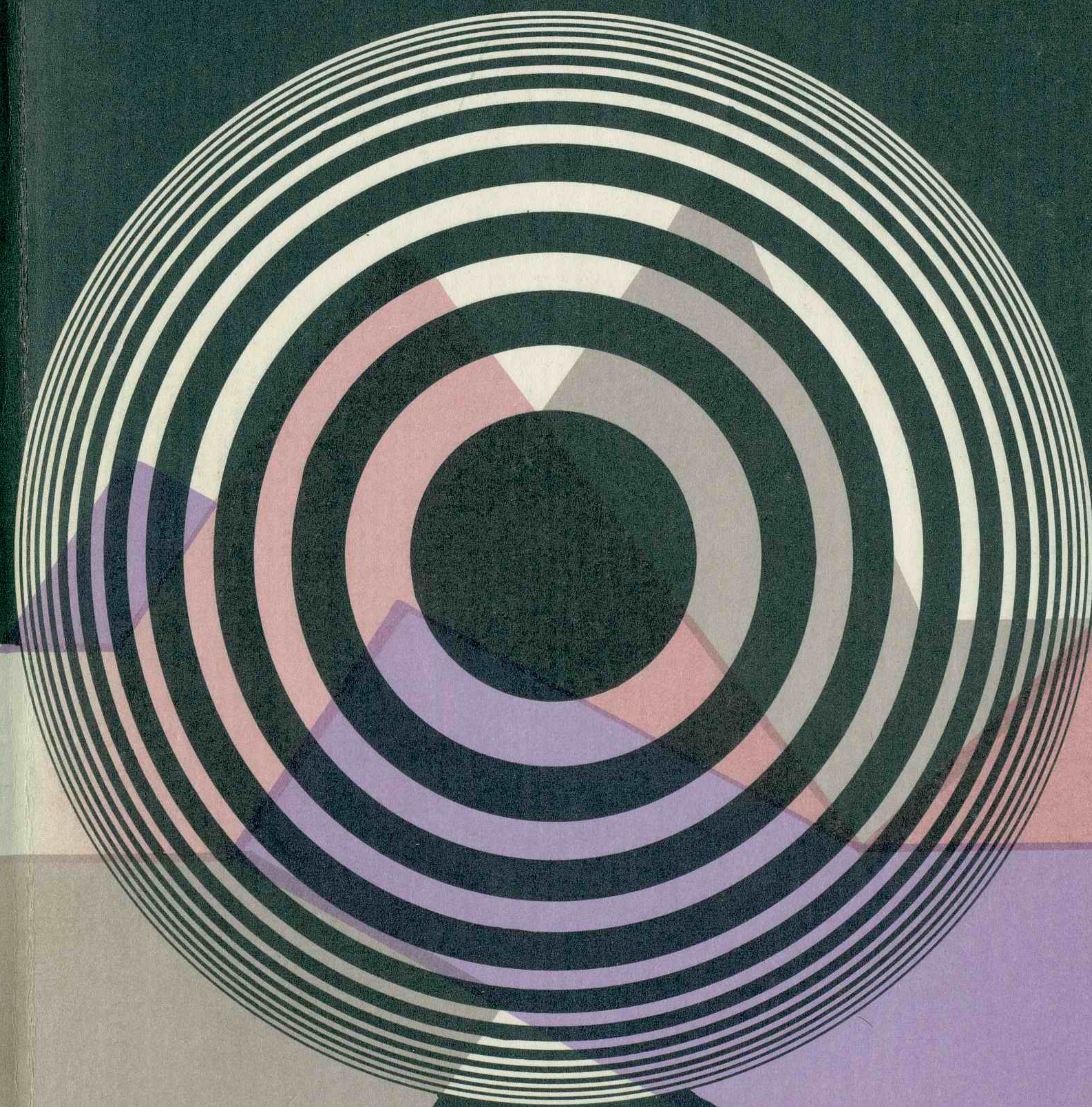


1 р. 70 к.

Эта книга предназначена для читателей, желающих получить представление о принципах Фурье, используемых в современной оптике, о той роли, которую они играют в других областях науки и техники. Преобразования Фурье, связанные с операциями свертки и корреляции, составляют основу не только формирования и обработки изображений в оптических системах, но также и научных исследований самых различных объектов, начиная с атомной структуры вещества до крупномасштабной структуры Вселенной, в том числе и некоторых современных методов спектроскопии. Они применяются также в теории связи и информации, включая обработку информации, которая не является оптической по своей природе. Дополняя учебники по оптике для студентов старших курсов, эта книга обеспечивает переход к более углубленному изучению предмета в различных специализированных приложениях.

И. Г. СТЮАРД

ВВЕДЕНИЕ В ФУРЬЕ-ОПТИКУ



FOURIER OPTICS: An Introduction

E. G. STEWARD, D. SC., F. INST. P.

Professor of Physics and Molecular Medicine
The City University, London

ELLIS HORWOOD LIMITED
Publishers Chichester

Halsted Press: a division of
JOHN WILEY & SONS
New York · Brisbane · Chichester · Ontario

И. Г. СТЮАРД

ВВЕДЕНИЕ В ФУРЬЕ-ОПТИКУ

Перевод с английского
Г. Д. КОПЕЛЯНСКОГО
и канд. физ.-мат. наук
В. И. КОСТЕНКО
под редакцией
д-ра физ.-мат. наук
Л. И. МАТВЕЕНКО

ББК 22.34
С 88
УДК 52 + 53

Стюард И. Г.

С 88 Введение в фурье-оптику: Пер. с англ.—М.: Мир, 1985.—182 с., ил.

В книге профессора Лондонского университета изложены основы нового направления, получившего название «фурье-оптика». Методы фурье-оптики, основанные на применении анализа Фурье к проблеме построения и обработки изображения, используются в физике, астрономии, при обработке информации и в других приложениях. Книгу отличают ясность и четкость изложения, многочисленные примеры и иллюстрации из различных областей астрономии и физики облегчают понимание содержания.

Предназначена для астрономов, оптиков, приборостроителей, специалистов по томографии, кристаллографии, голографии, в том числе для студентов и аспирантов соответствующих специальностей.

С $\frac{1704050000-138}{041(01)-85}$ 66-85, ч. 1

ББК 22.34
535

*Редакция литературы по космическим исследованиям,
астрономии и геофизике*

© 1983 E. G. Steward/Ellis Horwood Limited
© Перевод на русский язык, «Мир», 1985

От редактора перевода

Использование физической оптики в самых различных приложениях — от астрономии до кристаллографии — сводится к построению и обработке изображения путем анализа спектрального и пространственного распределения излучения. Для такой обработки может быть использован фурье-анализ, основанный на применении преобразования Фурье к полученной информации. Это обстоятельство было осознано достаточно давно, но только появление современной мощной вычислительной техники позволило сформироваться новому направлению, получившему название «фурье-оптика». Изложению основ этого нового направления и посвящена предлагаемая вниманию советского читателя книга профессора Лондонского университета И. Г. Стюарда.

Автор начинает с рассмотрения принципов построения и анализа изображений, полученных посредством оптических устройств, при помощи преобразований Фурье. Изложение ведется на достаточно простом математическом и физическом уровне, но в то же время постоянно подчеркивается, что излагаемые методы построения и анализа изображений имеют общее значение и применимы к информации, которая не является оптической по своей природе. Особенности изложения и тематика отдельных глав подробно освещены автором в предисловии, и потому нет необходимости их здесь касаться. Отметим лишь, что многие из проблем, затронутых в книге, активно разрабатываются и в нашей стране. В частности, метод сверхдальней радиоинтерферометрии был предложен в Советском Союзе (Л. И. Матвеев, Н. С. Кардашев, Г. Б. Шоломицкий. Известия ВУЗ, радиофизика, 8, 651, 1965). Эта книга поможет как студенту, так и начинающему исследователю познакомиться с основами фурье-оптики и облегчит переход к более углубленному изучению предмета в различных приложениях, будь то астрономия, радиоастрономия, физика или биология.

При переводе мы стремились передать своеобразие «лекционного» стиля автора и используемой им терминологии; в необходимых случаях дополнительные разъяснения сделаны в виде примечаний редактора.

В заключение отметим, что книга, несомненно, представит интерес для широкого круга специалистов, занимающихся проблемами

обработки данных наблюдений, исправления данных за аппаратную функцию, выделения образа на фоне шумов, представления образа в удобном для последующей работы виде. В частности, книга окажется полезной и для специалистов, занимающихся томографией, изучением периодических структур в биологии и кристаллографии, в радиолокационных исследованиях, при проектировании многоэлементных оптических и радиоастрономических систем. Она послужит полезным учебным пособием для студентов старших курсов и аспирантов соответствующих специальностей.

Л. И. Матвеев

Апрель 1984 г.

Предисловие

Эта книга предназначена для читателей, желающих получить представление о принципах Фурье, используемых в современной оптике, о той роли, которую они играют в других областях науки и техники. Преобразования Фурье, связанные с операциями свертки и корреляции, составляют основу не только формирования и обработки изображений в оптических системах, но также и научных исследований самых различных объектов, начиная с атомной структуры вещества до крупномасштабной структуры Вселенной, в том числе и некоторых современных методов спектроскопии. Они применяются также в теории связи и информации, включая обработку информации, которая не является оптической по своей природе. Дополняя учебники по оптике для студентов старших курсов, эта книга обеспечивает переход к более углубленному изучению предмета в различных специализированных приложениях.

Некогерентное формирование изображения описано, исходя из понятий свертки и передаточной функции, причем особое внимание уделяется свойствам линейности и инвариантности, которые присущи многим типам электрических цепей (нелинейные цепи в данной книге не затрагиваются). Процесс когерентного формирования изображения на основе двойного преобразования Фурье иллюстрируется его специфическим применением в рентгеновской кристаллографии.

В качестве введения в обширную сферу приложений оптической фильтрации и обработки изображений изложены основные идеи в области амплитудной, фазовой и голографической фильтрации, иллюстрированные примерами из оптической и электронной микроскопии и быстро развивающейся области распознавания образов. Также кратко описана обработка на основе корреляции спектров мощности и геометрической оптики.

Исторические сведения приводятся в тех случаях, когда они помогают уяснить научный смысл и иллюстрируют прогресс научных исследований. Такие отступления необходимы, ибо для понимания существующего положения дел и характера прогресса в данной области, следует иметь представление о ее истоках. Например, в гл. 6 объясняется, каким образом развитие современных методов интерферометрии в оптической и радиоастрономии связано с исходными принципами звездного и спектрального интерферометров Майкельсона.

Описание методов, предложенных Майкельсоном, который предвидел их огромные потенциальные возможности, служит отправной точкой для понимания сути подобных методов, применяемых в наше время.

С целью стимулировать читателя и помочь ему разобраться во всем многообразии литературы по обсуждаемым вопросам помимо дополнительной библиографии в тексте приводятся ссылки.

Предполагается, что читатель уже знаком с математикой в объеме вводного университетского курса. Все дополнительные сведения даются в тексте по мере их использования. В основном изложение ограничено одномерным случаем, чтобы обеспечить по возможности быстрое восприятие математических формул.

С удовольствием выражаю свою искреннюю признательность д-ру Р.С. Лонгхерсту за полезные беседы и предложения, высказанные на разных стадиях подготовки конечного варианта рукописи. Его опыт автора основополагающего вводного курса оптики оказался неоценимым. Я признателен также проф. Генри Липсону, который прочитал и прокомментировал гл. 1, 2 и 5 и служил для меня источником поддержки на протяжении многих лет, а также проф. Мартину Райлу за полезные советы по гл. 6, касающиеся астрономии. Хотелось бы поблагодарить также моего коллегу проф. М.А. Джасвона (декана факультета математики Лондонского университета) за полезные замечания, касающиеся некоторых математических формул.

Я благодарен Роджеру Меггсу (физический факультет) за помощь в подготовке рисунков и Питеру Кловеру (группа звуковизуальных средств) за подготовку графиков.

И. Г. Стюард

Февраль 1983 г.

Глава 1

Вступление

1.1. ВВЕДЕНИЕ

Данная книга посвящена прежде всего фурье-принципам в двух взаимосвязанных областях физической оптики: 1) формирование и обработка изображений и 2) изучение пространственного распределения и спектрального состава источников излучения. Они объединены вместе, поскольку во многих отношениях взаимосвязаны, а когда такая связь существует, совместное изучение предметов, как правило, обеспечивает определенные преимущества. Разнообразные применения этих принципов, упомянутые в предисловии, будут обсуждаться в процессе изложения.

Читая эти страницы, мы можем рассматривать обе указанные темы в контексте связи с видимым светом, но вместе с тем мы будем обращаться к некоторым важным приложениям в других областях электромагнитного спектра. Обсуждая формирование изображений, мы будем переходить из оптической области к методу использования рентгеновских лучей для расшифровки атомной структуры вещества, а в другом предельном случае к астрономии и структуре Вселенной. Что же касается спектроскопии, то здесь методы Фурье применяются теперь в широком диапазоне частот.

Математические методы, получившие свое название в честь Ж.Б. Фурье, чрезвычайно эффективны в указанных областях. Они вводятся в гл. 3 и 4 и широко применяются в заключительных главах. Описанная в учебниках общей физики *скалярная аппроксимация* полностью пригодна для математического представления света и использована по всей книге. В приложении А дается сводка обозначений и основных уравнений, напоминается смысл таких терминов, как разность путей и разность фаз, и рассматривается использование фазовых диаграмм для суммирования волн с различными амплитудами и фазами.

В данной главе мы на общепринятом языке определим некоторые физические процессы, характерные для указанных двух главных областей оптики. Эти определения очень удобно ввести на основе анализа опыта Юнга и явлений типа колец Ньютона.

Читатель, по-видимому, уже знает, что исторический эксперимент, выполненный Томасом Юнгом в 1801 г., послужил решительным до-

и определяемая выражением

$$\delta = \frac{2\pi}{\lambda} l.$$

Интенсивность освещенности¹⁾ в Р равна

$$I = R^2 = A_1^2 + A_2^2 + 2A_1A_2 \cos \delta = A_1^2 + A_2^2 + 2A_1A_2 \cos \frac{2\pi l}{\lambda} \quad (1.01)$$

Для сохранения общности мы не будем считать, что $A_1 = A_2$, хотя в данном случае это равенство имеет место.

Третий член является «интерференционным членом» и позволяет вычислить положение максимумов и минимумов интерференционных полос. В точках Р, для которых $l = 0, \lambda, 2\lambda, \dots$, дифрагировавшие на двух апертурах волновые фронты находятся в фазе, и мы получаем максимум интенсивности

$$I_{\max} = A_1^2 + A_2^2 + 2A_1A_2 = (A_1 + A_2)^2. \quad (1.02)$$

Между ними при $l = \lambda/2, 3\lambda/2, 5\lambda/2, \dots$ волновые фронты имеют противоположные фазы («противофазны») и на эти точки приходится минимум интенсивности

$$I_{\min} = (A_1 - A_2)^2. \quad (1.03)$$

(С увеличением расстояния между точками Р и О интенсивность картины все более падает, так как амплитуды вторичных волн уменьшаются при увеличении угла. В противном случае существовал бы волновой фронт, распространяющийся в обратном направлении. Для учета этого обстоятельства в количественных соотношениях для распределений амплитуд интенсивностей вводится коэффициент наклона (наклона), явно входящий в анализ Кирхгофа. В них также должен присутствовать коэффициент, учитывающий закон *обратной квадратичной зависимости*, поскольку расстояние между точкой Р и апертурами меняется с изменением положения точки Р на экране. Следуя общепринятой практике, в данной книге эти поправочные коэффициенты исключены из уравнений.)

В центральной области расстояние между полосами Δ (см. рис. 1.1, а) приблизительно можно определить, учитывая, что для первой полосы $\Delta \approx L\theta_1$, $D\theta_1 \approx l = \lambda$,

следовательно,

$$\Delta \approx \frac{L\lambda}{D}. \quad (1.04)$$

¹⁾ Поток (мощность), падающий на единицу площади. Другие названия: *освещенность*, *блеск*.

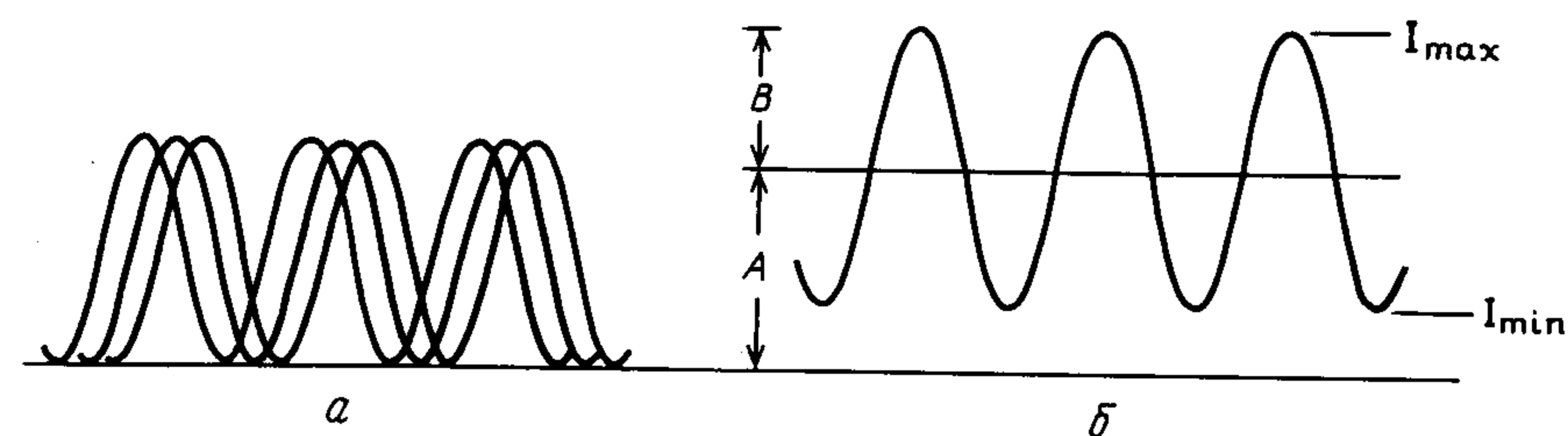


Рис. 1.3. Размытие полос дает картину с тем же периодом, но меньшей видностью.

Важно отметить, что расстояние между полосами пропорционально λ и обратно пропорционально D , а также, что реальная интенсивность в любой точке картины определяется амплитудами света, приходящего от каждого точечного отверстия и характера интерференции двух составляющих.

Существует несколько аспектов, в которых опыт Юнга иллюстрирует типы физических явлений, связанных с двумя главными темами этой книги, упомянутыми в начале данного раздела. Рассмотрим вопрос, относящийся к спектральному и пространственному распределению источников излучения. Чтобы полосы в опыте Юнга имели хорошую «видность» — достаточную четкость, — для освещения апертур важно использовать весьма малые источники. Группы полос, полученных от пространственно разнесенных точек, не очень малых источников, смещены одна относительно другой, так что результирующая интерференционная картина имеет низкую видность.

Вопросы видности полос являются основой этой книги, и нам необходимо показать, как аналитически просто определяется видность. Рис. 1.3 иллюстрирует пример явления, которое мы только что описали. На рис. 1.3, а изображено смещение лепестков в малой области. Хотя здесь представлены только три отдельные интерференционные картины, в действительности обычно наблюдается непрерывное распределение в некоторой области (конкретные детали будут зависеть от размера и распределения интенсивности в источнике). Окончательный результат показан на рис. 1.3, б, а видность определяется как

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{B}{A} \quad (1.05)$$

Другие названия этого параметра — *модуляция* или *контраст*. (Заметим, однако, что «контраст» определяется иногда как $(I_{\max} - I_{\min})/I_{\min}$.)

Если нужно получить полосы с хорошей видностью, то помимо источника малых размеров освещенность в опыте Юнга должна обладать свойством, которое иногда неправильно называется «монокromaticностью».

Таким образом, видность полос содержит в себе информацию как о пространственной, так и спектральной природе источника S . Получение этой информации по эффектам интерференции и составит главную часть нашего анализа и явится основой содержания гл. 6.

Переходя к другому важному вопросу, который касается формирования изображения, заметим, что в этом случае излучение используется как средство передачи информации об объекте в то место, где происходит распределение излучения, формирующее изображение объекта. С помощью подходящих линз свет, испущенный в опыте Юнга двумя точечными отверстиями, может быть использован для получения изображения, которое выглядит как сами точечные отверстия. (Подобным же образом при наблюдении точечных отверстий непосредственно глазом хрусталик воспроизводит изображение на сетчатке задней стенки глаза.) Как работает линза? Почему так происходит, что от протяженного источника S полосы на экране не видны, однако введение линзы позволяет получить их изображение?

Ответы на эти вопросы и общие аспекты затронутых выше взаимосвязей между видностью полос и природой источника касаются не только дифракции и интерференции, но также «когерентности» излучения, на которую мы теперь должны обратить наше внимание.

1.2. КОГЕРЕНТНОСТЬ И ИСТОЧНИКИ СВЕТА

Для наблюдения интерференционных эффектов на экране в опыте Юнга необходимо, чтобы пути волн света, приходящие к нему от двух апертур (B и C на рис. 1.1, a), налагались и имели одну и ту же частоту и чтобы разность фаз между ними была постоянной. Если бы эти условия выполнялись идеально, то можно было бы сказать, что освещенности апертур когерентны.

Использование общего источника света, S на рисунке, для освещения обоих точечных отверстий является в определенной мере способом обеспечения требований когерентности. Каждая пара волновых фронтов, испущенных в B и C , обусловлена одним волновым фронтом от S . И если S является точечным источником, то все фронты испущенных им волн должны проходить определенные расстояния до B и C , сохраняя постоянную разность фаз между дифрагировавшими фронтами, исходящими из B и C .

Однако обычные источники света несовершенны и дают освещенность, которая когерентна лишь в большей или меньшей мере, т.е. частично когерентна. По самой своей природе излучение *фотонов* (квантов света) атомами означает, что каждый волновой пучок связан с фотоном, излученным за конечное время, и влияет на так называемую *временную когерентность* излучения. Более того, поскольку каждый реальный источник имеет конечный размер, пути волн, испущенные в пространственно разнесенных точках, влияют на так на-

зываемую *пространственную когерентность* поля освещенности, определяемую источником. В последующих разделах оба эти свойства когерентности рассматриваются в вводном плане в контексте опыта Юнга.

1.2.1. Временная когерентность

Конечная продолжительность излучения атомом отдельного волнового пучка света означает, что он не может быть бесконечно длинным (мы проанализируем это более подробно в разд. 4.6). В результате он занимает некоторую (хотя и узкую) область частот, т.е. имеет «полосу частот». Даже свет лазера обладает конечной полосой частот, хотя и предельно узкой, с соответствующей длиной пучков в несколько десятков километров. В типичных нелазерных источниках, называемых обычно *тепловыми источниками*, тепловые колебания излучающих атомов наряду с другими эффектами ухудшают когерентность света и ограничивают время, в течение которого волновой пучок можно рассматривать как аппроксимацию простого гармонического колебания. По этим причинам *монохроматический* свет от таких источников, как газоразрядные трубки, более правильно называть *квазимонохроматическим*. *Белый свет* является полной противоположностью лазерному и имеет столь короткие волновые пучки, что его нельзя отождествить ни с одной определенной частотой.

Указанное свойство света от любого источника связано с его временной когерентностью, которую можно качественно определить как интервал времени, в течение которого фаза волнового движения последовательно изменяется предсказуемым образом при прохождении фиксированной точки в пространстве; чем больше этот интервал, тем выше степень временной когерентности. Длина волнового пучка, удовлетворяющего указанному требованию, представляет собой *длину когерентности*, которая равна произведению времени когерентности на скорость света. (При использовании термина «длина когерентности» следует помнить, что имеется в виду спектральная частота света, а не какое-то свойство его пространственного распределения.)

Если даже предположить, что источник S на рис. 1.1, a действительно является точечным, конечная длина когерентности означает, что для точек P , достаточно удаленных от оси, длину l можно принять сравнимой с длиной волнового пучка. Волновые пучки, которые одновременно исходят из B и C (и своим происхождением обязаны одному волновому пучку, испущенному S), не должны тогда полностью накладываться в P и, следовательно, видность полос должна снижаться. Еще дальше вдоль экрана полосы должны полностью исчезать, образуя непрерывный уровень освещенности, обусловленный отдельными независимыми вкладами от двух апертур.

Как мы отметили в разд. 1.1, конечный размер источника в опыте Юнга также приводит к снижению видности полос и мы рассмот-

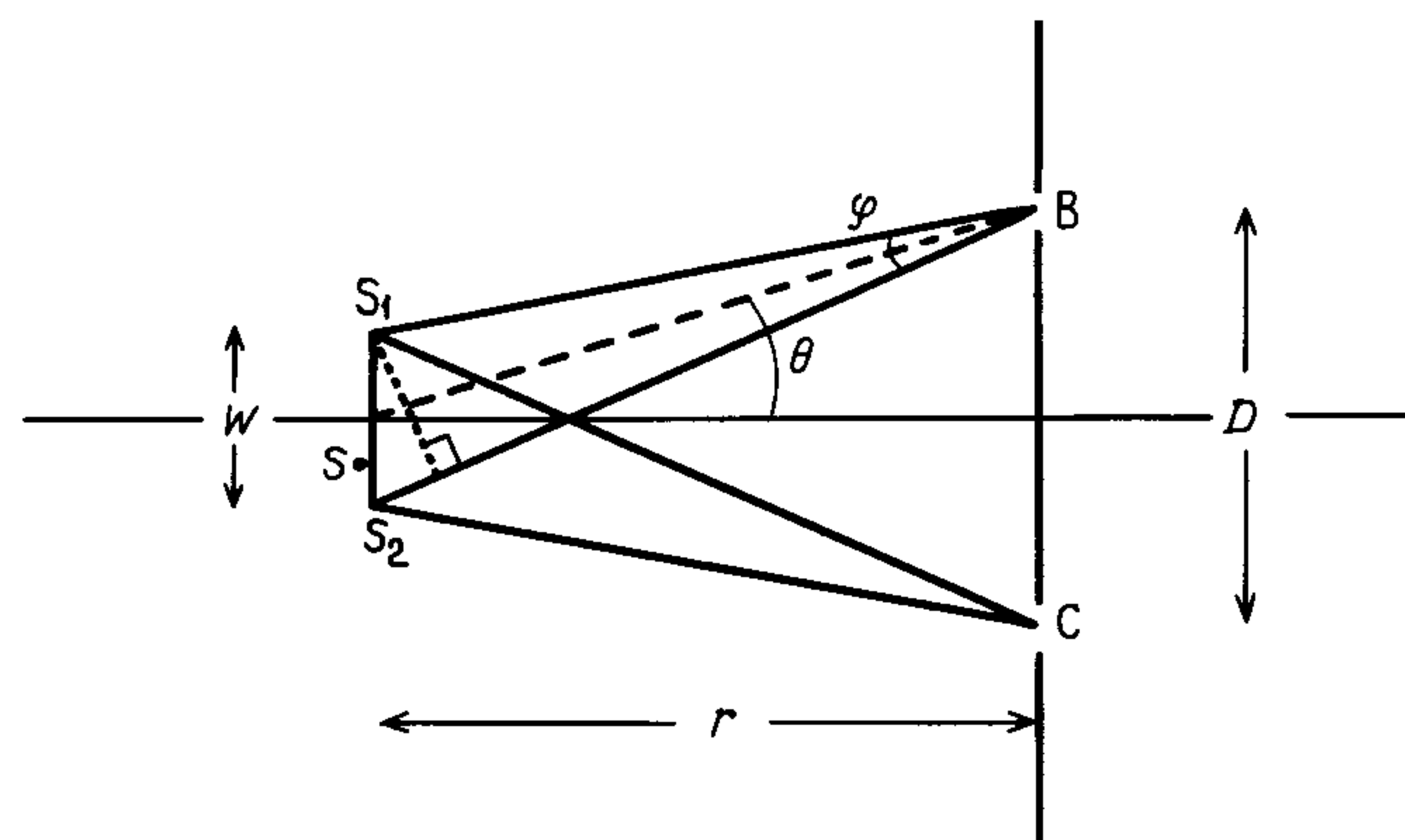


Рис. 1.4. Пространственная когерентность. Влияние размера источника в опыте Юнга.

рим этот эффект в следующем разделе. (В разд. 1.4 описана схема, более чувствительная к временной когерентности, а в разд. 6.4.2 обсуждается характер связи видности и временной когерентности.)

1.2.2. Пространственная когерентность

Процессы излучения в различных точках обычного теплового источника света не зависят один от другого, и в этом смысле такие источники можно считать некогерентными. Но мы интересуемся не столько природой самого источника, сколько свойствами формируемого им поля освещенности, например в плоскости на некотором расстоянии от источника. Таким образом, в опыте Юнга нас интересует масштаб, на котором между В и С имеется постоянное фазовое соотношение (рис. 1.1, а) и где можно наблюдать эффекты интерференции. Мы уже отмечали влияние ограниченной временной когерентности, связанное с конечной полосой частот света, излучаемого источником. Какое же влияние оказывает конечный размер источника?

На рис. 1.4 схематично и в сильно увеличенном масштабе показаны две апертуры В и С в опыте Юнга с источником света шириной W , находящимся на расстоянии r . Предположим, что свет имеет неограниченную длину когерентности. Свет от некоторой точки S в источнике освещает апертуры, и на экране возникают интерференционные полосы. В зависимости от положения S существует определенная задержка между моментами прихода некоторого волнового фронта на апертуру В и апертуру С. Величина этой задержки определяет положения максимумов и минимумов интенсивности полос на экране (на рисунке не показаны), обусловленных светом, приходящим от точки S . Если бы источник состоял только из этой единственной точки (как в идеализированном опыте Юнга), то наблюдались бы полосы

с максимальной видностью. Однако реальный источник имеет конечный размер, и полосы, обусловленные излучением из других точек, смещены относительно полос, обусловленных излучением из S . Более того, при использовании обычных тепловых источников света вне зависимости от того, как велика временная когерентность, интерференционные картины, обусловленные светом, испущенным различными точками источника, являются полностью аддитивными по интенсивности (т.е. интерференция между ними отсутствует), поскольку они совершенно не связаны. Поэтому свет от всего источника вызывает размытие полос, как показано на рис. 1.3, с вытекающим отсюда уменьшением видности картины полос.

Таким образом, в общем случае поле освещенности от источника конечных размеров, если он даже полностью монохроматичен, вновь оказывается лишь частично когерентным, но на этот раз пространственно благодаря пространственному распределению яркости источника. Степень же когерентности между двумя точками поля физически проявляется в видности интерференционных полос, образуемых светом от этих двух точек. Фактическое соотношение между пространственной когерентностью и видностью полос рассматривается в разд. 6.4.

В отношении практического применения можно сказать, что высокая степень когерентности на апертурном экране, показанном на рис. 1.4, получается, только когда разность $BS - CS$ изменяется лишь на малую долю длины волны при всех положениях S относительно S_1 и S_2 . Это обстоятельство можно выразить в виде

$$(BS_2 - BS_1) - (CS_2 - CS_1) \ll \lambda.$$

Так как размер W , очевидно, очень мал по сравнению с r , вышеприведенное неравенство может быть приближенно записано в виде

$$2W\theta \ll \lambda.$$

Поскольку $r\theta \approx D/2$ и $r\phi = W$, мы получаем

$$\phi \ll \lambda/D. \quad (1.06)$$

Это соотношение говорит нам о том, что для пространственной когерентности освещенности на точечных отверстиях В и С угол ϕ (в радианах), стягиваемый источником, должен быть существенно меньше, чем λ/D . С другой стороны, мы можем сказать, что в плоскости точечных отверстий освещенность обладает существенной пространственной когерентностью на участке длиной D , составляющем малую долю λ/ϕ , долю, определяемую в соответствии с теми или иными экспериментальными требованиями. Это конкретное значение D называется (поперечной) шириной когерентности для освещенности на рассматриваемой плоскости. При достаточной информации об источнике зона когерентности в плоскости на заданном расстоянии от него может быть определена таким же образом.

Здесь уместно задержаться и постараться понять, что вне зависимости от того, сколь сильно может различаться когерентность поля освещенности от одной точки к другой, такие различия незаметны глазом; глаз чувствителен только к интенсивности.

Возвращаясь к снижению видности полос, вызванному увеличением размера источника, заметим, что этот эффект служит основой измерения с помощью *звездного интерферометра Майкельсона* угловых диаметров звезд, слишком малых для измерений обычным способом на телескопе. Этот метод описывается в гл. 6, где показано также, каким образом изменения видности полос в зависимости от расстояния между двумя апертурами позволяют получить информацию о распределении яркости источника.

Наконец, мы должны заметить, что в отношении пространственной когерентности свет лазерных источников является исключением точно так же, как и для временной когерентности (разд. 1.2.1). Пространственная когерентность сохраняется непосредственно поперек лазерного пучка (причем его ширина для практических задач может быть увеличена без потери когерентности с помощью изготовленной соответствующим образом системы линз — *расширителя пучка*). Очень похожими на свет лазера с точки зрения пространственной и временной когерентности являются *радиоволны*, излучаемые радиопередатчиками.

1.3. ФОРМИРОВАНИЕ ОПТИЧЕСКОГО ИЗОБРАЖЕНИЯ

Мы подошли теперь к тому, чтобы определить и оценить некоторые общие аспекты формирования оптического изображения, прежде чем заняться более детальным их изучением в гл. 5.

В разд. 1.1 мы отметили, что применение линз в опыте Юнга должно обеспечить формирование изображения двух точечных отверстий. В этом случае очевидно, что дифракция на точечных отверстиях является первым шагом в формировании их оптического изображения, а вторым шагом будет «рекомбинация» этого света линзами в изображение.

В разд. 1.1 мы напомнили также, что полосы в опыте Юнга не видны при использовании протяженного источника света. Однако общеизвестно, что линзы тем не менее будут давать изображение апертур.

Отличается ли действие линз в этих двух случаях? Если говорить кратко, то нет, по причинам, которые можно достаточно легко понять из качественных соображений.

Рассмотрим первый случай, когда опыт Юнга выполняется классическим путем, который теперь мы можем описать как случай достаточно высокой когерентной освещенности апертурного экрана. Источник весьма мал, что обеспечивает сохранение одного и того же фазо-

вого соотношения между каждой парой дифрагировавших волновых фронтов, испущенных апертурами В и С. Каждая пара дает вклад в одну и ту же последовательность полос, и положение максимумов и минимумов интенсивности в интерференционной картине можно определить с помощью уравнения (1.01).

Во втором случае с использованием протяженного источника пространственная когерентность освещенности на апертурном экране отсутствует (мы можем предполагать, что временная когерентность имеет место в обоих случаях). Как мы видели, полосы, обусловленные волновыми фронтами из одной точки источника, смещены относительно полос от фронтов из других точек источника. Если источник достаточно большой, то окончательным результатом является пятно на экране со сравнительно равномерной освещенностью. Тем не менее здесь полосы образуются так же, как и в первом случае, но они являются мгновенными, и их положение непрерывно изменяется, давая однородную освещенность даже за минимально возможное короткое время. Из наличия информации об индивидуальных фазовых соотношениях в свете, который будет поступать на внесенные линзы, следует, что по-прежнему существует возможность сформировать изображение апертуры с помощью линз. (При когерентном освещении стабильность фазовых соотношений при некоторых условиях позволяет записать полную информацию, включая фазы, о волновых фронтах от апертурного экрана; в этом состоит основа голографии (разд. 5.4.)

Вышеприведенные замечания о формировании изображений при когерентном освещении (или, сокращенно, о когерентном формировании изображений) объекта в виде точечной маски в опыте Юнга равным образом применимы к 1) более сложным маскам, таким, например, как 35-мм слайды в диапроекторах, 2) непрозрачным объектам, освещаемым обычными тепловыми источниками света, и 3) самосветящимся объектам, которые люминесцируют (например, телевизионное изображение) или нагреты (например, инфракрасная фотография горячих тел). В каждой из этих категорий существуют те же мгновенные фазовые соотношения, какие мы описывали раньше.

В этом широком контексте формирования изображений вместо термина *дифракция* часто используется название *рассеяние*. В литературе существует определенная путаница в этих понятиях, но такие детали не влияют на наши рассуждения, и мы можем считать указанные термины синонимами.

В формировании изображений существуют два аспекта, которые сейчас следует отметить. Один из них связан с образованием дифракционной картины Фраунгофера в задней фокальной плоскости линзы, формирующей изображение, другой — с влиянием конечных размеров апертуры линзы на изображение. (Рассмотрение аберраций в линзах можно отложить до гл. 5.)

1.3.1. Дифракция Фраунгофера

Рассмотрим рис. 1.5, на котором изображена объектная маска с двумя очень малыми апертурными отверстиями В и С, однородно освещенными квазимонохроматическим светом от удаленного источника. Плоские волны поступают по нормали к маске, а сферические волновые фронты расходятся из В и С. Схема такая же, как и в опыте Юнга, за тем исключением, что теперь дополнительно у нас есть линза, которая создает изображение точечных отверстий в плоскости, расположенной, как показано на рисунке. Непосредственный интерес представляет, однако, задняя фокальная плоскость линзы. Рассмотрим любую точку Р, лежащую в направлении под углом θ к оси линзы; в ней складываются вместе и интерферируют только составляющие, распространяющиеся от В и С в направлении θ (сравните с опытом Юнга, где интерференция в точке Р на рис. 1.1 происходит между светом, распространяющимся от апертур в разных направлениях). Мы увидим, что конкретная дифракционная картина (определяемая ниже как фраунгоферовская) в задней фокальной плоскости отображающей линзы является особенно важным промежуточным шагом в формировании изображения, выполняемом линзой. Это позволяет оценить конечную стадию формирования изображения и предоставляет единственную и особую по своей важности возможность для преобразования изображения. Указанное обстоятельство подробно обсуждается в гл. 5, но здесь мы исследуем некоторые свойства картины, сформированной в описанном выше примере. Прежде, однако, отметим, что для экспериментального получения таких дифракционных картин Фраунгофера необходимо обеспечить существование статистических фазовых соотношений, обусловленных когерентным освещением (см. замечания в предыдущем разделе о различиях между когерентным и некогерентным формированием изображения). До гл. 5, где вновь обсуждается эта разница, мы будем (если не указано особо) предполагать, что условия когерентности выполняются.

Для вычисления результата сложения двух составляющих в точке Р на рис. 1.5 необходимо учесть разность фаз δ между ними, обусловленную разностью путей l , так, как мы поступали в опыте Юнга. Мы имеем

$$\delta = \frac{2\pi l}{\lambda}, \quad (1.07)$$

где в данном случае

$$l = D \sin \theta. \quad (1.08)$$

Интенсивность I в точке Р можно вычислить с помощью фазовой диаграммы так же, как на рис. 1.1. Это дает (ср. с уравнением (1.01))

$$I = A_1^2 + A_2^2 + 2A_1A_2 \cos \frac{2\pi l}{\lambda},$$

где l определяется теперь уравнением (1.08).

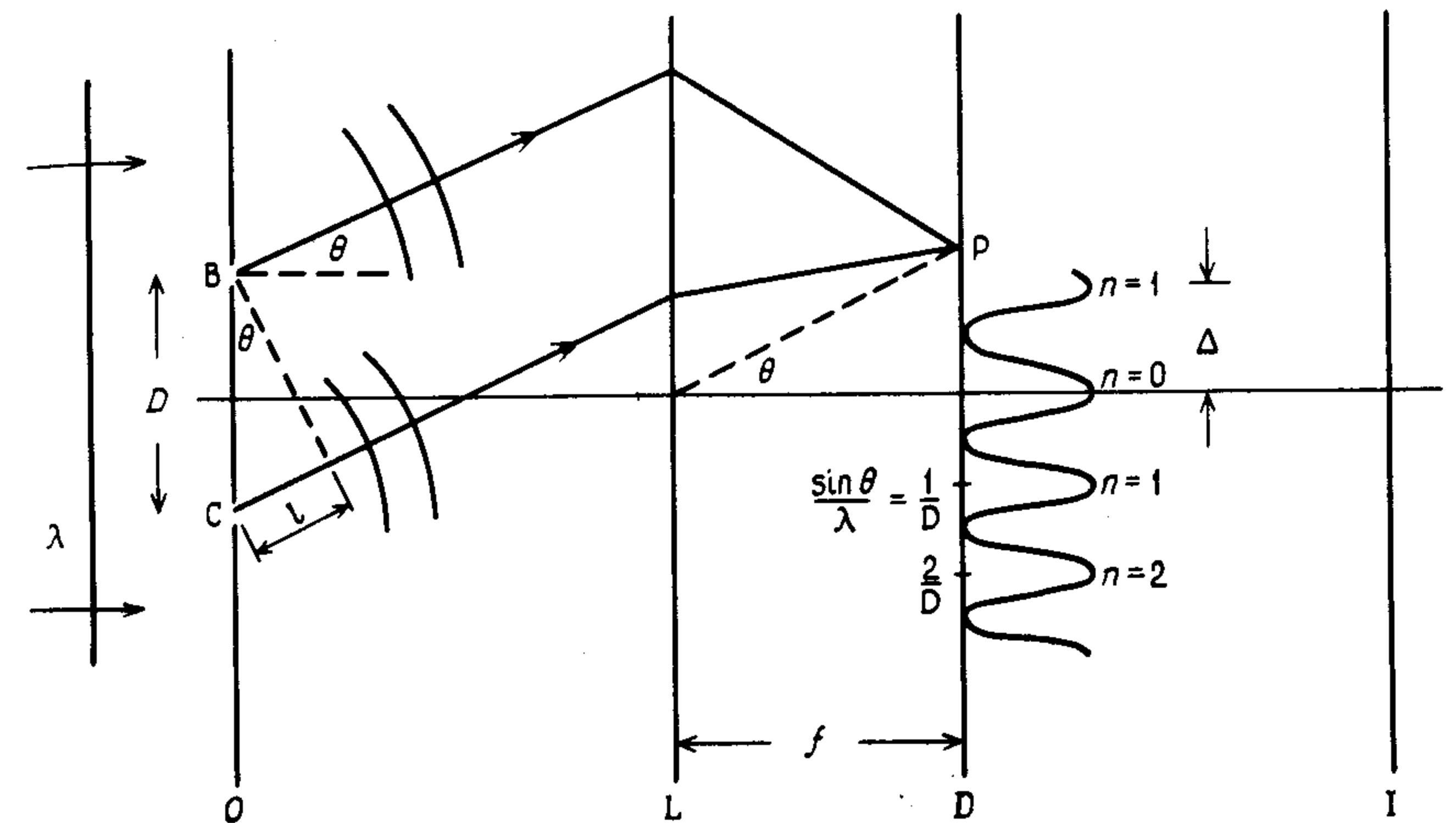


Рис. 1.5. Дифракция как промежуточный шаг в формировании изображения (О—объектная маска, L—линза, D—экран в плоскости дифракции, I—плоскость изображения).

Две амплитуды равны, поскольку обе составляющие приходят под одним и тем же углом. При подстановке $A_1 = A_2 = A$ вышеприведенное уравнение принимает вид

$$I = 2A^2 \left[1 + \cos \left(\frac{2\pi}{\lambda} D \sin \theta \right) \right] = 4A^2 \cos^2 \left(\frac{\pi}{\lambda} D \sin \theta \right). \quad (1.09)$$

Эта картина часто упоминается как содержащая « $\cos^2$ -полосы».

В направлении максимума интенсивности полос освещенность от каждой апертуры изменяется ступенчато с шагом

$$l (= D \sin \theta) = n\lambda, \quad (1.10)$$

где порядок дифракции (или интерференции) n имеет целое или нулевое значение. Тогда интенсивность максимумов равна

$$I_{\max} = 4A^2.$$

Следует вновь отметить, что, хотя направления максимумов определяются D и λ , их значения зависят от амплитуд, излучаемых в этих направлениях отдельными апертурами (см. разд. 2.3).

Между максимумами лежат минимумы, которые равны нулю только в случае, когда (как обычно бывает) $A_1 = A_2$.

Расстояние между полосами Δ легко вычислить с помощью рис. 1.5. Если максимум первого порядка расположен, например, при $\theta = \pm \theta_1$ и если θ_1 мало (масштаб рисунка, конечно, намеренно искажен), мы получаем

$$\Delta \approx f\theta_1, \quad D\theta_1 \approx \lambda,$$

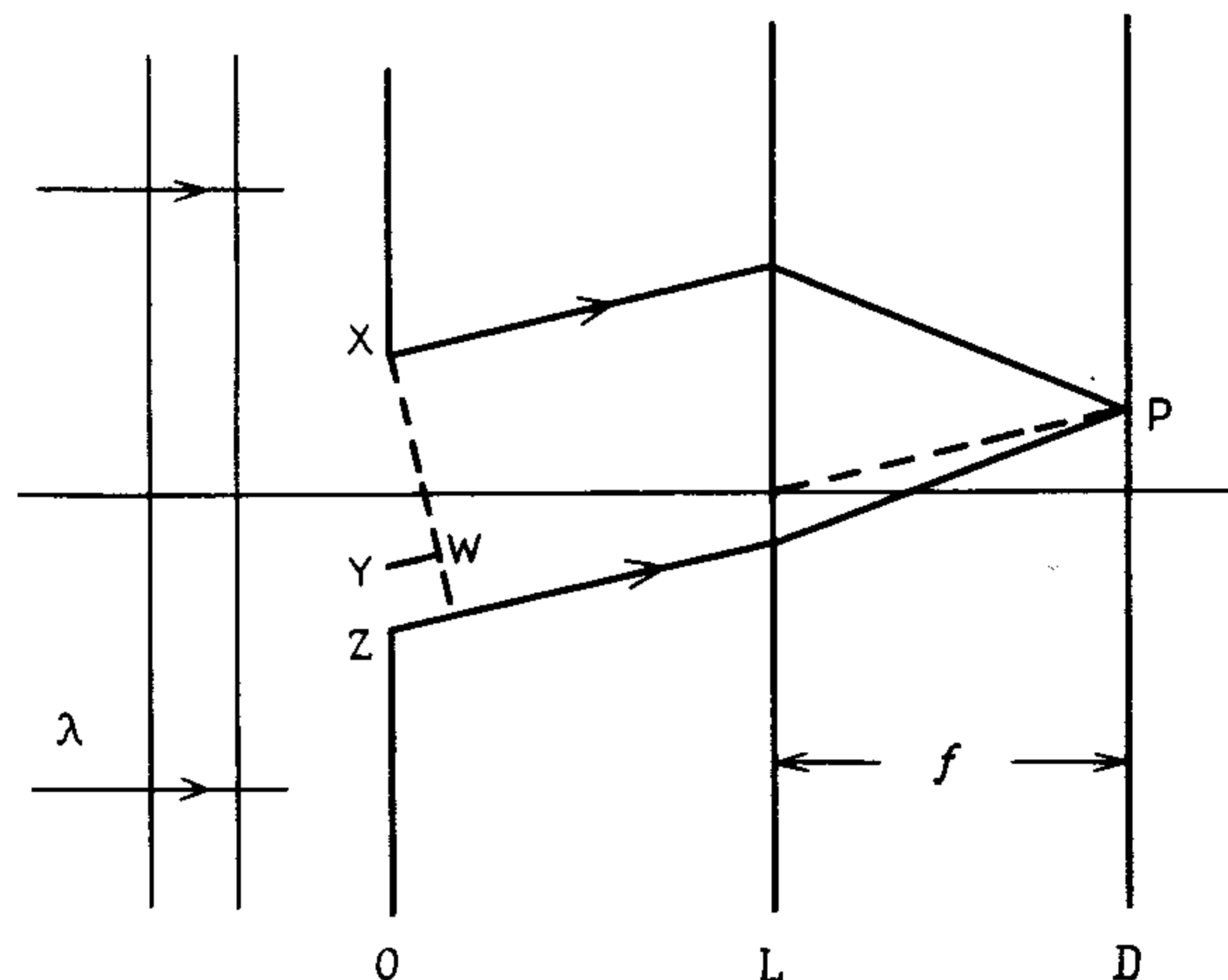


Рис. 1.6. Дифракция Фраунгофера (О – объектная маска, L – линза, D – плоскость дифракции).

откуда

$$\Delta \approx \frac{f\lambda}{D} \quad (1.11)$$

Независимо от формы, числа и прочих свойств апертур в маске, все полученные таким образом картины считаются принадлежащими к *фраунгоферовскому* типу. На рис. 1.6 вновь показана схема с объектной маской, содержащей одиночную апертуру конечного размера; детали получаемой при этом картины Фраунгофера рассматриваются в разд. 2.2. Во всех примерах дифракции Фраунгофера существует линейное изменение оптической длины пути, проходимого дифрагировавшим светом от точек объекта до конкретной точки дифракционной картины. Таким образом, разность оптических длин $YP - XP = YW$ на рис. 1.6 пропорциональна XY . В противоположность этому соответствующее изменение на рис. 1.2 является нелинейным и образующиеся при этих условиях картины принадлежат к картинам *типа Френеля*.

Поскольку дифракционная картина Фраунгофера представляет собой ту же самую картину, которая получалась бы на бесконечности в отсутствие линз, другой часто используемой альтернативной характеристикой является *дифракция в дальней зоне*. В противоположность ей дифракция Френеля называется *дифракцией в ближней зоне*, хотя следует отметить, что к категории френелевских (ближней зоны) относится большое многообразие картин, в то время как фраунгоферовская дифракция возникает только в одном предельном случае. Например, когда опыт Юнга проводится при достаточно большом расстоянии источника и экрана (на котором наблюдаются полосы) от апертурной маски, картина практически не отличается от фраунгоферовской. Если расстояния существенно меньше (как показано в увели-

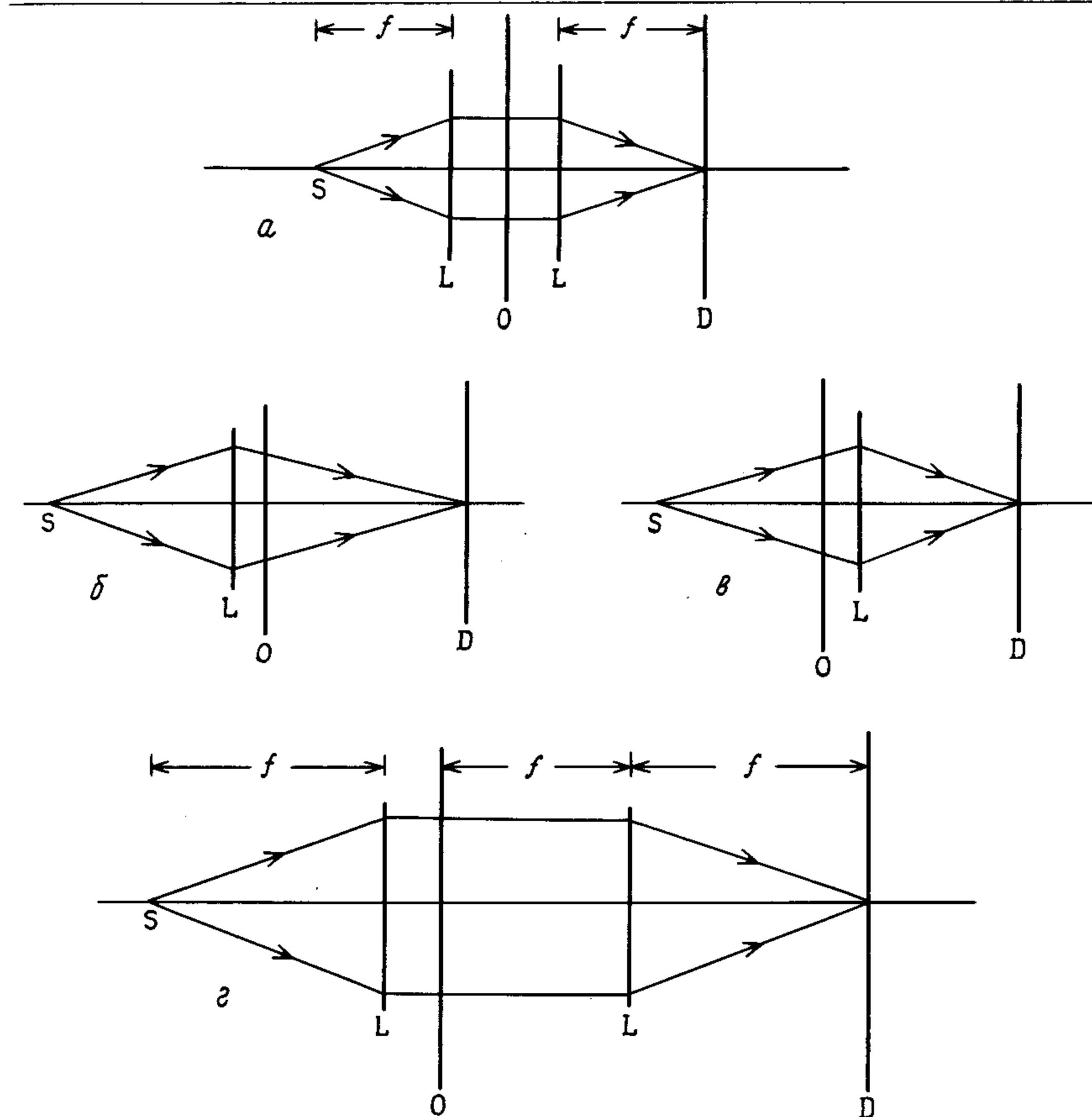


Рис. 1.7. Схемы получения картин дифракции Фраунгофера (О – объектная маска, L – линза, D – плоскость дифракции).

ченном масштабе на рис. 1.1, а), то картина Френеля будет совершенно иной.

Здесь нет необходимости углубляться в обширную область методов создания освещенностей, применяемых в оптической микроскопии. Однако полезно показать на рис. 1.7 некоторые приемы, специально используемые для наблюдения и регистрации дифракционных картин Фраунгофера, как правило, в целях «оптической фильтрации» (гл. 5). Из них на рис. 1.7, а представлен еще один способ получения той же освещенности, как на рис. 1.5 и 1.6, в то время как схемы б и в дают почти такие же картины как а. Если необходима детальная передача фазы и амплитуды, требования к схеме являются более строгими. Полный учет свойств линз, формирующих изображения, показывает, что в этом случае требуется схема, представленная на

рис. 1.7,г. Здесь объектный экран помещен в передней фокальной плоскости линзы, в задней фокальной плоскости которой формируется дифракционное изображение.

