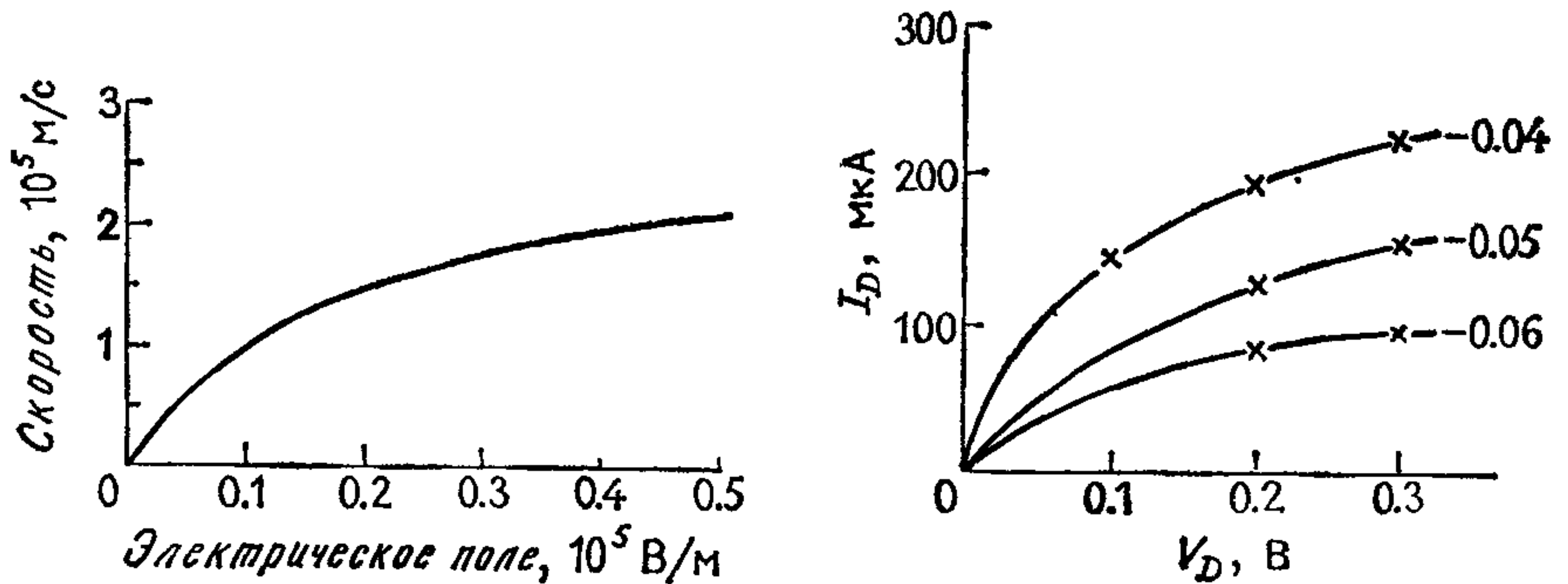


Рис. 10.23. Кривая дрейфовой скорости для GaAs при $T = 80$ К, плотность легирования $5 \times 10^{20} \text{ м}^{-3}$ (слева). Статическая характеристика I_D/V_D для охлаждаемого ПТ (справа). ([323], с любезного разрешения Electronics Letters, © Institution of Electrical Engineers)

выходной импеданс (чтобы дать большее усиление). Первая примерно пропорциональна максимальной дрейфовой скорости, а последний улучшается с помощью лучшего насыщения кривой \bar{v}/E . При комнатной температуре насыщение происходит в связи с рассеянием на верхних долинах, но для этого требуется потенциал в 3В. Слабое насыщение можно получить при одной десятой этого напряжения использованием температур жидкого азота (~ 80 К) и уровней легирования около $5 \times 10^{20} \text{ м}^{-3}$. При этой температуре и легировании имеет значение только рассеяние, обусловленное ионизованными примесями. Эффект насыщения получается посредством увеличения рассеяния, соответствующего порогу эмиссии полярно-оптических фононов. Это выделяется как выпуклость на кривой рассеяния частиц долины I при $k = 0.022 \times 10^{10} \text{ м}^{-1}$ на рис. 10.6 Все электронные энергии низки и недостаточны для рассеяния электронов в долину II. Таким образом, при этих температурах GaAs действует подобно материалу с одиночной долиной, например кремнию.

Кривая дрейфовой скорости для GaAs при температуре 80 К показана на рис. 10.23 (слева). Полученные скорости насыщения аналогичны кривой при комнатной температуре на рис. 10.7; следовательно, прибор будет иметь похожую частоту отсечки. Однако подвижность в слабом поле составляет $10 \text{ м}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$ и приблизительно в 25 раз больше, чем при комнатной температуре, из-за очень малой амплитуды колебаний решетки и, следовательно, малого рассеяния. Особенно важным является низкий уровень электрического поля, требуемый для нормальной работы, а именно $0,6 \times 10^5 \text{ В/м}$ при температуре 80 К по сравнению с $30 \times 10^5 \text{ В/м}$ при температуре 300 К.

Моделируемый прибор имеет ту же самую геометрию, что ПТ, изображенный на рис. 10.2, за исключением того, что эпи-

таксиальный слой и подложка оба имеют глубину 0.32 мкм. Статическая характеристика I_D/V_D для этого прибора показана на рис. 10.23 (справа). Умеренное насыщение получается на кривой для $V_G = -0,06$ В при $V_D = 0,3$ В и выходном токе $I_D = 100$ мкА. Диссипация выходной мощности на этой рабочей точке поэтому равна 30 мкВт, что в 2000 раз меньше, чем значение диссипации при комнатной температуре. Таким образом, численное моделирование смогло подтвердить исходную идею разработчика прибора. Были также измерены параметры Y , что

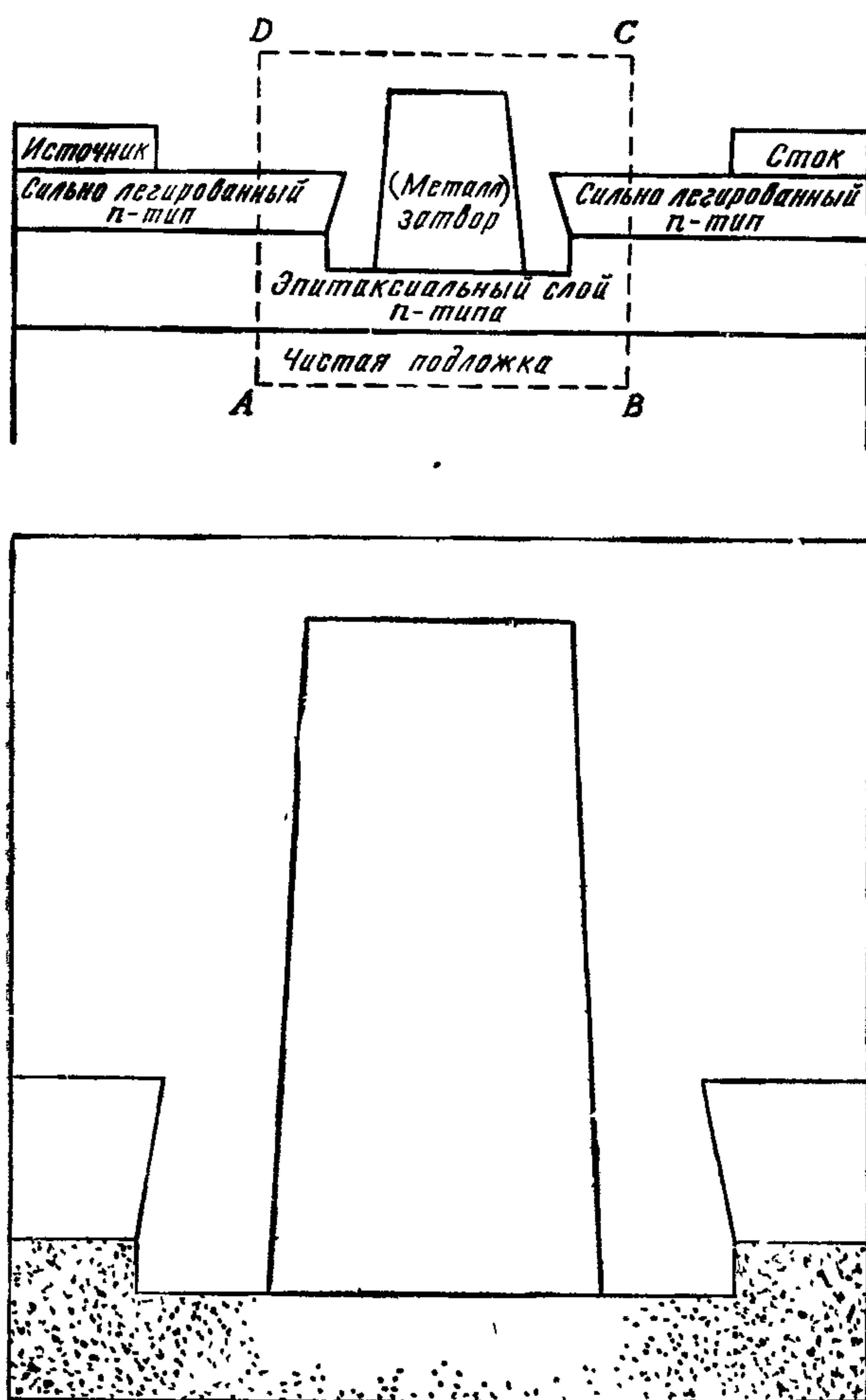


Рис. 10.24. Геометрия ПТ с вытравленным затвором, показывающая моделируемую на компьютере область ABCD (вверху); результат численного моделирования (внизу). (Любезно предоставлено С. Моглестью и С. Бердом).

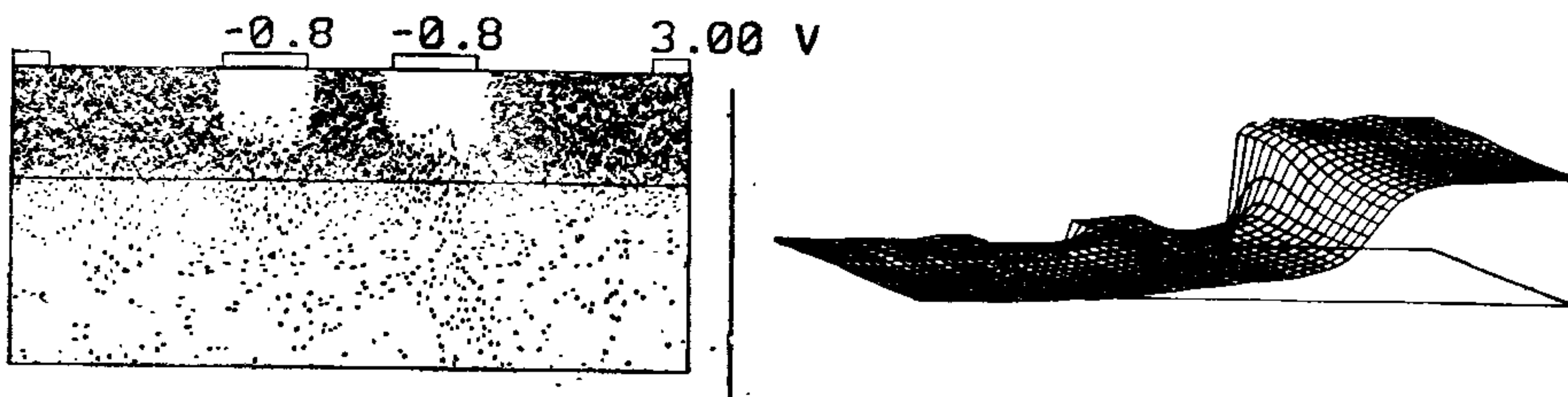


Рис. 10.25. Моделирование ПТ на арсениде галлия с двойственным затвором или на тетраде. Положение всех моделируемых электронов, показывающее две области обеднения под двумя затворами (слева); изометрическая проекция распределения потенциала в приборе (справа) ([287], с любезного разрешения Solid-State and Electron Devices, © Institution of Electrical Engineers.)

позволило предсказать поведение охлаждаемого ПТ в логической схеме [323]. Потребуется продолжительная программа лабораторных работ и дальнейшее численное моделирование перед тем, как прибор станет товарным продуктом. Если это удастся сделать, то оно станет классическим примером значения и роли численного моделирования для конструирования и усовершенствования электронных приборов.

10.4.7. Сложные геометрии

В моделировании ПТ, описанных к настоящему моменту, геометрия упрощена до прямоугольной, как показано на рис. 10.2. Это разумное приближение для многих приборов и имеет преимущество в возможности использования ранних методов быстрого решения эллиптических уравнений (например, PLOT4) для решения уравнения Пуассона (см. разд. 6.5.1—6.5.6). В некоторых реальных приборах улучшенные характеристики получаются с помощью помещения электрода затвора в вытравленный желоб, как показано на рис. 10.24 (вверху). Плоскую аппроксимацию проводить больше нельзя, и необходимо моделировать геометрию разреза $ABCD$. Распространение на метод матрицы емкости выполнено С. Дж. Бердом в программе PLOT4A, которая позволяет изменять диэлектрическую постоянную, а также задавать электроды внутри прямоугольной области. Сильно легированные n -области рассматриваются как эффективные электроды источника и стока из-за их высокой проводимости, и столбообразный затвор моделируется как электрод. Граница раздела между воздухом и полупроводником моделируется как поверхность, через которую меняется диэлектрическая постоянная. Вокруг левой и правой границ вычислительной области ставится условие равенства нулю нормальной компоненты электрического поля. На верхней и нижней поверхностях можно налагать условие, соответствующее бесконечно про-

тяженной области. Результаты этого моделирования позволяют проверить справедливость условия нормального поля, которое налагается в плоском приближении. Численное моделирование, показанное на рис. 10.24 (внизу), представляет прибор типа А, описанный в [378, рис. 2а].

Другая изученная геометрия представляет двухзатворный ПТ или тетрод [287, 288]. На рис. 10.25 демонстрируются положение частиц и электростатический потенциал при моделировании такого прибора. Это моделирование проведено посредством метода нахождения потенциала РОТ4, использованного для ПТ с одним затвором. Принципиальными проблемами являются проблемы, возникающие из-за большого размера прибора и, следовательно, большого числа требуемых в моделировании частиц.

Астрофизика

11.1. ВВЕДЕНИЕ

Астрофизика является особенно плодотворной областью для применения численного моделирования, поскольку изучаемые системы — звезды и галактики — не поддаются управляемому лабораторному эксперименту. При масштабах времени 10^8 лет и длины порядка 10^{23} м астрофизики и астрономы выступают только в роли пассивных наблюдателей изучаемых систем. Единственный способ выполнения возможных экспериментов заключается в моделировании системы на компьютере и проведении экспериментов с помощью этого моделирования.

11.1.1. Эволюция звезд

Возможно, наиболее важный вклад вычислительного эксперимента в астрофизику был сделан в изучении эволюции звезд (см., например, [204, гл. 8 и 9]). Этот процесс включает в себя решение уравнений в частных производных для переноса тепла внутри звезды одновременно с уравнениями, определяющими горение ядерного топлива. Это исследование было настолько успешным, что, как полагают, ныне вполне выяснена история жизни большинства типов звезд, от рождения через главную последовательность и стадию гиганта к угасанию, как белый карлик или нейтронная звезда. До сих пор модели частиц не играли роли в этих исследованиях, за исключением того, что они использовались при определении уравнения состояния однокомпонентной плазмы (см. гл. 12), которая представляет собой хорошую модель для определенных классов звезд типа белых карликов [203].

11.1.2. Гравитационная задача N тел

Роль модели частиц проявляется при моделировании систем, в которых звезду можно рассматривать как точечную массу, имеющую лишь свойство притяжения и массу, другими словами, такие важные при изучении эволюции звезд процессы, как ядерное горение, излучение и конвекция, не рассматриваются. Задача определения поведения совокупности N точеч-

ных масс, когда они движутся под действием их взаимных гравитационных сил в соответствии с законом движения Ньютона, носит название классической гравитационной задачи N тел. Эволюция скоплений из звезд или галактик (каждая рассматривается как точечная масса) и развитие спиральных и барообразных структур в одиночной галактике — все это всевозможные задачи, которые можно аппроксимировать таким образом.

11.1.3. Столкновительные и бесстолкновительные системы

Задачи N тел и методики, используемые при их решении, разбиваются на три класса в зависимости от важности парных столкновений в эволюции системы. Интенсивность двойных столкновений измеряется величиной, на которую возмущается орбита звезды в среднем локальном гравитационном поле из-за присутствия отдельных соседних звезд.

Для систем с фиксированной полной массой интенсивность двойных столкновений ν_D обратно пропорциональна числу звезд N , на которое эта масса делится (см. разд. 11.3.4). Если N удваивается, то, хотя число рассеивающих центров тоже удваивается, сила взаимодействия между парой звезд делится на четыре (сила пропорциональна квадрату массы). В результате получается, что полный эффект столкновений уменьшается наполовину.

Чандрасекхар [74, с. 73] предлагает следующую формулу для времени двойного соударения $T_D = \nu_D^{-1}$ в трехмерной системе точечных масс:

$$T_D = \frac{v^3}{8\pi n G^2 m^2 H \ln(Dv^2/2Gm)}, \quad (11.1)$$

где n — объемная плотность звезд; m — масса звезды; v — относительная скорость звезд; D — расстояние между звездами; G — гравитационная постоянная; $H = 0.4$ приближенно. Поскольку $m \sim N^{-1}$ и $n \sim N$, то, пренебрегая слабым изменением логарифма, имеем $\nu_D = T_D^{-1} \sim N^{-1}$, как утверждалось ранее.

Если система состоит из нескольких звезд, то орбита любой отдельной звезды зависит в основном от точного положения и массы ее локальных соседей, и говорят, что система становится системой с преобладающим влиянием этих двойных взаимодействий. Кульминацией столкновительной орбиты, конечно, является система двойных звезд, в которой две звезды вращаются вокруг друг друга и только их среднее движение реагирует на среднее поле системы. Эволюция системы с доминирующим столкновением сильно зависит от масс индивидуальных звезд, которые должны учитываться должным образом.

Для системы, состоящей из большого числа звезд, как в галактике, где $N \approx 10^{11}$, Чандрасекхар [74] оценивает, что время

отклонения орбиты звезды на одну десятую часть прямого угла (или 9°) составляет около 100 вращений галактики. Это время есть примерно предел времени, в течение которого эффекты столкновений могут рассматриваться как пренебрежимо малые. Оно также соответствует расчетному возрасту Вселенной, а следовательно, и максимальному возрасту галактики. По этой причине галактики описываются как бесстолкновительные системы.

При моделировании галактики представляют интерес следующие величины — средние значения массовой плотности, потенциала и поля, усредненные по сотням или тысячам межзвездных расстояний. Следовательно, частицы, используемые в моделировании галактики, переносят массу около 10^6 звезд и их движение представляет собой скорее перенос средней массовой плотности в их окрестности, чем орбиты любой отдельной звезды. Системы двойных звезд, несомненно, образуют значительную долю звезд в нашей галактике, но их подробное орбитальное движение не отображается в моделировании галактики. Такие звезды представляются лишь в той степени, насколько их движение влияет на локальные средние величины. Следовательно, задачи, связанные с интегрированием по времени тесных двойных систем, которые имеют преобладающее значение в системах с малыми столкновениями, не важны в бесстолкновительных системах.

Движение частицы в бесстолкновительной системе в значительной степени управляется локальным средним гравитационным полем и вообще вряд ли подвергается влиянию частного распределения других звезд в ее окрестности. Действительно, если наблюдается движение частиц в бесстолкновительной системе, то может показаться, что они дрейфуют друг через друга подобно теням, когда они реагируют на среднее поле. Кажется, что частицы игнорируют друг друга, и не имеется ни резких изменений, ни образования двойных систем.

Таким образом, эволюция бесстолкновительной системы определяется средней массовой плотностью, а не массами отдельных звезд. Поэтому моделирование для удобства обычно проводится с большим числом модельных звезд равной массы, даже несмотря на то, что в реальной галактике звезды имеют распределение по массам. Однако в бесстолкновительной системе это распределение не влияет на эволюцию локальных средних величин.

11.1.4. Скучивание звезд и галактик

Реальные системы звезд могут варьировать от систем с преобладанием столкновений до бесстолкновительных. Рассеянные, или галактические, звездные скопления представляют собой группы до нескольких сотен звезд (например, скопления Плеяды

и Гиады, каждое из которых содержит около 100 звезд), расположенные близко к плоскости галактического диска. Очевидно, что такие системы являются системами с преобладанием столкновений, и при моделировании следует точно интегрировать орбиту каждой звезды, принимая во внимание ее отдельную массу. Рассеянные скопления состоят из молодых звезд с характерным возрастом $\sim 2 \times 10^8$ лет и имеют возраст около одной десятой возраста галактики. Эффекты столкновения приводят к значительной потере звезд из этих скоплений.

В другом предельном случае имеется полная галактика из $10^{10} - 10^{12}$ звезд, которую можно рассматривать как бесстолкновительную систему, и в этом случае частицы при моделировании являются макрозвездами с массой, соответствующей примерно 10^6 реальным звездам. Это приближение является разумным для бесстолкновительных систем, поскольку, как было видно выше, массы частиц в бесстолкновительной системе не влияют на эволюцию средних величин. Где-то между этими крайними случаями лежат шаровые скопления. Эти скопления представляют собой скопления из $10^4 - 10^6$ звезд, сферической формы, расположенные вне галактической плоскости. Они образуют наиболее трудную для моделирования группу, поскольку N не является достаточно большим для того, чтобы пренебречь столкновениями, но является слишком большим, чтобы интегрировать орбиту каждой звезды тем же самым образом, как это делается для малого галактического звездного скопления. Шаровые скопления имеют приблизительно такой же возраст, как галактика.

Если мы рассмотрим Вселенную в космологическом масштабе, то опять имеем иерархию размеров скоплений (см. [152], [204, гл. 11]). Считая теперь, что частица представляет целую галактику, имеем, что наша собственная Галактика Млечный Путь принадлежит к местной группе из нескольких дюжин галактик, которые связаны гравитационными силами как малое скопление, имеющей размер приблизительно 16 галактических диаметров¹⁾. Наша собственная местная группа галактик находится вблизи края скопления, состоящего из малых скоплений и называемого локальным сверхскоплением. Его протяженность составляет 800 галактических диаметров, и оно содержит много тысяч галактик. На расстояниях, превышающих приблизительно 1000 галактических диаметров, густота скоплений падает. Полное число галактик в наблюдаемой части Вселенной оценивается по-разному, но должно составлять около 10^9 .

¹⁾ В этом обсуждении один галактический диаметр составляет 40 килопарсек (кпс), или 130 000 световых лет, что примерно равно диаметру нашей Галактики.

11.1.5. Большой взрыв

Чтобы понять истоки иерархии скоплений, необходимо рассмотреть космологические теории происхождения Вселенной. В сценарии большого взрыва все образовалось во время колоссального взрыва около 15—20 миллиардов лет назад¹⁾, когда все вещество Вселенной, главным образом в форме излучения, было выброшено с большой скоростью во всех направлениях из малого объема пространства. Затем гравитационное притяжение замедлило расширение, локальные изменения плотности были усилены локальным гравитационным притяжением и вещество сверхскоплений было выделено. В свою очередь, локальные изменения плотности в каждом сверхскоплении приводили к выделению скоплений, которые сами разделились на отдельные галактики.

Если средняя массовая плотность Вселенной меньше некоторой критической величины, то гравитационное притяжение не в состоянии остановить расширение, и говорят, что Вселенная открытая и будет всегда расширяться. Если, с другой стороны, плотность превышает это критическое значение, то гравитационное притяжение в конце концов обратит направление расширения и Вселенная будет сжиматься обратно к экстремально горячему и плотному состоянию, из которого процесс начнется опять с другого большого взрыва. Тогда говорят, что Вселенная замкнутая. Мы упоминаем сейчас эту точку зрения, поскольку оба примера моделирования, данные ниже в этой главе, дают результаты, которые влияют на оценки массовой плотности Вселенной и, следовательно, имеют отношение к важному космологическому вопросу. Наша Вселенная является замкнутой или открытой системой? [153, с. 96].

Возвращаясь теперь к веществу, которое выделилось как зародышевая галактика или протогалактика, следует представить себе это как вращающийся газовый шар, развивающийся как независимая система. Численное моделирование, несомненно, в конце концов, покажет, каким образом радиус, масса и момент импульса определяют последующую эволюцию протогалактики и приводят к различным типам наблюдаемых эллиптических, спиральных и барообразных спиральных галактик (см. [340] или [204, гл. 11]). Вероятно, что большие массивные системы с малым вращением развиваются в бесструктурные эллиптические галактики и что системы с большими моментами вращения развиваются в тонкие дискообразные галактики, но это еще не проверено с помощью моделирования. Область широко открыта для астрономов-вычислителей. Случай, изучаемый далее в этой главе, связан с эволюцией спиральных структур в

¹⁾ Один миллиард = 10^9 .

тонкую дискообразную галактику, поэтому сейчас мы дадим возможный сценарий для эволюции такой галактики.

Первые звезды, которые конденсируются из вращающейся протогалактики, были так называемые «старые» или звезды «сферической составляющей». Они конденсировались в то время, когда газ был все еще грубо сферическим, и многие звезды сформировались в больших шаровых скоплениях, состоящих из 10^4 — 10^6 звезд. Эти скопления распределены более или менее равномерно по сфере и образуют часть галактического гало. Звезды сферической составляющей гало имеют очень малую скорость вращения и осциллируют случайным образом от стороны к стороне сферического гало. В нашем моделировании предполагается, что они обеспечивают фиксированное внешнее гравитационное поле, через которое движется остальная часть галактики.

Остальная масса галактики является все еще газообразной и служит причиной вязкого затухания любых относительных движений. Когда газ сжимается под действием гравитационного притяжения, сохранение момента количества движения ограничивает сжатие в перпендикулярном к оси вращения направлении некоторым минимальным радиусом. Однако может иметь место сжатие, параллельное оси вращения в направлении к экваториальной плоскости и от нее. Эффекты вязкости в конце концов подавят осцилляции относительно экваториальной плоскости протогалактики и в результате останется состояние тонкого вращающегося газового диска с толщиной, примерно равной одной двенадцатой его радиуса. Звезды, родившиеся из этого вращающегося диска, являются относительно молодыми и называются «плоской составляющей» звезд. Именно в этой составляющей гравитационные неустойчивости создают красивые спиральные и барообразные формы, характеризующие тонкие дискообразные галактики.

Большинство вычислительных экспериментов посвящено изучению этих неустойчивостей. Некоторые звезды плоской составляющей образуются в малых открытых скоплениях, которые располагаются близко к плоскости вращающегося галактического диска. По этой причине они называются «рассеянные», или «галактические», скопления и обычно содержат несколько сотен звезд.

Упомянутую выше структуру тонкой дискообразной галактики можно увидеть на рис. 11.1 (вверху), который демонстрирует галактику M104, полученную с помощью 200-дюймового отражающего телескопа Хейла в обсерватории Mount Palomar. Эта галактика видна почти с ребра, и диск виден, главным образом, как тонкая темная линия, появляющаяся из-за поглощения света частицами пыли, которые являются важной компонентой диска. На диске можно также рассмотреть слабосветящиеся

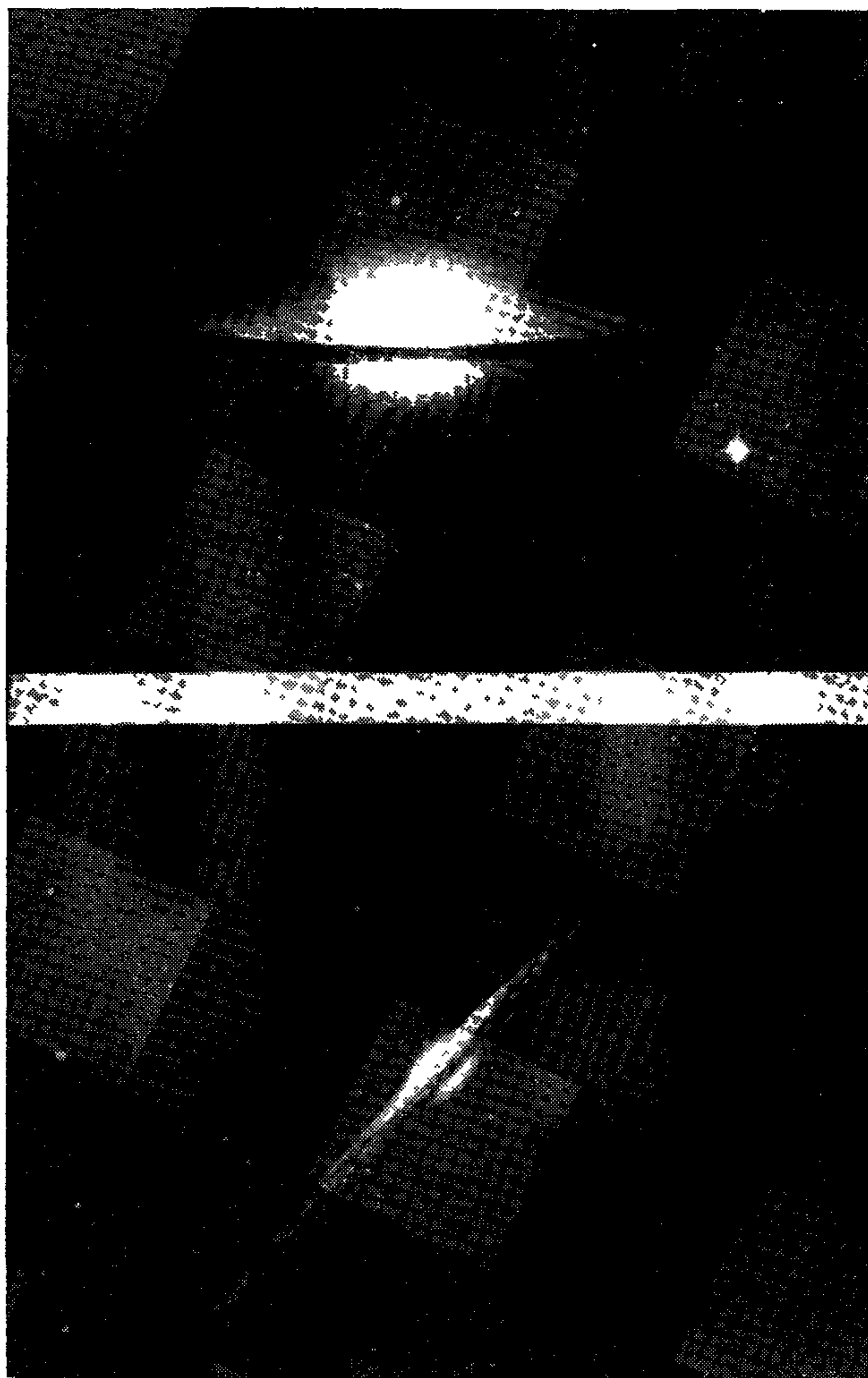


Рис. 11.1. Две дискообразные галактики, показывающие диффузное гало (M104, NGC 4594) (вверху) и центрально конденсированное гало (NGC 4565) (внизу). (С любезного разрешения Hale Observatories, 200-дюймовый телескоп Хейла.)

сгустки звезд, образующие очень плотные спирали или кольцеобразные структуры. Гало, которое является в этом примере большим и диффузным, выглядит как сферическое свечение. Гало не является однородным и во внешних областях вполне можно различить многочисленные расплывчатые сгущения звезд. Это шаровые скопления. Они образуют подобно остальной части гало сферическое распределение относительно центра.

Звезды нашей собственной Галактики, которые также появляются на пластинке, различаются с помощью горизонтальных и вертикальных линий, исходящих из их изображений. (Нарис. 11.1 (внизу) показана дискообразная галактика с малым центрально конденсированным гало.)

11.1.6. Численное моделирование звездных систем

Если мы рассмотрим компьютеры, имеющие время выполнения арифметической операции с плавающей запятой, сравнимое либо с CDC 7600, либо IBM 360/195 (у которых для оценки машинного времени можно использовать 1/2 мкс, затраченных на одну операцию с плавающей запятой), то для моделирования малых скоплений, состоящих не более чем из 1000 звезд, шаровых скоплений, состоящих из 10^5 звезд, и галактик, состоящих из 10^{11} звезд, наиболее подходящими являются методы РР, Р³М и РМ соответственно (см. гл. 1). Подобное деление, основанное на количестве взаимодействующих тел, можно провести для моделирования скоплений галактик различных размеров. Поэтому при моделировании звездных систем довольно удобно использовать все описанные в этой книге методы. Отличный обзор численного моделирования звездных систем дан Арсетом и Лекаром [7].

Метод РР для моделирования малых скоплений обширно разрабатывался Арсетом и др. в начале 1960 г. [1, 168, 194]. Сила, действующая на звезду, конечно, вычисляется прямым суммированием взаимодействий этой звезды со всеми другими звездами, и для получения силы, действующей на все N звезд из скопления, требуется приблизительно $5N^2$ операций. Поскольку мы имеем дело с меньшим чем 1000 количеством звезд, то может использоваться метод РР; он имеет важное преимущество, определяя силы с такой же точностью, что и арифметическая точность компьютера. Некоторые специальные методики интегрирования по времени, развитые для точного интегрирования звездных орбит в малых скоплениях, описываются в разд. 11.2.

Цель моделирования состоит в получении количественных результатов по эволюции звездных скоплений, включая частоту вылета звезд и частоту образования двойных звезд, которые можно использовать для определения возраста. Механизмы образования и распада двойных звезд изучаются Арсетом [5] с помощью системы, состоящей из 250 и 500 моделируемых звезд. Если начальные положения распределены случайно и равномерно внутри сферы, то на эволюцию оказывает доминирующее влияние сегрегация объектов по массе, которая приводит к образованию высокоплотного ядра из более тяжелых звезд, окруженного диффузным гало из более легких звезд. Эффект равномерного распределения энергии, обусловленный столкнове-

ниями, состоит в передаче высоких скоростей более легким звездам, причем некоторые из них вполне могут достигнуть скорости вылета и покинуть скопление. Другие описывают вытянутые орбиты и образуют гало. Центральное ядро теряет частицы посредством этого механизма и сокращается в размерах. Именно в этом высокоплотном ядре вероятность образования двойных звезд становится большой. Двойные звезды образуются либо при многократных столкновениях (например, тройные столкновения), либо посредством испарения звезд из ядра до тех пор, пока не останутся две звезды. Имеется долгосрочная тенденция к концентрации большей части энергии в одну тяжелую тесно связанную двойную звезду. Шкала по времени для этого развития составляет от 10 до 40 времен пересечения, зависящих от крутизны спектра масс. В этом механизме стока энергии более 50 % полной энергии поглощается одной тяжелой двойной звездой после 6—18 времен пересечения.

Очевидно, что методики, используемые для моделирования малых скоплений звезд, можно в равной степени применить к эволюции малых скоплений галактик. Арсет [1—3] сообщил о серии численных экспериментов для системы из 100 галактик как с первоначальным вращением, так и без него. Включение вращения укорачивает среднее время релаксации и вызывает значительную сплюснутость. Большая часть момента импульса переносится к гало, и сплюснутость остается «вмороженной». Найдено, что имеется значительная корреляция между величиной сплюснутости и моментом импульса. Модель Арсета использовалась также Эфстатиу и Джонсом [125] для изучения эффективности механизма приливного вращающего момента для генерации упомянутого выше галактического вращения во время расширения Вселенной.

Метод РР можно применить для моделирования систем, содержащих до 4000 тел, которые представляют звезды или галактики. В работах [6, 153] сообщается о моделях, содержащих от 2000 до 4000 галактик соответственно, которые используются для изучения образования скоплений в расширяющейся Вселенной. Подобные эксперименты на больших системах, включающих от 10 000 до 20 000 галактик, стали возможными с помощью алгоритма Р³М [189, 120] и предприняты в работе [124] (см. разд. 11.4). Обзор расчетов N тел по сгущиванию галактик и их влияние на космологию даны в [131].

Другие приближенные методы были развиты для больших систем, таких как шаровые скопления. Они основаны на взятии первых четырех моментов от распределения скоростей и получении системы уравнений гидродинамического типа Фоккера — Планка (см. [240, 241]). Основное упрощение состоит в том, что одночастичная функция распределения адекватно описывает такую систему (т. е. двойная корреляция, представляющая

образование двойных систем, игнорируется) и что двухчастичные столкновения определяют релаксацию. Уравнения Фоккера — Планка можно также моделировать статистическими методами, и два вида таких методов Монте-Карло были развиты в работах [169, 170, 350—353]. Было получено отличное согласие между расчетами N тел, гидродинамическими уравнениями и методами Монте-Карло [7]. Эти методы приближенного моделирования выходят за рамки книги, и с дальнейшими подробностями читатель может ознакомиться в упомянутых выше работах.

Моделирование бесстолкновительных галактик обсуждалось неоднократно, поскольку метод РМ допускает моделирование до нескольких сот тысяч частиц. Даже в таких ранних работах Линдблада [261, 262] с 50 частицами и более современной работе Острайкера и Пиблса [304] с 200 частицами использовался метод РР. Метод РМ ввел Хокни [178], используя модификацию программы, написанной для моделирования горячей газовой плазмы (см. [176, 177], а также гл. 9), и приспособил системы из 2000 частиц для моделирования на компьютере IBM 7090. Эта модель не имела физической интерпретации, поскольку каждая частица представляла массу в виде бесконечно длинного стержня, но она служила для демонстрации возможностей метода РМ. Распространение метода РМ на моделирование точечных масс, ограниченное движением в плоскости, было выполнено Холлом и Хокни [201], которые моделировали от 50 000 до 200 000 частиц на CDC 6600.

Распространение на трехмерный случай было выполнено Хокни и Браунриггом [188]. Две последние модификации стали возможными благодаря использованию алгоритма быстрого преобразования Фурье [88], который позволяет применять технику преобразования Фурье для получения гравитационного потенциала [180]. Дальнейшие улучшения были проведены для методов нахождения потенциала Джеймсом [217] и описаны в разд. 6.5.5. О других трехмерных моделях сообщалось в работах [200, 218, 284]. Использование моделей РМ для изучения спиральной структуры выбрано в качестве примера исследования в разд. 11.3.

Совершенно отличный подход для моделирования бесстолкновительной системы был выбран Миллером и Прендергастом [285]. Они дискретизировали четырехмерное фазовое пространство (x, y, v_x, v_y) в растр ячеек (256, 256, 63, 63). Каждая ячейка локализуется одним битом памяти и если этот бит установлен на единицу, то звезда имеется в этом местоположении фазового пространства. Аналогично, нуль означает отсутствие звезды. Таким образом, в этом методе две звезды не могут занимать одно и то же положение в фазовом пространстве. Уравнения Ньютона определяют движение через фазовое пространство

битов, которые представляют звезды. Следовательно, иногда этот метод описывается как проталкивание битов в фазовом пространстве и имеет очевидную привлекательность на таких компьютерах, как ICL с распределенным векторным процессором [134, 185], который можно запрограммировать на уровне битов. Преимущество состоит в том, что орбиты частиц в дискретизованном фазовом пространстве можно сделать достаточно точными, и в этом смысле они могут описываться как в точности бесстолкновительные. Однако очевидным ограничением является грубость дискретизации. Описанное моделирование также связано с периодическими граничными условиями [286], которые, очевидно, налагают на решение нереальную квадратную симметрию. Однако это ограничение можно легко устранить путем использования одного из методов нахождения потенциала изолированной системы, упомянутых выше и в гл. 6. Миллер [284] также сообщил о модели галактики на базе параллельного компьютера ILLIAC IV (182), на котором промоделировано 112 000 звезд на пространственной сетке $(64 \times 64 \times 64)$.

11.2. МАЛЫЕ СКОПЛЕНИЯ

Теперь мы опишем методы, пригодные для моделирования малых скоплений, содержащих менее 1000 звезд, основанные главным образом на работе Арсета.

11.2.1. Закон взаимодействия

Парный потенциал и парная сила между двумя частицами массы m , разделенными на расстояние r , задаются в следующем виде:

$$\varphi(r) = - \frac{Gm^2}{(r^2 + \epsilon^2)^{1/2}}, \quad (11.2a)$$

$$F(r) = - \frac{\partial \varphi}{\partial r} = - \frac{Gm^2 r}{(r^2 + \epsilon^2)^{3/2}}. \quad (11.2b)$$

Параметр ϵ вводится для ослабления взаимодействия на коротких радиусах. В этих выражениях ни потенциал, ни сила не стремятся к бесконечности при нулевом расстоянии, как в случае взаимодействия точечных масс. Потенциал стремится к конечной величине $-Gm^2/\epsilon$, а сила приближается к нулю подобно $-Gm^2 r/\epsilon^3$ при $r \rightarrow 0$. Такое взаимодействие характерно для двух облаков массы m и радиуса ϵ . Сначала ϵ было введено в моделирование малых скоплений галактик [1] и представляло конечный размер галактик. Однако вскоре было найдено, что необходимо ввести ненулевое значение ϵ в моделирование системы звезд, хотя моделирование точечных масс ($\epsilon = 0$) было бы лучшей физической моделью. Это происходит из-за необосно-

ванной вычислительной практики позволять любым переменным достигать исключительно больших значений, когда интегрирование по времени становится неточным и в этих случаях может произойти переполнение арифметического устройства.

Чтобы оценить максимальное значение ε , которое можно взять без нарушения физики явления, рассмотрим систему с полной энергией TE , достигающую конечного состояния, в котором все частицы, за исключением двух, улетели на бесконечность, не унося при этом энергии, а оставшиеся две частицы образовали тесную двойную систему с главной полуосью a . Тогда, поскольку частицы на бесконечности обладают нулевой энергией, полная энергия системы сосредоточена в одной двойной звезде, следовательно,

$$\frac{Gm^2}{2a} = -TE. \quad (11.2c)$$

Если ε выбрано меньше a , то компоненты самой тесной двойной звезды не могут сблизиться на расстояние, для которого закон взаимодействия (11.2b) значительно отклоняется от закона взаимодействия точечных масс.

11.2.2. Интегрирование по времени

Основная проблема при расчете малых скоплений заключается в точном интегрировании орбит звезд, особенно когда формируются тесные двойные системы. Поскольку приходится следить за относительно немногими частицами, то можно использовать схемы более высокого порядка, чем схема с перешагиванием (см. гл. 4). Было использовано много схем, но наиболее эффективными оказались методы, использующие полиномы высокого порядка [249]. Здесь мы суммируем то, что применялось Арсетом [4]. Поскольку звезды могут иметь сильно отличающиеся скорости и ускорения, каждой звезде приписывается ее собственный переменный шаг по времени, который выбирается для достижения заданной точности при расчете ее орбиты. Пара сил пересчитывается только тогда, когда включенные в рассмотрение звезды передвинулись с момента последнего вычисления и экономия на вычисление силы значительна.

При интегрировании звездных орбит используется полином четвертого порядка для силы, действующей на единицу массы каждой частицы в момент времени t , в виде

$$\mathbf{F} = \mathbf{F}_0 + \mathbf{B}t_r + \mathbf{C}t_r^2 + \mathbf{D}t_r^3 + \mathbf{E}t_r^4 \quad (11.3a)$$

где t_r — время после последнего расчета силы, действующей на рассматриваемую частицу. Если это происходило в момент t_0 , то

$$t_r = t - t_0, \quad (11.3b)$$

\mathbf{F}_0 — сила, вычисленная для частицы в момент времени $t = t_0$, а векторные значения \mathbf{B} , \mathbf{C} и \mathbf{D} вычисляются для того, чтобы согласовать \mathbf{F} с силой, полученной из трех предыдущих вычислений. Коэффициент \mathbf{E} не определяется до следующего вычисления силы в нижнем цикле (т. е. он действует как корректирующий член).

Каждая частица i несет запись времени t_i последнего расчета действующей на нее силы и текущее значение индивидуального шага по времени Δt_i . Частицы перемещаются на один шаг по времени, как следует ниже:

1. Находим следующую частицу, скажем α , для которой необходимо рассчитать силу. Проводим поиск

$$\min_i (t_i + \Delta t_i), \quad (11.4a)$$

где $\alpha = i$ в минимуме.

2. Находим время продвижения

$$t^* = t_\alpha + \Delta t_\alpha. \quad (11.4b)$$

3. Изменяем положение всех частиц, кроме α , на новое время t^* , используя только два первых члена силового полинома (11.3a):

$$\mathbf{F} = \mathbf{F}_0 + \mathbf{B}t_r.$$

Два интегрирования по времени дают изменение положения частицы в виде

$$\Delta \mathbf{r} = \dot{\mathbf{r}}_0 t_r + \frac{1}{2} \mathbf{F}_0 t_r^2 + \frac{1}{6} \mathbf{B} t_r^3. \quad (11.5)$$

4. Предсказываем новое положение частицы α с помощью двух интегрирований по времени кубического выражения для действующей на нее силы

$$\mathbf{F} = \mathbf{F}_0 + \mathbf{B}t_r + \mathbf{C}t_r^2 + \mathbf{D}t_r^3. \quad (11.6)$$

5. Вычисляем силу на единицу массы, действующую на частицу α , с помощью суммирования вкладов от всех других частиц, используя их положения, полученные из шага 3:

$$\mathbf{F}_\alpha = - \sum_{\substack{j=1 \\ j \neq \alpha}}^N \frac{Gm_j (\mathbf{r}_\alpha - \mathbf{r}_j)}{(|\mathbf{r}_\alpha - \mathbf{r}_j|^2 + \varepsilon^2)^{3/2}}. \quad (11.7)$$

6. Используем новое значение силы, действующей на частицу α , для оценки коэффициента \mathbf{E} в квадратичном выражении для силы (11.3a). Корректируем новое положение частицы α посредством суммирования в перемещении, обусловленном коэффициентом \mathbf{E} .

7. Вычисляем новую скорость частицы α посредством одного интегрирования по времени квадратичного выражения для силы.

Если $\dot{\Delta r}_0$ есть изменение скорости, то

$$\dot{\Delta r}_0 = F_0 t_r + \frac{1}{2} B t_r^2 + \frac{1}{3} C t_r^3 + \frac{1}{4} D t_r^4 + \frac{1}{5} E t_r^5. \quad (11.8)$$

8. Вычисляем время до следующего расчета силы для частицы α из выражения

$$\Delta t_\alpha = \left[\eta \left(\frac{F_0 + \dot{F}_0 \Delta t}{\frac{1}{6} \ddot{F}_0 + \frac{1}{24} F_0^{(IV)} \Delta t} \right) \right]^{1/3} = \left[\eta \left(\frac{F_0 + B \Delta t}{D + E \Delta t} \right) \right]^{1/3}, \quad (11.9)$$

где Δt — величина последнего шага по времени, выбранного частицей. В качестве предохранения от неустойчивости не допускается, чтобы новый шаг по времени превышал в $\sqrt{2}$ раз значения старого шага по времени. Уравнение (11.9) вычисляется отдельно для каждой компоненты векторов, и выбирается минимум. Точки (или римские цифры в верхнем индексе) обозначают порядок производной по времени.

Этот цикл точно переместил одну частицу (номер α) и перешел от времени t к t^* . Теперь цикл повторяется на этапе 1 выбора следующей частицы для продвижения, и ею может быть опять частица α . Требуемая память равна $28N$, и рекомендуется точность в 10 десятичных разрядов (это означает двойную точность на IBM 370) для координат, скоростей и времени. Сила вычисляется с одинарной точностью на IBM 370. Количество операций, требуемых для однократного исполнения цикла, т. е. для продвижения одной частицы на один шаг, составляет примерно $250N$. Полное необходимое время будет зависеть от требуемого числа перемещений звезд на одну секунду, которое в свою очередь будет зависеть от природы орбит звезд. Полное машинное время для воспроизведения заданного физического времени t_{ph} пропорционально $N^2 t_{ph}$. Параметр отбора временного шага n определяется эмпирически, и значение, близкое к 2×10^{-4} , приводит к удовлетворительному интегрированию в отсутствие экстремально тесных сближений. Относительные ошибки в расчете энергии для двойной звезды с эксцентриситетом 0.92 составляют около 10^{-5} .

Улучшения упомянутой выше основной схемы интегрирования N тел были проведены в работе [13], где включены эффекты близких и дальних объектов отдельно и на различных временных масштабах. В этом подходе интегрирование ускоряется, и метод является более эффективным, чем основной метод при $N > 50$. Используя улучшенную версию этого метода, Арсет получает модель, способную описать движение системы из 4000 галактик.

Если связанные пары звезд, называемые двойными звездами, образуются во время эволюции звездной системы, как это обычно бывает, то в рассмотренном выше методе интегрирования становятся преобладающими расчеты очень коротких

временных шагов, необходимых для точного определения орбит двойных систем. Эффективная и всеобъемлющая программа моделирования должна поэтому делать специальное обеспечение для распознавания таких двойных звезд и для расчета их орбит. Подходящая методика для специального рассмотрения орбит двойных звезд состоит в регуляризации двух тел Леви-Чивита, обобщенной на трехмерный случай в работе [230]. Она была использована для проблем трех тел Петерсом [310] и модифицирована для проблем N тел Арсетом [4].

В процедуре регуляризации уравнения движения записываются отдельно для движения центра масс и относительного движения пары. В выражении для движения центра масс исключается аналитически очень большая сила внутреннего взаимодействия между компонентами пары. Собственная сила взаимодействия преобладает в уравнениях относительного движения, и внешние силы, приложенные к значительно разделенным двойным звездам, действуют как относительно малое возмущение. Введение фиктивной временной переменной τ , подчиняющейся дифференциальному соотношению $d\tau = dt/r$, где r — расстояние разделения пары, устраняет бесконечность в уравнениях движения Ньютона, обусловленную бесконечной силой, соответствующей нулевому разделению. На языке переменной фиктивного времени уравнения относительного движения являются несингулярными и интегрирование их по времени можно выполнить с равным шагом $\Delta\tau$. По мере того как разделение уменьшается, эти временные шаги, конечно, соответствуют все меньшим и меньшим интервалам Δt физического времени. Около пятидесяти временных шагов выбрано для полного оборота в относительном движении, период которого равен двум периодам исходных нерегуляризованных уравнений движения Ньютона. Описанная ранее процедура четвертого порядка интегрирования по времени используется для интегрирования регуляризованных уравнений движения.

Необходимо установить критерии допуска пары частиц для рассмотрения с помощью регуляризации двух тел и для ограничения такого рассмотрения. Частица опознается для регуляризации, если ее временной шаг уменьшается и становится меньше $\Delta\tau_{\min}$. Затем выполняется поиск других частиц в окрестности и принимается другой компаньон пары, если их разделение меньше R_{\min} и вклад в силу превосходит все другие. Последнее условие необходимо для гарантии того, что мы не имеем дело, например, с триплетом. Регуляризация отдельной пары прекращается, как только сила внутреннего взаимодействия пары перестает быть основной силой, приложенной к обоим партнерам, и члены пары могут менять своих партнеров в результате этого теста. Это случится, если двойная звезда разрушается при прохождении третьей частицы.

11.3. СПИРАЛЬНЫЕ ГАЛАКТИКИ

11.3.1. Теории спиральной структуры

Численное моделирование спиральных структур в тонких дискообразных галактиках дает хороший пример конструктивного взаимодействия между моделированием, теорией и наблюдением, которое может иметь место во время исследования сложной проблемы. В морфологической классификации наблюдаемых галактик, данной Хабблом [205, 207] и показанной на рис. 11.2, эллиптическая, спиральная и спиральная с баром формы, наблюдавшиеся на небе, приведены в последовательности, основанной на постепенном изменении формы. Хотя Хаббл применил термины «ранний» и «поздний» в описании форм в левой и правой частях его классификации, все еще остается открытым вопрос, подразумевает ли классификация временную эволюцию или форма галактики определяется в основном начальными условиями (т. е. моментом импульса, энергией и массой), заданными, когда протогалактика конденсировалась в расширяющейся Вселенной (см. разд. 11.1.5). Спиральные галактики образуют большую долю от наблюдаемых галактик, и, следовательно, считается, что они являются долгоживущими структурами с временами жизни, сравнимыми с временем жизни галактики и, следовательно, самой Вселенной. Поэтому типичная оценка времени жизни спирали должна составлять много десятков галактических вращений.

Центральный вопрос, обращенный к любой теории эволюции галактики, состоит в том, как долгоживущие спиральные особенности могут сохраняться, когда звезды на различных радиусах вращаются со значительно отличающимися периодами вращения. Очевидно, что если ветви спирали всегда содержат одни и те же звезды, то это дифференциальное вращение будет превращать любую конфигурацию во все более плотно закрученную спираль в пределах нескольких вращений. Такие «ма-

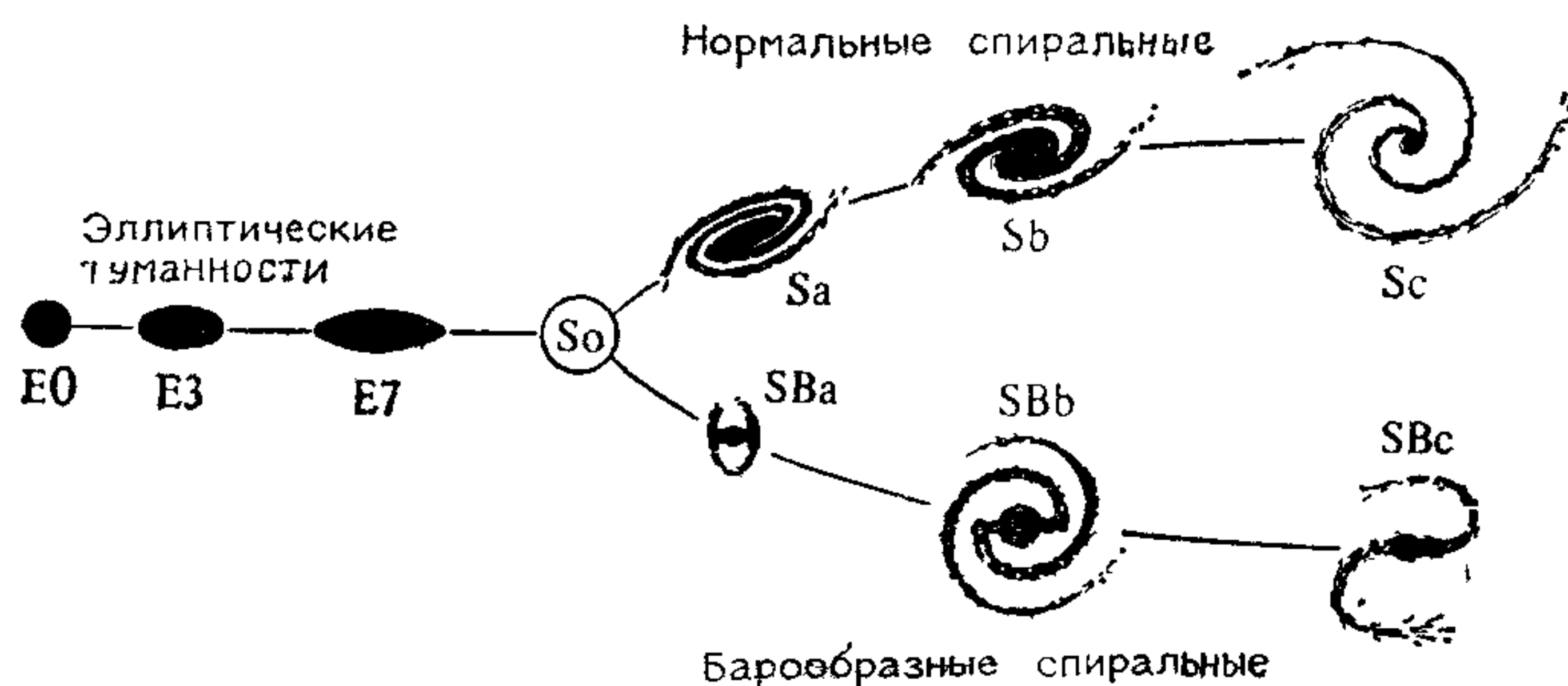


Рис. 11.2. Морфологическая классификация галактик Хаббла [207]. (С любезного разрешения Yale University Press.)

териализованные спирали» не могут быть долгоживущими, и более удовлетворительная точка зрения принадлежит Лину и его коллегам, которые рассматривают спиральные особенности как волны плотности в галактике, возникающие полностью из-за гравитационных эффектов. Эти волны плотности могут распространяться со скоростями, отличными от скоростей дифференциально вращающихся звезд, во многом таким же образом, как звуковые волны движутся через молекулы воздуха. Следовательно, проблема закручивания не возникает, поскольку звезды, составляющие заданную спираль, непрерывно меняются.

Теория спиральной структуры, основанная на волнах плотности, была дана в сериях статей (см., например, [242, 258, 259, 273, 274]), исчерпывающий обзор которых дан в [372]. Разъяснительная статья [257], вероятно, является лучшей исходной ссылкой. Теория спиральной структуры Лина — Шу основана на асимптотическом расчете, который показывает, каким образом спиральные моды с определенными характеристиками (т. е. характерная скорость, питч-угол) являются самосогласованными решениями для дифференциально вращающейся системы газ — звезды. Эта теория применяется к плотно намотанным спиральям с питч-углами, меньшими 30° , и дает дисперсионное соотношение и скорость роста волн.

Возможное объяснение механизма образования волн плотности, данное Линден-Бэллом и Калнайсом [271], основано на том факте, что запаздывающая спиральная ветвь создает вращающий момент, который переносит момент импульса наружу. Для большинства галактик, в которых угловая скорость уменьшается с радиусом, такой перенос означает понижение полной энергии вращения и, следовательно, увеличение энергии случайного движения при сохранении энергии. Это последнее изменение энергии представляет увеличение энтропии системы, и теория обращается к утверждению, как в термодинамике, что изолированная динамическая система стремится развиваться в направлении возрастания энтропии. Другой механизм, основанный на резонансном взаимодействии орбит частиц с барообразным возмущением, был дан Линден-Бэллом [270]. Поскольку динамические эксперименты нельзя выполнить в реальных галактиках, роль численного моделирования состоит в обеспечении экспериментов с вычислительными моделями, с которыми можно сравнить упомянутые выше теории. Далее будет описана серия таких экспериментов.

11.3.2. Модель

Идеализированная физическая модель тонкой дискообразной галактики показана на рис. 11.3. Тонкий диск (массы m_d), состоящий из звезд, с параметром расстояния a_d , определяющим

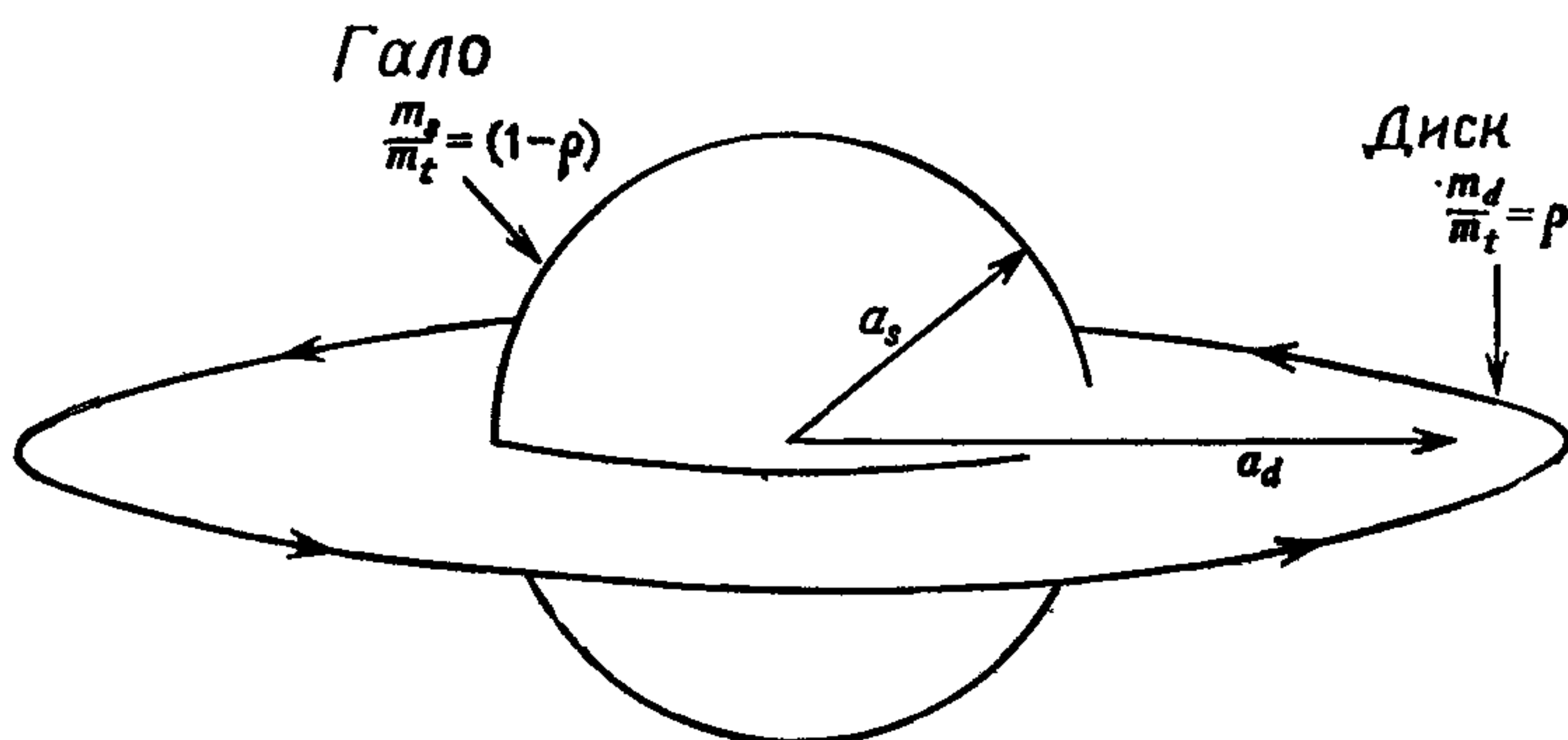


Рис. 11.3 Идеализированный вариант структуры дискообразной галактики, показывающий компоненты диска и гало и переменные $\rho = m_d/m_t$ и $\alpha = a_d/a_s$; $m_t = m_d + m_s$ — полная масса. ([185], с любезного разрешения Contemporary Physics, © Taylor and Francis Ltd)

его радиальное протяжение, вращается через невращающееся сферически симметричное гало (массы m_s) с размером, характеризующимся параметром расстояния a_s . Параметр расстояния может быть радиусом распределения звезд, если оно конечно, или параметром, определяющим скорость затухания плотности с расстоянием. Точное значение будет зависеть от аналитической формы распределения масс, но в сущности плотность звезд можно считать нулем на расстояниях, больших чем несколько характерных расстояний. Важными физическими параметрами являются $\rho = m_d/(m_d + m_s)$ — доля полной массы в диске и $\alpha = a_d/a_s$ — степень, до которой гало конденсируется к центральному балджу. Мы приведем результаты для тяжелого диффузного гало ($\rho = 0.2$, $\alpha = 1$) и легкого центрально сконденсированного гало ($\rho = 0.7$, $\alpha = 3$).