Далее в этой книге предполагается, если не оговорено особо, что обсуждаемые дифракционные картины относятся к фраунгоферовскому типу либо в определенной степени являются его приближением, поскольку для рассматриваемых задач отличия пренебрежимо малы (так называемое приближение дальней зоны или плоской волны).

Как уже отмечалось, роль дифракции в оптических системах формирования изображения как его промежуточного шага составляет основное содержание гл. 5. В связи с этим следует заметить, что картина дифракции Фраунгофера, определяемая объектом (таким, как рассматриваемая здесь апертурная маска), наблюдается в плоскости, где в качестве объекта формируется изображение источника (см. рис. 1.7); типичным примером является картина, получаемая при наблюдении уличного фонаря через занавеску. С помощью понятия *сопряжения*, имеющего тот же смысл, что и в геометрической оптике (где рассматривается изображение, формируемое в плоскости, сопряженной объекту) картина дифракции Фраунгофера рассматривается в плоскости, сопряженной источнику излучения. (Схематично это показано на рис. 5.5.)

1.3.2. Апертура линзы

Поскольку семейство волнового фронта, испускаемое объектом, является единственным источником информации, на которой основано формирование изображения, можно ожидать, и это действительно так, что чем больше семейство волнового фронта, поступающее на линзу, тем лучше качество изображения. Сформулированное несколько иначе, это положение сводится к хорошо известной аксиоме, гласящей, что чем больше апертура линзы, тем лучше определено изображение (в предположении, что аберрации не являются ограничивающим фактором).

Другая, эквивалентная модель рассматривает, каким образом конечная апертура линзы, служащей для построения изображения, будет ухудшать формирование изображения каждой точки объекта в отдельности. Читатель поймет, что здесь используется историческая работа о разрешающей способности телескопов, где отмечается, что изображение звезды (близко аппроксимирующей точечный источник) размывается дифракцией на апертуре линзы в диск, окруженный кольцами. Диск носит название *картины Эри* в честь члена Британского астрономического общества сэра Джорджа Эри, который исследовал детали этой картины в 1835 г. (разд. 2.3). Размеры картины Эри обратно пропорциональны диаметру дифракционной апертуры. Поэтому каждая точка объекта будет представлена в виде точки только при бесконечно большом размере апертуры.

Оба этих подхода к формированию изображения нашли свое применение и будут рассмотрены детально в гл. 5.

1.4. ИНТЕРФЕРЕНЦИЯ ПРИ ДЕЛЕНИИ АМПЛИТУД

Кольца Ньютона и полосы, наблюдаемые на тонких пленках, таких, например, как мыльные пузыри, нефть на поверхности воды и т.п., обусловлены интерференцией, возникающей при частичном отражении света от двух (или более) последовательных границ между средами с различными показателями преломления. Если волновой пучок падающего света частично отражается на первой границе (воздух/нефть в случае нефтяной пленки на воде), то уменьшенная амплитуда того же пучка передается дальше и затем частично отражается на следующей границе (нефть/вода). Интерференция возникает, если два отражения складываются вместе, как, например, при наблюдении глазом, а результат зависит от разности пути, которая появляется между ними из-за разности поверхностей. (Цветовые эффекты в белом свете наблюдаются, когда разность пути – функция толщины пленки и угла наблюдения – такова, что интерференция приводит к усилению для одних длин волн и к ослаблению для других.)

В противоположность интерференции, возникающей при дифракции за счет деления волновых фронтов апертурами, приведенные выше эффекты классифицируются как интерференция при делении амплитуд, а устройства, построенные на этом типе интерференции, называются «интерферометрами с расщеплением амплитуды». Представленный на рис. 1.8,а пример относится к интерференции между частично отраженными лучами от двух поверхностей тонкой параллельной пластинки. Каждый проходящий волновой пучок частично отражается на двух поверхностях воздух/стекло в точке О и стекло/воздух в точке В. Если μ – показатель преломления стекла, то оптическая разность пути l между двумя отраженными лучами (1, 2) в точках О и С дается выражением

$$l = \mu(OB + BC) - OD.$$

Поскольку $\mu OA = OD$, это выражение сводится к

$$l = \mu(AB + BC) = \mu BC(\cos 2\theta + 1) = \frac{\mu h}{\cos \theta} 2 \cos^2 \theta,$$

$$l = 2\mu h \cos \theta. \quad (1.12)$$

Усиливающая интерференция между отраженными лучами 1 и 2 возникает, когда

$$2\mu h \cos \theta = \left(m + \frac{1}{2}\right)\lambda, \quad (1.13)$$

где m – целое или ноль. Член $\lambda/2$ в правой части уравнения учитывает изменение фазы на π , происходящее при отражении в точке О (см. трактовку отражения и преломления Стокса в приложении Б).

Если на рис. 1.8,а падающий свет не ограничен плоскостью рисунка,

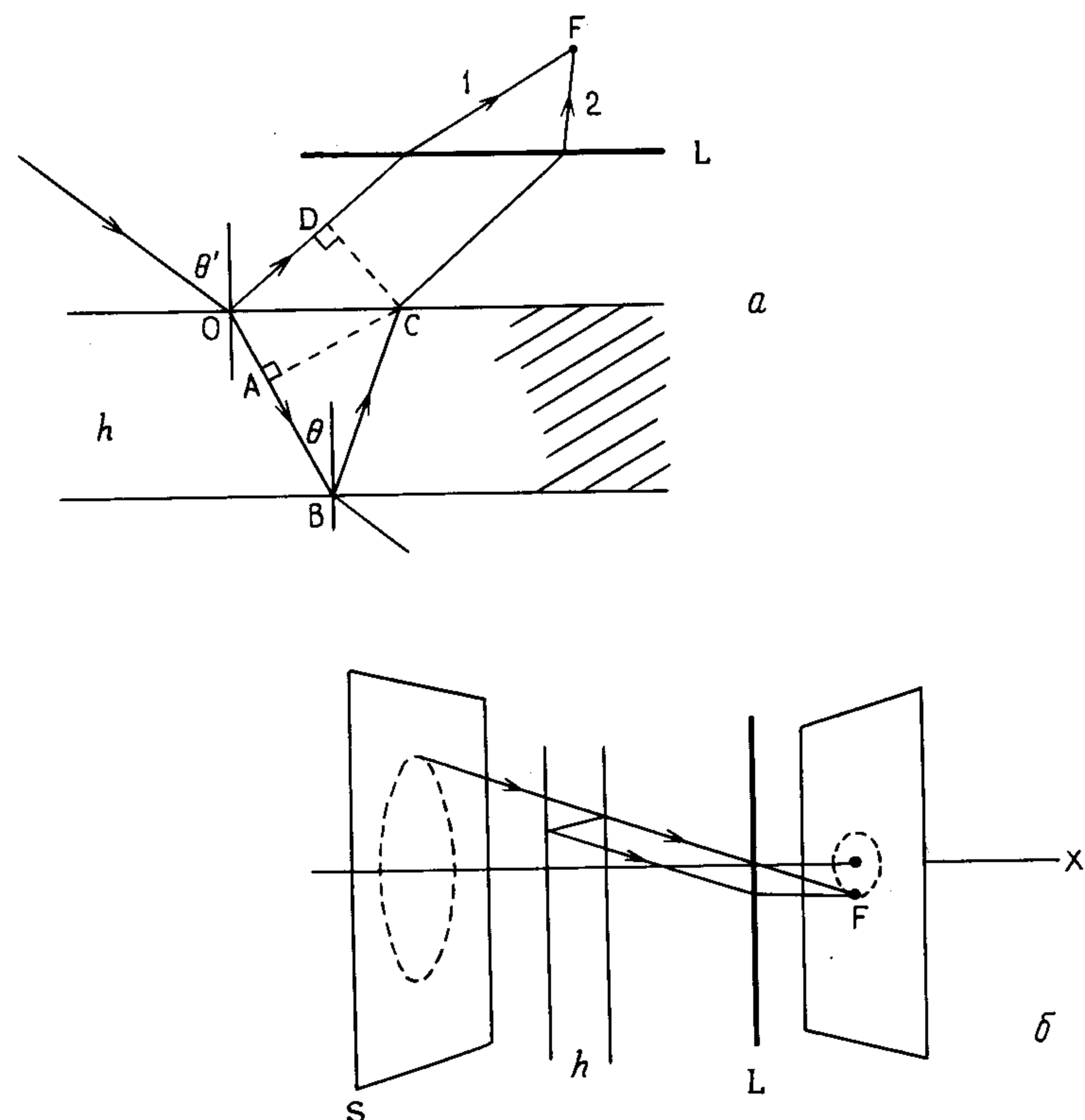


Рис. 1.8. Интерференция путем деления амплитуды (S—протяженный источник, L—линза).

могут наблюдаться концентрические круговые интерференционные полосы, как схематично показано на рис. 1.8, б. Это вызвано тем, что приведенное выше условие для максимума интенсивности, например в F, равным образом выполняется для всех точек окружности, проходящей через F, с центром на оси X.

Из замечаний по поводу когерентности в разд. 1.2 очевидно, что поскольку h может быть достаточно большим, влияние на видность полос временной (но не пространственной) когерентности (и соответственно спектрального состава излучения) в этой схеме может быть существенным. Для практического изучения этого явления используется плоскопараллельный слой воздуха изменяемой толщины. В гл. 6 мы встретимся с такой схемой в спектральном интерферометре Майкельсона и увидим, каким образом изменение видности полос при изменении расстояния между пластинами связано фурье-преобразованием со спектральным составом света. Как упоминалось ранее, эта зависимость служит основой некоторых современных методов спектроскопии.

Глава 2

Дифракция Фраунгофера

2.1. ВВЕДЕНИЕ

В этой главе в качестве подготовки последующего изложения описываются детали дифракционных картин Фраунгофера, возникающих на некоторых простых апертурах и решетках. Большая часть этого материала уже должна быть знакома читателю, но некоторые свойства, ниже отмечаемые особо, должны быть определены и выделены.

Глава завершается рассмотрением тесной аналогии между оптической дифракцией и дифракцией рентгеновских лучей на кристаллах. Косвенный способ, при котором можно получить изображения атомов в кристаллах на основе данных рентгеновской дифракции (косвенный, поскольку в рентгеновских лучах нельзя использовать линзы), будет кратко описан в гл. 5 в качестве приложения теории формирования изображения.

2.2. КАРТИНА ДИФРАКЦИИ НА ОДИНОЧНОЙ ЩЕЛИ

На рис. 2.1, а показано сечение щели шириной a и длиной l в направлении, перпендикулярном плоскости рисунка. Щель равномерно освещена монохроматическим светом с длиной волны λ , удовлетворяющим требованиям когерентности, сформулированным в разд. 1.2 и такой схемой опыта, где плоские волновые фронты нормально падают на щель. При этом дифракционная картина Фраунгофера, определяемая щелью, образуется в задней фокальной плоскости линзы. Предположим, что $l \gg a$, тогда события в плоскостях, параллельных плоскости рисунка, могут считаться одинаковыми; такая картина является *одномерной дифракционной картиной*. Ее детальный вид можно получить с помощью модели волновых цугов Гюйгенса.

Представим, что щель разделена параллельно ее длине на большое число узких полосок одинаковой ширины. Вся щель равномерно освещена, полоски имеют одинаковую площадь и поэтому амплитуды составляющих в данном направлении θ от каждой полоски одинаковы. Расчет комбинации всех составляющих поперек всей щели легче всего выполнять с помощью векторной диаграммы. Фаза света от центральной полоски, которой соответствует точка C, может служить удобной линией отсчета на векторной диаграмме. На рис. 2.2, а вектор света от этой полоски обозначен как C' . Фаза света от соседней полоски, распо-

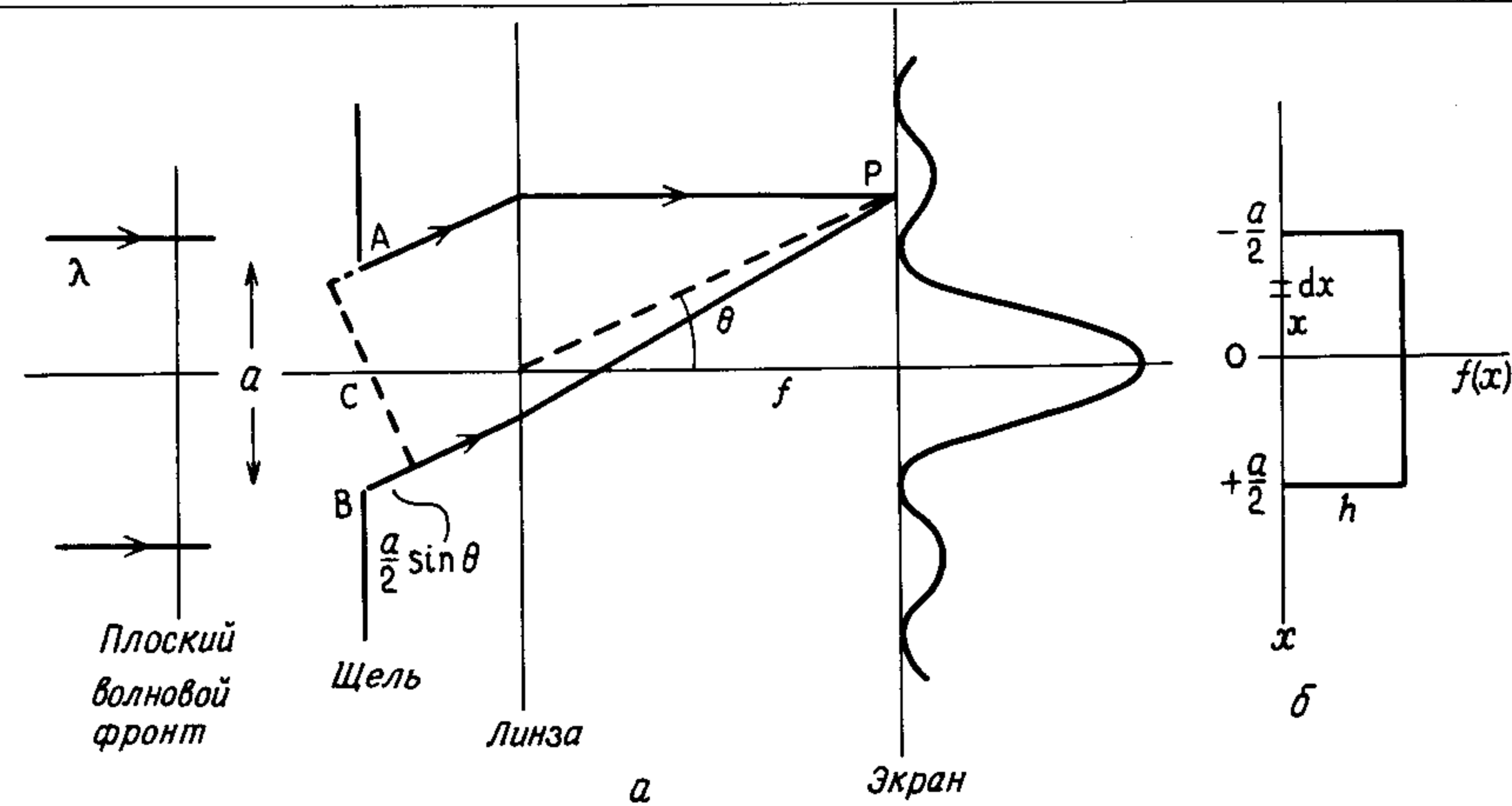


Рис. 2.1. Дифракция Фраунгофера на одиночной щели.

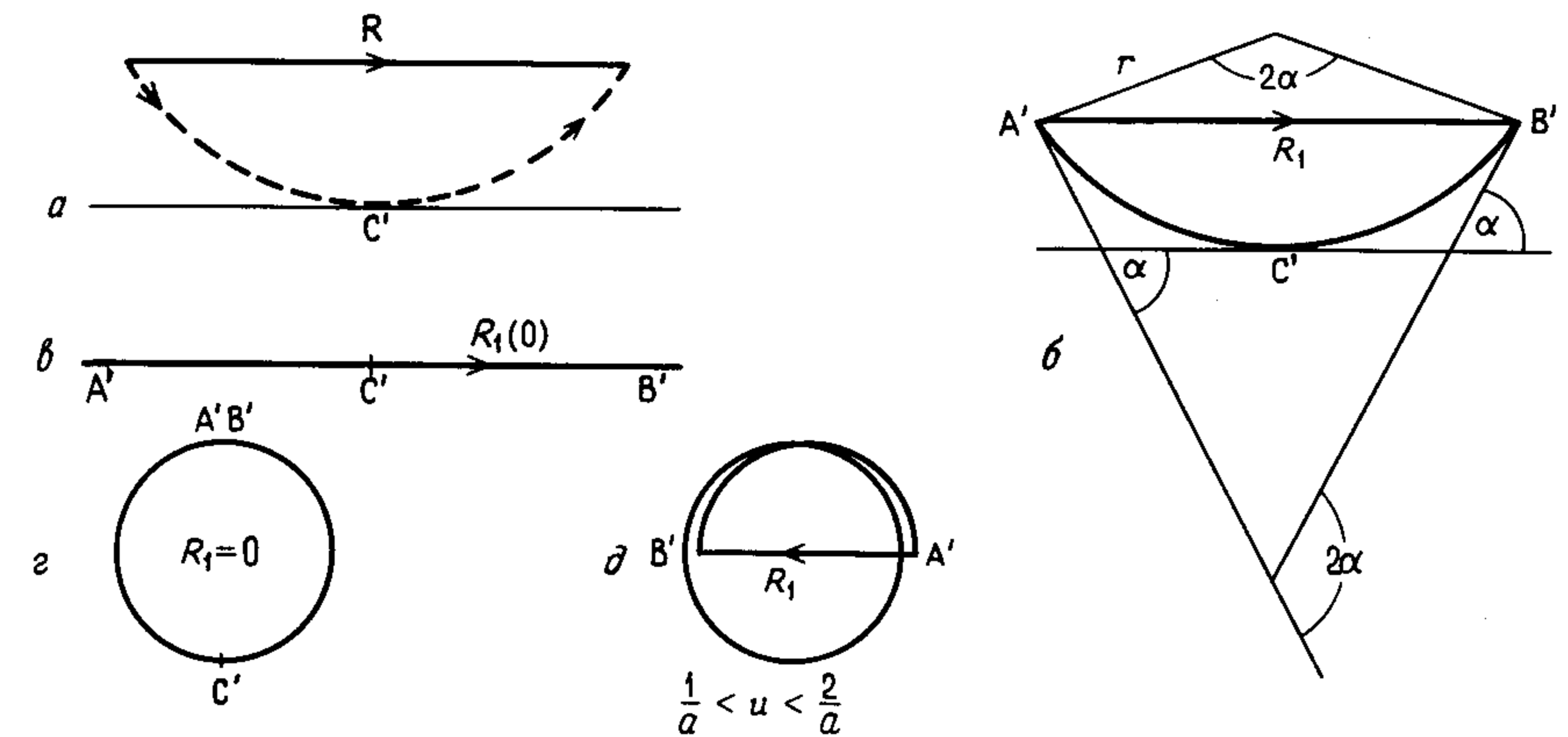
ложенной ниже С на рис. 2.1, а, отстает от фазы в С, а для соседней полоски выше С опережает ее. На векторной диаграмме соответствующие векторы добавляются к С по мере того, как пары полосок оказываются все дальше от С, пока не будут учтены вклады от всех полосок. Результатом этого построения является вектор R , но для получения точного значения нам требуется векторная диаграмма, которая представляет собой предел для полосок бесконечно малой ширины. Предельный переход выполнить нетрудно. Вначале заметим, что представленная векторная диаграмма напоминает часть правильного многоугольника, поскольку 1) векторы имеют равную длину и 2) они поворачиваются на одинаковые углы (ввиду линейного изменения фазы между составляющими от последовательных полосок). Свет от точки А на рис. 2.1, а опережает по фазе свет от С на величину

$$\alpha = \frac{2\pi}{\lambda} \cdot \frac{a}{2} \sin \theta, \quad (2.01)$$

а свет от точки В отстает на такую же величину α . Для света из точки посередине между С и В фазовая задержка равна $\alpha/2$ и т.д. В пределе, когда ширина полос уменьшается до нуля, векторы образуют дугу круга радиуса r , как показано на рис. 2.2, б, а результирующая амплитуда R_1 определяется выражением

$$R_1 = 2r \sin \alpha. \quad (2.02)$$

Индекс поставлен с целью напомнить нам, что этот результат относится к одиночной щели. Дуга $A'C'B'$ длиной $r \cdot 2\alpha$ равна полной длине

Рис. 2.2. Дифракция на одиночной щели: векторные диаграммы. (Небольшое смещение на рис. д показывает, что полная разность фаз равна 3π .)

всех элементарных векторов, вытянутых в прямую линию, как показано на рис. 2.2, в. Теперь эта прямая линия представляет векторную диаграмму в направлении $\theta = 0$, так как в этом особом направлении (см. рис. 2.1, а) разность пути отсутствует и все составляющие находятся в фазе с вектором в точке С. Поэтому мы можем записать

$$R_1(0) = r \cdot 2\alpha,$$

где $R_1(0)$ обозначает амплитуду в направлении $\theta = 0$. Подстановка в уравнение (2.02) дает

$$R_1 = R_1(0) \left[\frac{\sin \alpha}{\alpha} \right] \quad (2.03)$$

Свойства членов в скобках показаны на рис. 2.3, а.

Экспериментально наблюдаемая картина интенсивности определяется в виде

$$I_1 = R_1^2,$$

как показано на рис. 2.3, б и рис. 2.12, в.

Переменная α является функцией $(\sin \theta)/\lambda$ и a [см. уравнение (1.1)]. Интерпретация дифракционной картины упрощается, если заменить $(\sin \theta)/\lambda$ одной переменной

$$u = \frac{\sin \theta}{\lambda}. \quad (2.04)$$

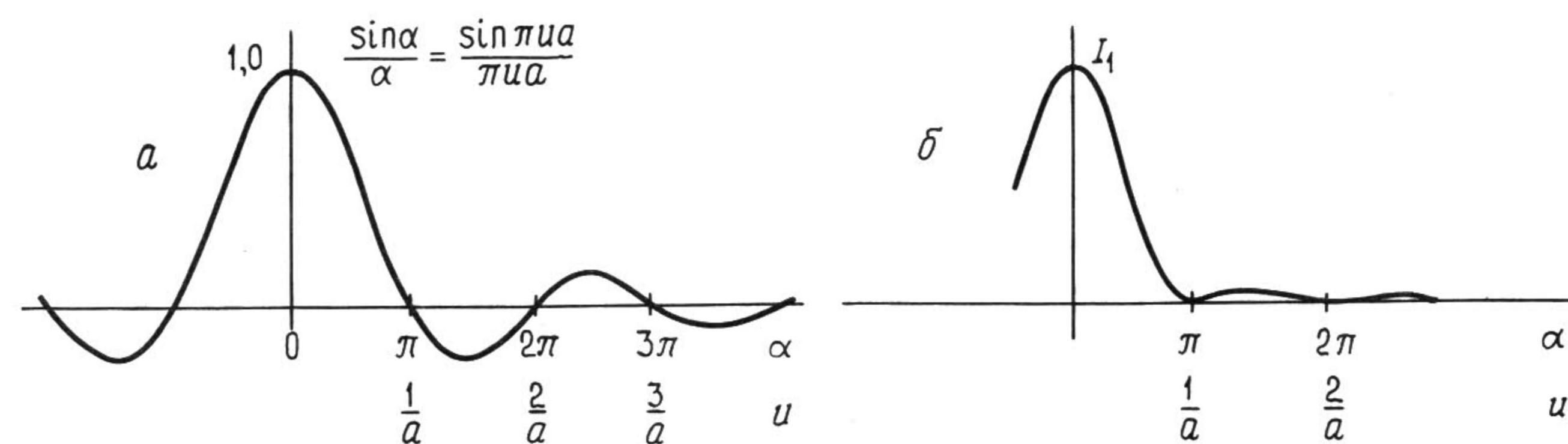


Рис. 2.3. Дифракционная картина от одиночной щели: *a* – амплитуда (sinc-функция); *b* – интенсивность.

Тогда уравнение (2.03) принимает вид

$$R_1 = R_1(0) \left[\frac{\sin \pi u a}{\pi u a} \right]. \quad (2.05)$$

Член в скобках в этом уравнении является примером sinc-функции вида $(\sin \pi x)/\pi x = \text{sinc } x$, которая широко употребляется в физике.

Рис. 2.3 иллюстрирует существование обратной зависимости между масштабом картины, выраженным через u , и шириной щели a . Теперь u имеет размерность обратной длины, и в результате мы приходим к идее описания дифракционных картин Фраунгофера как существующих во взаимном пространстве. Эта идея будет развернута ниже, при рассмотрении картин от многоэлементных апертур (разд. 2.4 и 2.5), и приобретет дополнительный смысл, когда мы познакомимся с применением метода Фурье к дифракции (гл. 3 и 4).

В то же время на рис. 2.3, *a* следует отметить изменение фазы амплитудной картины в зависимости от u . При изменении u от нуля результирующая амплитуда уменьшается, что соответствует искривлению векторной диаграммы *b* на рис. 2.2. Однако результирующий вектор остается параллельным базовой линии фазы, т.е. его фаза равна фазе освещенности в направлении $u = 0$. Когда $u = 1/a$, векторная диаграмма образует полный круг (рис. 2.3, *г*) и результирующая амплитуда равна нулю. [Соотношение $u = 1/a$ означает, что $a \sin \theta = \lambda$, т.е. разность пути между составляющими от точек А и В составляет целую длину волны (рис. 2.1, *a*). Как может показаться, это противоречит нулевой амплитуде от всей щели, но заметим, что существует разница пути в $1/2\lambda$, вызывающая взаимное подавление составляющих от точек А и С и аналогично от точек ниже А и несколько ниже С и так далее вдоль щели.] После того как u возрастает до $u = 2/a$, векторная диаграмма (рис. 2.2, *д*) проходит больше полного круга, и потому результирующий вектор вновь параллелен базовой линии. Однако теперь фаза равна π , что проявляется в отрицательной амплитуде на рис. 2.3, *a*. Последующие нули и обращения фазы соответствуют $u = 2/a, 3/a, 4/a, \dots$, и век-

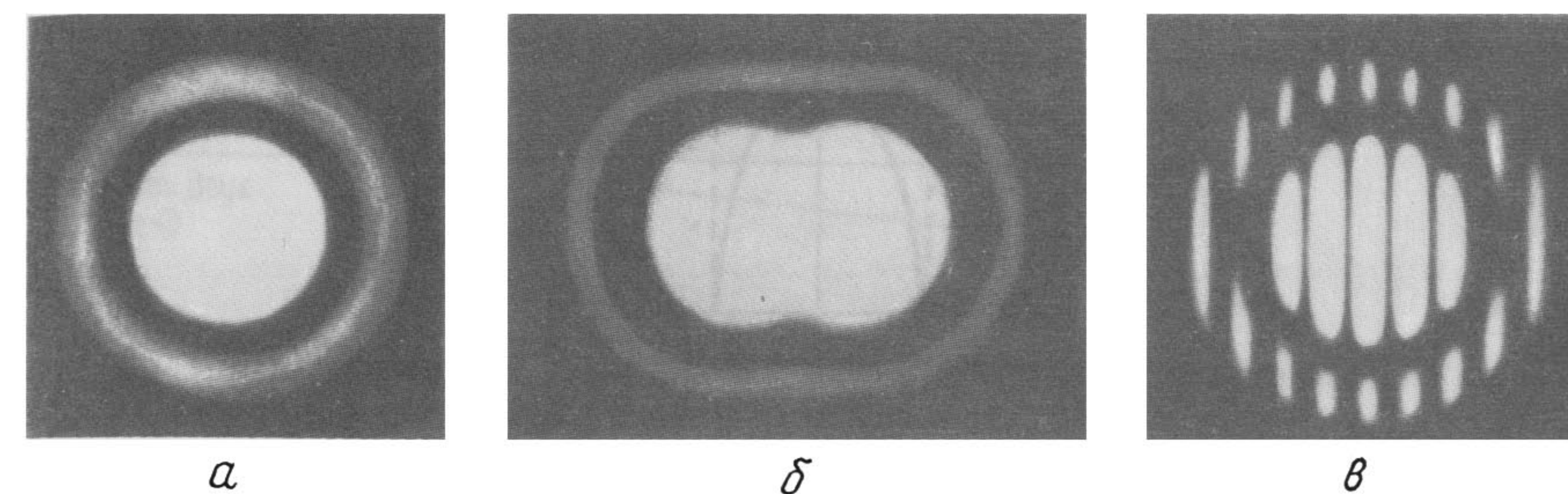


Рис. 2.4. Дифракционная картина от круглой апертуры: *a* – одиночная апертура (картина Эри); *б* – две отдельно освещенные апертуры, когерентность отсутствует; *в* – две апертуры, освещенные общим когерентным источником [рис. *a* и *в* взяты из книги Гарбурна и др. (1975)].

торная диаграмма образует все новые окружности постепенно уменьшающегося диаметра.

Амплитудное распределение освещенности по любой апертурной системе, такой, как рассматриваемая выше щель, – амплитудное пропускание системы – характеризуется *апертурной функцией*. На рис. 2.1, *б* апертурная функция $f(x)$ имеет простую прямоугольную форму («цилиндр»), т.е. равна постоянному значению h по всей ширине одиночной щели и вне ее. (Вообще $f(x) dx$ представляет собой амплитуду волнового процесса, обусловленного элементом dx при данном значении x .)

В контексте этой книги апертурную функцию любой системы апертур (не обязательно лишь одной апертуры) полезно рассматривать как представление (*оптической*) *структуры* системы.

2.3. КРУГЛАЯ АПЕРТУРА

Картина дифракции Фраунгофера от круглой апертуры особенно важна в связи с требованиями к качеству большинства оптических приборов.

К сожалению, детали картины трудно получить графическим векторным методом, который использовался в предыдущем разделе для апертуры в виде щели. Причина состоит в том, что не все полосы, на которые предполагалась разделенной апертура, имеют теперь одинаковую длину (см. апертурные функции на рис. 2.5, *a*). Их размер постепенно увеличивается, а затем уменьшается по апертуре и векторная диаграмма уже не имеет формы правильного многоугольника. Решение для этого примера лучше всего получается аналитически, а детали можно найти в обычных учебниках. Дифракционная картина (рис. 2.4, *a*) представляет собой диск в центре, окруженный круглыми концентрическими полосами, и известна как *картина Эри* по имени сэра Джорджа Эри, члена Британского астрономического общества, который подробно исследовал ее детали в 1835 г.

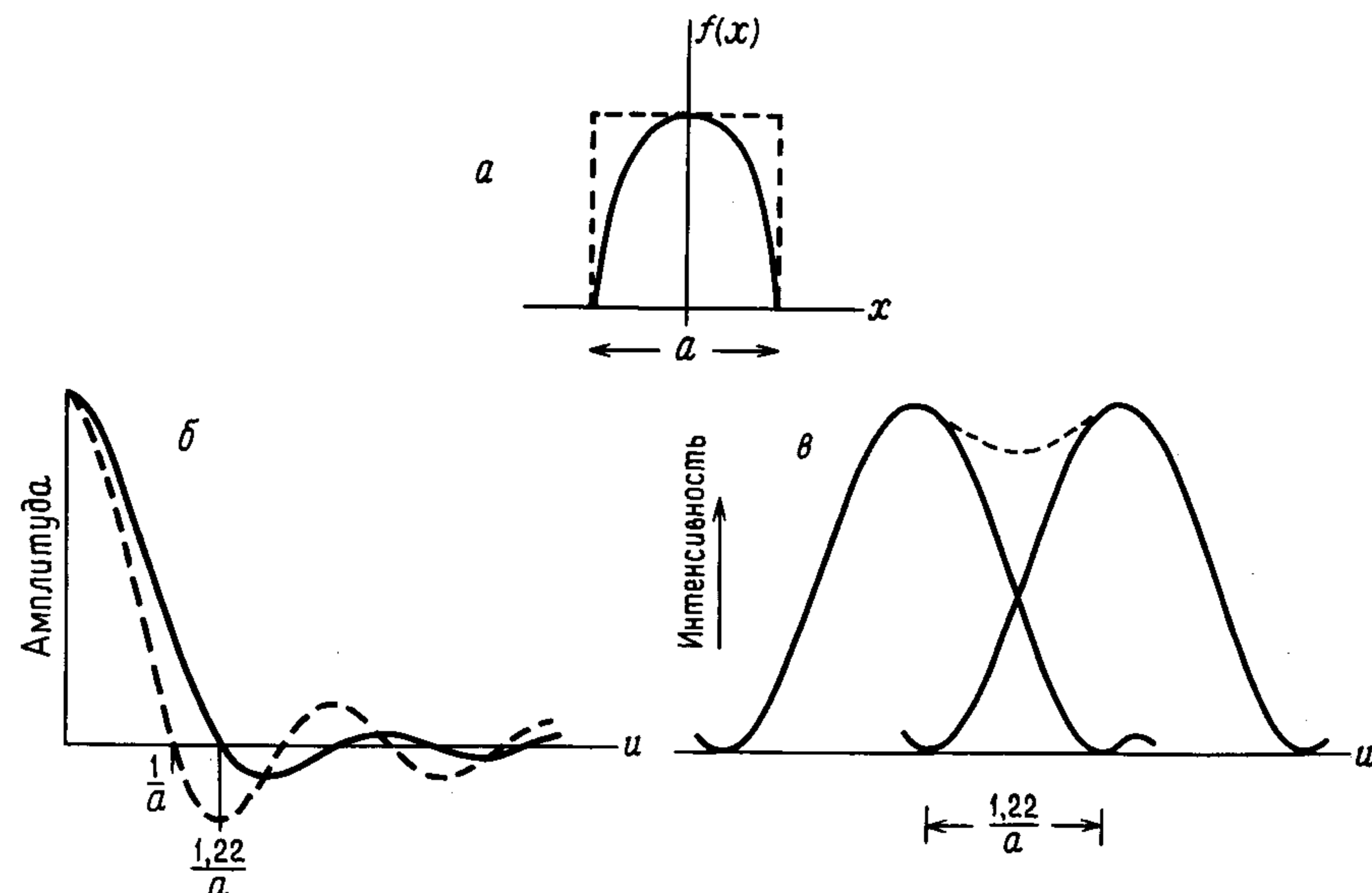


Рис. 2.5. a – апертурные функции $f(x)$ (сплошная линия – однородно освещенная круглая апертура диаметром a , пунктир – однородно освещенная щель шириной a); b – амплитудные дифракционные картины (сплошная линия – круглая апертура диаметром a ; пунктир – щель шириной a); c – критерий Рэлея (картины от круглой апертуры).

Изменение амплитуды по диаметру картины Эри показано сплошной линией на рис. 2.5, б. Хотя она и напоминает картину в случае однощелевой апертуры (показана пунктиром), характер ее описывается не sinc -функцией, а функцией Бесселя. Для нас существенной разницей между ними является то, что первый нуль интенсивности после максимума в центре картины Эри достигается не раньше, чем будет выполнено условие

$$a \sin \theta \approx 1,22\lambda, \quad (2.06)$$

где a – диаметр апертуры. Смещены также и последующие нули.

Сразу же становится очевидным пример возникновений этой картины при получении изображения звезды с помощью астрономического телескопа, и Эри исследовал ее детали именно для такого случая. Фактически для земного наблюдателя звезда находится на бесконечности, и поэтому ее изображение образуется в задней фокальной плоскости объектива телескопа, где оно изучается с помощью окуляра. Поскольку ее угловой размер (угол, под которым ее диаметр виден с Земли) чрезвычайно мал, ее изображение должно быть близким к точечному. Однако фронты световых волн от звезды прерываются апертурой

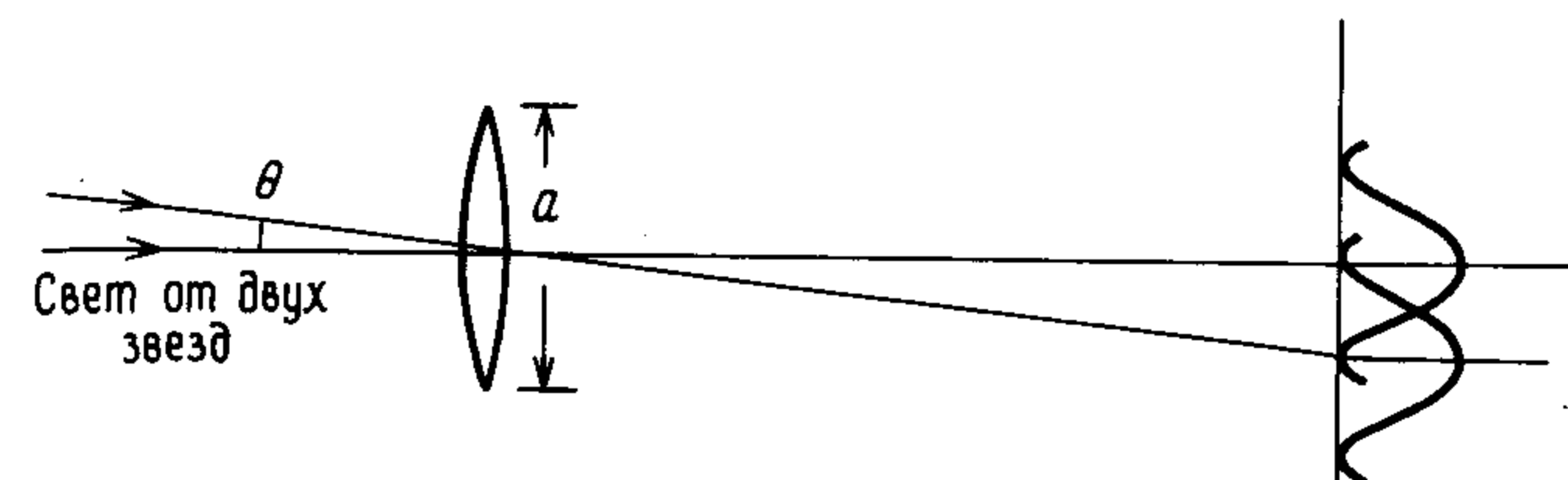


Рис. 2.6. Астрономический телескоп. Критерий Рэлея в применении к изображению двух звезд на малом угловом расстоянии.

объектива, имеющей конечный размер, возникает дифракция, и в результате изображением звезды оказывается дифракционная картина Фраунгофера от этой апертуры. Для иллюстрации возникающей при этом дифракции Фраунгофера служит схема на рис. 1.6, где показана апертура объектива, действующая как дифракционная маска.

Независимо от погрешностей объектива (линзы или зеркала) астрономического телескопа он даже в самом лучшем случае дает не точечное изображение звезды, а лишь картину Эри распределения интенсивности, обусловленного апертурой объектива телескопа (такую линзу называют *дифракционно ограниченной*). В более широком контексте гл. 5 эта картина – отклик системы на точечное (импульсное) воздействие – является *функцией рассеяния точки* (ФРТ) этой системы.

Уравнение (2.06) показывает зависимость диаметра центрального диска (*диска Эри*) от диаметра апертуры и длины волны света. Размер этого диска по существу и определяет предельное разрешение телескопа. Рассмотрим изображение двух звезд с малым угловым расстоянием θ (рис. 2.6). Поскольку они являются некогерентными по отношению друг к другу источниками, изображение состоит из двух картин интенсивности Эри. Поэтому возможность разрешения двух звезд зависит от размера дисков Эри и расстояния, на котором они перекрываются. Общепринятое граничное условие, *критерий Рэлея*, представляет собой расстояние, показанное на рис. 2.4, б и 2.5, в. Согласно этому критерию, две картины разрешаются, если центр диска Эри одной из них налагается на темное кольцо другой. Это обеспечивает провал на 20% в суммарной кривой интенсивности между пиками (которые предполагаются нами одинаковыми по интенсивности). Величина этого провала, хотя и выбрана весьма произвольной, тем не менее является во многих случаях удобным критерием разрешения.

Отсюда на основе критерия Рэлея уравнение (2.06) дает предел углового разрешения для *телескопа* с размером апертуры a в виде

$$\theta \approx 1,22 \lambda/a, \quad (2.07)$$

или, приближенно для желтого света,

$$\theta \approx \frac{0,15}{a(\text{м})} (\text{сек. дуги}).$$

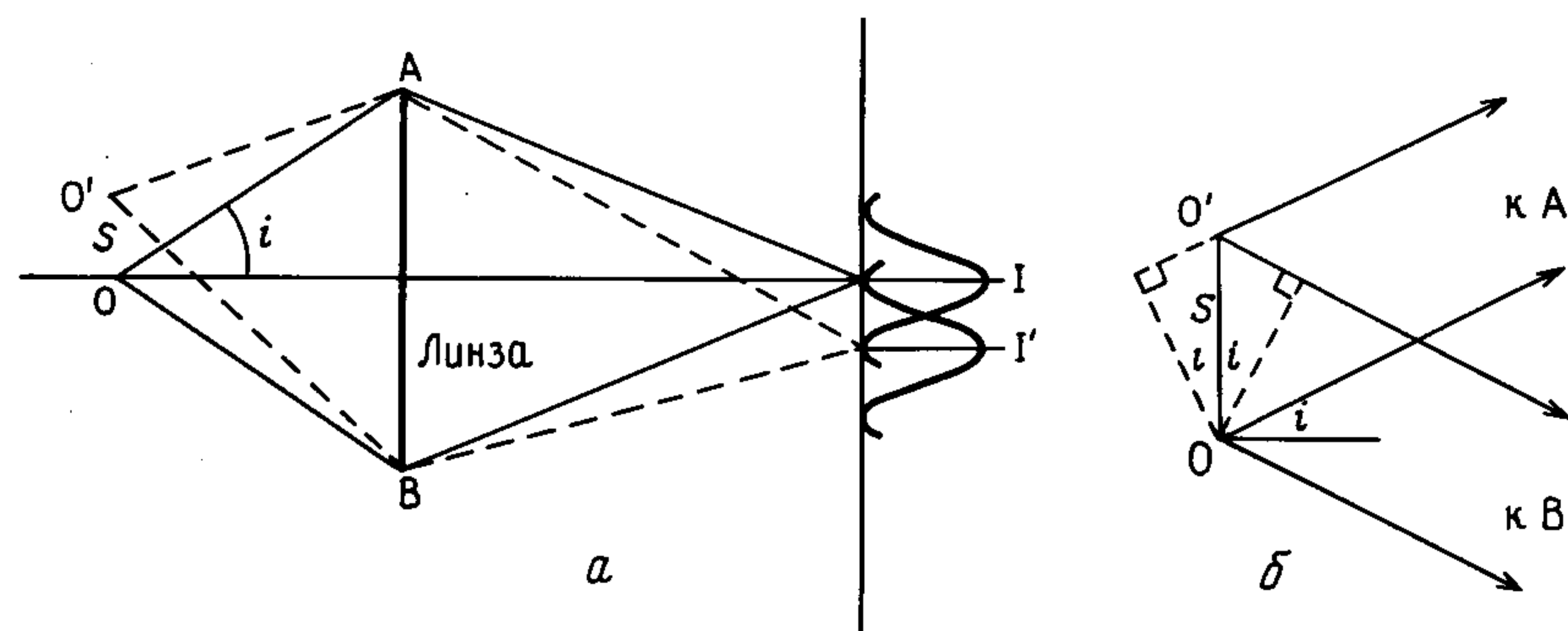


Рис. 2.7. Оптический микроскоп (случай самосветящегося объекта). а – изображения (картины Эри) двух точек объекта O и O', разделенных расстоянием s , согласно критерию Рэлея, разрешаются линзами объектива, если $O'B - O'A = 1,22\lambda$, так как $OA = OB$. Из схемы б находим: $O'B = OB + s \sin i$, $O'A = OA - s \sin i$, следовательно, $s = 0,6\lambda/\sin i$.

(Для улучшения способности телескопа разрешать очень близкие изображения звезд можно подавить кольца, окружающие диск Эри. Это достигается путем *аподизации* – видоизменения апертурной функции объектива с помощью, например, стеклянной пластинки, пропускание которой изменяется от точки к точке наперед известным способом. Однако за ограничение пропускания объектива приходится расплачиваться уширением самого центрального диска.)

Те же самые факторы определяют предел разрешения зрительных труб или фотокамер, предназначенных для наблюдения земных объектов. При нормальных условиях освещенности каждая точка наземного объекта рассеивает свет и участвует в формировании изображения независимо от соседних точек. Ситуация здесь фактически такая же, как при построении изображения звездного скопления. По этой причине термин «самосветящийся объект» зачастую с определенной степенью вольности используется в обоих контекстах для краткого указания на объекты, изображения которых строятся при некогерентных условиях. В случае зрительной трубы или фотокамеры изображение каждой точки объекта, служащей источником, также не является точкой, а представляет собой дифракционную картину апертуры объектива (ср. с разд. 1.3.1). (Мы не будем рассматривать роль окуляра при формировании изображения телескопом или микроскопом, о котором речь идет ниже, поскольку он представляет собой вторичный элемент оптической схемы и не является главным источником искажений.)

Аналогичная ситуация возникает и при использовании оптического микроскопа (простого или сложного) для изучения самосветящегося объекта при условиях, подобных рассмотренным выше. На основе критерия Рэлея, как и в телескопе, минимальное расстояние s на рис. 2.7, а

между двумя точками объекта, которое можно разрешить микроскопом, определяется соотношением

$$s \approx 0,6\lambda/\sin i, \quad (2.08)$$

где i – половина угла, стягиваемого объективом, с вершиной у объекта. Если пространство объекта заполнено средой с показателем преломления μ (используется масляно-иммерсионный объектив), то тогда длина волны света уменьшается до λ/μ и приведенное выше выражение принимает вид:

$$s \approx 0,6\lambda/\mu \sin i, \quad (2.09)$$

где $\mu \sin i$ – *числовая апертура* объектива микроскопа. Максимально возможное значение i равно 90° , что примерно определяет *предел разрешения микроскопа* $(1/2) (\lambda/\mu)$.

В общем случае в оптических системах формирования изображения имеется диафрагма, которая регулирует способность системы собирать свет. Эта *апертурная диафрагма*, нередко помещаемая между различными линзовыми элементами систем, неизбежно приводит к возникновению дифракции. Со стороны объекта (т.е. источника) эта апертура называется *входным зрачком*, а со стороны изображения – *выходным зрачком*. На языке инструментальной (приборной) оптики зрачки являются, таким образом, изображениями апертурной диафрагмы, построенными в пространствах объекта и изображения. А определенная уже в разд. 2.2 апертурная функция, представленная в координатной системе пространства изображения, называется (*выходной*) *функцией зрачка*.

2.4. ДВОЙНАЯ АПЕРТУРА

2.4.1. Две щели

Выяснение свойств картины дифракции Фраунгофера, получаемой в случае двух щелей, является полезным введением при переходе к картине со многими щелями (разд. 2.5), где действуют аналогичные принципы.

Схема на рис. 2.8, а похожа на рассмотренную выше, но в ней имеются две параллельные друг другу щели с расстоянием D между их центрами. Пространственное распределение комплексных амплитуд от каждой щели в точности такое же, как прежде, но окончательный результат в любом направлении зависит от разности пути между двумя составляющими в этом направлении. Например, полное усиление происходит только для таких θ , при которых дифрагировавший от двух щелей свет приходит в фазе. Из рисунка следует, что это условие выполняется, когда

$$D \sin \theta = n\lambda, \quad (2.10)$$

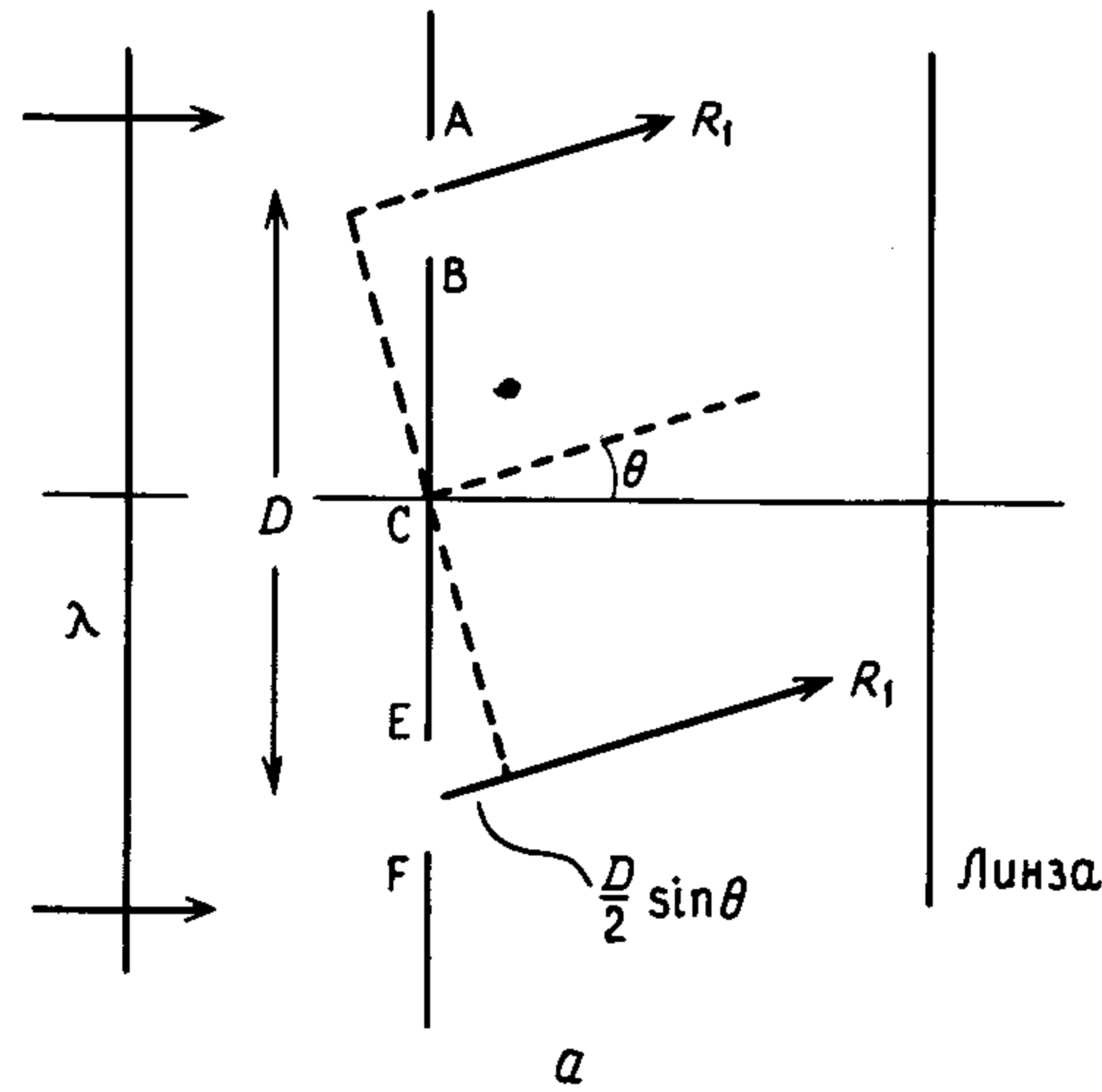


Рис. 2.8. Дифракция на двух щелях.

т.е., учитывая (2.04),

$$\frac{\sin \theta}{\lambda} = u = \frac{n}{D},$$

где порядок дифракции n — нуль или целое число. Во всех других направлениях интерференционная картина в разной степени размывается.

Векторная диаграмма позволяет получить аналитическое выражение для всей картины. В зависимости от начала отсчета фаз, которое на рисунке для удобства находится в точке С, свет от щели АВ дифрагирует в направлении θ с фазой $+\beta$, а от щели EF — с фазой $-\beta$, где

$$\beta = \frac{2\pi}{\lambda} \frac{D}{2} \sin \theta, \quad (2.11)$$

т.е.

$$\beta = \pi u D.$$

Из векторной диаграммы (рис. 2.8, б) следует, что результирующая амплитуда равна

$$R_2 = 2R_1 \cos \beta.$$

Используя R_1 , из уравнения (2.05) и β из уравнения (2.11) получаем, что

$$R_2 = 2R_1(0) \left(\frac{\sin \pi u a}{a} \right) \cos \pi u D. \quad (2.12)$$

Это выражение с точностью до постоянного множителя представлено на рис. 2.9. Изменение фазы в амплитудной картине становится ясным, если нарисовать для разных значений β векторные диаграммы, подобные использованной выше. Окончательный график на рис. 2.9 дает распределение интенсивности, полученное при возведении в квадрат амплитудной картины, т.е.

$$I_2 = 4I_1(0) \left(\frac{\sin \pi u a}{\pi u a} \right)^2 \cos^2 \pi u D. \quad (2.13)$$

Пример наблюдаемой картины показан на рис. 2.12, б.

В связи с двухщелевой дифракционной картиной следует отметить несколько важных обстоятельств. Как свидетельствуют уравнение (2.13) и рис. 2.9, эта картина определяется произведением двух дифракционных членов: дифракционной картины для одиночной щели и интерференционной картины вида $\cos^2$, связанной с интерференцией света, дифрагировавшего на двух щелях.

Максимумы полос картины вида $\cos^2$ соответствуют направлениям, в которых наблюдается интерференция с усилением интенсивности света от двух щелей. В этом случае мы имеем $D \sin \theta = n\lambda$, т.е. $u = n/D$ [см. уравнение (2.10)], член $\cos^2$ равен единице и наблюдаемая интенсивность, как и следовало ожидать, равна квадрату суммы отдельных амплитуд от двух щелей. Подобным же образом между этими направлениями происходит интерференция со снижением интенсивности, член $\cos^2$ равен нулю и наблюдаемая освещенность равна также нулю. Таким образом, максимумы и минимумы наблюдаются в направлениях, определяемых расстоянием D между щелями. Заметим, что их интенсивности определяются тем не менее амплитудами света, дифрагировавшего на щелях в тех же направлениях. В этом смысле наблюдаемая дифракционная картина может рассматриваться как усиленные отсчеты из однощелевой дифракционной картины, которые ограничены направлениями, определяемыми расстояниями между щелями. Это с очевидностью следует из рис. 2.9 и 2.12, б, где показано, что зависимость для одной щели является огибающей полос $\cos^2$. Подстановка для большей убедительности условия $u = n/D$ в уравнение (2.13) дает

$$(I_2)_{\max} = 4I_1(0) \left[\frac{\sin \frac{\pi n a}{D}}{\frac{\pi n a}{D}} \right]^2, \quad (2.14)$$

и член в квадратных скобках можно охарактеризовать как *выборочную sinc-функцию*.

В соответствии с интерпретацией u как параметра во взаимном пространстве (разд. 2.2) можно считать, что эта выборка из однощелевой

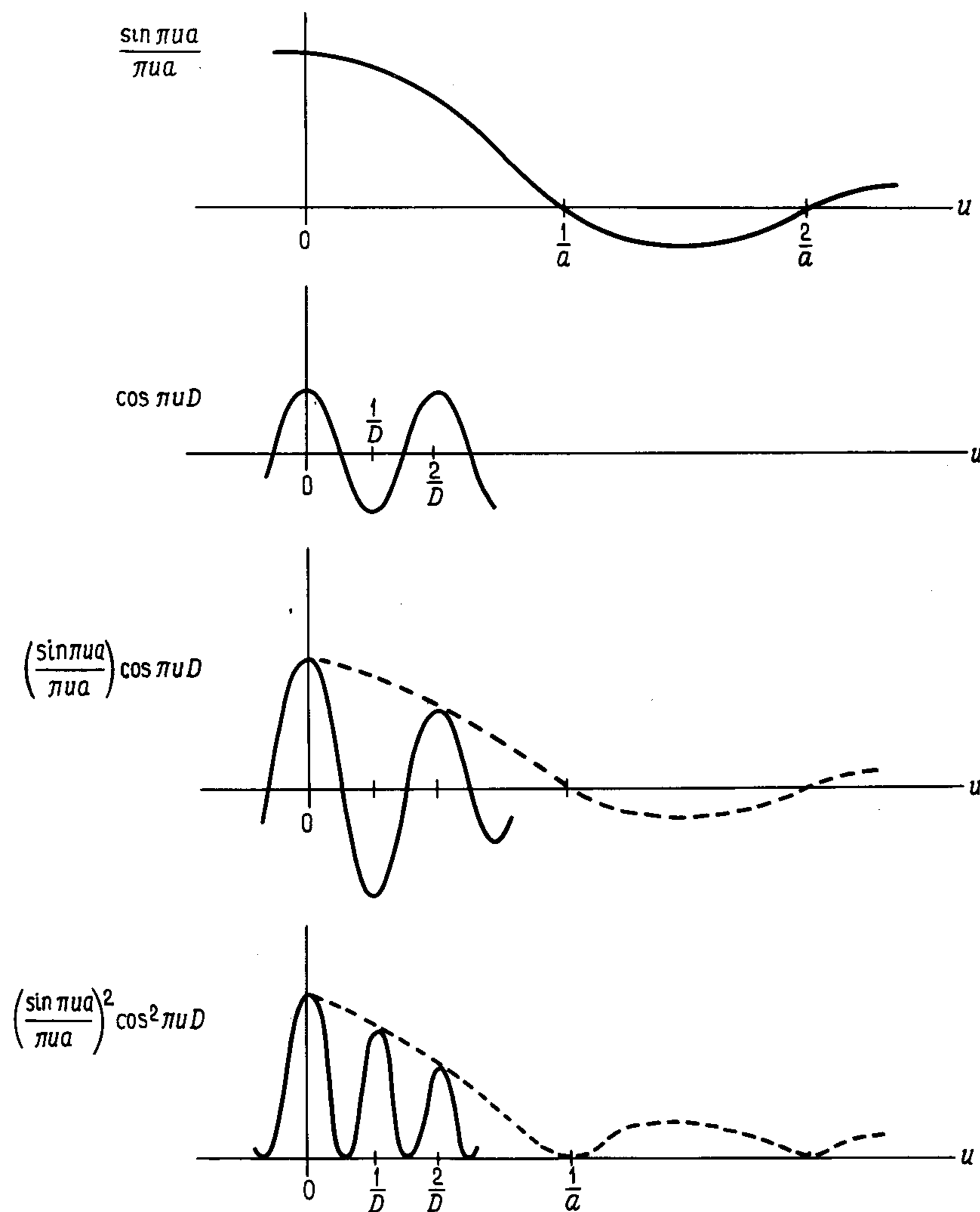
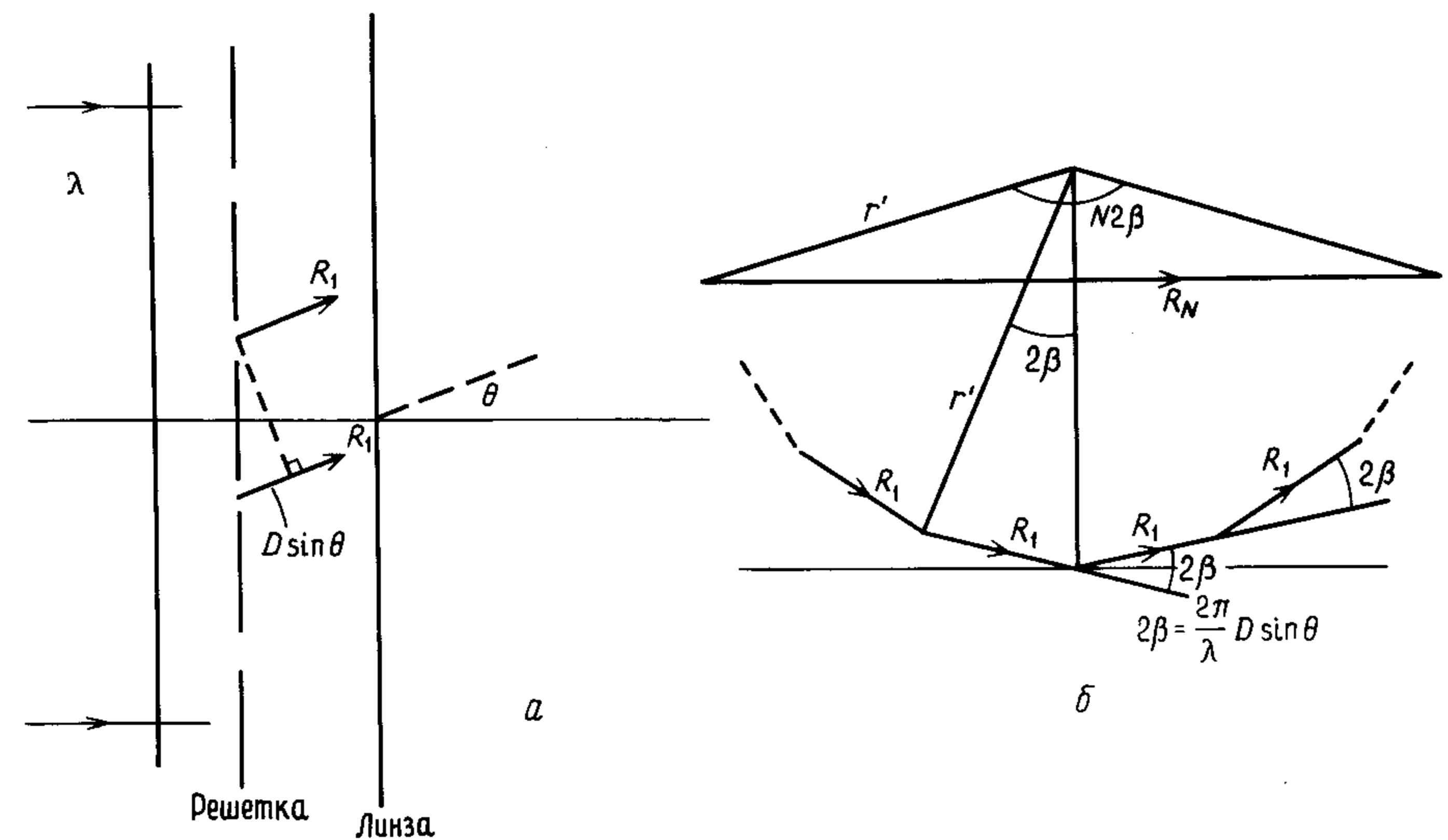


Рис. 2.9. Картина дифракции на двух щелях.

картины имеет место в этом пространстве. Мы можем сказать, что картина от двух щелей представляет собой выборку из однощелевой картины во взаимном пространстве; эта выборка осуществляется для значений u , которые обратно пропорциональны расстоянию D между щелями, т.е. $u = n/D$. Таким образом, чем меньше D , тем шире разнесены интерференционные полосы.

Рис. 2.10. Дифракция на N щелях (период решетки D).

2.4.2. Две круглые апертуры

Получение дифракционной картины от двух круглых апертур, расположенных на расстоянии D друг от друга, во многом аналогично двум щелям. Здесь нет нужды разбирать все подробности. Как показано на рис. 2.4, в, снова мы имеем дифракционную картину одиночной апертуры (в данном случае это картина Эри), умноженную на тот же, что выше, член $\cos^2$, т.е. член, который обусловлен расстоянием D между апертурами. Обратите внимание на различие между картинами б и в на рис. 2.4. Напомним, что картина б представляет результат суммирования по интенсивности двух различных картин Эри, полученных при некогерентном освещении от различных источников. На рис. 2.4, в, где имеется когерентность, происходит суммирование по фазе двух амплитудных картин Эри.

2.5. РЕШЕТКА ИЗ N ЩЕЛЕЙ

Векторный метод, аналогичный использованному для двойной щели в разд. 2.4.1 (рис. 2.10), позволяет легко определить детали дифракционной картины для работающей на пропускание одномерной плоской многощелевой решетки, используемой в лаборатории.

Если решетка состоит из N щелей, то имеется N векторов, каждый длиной R , с такими же, как и раньше, амплитудами от каждой отдельной щели [см. уравнение (2.05)]. Вновь последовательные векторы обра-

зуют между собой угол 2β , где $\beta = \pi u D$ [см. уравнение (2.11)]. Как показано на рис. 2.10, концы векторов лежат на окружности радиуса $r' = R_1/2\sin\beta$. Для N щелей результирующий вектор R_N в итоге равен

$$R_N = 2r'\sin N\beta = R_1 \frac{\sin N\beta}{\sin \beta}$$

Подстановка β дает

$$R_N = R_1 \frac{\sin N\pi u D}{\sin \pi u D}$$

Здесь D – расстояние между соседними щелями (период решетки). Воспользовавшись, как и прежде, уравнением (2.05) для R_1 , получаем

$$R_N = R_1(0) \left(\frac{\sin \pi u a}{\pi u a} \right) \left(\frac{\sin N\pi u D}{\sin \pi u D} \right). \quad (2.15)$$

Интенсивность определяется квадратом этого выражения, а именно

$$I_N = I_1(0) \left(\frac{\sin \pi u a}{\pi u a} \right)^2 \left(\frac{\sin N\pi u D}{\sin \pi u D} \right)^2 \quad (2.16)$$

Вновь мы получили (возведенное в квадрат) произведение двух членов, из которых первый соответствует одной щели, а второй, представленный здесь как член решетки, соответствует $\cos^2 \pi u D$ для $N=2$ в уравнении (2.13).

На рис. 2.11 графически показано формирование члена решетки. Интерференционные полосы, образующие основные пики, являются главными максимумами. Они возникают при полном усилении между светом, дифрагировавшим от последовательных щелей, и, так же как в случае двух щелей, это происходит при $D \sin \theta = n\lambda$, т.е. $u = n/D$, где n – порядок дифракции [ср. с уравнением (2.10)].

В направлениях главных максимумов значение члена решетки равно N^2 , поскольку при $u \rightarrow n/D$ мы получаем

$$\left(\frac{\sin N\pi u D}{\sin \pi u D} \right)^2 \rightarrow \left(\frac{N\pi u D}{\pi u D} \right)^2 = N^2.$$

Поскольку для большинства используемых типов решеток N велико (как правило, десятки тысяч), главные максимумы и более слабые вторичные максимумы становятся очень острыми, как видно на рис. 2.11. Рисунок также показывает, что при больших N вторичные максимумы сближаются. Полное воздействие на вторичные максимумы состоит в том, что они становятся несущественными для дифракционной кар-

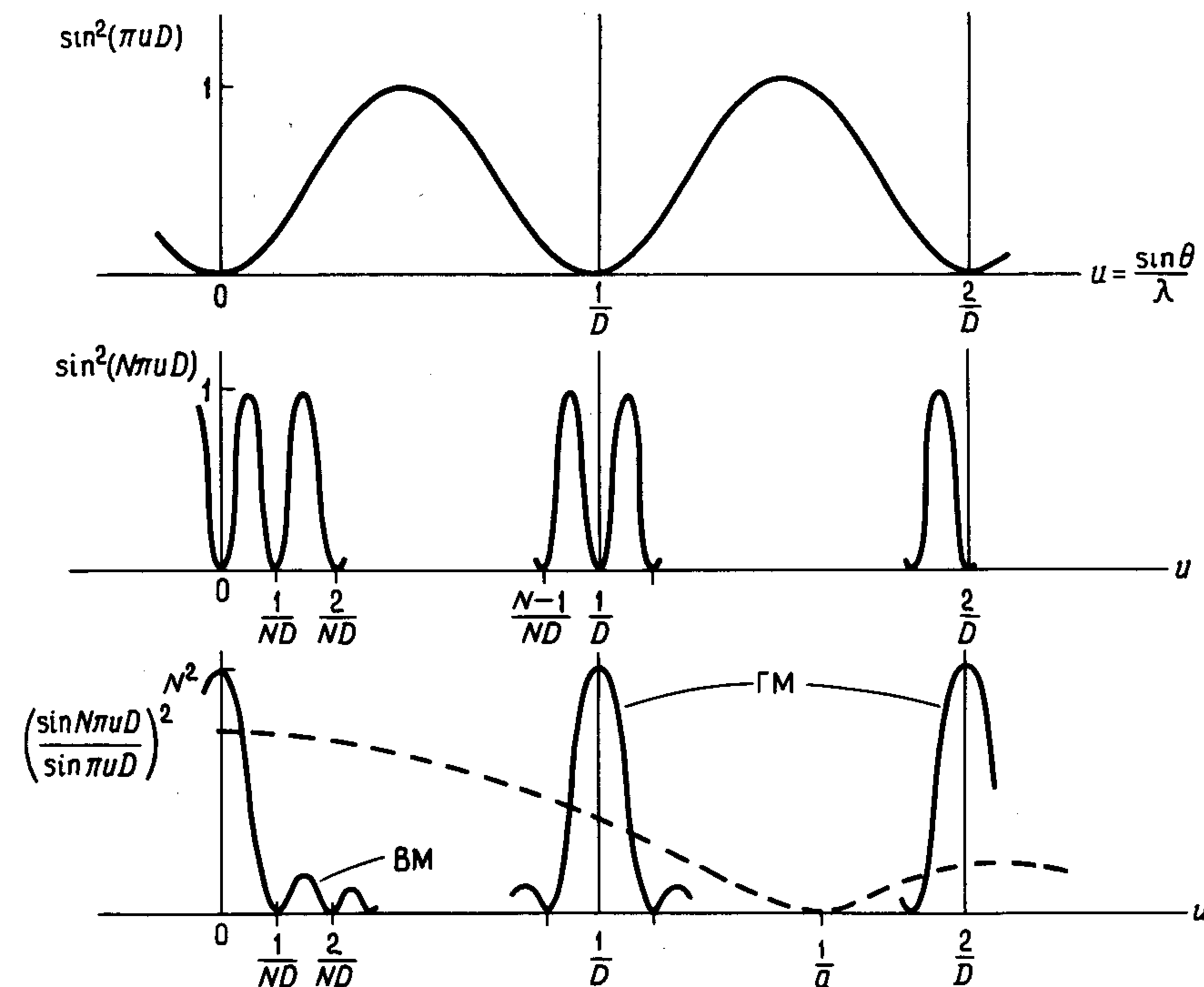


Рис. 2.11. Картина дифракции для N щелей (ГМ – главные максимумы; BM – вторичные максимумы; пунктиром показан член одиночной щели).

тины; поэтому в дальнейшем мы не будем принимать их во внимание.

Интенсивность освещенности в направлениях главных максимумов выражается как квадрат произведения N на амплитуду дифракции от каждой щели в этих направлениях. Для таких направлений уравнение (2.16) принимает вид

$$(I_N)_{\max} = I_1(0) \left[\frac{\sin(\pi n a/D)}{(\pi n a/D)} N \right]^2, \quad (2.17)$$

что представляет собой усиленную выборочную функцию sinc^2 , как для случая двух щелей ($N=2$). На рис. 2.12 приведены дифракционные картины от решеток с числом щелей N от 1 до 6. Здесь показано, как главные максимумы представляют выборку дифракционной картины одной щели.

Член одиночной щели и член решетки оба могут быть описаны во взаимном пространстве (разд. 2.2). Пространственное распределение главных максимумов в плоскости дифракции Фраунгофера может быть описано как взаимная структура решетки, например одномерная струк-

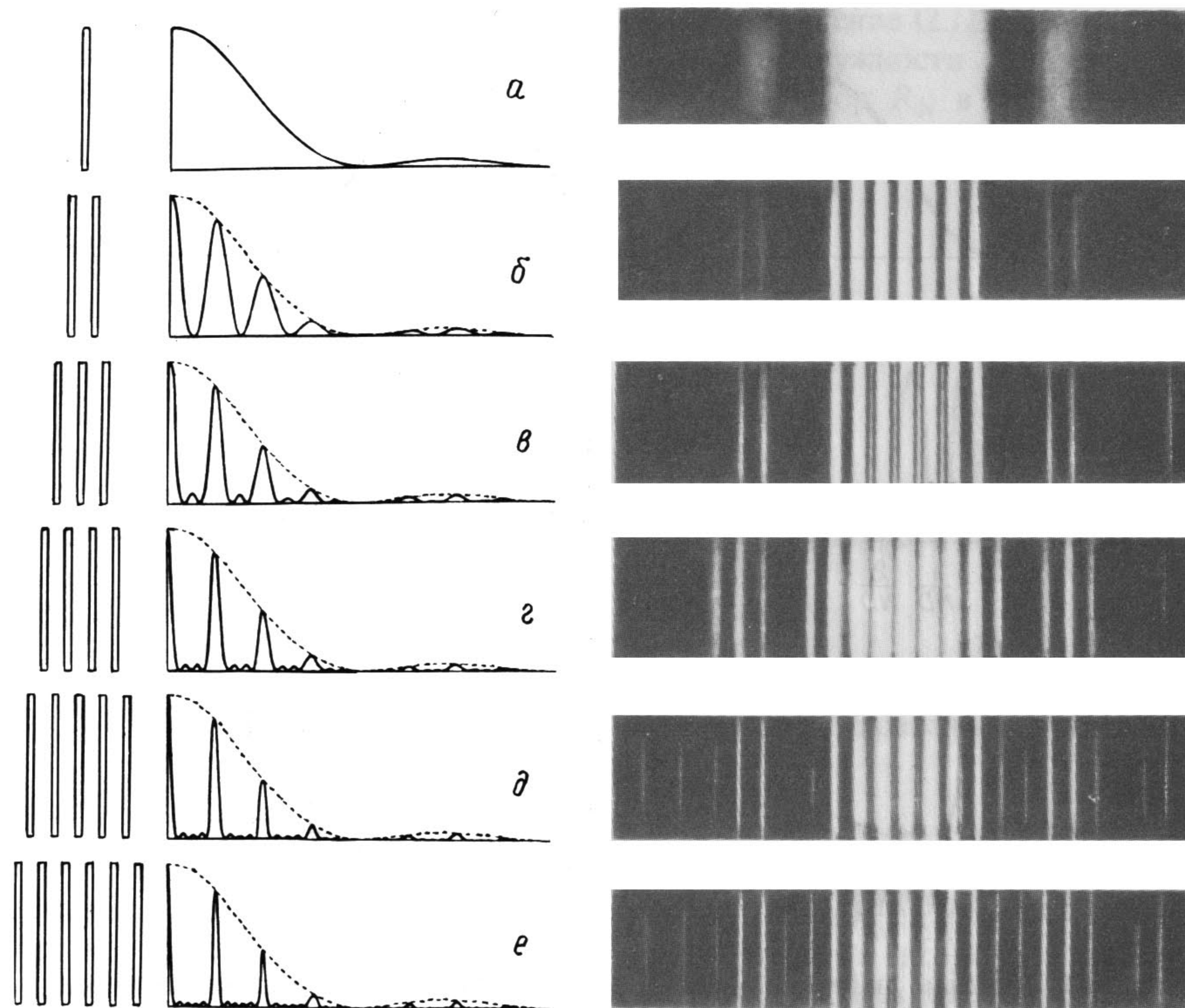


Рис. 2.12. Дифракционные картины от решеток с различным числом щелей [43].

тура решетки в представленном примере. Таким образом, можно сказать, что для больших N , когда главные максимумы являются острыми, полученную картину можно рассматривать как усиленную выборку дифракционной картины от одиночной щели в точках взаимной структуры решетки. Это представление является весьма важным.

Наконец, следует заметить, что ввиду симметричности картины относительно центрального (нулевого порядка) максимума все другие максимумы (имеются в виду главные максимумы) дифракционных порядков $\pm n$ расположены парами симметрично относительно центрального максимума.

2.6. ДВУХМЕРНЫЕ РЕШЕТКИ

Распространение результатов предыдущих разделов на двумерный случай не требует введения новых принципов. Единственный результат, который здесь нам нужно отметить, схематично обобщен на рис. 2.13. Апертурная маска, содержащая большое число точечных отверстий,

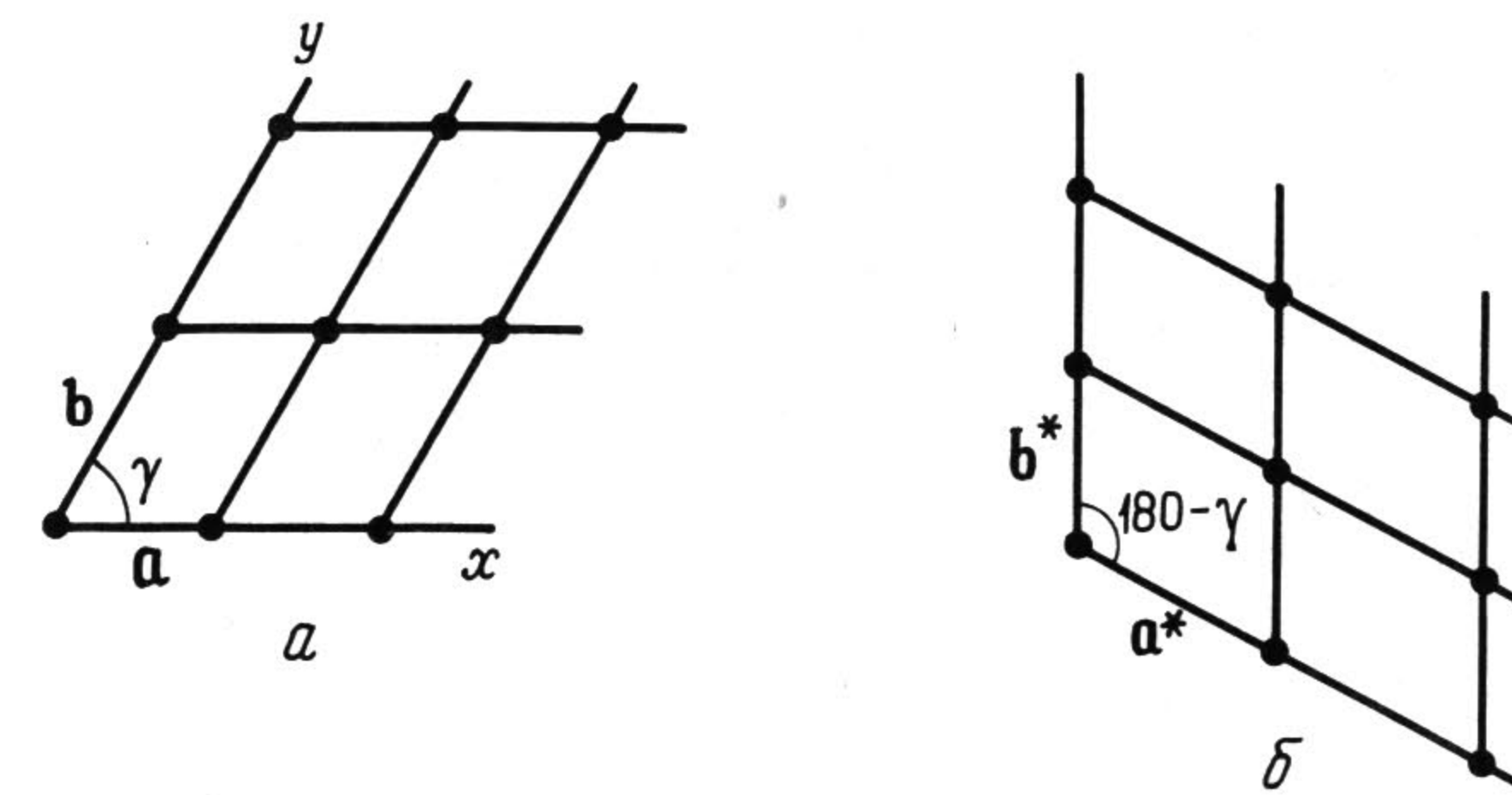


Рис. 2.13. Двухмерная дифракция.

образующих двумерную структуру, лишь небольшая часть которой показана на рис. 2.13,а, дает дифракционную картину из набора пятен, расположенных так, как изображено на рис. 2.13,б. В соответствии с предыдущими результатами масштаб дифракционной картины взаимно связан с масштабом решетки из точечных отверстий и ряды пятен в первой перпендикулярны рядам отверстий во второй. Если $\mathbf{a}$, $\mathbf{b}$ — направляющие векторы дифракционной решетки на рис. 2.13,а, а $\mathbf{a}^*$, $\mathbf{b}^*$ — соответствующие векторы дифракционной картины на рис. 2.13,б, то мы имеем

$$|\mathbf{a}^*| = K/|\mathbf{a}|, \quad |\mathbf{b}^*| = K/|\mathbf{b}|,$$

где K — константа. Кроме того, можно написать

$$\mathbf{a}^* \cdot \mathbf{b} = 0, \quad \mathbf{b}^* \cdot \mathbf{a} = 0.$$

В соответствии с концепцией взаимности маска из точечных отверстий, расположенных как пятна на рис. 2.13,б, должна давать дифракционную картину, подобную показанной на рис. 2.13,а. (Звездочка используется здесь для обозначения направлений во взаимной структуре и не относится к комплексному сопряжению.)

2.7. КРИСТАЛЛЫ В КАЧЕСТВЕ ТРЕХМЕРНЫХ РЕШЕТОК

В кристаллах группа атомов (иногда, как в некоторых металлах, только один атом) тождественно связана с каждым узлом регулярной пространственной решетки. Период решетки обычно составляет от 1 до 10 нм, хотя в кристаллах органических веществ, например, это значение существенно больше, поскольку в них с каждым узлом решетки связано очень много атомов. Поскольку длины волн рентгеновских лучей лежат в пределах приблизительно от 10^{-2} до 10 нм (см. приложение Г) и они рассеиваются (или дифрагируют, эти термины, как указано в разд. 1.3, используются на равных правах) электронными облаками атомов, кристаллы действуют подобно трехмерным дифракционным решеткам

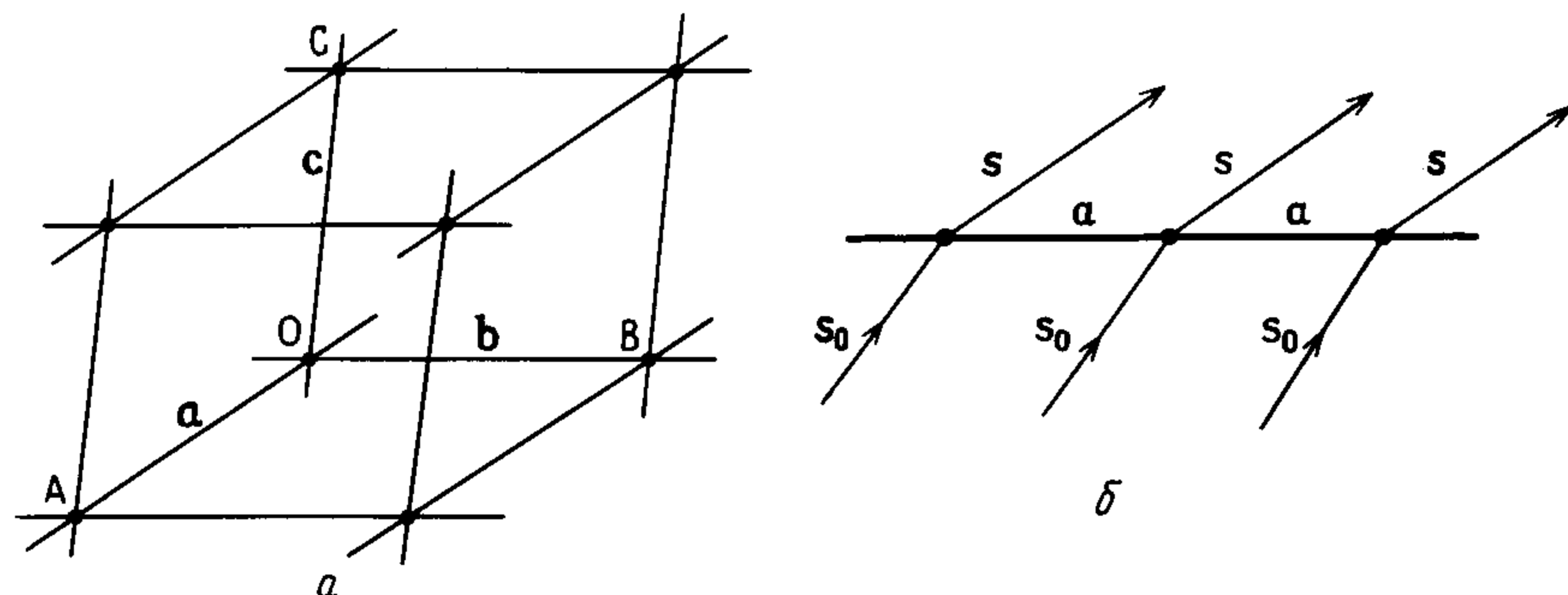


Рис. 2.14. Кристаллическая структура в качестве примера трехмерной дифракционной решетки.

для рентгеновских лучей. Опыт Лауэ, поставленный в 1912 г., продемонстрировал реальность такой интерпретации, позволяя предположить в противовес некоторым воззрениям того времени, что рентгеновские лучи являются волнами или по крайней мере обладают волновыми свойствами. За это Макс фон Лауэ в 1914 г. была присуждена Нобелевская премия по физике. В следующем году премия была присуждена У. Л. Брэггу и его отцу У. Г. Брэггу за то, что они незамедлительно применили эффекты дифракции к анализу кристаллической структуры рентгеновскими лучами, который стал в настоящее время предметом *рентгеновской кристаллографии*. Как мы уже отмечали, изображение атомов нельзя построить непосредственно путем фокусировки рентгеновских лучей. Этот процесс должен выполняться косвенно на основе информации о дифракционных эффектах.

Поскольку кристалл подобен трехмерной решетке, а не одно- или двумерной, то условия, необходимые для возникновения эквивалента главных максимумов в оптической дифракции, удовлетворяются не столь легко. Рассмотрим единичную ячейку кристаллической решетки, изображенную на рис. 2.14, а. Представим, что кристалл пронизывается пучком квазимонохроматических волн с длиной волны λ . Каково основное требование, необходимое для получения дифракционного максимума в некотором направлении? Оно состоит в том, что рентгеновские лучи, рассеянные в данном направлении (идентичными) ансамблями атомов с центрами в узлах решетки А, В и С, должны совпадать по фазе с лучами, рассеянными ансамблем в точке О. Тогда рассеянные этими центрами волны будут находиться в фазе с рассеянными от соседних узлов и так далее по кристаллу. Совсем не обязательно, чтобы в узле решетки располагался только один атом. Это требование не влияет на возможность существования дифракционного максимума, так как все связано с периодом решетки — расстоянием между соответствующими атомами, расположенными одинаково по отношению к последовательным узлам кристаллической решетки. Разумеется, узел решетки,

о которой идет речь, является чисто воображаемой точкой и ее расположение относительно ансамбля атомов произвольно, хотя другие условия, такие, как симметрия, влияют на ее выбор. С другой стороны, *интенсивность* дифракционного максимума зависит от взаимного расположения атомов в ансамбле, и мы обсудим этот вопрос в данном разделе ниже.

Рассмотрим повторяемость узлов решетки в направлении a кристалла (рис. 2.14, б). Для рентгеновских лучей, падающих в направлении, характеризуемом единичным вектором s_0 , излучение, рассеянное от ансамбля, связанного с одним из узлов решетки, в направлении с единичным вектором s , будет находиться в фазе с рассеянием от следующего узла, если

$$a \cdot s - a \cdot s_0 = (s - s_0) \cdot a = h\lambda,$$

где h — целое число или нуль, а a — вектор смещения решетки. Не обязательно, чтобы s или s_0 лежали в плоскости рисунка.

Для идеальной трехмерной решетки одновременно должны выполняться следующие три условия:

$$\left. \begin{aligned} (s - s_0) \cdot a &= h\lambda \\ (s - s_0) \cdot b &= k\lambda \\ (s - s_0) \cdot c &= l\lambda \end{aligned} \right\} \text{ (уравнения Лауэ),} \quad (2.18)$$

где h, k, l — целые числа или нуль, a, b, c — векторы смещения решетки. Заметим, что решетка необязательно должна быть ортогональной.

При фиксировании ориентации кристалла относительно направления падения рентгеновского пучка выполнение вышеприведенных условий более чем для нескольких дифракционных максимумов маловероятно, если только, как в первоначальном опыте, не используется непрерывный спектр рентгеновских волн. Однако отсутствие информации о длине волны, ответственной за какой-либо конкретный дифракционный максимум, является очевидным недостатком. Если не считать исследований определенных типов, теперь в практике повсеместно используется квазимонохроматическое излучение и при просвечивании рентгеновским пучком наклон кристалла постепенно меняется с тем, чтобы обеспечить выполнение условий Лауэ. Здесь нет необходимости касаться детально того, как это осуществляется на практике. Достаточно сказать, что существует возможность получения трехмерной дифракционной картины от кристалла. Она образует трехмерную структуру, которая взаимосвязана со структурой кристалла, как и в случае двумерных решеток, рассмотренных в предыдущем разделе.

Направления дифракционных рентгеновских максимумов от кристалла определяются для заданной длины волны рентгеновских лучей размерами кристаллической решетки. Поэтому измерения этих направле-

ний позволяют выявить геометрию кристаллической структуры. Аналогично *интенсивности* дифракционных максимумов определяются интенсивностью рассеяния в направлениях максимумов от групп атомов, тождественно связанных с каждым узлом решетки кристалла. Таким образом, эти интенсивности содержат информацию о составе и расположении атомов в группе, связанной с каждым узлом кристаллической решетки.

Все сказанное позволяет провести аналогию с процессом, благодаря которому максимумы оптической дифракционной решетки являются усиленной выборкой дифракционной картины только от одной апертуры решетки. Картина же от одной апертуры определяется ее апертурной функцией – оптической структурой апертуры.

На наше счастье мы можем получить эту усиленную выборку рентгеновских лучей, рассеянных от одиночной ячейки кристалла. Если бы даже и удалось изолировать одну ячейку кристалла и работать с ней, рассеянное ею рентгеновское излучение было бы слишком слабым для измерений. В действительности мы имеем дело с необычайно усиленными сигналами реального кристалла. Цена, которую мы платим за это, состоит в том, что сигналы ограничены некоторыми направлениями, определяемыми геометрией кристаллической решетки. Однако число этих направлений достаточно велико и позволяет построить очень детальную картину расположения атомов и распределения электронной плотности в кристаллической структуре. (Дифракция рентгеновских лучей на некристаллических веществах, таких, как стекло и жидкости, также дает информацию об их структуре, но детальное обсуждение этого вопроса остается за рамками настоящей книги.)

В гл. 5 мы увидим, что использование результатов дифракции рентгеновских лучей для определения расположения атомов в кристаллах оказывается, по существу, приложением теории формирования оптического изображения Аббе – Портера. Тесная аналогия между принципами построения изображения кристаллической структуры рентгеновскими лучами и формированием обычного оптического изображения было основным предметом интереса для сэра Лоуренса Брэгга на протяжении почти всей его жизни, предметом, в который он внес значительный вклад.

В связи с вопросами фурье-анализа, которые мы частично связывали с этим, было бы полезно получить выражение для амплитуды максимума рентгеновской дифракции от единичной ячейки кристалла. Это можно сделать с помощью концепции отражения Брэгга, с которой читатель должен быть знаком. (В приложении В показано, что дифракционный максимум, соответствующий определенному набору величин h , k , l в уравнениях Лауэ, можно рассматривать как отражение падающего пучка рентгеновских лучей плоскостями структуры кристалла, заданными теми же значениями h , k , l .)

На рис. 2.15 показан вид с ребра ряда плоскостей решетки, рассеянных на расстояние d . Рентгеновские лучи падают на плоскости под

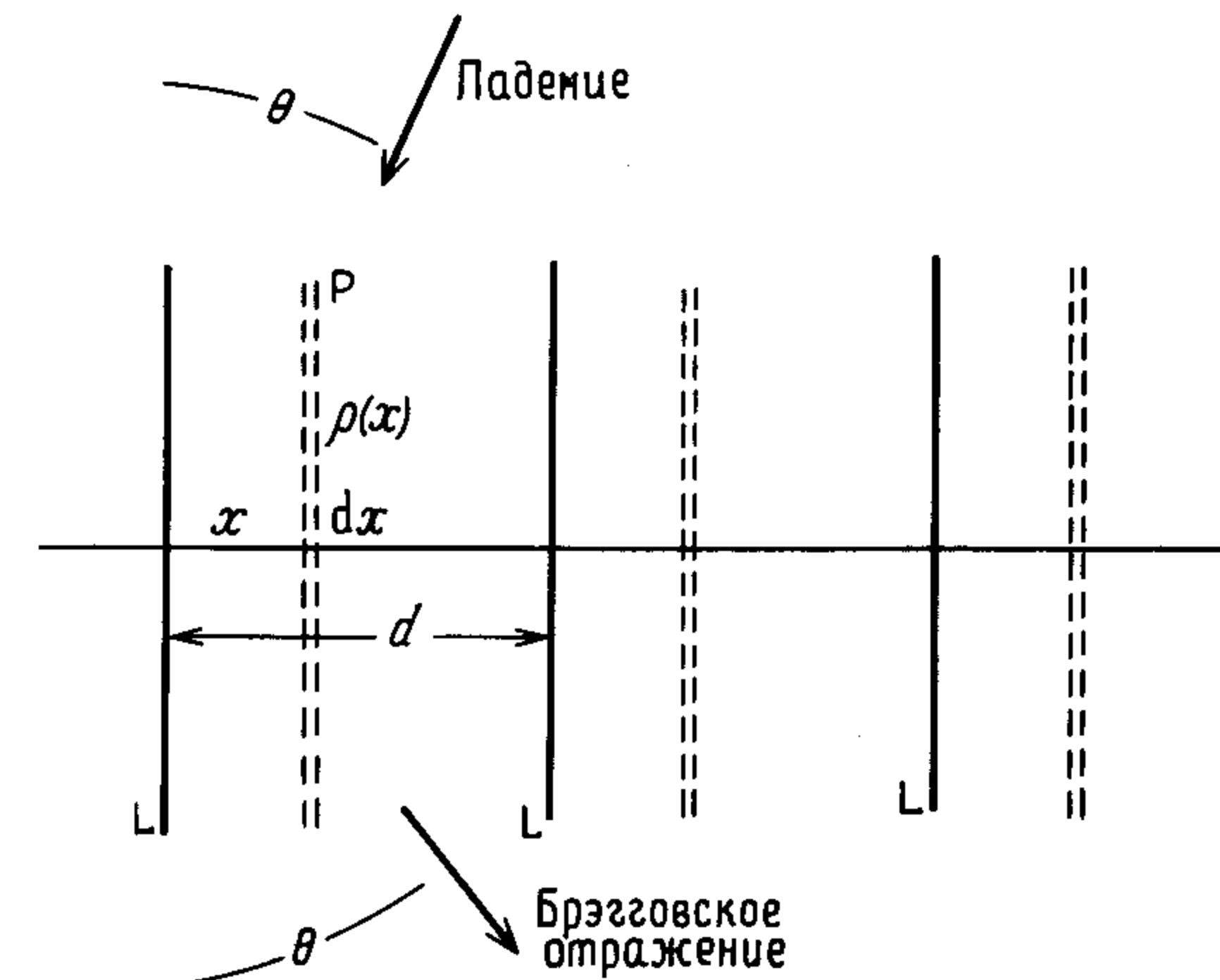


Рис. 2.15. Отражение рентгеновских лучей Брэгга (L – сечение плоскости кристаллической решетки).

углом Брэгга θ , соответствующим используемой длине волны. Рассуждая в данный момент с позиции представлений о точечных атомах, можно заключить, что в любой плоскости структуры атомы синфазно рассеивают в направлении θ ; это и является основой «отражения». Чтобы отражение от последовательных плоскостей кристаллической решетки происходило однофазно, должно выполняться условие Брэгга, а именно

$$2d \sin \theta = n\lambda, \quad (2.19)$$

где n – порядок отражения.

Однако, как мы уже отмечали, атомы не являются точечными рассеивающими элементами и область их распределения не ограничена узлами решетки. Картину решетки следует представлять в виде непрерывного распределения рассеивающего вещества, которое регулярно повторяется по кристаллу. (Вопросы симметрии кристаллической структуры мы здесь не обсуждаем.) Рассмотрим, воспользовавшись рис. 2.15, отражение первого порядка ($n = 1$), связанное с параметром d кристаллической решетки. Как мы отметили выше, рассеяние от части распределения электронной плотности непосредственно в плоскостях кристаллической решетки полностью синфазно в направлении отражения, но мы должны принять во внимание и рассеяние веществом между этими плоскостями. Для отражения первого порядка получается разница на длину волны (разность фаз 2π) между излучением, рассеянным последовательными плоскостями решетки. Поэтому вклад в то же самое рассеянное излучение от рассеивающего вещества в плоскости P будет отличаться по фазе на $(x/d)2\pi$. Если $\rho(x)$ представляет собой амплитуду рассеянного излучения от слоя толщиной dx на расстоянии x , то полная

амплитуда такого излучения, отраженного от всего рассеивающего вещества на одном периоде структуры решетки, дается выражением

$$A_1(\theta) = \int_0^d \rho(x) \exp[i(x/d)2\pi] dx. \quad (2.20)$$

Для отражения второго порядка при отражении от последовательных плоскостей получается разность длины пути в две длины волны, и тогда разность фаз для плоскости Р будет равна $(2x/d)2\pi$. Аналогично для отражения n -го порядка будет существовать разность фаз $(nx/d)2\pi$ и соответствующее уравнение принимает вид

$$A_n(\theta) = \int_0^d \rho(x) \exp[i(nx/d)2\pi] dx, \quad (2.21)$$

где θ связано с d уравнением Брэгга (2.19).

Конечно, «отражение» может иметь место с каждой стороны последовательности плоскостей и амплитуды обычно равны (закон Фриделя). Поэтому можно считать, что это уравнение описывает амплитуду пары отражений, или, что то же самое, амплитуду пар дифракционных максимумов порядка $\pm n$.

Глава 3

Ряды Фурье и периодические структуры

«Теорема Фурье не только является одним из самых прекрасных результатов современного анализа, но и, можно сказать, представляет собой совершенно незаменимый инструмент для решения почти всех неясных проблем новейшей физики».

Лорд Кельвин

3.1. ВВЕДЕНИЕ

В гл. 2 на частном примере было показано, что результат дифракции от периодических объектов в форме оптической решетки определяется «структурой» решетки, характеризующей ее апертурной функцией. То же самое оказывается верным и для результатов по дифракции рентгеновских лучей, полученных из исследования расположения атомов, образующих периодическую структуру кристалла, подобную решетке. Мы отмечали также, что оптическая дифракция является промежуточным шагом в формировании изображений с помощью линзы. При этом линза выполняет задачу сведения дифрагированного света в плоскости изображения. При работе с рентгеновскими лучами линза непригодна, и для формирования изображения структурного расположения атомов в кристалле при воздействии рентгеновских лучей должны использоваться другие, нежели дифракция, способы построения изображения.

Для более глубокого понимания этих вопросов, а также по другим причинам, описанным в гл. 5, нам нужно больше узнать о соотношении между дифракционной картиной и структурой создающего ее объекта. На этой стадии методы фурье-анализа начинают играть особенно важную роль.

Ограничимся пока периодическими структурами, такими, как многоапертурные решетки или кристаллы, полную апертурную функцию (или «структуру») которых математически можно построить путем суммирования бесконечных рядов из синусоидальных гармоник – рядов Фурье, названных в честь Ж.Б.Ж. Фурье, пионера этого математического метода.

Утверждение, что периодическая функция может быть выражена в виде суммы последовательности гармоник, известно, как теорема Фурье. Большинство функций подчиняется ей, и отдельные исключения, хотя ни одно из них и не упоминается в этой книге, могут быть найдены в учебниках математики.

Строго говоря, дифракционная картина, которую можно представить рядами Фурье, должна повторяться бесконечное число раз. В этой связи в приложениях, которые нам будут встречаться, число повторений должно быть достаточным для правомочности использования рядов. На практике в большинстве используемых решеток число повторов до-

статочно велико. Например, в кристалле с линейным размером 1 мм присутствует более одного миллиона повторений «элементарной единичной картины», характерной для структуры расположения атомов.

В этой главе мы рассмотрим замечательный факт, состоящий в том, что дифракционная картина, создаваемая решеткой, представляет собой проявление, почти физическое воплощение, тех гармоник, которые составляют математическое описание структуры решетки.

Дифракция является промежуточной стадией в формировании оптического изображения. Это означает, в частности, что на стадии дифракции мы можем путем расчета управлять процессом образования изображения. Указанное обстоятельство определяет многие аспекты оптической обработки (гл. 5). Другим главным следствием стала разработка методов определения атомной структуры кристаллов независимо от их сложности по результатам рентгеновской дифракции.

В гл. 4 соотношения, относящиеся к периодическим структурам, распространяются также и на непериодические. Но сначала о некоторых свойствах рядов Фурье (разд. 3.2 и 3.3) и их связи с дифракцией на непериодических структурах (разд. 3.4).

3.2. РЯДЫ ФУРЬЕ

Использование рядов Фурье для представления периодической картины, такой, как оптическая структура решетки, можно проиллюстрировать рис. 3.1, а, где показано только два периода. Этот частный случай выбран оттого, что он дает возможность ясно продемонстрировать, как его можно представить рядами Фурье, а также с учетом возможности использовать его для объяснения в дальнейшем интересного примера (разд. 3.4.2).

Ряд Фурье для одномерной периодической функции $f(x)$ может быть записан как

$$\begin{aligned} f(x) &= \frac{A_0}{2} + A_1 \cos\left(\frac{2\pi x}{D} - \alpha_1\right) + A_2 \cos\left(\frac{2\pi x}{D/2} - \alpha_2\right) + \\ &= \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2\pi n x}{D} - \alpha_n\right) \end{aligned} \quad (3.01)$$

т.е. в виде бесконечного ряда гармонических «волн» с амплитудами A_n (коэффициенты Фурье), фазами α_n и постепенно уменьшающимися «длинами волн» D/n , называемыми в дальнейшем «периодами» во избежание путаницы с оптическими длинами волн. Величина $A_0/2$ является постоянной, а множитель $1/2$ служит только для упрощения в дальнейшем записи некоторых уравнений.

Аналитическая процедура определения коэффициентов для любой заданной функции $f(x)$ описана в разд. 3.3. Между тем рис. 3.1, б по-

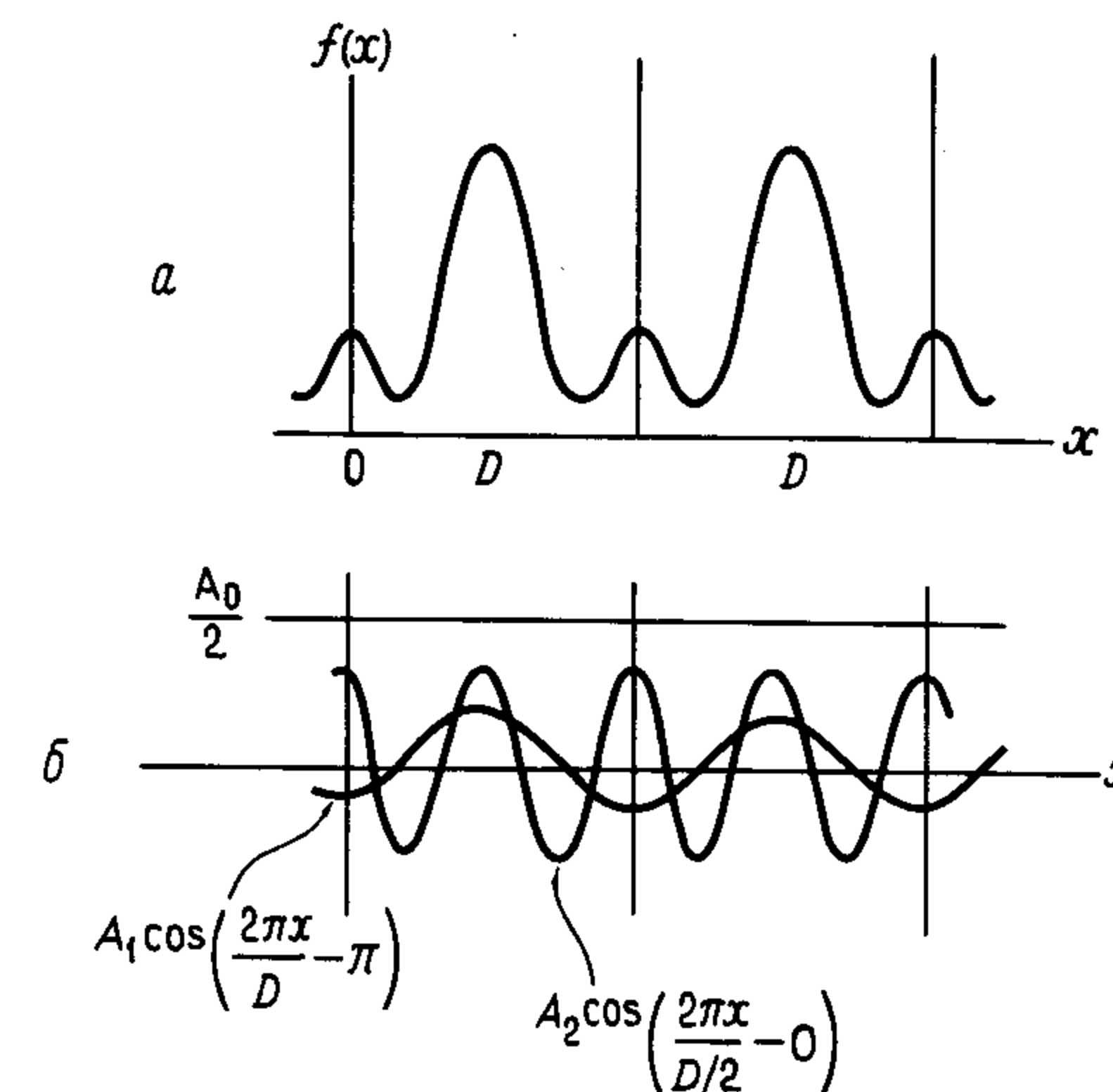


Рис. 3.1. Представление периодической функции $f(x)$ рядом Фурье.

казывает графически, как можно распознать, по крайней мере качественно, первые несколько членов *фурье-разложения* конкретной картины. На этом рисунке мы видим, что член с $n = 1$, т.е. $A_1 \cos((2\pi x/D) - \alpha_1)$, подобен до некоторой степени $f(x)$ при $\alpha_1 = \pi$. Величина α_1 может принимать значения 0 или π , однако $f(x)$ в этом примере является четной функцией, и $\alpha_1 = 0$ ухудшило бы сходство. В грубом приближении представлены по крайней мере наибольшие максимумы в картине, но не слабые. Член, соответствующий $n = 2$ с $\alpha_2 = 0$, помогает существенно улучшить сходство. Добавленный к члену с $n = 1$ со значением своей амплитуды A_2 он начинает выделять слабые максимумы и усиливать сильные. Добавление более высоких гармоник постепенно улучшает соответствие. Однако до этого момента суммирование распределяется равным образом выше и ниже оси x . Постоянный член $A_0/2$ в уравнении (3.01) как раз и поднимает характеристику над осью x до истинного уровня кривой $f(x)$.

Значения α_n в уравнении (3.01) являются «фазами» гармоник. Например, фаза $\alpha_1 = \pi$ означает, что косинусный член при $n = 1$ сдвигается вдоль оси x на величину $D/2$ по сравнению с косинусоидой с нулевой фазой.

Суммирование гармоник, *фурье-синтез*, формально справедливо до $n = \infty$. Практически, однако, менее сотни членов в большинстве случаев вполне достаточно, хотя можно привести примеры, когда их требуется больше.

Структура оптической дифракционной решетки может быть представлена рядами Фурье таким же образом, как и для картины, которую мы только что рассматривали. Например, как расширение понятия «апертурная функция» для одиночной щели (разд. 2.2) на рис. 3.3, а апертур-

ная функция однородно освещенной многослойной решетки может быть представлена функцией $f(x)$. Представление ее в виде рядов Фурье и чрезвычайно важное соотношение между рядами и дифракционной картиной, создаваемой решеткой, рассматриваются в разд. 3.4. С этой целью в следующем разделе вводится метод, используемый для количественной оценки членов рядов Фурье.