Чтобы построить вычислительную модель, мы далее упростим ситуацию, рассматривая звезды, образующие гало, как неподвижные, и представим их влияние только в виде неподвижного внешнего гравитационного поля, сквозь которое начинает вращаться диск. Звезды, образующие гало, обладают малым вращательным движением, и они не образуют различимую структуру, так что это приближение является хорошим в случае наличия малого обмена энергией между гало и диском. Для того чтобы подкрепить это утверждение, Хол [200] провел некоторые расчеты, в которых гало и диск представляются 50 000 движущимися частицами каждый. Он обнаружил, что гало остается по существу неподвижным, подтвердив справедливость данного выше приближения.

Вычислительная модель, использованная Хокни и Браунриггом [188], иллюстрируется на рис. 11.4 и состоит обычно из 20 000—25 000 моделируемых макрозвезд, движущихся по сетке

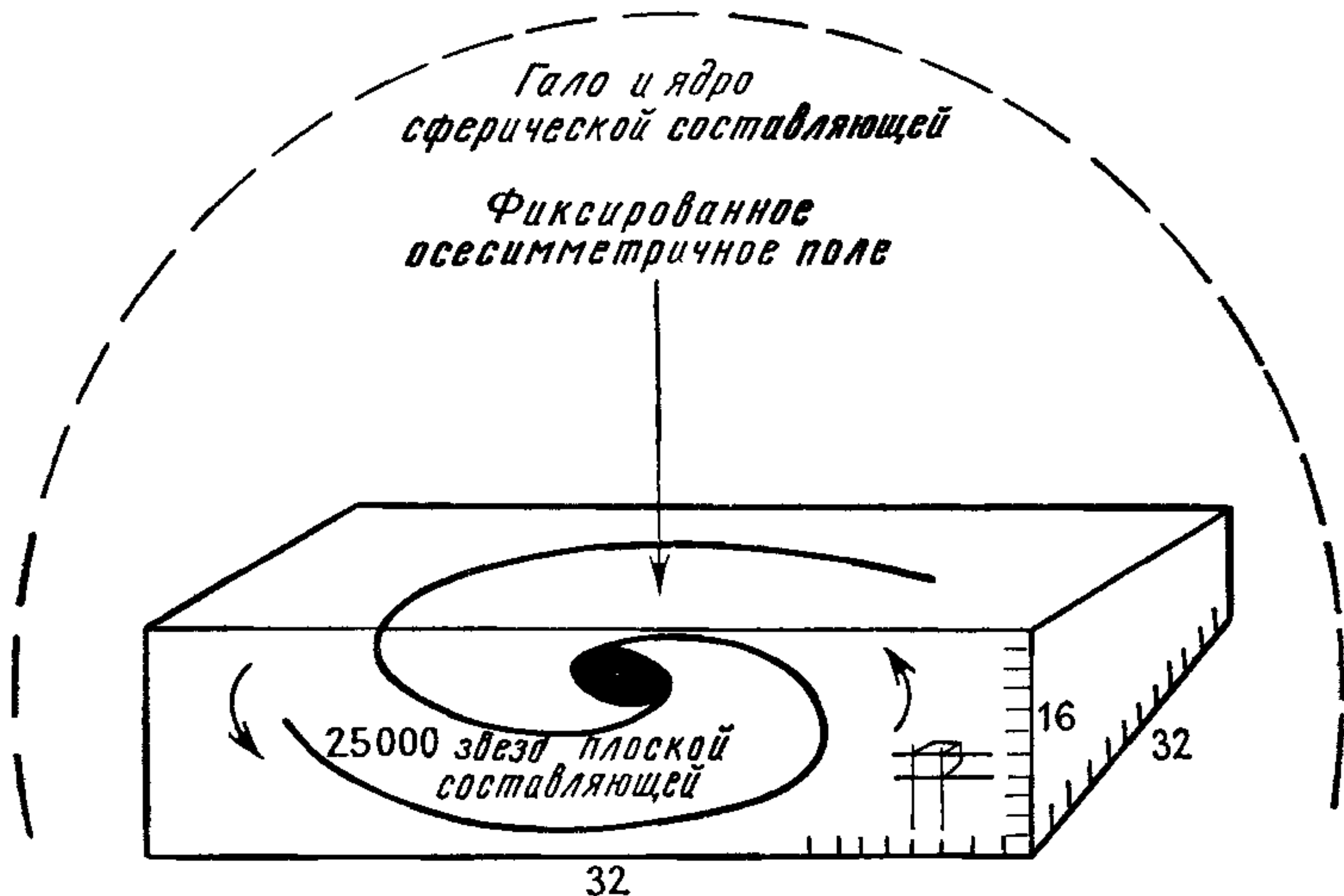


Рис. 11.4. Трехмерная вычислительная модель тонкой дискообразной галактики, разработанная Хокни и Браунриггом [188] и демонстрирующая типичную используемую сетку.

($32 \times 32 \times 16$), на которой рассчитывается гравитационный потенциал. Используется процедура моделирования РМ (см. гл. 1) с распределением массы и интерполяцией поля по схеме СИС. Шаг по времени на IBM 360/195 занимает около 10 с, причем делится следующим образом: нахождение потенциала 4 с, интегрирование по времени параметров частицы 2 с, проведение анализа спирали 4 с. Расчет проводится полностью в основной высокоскоростной памяти компьютера и занимает от 1 до 1.2 Мбайт памяти. Аналогичная модель развита на CDC 7600 Джеймсом и Селлвудом [218], которые используют улучшенный метод нахождения потенциала Джеймса [217]. Получается отличное согласие между этими двумя запрограммированными независимо моделями [218]. Ясно, что бесстумная модель «частица — сетка» (QPM), использующая облако треугольной формы (TSC) для распределения массы и придающая форму функции Грина для понижения далее сеточных флуктуаций (см. гл. 5), имеет приложение к галактическим моделям РМ. Любые новые модели должны использовать эти методы.

Необходимо выбрать три параметра моделирования: приращение пространственной сетки H , шаг по времени DT и полное число частиц N . Пространственная сетка должна быть связана с требуемым пространственным разрешением, а шаг по времени — с характерными физическими частотами; число частиц должно быть достаточным для создания бесстолкновительной системы. Ясно, мы хотим сделать H , DT настолько малым, а N настолько большим, насколько возможно при совместности с

требованиями частной задачи и объемом доступного машинного времени. Как всегда, тут необходим компромисс между точностью моделирования и стоимостью затраченного на расчет машинного времени.

11.3.3. Выбор временного шага и размера сетки

Типичную спиральную галактику, такую, как наша собственная, можно рассматривать как имеющую радиус 20 килопарсек (кпк) ($1 \text{ кпк} = 3.3 \times 10^3$ световых лет $= 3.1 \cdot 10^{16}$ км) и особенности, такие как ширина спиральной ветви, в пределах от одного до десяти килопарсек в зависимости от плотности спирали. Поэтому используется наиточнейшая пространственная сетка, допустимая с точки зрения расхода машинного времени. Сетка $(32 \times 32 \times 16)$ с интервалом $H = 25$ кпк допускает 16 точек поперек диаметра галактики и границу из 8 точек между краем галактики и краем вычислительной области. Поскольку особенность можно адекватно разрешить, если только имеется несколько сеточных точек поперек ее, то ясно, что эта сетка будет разрешать только более широкие спиральные особенности галактики. Пространственное разрешение в галактической плоскости можно удвоить посредством использования сетки $(64 \times 64 \times 8)$ и увеличить далее с помощью применения алгоритма Р³М для добавления межсеточного разрешения. Это возрастание пространственного разрешения увеличивает машинное время по нескольким причинам. Во-первых, время нахождения потенциала будет удваиваться, поскольку удваивается количество сеточных узлов. Во-вторых, найдено, что интенсивность столкновения увеличивается по мере того, как понижается H , потому что максимальная сила взаимодействия между парой частиц возрастает (см. разд. 11.3.4). Следовательно, должно использоваться больше частиц для удержания моделирования в бесстолкновительном режиме. Это опять увеличивает машинное время.

Чтобы определить соответствующие шкалы по времени, необходимо рассмотреть движение звезды, когда оно возмущено на фоне круговой равновесной орбиты. Если мы используем цилиндрические координаты (r, θ) , лежащие в плоскости галактического диска, то уравнение движения звезды на радиусе r имеет вид

$$\ddot{r} - r\dot{\theta}^2 = F(r), \quad (11.10)$$

в котором радиальное и центробежное ускорения в левой части приравниваются силе $F(r)$, действующей на единицу массы и обусловленной гравитационным притяжением. Если мы рассматриваем осесимметричное распределение звезд, то эта сила не

зависит от осевой координаты θ и момент импульса на единицу массы h сохраняется:

$$r^2\dot{\theta} = h = \text{const.} \quad (11.11)$$

Исключая $\dot{\theta}$ из (11.10) и (11.11), получаем

$$\ddot{r} - h^2/r^3 = F(r). \quad (11.12)$$

Установившееся решение уравнения (11.12) представляет собой круговую орбиту с $r = r_0$ и $\dot{r}_0 = 0$, для которой

$$-h^2 = r_0^3 F(r_0), \quad (11.13)$$

откуда, исключая h из (11.12) и (11.13), получаем

$$\ddot{r} + F(r_0) \left(\frac{r_0}{r}\right)^3 = F(r). \quad (11.14)$$

Если теперь представить возмущение r' относительно r_0 так, что $r = r_0 + r'$, и провести разложение Тейлора до двух членов относительно $r = r_0$, то

$$\ddot{r}_0 + \ddot{r}' + F(r_0) r_0^3 \left(\frac{1}{r_0^3} - \frac{3}{r_0^4} r'\right) = F(r_0) + r' \frac{dF}{dr} \Big|_{r=r_0}. \quad (11.15)$$

Учитывая, что $\dot{r}_0 = 0$, и исключая члены с $F(r_0)$, получаем

$$\ddot{r}' = -\omega_e^2 r', \quad (11.16)$$

где

$$\omega_e = \left(-\frac{3F(r_0)}{r_0} - \frac{dF}{dr} \Big|_{r=r_0} \right)^{1/2}. \quad (11.17)$$

Уравнение (11.16) является уравнением для простого гармонического колебания, а частота ω_e называется эпициклической частотой. В уравнении (11.17) как $F(r_0)$, так и ее производные отрицательны, поэтому эпициклическая частота — вещественное число. Движение возмущенной звезды можно визуализировать как круговое вращение с частотой ω_e относительно ведущего центра, который движется вдоль невозмущенной круговой орбиты. Для устойчивого интегрирования по схеме с перешагиванием мы должны потребовать, чтобы

$$\begin{aligned} \omega_e DT &< 2 \text{ для устойчивости,} \\ \omega_e DT &< 0,25 \text{ для разумной точности.} \end{aligned} \quad (11.18)$$

Если $\Omega(r)$ — угловая скорость звезд на радиусе r , то баланс центробежной и гравитационной сил дает

$$F(r) = -r\Omega^2, \quad (11.19)$$

$$\omega_e = \left(4\Omega^2 + 2\Omega r \frac{d\Omega}{dr} \right)^{1/2}. \quad (11.20)$$

Если галактика вращается как твердое тело, то $\Omega(r) = \Omega_0$ есть постоянная и

$$\omega_e = 2\Omega_0. \quad (11.21)$$

С другой стороны, если звезды движутся с постоянной азимутальной скоростью v_0 , то

$$\Omega = v_0/r, \quad d\Omega/dr = -v_0/r^2, \quad \omega_e = \sqrt{2}\Omega = \sqrt{2}v_0/r. \quad (11.22)$$

Кривая вращения $\Omega(r)$ большинства галактик лежит между этими двумя экстремумами, и можно заключить, что эпициклическое движение будет удовлетворительно интегрироваться с тем же временным шагом, с которым удовлетворительно интегрируется круговое орбитальное движение. Наиболее неприятная ситуация может возникать в случае почти постоянной азимутальной скорости для малых радиусов, когда уравнение (11.22) предсказывает очень большую эпициклическую частоту. По этой причине иногда может быть необходимо отказаться от интегрирования звездных орбит внутри определенного расстояния от галактического центра до звезды — внутри этого расстояния модель рассматривается как часть неподвижного гало. Поскольку в центральной области дискообразных галактик структура неразличима, это приближение не искажает физики явления. При типичном моделировании следовало бы выбирать от 100 до 200 временных шагов на один средний период вращения, поэтому $\omega_e DT \simeq \Omega DT \simeq 0.05$.

Более строгое ограничение на временной шаг, вероятно, должно возникать из требования, что частица за один шаг по времени не должна перемещаться на большее расстояние, чем необходимое пространственное разрешение. Если это произойдет, то пространственные особенности, конечно, будут расплываться. Если мы пытаемся достигнуть разрешения, сравнимого с размером сетки H , то это означает, что частица не должна перемещаться на расстояние, большее, чем размер сеточной ячейки за один шаг по времени. Поскольку метод РМ не может разрешать изменения на меньшем, чем ячейка, размере, то также верно, что нет выигрыша в разрешении, когда частица перемещается на расстояние, существенно меньшее, чем сеточная ячейка за один временной шаг. Следовательно, естественный выбор временного шага, вероятно, должен быть таким: $DT = H/v_0$, где v_0 — средняя скорость частицы. Заметим, что это условие означает, что временной шаг уменьшается вдвое, если сеточная ячейка уменьшается вдвое, что далее увеличивает вычислительную стоимость возросшего пространственного разрешения. Если применить этот критерий к галактике, имеющей 8 ячеек по радиусу, то увидим, что 100 шагов на вращение соответствуют двум шагам на прохождение ячейки, и, следовательно, условие удовлетворяется.

11.3.4. Время столкновения и число частиц

Число частиц, используемых в моделировании, должно быть достаточным, чтобы эффекты парных столкновений были не существенны при продолжительности вычислительного эксперимента. Для предсказания интенсивности парных столкновений как функции параметров N , DT и N вычислительной модели удовлетворительной теории не существует. Однако мы сделаем простую оценку, основанную на импульсном приближении. Если две частицы массы m проходят друг мимо друга с прицельным расстоянием p и относительной скоростью v_{rel} , как показано на рис. 11.5, то величину перпендикулярной скорости v_{\perp} , приобретенную во время парного столкновения, можно получить приравниванием перпендикулярного изменения количества движения перпендикулярному импульсу (сила, умноженная на время приложения). В качестве времени столкновения выбирается время прохождения расстояния $2p$; следовательно,

$$mv_{\perp} = Ft = F(p) \frac{2p}{v_{\text{rel}}}, \quad (11.23)$$

и отклонение $\delta\theta$, вызванное столкновением, равно

$$\delta\theta = \frac{v_{\perp}}{v_{\parallel}} = F(p) \frac{2p}{mv_{\text{rel}}v_{\parallel}}, \quad (11.24)$$

где v_{\parallel} — абсолютная скорость частицы.

Если мы считаем, что диск очень тонкий и имеет поверхностную плотность n_2 , то поток частиц с прицельными расстояниями, лежащими между p и $p + dp$, равен $2n^2v_{\text{rel}}dp$. Это выражение представляет собой число парных столкновений, возникающих за одну секунду, и если эти столкновения предполагаются статистически независимыми, то можно рассчитать среднеквадратичное полное отклонение $\langle \Delta\theta^2 \rangle$ за время Δt , суммируя квадраты от отдельных отклонений и интегрируя по всем

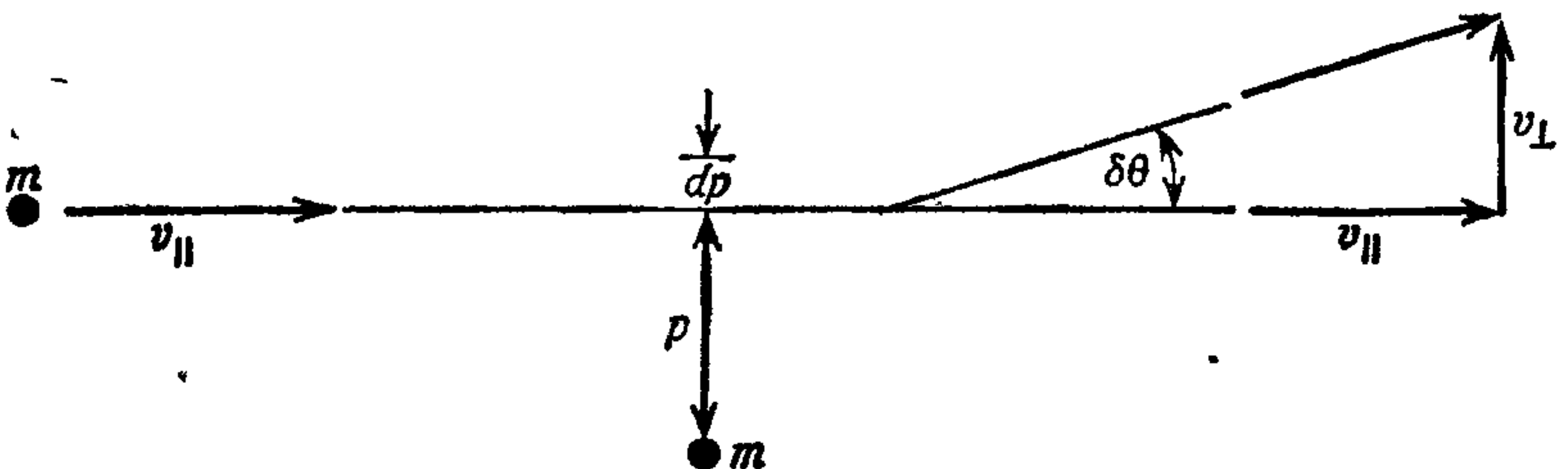


Рис. 11.5. Переменные, используемые в импульсном приближении,

прицельным расстояниям

$$\begin{aligned} \langle \Delta\theta^2 \rangle &= \int (2n_2 v_{\text{rel}} dp) (2Fp/mv_{\text{rel}}v_{\parallel})^2 \Delta t = \\ &= \frac{8n_2}{m^2 v_{\text{rel}}} \frac{\Delta t}{v_{\parallel}^2} \int F^2 p^2 dp. \end{aligned} \quad (11.25)$$

Время столкновения τ_{coll} определяется как время, за которое среднее отклонение достигает 90° или $\pi/2$ рад. Следовательно,

$$\tau_{\text{coll}} = \frac{\pi^2 m^2 v_{\text{rel}} v_{\parallel}^2}{32n_2} \left(\int F^2 p^2 dp \right)^{-1}. \quad (11.26)$$

При вычислении (11.26) необходимо принять во внимание, что сетка сглаживает взаимодействие между частицами на коротком радиусе действия и что моделируемые макрозвезды взаимодействуют подобно облакам массы шириной W . Взаимодействие между двумя такими облаками приблизительно равно

$$F(p) = \begin{cases} -\frac{Gm^2}{W^3} p, & p < W, \\ -\frac{Gm^2}{p^2}, & p \geq W. \end{cases} \quad (11.27)$$

Подставляя это взаимодействие в уравнение (11.26) и выполняя интегрирование от нуля до ρ_{max} , получаем

$$\tau_{\text{coll}} = \frac{5\pi^2}{192} \frac{v_{\text{rel}} v_{\parallel}^2 W}{n_2 G^2 m^2} \left[1 - \frac{5}{6} \left(\frac{W}{\rho_{\text{max}}} \right) \right]^{-1}. \quad (11.28)$$

Время столкновения представляет собой локальное свойство распределения частиц и должно рассчитываться как функция радиуса для конкретного используемого распределения плотности и конкретных подходящих значений относительной и абсолютной скоростей. Чтобы получить типичные величины, рассмотрим частицы на краю диска с постоянной плотностью радиуса R с азимутальной скоростью v_{rot} . Тогда приближенно имеем

$$\frac{GmN}{\rho R^2} = \frac{v_{\text{rot}}^2}{R}, \quad (11.29a)$$

где ρ — доля массы в диске, а период вращения равен

$$\tau_{\text{rot}} = \frac{2\pi R}{v_{\text{rot}}}. \quad (11.29b)$$

Мы также принимаем

$$\rho_{\text{max}} = R, \quad v_{\text{rel}} = \alpha v_{\text{rot}}, \quad v_{\parallel} = v_{\text{rot}} \quad (11.29c)$$

и замечаем, что $N = n_2 \pi R^2$. Тогда

$$\frac{\tau_{\text{coll}}}{\tau_{\text{rot}}} = \frac{5\pi^2}{384} \frac{\alpha N}{\rho^2} \left(\frac{W}{R}\right) \left[1 - \frac{5}{6} \left(\frac{W}{R}\right)\right]^{-1}. \quad (11.30)$$

Ширину частицы, как и в случае моделирования плазмы (см. гл. 9), можно выбрать равной парному разделению при максимальном значении силы взаимодействия. Как для схемы NGP, так и для схемы SIC имеем $W = H$; с другой стороны, если используются методы распределения заряда TSC или QPM, величина W будет большей (от $2H$ до $3H$). Если допустить, что используется метод SIC, и взять типичные значения для других переменных $\rho = 0.5$, $R = 8H$ и $\alpha = 0.1$, то

$$\frac{\tau_{\text{coll}}}{\tau_{\text{rot}}} = \frac{5\pi^2}{384} \frac{4}{80} \frac{N}{0.9} = \frac{N}{140}. \quad (11.31)$$

Поэтому мы заключаем, что в типичном случае для каждого бесстолкновительного вращения требуется около 140 частиц. Если выберем $N = 25\,000$, как это сделано в работах [180, 218], то получим время столкновения, равное примерно 200 вращений. Это означает, что после 200 вращений движения частиц полностью определяются нефизическими столкновениями макрочастиц, поскольку среднее отклонение, обусловленное этой причиной, будет составлять 90° . Более реалистично рассматривать систему как бесстолкновительную только для одной десятой этого времени, т. е. примерно для 20 вращений, когда — поскольку среднее значение квадрата отклонения пропорционально времени, — среднее отклонение будет составлять около 30° . К счастью, это соответствует моменту окончания большинства численных экспериментов. Если требуется проведение численных экспериментов на больших временах, то должно использоваться значительно больше частиц. Уравнение (11.30) также показывает, что время столкновения возрастает, если доля ρ массы в диске уменьшается. Следовательно, для задач с тяжелым гало проблема вычислительных столкновений несколько упрощается.

Хотя сделанная выше оценка времени столкновения является лишь приближенной, она показывает важность удержания полного числа частиц N настолько большим, насколько это возможно. Заметим, что, поскольку $W = H$ как для NGP, так и для SIC, время столкновения, рассчитанное по (11.30), одно и то же. Следовательно, метод NGP должен иметь преимущество, поскольку он дает возможность промоделировать большее количество частиц за то же самое машинное время. Моделирование с несколькими миллионами частиц будет возможным при использовании метода на параллельных компьютерах. Однако большинство авторов используют интерполяцию SIC или ее модификацию из за превосходного сохранения энергии.

Имеется немного измерений времени столкновения при моделировании галактик, но всестороннего изучения, сравнимого с исследованием Хокни [181] для электростатических моделей плазмы (см. гл. 9), не проведено. Хол [197] измерил время релаксации энергии при моделировании дискообразной галактики и нашел, что оно равно 560—1700 периодов вращения для системы, состоящей из 100 000 макрозвезд, разделенной на две различающиеся по массе группы. Время релаксации энергии для системы частиц с различной массой есть время, необходимое для достижения средней кинетической энергией, приходящейся на одну частицу в каждой различающейся по массе группе, значения средней кинетической энергии, приходящегося на одну частицу для всей системы. Равное распределение энергии между отдельными группами, различающимися по массе, возможно из-за обмена энергией во время парного столкновения, и поэтому ожидается, что время релаксации энергии будет того же порядка величины, что и время столкновения, определенное выше. Если условия для эксперимента Хола подставляются в уравнение (11.30), то получаем около 400 вращений ($\rho = 1.0$, $N = 10^5$, $\alpha = 1$, $W/R = 1/32$), что находится в разумном согласии с измеренными значениями.

11.3.5. Универсальная барообразная неустойчивость

Численное моделирование тонких дискообразных галактик широко развивается примерно с 1967—1968 гг.; прекрасная обзорная статья, описывающая исследования вплоть до 1974 г., написана Холлом [198]. Наиболее ранняя работа касалась эволюции «холодных» бесконечно тонких дискообразных звезд с вращением, вполне достаточным для уравновешивания центробежной силой гравитационного притяжения. Диск считается «холодным», поскольку движение всех звезд первоначально является чисто азимутальным, и нет случайного разброса по скоростям относительно этого кругового движения. Если поверхностная массовая плотность равна

$$n_2(r) = n_2(0) \sqrt{1 - r^2/R^2}, \quad r \leq R, \quad (11.32a)$$

где R — радиус диска, то диск находится в равновесии, когда он вращается подобно твердому телу с угловой скоростью

$$\omega_0 = \sqrt{Gn_2(0)/2R}. \quad (11.32b)$$

Хотя первоначально диск находится в равновесии, это равновесие является неустойчивым, и диск разбивается на несколько кусков лишь за несколько вращений. Этот процесс показан на рис. 11.6 и связан с неустойчивостью Джинса. Он представляет собой явление, посредством которого гравитационное притяже-

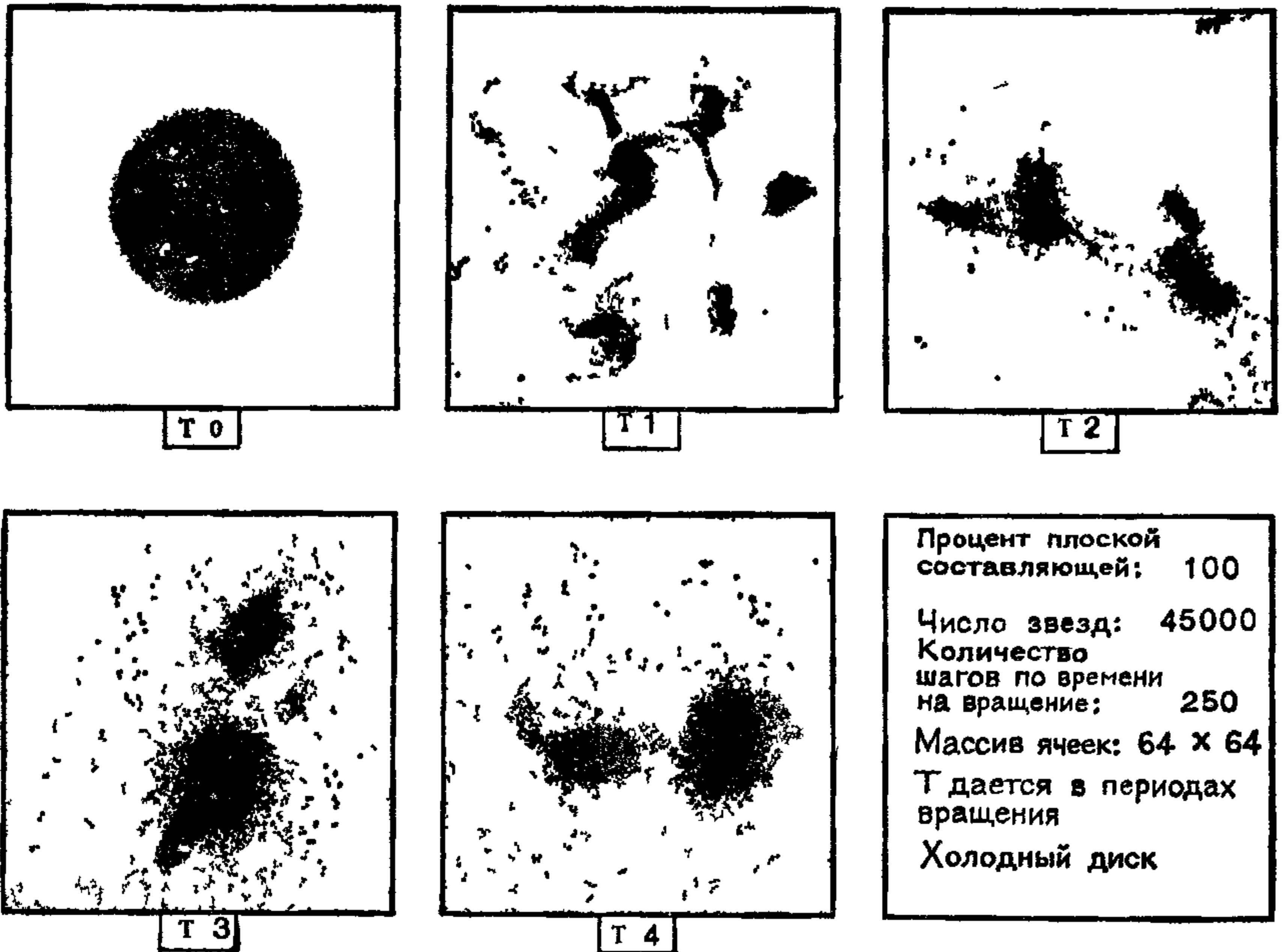


Рис 116 Быстрое разрушение холодного диска без гало ($\rho = 10$, $Q = 0$), обусловленное неустойчивостью Джинса. Время T дается в единицах периода вращения начального диска ([188], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society)

ние усиливает любой локальный максимум в плотности звезд. Тумре [371] показал, что осесимметричные неустойчивости такого рода можно подавить, если существует минимальная дисперсия радиальной скорости, заданная соотношением

$$\sigma_{r, \min} = 3.36 G n_2 / \omega_e, \quad (11.33)$$

где (поскольку $\langle v_r \rangle = 0$) $\sigma_r = \langle (v_r - \langle v_r \rangle)^2 \rangle^{1/2}$ есть среднеквадратичная радиальная скорость звезд на радиусе r . Если ввести дисперсию скорости, то диск считается «теплым» и степень выполнения критерия Тумре уравнения (11.33) выражается посредством локального значения Q — отношения действительной радиальной дисперсии к минимальному значению, требуемому для устойчивости:

$$Q = \sigma_r / \sigma_{r, \min}. \quad (11.34)$$

Многие вычислительные эксперименты были выполнены с $Q = 1$ для ряда начальных распределений плотности и кривых вращения; найдено, что диск стабилизируется относительно быстрого развала, показанного на рис. 116 (см [191, 195, 196, 198, 282, 283, 30]). Однако во всех случаях в системе преобла-

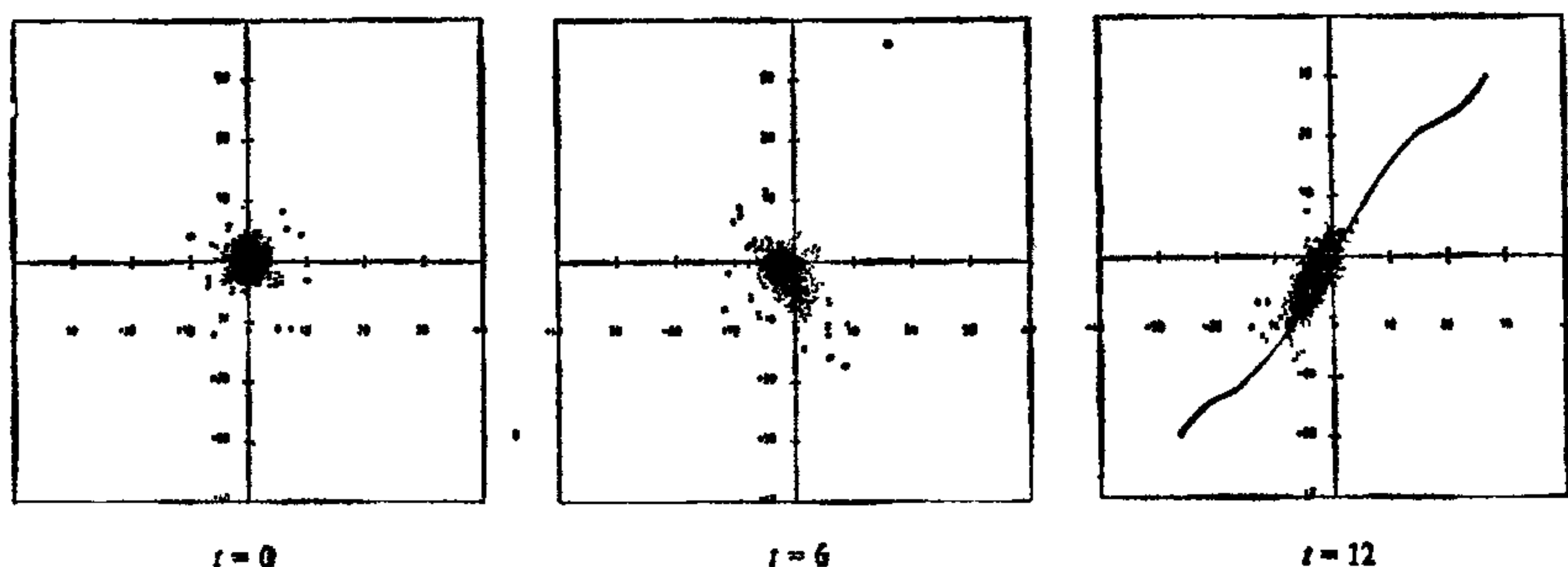


Рис 11.7 Стабилизация диска без гало ($\rho = 1.0$) посредством случайных движений звезд ($Q = 1$), однако замечаем окончательную стадию барообразной моды и отсутствие спиралей. Время дано в единицах, равных 10^8 лет. ([30], с любезного разрешения Astronomy & Astrophysics, © Springer Verlag.)

дает в конце концов медленно растущая барообразная мода, как показано на рис 11.7. Хол [198] исследовал эволюцию массовой плотности и момента импульса с радиусом, по мере того как растет барообразная мода. Он нашел, что бар переносит момент импульса наружу и что конечное распределение плотности удовлетворительно представляется в виде суммы двух экспонент:

$$n_2(r) = c_s \exp(-r/a_s) + c_d \exp(-r/a_d), \quad (11.35)$$

где $\alpha = a_d/a_s = 6$. Хотя этот эксперимент проводился путем двумерного моделирования, мы отождествляем слагаемые как ядро с масштабом длины a_s и диск с масштабом длины a_d и берем это в качестве примера стремления гравитационной системы к центрально конденсированным (α велико) распределениям массы. Этот результат находится в соответствии с наблюдениями Вокулера [375] и Фримана [136]. Барообразная неустойчивость меняет распределение величины Q , которое первоначально повсюду было равно 1, на распределение, изменяющееся приблизительно от 1 в центре до 7 во внешних областях. Период вращения бара приблизительно равен двум периодам вращения исходного диска.

По-видимому, неизбежной судьбой дискообразной галактики без гало ($\rho = 1$) должен стать горячий вращающийся бар, и нетрудно понять, что барообразные галактики образуют значительную группу галактик, наблюдаемых на небе (см классификацию Хаббла; рис. 11.2). Ни в каких этих экспериментах не было каких-либо данных о долгоживущей спиральной структуре, и окончательные дисперсии скорости в вычислительных системах превышают значения, реально наблюдаемые в спиральных галактиках.

Чтобы понять спиральную структуру, очевидно, необходимо найти некоторое воздействие, отличное от дисперсии скорости, для стабилизации неустойчивости Джинса и стабилизирующее, кроме того, диск относительно образования бара. Если результирующий диск является относительно холодным, $Q \approx 1$, то появится некоторый шанс, что спиральные волны могут расти. Это новое воздействие порождается влиянием звезд гало, которое мы игнорировали до настоящего момента. Далее мы увидим, что можно получить спирали как с тяжелыми диффузными гало (рис. 11.1, сверху), так и с легкими центрально сконденсированными гало (рис. 11.1, внизу).

11.3.6. Условия устойчивости спиральной структуры

Условие устойчивости вращающейся звездной системы относительно роста барообразной моды было постулировано Острайкером и Пиблсом [304] зависящим от отношения t вращательной кинетической энергии к абсолютному значению потенциальной энергии

$$t = KE_{\text{rot}} / |PE|. \quad (11.36)$$

Последовательность Маклорена однородных по плотности равномерно вращающихся жидких тел [232, гл. XII] и многие численные моделирования N тел неустойчивы, если $t > 0.14$. Острайкер и Пиблс формулируют это условие следующим образом:

При отсутствии контрпримеров кажется, что случай $t = 0.14$ представляет собой приблизительно максимум вращательной энергии, при которой осесимметричная звездная система может удерживаться и оставаться устойчивой относительно образования бара.

В холодном вращающемся диске ($Q = 0$) вращательная кинетическая энергия составляет половину потенциальной энергии, следовательно, $t = 0.5$, а для теплого диска, удовлетворяющего критерию Тумре для устойчивости относительно мелко-масштабных конденсаций Джинса ($Q = 1$), $t = 0.35$. Оба случая являются неустойчивыми согласно условию Острайкера и Пиблса, и за одно вращение барообразная неустойчивость возрастает и переводит в достаточном количестве энергию организованного вращения в случайное движение, понижая t до ~ 0.14 .

Для того чтобы иметь систему, в исходном положении являющейся устойчивой относительно возбуждения барообразной моды, необходимо иметь вначале $t < 0.14$. Этого можно достигнуть путем введения невращающегося гало звезд, которое не дает вклада в KE_{rot} , но в значительной степени увеличивает абсолютное значение полной потенциальной энергии. Острайкер

и Пиблс вводят диффузное гало — компоненту с масштабом длины, сравнимым с радиусом вращающегося диска из звезд, т. е. $\alpha \approx 1$, и находят, что если масса гало превышает приблизительно 67 % от полной массы ($\rho < 0.33$), то в исходном положении $t < 0.14$ и не наблюдаются барообразные моды. Гало, имеющее такую массу, описывается как тяжелое или массивное. Эксперименты Острайкера и Пиблса проводились в трехмерном случае со 150—300 моделируемыми звездами и компонента гало представлялась фиксированной приложенной внешней силой. Условие устойчивости также соответствует результатам моделирования Хола [195] и Миллера [282], в которых использовалось около 100 000 макрозвезд.

На рис. 11.8 показана эволюция холодного в исходном положении диска с тяжелым диффузным гало ($Q = 0$, $\rho = 0.2$, $\alpha = 1$) [29, 188]. Сначала заметим, что нет доказательства барообразного образования даже после 13 вращений. «Холодное» начальное условие ($Q = 0$) делает систему неустойчивой относительно мелкомасштабных конденсаций Джинса, и они принимают форму плотно намотанных спиралей, которые продолжают существовать все время эксперимента. Эти спиральные особенности являются не спиральными волнами, а конденсатами вещества, которые образуются, разрушаются и заново образуются сложным образом, в то время как они закручиваются дифференциальным вращением галактики. Хотя качественная полная спиральная картина является долгоживущей, детальные особенности внутри нее быстро меняются. Многие галактики Sa и Sb (см. рис. 11.2) имеют волокнистые особенности, аналогичные приведенным на рис. 11.8, и поэтому в холодном распределении звезд могут вызываться комбинацией конденсации Джинса и дифференциального вращения.

Чтобы определить спиральную волну, проводится анализ Фурье для распределения плотности по азимуту θ для ряда колец при различных радиусах r ; таким образом,

$$n_2(r, \theta) = \sum_{m=0}^{\infty} A_m(r) \cos(\Phi_m - m\theta), \quad (11.37)$$

где Φ_m — фаза спиральной моды с m ветвями и A_m/A_0 — ее относительная амплитуда. Характерная скорость спирали равна

$$\Omega_p = \frac{1}{m} \frac{\partial \Phi_m}{\partial t}, \quad (11.38a)$$

а волновое число k задается выражением

$$k = \frac{\partial \Phi_m}{\partial r}. \quad (11.38b)$$

Говорят, что спиральные волны существуют, если характерная скорость Ω_p является существенно постоянной на значительной

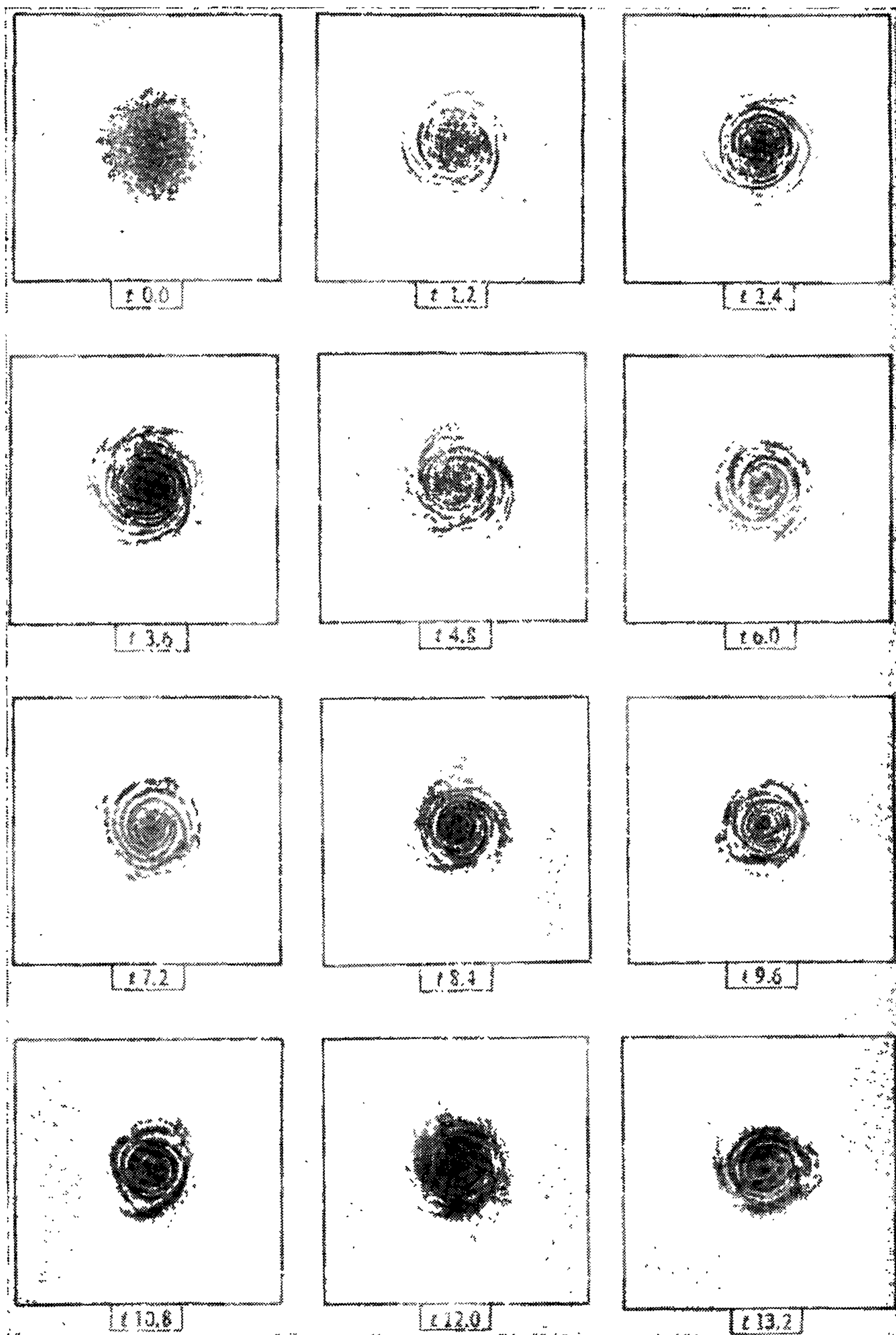


Рис. 11.8. Эволюция «холодной» дискообразной галактики, стабилизированная тяжелым диффузным гало ($Q = 0$; $\rho = 0.2$; $\alpha = 1$), показывающая развитие волокнообразного кольца и спиральных структур. ([29], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

части радиуса R галактики. Анализ обычно выполняется как на гравитационном потенциале, так и на плотности. Волновой анализ данных на рис. 11.8 демонстрирует когерентную волну с $m = 2$ между 4 и 36 кпк после 13 вращений. Спираль является полностью открытой и едва видна во внешних областях.

Спиральные волны плотности можно отчетливо видеть в «теплых» дисках, которые имеют достаточную случайную скорость для предотвращения конденсации Джинса ($Q > 1$). На рис. 11.9 представлен результат анализа спирали из трех ветвей в системе с $Q = 1$, $\rho = 0.2$, $\alpha = 4$ [29]. Сплошная и частично зигзагообразная линия является геометрическим местом точек пиков волны плотности. Относительная амплитуда волны составляет 36%. Радиальная протяженность волны показана на рис. 11.10, на котором видно, что характерная скорость является приемлемо постоянной приблизительно от 7 до 22 кпк.

В теории волн плотности являются важными три резонанса между звездными орбитальными движениями и волновой характерной скоростью. Они представляют собой внутренний и внешний резонансы Линдблада (ILR и OLR) и коротационный

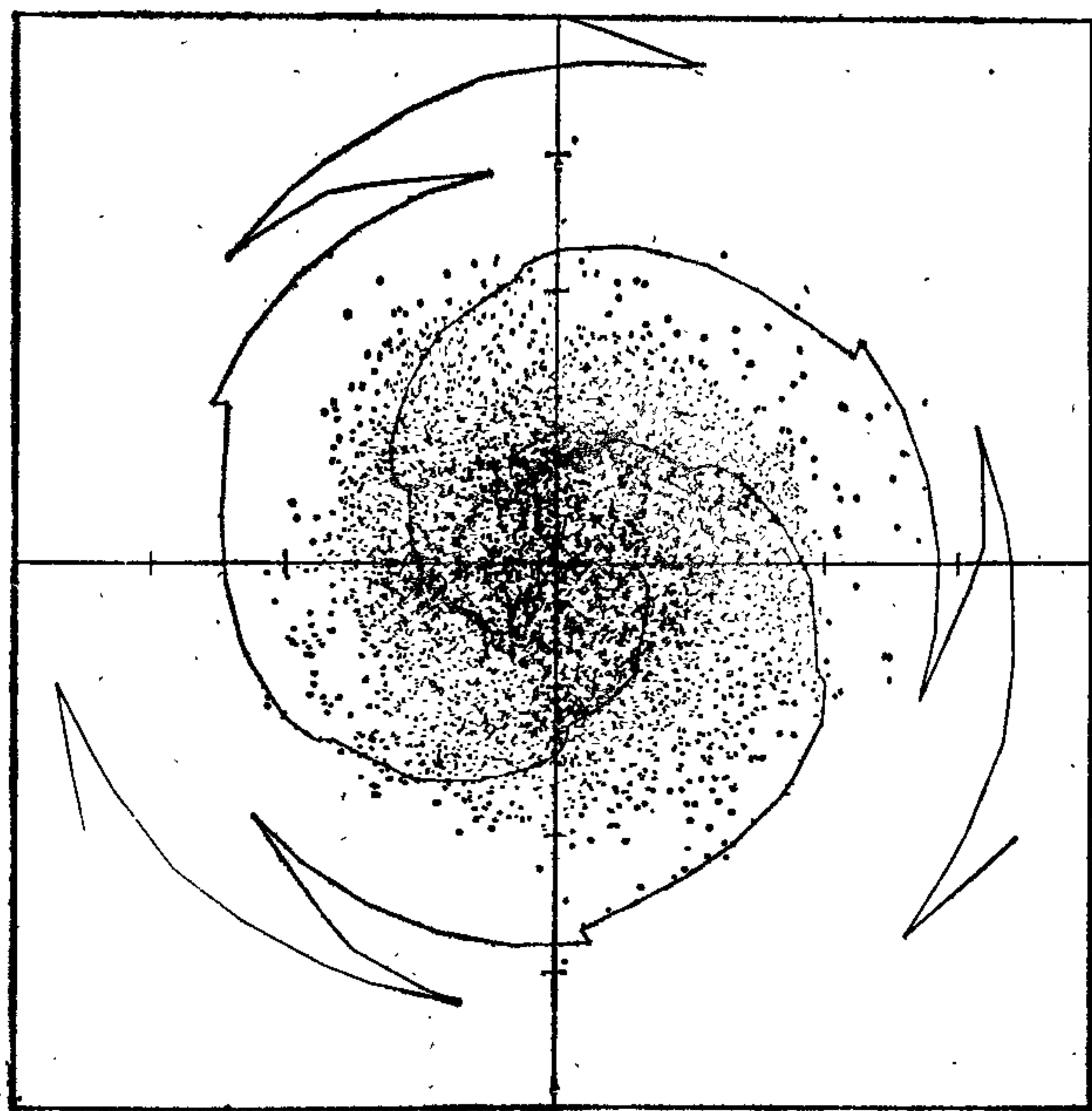


Рис. 11.9. Анализ слабо выраженной спирали с тремя ветвями в «теплой» галактике с тяжелым гало ($Q = 1$, $\rho = 0.2$, $\alpha = 4$). Линии показывают геометрическое место точек максимума плотности, полученное анализом Фурье, ([29], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

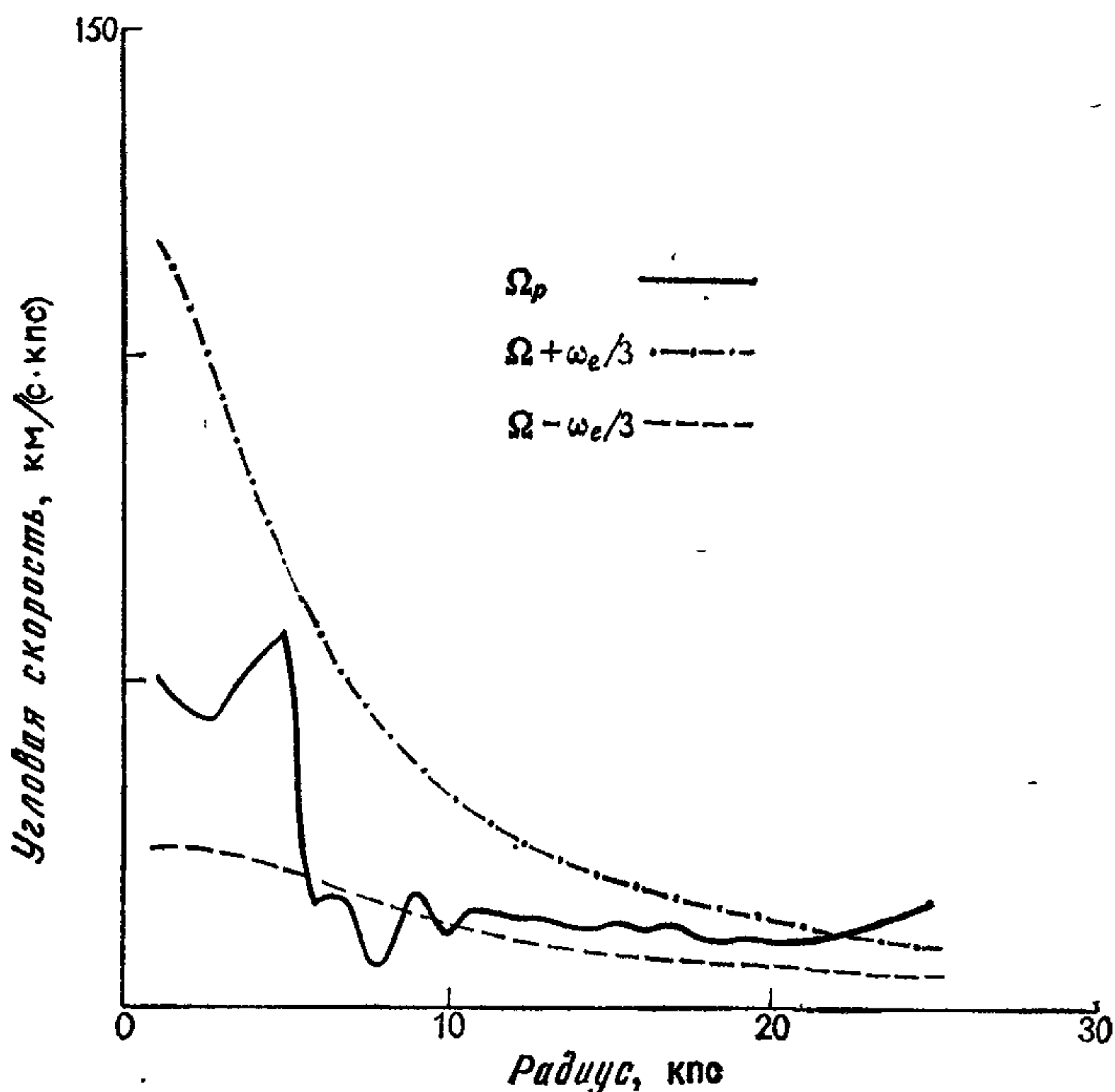


Рис. 11.10. Модельная скорость спирали с тремя ветвями из рис. 11.9 как функция радиуса, демонстрирующая внутренний и внешний резонансы Линдблада при 10 и 22 кпс соответственно. Коротационный резонанс располагается при 17 кпс; ω_e — эпициклическая частота, Ω — частота вращения звезд, Ω_p — частота вращения волны ([29], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

резонанс (CR). При коротации звезды и волна вращаются с одинаковой угловой скоростью. При внутреннем (внешнем) резонансе Линдблада угловая скорость звезд больше (меньше), чем угловая скорость волны; однако звезды, перемещающиеся во время их эпициклического движения, всегда проходят пик волны в той же самой фазе и имеют когерентное воздействие на нее. Условие резонансов имеет вид [260, 263]:

$$\Omega_p = \Omega + s\omega_e/m, \quad (11.39)$$

где $s = 0$ при коротации; $s = -1$ при внутреннем резонансе Линдблада; $s = +1$ при внешнем резонансе Линдблада; Ω_p и Ω — угловая скорость волновой модели и звезд соответственно.

На рис. 11.10 показаны кривые $\Omega \pm \omega_e/m$, из которых мы находим внутренний и внешний резонансы Линдблада при 10 и 22 кпс соответственно. Кривая ротации звезд $\Omega(r)$ не нанесена, но коротация происходит при 17 кпс. Отчетливо видно, что

волна наблюдается в области, заключенной между внутренним и внешним резонансами и включающей радиус коротации.

О подобных моделированиях N тел с массивными гало и разнообразными распределениями масс гало сообщили несколько авторов (см. [199]). Джеймс и Селлвуд [218] исследовали первоначально холодную систему с изменяющейся толщиной в трехмерном пространстве с двумя распределениями гало. В обоих случаях гало составляет 80 % от общей массы ($\rho = 0.2$), а масштаб длины гало — около $1/5$ галактического радиуса ($\alpha = 5$). Такие распределения, для которых $\alpha \gg 1$, описываются как центрально конденсированные. Образование бара успешно подавляется, и возникают ясно видимые долгоживущие спирали. Селлвуд и Джеймс [345] рассмотрели спирали с точки зрения теории механизма генерации волн плотности Линден-Белла и Калнайса (ЛБК) [271]. Этот механизм зависит от внешнего потока момента количества движения от внутреннего резонанса Линдблада (ILR) к коротационному резонансу (CR) или внешнему резонансу Линдблада (OLR), который Селлвуд и Джеймс наблюдали непосредственно в их модели. Вращающий момент, необходимый для потока наружу, обеспечивается отставанием спирали в нормальном случае, когда угловая скорость звезд уменьшается с радиусом, и опережением спирали в противоположном случае. Корректируя распределение массы гало, можно получить оба типа кривой вращения; тогда наблюдаются как отставание спирали, так и опережение, что предсказано теорией ЛБК. Отсутствие спирали, когда нет дифференциального вращения, также находится в согласии с теорией. Селлвуд и Джеймс нашли, что нет переноса момента количества движения, следовательно, нет спиральных волн, если нет как внутреннего резонанса Линдблада, так и коротационного резонанса. Присутствие или отсутствие внешнего резонанса Линдблада, по-видимому, не имеет отношения к этому критерию.

Сравнение волн плотности, наблюдавшихся в вычислительных экспериментах, с асимптотической теорией наименьшего порядка Лина и Шу (LS) [258], пренебрегающей азимутальными эффектами, было выполнено в работе [29], где измерялось дисперсионное соотношение для волн с тремя ветвями, показанных на рис. 11.9. Они нашли, что наблюдаемая дисперсия волн существует в основном в областях дисперсионной диаграммы, запрещенных для волн Лина — Шу. Селлвуд и Джеймс также обнаружили, что теория наименьшего порядка Лина — Шу не в состоянии объяснить их эксперименты. Не только наблюдаются спирали, существенно более открытые, чем предсказывает эта теория Лина — Шу, но они также гладко проходят через область коротации, которая является сингулярной точкой в теории и должна содержать спирали с бесконечной плотностью.

Ясно, что численное моделирование подтверждает теоретическую точку зрения, состоящую в том, что крупномасштабная когерентная спиральная структура представляет собой волновое явление в распределении плотности. Также ясно, однако, что спиральная структура, наблюдаемая в вычислительных экспериментах, является более сложной и совершенно отличной от структуры, допускаемой в асимптотических теориях спиральных волн. Детальное объяснение волн, наблюдаемых в вычислительных экспериментах, все еще остается нерешенной проблемой теории.

Описанные выше спиральные явления получаются с массивными гало, содержащими приблизительно 80 % полной массы галактики. Распределение массы гало изменяется от диффузного ($\alpha = 1$) до центрально сконденсированного ($\alpha = 4$). Это привело исследователей к заключению, что массивные гало являлись необходимым условием для долгоживущей спиральной структуры, и, следовательно, к уверенности, что полная масса типичной галактики в 4 или 5 раз больше, чем допускалось ранее на основании того, что большая часть массы располагается в диске. Это означает в свою очередь, что средняя массовая плотность Вселенной была также недооценена. Это является важным моментом для космологов, поскольку плотность Вселенной определяет, является ли Вселенная «замкнутой» или «открытой» (см. разд. 11.1.5). Оценки массовой плотности Вселенной, основанные на видимых звездах, в 10—100 раз меньше, чем критическая плотность, требуемая для замкнутой Вселенной [147]. Если широко распространенное существование спиральных галактик означает значительно более массивные гало, чем допускалось в сделанных выше оценках, то космологи — приверженцы представления о замкнутой Вселенной — могут учесть большую часть недостающей массы, необходимой для того, чтобы получить плоскую Вселенную¹⁾.

Заключение, состоящее в том, что широко распространенное существование спиральных галактик означает массивные гало, оспаривалось Берманом и Марком [30] на основе дальнейшего численного моделирования. Они заметили, что если параметр t вычислен для диска с поверхностной плотностью, меняющейся как $(r^2 + a_d^2)^{-3/2}$, и сферического гало с объемной плотностью, меняющейся как $(r^2 + a_s^2)^{-5/2}$, то $t \rightarrow 0$ не только для случая массивных гало ($\rho \rightarrow 0$), но также для случая центрально конденсированных гало с $\alpha = a_d/a_s \rightarrow \infty$. Действительно, $t = 0.14$ получается не только для массивного диффузионного гало $(\rho, \alpha) = (0.3, 0.7)$, но также для относительно легкого центрально конденсированного гало $(\rho, \alpha) = (0.5, 10)$. Вторая альтернатива следует из предположения Лиана, что устойчивость отно-

¹⁾ То есть предельный случай замкнутой Вселенной, — Прим. перев.

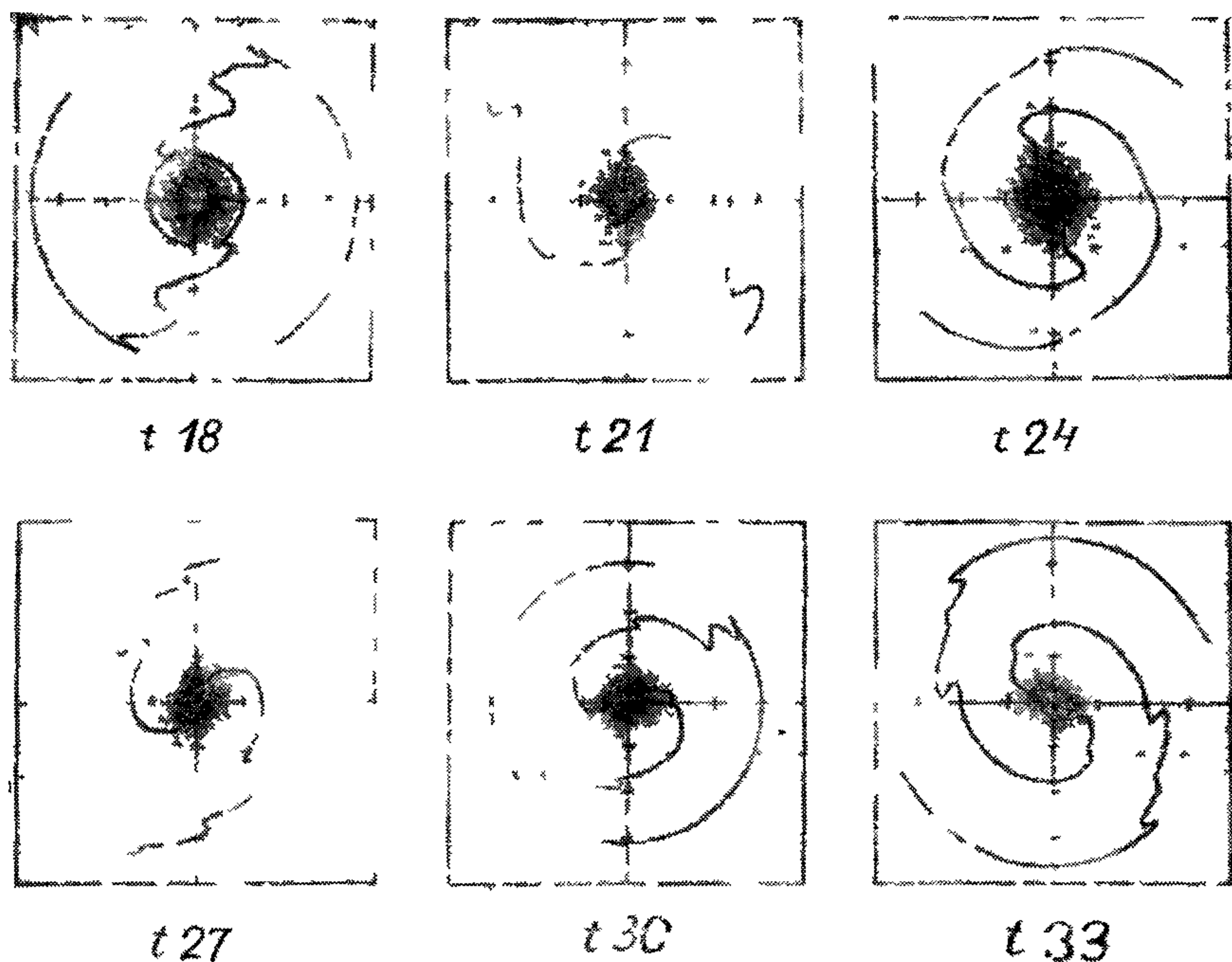


Рис 11.11 Эволюция галактики с легким центрально конденсированным гало ($\rho = 0.7$, $\alpha = 3$, $Q = 1$) Время дано в единицах, равных 10^8 лет ([30], с любезного разрешения Astronomy & Astrophysics, © Springer-Verlag)

сительно образования бара можно найти у ядра с относительно легким балджем, как видно в галактике NGC 4565 (рис. 11.1, внизу).