В связи с использованием нами рядов Фурье при исследовании дифракции полезно расширить наше представление о частотах гармоник, а также об их периодах. Подобно тому, как периодичность во времени T соответствует временной частоте $1/T$, так и период пространственной гармоники D/n соответствует пространственной частоте n/D , которая представляет собой число повторений на единице масштаба картины.

3.3. НАХОЖДЕНИЕ КОЭФФИЦИЕНТОВ ФУРЬЕ: ЧЕТНЫЕ ФУНКЦИИ

Для четных функций, рассмотренных в предыдущем разделе, уравнение (3.01) можно записать в виде

$$f(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos \frac{2\pi nx}{D}, \quad (3.02)$$

если считать, что A_n необязательно положительно.

Умножая обе части уравнения на $\cos(2\pi hx/D)$, где $h = 0$ или является целым числом, и интегрируя по периоду D , получаем

$$\begin{aligned} \int_D f(x) \cos \frac{2\pi hx}{D} dx &= \frac{A_0}{2} \int_D \cos \frac{2\pi hx}{D} dx + \int_D \sum_{n=1}^{\infty} A_n \cos \frac{2\pi nx}{D} \cos \frac{2\pi hx}{D} dx = \\ &= \frac{A_0}{2} \int_D \cos \frac{2\pi hx}{D} dx + \frac{1}{2} \left[\int_D \sum_{n=1}^{\infty} A_n \cos \frac{2\pi(n+h)x}{D} dx + \right. \\ &\quad \left. + \int_D \sum_{n=1}^{\infty} A_n \cos \frac{2\pi(n-h)x}{D} dx \right]. \end{aligned}$$

Интеграл от каждого члена в первой сумме в правой части этого уравнения равен нулю, поскольку n — целое число и $h = 0$, или является целым числом. То же самое относится и ко второй сумме, за исключением одного члена при $h = n$, который равен $\int_D A_n \cos 0 dx = A_n D$. Первый

член в правой части также равен нулю при $h = n \neq 0$, поэтому все уравнение сводится к выражению вида

$$\int_D f(x) \cos \frac{2\pi nx}{D} dx = \frac{1}{2} [0 + A_n D],$$

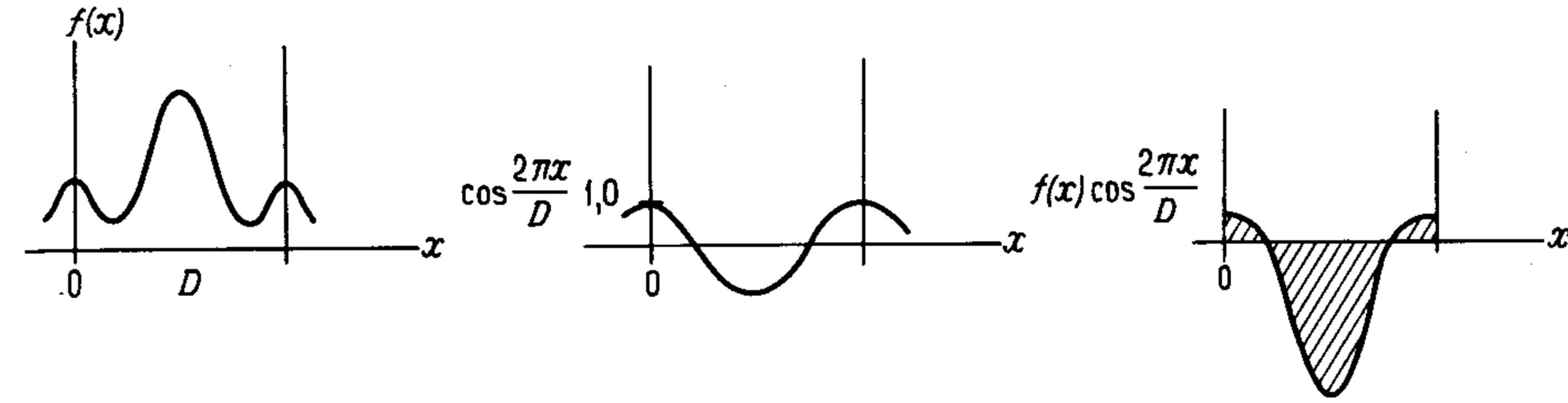


Рис. 3.2. Графическое представление выражения $A_1 = \int_D f(x) \cos(2\pi x/D) dx$.

откуда получаем

$$A_{n \neq 0} = \frac{2}{D} \int_D f(x) \cos \frac{2\pi nx}{D} dx.$$

При $h = 0$ только первый член не равен нулю, откуда

$$\int_D f(x) dx = \frac{1}{2} A_0 D,$$

следовательно,

$$A_0 = \frac{2}{D} \int_D f(x) dx.$$

Теперь понятна причина выбора произвольного множителя $1/2$ в постоянном члене уравнения (3.01): это позволяет объединить два приведенных выше выражения для коэффициентов

$$A_n = \frac{2}{D} \int_D f(x) \cos \frac{2\pi nx}{D} dx, \quad (3.03)$$

где $n = 0, 1, 2, \dots$

Выражение для постоянного члена $A_0/2$ является просто средним значением $f(x)$ по всему единичному отрезку картины с периодом D . Этим подтверждается и количественно определяется полученное в разд. 3.2 заключение, что $A_0/2$ представляет собой постоянную величину, на которую сумма должна подняться над осью x , чтобы достигнуть соответствующей высоты профиля $f(x)$.

Любопытна интерпретация уравнения (3.03) для случая $n \neq 0$. Как показано графически на рис. 3.2 (хотя обычно это иллюстрируется расчетом), если мы хотим выяснить, присутствует ли гармоника определенной пространственной частоты n/D в представлении данной функции рядом Фурье, то надо умножить единичную картину функции на гармо-

нику этой частоты (единичной амплитуды и нулевой фазы) и затем проинтегрировать по D . Нулевой результат означает, что гармоника в функции отсутствует. Ответ, отличный от нуля, указывает на то, что она есть, и значение, полученное при интегрировании (то есть площадь под кривой произведения) после умножения на $2/D$ в соответствии с уравнением (3.03), дает ее амплитуду. Фаза гармоники определяется знаком полученной величины и зависит от того, является результирующая площадь под кривой произведения положительной или же отрицательной.

Для наших целей удобнее переписать уравнение (3.02) в следующем виде:

$$f(x) = \sum_{n=-\infty}^{+\infty} \frac{A_n}{2} \cos \frac{2\pi n x}{D}. \quad (3.04)$$

Член при $n=0$ тот же, что и раньше. Для других значений коэффициенты образуют пары $(1/2) A_{+n}$ и $(1/2) A_{-n}$, члены которых равны $[\cos(-x) = \cos(+x)]$ и имеют одинаковые пространственные частоты. Следовательно, они складываются, давая ту же амплитуду гармоники, как и в случае, когда имелись только положительные значения n . Отсюда следует, что уравнение (3.03) дает коэффициенты $(1/2) A_n$ и в уравнении (3.04) для всех n , т.е.

$$\frac{A_n}{2} = \frac{1}{D} \int_D f(x) \cos \frac{2\pi n x}{D} dx \quad (3.05)$$

Приведение ряда к форме, где представлены пары гармоник с амплитудой, равной половине первоначальной, но с учетом как отрицательных, так и положительных значений n , вовсе не случайно. В следующем разделе показано, что этим парам гармоник физически соответствуют пары дифракционных максимумов порядка $\pm n$, определяемые решеткой.

Общие выражения $f(x)$ для случая нечетной функции приведены в разд. 3.5.

3.4. ОПТИЧЕСКИЕ И КРИСТАЛЛИЧЕСКИЕ

ДИФРАКЦИОННЫЕ РЕШЕТКИ:

ФИЗИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ ЧЛЕНОВ ФУРЬЕ

3.4.1. Оптическая дифракция

Решетка, апертурная функция (или, как мы ее называли, оптическая структура) которой представлена функцией $f(x)$ на рис. 3.3, а, имеет щели шириной a и период повторения D . Амплитуда освещенности на щелях (постоянная) равна h .

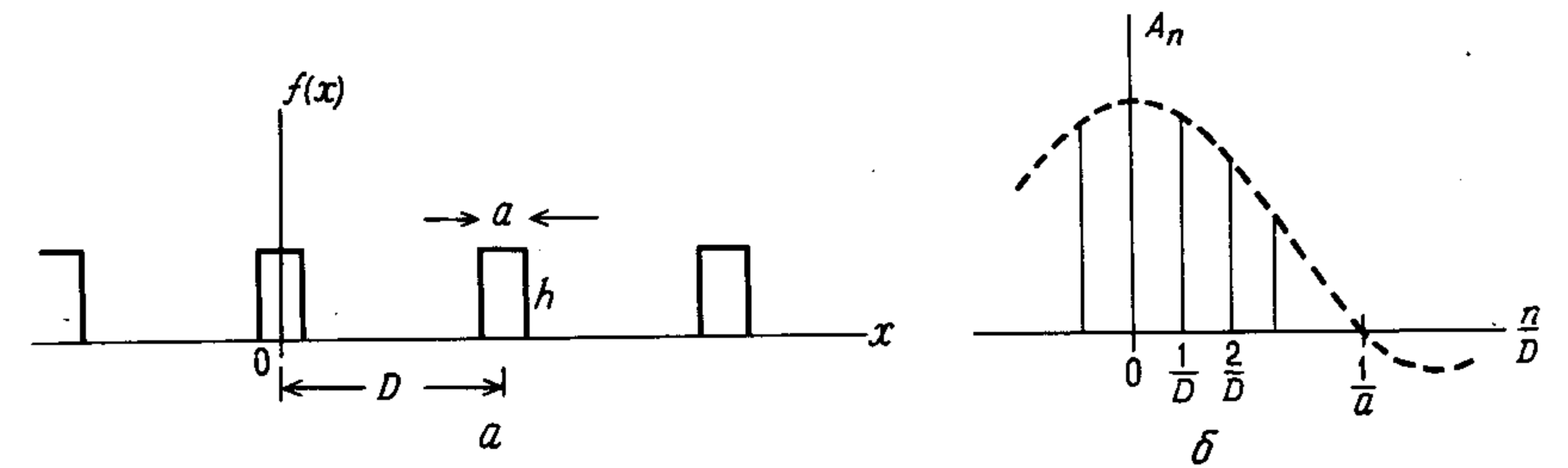


Рис. 3.3. Фурье-анализ периодической прямоугольной функции.

Фурье-разложение $f(x)$ получается подстановкой этих параметров в уравнение (3.05) следующим образом:

$$\frac{A_n}{2} = \frac{1}{D} \left[\int_0^{a/2} h \cos \frac{2\pi n x}{D} dx + \int_{D-a/2}^D h \cos \frac{2\pi n x}{D} dx \right],$$

откуда

$$\frac{A_n}{2} = \frac{ha}{D} \left[\frac{\sin \frac{\pi na}{D}}{\frac{\pi na}{D}} \right]. \quad (3.06)$$

Будучи нанесенными на рис. 3.3, б, они дают амплитуды и фазы пар $(\pm n)$ гармоник $f(x)$ с периодом D/n (пространственная частота n/D). И на графике, и в уравнении (3.06) в них также легко распознать выборочную *sinc*-функцию, которая несет информацию об амплитудах и фазах пар дифракционных максимумов или о порядке $\pm n$, определяемых решеткой, апертурная функция которой равна $f(x)$ (сравните значения интенсивности в уравнении (2.17)). Направления же этих дифракционных максимумов даются пространственными частотами гармоник Фурье, так как

$$\frac{n}{D} = \frac{\sin \theta}{\lambda} (= u). \quad \{2.10\}$$

Приведенные выше соображения применимы практически к любому типу апертурной функции, хотя огибающая коэффициентов, как правило, не является *sinc*-функцией.

Итак, мы имеем общий важный результат, что амплитуда пары n -го порядка дифракционных максимумов решетки является мерой амплитуды пары гармоник n -го порядка, составляющих ее полную апертурную функцию — то, что мы называли ее оптической структурой. Направление

ние этих дифракционных максимумов $(\sin \theta)/\lambda$, дается гармоникой пространственной частоты (n/D) (или периодом гармоник D/n).

По этим причинам плоскость дифракционной картины называют *фурье-плоскостью* (или *фурье-пространством*), или, иначе, *плоскостью частот* (пространственных). Кроме того, как впервые отмечено в гл. 2, можно также использовать термин *взаимное пространство* ввиду существования соотношения взаимности между масштабом дифракционной системы и создаваемой ею картиной. Каждая из этих интерпретаций имеет свои специфические особенности и область применения, например «частотное пространство», широко используется в обработке оптических данных (разд. 5.5).

3.4.2. Дифракция на кристаллах

В разд. 2.7 было показано, что амплитуда $A_n(\theta)$ брэгговского отражения n -го порядка на плоскостях кристаллической решетки, расположенных на расстоянии d , определяется соотношением

$$A_n(\theta) = \int_{x=0}^{x=d} \rho(x) \exp[i(n\pi x/d) 2\pi] dx, \quad (2.21)$$

где θ определяется d согласно уравнению Брэгга.

Применим это соотношение к кристаллической структуре хлорида цезия CsCl (рис. 3.4). Распределение плотности электронов структуры в направлении одного из ребер куба может быть представлено кривой $f(x)$, показанной на рис. 3.1, а; она была выбрана специально для этой цели. Слабые пики на этой кривой соответствуют электронному облаку ионов хлора, а большие пики – ионов цезия. (В структуре одинаковое число ионов каждого типа.)

Если мы отбросим разные факторы, зависящие от угла (или других причин), то в этом приближении можем предположить, что амплитуда рассеяния рентгеновских лучей $\rho(x)$ в вышеприведенном уравнении равна $f(x)$, и, поскольку данная конкретная функция $f(x)$ является четной, мы можем переписать это уравнение в виде

$$A_n(\theta) = \int_{x=0}^{x=d} f(x) \cos \frac{2\pi n x}{d} dx \quad (3.07)$$

Сравнение с уравнением (3.05) показывает, что эти значения амплитуд дифракции пропорциональны амплитудам гармонических членов фурье-разложения распределения плотности электронов между отражающими плоскостями в кристаллической структуре.

Такое же соотношение применимо к трехмерной кристаллической структуре, и оно легко распространяется на случай, когда центральная симметрия отсутствует. Таким образом, мы имеем весьма важный ре-

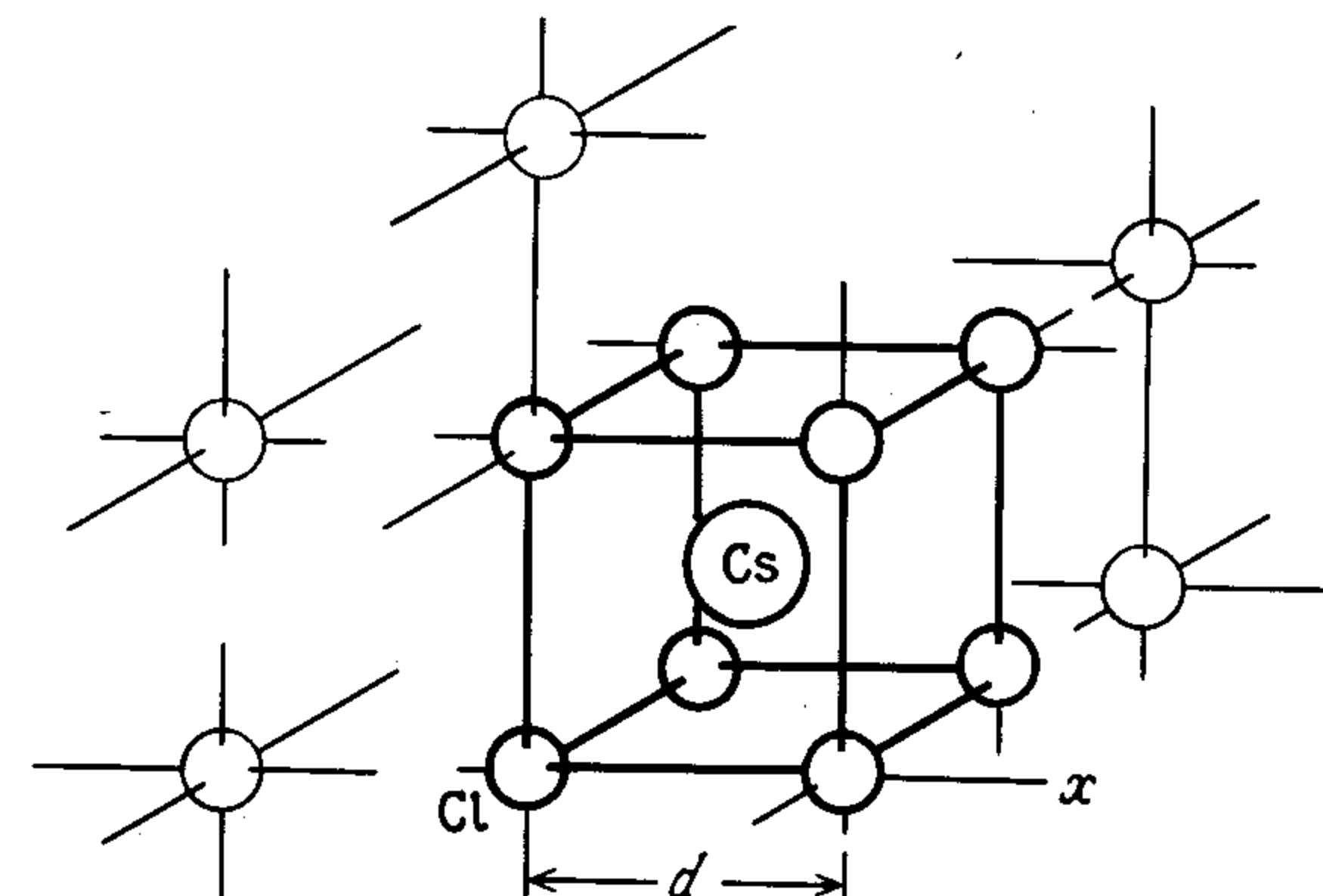


Рис. 3.4. Кристаллическая структура хлорида цезия CsCl.

зультат фундаментального значения, состоящий в том, что измерение рентгеновских дифракционных максимумов – более удачно названных брэгговскими отражениями – обеспечивает информацию о членах разложения Фурье кристаллической структуры точно так же, как было со «структурой» оптической дифракционной решетки. Экспериментальные трудности заключаются в том, что, поскольку могут быть измерены только интенсивности дифрагированных рентгеновских лучей, фазы фурье-членов остаются неизвестными и фурье-синтез нельзя осуществить прямым путем (разд. 5.3.3).

Замечания в конце предыдущего раздела, касающиеся способов описания пространства, в котором образуется дифракционная картина, в равной степени применимы и здесь. Понятие взаимного пространства особенно полезно в рентгеновской кристаллографии.

3.5. РЯДЫ ФУРЬЕ: ОБЩИЕ ОПРЕДЕЛЕНИЯ

Если $f(x)$ является периодической функцией общего вида, α_n в уравнении (3.01) может принимать значения от 0 до 2π . В таком случае наиболее часто используются следующие две альтернативные формы представления рядов.

3.5.1. Ряды по синусам и косинусам

Разложим уравнение (3.01) следующим образом:

$$f(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty} \left[A_n \cos \alpha_n \cos \frac{2\pi n x}{D} + A_n \sin \alpha_n \sin \frac{2\pi n x}{D} \right].$$

Результат можно записать в виде

$$f(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi n x}{D} + \sum_{n=1}^{\infty} b_n \sin \frac{2\pi n x}{D}, \quad (3.08)$$

где $a_n = A_n \cos \alpha_n$, $b_n = A_n \sin \alpha_n$. Исходный ряд косинусов с разными фазами заменен комбинацией косинусов и синусов [сравните $2\cos(2\pi nx/D - \pi/3) = 1\cos 2\pi nx/D + \sqrt{3}\sin 2\pi nx/D$].

Следующие ниже выражения для новых коэффициентов получены таким же способом, как ранее в одиночных рядах:

$$\begin{aligned} A_0 &= \frac{2}{D} \int_D f(x) dx, \\ a_n &= \frac{2}{D} \int_D f(x) \cos \frac{2\pi nx}{D} dx, \\ b_n &= \frac{2}{D} \int_D f(x) \sin \frac{2\pi nx}{D} dx \end{aligned} \quad (3.09)$$

Читатель может убедиться, что для четной функции все $b_n = 0$ и аналогично для нечетной функции все $a_n = 0$.

3.5.2. Выражение в экспоненциальном виде

Для действительной функции $f(x)$ мы можем воспользоваться уравнением (3.01) и переписать его в виде

$$\begin{aligned} f(x) &= \frac{A_0}{2} + \sum_{n=1}^{\infty} \frac{1}{2} A_n \exp \left[i \left(\frac{2\pi nx}{D} - \alpha_n \right) \right] + \\ &+ \sum_{n=1}^{\infty} \frac{1}{2} A_n \exp \left[-i \left(\frac{2\pi nx}{D} - \alpha_n \right) \right]. \end{aligned}$$

Знак минус перед i во второй сумме может быть приписан n , что позволяет выполнять суммирование для отрицательных значений n . Следовательно, уравнение может быть преобразовано к виду

$$\begin{aligned} f(x) &= \frac{A_0}{2} + \sum_{n=1}^{\infty} \frac{1}{2} A_n \exp \left[i \left(\frac{2\pi nx}{D} - \alpha_n \right) \right] + \\ &+ \sum_{n=-1}^{-\infty} \frac{1}{2} A_n \exp \left[i \left(\frac{2\pi nx}{D} + \alpha_n \right) \right], \end{aligned}$$

что кратко можно записать, как

$$f(x) = \sum_{n=-\infty}^{\infty} C_n \exp \left(\frac{i2\pi nx}{D} \right). \quad (3.10)$$

Комплексные коэффициенты C_n содержат в себе A_n и α_n , как показано ниже:

$$\left. \begin{aligned} C_{n>0} &= \frac{1}{2} A_n \exp(-i\alpha_n), \\ C_{n<0} &= \frac{1}{2} A_n \exp(i\alpha_n), \\ C_{n=0} &= \frac{1}{2} A_0. \end{aligned} \right\} \quad (3.11)$$

В результате следует, что

$$\left. \begin{aligned} C_{n<0} &= C_{n>0}^*, \\ \text{и поэтому} \\ |C_{n<0}| &= |C_{n>0}|, \end{aligned} \right\} \quad (3.12)$$

где звездочка означает комплексное сопряжение.

Коэффициенты, полученные таким же образом, как в рядах по косинусам, дают нам пару уравнений

$$f(x) = \sum_{n=-\infty}^{\infty} C_n \exp \left(\frac{i2\pi nx}{D} \right) \quad (3.10)$$

$$C_n = \frac{1}{D} \int_D f(x) \exp \left(-\frac{i2\pi nx}{D} \right) dx. \quad (3.13)$$

Для примера рассмотрим ряды Фурье в экспоненциальном виде для функции $f(x) = \cos(2\pi 5x/D)$. Конечно, это выражение можно было бы подставить в уравнение (3.13) и получить коэффициенты Фурье C_n . Однако ответ можно получить и проверкой, поскольку

$$\cos \frac{2\pi 5x}{D} = \frac{1}{2} \left[\exp \left(\frac{i2\pi 5x}{D} \right) + \exp \left(-\frac{i2\pi 5x}{D} \right) \right].$$

Экспоненциальное представление ряда для $f(x)$ включает только эти два члена, а именно члены, для которых $n = \pm 5$ с коэффициентами $1/2$.

Аналогично если $f(x) = \sin(2\pi 5x/D)$, то только члены с $n = \pm 5$ не равны нулю, а коэффициенты соответственно равны $\pm 1/2i$. Рис. 3.5 показывает эти два примера в графическом виде.

Уравнение (3.13) было получено из уравнения (3.01), т.е. при действительной $f(x)$, когда мнимые члены, возникают только таким образом,

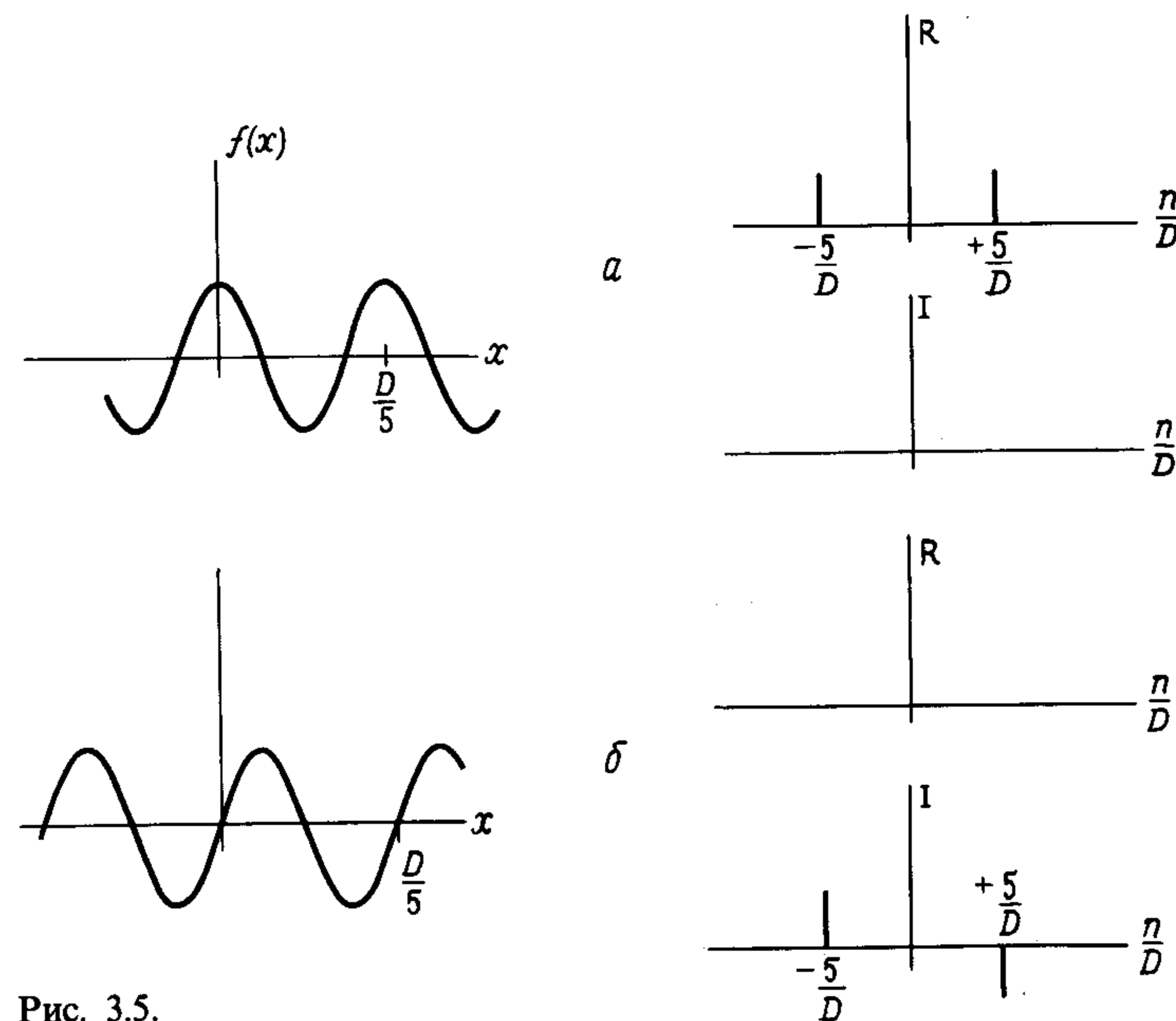


Рис. 3.5.

как было показано выше, и полное выражение все еще оказывается действительным.

Приведенные выше примеры иллюстрируют также условие, налагаемое уравнением (3.12). Таким образом, для четных действительных функций пары коэффициентов действительны и имеют одинаковый знак, тогда как для действительных нечетных функций пары комплексно сопряжены и имеют противоположные знаки.

Для комплексных функций уравнение (3.12) не выполняется. Это легко иллюстрируется при рассмотрении функции вида

$$\sin \frac{2\pi 5x}{D} + \frac{i}{2} \cos \frac{2\pi 5x}{D},$$

коэффициенты которой равны $C_5 = -i/4$, $C_{-5} = 3i/4$.

В качестве последнего примера экспоненциальных рядов Фурье, читатель может получить коэффициенты для периодической прямоугольной функции, показанной на рис. 3.3, а. Можно показать, что подстановка параметров этой функции в уравнение (3.13) дает

$$C_n = \frac{ha}{D} \left[\frac{\sin \frac{\pi na}{D}}{\frac{\pi na}{D}} \right] \quad (3.14)$$

Здесь, как и раньше, получается дискретная sinc-функция [уравнение (3.06)]. Поскольку функция действительная и четная, коэффициенты действительны и попарно равны, как получено в разд. 3.3 с помощью разложения по косинусам.

3.5.3. Пространство и время

Описанные в этой главе ряды Фурье легко распространяются на двух- и трехмерные случаи. Аналогичным образом, когда анализируемые картины являются функциями времени, а не пространства, коэффициенты разложения Фурье относятся к временным частотам, а не к пространственным. Соответствующие уравнения в экспоненциальном виде выглядят тогда следующим образом:

$$f(t) = \sum_{n=-\infty}^{\infty} C_n \exp \left(-\frac{i2\pi nt}{T} \right), \quad (3.15)$$

$$C_n = \frac{1}{T} \int_T f(t) \exp \left(-\frac{i2\pi nt}{T} \right) dt. \quad (3.16)$$

Например, если использованная нами периодическая прямоугольная функция представляет собой поток импульсов с длительностью a и временем повторения T , фурье-члены образуют спектр частот, необходимый для генерации импульсов.

В разд. 4.6 использование методов анализа Фурье в применении к одиночному временному импульсу позволяет интерпретировать спектр частот для цугов волн, связанных с фотонами света.

4.1. ВВЕДЕНИЕ

В гл. 3 гармонические члены фурье-разложения оптической структуры объекта в виде многоапертурной решетки отождествлены с деталями создаваемой ею дифракционной картины. Аналогичная связь существует между единичной апертурой и ее дифракционной картиной, и эта связь нуждается в дальнейшем исследовании. Речь идет о *фурье-преобразовании*, включающем в себя, как и следовало ожидать для непериодической картины, не ряды, а интеграл. Этот вопрос рассматривается в разд. 4.2.

Преобразование Фурье играет также другую важную роль в физической оптике. Трудно переоценить его значение и для физики в целом. Эта глава посвящена возможностям, которые открывает преобразование Фурье, обеспечивающее более глубокое понимание соотношения между дифракционной картиной, создаваемой многоапертурной дифракционной системой, такой, как решетка или кристалл, и ее (полной) апертурной функцией или структурой. Основные идеи этого подхода представлены в разд. 4.3–4.5 для различных применений в гл. 5 в связи с формированием и обработкой изображения. В разд. 4.3 мы рассматриваем дифракционную картину решетки и в разд. 4.4 – ее апертурную функцию. Последняя обсуждается на языке свертки – т. е. на основе другой концепции и математической процедуры, широко используемой в физике. В разд. 4.5 как пример теоремы свертки совместно представлены две стороны соотношения – апертурная функция решетки и дифракционная картина, создаваемая ею.

В разд. 4.6 обсуждается нечто совершенно иное, а именно роль преобразования Фурье в соотношении длины волнового цуга света с его спектральным составом. Это становится важным, когда в гл. 6 мы обращаемся к анализу источников излучения с помощью интерферометрии.

Наконец, в разд. 4.7 мы вводим понятие корреляции – операции, занимающей существенное место в содержании гл. 5 и 6.

4.2. ПРЕОБРАЗОВАНИЕ ФУРЬЕ
И ДИФРАКЦИЯ НА ОДИНОЧНОЙ ЩЕЛИ

Вначале обратимся вновь к периодической прямоугольной функции (рис. 4.1, а). При выбранном начале отсчета она является четной функ-

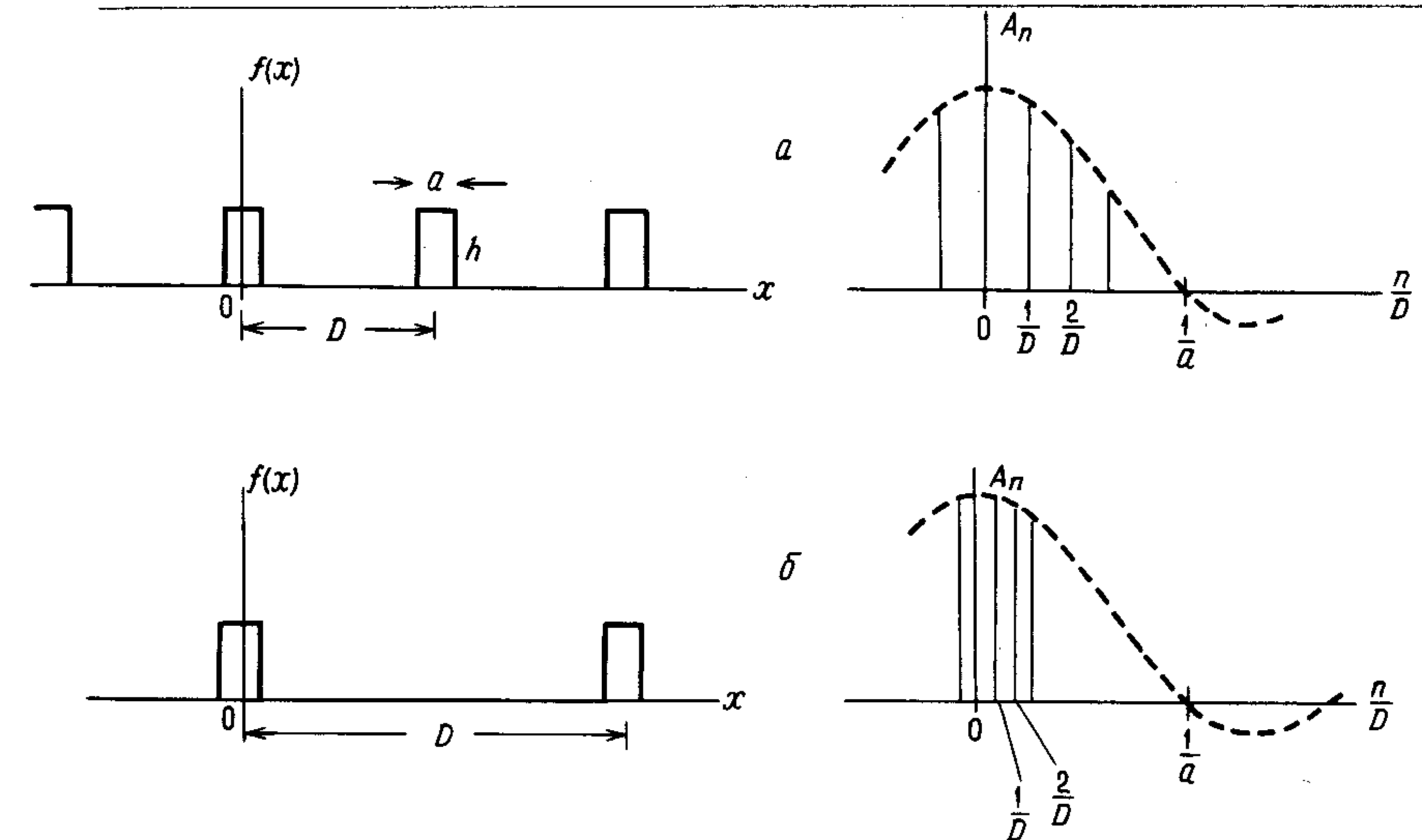


Рис. 4.1. Фурье-анализ периодической прямоугольной функции.

цией, и, как было отмечено в разд. 3.4.2, если ее представить рядом Фурье

$$f(x) = \sum_{n=-\infty}^{+\infty} \frac{A_n}{2} \cos \frac{2\pi n x}{D}, \quad \{3.04\} \quad (4.01)$$

то в этом случае коэффициенты

$$\frac{A_n}{2} = \frac{1}{D} \int_D f(x) \cos \frac{2\pi n x}{D} dx \quad \{3.05\} \quad (4.02)$$

принимают вид

$$\frac{A_n}{2} = \frac{ha}{D} \left[\frac{\sin(\pi na/D)}{\pi na/D} \right]. \quad \{3.06\} \quad (4.03)$$

Это последнее выражение содержит член, который мы называли sinc-функцией, и она снова показана на рис. 4.1, а. Кроме того, если $f(x)$ представляет апертурную функцию всей дифракционной решетки с периодом D , то мы отождествляем ее с амплитудой света от каждой щели только в направлении $(\pm \theta)$ максимумов решетки (порядка $\pm n$). Это дифракционная картина на одиночной щели $(\sin \pi u)/(\pi u)$, причем u ограничено характерными значениями, задаваемыми выражением

$$u = \frac{\sin \theta}{\lambda} = \frac{n}{D}, \quad \{2.10\} \quad (4.04)$$

где $D \sin \theta = n\lambda$ будет условием для максимумов решетки (разд. 2.5).

Следует также напомнить, что гармоники в фурье-разложении (уравнение 4.03) апертурной функции решетки интерпретируются как пространственные частоты n/D . Уравнение 4.04 определяет направления $(\pm \theta)$ главных дифракционных максимумов решетки.

Рассмотрим теперь, что произойдет, если D увеличится, как показано на рис. 4.1, б. Пространственные частоты и, следовательно, направления дифрагировавших лучей становятся ближе друг к другу. Это означает, что для предельного случая неперiodической функции возможны все частоты. Разумеется такой вывод согласуется с нашими представлениями о дифракции: дифракционная картина одиночной апертуры является непрерывной функцией от u .

В уравнениях (4.01) и (4.02) n и D тогда теряют свой смысл, и мы переходим к интегралам

$$f(x) = \int_{-\infty}^{+\infty} F(u) \cos 2\pi u x \, du, \quad (4.05)$$

$$F(u) = \int_{-\infty}^{+\infty} f(x) \cos 2\pi u x \, dx. \quad (4.06)$$

В рассматриваемом нами примере $f(x)$ является четной функцией, и говорят, что $F(u)$ — косинусное фурье-преобразование $f(x)$. Уравнение (4.06) соответствует уравнению (4.02), и дискретное суммирование в уравнении (4.01) заменено интегрированием в уравнении (4.05). Уравнение (4.03) принимает вид

$$F(u) = ha \left(\frac{\sin \pi u a}{\pi u a} \right) \quad (4.07)$$

в согласии с выводом векторной диаграммы дифракционной картины для одиночной щели [ср. уравнение (2.05), где величина R_1 означает полную амплитуду интенсивности на щели, а здесь соответствует ha].

Если $f(x)$ нечетная функция, мы можем точно так же, как в разд. 3.5.2 рассмотреть общий случай в виде интегралов по синусу и косинусу. Однако экспоненциальное представление является в общем более удобным и обладает тем преимуществом, что может быть применено как для комплексных, так и для действительных функций. Используя уравнения (3.10) и (3.13), мы можем непосредственно написать

$$f(x) = \int_{-\infty}^{+\infty} F(u) \exp(2\pi i u x) \, du, \quad (4.08)$$

$$F(u) = \int_{-\infty}^{+\infty} f(x) \exp(-2\pi i u x) \, dx. \quad (4.09)$$

(Несущественно, какое из этих выражений содержит отрицательный знак.)

В конце данного раздела мы рассмотрим приложение такой экспоненциальной трактовки к простому случаю прямоугольной функции и увидим, что это приводит к результату, аналогичному полученному нами выше. Однако сначала сделаем некоторые общие замечания.

Произведение x и u в этих уравнениях и в уравнениях (4.05) и (4.06) является безразмерным: x и u — сопряженные параметры. И в связи с явной «симметрией» в пределах каждой из этих пар уравнений функции $f(x)$ и $F(u)$ описываются, как преобразование Фурье (или «интеграл») одна от другой. Они образуют так называемую пару преобразования Фурье. (Условно принято разделять «преобразование» и «обратное преобразование». Это различие связано с отрицательным знаком в одной из экспонент, но в данном случае для нас оно не имеет физического смысла.)

Чтобы понять, как используются уравнения (4.08) и (4.09) в случае комплексных функций, допустим, что мы имеем комплексную функцию $f(x)$, и хотим рассчитать $F(u)$. Функцию $f(x)$ можно выразить как сумму действительной и мнимой частей следующим образом:

$$f(x) = v(x) + i w(x). \quad (4.10)$$

С использованием формулы Эйлера $e^{i\theta} = \cos \theta + i \sin \theta$ уравнение (4.09) принимает вид

$$F(u) = \int_{-\infty}^{+\infty} v(x) \cos(2\pi u x) \, dx + \int_{-\infty}^{+\infty} w(x) \sin(2\pi u x) \, dx + \\ + i \left[\int_{-\infty}^{+\infty} w(x) \cos(2\pi u x) \, dx - \int_{-\infty}^{+\infty} v(x) \sin(2\pi u x) \, dx \right]. \quad (4.11)$$

Каждое подынтегральное выражение действительно, и мы видим, что $F(u)$ может быть рассчитано обычным способом путем комбинации первых двух членов с двумя вторыми на комплексной плоскости (диаграмма Арганда).

Прежде чем продолжить наше изложение, покажем, как экспоненциальная форма приводится к простым формам, если функция $f(x)$ действительная, и, в частности, к уравнению (4.05), если она также четная.

Для этого положим

$$\left. \begin{aligned} F(u) &= A(u) + iB(u), \\ \text{где } A(u) &= F(u) \cos \alpha(u), \\ B(u) &= F(u) \sin \alpha(u) \\ \text{и } \alpha(u) &= \arctg \frac{B(u)}{A(u)} \end{aligned} \right\} \quad (4.12)$$

Тогда мы можем представить уравнение (4.08) в виде

$$\begin{aligned}
 f(x) &= \int_{-\infty}^{+\infty} (A(u) + iB(u)) [\cos(2\pi ux) + i \sin(2\pi ux)] du = \\
 &= \int_{-\infty}^{+\infty} [(A(u) \cos(2\pi ux) - B(u) \sin(2\pi ux))] du + \\
 &+ \int_{-\infty}^{+\infty} [(B(u) \cos(2\pi ux) + A(u) \sin(2\pi ux))] du.
 \end{aligned} \quad (4.13)$$

Если $f(x)$ — действительная функция, то в данном случае мнимый член должен быть равен нулю. Для этого требуется, чтобы выполнялись условия

$$B(-u) = -B(+u)$$

$$A(-u) = A(+u),$$

которые означают, что

$$F(u) = A(u) + iB(u), \quad u > 0$$

$$\text{и } F(u) = A(u) - iB(u), \quad u < 0,$$

$$\left. \begin{aligned} \text{т. е. } F(u) &= F^*(-u) \\ \text{и, следовательно, } |F(u)| &= |F(-u)|. \end{aligned} \right\} \quad (4.14)$$

Таким образом, для действительной функции $f(x)$ мы имеем только первый член в уравнении (4.13), т. е.

$$f(x) = \int_{-\infty}^{+\infty} A(u) \cos(2\pi ux) du - \int_{-\infty}^{+\infty} B(u) \sin(2\pi ux) du, \quad (4.15)$$

или, с помощью уравнения (4.12), можем прийти к следующему выражению

$$\left. \begin{aligned} f(x) &= \int_{-\infty}^{+\infty} F(u) \cos[2\pi ux + \alpha(u)] du \\ &= 2 \int_0^{+\infty} F(u) \cos[2\pi ux + \alpha(u)] du \end{aligned} \right\} \quad (4.16)$$

где на $F(u)$ наложены условия (4.14). Исходное выражение (4.16) сводится к уравнению (4.05), если $f(x)$ четная функция. Следует отметить, большое сходство между вышеприведенными выражениями и соотношениями, в которых используется представление рядов Фурье в экспоненциальном виде (разд. 3.5.2).

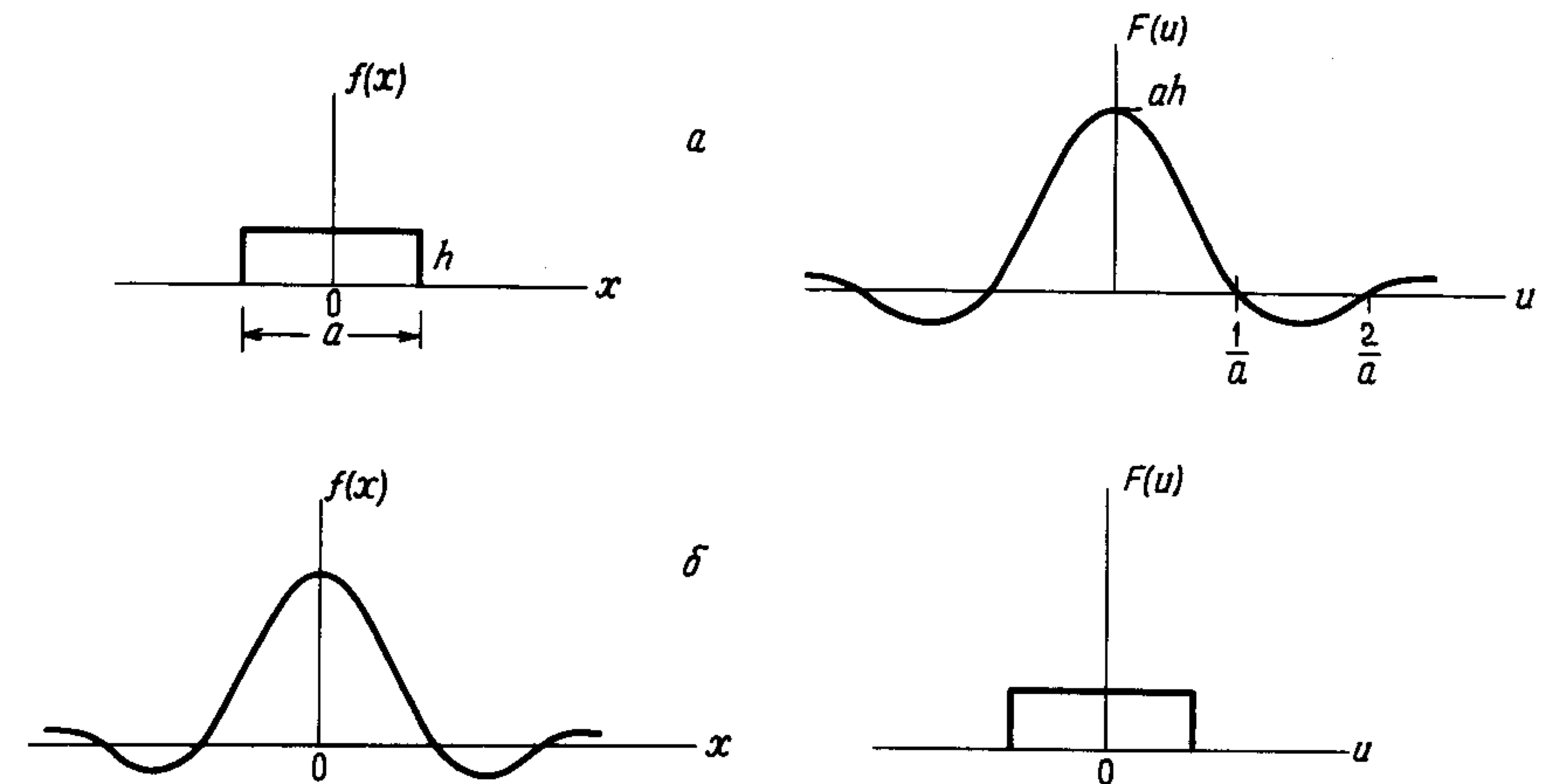


Рис. 4.2. Пара Фурье-преобразований $f(x)$ и $F(u)$.

Как и ряды Фурье, преобразования Фурье в равной мере применимы для двух и трехмерных случаев. Они также используются для сопряженной пары время — частота (разд. 4.6).

Возвратимся к случаю одиночной прямоугольной функции (рис. 4.2, а). Уравнение (4.09) позволяет достаточно просто получить ожидаемую sinc-функцию. Запишем

$$F(u) = \int_{-a/2}^{+a/2} h \exp(-2\pi i u x) dx \quad (4.17)$$

$$= ha \left(\frac{\sin \pi u a}{\pi u a} \right) \quad (4.18)$$

в соответствии с предыдущими выкладками.

Отметим, что уравнение (4.17) само по себе можно непосредственно связать с амплитудой света, дифрагировавшего в направлении u от однородно освещенной щели шириной a . Это подтверждается рис. 4.3, где видно, что вклад от элемента dx волнового фронта на расстоянии x от начала координат (0) имеет амплитуду $h dx$ и фазу $(2\pi/\lambda) x \sin \theta = 2\pi u x$, так что для всей щели полная амплитуда равна

$$\int_{-a/2}^{+a/2} h \exp(2\pi i u x) dx.$$

Дифракционную картину Фраунгофера, создаваемую одиночной апертурой, в общем случае можно считать преобразованием Фурье от ее апертурной функции. Отметим, что симметрия в соотношении между $f(x)$ и $F(u)$ такова, что если $f(x)$ была бы апертурной функцией типа sinc-функции, то ее дифракционная картина $F(u)$ должна быть прямоугольной функцией. Эта взаимная перестановка функций показана на

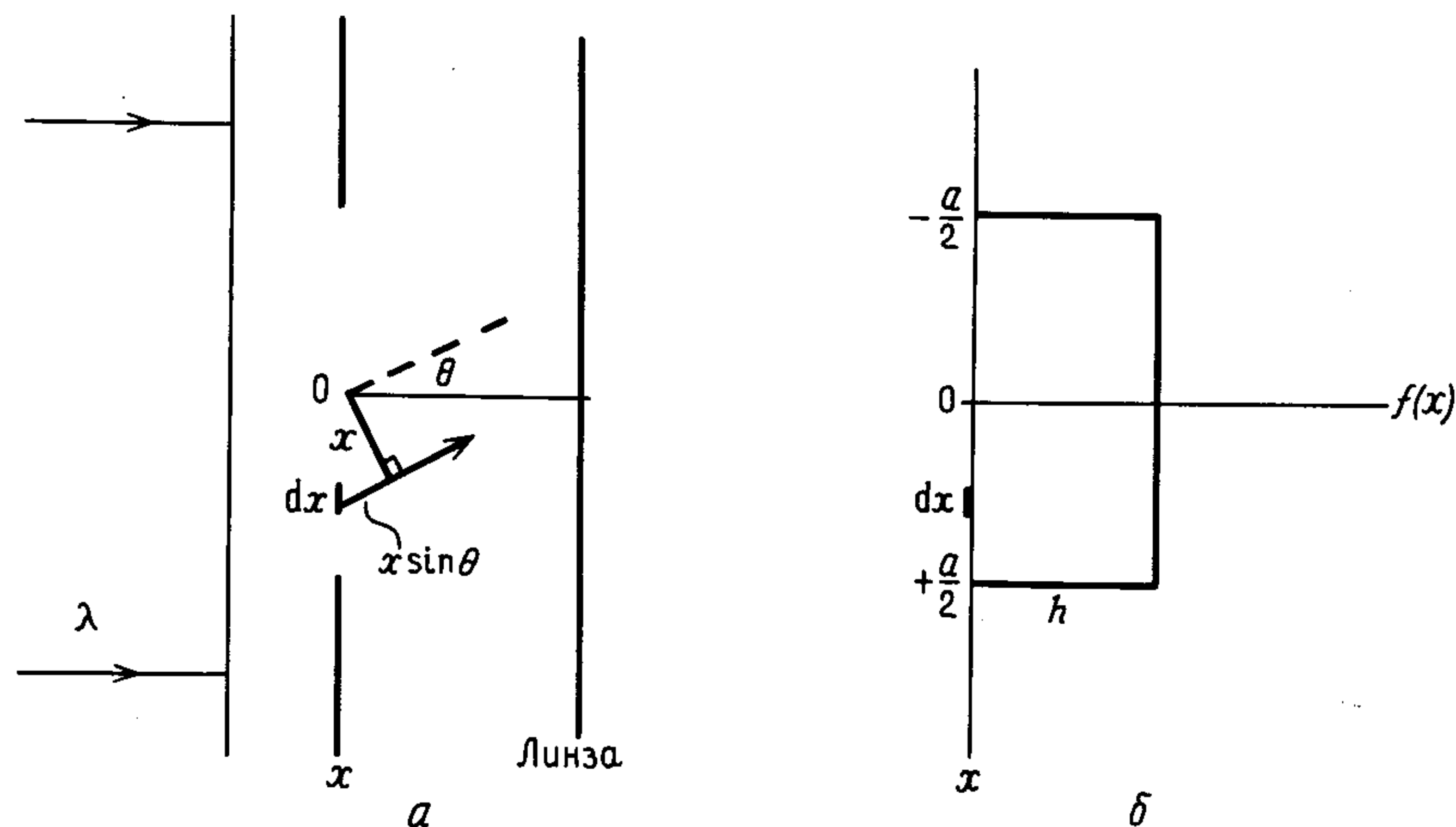


Рис. 4.3. Дифракция Фраунгофера на одиночной щели.

рис. 4.2, б.

Полученные ранее выводы (разд. 2.4 и 2.5), касающиеся соотношения дифракционных картин, создаваемых решеткой и единичной щелью, могут быть теперь расширены для описания картины решетки как выборки преобразования Фурье от апертурной функции единичной щели.

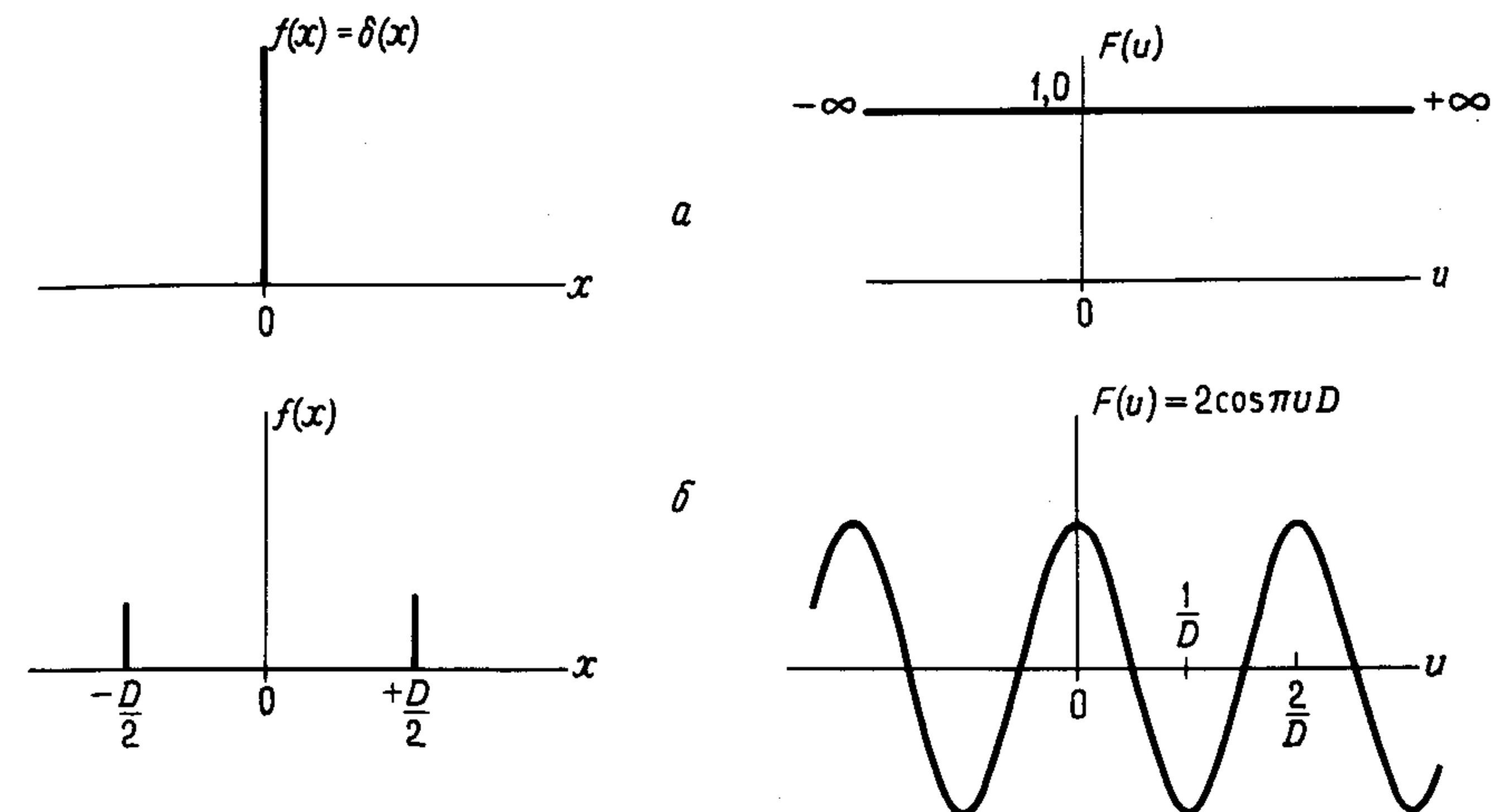
Как преобразование Фурье от единичной щели, так и ряды Фурье для решетки пространственно определены через u в одном случае непрерывно, а в другом дискретно. Следовательно, оба представления могут быть описаны как существующие в *пространстве Фурье* или *частотном пространстве*, как показано в разд. 3.4.1 в связи с дифракционной решеткой. Это очень полезное обобщение интерпретации дифракции, и оно является верным для любой апертурной функции.

Отождествление картины дифракции Фраунгофера для апертурной функции с преобразованием Фурье от этой апертурной функции приводит к трактовке линзы как устройства для выполнения преобразования Фурье. Мы настолько привыкли к линзам, что легко забываем о том, какие замечательные качества свойственны этому в сущности простому куску стекла.

4.3. ДИФРАКЦИОННАЯ КАРТИНА РЕШЕТКИ КАК ПРОИЗВЕДЕНИЕ ПРЕОБРАЗОВАНИЙ

Ниже приводится выражение для дифракционной картины, создаваемой решеткой из N щелей:

$$R_N = R_1(0) \left(\frac{\sin \pi u a}{\pi u a} \right) \left(\frac{\sin N \pi u D}{\sin \pi u D} \right). \quad \{2.15\} \quad (4.19)$$

Рис. 4.4. δ -функции и их фурье-преобразования. а — δ -функция в начале координат; б — δ -функции при $x = \pm D/2$.

Как показано в предыдущем разделе, член «одиночной щели» ($\sin \pi u a / \pi u a$) представляет собой преобразование Фурье от апертурной функции щели.

Обращая теперь внимание на «член решетки» в указанном выражении, мы найдем, что он также является преобразованием Фурье, но на этот раз от «решетчатой структуры», на которой распределены отдельные щели. Полученный результат, состоящий в том, что дифракционная картина (уравнение 4.19) есть произведение двух преобразований, применим и для других форм апертурной функции, а также для одно-, двух- и трехмерных решеток.

Для установления природы преобразования Фурье для члена решетки регулярная структура, образующая решетку, представляется последовательностью «маркеров» с идентичными апертурами — в данном случае щелями. Для такого маркера мы используем так называемую δ -функцию, математическое представление (как функция она не имеет реального математического смысла), определяемое, как предельная форма прямоугольной функции (рис. 4.2, а), у которой площадь (выбирается обычно равной единице) сохраняется постоянной, ширина стремится к нулю, в то время как высота уходит в бесконечность. Таким образом, δ -функция равна нулю всюду, за исключением одной характерной точки, где она бесконечна. В некоторых случаях она описывается как (единичная) *импульсная функция*. (Ни одна из известных обычных функций не ведет себя подобным образом, и потому ее относят к обобщенным функциям; обычно она характеризуется своими интегральными свойствами.)

Рассмотрим, во-первых, единичную δ -функцию, расположенную в на-

чале координат на оси x (рис. 4.4, а). Она обозначена как функция $f(x) = \delta(x)$, и уравнения (4.17) и (4.18) показывают, что ее преобразование Фурье равно единице (если $ha = 1$) для всех u . Такого результата можно ожидать в случае дифракционной картины от «идеальной узкой» щели (щели бесконечно малой ширины), поскольку здесь не будет разности пути, которая вызывает интерференцию света, дифрагировавшего на щели конечной ширины. Практически такая щель должна пропускать бесконечно малое количество света, и именно в этом смысле она упоминается выше как маркер.

Для построения последовательности δ -функций, представляющих решетку, необходимо определить δ -функции при других значениях x . Если δ -функция смещена от начала координат в положение x_0 , то ее можно записать как

$$f(x) = \delta(x - x_0)$$

и ее преобразование Фурье равно

$$F(u) = \int_{-\infty}^{+\infty} \delta(x - x_0) \exp(2\pi i u x) dx.$$

Положим $x - x_0 = x'$, тогда $dx = dx'$, и мы имеем

$$\begin{aligned} F(u) &= \int_{-\infty}^{+\infty} \delta(x') \exp[2\pi i u (x' + x_0)] dx' = \\ &= \exp(2\pi i u x_0) \int_{-\infty}^{+\infty} \delta(x') \exp(2\pi i u x') dx' = \exp(2\pi i u x_0), \end{aligned} \quad (4.20)$$

т.е. комплексные числа с единичными модулями и фазой, зависящей от x_0

Равномерная одномерная решетка записывается подобным же образом в виде

$$f(x) = \sum_N \delta(x - x_N), \quad (4.21)$$

а ее фурье-преобразование –

$$F(u) = \sum_N \exp(2\pi i u x_N). \quad (4.22)$$

В качестве примера рассмотрим две δ -функции при $x = \pm D/2$. Их фурье-преобразование выглядит следующим образом:

$$F(u) = \exp(\pi i u D) + \exp(-\pi i u D) = 2 \cos \pi u D.$$

Эта конкретная пара Фурье $f(x)$ и $F(u)$ показана на рис. 4.4, б. Здесь $F(u)$ легко интерпретируется как дифракционная картина для двух идеальных узких щелей, расположение которых обусловлено маркерами

типа δ -функций, определяемыми $f(x)$. Это член решетки в уравнении (4.19) для $N = 2$.

Таким же образом, преобразование Фурье от линейной последовательности N одинаково расположенных δ -функций является членом решетки для N одинаково расположенных щелей.

Итак, в итоге мы получаем, что «член одиночной щели» и «член решетки» являются соответственно преобразованием Фурье одноапертурной функции и последовательности δ -функций, определяющих структуру решетки. Дифракционная картина является произведением указанных двух преобразований, и этот результат не ограничивается приведенным здесь частным примером.

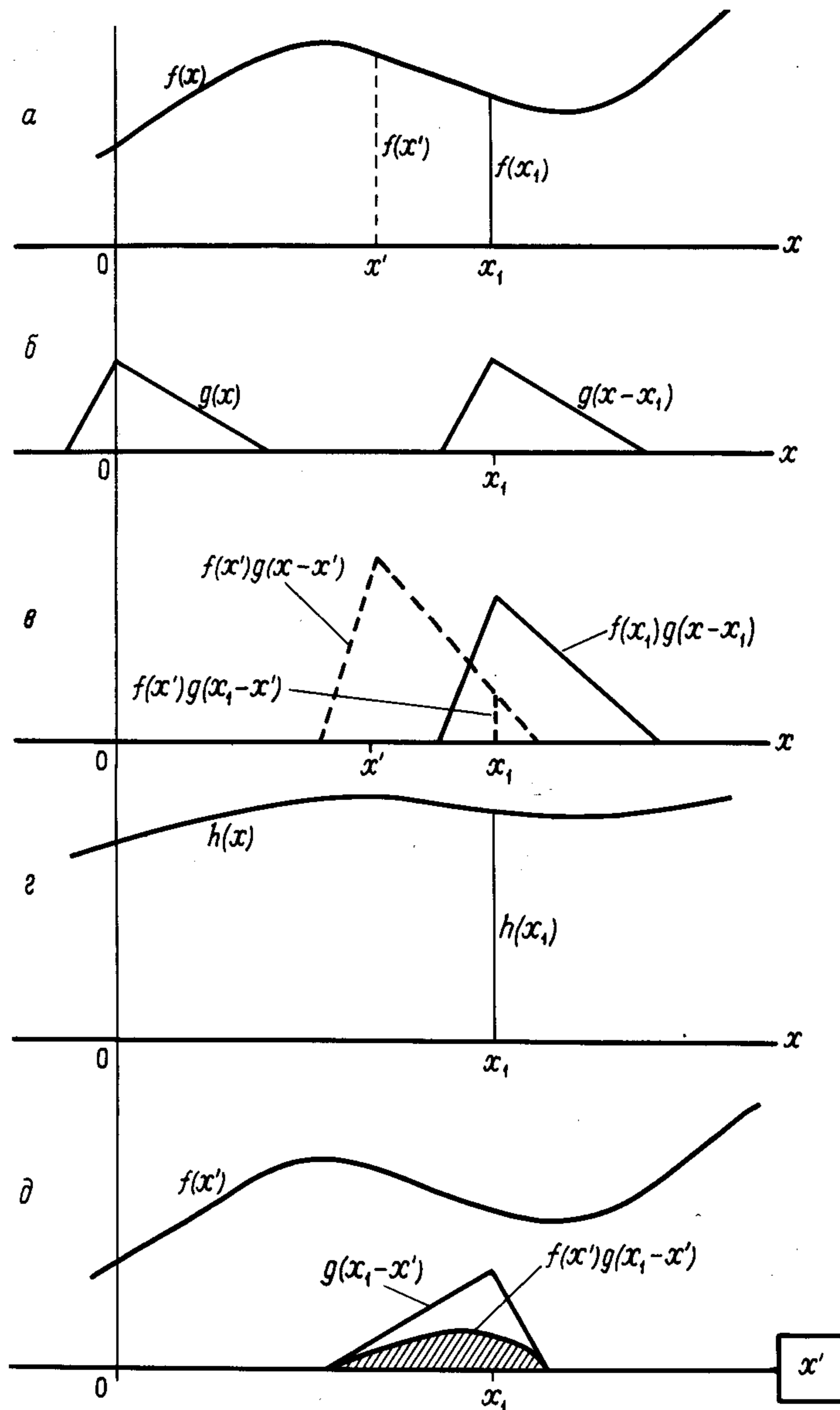
4.4. СВЕРТКА

4.4.1. Введение

В предыдущем разделе апертурная функция решетки в целом была описана как распределение составляющей ее функции одиночной апертуры и соответствующей последовательности δ -функций, определяющей структуру решетки. Такое распределение одного явления, определяемого в некотором смысле другим, представляет собой пример «свертки». Это процесс, который проявляется в разных формах и в многочисленных контекстах. Представление свертки, содержащееся в теореме свертки (разд. 4.5), чрезвычайно важно и полезно для решения целого ряда проблем. Однако наибольшее значение свертка имеет в областях формирования и обработки изображения, которые мы здесь и рассматриваем.

Имеются особые свойства свертки, которые не очень четко иллюстрируются на примере из-за разрывной природы δ -функций. В качестве введения мы рассмотрим более общий случай, когда непрерывный «входной» сигнал обрабатывается системой, создавая «выход».

Представим себе, что функция $f(x)$ на рис. 4.5, а является одномерным входным сигналом для какого-нибудь устройства. Это может быть, например, подаваемое на усилитель напряжение, изменяющееся во времени; тогда вместо $f(x)$ мы введем обозначение $f(t)$. Или же это может быть пространственное распределение интенсивности света в определенном направлении x по экрану. В первом из этих примеров, очевидно, потребуется усилитель, для увеличения (усиления) каждой ординаты (см. рис. 4.5, а) в несколько раз, чтобы сформировать выходной сигнал. Во втором примере можно представить, что фотометр обеспечивает точное сканирование и измерение распределения интенсивности $f(x)$ и определенным образом воспроизводит его. В обоих случаях, однако, обработка «входа» практически оказывается весьма несовершенной. Каждая ордината входного сигнала не воспроизводится чисто, а размывается. Таким образом, входная δ -функция при $x = 0$ создает на выходе некоторую функцию $g(x)$, а не δ -функцию.

Рис. 4.5. Свертка $f(x)$ и $g(x)$.

Мы принимаем, что «форма» отклика системы, включая нежелательное размытие, является инвариантной, т.е. она одна и та же для каждого значения x . Отклик представлен треугольником на рис. 4.5,б при двух значениях x , где он соответствует $g(x)$ и $g(x-x_1)$ для единичного сигнала на входе. Рассмотрим, что происходит на входе в точке x_1 (рис. 4.5,а). В этом случае ордината на входе равна $f(x_1)$ и она размывается откликом системы, становясь кривой $f(x_1)g(x-x_1)$, показанной сплошной линией на рис. 4.5,в. Таким образом, $f(x_1)$ является «весом» единичного отклика $g(x-x_1)$. В этом размытии только ордината при x_1 создает реальный вклад в выходной сигнал при x_1 . Эта выходная ордината не показана на рисунке, поскольку следует иметь в виду, что здесь имеются другие вклады, которые порождаются размытием других ординат в $f(x)$. Один из таких вкладов, возникающий из-за размытия ординаты при x' на рис. 4.5,а, показан пунктирной кривой на рис. 4.5,в. Этот вклад, создаваемый на выходе при x_1 , является ординатой $f(x')g(x_1-x')$.

В этом случае общий выходной сигнал при x_1 , показанный как h на рис. 4.5,г, является суммой воздействия размытия при x_1 на все ординаты $f(x)$. В приведенном примере только относительно близкие ординаты будут оказывать влияние, потому что $g(x)$ довольно узкая. Общий выходной сигнал при x_1 определяется следующим интегралом:

$$h(x_1) = \int_{x'=-\infty}^{x'=\infty} f(x') g(x_1 - x') dx'. \quad (4.23)$$

То же самое применимо к любому значению x_1 , поэтому выходной сигнал $h(x)$ можно выразить в виде

$$h(x) = \int_{x'=-\infty}^{x'=+\infty} f(x') g(x - x') dx'. \quad (4.24)$$

Этот интеграл известен как *свертка* $f(x)$ и $g(x)$, где обе функции могут быть комплексными. Кратко эта запись представляется в виде

$$h(x) = f(x) * g(x),$$

или просто $h(x) = f * g$.

Хотя название *свертка* (convolution) для этого интеграла является общепринятым, имеются и другие, например *свернутое произведение* (folding product, ср. немецкое Faltung—складка), *составное произведение* (composition product), *интеграл суперпозиции* (superposition integral). В соответствующем контексте $g(x)$ будет называться *сглаживающей функцией* (или *размытием*), *функцией рассеяния линии* (ФРЛ) или в двумерном случае *функцией рассеяния точки* (ФРТ).

Рассмотренный нами пример иллюстрирует, каким образом свертка

обуславливает распределение одной функции в соответствии с законом, определяемым другой функцией. Эта операция предполагает умножение каждой ординаты функции целиком на другую функцию и суммирование результатов. Следует отметить, что, как нередко бывает на практике, мы имеем дело с *линейной системой*, т.е. подразумевается, что ее выход является линейной суперпозицией выходов, получаемых от каждого отдельного компонента на входе. Отметим также, что мы предполагаем «инвариантность» функции $g(x)$. Мы вернемся к этим проблемам в разд. 5.1.

До сих пор нами рассматривалась преимущественно физическая сторона картины. Для расчета выхода при любом частном значении x , таком, как x_1 , необходимо выполнить интегрирование согласно уравнению (4.24). Эта процедура представлена графически на рис. 4.5, δ , где x' является вспомогательной переменной интегрирования. Функция $f(x')$ умножается на $g(x_1 - x')$, и область (заштрихованная) под кривой произведения представляет собой величину выхода $h(x_1)$ при x_1 . Отметим, что перед умножением и интегрированием $g(x)$ как смещается, так и переворачивается (ср. термин «свернутое произведение» как альтернативу «свертке»). Вследствие этого «перевертывания» свертка является коммутативной, т.е.

$$f * g = g * f.$$

Она также ассоциативна:

$$f * (g * h) = (f * g) * h$$

и дистрибутивна по сложению:

$$f * (g + h) = (f * g) + (f * h).$$

4.4.2. Представление решетки сверткой

Совокупность δ -функций, определяющих решетчатую структуру, может быть представлена функцией $f(x)$ на рис. 4.6 и функцией одиночной апертуры $g(x)$, хотя они могли бы быть выбраны обратным образом. Тогда в результате оказывается, что апертурная функция всей решетки является сверткой $h(x)$ одной функции с другой.

Этот пример, подобный приведенному в предыдущем разделе, показывает, каким образом свертка может быть представлена как суммарное влияние распределения одной функции в соответствии с законом, определяемым другой функцией. В показанном ранее примере (рис. 4.5) это распределение приводит до некоторой степени к перекрытию ординат, в то время как в примере с решеткой этого не происходит. Совершенно другой пример свертки приводится в гл. 5 в связи с формированием изображения.

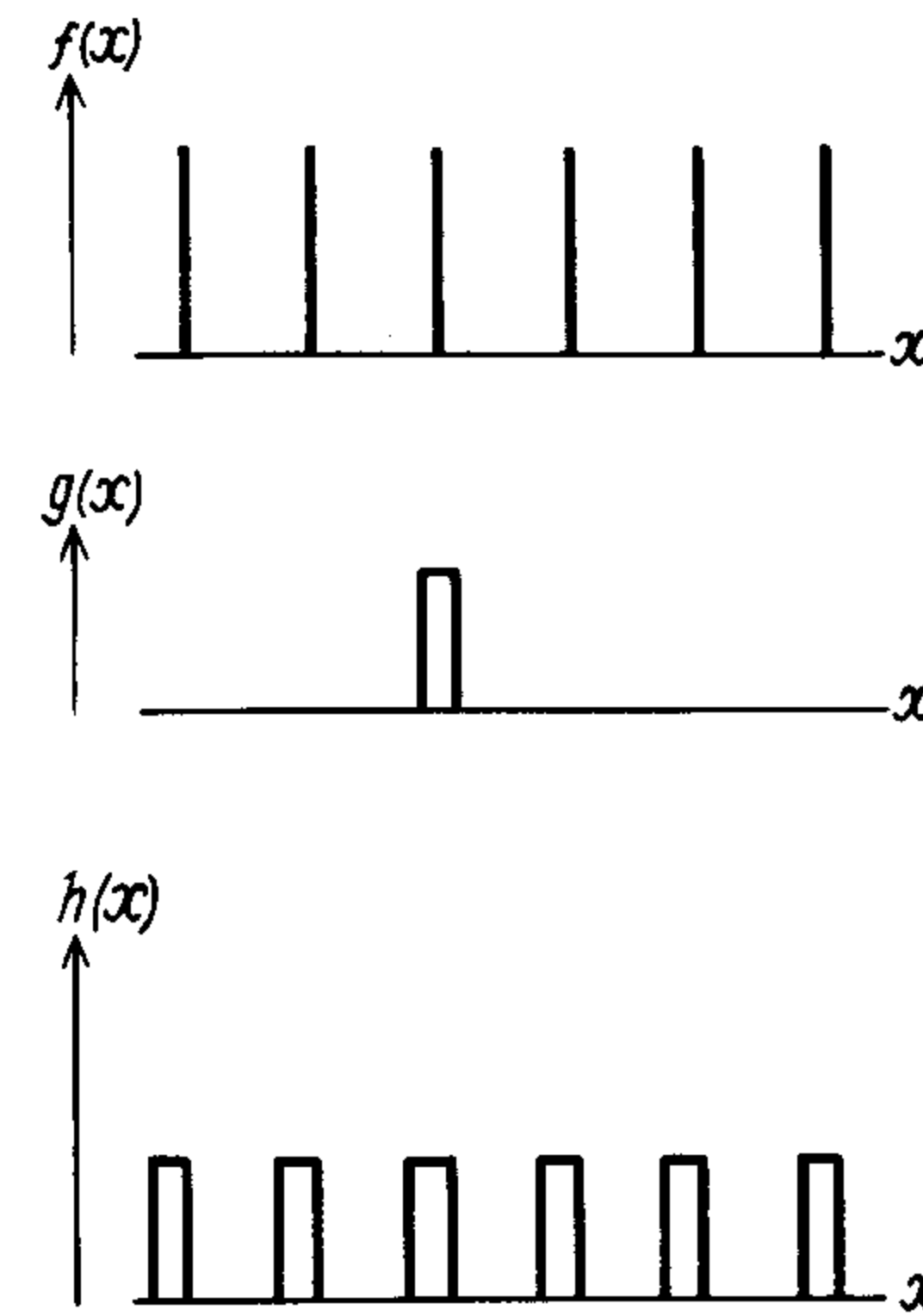


Рис. 4.6. Апертурная функция решетки $h(x)$ как свертка последовательности δ -функций $f(x)$ с функцией одиночной апертуры $g(x)$.

4.5. ТЕОРЕМА О СВЕРТКЕ И ДИФРАКЦИЯ

Теперь мы можем сопоставить вместе некоторые данные двух предыдущих разделов и получить результат, который имеет исключительную важность.

В разд. 4.4.2 было показано, что апертурная функция решетки в целом может быть описана как свертка функции одиночной апертуры с последовательностью δ -функций, которая определяет распределение этой апертуры в решетке.

Как мы видели в разд. 4.4.3, дифракционная картина решетки равна произведению фурье-преобразования функции одиночной апертуры и фурье-преобразования последовательности δ -функций, определяющих решетку.

Поскольку дифракционная картина решетки является преобразованием Фурье от ее полной апертурной функции, мы можем, следовательно, сказать, что фурье-преобразование свертки функции одиночной апертуры с последовательностью δ -функций равно произведению отдельных преобразований. Это пример *теоремы о свертке*, которая утверждает, что фурье-преобразование свертки двух функций равно произведению их собственных преобразований.

В данном контексте эта теорема приводит к очень важному результату, состоящему в том, что свертка в пространстве объекта (физическом пространстве) соответствует умножению в дифракционном пространстве (т.е. пространстве Фурье или взаимном пространстве). Это следствие не только позволяет наглядно объяснить процесс формирования изображения, но и служит мощным инструментом с точки зрения его численной обработки (разд. 5.5).

Поскольку теорема свертки представлена здесь весьма специфиче-

ским способом, у читателя может возникнуть желание ознакомиться с прямым ее выводом.

Мы использовали $h(x)$ для обозначения свертки двух функций $f(x)$ и $g(x)$. Их собственные преобразования Фурье записываются соответственно $H(u)$, $F(u)$ и $G(u)$, где x и u – обычные сопряженные переменные. Таким образом, теорема свертки может быть выражена в форме утверждения, что если

$$\left. \begin{aligned} h(x) &= f(x) * g(x), \\ \text{то тогда } H(u) &= F(u) G(u). \end{aligned} \right\} \quad (4.25)$$

Для доказательства запишем, используя определение преобразования Фурье:

$$H(u) = \int_{-\infty}^{+\infty} h(x) \exp(2\pi i u x) dx. \quad (4.26)$$

Заменяя $h(x)$ сверткой $f(x)$ и $g(x)$ и используя уравнение (4.34), получаем

$$H(u) = \int \exp(2\pi i u x) \left[\int f(x') g(x - x') dx' \right] dx. \quad (4.27)$$

Изменяя порядок интегрирования, запишем

$$H(u) = \int f(x') \left[\int \exp(2\pi i u x) g(x - x') dx \right] dx'. \quad (4.28)$$

Во внутреннем интеграле, в котором x' является постоянной величиной, проводим замену переменной $x - x' = X$. Тогда $dx = dX$ и

$$H(u) = \int f(x') \left[\int \exp(2\pi i u (X + x')) g(X) dX \right] dx'. \quad (4.29)$$

Множитель $\exp(2\pi i u x')$ является постоянным в пределах внутреннего интеграла, и, поскольку остальная часть внутреннего интеграла не содержит x' , она является постоянной для внешнего интегрирования по отношению к x' . Таким образом, полное выражение разделяется на произведение двух отдельных интегралов:

$$H(u) = \int_{-\infty}^{+\infty} f(x') \exp(2\pi i u x') dx' \int_{-\infty}^{+\infty} g(X) \exp(2\pi i u X) dX, \quad (4.30)$$

т.е.

$$H(u) = T[f(x)] \times T[g(x)], \quad (4.31)$$

или

$$H(u) = F(u) G(u), \quad (4.32)$$

где T обозначает преобразование Фурье.

4.6. ПРЕОБРАЗОВАНИЕ ФУРЬЕ И СВЕТОВЫЕ ВОЛНЫ

Замечания, сделанные в разд. 1.2 относительно природы световых волн и дискретной природы их формирования, могут быть теперь подтверждены. Преобразование Фурье используется для этого таким же образом, как при исследовании соотношения между апертурной функцией и распределением составляющих ее пространственных частот. Вместо того, чтобы рассматривать «пространственный импульс», мы обратимся теперь к импульсу во времени, связанному с распределением (т.е. спектром) временных частот. Сопряженные переменные x и u заменяются временем t и частотой ν . Уравнения для преобразования Фурье могут быть записаны путем простой замены символов в уравнениях (4.08) и (4.09), как следует ниже:

$$f(t) = \int_{-\infty}^{+\infty} F(\nu) \exp(2\pi i \nu t) d\nu, \quad (4.33)$$

$$F(\nu) = \int_{-\infty}^{+\infty} f(t) \exp(-2\pi i \nu t) dt. \quad (4.34)$$

Чтобы обеспечить аналогию между этим новым «сценарием» и дифракцией, на рис. 4.7,а представлены прямоугольная функция и преобразование от нее, обозначенные теперь в соответствии с новой переменной. Однако, как мы уже знаем, основная компонента прямоугольной функции не периодическая (т.е. нулевой частоты) с постоянной амплитудой, вследствие чего функция полностью положительна. Более подходящим примером для рассмотрения световых волн является пара преобразований на рис. 4.7,б. Здесь показана чистая синусоидальная волна с частотой ν_1 , представленная в виде цуга конечной продолжительности и длины. Она имеет амплитудно-частотное распределение, размытое около ν_1 так, что суммирование дает *группу волн* (или *волновой пакет*), которая представляет собой профиль в пределах цуга, но суммарная амплитуда равна нулю с любой стороны от него. Если цуг длинный, то частотное размытие невелико и наоборот, т.е. взаимосвязь здесь такая же, как в случае с парой пространственного преобразования Фурье. Строго говоря, монохроматический свет предполагает наличие цугов бесконечной длины, но это условие физически не выполнимо, поскольку свет излучается атомами дискретно, в виде фотонов; в результате все спектральные линии имеют конечную ширину. Если на рис. 4.7,б «ширина» частотного распределения взята в основном в пределах $\nu_1 \pm \delta\nu$, то мы имеем

$$\delta\nu \delta t \approx 1. \quad (4.35)$$

Эта мера ширины, называемая *шириной полосы*, использована, поскольку большая часть энергии связана с центральным пиком, что можно

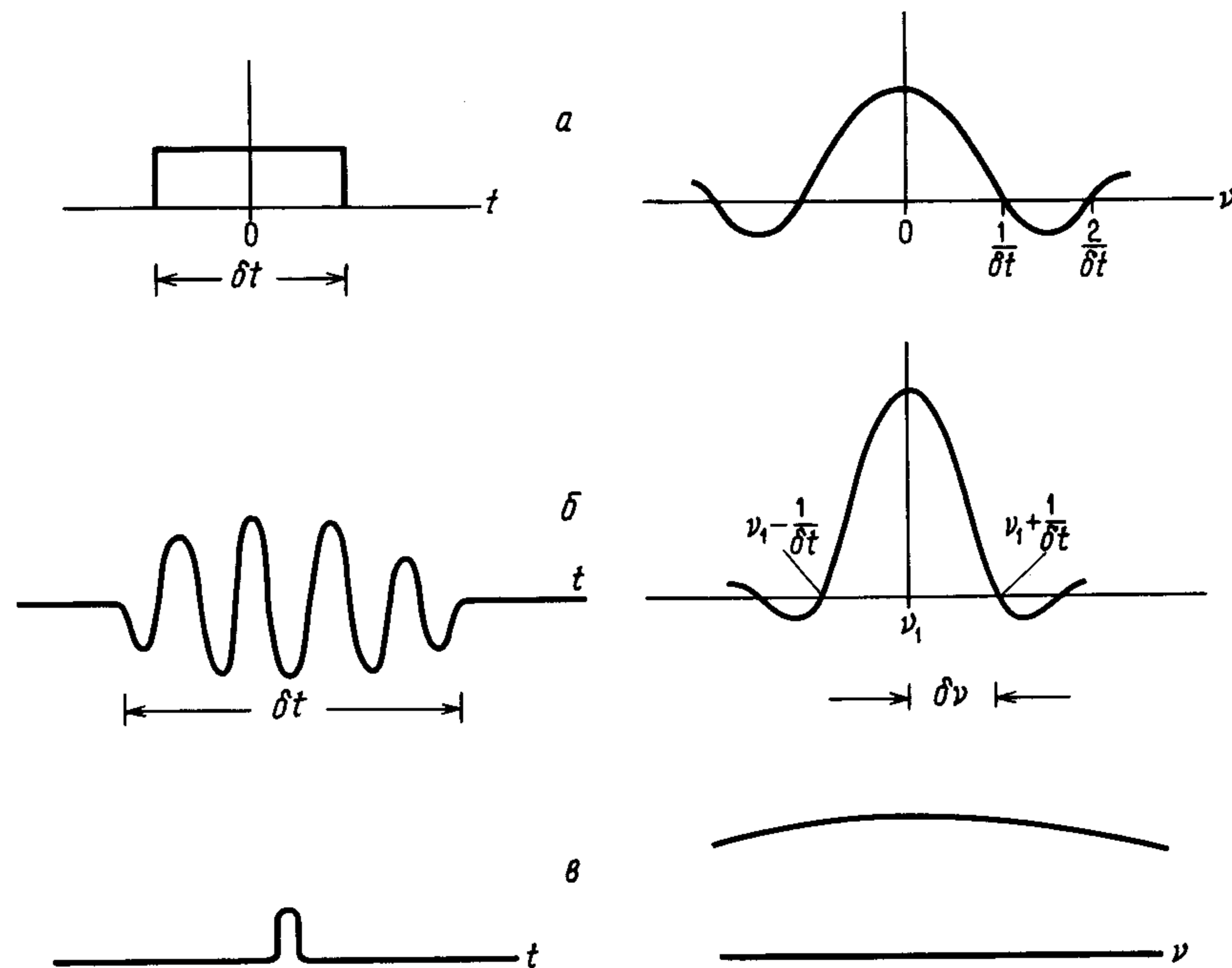


Рис. 4.7. Фурье-преобразование в пространстве время–частота.

увидеть, если возвести амплитуду в квадрат и получить распределение по энергии (*спектр мощности*).

Поскольку волновые цуги, излучаемые тепловыми (не лазерными) источниками света, не одинаковы по продолжительности и частоты изменяются вследствие теплового движения, влияния полей и т.д., целесообразно принять среднюю продолжительность времени Δt и спектральное распределение, отличающееся до некоторой степени от рассмотренного выше. Однако взаимосвязь между продолжительностью импульса и частотным размытием имеет фундаментальный характер, и если полоса частот равна $\Delta \nu$, то мы имеем так называемую *теорему о ширине полосы*

$$\Delta \nu \Delta t \approx 1. \quad (4.36)$$

В этом уравнении Δt является временем когерентности света. Соответствующая *длина когерентности* l_c равна

$$l_c = c \Delta t \approx c / \Delta \nu. \quad (4.37)$$

С учетом равенства $c = \lambda \nu$ это уравнение принимает вид

$$l_c \approx \lambda^2 / \Delta \lambda. \quad (4.38)$$

Источники света, такие, как газоразрядные трубки, дают спектральную ширину линии порядка 0,1 нм, так что, принимая $\lambda = 550$ нм, получаем $l_c \approx 5500$ длин волн. Хотя это количество длин волн велико, физически оно соответствует малому отрезку (около 3 мм), сравнимому с размерами элементов и смещениями, используемыми в интерферометрии.

В гл. 6 мы увидим, как методы Фурье в спектроскопии основываются на измерениях временной когерентности в интерферометрии.

В противоположность вышеприведенному примеру ширина линии лазерного излучения настолько мала, что l_c может достигать десятков километров, в то же время «белый свет» от лампы накаливания с вольфрамовой нитью имеет весьма широкий частотный диапазон (рис. 4.7, в), поскольку состоит из импульсов чрезвычайно короткой длительности и ни одна из частот не является доминирующей (условие, использованное некоторыми авторами как определение *импульса*).

4.7. КОРРЕЛЯЦИЯ

В настоящее время корреляция является наиболее широко распространенным методом обработки различных сигналов (оптических и других) и данных. Мы можем видеть примеры применения корреляции в гл. 5 и 6. При всех различных проявлениях корреляция является, по существу, методом оценки и определения взаимных связей, имеющих форму подобий или совпадений. В качестве примера можно привести корреляцию фаз Луны с приливами на Земле: если изобразить эти два периодических процесса на графике в зависимости от времени, то видно, что они взаимосвязаны.

Процесс поиска корреляции по существу сводится к сравнению двух картин, но это далеко не так просто, как может показаться вначале. Поясним возникающие при этом проблемы на примере. На рис. 4.8, а на плоскости x, y расположена некоторая фигура. Допустим, она описывается функцией $f(x, y)$, где f является некоторой заданной характеристикой фигуры, например ее цвет, композиция, отражательная способность, прозрачность и т.д.; вне фигуры f равна нулю. Эта функция неприменима для описания фигуры независимо от ее положения на плоскости x, y . Но во многих случаях необходим способ определения функции, «инвариантный по отношению к переносу». В таком способе предусматриваются координаты, которые являются внутренними для функции, но определяют ее так же, как $f(x, y)$. Рассмотрим функцию $P_{ff}(u, v)$, которая представляет собой произведение значения f при x, y на соответствующее значение при $x + u, y + v$, причем эти произведения суммируются (интегрируются) для всех x, y . Такую функцию можно выразить, как

$$P_{ff}(u, v) = \iint_{-\infty}^{+\infty} f(x, y) f(x + u, y + v) dx, dy. \quad (4.39)$$

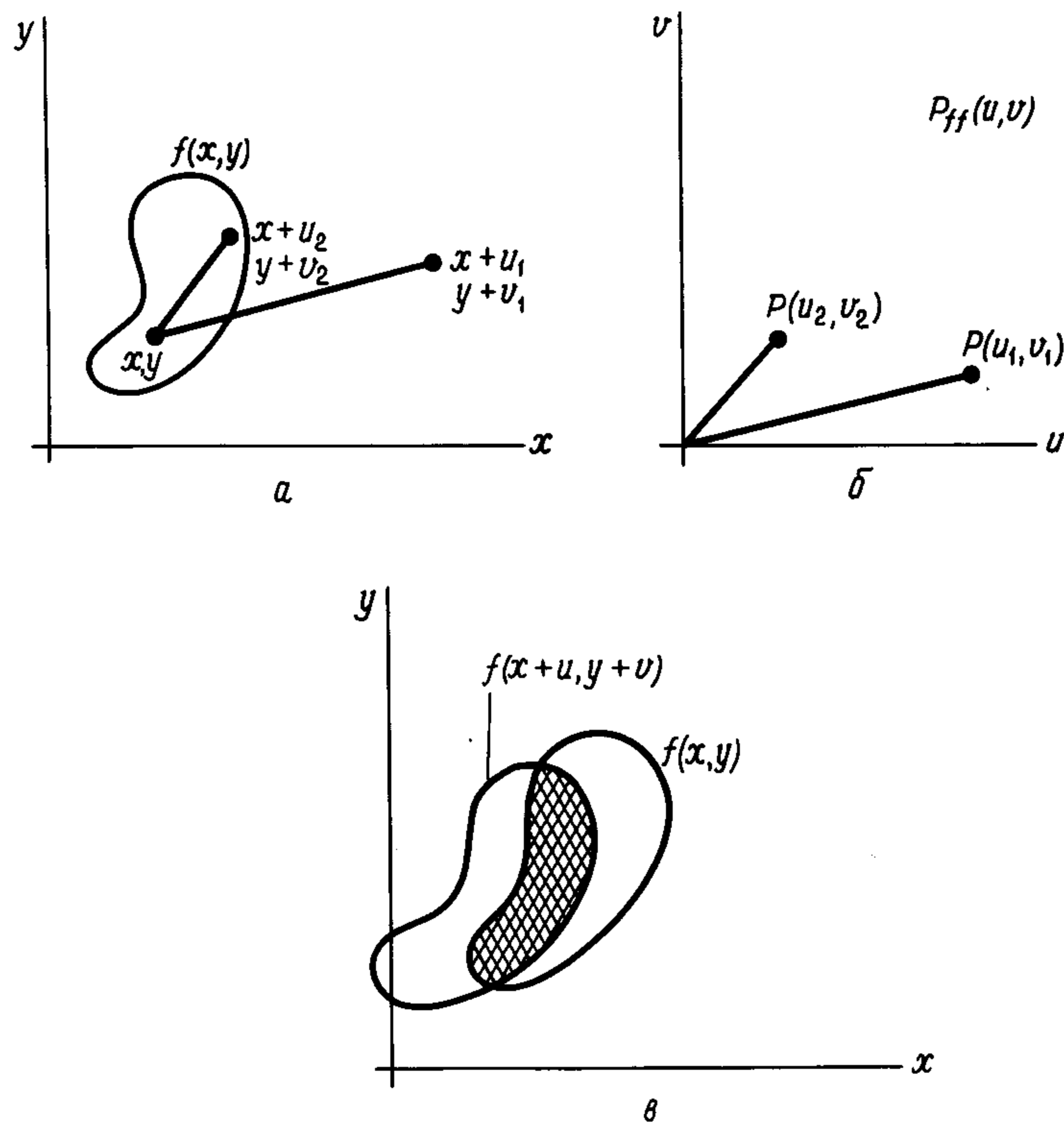


Рис. 4.8. Автокорреляция.

Как показано на рисунке, эта функция представляет векторное описание фигуры. Так как f равна нулю для всех точек вне фигуры, то на рис. 4.8, а произведение значений f для всех пар точек на плоскости x, y , разнесенных на u_1, v_1 , суммированное по всем x, y , равно нулю. По этой причине P для u_1, v_1 на рис. 4.8, б имеет нулевое значение. С другой стороны, величина $P(u_2, v_2)$ явно не равна нулю.

Таким образом функция P предопределяет картину, которая описывает f независимо от ее положения на плоскости x, y . Поскольку здесь функция f сравнивается сама с собой, выражение (4.39) называют *автокорреляционной функцией*. (Очевидно, $P(0, 0)$ соответствует сильному пику на P -карте и можно доказать, что она представляет максимальное значение.)

Важно отметить, что уравнение (4.39) также дает нам другой способ визуализации вычислений. Видно, что значение P для любых выбранных u, v получается смещением функции f относительно себя самой на $-u, -v$ и определением площади перекрытия (рис. 4.8, в) (сравните выше-

приведенные замечания относительно $P(0, 0)$, представляющей максимальное значение P).

Для сравнения двух разных фигур $f(x)$ и $g(x)$ можно записать *кросс-корреляционную функцию*

$$P_{12}(u, v) = \iint_{-\infty}^{+\infty} f(x, y) g(x+u, y+v) dx dy, \quad (4.40)$$

где P теперь относится к произведению f и g .

Обе корреляционные функции могут быть выражены для одно-, двух- и трехмерных случаев, а также для времени. С целью их детального изучения вернемся к одномерному случаю

$$P_{12}(u) = \int_{-\infty}^{+\infty} f(x) g(x+u) dx \quad (4.41)$$

и так же для P_{11} .

Чтобы вышеприведенное уравнение можно было описать сдвигом u в положительном направлении x , следует переписать его следующим образом:

$$P_{12}(u) = \int_{-\infty}^{+\infty} f(x) g(x-u) dx. \quad (4.42)$$

Простой одномерный пример автокорреляции показан на рис. 4.9, а и в. (Рис. 4.9, б и г относятся к разд. 4.7.1.) Читатель может начертить несколько графиков с целью убедиться, что, когда вторая прямоугольная функция, идентичная первой на рис. 4.9, а, сдвигается, принимая разные положения вдоль оси x , площадь под произведением для каждого сдвига соответствует результату, показанному на рис. 4.9, в.

Определение корреляции нередко выражается в несколько иной форме, а именно:

$$c(x) = f(x) \odot g(x) = \int_{-\infty}^{+\infty} f(x') g(x' - x) dx' \quad (4.43)$$

и аналогично для автокорреляции. В этом выражении x является смещением, а x' — вспомогательной переменной интегрирования.

Таким образом, корреляция подобна свертке [ср. уравнение (4.24)], но отличается тем, что функция g сдвигается, а не переворачивается (перегибается). Символ $\odot$ используется довольно широко, чтобы отличить корреляцию от свертки, хотя, к сожалению, для обозначения этих операций нет общепринятых символов.

Если функции комплексные, то принято определять *комплексную*

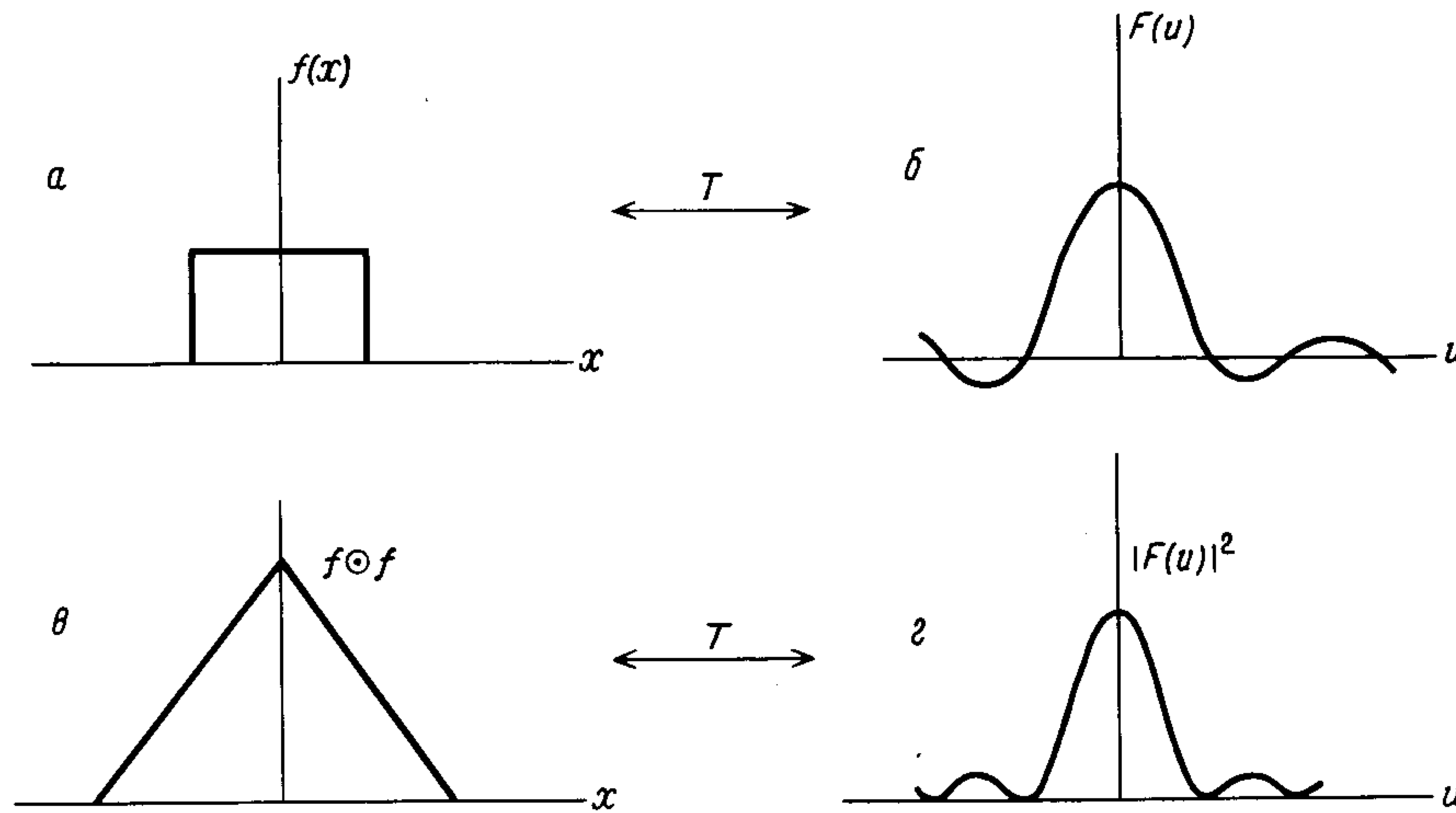


Рис. 4.9. Пример автокорреляции (а, в) и теоремы Винера–Хинчина (б, г); буква T обозначает фурье-преобразование.

кросс-корреляционную функцию

$$\gamma_{fg}(x) = f(x) \odot g^*(x) = \int_{x'=-\infty}^{x'+\infty} f(x') g^*(x' - x) dx' \quad (4.44)$$

и комплексную автокорреляционную функцию

$$\gamma_{ff}(x) = f(x) \odot f^*(x) = \int_{x'=-\infty}^{x'+\infty} f(x') f^*(x' - x) dx. \quad (4.45)$$

(Отметим, что, если звездочка связана с другой функцией, то получается сопряженный результат.) Для комплексных $f(x)$ и $g(x)$ отмеченное выше различие между сверткой и корреляцией может быть представлено в общем виде как

$$f(x) \odot g^*(x) = f(x) * g^*(-x). \quad (4.46)$$

Из этого следует, что корреляция не коммутативна, т.е.

$$f(x) \odot g^*(x) \neq g^*(x) \odot f(x). \quad (4.47)$$

Иногда бывает удобно нормировать корреляционные функции. Нормировка достигается делением выражений (4.44) и (4.45) на центральное значение корреляции, т.е. на значение для $x=0$ (нулевой сдвиг). Это дает

$$\text{нормированная } \gamma_{fg}(x) = \frac{\int f(x') g^*(x' - x) dx'}{\int f(x') g^*(x') dx'}, \quad (4.48)$$

$$\text{нормированная } \gamma_{ff}(x) = \frac{\int f(x') f^*(x' - x) dx'}{\int f(x') f^*(x') dx'} \quad (4.49)$$

$$= \frac{\int f(x') f^*(x' - x) dx'}{\int |f(x)|^2 dx'}, \quad (4.50)$$

где интегрирование выполняется в пределах от $-\infty$ до $+\infty$.

4.7.1. Теорема автокорреляции (теорема Винера–Хинчина)

Эта теорема утверждает, что преобразование Фурье автокорреляции от функции $f(x)$ равно квадрату модуля ее преобразования.

В общем случае если преобразование комплексной функции есть $F(u)$, то, согласно этой теореме преобразование, $|F(u)|^2$ является комплексной автокорреляцией от $f(x)$.

Используя T для обозначения преобразования Фурье, мы имеем

$$\begin{aligned} T[|F(u)|^2] &= \int_{-\infty}^{+\infty} |F(u)|^2 \exp(2\pi i u x) du = \\ &= \int_{-\infty}^{+\infty} (F(u) F^*(u)) \exp(2\pi i u x) du = f(x) * f^*(-x)^{1)}, \end{aligned} \quad (4.51)$$

откуда, используя уравнение (4.46), получаем

$$T[|F(u)|^2] = f(x) \odot f^*(x), \quad (4.52)$$

т.е.

$$F(u) F^*(u) \xrightarrow{T} f(x) \odot f^*(x). \quad (4.53)$$

Это соотношение иллюстрируется на рис. 4.9 для случая действительной функции $f(x)$. Заменяя переменные на время и частоту, получим

$$T[|F(v)|^2] = f(t) \odot f^*(t), \quad (4.54)$$

где $|F(v)|^2$ представляет собой *спектр мощности*, выраженный через квадрат амплитуды как функции (временной) частоты. По аналогии $|F(u)|^2$ в уравнении (4.52) называют *спектром мощности* в зависимости от *пространственной частоты*. Приложение обоих случаев будет приведено в последующих главах.

¹⁾ Отрицательный знак в $f^*(-x)$ появляется от фурье-преобразования сопряженного

$$\text{т.е. } \int_{-\infty}^{+\infty} F^*(u) \exp(2\pi i u x) du = \left[\int_{-\infty}^{+\infty} F(u) \exp(-2\pi i u x) du \right]^*,$$

$$T[F^*(u)] = f^*(-x).$$

Формирование и обработка оптического изображения

5.1. Введение

В этой главе в общих чертах показаны главные положения фурье-анализа при формировании оптического изображения и его обработке в условиях когерентного и некогерентного освещения. Они включают как одностороннее преобразование Фурье, так и преобразование в сочетании со сверткой и корреляцией. Следует, однако, сразу же привлечь внимание к тому факту, что важность этих положений не ограничивается обработкой данных, имеющих оптическое происхождение. В настоящее время можно привести большое число примеров, когда методы оптической обработки используются для данных, по своей природе не являющихся оптическими. Основная причина кроется в том, что математические операции, которые применяются для большинства оптических систем, часто используются также в системах связи. Оптический аналог весьма привлекателен, поскольку ему свойственно преимущество двумерного представления и параллельной обработки данных. Этот способ во все увеличивающейся степени внедряется в практику в связи с разработкой электронно-оптических устройств сопряжения в сочетании с ЭВМ. Когда по каким-то причинам оптические методы не употребляются, ЭВМ может применяться изолированно в целях использования тех же фундаментальных принципов для «цифрового изображения» и обработки.

Однако цель данной главы состоит в рассмотрении лишь существа преобразования Фурье, свертки и корреляции, о которых мы уже говорили в «оптическом» контексте в предыдущих главах. Эти операции образуют основной инструментарий в области формирования и обработки оптического изображения. Для более детальной информации по конкретным вопросам и многочисленным приложениям в областях связи следует обратиться к соответствующей специальной литературе.

Перед детальным рассмотрением указанных вопросов было бы полезно сделать предварительный обзор разд. 5.2, 5.3, посвященных собственно формированию оптического изображения. Оптическая обработка в противоположность построению изображения связана с вмешательством в процесс разными способами и для разных целей и является предметом разд. 5.5.

На современном этапе известны две полезные модели формирования оптического изображения, каждая из которых наиболее удачно применяется в совершенно определенных условиях. Обе они были затронуты в гл. 1 и упоминались в той или иной степени в промежуточных главах.

В первой модели делается акцент на общий характер дифракции (рассеяние) света от объекта, когда условия по крайней мере частично когерентны, и на способ сведения света для формирования изображения. Аспекты анализа Фурье, относящиеся к первой части этого вопроса, уже знакомы нам по гл. 3 и 4. В разд. 5.3 мы рассматриваем их снова на этот раз с учетом второго этапа формирования изображения. Эта модель первоначально была сформулирована (в основном качественно) в 1873 г. Э. Аббе [1], который занимался проблемами наблюдений периодических объектов под микроскопом. Как можно сказать, пользуясь современной терминологией, он выяснил, что при способах освещения, используемых обычно в оптической микроскопии, формирование изображения вовсе не является полностью некогерентным процессом, как иногда полагают; в действительности в некоторых современных системах он может быть почти когерентным.

Для когерентно освещенного объекта количественное отождествление порядков дифракции с членами рядов Фурье, описывающих структуру объекта, и процесс их восстановления в изображение с помощью линзы, подобный фурье-синтезу, довольно подробно обсуждались А.Б. Портером в 1906 г. [47]. (Тогда не предполагали, что значительный шаг в этом направлении уже был сделан лордом Рэлеем [49] в 1874 г. и особенно в его статье [51].)

Расширение интерпретации Аббе–Портера с указанием на аналогичную роль преобразований Фурье в формировании когерентного изображения непериодических объектов было осуществлено главным образом специалистами по рентгеновской кристаллографии в начале 40-х годов.

Важность подхода с использованием этой модели состоит в ее чувствительности к процессу, при котором пространственные частоты «структуры» объекта (периодической и непериодической) выражаются дифрагированными волновыми фронтами и восстанавливаются для формирования изображения. Использование когерентного освещения позволяет воздействовать на дифракционную плоскость (плоскость пространственных частот) таким образом, что формирование изображения может управляться посредством «фильтрации». Это один из аспектов оптической обработки, другие упомянуты в разд. 5.5.