Вычислительный эксперимент Бермана и Марка был даже более успешным, поскольку устойчивая спираль с двумя ветвями сохранялась в течение 33×10^8 лет (≈ 15 вращений) для гало, содержащего только 30% общей массы ($\rho = 0.7$, $\alpha = 3$, $Q = 1$). Это демонстрируется на рис. 11.11. Для противопоставления эволюция системы без гало показывается на рис. 11.7, когда резко выраженный бар развивается только за несколько вращений. В эксперименте с балджем ядра значение параметра t в исходный момент равно ~ 0.4 и падает в процессе вращения до 0.26, после чего оно медленно снижается примерно до 0.25 к концу эксперимента. Из этого ясно, что Берман и Марк нашли контрпример к условию устойчивости Острайкера и Пиблса, в котором они имели систему, устойчивую к образованию бара с $t > 0.14$. Они также нашли условия, когда относительно лег-

кий балдж ядра поддерживает устойчивую спиральную структуру, которая является контрпримером того, что массивные гало — необходимое условие для устойчивой спиральной структуры. Этот вычислительный эксперимент будет, несомненно, оказывать важное влияние на развитие теории спиральной структуры и косвенно на космологию.

В вычислительных экспериментах этого раздела звезды гало были представлены как фиксированная внешняя сила, через которую движутся звезды диска. Хотя это есть естественная аппроксимация, однако следует проверить то, что результаты изменятся незначительно, если гало само моделируется корректно как совокупность макрозвезд. Противопоток частиц приводит к неустойчивостям других полей (например, двухпоточковой неустойчивости плазмы), и может случиться так, что приближение неподвижного гало подавляет важные эффекты. Эта возможность исследовалась Холлом [200] в трехмерном моделировании с 50 000 звезд, представляющими диск, и, кроме того, с 50 000 звезд, представляющими гало. Сетка потенциала содержит массив ячеек ($64 \times 64 \times 16$). Массы гало и диска были равны ($\rho = 0.5$), а гало было выпуклым, как в NGC 4565 ($\alpha = 4$). В сферическом гало не имелось заметных эффектов, а само гало осталось сферически симметричным. Диск оставался плоским и поддерживал спираль с двумя ветвями, которая все еще преобладала в конце эксперимента (3 вращения). Бар не образовывался. Этот эксперимент подтвердил, что использование неподвижного гало является хорошим приближением.

11.3.7. Протогалактика

В вычислительных экспериментах, описанных выше, предполагалось, что галактика уже прошла эволюцию от исходного вращающегося газового шара в горячее сферическое стационарное гало (состоящее из звезд сферической составляющей), через которое вращается тонкий диск (состоящий из звезд плоской составляющей). Полное трехмерное моделирование из 100 000 звезд, представленное в конце раздела, использовалось для моделирования всей системы. Следовательно, его можно применить для моделирования эволюции самой протогалактики, что было выполнено Холлом и Зангом [202]. Они использовали сетку ($33 \times 33 \times 33$), 100 000 макрозвезд и следили за эволюцией в течение 3—5 вращений. Другие вычислительные эксперименты протогалактики были проведены Готтом [146] в предположении осевой симметрии и Миллером [284] при помощи параллельного компьютера ILLIAC IV.

Если начать с вращающегося сферического газового шара и допустить, что большая часть газа сконденсировалась в звезды за время первого коллапса, то последующая эволюция будет

хорошо представлена бесстолкновительной моделью частиц использованного выше типа Хол и Занг рассматривали сферу с равномерной плотностью и изменяли величину вращения от 0 до 1.15 от уравнивающего гравитационное притяжение в экваториальной плоскости. Это охватывает диапазон пяти моделей для изменения переменной t от 0 до 0.44, введенной Острайкером и Пиблсом. Величина случайной кинетической энергии заключалась в диапазоне от 0.25 до 0.05 значения потенциальной энергии. Продолжительность каждого эксперимента была равна времени трех оборотов системы, находящейся в экваториальном равновесии. В конце эксперимента состояние без вращения все еще сферически симметричное, но исходная постоянная плотность релаксировала к экспоненциальному убыванию с одним масштабом длины ($\sim e^{-r/a}$). Это состояние напоминает галактику E0, такую, как NGC 3379. Состояния с величиной вращения, равной 0.5 и 0.7 от вращения, необходимого для равновесия, обнаруживают конечную чечевицеобразную форму с укорочением, приблизительно равным 25 % вдоль оси вращения. Конечное распределение все еще является осесимметричным и напоминает галактику E2, такую, как M32. Эти три случая имеют значения t , равные 0, 0.08 и 0.16 соответственно, и являются устойчивыми относительно образования бара или в последнем случае нейтрально устойчивыми. В любом случае отсутствуют какие-либо признаки образования бара за время трех вращений.

В двух оставшихся случаях брались значения скорости вращения 0.87 и 1.15 от равновесной и значения t 0.25 и 0.44. Ожидалось, что эти случаи должны быть неустойчивыми относительно образования бара согласно критерию Острайкера и Пиблса. Обе системы образуют барообразные структуры после трех и двух вращений соответственно. Барообразные структуры вращаются примерно в экваториальной плоскости, и в случае двух вращений они возбуждают короткую спиральную структуру с двумя ветвями, подобную барообразной SBa-галактике NGC 175.

Эксперименты с равномерно вращающейся бесстолкновительной протогалактикой показали эволюцию обыкновенной эллиптической и барообразной эллиптической форм. Полная модель должна обязательно по отдельности включать газообразную и звездную компоненты и подходящий механизм для рождения звезд из газа, как это было сделано намного раньше для дискообразных галактик Миллером, Прендергастом и Квирком [286]. Когда учитываются такие эффекты, то появляется некоторая надежда моделировать различные условия в протогалактике, что приводит к образованию тонких дисков, гало и нормальной спиральной структуры.

11.4. СКУЧИВАНИЕ ГАЛАКТИК

Теперь мы вернемся к явлению сгущивания галактик, упомянутому в разд. 11.1.4 и 11.1.6. Ожидается, что существующее в настоящее время распределение галактик содержит ключ к пониманию состояния Вселенной, когда галактики конденсировались, и к природе последующего расширения Вселенной. В частности, есть надежда, что подробное изучение сгущивания подтвердит либо замкнутую, либо открытую модель Вселенной. Численное моделирование N тел сыграло большую роль в исследовании выводов различных космологических теорий относительно начальных условий и уравнений расширяющейся Вселенной. Ниже мы представляем уравнения (разд. 11.4.1 и 11.4.2), численную модель (разд. 11.4.3) и некоторые результаты (разд. 11.4.4) из работ [123, 124] при изучении моделей Вселенной Фридмана при помощи кода P^3M с 20 000 моделируемыми галактиками. Предыдущая работа, использующая метод РР, была ограничена 1000 или самое большее 4000 моделируемыми галактиками.

11.4.1. Уравнения расширяющейся Вселенной

В космологии Фридмана [137] Вселенная предполагалась однородной и изотропной, если проводилось усреднение по достаточно большим расстояниям, чтобы размазать явные неоднородности, возникающие от сгущивания галактик, отдельных галактик и звезд. При рассмотрении общего расширения мы говорим поэтому о средней плотности материи $\bar{\rho}(t)$, которая является функцией времени, но не пространства. Рассматривая Вселенную как совокупность точечных масс, любую из которых можно взять в качестве начала координат, исходное состояние Вселенной описывают заданием положений $r_i(0)$ и скоростей $v_i(0)$ при заданном времени $t = 0$ по отношению к выбранному началу координат. Будучи равномерным, расширение приводит только к радиальным движениям и скоростям, если оно наблюдалось из какого-либо центра. Следовательно, r_i и v_i — оба скаляры — радиальное расстояние и радиальная скорость соответственно. Мы также ограничиваем обсуждение моделями Фридмана с нулевым значением космологической постоянной Эйнштейна Λ (см., например, [369]).

В последующее время, начиная с $t = 0$, расширение описывается параметром расширения

$$a(t) = r_i(t)/r_i(0), \quad a(0) = 1. \quad (11.40)$$

Из-за предположения однородности параметр $a(t)$ является тем же самым для всех частиц и промежутков. При расширении все объемы увеличиваются на множитель $a^3(t)$, следовательно,

средняя плотность вещества уменьшается согласно выражению

$$\bar{\rho}(t) = \bar{\rho}(0)/a^3(t). \quad (11.41)$$

В любой момент времени справедлив закон Хаббла [206; 207, гл. VII], связывающий скорость и расстояние, а именно что радиальная скорость расширения пропорциональна расстоянию от частицы:

$$v_i(t) = H(t) r_i(t). \quad (11.42)$$

Заметим, что «постоянная» Хаббла $H(t)$, хотя и одна и та же для всех частиц, обычно является функцией времени. Ее обратная величина представляет собой естественную единицу времени для космологических задач, которую мы используем ниже. Она является временем, которое прошло бы после большого взрыва, если бы скорость расширения была постоянной. Подставляя уравнение (11.40) в (11.42), получаем

$$H(t) = \frac{v_i(t)}{r_i(t)} = \frac{\dot{a}(t)}{a(t)}, \quad (11.43)$$

где точка используется для обозначения производной по времени.

Из теоремы Гаусса (см. [165]) можно заключить, что сила на радиусе r в сферически симметричной системе создается только массой, расположенной внутри этого радиуса и действующей так, как будто вся она находилась в центре. Следовательно, уравнение движения Ньютона для i -й частицы получается приравниванием ускорения к гравитационному притяжению всей массы, расположенной внутри радиуса r_i :

$$\ddot{r}_i = -\frac{4}{3}\pi r_i^3 \bar{\rho} (G/r_i^2) = -\frac{4}{3}\pi G \bar{\rho}(t) r_i(t). \quad (11.44)$$

Разделив уравнение (11.44) на $r_i(t)$, получаем уравнение Фридмана для параметра расширения

$$\ddot{a}(t) = -\frac{4}{3}\pi G \bar{\rho}(t) a(t), \quad (11.45a)$$

или, используя уравнение (11.41),

$$\ddot{a}(t) = -\frac{4}{3}\pi \frac{G \bar{\rho}(0)}{a^2(t)}. \quad (11.45b)$$

Умножая уравнение (11.45b) на $2\dot{a}$ и интегрируя, получаем

$$\frac{d}{dt} \left[\dot{a}^2 - \frac{8}{3} \pi \frac{G \bar{\rho}(0)}{a(t)} \right] = 0, \quad (11.46)$$

следовательно,

$$\dot{a}^2 - \frac{8}{3} \pi \frac{G \bar{\rho}(0)}{a(t)} = -k, \quad (11.47)$$

где k — постоянная, которая, согласно терминологии общей теории относительности, представляет собой кривизну пространства. Уравнение (11.47) есть уравнение энергии, первый член

пропорционален кинетической энергии частицы, а второй член — ее потенциальной энергии. Если кинетическая энергия превышает потенциальную ($k < 0$), то расширение продолжается неограниченно и Вселенная открытая. Обратное, если потенциальная энергия превышает кинетическую ($k > 0$), то расширение в конце концов остановится и Вселенная замкнутая. Предельный случай равенства между кинетической и потенциальной энергиями ($k = 0$) означает непрерывное расширение, хотя обычно его рассматривают как предельный случай замкнутой Вселенной.

Уравнение (11.47) с $k = 0$ позволяет нам вычислить критическую плотность вещества при $t = 0$, которая соответствует случаю плоской Вселенной, в виде

$$\bar{\rho}_c(0) = \frac{3}{8\pi G} a(t) \dot{a}^2(t), \quad (11.48)$$

или, используя уравнение (11.41), получаем, что критическая плотность в момент времени t равна

$$\bar{\rho}_c(t) = \frac{3}{8\pi G} \left[\frac{\dot{a}(t)}{a(t)} \right]^2 = \frac{3H^2(t)}{8\pi G}. \quad (11.49)$$

Обычно Вселенную описывают через безразмерный параметр плотности Ω , который представляет собой отношение средней массовой плотности к плотности, требуемой для плоской Вселенной

$$\Omega(t) = \frac{\bar{\rho}(t)}{\bar{\rho}_c(t)} = \frac{8\pi G \bar{\rho}(t)}{3H^2(t)} = \quad (11.50a)$$

$$= \frac{8\pi G \bar{\rho}(0)}{3a(t) \dot{a}^2(t)}. \quad (11.50b)$$

Комбинируя уравнение (11.50b) с (11.47), видим, что если $\Omega(0) = 1$, то $k = 0$ и $\Omega(t)$ остается равным единице в процессе расширения. Этот случай называется Вселенной Эйнштейна — де Ситтера. Если, однако, $\Omega < 1$, что соответствует открытой Вселенной, то Ω будет уменьшаться в процессе расширения. Параметр Ω представляет собой также отношение потенциальной энергии к кинетической и играет роль, аналогичную той, которую играет параметр Γ_0 из гл. 12.

Если время измеряется в единицах начального времени Хаббла $H^{-1}(0)$, то уравнение Фридмана (11.47) можно выразить через единственный параметр $\Omega(0)$:

$$\frac{da(t')}{dt'} = \left\{ \frac{\Omega(0)}{a(t')} + [1 - \Omega(0)] \right\}^{1/2}, \quad (11.51)$$

где $t' = H(0)t$. При заданной начальной безразмерной плотности $\Omega(0)$ и начальном условии $a(0) = 1$ это уравнение однозначно определяет последующее расширение Вселенной. Расширение, таким образом, определяется единственным параметром $\Omega(0)$.

11.4.2. Сопутствующие координаты

При изучении скучивания галактик полезно выделить общее расширение Вселенной путем введения сопутствующих координат, обозначенных штрихованными переменными

$$\mathbf{r}'(t) = \mathbf{r}_i(t)/a(t), \quad (11.52a)$$

$$\mathbf{r}'(0) = \mathbf{r}_i(0). \quad (11.52b)$$

Тогда, если бы гравитационного взаимодействия между частицами Вселенной не было, переменные должны были оставаться стационарными в сопутствующих координатах. Поэтому любое движение частиц в сопутствующих координатах является скучиванием, возникающим из-за гравитационных сил. Дополнительно используется факт, что средняя плотность вещества в сопутствующей системе координат является постоянной и равна ее начальному значению $\bar{\rho}(0)$.

Уравнения Ньютона для скорости и положения частицы в исходных или «собственных» координатах выражаются в виде

$$\mathbf{v} = -\nabla\varphi, \quad (11.53a)$$

$$\dot{\mathbf{r}} = \mathbf{v}, \quad (11.53b)$$

$$\nabla^2\varphi = 4\pi G\rho(\mathbf{r}, t), \quad (11.53c)$$

где φ — гравитационный потенциал; для простоты нижний индекс, соответствующий номеру частицы, у \mathbf{v} и \mathbf{r} опущен. Массовая плотность $\rho(\mathbf{r}, t)$ по сравнению с $\bar{\rho}$ представляет эффекты скучивания и, следовательно, включает изменения в масштабе отдельной галактики.

Подставляя уравнение (11.52a) в (11.53a) и учитывая, что $H = \dot{a}/a$, получаем

$$\dot{\mathbf{v}}' + 2H(t)\mathbf{v}' = -\frac{1}{a}\nabla\varphi - \frac{\ddot{a}}{a}\mathbf{r}', \quad (11.54a)$$

где по определению

$$\dot{\mathbf{r}}' = \mathbf{v}'. \quad (11.54b)$$

Правую часть уравнения (11.54a) можно записать в виде градиента, замечая, что $\nabla\varphi = a^{-1}\nabla'\varphi$ и $\nabla'(1/2)\mathbf{r}^2 = \mathbf{r}'$, где штрихи показывают, что пространственные производные взяты относительно сопутствующих координат. Тогда имеем

$$\dot{\mathbf{v}}' + 2H(t)\mathbf{v}' = -\frac{1}{a^3}\nabla'\varphi', \quad (11.55a)$$

где

$$\varphi' = a\varphi + \frac{a^2\ddot{a}}{2}\mathbf{r}'^2. \quad (11.55b)$$

Подставляя уравнение (11.55b) в (11.53c) и используя $\nabla^2 = a^{-2}\nabla'^2$ и $\nabla'^2(1/2)r'^2 = 3$, получаем сопутствующую форму уравнения Пуассона

$$\nabla'^2\varphi' = 4\pi G\rho a^3 + 3\bar{a}a^2 = \quad (11.56a)$$

$$= 4\pi G(\rho'(r', t) - \bar{\rho}(0)), \quad (11.56b)$$

где плотность в сопутствующей системе координат $\rho' = a^3\rho$, и мы использовали уравнение (11.45b) для первоначальной средней массовой плотности $\bar{\rho}(0)$.

Поэтому уравнения Ньютона в сопутствующей системе координат, соответствующие уравнению (11.53), имеют вид

$$\dot{\mathbf{v}}' + 2H(t)\mathbf{v}' = -\frac{1}{a^3}\nabla'\varphi = \mathbf{F}', \quad (11.57a)$$

$$\dot{\mathbf{r}}' = \mathbf{v}', \quad (11.57b)$$

$$\nabla'^2\varphi' = 4\pi G(\rho'(r', t) - \bar{\rho}(0)). \quad (11.57c)$$

Можно видеть, что замена координат привела к двум эффектам. Во-первых, в уравнение для ускорения (11.57a) введен член, описывающий вязкое трение и пропорциональный постоянной Хаббла и скорости, и, во-вторых, в правой части уравнения Пуассона (11.57c) появилась постоянная отрицательная массовая плотность, равная средней массовой плотности $\bar{\rho}(0)$ с обратным знаком. Этот последний член обеспечивает равенство нулю полной массы любой системы, содержащей много галактик. Фактически мы будем моделировать кубическую область Вселенной с тройной периодичностью, содержащей 20 000 моделируемых галактик. Заметим, что решение уравнения Пуассона с граничными условиями, обладающими тройной периодичностью, математически отсутствует, если полная масса в системе не равна нулю. Поэтому член с отрицательной массой, который появился из-за изменения координат, является необходимостью и не должен рассматриваться каким-либо образом как искусственный или нефизический.

Физические уравнения дополняются уравнением энергии, выведенным в сопутствующих координатах Лейзером [243] и Ирвином [215]

$$\frac{d}{dt}(a^4T) + a\frac{dW}{dt} = 0, \quad (11.58a)$$

где кинетическая энергия

$$T = \sum_i \frac{1}{2} m_i a v_i'^2 \quad (11.58b)$$

и потенциальная энергия

$$W = \sum_i m_i \varphi'(r'_i). \quad (11.58c)$$

Эти энергии связаны с соответствующими энергиями, полученными в собственных координатах, соотношениями $\hat{T} = a^2 T$, $\hat{W} = W/a$.

11.4.3. Вычислительная модель

При помощи этой модели мы исследуем кубическую область расширяющейся Вселенной и налагаем граничные условия с тройной периодичностью для приближенного описания остальной части Вселенной. В кубе движутся 20 000 галактик, причем каждая представляется как облако массы конечного размера диаметром, приблизительно равным $1/5$ среднего межчастичного промежутка. Периодичность гарантирует, что все галактики, покидающие куб, вводятся обратно, и поэтому средняя массовая плотность в кубе остается постоянной и равной $\bar{\rho}(0)$. Потенциал и силы подсчитываются с помощью алгоритма P^3M (см. гл. 8) и к каждому узлу сетки добавляется подходящим образом постоянная отрицательная массовая плотность $-\bar{\rho}(0)$ во время этапа раздачи массы алгоритма. В приводимых здесь расчетах использовалась кубическая сетка из 32^3 ячеек.

Используемая программа была такой же, как программа, написанная Иствудом, Хокни и Лоуренсом [120] для моделирования ионных микрокристаллов (см. гл. 12). Программа была модифицирована, как описано ниже, Эфстатиу и Иствудом [124] для адаптации ее к проблеме скучивания галактик. Она включает введение вязкого члена в уравнения движения, модификацию вида короткодействующих сил и дополнение программы для интегрирования уравнений Фридмана. Требуются также новые программы для измерения энергии, корреляционных функций, спектра флуктуаций и усреднений скорости, а также для формулировки начальных условий. Относительная легкость, с которой программа численного моделирования могла быть адаптирована для использования в совершенно иной области науки, придает особое значение большой гибкости численного моделирования как инструмента в научном исследовании.

Уравнения движения (11.57а, б) представляются центральными разностями

$$v^{n+1/2} = v^{n-1/2} \frac{(1 - H(t) DT)}{(1 + H(t) DT)} + \frac{F'^n DT}{(1 + H(t) DT)}, \quad (11.59a)$$

$$r^{n+1} = r^n + v^{n+1/2} DT, \quad (11.59b)$$

где верхние индексы обозначают временные уровни, например $v^{n+1/2} = v((n + 1/2) DT)$. Уравнение энергии Лейзера — Ирвина (11.58а) представляется в разностной форме выражением

$$\frac{(a^{n+1})^4 T^{n+1} - (a^n)^4 T^n}{DT} + \frac{a^{n+1} W^{n+1} - a^n W^n}{DT} - W^{n+1} \dot{a}^{n+1} = 0 \quad (11.60a)$$

и просуммировано для получения соотношения

$$(a^4 T)^m + (aW)^m - \sum_{n=1}^m DTW^n \dot{a}^n = C, \quad (11.60b)$$

где постоянная C зависит от начальных условий:

$$C = (a^4 T)^0 + (aW)^0. \quad (11.60c)$$

Вычисляется левая часть уравнения (11.60b) и в качестве контроля энергии используется ее постоянство во время счета. Выбранный шаг по времени был равен 0.022 от начального времени Хаббла $H^{-1}(0)$ и в типичном режиме приблизительно для 800 временных шагов $\Delta C / \Delta(aW) \leq 2\%$.

Уравнения Фридмана для параметра расширения (11.45) и (11.47) интегрируются при помощи ряда Тейлора. Вообще говоря, значение a на следующем временном шаге равно

$$a^{n+1} = a^n + \dot{a}^n DT + \ddot{a}^n \frac{DT^2}{2} + \frac{\dddot{a}^n DT^3}{6} + \frac{\overset{\cdot\cdot\cdot}{a}^n}{24} DT^4 + \dots, \quad (11.61a)$$

и аналогично значение \dot{a} равно

$$\dot{a}^{n+1} = \dot{a}^n + \ddot{a}^n DT + \dot{\ddot{a}}^n \frac{DT^2}{2} + \overset{\cdot\cdot\cdot}{\ddot{a}}^n \frac{DT^3}{6} + \dots \quad (11.61b)$$

Определяя из последнего шага значения a^n и \dot{a}^n , производные более высокого порядка можно найти дифференцированием уравнения (11.51). Используя $H^{-1}(0)$ в качестве единицы времени, получаем

$$\ddot{a} = -\Omega(0)/2a^2, \quad (11.61c)$$

$$\dot{\ddot{a}} = +\Omega(0)\dot{a}/a^3, \quad (11.61d)$$

$$\overset{\cdot\cdot\cdot}{\ddot{a}} = \Omega(0)(\ddot{a}/a^3 - 3\dot{a}/a^4). \quad (11.61e)$$

О точности интегрирования можно судить путем проверки, насколько хорошо полученное решение удовлетворяет уравнению (11.51). Если ряд для a выбирается содержащим члены вплоть до DT^4 , а ряд для \dot{a} — члены до DT^3 , то интегрирование определяется с точностью приблизительно до шести значащих цифр. При интегрировании уравнений Фридмана используется тот же временной шаг, что и при интегрировании по времени уравнения движения частиц (11.59).

Основные диагностики, используемые в описании сгущения, представляют собой двухточечные и многоточечные корреляционные функции. Хотя были измерены трех- и четырехточечные корреляции, мы представили только результаты для двухточечных корреляций. Оценка для двухточечной корреля-

ции определяется формулой

$$\xi(r) = \frac{N_p}{\bar{n}N_cV} - 1, \quad (11.62)$$

где N_p — число пар частиц с расстоянием между частицами от r до $r + \Delta r$, V — объем этого слоя, N_c — число частиц, выбранных в качестве центров, и \bar{n} — средняя плотность частиц. Согласно этому определению, наиболее нескученное распределение — равномерное случайное распределение — имеет корреляцию $\xi(r) \equiv 0$. Следовательно, двухточечная корреляционная функция является мерой скучивания. В типичной ситуации $\xi(r)$ измеряется до расстояний, равных $1/7$ стороны кубического ящика, и накапливается число пар галактик внутри равных интервалов $\log r$. Двухточечная корреляционная функция, определяемая уравнением (11.62), равна радиальной (или парной) функции распределения $g(r)$ молекулярной динамики минус единица (см. гл. 12).

При использовании этой программы для эволюции модели с $\Omega = 1$ при параметре расширения, равном 10, требуется около 820 временных шагов. Временной шаг слабо зависит от степени скучивания, но обычно занимает (для $a \sim 5$) около 18 с на IBM 360/195. Это время делится между различными частями алгоритма Р³М следующим образом:

Раздача массы	4.8 с
Расчет потенциала (сетка 32 ³)	3.5 с
Вычисление сеточной силы	1.4 с
Вычисление короткодействующей силы	6.8 с
Продвижение положений	1.3 с
<hr/>	
Полный цикл	17.8 с

Вычисление двухточечной корреляционной функции по положениям частиц проводится отдельной программой и занимает 2—5 мин в зависимости от степени скучивания. Все 20 000 частиц выбираются в качестве центров и объем расчетов ограничивает ее использование до 4 значений параметра расширения ($a = 4, 6, 9, 19.3$). Интересно заметить, что расчет корреляций требует в десять раз больше времени, чем основной цикл временного шага. Это объясняется тем, что используемые усложненные алгоритмы, применяемые для избежания вычислений объемом N^2 на цикле временного шага — а именно алгоритма Р³М — не использовались при анализе результатов!

11.4.4. Результаты и выводы

Основным результатом наблюдения является результат Тотсуи и Кихары [374], которые нашли эмпирическое соотношение $\xi(r) \sim r^{-1.8}$ для данных, полученных Ликской обсерваторией.

Позднее Пиблс и его коллеги в университете Принстона более широко проанализировали данные трех наблюдательных групп: каталог Цвикки из 5000 галактик, собранный в Калифорнийском технологическом институте, каталог из 10^6 галактик, полученный Шайном и Виртаненом в Ликской обсерватории, и каталог из 10 000 галактик, составленный Рудницким в Ягеллонском университете Кракова. Пиблс [307] нашел в соответствии с Тотсуи и Кихарой, что двухточечная корреляционная функция приблизительно следует степенному закону в широком диапазоне масштабов

$$\xi(r) = \left(\frac{r_0}{r}\right)^\gamma, \quad \gamma = 1.77 \pm 0.066, \quad (11.63a)$$

где $r_0 = 5.3 \times (1.5)^{\pm 1} h^{-1} \text{ Мпк}^1$ в диапазоне $0.1 h^{-1} \text{ Мпс} \leq r \leq \leq 9 h^{-1} \text{ Мпк}$. Позднее повторный анализ ликских данных Гротом и Пиблсом [152] показал резкое изменение наклона или излома функции $\xi(r)$ при

$$\xi_{\text{break}} \sim 0.3, \quad r_{\text{break}} \sim 9 h^{-1} \text{ Мпк}. \quad (11.63b)$$

Задача теории вместе с численным моделированием состоит в том, чтобы установить условия, которые могли привести к наблюдаемым результатам, суммированным в простых соотношениях (11.63).

За проблему взялись Дэвис, Грот и Пиблс [101] и Дэвис и Пиблс [102], которые выполнили численное интегрирование кинетических уравнений БГКИ, основанное на правдоподобных допущениях для замыкания иерархии уравнений и начальных условий. Эти допущения включают использование наблюдаемого вида трехточечной корреляционной функции. Они получили хорошее согласие со значением γ в уравнении (11.63) и наблюдали излом функции $\xi(r)$ в примерно правильном месте. Однако этот результат теперь оказался под сомнением в связи с результатами численного моделирования N тел, описанного ниже. Пренебрегая численными погрешностями, которые, как можно показать, малы, в экспериментах с N телами точно интегрируется движение 20 000 галактик и получается совершенно другой результат.

Другой важный набор теоретических результатов получен Пиблсом [307] и связывает спектр мощности начальных флуктуаций плотности со значением γ . При рассмотрении галактик как точечных масс плотность числа частиц равна

$$n(\mathbf{r}) = \sum_{\text{галактики}} \delta(\mathbf{r} - \mathbf{r}_i), \quad (11.64a)$$

¹⁾ $h^{-1} \text{ Мпк}$ является астрономической единицей расстояния; h^{-1} — постоянная $H/(100 \text{ км} \cdot \text{с}^{-1} \cdot (\text{Мпк}^{-1}))$, а Мпк есть мегапарсек.

где δ — трехмерная дельта-функция Дирака, а относительная флуктуация плотности равна

$$\Delta(\mathbf{r}) = n(\mathbf{r})/\bar{n} - 1, \quad (11.64b)$$

где средняя плотность

$$\bar{n} = \langle n(\mathbf{r}) \rangle \quad (11.64c)$$

и корреляционная функция

$$\xi(\mathbf{r}') = \langle \Delta(\mathbf{r}) \Delta(\mathbf{r} + \mathbf{r}') \rangle. \quad (11.64d)$$

Ломаные скобки означают усреднение, взятое по большому ансамблю. Спектр мощности, или структурный фактор $S(k)$, получается преобразованием Фурье флуктуации $\Delta(\mathbf{r})$ в гармоники \mathbf{k} с амплитудой Δ_k ; тогда

$$S(\mathbf{k}) = |\Delta_k|^2 \quad (11.65)$$

пропорционально энергии или мощности k -й моды. Из-за изотропии пространства $S(\mathbf{k})$ рассматривается как функция только модуля вектора \mathbf{k} .

В случае если начальный спектр мощности флуктуаций имеет вид степенного закона

$$|\Delta_k|^2 \sim |k|^n, \quad (11.66a)$$

Дэвис и Пиблс [102] находят, что

$$\gamma = \frac{3n + 9}{n + 5}, \text{ если } r \ll \lambda, \xi \gg 1, \quad (11.66b)$$

$$= n + 3, \text{ если } r \gg \lambda, \xi \ll 1, \quad (11.66c)$$

где $\lambda = (\bar{n})^{-1/3}$ — среднее расстояние между галактиками. Уравнение (11.66c) выводится из линейной теории и справедливо, когда галактики находятся далеко друг от друга. Уравнение (11.66b) основано на линейной теории для случая $\Omega = 1$ и справедливо, если галактики близки друг к другу.

Ясно, что значение $n = 0$, которое дает наблюдаемую величину $\gamma = 1.8$ для малых расстояний и спектр до $\gamma = 3$ для больших расстояний, является заманчивым, поскольку с правильными амплитудами оно могло дать хорошее согласие с наблюдаемыми данными (11.63a) и (11.63b). Случай $n = 0$ означает, что флуктуации имеют равные амплитуды для всех волновых чисел (т. е. это белый шум), и соответствует равномерному случайному распределению галактик, для которого скучивание отсутствует и $\xi(r) \equiv 0$. Оно также известно как распределение Пуассона.

Вычислительные эксперименты с 20 000 тел начинаются от распределения с $n = 0$, показанного на рис. 11.12, который пред-



Рис. 11.12. Начальное условие для эксперимента по галактическому скучиванию, использующего 20 000 частиц. Каждая точка представляет галактику, и эти точки первоначально расположены согласно равномерному случайному распределению в квадрате (распределение Пуассона). ([124], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

ставляет проекцию трехмерного распределения на плоскость. Каждая точка обозначает моделируемую галактику и имеет начальную скорость, соответствующую чистому расширению Хаббла (11.42). На рис. 11.13 показано распределение галактик после эволюции с параметром расширения $a = 9.9$. Внизу представлен случай открытой Вселенной, которая стартовала с плотности $\Omega = 0.775$ и эволюционировала к плотности $\Omega = 0.26$. Более комковатое распределение наверху рисунка соответствует распределению для случая плоской Вселенной, начинающей с тех же начальных положений, для которой $\Omega = 1$ при всех значениях времени. Рисунки выполнены в одном масштабе в сопутствующих координатах, но расчетные области на рис. 11.13, конечно, изображают расстояние в пространстве, в 9.9 раз большее, чем расстояние на рис. 11.12, вследствие расширения Вселенной.

Двухточечная корреляционная функция для открытой Вселенной показана на рис. 11.14. Она получена в некоторый более поздний момент времени, чем на рис. 11.13, после расширения при $a = 19.3$, когда плотность понизилась до $\Omega = 0.15$. Темные кружки (ансамбль 3) представляют усреднение по трем вычислительным экспериментам с 1000 телами, использующим машинный код Арсета (см. разд. 11.2.1 и 11.2.2, а также [6]), модифицированный Эфстатию. Светлые кружки (ансамбль 5) дают усреднение по двум вычислительным экспериментам с 20 000 телами, использующим описанный здесь код Р³М Иствуда, Хокни, Лоуренса и Эфстатию. Данные моделирования с 1000 телами дают хорошую статистику для расстояний от $0.002R_s$ до $0.05R_s$ ($R_s = 1/4.4$ стороны области). С другой стороны, моделирование 20 000 тел дает результаты для расстоя-

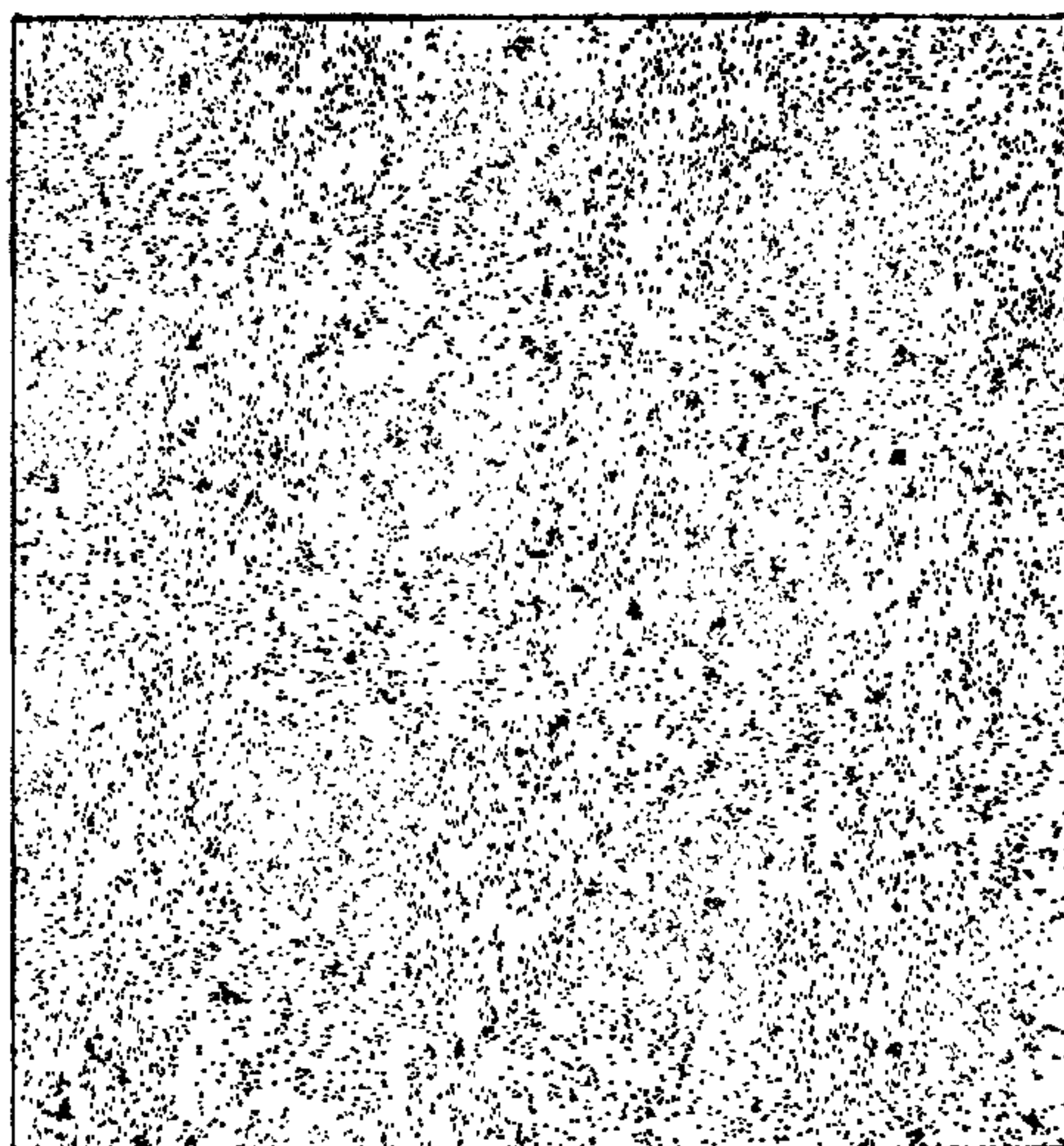
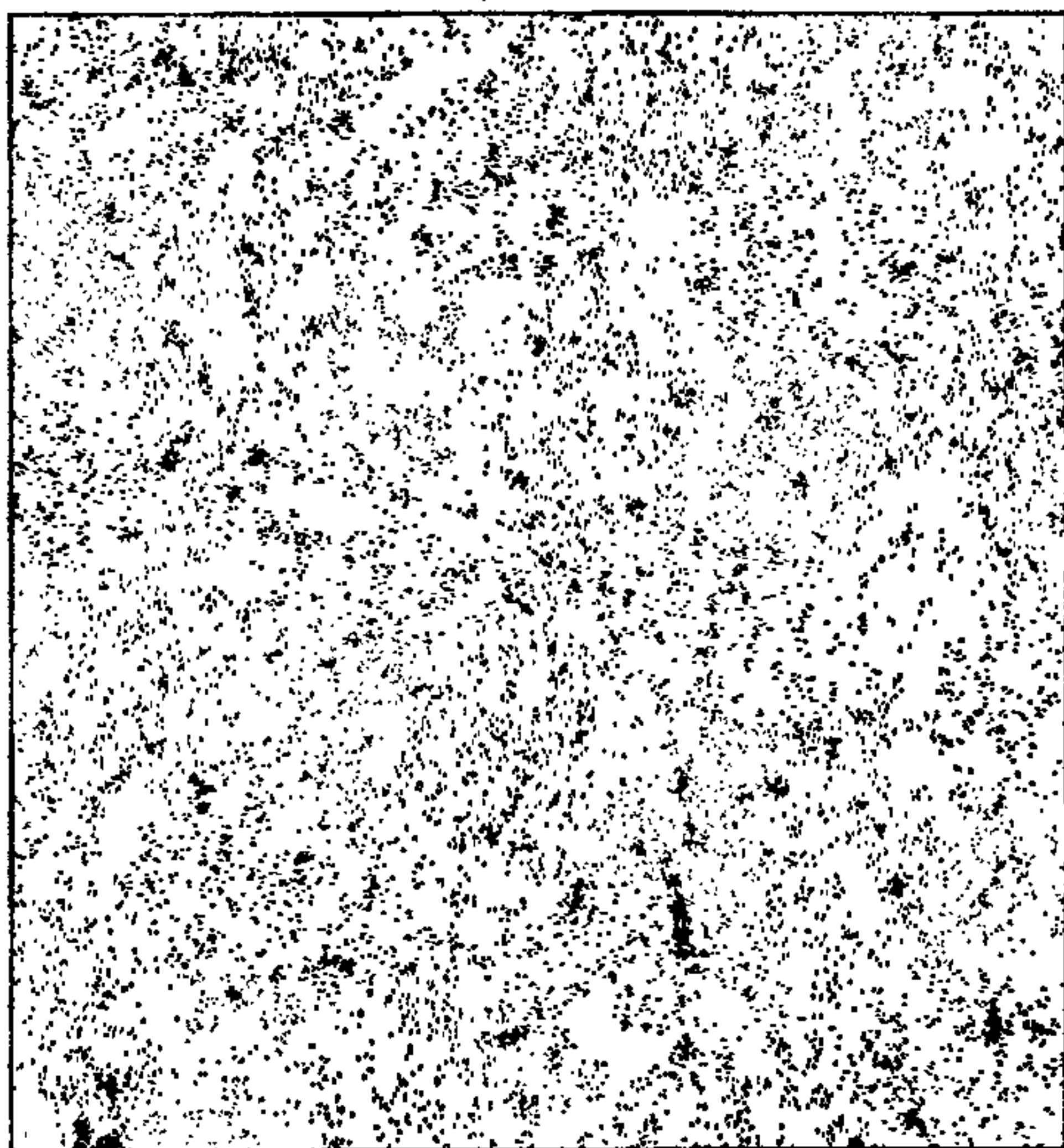


Рис. 11.13. Распределение галактик после расширения с коэффициентом $a = 9.9$, начавшегося из распределения на рис. 11.12. Внизу — открытая Вселенная, когда параметр плотности $\Omega = 0.26$, и наверху — плоская Вселенная с $\Omega = 1.0$. ([124], с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

ний до 0.5 стороны области. При расстояниях, меньших чем $0.025R_s$, на данные моделирования с 20 000 телами влияет конечный диаметр $d_m \approx 0.007R_s$, заданный для каждой галактики, и поэтому они игнорируются. Действительно, расчеты 20 000 тел можно провести с $d_m = 0.0024R_s$ (рис. 11.14). Данные о моделировании 1000 и 20 000 тел хорошо стыкуются, и Эфстатию и Иствуд [124] получают следующее приближение наименьших

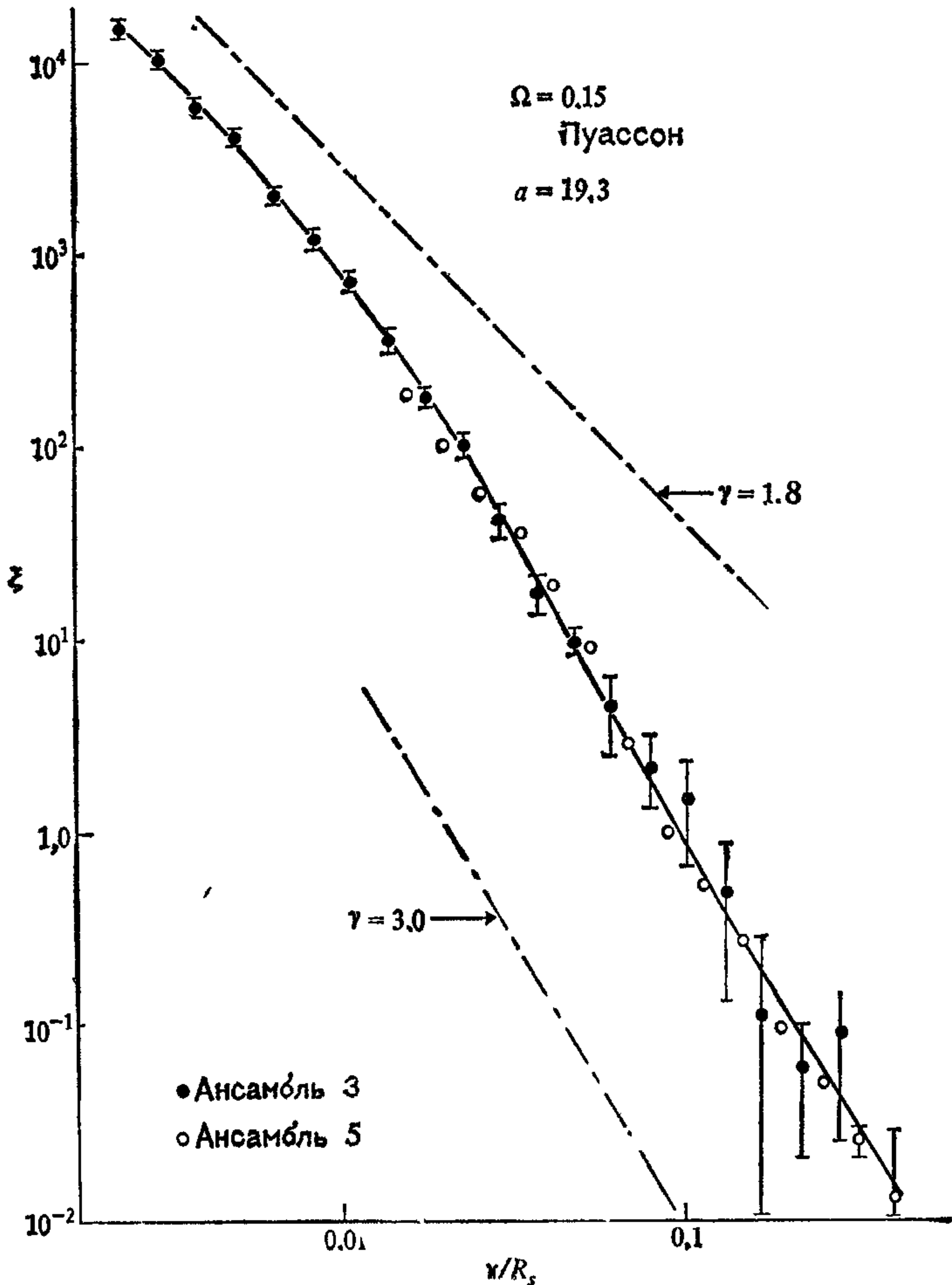


Рис 11.14 Двухточечная корреляционная функция $\xi(x)$ как функция расстояния x для ансамбля открытых вселенных после расширения с $a = 19.3$, когда $\Omega = 0.15$. Темные кружки — результат экспериментов с 1000 телами, а светлые кружки — эксперименты с 20 000 телами. Пунктирные линии показывают изменения вида $\xi(x) \sim x^{-\gamma}$, сплошная линия является кубическим приближением к результатам моделирования [124] (С любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society)

квадратов к кубическому уравнению, которое изображено жирной кривой на рис. 11.14:

$$\log \xi(r) = -3.000 - 2.550 \log y + 0.577 (\log y)^2 + 0.200 (\log y)^3, \quad (11.67a)$$

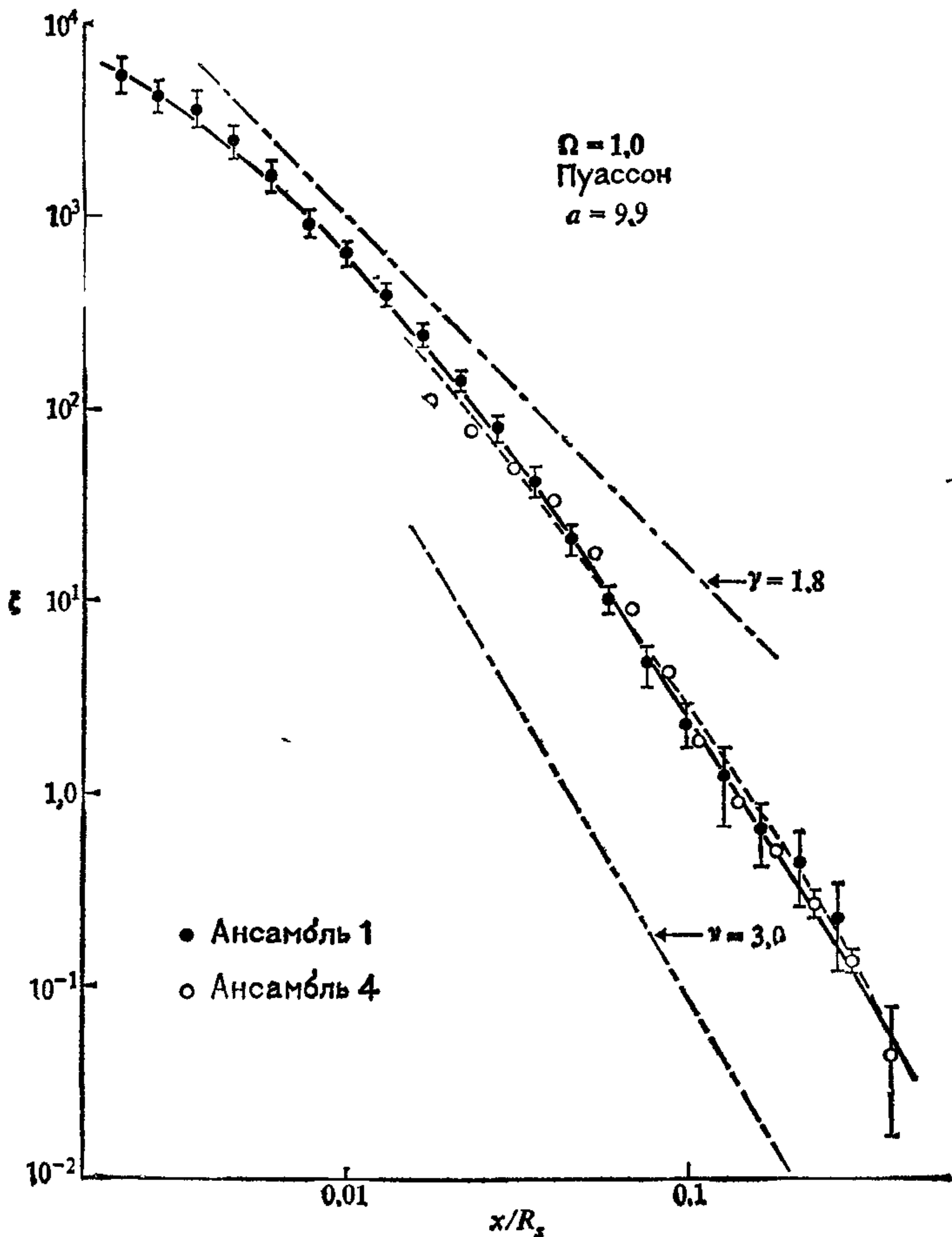


Рис 11.15 Двухточечная корреляционная функция для ансамбля плоских вселенных после расширения с $a = 9.9$. Другие параметры см. на рис. 11.14. ([124] с любезного разрешения Royal Astronomical Society Monthly Notices, © Royal Astronomical Society.)

где $0.002 \ll y = r/R_s \leq 0.5$, откуда

$$\gamma = -d(\log \xi)/d(\log y) =$$

$$= +2.550 - 1.154 \log y - 0.600 (\log y)^2. \quad (11.67b)$$

Аналогичные результаты для плоской Вселенной Эйнштейна — де Ситтера ($\Omega = 1$) показаны на рис. 11.15. Ансамбль 1 представляет собой усреднение по шести вычислительным экспериментам с 1000 телами, включая результаты, полученные Арсетом и Фоллом, которые использовали различные версии кода RR. Ансамбль 4 есть усреднение по двум вычислительным экспери-

ментам с 20 000 телами. Для корреляционной функции был получен следующий результат [123]:

$$\log \xi(r) = -2.537 - 2.957 \log y + 0.092 (\log y)^2 + 0.119 (\log y)^3, \quad (11.68a)$$

$$\gamma = +2.957 - 0.184 \log y - 0.357 (\log y)^2. \quad (11.68b)$$

Такое согласие между многими различными вычислительными режимами, начинающимися с различных псевдоравномерных случайных положений и использующими три вычислительных кода, основанных на двух совершенно различных алгоритмах и написанных независимо различными программистами, вызывает значительное доверие к результату, выраженному соотношениями (11.67) и (11.68).

Основное различие между случаями открытой и плоской Вселенной состоит в том, что корреляция уменьшается более быстро в случае открытой Вселенной с γ , принимающим значения 2.5, 3.1 и 2.5 для $y = 0.01$, 0.1 и 1 соответственно по сравнению с $\gamma = 1.9$, 2.8 и 3.0 для случая плоской Вселенной. Это также убеждает в том, что последний случай $\Omega = 1$ удовлетворяет теоретическим предельным значениям $\gamma = 1.8$ и 3.0, данным уравнениям (11.66b, c). Однако положение излома между этими двумя градиентами должно быть оценено при $\xi \approx 50$, а не при наблюдаемом значении $\xi = 0.3$.

Оба соотношения (11.67) и (11.68) для двухточечной пространственной корреляции можно превратить в эквивалентную двухточечную угловую корреляционную функцию $W(\theta)$, измеренную наблюдателями. Это требует численного решения уравнения Лимбера [256]. После преобразования результаты можно нанести на тот же самый график, где изображены экспериментальные данные, и это демонстрируется на рис. 11.16. Из рисунка ясно, что случай открытой Вселенной согласуется с наблюдениями только для самых малых угловых расстояний. На большем диапазоне расстояний модель открытой Вселенной явным образом противоречит наблюдениям. Можно сказать, что плоская Вселенная должна согласовываться с наблюдениями для угловых расстояний, меньших чем 1° . Этот случай грубо отвечает пределу ранних расчетов 1000 тел, которые в любом случае обладали большим шумом. Поэтому некоторые авторы [122, 131, 148] пришли к выводу, что Вселенная с $\Omega = 1$ совместима с наблюдениями. Изложенные здесь результаты Эфстату и Иствуда [124] с 20 000 телами увеличили диапазон и точность результатов с $\Omega = 1$ до момента, когда имеется явное разногласие между случаем $\Omega = 1$ и наблюдением. Кроме того, рассмотрение распределений скоростей вне галактических расстояний, равных половине стороны расчетной области, показывает отличное согласие с линейной теорией. Можно сказать

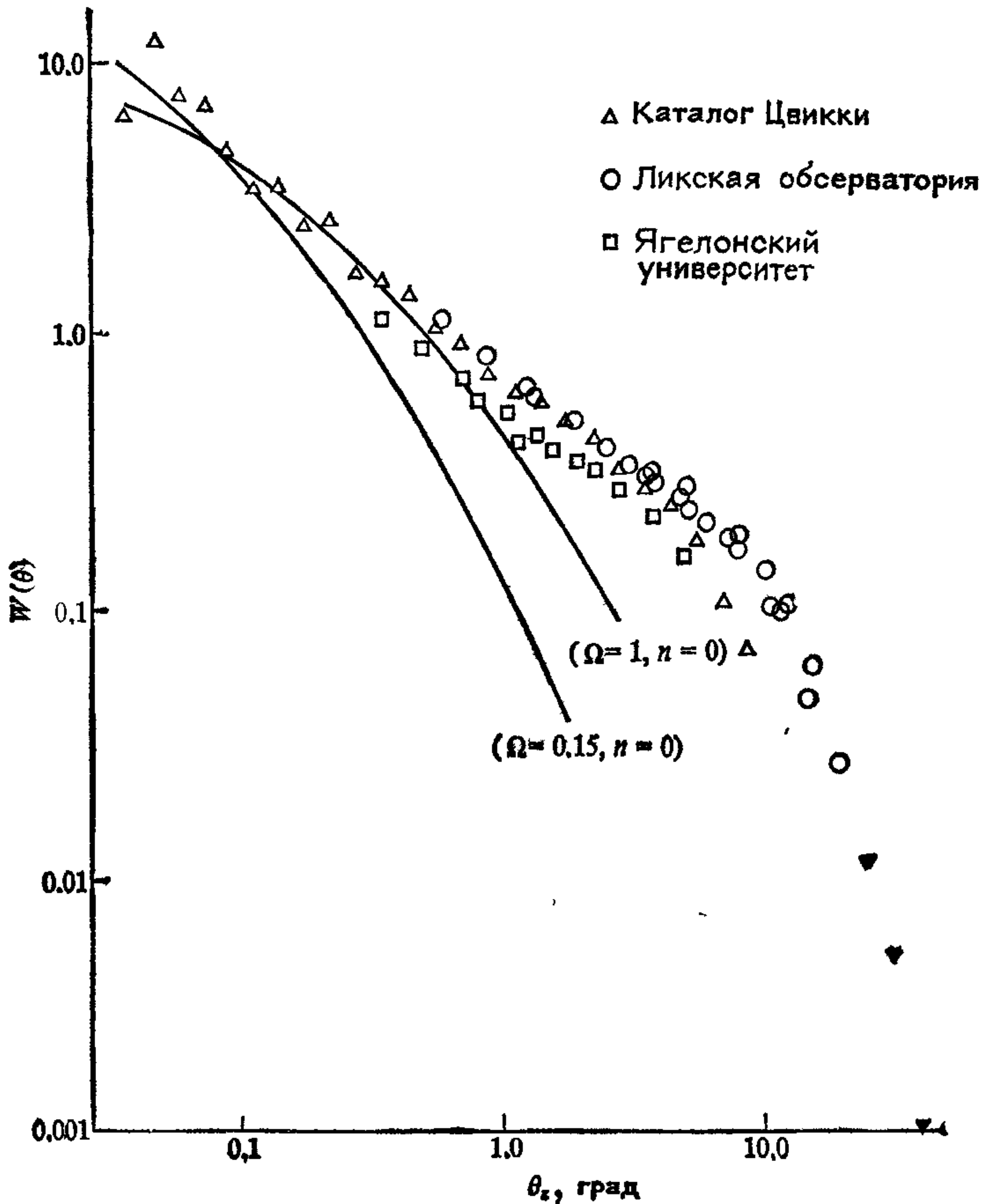


Рис. 11.16. Сравнение измерений двухточечной корреляционной функции (сплошные линии) в численном моделировании, выраженной через эквивалентное угловое расстояние на небе, с наблюдениями. ([124, 152], *The Astrophysical Journal*, published by the University of Chicago Press; © 1977 The American Astronomical Society; and *Royal Astronomical Society Monthly Notices*, © Royal Astronomical Society.)

совершенно уверенно, что вычислительные эксперименты не показывают излома в наклоне вблизи 10° , который является такой резко выраженной особенностью наблюдений. С другой стороны, обнаружилось, что трехточечная корреляционная функция ζ должна быть связана с двухточечной корреляционной функцией соотношением $\zeta = Q\xi^2$ с $Q \sim 1$. Это хорошо согласуется с наблюдаемым значением $Q = 1.3 \pm 0.2$ [152] и усиливает аргумент в пользу того, что галактики скучиваются иерархично.

Возможно, что другой спектр для начальных флуктуаций плотности, например отрицательные значения n или другой вид

$S(k)$ (например, функция $\exp(k/k_m)k^n$, которая учитывает адиабатическую составляющую), может привести результаты моделирования N тел в лучшее согласие с наблюдением. За неимением этого, по-видимому, уравнения Фридмана с $\Lambda = 0$ не дают удовлетворительного описания расширяющейся Вселенной. Введение отличной от нуля космологической постоянной ($\Lambda \neq 0$), приводит к бесконечным возможностям (см. [369]). Задача вычислительного эксперимента на модели с $\Lambda \neq 0$ не является более сложной, чем с $\Lambda = 0$, но дополнительный параметр делает число случаев, которые необходимо рассмотреть, обескураживающе большим. Если мы ставим условие $\Lambda = 0$, то оказывается, что Вселенная с большой плотностью ($\Omega \approx 1$) имеет наибольший шанс удовлетворить наблюдениям. Поскольку, однако, наблюдаемое вещество во Вселенной достигает только $\Omega \approx 0.1$ [369, рис. 3], то это оставляет значительную проблему объяснить природу «скрытой» массы.

Мы должны закончить этот раздел предостерегающим замечанием о том, что никоим образом не установлено, что представление Вселенной множеством дискретных масс, как это сделано в вычислительных экспериментах N тел, является хорошей моделью для реальной Вселенной. Этот вопрос и другие, связанные со справедливостью вычислительных экспериментов N тел, подробно обсуждены Пиблсом [308] и Фоллом [130]. Вычислительные эксперименты N тел должны стать хорошим описанием скучивания, если в некоторой эпохе галактики были слабо скучены и поэтому действовали как точечные частицы. Однако возможно также, что скучивание было сильно нелинейным на достаточно малых масштабах даже в ранних эпохах до образования галактик. К тому же если (как это допускается большей частью данных) большая часть массы Вселенной содержится в невидимой компоненте (скрытая масса), то нет гарантии того, что галактики когда-либо действовали как точечные частицы. Если это так, то результаты, полученные из экспериментов N тел, неприменимы к реальной Вселенной.

Твердое тело, жидкости и фазовые превращения

12.1. ВВЕДЕНИЕ

12.1.1. Молекулярная динамика

К числу главных достижений вычислительного эксперимента относится изучение конденсированных веществ, особенно жидкостей и фазовых превращений. Соответствующие результаты получены путем интегрирования движения некоторого числа характерных частиц рассматриваемого вещества, а метод назван «молекулярной динамикой» (МД), хотя большая часть работы и проводится с системами, о которых никак не скажешь, что они образуют молекулы. Движение таких частиц обычно предполагается классическим и подчиняющимся законам Ньютона, при этом силы, как правило, считают центральными, действующими вдоль линии, соединяющей пары частиц. Ни одно из этих ограничений никак не связано с самой идеей молекулярной динамики, и естественно ожидать, что этот метод можно разработать для неклассического движения и нецентральных многочастичных сил. Круг вопросов, обсуждаемых в настоящей главе, будет, однако, ограничен моделированием классических центральных сил.

Веществами, которые с самого начала обратили на себя наибольшее внимание, были простейшие атомные системы (инертные газы аргон, неон, криптон и др.), простейшие заряженные системы (галоидные соли щелочных металлов KCl, NaCl, KI и т. д.) и простейшая плотная плазма, состоящая только из одного сорта заряженных частиц, например электронов. В этой главе в качестве примеров будут подробно рассмотрены последние два случая. С точки зрения вычислителя моделирование полностью определено, если задана зависимость силы взаимодействия частиц от расстояния. Как и математику-прикладнику или инженеру, ему совершенно ни к чему подробно разбираться в физическом обосновании этого закона изменения силы — для этого есть физик. Если закон взаимодействия известен, роль вычислителя сводится к тому, чтобы разработать робастную модель на компьютере с известной точностью. Поскольку характер задачи моделирования определяется, главным образом, характерными особенностями силы взаимодействия, мы рассмотрим сейчас кратко те законы взаимодействия, которые используются при моделировании методом молекулярной дина-

мики. Фигурирующие в тексте времена́ счета и возможности моделей будут относиться к компьютерам с производительностью, как у IBM 360/195 или CDC 7600. Библиография работ по молекулярной динамике и смежным областям вплоть до 1973 г., насчитывающая 392 наименования, имеется в работе [144].