Вторая модель формирования изображения, которую мы рассматриваем в разд. 5.2, применима к условиям как когерентного, так и некогерентного освещения. И здесь Рэлей внес важный вклад [51], на этот раз под влиянием более ранних работ Эри и Гельмгольца. Модель представляет изображение как комбинацию картин Эри (или более сложных картин, если присутствуют aberrации), которые оптическая система должна создавать отдельно для света из каждой точки объекта. Если освещение некогерентно, то интенсивности картин Эри, определяемые всеми точками объекта, являются просто аддитивными. Если же оно когерентно, то присутствует интерференция и тогда изображение математически представляет собой комбинацию картин Эри с комплексными амплитудами. Рэлей рассматривал оба предельных случая. При пред-

ставлении объекта в виде последовательности равно разнесенных точечных источников он использовал ряды Фурье для описания суммирования регулярно расположенных картин интенсивности Эри, формирующих изображение согласно этой модели. Он показал, что в условиях когерентности члены рядов соответствуют порядкам дифракции модели Аббе, упомянутой ранее.

Важным развитием этой модели свертки для формирования изображения является результат Дюфио (1946), который в своей книге «Интеграл Фурье и его приложение в оптике» выразил признательность Майкельсону и Рэлею как пионерам использования методов Фурье в физической оптике. Далее, в гл. 6 мы остановимся более подробно на вкладе Майкельсона. Что касается Рэлея, ученого с чрезвычайно разносторонними интересами, внесшего большой вклад почти во все области физики, следует отметить, что он получил свою Нобелевскую премию по физике в 1904 г. за исследования плотности газов и за открытие аргона.

Дюфио, по-видимому, первым выразил аналитически изображение непрерывного распределения интенсивности по объекту в виде свертки (он применял термин *faltung* – *складка*) этого распределения с функцией которую теперь мы называем функцией рассеяния точки, или импульсным откликом системы (картина Эри в простейшем случае). При этом он использовал комплексные амплитуды или интенсивности, соответствующие условиям освещенности. Применение теоремы свертки показало ему, что спектр пространственных частот изображения является произведением частотного спектра распределения объекта и отклика системы. Поэтому оптическую систему можно рассматривать как «переносящую» в той или иной степени компоненты пространственных частот объекта на плоскость изображения. Эта работа представляет нам один из наиболее важных в настоящее время аспектов рассматриваемого предмета, состоящий в том, что система формирования изображения обладает *передаточной функцией* (по Дюфио *facteur de transmission* – *коэффициент передачи*) для каждой частоты, поступающей на ее вход. Такое представление (к этому вопросу мы и переходим), очень удобно при конструировании и анализе оптических систем. Важность этого представления состоит в том, что, как можно убедиться, оно во многом дополняет модель Аббе.

Немалое значение для разработки модели формирования изображения на основе свертки имели работы [57, 15, 16] и др., где отмечалась взаимозависимость между этими аспектами оптики и идеями и методами, применяемыми при анализе электрических цепей различных сетей связи типа *линейных систем*. Краткое отступление здесь в другую область оправдывается ценным взаимным обогащением идеями, которое обеспечивается пониманием того, что в основе указанных двух направлений лежат одни и те же фундаментальные принципы.

В общем смысле «систему» можно определить как любое устройство, которое «отображает последовательность входных функций в пос-

ледовательность выходных функций». Как это отмечалось в разд. 4.4.1 в несколько иных выражениях, система является «линейной», если ее отклик на ряд входных сигналов является суммой откликов, полученных при отдельном воздействии. В приложении к оптике система является линейной, если, как это обычно бывает, пространственные частоты передаются системой с изменением только амплитуды и фазы, но без изменения частот.

При проектировании и анализе линейных электрических цепей один из методов состоял в исследовании выходного сигнала, полученного способом, описанным выше, для случая формирования оптического изображения, т.е. путем свертки входного сигнала (представленного последовательностью импульсов с изменяющейся амплитудой) с единичным импульсным откликом системы. Однако интегрирование, необходимое для исследования влияния различных фильтров, при этом становилось очень сложным. Еще более трудным было обращение свертки, применяемое при проектировании фильтров с условием создания определенных выходных сигналов по заданным входным. Именно применение теоремы свертки обеспечило во многих случаях столь необходимые упрощения. Из этой теоремы следует, что спектр временных частот на выходе линейной электрической системы является просто произведением входного частотного спектра и частотного спектра единичного импульсного отклика системы (ее передаточной функции). Интегрирование во временной области заменяется более простой операцией перемножения в частотной области. Более того, полная частотная характеристика нескольких последовательно включенных фильтров является просто произведением их собственных передаточных функций. Поэтому неудивительны замечания о том, что если бы теория цепей была ограничена временным подходом, то она никогда не получила бы такого развития.

Однако описанное выше прямое использование теоремы свертки как в системах связи, так и в системах формирования оптического изображения выдвинуло дополнительное требование, а именно *инвариантность* (или стационарность). Строго говоря, оно означает, что, например, в электрической цепи отклик на единичный импульс должен не зависеть от момента его подачи на вход, т.е. это должна быть система, *инвариантная во времени*. Таким же образом в системе формирования оптического изображения представление точечного объекта – функция рассеяния точки – должно быть одинаково по всему полю; это должна быть система, *инвариантная в пространстве* (ср. разд. 4.4.1). В начале следующего раздела будут обсуждаться следствия этого требования в обработке оптического изображения. (Рассматривается ситуация, при которой система не является *инвариантно линейной*. В целом же проблемы нелинейных систем выходят за рамки этой книги.)

Влияние электротехники на оптику, проявившееся и во многих других аспектах, сильно возросло в результате исследований Винера [63] и Шеннона [58] по теории статистических цепей и теории информации.

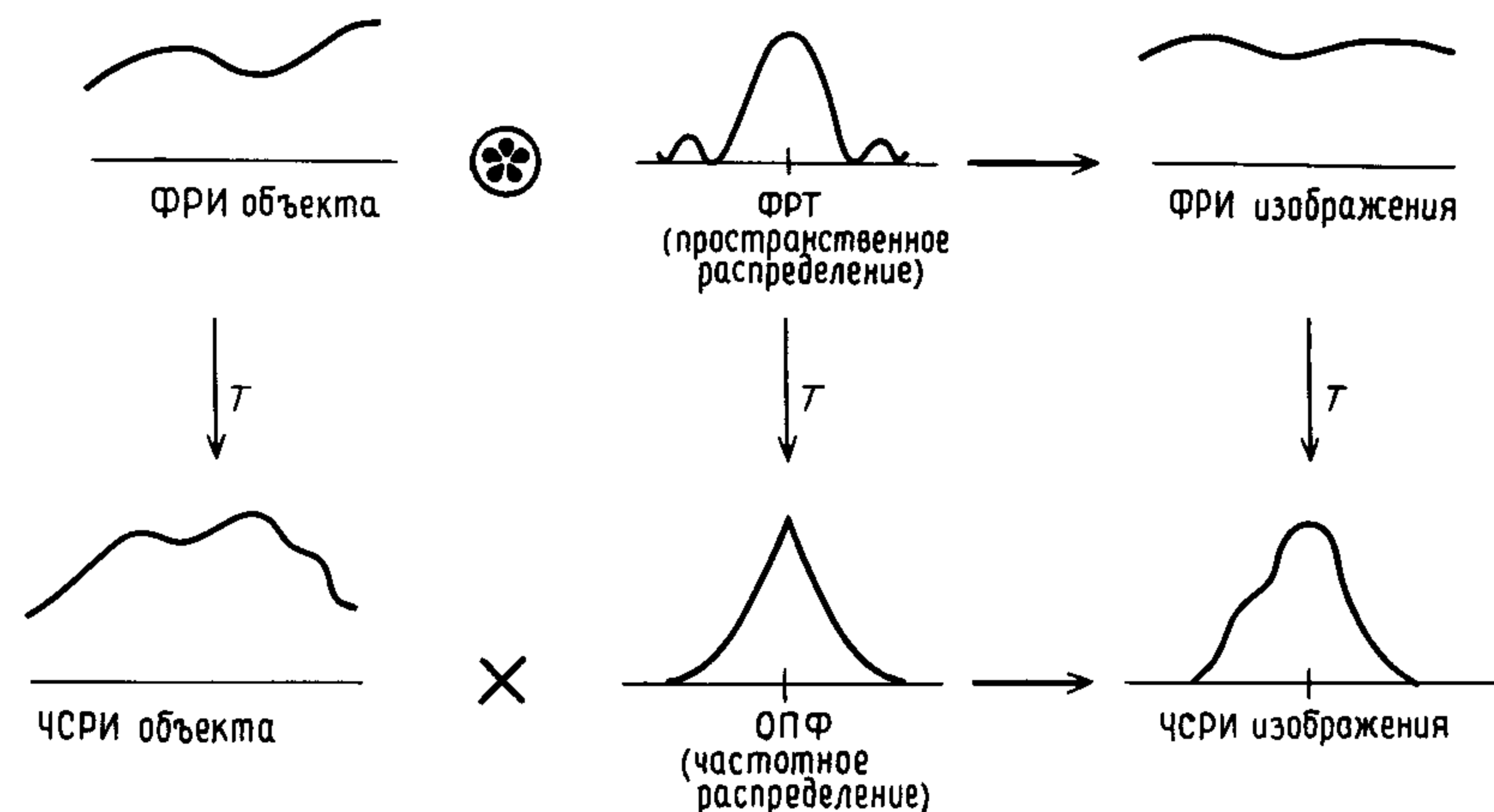


Рис. 5.1. Функция рассеяния и передаточная функция. Схема функциональных связей при некогерентном освещении. (T – преобразование Фурье; ФРИ – функция распределения интенсивности; ЧСРИ – частотный спектр распределения интенсивности; ФРТ – функция рассеяния точки; ОПФ – оптическая передаточная функция; $*$ – свертка; $\times$ – умножение.)

Это привело, например, к использованию корреляционных функций при обработке оптических изображений в реальном пространстве, с некоторыми аспектами которого мы познакомимся в разд. 5.5.

5.2. ФОРМИРОВАНИЕ НЕКОГЕРЕНТНОГО ОПТИЧЕСКОГО ИЗОБРАЖЕНИЯ

Описание в предыдущем разделе процесса формирования изображения как свертки обычно можно применять для практического использования в телескопах, камерах и т.д. в виде свертки интенсивностей. Для большинства фотографических, телевизионных и других систем было определено хорошее согласие между экспериментальными измерениями и расчетами, основанными на предположении, что освещенность при нормальных условиях, по существу, носит некогерентный характер (Барнс, 1971). К счастью, в этом есть много преимуществ, включая удобство использования телевизионных кадров, дисплеев на светодиодах и др. в качестве входных устройств систем обработки.

Однако если в системе присутствуют оптические aberrации, то отклик системы на единичный импульс, который в данном случае является функцией рассеяния точки системы (разд. 2.3), для разных точек в объектном поле может различаться. Такие изменения могут, как мы видели, сделать невозможным применение теоремы свертки. К счастью, если система хорошо скорректирована, остаточные эффекты aberrаций постоянны по области, где изображение любой точки в объектном поле достаточно интенсивно. В этом случае система назы-

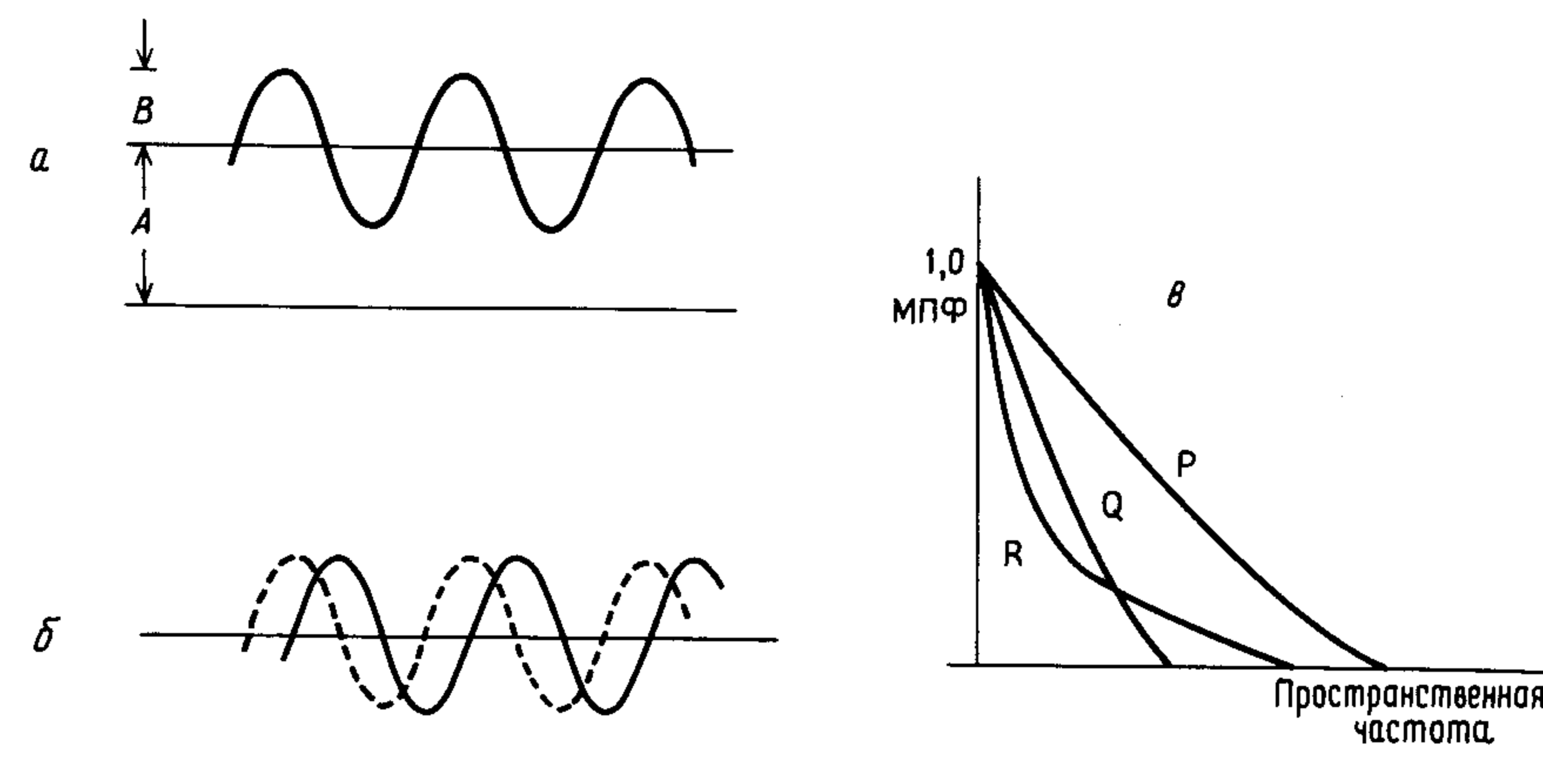


Рис. 5.2. Передаточные функции (некогерентное освещение). a – модуляция отдельного частотного компонента, определяемая величиной B/A ; b – соответствующий a фазовый сдвиг; c – примеры модуляционных передаточных функций (МПФ).

вается *изопланатической* (аналогично она называется *ахроматической*, когда положение точек изображения не зависит от длины волны света в интересующем нас диапазоне длин волн). При таких условиях практически система является пространственно инвариантной, и теорема свертки может быть использована непосредственно.

Как показано на рис. 5.1, хотя и чисто символически в одном измерении, приложение теоремы свертки создает частотный спектр распределения интенсивности изображения в виде произведения спектра частот распределения интенсивности (ЧСРИ) по объекту и преобразования Фурье от ФРТ. Преобразование от ФРТ является оптической передаточной функцией (ОПФ) системы.

Это означает, что каждая синусоидальная компонента пространственного распределения «преобразуется» в изображение без изменения ее частоты; меняются только ее амплитуда и фаза. Рассмотрим сначала амплитуду. На рис. 5.2, a представлено распределение интенсивности по объекту, состоящее из одного синусоидального компонента определенной пространственной частоты. «Эффективный размер» этого компонента выражается в зависимости от модуляции M , представленной выражением

$$M = B/A \quad (5.01)$$

[ср. определения в разд. 1.1 и уравнение (1.05)].

Отношение модуляций в изображении и непосредственно в самом объекте именуется *модуляционным коэффициентом передачи* для этой частоты; этот коэффициент является показателем эффективности си-

стемы формирования изображения на этой частоте и обычно меньше единицы (кроме нулевой частоты). Выраженный как функция частоты в соответствующем диапазоне он представляет собой *модуляционную передаточную функцию* (МПФ) системы.

Любые частотно-зависимые изменения фазы, вносимые системой, должны проявляться на изображении в виде горизонтальных смещений синусоидальных компонентов, составляющих изображение, как показано на рис. 5.2, б. В дополнение к МПФ имеется, следовательно, *фазовая передаточная функция* (ФПФ), хотя она часто не принимается в расчет.

Указанные две функции связаны с оптической передаточной функцией (ОПФ) (или *частотным откликом*) системы следующим образом:

$$(ОПФ) = (МПФ) \exp [i (ФПФ)]. \quad (5.02)$$

5.2.1. Определение передаточных функций

1) Расчетное определение передаточных функций

Вычисление ОПФ оптической системы по ее техническим данным производится несколькими методами. В одном из них для учета вклада aberrаций предусматривается расчет прохождения большого числа лучей через систему от единичной точки объекта. При равномерном разнесении лучей по апертуре линзы, распределение плотности точек, получившихся в плоскости изображения, дает распределение интенсивностей, соответствующее функции рассеяния точки. Затем преобразование Фурье определяет *геометрическую ОПФ* системы. Если система свободна от aberrаций, геометрическая ОПФ равна единице для всех частот; каждая точка объекта будет изображаться точкой. Поправка за дифракцию вносится умножением этой геометрической передаточной функции на передаточную функцию для эквивалентной дифракционно-ограниченной системы, т.е. идеальной системы, свободной от всех недостатков.

Другие методы связаны с детальным расчетом апертурной функции, включая эффекты aberrации. Это распределение комплексной амплитуды по апертуре мы будем обозначать $f(x)$, как и апертурную функцию в предыдущих главах. Его преобразование Фурье $F(u)$ является комплексной амплитудой дифракционной картины изображения точечного источника. Квадрат модуля соответствует ФРТ, а преобразование Фурье от него представляет собой ОПФ. В одном измерении это иллюстрируется на рис. 4.9 на хорошо известном примере $f(x)$, являющейся единичной прямоугольной функцией. Схема вычисления записывается в виде $a \rightarrow b \rightarrow z \rightarrow v$.

Иначе можно записать

$$ОПФ = T \{ |F(u)|^2 \},$$

где T обозначает преобразование Фурье, и, используя теорему автокор-

реляции (разд. 4.7.1), получаем

$$ОПФ = f(x) \odot f^*(x), \quad (5.03)$$

т.е. ОПФ определяется непосредственно путем вычисления автокорреляции апертурной функции. На рис. 4.9 это соответствует процессу $a \rightarrow b$.

2) Экспериментальное определение передаточных функций

Для экспериментального определения передаточных функций разработаны различные методы, некоторые из них достаточно сложные. Здесь следует рассмотреть только один простой метод, выбранный потому, что в нем еще раз наглядно подчеркивается смысл производимых операций. В этом методе МПФ определяется для одной частоты в каждый момент времени. Используется последовательность объектных экранов; каждый экран создает синусоидальную картину, соответствующую отдельной пространственной частоте; они напоминают несфокусированную тень от ряда равноотстоящих вязальных спиц. Отношение модуляции в изображении к модуляции в объекте определяется для картин пространственных частот, охватывающих необходимый частотный диапазон. Частотно-зависимые фазовые сдвиги, составляющие ФПФ, даются относительными положениями поперечных полос изображения и объекта на каждой калибровочной частоте. По своей природе этот метод скорее дает «функцию рассеяния линии» (ФРЛ), чем функцию рассеяния точки (ФРТ).

Рассмотренные передаточные функции обеспечивают более информативную оценку системы линз, чем простое измерение ее предела разрешения. На рис. 5.2, в это иллюстрируется кривыми МПФ. Кривая R соответствует линзе, свободной от всех aberrаций; относительная контрастность уменьшается с увеличением частоты до тех пор, пока не достигнет нулевого значения на пределе разрешения линзы (ср. с рис. 5.1). Кривые Q и R представляют линзы с aberrациями. Они показывают, что пока кривая R имеет частотный предел, превосходящий Q , она дает контраст (модуляцию) изображения меньше, чем на низких частотах. Выбор между двумя кривыми может быть сделан в соответствии с характером применения. Оптические передаточные функции не дают полного ответа на проблему оценки качества системы, особенно если в окончательном формировании изображения участвует глаз, хотя и являются более совершенными по сравнению с устаревшим и даже ошибочным измерением предела разрешения как критерия оптического качества. Глаз является плохой системой формирования изображения, но он связан со сложной обработкой данных в сетчатке и мозге. Это делает очень трудным предсказание и определение полного отклика в какой-либо конкретной ситуации.

5.3. ФОРМИРОВАНИЕ КОГЕРЕНТНОГО ОПТИЧЕСКОГО ИЗОБРАЖЕНИЯ

Экспериментальные исследования Эрнста Аббе по улучшению качества объективов для микроскопов, выполненные в 70-е годы прошлого века, легли в основу используемого сегодня подхода при рассмотрении когерентного изображения. Работая в Йенском университете как сотрудник (а позднее и как компаньон) Карла Цейса, Аббе обнаружил, что объективы микроскопов, изготовленные с наиболее тщательной коррекцией aberrаций, дают разрешение худшее, чем менее тщательно скорректированные объективы большей апертуры. В экспериментах с образцами, имеющими периодическую структуру, такими, как чешуйки насекомых и скелеты диатомей, он показал, что влияние большой апертуры связано с дифракцией на самом образце (свидетельство некоторой когерентности в условиях освещения).

Он продемонстрировал, каким образом дифракционные максимумы, сформированные в задней фокальной плоскости объектива, влияют на построение изображения; при этом максимумы высших порядков (или высшие пространственные частоты, как мы теперь считаем) ответственны за более тонкие детали в изображении. Аббе ввел волновую теорию («волновую оптику») в инструментальную оптику, бывшую ранее исключительно сферой применения геометрической (или «лучевой») оптики.

Интерпретация работы Аббе в терминах рядов Фурье была описана в важной статье А. Б. Портера в 1906 г. [46]. Он также показал их физическую суть серией изящных демонстрационных опытов. Например, он показал, как изменяется качество изображения объекта с периодической структурой в зависимости от потери дифракционных порядков в разных комбинациях.

Дополненный впоследствии применением преобразования Фурье к формированию изображения объектов с непериодической структурой, подход Аббе проявился в создании многих исключительно важных методов. Как уже упоминалось, они зависят главным образом от рассмотрения френгоферовой дифракции с точки зрения пространственных частот и доступности дифракционной картины как математически, так и экспериментально в случае использования когерентных условий.

5.3.1. Объекты с периодической структурой

Рассмотрим одномерную многоселевую прозрачную дифракционную решетку, показанную в предыдущих главах, которая играет роль объекта, изображение которого строится линзой (рис. 5.3). Волновые фронты, образующие разные дифракционные максимумы, приходят в разные фокальные точки задней фокальной плоскости линзы. Свет, проходящий

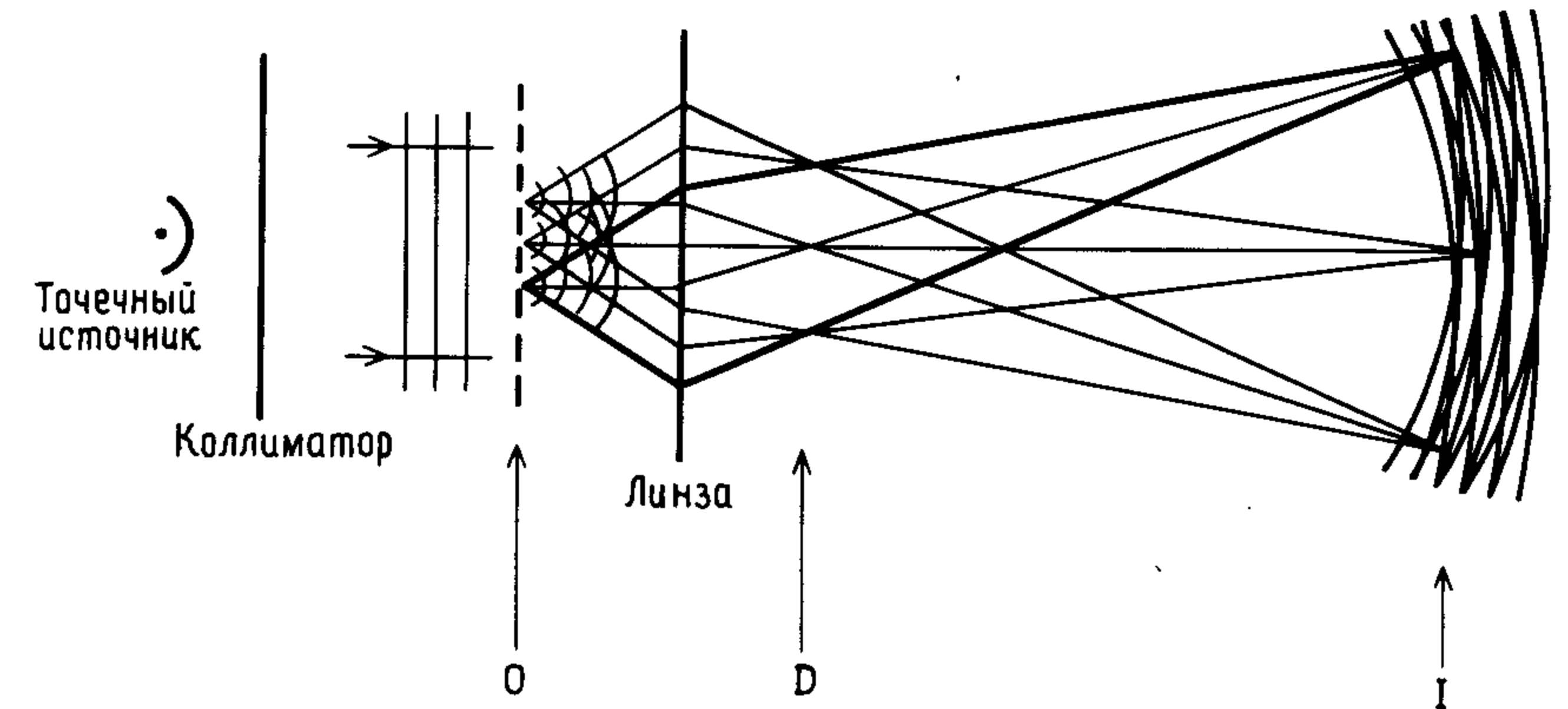


Рис. 5.3. Формирование изображения в микроскопе согласно Аббе (O – плоскость объекта, D – дифракционная плоскость; I – плоскость изображения) [5].

через эти фокусы на дифракционной плоскости, складывается на плоскости изображения, восстанавливая оптическое изображение объекта. Предположим условно, что линза совершенна и обладает бесконечной апертурой. Отметим, что, как обозначено двумя жирными линиями на рисунке, здесь отсутствует противоречие между этим волновым описанием и методом лучей геометрической оптики.

Взятая сама по себе *любая* пара фокусов создает ряд синусоидальных групп на плоскости изображения. Это напоминает опыт Юнга (разд. 1.1), где пара апертур действует таким же образом. В этом смысле формирование изображения можно рассматривать как двойной процесс дифракции (идея, выдвинутая Цернике около 1935 г. [64]).

Чтобы увидеть, каким способом строится изображение объекта типа решетки, показанной на рис. 5.3, рассмотрим вклады в изображение, вносимые парами дифрагированных лучей, включающих разные порядки дифракции.

Рис. 5.4 показывает схематически образование пары дифрагированных лучей n -го порядка, испускаемых объектом в плоскости изображения, они интерферируют, создавая распределение гармоник освещенности с периодом D'_n , задаваемым выражением

$$D'_n \sin \theta'_n = \lambda. \quad (5.04)$$

Условием для формирования дифракционных максимумов n -го порядка от объекта является выражение вида

$$D \sin \theta_n = n \lambda, \quad (5.05)$$

откуда

$$D'_n = \frac{D}{n} \left(\frac{\sin \theta_n}{\sin \theta'_n} \right)$$

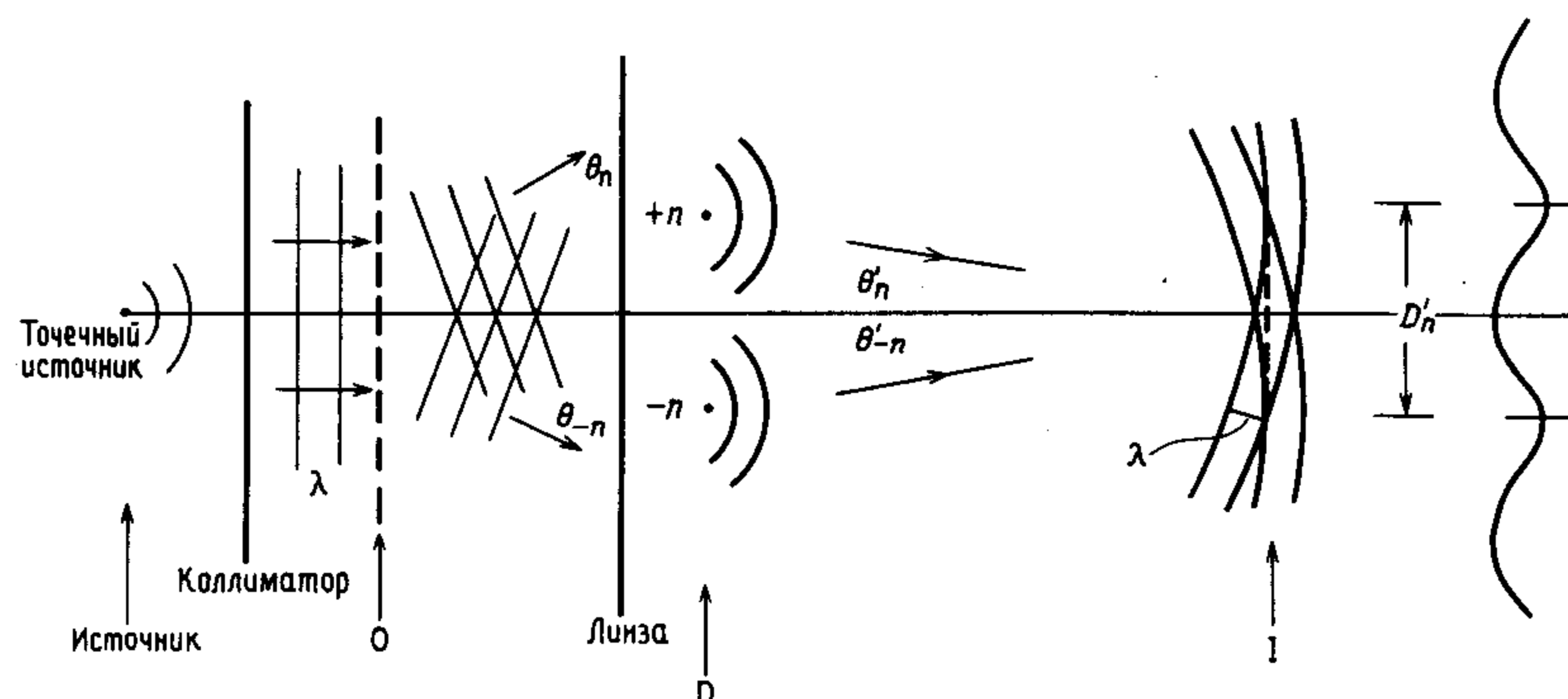


Рис. 5.4. Формирование изображения: дифракция и восстановление [О—объект (дифракционная решетка с периодом D); D —дифракционная плоскость; I —изображение].

Член в скобках является постоянным по величине и связан с увеличением линзы. Итак, мы имеем

$$D_n' \propto D/n. \quad (5.06)$$

Пара максимумов первого порядка интерферирует в плоскости изображения, создавая простые гармонические вариации освещенности, которые соответствуют основному периоду решетки. Этот период представляет собой минимальную информацию об объекте без тонких деталей его оптической структуры. Каждая пара последующих максимумов более высокого порядка добавляет последовательно к общей освещенности гармоники более короткого периода ($\propto D/n$), которые формируют изображение. Все детали изображения строятся способом, вполне аналогичным фурье-синтезу. В разд. 3.4.1 было показано, что дифракционные максимумы сами заключают в себе фурье-анализ решетчатого объекта, и была сделана ссылка на дифракционную плоскость, описываемую как фурье-плоскость. Поэтому процесс формирования изображения в рассматриваемом нами примере можно интерпретировать как двойную фурье-обработку: с дифракционной картиной в качестве фурье-анализа решетки и изображением в качестве фурье-синтеза данного фурье-анализа. Такая интерпретация особенно очевидна, если вспомнить принцип обратимости. Все порядки дифракции, которые создают изображение путем суммирования гармоник, возвращают к решетчатому объекту, где они рекомбинируют, образуя первоначальное распределение освещенности (апертурной функции) на решетке.

Важно понять, что это описание двойного преобразования Фурье полностью согласуется с интерпретацией двойной дифракции Цернике, и мы имеем дело с разными способами выражения одного и того же явления.

Для построения идеального изображения необходим бесконечный фурье-синтез, требующий не только образования бесконечного числа порядков дифракции, но и сплошного заполнения оптической системы. Очевидно, что оба этих требования невыполнимы. Уравнение (5.05) показывает, как величины D и λ накладывают ограничение на число возможных дифракционных максимумов, которые могут быть созданы, и, разумеется, любой объектив имеет конечную апертуру ($\sin\theta < 1$).

В другом предельном случае, если условия допускают только пары дифракционных максимумов нулевого и первого порядков, поступающих в объектив, в изображении может быть разрешен только основной период D решетки, и при том, как указано выше, лишь как простое синусоидальное изменение освещенности. Уравнение (5.05) показывает, что сообразно с этим наименьший период объекта, поддающийся разрешению, определяется условием

$$D \sin\theta = 1\lambda,$$

т. е.

$$D = \lambda/\sin\theta, \quad (5.07)$$

где $\sin\theta$ —уже упоминавшаяся нами *численная апертура* объектива (разд. 2.3). Это дает *предел разрешения*, равный примерно длине волны используемой освещенности.

Предел разрешения может быть сделан меньше указанного, так как комбинация нулевого порядка с одним только максимумом первого порядка вполне достаточна для формирования изображения с основным периодом решетки. Используя сходящийся пучок, можно добиться того, что лучи будут приходить на противоположные края апертуры объектива и тогда предел разрешения $\sim \lambda/2$. Это минимальное условие присутствия максимумов нулевого и первого порядков выражает *принцип Аббе*.

Точно такие же соображения применяются в случае двумерной решетки, т. е. решетки в форме двумерной структуры, каждая точка которой имеет идентичную апертуру (разд. 2.6). В случае сборной решетки каждая «апертура» может быть, например, малым отверстием или группой отверстий. Дифракционная картина в этом случае представляет собой двумерную решетку пятен освещенности, причем порядок дифракции каждого пятна определяется двумя целыми числами (сравните три числа в уравнениях (2.18) для трехмерной решетки). Рекомбинация, вторая стадия в формировании изображения, выполняется точно таким же способом, как и в одномерном случае.

В разд. 5.3.3 описаны экспериментальные наблюдения этих эффектов.

5.3.2. Непериодические объекты

Подход, рассмотренный в предыдущем разделе, можно применить и к случаю непериодических объектов, потому что *дискретные* порядки дифракции не являются его необходимой предпосылкой. Непериодический объект можно считать эквивалентным одной апертуре (щели) решетки, и мы знаем, что в этом случае используется преобразование Фурье вместо рядов Фурье. Дифракционная картина в фокальной плоскости линзы представляет собой картину непрерывного рассеяния с угловым изменением амплитуды и фазы, зависящим от апертурной функции; это – преобразование Фурье от функции амплитудного распределения по объекту (ср. оценку линзы как «преобразователя Фурье» в разд. 4.2). Восстановление этой картины в плоскости изображения сводится к суммированию интерференционных полос, создаваемых парой дифрагированных лучей (под углом $\pm \theta'$ на рис. 5.4), но с непрерывным диапазоном разнесения полос и ориентаций. Формирование изображения может быть описано как процесс двойного преобразования Фурье. Это описание в общем применимо как к периодическим, так и к непериодическим объектам, поскольку даже первые из них имеют конечный размер, что позволяет говорить об изображении как о преобразовании дифракционной картины, независимо от природы объекта. Мы уже использовали эту идею в разд. 4.5.

5.3.3. Примеры преобразований в оптике

Схема оптического дифрактометра для демонстрации и использования принципов, описанных ранее, показана в упрощенном виде на рис. 5.5. Обычно используется гелий-неоновый лазер с расширителем пучка для обеспечения освещенности с почти идеальной когерентностью (временной и пространственной) по всему плоскому волновому фронту в положении О, где расположены объектные маски. Дифракционная картина (преобразование Фурье), создаваемая маской в положении О, формируется в фокальной плоскости D объектива L_1 , а изображение (двойное преобразование) от О формируется на плоскости I. На практике для получения дифракционных картин приемлемого размера L_1 должен быть длиннофокусной линзой или соответствующей эквивалентной системой (например, комбинация фотографии и телевидения). Вторая линза L_2 (ее положение обозначено на рисунке пунктирной линией) нужна для формирования действительного изображения на приемлемом расстоянии от объектной маски.

Фоторегистрация дифракционных картин и изображений по своей природе является записью только интенсивностей (кроме голографии). Иллюстрации в следующих примерах описывают, следовательно, *оптические преобразования*, а не преобразования Фурье. Они широко представлены в полном собрании сочинений по оптическим преобразованиям Х. Липсона и его коллег.

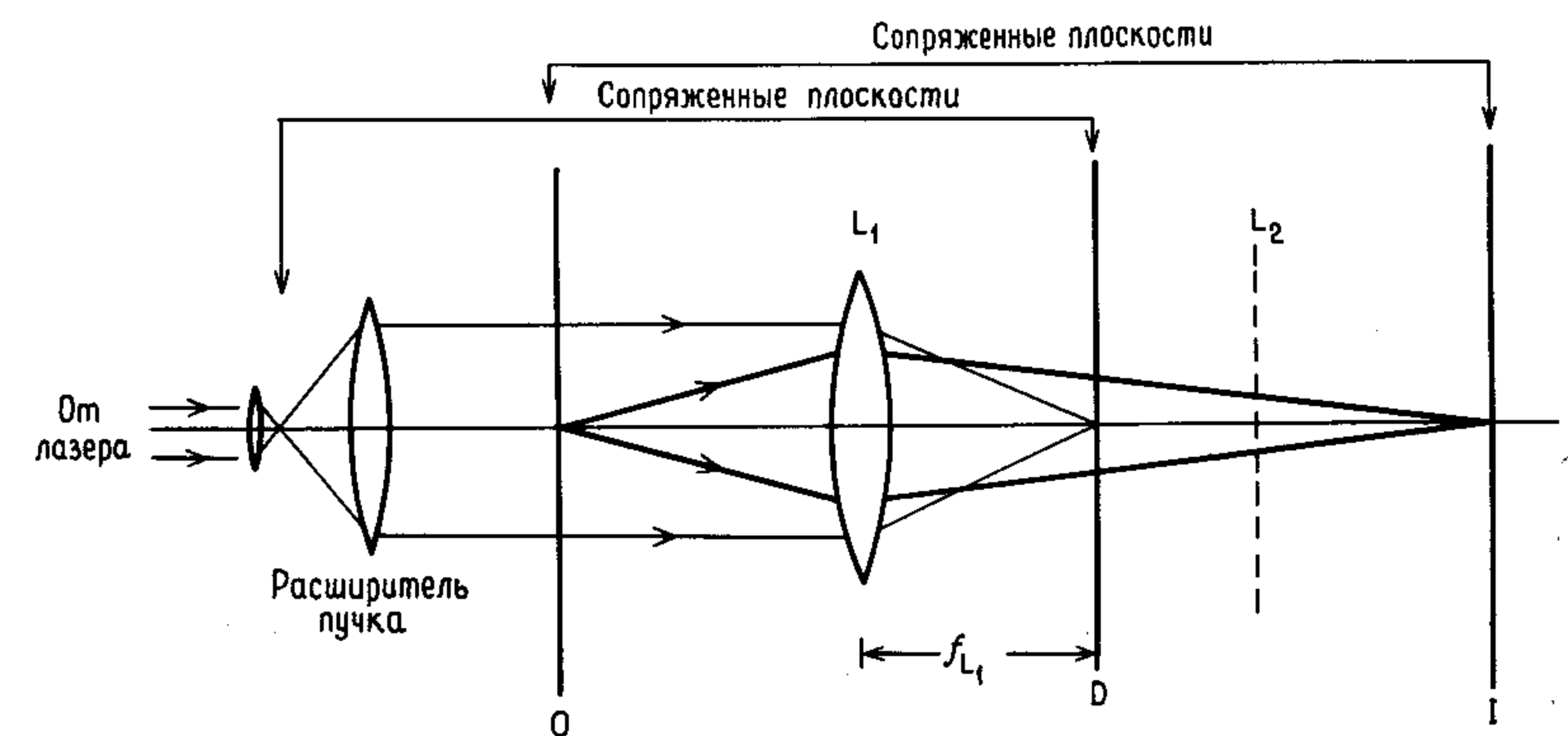


Рис. 5.5. Основная схема для наблюдения оптических преобразований (О – плоскость объекта; D – дифракционная плоскость; I – плоскость изображения).

1) Дифракция на объектных масках

Последовательность а–д рис. 5.6 иллюстрирует некоторые из основных положений, которые упоминались в связи с ролью дифракции на первой стадии формирования изображения в когерентном свете. В каждом примере представлены объектная маска и создаваемое ею оптическое преобразование.

- а) *Объектная маска*: двухмерная решетка из точечных отверстий. Ее можно рассматривать как решетчатую структуру из δ -функций. Маска содержит гораздо больше повторов, чем показано на рисунке. *Преобразование* (почти всегда двухмерные ряды Фурье, поскольку объект содержит много повторов): обратная структура – ее геометрия и размеры находятся в обратной зависимости с аналогичными параметрами структуры объекта. Интенсивности примерно постоянны, как это и следует ожидать для малых диафрагм, аппроксимируемых δ -функциями (т.е. большой диск Эри).
- б) *Объектная маска*: одна апертура, содержащая неповторяющуюся, нерегулярную группу отверстий. *Преобразование*: нерегулярная картина, в которой освещенность изменяется с направлением в зависимости от расположения и размеров отверстий в объектной маске.
- в) *Объектная маска*: свертка вариантов а и б. Как и в случае а, маска содержит гораздо больше повторов, чем здесь показано. *Преобразование*: произведение отдельных преобразований а и б. Этот пример иллюстрирует теорему свертки, а также показывает, что картина, создаваемая решеткой, является выборкой картины, создаваемой единичным элементом решетки (т.е. ее преобразование). Выборка проводится в направлении, определяемом размерами решетчатой структуры и, разумеется, длиной волны света.

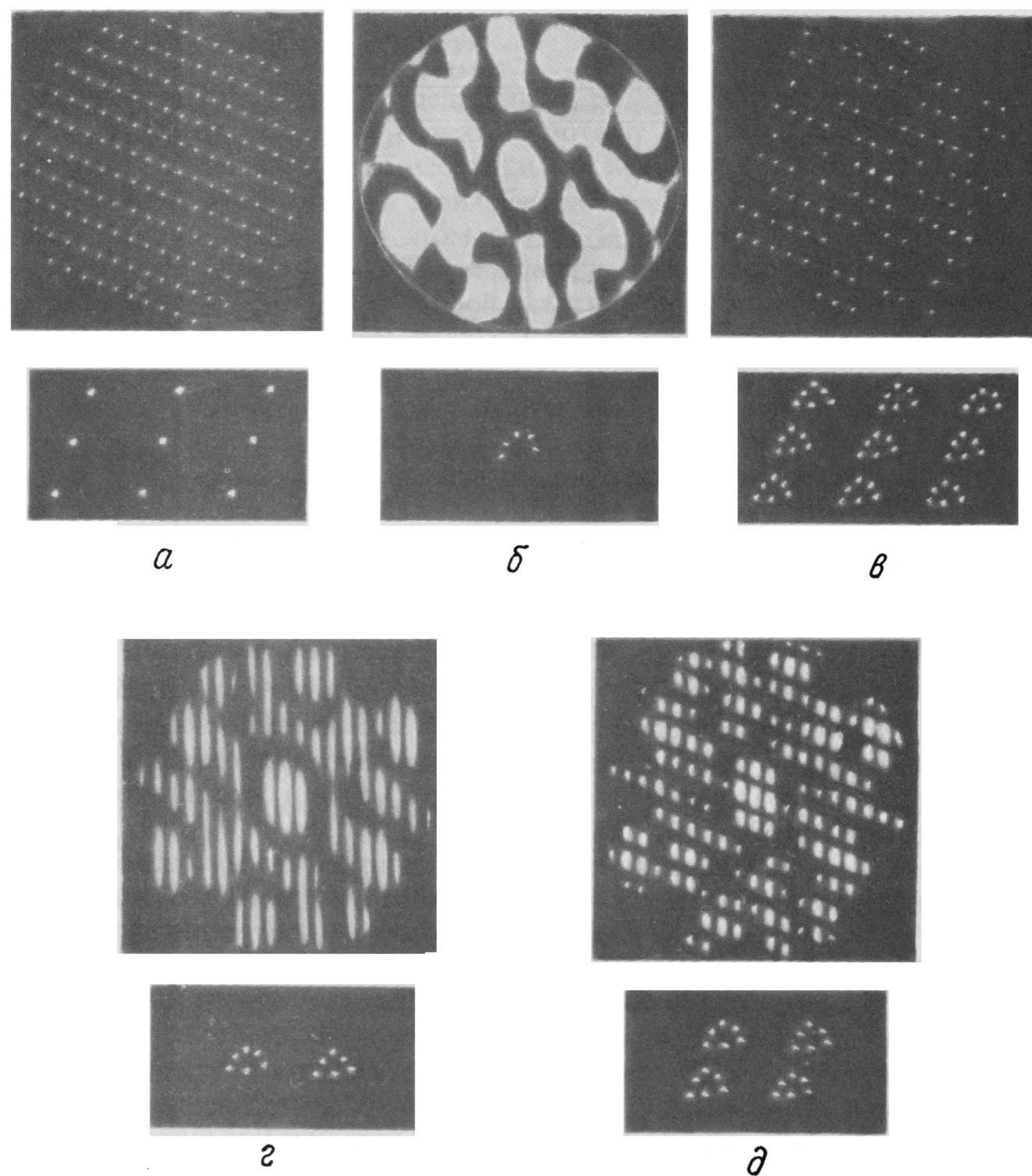


Рис. 5.6. Оптические преобразования. Каждому преобразованию соответствует показанная внизу маска, за исключением случаев *a* и *в*, для которых маски содержат намного больше периодов, чем показано на рисунке [60].

- г) *Объектная маска*: двухапертурная «решетка», которая является сверткой одной апертуры в случае *б* с двумя δ -функциями.
Преобразование: это опять произведение двух отдельных преобразо-

- д) *Объектная маска*: свертка одной апертуры в случае *б* с четырьмя δ -функциями, определяющими решетчатую структуру в случае *a*.
Преобразование: преобразование одной апертуры в этом случае умножается на две последовательности полос $\cos^2$, которые определяют обратную структуру в случае *a*. Снова демонстрируется наличие выборки.

Изыщные примеры использования оптических преобразований были обнаружены в рентгеновской кристаллографии, где, как отмечено в гл. 2, формирование изображений атомов не может быть выполнено непосредственно, потому что отсутствуют линзы, которые могут быть использованы для сведения дифрагированных рентгеновских лучей. Отметим, что если зарегистрированы только интенсивности, то фурье-суммирование не может быть выполнено ни аналитически, ни экспериментально из-за отсутствия данных о фазах. В годы формирования указанного направления исследований У.Л. Брэгг сыграл ключевую роль в разработке методов оптического фурье-анализа для рассмотрения и решения этой и других проблем рентгеновской кристаллографии. Несмотря на то что развитие ЭВМ привело к машинным методам решения «фазовой проблемы», работа Брэгга явилась важным вкладом в широкую область оптической обработки. В качестве основной литературы по развитию и применениям оптических методов к дифракции рентгеновских лучей, читатель может обратиться к работам, упомянутым в начале этого раздела.

Пример аналогии между дифракцией рентгеновских лучей на кристаллах и первой стадией формирования оптического изображения решетчатого объекта показан на рис. 5.7. На рис. 5.7, *a* изображена часть оптической маски, представляющей собой двухмерную проекцию кристаллической структуры фталоцианина; на рис. 5.7, *б* показана оптическая дифракционная картина, создаваемая ею [10]. Рис. 5.7, *б* согласуется с данными рентгеновских исследований не только в отношении геометрии расположения пятен но и по соответствию экспериментально наблюдаемой рентгеновской интенсивности картине на рисунке. Исторически этот метод вначале применялся для определения неизвестной кристаллической структуры путем изготовления «пробных» масок на основе химических и других соображений. Он был существенно упрощен при дальнейшем развитии техники (см. конец раздела 2), когда было показано, что основная ячейка и только три периода вполне достаточны в качестве маски, поскольку они определяют структуру, на которой основана двухмерная проекция кристалла. Это иллюстрируется рис. 5.6, где в случае *в* в качестве маски было использовано большое

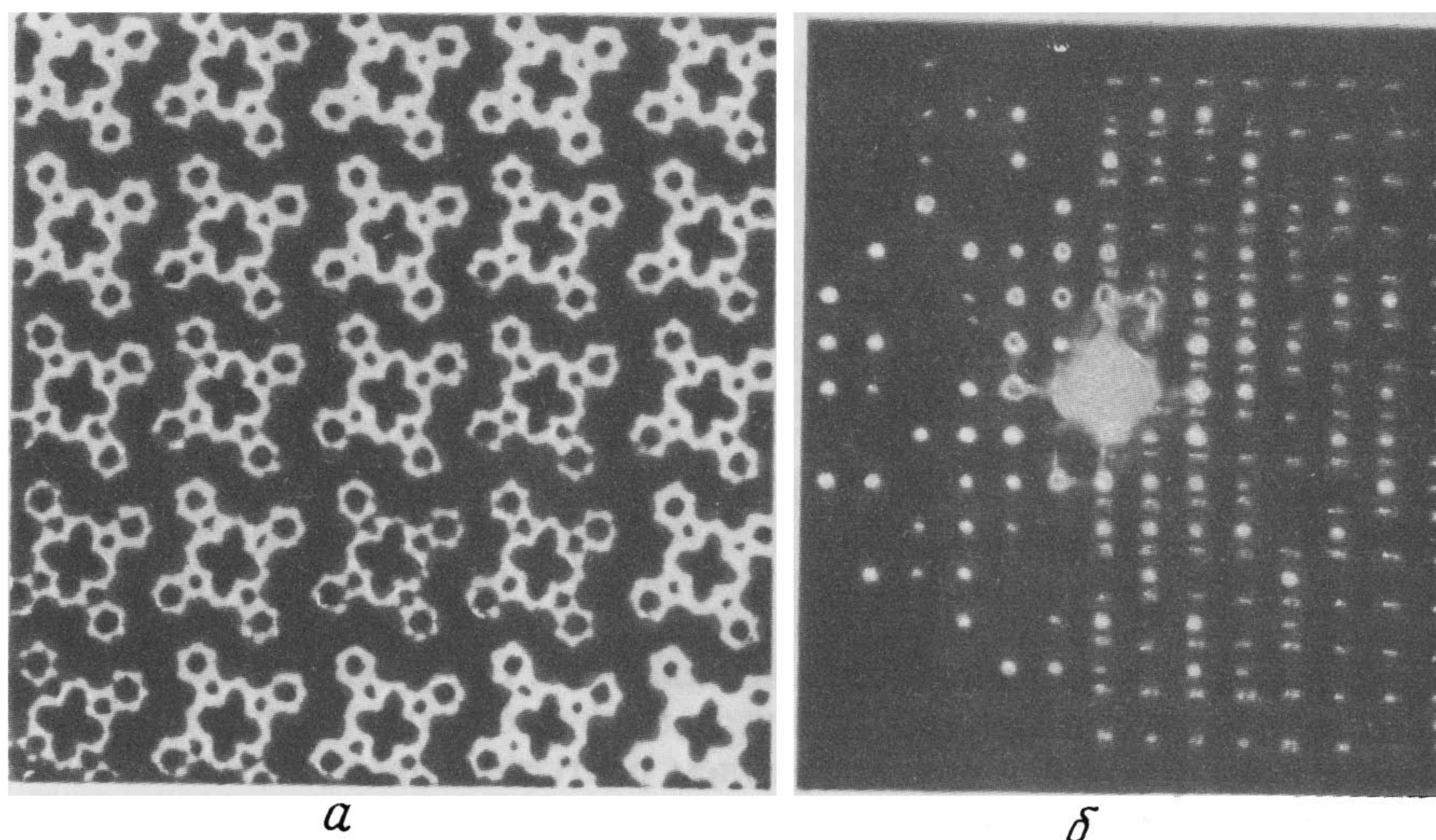


Рис. 5.7. Первая стадия формирования изображения. *a* – маска, представляющая двухмерную проекцию кристаллической структуры фталоцианина; *б* – картина оптической дифракции, полученная с маской *a*. (Nature, 154, 69. Copyright Macmillan Journals Limited, 1944).

2) Восстановление изображения

Рис. 5.8 представляет другой пример из области рентгеновской кристаллографии, на этот раз относящийся ко второй стадии формирования изображения – его восстановлению. Рис. 5.8, *a* является схематическим «видом» вниз по оси кристаллической структуры минерала диопсида $\text{CaMg}(\text{SiO}_3)_2$, который У. Л. Брэгг использовал в нескольких случаях для иллюстрации оптических принципов, лежащих в основе «рентгеновского анализа» кристаллических структур [6, 8, 9, 10].