12.1.2. Закон взаимодействия

Самые первые вычислительные эксперименты проводились Ордером и Вейнрайтом [14] с абсолютно твердыми сферами, ведущими себя подобно системе бильiardных шаров. В этом случае, пока шары не соприкасаются, они никак не взаимодействуют; если же они соприкасаются, их движение мгновенно изменяется. Для моделирования твердых сфер применяются специальные методы, включающие в себя расчет времени следующего «атомного» столкновения, и читатель может ознакомиться с ними подробно в указанной выше статье. Методы частица — сетка, которые предназначены для моделирования гладко меняющихся дальнедействующих сил, не пригодны для моделирования систем твердых сфер.

Для моделирования реальных веществ требуется более реалистичный потенциал парного взаимодействия V , и, как было установлено в работах [318] и [279], свойства конденсированных инертных газов хорошо предсказывает потенциал Леннарда-Джонса, содержащий обратные степени расстояния. Хотя в первоначальной публикации Джонс [224] остановился на других показателях степени, в более поздних работах был принят стандартный потенциал с 6-й и 12-й обратными степенями:

$$V(r_{ij}) = \epsilon \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_0}{r_{ij}} \right)^6 \right] = \quad (12.1a)$$

$$= 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (12.1b)$$

где r_{ij} — расстояние между i -м и j -м атомами, а ϵ и r_0 обозначают соответственно глубину потенциальной ямы и равновесное расстояние между двумя атомами. Обе формы записи (12.1a) и (12.1b) эквивалентны, при этом $r_0 = 2^{1/6}\sigma = 1.12\sigma$. Сила обращается в нуль при $r = r_0$, а потенциал равен нулю при $r = \sigma$. На больших расстояниях этот потенциал является притягивающим (как r^{-6}), а на малых расстояниях — отталкивающим (r^{-12}). В большинстве модельных расчетов взаимодействие не учитывается, когда расстояние между атомами превышает радиус обрезания $r_1 \simeq 3r_0$. Если силу, действующую на частицу, находят путем суммирования вкладов от всех взаимодействующих

с ней частиц, как, например, в традиционном РР-методе, то за счет такого обрезания объем вычислений будет пропорционален полному числу частиц N . В рамках этого приближения можно моделировать системы, состоящие из нескольких тысяч атомов (например, [343]).

Если система содержит электрически заряженные атомы (т. е. ионы), потенциал взаимодействия на больших расстояниях является кулоновским с законом пропорциональности r^{-1} . В случае такого медленно меняющегося взаимодействия обрезание недопустимо и при вычислении силы нужно учитывать взаимодействие между всеми N^2 парами частиц. Кроме того, при использовании периодических граничных условий необходимо учитывать влияние всех зеркальных зарядов. Обычно это делается с помощью эффективного метода суммирования [121]; (см. [341, с. 298—305]). Тем не менее учет всех взаимодействий и их изображений РР-методом представляет собой тяжелую вычислительную задачу, при этом размеры системы ограничивают несколькими сотнями частиц (например, [319] и [255]). Наличие в ионной задаче и дальнедействующей, и короткодействующей сил делает ее идеально подходящей для решения Р³М-методом ([189] и гл. 8), и это позволяет моделировать системы, состоящие из 10 648 ионов, с шагом по времени 26 с. Метод Р³М использован в тех задачах, которые рассматриваются далее в этой главе. Помимо статических характеристик, изучавшихся в вышеуказанных работах, в численных экспериментах исследованы коллективные динамические свойства жидкого RbBr [89] и жидкого NaCl [11]. Кроме того, в численных экспериментах измерялись [83] характеристики переноса в расплаве галоидов щелочных металлов. Обзор работ по структуре и динамике простых ионных жидкостей содержится в [305].

Простейший потенциал взаимодействия, который применяется в изучении ионных систем, содержит обратные степени расстояния [162]:

$$V(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left[1 + \frac{\text{sgn}(q_i q_j)}{p+1} \left(\frac{s_i + s_j}{r_{ij}} \right)^p \right], \quad (12.2)$$

где q_i и s_i — соответственно заряд и радиус i -го иона, а показатель p определяет жесткость отталкивающей части потенциала. Эта форма потенциала, как и потенциала Леннарда — Джонса, обладает тем достоинством, что она масштабируется (см. разд. 12.3.1) и ее можно изучать теоретически, сохранив при этом все существенные свойства конденсированной ионной системы. Если требуется очень близкое согласие между физической системой и моделью, следует отдать предпочтение более сложному потенциалу в форме Борна — Хаггинза — Мейера [44,

208, 43]:

$$V(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + B_{ij} \exp(-\alpha_{ij} r_{ij}) - \frac{C}{r_{ij}^6} - \frac{D_{ij}}{r_{ij}^8}. \quad (12.3)$$

Для галоидов щелочных металлов значения постоянных B , C , D и α получены в работах [138, 313]. Эти постоянные подбирались так, чтобы получалось наилучшее согласие с упругими постоянными соответствующих веществ.

Обе вышеприведенные формы взаимодействия предполагают, что электронные оболочки иона жестко связаны с положительным ядром и что центр всех электронных оболочек строго совпадает с центром ядра. В реальных ионах, в особенности в больших ионах типа фтора и иода, электрическое поле других ионов вызывает смещение центра отрицательного заряда относительно центра положительного заряда и ион, как говорят, становится поляризованным. Для того чтобы учесть этот эффект, необходимо рассчитывать движение легких электронных оболочек и движение тяжелых положительных ядер отдельно. С этой целью была разработана [341] модель оболочек, в которой безмассовая электронная оболочка связана с ядром посредством «пружинки». На каждом шаге по времени ядра движутся по законам Ньютона под действием принятых сил, как в модели жестких ионов. Затем методом последовательных приближений электронные оболочки перемещают в положение, в котором действующие на них силы обращаются в нуль. Такой способ введения поляризации обычно удлиняет время счета одного шага в несколько раз. В данной оболочечной модели ядра рассматриваются как точечные заряды без какого бы то ни было отталкивания на малых расстояниях, тогда как оболочки взаимодействуют в соответствии с потенциалом типа Фами — Този, т. е. (12.3). Численные значения параметров, входящих в эти выражения, и силовые постоянные получают путем подгонки частотных характеристик соответствующих кристаллов. Для галоидов щелочных металлов эти значения даны в работе [341].

Однокомпонентная плазма (ОКП) представляет собой систему, в которой заряженные частицы одного сорта движутся на однородном заряженном фоне противоположного знака. Системы такого рода существуют в плотном звездном веществе белых карликов [149]. В этом случае движение электронов ограничивается принципом запрета Паули, что выглядит как движение на однородном заряженном фоне. При определенных условиях плотности и температуры движение тяжелых положительных ионов (например, ^{56}Fe) можно трактовать как классическую ОКП [55]. Предполагается, что при достаточно высоких плотностях и давлениях эта жидкость кристаллизуется, и условия, при которых этот переход происходит, имеют немаловажное значение для теории таких звезд. Очевидно, что в данном слу-

чае лабораторные эксперименты невозможны, и единственным способом экспериментального изучения является моделирование на машине. Потенциал взаимодействия берется самый простой из рассмотренных ранее, а именно отталкивающий потенциал одинаковых ионов с зарядом q :

$$V(r) = q^2 / (4\pi\epsilon_0 r). \quad (12.4)$$

Такая система была промоделирована [313] с помощью метода Монте-Карло [281]. В этом методе с помощью датчика случайных чисел разыгрываются многие тысячи возможных конфигураций системы, и путем усреднения по конфигурациям находят равновесные характеристики. В отличие от молекулярной динамики динамика системы не моделируется и зависящие от времени характеристики не могут быть измерены.

В работе [187] был описан метод молекулярной динамики для моделирования ОКП в двумерной системе точечных зарядов, движение которых возможно только в одной плоскости. Такая система может быть реализована в лабораторных условиях двумя способами: либо на поверхности жидкого гелия (что было предсказано в работе [85] и подтверждено экспериментально [53]), либо в виде инверсионного слоя в кремниевом МОП-транзисторе [354]. История исследований двумерной ОКП интересна сама по себе как пример взаимодействия вычислительной науки с теорией и лабораторным экспериментом. В 1974 г. Платцман и Фукуяма [312] предсказали, причем на основании весьма правдоподобных соображений, что переход из твердой фазы в жидкую будет происходить при отношении потенциальной энергии к кинетической Γ_0 , равном примерно трем. Были предприняты энергичные усилия по экспериментальной проверке этой теории, однако первые результаты были получены не в лаборатории. Они поступили из вычислительных экспериментов Хокни и Брауна [187], проведенных с помощью новой двумерной P³M-программы [189], модифицированной для точечных зарядов в плоскости. В этом молекулярно-динамическом эксперименте был обнаружен фазовый переход при $\Gamma_0 = 95 \pm 2$, т. е. при значении, почти в 34 раза большем, чем предсказывалось теорией, что и побудило теоретиков вновь вернуться к этой задаче.

В итоге в 1978 г. Таулесс представил новую теорию [367], основанную на другом физическом механизме плавления, которая предсказывала фазовый переход при $\Gamma_0 = 78$. С учетом неточностей в значениях некоторых физических констант это вполне согласуется с молекулярно-динамическим экспериментом. Данная история закончилась уже после того, как возможности лабораторных экспериментов возросли до такой степени, что позволили провести физические измерения. Тем не менее тот эксперимент, который было просто осуществить на машине,

оставался еще очень трудным в лаборатории. И только в 1979 г., почти через четыре года после первого наблюдения плавления в вычислительных экспериментах, фазовый переход впервые наблюдался в лаборатории [151] при $\Gamma_0 = 137 \pm 15$. Примерно в то же время в расчетах методом Монте-Карло [140] был обнаружен фазовый переход плавления первого рода при $\Gamma_0 = 125 \pm 15$. Указанные результаты вполне согласуются между собой, но значительно отличаются от прежних теоретических и молекулярно-динамических результатов. Очевидно, что для объяснения этих расхождений необходимы дальнейшие исследования и в лабораторном эксперименте, и в вычислительном. В разд. 12.2 в качестве примера конкретной задачи взят численный эксперимент по двумерной однокомпонентной плазме.

Молекулярная динамика широко применяется также для изучения фазовых превращений, в частности плавления и стеклования, в нейтральных и ионных системах. За моделированием плавления двумерного кристалла Леннарда-Джонса [93] последовало изучение роли дислокаций в фазовых переходах в трехмерных системах Леннарда-Джонса, состоящих из 396 атомов [219, 99, 91]. Это молекулярно-динамическое исследование существенно способствовало развитию единой теории плавления и стеклования [92], поскольку моделирование частицами давало возможность проверять на микроскопическом уровне многие из основных идей теории таким способом, который невозможен в лаборатории. В разд. 12.3 в качестве конкретной задачи мы рассмотрим исследование методом молекулярной динамики влияния относительного радиуса и жесткости на плавление ионных микрокристаллов [18, 17]. В этом исследовании численный эксперимент дает возможность менять жесткость ионов независимо от их радиусов, что опять-таки невозможно в лаборатории. Кроме того, поскольку различные теории предсказывают разную зависимость точки плавления от жесткости, такие численные эксперименты обеспечивают новую и независимую проверку теорий плавления (см. разд. 12.3.7).

12.1.3. Интегрирование по времени

После того как мы обсудили вычисление сил, рассмотрим интегрирование по времени. Для интегрирования уравнений движения Ньютона чаще всего используют центрированную разностную схему с перешагиванием ([180, 57] и гл. 4), иногда называемую применительно к рассматриваемым задачам алгоритмом Верлета [377], либо некоторую ее разновидность, принадлежащую Биману [27]. Последняя схема дает те же самые орбиты, что и схема с перешагиванием, но имеет другое определение скорости; однако вычисления по этой схеме существенно дороже, чем по простой схеме с перешагиванием. (Алго-

ритм Бимана анализировался в разд. 4.5.) Некоторые исследователи пользуются методами предиктор-корректор более высокого порядка [318], что значительно дороже по машинному времени и памяти. По этой причине они целесообразны лишь для достаточно небольших систем, состоящих из примерно 100 частиц. Для больших систем из многих тысяч частиц единственной практически реальной схемой является схема с перешагиванием, поскольку она требует минимум как машинного времени, так и памяти.

12.2. ДВУМЕРНАЯ ЭЛЕКТРОННАЯ ПЛЕНКА

Однокомпонентная плазма, движение которой удерживается в одной плоскости, называется по-разному — двумерной электронной жидкостью, двумерной моделью твердого тела или пленкой. В данном случае задачу моделирования можно сформулировать одним предложением: создать модель системы одинаковых точечных зарядов, движущихся в двух измерениях сквозь однородно заряженный нейтрализующий фон и отталкивающихся друг от друга по кулоновскому закону обратных квадратов. Это выражается следующим уравнением:

$$m \frac{d^2 \mathbf{r}_i}{dt^2} = - \sum_j' \frac{q^2}{4\pi\epsilon_0 |\mathbf{r}_{ij}|^2} \mathbf{r}_{ij}, \quad (12.5)$$

где m и q — соответственно масса и заряд частиц, а \mathbf{r}_{ij} — вектор, соединяющий i -ю частицу с j -й. Суммирование проводится по всем частицам j , кроме частицы i ¹⁾, и $\hat{\mathbf{r}}_{ij}$ обозначает *единичный* вектор в направлении от i к j . В уравнение (12.5) нет необходимости включать однородный заряженный фон, поскольку в силу симметрии это никак не скажется на силе, действующей на частицу.

12.2.1. Безразмерные уравнения

Если n представляет собой двумерную (или поверхностную) плотность, то единицей характерного расстояния будет

$$r^* = (\pi n)^{-1/2}. \quad (12.6)$$

Это расстояние равняется радиусу такого диска, который — если его связать с каждой частицей — даст полную площадь рассматриваемой системы. Следовательно, оно приблизительно равно половине среднего расстояния между частицами. Для того чтобы получить единицу характерного времени, рассмотрим среднюю скорость частиц v^* , находящихся в тепловом равновесии при

¹⁾ Это отмечается штрихом у знака суммы. — Прим. перев.

температуре T . Приравнивая тепловую энергию, приходящуюся на частицу, которая равна $1/2 k_B T$, кинетической энергии в одном измерении, имеем

$$\frac{1}{2} m v^{*2} = \frac{1}{2} k_B T, \quad (12.7a)$$

$$v^* = (k_B T / m)^{1/2}. \quad (12.7b)$$

Поскольку в данном случае движение двумерно, среднеквадратичная скорость частиц равна $\sqrt{2} v^*$. Тем не менее в качестве характерной для данной задачи единицы скорости принимается v^* . Тогда характерное время t^* равняется

$$t^* = r^* / v^* \quad (12.8)$$

и представляет собой среднее время, за которое частица с тепловой скоростью проходит примерно половину межчастичного расстояния.

Если расстояние и время, т. е. r и t , выразить теперь в единицах r^* и t^* , то получим уравнение (12.5) в безразмерном виде:

$$\frac{d^2 \mathbf{r}_{ij}}{dt^2} = -\Gamma_0 \sum_j' \frac{1}{|\mathbf{r}_{ij}|^2} \hat{\mathbf{r}}_{ij}, \quad (12.9a)$$

где

$$\Gamma_0 = \left(\frac{q^2}{4\pi\epsilon_0 r^*} \right) / k_B T = \frac{\pi^{1/2} q^2}{4\pi\epsilon_0 k_B} \frac{n^{1/2}}{T}. \quad (12.9b)$$

Параметр Γ_0 является мерой отношения потенциальной энергии системы к кинетической. Теперь ясно, почему важно выразить исходное уравнение (12.5) в безразмерном виде (12.9) — выясняется, что для описания системы в действительности необходимо не несколько параметров (m, q, n, T), как можно было бы думать, а всего лишь один — величина Γ_0 . Вычислительный эксперимент, выполненный при конкретных значениях q, n, T , можно с равным успехом считать проведенным с любыми другими q, n, T , имеющими то же самое значение Γ_0 . Особую важность имеет взаимозаменяемость плотности и температуры. Выражение (12.9b) показывает, что, например, уменьшить плотность в 100 раз — это все равно, что увеличить в 10 раз температуру. Большинство проводимых лабораторных экспериментов приходится проводить при постоянной температуре (жидкий гелий находится при ~ 1 К), варьируя плотность, однако в вычислительном эксперименте удобнее изменять температуру при постоянной плотности (мы сохраняем неизменным число частиц в постоянной расчетной области). Выражение (12.9b) показывает, что не имеет значения, как именно поступать, и что конечный результат надлежит оценивать, исходя из значения Γ_0 . Конечно, сам факт, что система характеризуется только одной без-

размерной переменной, очень сильно упрощает вычислительную задачу, поскольку нам нужно только прогнать программу достаточное число раз, чтобы определить соответствующие характеристики в зависимости от одного параметра, а не нескольких. Говорят, что взаимодействия, которые можно трактовать таким образом, являются масштабируемыми.

12.2.2. Выбор шага по времени

Для того чтобы выбрать в модели шаг по времени, необходимо выделить в физической системе самую высокую частоту колебаний ω_1 и удовлетворить условию устойчивости (см. разд. 4.4.1, формула (4.42)):

$$\omega_1 DT < 2. \quad (12.10)$$

Полный анализ мод колебаний электронной жидкости весьма сложен (см., например, [302, 26]) и для нашей конкретной цели фактически не нужен. Нам необходимы лишь верхние оценки, которых достаточно для выбора разумного шага по времени и выявления его зависимости от основных параметров. В конечном счете выбор шага по времени определяется орбитами частиц и сохраняемостью энергии.

Для того чтобы оценить наибольшую частоту, которая будет возникать под действием максимального градиента силы, представим себе следующую простую одномерную картину. Градиент максимален в случае, когда соседние частицы движутся навстречу друг другу в точку наибольшего сближения. Уравнения движения такой пары частиц имеют вид (считаем $x_2 > x_1$):

$$m_1 \frac{d^2 x_1}{dt^2} = -F(x_2 - x_1), \quad (12.11a)$$

$$m_2 \frac{d^2 x_2}{dt^2} = F(x_2 - x_1), \quad (12.11b)$$

где F — абсолютная величина силы отталкивания.

Умножая первое уравнение на m_2 , второе на m_1 и вычитая одно из другого, получаем уравнение для относительного движения:

$$m^* \frac{d^2 r}{dt^2} = F(r), \quad (12.12)$$

где $r = x_2 - x_1$ — расстояние между частицами, а $m^* = m_1 m_2 / (m_1 + m_2)$ — приведенная масса. Вообще говоря, сила $F(r)$ нелинейна, и для получения эффективной частоты мы должны рассмотреть уравнение относительно малого возмущения, введенного в (12.12). Положим $r = r + r'$, где r' мало; тогда

$$m^* \frac{d^2 (r + r')}{dt^2} = F(r + r') \simeq F(r) + r' \frac{dF}{dr} + \dots, \quad (12.13)$$

где мы разложили правую часть в ряд Тейлора, удержав только первые два члена. Вычитая (12.12) из (12.13), получаем для возмущения линеаризованное уравнение

$$\frac{d^2 r'}{dt^2} = -\omega^2 r', \quad \omega = \left(-\frac{1}{m^*} \frac{dF}{dr} \right)^{1/2}. \quad (12.14)$$

Если dF/dr отрицательно, как это имеет место для отталкивания на малых расстояниях, это уравнение описывает простые гармонические колебания с частотой ω . Максимальная частота будет наблюдаться для малых масс и более крутых (другими словами, более жестких) зависимостей отталкивания от расстояния между частицами.

В случае электронной пленки имеем $m^* = m/2$ и $F(r) = q^2/(4\pi\epsilon_0 r^2)$, и, следовательно,

$$\omega = \left(\frac{q^2}{\pi\epsilon_0 m r^3} \right)^{1/2}. \quad (12.15)$$

Это выражение показывает, что, по мере того как расстояние наибольшего сближения r стремится к нулю, частота колебаний стремится к бесконечности. Поскольку устойчивость требует, чтобы $\omega DT < 2$, это означает, что любой шаг по времени в принципе неустойчив. Однако вероятность малых расстояний между частицами зависит от температуры и очень мала. В одномерном максвелловском распределении 98 % частиц обладают энергией, меньшей $3k_B T$, и подходят друг к другу не ближе, чем на расстояние $r = b$, которое можно примерно оценить, приравняв потенциальную энергию парного взаимодействия и кинетическую энергию:

$$\frac{q^2}{4\pi\epsilon_0 b} = \beta k_B T, \quad \beta = 3. \quad (12.16a)$$

Подставляя в (12.15), получаем

$$\omega_1 = \left(\frac{q^2}{\pi\epsilon_0 m b^3} \right)^{1/2} = \frac{8\pi\epsilon_0}{m^{1/2} q^2} (\beta k_B T)^{3/2}, \quad (12.16b)$$

и шаг по времени может быть выбран из условия

$$\omega_1 DT = \alpha_1, \quad \alpha_1 < 2, \quad (12.17a)$$

$$DT = \alpha_1 / \omega_1. \quad (12.17b)$$

Чтобы не допустить возникновения численной неустойчивости для небольшой группы частиц с энергиями больше $\beta k_B T$, необходимо ограничить силу отталкивания ее значением при $r = b$. В результате используемый закон для силы принимает вид

$$F(r) = q^2/(4\pi\epsilon_0 b^2), \quad r < b \quad (12.18a)$$

$$= q^2/(4\pi\epsilon_0 r^2), \quad r \geq b. \quad (12.18b)$$

Второе условие, которому необходимо удовлетворить, связано с тем, что интегрирование орбит для большинства частиц должно производиться достаточно точно. Большая часть частиц колеблется с частотой ω_0 , отвечающей среднему относительному расстоянию $2r^*$, где

$$\omega_0 = \left(\frac{q^2}{8\pi\epsilon_0 m r^{*3}} \right)^{1/2}. \quad (12.19)$$

Если для интегрирования уравнений Ньютона применяется схема с перешагиванием, уравнение (12.14) аппроксимируется следующим образом:

$$\frac{r'(t+DT) - 2r'(t) + r'(t-DT)}{DT^2} = -\omega^2 r'(t) + \frac{DT^2}{12} \frac{d^4 r'}{dt^4}. \quad (12.20)$$

Для колебания с частотой ω имеем

$$\frac{d^4 r'}{dt^4} = \frac{d^2}{dt^2} (-\omega^2 r') = \omega^4 r', \quad (12.21)$$

и отношение погрешности аппроксимации (последний член в (12.20)) к истинной силе (первый член в правой части) равно

$$\frac{DT^2}{12} \frac{d^4 r'}{dt^4} / (-\omega^2 r') = -\frac{1}{12} \omega^2 DT^2. \quad (12.22)$$

Если в качестве условия хорошего интегрирования средней орбиты мы примем $\omega_0 DT = \alpha_0 = 0.2$, то, согласно (12.22), относительная погрешность аппроксимации будет равна $\alpha_0^2/12 = 1/3\%$. Это значение α_0 соответствует $10\pi \simeq 31$ шагу интегрирования на период колебания, что, как установлено, является вполне приемлемым компромиссом между точностью и экономичностью (в смысле машинных ресурсов).

Условия на временной шаг можно выразить через безразмерные переменные, в результате чего они принимают следующий вид:

1. Точность при среднем расстоянии:

$$\omega_0^2 DT^2 = \frac{1}{2} \Gamma_0 \left(\frac{DT}{t^*} \right)^2 = \alpha_0^2, \quad \alpha_0 = 0.2. \quad (12.23a)$$

2. Устойчивость при самой высокой частоте

$$\omega_1^2 DT^2 = \omega_0^2 DT^2 \left(\frac{2r^*}{b} \right)^3 = \alpha_1^2, \quad \alpha_1 = 1. \quad (12.23b)$$

3. Очень мало частиц с $r < b$:

$$\left(\frac{b}{r^*} \right) = \frac{\Gamma_0}{\beta}, \quad \beta = 3 \quad (12.23c)$$

Можно показать, что эти величины связаны следующими соотношениями:

$$\left(\frac{\omega_1}{\omega_0}\right)^2 = \left(\frac{a_1}{\alpha_0}\right)^2 = \left(\frac{2r^*}{b}\right)^3 = \left(\frac{2\beta}{\Gamma_0}\right)^3. \quad (12.24)$$

Способ удовлетворить этим условиям зависит от величины Γ_0 , и можно предложить следующую процедуру:

4. Выбираем b :

$$\left(\frac{b}{r^*}\right) = \left(\frac{\Gamma_0}{\beta}\right), \quad \beta = 3. \quad (12.25a)$$

5. Выбираем α_0 :

$$\alpha_0 = \min\left[0.2, \alpha_1 \left(\frac{\Gamma_0}{2\beta}\right)^{3/2}\right], \quad \alpha_1 = 1. \quad (12.25b)$$

6. Выбираем DT :

$$DT = \frac{\alpha_0 t^*}{(\Gamma_0/2)^{1/2}}. \quad (12.25c)$$

Для случаев низких температур и высоких плотностей (большие Γ_0) шаг по времени определяется условием 1, тогда как в случаях высоких температур и низких плотностей (малые Γ_0) шаг определяется условиями 2 и 3. Перейдя к основным физическим переменным, имеем, что в первом случае шаг

$$DT = \alpha_0 8^{1/2} \varepsilon_0^{1/2} m^{1/2} / (\pi^{1/4} q n^{3/4}) \quad (12.26a)$$

зависит только от плотности, а во втором случае

$$DT = \alpha_1 q^2 m^{1/2} / (8\pi \varepsilon_0 \beta^{3/2} k_B^{3/2} T^{3/2}) \quad (12.26b)$$

зависит только от температуры.

До сих пор мы рассматривали частоты, обусловленные парным взаимодействием зарядов. В то же время во всякой системе с дальнедействующими силами возможны также частоты, отвечающие коллективному поведению всей системы. Примером такой коллективной моды служат плазменные колебания, описанные в гл. 2. Если движение зарядов трехмерно, как это имеет место в горячей плазме, о которой говорилось в гл. 9, и в расплаве галоидов щелочных металлов, частота плазменных колебаний приблизительно равна

$$\omega = \left(\frac{nq^2}{\varepsilon_0 m}\right)^{1/2}, \quad (12.27)$$

где n — число зарядов в единице объема. Эта частота не зависит от волнового числа k . Если же движение происходит в одной плоскости, как в случае двумерной электронной пленки, частота сильно зависит от k [302] и

$$\omega = \left(\frac{nq^2}{2\varepsilon_0 m}\right)^{1/2} k^{1/2}, \quad (12.28)$$

где теперь n обозначает число зарядов на единицу площади. При анализе коллективных мод колебания считаются синусоидальными и пропорциональными $\exp[i(kx - \omega t)]$. Следовательно, пространственная длина волны равна $\lambda = 2\pi/k$. Самой высокочастотной плазменной модой будет мода с длиной волны, в два раза превышающей среднее межионное расстояние ($\lambda = 4r^*$), а самую низкую частоту будет иметь волна, которая укладывается ровно один раз в расчетной области длины L . Исходя из этих двух предельных случаев, получим наименьшую возможную частоту

$$\omega_2 = \left(\frac{nq^2}{2\epsilon_0 m} \right)^{1/2} \left(\frac{2\pi}{L} \right)^{1/2} \quad (12.29a)$$

и самую высокую частоту (с $k = 2\pi/\lambda = \pi/2r^*$)

$$\omega_3 = \left(\frac{nq^2}{2\epsilon_0 m} \right)^{1/2} \left(\frac{\pi}{2r^*} \right)^{1/2}. \quad (12.29b)$$

Выражая через безразмерное время, имеем

$$\omega_3 t^* = \pi^{1/2} \Gamma_0^{1/2}, \quad (12.30a)$$

$$\omega_2 t^* = \omega_3 t^* \left(\frac{4r^*}{L} \right)^{1/2}. \quad (12.30b)$$

Из (12.30a) и (12.23a) видно, что более жесткие ограничения на временной шаг будет накладывать самая высокочастотная плазменная мода, а не ω_0 , т. е. частота колебаний, отвечающая среднему расстоянию.

12.2.3. Характерные масштабы задачи

Эксперимент на двумерной электронной пленке проводился при постоянном объеме и из соображений удобства был пересчитан на плотность $n = 10^{14}$ элек. рон/м². Фазовый переход изучается путем изменения температуры, а получаемый окончательный результат выражается через безразмерный коэффициент Γ_0 . Значения остальных параметров были следующими:

1. Число частиц $N = 10^4$.
2. Расчетная область ($L \times L$); $L = 10^{-5}$ м.
3. Пространственная сетка (256×256) с $H = 3.9 \times 10^{-8}$ м.
4. Шаг по времени $DT = 0.5 \times 10^{-12}$ с.
5. Короткодействующая поправка к силе:
внутренняя граница $b = 1.5H = 5.86 \times 10^{-8}$ м,
внешняя граница $r_c = 4H = 15.62 \times 10^{-8}$ м.
6. Размер цепочечной ячейки $= 4H$.
7. Количество соседей в круге радиуса $r_c \approx 6$.
8. Период треугольной решетки $a = 2.74H = 10.75 \times 10^{-8}$ м.
9. Единица расстояния $r^* = 1.44H = 5.64 \times 10^{-8}$ м.

Как в жидкой, так и в твердой фазах электроны расположены в приблизительно треугольной плотно упакованной решетке, при этом характерное расстояние между ними составляет около 3 ячеек сетки. Именно поэтому нужно вычислять силу с пространственным разрешением гораздо более высоким, чем шаг сетки, и простой расчет методом частица — сетка с разрешением примерно в одну ячейку сетки был бы безнадежно неточен. Поэтому используется метод Р³М, в котором к сеточной силе добавляется короткодействующая поправка. Эта сила аккуратно подправляется от внутренней границы, т. е. $b = 1.5H$, до внешней границы $r_c = 4H$, за пределами которой уже достаточную точность обеспечивает одна сеточная сила. Так что парное взаимодействие передается правильно для всех расстояний, больших b , к тому же для $r < r_c$ разрешение не зависит от сеточного шага.

Шаг по времени должен выбираться с учетом частот, введенных в предыдущем разделе. Исходя из приведенных выше данных, получаем:

Парное колебание около среднего расстояния при $r = 2r^*$

$$\omega_0 DT = \frac{\pi^{1/4} q n^{3/4}}{2^{3/2} m^{1/2} \epsilon_0^{1/2}} DT = 0.42. \quad (12.31a)$$

Максимальное парное колебание около $r = b$

$$\omega_1 DT = \left(\frac{2r^*}{b}\right)^{3/2} \omega_0 DT = 1.12. \quad (12.31b)$$

Минимальная плазменная частота при $k = 2\pi/L$

$$\omega_2 DT = \left(\frac{nq^2}{2\epsilon_0 m} \frac{2\pi}{L}\right)^{1/2} DT = 0.157. \quad (12.31c)$$

Максимальная плазменная частота при $k = \pi/(2r^*)$

$$\omega_3 DT = \omega_2 DT \left(\frac{L}{4r^*}\right)^{1/2} DT = 1.05. \quad (12.31d)$$

Видно, что шаг по времени удовлетворяет условию устойчивости ($\omega DT < 2$) для обеих максимальных частот — и парного колебания, и плазменной. При $\omega DT \approx 1$ орбиты имеют осцилляторный характер и передаются качественно правильно при почти шести шагах интегрирования на период, но погрешность аппроксимации в 8 % является большой. Тем не менее большая часть частиц колеблется с частотой порядка ω_0 и представляется 15 шагами интегрирования на периоде с погрешностью аппроксимации около 1.5 %. Максимальная длина плазменной волны равна размеру расчетной области, и, следовательно, минимальное

волновое число $k_{\min} = 2\pi/L$. Период такой волны интегрируется примерно за 40 шагов с погрешностью аппроксимации около 0.2 %. Если мы считаем такое интегрирование хорошим вплоть до частоты $\omega DT \approx 0.5$, то в этот интервал попадают первые десять волновых чисел до $k = 10k_{\min}$. Большие волновые числа, несмотря на устойчивость и качественно правильное поведение, интегрируются неточно. Вышеуказанный шаг по времени является, пожалуй, наибольшим допустимым шагом, который можно еще использовать при данных условиях. Если бы на первый план не выступали соображения экономии машинного времени, то было бы лучше взять шаг по времени в два раза меньше. Тогда все частоты интегрировались бы с погрешностью аппроксимации не хуже примерно 2 %.

12.2.4. Машинное время и память

При описанном выше соотношении масштабов задачи время счета на IBM 360/195 одного шага по времени распределяется между различными этапами расчета следующим образом:

Раздача заряда на сетку	0.86 с
Нахождение потенциала	2.00 с
Добавление короткодействующей силы и ускорение	5.94 с
	<hr/>
Суммарное время счета одного шага	8.80 с

Видно, что короткодействующая коррекция является самой дорогой частью расчета, хотя количество соседей, для которых $r < r_c$, и составляет всего около шести. Теоретически в Р³М-методе минимум в числе операций достигается при одинаковом времени на расчет короткодействующей силы и на нахождение потенциала. Так как в данном случае это условие не выполнено, необходимо посмотреть, не приведет ли использование более мелкой сетки, скажем (512×512) , к ускорению счета. Это увеличит в 4 раза время нахождения потенциала (до 8 с), но уменьшит в 4 раза время расчета короткодействующих сил (до 1.5 с). Полное время составит 10 с и основную часть займет расчет потенциала. Видно, что другой выбор масштабов задачи не только хуже по времени, но и нежелателен потому, что для потенциала используется в 4 раза больше памяти. Из-за того что полное число узлов сетки можно изменять только в число раз, кратное четырем, достигнуть теоретического минимума невозможно. Объем памяти, используемый при первом варианте параметров, включая место для программы и всех данных, составляет 802 кбайт. С учетом программ диагностики для типичного варианта в 300 шагов по времени требуется 48 мин машинного времени на IBM 360/195.

12.2.5. Плавление электронной пленки

Вычислительный эксперимент выполняется путем задания начальных положений и скоростей всех частиц, после чего система предоставляется сама себе. Чтобы гарантированно не привносить в рассматриваемую систему никаких искусственных структур, вычислительный эксперимент начинался со случайного распределения электронов. Эта система могла релаксировать, в то же время ее температура поддерживалась постоянной (~ 1 К) путем перенормировки скоростей на каждом шаге по времени. Затем температурное ограничение отключалось и посредством вычисления на протяжении 200 шагов по времени без подвода энергии устанавливалось тепловое равновесие. Тепло вводилось в систему постепенно, для чего на протяжении 25 шагов скорости умножались на каждом шаге по времени на коэффициент чуть больше единицы (1.004), затем система успокаивалась до достижения нового равновесного состояния за время еще 50 шагов, просчитываемых без подвода энергии. На каждом временном шаге вычислялись кинетическая, потенциальная и полная энергии, что давало соответственно мгновенные значения температуры, потенциальной энергии и внутренней энергии системы. Было обнаружено, что мгновенная температура колеблется с периодом примерно 14 шагов по времени и амплитудой колебаний около 2.5%. После вычисления скользящего среднего от энергии по 28 шагам колебания в измерениях температуры уменьшились до $\sim 0.5\%$. Было установлено, что время достижения системой нового равновесия после такого подвода тепла составляет около 25 шагов по времени, так что в качестве термодинамических переменных нового состояния можно было принять скользящие средние от температуры T и внутренней энергии U , вычисленные через 75 шагов. Этот процесс повторялся, давая значения T и U в последовательности равновесных состояний. Если параметры состояния (T_1, U_1) и (T_2, U_2) относятся к двум последовательным состояниям, то удельная теплоемкость в точке $\bar{T} = (T_1 + T_2)/2$ подсчитывалась по формуле $C_v = (U_2 - U_1)/(T_2 - T_1)$. Для измерений в окрестности фазового перехода коэффициент нагрева был уменьшен до 1.001 и для приведения системы в равновесие расчет без нагрева проводился на протяжении 75 шагов по времени.

На рис. 12.1 показана зависимость внутренней энергии от температуры. Асимптотическое поведение при высоких и низких температурах с очевидностью показывает, что фазовый переход происходит в области температур 3 К для $n = 10^{14}$ м⁻². Наблюдаемая с шагом по температуре ~ 0.05 К зависимость внутренней энергии от температуры ведет себя достаточно гладко при прохождении через точку фазового перехода. Если при 3 К внутренняя энергия и испытывает скачок, то он меньше $0.1 k_B$,

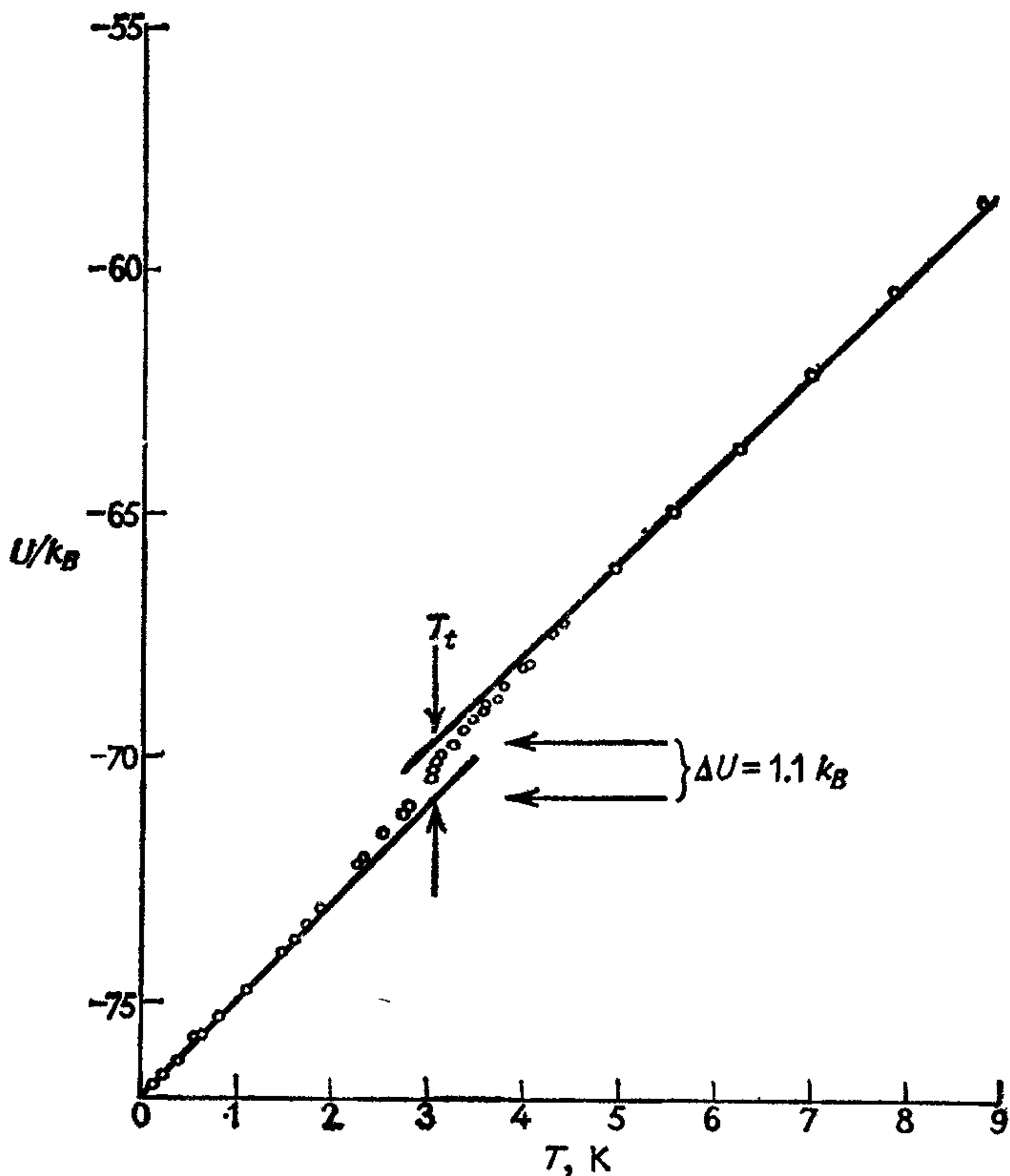


Рис. 12.1 Изменение полной внутренней энергии U в зависимости от температуры T в двумерной электронной пленке. Плотность электронов 10^{14} м^{-2} . ([187], с любезного разрешения Journal of Physics C: Solid-State Physics, © Institute of Physics.)

что надо сравнить с полным изменением этой величины $\Delta U = 1.1 k_B/\text{электрон}$, связанным с фазовым переходом. Характер перехода более четко проявляется на зависимости удельной теплоемкости C_v от температуры T (рис. 12.2). Мы наблюдаем постепенное увеличение C_v с температурой, начиная от значения вблизи $2 k_B$ при 0 К до максимума примерно $3.8 k_B$ при $T_t = 3.1$ К. После чего наблюдается почти скачкообразный спад до значения в окрестности $2 k_B$. В течение остального времени эксперимента, за которое температура возросла до 8 К, удельная теплоемкость больше почти не меняется.

Интегрируя остаточную удельную теплоемкость, находим, что полное изменение внутренней энергии равно $1.1 \pm 0.1 k_B/\text{электрон}$. Считая, что температуре фазового перехода отвечает максимум на кривой C_v от T , получим $T_t = 3.1 \pm 0.05$ К при $n =$

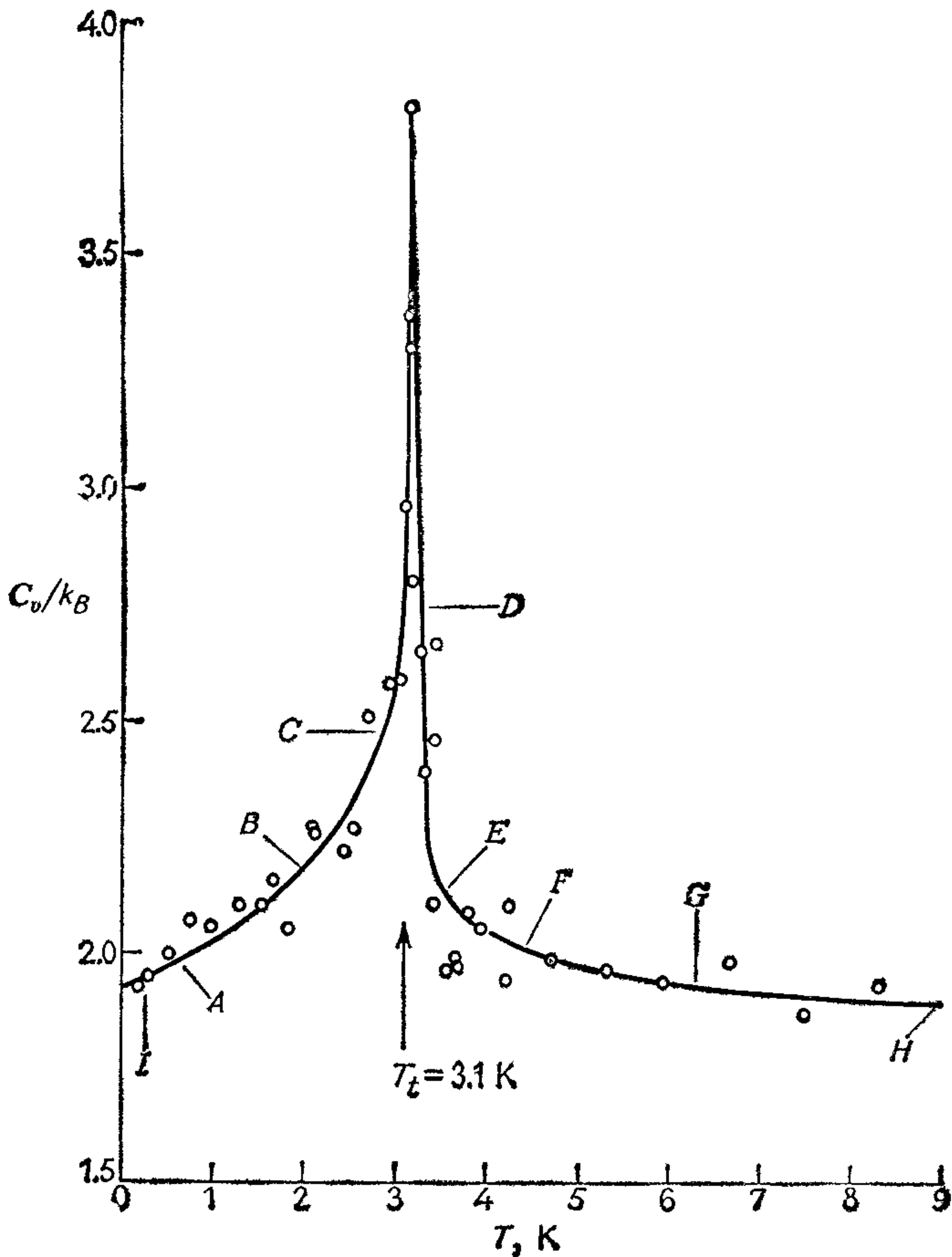
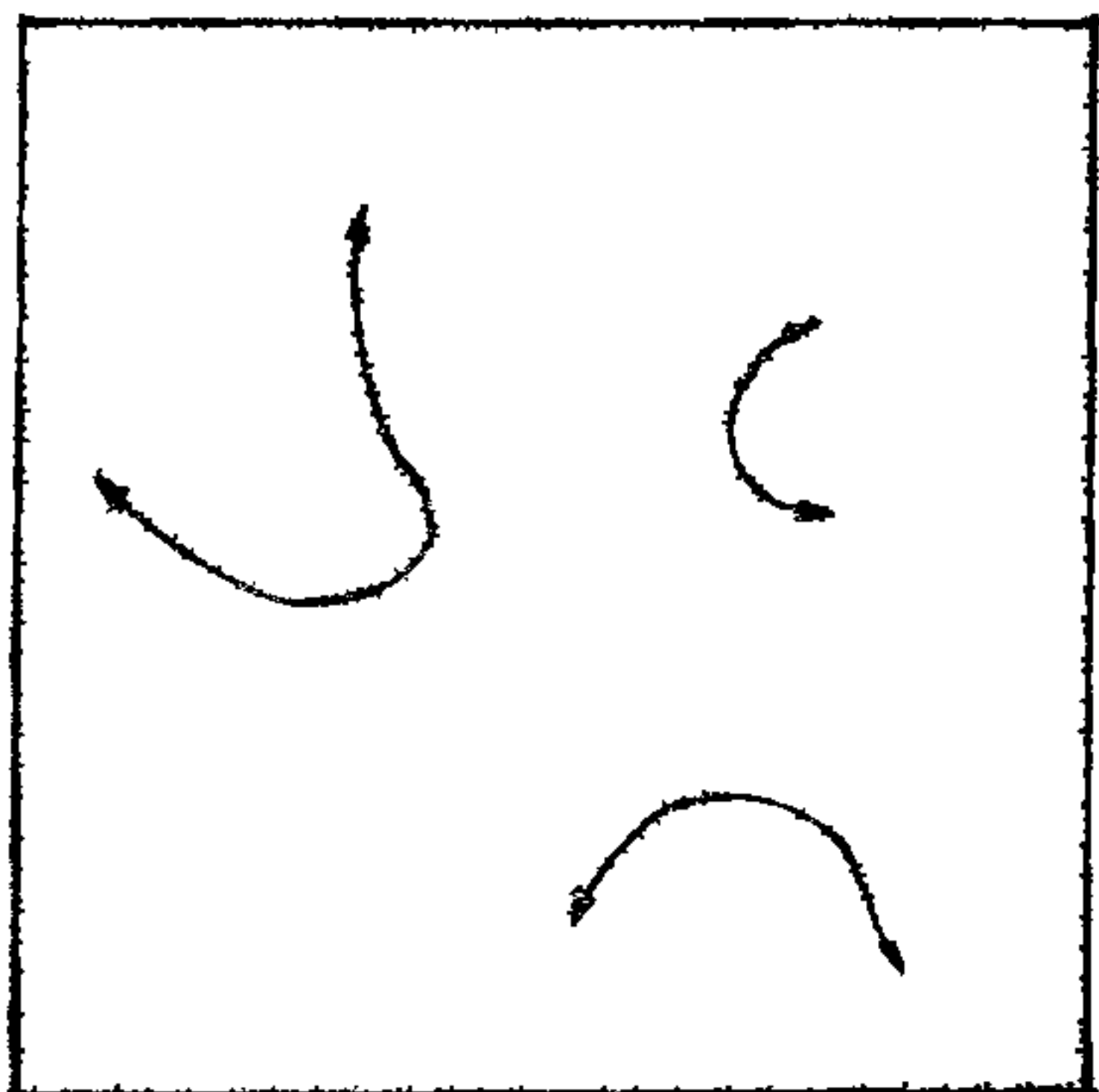
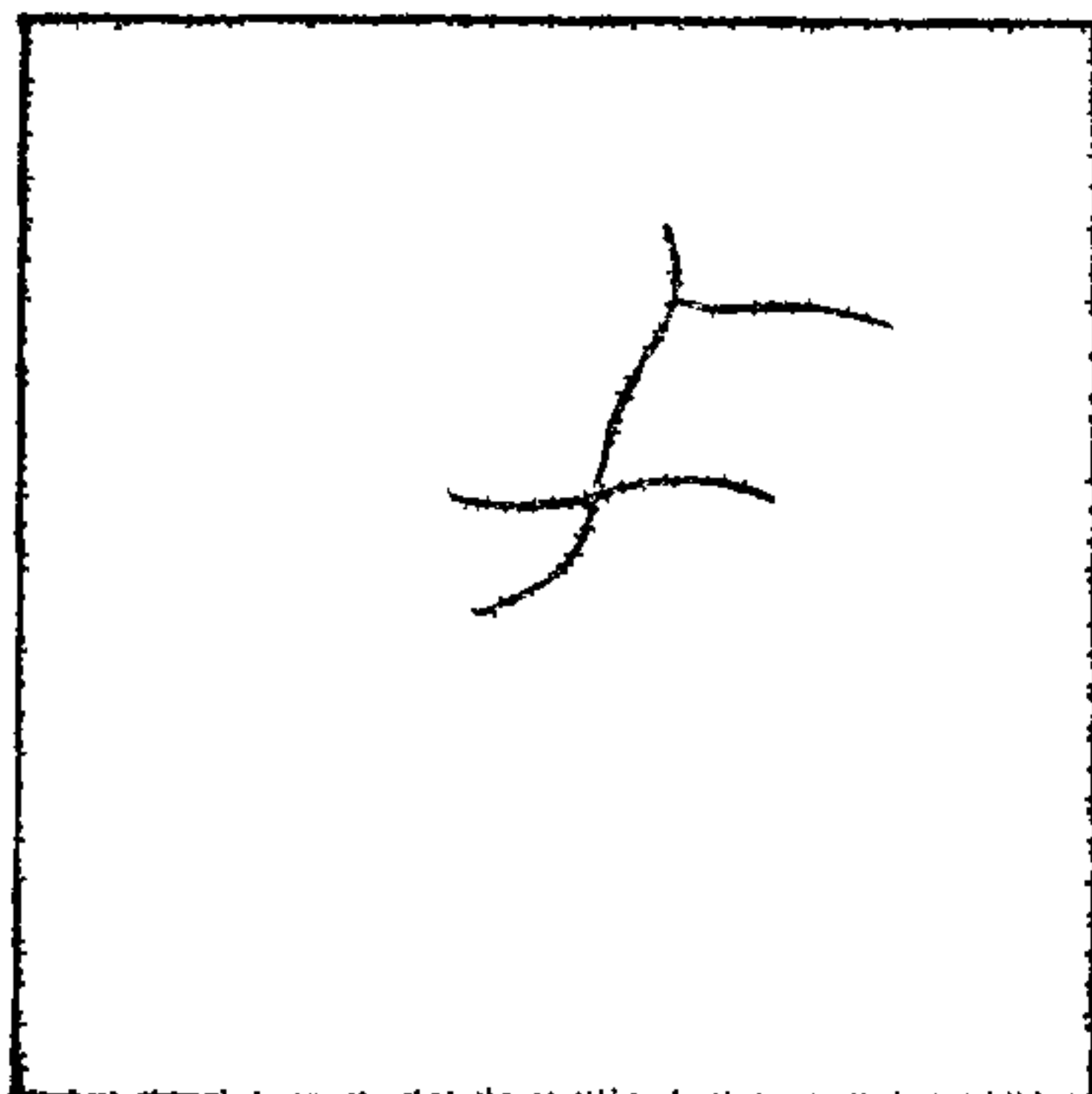


Рис 122. Аномалия удельной теплоемкости λ типа, наблюдаемая в двумерной электронной пленке. Плотность электронов 10^{14} м^{-2} ([187], с любезного разрешения Journal of Physics C: Solid State Physics, © Institute of Physics.)

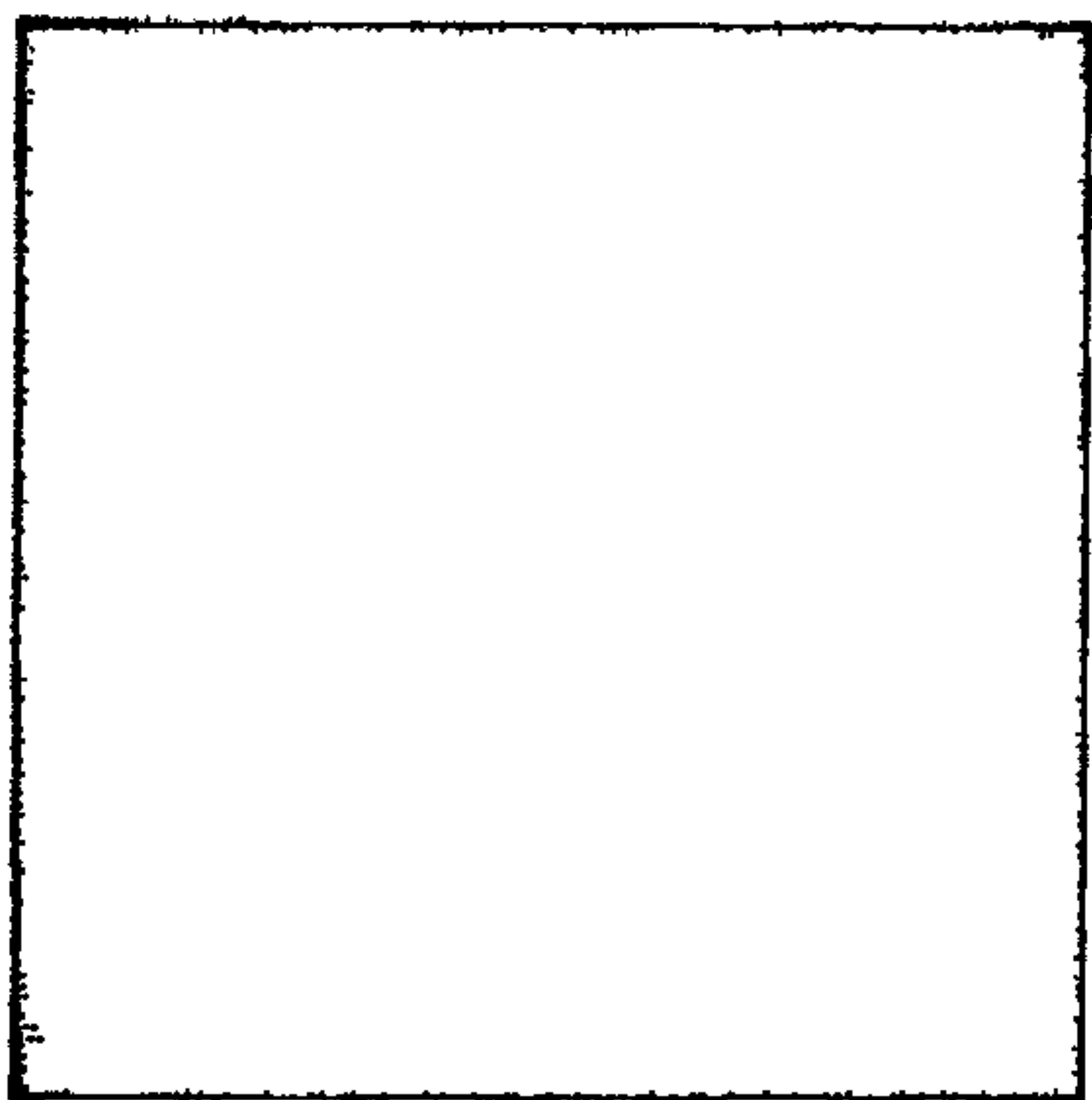
$= 10^{14} \text{ м}^{-2}$. В результате находим, что при фазовом переходе в классической двумерной электронной жидкости $\Gamma_0 = 95 \pm 2$. Для сравнения значения, найденные для соответствующей трехмерной переменной $\Gamma_c = (4\pi n/3)^{1/3} q^2 / (4\pi \epsilon_0 k_B T)$ в классической трехмерной электронной жидкости, составляют 125 [55] и 155 ± 10 [313]. Следовательно, двумерная система в каком-то смысле более устойчива, чем трехмерная, что совпадает с оценкой, данной в работе [312]. Однако абсолютное значение Γ_0 , полученное в данном вычислительном эксперименте, не согласуется со значением 28, найденным теоретически в работе [312] с использованием самосогласованного гармонического приближения.



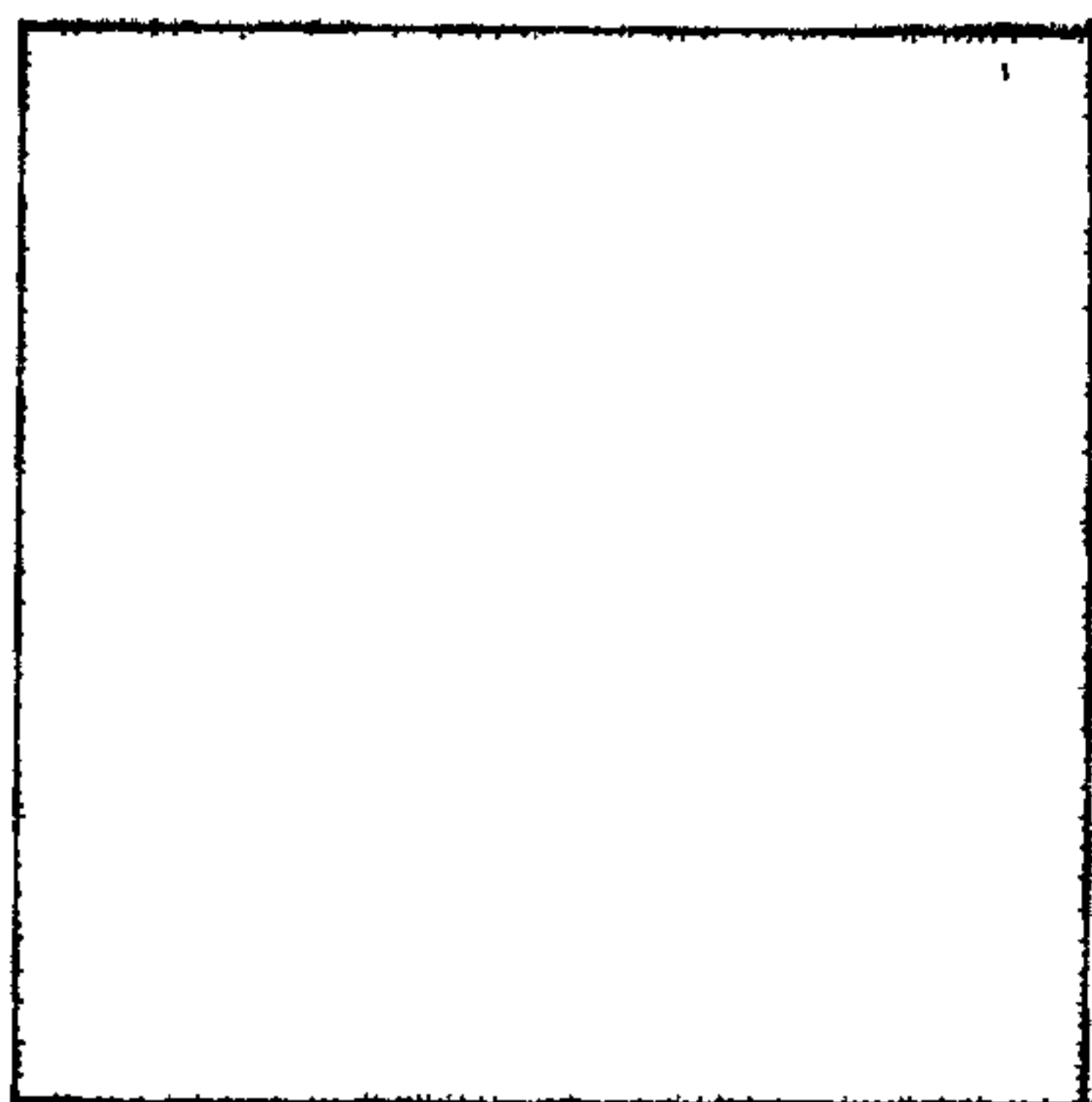
(A)



(B)



(C)



(D)

Рис. 123. Исчезновение дальнего порядка при прохождении через λ -переход. Плотность электронов 10^{14} м^{-2} . Жирными кривыми намечены некоторые из границ доменов (A) 0.65 К; (B) 1.90 К; (C) 281 К, (D) 330 К ([187], с любезного разрешения Journal of Physics C. Solid-State Physics, © Institute of Physics)

Эксперимент проведен с системой из 10 000 частиц и с системой из 2500 частиц. В области температур 2—8 К, для которой размер домена (см. ниже) меньше размера системы из 2500 частиц, результаты для большой и малой систем находятся в приемлемом согласии. В частности, с увеличением размера образца не наблюдается обострения пика, что свидетельствует о том, что конечная ширина аномалии в удельной теплоемкости обусловлена физическим явлением, а не является неким искусственным эффектом размера образца. При температурах ниже

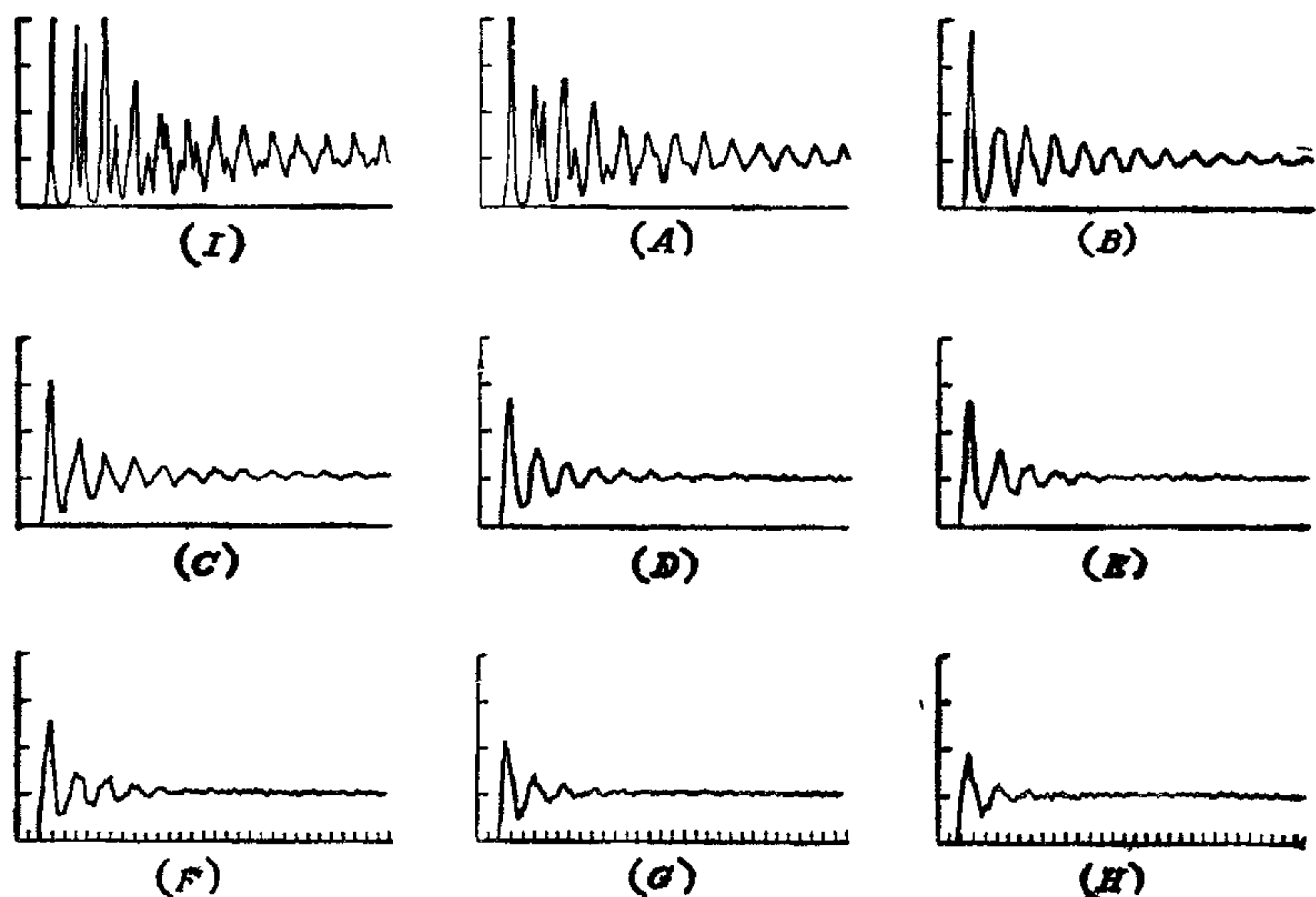


Рис. 12.4. Изменение функции радиального распределения при прохождении через λ -переход. Плотность электронов 10^{14} м^{-2} . Деления на оси x соответствуют шагу сетки H . Совершенная треугольная решетка имела бы пики в точках 2.74, 4.75, 5.48 и 7.3 делений. Деление на оси y соответствует единице. Буквы в скобках отвечают соответствующим точкам на рис. 12.2. ([187], с любезного разрешения Journal of Physics C: Solid-State Physics, © Institute of Physics.)

2 К размер домена превышает размер образца из 2500 частиц. В этой области маленький образец не может правильно передать рассматриваемое явление. Удельная теплоемкость, наблюдаемая при этих температурах для небольшого образца, почти постоянна и составляет $2.2\text{--}2.3 k_B$.

Все представленные до сих пор величины — это величины, обычно доступные и в физическом эксперименте. Вычислительный же эксперимент обладает тем достоинством, что, кроме того, он позволяет детально исследовать микроскопическую структуру вещества в процессе фазового превращения. Приведем теперь эти результаты.

На рис. 12.3 изображены все 10 000 частиц в четырех состояниях, отвечающих точкам A , B , C и D на кривой удельной теплоемкости. Видно, что по мере прохождения точки фазового перехода постепенно уменьшаются размеры микроструктурных доменов и утрачивается дальний порядок. В точке A (0.65 К) четко видны два крупных домена неправильной формы и один кольцевой, причем во всех из них просматривается регулярная треугольная плотноупакованная решетка. Ясно видны границы доменов, на которых меняется ориентация решетки. Такая структура отвечает структуре полимикроструктур-

лического твердого тела, причем размер кристаллита больше размера образца, равного 10^{-5} м. В точке *B* (1.9 К) структура в основном не изменилась, за тем исключением, что теперь размер кристаллита меньше образца на $\approx 0.25 \times 10^{-5}$ м. В точке *C* (2.8 К), как раз перед фазовым переходом, микрокристаллическая структура почти полностью исчезла, тогда как в точке *D* (3.3 К), сразу после перехода, дальний порядок кристаллической структуры уже полностью отсутствует, и наблюдается структура, которая ассоциируется с жидкостью.

Более отчетливо ближний порядок выявляют путем вычисления $g(r)$ — парной функции распределения для данного вещества. Это показано на рис. 12.4 для девяти точек на кривой удельной теплоемкости. В точках *I* и *A* (< 0.65 К) наблюдаются весьма мощные и устойчивые пики, соответствующие кристаллической структуре с треугольной решеткой. Между точками *B* и *C* количество пиков в $g(r)$ уменьшается примерно с 30 до 12, а затем между точками *D* и *E* резко падает до 6 и 4, соответственно. От точки *F* до *H* функция $g(r)$ имеет вид, характерный для жидкости и количество пиков по мере роста температуры постепенно уменьшается от 4 до 2.

12.3. ИОННЫЕ МИКРОКРИСТАЛЛЫ

12.3.1. Безразмерные уравнения

Изучение фазовых превращений в ионных микрокристаллах проводилось при помощи трехмерной программы P3M3DP, созданной Иствудом, Хокни и Лоуренсом [120], реализующей метод P³M.

Для описания отталкивания электронных оболочек ионов на малых расстояниях мы, следуя Полингу [306], использовали взаимодействие кулоновского типа с некоторой обратной степенью расстояния. Сила взаимодействия пары ионов, обозначаемых индексами *i* и *j*, находящихся на расстоянии r_{ij} , имеет вид

$$F(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}^2} \left[1 + \operatorname{sgn}(q_i q_j) \left(\frac{s_i + s_j}{r_{ij}} \right)^p \right]. \quad (12.32)$$

где q — ионный заряд, а s , которое характеризует размер иона, пропорционально ионному радиусу. Параметр p определяет «жесткость» отталкивания между ионами.

Если указанное взаимодействие заложить в уравнения движения Ньютона и использовать соответствующие единицы, то можно показать, что эти уравнения зависят от четырех параметров. Для системы, состоящей из одного сорта положительных ионов и одного сорта отрицательных, что обозначается ин-

дексами $+$ и $-$, такими параметрами являются отношение зарядов $\sigma (=q_+/q_-)$, отношение масс $\mu (=m_+/m_-)$, отношение радиусов $\rho (=s_+/s_-)$ и параметр жесткости p . Если мы ограничимся рассмотрением эквивалентных веществ, например КСl, СаО, для которых $\sigma = -1$, то параметров будет только три: μ , ρ и p . Можно показать, что зависимость от отношения масс очень слабая, поэтому основное внимание направлено на изучение влияния отношения радиусов и параметра жесткости. Вышеуказанная сила взаимодействия является, как говорят, масштабируемой, поскольку вычислительный эксперимент, выполненный для одного набора из восьми основных данных $(\epsilon_0, m_{\pm}, q_{\pm}, s_{\pm}, p)$, можно пересчитать и рассматривать как новый результат, полученный для какой-то другой системы с такими же σ , μ , ρ и p .

Зависимость от основных данных фигурирует в явном виде при выборе подходящей системы единиц. Следуя общей процедуре разд. 12.2.1, будем отмечать характерные для данной задачи величины звездочками. Такими величинами являются:

- $r^* = s_+ + s_-$, равновесное расстояние для изолированной молекулы газа;
- $m^* = m_+m_-/(m_+ + m_-)$, приведенная масса изолированной молекулы газа;
- $t^* = (4\pi\epsilon_0 r^{*3} m^*/|q_+q_-|)^{1/2}$, приближенно период самой высокочастотной моды оптических колебаний кристалла;
- $F^* = |q_+q_-|/(4\pi\epsilon_0 r^{*2})$, сила кулоновского взаимодействия ионов изолированной молекулы, находящихся на равновесном расстоянии;
- $E^* = [q_+q_-|/(4\pi\epsilon_0 r^*)$, кулоновская потенциальная энергия на равновесном расстоянии изолированной молекулы.

Если выразить все переменные в указанных единицах, то получим следующие безразмерные уравнения, в которых, как и в оставшейся части этого раздела и в разд. 12.3.2, переменные r , t , F и т. д. обозначают теперь безразмерные переменные r/r^* , t/t^* , F/F^* и т. д. Мы не вводим для безразмерных переменных новых значков, чтобы не загромождать индексные выражения. Итак, парные взаимодействия принимают вид

$$F_{+-}(r_{ij}) = -\frac{1}{r_{ij}^2} \left[1 - \left(\frac{1}{r_{ij}} \right)^p \right], \quad (12.33a)$$

$$F_{++}(r_{ij}) = +\frac{1}{r_{ij}^2} \left[1 + \left(\frac{2/(1+\rho^{-1})}{r_{ij}} \right)^p \right], \quad (12.33b)$$

$$F_{--}(r_{ij}) = +\frac{1}{r_{ij}^2} \left[1 + \left(\frac{2/(1+\rho)}{r_{ij}} \right)^p \right], \quad (12.33c)$$

а уравнения движения превращаются в уравнения

$$(1 + \mu) \frac{d^2 \mathbf{r}}{dt^2} = \sum_{\text{пары}} \mathbf{F} \quad (12.34a)$$

для положительных ионов и

$$(1 + \mu^{-1}) \frac{d^2 \mathbf{r}}{dt^2} = \sum_{\text{пары}} \mathbf{F} \quad (12.34b)$$

для отрицательных ионов.

Для исследования процессов плавления мы выбрали для начала три реальных вещества и одно гипотетическое, что отмечено в табл. 12.1.

Таблица 12.1. Отношение масс и радиусов для четырех солей, выбранных в качестве теста

Соль	μ	ρ	$\frac{DT}{10^{-15} \text{ с}}$ в единицах	$\omega_0 DT$
KCl	1.10	0.73	7.5	0.42
RbCl	2.40	0.82	7.5	0.33
(LiCl)*	0.20	0.82	3.5	0.51
NaCl	0.65	0.55	6.5	0.52

Данные о радиусах из работы [306]; $\rho = 8$; DT — шаг по времени.

Таблица 12.2. Данные по отдельным ионам, использованные в (12.32)

Ион	Масса m в единицах 10^{-27} кг	Заряд q в единицах e	Размер s в единицах 10^{-10} м
Cl	58.8	-1	1.55
Li	11.8	+1	0.51
Na	38.0	+1	0.77
K	64.7	+1	1.13
Rb	141.2	+1	1.27

$e = +1.6 \times 10^{-19} \text{ Кл}$ — абсолютная величина заряда электрона.

Большинство галоидов щелочных металлов и многие другие представляющие интерес соли лежат внутри треугольника в (μ, ρ) -плоскости, в вершинах которого расположены NaCl, (LiCl)* и RbCl. (LiCl)* представляет собой гипотетическое ве-

щество с отношением масс, как у LiCl, и отношением радиусов, как у RbCl, и было введено для того, чтобы определить влияние отношения масс. Включение в рассмотрение KCl обусловлено тем, что это вещество наиболее широко используется в качестве стандартного образца и для него имеются как расчетные, так и физические данные. Чтобы легче сравнивать получаемые результаты с экспериментом, моделирование проводится в единицах СИ, что отражено в табл. 12.2.

12.3.2. Выбор шага по времени

Как и в случае электронной пленки, выбор шага по времени должен производиться с учетом частот парных колебаний и коллективной плазменной частоты.

Частоту колебаний изолированной молекулы, образованной одним положительным и одним отрицательным ионами, можно определить, исходя из (12.33а) и (12.34а, б). Следуя схеме исследования уравнений (12.11) и (12.12), имеем в безразмерном виде

$$\frac{d^2r}{dt^2} = -\frac{1}{r^2} \left[1 - \left(\frac{1}{r} \right)^p \right]. \quad (12.35)$$

Проводя обычное разложение около равновесного положения $r = 1$, для чего положим $r = 1 + r'$ (см. (12.13) и (12.14)), имеем

$$\frac{d^2r'}{dt^2} = -\omega_0^2 r', \quad (12.36a)$$

где

$$\omega_0 = p^{1/2}. \quad (12.36b)$$

Возвращаясь к единицам СИ, получаем

$$\omega_0 DT = p^{1/2} DT/t^* = \left[\frac{p |q_+ q_-|}{4\pi\epsilon_0 (s_+ + s_-)^3} \left(\frac{1}{m_+} + \frac{1}{m_-} \right) \right]^{1/2} DT. \quad (12.37)$$

Этот результат отражает зависимость частоты колебаний от всех основных параметров. В частности, частота пропорциональна корню квадратному из параметра жесткости p и обратно пропорциональна равновесному расстоянию $(s_+ + s_-)$ в степени $3/2$. Зависимость от массы показывает, что частота определяется массой более легкого иона.

Частота колебаний в кристаллической решетке будет отличаться от частоты молекулы ω_0 по двум причинам. Во-первых, потому, что равновесное расстояние, иначе говоря, параметр решетки, в кристалле больше, чем $(s_+ + s_-)$, и, во-вторых, из-за влияния других ионов решетки. Эти эффекты, однако, стремятся компенсировать друг друга и сравнительно малы. Как правило,

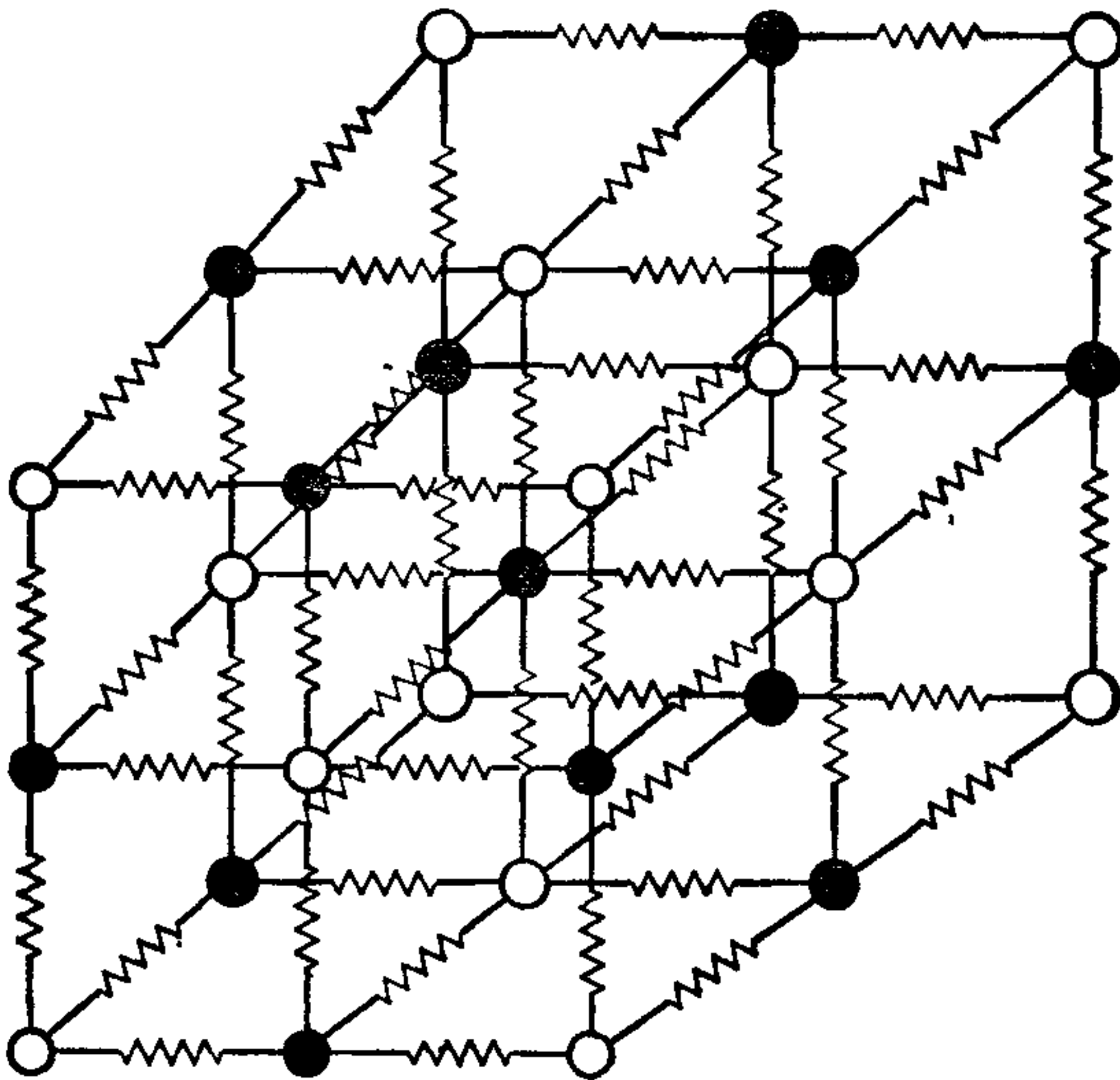


Рис. 12.5. Расположение ионов в кристалле галоида щелочного металла. Темные кружки изображают положительные ионы, например K^+ ; светлые кружки — отрицательные ионы, например Cl^- . Зигзагообразными линиями показаны воображаемые пружинки, соединяющие соседние ионы, которые используются при оценке частот колебаний.

бывает достаточно в качестве первоначального варианта шага по времени опираться на величину ω_0 . Поскольку большая часть ионов колеблется примерно с этой частотой и нужно обеспечить достаточную точность интегрирования, шаг следует выбирать приблизительно из условия

$$DT \leq 0.5/\omega_0. \quad (12.38)$$

Значения $\omega_0 DT$ для $p = 8$ приведены в табл. 12.1 и, как легко видеть, этому условию удовлетворяют.

Равновесное расстояние a между соседними положительными и отрицательными ионами в кристалле галоида щелочного металла можно определить из минимума потенциальной энергии решетки. Все галоиды щелочных металлов, за исключением $CsCl$, образуют кристаллы с кубической решеткой, в которой положительные и отрицательные ионы занимают по очереди положения в вершинах каждого куба, как показано на рис. 12.5. Ближайшее окружение каждого иона составляют шесть соседей. Это называется «первой координационной оболочкой», а количество ионов в ней (в данном случае шесть) называют «координационным числом». Следующие ближайшие соседи образуют оболочку из 12 ионов того же знака на расстоянии $\sqrt{2}a$. Следуя Полингу [306], получим, что энергия, приходящаяся на

молекулу такой бесконечной решетки, выраженная в безразмерных единицах, равняется

$$U(r) = -\frac{A}{r} + \frac{C}{p+1} \frac{1}{r^{p+1}} G(\rho), \quad (12.39)$$

где r — расстояние между соседними положительным и отрицательным ионами в решетке, а

$$G(\rho) = 1 + \frac{1}{2^{(p+1)/2}} \left\{ \left[\frac{2}{1+\rho^{-1}} \right]^p + \left[\frac{2}{1+\rho} \right]^p \right\}. \quad (12.40)$$

В этом выражении $A = 1.747$ представляет собой постоянную Маделунга, которая учитывает сумму электростатической энергии всех пар ионов в решетке, а $C = 6$ — координационное число. Первый член в (12.39) отвечает электростатической энергии, обусловленной бесконечной решеткой ионов, а второй член — энергии, соответствующей отталкиванию между ионами первой и второй координационных оболочек. Влиянием отталкивания более удаленных ионов пренебрегается. Первый член в выражении для $G(\rho)$ обусловлен отталкиванием разноименных ионов на расстоянии r . Вторым членом соответствует отталкиванию пар положительных ионов, разделенных расстоянием $\sqrt{2}r$, и третий член связан с отталкиванием отрицательных ионов, отстоящих также на $\sqrt{2}r$.

Равновесное расстояние решетки a при нулевой температуре получается путем нахождения такого расстояния r , при котором $U(r)$ минимальна. Это имеет место, когда

$$\frac{dU}{dr} \Big|_{r=a} = 0, \quad (12.41a)$$

откуда
$$\frac{A}{a^2} - \frac{C}{a^{p+2}} G(\rho) = 0 \quad (12.41b)$$

и
$$a = \left[\frac{C}{A} G(\rho) \right]^{1/p}. \quad (12.41c)$$

Поскольку в качестве единицы расстояния r^* выступает равновесное расстояние одной молекулы, величина a , определяемая выражением (12.41c), показывает, во сколько раз равновесное расстояние решетки больше, чем у изолированной молекулы. В случае $\rho = 1$ и $p = 8$ функция $G(\rho) = 1.088$ и $a = 1.167$.

Подставляя (12.41c) в (12.39), получаем более простое выражение для энергии решетки

$$U(r) = -\frac{A}{r} \left[1 - \frac{1}{p+1} \left(\frac{a}{r} \right)^p \right]. \quad (12.42a)$$

Следовательно, при равновесном расстоянии решетки $r = a$ на молекулу приходится энергия

$$U(a) = -\frac{p}{p+1} \frac{A}{a}. \quad (12.42b)$$

Эта энергия связи представляет собой энергию, которая требуется для разделения кристалла на составляющие его отдельные ионы.

Для того чтобы получить выражения для частот колебаний кристаллической решетки, необходимо перевести выражение для зависимости энергии решетки от периода решетки на язык пружинной модели¹⁾. В этой модели взаимодействие осуществляется только посредством пружин, соединяющих соседние ионы (см. рис. 12.5), причем их длина в ненапряженном состоянии равна a . Если постоянная пружинки²⁾ равна λ и решетка растянута так, что ее период равен $a + \delta r$, то энергия решетки, приходящаяся на пару ионов (или молекулу), составляет

$$U(a + \delta r) = U(a) + \frac{C}{2} \lambda \delta r^2, \quad (12.43a)$$

где $C = 6$ — связанное с каждой молекулой количество пружин. С другой стороны, фактическое изменение энергии решетки получается из разложения (12.42a) в ряд Тейлора:

$$U(a + \delta r) = U(a) + \left. \frac{dU}{dr} \right|_a \delta r + \left. \frac{d^2U}{dr^2} \right|_a \frac{\delta r^2}{2}, \quad (12.43b)$$

где $\left. \frac{dU}{dr} \right|_a = 0$ в силу (12.41a)

$$\text{и} \quad \frac{d^2U}{dr^2} = \frac{A}{r^3} \left[(\rho + 2) \left(\frac{a}{r} \right)^\rho - 2 \right], \quad (12.43c)$$

$$\left. \frac{d^2U}{dr^2} \right|_a = \frac{\rho A}{a^3}. \quad (12.43d)$$

Сравнивая (12.43b) с (12.43a), имеем

$$\lambda = \frac{1}{C} \left. \frac{d^2U}{dr^2} \right|_a = \frac{\rho A}{Ca^3}. \quad (12.44)$$

Знание эффективной силовой постоянной λ позволяет вычислить частоту ряда мод колебаний. Во всех случаях частота определяется по формуле

$$\omega = \left(\frac{N\lambda}{m} \right)^{1/2} = \left(\frac{N}{C} \frac{\rho A}{ma^3} \right)^{1/2} = \left(\frac{N}{C} \frac{A}{ma^3} \right)^{1/2} \omega_0, \quad (12.45)$$

где N — приходящееся на одну молекулу число пружин, которые изменяют свою длину, иначе говоря, «задействованы» в данной моде, а m — соответствующая масса в единицах приведенной массы.

Если кристалл «дышит» и все ионы движутся в унисон, задействованы все пружинки ($N = C$) и $m = 1$, отсюда частота

¹⁾ Так называемое гармоническое приближение. — Прим. перев.

²⁾ Чаще используется термин «силовая постоянная». — Прим. перев.

бризерной моды равна

$$\omega_4 = \left(\frac{d^2U}{dr^2} \Big|_a \right)^{1/2} = \left(\frac{pA}{a^3} \right)^{1/2} = \left(\frac{A}{a^3} \right)^{1/2} \omega_0. \quad (12.46a)$$

Сравнивая с частотой колебаний изолированной молекулы, т. е. с ω_0 в (12.36b), имеем для случая $\rho = 1$, $p = 8$

$$\frac{\omega_4}{\omega_0} = \left(\frac{A}{a^3} \right)^{1/2} = 1.048. \quad (12.46b)$$

Тем самым продемонстрирована близость частот решетки и молекулы.

Если один ион с безразмерной массой m колеблется в своей потенциальной яме, тогда как все остальные неподвижны, то $N = 2$ и частота одночастичного колебания равняется

$$\omega_5 = \left(\frac{2}{C} \frac{pA}{ma^3} \right)^{1/2}. \quad (12.47a)$$

Если ионы имеют примерно одинаковую массу, то приведенная масса составляет половину массы одного иона, и, следовательно, $m = 2$. Отсюда

$$\omega_5 = \left(\frac{pA}{Ca^3} \right)^{1/2} = C^{-1/2} \omega_4. \quad (12.47b)$$

Если соседние плоскости ионов движутся в противоположных направлениях, то имеем «оптическую» моду колебаний, у которой $N = 2$ и $m = 1$, и, следовательно, оптическая частота равна

$$\omega_6 = \left(\frac{2}{C} \frac{pA}{a^3} \right)^{1/2} = \left(\frac{2}{C} \right)^{1/2} \omega_4. \quad (12.48)$$

Суммируя полученные результаты и используя данные для KCl, имеем

молекулярные колебания	$\omega_0 DT = 0.42$ (15 шагов на период),
бризерная мода	$\omega_4 DT = 0.44$ (14 шагов на период),
одночастичные колебания	$\omega_5 DT = 0.19$ (33 шага на период),
оптическая мода	$\omega_6 DT = 0.27$ (24 шага на период).

Таким образом, хотя выбор $\omega_0 DT \approx 0.4$, быть может, и кажется не совсем верным, мы считаем, что большая часть частиц колеблется в кристалле с более низкими частотами вблизи ω_5 и ω_6 , а они интегрируются аккуратно. Бризерную моду, хотя она и присутствует, не нужно интегрировать особенно точно, поскольку ее амплитуда мала. Однако важно выбрать такой шаг по времени, при котором бризерная мода устойчива (т. е. $\omega_4 DT < 2$), поскольку в противном случае ее амплитуда будет экспоненциально нарастать и очень быстро этот тип движения станет преобладающим. На рис. 12.6 показаны орбиты, полученные для

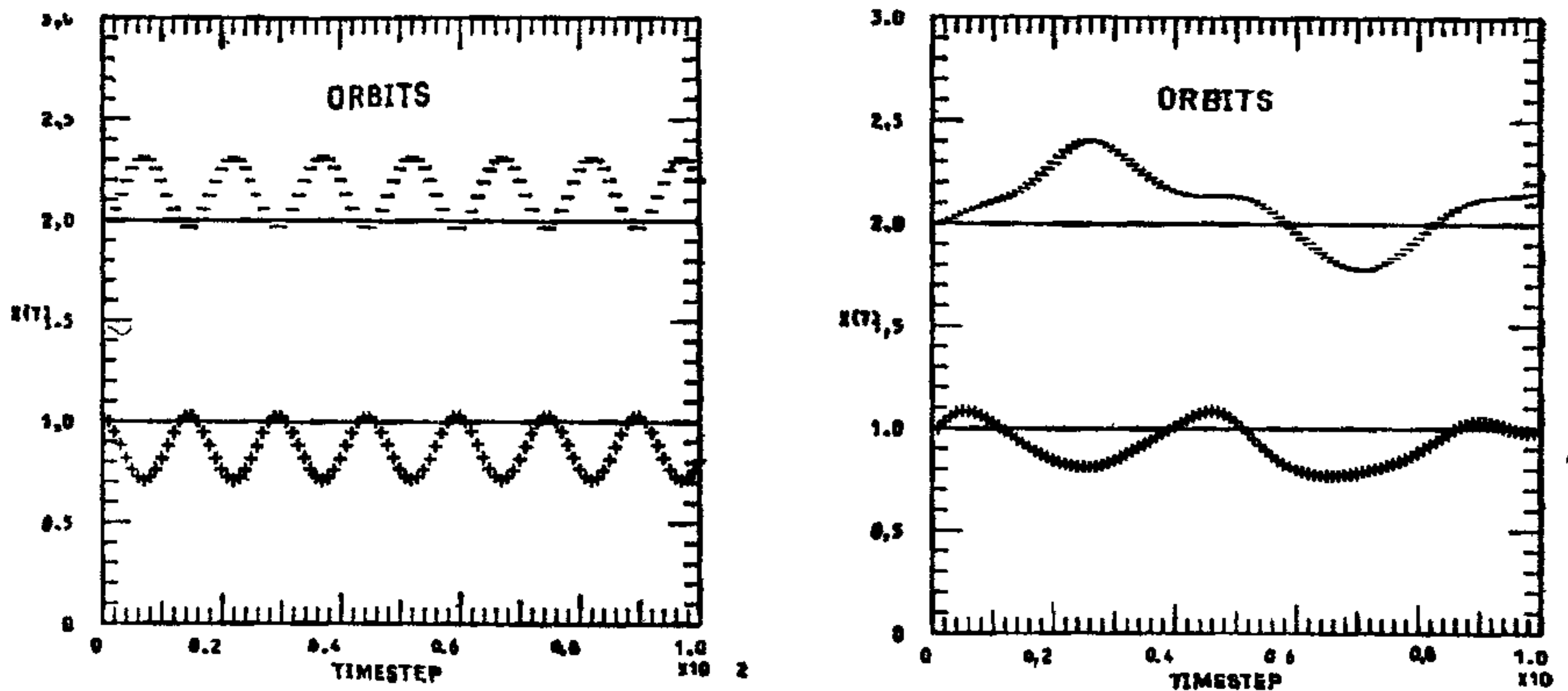


Рис. 12.6. Слева — колебание изолированной пары ионов, т. е. молекулы, при $\omega_0 DT = 0.42$. Справа — типичные колебания ионов в кристалле при том же шаге по времени. Характерна более низкая частота, отвечающая $\omega_0 DT = 0.3$. ([120], с любезного разрешения Computer Physics Communications, © North-Holland Publishing Co.)

молекулы KCl (слева) и кристалла KCl (справа) при температуре 600 К, что, как можно увидеть, находится в согласии с указанными оценками для мод колебаний.

В программе P3M3DP сделано так, что при всех $r < b$, где b — радиус, для которого

$$\left[2 \left(\frac{1}{m_+} + \frac{1}{m_-} \right) \left| \frac{dF}{dr} \right| \right]^{1/2} DT = 1.8, \quad (12.49)$$

сила остается постоянной, т. е. $F(r) = F(b)$. Это служит гарантией того, что связанная с градиентом силы частота никогда не превысит границу устойчивости $\omega DT = 2$. Если выясняется, что слишком много ионов обладают энергией, позволяющей им подходить друг к другу ближе, чем на расстояние b , то шаг по времени необходимо уменьшить. На рис. 12.7 показана зависимость силы от расстояния, которая использовалась при моделировании KCl. Сплошной линией показана полная сила; пунктирная линия, идущая в начало координат, представляет собой сеточную силу, а пунктирная линия, стремящаяся к нулю на больших расстояниях, отвечает короткодействующей поправке.

Коллективной модой колебаний является плазменная частота, которая для конкретного иона равняется (в единицах СИ):

$$\omega_p = \left(\frac{nq^2}{\epsilon_0 m} \right)^{1/2}, \quad (12.50)$$

где n — плотность этого иона, а m — масса иона. Из (12.50) ясно, что самый легкий ион имеет самую высокую частоту, и он

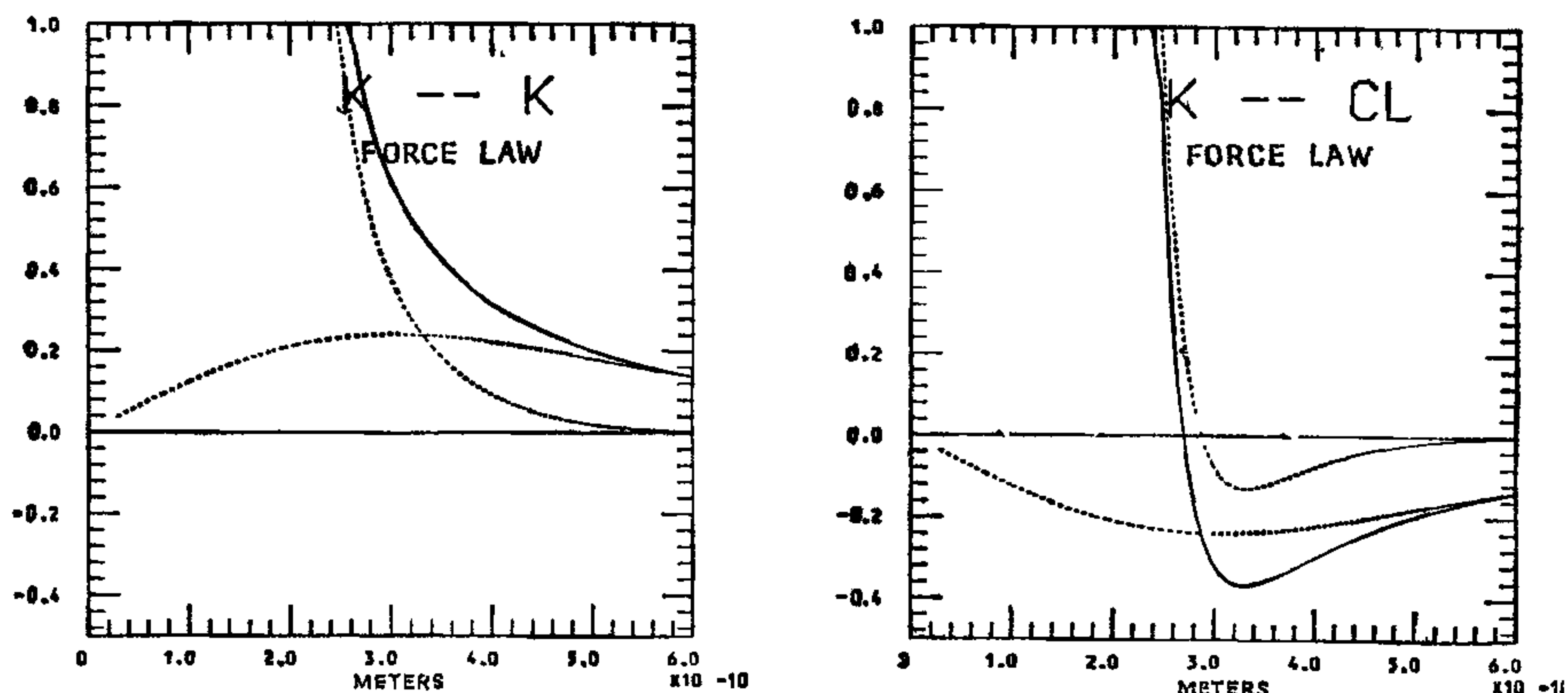


Рис. 12.7. Закон взаимодействия одноименных и разноименных ионов в модели KCl. Шаг сетки $H = 1.57 \times 10^{-10}$ м и короткодействующая поправка распространяется на расстояния до $r_c = 6.28 \times 10^{-10}$ м. Полная сила (сплошная линия), сеточная сила (пунктирная линия, идущая в начало координат), короткодействующая поправка (другая пунктирная линия). ([120], с любезного разрешения Computer Physics Communications, © North-Holland, Publishing Co.).

будет определять шаг по времени. Переходя к безразмерным переменным, имеем

$$\omega_p = (4\pi n/m)^{1/2}, \quad (12.51)$$

а по отношению к частоте молекулы

$$\frac{\omega_p}{\omega_0} = \left(\frac{4\pi n}{m\rho} \right)^{1/2}. \quad (12.52)$$

Для случая ионов примерно одинаковой массы (т. е. K^+ и Cl^-) безразмерная масса $m = 2$. Безразмерная плотность n — число положительных ионов в объеме $r^* \times r^* \times r^*$. Для решетки KCl $n = 1/2$, отсюда для случая $\rho = 8$

$$\frac{\omega_p}{\omega_0} = \frac{\omega_p DT}{\omega_0 DT} = \left(\frac{\pi}{\rho} \right)^{1/2} = 0.6. \quad (12.53)$$

Таким образом, мы показали, что для типичного галоида щелочного металла плазменная частота меньше частоты колебаний молекулы. Любой шаг по времени, выбранный, исходя из (12.38), для качественного интегрирования молекулярных колебаний, будет столь же хорош для интегрирования плазменной частоты. Из выражения (12.52) очевидно, что, когда отталкивание становится слабее ($\rho < 3$), плазменная частота может превысить частоту молекулы, и тогда уже она будет определять временной шаг. Такая ситуация имеет место для трехмерной однокомпонентной плазмы, для которой $\rho = 0$. Однако в случае галоидов щелочных металлов найдено, что $\rho \geq 8$ и, значит, шаг по времени всегда определяется колебательными модами.

12.3.3. Пространственная сетка и экономия машинного времени

При выборе пространственной сетки руководствуются желанием минимизировать время счета и объем машинной памяти. Вычислительные эксперименты, о которых здесь сообщается, выполнялись с пространственной сеткой $16 \times 16 \times 16$, насчитывающей 4096 ячеек. Ширина одной ячейки $H = 3.36 \times 10^{-10}$ м, а внешняя граница короткодействующей силы $r_c = 3.2H$. Цепочечная сетка имеет размер $5 \times 5 \times 5$ ячеек, при этом шаг сетки равен $3.2H$. При этих параметрах погрешности в вычислении силы составляют (см. разд. 8.3 и 8.6):

Среднеквадратичная погрешность силы $Q = 1.0 \times 10^{-4}$.

Среднеквадратичные флуктуации силы $P = 7.7 \times 10^{-5}$.

Среднеквадратичное отклонение силы от ее среднего $Z = 2.3 \times 10^{-5}$.

Короткодействующая сила протабулирована как функция r^2 , чтобы не надо было извлекать квадратный корень. Для каждой пары ионов используется линейная интерполяция по таблице из 512 значений. Функция радиального распределения (ФРР) накапливается вплоть до $r = r_c$ в ходе вычисления короткодействующей поправки. Давление и потенциальная энергия вычисляются как интегралы от ФРР, при этом в результате обрывания интегрирования на $r = r_c$ возникают следующие погрешности (см. разд. 12.3.4(4)):

Аппроксимация давления 1.6×10^4 Н/м²;

Аппроксимация энергии 3.4×10^1 Дж/кг.

Для случая системы из 512 ионов на IBM 360/195 получены следующие времена работы программы:

(1) Заполненная область	(2) Микрокристалл
QSHARE = 0.07 с	0.07 с
POTP3M = 0.30 с	0.30 с
MESHFC = 0.10 с	0.10 с
SRFORC = 0.40 с	0.60 с
NEWENX = 0.02 с	0.02 с
ВРЕМЯ ШАГА = 0.89 с	1.09 с

Измеренные по отдельности времена относятся соответственно к задаче заряда на сетку (QSHARE), нахождению потенциала (POTP3M), вычислению сеточной силы (MESHFC), вычислению короткодействующей поправки (SRFORC) и изменению положений частиц (NEWENX). В обоих вариантах время вычисления короткодействующей поправки превышает время расчета потенциала, что говорит о том, что соотношение масштабов задачи не является оптимальным. Нетрудно, однако, понять, что

выбор более мелкой сетки размером $32 \times 32 \times 32$ привел бы к еще худшему балансу. Время на РОТРЗМ увеличилось бы при этом в 8 раз и составило 2.4 с, а время на SRFORC уменьшилось бы в 8 раз до ~ 0.05 с. Общее время счета шага возросло бы до примерно 2.6 с. Требуемый для программы и данных объем памяти при сетке 16^3 составляет 476 кбайт.

12.3.4. Измерения термодинамических характеристик

В программе РЗМЗДР на каждом шаге по времени предусмотрена следующая диагностика:

1. *Потенциальная энергия.* Полная потенциальная энергия

$$V = \frac{1}{2} \sum_{\substack{i, s \\ (i', s') \neq (i, s)}} \sum_{i', s'} \varphi_{i, i'} \quad (12.54)$$

по аналогии с силой расщепляется на короткодействующую и сеточную части. Потенциальная энергия $\varphi_{i, i'}$ парного взаимодействия частицы i сорта s с частицей i' сорта s' записывается в виде

$$\varphi_{i, i'} = \varphi_{i, i'}^{sr} + \varphi_{i, i'}^m \quad (12.55)$$

Потенциальная энергия поля короткодействующей силы V_{sr} вычисляется, исходя из парной корреляционной функции короткодействия, получаемой в процессе вычисления короткодействующей силы

$$V_{sr} = \sum_p \sum_{(i, i') \in P(p)} \varphi_{i, i'}^{sr} \quad (12.56)$$

Сумма по p берется по всем цепочечным ячейкам, а $P(p)$ представляет собой множество пар, в каждой из которых хотя бы один элемент принадлежит цепочечной ячейке p .

Сеточная потенциальная энергия получается посредством аппроксимации $\varphi_{i, i'}^m$ ее сеточными значениями

$$\varphi_{i, i'}^m \approx \frac{q_s q_{s'}}{\epsilon_0} \sum_{p, p'} W(\mathbf{x}_i - \mathbf{x}_p) G(\mathbf{x}_p - \mathbf{x}_{p'}) W(\mathbf{x}_{i'} - \mathbf{x}_{p'}), \quad (12.57)$$

что дает
$$V_m = \frac{V_c}{2} \sum_p \rho(p) \varphi(p) - V_{сам}, \quad (12.58)$$

где $V_c = H^3$ — объем ячейки.

Суммы по p и p' , входящие в (12.57) и (12.58), берутся по всем сеточным узлам на сетке заряда-потенциала. Член $V_{сам}$ появляется в (12.58) из-за того, что в сумму по узлам сетки входит и член с $(i, s) = (i', s')$.

Первый член в выражении (12.58) вычисляется по гармоникам сеточной плотности заряда

$$\frac{V_c}{2} \sum_{\mathbf{p}} \rho(\mathbf{p}) \varphi(\mathbf{p}) \equiv \frac{1}{2V_b} \sum_{\mathbf{k}} \hat{G}(\mathbf{k}) |\hat{\rho}(\mathbf{k})|^2, \quad (12.59)$$

где $V_b = L^3$ — объем области.

Второй член находится по формуле

$$V_{\text{сам}} = \left(\frac{13}{35\pi\epsilon_0} \right) \sum_s N_s q_s^2, \quad (12.60)$$

где N_s — число частиц сорта s , а выражение в скобках представляет собой собственную энергию единичного заряда, имеющего форму S_2 .

2. *Кинетическая энергия.* Кинетическая энергия \mathcal{T} вычисляется по импульсам частиц p_i :

$$\mathcal{T} = \sum_{i,s} \frac{p_i^2}{2m_s} \quad (12.61)$$

и пересчитывается на те же временные слои n , на которых известна потенциальная энергия, с помощью аппроксимации

$$\mathcal{T}^n = \sum_{i,s} \left[\frac{(p_i^{n+1/2})^2 + (p_i^{n-1/2})^2}{4m_s} \right] + \frac{V^{n+1} - 2V^n + V^{n-1}}{8}. \quad (12.62)$$

3. *Давление.* Давление вычисляется как

$$P = P_K + P_\varphi = \quad (12.63)$$

$$= nk_B T + \frac{1}{6V_b} \sum_{\substack{(i,s) \\ (i',s') \neq (i,s)}} \sum_{(i',s')} (\mathbf{x}_i - \mathbf{x}_{i'}) \cdot \mathbf{f}_{ii'}, \quad (12.64)$$

где n — плотность числа частиц, k_B — постоянная Больцмана, T — температура, $\mathbf{f}_{ii'}$ — сила парного взаимодействия, и сумма по (i, s) берется по всем частицам в расчетной области, а сумма по (i', s') — по всем частицам в расчетной области и ее периодических изображениях.

Кинетическое давление¹⁾ P_K вычисляется через кинетическую энергию

$$P_K = nk_B T = \frac{2\mathcal{T}}{3V_b}. \quad (12.65)$$

¹⁾ Иначе, газокинетическое давление. — Прим. перев.

Давление P_Φ , обусловленное взаимодействием, находится путем расщепления суммы в правой части (12.64) на полную потенциальную энергию¹⁾ и поправку

$$P_\Phi = \frac{V}{3V_b} + \frac{1}{3V_b} \sum_p \sum_{(ij) \in P(p)} [(\mathbf{x}_i - \mathbf{x}_j) \cdot \mathbf{f}_{ij}^+ - \Phi_{i,j}^+]. \quad (12.66)$$

Для вычисления этой поправки используются значения парной корреляционной функции короткодействия, сосчитанные при расчете короткодействующей силы. $\Phi_{i,j}^+$ представляет собой потенциальную энергию парного взаимодействия, отвечающую некулоновской части \mathbf{f}_{ij}^+ короткодействующей силы.

4. Погрешности аппроксимации энергии и давления. В случае когда зависимость силы от расстояния имеет вид кулоновской силы с отталкивающим ядром, обрывание таблиц короткодействующей силы на $r = r_c$ может приводить к заметным отличиям измеренных значений давления (в части, обусловленной взаимодействием) и потенциальной энергии от ожидаемых даже в случаях, где динамика моделируется надежно. Причина этого расхождения связана с тем, что в численной схеме малые значения силы отталкивания при $r \geq r_c$ заменяются нулем.

Если короткодействующие поправки обрываются для пар, разделенных расстоянием, большим r_c , то погрешность, привносимая в вычисление потенциальной энергии, равняется

$$\Delta V = \frac{1}{2} \sum_{(i,s)} \sum_{\substack{(i',s') \\ r_{i,i'} > r_c}} \Phi_{i,i'}^+ \quad (12.67)$$

где $\Phi_{i,i'}^+$ — вклад отталкивающего ядра в потенциальную энергию парного взаимодействия частиц i и i' . Перепиывая (12.67) через функцию радиального распределения $g_{s,s'}(r)$ для сортов s и s' , получаем

$$\Delta V = \frac{1}{2} \sum_{s,s'} N_s \frac{N_{s'}}{V_b} \int_{r_c}^{\infty} 4\pi r^2 g_{s,s'}(r) \Phi_{s,s'}^+(r) dr. \quad (12.68)$$

Оценка величины ΔV получается посредством замены $g_{s,s'}$ ее асимптотическим значением, равным единице, с последующим вычислением интеграла:

$$\Delta V \simeq \frac{1}{2} \sum_{s,s'} N_s \frac{N_{s'}}{V_b} \int_{r_c}^{\infty} 4\pi r^2 \Phi_{s,s'}^+(r) dr. \quad (12.69)$$

¹⁾ В единице объема. — Прим. перев.

Аналогичным образом получается оценка погрешности давления:

$$\Delta P_{\Phi} \simeq \frac{\Delta V}{3V_b} + \frac{1}{6V_b} \sum_{s,s'} N_s \frac{N_{s'}}{V_b} \int_{r_c}^{\infty} 4\pi r^2 [rf_{s,s'}^+(r) - \varphi_{s,s'}^+(r)] dr = \quad (12.70)$$

$$= \frac{1}{6V_b^2} \sum_{s,s'} N_s N_{s'} \int_{r_c}^{\infty} 4\pi r^3 f_{s,s'}^+(r) dr \quad (12.71)$$

Существуют две возможности использования этих поправочных членов. Если мы рассматриваем численную процедуру как средство точно промоделировать «приближенный» закон изменения силы с расстоянием (а именно закон, в котором $f^+ \equiv 0$ для $r > r_c$), то ΔV и ΔP_{Φ} не надо включать в расчет энергии и давления соответственно и использовать только как меру отличия «приближенного» и «точного» законов взаимодействия. Если, однако, эта процедура рассматривается в качестве приближенного метода моделирования «точного» закона взаимодействия (т. е. закона, в котором f^+ очень мала при $r \geq r_c$, но не нуль), то ΔV и ΔP_{Φ} следует добавить к V и P_{Φ} соответственно. В обоих случаях желательно выбирать r_c так, чтобы $\Delta V/V$ и $\Delta P_{\Phi}/P_{\Phi}$ были малы. В данной программе ΔV и ΔP_{Φ} подсчитываются и включаются в выходные данные на стадии инициализации, но не прибавляются к V и P_{Φ} .

5. Парная корреляционная функция. На каждом шаге по времени для всех пар сортов (s, s') вычисляется функция радиального распределения $g_{s,s'}(r_k)$ для расстояний между частицами вплоть до радиуса обрезания r_c таблиц короткодействующей силы. Значения $g_{s,s'}(\mathbf{x}_i; r_k)$ получаются при помощи обратной линейной интерполяции по частицам i' сорта s' на сетке концентрических сферических ячеек с центром в частице i сорта s , где узел ячейки k расположен в точке $r_k^2 = k \Delta r^2$; границы этих ячеек определяются точками $r_-^2 = r_k^2 - \Delta r^2/2$ и $r_+^2 = r_k^2 + \Delta r^2/2$ и они имеют объем V_k . Функция радиального распределения $g_{s,s'}(r_k)$ получается усреднением $g_{s,s'}(\mathbf{x}_i; r_k)$ по всем частицам i сорта s . Если $s = s'$, такие члены не включаются в функцию распределения.

Функцию радиального распределения $g_{s,s'}(r_k)$ можно записать в виде сумм по парам частиц (i, i') , где i обозначает частицы сорта s , а i' — частицы сорта s' :

$$g_{s,s'}(r_k) = \frac{2V_b}{N_s N_{s'} V_k} \sum_{\mathfrak{p}} \sum_{(i, i') \in P(\mathfrak{p})} U(|\mathbf{x}_i - \mathbf{x}_{i'}|^2 - r_k^2). \quad (12.72)$$

Функция U — функция обратной линейной интерполяции:

$$U(\xi) = \begin{cases} 1 - \frac{|\xi|}{\Delta r^2}, & \text{если } |\xi| \leq \Delta r^2, \\ 0, & \text{в других случаях,} \end{cases} \quad (12.73)$$

и V_b — объем расчетной области, N_s — число частиц сорта s , V_k — объем сферической ячейки k . Сумма по p берется по цепочечным ячейкам, а $P(p)$ обозначает множество пар, в каждой из которых хотя бы один элемент принадлежит цепочечной ячейке p .

Потенциальная энергия поля короткодействующей силы V_{sr} и часть давления, связанная с взаимодействием, вычисляются по $g_{s,s'}(r_k)$ и табулированным значениям $\varphi_{s,s'}(r_k)$ и $[r_k f_{s,s'}^+(r_k) - \varphi_{s,s'}^+(r_k)]$ с помощью следующих выражений:

$$V_{sr} = \sum_{s,s'} \sum_k \frac{N_s N_{s'} V_k}{2V_b} g_{s,s'}(r_k) \varphi_{s,s'}(r_k), \quad (12.74)$$

$$P_\varphi = \frac{1}{3V_b} \left\{ V + \sum_{s,s'} \sum_k \frac{N_s N_{s'} V_k}{2V_b} g_{s,s'}(r_k) [r_k f_{s,s'}^+(r_k) - \varphi_{s,s'}^+(r_k)] \right\}. \quad (12.75)$$

Суммы по k в (12.74) и (12.75) берутся по всем позициям таблиц.

6. Коэффициент диффузии. Несмотря на то что измерение коэффициента диффузии не включено в опубликованную версию программы РЗМЗДР, это важное измерение можно легко добавить. Уравнение, описывающее постоянную анизотропную диффузию в d -мерном пространстве, имеет вид

$$\sum_{i=1}^d D_i \frac{\partial^2 n}{\partial x_i^2} = \frac{\partial n}{\partial t}, \quad (12.76)$$

где $n(x_i, t)$ — плотность числа частиц, при этом вместо обычных обозначений координат x, y, z используются x_1, x_2, x_3 . Подстановкой можно показать, что решением этого уравнения, отвечающим начальному условию $x_i = 0$ при $t = 0$ для всех N частиц, является функция

$$n(x_i, t) = N A \exp \left[- \sum_{i=1}^d (x_i^2 / 4D_i t) \right], \quad (12.77a)$$

где $A = \prod_{i=1}^d (4\pi D_i t)^{-1/2}. \quad (12.77b)$

Через время t частица будет находиться в точке x_i и ее смещение Δx_i относительно начального положения равно x_i ; отсюда среднеквадратичное смещение составляет

$$\langle (\Delta x_i)^2 \rangle = N^{-1} \int_{-\infty}^{\infty} n(x_i, t) x_i^2 dx_i = \quad (12.78)$$

$$= 2D_i t. \quad (12.79)$$

Следовательно, для одномерной диффузии коэффициент диффузии определяется выражением

$$D_i = \frac{1}{2} \langle (\Delta x_i)^2 \rangle / t. \quad (12.80)$$

Если диффузия изотропна, т. е. $D_1 = D_2 = D_3 = D$, то ее можно также определить из квадрата длины вектора смещения

$$|\Delta \mathbf{r}|^2 = \sum_{i=1}^d (\Delta x_i)^2,$$

среднее которого равно

$$\langle |\Delta \mathbf{r}|^2 \rangle = \sum_{i=1}^d \langle (\Delta x_i)^2 \rangle = \quad (12.81a)$$

$$= 2dDt, \quad (12.81b)$$

откуда
$$D = \frac{1}{2d} \langle |\Delta \mathbf{r}|^2 \rangle / t. \quad (12.82)$$

В случае ансамбля частиц измеряется смещение каждой частицы от ее начального положения

$$\Delta x_{i,j} = x_{i,j}^n - x_{i,j}^0, \quad (12.83)$$

где $x_{i,j}^n$ — i -я компонента координаты j -й частицы в момент времени $t = nDT$. Тогда среднеквадратичное смещение равняется

$$\langle (\Delta x_i)^2 \rangle = N^{-1} \sum_{j=1}^N (\Delta x_{i,j})^2. \quad (12.84)$$

Для проведения этих вычислений в программе можно один раз запомнить $x_{i,j}^0$ и далее просчитывать (12.83), (12.84), (12.81) и (12.82). С другой стороны, можно найти смещение, накапливая на каждом шаге по времени скорость для каждой частицы отдельно следующим образом:

$$\Delta x_{i,j} = \sum_{k=1}^N v_{i,j}^{k-1/2} DT, \quad (12.85)$$

где $v_{i,j}^{k-1/2}$ — i -я компонента скорости j -й частицы на k -м временном шаге.

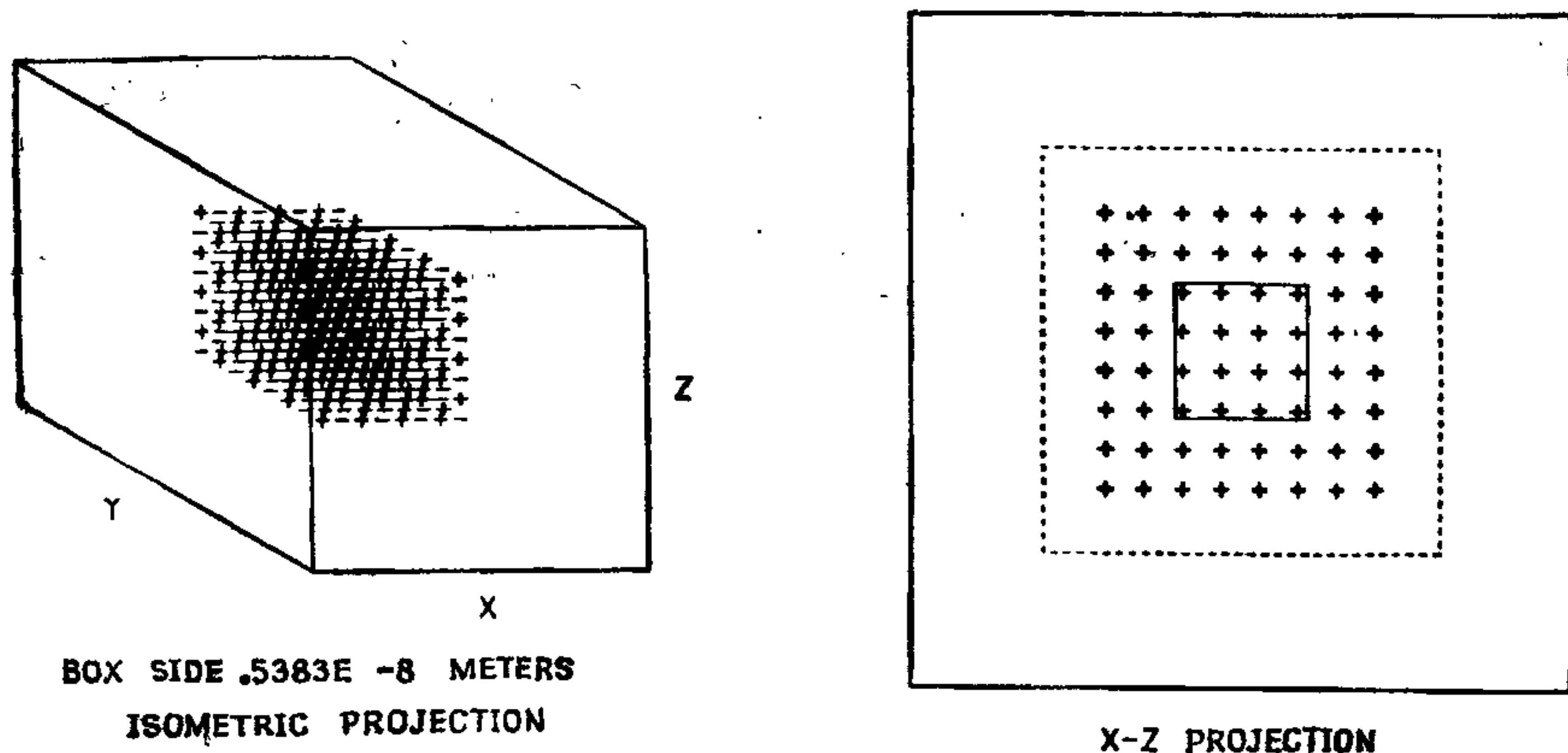


Рис. 12.8 Изометрическая проекция (слева) микрокристалла, заполняющего часть расчетной области, и (справа) сечение $X-Z$ плоскостью, показывающее расположение области 1 (сплошные линии) и области 2 (пунктирные линии), применяемые при измерении ФРП и давления в центре микрокристалла. ([120], с любезного разрешения Computer Physics Communications, © North-Holland Publishing Co.)

В любом случае требуется дополнительная память — по одной ячейке на каждую компоненту координаты каждой частицы. Если это больше, чем можно себе позволить, то измерения можно проводить только для какой-то части частиц. В то же время, если известно, что диффузия изотропна, ее достаточно вычислять только по одной координате и пользоваться формулой (12.80). Тогда требуется только одна дополнительная ячейка памяти на частицу. Если необходимо найти коэффициент диффузии, исходя из множества начальных координат, которые вычисляемы, то никакой дополнительной памяти выделять не нужно, поскольку $x_{i,j}^0$ можно пересчитывать всякий раз, когда необходимо измерить диффузию.

12.3.5. Измерения с разными областями

Измерения, перечисленные в разд. 12.3.4, представляют собой средние по всем частицам системы, и наиболее удобны, если вещество полностью заполняет расчетную область. Однако в случае микрокристалла, занимающего только часть рассматриваемой системы, измерения радиального распределения и давления требуется производить по-другому. Это достигается тем, что определяются две области, как, например, показано на рис. 12.8. Область 1 представляет собой прямоугольный параллелепипед, который задает изучаемую область, а область 2 полностью ее охватывает и содержит все частицы в заданных пределах. Для расчета ФРП, $g_{ss'}(r)$, все частицы сорта s из об-

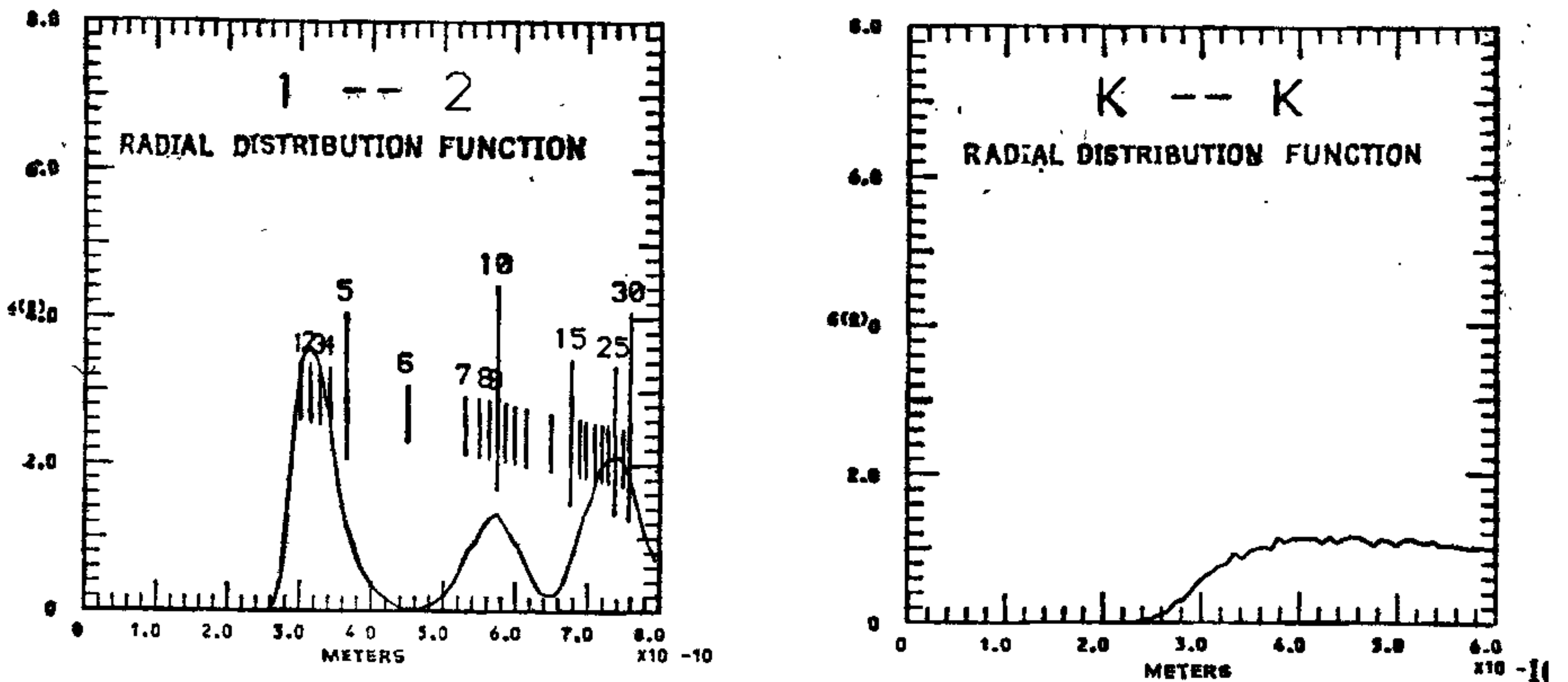


Рис. 12.9. Функция радиального распределения (слева) для микрокристалла KCl при температуре около 600 К и (справа) жидкий KCl, нагретый до 400 К. Ионы типа 1 K^+ , типа 2 Cl^- . ([120], с любезного разрешения Computer Physics Communications, © North-Holland Publishing Co.)