Атомы показаны в виде черных кружков; наибольшие из них представляют атомы Ca и Mg, которые накладываются друг на друга, а наименьшие – атомы Si и O (которые образуют цепочки, тянущиеся через структуру). На рис. 5.8, *б* мы имеем маску, содержащую небольшие отверстия для представления экспериментальных данных, полученных при наблюдении рентгеновской дифракции; площадь отверстий пропорциональна квадратному корню из интенсивности рентгеновских лучей. Рисунок *в* представляет изображение, сформированное как оптическое преобразование *б*, и оно хорошо согласуется с *a*. Восстановление достигло цели, так как почти все пучки дифрагированных рентгеновских лучей имеют одинаковую фазу, в основном определяемую очень сильным рассеянием на атомах Ca и Mg, которые являются центрами симметрии. Примеры пучков с разными фазами можно получить, если поместить перед каждым отверстием кусочек слюды и путем вращения отдельных

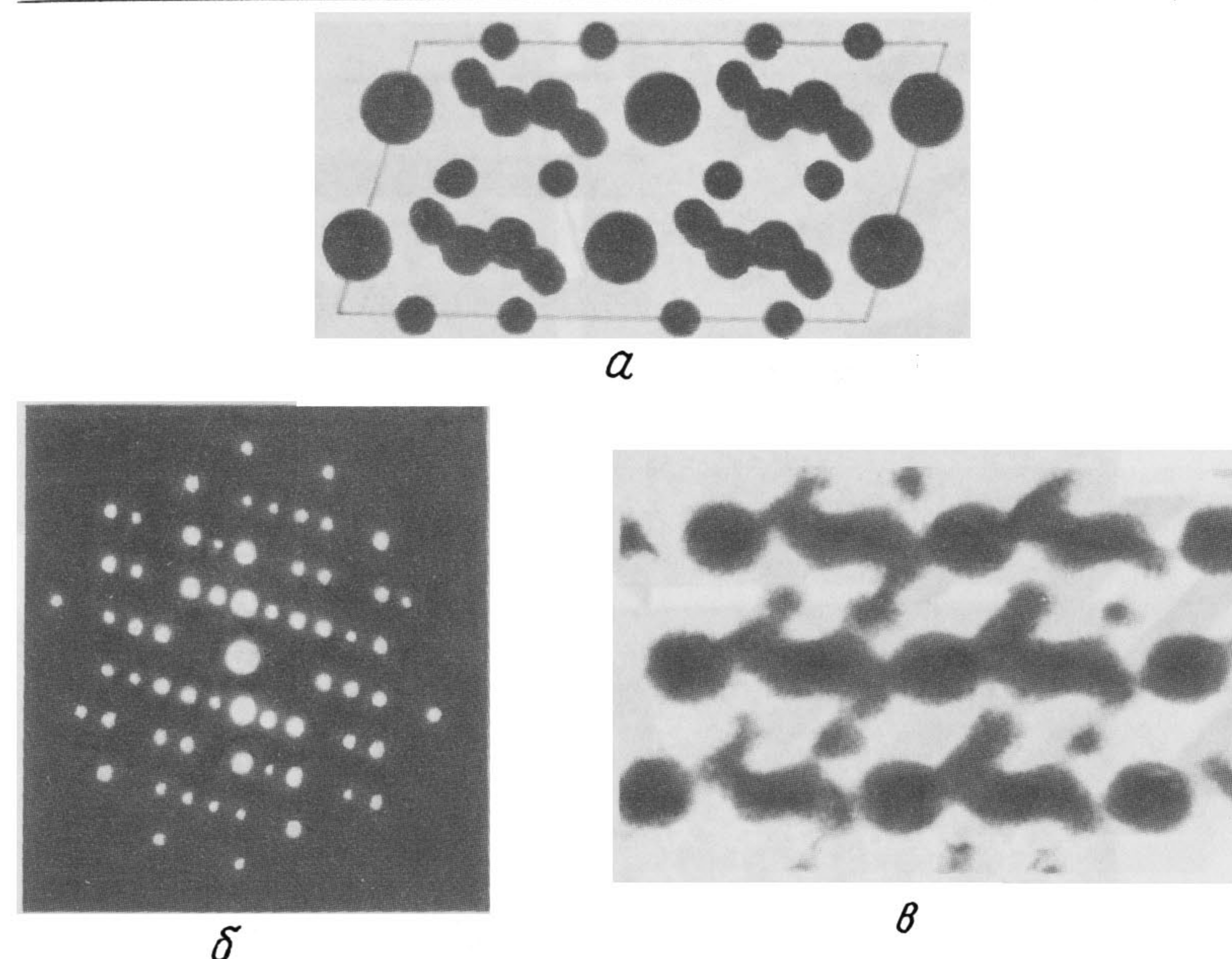


Рис. 5.8. Вторая стадия формирования изображения. *a* – представление двухмерной проекции кристаллической структуры диопсида, $\text{CaMg}(\text{SiO}_3)_2$; *б* – оптическая маска, соответствующая дифракционной картине рентгеновских лучей; *в* – оптическое восстановление изображения с использованием маски *б*. (Nature, 143, 678; 149, 470. Copyright Macmillan Journals Limited, 1939, 1942.)

кусочков добиться соответствующих фазовых сдвигов. (Вопреки первому впечатлению, этот метод не требует использования поляризованного света для восстановления оптического изображения.)

Брэгг продемонстрировал, что такие изображения можно считать состоящими из синусоидальных гармоник, создаваемых парами дифракционных максимумов, как описано в разд. 5.3.1. На рис. 5.9 показаны наборы групп синусоид, подобных тем, что создавались отдельно парами разных дифракционных максимумов; их фазы проявляются в расположении «гребней» и «впадин» относительно углов единичной ячейки. Брэгг последовательно экспонировал лист фотобумаги для получения высококачественных картин, подобных тем, что представляют дифракционные эффекты, производимые маской, показанной на рис. 5.8, *б*. Принимая выдержки ниже уровня, приводящего к излишнему потемнению фотобумаги, он получил картину, подобную показанной на рис. 5.8, *в*.

Использование рентгеновских лучей на первом этапе формирования изображения в сочетании с применением видимого света и линзы для завершения второго этапа известно [11], как *двухволновая микроскопия*, хотя Брэгг предпочитал термин «рентгеновская микроскопия». Этот ме-

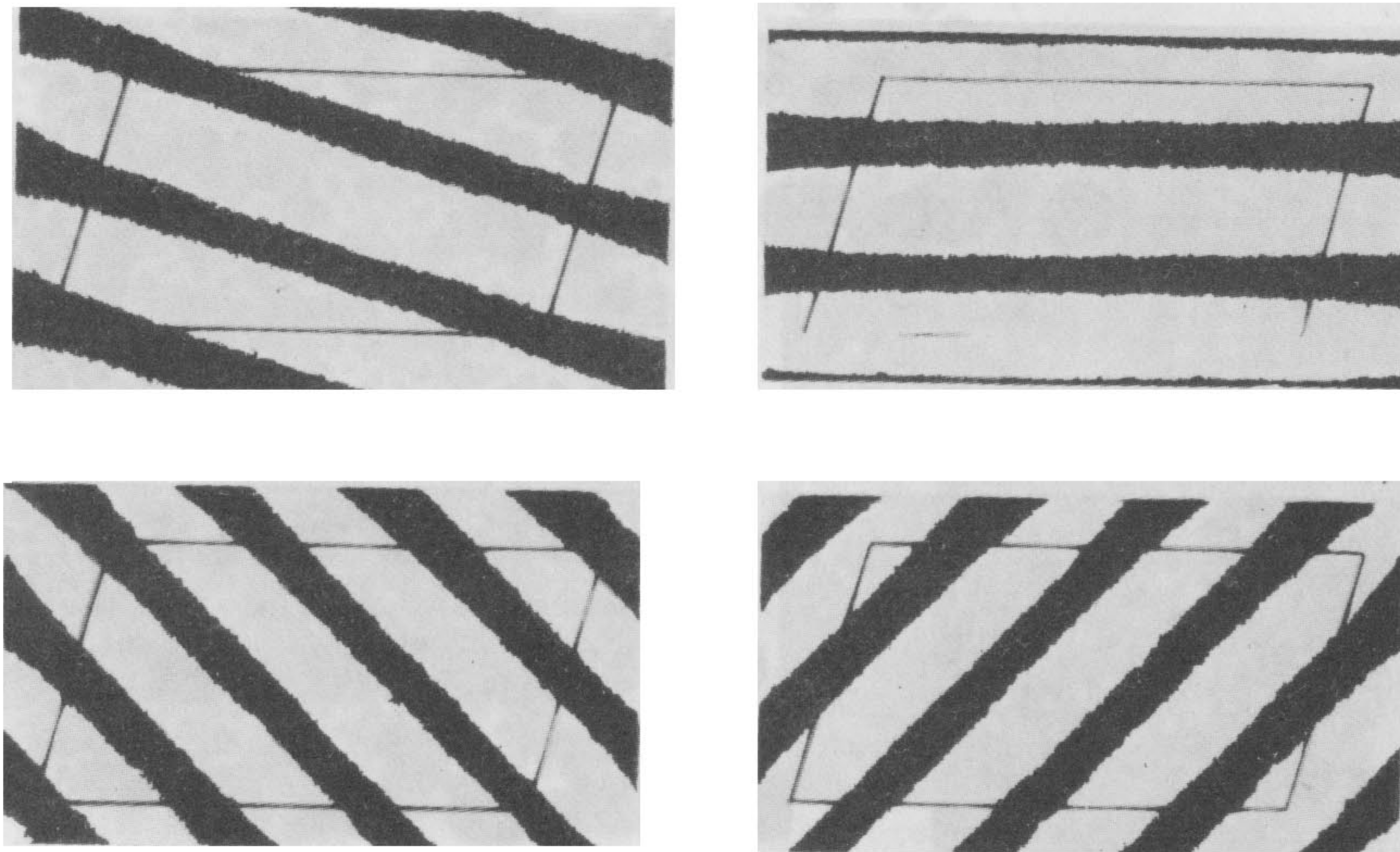


Рис. 5.9. Синусоидальные полосы света и тени, сформированные в плоскости изображения парой дифракционных максимумов [7].

тод стал основным источником вдохновения для Д. Габора при изобретении им голографии (разд. 5.4) и стал основой метода использования видимого света в обработке изображений электронного микроскопа, разработанного А. Клягом и его коллегами (разд. 5.5.1).

Как упоминалось выше, отсутствие экспериментальных данных о фазе рентгеновских лучей, которые необходимы, чтобы формирование изображения завершилось оптическим восстановлением, не позволяет использовать этот метод для обычного определения кристаллической структуры. Решение указанной проблемы – «проблемы фазы» рентгеновской кристаллографии – было достигнуто в основном благодаря развитию математических методов, позволяющих делать выводы о фазах с помощью ЭВМ, исходя из данных об интенсивности рентгеновских лучей и существования определенных граничных условий, таких, как конечные размеры атомов и отсутствие отрицательной электронной плотности.

Ценность оптической аналогии состоит в том, что она и теперь обеспечивает возможность определения структуры, кроме того она полезна и в различных специальных приложениях, таких, как изучение хаотичности и других эффектов в полу- и поликристаллических материалах. Современная методика, используемая в работе оптического преобразователя (примеры представлены на рис. 5.6 и 5.10), в основном была разработана Липсоном, Тейлором и др. в Манчестере (Великобритания) начиная с 50-х годов. В книге [27] содержится большой набор иллюстраций по этой теме.

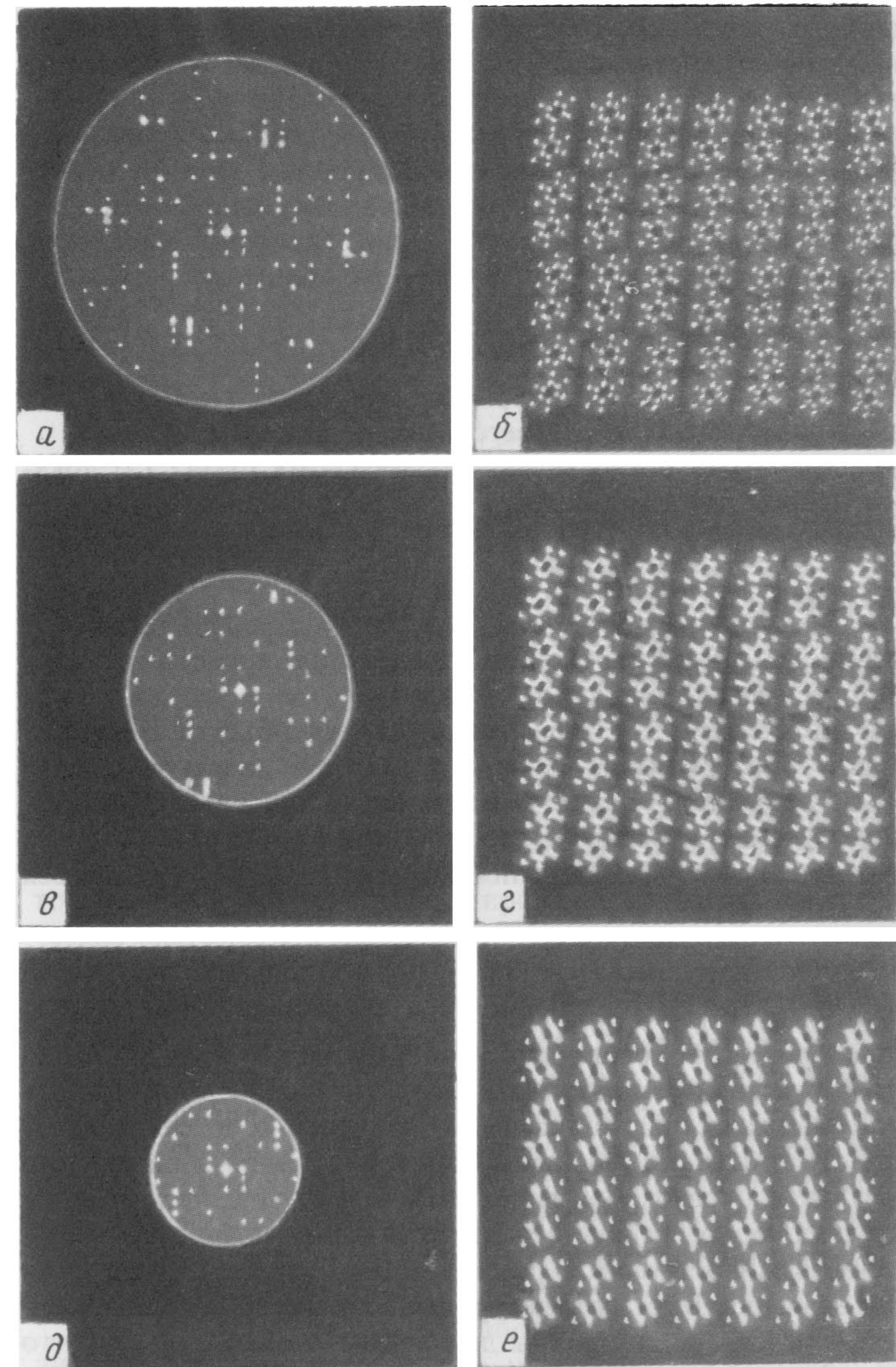


Рис. 5.10. Ухудшение качества изображения при удалении высоких пространственных частот путем фильтрации в плоскости дифракции [60].

3) Апертура линзы и aberrации

На рис. 5.10, а, в, д круглые апертуры уменьшающегося размера были введены в центр плоскости дифракции для ограничения количества информации, определяющей формирование изображения. Удаление внеш-

них дифракционных пятен маски существенно снижает вклад в изображение высоких пространственных частот. Результаты, показанные соответственно на рис. 5.10, б, г, е, лишены тонких деталей.

Для расчета влияния апертуры мы можем утверждать, что имеется апертурная функция, которая умножает дифракционную картину объекта на 1 внутри ее границ и на 0 снаружи. Поскольку изображение является преобразованием той части картины в дифракционной плоскости, которая обеспечивает вклад в изображение, то мы можем записать

$$\text{Изображение} = T[T(\text{объект}) \times \text{АФ}], \quad (5.08)$$

где T обозначает преобразование Фурье, а для апертурной функции (АФ) должны быть определены соответствующие пределы.

Эквивалентность этого представления рэлеевскому (разд. 5.2) доказывается применением теоремы свертки к вышеприведенному уравнению, которое сразу дает

$$\text{Изображение} = \text{Объект} * T(\text{АФ}) \quad (5.09)$$

т.е. изображение является сверткой распределения яркости объекта с картиной Эри, даваемой апертурой линзы.

В соответствии с моделью Рэлея каждая точка объекта считается здесь источником и изображается как картина Эри апертуры линзы.

Относительно aberrаций можно сказать, что их влияние видоизменяет апертурную функцию. Однако принципы, о которых речь шла выше, остаются в силе, и для рассмотрения aberrаций следует обратиться к специальной литературе.

5.4. ГОЛОГРАФИЯ

Изобретение в 1948 г. голографии Д. Габором, за которое ему была присуждена Нобелевская премия по физике 1971 г., основано на его работе по улучшению качества изображений, получаемых в электронной микроскопии. Результаты, полученные в 40-х годах с электронными микроскопами, оказались разочаровывающими, поскольку, несмотря на стократное улучшение в разрешающей способности по сравнению с лучшими оптическими микроскопами, разрешение оставалось далеким от теоретического значения. Быстрые электроны, используемые в электронной микроскопии, имеют длину волны де Бройля около $1/20 \text{ \AA}$, так что атомы должны разрешаться; однако практически предел в то время составлял около $12 \text{ \AA}$. Основной причиной неудачи было наличие aberrаций, связанных с использованием электронных линз. Именно при поиске путей решения этой проблемы Габором был создан метод, названный им *восстановлением волнового фронта*. Частично его идея исходила из принципов, заложенных в двухволновой микроскопии. У.Л. Брэгга (разд. 5.3.3). Он полагал, что если ему удастся зарегистрировать фазы так же, как и интенсивности в изображении электронного микроскопа,

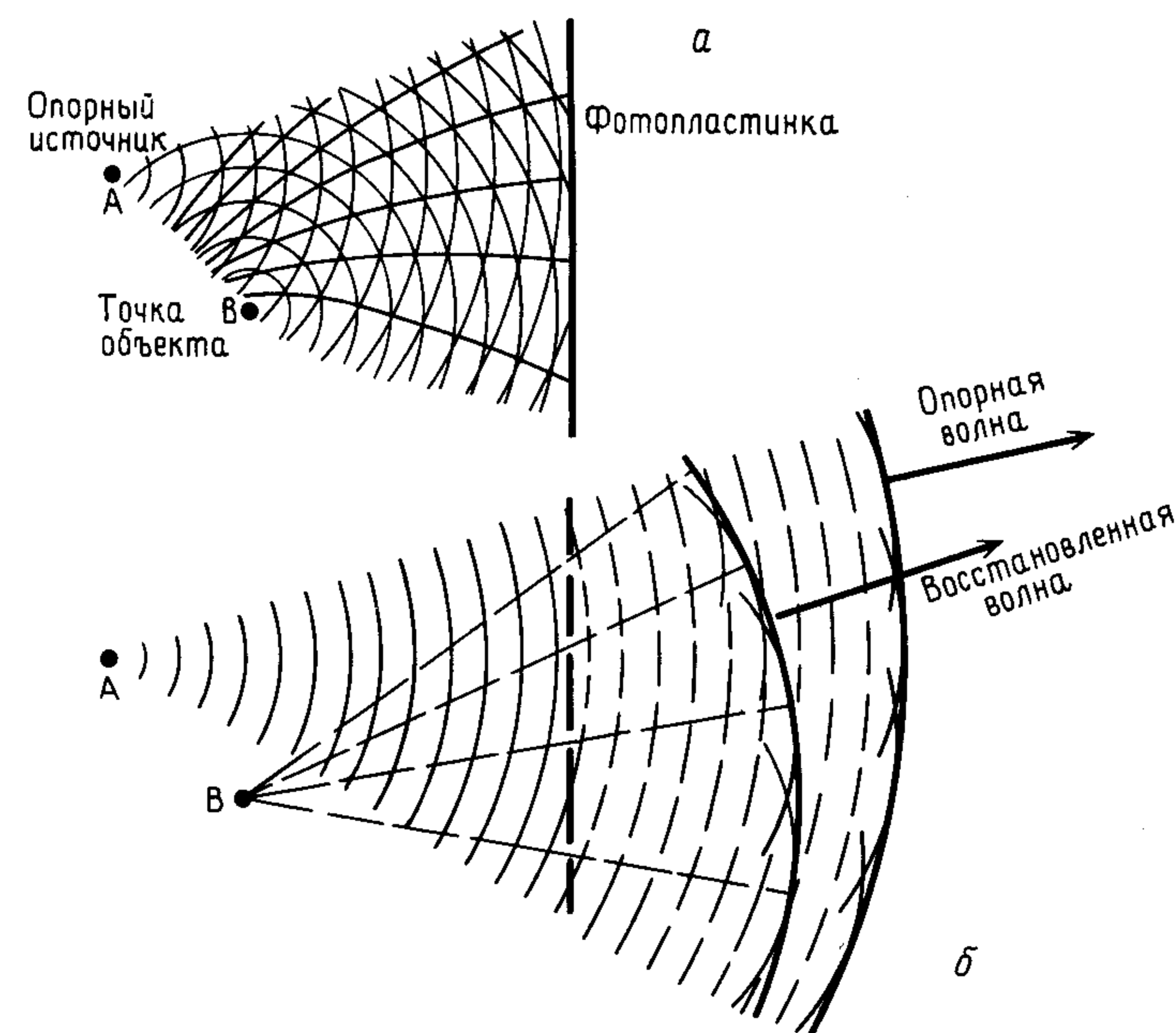


Рис. 5.11. Схематическое изображение Габором восстановления волнового фронта. а – регистрация; б – восстановление.

то он сможет закончить формирование изображения с помощью оптической системы, которая в то же время может быть предназначена для коррекции aberrаций в электронной оптике. Испытанный способ регистрации фаз заключался в замене их эффектами интерференции. Обдумывая способ осуществления этой операции в данном конкретном примере, Габор учитывал успешное использование Цернике «когерентного фона» для исследования aberrаций линз способом, который позволял выявить фазы и интенсивности. При этом на первой стадии появлялась возможность использовать электронный микроскоп для получения интерференционной картины между излучением, рассеянным объектом и неиспользованным «фоном» (т.е. нерассеянной частью освещающего пучка электронов).

Исходные эксперименты Габора выполнялись исключительно с видимым светом и основывались на методе, от которого вскоре отказались, и потому мы не будем описывать его здесь подробно. Принцип метода в противоположность применяемой на практике схеме показан на рис. 5.11. Рассматривается отдельная точка объекта В. Волновые фронты, рассеянные В, интерферируют с волновыми фронтами когерентного фона или показанного на рисунке когерентного опорного источника А. Предположим, что интерференционная картина регистрируется фотоспособом в виде позитива, который после обработки будет

пропускать свет только в положениях интерференционных максимумов. Если осветить теперь опорным источником саму фотографию, то прошедшие волновые фронты остаются с ним в фазе. Однако, как и интерференционные максимумы, они также совпадают с волновыми фронтами, которые расходятся от В, если В будет по-прежнему неподвижным. Следовательно, волновые фронты от В, несмотря на их отсутствие, могут быть названы «восстановленными». Фотография проявляет себя в виде «голограммы», как назвал ее Габор (от греческого слова *holos*, означающего «целое»), указывая на то, что фотография содержит всю информацию об объекте как амплитуду, так и фазу.

Однако, подтверждая основную идею, результаты Габора ухудшались недостаточной длиной когерентности (только 0,1 мм) света от использованной ртутной лампы высокого давления и низким уровнем освещенности, получаемой после введения малой диафрагмы (диаметром 3 мкм) для обеспечения достаточной пространственной когерентности. Из-за этой и ряда других причин применение указанного метода в электронной микроскопии было неудачным. Как отмечал Габор, голография была надолго заброшена. Возрождение наступило после работы Е. Н. Лейта и Дж. Упатника [33]. Успех их был обусловлен тем, что они обнаружили сходство процесса восстановления волнового фронта Габора с принципами теоретической работы, выполненной Лейтом с сотрудниками по локатору бокового обзора. В них предусматривалось применение «бокового» опорного освещения, что обеспечивало существенное улучшение характеристик [34, 35]. Затем в этих разработках были использованы незадолго до того созданные лазеры и сочетание этих двух достижений привело к более универсальному и улучшенному процессу голографии.

Одна из геометрических схем для записи «голограммы Лейта–Упатника» показана на рис. 5.12, а. Когерентное излучение с плоским волновым фронтом рассеивается (в этом примере) прозрачным объектом, и голограмма образуется при условии, что рассеянный пучок интерферирует с опорным лучом, создаваемым из подходящим образом отведенной неиспользованной части падающего излучения. Чтобы понять, каким образом голограмма, полученная при фоторегистрации этой интерференционной картины, несет информацию об амплитуде и фазе, необходимо для «восстановления» изображения объекта, достаточно рассмотреть процесс лишь в одном измерении (ось x на рис. 5.12, а).

Пусть A_x и δ_x являются соответственно амплитудой и фазой освещенности на оси x , которые поступают от объекта. Амплитуда A_0 опорной освещенности на голограмме может быть принята постоянной, и ее фазу в любой точке на оси x положим равной $(2\pi/\lambda)x \sin \theta = 2\pi\alpha x$, где θ – угол падения опорного луча и $\alpha = (\sin \theta)/\lambda$. Общая амплитуда R_x как функция от x может быть поэтому выражена как

$$R_x = A_0 \exp(i2\pi\alpha x) + A_x \exp(i\delta) \quad (5.10)$$

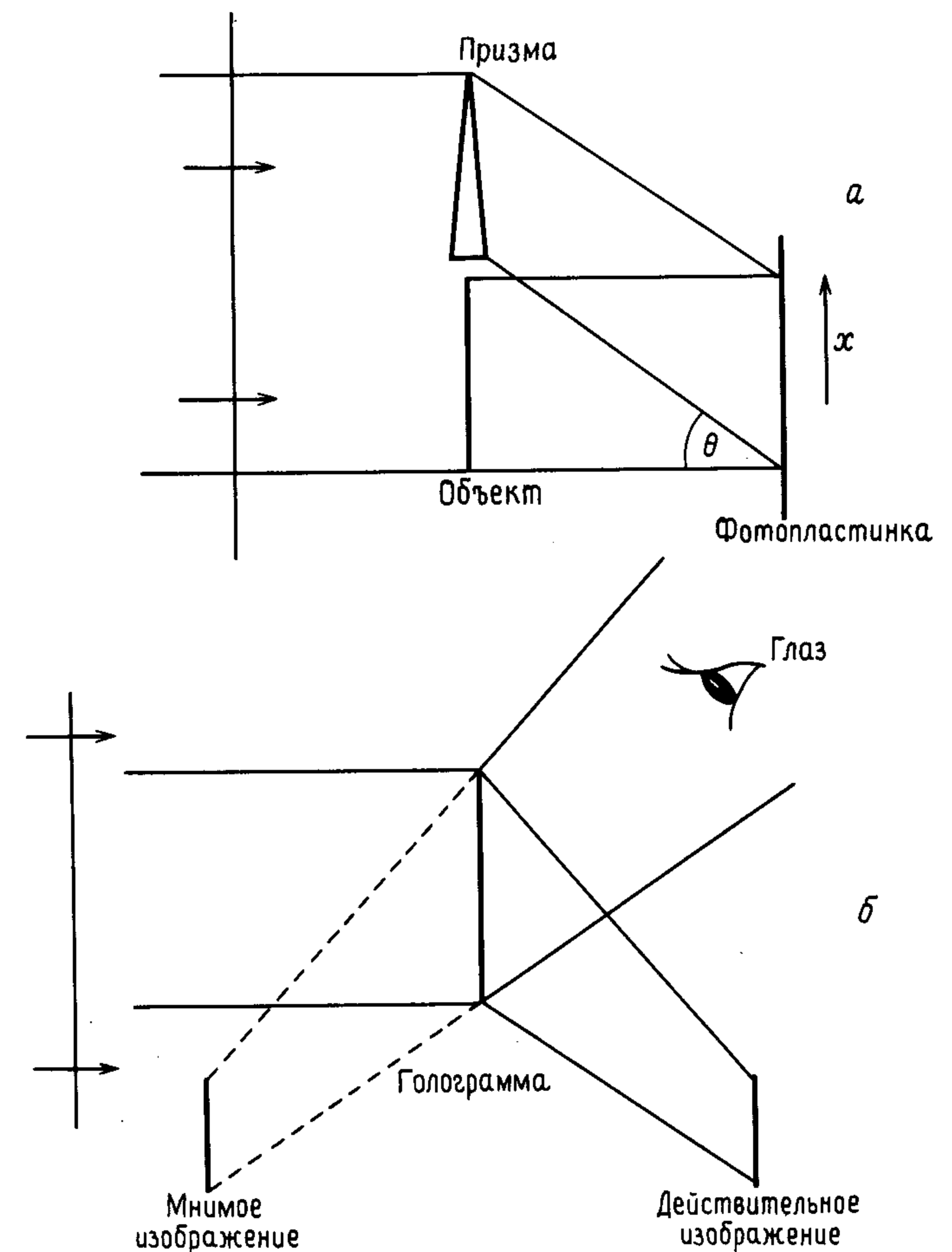


Рис. 5.12. Метод голографии Лейта–Упатника. а – регистрация (как обычно, в схемах этого типа изображена призма, хотя на практике для отклонения опорного пучка обычно используют зеркало); б – восстановление.

Интенсивность I_x в точке x равна

$$I_x = |R_x|^2 = A_0^2 + A_x^2 + A_0 A_x \exp(-i2\pi\alpha x - \delta_x) + A_0 A_x \exp(i2\pi\alpha x - \delta_x) \quad (5.11)$$

$$= A_0^2 + A_x^2 + 2A_0 A_x \cos(2\pi\alpha x - \delta_x). \quad (5.12)$$

Слагаемое с косинусом представляет систему полос, записанную в голограмме. Она несет (отсюда и «несущие полосы») всю информацию о системе волнового фронта объекта, поскольку амплитуда полос $A_0 A_x$

пропорциональна $A_x(A_0 - \text{постоянная})$ и положение полос (расстояние $1/\alpha$) по отношению к x определяется δ_x .

Используемая для восстановления голограмма может быть освещена параллельным когерентным пучком света, падающим по нормали (рис. 5.12, б). При подходящем процессе фотообработки амплитудный коэффициент пропускания голограммы линейно связан с I_x в уравнениях (5.11 и 5.12). Следовательно, амплитуда $(A_x)_{\text{восст}}$ восстановленного волнового фронта определяется непосредственно правыми частями этих уравнений, если мы пренебрежем множителем, который не влияет на интерпретацию. Используя уравнение (5.11), получим

$$(A_x)_{\text{восст}} = A_0^2 + A_x^2 + A_0 A_x \exp(-i(2\pi\alpha x - \delta_x)) + A_0 A_x \exp(i(2\pi\alpha x - \delta_x)). \quad (5.13)$$

Первые два члена представляют освещенность в направлении восстанавливающего пучка. В основном она определяется A_0^2 , поскольку A_x^2 бывает обычно много меньше, чем A_0^2 . Третий и четвертый члены могут быть переписаны соответственно в виде

$$A_0 \exp(-i2\pi\alpha x) A_x \exp(i\delta_x) + A_0 \exp(i2\pi\alpha x) A_x \exp(-i\delta_x).$$

В первом из них $A_x \exp(i\delta_x)$ является комплексной амплитудой распределения волнового фронта, который первоначально был испущен объектом при формировании голограммы. Сам по себе он должен обеспечить построение (мнимого) изображения объекта в его исходном положении. Однако умножение на $A_0 \exp(-i2\pi\alpha x)$ вызывает фазовый сдвиг, эквивалентный вращению, вызывая необходимость просмотра голограммы в направлении $-\theta$ (рис. 5.12, б). Во втором члене $A_x \exp(-i\delta_x)$ соответствует (действительному) изображению, комплексно сопряженному объекту, и умножение на $A_0 \exp(i2\pi\alpha x)$ означает, что он наблюдается с направления θ . Будучи комплексно сопряженным объекту (обратите внимание на знак минус в экспоненте), это последнее изображение обращено таким образом, что становится видимым изнутри (специалисты называют его *псевдоскопическим*).

Следует отметить, что в приведенном выше простом анализе восстановления использовался нормально падающий свет, а не дубликат исходного опорного пучка. Если учитывать толщину регистрирующей эмульсии, то условия становятся более критичными.

В настоящее время обычно используются другие схемы для создания голограмм непрозрачных и прозрачных объектов, для трехмерного цветного изображения и для различных применений в СВЧ-технике, акустике, «некогерентной фотографии», неразрушающих испытаниях, исследованиях движения, хранения информации и т.д. Их описание можно найти в большинстве современных учебников по физике. Многие из полезных свойств голограмм не связаны, однако, с этими усовершенствованиями в технике построения изображения, так что описывать их

здесь нет необходимости, и мы ограничимся лишь рассмотрением некоторых аспектов фурье-анализа применительно к голографии в разд. 5.5.1 (3).

5.5. ОПТИЧЕСКАЯ ОБРАБОТКА

5.5.1. Когерентная обработка

При описании двухступенчатого процесса формирования когерентного оптического изображения было показано, что качество изображения определяется информацией, получаемой от объекта. В частности, оно зависит от того, насколько верно составляющий его спектр пространственных частот воспроизводит спектр объекта. Мы также видели, что спектр пространственных частот доступен в плоскости дифракции. Рис. 5.10 представляет пример, иллюстрирующий ухудшение качества изображения при удалении высоких частот, которое осуществлялось путем простого размещения в плоскости дифракции апертуры (фильтра низких частот), ограничивающей вклад в изображение частот выше определенного значения.

Это был самый простой пример когерентной *пространственной фильтрации*, который теперь необходимо рассмотреть в более общем плане, поскольку имеются другие типы и методы фильтрации в частотной области, диапазон применимости которой чрезвычайно широк.

Основная схема воспроизводится в упрощенном виде на рис. 5.13. Объект представлен в виде транспаранта, который передает комплексную амплитуду $f(x)$; для упрощения математических выкладок взято только одно измерение. Транспарант освещается падающим на него перпендикулярно светом, например от системы лазер-расширитель пучка, создающей когерентное освещение плоской волной. На дифракционной плоскости мы имеем преобразование Фурье от транспаранта

$$F(u) = \int f(x) \exp(-2\pi i u x) dx. \quad (5.14)$$

Изображение формируется как очередная стадия преобразования Фурье.

Одним из вариантов пространственной фильтрации изображения в дифракционной плоскости и примере (рис. 5.10), приведенном в начале этого раздела, является *амплитудная фильтрация* (в частности, *заграждающий фильтр*, именуемый так оттого, что он либо полностью пропускает, либо блокирует любую данную частоту).

Применяется также *комплексная фильтрация*, т.е. фильтрация по фазе или по комбинации амплитуды и фазы.

Таким образом, *оптическая фильтрация* является особой формой оптической обработки, при которой пространственное фурье-преобразование объекта обеспечивает заранее определенное воздействие на изображение. Основы этого метода были заложены в теории формирования изображения, разработанной Аббе.

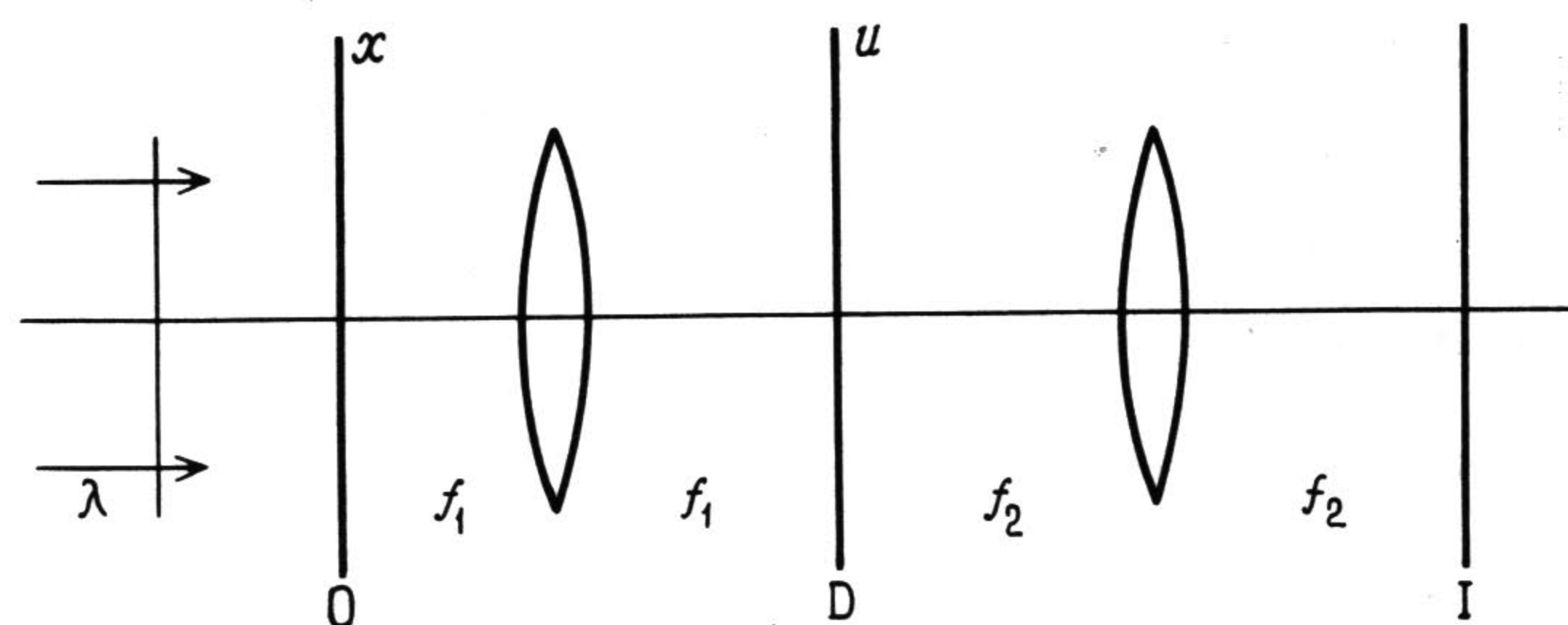


Рис. 5.13. Основная схема когерентной оптической фильтрации (O – плоскость объекта; D – дифракционная плоскость; I – плоскость изображения).

Приведенные ниже примеры трех типов фильтрации помогут получить представление относительно диапазона их применения.

1) Амплитудные фильтры

Ценным применением фильтрации этого типа стал разработанный в 1964 г. А. Клюгом и Дж. Е. Бергером [31] анализ электронных микрофотографий больших биологических молекул на основе оптической дифракции с использованием электронной микрофотографии в качестве дифракционной маски.

Многие биологические образцы содержат представляющие интерес компоненты с регулярной структурой, например системы белковых клеточных комплексов в вирусах. Однако зачастую в электронных микрофотографиях детали просматриваются недостаточно четко, например, из-за наличия в образцах материала с неупорядоченной структурой. Процедуру, разработанную Клюгом и его сотрудниками, можно описать, ссылаясь на рис. 5.13. Электронная микрофотография располагается в объектной плоскости, а ее картина в плоскости оптической дифракции фотографируется. Регулярную дифракционную картину создают только те компоненты объекта, которые имеют регулярную структуру. Другие дифракционные эффекты, вызванные наличием неупорядоченного материала, можно распознать и исключить с помощью соответствующей маски. Если электронная микрофотография по-прежнему находится в объектной плоскости, а в плоскости дифракции размещается фильтрующая маска, позволяющая использовать для обработки лишь выбранные дифрагированные лучи, то получается изображение, на котором гораздо яснее выделяется картина периодической компоненты объекта.

Другой пример иллюстрирует рис. 5.14. На рис. 5.14, а показана электронная микрофотография трубчатой структуры, обнаруженной у одного из видов бактериофага. Внешний слой трубки состоит из белковых молекул с молекулярным весом примерно 50 000, расположенным по спирали подобно виткам пружины. Для получения этой картины

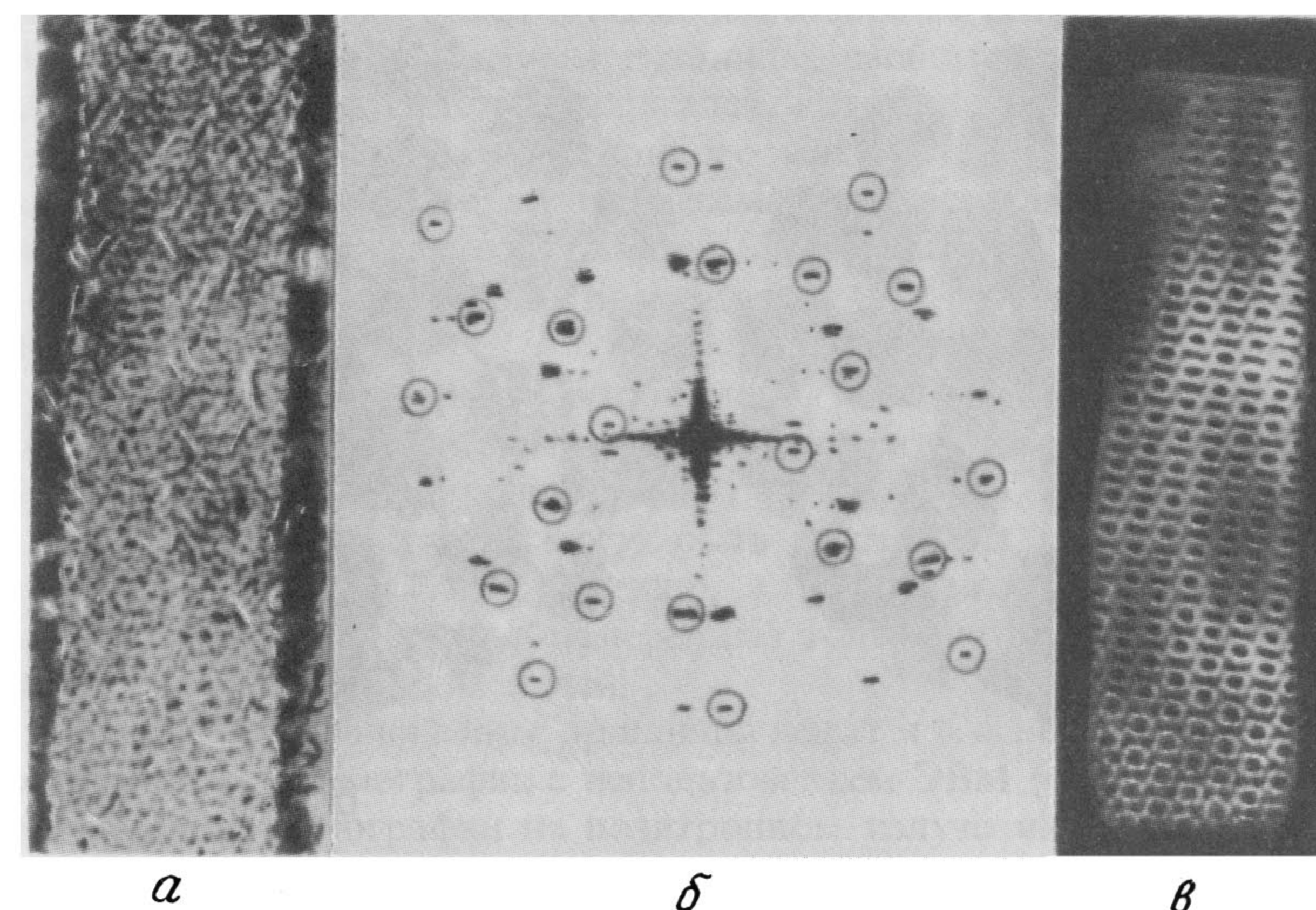


Рис. 5.14. Оптическая фильтрация в электронной микроскопии. а – электронная микрофотография негативно окрашенной уплощенной трубчатой структуры в бактериофаге ($\times 200\,000$); б – картина оптической дифракции от а (кружками обведены дифракционные пятна, связанные со структурой на одной стороне уплощенной трубки. Маска сделана так, что только эти пятна формируют изображение); в – результирующее отфильтрованное изображение, показывающее спиральную структуру из молекул, объединенных в шестигранники [14].

образец сначала был помещен в среду с высокой электронной плотностью (муравьинокислый уранил). Такая обработка является обычной процедурой. Среда рассеивает электроны более эффективно, чем белок, поэтому, заполняя отверстия и щели на поверхности, она подчеркивает ее структуру. Однако трубчатые образцы, подобные указанному, выравниваются при подготовке к электронной микроскопии, и детали изображений задней и передней сторон трубки взаимно налагаются. На рис. 5.14, б и в можно увидеть, каким образом эти изображения могут быть разделены. Оптическая дифракционная картина от образца а, полученная, как описано выше, показана на рис. б. Кружками обведены дифракционные пятна, которые можно отождествить со структурой одного определенного (например, заднего) слоя трубки. Картина, показанная на рис. 5.14, в, была получена с маской, которая пропускает на плоскость изображения лишь составляющую от этих пятен. Размещение отдельных молекул (в шестигранниках) можно теперь наблюдать более отчетливо. Разумеется, эту работу следует выполнять с особой осторожностью, так как неправильный выбор дифракционных пятен создает ложное изображение.

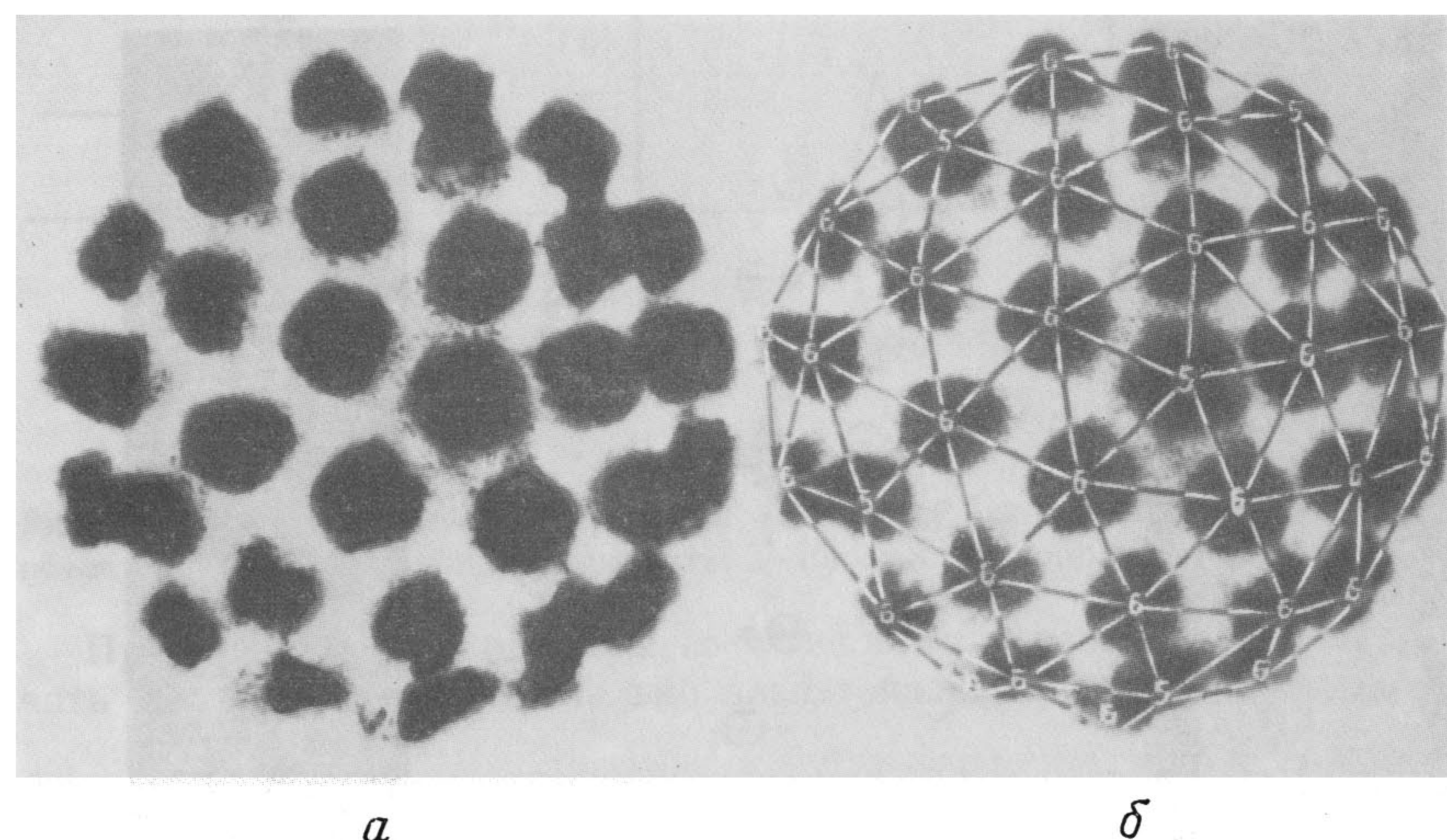


Рис. 5.15. *a* – трехмерное восстановленное изображение вируса человеческой бородавки (диаметр примерно 500 Å); *б* – показана внутренняя двадцатигранная решетчатая поверхность, помечены вершины с 5 и 6 складками ($1 \text{ Å} = 10^{-10} \text{ м}$) [32].

Те же принципы используются теперь для обработки электронных микрофотографий на ЭВМ. Фотографическое изображение преобразуется в цифровую форму путем измерения оптической плотности, а для выполнения преобразований Фурье и фильтрации используется ЭВМ. При применении этого метода сохраняется информация как о фазах, так и о интенсивностях, и в общем он обеспечивает более широкие возможности, чем оптический метод для коррекции aberrаций и других нежелательных эффектов, связанных с электронной оптикой микроскопа. Если рассматривать электронную микрофотографию как апертурную функцию, хотя и очень сложную, то ее преобразование Фурье может быть рассчитано полностью с учетом всех деталей распределения амплитуды и фазы. (Поскольку фазы «не видны», то, как правило, в оптической обработке о них с легкостью забывают, хотя в приложениях, подобных описанному, они могут быть столь же или даже более важными, чем амплитуда. Однако, как мы уже отмечали, оптические методы имеют свои преимущества.)

Следующим большим успехом Ключа и его сотрудников была разработка метода получения трехмерного изображения больших биологических молекул с высоким разрешением [13]. Рис. 5.15 иллюстрирует пример «восстановления» трехмерного изображения вируса человеческой бородавки. Рассчитанное путем комбинаций членов Фурье для ряда обычных двухмерных фотографий в электронном микроскопе, полученных при разных углах наклона образца, трехмерное изображение

оказывается гораздо более четким, чем любое из исходных двухмерных. Это достигнуто исключением путаницы, свойственной отдельным изображениям, из-за наложения деталей, которые лежат в пределах глубины фокуса объектива электронного микроскопа. До этих исследований, глубина фокуса объектива рассматривалась как фундаментальное ограничение, присущее электронной микроскопии.

За эту работу Ключ был награжден в 1982 г. Нобелевской премией в области химии.

Получение трехмерного изображения этим методом представляет собой развитие концептуальной основы рентгеновской кристаллографии. Главное различие здесь состоит в том, что в электронной микроскопии амплитуды и фазы могут быть рассчитаны, как отмечалось выше, по имеющемуся изображению, в то время как в рентгеновской кристаллографии доступна измерениям только интенсивность дифрагированных рентгеновских лучей.

Во многом аналогичные принципы лежат и в основе новых методов: рентгеновской томографии с использованием ЭВМ («рентгеновское сканирование»), томографии на позитронном излучении и томографии на основе ядерного магнитного резонанса (см. описание в [48]).

Отмечая важную роль фазы и необходимость сохранения информации о ней, мы отчасти затронули содержание следующих разделов. Другие приложения амплитудной фильтрации включают удаление следов развертки из телевизионных кадров, удаление точечной структуры полутонных изображений (двумерный эквивалент удаления развертки растра), исключение аддитивного шума и управление контрастностью изображения. Последняя операция из перечисленных производится путем изменения баланса вкладов высоких и низких частот в изображение.

2) Фазовые фильтры

Мы упоминали в предыдущем подразделе о важности, а зачастую и необходимости данных, касающихся фазы. Приложение фильтрации, в котором фазы играют ведущую роль, относится к области оптической микроскопии.

Срезы биологических материалов, проверяемые в оптической микроскопии, нередко бывают в значительной степени или даже полностью прозрачными. Говоря без преувеличения, это создает трудность в наблюдении их структуры, если не использован какой-либо особый прием. Для таких материалов характерно изменение от точки к точке показателя преломления и, следовательно, их оптической толщины. Так как это явление создает только фазовое различие между светом, прошедшим через разные области, оно не оказывает влияния на амплитуду прошедшего света и, следовательно, не воспринимается человеческим глазом. По очевидным причинам материалы этого типа называют *фазовыми объектами* в противоположность *амплитудным объектам*, которыми мы в основном интересовались.