ласти 1, называемые опорными частицами, связываются попарно со всеми частицами сорта s' из области 2. Если N_i — количество ss' пар, приходящихся на опорную частицу сорта s в i -м кубике гистограммы объема V_i , то

$$g_{ss'}(r_i) = N_i / (n_s V_i), \quad (12.86)$$

где ФРР нормирована на среднюю плотность числа частиц n_s сорта s' в области 1. Вычисленная таким образом ФРР для одноименных и разноименных ионов показана на рис. 12.9. На корреляции $A-B$ числа над вертикальными координационными отметками указывают количество ионов B , окружающих опорный ион A на уровне положения соответствующей отметки. На рис. 12.9 (слева), который отвечает кристаллу KCl, из положения координационных чисел видно, что первая оболочка из ионов хлора вокруг иона калия содержит 6 ионов и что вторая оболочка из хлора имеет 8 ионов, что можно было предположить, исходя из строения кристалла (рис. 12.5). На правом рисунке показана корреляция между одноименными ионами в жидком KCl при очень высокой температуре 4000 К. Эта температура столь высока, что в ФРР не появляется ни одного пика.

Давление в области 1 получается из измерения вириала, обусловленного всеми взаимодействующими парами, в которых оба иона находятся в области 1. Вклад в вириал в области 1 от взаимодействий с частицами, находящимися вне области 1, хотя он и возникает из-за дальнедействующих сил, трактуется как эквивалентное внешнее давление на поверхность области 1. Тогда теорема вириала дает

$$P_1 = \frac{1}{V_1} \left(\frac{2}{3} \mathcal{T} + \frac{1}{3} \sum r_{ij} f_{ij} \right). \quad (12.87)$$

где r_{ij} — расстояние между парой ионов, взаимодействующих с силой f_{ij} , а сумма берется по всем парам, целиком лежащим в области 1; V_1 — объем области 1. Черта обозначает усреднение по времени, как правило, по 150 шагам.

12.3.6. Плавление, переохлаждение и стеклование

В работе [18] детально исследовано поведение одной соли, а именно модельной KCl, в процессе всех ее фазовых превращений. В проведенном вычислительном эксперименте в центре расчетной области помещался кубический микрокристалл из 512 ионов, находящийся в равновесном состоянии при температуре около 40 К, и подвергался циклическому нагреванию с периодом 200 шагов по времени. В течение первых 25 шагов тепло подводится посредством умножения на каждом шаге всех скоростей на коэффициент $\text{HTFAC} = 1.001$. На протяжении остальных 175 шагов нагрев отключен и система имеет возможность прийти в равновесное состояние. Внутренняя энергия U и температура T получают усреднением полной энергии системы и кинетической энергии по последним 32 шагам цикла.

На рис. 12.10 показана зависимость U от T , снятая в вычислительном эксперименте с моделью KCl. Кристалл нагревается от точки A до B , плавится между B и C и уже жидкость нагревается от C до D и далее до 1600 К. Эти точки показаны темными кружками на рис. 12.10. Затем система охлаждается до примерно 14 К посредством задания $\text{HTFAC} = 0.999$. Полученные точки показаны светлыми кружками. Заметим, что при нагревании фазовый переход 1-го рода происходит между точками B и C и по ним определяются точка плавления $T_f = T(B \text{ или } C)$ и скрытая теплота плавления $L_f = U(C) - U(B)$. Теплоемкости твердой и жидкой фаз получают из наклона прямых, проведенных из A в B и из C в D . Теплоемкость жидкости заметно превышает теплоемкость твердого тела. При охлаждении возврата к кристаллической форме не происходит, и участок от C до E соответствует переохлажденной жидкости. Обнаружено, что в окрестности $T_f/3 \approx 310$ К теплоемкость уменьшается до значения теплоемкости кристаллической фазы. Такое поведение характерно для стекол, и мы отождествляем эту температуру с температурой стеклования T_g . Насколько нам известно, это является первым наблюдением стеклования в микрокристаллах, хотя похожие результаты получены в расчете [388] для трижды периодической бесконечной системы с экспоненциальным отталкиванием.

Характеристики микрообразца, измеренные в вычислительном эксперименте, представлены ниже вместе с физическими значениями для макроскопического образца, которые указаны в скобках:

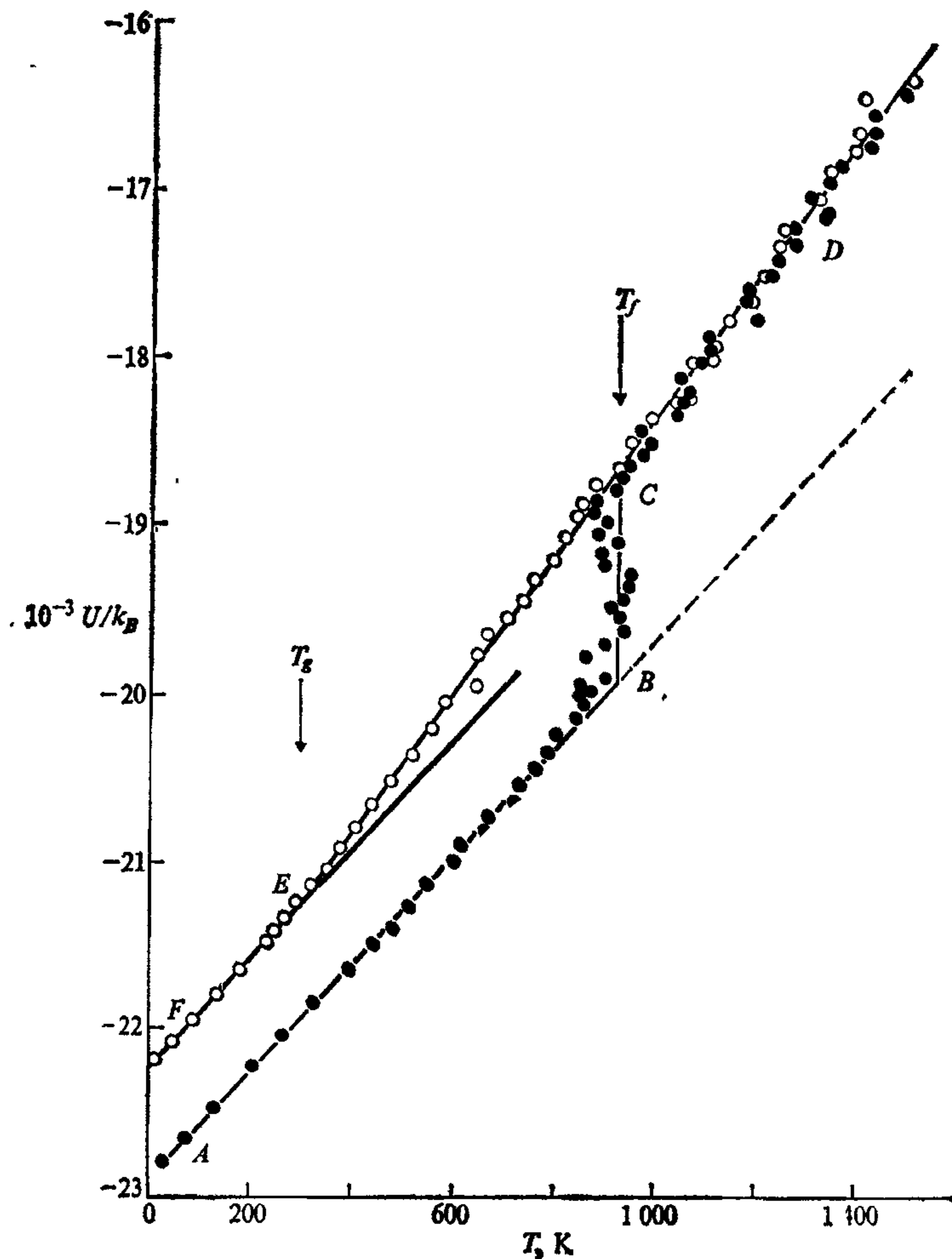


Рис. 12.10. Зависимость U от T , полученная в процессах плавления и стеклования на модельном KCl. T_f — точка плавления, T_g — точка стеклования. По обеим осям — градусы Кельвина. Начало отсчета U произвольное. ([18], с любезного разрешения Journal of Non-Crystalline Solids, © North-Holland Publishing Co.)

Фазовые превращения

- Точка плавления T_f = 950 ± 20 (1045) К,
 Скрытая теплота плавления L_f/k_B = 1240 ± 50 (1580) К,
 Точка стеклования T_g = 310 ± 50 К,
 Изменение энтропии $S_f = L_f/k_B T_f$ = 1.3 ± 0.08 (1.5)

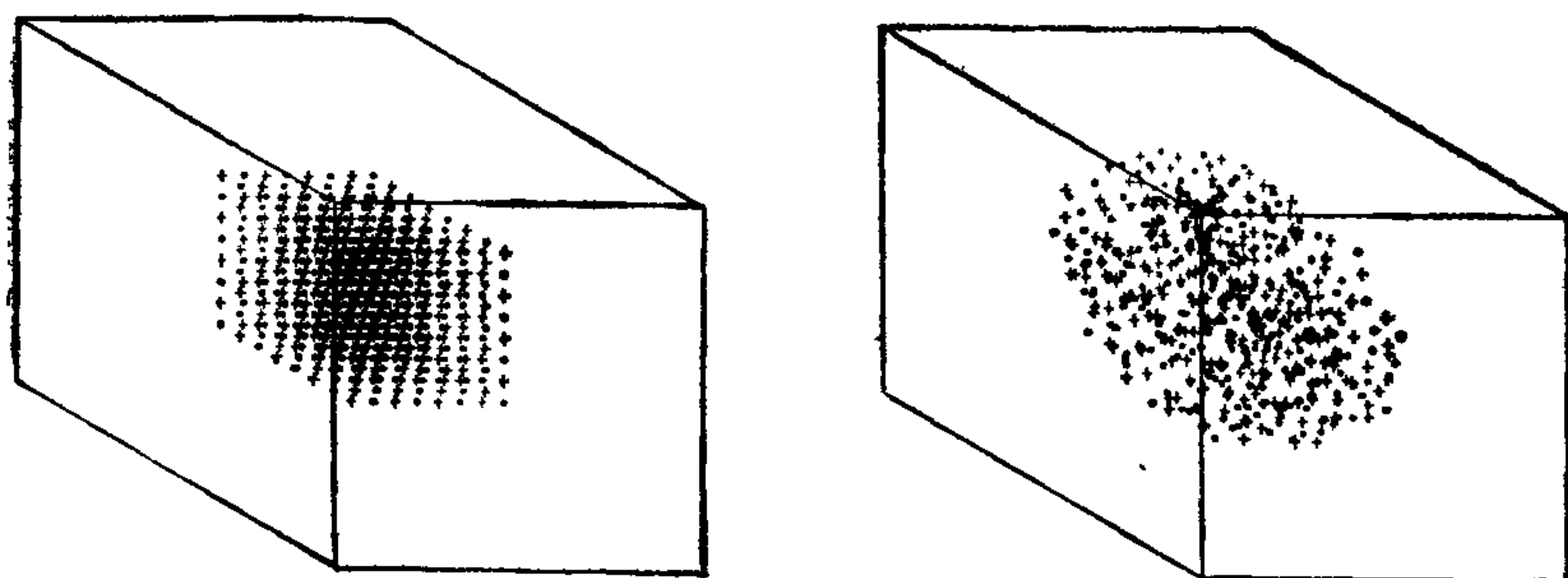


Рис. 12.11. Модель кристалла KCl из 512 ионов; (слева) до плавления, точка A на рис. 12.10; (справа) после плавления, точка C на рис. 12.10. Центры ионов калия отмечены (+), ионов хлора [•] ([18], с любезного разрешения Journal of Non-Crystalline Solids © North-Holland Publishing Co.)

Теплоемкость на ион при нулевом давлении

$$\begin{aligned}
 C_p \text{ твердой фазы} &= 3.24 \pm 0.05 \text{ (3.03) } k_B, \\
 C_p \text{ жидкости} &= 4.14 \pm 0.09 \text{ (4.03) } k_B, \\
 C_p \text{ стекла} &= 3.24 \pm 0.03 k_B.
 \end{aligned}$$

Наблюдаемые в численном эксперименте точка плавления и скрытая теплота плавления микрокристалла меньше соответствующих значений для реального макроскопического кристалла. Этого следовало ожидать, поскольку ионы, находящиеся на краю микрокристалла, не имея полного набора соседей, в каком-то смысле уже частично расплавлены. Можно предполагать, что по той же причине удельная теплоемкость микрокристалла будет заключена между значениями для реального кристалла и жидкости. Это также обнаружено.

На рис. 12.11 показано распределение ионов в микрокристалле и в жидкой капельке после плавления. Распределение в стекловидном состоянии похоже на распределение в капле. Мы определяем координационное число иона как количество ионов, его окружающих, на расстоянии вплоть до радиуса обрезания, который выбирается близким к первому минимуму функции радиального распределения. Координация ионов хлора вокруг калия неотличима от координации калия вокруг хлора. К счастью, результаты совершенно нечувствительны к выбору радиуса обрезания. На рис. 12.12 показано изменение координационного числа, усредненного по всем ионам, кроме тех, которые лежат в приграничном слое толщиной $\sim 6 \times 10^{-10}$ м. Видно резкое уменьшение координационного числа от 6 до 4 при плавлении и слабое увеличение координационного числа до 4.6 при охлаждении до стеклообразного состояния. Расстояние до ближайшего соседа уменьшается при плавлении на 8 %.

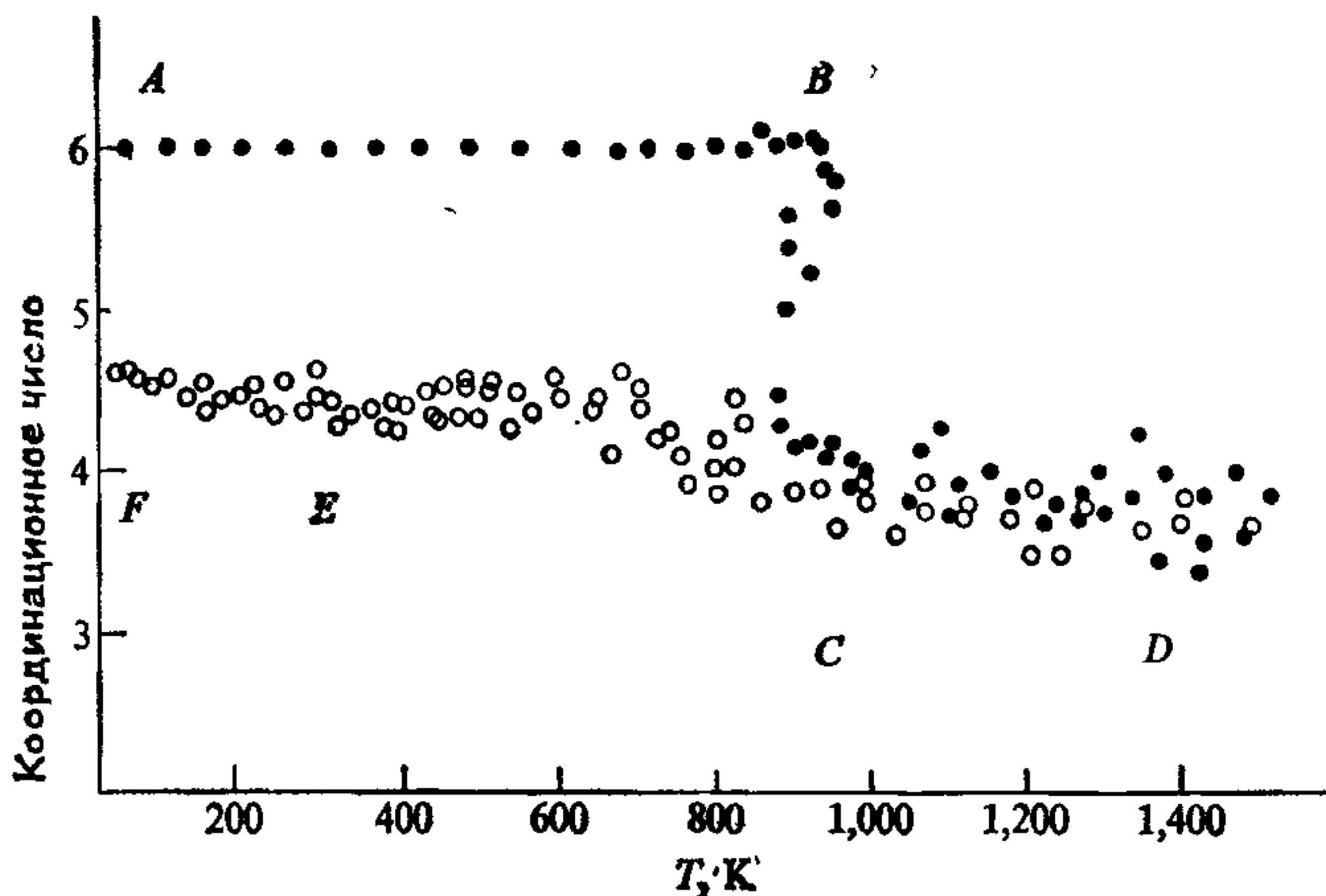


Рис. 12.12. Изменение координационного числа в ходе плавления и стеклования в модельном КСl. Буквы указывают положение на кривой $U(T)$ на рис. 12.10. ([18], с любезного разрешения Journal of Non-Crystalline Solids, © North-Holland Publishing Co.)

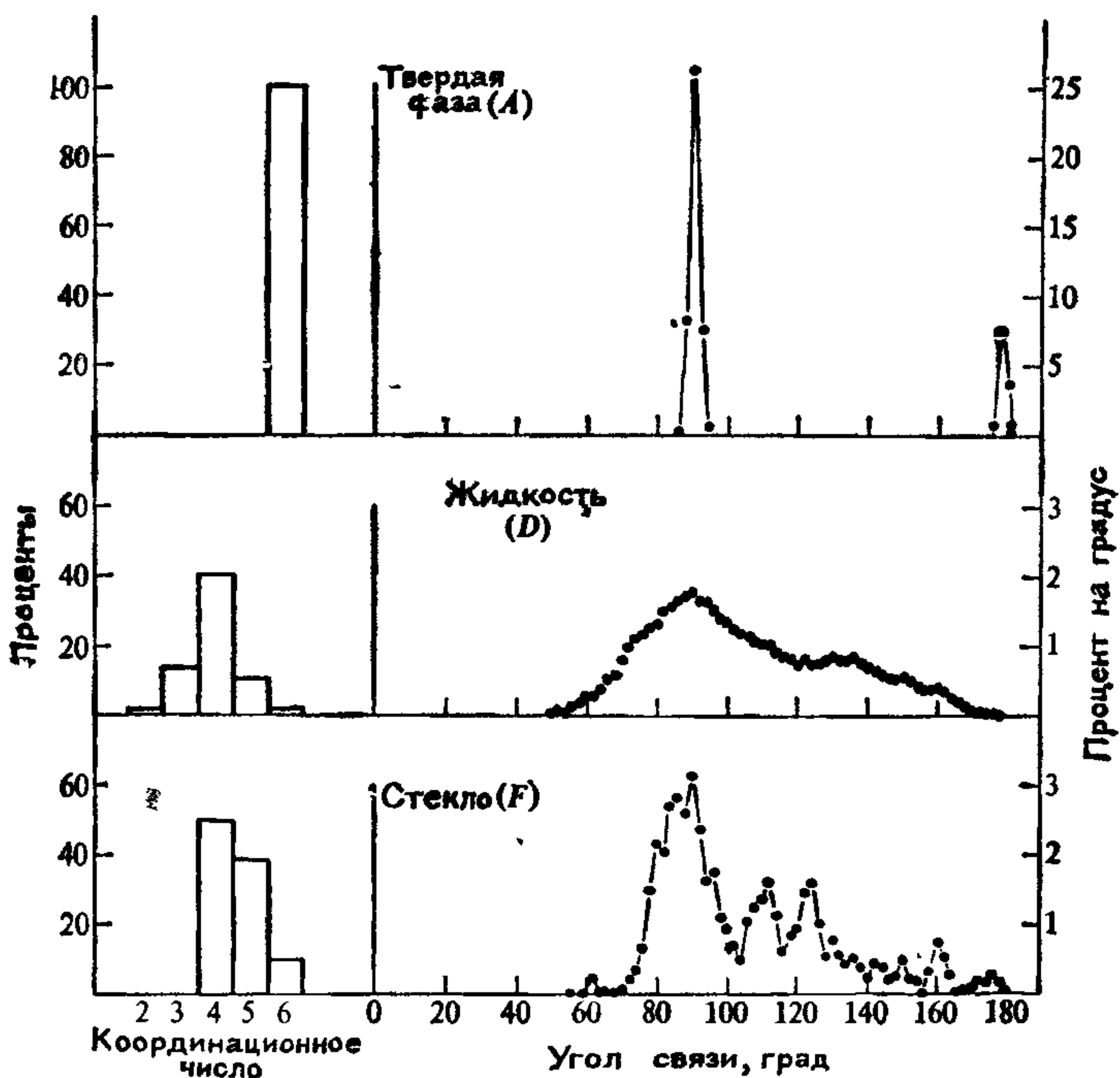


Рис. 12.13. Распределение по координационному числу и углу связи $SiKSi$ в твердом, жидком и стеклообразном состояниях. Буквы соответствуют положениям на кривой $U(T)$ на рис. 12.10. ([18], с любезного разрешения Journal of Non-Crystalline Solids, © North-Holland Publishing Co.)

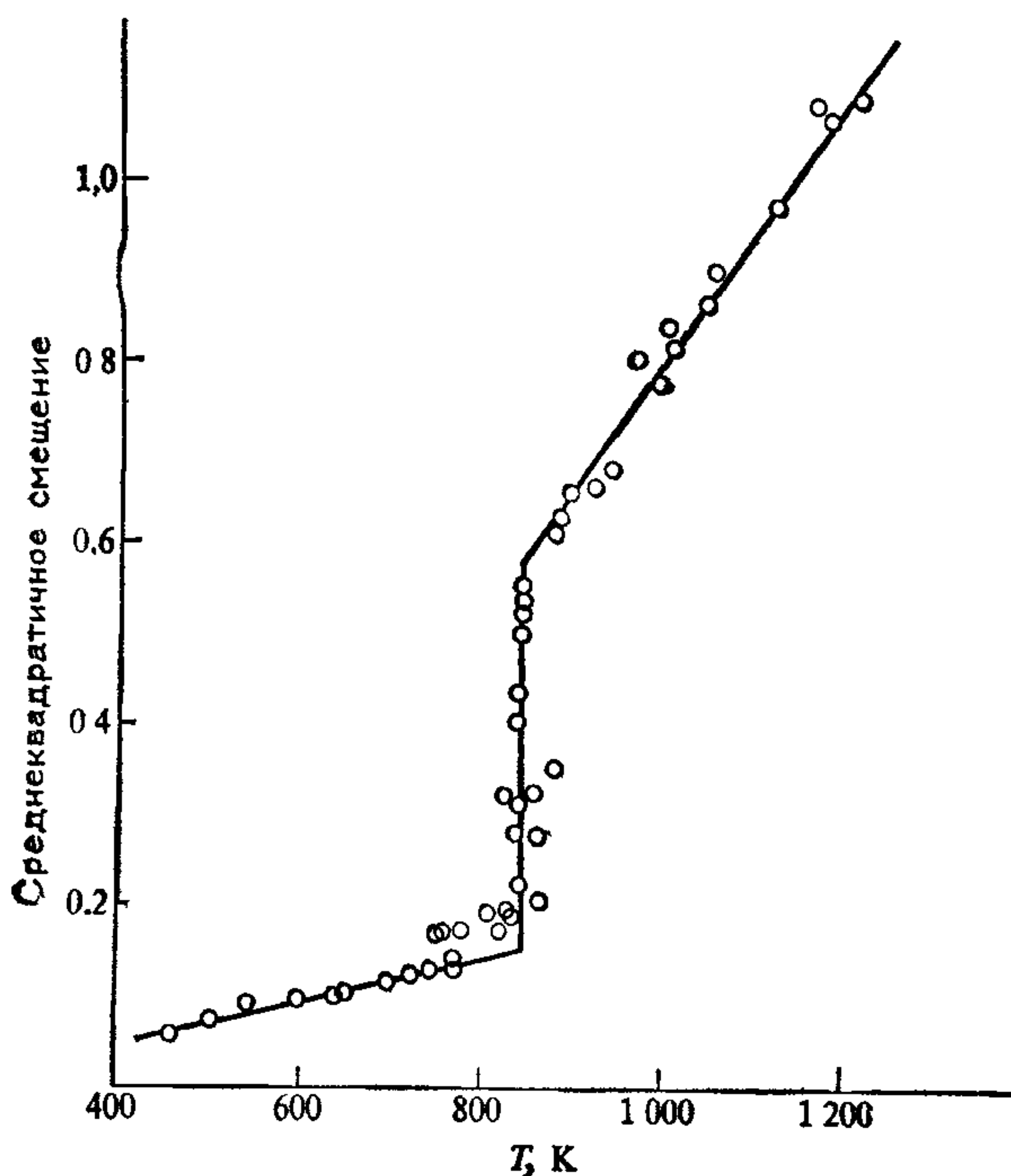
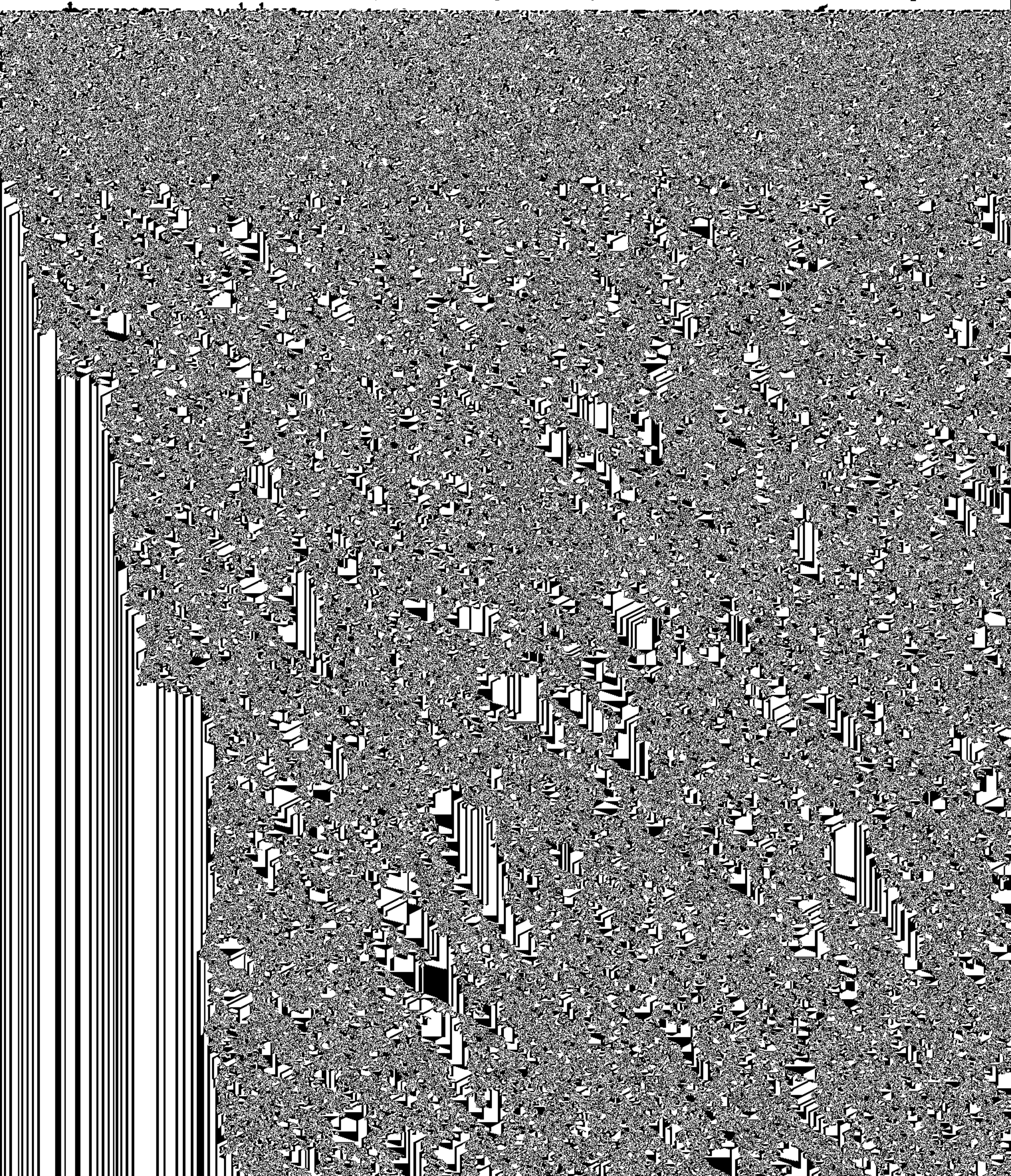


Рис. 12.14. Среднеквадратичное смещение (СКС) ионов за 175 шагов как функция температуры в вычислительном эксперименте с NaCl. ([17], с любезного разрешения J. Phys C Solid State Phys, © Institute of Physics.)

Мы получили также (рис. 12.13) распределение по координационному числу и углу связи Cl—K—Cl в твердом, жидком и стеклообразном состояниях. В твердой фазе координационное число всех ионов равно 6 и угол связи имеет пики при 90° и 180° , как и ожидалось, исходя из строения кристалла. В жидкости наблюдается симметричное распределение по координационным числам 3, 4 и 5, при этом угол связи имеет максимум в окрестности 90° и монотонно спадает до нуля при 180° . Самая поразительная особенность стекла состоит в полном исчезновении ионов с координационным числом 3 и возникновении трех пиков в распределении по углу связи при 90° , вблизи тетраэдрического угла 109° и при 125° . Все ионы дают примерно одинаковый вклад в первый пик, но второй и третий пики обязаны своим появлением в основном ионам с координационным числом, равным четырем.

Важным диагностическим средством в изучении плавления и стеклования является измерение диффузии (см. разд. 12.3.4(6)). В фазовом переходе плавления коэффициент диффузии резко возрастает. Это показано на рис. 12.14, относящемся к моделированию NaCl, на котором нанесена температурная зависимость

среднеквадратичного смещения, измеренная через 175 шагов по времени. Интересно отметить, что приблизительно за 100° до фазового перехода плавления диффузия заметно возрастает. Наблюдая за движением всех ионов в течение этой фазы «предплавления», можно обнаружить, что указанное явление обусловлено диффузией (или плавлением) ионов поверхностного слоя, которая может происходить при более низкой температуре, чем диффузия ионов во внутренних областях микрокристалла. Если измерять также коэффициент диффузии во время охлаждения жидкости, то можно идентифицировать точку стеклования как температуру, при которой экстраполированное значение коэф-



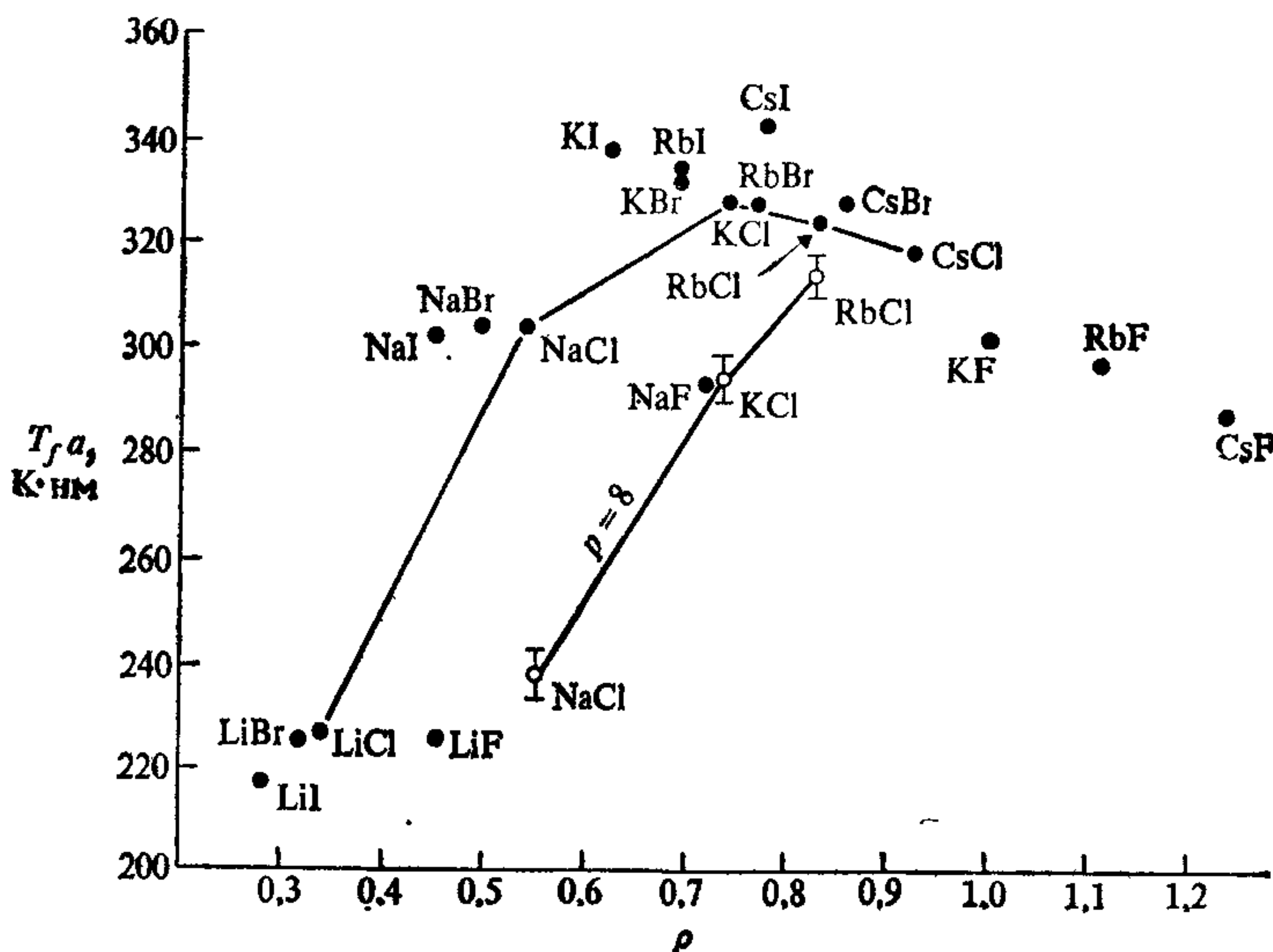


Рис. 12.15. Изменение точки плавления в зависимости от отношения радиусов ρ . Экспериментальные значения (●), численное моделирование (○). ([17], с любезного разрешения J. Phys. C: Solid-State Phys., © Institute of Physics.)

к RbCl, которое обнаружено в эксперименте. Кроме того, результаты моделирования все лежат ниже экспериментальных значений. Следовательно, при постоянном значении жесткости не удастся добиться удовлетворительного согласия между моделированием и экспериментом. Далее мы увидим (рис. 12.18), что точка плавления растет вместе с жесткостью, и согласие с экспериментальными значениями температур плавления может быть достигнуто путем задания для каждого вещества своей жесткости.

Нельзя, кроме того, забывать, что в данном случае сравниваются результаты моделирования микрокристалла из 512 ионов с экспериментальными измерениями на больших образцах. Имеются независимые данные о том, что точка плавления понижается с уменьшением размера образца [90, 96]. Это влияние размера подтверждено также в численных расчетах [16], в которых моделировалось плавление микрокристаллов, содержащих всего лишь 8 ионов. Эти результаты показаны на рис. 12.16, на котором нанесены значения температуры плавления модельного KCl в зависимости от числа ионов на ребре кубического микрокристалла.

Если считать, что влияние отношения радиусов, жесткости и размера можно рассматривать независимо, то температуру плавления можно представить в виде произведения трех множите-

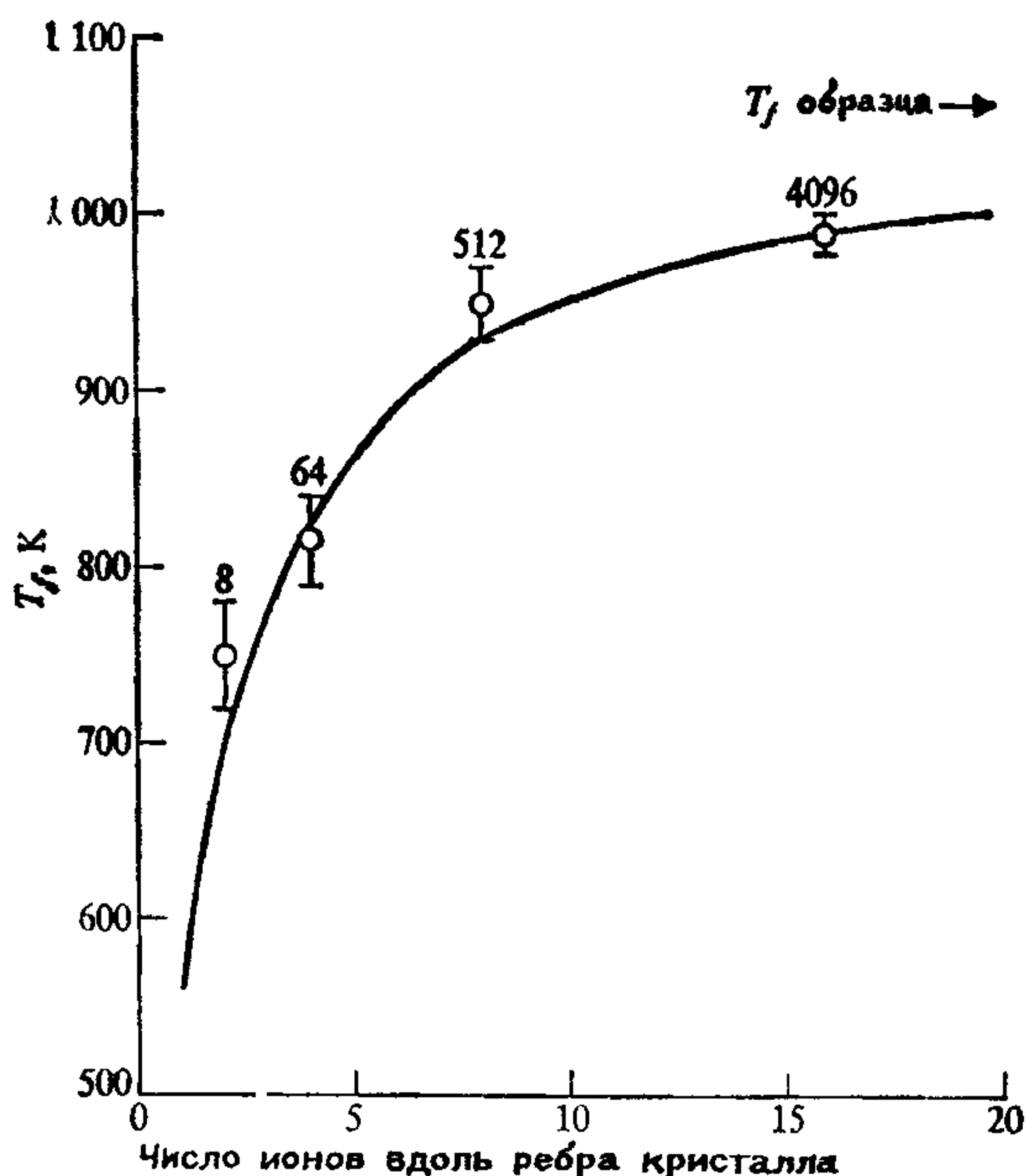


Рис. 12.16. Изменение точки плавления в зависимости от размера кристалла для модельного KCl с $p = 8$. Около точек указано полное число ионов в кристалле. Непрерывная кривая соответствует теоретическому предсказанию [94]. (С любезного разрешения М. Амини [16].)

лей, каждый из которых зависит только от одной из указанных переменных:

$$T_{fa} = R(\rho) H(\rho) S(N), \quad (12.88)$$

где N — число ионов в микрокристалле. В таком случае кривая на рис. 12.16 пропорциональна $S(N)$. Теоретические формулы, учитывающие влияние размера образца, были получены для плоского кристалла [96] и кубического микрокристалла [94]. Последней формуле отвечает на рис. 12.16 сплошная кривая, которая хорошо согласуется с результатами моделирования, показанными белыми кружками. Имеется подтверждение этой зависимости от размеров для реальных металлов [95].

Если результаты моделирования с 512 ионами сначала пересчитать с помощью множителя $S(N)$ на бесконечный кристалл, то удастся добиться согласия с экспериментальными данными при следующих параметрах жесткости:

$$\begin{aligned} \text{NaCl} & \quad p = 8.9, \\ \text{KCl} & \quad p = 7.9, \\ \text{RbCl} & \quad p = 7.4. \end{aligned} \quad (12.89)$$

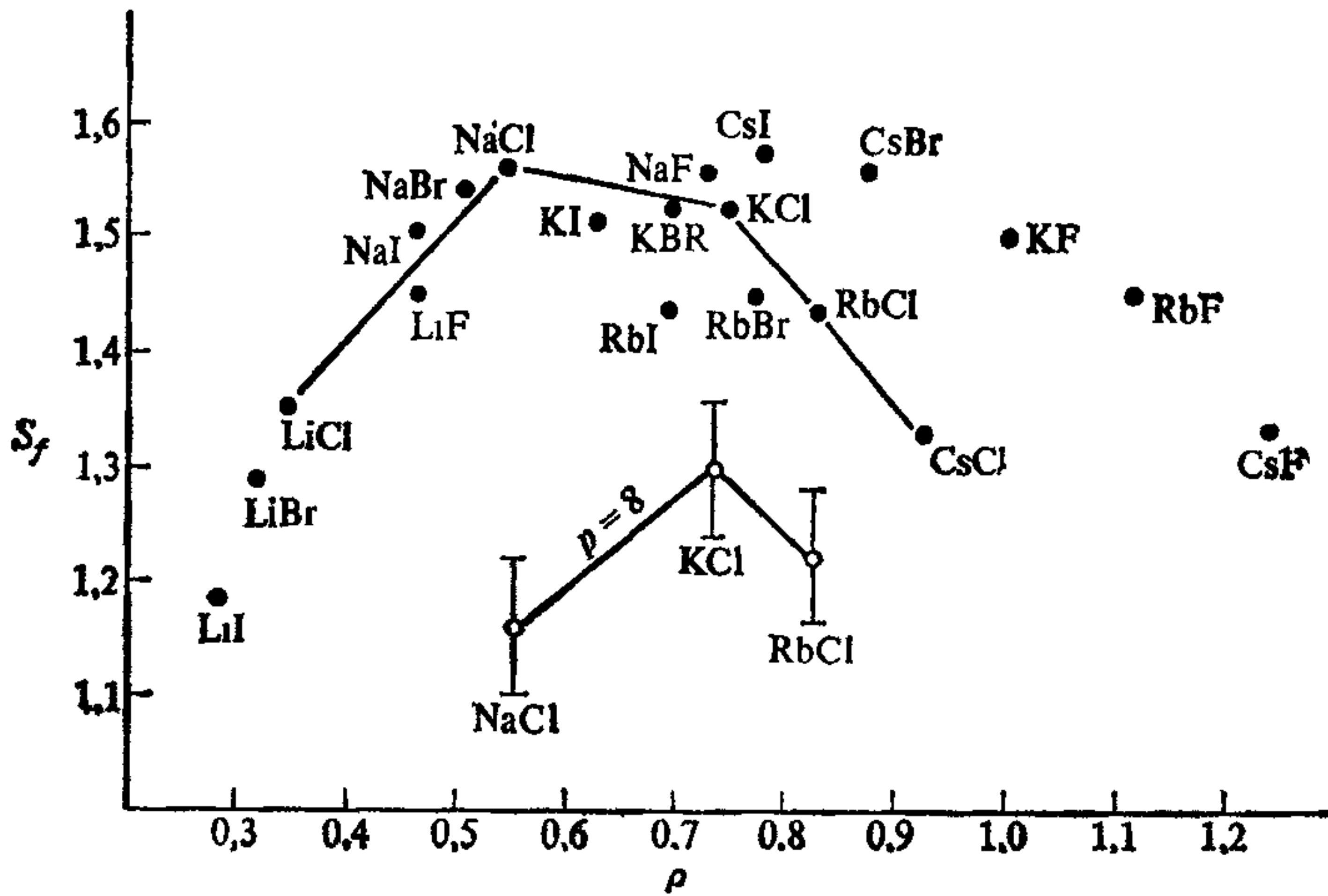


Рис. 12.17. Изменение энтропии плавления $S_f = L_f/k_B T_f$ в зависимости от отношения радиусов ρ . Экспериментальные значения (\bullet), численное моделирование (\circ). ([17], с любезного разрешения J. Phys. C: Solid-State Phys, © Institute of Physics)

На необходимость использовать переменную жесткость в потенциале взаимодействия галоидов щелочных металлов указывалось также в работе [17] при рассмотрении потенциала Фами — Този со значениями параметров из [341]. В этом случае параметр отталкивания α_{ij} в формуле (12.3) принимает разные значения для каждой соли.

Результаты для энтропии плавления представлены на рис. 12.17, из которых вновь с очевидностью следует, что постоянная жесткость не объясняет удовлетворительно экспериментальные результаты. Хотя общий характер поведения результатов моделирования качественно такой же, все полученные численно точки проходят ниже эксперимента и их максимум сдвинут в сторону больших значений ρ . Если результаты моделирования пересчитать на бесконечный кристалл и воспользоваться параметрами жесткости (12.89), то для энтропии плавления получим следующую связь модельных и экспериментальных результатов:

$$\begin{aligned} \text{NaCl} \quad S_f &= 1.35 (1.55), \\ \text{KCl} \quad S_f &= 1.40 (1.51), \\ \text{RbCl} \quad S_f &= 1.27 (1.44), \end{aligned} \tag{12.90}$$

где в скобках указаны экспериментальные значения.

Из приведенных результатов видно, что если параметр жесткости, входящий в простой закон изменения силы с расстоя-

нием (12.32), подобрать так, чтобы получить согласие с экспериментальными точками плавления, то значения энтропии плавления также лучше согласуются с экспериментом. Однако и в этом случае результаты численного моделирования все же неизменно меньше экспериментальных значений на 10—15%. Это наводит на мысль, что необходимо учесть какой-то дополнительный физический механизм, и тогда можно будет рассчитывать на получение хорошего согласия как по температуре плавления, так и по энтропии. Одной из возможностей является учет поляризации ионов с помощью модели оболочек (см. разд. 12.1.2 и работу [341, с. 306]).

12.3.8. Проверка теорий плавления

Результаты моделирования служат также хорошим пробным камнем для теорий плавления. Эти теории предсказывают разные зависимости от параметра жесткости. В лабораторных условиях эту зависимость удовлетворительно проверить нельзя, поскольку экспериментатор не в состоянии регулировать жесткость ионов. С другой стороны, вычислитель как раз очень точно контролирует этот параметр, и, значит, результаты вычислительного эксперимента могут быть использованы для проверки теории таким способом, который невозможен в лаборатории.

Вот три теории и их предсказания:

1. *Теория соответственных состояний*, в которой как $T_f a$, так и L_f пропорциональны $U(a)$. Из (12.42b) получаем

$$T_f a \sim p/(p+1), \quad S_f - \text{константа.} \quad (12.91)$$

2. *Критерий Линдемана* [264], согласно которому плавление кристалла происходит в случае, когда амплитуда тепловых колебаний достигает определенной части θ периода решетки. Считая силовую постоянную колебания λ пропорциональной $d^2U/dr^2 = pA/a^3$ (см. (12.44)) и приравняв энергию колебания $k_B T_f$, получим

$$k_B T_f = \frac{1}{2} \lambda (\theta a)^2 = \frac{1}{2} \frac{pA}{a^3} \theta^2 a^2, \quad (12.92a)$$

и отсюда

$$T_f a \sim p. \quad (12.92b)$$

Эта теория ничего не говорит относительно поведения скрытой теплоты и энтропии плавления.

3. *Теория дислокаций*, в соответствии с которой кристалл плавится, когда свободная энергия $U_d - TS_d$ для термической

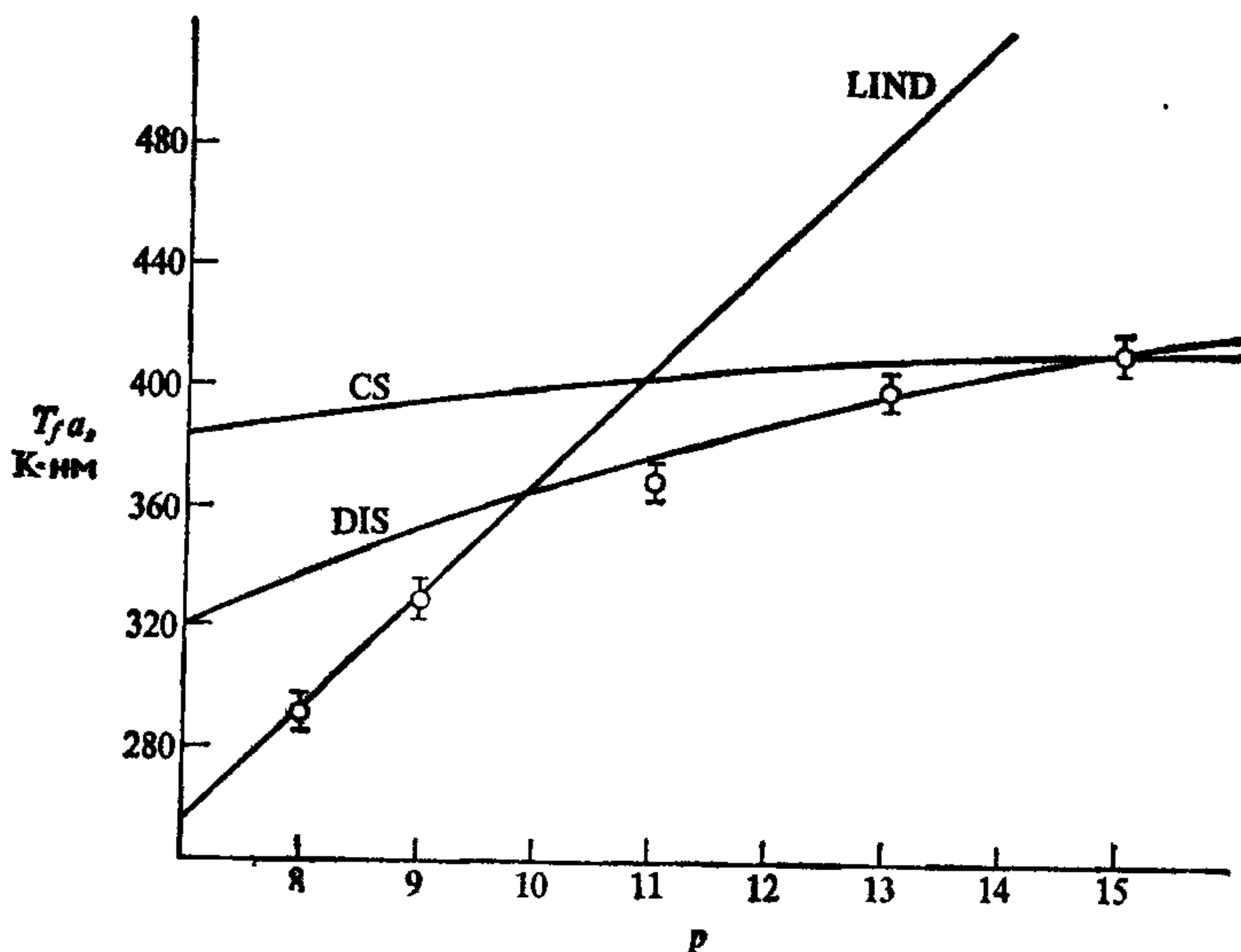


Рис. 12.18. Зависимость точки плавления T_{fa} от параметра жесткости ρ для модельного KCl. Сравниваются теория дислокаций (DIS), теория соответственных состояний (CS) и критерий Линдемана (LIND) с моделированием методом молекулярной динамики (\circ). ([17], с любезного разрешения J. Phys. C: Solid-State Phys., © Institute of Physics.)

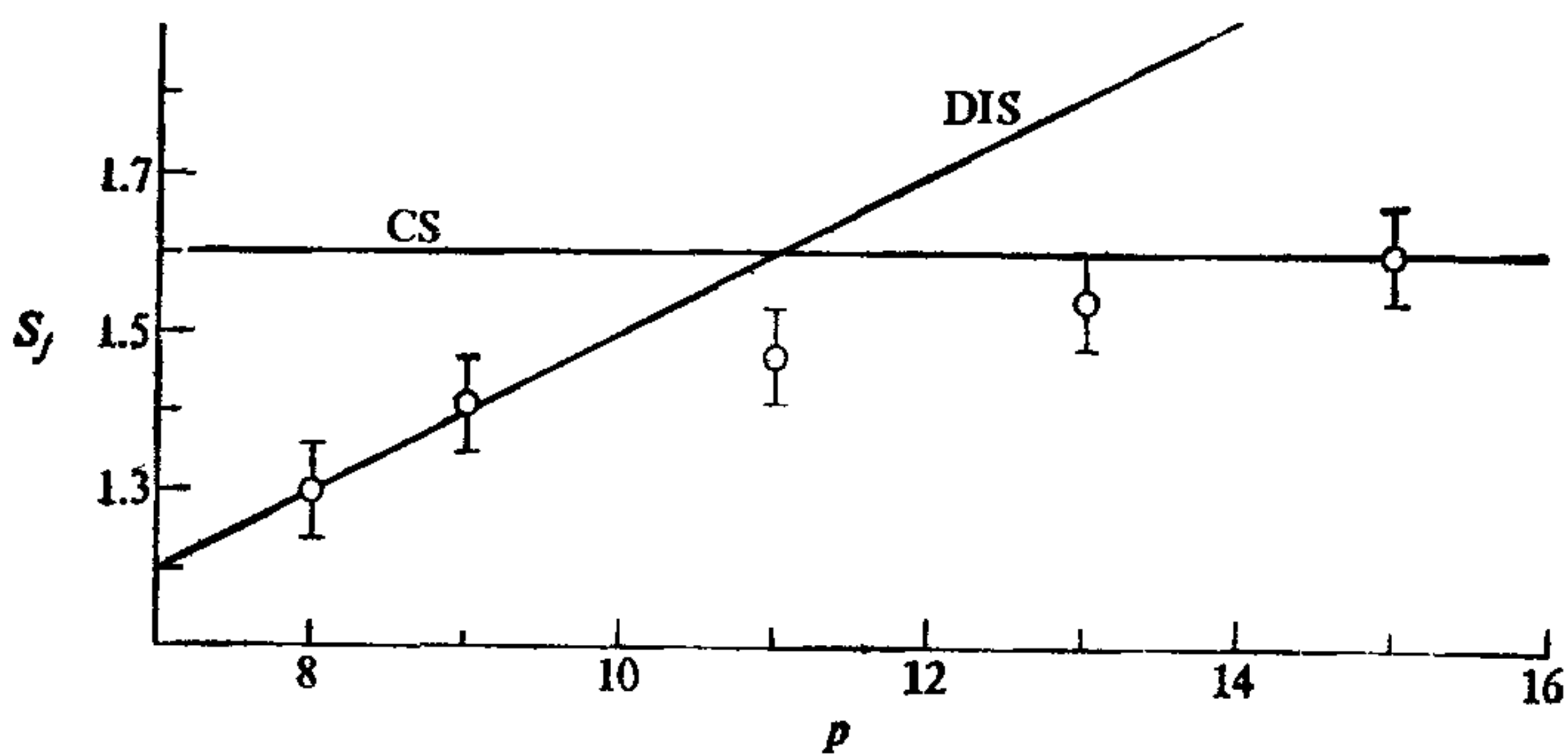


Рис. 12.19. Зависимость энтропии плавления S_f от параметра жесткости ρ для модельного KCl. С молекулярной динамикой (\circ) сравниваются теория дислокаций (DIS) и теория соответственных состояний (CS). ([17], с любезного разрешения J. Phys. C: Solid-State Phys., © Institute of Physics.)

генерации дислокаций обращается в нуль. В работе [17] в рамках упрощенной теории дислокаций [229] получено, что

$$T_{fa} \sim \rho/(\rho + 5) \text{ и } S_f \sim (\rho + 5). \quad (12.93)$$

Приведенные выше предсказания совершенно разные и поэтому зависимость характеристик плавления от параметра жесткости обеспечивает хороший способ провести различие между этими теориями. На рис. 12.18 показано изменение $T_f a$ в зависимости от ρ для трех теорий и вычислительного эксперимента. Видно, что результаты моделирования находятся в хорошем согласии с критерием Линдемана при малых ρ и теорией дислокаций при больших ρ . Предсказания теории соответственных состояний не согласуются с вычислительными экспериментами. Что касается энтропии плавления, представленной на рис. 12.19, то результаты моделирования согласуются с теорией дислокаций при малых ρ и теорией соответственных состояний при больших ρ . Ясно, что ни одна из указанных теорий удовлетворительно не объясняет всех результатов.

ПРИЛОЖЕНИЕ

Преобразование Фурье, ряд Фурье и дискретное преобразование Фурье

Преобразование Фурье и ряд Фурье широко применяются при построении и анализе моделей частиц. Сочетание в численных моделях непрерывных и дискретных величин с неизбежностью влечет как использование обобщенных функций (например, функции выборки Π (см. разд. П.4.5)), так и совместное использование преобразований и рядов. Представленные ниже краткие сведения о преобразовании Фурье (FT), ряде Фурье (FS) и дискретном преобразовании Фурье (FFT) преследуют двойную цель: (1) показать, как связаны между собой все эти преобразования и (2) ввести принятые в этой книге обозначения и нормировки. Тем читателям, которых интересует более подробное рассмотрение, следует обратиться к книгам [46] или [51].

Далее везде величины с крышкой вверху обозначают результат преобразования, символ « \supset » обозначает любое из двух (прямое или обратное) преобразований (для FT, FS и FFT). Так, например, $A \supset \hat{A}$ расшифровывается как «преобразование Фурье функции A равняется \hat{A} », и т. д.

П.1. ОПРЕДЕЛЕНИЯ ПРЕОБРАЗОВАНИЙ

П.1.1. Преобразование Фурье (FT)

Пусть A — непрерывная функция x и пусть \hat{A} — непрерывная функция k ; тогда пара преобразований FT определяется соотношениями ¹⁾

$$A(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{A}(k) e^{ikx}, \quad (\text{П.1})$$

$$\hat{A}(k) = \int_{-\infty}^{\infty} dx A(x) e^{-ikx}. \quad (\text{П.2})$$

¹⁾ Принято называть (П.2) преобразованием Фурье функции A , а (П.1) обратным преобразованием. — Прим перев.

Если x отвечает координате, то k можно отождествить с волновым числом; k связано с длиной волны гармоника λ соотношением

$$k = \frac{2\pi}{\lambda}. \quad (\text{П.3})$$

П.1.2. Ряд Фурье (FS)

Если B является периодической функцией непрерывной переменной x с длиной периода L , то пара интегральных преобразований заменяется разложением в ряд Фурье (FS (i)), определяемым соотношениями

$$\hat{B}(k) = \int_L dx B(x) e^{-ikx}, \quad (\text{П.4})$$

$$B(x) = \frac{1}{L} \sum_{l=-\infty}^{\infty} \hat{B}(k) e^{ikx}. \quad (\text{П.5})$$

Волновое число k принимает теперь только такие значения, чтобы на длине периода L укладывалось целое число длин волн:

$$k = \frac{2\pi l}{L}, \quad l - \text{целое}. \quad (\text{П.6})$$

Подобным же образом, если непрерывную переменную x заменить набором дискретных значений $\{x_p = pH; p \text{ целое}\}$, то периодическим станет k -пространство с периодом $k_g = 2\pi/H$, при этом соотношения для ряда Фурье (FS (ii)) имеют вид

$$C(x_p) = \int_{k_g} \frac{dk}{2\pi} \hat{C}(k) e^{ikx_p}, \quad (\text{П.7})$$

$$\hat{C}(k) = H \sum_{p=-\infty}^{\infty} C(x_p) e^{-ikx_p}. \quad (\text{П.8})$$

П.1.3. Дискретное преобразование Фурье (FFT)

Если периодическая функция D задана в дискретной последовательности точек — узлов сетки $x_p = pH$, и длина периода этой дискретной функции равна L , тогда можно определить дискретное преобразование (FFT), отображающее D на интервал $k_g = 2\pi/H$. Если N — число узлов сетки на длине одного периода, то

$$D(x_p) = \frac{1}{L} \sum_{l=0}^{N-1} \hat{D}(k) e^{ikx_p}, \quad (\text{П.9})$$

$$\hat{D}(k) = H \sum_{p=0}^{N-1} D(x_p) e^{-ikx_p}. \quad (\text{П.10})$$

Таблица П.1. Прямые и обратные преобразования

Преобразование	x -пространство	k -пространство	Пары преобразований
FT	x непрерывное	k непрерывное	$A(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{A}(k) e^{ikx} \quad \hat{A}(k) = \int_{-\infty}^{\infty} dx A(x) e^{-ikx}$
FS (i)	x непрерывное. Периодичность с периодом L	k дискретное, шаг $= k_0 = 2\pi/L$	$B(x) = \frac{1}{L} \sum_{l=-\infty}^{\infty} \hat{B}(k) e^{ikx} \quad \hat{B}(k) = \int_L dx B(x) e^{-ikx}$
FS (ii)	x дискретное, шаг H	k непрерывное. Периодичность с периодом $kg =$ $= 2\pi/H$	$C(x_p) = \int_{kg} \frac{dk}{2\pi} \hat{C}(k) e^{ikx_p} \quad \hat{C}(k) = H \sum_{p=-\infty}^{\infty} C(x_p) e^{-ikx_p}$
FFT	x дискретное Период L , шаг H	k дискретное Период kg , шаг $= k_0 = 2\pi/L$	$D(x_p) = \frac{1}{L} \sum_{l=0}^{N-1} \hat{D}(k) e^{ikx_p} \quad \hat{D}(k) = H \sum_{p=0}^{N-1} D(x_p) e^{-ikx_p}$

Примечание для дискретных x и x принимает значения $x_p = pH$, p целое, а для дискретных k k принимает значения $k = lk_0$, l целое.

Для вычисления дискретного преобразования Фурье существует эффективный алгоритм быстрого преобразования Фурье [51], который удобно обозначается той же аббревиатурой¹⁾.

П.2. СВОЙСТВА СИММЕТРИИ

Все четыре пары преобразований обладают одинаковыми свойствами симметрии. Например, если $f \supset \hat{f}$ и f — вещественная и четная функция, то \hat{f} — также вещественная и четная функция. В более общем виде эти свойства симметрии функции f можно схематически изобразить диаграммой.

$$\begin{array}{r}
 f(x) = E_r(x) + iE_i(x) + O_r(x) + iO_i(x) \\
 \uparrow \qquad \uparrow \qquad \uparrow \qquad \swarrow \searrow \\
 \hat{f}(k) = \hat{E}_r(k) + i\hat{E}_i(k) + \hat{O}_r(k) + i\hat{O}_i(k)
 \end{array}
 \tag{П.11}$$

Функции E и O изображают четную и нечетную части f , а индексы r и i относятся к вещественной и мнимой частям соответственно. Стрелками в (П.11) указано соответствие между частями f и \hat{f} . Эти соответствия перечислены в табл. П.2.

Таблица П.2. Свойства симметрии функции и ее преобразования

$f(x) \supset \hat{f}(k)$ или $\hat{f}(k) \supset f(x)$	
Вещественная + четная	Вещественная + четная
Вещественная + нечетная	Мнимая + нечетная
Мнимая + четная	Мнимая + четная
Мнимая + нечетная	Вещественная + нечетная
Вещественная	Эрмитова
Мнимая	Антиэрмитова
Четная	Четная
Нечетная	Нечетная

П.3. ТЕОРЕМЫ

Теорема о свертке занимает центральное место в применении спектральных методов к моделям частиц. Свертка $h(x)$ двух функций $f(x)$ и $g(x)$ определяется как

$$h(x) = \int_{-\infty}^{\infty} f(x') g(x - x') dx'.
 \tag{П.12}$$

¹⁾ FFT — Fast Fourier Transform — быстрое преобразование Фурье, FFT — Finite Fourier Transform — дискретное преобразование Фурье (англ). — Прим перев

Существует удобная сокращенная запись для свертки

$$h(x) = f * g, \quad (\text{П.13})$$

где звездочка означает вычисление интеграла, приведенного в правой части (П. 12), или соответствующей суммы для дискретных величин, смотря по тому, что требуется (см. табл. П. 3).

Таблица П.3. Свертки и их преобразования

Преобразование	x-пространство	k-пространство
FT	$\int_{-\infty}^{\infty} dx' f(x') g(x-x')$	$\hat{f}(k) \hat{g}(k)$
	$f(x) g(x)$	$\int_{-\infty}^{\infty} \frac{dk'}{2\pi} \hat{f}(k') \hat{g}(k-k')$
FS (i)	$\int_L dx' f(x') g(x-x')$	$\hat{f}(k) \hat{g}(k)$
	$f(x) g(x)$	$\frac{1}{L} \sum_{k'=-\infty}^{\infty} \hat{f}(k') \hat{g}(k-k')$
FS (ii)	$H \sum_{p'=-\infty}^{\infty} f(x_{p'}) g(x_p - x_{p'})$	$\hat{f}(k) \hat{g}(k)$
	$f(x_p) g(x_p)$	$\int_{k_g} \frac{dk'}{2\pi} \hat{f}(k') \hat{g}(k-k')$
FFT	$H \sum_{p'=0}^{N-1} f(x_{p'}) g(x_p - x_{p'})$	$\hat{f}(k) \hat{g}(k)$
	$f(x_p) g(x_p)$	$\frac{1}{L} \sum_{k'=0}^{N-1} \hat{f}(k') \hat{g}(k-k')$

Оператор свертки «*» является коммутативным, ассоциативным и дистрибутивным:

$$f * g = g * f, \quad (\text{П.14})$$

$$f * (g * h) = (f * g) * h, \quad (\text{П.15})$$

$$f * (g + h) = f * g + f * h. \quad (\text{П.16})$$

Теорема о свертке¹⁾ утверждает, что если $f \supset \hat{f}$ и $g \supset \hat{g}$, то $f * g \supset \hat{f}\hat{g}$ и $\hat{f} * \hat{g} \supset fg$. Свертки и отвечающие им произведения фурье-образов для преобразований FT, FS и FFT приведены в табл. П.3.

В табл. П.4 перечислены некоторые полезные теоремы, которые (за исключением теоремы о дифференцировании) с равным успехом годятся для FT, FS или FFT, когда соответствующее преобразование используется для связи величин в x - и k -пространствах.

Таблица П.4. Теоремы

Теорема о свертке	Если $f \supset \hat{f}$ и $g \supset \hat{g}$, то $f * g \supset \hat{f}\hat{g}$ и $\hat{f} * \hat{g} \supset fg$
Теорема подобия	Если $f(x) \supset \hat{f}(k)$, то $f\left(\frac{x}{a}\right) \supset a \hat{f}(ka)$
Умножение на константу	Если $f \supset \hat{f}$, то $bf \supset b\hat{f}$
Аддитивность	Если $f \supset \hat{f}$ и $g \supset \hat{g}$, то $f + g \supset \hat{f} + \hat{g}$
Теорема о сдвиге	Если $f(x) \supset \hat{f}(k)$, то $f(x + a) \supset \hat{f}(k) e^{ika}$
Теорема Парсеваля	Если $f \supset \hat{f}$ и $g \supset \hat{g}$, то $\int_{-\infty}^{\infty} f(x) g^*(x) dx =$ $= \int_{-\infty}^{\infty} \hat{f}(k) \hat{g}^*(k) \frac{dk}{2\pi}$ (аналогичные результаты справедливы для FS и FFT — см. табл. П.3)
Теорема о дифференцировании	Если $f \supset \hat{f}$, то $\frac{df}{dx} \supset ik\hat{f}$
Теорема взаимности	Если $f \supset \hat{f}(k) = g(k)$, то $g(x) \supset \hat{g}(k) = 2\pi \hat{f}(-k)$

Еще одна важная теорема, касающаяся преобразований, а именно теорема о выборке, обсуждается в гл. 5 в связи с ее использованием для анализа моделей частиц.

¹⁾ Называемая также теоремой Бореля о свертке. — Прим. перев.

П.4. СПЕЦИАЛЬНЫЕ ФУНКЦИИ

На рис. П.1 схематически изображены некоторые конкретные функции и их преобразования Фурье:

П.4.1. Функция прямоугольного импульса $\Pi(x)$:

$$\Pi(x) = \begin{cases} 0 & |x| > \frac{1}{2} \\ \frac{1}{2} & |x| = \frac{1}{2} \\ 1 & |x| < \frac{1}{2} \end{cases} \quad (\text{П.17})$$

$$\Pi(x) \supset \text{sinc}\left(\frac{k}{2\pi}\right) = \frac{\sin k/2}{k/2}. \quad (\text{П.18})$$

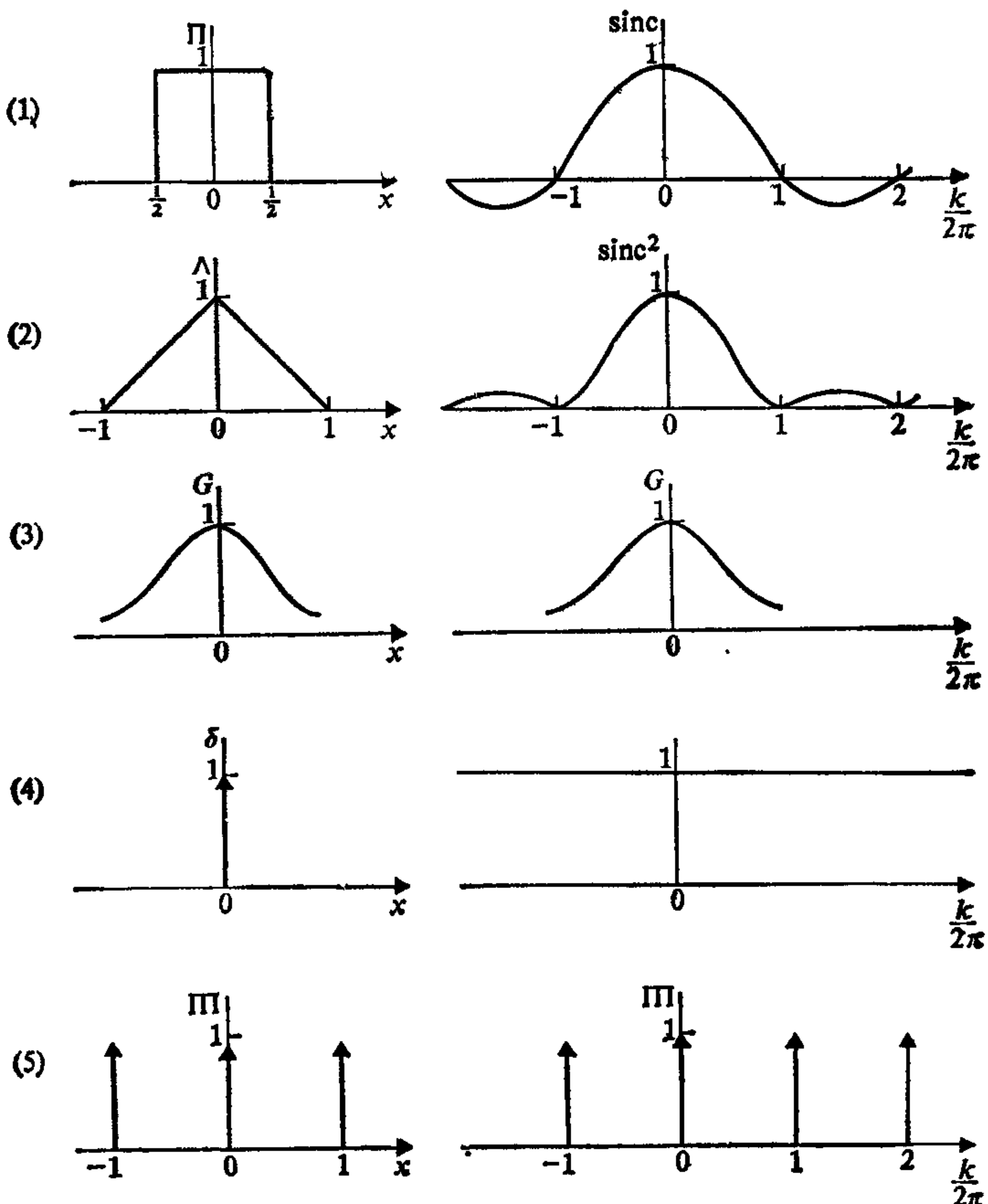


Рис. П.1. Специальные функции и их преобразования. Амплитуда импульсов у $\delta(x)$ и $\Pi(x)$ отражает площадь под импульсами,

П.4.2. Функция треугольного импульса $\Lambda(x)$:

$$\Lambda(x) = \begin{cases} 0, & |x| > 1 \\ 1 - |x|, & |x| < 1, \end{cases} \quad (\text{П.19})$$

$$\Lambda(x) \doteq \Pi * \Pi \supset \text{sinc}^2\left(\frac{k}{2\pi}\right). \quad (\text{П.20})$$

П.4.3. Гауссоида $G(x)$:

$$G(x) = e^{-\pi x^2}, \quad (\text{П.21})$$

$$G(x) \supset \hat{G}(k) = G\left(\frac{k}{2\pi}\right). \quad (\text{П.22})$$

П.4.4. Дельта-функция Дирака $\delta(x)$:

$$\delta(x) = 0, \quad x \neq 0, \quad (\text{П.23})$$

$$\int_{-\varepsilon}^{\varepsilon} \delta(x) dx = 1, \quad \varepsilon > 0, \quad (\text{П.24})$$

$$\delta(x) = \int_{-\infty}^{\infty} \frac{dy}{2\pi} e^{\pm ixy}, \quad (\text{П.25})$$

$$\delta(x) \supset 1. \quad (\text{П.26})$$

Дельта-функцию можно рассматривать как предел последовательности все более узких и высоких гауссоид, причем площадь под каждой гауссоидой этой последовательности равна единице:

$$\delta(x) = \lim_{a \rightarrow 0} \frac{1}{(2\pi a^2)^{1/2}} \exp\left(-\frac{x^2}{2a^2}\right). \quad (\text{П.27})$$

В дискретных преобразованиях в роли δ -функции выступает величина, называемая дельта-символом Кронекера:

$$\delta_{l,0} = \begin{cases} 1, & l = 0, \\ 0 & \text{в других случаях.} \end{cases} \quad (\text{П.28})$$

П.4.5. Функция выборки $\Pi(x)$

Функция выборки представляет собой бесконечную гребенку δ -функций, расположенных друг от друга на расстоянии, равном единице:

$$\Pi(x) = \sum_{n=-\infty}^{\infty} \delta(x - n), \quad (\text{П.29})$$

$$\Pi(x) \supset \hat{\Pi}(k) = \Pi\left(\frac{k}{2\pi}\right). \quad (\text{П.30})$$

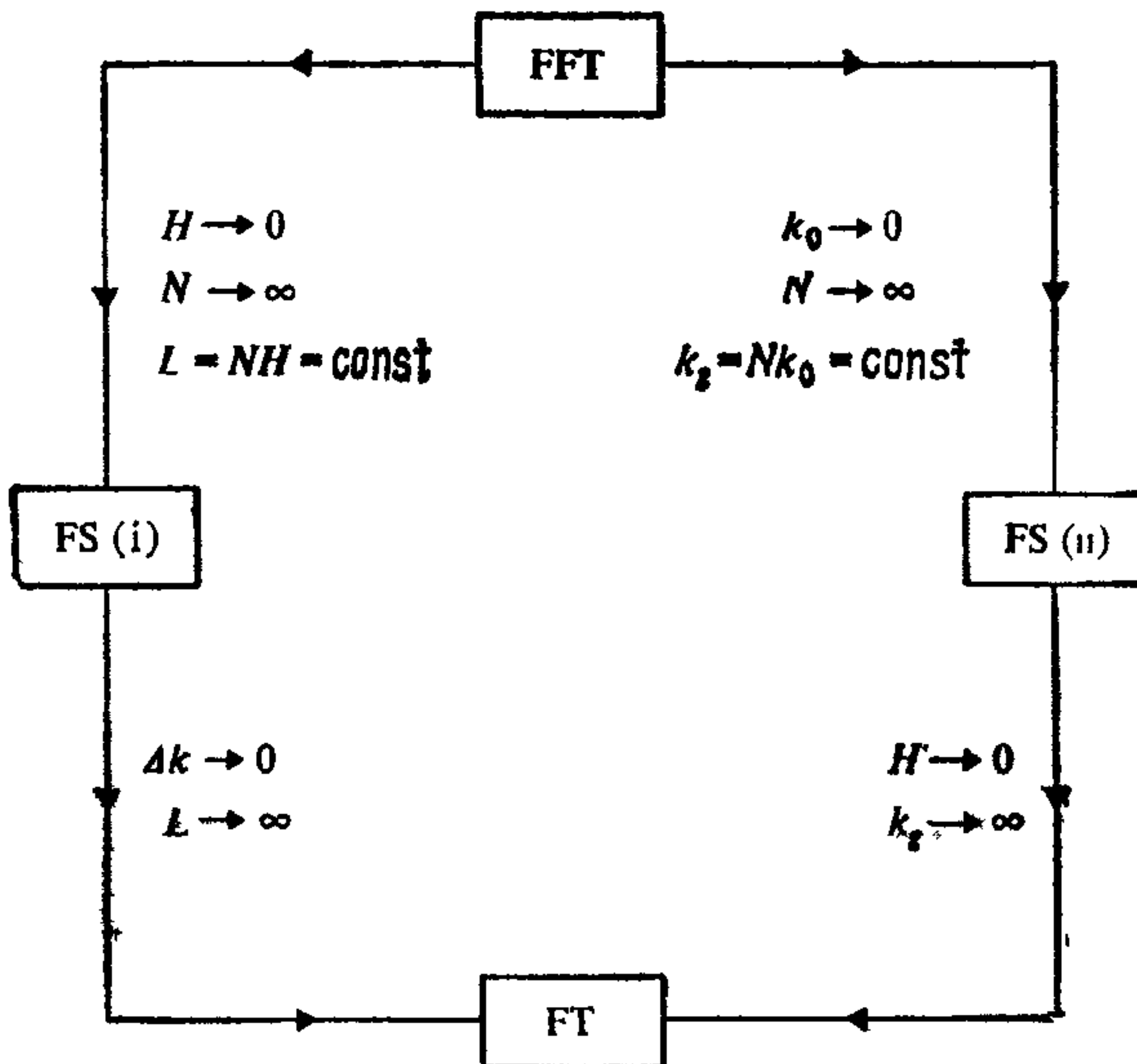


Рис. П 2. Преобразования FFT, FS и FT связаны предельными переходами, указанными около стрелок.

Аналогично можно определить функции выборки для FS и FFT. Например, для FS (i)

$$\Pi(x) \supset \Pi_s\left(\frac{k}{2\pi}\right), \quad (\text{П.31})$$

где для совместимости длина периода L кратна расстоянию между δ -функциями в x -пространстве и

$$\Pi_s\left(\frac{k}{2\pi}\right) = L \sum_{n=-\infty}^{\infty} \delta_{l, nL} \quad (\text{П.32})$$

$$\left(\text{сравните с } \Pi\left(\frac{k}{2\pi}\right) = L \sum_{n=-\infty}^{\infty} \delta(l - nL) \text{ для FT} \right).$$

П.5. СВЯЗЬ МЕЖДУ ПРЕОБРАЗОВАНИЯМИ

Преобразование Фурье можно получить из ряда Фурье, а ряд Фурье можно получить из дискретного преобразования Фурье путем предельного перехода, что отражено в схематическом виде на рис П. 2. Обратный переход получается с помощью функции выборки, как показано на рис. П. 3. Выражения FS и FFT для одного периода получаются дискретизацией бесконечных систем путем умножения на функцию прямоугольного импульса, ширина которой приводится к длинам периодов L или k_g , смотря что требуется.

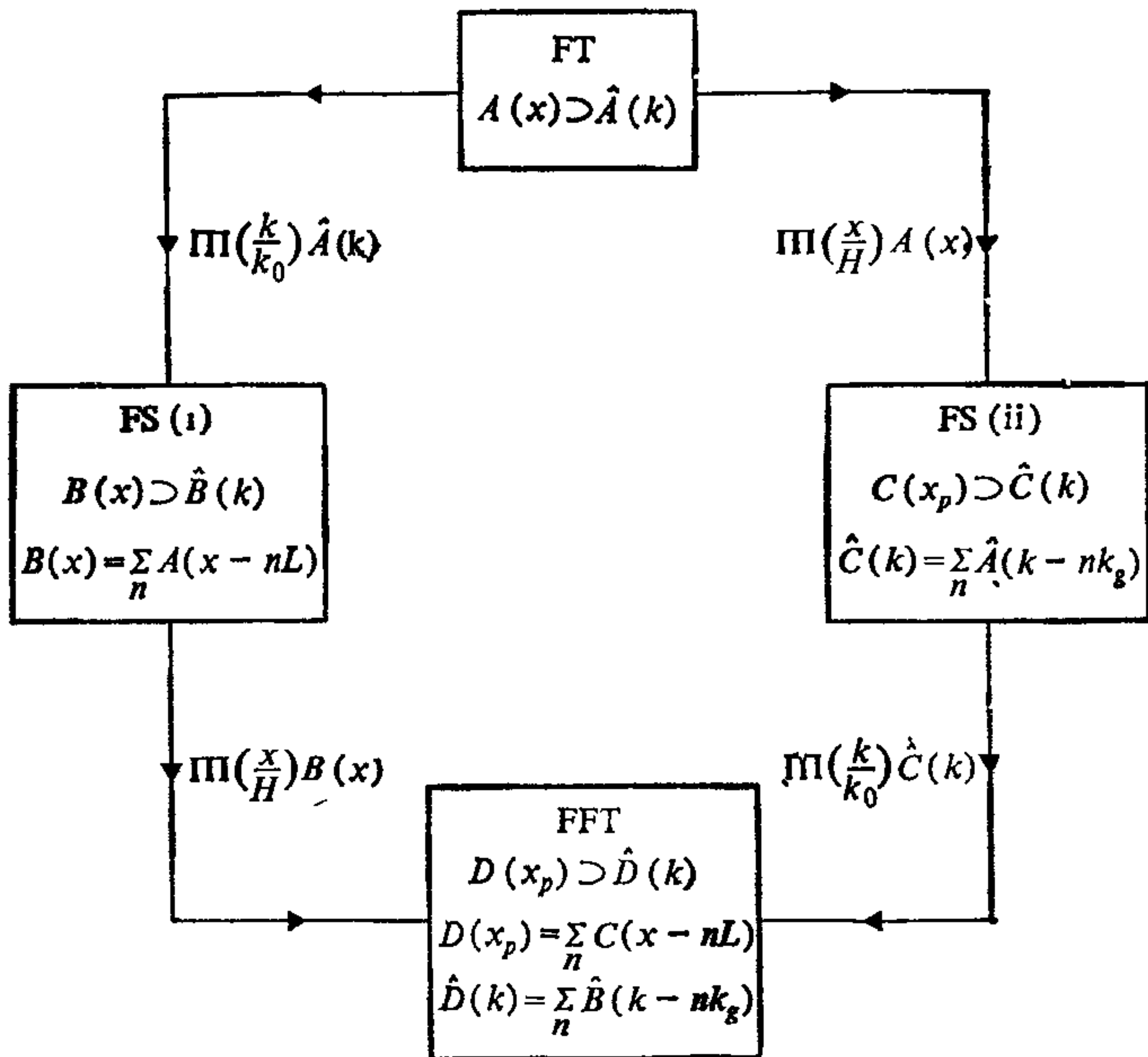


Рис. П.3. Преобразования FT, FS и FFT связаны посредством дискретизации, указанной около стрелок. Изображенные в квадратах суммы по n называются «суммами по сеточным гармоникам»

П.6. МНОГОМЕРНЫЕ ПРЕОБРАЗОВАНИЯ

Все одномерные преобразования и их свойства непосредственно обобщаются на случай двух, трех и более измерений. Трехмерные результаты получаются заменой

1. x на $\mathbf{x} = (x_1, x_2, x_3)$,
 2. k на $\mathbf{k} = (k_1, k_2, k_3)$,
 3. kx на $\mathbf{k} \cdot \mathbf{x}$,
 4. dx на $d\mathbf{x}$, $\frac{dk}{2\pi}$ на $\frac{d\mathbf{k}}{(2\pi)^3}$,
 5. L на $V_b = L_1 L_2 L_3$,
 6. H на $V_c = H_1 H_2 H_3$,
- и т. д. Индексы 1, 2 и 3 указывают направления декартовых осей.

Литература

1. Aarseth, S. J. (1963): Dynamical Evolution of Clusters of Galaxies — I, *Mon. Not. R. Astron. Soc.*, vol. 126, pp. 223—255.
2. Aarseth, S. J. (1966): Dynamical Evolution of Clusters of Galaxies — II, *Mon. Not. R. Astron. Soc.*, vol. 132, pp. 35—65.
3. Aarseth, S. J. (1969): Dynamical Evolution of Clusters of Galaxies — III, *Mon. Not. R. Astron. Soc.*, vol. 144, pp. 537—548.
4. Aarseth, S. J. (1971a): Direct Integration Methods of the N-body Problem, *Astrophys & Space Sci.*, vol. 14, pp. 118—132.
5. Aarseth, S. J. (1971b): Binary Evolution in Stellar Systems, *Astrophys. & Space Sci.*, vol. 13, pp. 324—334.
6. Aarseth, S. J., J. R. Gott, and E. L. Turner (1979): N-Body Simulations of Galaxy Clustering I: Initial Conditions and Galaxy Collapse Times, *Astrophys. J.*, vol. 228, pp. 664—683.
7. Aarseth S. J. and M. Lecar (1975): Computer Simulations of Stellar Systems, *Annu. Rev. Astron. & Astrophys.*, vol. 13, pp. 1—21.
8. Abe, H., J. Miyamoto, and R. Itatani (1975); Grid Effects on the Plasma Simulation by the Finite-Sized Particle, *J. Comput. Phys.*, vol. 19, pp. 134—149.
9. Abernathy, F. H., and R. E. Kronauer (1962): The Formation of Vortex Streets, *J. Fluid Mech.*, vol. 13, pp. 1—20.
10. Abramowitz, M., and I. A. Stegun (1964): «Handbook of Mathematical Functions», U. S. Department of Commerce, National Bureau of Standards, AMS 55, Washington D. C. [Имеется перевод: Справочник по специальным функциям. — М.: Наука, 1979.]
11. Adams, E. M., I. R. McDonald, and K. Singer (1977): Collective Dynamical Properties of Molten Salts: Molecular Dynamics Calculations on Sodium Chloride, *Proc. R. Soc. London, ser. A*, vol. 357, pp. 37—57.
12. Agajianian, A. H. (1975): A Bibliography on Semiconductor Device Modeling, *Solid-State Electron.*, vol. 18, pp. 917—929.
13. Ahmad, A., and L. Cohen (1973): A Numerical Integration Scheme for the N-Body Gravitational Problem, *J. Comput. Phys.*, vol. 12, pp. 389—402.
14. Alder, B. J., and T. E. Wainwright (1959): Studies in Molecular Dynamics. I. General Method, *J. Chem. Phys.*, vol. 31, pp. 459—466.
15. Alwin, V. C., D. H. Navon, and L. J. Turgeon (1977): Time-dependent Carrier Flow in a Transient Structure under Nonisothermal Conditions, *IEEE Trans. Electron. Devices*, vol. ED-24, pp. 1297—1304.
16. Amini, M. (1979): «Computer Simulation of Phase Changes», Ph. D Thesis, Reading University.
17. Amini, M., D. Fincham, and R. W. Hockney (1979): A Molecular Dynamics Study of the Melting of Alkali-Halide Crystals, *J. Phys. C: Solid State Phys.*, vol. 12, pp. 4707—4720.
18. Amini, M., and R. W. Hockney (1979): Computer Simulation of Melting and Glass Formation in a Potassium Chloride Microcrystal, *J. Non-Cryst. Solids*, vol. 31, pp. 447—452.

19. Andrew, R. (1972): Improved Formulation of Gummel's Algorithm for Solving the 2-Dimensional Current Flow Equations in Semiconductor Devices, *Electron Lett.*, vol. 8, pp. 536—538.
20. Balescu, R. (1960): Irreversible Processes in Ionised Gases, *Phys. Fluids*, vol. 3, pp. 52—63.
21. Bank, R. E. (1977): Marching Algorithms for Elliptic Boundary Value Problems, II: The Variable Coefficient Case, *SIAM J. Numer. Anal.*, vol. 14, pp. 950—970.
22. Bank, R. E., and D. J. Rose (1977): Marching Algorithms for Elliptic Boundary Value Problems, I: The Constant Coefficient Case, *SIAM J. Numer. Anal.*, vol. 14, pp. 792—829.
23. Barnes, C. W. (1972): Virtual Memory Utilisation in Large Plasma Simulation Codes, SUIPR report No. 497, Stanford University.
24. Barnes, J. J. and R. J. Lomax (1974): Two-Dimensional Finite Element Simulation of Semiconductor Devices, *Electron. Lett.*, vol. 10, pp. 341—343.
25. Barnes, J. J., R. J. Lomax, and G. I. Haddad (1976): Finite-Element Simulation of GaAs MESFET's with Lateral Doping Profiles and Submicron Gates, *IEEE Trans. Electron Devices*, vol. ED-23, pp. 1042—1048.
26. Baus, M. (1978): Collective Modes and Dynamic Structure Factor of a Two-Dimensional Electron Fluid, *J. Stat. Phys.*, vol. 19, pp. 163—176.
27. Beeman, D. (1976): Some Multistep Methods for use in Molecular Dynamics Calculations, *J. Comput. Phys.*, vol. 20, pp. 130—139.
28. Berk, H. L., and K. V. Roberts (1970): The Water-Bag Model, *Methods Comput. Phys.*, vol. 9, pp. 88—134. [Имеется перевод: Бэрк Г., Робертс К. Модель «водяного мешка». — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга. — М.: Мир, 1974, с. 96.]
29. Berman, R. H., D. R. K. Brownrigg, and R. W. Hockney (1978): Numerical Models of Galaxies I. The Variability of Spiral Structure, *Mon. Not R. Astron. Soc.*, vol. 185, pp. 861—875.
30. Berman, R. H., and J. W.-K. Mark (1979): Are Massive Haloes Necessary to Prevent Rapid, Global Bar Formation? *Astron. & Astrophys.*, vol. 77, pp. 31—36.
31. Bernstein, I. B., and F. Engelmann (1966): Quasi-Linear Theory of Plasma Waves, *Phys. Fluids*, vol. 9, pp. 937—952.
32. Birdsall, C. K., and W. B. Bridges (1961): Space-Charge Instabilities in Electron Diodes and Plasma Converters, *J. Appl. Phys.*, vol. 32, pp. 2611—2618.
33. Birdsall, C. K., and W. B. Bridges (1966): «Electron Dynamics of Diode Regions», Academic Press, New York.
34. Birdsall, C. K., and J. M. Dawson (1970): Plasma Physics, in «Computers and their Role in the Physical Sciences» (ed. S. Fernbach and A. Taub), University of California, Gordon and Breach, Science Publishers, Inc. New York, pp. 247—310.
35. Birdsall, C. K., and D. Fuss (1969): Clouds-in-Clouds, Clouds-in-Cells Physics for Many-Body Plasma Simulation, *J. Comput. Phys.*, vol. 3, pp. 494—511.
36. Birdsall, C. K., and A. B. Langdon (1985); «Plasma Physics via Computer Simulation», McGraw-Hill, New York.
37. Birdsall, C. K., and N. Maron (1980); Plasma Self-Heating and Saturation due to Numerical Instability, *J. Comput. Phys.*, vol. 36, pp. 1—19.
38. Birdsall, C. K., A. B. Langdon, and H. Okuda (1970): Finite-Size Particle Physics Applied to Plasma Simulation, *Methods Comput. Phys.*, vol. 9, pp. 241—258. [Имеется перевод: Бэрдсол Ч., Лендон А., Окуда Х. Физика системы частиц конечных размеров и ее применение к моделированию плазмы. — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга. — М.: Мир, 1974, с. 242.]

39. Blakemore, J. S. (1970): «Solid State Physics», W. B. Saunders Co., London. [Имеется перевод: Блейкмор Дж. Физика твердого состояния. — М.: Металлургия, 1972.]
40. Blum, E. K. (1962): A Modification of the Runge-Kutta Fourth Order Method, *Math. Comput.*, vol. 16, pp. 176—187.
41. Boris, J. (1970): Relativistic Plasma Simulation — Optimization of a Hybrid Code, *Proc. 4th Conf. on Numer. Simulation of Plasmas Office of Naval Research, Arlington, Va.*, pp. 3—67.
42. Boris, J., and K. V. Roberts (1969): The Optimization of Particle Calculations in 2 and 3 Dimensions, *J. Comput. Phys.*, vol. 4, pp. 552—571
43. Born, M., and K. Huang (1954): «Dynamical Theory of Crystal Lattices», Oxford University Press, London. [Имеется перевод: Борн М., Хуан Кунь. Динамическая теория кристаллических решеток. — М.: ИЛ, 1959.]
44. Born, M., and J. E. Mayer (1932): Zur Gittertheorie der Ionen Kristalle, *Z. Phys.*, vol. 75, pp. 1—18.
45. Bott, I. B., and W. Fawcett (1968): The Gunn Effect in Gallium Arsenide, *Adv. Microwaves*, vol. 3, pp. 223—300.
46. Bracewell, R. (1965): «The Fourier Transform and its Application», McGraw-Hill, New York.
47. Brackbill, J. U. (1976): Numerical Magnetohydrodynamics for High-Beta Plasmas, *Methods Comput. Phys.*, vol. 16, pp. 1—41. [Имеется перевод: Брекбилл Дж. Численная магнитная гидродинамика для плазмы с большим бета. — В кн.: Управляемый термоядерный синтез, под ред Дж. Киллина. — М.: Мир, 1980, с. 11.]
48. Brewitt-Taylor, C. R. (1979): Velocity Overshoot Effects in 2-dimensional Device Simulation, in «Numerical Analysis of Semiconductor Devices» (eds. B. T. Browne and J. J. H. Miller), Boole Press, Dublin, pp. 191—193.
49. Brewitt-Taylor, C. R., P. H. Robson, and J. E. Sitch (1978): New Estimate of the Minimum Noise Figure of a MESFET, *Electron. Lett.*, vol. 14, pp. 818—820.
50. Brewitt-Taylor, C. R., P. N. Robson, and J. E. Sitch (1980): The Noise Figure in MESFETs, *IEE Proc.*, part I, vol. 127, pp. 1—8.
51. Brigham, E. O. (1974): «The Fast Fourier Transform», Prentice-Hall Inc., Englewood Cliffs, N. J.
52. British Standards Institution (1975): BS4148: «The Abbreviation of Titles of Periodicals. Pt 2, Word-Abbreviation List», British Standards Institution, London.
53. Brown, T. R., and C. C. Grimes (1972): Observation of Cyclotron Resonance in Surface-Bound Electrons on Liquid Helium, *Phys. Rev. Lett.*, vol. 29, pp. 1233—1236.
54. Brownrigg, D. K. K. (1975): «Computer Modelling of Spiral Structure in Galaxies», Ph. D. Thesis, University of Reading.
55. Brush S. G., H. L. Sahlin, and E. Teller (1966): Monte-Carlo Study of a One-Component Plasma. I, *J. Chem. Phys.*, vol. 45, pp. 2102—2118.
56. Buneman, O. (1959): Dissipation of Currents in Ionised Media, *Phys. Rev.*, vol. 115, pp. 503—517.
57. Buneman, O. (1967): Time-Reversible Difference Procedures, *J. Comput. Phys.*, vol. 1, pp. 517—537.
58. Buneman, O. (1969): A Compact Non-Iterative Poisson-Solver, SUIPR report No. 294, Stanford University.
59. Buneman, O. (1971): Analytic Inversion of the Five-Point Poisson Operator, *J. Comput. Phys.*, vol. 8, pp. 500—505.
60. Buneman, O. (1973): Analytic Inversion of Nine-Point Poisson Operator, *J. Comput. Phys.*, vol. 11, pp. 447—448.
61. Buneman, O. (1976): The Advance from 2D Electrostatic to 3D Electromagnetic Particle Simulation, *Comput. Phys. Commun.*, vol. 12, pp. 21—31.

62. Buneman, O. (1979): Vectorization and Parallelism in Stanford's 3D, EM Particle Code, MFECC Buffer, vol. 3, No. 7, pp. 7—12. Magnetic Fusion Energy Computer Center Monthly Bulletin, Livermore, Cal.
63. Buneman, O. (1980): Tetrahedral Finite Elements for Interpolation, SIAM J. Sci. & Stat. Comput., vol. 1, pp. 223—248.
64. Buneman, O. C. W. Barnes, J. C. Green, and D. E. Neilsen (1980): Principles and Capabilities of 3-D, E-M Particle Simulations, J. Comput. Phys., vol. 38, pp. 1—44.
65. Buneman, O., and D. A. Dunn (1966): Computer Experiments in Plasma Physics, Sci. J. (U. K.), vol. 2, No. 7, pp. 34—40.
66. Burger, P. (1964): Nonexistence of DC States in Low-Pressure Thermionic Converters, J. Appl. Phys., vol. 35, pp. 3048—3049.
67. Burger, P. (1965): Theory of Large-Amplitude Oscillation in the One-Dimensional Low-Pressure Cesium Thermionic Converter, J. Appl. Phys., vol. 36, pp. 1938—1943.
68. Burger, P. (1967) Elastic Collisions in Simulating One-Dimensional Plasma Diodes on the Computer, Phys. Fluids, vol. 10, pp. 658—666.
69. Burger, P., D. A. Dunn and A. S. Halstead (1966): Computer Experiments on the Randomization of Electrons in a Collisionless Plasma, Phys. Fluids, vol. 8, pp. 2263—2272.
70. Butcher, J. C. (1964): Implicit Runge-Kutta Processes, Math. Comput., vol. 18, pp. 50—64.
71. Buzbee, B. L., F. W. Dorr, J. A. George, and G. H. Golub (1971): The Direct Solution of the Discrete Poisson Equation on Irregular Regions, SIAM J. Numer. Anal., vol. 8, pp. 722—736.
72. Buzbee, B. L., G. H. Golub, and C. W. Nielson (1970): On direct Methods for Solving Poisson's Equation, SIAM J. Numer. Anal., vol. 7, pp. 627—656
73. Byers, J. A., B. I. Cohen, W. C. Condit, and J. D. Hanson (1978): Hybrid Simulations of Quasineutral Phenomena in Magnetised Plasma, J. Comput. Phys., vol. 27, pp. 363—396.
74. Chandrasekhar, S. (1942): «Principles of Stellar Dynamics», Dover Publications Inc., New York. [Имеется перевод: Чандрасекар С. Принципы звездной динамики. — М.: ИЛ, 1948.]
75. Chen, L., A. B. Langdon, and C. K. Birdsall (1974): Reduction of Grid Effects in Simulation Plasmas, J. Comput. Phys., vol. 14, pp. 200—222
76. Chen, L., and H. Okuda (1975): Theory of Plasma Simulation Using Multipole-Expansion Scheme, J. Comput. Phys., vol. 19, pp. 339—352.
77. Christiansen, J. P. (1973a): «The Non-Linear Dynamics of Vortex Flows by Numerical Methods», Ph. D. Thesis, Warwick University.
78. Christiansen, J. P. (1973b): Numerical Simulation of Hydrodynamics by the Method of Point Vortices, J. Comput. Phys., vol. 13, pp. 363—379.
79. Christiansen, J. P., and R. W. Hockney (1971): DELSQPHI, a 2D Poisson Solver Program, Comput. Phys. Commun, vol. 2, pp. 139—155.
80. Christiansen, J. P., and K. V. Roberts (1974): OLYMPUS — A Standard Control and Utility Package for Initial Value FORTRAN Programs, Comput Phys Commun., vol. 7, pp 245—270.
81. Christiansen, J. P., and N. J. Zabusky (1973): Instability, Coalescence and Fission of Finite-Area Vortex Structures, J. Fluid Mech., vol. 61, pp. 219—243
82. Chryssafis, A., and W. Love (1979): A Computer-Aided Analysis of One-Dimensional Thermal Transients in n-p-n Power Transistors, Solid-State Electron, vol. 22, pp. 249—256.
83. Ciccotti, G., J. Jacucci, and I. R. McDonald (1976): Transport Properties of Molten Alkali Halides, Phys. Rev., ser. A, vol. 13, pp. 426—436.
84. Clemmow, P. C., and J. P. Dougherty (1969): «Electrodynamics of Particles and Plasmas», Addison-Wesley, Reading, Mass.

85. Cole, M. W., and M. H. Cohen (1969): Image-Potential-Induced Surface Bands in Insulators, *Phys. Rev. Lett.*, vol. 23, pp. 1238—1241.
86. Concus, P., and G. H. Golub (1973): Use of Fast Direct Methods for the Efficient Numerical Solution of Non-Separable Elliptic Equations, *SIAM J. Numer. Anal.*, vol. 10, pp. 1103—1120.
87. Concus, P., G. H. Golub, and D. P. O'Leary (1976): A Generalized Conjugate Gradient Method for the Numerical Solution of Elliptic Partial Differential Equations, in «Sparse Matrix Computations» (eds. J. R. Bunch and D. J. Rose), Academic Press, New York, pp. 309—332.
88. Cooley, J. W., and J. W. Tukey (1965): An Algorithm for the Machine Calculation of the Complex Fourier Series, *Math. Comput.*, vol. 19, pp. 297—301.
89. Copley, J. R. D., and A. Rahman (1976): Density Fluctuations in Molten Salts: Molecular Dynamics Study of Liquid RbBr, *Phys. Rev.*, ser. A, vol. 13, pp. 2276—2286.
90. Cotterill, R. M. J. (1975): Melting Point Depression in Very Thin Lennard-Jones Crystals. A Molecular Dynamics Study, *Philos. Mag.*, vol. 32, pp. 1283—1288.
91. Cotterill, R. M. J., W. Damgaard Kristensen, and E. J. Jensen (1974): Molecular Dynamics Studies of Melting. III. Spontaneous Dislocation Generation and the Dynamics of Melting, *Philos. Mag.*, vol. 30, pp. 245—263.
92. Cotterill, R. M. J., E. J. Jensen, W. Damgaard Kristensen, R. Paetsch, and P. O. Esbjorn (1975): A Unified Theory of Melting, Crystallization and Glass Formation, *J. Phys. (Fr.)*, Coll. C2, Suppl. 4, vol. 36, pp. 35—48.
93. Cotterill, R. M. J., and L. B. Pedersen (1972): A Molecular Dynamics Study of the Melting of a Two-Dimensional Crystal, *Solid State Commun.*, vol. 10, pp. 439—441.
94. Couchman, P. R. (1979): The Lindemann Hypothesis and the Size-Dependence of Melting Temperature II, *Philos. Mag.*, ser. A., vol. 40, pp. 637—643.
95. Couchman, P. R., and W. A. Jesser (1977): Thermodynamic Theory of Size Dependence of Melting Temperature in Metals, *Nature*, vol. 269, pp. 481—483.
96. Couchman, P. R., and C. L. Ryan (1978): The Lindemann Hypothesis and the Size-Dependence of Melting Temperature, *Philos. Mag.*, vol. 37, pp. 369—373.
97. Crank, J., and P. Nicolson (1947): A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat Conduction Type, *Proc. Cambridge Philos. Soc.*, vol. 43, pp. 50—67.
98. Curtis, A. R., and J. K. Reid (1971): FORTRAN Subroutines for the Solution of Sparse Sets of Linear Equations, UKAEA report R6844, H. M. Stationery Office, London.
99. Damgaard Kristensen, W., E. J. Jensen, and R. M. J. Cotterill (1974): Molecular Dynamics Studies of Melting: II. Dislocation Density and Thermodynamic Functions, *Philos. Mag.*, vol. 30, pp. 229—243.
100. Daniel, J. W., and R. E. Moore (1970): «Computation and Theory in Ordinary Differential Equations», W. H. Freeman & Co., San Francisco.
101. Davis, M., E. J. Groth, and P. J. E. Peebles (1977): Study of Galaxy Correlations: Evidence for the Gravitational Instability Picture in a Dense Universe, *Astrophys. J.*, vol. 212, pp. L107—L111.
102. Davis, M., and P. J. E. Peebles (1977): On the Integration of the BBGKY Equations for the Development of Strongly Nonlinear Clustering in an Expanding Universe, *Astrophys. J. Suppl.*, vol. 34, pp. 425—450.
103. Dawson, J. M. (1962): One-Dimensional Plasma Model, *Phys. Fluids*, vol. 5, pp. 445—459.
104. Dawson, J. M. (1970): The Electrostatic Sheet Model for a Plasma and its Modification to Finite-Size Particles, *Methods Comput. Phys.*, vol. 9, pp. 1—28, [Имеется перевод: Доусон Дж. Электростатическая модель плоских

- листов для плазмы и ее модификация для частиц конечного размера.— В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга. — М.: Мир, 1974, с. 11.]
105. Dawson, J. M., H. Okuda, and B. Rosen (1976): Collective Transport in Plasmas, *Methods Comput. Phys.*, vol. 16, pp. 281—325. [Имеется перевод: Доусон Дж., Окуда Х., Розен Б. Коллективный перенос в плазме.— В кн.: Управляемый термоядерный синтез, под ред. Дж. Киллина. — М.: Мир, 1980, с. 296]
 106. Dekker, A. J. (1958); «Solid State Physics», Macmillan & Co., Ltd, London.
 107. Detyna, E. (1979): Point Cyclic Reduction for Elliptic Boundary Value Problems I: The Constant Coefficient Case, *J. Comput. Phys.*, vol. 33, pp. 204—216.
 108. Dorr, F. W. (1970): The Direct Solution of the Discretised Poisson Equation on a Rectangle, *SIAM Rev.*, vol. 12, pp. 248—263
 109. Drangeid, K. E., and R. Sommerhalder (1970): Dynamic Performance of Schottky-Barrier Field-Effect Transistors, *IBM J. Res. & Dev.*, vol. 14, pp. 82—94.
 110. Drummond, W. E., and D. Pines (1964): Nonlinear Plasma Oscillations, *Ann. Phys. (U. S. A.)*, vol. 28, pp. 478—499.
 111. Duff, I. S., A. M. Erisman, and J. K. Reid (1976): On George's Nested Dissection Method, *SIAM J. Numer. Anal.*, vol. 13, pp. 686—695.
 112. Dungey, J. W. (1961): Interplanetary Magnetic Field and the Auroral Zones, *Phys. Rev. Lett.*, vol. 6, pp. 47—48.
 113. Eastwood, J. W. (1972): Consistency of Fields and Particle Motion in the Speiser Model of the Current Sheet, *Planet. & Space Sci.*, vol. 20, pp. 1555—1568.
 114. Eastwood, J. W. (1974): The Warm Current Sheet Model, and its Implications on the Temporal Behaviour of the Geomagnetic Tail, *Planet. Space Sci.*, vol. 22, pp. 1641—1668.
 115. Eastwood, J. W. (1975): Optimal Particle Mesh Algorithms, *J. Comput. Phys.*, vol. 18, pp. 1—20
 116. Eastwood, J. W. (1976a): Value for Money in Particle-Mesh Plasma Simulations, in «Computational Methods in Classical and Quantum Physics» (ed. M. B. Hooper), Advance Publications Ltd, London, pp. 196—205.
 117. Eastwood, J. W. (1976b): Optimal P³M Algorithms for Molecular Dynamics Simulations, in «Computational Methods in Classical and Quantum Physics» (ed. M. B. Hooper), Advance Publications Ltd, London, pp. 206—228
 118. Eastwood, J. W., and D. R. K. Brownrigg (1970): Remarks on the Solution of Poisson's Equation for Isolated systems, *J. Comput. Phys.*, vol. 32, pp. 24—38.
 119. Eastwood, J. W., and R. W. Hockney (1974): Shaping the Force Law in Two-Dimensional Particle-Mesh Models, *J. Comput. Phys.*, vol. 16, pp. 342—359.
 120. Eastwood, J. W., R. D. Hockney, and D. N. Lawrence (1980): P3M3DP—the Three-Dimensional Periodic Particle-Particle/Particle-Mesh Program, *Comput. Phys. Commun.*, vol. 19, pp. 215—261.
 121. Eastwood, J. W., and K. V. Roberts (1979): The Authentication of Computer Programs, *Computer. Phys. Commun.*, to be submitted.
 122. Efstathiou, G. (1979a): The Clustering of Galaxies and its Dependence upon Ω , *Mon. Not. R. Astron. Soc.*, vol. 187, pp. 117—127.
 123. Efstathiou, G. (1979b): «On the Rotation and Clustering of Galaxies», Ph. D. Thesis, University of Durham.
 124. Efstathiou, G., and J. W. Eastwood (1981); On the Clustering of Particles in an Expanding Universe, *Mon. Not. R. Astron. Soc.*, vol. 194, pp. 503—526.

125. Efstathiou, G., and B. J. T. Jones (1979): The Rotation of Galaxies: Numerical Investigations of the Tidal Torque Theory, *Mon. Not. R. Astron. Soc.*, vol. 186, pp. 133—144.
126. Ehrlich, L. W. (1979): A Marching Technique for Non-Separable Equations, *Math. Comput.*, vol. 33, pp. 881—890.
127. Eldridge, O. C., and M. Feix (1962): One-Dimensional Plasma Model at Thermodynamic Equilibrium, *Phys. Fluids*, vol. 5, pp. 1076—1080.
128. Ewald, P. P. (1921): Die Berechnung Optischer und Elektrostatische Gitterpotentiale, *Ann. Phys. (Ger.)*, vol. 64, pp. 253—287.
129. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. — М.: Физматгиз, 1960.
130. Fall, S. M. (1978): On the Evolution of Galaxy Clustering and Cosmological N-Body Simulations, *Mon. Not. R. Astron. Soc.*, vol. 185, pp. 165—177.
131. Fall, S. M. (1979): Galaxy Correlations and Cosmology, *Rev. Mod. Phys.*, vol. 51, pp. 21—42.
132. Faulkner, E. A. (1966): «Principles of Linear Circuits», Chapman & Hall Ltd., London.
133. Fawcett, W., A. D. Boardman, and S. Swain (1970): Monte Carlo Determination of Electron Transport Properties in Gallium Arsenide, *J. Phys. Chem. Solids*, vol. 31, pp. 1963—1990.
134. Flanders, P. M., D. J. Hunt, S. F. Reddaway, and D. Parkinson (1977): Efficient High Speed Computing with the Distributed Array Processor, in «High Speed Computer and Algorithm Design», Academic Press Ltd., London, pp. 113—128.
135. Forsythe, G. E., and W. R. Wasow (1960): «Finite-Difference Methods for Partial Differential Equations», John Wiley & Sons, Inc., New York. [Имеется перевод: Вазов В., Форсайт Дж. Разностные методы решения дифференциальных уравнений в частных производных. — М.: ИЛ, 1963.]
136. Freeman, K. C. (1970): On the Disks of Spiral and SO Galaxies, *Astrophys J.*, vol. 160, pp. 811—830.
137. Friedman, A. (1922): Uber die Krümmung des Raumes, *Z. Phys.*, vol. 10, pp. 377—386.
138. Fumi, F. G., and M. P. Tosi (1964): Ionic Sizes and Born Repulsion Parameters in the NaCl-Type Alkali Halides — I, *J. Phys. Chem. Solids*, vol. 25, pp. 31—43.
139. Fyfe, D. J. (1966): Economic Evaluation of Runge-Kutta Formulae, *Math. Comput.*, vol. 20, pp. 392—398.
140. Gann, R. C., S. Chakravarty, and G. V. Chester (1979): Monte-Carlo Simulation of the Two-Dimensional One Component Plasma; *Phys. Rev.*, ser. B. vol. 20, pp. 362—374.
141. Gaur, S. P., and D. H. Navon (1976): Two-Dimensional Carrier Flow in a Transistor Structure under Nonisothermal Conditions, *IEEE Trans. Electron Devices*, vol. ED-23, pp. 50—57.
142. George, J. A. (1973): Nested Dissection of a Regular Finite-Element Mesh, *SIAM J. Numer. Anal.*, vol. 10, pp. 345—363.
143. Gill, S. (1951): A Process for the Step-by-Step Integration of Differential Equations in an Automatic Digital Computing Machine, *Proc. Cambridge Philos. Soc.*, vol. 47, pp. 96—108.
144. Goel, S. P., and R. W. Hockney (1974): A Resource Letter CSSMD-I, Computer Simulation Studies by the Method of Molecular Dynamics, *Rev. Bras. Fis.*, vol. 4, pp. 121—157.
145. Goldstein, H. (1959): «Classical Mechanics», Addison-Wesley, Inc., Reading, Mass. [Имеется перевод: Голдстейн Г. Классическая механика. — М.: Наука, 1975.]
146. Gott, J. R. (1977): Recent Theories of Galaxy Formation, *Annu. Rev. Astron. & Astrophys.*, vol. 15, pp. 235—266.

147. Gott, J. R., J. E. Gunn, and B. M. Tinsley (1974): An Unbound Universe? *Astrophys. J.*, vol. 194, pp. 543—553.
148. Gott, J. R., E. L. Turner, and S. J. Aarseth (1979): N-Body Simulations of Galaxy Clustering III: The Covariance Function, *Astrophys. J.*, vol. 234, pp. 13—26.
149. Greenstein, J. L. (1959): Dying Stars, *Sci. Am.*, vol. 200, No. 1, 46—53.
150. Griffiths, R. J. M., I. D. Blenkinsop, and D. R. Wight (1979): Preparation and Properties of GaAs Layers for Novel F. E. T. Structures, *Electron Lett.*, vol. 15, pp. 629—630.
151. Grimes, C. C., and G. Adams (1979): Evidence for a Liquid-to-Crystal Phase Transition in a Classical Sheet of Electrons, *Phys. Rev. Lett.*, vol. 42, pp. 795—798.
152. Groth, E. J., and P. J. E. Peebles (1977): Statistical Analysis of Catalogs of Extra-Galactic Objects. VII Two- and Three-Point Correlation Functions for the High-Resolution Shane-Wirtanen Catalog of Galaxies, *Astrophys. J.*, vol. 217, pp. 385—405.
153. Groth, E. J., P. J. E. Peebles, M. Seldner, and R. M. Soneira (1977): The Clustering of Galaxies, *Sci. Am.*, vol. 237, No. 5, pp. 76—98.
154. Guernsey, R. L. (1962): Kinetic Equation for a Completely Ionized Gas, *Phys. Fluids*, vol. 5, pp. 322—328.
155. Gula, W. P., and C. K. Chu (1973): Effect of Krook Model Collisions on Two-Stream Instability, *Phys. Fluids*, vol. 16, pp. 1135—1141.
156. Gummel, H. K. (1964): A Self-Consistent Iterative Scheme for One Dimensional Steady-State Transistor Calculations, *IEEE Trans. Electron Devices*, vol. ED-11, pp. 455—465.
157. Gunn, J. B. (1963): Microwave Oscillation of Current in III-V Semiconductors, *Solid State Commun.*, vol. 1, pp. 88—91.
158. Hamilton, D. J., F. A. Lindholm, and A. H. Marshak (1971): «Principles and Applications of Semiconductor Device Modeling», Rinehart and Winston, New York.
159. Hamilton, J. M. (1976): The Numerical Solution of Ampere's Equation on a 2-Dimensional Mesh. Part I — Theoretical Discussion, Reading University Computer Science Report, RCS-49.
160. Hamilton, J. M. (1977): COIRCE: The Numerical Solution of Ampere's Equation on a 2-Dimensional Mesh. Part II — Further Theoretical Results, Timing and Convergence Properties, Reading University Computer Science Report, RCS50.
161. Hamilton, J. M. (1980): «Computer Simulation of the Geomagnetic Current Sheet», Ph. D. Thesis, University of Reading.
162. Hansen, J. P., and I. R. McDonald (1975): Statistical Mechanics of Dense Ionized Matter. IV. Density and Charge Fluctuations in a Simple Molten Salt, *Phys. Rev.*, ser. A, vol. 11, pp. 2111—2123.
163. Hansen, J. P., and I. R. McDonald (1976): «Theory of Simple Liquids», Academic Press Ltd., London.
164. Harlow, F. H. (1964): The Particle-in-Cell Computing Method in Fluid Dynamics, *Methods Comput. Phys.*, vol. 3, pp. 319—343. [Имеется перевод: Харлоу Ф. К. Численный метод частиц в ячейках для задач гидродинамики. — В кн.: Вычислительные методы в гидродинамике под ред. Б. Олдера, С. Фернбаха и М. Потенберга. — М.: Мир, 1967, с. 316.]
165. Harnwell, G. P. (1949): «Principles of Electricity and Magnetism», McGraw-Hill, New York.
166. Hartree, D. R. (1950): Some Calculations of Transients in an Electronic Valve, *Appl. Sci. Res.*, col. B1, pp. 379—390.
167. Hartree, D. R., and P. Nicolson (1941—1944); *CVD Reports Mag.* 3, 12, 18, 23, 36, British Admiralty, London.
168. Henon, M. (1964): L'Evolution Initiale d'un Amas Spherique, *Ann. Astrophys. (Fr.)*, vol. 27, pp. 83—91.

169. Henon, M. (1971a): Monte-Carlo Models of Star Clusters, *Astrophys. & Space Sci.*, vol. 13, pp. 284—299.
170. Henon, M. (1971b). The Monte-Carlo Method, *Astrophys. & Space Sci.*, vol. 14, pp. 151—167.
171. Henrici, P. (1962): «Discrete Variable Methods in Ordinary Differential Equations», John Wiley & Sons, Inc., New York.
172. Hewett, D. W., and C. W. Nielson (1978): A Multidimensional Quasineutral Plasma Simulation Model, *J. Comput. Phys.*, vol. 29, pp. 219—236.
173. Hill, G., P. N. Robson, A. Majerfeld, and W. Fawcett (1977): Effect of Ionized Impurity Scattering on the Electron Transit Time in GaAs and InP FETs, *Electron. Lett.*, vol. 13, pp. 235—236.
174. Himsworth, B. (1972): A Two-Dimensional Analysis of Gallium Arsenide Junction Field Effect Transistors with Long and Short Channels, *Solid-State Electron*, vol. 15, pp. 1353—1361.
175. Hockney, R. W. (1965): A Fast Direct Solution of Poisson's Equation using Fourier Analysis, *J. Assoc. Comput. Mach.*, vol. 12, pp. 95—113.
176. Hockney, R. W. (1966a): «The Computer Simulation of Anomalous Plasma Diffusion and the Numerical Solution of Poisson's Equation», Ph. D. Thesis, Stanford University.
177. Hockney, R. W. (1966b): Computer Experiment of Anomalous Diffusion, *Phys. Fluids*, vol. 9, pp. 1826—1835.
178. Hockney, R. W. (1967): Gravitational Experiments with a Cylindrical Galaxy, *Astrophys. J.*, vol. 150, pp. 797—806.
179. Hockney, R. W. (1968): Formation and Stability of Virtual Electrodes in a Cylinder, *J. Appl. Phys.*, vol. 39, pp. 4166—4170.
180. Hockney, R. W. (1970): The Potential Calculation and Some Applications, *Methods Comput. Phys.*, vol. 9, pp. 135—211. [Имеется перевод: Хокни Р. Методы расчета потенциала и их предложения. — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга. — М.: Мир, 1974, с. 143.]
181. Hockney, R. W. (1971): Measurements of Collision and Heating Times in a Two-Dimensional Thermal Computer Plasma, *J. Comput. Phys.*, vol. 8, pp. 19—44.
182. Hockney, R. W. (1977): Super-Computer Architecture, in «Infotech State of the Art Report: Future Systems I», Infotech International, Maidenhead, Berks, pp. 277—305.
183. Hockney, R. W. (1978a): Computers, Compilers and Poisson-Solvers, in «Computers, Fast Elliptic Solvers and Applications» (Ed. U. Schumann), Advance Publications Ltd., London, pp. 75—97.
184. Hockney, R. W. (1978b): POT4-A FACR(1) Algorithm for Arbitrary Regions, in «Computers, Fast Elliptic Solvers and Applications» (ed. U. Schumann), Advance Publications Ltd., London, pp. 141—169
185. Hockney, R. W. (1979): The Large Parallel Computer and University Research, *Contemp. Phys.*, vol. 20, pp. 149—185.
186. Hockney, R. W. (1980): Rapid Elliptic Solvers, in «Numerical Methods in Applied Fluid Dynamics» (ed. B. Hunt), Academic Press Ltd., London.
187. Hockney, R. W., and T. R. Brown (1975): A Lambda Transition in a Classical Electron Film, *J. Phys. C: Solid State Phys.*, vol. 8, 1813—1822.
188. Hockley, R. W., and D. R. K. Brownrigg (1974): Effects of Population II Stars and Three-Dimensional Motion on Spiral Structure, *Mon. Not. R. Astron. Soc.*, vol. 167, pp. 351—357.
189. Hockney, R. W., S. P. Goel, and J. W. Eastwood (1973): A 10 000 Particle Molecular Dynamics Model with Long-Range Forces, *Chem. Phys. Lett.*, vol. 21, pp. 589—591.
190. Hockney, R. W., S. P. Goel, and J. W. Eastwood (1974): Quiet High-Resolution Computer Models of a Plasma, *J. Comput. Phys.*, vol. 14, pp. 148—158.
191. Hockney, R. W., and F. Hohl (1969): Effects of Velocity Dispersion on the

- Evolution of a Disk of Stars, *Astron. J.*, vol. 74, pp. 1102—1104 & 1119—1124.
192. Hockney, R. W., and C. R. Jesshope (1981): «Parallel Computers — Architecture, Programming and Algorithms», Adam Hilger, Bristol.
 193. Hockney, R. W., R. A. Warriner, and M. Reiser (1974): Two-Dimensional Particle Models in Semiconductor Device Analysis, *Electron. Lett.*, vol. 10, pp. 484—486.
 194. Hoerner, S. von (1960): Die Numerische Integration des N-Körper-Problems für Sternhaufen I, *Z. Astrophys.*, vol. 50, pp. 184—214.
 195. Hohl, F. (1971): Numerical Experiments with a Disk of Stars, *Astrophys. J.*, vol. 168, pp. 343—359.
 196. Hohl, F. (1972): Evolution of a Stationary Disk of Stars, *J. Comput. Phys.*, vol. 9, pp. 10—25.
 197. Hohl, F. (1973): Relaxation Time in Disk Galaxy Simulations, *Astrophys. J.*, vol. 184, pp. 353—359.
 198. Hohl, F. (1975): N-Body Simulations of Disks, in «Dynamics of Stellar Systems» (ed. A. Hayli), Reidel Co., Dordrecht, pp. 349—366.
 199. Hohl, F. (1976): Suppression of Bar Instability by a Massive Halo, *Astron. J.*, vol. 81, pp. 30—36.
 200. Hohl, F. (1978): Three-Dimensional Galaxy Simulations, *Astron. J.*, vol. 83, pp. 768—778.
 201. Hohl, F., and R. W. Hockney (1969): A Computer Model of Disks of Stars, *J. Comput. Phys.*, vol. 4, pp. 306—323.
 202. Hohl, F., and T. A. Zang (1979): Collapse and Relaxation of Rotating Stellar Systems, *Astron. J.*, vol. 84, pp. 585—600.
 203. Horn, H. M. van (1971): «White Dwarfs» (ed. W. J. Luyten), Springer-Verlag, Berlin, p. 96.
 204. Hoyle, F. (1962): «Astronomy», Doubleday and Co., Inc., New York.
 205. Hubble, E. (1926): Extra-Galactic Nebulae, *Astrophys. J.*, vol. 64, pp. 321—369.
 206. Hubble, E. (1929): A Relation between Distance and Velocity among Extra Galactic Nebulae, *Proc. Natl. Acad. Sci. USA*, vol. 15, pp. 168—173.
 207. Hubble, E. (1936): «The Realm of the Nebulae», Yale University Press, reprinted (1958) by Dover Publications, Inc., New York.
 208. Huggins, M. L., and J. E. Mayer (1933): Interatomic Distances in Crystals of the Alkali Halides, *J. Chem. Phys.*, vol. 1, pp. 643—646.
 209. Hughes, M. H. (1971): Solution of Poisson's Equation in Cylindrical Coordinates, *Comput. Phys. Commun.*, vol. 2, 157—167.
 210. Hughes, M. H. (1980a): COMPOS — the OLYMPUS Fortran Compositor, *Comput. Phys. Commun.*, submitted.
 211. Hughes, M. H. (1980b): GENESIS — A Program for Automatic Construction of Standard OLYMPUS Subprograms, *Comput. Phys. Commun.*, to be submitted
 212. Hughes, M. H., and A. P. V. Roberts (1974a): TIMER — A Software Instrumentation Routine for Making Timing Measurements, *Comput. Phys. Commun.*, vol. 8, pp. 118—122.
 213. Hughes, M. H., and K. V. Roberts (1974b): OLYMPUS Restart Facilities, *Comput. Phys. Commun.*, vol. 8, pp. 123—129.
 214. Hughes, M. H., K. V. Roberts, and P. D. Roberts (1975): OLYMPUS and Preprocessor Package for an IBM 370/165, *Comput. Phys. Commun.*, vol. 9, pp. 51—58.
 215. Irvine, W. M. (1961): «Local Irregularities in a Universe Satisfying the Cosmological Principle», Ph. D. Thesis, Harvard University.
 216. Isihara, A. (1971): «Statistical Physics», Academic Press, Inc., New York. [Имеется перевод: Исихара А. Статистическая физика. — М.: Мир, 1973]
 217. James, R. A. (1977): The Solution of the Poisson-Equation for Isolated Source Distributions, *J. Comput. Phys.*, vol. 25, pp. 71—93.

218. James, R. A., and J. A. Sellwood (1978): *Galactic Models with Variable Spiral Structure*, *Mon. Not. R. Astron. Soc.*, vol. 182, pp. 331—344.
219. Jensen, E. J., W. Damgaard Kristensen, and R. M. J. Cotterill (1973): *Molecular Dynamics Studies of Melting: I. Dislocation Density and the Pair Distribution Function*, *Philos Mag*, vol. 27, pp. 623—632.
220. Jensen, K., and N. Wirth (1978): «Pascal — User Manual and Report», 2nd ed., Springer-Verlag, New York. [Имеется перевод: Йенсен К., Вирт В. ПАСКАЛЬ: Руководство для пользователя и описание языка. — М.: Финансы и статистика, 1982.]
221. Jesshope, C. R. (1975): *Numerical Solutions to the 2 Dimensional Time-Dependent Semiconductor Equations*, *Electron Lett.*, vol. 11, pp. 431—433.
222. Jesshope, C. R. (1979): *SIPSOL — A suite of Subprograms for the Solution of the Linear Equations Arising from Elliptic Partial Differential Equations*, *Comput. Phys. Commun.*, vol. 17, pp. 383—391.
223. Johnson, F. S. (1960): *The Gross Character of the Geomagnetic Field in the Solar Wind*, *J. Geophys. Res.*, vol. 65, pp. 3049—3051.
224. Jones, J. E. (1924): *On the Determination of Molecular Fields — II. From the Equation of State of a Gas*, *Proc. R. Soc., ser. A*, vol. 106, pp. 463—477.
225. Kennedy, D. P., and R. R. O'Brien (1970): *Computer-Aided Two-Dimensional Analysis of the Junction Field-Effect Transistor*, *IBM J. Res. & Dev.*, vol. 14, pp. 95—116.
226. Kilpatrick, J. A., and W. D. Ryan (1971): *Two-Dimensional Analysis of Lateral-Base Transistors*, *Electron Lett.*, vol. 7, pp. 226—227.
227. Kittel, C. (1971): «*Introduction to Solid State Physics*», 4th ed., John Wiley and Sons, Inc., New York. [Имеется перевод: Киттель Ч. Введение в физику твердого тела. — М.: Наука, 1978.]
228. Krueer, W. L., J. M. Dawson, and B. Rosen (1973): *The Dipole Expansion Method in Plasma Simulation*, *J. Comput. Phys.*, vol. 13, pp. 114—129.
229. Kuhlman-Wilsdorf, D. (1965): *Theory of Melting*, *Phys. Rev., ser. A*, vol. 140, pp. 1599—1610.
230. Kustaanheimo, P., and E. Stiefel (1965): *Perturbation Theory of Kepler Motion Vased on Spinor Regularization*, *J. Reine Angew. Math.*, vol. 218, pp. 204—219.
231. Kutta, W. (1901): *Beitrag zur Naherungsweise Integration Totaler Differentialgleichungen*, *Z. Math. Phys.*, vol. 46, pp. 435—453.
232. Lamb, H. (1932): «*Hydrodynamics*», Cambridge University Press, republished (1954) by Dover Publications Inc., New York. [Имеется перевод: Ламб Г. Гидродинамика. — М. — Л.: Гостехиздат, 1947.]
233. Langdon, A. B. (1970a): *Non-Physical Modification to Oscillations, Fluctuations and Collisions due to Space Time Differencing*, *Proc. 4th Annu. Conf. Numer. Simulation of Plasmas*, Office of Naval Research, Arlington, Va., pp. 467—495.
234. Langdon, A. B. (1970b): *Effect of the Spatial Grid in Simulation Plasmas*, *J. Comput. Phys.*, vol. 6, pp. 247—267.
235. Langdon, A. B. (1973): *Energy Conserving Plasma Simulation Algorithms*, *J. Comput. Phys.*, vol. 12, pp. 247—268.
236. Langdon, A. B. (1979a): *Kinetic Theory for Fluctuations and Noise in Computer Simulation of Plasma*, *Phys. Fluids*, vol. 22, pp. 163—171.
237. Langdon, A. B. (1979b): *Analysis of the Time Integration in Plasma Simulation*, *J. Comput. Phys.*, vol. 30, pp. 202—221.
238. Langdon, A. B., and C. K. Birdsall (1970): *Theory of Plasma Simulation Using Finite-Size Particles*, *Phys. Fluids*, vol. 13, pp. 2115—2122.
239. Langdon, A. B., and F. Lasinski (1976): *Electromagnetic and Relativistic Plasma Simulation Models*, *Methods Comput. Phys.*, vol. 16, pp. 327—366. [Имеется перевод: Лэнгдон А., Лазински Б. Электромагнитные и релятивистские вычислительные модели плазмы — В кн.: Управляемый термоядерный синтез. под. ред. Дж. Киллина. — М.: Мир, 1980, с. 347.]

240. Larson, R. B. (1970a): A Method for Computing the Evolution of Star Clusters, *Mon. Not. R. Astron. Soc.*, vol. 147, pp. 323—337.
241. Larson, R. B. (1970b): The Evolution of Star Clusters, *Mon. Not. R. Astron. Soc.*, vol. 150, pp. 93—110.
242. Lau, Y. Y., C. C. Lin, and J. W.-K. Mark (1976): Unstable Spiral Modes in Disk-Shaped Galaxies, *Proc. Natl. Acad. Sci. USA*, vol. 73, pp. 1379—1381.
243. Layzer, D. (1963): A Preface to Cosmology. I. The Energy Equations and the Virial Theorem for Cosmic Distributions, *Astrophys. J.*, vol. 138, pp. 174—184.
244. Le Bail, R. C. (1971): Use of Fast Fourier Transforms for Solving Partial Differential Equations, *J. Comput. Phys.*, vol. 9, pp. 440—465.
245. Leboeuf, J. N., T. Tajima, and J. M. Dawson (1979): A Magnetohydrodynamic Particle Code for Fluid Simulation of Plasmas, *J. Comput. Phys.*, vol. 31, pp. 379—408.
246. Leboeuf, J. N., T. Tajima, C. F. Kennel, and J. M. Dawson (1978): Global Simulation of the Time-Dependent Magnetosphere, *Geophys. Res. Lett.*, vol. 5, pp. 609—612.
247. Leboeuf, J. N., T. Tajima, C. F. Kennel, and J. M. Dawson (1979): Global Magnetohydrodynamic Simulation of the Two-Dimensional Magnetosphere, *Geophys. Monogr. Ser.*, vol. 21, pp. 536—556.
248. Lebowitz, P. A., and P. J. Price (1971): Direct Microscopic Simulation of Gunn Domain Phenomena, *Appl. Phys. Lett.*, vol. 19, pp. 530—532.
249. Lecar, M. (1968): A Comparison of Eleven Numerical Integrations of the same Gravitational 25-Body Problem, *Bull. Astron. (Fr.)*, 3e ser., t. 3, pp. 91—104.
250. Lenard, A. (1960): On Bogoliubov's Kinetic Equation for Spatially Homogeneous Plasma, *Ann Phys. (USA)*, vol. 10, pp. 390—400.
251. Levy, R. H., and R. W. Hockney (1968): Computer Experiments of Low-Density Crossed-Field Electron Beams, *Phys Fluids*, vol. 11, pp. 766—771.
252. Lewis, H. R. (1970a): Energy-Conserving Numerical Approximations for Vlasov Plasmas, *J. Comput. Phys.*, vol. 6, pp. 136—141.
253. Lewis, H. R. (1970b): Application of Hamilton's Principle to the Numerical Analysis of Vlasov Plasmas, *Methods Comput. Phys.*, vol. 9, pp. 307—338. [Имеется перевод: Льюис Х. Применение принципа Гамильтона к анализу процессов в плазме в приближении Власова — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга — М.: Мир, 1974, с. 143.]
254. Lewis, H. R., A. Sykes, and J. A. Wesson (1972): A Comparison of Some Particle-in-Cell Plasma Simulation Methods, *J. Comput. Phys.*, vol. 10, pp. 85—106.
255. Lewis, J. W. E., K. Singer, and L. V. Woodcock (1975): Thermodynamic and Structural Properties of Liquid Ionic Salts Obtained by Monte-Carlo Computation, *J. Chem. Soc., Faraday II*, vol. 71, pp. 301—312.
256. Limber, D. N. (1953): The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field, *Astrophys. J.*, vol. 117, pp. 134—144.
257. Lin, C. C., and Y. Y. Lau (1979): Density Wave Theory of Spiral Structure of Galaxies, *Stud. Appl. Math.*, vol. 60, pp. 97—163.
258. Lin, C. C., and F. H. Shu (1964): On the Spiral Structure of Disk Galaxies, *Astrophys. J.*, vol. 140, pp. 646—655.
259. Lin, C. C., C. Yuan, and F. H. Shu (1969): On the Spiral Structure of Disk Galaxies III. Comparison with Observation, *Astrophys. J.*, vol. 155, pp. 721—746.
260. Lindblad, B. (1961): On the Formation of Dispersion Rings in the Central Layer of a Galaxy, *Stockholms Obs. Ann.*, vol. 21, No. 8.
261. Lindblad, P. O. (1960a): On Tidal Interaction between Galaxies, *Stockholms Obs. Ann.*, vol. 21, No. 3.

262. Lindblad, P. O. (1960b): The Development of Spiral Structure in a Galaxy Approached by Numerical Computations, *Stockholms Obs. Ann.*, vol. 21, No. 4.
263. Lindblad, P. O. (1962): Gravitational Resonance Effects in the Central Layer of a Galaxy, in «The Distribution and Motion of Interstellar Matter in Galaxies» (ed. L. Woltjer), W. A. Benjamin Inc., New York, pp. 222—233.
264. Lindemann, F. A. (1910): Über die Berechnung Molekularer Eigenfrequenzen, *Phys. Z.*, vol. 11, pp. 609—612.
265. Lindsey, C. H., and S. G. Van der Meulen (1971): «Informal Introduction to Algol 69», North-Holland Publ. Co., Amsterdam.
266. Loeb, H. W., R. Andrew, and W. Love (1968): Application of Two-Dimensional Solutions of the Shockley-Poisson Equation to Inversion-Layer MOST Devices, *Electron. Lett.*, vol. 4, pp. 352—354.
267. Lomax, R. J. (1960): Transient Space Charge Flow, *J. Electron & Control*, vol. 9, pp. 127—140.
268. Longmire, C. L. (1963): «Elementary Plasma Physics» John Wiley and Sons, Inc., New York. [Имеется перевод: Лонгмайр К. Л. Физика плазмы. Элементарный курс. — М.: Атомиздат, 1966.]
269. Lorenz, E. N. (1976): A Rapid Procedure for Inverting Del-Square with Certain Computers, *Mon. Weather Rev.*, vol. 104, pp. 961—966.
270. Lyden-Bell, D. (1979): On a Mechanism that Structures Galaxies, *Mon. Not. R. Astron. Soc.*, vol. 187, pp. 101—107.
271. Lynden-Bell, D., and A. J. Kalnajs (1972): On the Generating Mechanism of Spiral Structure, *Mon. Not. R. Astron. Soc.*, vol. 157, pp. 1—30.
272. Marder, R. M. (1975): GAP-A PIC-Type Fluid Code, *Math. Comput.*, vol. 29, pp. 434—446.
273. Mark, J. W.-K. (1974): On Density Waves in Galaxies. I. Source Terms and Action Conservation, *Astrophys. J.*, vol. 193, pp. 539—559.
274. Mark, J. W.-K. (1977): On Density Waves in Galaxies. V. Maintenance of Spiral Structure and Discrete Spiral Modes, *Astrophys. J.*, vol. 212, pp. 645—658.
275. Martin, E. D. (1974): A Generalised-Capacity-Matrix Technique for Computing Aerodynamic Flows, *Comput & Fluids*, vol. 2, pp. 79—97.
276. Maruhn, J. A., T. A. Welton, and C. Y. Wong (1976): Remarks on the Numerical Solution of Poisson's Equation for Isolated Charge Distributions, *J. Comput. Phys.*, vol. 20, pp. 326—335.
277. Matsuda, Y., and H. Okuda (1975): Collisions in Multidimensional Plasma Simulations, *Phys. Fluids*, vol. 18, pp. 1740—1747.
278. McCrory, R. L., R. L. Morse, and K. A. Taggart (1977): Growth and Saturation of Instability of Spherical Implosion Driven by Laser or Charged Particle Beams, *Nuclear Sci & Eng.*, vol. 64, pp. 163—176.
279. McDonald, I. R., and K. Singer (1967): Calculation of Thermodynamic Properties of Liquid Argon from Lennard-Jones Parameters by a Monte-Carlo Method, *Discuss. Faraday Soc.*, vol. 43, pp. 40—49.
280. Meijerink, J. A., and H. A. van der Vorst (1977): An Iterative Solution Method for Linear Systems of which the Coefficient Matrix is a Symmetric M-Matrix, *Math. Comput.*, vol. 31, pp. 148—162.
281. Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953): Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, vol. 21, pp. 1087—1092.
282. Miller, R. H. (1971): Numerical Experiments in Collisionless Systems, *Astrophys. & Space Sci.*, vol. 14, pp. 73—90.
283. Miller, R. H. (1976): Predominance of Two-Arm Spirals, *Astrophys. J.*, vol. 207, pp. 408—413.
284. Miller, R. H. (1978): Free Collapse of a Rotating Sphere of Stars, *Astrophys. J.*, vol. 223, pp. 122—128.

285. Miller, R. H., and K. H. Prendergast (1968): Stellar Dynamics in a Discrete Phase Space, *Astrophys. J.*, vol. 151, pp. 699—709.
286. Miller, R. H., K. H. Prendergast, and W. J. Quirk (1970): Numerical Experiments on Spiral Structure, *Astrophys. J.*, vol. 161, pp. 903—916.
287. Moglestue, C. (1979): Computer Simulation of a Dual Gate GaAs Field-Effect Transistor using the Monte-Carlo Method, *Solid State & Electron Devices*, vol. 3, pp. 133—136.
288. Moglestue, C., and S. J. Beard (1979): A Particle Model Simulation of Field Effect Transistors, in «Numerical Analysis of Semiconductor Devices». (eds. B. T. Browne and J. J. H. Miller), Boole Press, Dublin, pp. 232—236.
289. Montgomery, D. C., and D. A. Tidman (1964): «Plasma Kinetic Theory», McGraw-Hill Book Company, New York.
290. Morse, R. L. (1970); Multidimensional Plasma Simulation by the Particle — in — Cell Method, *Methods Comput. Phys.*, vol. 9, pp. 213—239. [Имеется перевод: Морз Р. Моделирование многомерной плазмы с помощью метода частиц в ячейке. — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Ротенберга. — М.: Мир, 1974, с. 213.]
291. Morse, R. L., and C. W. Nielson (1969): Numerical Simulation of Warm Two-Beam Plasma, *Phys. Fluids*, vol. 12, pp. 2418—2425.
292. N. C. C. (1970): Standard Fortran Programming Manual, Computer Standards Series, National Computing Centre Ltd., Manchester, UK.
293. Nielsen, D. (1978): «Three-Dimensional Electromagnetic Particle Simulation of Fusion Plasmas», Ph. D. Thesis, Stanford University.
294. Nielsen, D., J. Green, and O. Buneman (1978); Dynamic Evolution of a Z Pinch, *Phys. Rev. Lett.*, vol. 42, pp. 1274—1277.
295. Nielson, W. C., and H. R. Lewis (1976): Particle-Code Models in the Non-radiative Limit, *Methods Comput. Phys.*, vol. 16, pp. 367—388. [Имеется перевод: Нильсон К., Льюис Г. Модели укрупненных частиц в безызлучательном пределе. — В кн.: Управляемый термоядерный синтез, под ред. Дж. Киллина. — М.: Мир, 1980, с. 347.]
296. Okuda, H. (1970): Nonphysical Instabilities in Plasma Simulation Due to Small $\lambda_D/\Delta x$, *Proc. 4th Annu. Conf. on Numer. Simulation of Plasma*, Office of Naval Research, Arlington, Va., pp. 515—526.
297. Okuda, H. (1972a): Nonphysical Noises and Instabilities in Plasma Simulation due to a Spatial Grid, *J. Comput. Phys.*, vol. 10, pp. 475—486.
298. Okuda, H. (1972b): Verification of Theory for Plasma of Finite-Sized Particles, *Phys. Fluids*, vol. 15, pp. 1268—1274.
299. Okuda, H., and C. K. Birdsall (1970): Collisions in a Plasma of Finite-Size Particles, *Phys. Fluids*, vol. 8, pp. 2123—2134.
300. O'Leary, D. P., and O. Widlund (1979): Capacity Matrix Methods for the Helmholtz Equation on General Three-Dimensional Regions, *Math. Comput.*, vol. 33, pp. 849—879.
301. Oliphant, T. A., and C. W. Nielson (1970): Simulation of Binary Collision Processes in Plasmas, *Phys. Fluids*, vol. 13, pp. 2103—2107.
302. Onuki, A. (1977): Collective Modes of a Two-Dimensional Electron Fluid, *J. Phys. Soc. Jpn*, vol. 43, pp. 396—405.
303. Orens, J. H., J. P. Boris, and I. Haber (1970): Optimisation Techniques for Particle Codes, in *Proc. 4th Annu. Conf. Numer. Simulation of Plasmas*, Office of Naval Research, Arlington, Va., pp. 526—558.
304. Ostriker, J. P., and P. J. E. Peebles (1973): A Numerical Study of the Stability of Flattened Galaxies: or, Can Cold Galaxies Survive? *Astrophys. J.*, vol. 186, pp. 467—480.
305. Parrinello, M., and M. P. Tosi (1979): Structure and Dynamics of Simple Ionic Liquids, *Riv. Nuovo Cimento, Ser. 3*, vol. 2, No. 6, pp. 1—69.
306. Pauling, L. (1950): «The Nature of the Chemical Bond», Oxford Univer-

- sity Press, London. [Имеется перевод: Паулинг Л. Природа химической связи. — М.: ИЛ, 1947.]
307. Peebles, P. J. E. (1974): The Gravitational-Instability Picture and the Nature of the Distribution of Galaxies, *Astrophys. J.*, vol. 189, pp. L51—L53.
308. Peebles, P. J. E. (1978): Comment after paper by S.-J. Aarseth, in «The Large Scale Structure of the Universe», IAU Symposium No. 79 (eds. M. Longair and J. Einasto), Reidel Co., Dordrecht, p. 194.
309. Peiravi, A., and C. K. Birdsall (1978): «Self Heating in 1-D Thermal Plasma, Comparison of Weights: Optimal Parameter Choice» Memorandum No. UCB/ERL M78/32, ERL, College of Engineering, University of California, Berkeley.
310. Peters, C. F. (1968): Numerical Regularisation, *Bull. Astron. (Fr.)*, ser. 3, vol. 3, 167—175.
311. Petravic, M., and G. Kuo-Petravic (1979): An ILUCG Algorithm which Minimises in the Euclidean Norm, *J. Comput. Phys.*, vol. 32, pp. 263—269.
312. Platzman, P. M., and H. Fukuyama (1974): Phase Diagram of the Two-Dimensional Electron Liquid, *Phys. Rev. B*, vol. 10, pp. 3150—3158.
313. Pollack, E. L., and J. P. Hansen (1973): Statistical Mechanics of Dense Ionised Matter. II. Equilibrium Properties and Melting Transition of the Crystallised One-Component Plasma, *Phys. Rev.*, ser. A, vol. 8, pp. 3110—3122.
314. Potter, D. (1973): «An Introduction to Computational Physics», John Wiley and Sons, Ltd., London. [Имеется перевод: Поттер Д. Вычислительные методы в физике. — М.: Мир, 1974.]
315. Potter, D. (1976): Waterbag Methods in Magnetohydrodynamics, *Methods Comput. Phys.*, vol. 16, pp. 43—83 [Имеется перевод: Поттер Д. Метод водяного мешка в магнитной гидродинамике. — В кн.: Управляемый термоядерный синтез, под ред. Дж. Киллина — М.: Мир, 1980, с. 51.]
316. Proskurowski, W. (1977): «Numerical Solution of the Helmholtz's Equation by Implicit Capacity Matrix Methods», Report LBL-6402, Lawrence Berkeley Laboratory, University of California, Berkeley.
317. Proskurowski, W., and O. Widlund (1976): On the Numerical Solution of Helmholtz's Equation by the Capacity Matrix Method, *Math. Comput.*, vol. 30, pp. 433—468.
318. Rahman, A. (1964): Correlations in the Motion of Atoms in Liquid Argon, *Phys. Rev.*, ser. A, vol. 136, pp. 405—411.
319. Rahman, A., R. H. Fowler, and A. H. Narten (1972): Structure and Motion in Liquid BeF_2 , LiBeF_3 and LiF from Molecular Dynamics Calculations, *J. Chem. Phys.*, vol. 57, pp. 3010—3011.
320. Rathmann, C. E., J. L. Vomvoridis, and J. Denavit (1978): Long-Time-Scale Simulation of Resonant Particle Effects in Langmuir and Whistler Modes, *J. Comput. Phys.*, vol. 26, pp. 408—442.
321. Rees, H. D. (1968): Calculation of Steady-State Distribution Function by Exploiting Stability, *Phys. Lett.*, vol. 26A, pp. 416—417.
322. Rees, H. D. (1969): Calculation of Distribution Functions by Exploiting the Stability of the Steady State, *J. Phys. Chem. Solids*, vol. 30, pp. 643.
323. Rees, H. D., G. S. Sanghera, and R. A. Warriner (1977): Low Temperature FET for Low-Power High-Speed Logic, *Electron. Lett.*, vol. 13, pp. 156.
324. Reid, J. K. (1971): On the Method of Conjugate Gradients for the Solution of Large Sparse Systems of Equations, in «Large Sparse Sets of Linear Equations», Proc. of Conf. in Oxford, April, 1970, Academic Press Ltd., London.
325. Reid, J. K. (1976): Sparse Matrices, in Proc. Conf. on «The State-of-the-Art in Numerical Analysis», York, April, Academic Press Ltd, London.
326. Reiser, M. (1970): Two-Dimensional Analysis of Substrate Effects in Junction FETs, *Electron. Lett.*, vol. 6, pp. 493—494.

327. Reiser, M. (1971): Difference Methods for the Solution of the Time Dependent Semi-Conductor Flow Equations, *Electron Lett.*, vol. 7, pp. 353—355.
328. Reiser, M. (1972): Large Scale Numerical Simulation in Semi-Conductor Device Modelling, *Comput Methods Appl. Mech. & Eng.*, vol. 1, pp. 17—37.
329. Reiser, M. (1973): On the Stability of Finite Difference Schemes in Transient Semiconductor Problems, *Comput. Methods Appl. Mech. & Eng.*, vol. 2, pp. 65—68.
330. Reitz, J. R., and F. J. Milford (1962): «Foundations of Electromagnetic Theory», Addison-Wesley, Reading, Mass.
331. Rice, S. A., and P. Gray (1965): «The Statistical Mechanics of Simple Liquids», Interscience Publishers, John Wiley and Sons, Inc., New York.
332. Richtmyer, R. D., and K. W. Morton (1967): «Difference Methods for Initial-Value Problems», Interscience Publishers, John Wiley and Sons, Inc., New York. [Имеется перевод: Рихтмайер Р., Мортон К. Разностные методы решения краевых задач. — М.: Мир, 1972.]
333. Roberts, K. V. (1974): An Introduction to the OLYMPUS System, *Comput. Phys. Commun.*, vol. 7, pp. 237—243.
334. Roberts, K. V. (1975): The OLYMPUS Programming System, *АТОМ*, vol. 226, pp. 137—147.
335. Roberts, K. V., and D. E. Potter (1970): Magneto hydrodynamic Calculation, *Methods Comput. Phys.*, vol. 9, pp. 339—420. [Имеется перевод: Робертс К., Поттер Д. Магнитогидродинамические методы. — В кн.: Вычислительные методы в физике плазмы, под ред. Б. Олдера, С. Фернбаха и М. Потенберга. — М.: Мир, 1974, с. 143.]
336. Rosen, B., W. L. Kruer, and J. M. Dawson (1970): A New Version of the Dipole Expansion Scheme, *Proc. 4th Annu. Conf. Numer. Simulation of Plasmas*, Office of Naval Research, Arlington, Va., pp. 561—573.
337. Ross, S. (1976): «A First Course in Probability», Macmillan, New York.
338. Runge, C. (1895): Ueber die Numerische Auflösung von Differentialgleichungen, *Math. Ann. (Ger.)*, vol. 46, pp. 167—178.
339. Sagdeev, R. Z. (1979): The 1979 Oppenheimer Lectures: Critical Problems in Plasma Astrophysics. II. Singular Layers and Reconnection, *Rev. Mod. Phys.*, vol. 51, pp. 11—20.
340. Sandage, A. (1961): «The Hubble Atlas of Galaxies», Publication 618, Carnegie Institute of Washington, Washington, D. C.
341. Sangster, M. J. L., and M. Dixon (1976): Interionic Potentials in Alkali Halides and their use in Simulation of Molten Salts, *Adv. Phys.*, vol. 25, pp. 247—342.
342. Scharfetter, D. L., and H. K. Gummel (1969): Large-Signal Analysis of Silicon Read Diode Oscillator, *IEEE Trans. Electron Devices*, vol. ED-16, pp. 64—77.
343. Schofield, P. (1973): Computer Simulation Studies in the Liquid State, *Comput. Phys. Commun.*, vol. 5, pp. 17—23.
344. Schumann, U. (Ed.) (1978): «Computers, Fast Elliptic Solvers and Applications», Advance Publications Ltd, London.
345. Sellwood, J. A., and R. A. James (1979): Angular Momentum Redistribution by Spiral Waves in Computer Models of Disc Galaxies, *Mon. Not. R. Astron. Soc.*, vol. 187, pp. 483—496.
346. Shampine, L. F. (1979): Storage Reduction for Runge-Kutta Codes, *ACM Trans. Math. Software*, vol. 5, pp. 245—250.
347. Shur, M. (1976): Influence of Non-Uniform Field Distribution on Frequency Limits in GaAs FETs, *Electron. Lett.*, vol. 12, pp. 615—616.
348. Slotboom, J. W. (1969): Iterative Scheme for 1- and 2-Dimensional D. C.-Transistor Simulation, *Electron. Lett.*, vol. 5, pp. 677—678.
349. Slotboom, J. W. (1973): Computer-Aided Two-Dimensional Analysis of Bipolar Transistors, *IEEE Trans. Electron Devices*, vol. ED-20, pp. 669—679.

350. Spitzer, L., and M. H. Hart (1971a): **Random Gravitational Encounters and the Evolution of Spherical Systems. I. Method**, *Astrophys. J.*, vol. 164, pp. 399—409.
351. Spitzer, L., and M. H. Hart (1971b): **Random Gravitational Encounters and the Evolution of Spherical Systems. II. Models**, *Astrophys. J.*, vol. 166, pp. 483—511.
352. Spitzer, L., and M. H. Hart (1972): **Random Gravitational Encounters and the Evolution of Spherical Systems. III. Halo**, *Astrophys. J.*, vol. 173, pp. 529—547.
353. Spitzer, L., and T. X. Thuan (1972): **Random Gravitational Encounters and the Evolution of Spherical Systems. IV. Isolated Systems of Identical Stars**, *Astrophys. J.*, vol. 175, pp. 31—61.
354. Stern, F. (1972): **Self-Consistent Results for n-Type Si Inversion Layers**, *Phys. Rev. ser. B*, vol. 5, pp. 4891—4899.
355. Stix, T. H. (1962): **«The Theory of Plasma Waves»**, McGraw-Hill, New York. [Имеется перевод: Стикс Т. К. Теория плазменных волн. — М.: Госатомиздат, 1965.]
356. Stone, H. J. (1968): **Iterative Solution of Implicit Approximations of Multi-Dimensional Partial Differential Equations**, *SIAM J. Numer. Anal.*, vol. 5, pp. 530—558.
357. Strang, G., and G. Fix (1973): **«An Analysis of the Finite Element Method»**, Prentice-Hall, Inc., Englewood Cliffs, N. J. [Имеется перевод: Стренг Г., Фикс Дж. Теория метода конечных элементов. — М.: Мир, 1977.]
358. Swarztrauber, P. N. (1974): **A Direct Method for the Discrete Solution of Separable Elliptic Equations**, *SIAM J. Numer. Anal.*, vol. 11, pp. 1136—1150.
359. Swarztrauber, P. N. (1977): **The Methods of Cyclic Reduction, Fourier Analysis, and the FACR Algorithm for the Discrete Solution of Poisson's Equation on the Rectangle**, *SIAM Rev.*, vol. 19, pp. 490—501.
360. Swarztrauber, P. N., and R. A. Sweet (1975): **Efficient FORTRAN Subprograms for the Solution of Elliptic Partial Differential Equations**, NCAR Technical Note, 1A-109, National Center for Atmospheric Research, Boulder, Col.
361. Sweet, R. A. (1974): **A Generalised Cyclic-Reduction Algorithm**, *SIAM J. Numer. Anal.*, vol. 11, pp. 506—520.
362. Sweet, R. A. (1977): **A Cyclic Reduction Algorithm for Solving Block Tri-diagonal Systems of Arbitrary Dimension**, *SIAM J. Numer. Anal.*, vol. 14, pp. 706—719.
363. Sze, S. M. (1969): **«Physics of Semiconductor Devices»**, Wiley-Interscience, New York.
364. Temperton, C. (1979a): **A Fast Poisson-Solver for Large Grids**, *J. Comput. Phys.*, vol. 30, pp. 145—148.
365. Temperton, C. (1979b): **Direct Methods for the Solution of the Discrete Poisson Equation: Some Comparisons**, *J. Comput. Phys.*, vol. 31, pp. 1—20.
366. Temperton, C. (1980): **On the FACR(1) Algorithm for the Discrete Poisson Equation**, *J. Comput. Phys.*, vol. 34, pp. 314—329.
367. Thouless, D. J. (1978): **Melting of the Two-Dimensional Wigner Lattice**, *J. Phys. C: Solid State Phys.*, vol. 11, pp. L189—L190.
368. Tien, P. K., and J. Moshman (1956): **Monte Carlo Calculation of Noise Near the Potential Minimum of a High-Frequency Diode**, *J. Appl. Phys.*, vol. 27, pp. 1067—1078.
369. Tinsley, B. M. (1977): **The Cosmological Constant and Cosmological Change**, *Phys. Today*, vol. 30, No. 6, pp. 32—38.
370. Todd, A. M. M. (1975): **Numerical Simulation of Shock-Heated Plasma with a Magnetic Dam**, *Phys. Fluids*, vol. 18, pp. 453—457.
371. Toomre, A. (1964): **On the Gravitational Stability of a Disk of Stars**, *Astrophys. J.*, vol. 139, pp. 1217—1238.

372. Toomre, A. (1977): Theories of Spiral Structure, *Annu. Rev. Astron. & Astrophys.*, vol. 15, pp. 437—478.
373. Tosi, M. P., and F. G. Fumi (1964): Ionic Sizes and Born Repulsion Parameters in the NaCl-Type Alkali Halides — II, *Phys. Chem. Solids*, vol. 25, pp. 45—52.
374. Totsuji, H., and T. Kihara (1969): The Correlation Function for the Distribution of Galaxies, *Publ. Astron. Soc. Jpn.*, vol. 21, pp. 221—229.
375. Vacouleurs, G. de (1959): General Physical Properties of External Galaxies, *Handb. Phys. (Ger.)*, vol. 53, pp. 311—372.
376. Varga, R. S. (1962): «Matrix Iterative Analysis», Prentice-Hall, Inc., Englewood Cliffs, N. J.
377. Verlet, L. (1967): Computer Experiments on Classical Fluids. I. Thermodynamic Properties of Lennard-Jones Molecules, *Phys. Rev.*, vol. 159, pp. 98—103.
378. Vokes, J. C., B. T. Hughes, D. R. Wight, J. R. Dawsey, and S. J. W. Shrubbs (1979): Novel Microwave GaAs Field Effect Transistors, *Electron. Lett.*, vol. 15, pp. 627—629.
379. Wada, T., and J. Frey (1979): Physical Basis of Short-channel MESFET Operation, *IEEE, J. Solid State Circuits*, vol. SC-14, pp. 398—412.
380. Wadhwa, R. P., O. Buneman, and D. F. Brauch (1965): Two-Dimensional Computer Experiments on Ion-Beam Neutralization, *Am. Inst. Aeronaut. & Astronaut.*, vol. 3, pp. 1076—1081.
381. Warriner, R. A. (1976): «Computer Simulation of Gallium Arsenide Semiconductor Devices», Ph. D. Thesis, Reading University
382. Warriner, R. A. (1977a): Distribution Function Relaxation Times in Gallium Arsenide, *Solid State & Electron Devices*, vol. 1, pp. 92—96.
383. Warriner, R. A. (1977b): Computer Simulation of Negative-Resistance Oscillators Using a Monte-Carlo Model of Gallium Arsenide, *Solid State & Electron Devices*, vol. 1, pp. 97—104
384. Warriner, R. A. (1977c): Computer Simulation of Gallium Arsenide Field-Effect Transistors Using Monte-Carlo Methods, *Solid State & Electron Devices*, vol. 1, pp. 105—110.
385. Widlund, O. (1972): On the Use of Fast Methods for Separable Finite-Difference Equations for the Solution of General Elliptic Problems, in «Sparse Matrices and their Applications» (ed. D. J. Rose and R. A. Willoughby), Plenum Press, London, pp. 121—131.
386. Wijngaarden, A., B. J. Mailieux, J. E. L. Peck, C. H. A. Koster, M. Sintzoff, C. H. Lindsey, L. G. L. T. Meertens, and R. G. Fisker (1976): «Revised Report on Algol 68», Springer-Verlag, Berlin. [Имеется перевод: Пересмотренное сообщение об Алголе 68. — М.: Мир, 1979.]
387. Williamson, J. H. (1980): Low Storage Runge-Kutta Schemes. *J. Comput. Phys.*, vol. 35, pp. 48—56.
388. Woodcock, L. V., C. A. Angell, and P. Cheeseman (1976): Molecular Dynamic Studies of the Vitreous State: Simple Ionic Systems and Silica, *J. Chem. Phys.*, vol. 65, pp. 1565—1577.
389. Yamaguchi, K., S. Asai, and H. Kodera (1976): Two-Dimensional Numerical Analysis of Stability Criteria of GaAs FETs, *IEEE Trans. Electron Devices*, vol. ED-23, pp. 1283—1290
390. Young, D. (1962): The Numerical Solution of Elliptic and Parabolic Partial Differential Equations, in «Survey of Numerical Analysis» (ed. J. Todd), McGraw Hill, New York, pp. 380—418.
391. Yu, S. P., G. P. Kooyers, and O. Buneman (1965): Time-Dependent Computer Analysis of Electron-Wave Interaction in Crossed-Fields, *J. Appl. Phys.*, vol. 36, 2550—3559.

Дополнительная литература

1. Белоцерковский О. М., Давыдов Ю. М. Метод крупных частиц в газовой динамике. — М.: Наука, 1982.
2. Гайфулин С. А., Карпов В. Я., Мищенко Т. В. САФРА. Функциональное наполнение. Система OLYMPUS. — Препринт ИПМ АН СССР № 27. — М., 1980.
3. Годунов С. К., Рябенский В. С. Разностные схемы. Введение в теорию. — М.: Наука, 1977.
4. Двумерные численные модели плазмы: Сборник научных трудов под ред. К. В. Брушлинского. — М.: ИПМ АН СССР, 1979.
5. Марчук Г. И. Методы вычислительной математики. — М.: Наука, 1980.
6. Марчук Г. И., Шайдуров В. В. Повышение точности решений разностных схем. — М.: Наука, 1979.
7. Решение задач методом крупных частиц. Труды ВЦ АН СССР под ред. С. П. Ломнева. — М.: ВЦ АН СССР, 1970.
8. Самарский А. А. Теория разностных схем. — М.: Наука, 1977.
9. Самарский А. А., Андреев В. Б. Разностные методы для эллиптических уравнений. — М.: Наука, 1976.
10. Самарский А. А., Николаев Е. С. Методы решения сеточных уравнений. — М.: Наука, 1978.
11. Численное моделирование коллективных процессов в плазме. Сборник научных трудов под ред. М. В. Масленникова. — М.: ИПМ АН СССР, 1980.
12. Яненко Н. Н. Метод дробных шагов решения многомерных задач математической физики. — Новосибирск: Наука, 1967.

Предметный указатель

- Автоматическая генерация программ** 80
Адиабатический закон 372
Алгоритм сопряженных градиентов 228—231
Альфвеновские волны 424
Ампера уравнение 420, 422
— модель частиц 420—424
- Баробразная неустойчивость** 519, 529
ББГКИ иерархия 28, 540
Безразмерные единицы 52, 555, 569
Белый карлик 494, 552
Бесстолкновительные системы 22, 23, 29, 495. *См. также Моделирование бесстолкновительной плазмы, Моделирование галактик, Власова уравнение*
Бимана алгоритм 554
Ближний порядок 569
Близкодействие 27
Больцмана закон 212
— постоянная 30, 91, 383, 581
— распределение 401
— уравнение 31
Большой взрыв 498
Бризерная мода 576
Бунемана алгоритм 247
Буферный слой 487
Быстрое преобразование Фурье 252, 318, 397
Быстрые методы для эллиптических уравнений 245
- Векторный потенциал** 420
Взаимодействие мод 60, 289
Вириал 587
Вистлеры 373
Вихревая модель 20, 34, 462
Власова уравнение 21, 29, 42, 46, 273
Внутренняя энергия 564
Вмороженность 372, 414
Время отклонения 383
— — в модели галактик 496
- — — плазмы 298, 381, 386
— — — системе галактик 516
— — — звезд 495
Вселенной замкнутость, открытость 498, 528, 532, 534
— расширение 498, 509, 532
— — в сопутствующих координатах 535
Вычислительная наука 14
Вычислительный эксперимент 14, 16, 18
- Галактики барообразные** 498, 509, 531
— дискообразные 510, 519, 524
— спиральные 498, 509
— эллиптические 498, 509, 531
Галактическое звездное скопление 457
Гало 500, 511, 522, 527
Гармонический осциллятор 137
Гаусса метод исключения 205, 232
Гаусса — Зайделя метод 220
Гаусса теорема 533
Гельмгольца уравнение 247
Геоманитный хвост 416
Гибридные модели 373
Гидродинамические уравнения 339, 434
Гироорбита 405
Гирорадиус 392
Гирочастота 392
Гладкость 194
Граничные условия в прямоугольнике 206
— — для дифференциального уравнения в частных производных 206, 248, 253
— — — изолированной системы 246, 259
— — периодические 254, 419
— — простые 206, 248
— — с заданными градиентами 255
— — — значениями 254
— — смешанные 206
Группы общих блоков 74

- Давление 581
 — магнитное 416
 — обусловленное взаимодействием 582
 Давления тензор 33
 Дальний порядок 567
 Дальнодействие 27
 Данжи модель магнитосферы 414, 419
 Двухжидкостное приближение 399
 Двухпоточковая неустойчивость 61
 Дебаевская длина 30, 44, 91, 371, 386, 404
 Дейтериево-третиевая плазма 391
 Джеймса алгоритм 263
 Джинса неустойчивость 519, 525
 Джонсона шум 402
 Диагностика 69, 75, 410
 Дирихле граничные условия 245, 254
 Дисперсионное соотношение для бесстолкновительной плазмы 43, 277, 281, 283
 — — — спиральных волн плотности 510, 527
 Дифференциальное вращение 509, 523
 Диффузия аномальная 374
 — боровская 393
 — в микрокристаллах 592
 — отрицательная 418
 — плазмы 392, 406
 — столкновительная 418
 — частиц 383
 — численная 406, 417
 Диэлектрическая проницаемость 270, 298
 Долина 444
 Домен 568
 Дрейф 34, 395, 405

 Жесткость ионов 569
 Жидкости моделирование 32, 380, 419

 Завихренность 34
 Затухание Ландау в бесстолкновительной плазме 44, 402
 — — — модельной РМ-плазме 289
 Звездные системы 501
 Зонная структура в полупроводниках 442

 Изолированные системы 206, 259
 Интеграл ошибок 466
 Интерполяция силы 49
 — — SIC 158
 — — NGP 50, 156
 — — TSC 332, 349
 Ионные кристаллы 580

 Итерации по Пикару 434
 Итерационный параметр 241, 270

 Классы подпрограмм 74
 Ковалентная связь 593
 Колебания молекул 562, 576
 Коллективная мода колебаний 560, 561, 577
 Конкуса и Голуба метод 269
 Координационная оболочка 573
 Координационное число 573, 587, 590
 Корневой годограф 128
 Коротационный резонанс 525
 Коррекция частоты 139
 Коррелированные системы 22
 Корреляционные функции 445
 Космология 502, 532, 545, 548
 Кривая вращения 515, 520
 Кривизна пространства 533
 Кристаллит 569
 Кубическая решетка 573
 Куранта — Фридрикса — Леви условие 418

 Лакса — Вендроффа метод 373, 417
 Ларморовский радиус 380, 392. См. также Гирорадиус
 Леви — Чивита регуляризация 508
 Легирование 430
 Лейзера — Ирвина уравнение энергии 536
 Леннарда — Балеску — Гарнси уравнение 273
 Леннарда-Джонса закон взаимодействия 550
 Лина — Шу теория спиральных волн 527
 Линдблада резонансы 526
 Линдемана критерий 597
 Лиувилля уравнение 27
 Локальные сверхскопления 497
 Лоренца сила 141
 Лоусона критерий 391
 Льюиса метод 184
 Лямбда-переход 566

 Магнитная гидродинамика 371
 — суббуря 416, 419
 Магнитное зеркало 408
 — поле 370
 — — межпланетное 414
 Магнитный шар 409
 Маделунга постоянная 574
 Максвелла распределение 44, 385, 433, 465, 558
 — уравнения 370, 374
 Марковица критерий 234

- Матрица емкости** 264
 — перехода 133, 215
Маха конус 418
МГД 371, 380, 393
 — модель частиц 416, 424
 — неустойчивости 393
Мера качества модели РМ 312
 — — — РЗМ 333
Местная группа галактик 497
Метод переменных направлений 206, 224, 244
 — преобразования Фурье 251
 — предиктор-корректор 555
 — частица-сетка (РМ) 38, 153, 273, 376
 — частица-частица (РР) 36, 502, 545, 551
 — частица-частица—частица-сетка (РЗМ) 40—41, 326
 — SIC 378, 385, 389, 423, 512, 518
 — TSC 378, 387, 389, 512, 518
 — FACR 255
Методы для разреженных матриц 232
 — свертки 259—263
Млечный путь 497
Модель ведущего центра 34—35
 — водяного мешка 380
 — электростатической плазмы 380—390
 — частиц 15
 — QPM 378, 512, 518
Молекулярная динамика 549
Монте-Карло 437, 463, 503, 553, 554
Мультипольное разложение 168
- Нейтральность заряда** 401, 424
Нейтронные звезды 494
Неполное разложение 238—239
Неустойчивость плазмы 396
Ньютона метод 211—214, 435, 436
Ньютона — Рафсона метод 173
- Обобщенные функции** 193, 605
Однокомпонентная плазма (ОКП) 494, 552, 553, 555
Одночастичные колебания 576
Окаймляющий слой 358
Операторы обострения 172, 200
Оптимизация моделей РМ 311
 — — РЗМ 345
Оптическая мода 576
Острайкера и Пиблса условие устойчивости 522, 523, 529
- Паразитные корни** 133, 145
Параметр решетки 572
Парная функция распределения 569
Парсевала теорема 605
- Паули принцип запрета** 552
Пересоединение магнитных силовых линий 415
Плавление ионных систем 554, 588
 — однокомпонентной плазмы 553
 — систем Леннарда-Джонса 554
 — электронной пленки 553, 564
Плазма в электростатическом поле 373
Плоская составляющая 499, 530
Погрешности аппроксимации 124
 — — энергии 582
 — — в уравнении поля 183
 — — в операторе градиента 183
 — — давления 582
 — округления 124
 — уравнение эволюции 127
Подвижность 34, 448—450
Подложка 430, 483
Поляризация 552, 597
Полярные сияния 414
Приведенная масса 557, 570
Пробочное отношение 409
Программные спецификации 67—70
Пространственное взвешивание 181
Протогалактика 498, 499, 509, 530—531
Пуассона распределение 541
 — уравнение 35
- Равновесное расстояние** 570, 572
Равнораспределение по энергии 501, 519
Радиус обрезания 326
 — эффективный 343, 367
Разностная схема с перешагиванием 46, 121, 125, 327, 368
Распределение заряда 49, 155, 165, 168, 170, 189
Релаксация энергии в галактиках 519
Рунге — Кутты схемы 147—151
- Самодействие** 161, 176, 188, 274, 309
Связные списки 339, 359
Сегрегация объектов по массе в системе звезд 501
Сетка заряда — потенциала 48, 327, 338
Сеточные гармоники 189, 289
Силовая постоянная 575
Системы координат 247
Система двойной звезды 496
Скелетная программа (CRONUS) 74—79
Скрытая теплота плавления 588, 589
Скучивание галактик 501, 532
 — звезд 504, 508
Случайное блуждание 382, 392, 467

- Смешанные схемы 160
 Соглашение по начальным буквам 79
 Солнечный ветер 414
 Сопутствующие координаты 535
 Сохранение импульса 308
 — количества движения 310
 — энергии 302
 Спиральные структуры в галактиках 509
 Среднеквадратичное смещение 592, 593
 Стандартный программный файл (СПФ) 70—74
 Стеклование 554, 588
 Стохастический нагрев 385. *См. также* Сохранение энергии
 Стоуна строго неявная процедура 239—243
 Субстанциальная производная 372
 Сферическая составляющая 499, 530
 Твердые сферы 550
 Теорема о свертке 260, 605
 Теория соответственных состояний 597—599
 Тепловая скорость 91, 394, 404
 Тепловое равновесие 564
 Тетраэдрический угол 592
 Токовый слой в магнитосфере 415—416, 420, 426
 Томаса алгоритм 227—228
 Трехмерный электромагнитный код 379
 — электростатический код 410
 Тумре критерий 519, 520, 522
 Угловая анизотропия 182, 313
 — — в операторе градиента 183
 — — — силе 313
 — — — уравнении поля 182
 Угол связи 591
 Удельная теплоемкость 564
 Управляемый термоядерный синтез 379, 391
 Уравнение неразрывности 399, 434
 Уравнения в частных производных 205—272
 — поля 27, 47—49, 51, 200. *См. также* Пуассона уравнение
 — — методы решения 207—272
 — с вязкостью 144
 Ускорение по Чебышёву 222, 271
 Фазовые превращения 549
 Фазовый переход 553, 564
 Факторизованные схемы 167, 174, 298, 332
 Фоккера — Планка уравнение 173, 502
 Фридмана вселенная 532
 — уравнения 533
 — — численное интегрирование 538
 Функция влияния 198
 — — оптимальная 313, 335
 — — — для чередующихся сеток 337
 — выборки 607
 — Грина 183, 198, 260, 378, 390. *См. также* Функция влияния
 — импульса прямоугольного 606
 — — треугольного 607
 — радиального распределения (ФРР) 368, 579, 587. *См. также* Парная корреляционная функция
 — с конечной шириной спектра 195, 326, 336
 Фурье преобразование 600
 Хаббла время 534
 — классификация галактик 509
 — постоянная 533, 536
 — расширение 542
 Холецкого метод неполного разложения 227, 243
 Центральная предельная теорема 466
 Цепочечная сетка 326, 337
 Цикл шага по времени 14, 36
 — — — — в моделях РМ 38
 — — — — — РР 36
 — — — — — РЗМ 41
 Частица конечного размера 177, 287, 298, 377
 Частота колебаний 558, 572, 575
 — плазменная 44, 58, 91, 373, 404
 — столкновений 300, 386, 392
 — эпициклическая 514, 526
 Чередование 322
 Шаровые звездные скопления 497
 Эйнштейна—де Ситтера вселенная 534, 545
 Эйнштейна космологическая постоянная 532, 548
 — соотношение 435
 Электромагнитные коды 378
 Электронная жидкость 401
 — плазма 212
 Электростатическое приближение 43, 373
 Энергия решетки 574
 Энтропия плавления 596
 Эпитаксиальный слой 430
 Эталонная сила 328
 Эффективная масса 441
 Якоби метод 219
 Q-минимизация 378, 389. *См. также* Оптимизация

Оглавление

Предисловие редакторов перевода	5
Предисловие	7
От авторов	10
Глава 1. Вычислительные эксперименты с использованием моделей частиц	14
1.1. Введение	14
1.2. Вычислительный эксперимент	15
1.2.1. Роль вычислительного эксперимента	16
1.2.2. Постановка вычислительного эксперимента	18
1.3. Пространственные и временные масштабы	19
1.4. Физические системы	24
1.4.1. Системы с корреляциями	26
1.4.2. Системы без корреляций (бесстолкновительные)	29
1.4.3. Системы со столкновениями	31
1.4.4. Системы с сильными столкновениями	32
1.5. Модели частиц	36
1.5.1. Метод частица — частица	36
1.5.2. Метод частица — сетка	38
1.5.3. Метод частица-частица — частица-сетка	40
Глава 2. Одномерная модель плазмы	42
2.1. Физическая система	42
2.2. Построение дискретной математической модели	44
2.2.1. Уравнения движения макрочастиц	45
2.2.2. Уравнения для полей	47
2.2.3. Раздача заряда и интерполяция силы	49
2.2.4. Дискретная модель	50
2.3. Численные алгоритмы	52
2.3.1. Безразмерные единицы	52
2.3.2. Раздача заряда	53
2.3.3. Уравнение Пуассона	54
2.4. Вычислительные эксперименты	57
2.4.1. Двухчастичный тест	58
2.4.2. Дисперсия волн	59
2.4.3. Холодная плазма	59
2.4.4. Сохранение энергии	61
2.4.5. Двухпоточковая неустойчивость	61
Глава 3. Программа моделирования	64
3.1. Введение	64
3.2. Требования к программе и спецификации	65

3.2.1. Требования пользователя	65
3.2.2 Программные спецификации	67
3.3. Система программирования OLYMPUS	70
3.4. Программа ES1DIV	82
3.4.1 Управляющая структура программы	82
3.4.2 Главный указатель	84
3.4.3 Класс 1: подпрограммы пролога	92
3.4.4 Подпрограммы счета и вывода результатов	116
3.5. Заключительные замечания	118
Глава 4. Схемы интегрирования по времени	121
4.1. Введение	121
4.2. Согласованность	122
4.3. Точность	124
4.4. Устойчивость	125
4.4.1. Метод корневого годографа	128
4.4.2 Матрица перехода	133
4.5. Эффективность	135
4.6. Схема с перешагиванием для гармонического осциллятора	137
4.7. Примеры схем интегрирования	141
4.7.1. Интегрирование силы Лоренца	141
4.7.2 Схемы интегрирования уравнений с вязкими силами	144
4.7.3. Экономичные по памяти схемы Рунге — Кутты	147
Глава 5. Вычисление силы в методе частица — сетка	153
5.1. Введение	153
5.2. Силы в одномерном случае	153
5.2.1. Непрерывная система	153
5.2.2. Схема NGP	155
5.2.3. Схема CIC	157
5.2.4 Смешанные схемы	160
5.3. Иерархия схем распределения заряда	161
5.3.1 Условия на больших расстояниях	162
5.3.2 Условия гладкости	169
5.3.3 Условие сохранения импульса	175
5.3.4 Форма облака и форма функции распределения заряда	177
5.4. Погрешности аппроксимации	182
5.5. Схемы, сохраняющие энергию	184
5.6. Спектральный метод	188
5.6.1. Распределение заряда	188
5.6.2 Нахождение потенциала	198
5.6.3 Интерполяция силы	200
5.6.4 Сила парного взаимодействия частиц	202
Глава 6. Решение уравнений поля	205
6.1 Введение	205
6.1.1 Выбор метода	207
6.2. Нелинейные задачи	211
6.2.1 Итерации по Ньютону	211
6.3. Сеточная релаксация	215
6.3.1 Метод Якоби (J)	219
6.3.2. Метод Гаусса — Зайделя (GS)	220
6.3.3 Последовательная верхняя релаксация (SOR)	221
6.3.4. Ускорение по Чебышёву	222
6.3.5. Блочные методы	223
6.3.6. Неявный метод переменных направлений (ADI)	224
6.4. Матричные методы	227

6.4.1.	Трёхдиагональный алгоритм Томаса	227
6.4.2	Алгоритм сопряженных градиентов (CGA)	228
6.4.3	Методы исключения для разреженных матриц (SM)	232
6.4.4	Неполное разложение	238
6.4.5	Строго неявная процедура Стоуна (SIP)	239
6.4.6	Метод неполных сопряженных градиентов Холецкого (ICCG)	243
6.5.	Быстрое решение эллиптических уравнений (RES)	245
6.5.1.	Циклическая редукция (CR)	247
6.5.2.	Многомерное преобразование Фурье (MFT)	251
6.5.3.	Метод FACR	255
6.5.4.	Методы свертки	259
6.5.5.	Алгоритм Джеймса	263
6.5.6.	Метод матрицы емкости	264
6.5.7.	Итерация по Конкусу и Голубу	269
6.6.	Заключительные замечания	272
Глава 7. Бесстолкновительные модели частиц		273
7.1.	Введение	273
7.2.	Кинетические уравнения	274
7.2.1.	Предел малого шага по времени	275
7.2.2.	Конечный шаг по времени	276
7.3.	Дисперсионное соотношение	277
7.3.1	Предел малого шага по времени	277
7.3.2.	Конечный шаг по времени	281
7.3.3.	Приближение тепловой плазмы	283
7.3.4.	Взаимодействие мод	289
7.4.	Многомерные ограниченные системы	295
7.4.1	Периодичность	295
7.4.2.	Двумерные и трехмерные системы	296
7.5.	Столкновения	298
7.6.	Законы сохранения	301
7.6.1.	Энергия	302
7.6.2.	Импульс	309
7.6.3	Момент количества движения	310
7.7.	Оптимизация	311
7.7.1	Сила взаимодействия частиц	313
7.7.2	Одномерные схемы	314
7.8.	Чередование	318
7.8.1	Усреднение силы	321
7.8.2.	Усреднение гармоник	322
7.8.3.	Многомерные схемы	324
Глава 8. Алгоритмы частица-частица — частица-сетка (P^3m)		326
8.1.	Введение	326
8.2.	Расщепление силы	328
8.3.	Сеточная сила	330
8.3.1	Распределение заряда	331
8.3.2.	Вычисление силы	332
8.3.3	Погрешности в силе	333
8.4.	Короткодействующая сила	337
8.4.1	Цепочечная сетка	337
8.4.2	Связные списки	339
8.4.3.	Изменение импульса	340
8.5.	Временное уравнение	342
8.6.	Оптимизация	345
8.6.1.	Вычисление точности силы	345

8.6.2	Сравнение схем	348
8.6.3	Соотношение цена — качество	352
8.7.	Практические соображения	356
8.7.1.	Программа	356
8.7.2	Организация данных	357
8.7.3.	Распределение заряда и интерполяция силы	359
8.7.4.	Нахождение потенциала	361
8.7.5.	Короткодействующая сила	364
8.7.6.	Выбор параметров	366
Глава 9. Моделирование плазмы		370
9.1.	Введение	370
9.1.1.	Магнитная гидродинамика	371
9.1.2.	Плазма в электростатическом поле	373
9.1.3.	Исторический обзор	375
9.2.	Двумерная электростатическая модель	380
9.2.1.	Время столкновения	381
9.2.2	Время нагрева	383
9.2.3.	Эмпирические корреляции	386
9.3.	Аномальная диффузия	391
9.3.1	Диффузионный эксперимент	393
9.3.2.	Простая количественная теория	399
9.3.3.	Выбор временного шага и размера ячейки	404
9.3.4.	Обсуждение эксперимента	407
9.3.5.	Модели с размерностью $2^{1/2}$ и 3D	408
9.3.6.	Диагностика и визуализация	410
9.4.	Магнитосфера	414
9.4.1.	Магнитогидродинамическая модель частиц	416
9.4.2	Глобальная модель магнитосферы	418
9.4.3.	Модель частиц Ампера	420
9.4.4.	Геоманитный хвост	424
Глава 10. Моделирование полупроводниковых приборов		427
10.1.	Введение	427
10.1.1.	Цель моделирования	427
10.1.2.	Постановка задачи	428
10.1.3.	Типы моделей	433
10.2.	Перенос электронов в полупроводниках	440
10.2.1.	Уравнения движения	440
10.2.2.	Зонная структура арсенида галлия	443
10.2.3.	Процессы рассеяния	445
10.2.4.	Подвижность	448
10.2.5.	Релаксационные эффекты в переходных процессах	450
10.3.	Разработка вычислительной модели	452
10.3.1.	Расчеты по методу частица — сетка	452
10.3.2.	Выбор рассеяния с помощью процедуры Монте-Карло	456
10.3.3.	Модифицированный цикл временного шага	461
10.4.	Моделирование полевых транзисторов (ПТ)	464
10.4.1.	Статические характеристики	465
10.4.2.	Наблюдение внутренних процессов ПТ	472
10.4.3.	Динамические характеристики	475
10.4.4.	Эквивалентная цепь с сосредоточенными параметрами	480
10.4.5.	Шумы	483
10.4.6.	Охлаждаемый ПТ	489
10.4.7.	Сложные геометрии	492

Глава 11. Астрофизика	494
11.1. Введение	494
11.1.1. Эволюция звезд	494
11.1.2. Гравитационная задача N тел	494
11.1.3. Столкновительные и бесстолкновительные системы	495
11.1.4. Скучивание звезд и галактик	496
11.1.5. Большой взрыв	498
11.1.6. Численное моделирование звездных систем	501
11.2. Малые скопления	504
11.2.1. Закон взаимодействия	504
11.2.2. Интегрирование по времени	505
11.3. Спиральные галактики	509
11.3.1. Теории спиральной структуры	509
11.3.2. Модель	510
11.3.3. Выбор временного шага и размера сетки	513
11.3.4. Время столкновения и число частиц	516
11.3.5. Универсальная барообразная неустойчивость	519
11.3.6. Условия устойчивости спиральной структуры	522
11.3.7. Протогалактика	530
11.4. Скучивание галактик	532
11.4.1. Уравнения расширяющейся Вселенной	532
11.4.2. Сопутствующие координаты	535
11.4.3. Вычислительная модель	537
11.4.4. Результаты и выводы	539
Глава 12. Твердое тело, жидкости и фазовые превращения	549
12.1. Введение	549
12.1.1. Молекулярная динамика	549
12.1.2. Закон взаимодействия	550
12.1.3. Интегрирование по времени	554
12.2. Двумерная электронная пленка	555
12.2.1. Безразмерные уравнения	555
12.2.2. Выбор шага по времени	557
12.2.3. Характерные масштабы задачи	561
12.2.4. Машинное время и память	563
12.2.5. Плавление электронной пленки	564
12.3. Ионные микрокристаллы	569
12.3.1. Безразмерные уравнения	569
12.3.2. Выбор шага по времени	572
12.3.3. Пространственная сетка и экономия машинного времени	579
12.3.4. Измерения термодинамических характеристик	580
12.3.5. Измерения с разными областями	586
12.3.6. Плавление, переохлаждение и стеклование	588
12.3.7. Влияние отношения радиусов, жесткости и размера	593
12.3.8. Проверка теорий плавления	597
Приложение. Преобразование Фурье, ряд Фурье и дискретное преобразование Фурье	600
П. 1. Определение преобразований	600
П. 1.1. Преобразование Фурье (FT)	600
П. 1.2. Ряд Фурье (FS)	601
П. 1.3. Дискретное преобразование Фурье (FFT)	601
П. 2. Свойства симметрии	603
П. 3. Теоремы	603
П. 4. Специальные функции	606
П. 4.1. Функция прямоугольного импульса $\Pi(x)$	606

П. 4.2. Функция треугольного импульса $\Lambda(x)$	607
П. 4.3. Гауссоида $G(x)$	607
П. 4.4. Дельта-функция Дирака $\delta(x)$	607
П. 4.5. Функция выборки $\Pi(x)$	607
П. 5. Связь между преобразованиями	608
П. 6. Многомерные преобразования	609
Литература	610
Предметный указатель	629

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».