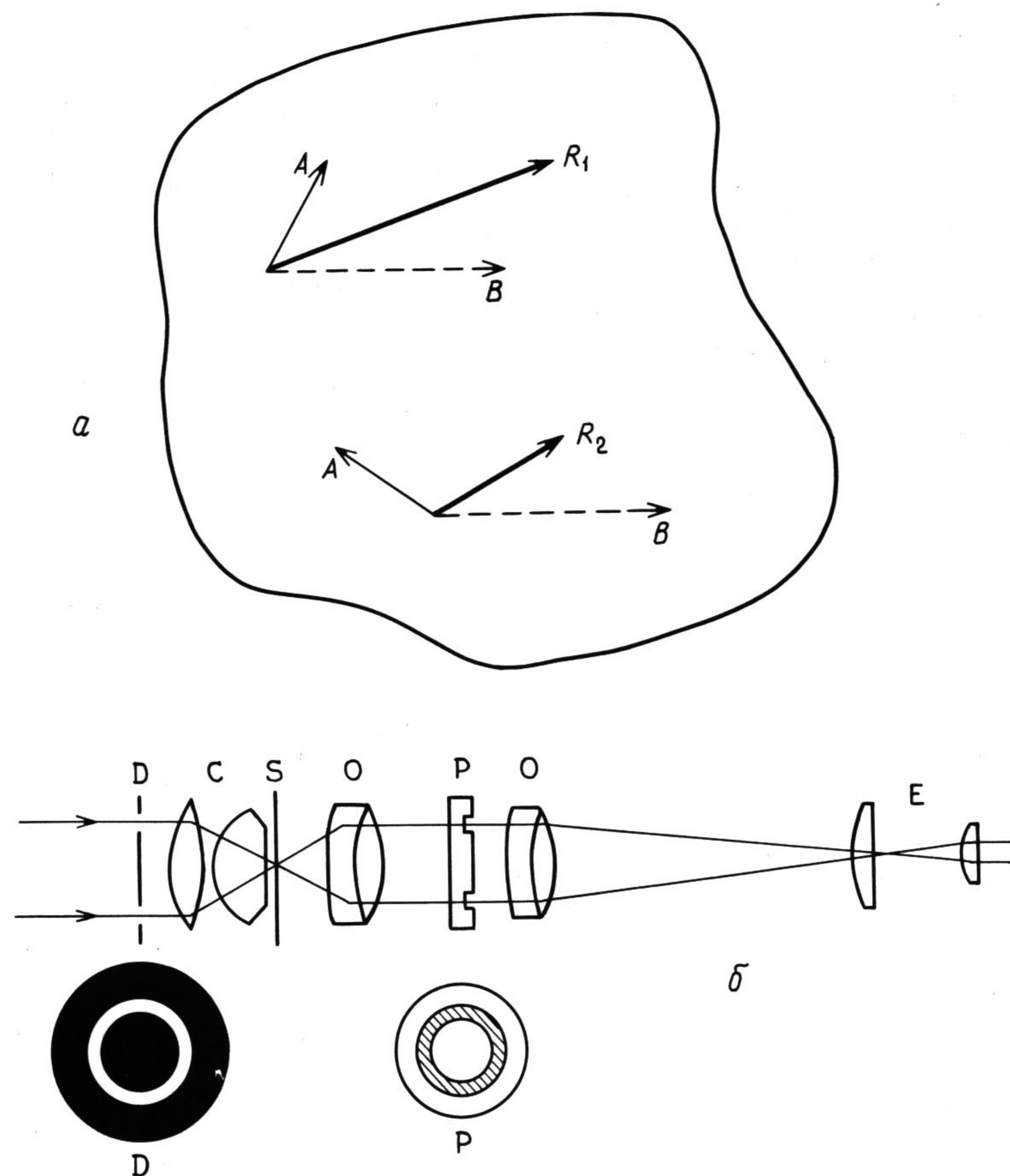


Рис. 5.16. Фазоконтрастная микроскопия (обозначения: D – диафрагма; C – конденсор; S – образец; O – детали объектива; P – фазовая пластинка; E – окуляр) [2].

Классическим способом выявления структуры материалов, подобных описанным, является использование селективного химического окрашивания, и значение этого метода нельзя недооценивать. Однако существует и более прямой метод определения различия между областями с различным показателем преломления (и, следовательно, разного биологического состава). Таким методом является *фазоконтрастная микроскопия*, предложенная первоначально Цернике приблизительно в 1935 г. [65], за что он был удостоен Нобелевской премии по физике 1953 г.

Рисунок 5.16, а иллюстрирует в упрощенном виде основные принципы этого метода. Векторы, обозначенные A, представляют освещенность в двух точках на плоскости изображения фазового объекта. Они

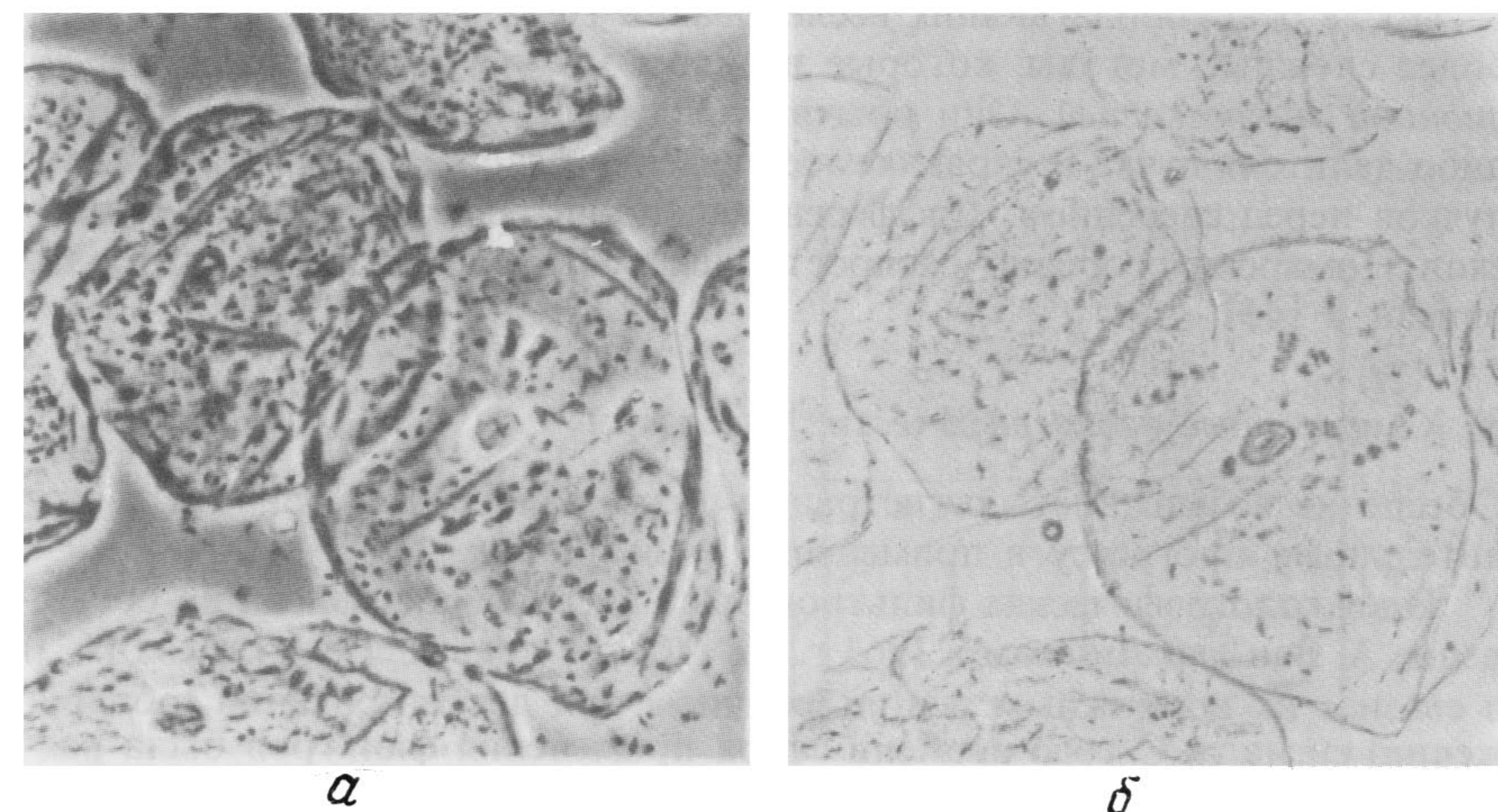


Рис. 5.17. Пример фазоконтрастной микроскопии. Оптические микрофотографии чешуйчатых клеток из слизистой рта (размер поля зрения примерно 100 мкм). а – фазовый контраст; б – обычный случай яркого поля (с разрешения д-ра М. Стюарда, Манчестерский университет).

имеют одинаковую амплитуду A , но их фазы отличаются из-за различия в оптической толщине объекта. Если к полю добавляется освещенность, имеющая постоянную амплитуду и фазу, как представлено пунктирным вектором, обозначенным B , то появляется интерференция и результирующие векторы R различаются как по амплитуде, так и по фазе. Следовательно, различие оптических длин пути оказывается «видимым». Одна из простейших установок, позволяющая достигнуть этого, показана схематически на рис. 5.16, б. Свет фокусируется на образце S линзами конденсора C. Диафрагма D в форме кольцевой щели расположена в фокальной плоскости конденсора, и свет, который не дифрагирует на образце, формирует ее изображение в виде светлого кольца на задней фокальной плоскости (P) объектива O. Эффект, показанный на рис. 5.16, а, может быть достигнут, следовательно, введением сдвига по фазе между этой «опорной» освещенностью и светом, дифрагирующим на образце. (Небольшая часть *дифрагированного* света, которая прошла через кольцо на задней фокальной плоскости, просто ослабляет эффект.) Фазовый сдвиг вводится посредством *фазовой пластинки* в P, состоящей из оптически параллельной стеклянной пластины, на которую нанесен тонкий слой диэлектрика либо по зоне кольца, либо всюду, исключая зону кольца, создавая соответственно «отрицательный» или «положительный» фазовый контраст. Пример, показанный на рис. 5.17, представляет чешуйчатые клетки из слизистой оболочки рта. Детали строения клетки видны более четко с фазовым контрастом (а), чем при нормальном поле яркой освещенности (б).

После основополагающих исследований Цернике были разработаны более сложные методы, которые в целом можно назвать *интерференционной микроскопией*. Они обычно предусматривают наличие какого-либо делителя пучка и управляемое введение фазового сдвига в один из пучков перед сведением. Дифференциально-интерференционный микроскоп Номарски особенно распространен в настоящее время и хорошо описан в [45].

3) Комплексные (голографические) фильтры

Обеспечив возможность регистрации фазы, голография привела к значительному прогрессу в применении фильтрации.

Идея голографических фильтров была впервые поставлена на обсуждение А. Ван дер Люттом в 1963 г. [61] (более доступна его статья [62]) в связи с их возможным использованием при детектировании (обнаружении) сигнала. С того времени сфера применения фильтров была расширена и включает коррекцию («выравнивание») aberrаций в оптических системах, компенсацию движения изображения и т.д. Прежде чем рассматривать применение, нам необходимо ознакомиться с основными принципами работы фильтра этого типа.

Голограммы того типа, который мы уже рассмотрели, называют френелевскими или фраунгоферовскими в зависимости от расположения регистрирующей фотопластинки в ближней или дальней зоне. Такие голограммы используются в «безлинзовой» фотографии трехмерных объектов и во многих других приложениях голографии. Однако голограмма может быть построена в любой плоскости и при схеме, показанной на рис. 5.18, она регистрируется в фокальной плоскости первой линзы, плоскости преобразования Фурье от объекта. Эта *голограмма преобразования Фурье* (или *обобщенная голограмма*) обладает свойствами, которые имеют особую ценность в определенных типах фильтрации. (Существует также «безлинзовая» геометрическая схема регистрации голограмм такого типа).

Ограничивая наше рассмотрение для простоты одним измерением, осью x , показанной на рис. 5.18, предположим, что f_x является комплексной амплитудой, пропускаемой при когерентном освещении объекта. В фокальной плоскости первой линзы мы имеем преобразование Фурье от f_x , которое можно обозначить F_X . Прописная буква F представляет собой обычное преобразование от f , а индекс X обозначает координату в фурье-плоскости. (Чтобы мы могли выявить зависимость по возможности яснее, пренебрежем разными факторами, такими, как длина волны освещения, фокусные расстояния линз и т.д., которые воздействуют только на геометрические стороны явления.)

Предположим, что $a_0 \exp(-i2\pi\alpha X)$ является комплексной амплитудой опорного освещения, поступающего на плоскость преобразования, где $\alpha = (\sin \theta)/\lambda$, как в разд. 5.4.

Распределение интенсивности I_X в плоскости преобразования за-

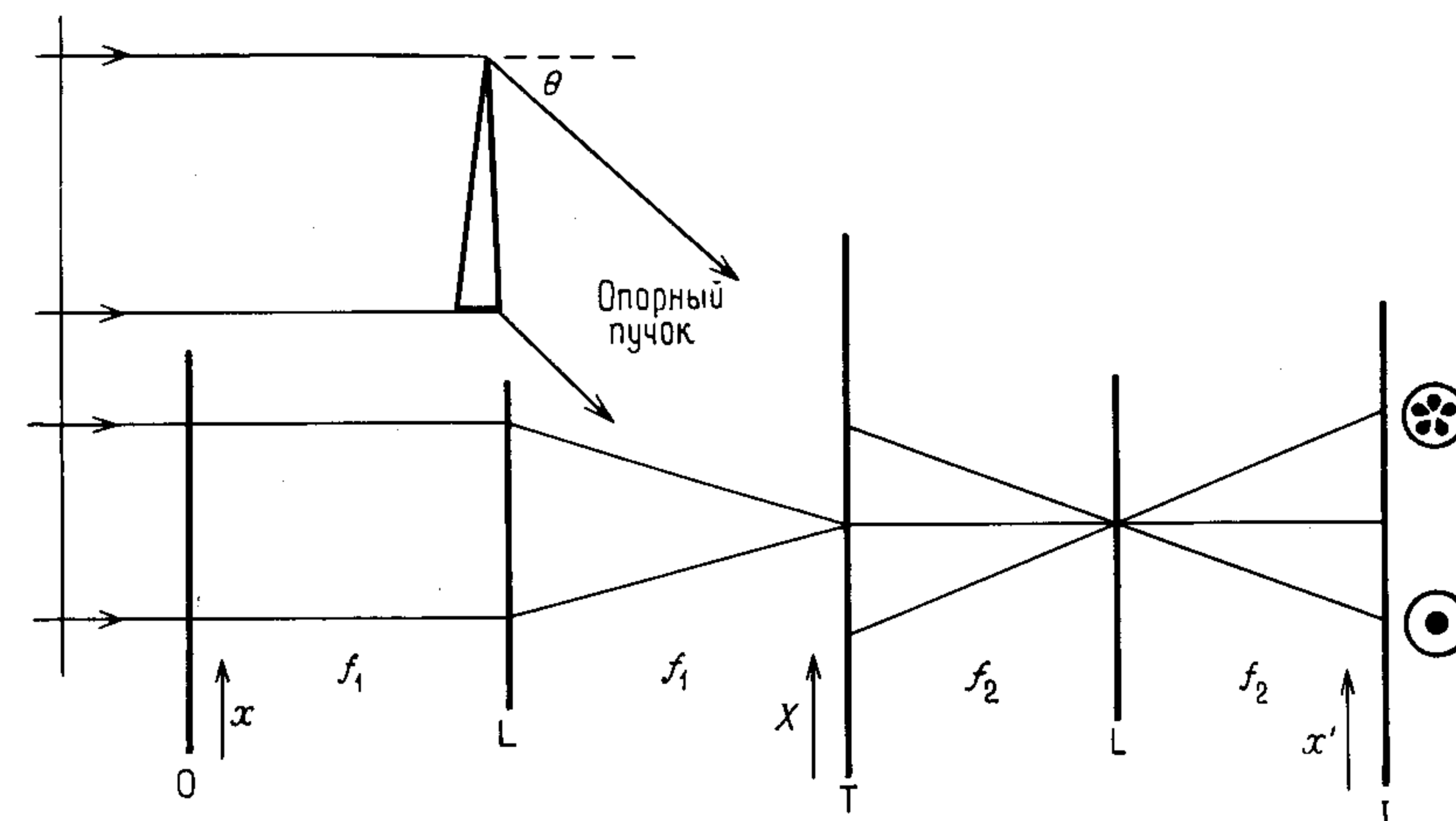


Рис. 5.18. Когерентная оптическая обработка. В оптическом корреляторе для распознавания образа используется согласованный пространственный фильтр (О – объект; L – линза; Т – преобразование Фурье; I – изображение).

дается уравнением

$$I_X = |a_0 \exp(-2\pi i \alpha X) + F_X|^2$$

$$= a_0 + |F_X|^2 + a_0 F_X \exp(2\pi i \alpha X) + a_0 F_X^* \exp(-2\pi i \alpha X). \quad (5.15)$$

Это распределение регистрируется фотоспособом в виде транспаранта, как описано ранее, так, что голограмма имеет амплитудное пропускание, пропорциональное этому распределению интенсивности.

Рассмотрим теперь, что произойдет, если эта голограмма расположена в плоскости преобразования, опорное освещение выключено и используется другой объект, пропускание которого характеризуется другой картиной комплексной амплитуды, скажем a_x . В фурье-плоскости последняя превращается в A_X . После «фильтрации» в фурье-плоскости с помощью голограммы, комплексное амплитудное пропускание которой задано уравнением (5.15), амплитуда на выходе плоскости преобразования, назовем ее U_X , будет равна

$$U_X = a_0^2 A_X + |F_X|^2 A_X + a_0 F_X A_X \exp(2\pi i \alpha X) + a_0 F_X^* A_X \exp(-2\pi i \alpha X). \quad (5.16)$$

В фокальной плоскости второй линзы мы имеем изображение, скажем, комплексной амплитуды $\psi_{x'}$, которое является преобразованием Фурье от U_X . Использование теоремы о свертке и преобразовании δ -

функции (уравнение 4.20) дает

$$\begin{aligned} \psi_{x'} = & a_0^2 a_{x'} + (f_{x'} * f_{x'} * a_{x'}) + \\ & + [(a_0 f_{x'}) * a_{x'} * \delta(x' + \alpha)] + \\ & + [(a_0 f_{x'}^*) * a_{x'} * \delta(x' - \alpha)] \end{aligned} \quad (5.17)$$

[относительно отрицательного знака в нижнем индексе $f_{x'}^*$ см. подстрочное примечание к уравнению (4.51)].

Первые два члена относятся к освещенности вблизи оси в плоскости изображения и обычно не используются. Третий член является сверткой $f_{x'}$ и $a_{x'}$ и смещен δ -функцией в положение $x' = -\alpha$; мы будем называть его *свернутым изображением*. Последний член также является сверткой $f_{x'}^*$ и $a_{x'}$ и смещен в положение $x' = \alpha$. Это также и кросс-корреляция $f_{x'}^*$ и $a_{x'}$ (см. уравнения 4.46), и называют ее *кросс-корреляционным изображением*.

Отметим, что сделанные выше выводы свидетельствуют о соответствии перемножения в плоскости Фурье (плоскости дифракции) свертке в плоскости изображения. Голограмма от f_x , о которой речь шла выше, известна как *фильтр Ван дер Люгта*, и следующие ниже примеры будут посвящены иллюстрации использования двух смещенных изображений.

а) *Кросс-корреляционное изображение. Пример:* распознавание образа (распознавание символа). Предположим, что мы хотим отождествить определенную «картину» в серии транспарантов «объекта». Голографический фурье-фильтр конкретной картины, такой, как упоминавшаяся выше f_x , создается описанным нами способом и размещается в плоскости преобразования. Идентичность ожидаемых картин объекта исследуется поочередно в объектной плоскости на рис. 5.18. Когда очередь доходит до искомой картины, в кросс-корреляционном изображении появляется яркое центральное пятно. Все картины, которые не соответствуют f_x , должны давать нулевую освещенность в этом изображении (см. разд. 4.7). Лишь в случае некоторого сходства картин получается слабый «паразитный отклик».

Фильтры этого типа, представляющие собой преобразование Фурье искомой картины, известны под общим названием *согласованные фильтры*.

б) *Свернутое изображение. Пример:* устранение размытия, компенсация aberrаций. Как показано в предыдущих главах, влияние aberrаций или какой-либо другой особенности системы формирования изображения, которая приводит к искажению изображения, можно рассматривать как свертку идеального изображения a_x с функцией размытия g_x (ср. разд. 5.3.3(3)). Практически полученное оптическое изображение должно быть сверткой $a_x * g_x$. Для устранения влияния g_x можно, следовательно, умножить (см. теорему свертки) преобразование этой свертки ($A_X G_X$)

в стадии дифракции на $1/G_X$. Для создания фильтра, выполняющего эту операцию, свет от точечного источника пропускается через систему линз с aberrациями и интерферирует с коллимированным опорным пучком. Однако здесь есть свои трудности, и за подробностями читатель должен обратиться к специальной литературе.

Недостатком обработки при когерентных условиях является то, что здесь мы имеем дело с комплексными амплитудами, не позволяющими использовать в качестве входных устройств систем обработки телевизионные дисплеи или дисплеи на светодиодах. Более того, когерентная обработка имеет тенденцию к помехам из-за «шума», создаваемого пылью, царапинами и другими дефектами оптических компонентов. К тому же выходной сигнал таких систем представляется обычно в форме интенсивностей, так что данные о фазе теряются.

5.5.2. Некогерентная обработка

Из-за присущих когерентной обработке недостатков, упомянутых в предыдущем разделе, в последнее десятилетие появилась тенденция к некогерентной обработке. Нет необходимости говорить о том, что она тоже имеет свои недостатки, и выбор должен зависеть от характера проблем. При некогерентной обработке входные и выходные величины обычно должны быть положительными и действительными. Правда, обработка биполярных и комплексных величин может быть выполнена путем мультиплексирования, но тогда обычно требуется некая гибридная форма системы.

Можно выделить два главных вида некогерентной обработки, и ниже приведено их краткое описание, но только в простейшем виде, потому что они выходят за рамки содержания данной книги. Первый метод снова имеет отношение к использованию дифракции, а второй основан исключительно на геометрических принципах, в которых дифракция не играет активной роли.

1) Обработка на основе дифракции

Пример: распознавание образов по корреляции энергетического спектра. На схеме, представленной на рис. 5.19, транспарант T_1 освещается плоской волной квазимонохроматического света. Комплексная амплитуда последнего может быть обозначена f_x , а фурье-преобразование от нее в фокальной плоскости линзы L_1 — соответственно F_X . Здесь для простоты снова используется одномерное представление. Рассеиватель в плоскости преобразования разрушает когерентность и создает некогерентное распределение интенсивности, в сущности подобное «самосвещающемуся», которое пропорционально величине $|F_X|^2$, являющейся энергетическим спектром f_x (ср. разд. 4.7.1).

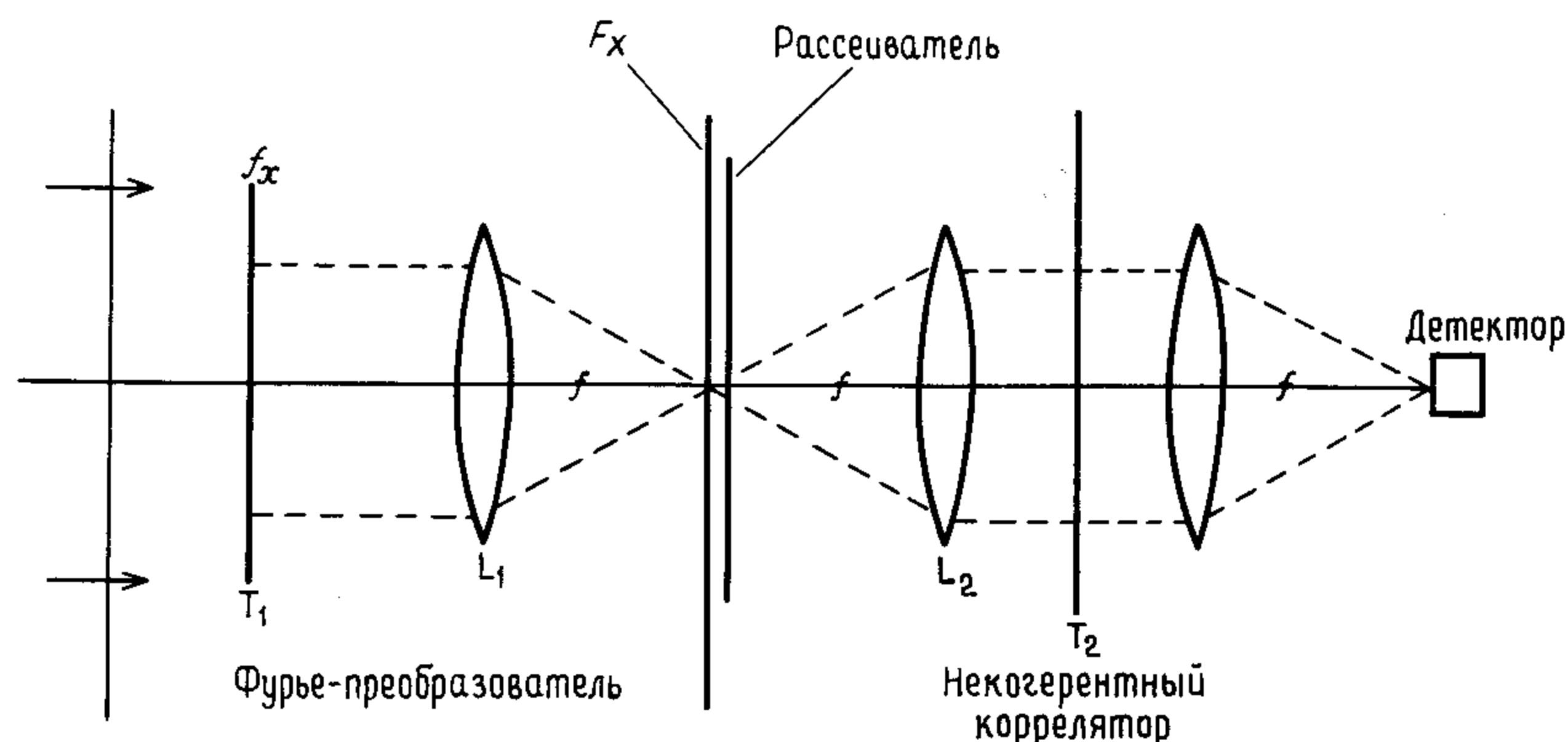


Рис. 5.19. Неогерентная оптическая обработка. Коррелятор энергетического спектра.

Другой транспарант расположен в позиции T_2 . Выходом системы является, следовательно, свертка $|F_X|^2$ с энергетическим спектром транспаранта в T_2 . Поскольку T_1 и T_2 обычно действительны, то любая инверсия свертки эквивалентна корреляции. Представленный схематически на рисунке прибор известен, как *коррелятор энергетического спектра*.

Для распознавания буквенно-цифровых образов применяют, например, метод, основанный на том, что кросс-корреляционный выход системы (измеряемый обычно фотодетектором, расположенным на оси) значительно отличается от нуля, если два транспаранта T_1 и T_2 имеют одинаковые спектры пространственных частот. Транспарант «требуемой» картины используется в положении T_2 , а неизвестные транспаранты располагают для наблюдения в T_1 . Успех зависит от того, насколько различаются разные спектры. Наряду с этим недостатком указанный метод имеет заметное преимущество перед описанным в разд. 5.5.1, состоящее в том, что измеренный кросс-корреляционный выход не зависит от положения и ориентации картин на входной плоскости, потому что их спектры ненаправленные.

Приложения этого метода включают в себя распознавание буквенно-цифровых символов при машинном переводе с одного языка на другой, поиске информации и т.д.

2) Обработка на основе геометрической оптики

Пример: распознавание образов. На рис. 5.20 свет от источника S сводится линзой L_1 и проходит через фотопластинку T_1 и затем через вторую пластинку T_2 , которая может быть смещена в своей плоскости. Линза L_2 фокусирует свет, прошедший через T_1 и T_2 на фотодетекторе, установленном в ее фокусе.

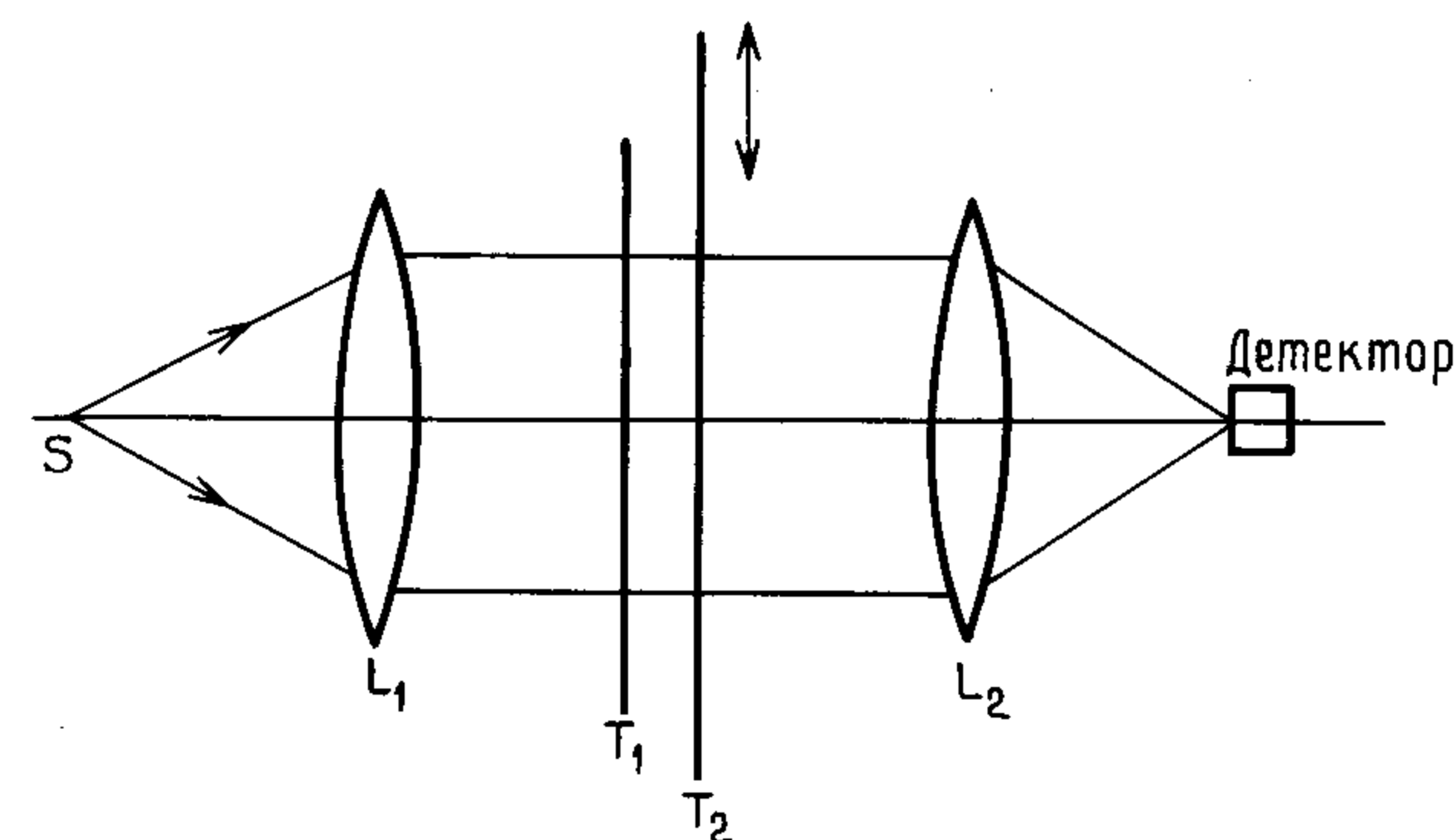


Рис. 5.20. Обработка на основе геометрической оптики. Транспаранты T_1 , T_2 расположены в плоскости (x, y) , перпендикулярной рисунку.

Пусть пропущенная интенсивность в T_1 и T_2 будет соответственно $t_1(x, y)$ и $t_2(x + u, y + v)$, где величины u, v определяют смещение T_2 относительно T_1 .

Детектор интегрирует свет, прошедший два транспаранта, и его выходной сигнал P как функция u, v , записывается уравнением

$$P(u, v) = \iint t_1(x, y) t_2(x + u, y + v) dx dy. \quad (5.18)$$

Это выражение является кросс-корреляцией $t_1(x)$ и $t_2(x)$ [ср. уравнение 4.40], и из разд. 4.7 должно быть ясно, что в общем случае сильный сигнал детектируется только в случае, если две картинки, представляющие f и g , одинаковы и если $u = v = 0$.

Следовательно, этот *оптический перемножитель и интегратор* может быть использован для распознавания образов, хотя в противоположность коррелятору энергетического спектра, обсуждавшемуся выше, он требует, чтобы картинки имели одну и ту же вращательную ориентацию.

Это краткое замечание имеет целью дать лишь самое поверхностное представление о типах обработки, которые возможны с применением неогерентных систем, основанных на геометрии изображения и им подобных. Они включают в себя пространственные и временные методы сканирования, многоканальную обработку и т.д. и проектируются на основе широкого диапазона устройств (светодиодов, матриц ПЗС и т.д.).

Статьи по практическому применению и разработкам в области электронной обработки данных можно найти в книгах, специально посвященных этой области: например, в серии, опубликованной издательством «Шпрингер» под общим названием «Вопросы прикладной физики» и «Вопросы современной физики», а также в серии «Достижения в оптике» под редакцией Е. Вольфа, опубликованной в Амстердаме (Голландия).

Интерферометрия и источники излучения

Пусть нам удалось бы сделать океан линзой и вырвать у неба разгадку его тайн, но и тогда, я думаю, многое осталось бы непознанным.

Ричард Джефрис, 1883
«Повесть моего сердца»

6.1. ВВЕДЕНИЕ

Мы переходим теперь от проблем использования излучения для исследования структуры, а также формирования и обработки изображений к вопросу изучения пространственной и спектральной структуры источников излучения.

В гл. 1 в самых общих чертах было показано, как видность интерференционных полос может поставлять информацию об источниках излучения. Для таких исследований особенно чувствительными являются два способа интерферометрии, впервые предложенных Майкельсоном.

Вначале Майкельсон использовал свой звездный интерферометр для измерения диаметров звезд, но он предвидел, как измерения видности полос могут давать информацию о распределении яркости источника, и продемонстрировал это на простых примерах.

С помощью своего спектрального интерферометра Майкельсон измерил длины волн спектральных линий, но он вновь предвидел и продемонстрировал в простых случаях его гораздо больший потенциал опять-таки в отношении использования видности полос для получения детальной информации о тонкой структуре спектра. Позже он воспользовался этим способом для эталонирования длины метра в единицах длины волны красной линии кадмия.

В 1907 г. Майкельсон был удостоен Нобелевской премии по физике за разработку этих оптических инструментов и за спектроскопические и метрологические исследования, которые он осуществил с их помощью.

В настоящее время оба способа считаются прототипами важных методов современной астрономии и спектроскопии. Положенные в их основу принципы сегодня так же важны, как и во времена Майкельсона. По этой причине они напоминаются в разд. 6.2 и 6.3, которые обеспечивают введение в разд. 6.4, посвященный современным представлениям, основанным на теории частичной когерентности.

Спектроскопия фурье-преобразования кратко описана в разд. 6.5, а в разд. 6.6 освещаются некоторые основные вопросы фурье-анализа при построении изображений в астрономии с помощью интерферометрии.

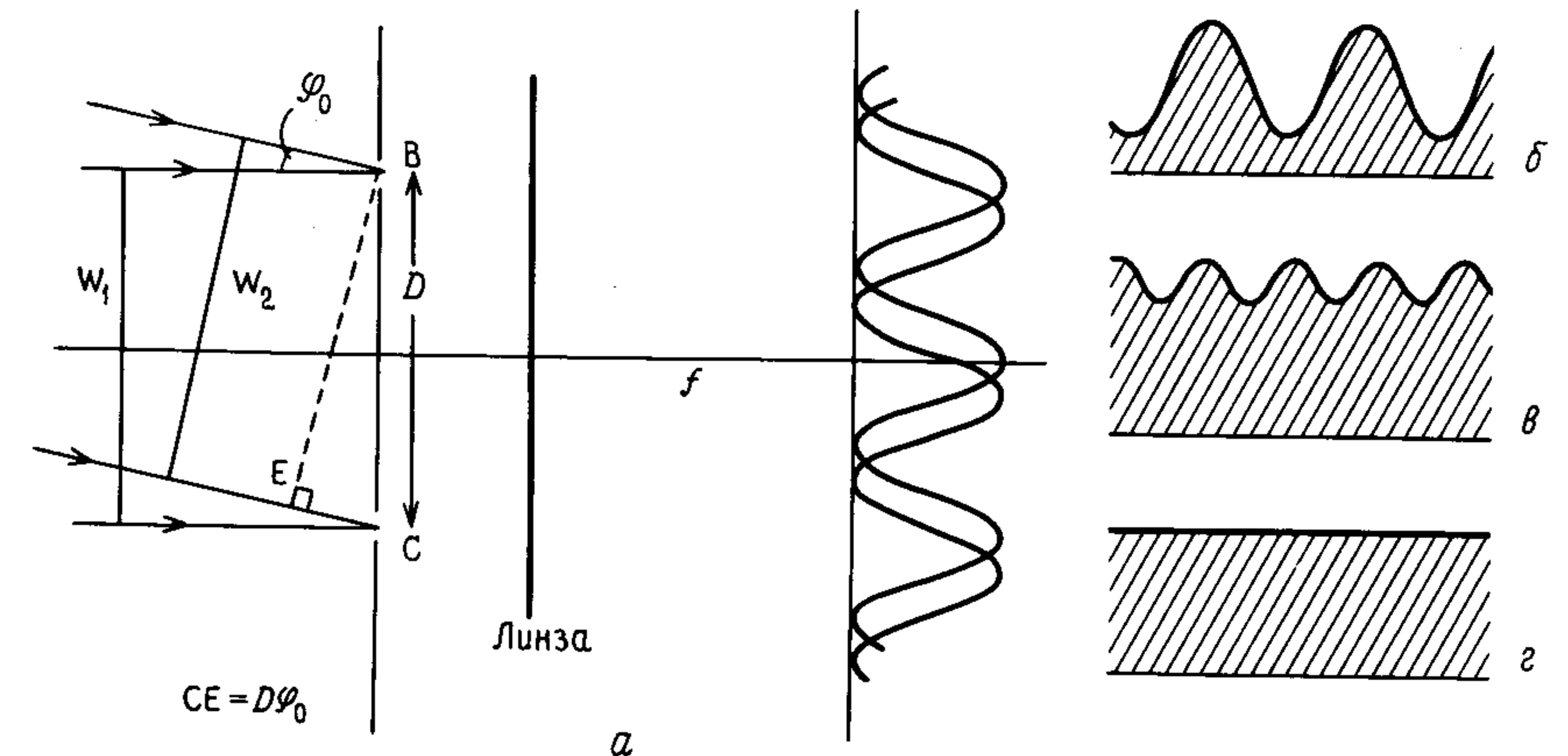


Рис. 6.1. Измерение диаметров звезд с помощью интерферометра.

6.2. ЗВЕЗДНЫЙ ИНТЕРФЕРОМЕТР МАЙКЕЛЬСОНА

6.2.1. Введение

Чтобы создать представление об использовании интерференции как непрямого способа применения телескопа для измерения угловых размеров астрономических объектов, рассмотрим рис. 6.1, а. На нем представлен апертурный экран, имеющий две щели, перпендикулярные рисунку и размещенные перед линзами телескопа (аналогичную схему нетрудно осуществить и для отражательного телескопа). Волновые фронты поступают от всех точек видимой части поверхности звезды, имеющей угловой диаметр ϕ_0 (стягиваемый ею угол с вершиной у Земли). На рисунке показаны только граничные фронты волн: W_1 , испущенный на одном краю диска, и W_2 от противоположного края. В фокальной плоскости линз образуется непрерывная система интерференционных полос типа $\cos^2$ (источник считается некогерентным) от полос, вызываемых W_1 , до полос, определяемых W_2 . Окончательным результатом является картина, показанная на рис. 6.1, б с видностью < 1 . Отметим, что расстояние между полосами остается таким же, как если бы источник был точечным, а именно $\Delta = f\lambda/D$ [уравнение (1.11)]. На практике интенсивность картины полос снижается с той и другой стороны от оси (ср. с выборкой на дифракционной картине от одиночной щели в разд. 2.4). Мы можем пренебречь этим понижением, если щели узкие и, в частности, если наблюдения, как случается на практике, ограничены центральной областью картины полос.

Теперь если D увеличивается, то расстояние между полосами уменьшается и видность полос снижается, как показано на рис. 6.1, в. При

достижении такого значения D , при котором расстояние между полосами становится в точности равным смещению полос, видность становится равной нулю, как на рис. 6.1,2. Теперь расстояние между полосами таково, что полосы от W_2 смещены в точности на одну относительно полос, связанных с W_1 . На схеме рис. 6.1,а край звезды, испускающий волновые фронты W_1 , равноудален от точек В и С. Таким образом, если бы звезда была равномерно светящимся прямоугольником, то волновые фронты W_2 от другого края должны были бы проходить в своем движении к точке С на одну длину волны больший путь, чем до точки В. При этом конкретном значении, скажем $D = D'$, мы имеем

$$D'\phi_0 = \lambda. \quad (6.01)$$

Если D увеличивается еще больше, то видность вновь не равна нулю до тех пор, пока не выполнится условие $D''\phi_0 = 2\lambda$, когда она еще раз становится нулем и так далее. (Естественно, что эти результаты не связаны со специфической ориентацией волновых фронтов W_1 по отношению к В и С, которая использовалась здесь для удобства.)

Поскольку звезда по существу представляется круглым светящимся диском, то полученный результат требует внесения поправки. Для этого можно воспользоваться методом, применявшимся при выводе разрешающей способности круглой апертуры (разд. 2.3). Конечный результат состоит в том, что первое исчезновение колец наступает не раньше, чем выполнится условие $D'\phi_0 = 1,22\lambda$. Таким образом, диаметр звезды определяется соотношением

$$\phi_0 = 1,22\lambda/D', \quad (6.02)$$

или, как грубый критерий для желтого света,

$$\phi_0 = \frac{0,15}{D'(\text{м})} \quad (\text{сек. дуги}). \quad (6.03)$$

Принцип этого метода, предложенного Физо в 1868 г. [20], был успешно использован Майкельсоном в 1890 г. [38] для измерения диаметров спутников Юпитера, для которых были получены значения около 1". При этом применялся 12-дюймовый (около 0,3 м) телескоп на Маунт-Гамильтон. В отношении звезд, однако, возник вопрос, позволяют ли атмосферные возмущения и механические вибрации телескопа проводить наблюдения, поскольку ожидалось, что их диаметры не в пример планетам должны составлять всего лишь несколько сотых долей секунды дуги. Частично эти опасения были рассеяны успешными наблюдениями интерференционных полос на 100-дюймовом (около 2,5 м) рефлекторе на Маунт-Вилсон. Однако было ясно, что для обеспечения разнеса апертур, необходимого, чтобы достичь исчезновения полос при таких малых угловых диаметрах звезд, потребовался бы весьма боль-

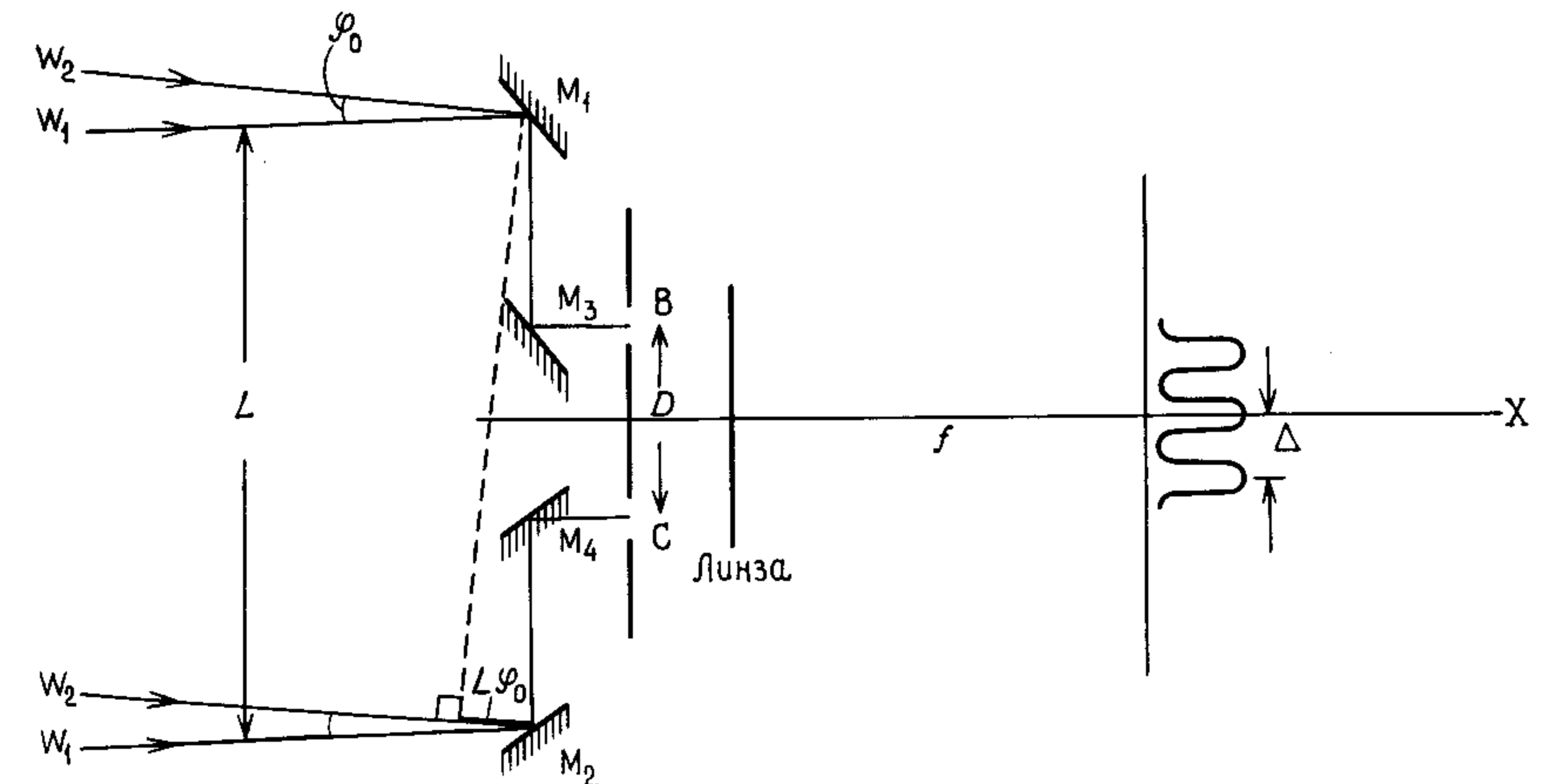


Рис. 6.2. Звездный интерферометр Майкельсона.

шой телескоп диаметром 50 футов (около 15 м) или более. Для щелей, разнесенных на такое расстояние, полосы должны располагаться столь тесно, что их невозможно наблюдать. Более того, в этом случае полученный результат мало отличался бы от разрешающей способности того же телескопа при обычном применении (см. разд. 2.3).

Для преодоления этих трудностей Майкельсон разработал звездный интерферометр, названный его именем. Он представлял собой линейную систему из четырех зеркал (рис. 6.2), каждое около 6 дюймов (около 15 см) в диаметре. Внешняя пара (M_1, M_2) действовала как «приемники» с переменным до 20 футов (около 6 м) расстоянием (которое теперь называли бы длиной базы). Зеркала внутренней пары (M_3, M_4) были зафиксированы и через две щели направляли поступающие сигналы на телескоп для получения полос обычным образом. Эта схема позволяла поддерживать расстояние между полосами равным постоянной и приемлемой величине ($\Delta = f\lambda/D$), определяемой расстоянием D между внутренними зеркалами, в то время как разность пути, соответствующая потере видности, определялась расстоянием между внешними зеркалами, достигающим величины L' , при которой

$$L'\phi_0 = 1,22\lambda,$$

т. е.

$$\phi_0 = 1,22\lambda/L'. \quad (6.04)$$

Каркас для крепления зеркал был установлен на 100-дюймовом (около 2,5 м) рефлекторе обсерватории Маунт-Вилсон, который был выбран на этот раз не из-за высокой оптической разрешающей способности, а оттого, что он мог служить в качестве жесткой опоры для несущей

щей фермы зеркал. Крайне важно обеспечить жесткую установку каждого зеркала по отношению к другим, так как путем тонкой подстройки расстояния между внешними зеркалами регистрировались очень малые разности оптических путей в системе. С помощью этого устройства Майкельсон и Физо выполнили измерения нескольких звезд-гигантов и сверхгигантов, диаметры которых не поддавались измерению обычным способом на телескопе обсерватории Маунт-Вилсон. Для Бетельгейзе, например, было получено значение $0,0047''$. [Для учета потемнения к краю диска звезды они использовали результаты наблюдений Солнца и на этом основании оценили, что действительные угловые размеры должны быть на 17% больше, чем значения, полученные с помощью уравнения (6.04).]

Новый метод по сути дал «увеличение» L/D по сравнению с первым вариантом прибора и теоретически был ограничен только L . Однако трудность обеспечения необходимой механической стабильности большой базы наряду с эффектами турбулентности в атмосфере препятствовала дальнейшему сколько-нибудь существенному прогрессу. Физо в 1930 г. построил 50-футовый вариант прибора, но он давал ненадежные результаты и в 1937 г. работа была приостановлена.

Однако Майкельсон ясно осознавал, что характер изменения видности при изменении расстояния между зеркалами содержит информацию не только о размере, но и о распределении яркости источника. Хотя в то время указанный метод не мог быть реализован практически, мы рассмотрим, каков был теоретический подход Майкельсона, учитывая важность этого метода в современной астрономии.

6.2.2. Проблемы видности полос

Вначале Майкельсон рассмотрел несколько примеров источников различной формы, но с однородным распределением яркости [37]. Он воспользовался схемой, которая в основном была аналогична показанной на рис. 6.3 (с другими обозначениями), и рассуждал следующим образом.

Некогерентный источник шириной W стягивает угол ϕ_0 с вершиной на интерферометре, показанном с перпендикулярными рисунку щелями, изменяемое расстояние между которыми равно D . (Величину D можно также рассматривать как расстояние между внешними зеркалами в предыдущем инструменте.) Рассмотрим свет от элементарной полоски шириной dx в точке S источника на расстоянии x от его оси. Пусть $y = f(x)$ — длина полоски, перпендикулярной рисунку. Поле освещенности от этой полоски выделяется в точках B и C и дифрагировавший от B и C свет образует обычную дифракционную картину. (Огибающая интерференционной картины от одиночной апертуры может быть опущена, поскольку требуется рассмотреть только самые ближайшие к оси полосы.) Для расчета интенсивности полос в направлении θ обратим

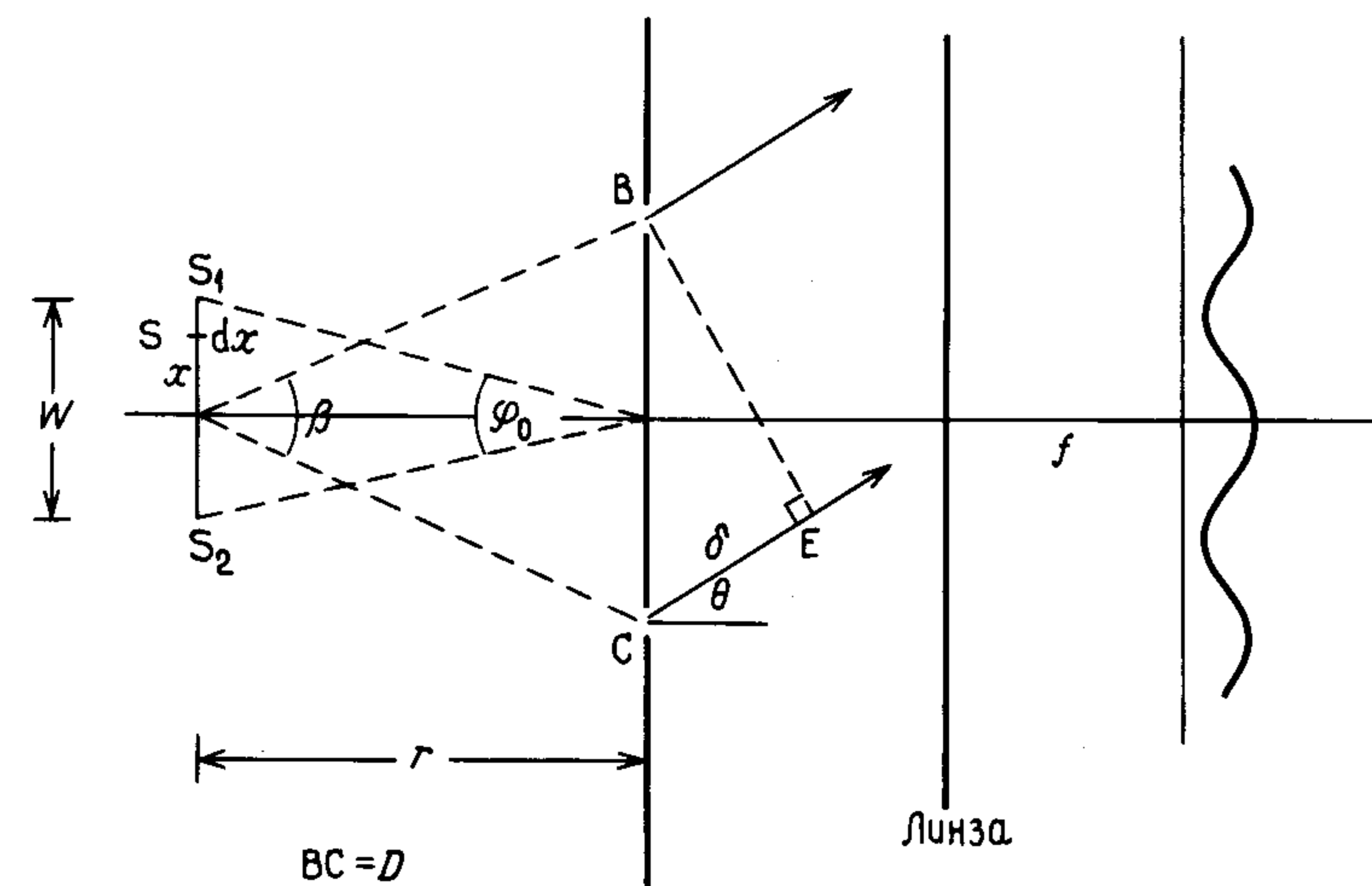


Рис. 6.3.

внимание, что разность пути между двумя расстояниями SCE и SB равна

$$= (SC + \delta) - SB =$$

$$= (SC - SB) + \delta =$$

$$= \beta x + \delta$$

(разность $SC - SB$ получается так же, как на рис. 2.7; параметр ι равен здесь $\beta/2$ и считается малым). Результат, получаемый для единичной амплитуды обычным образом с помощью векторной диаграммы, представляется в виде

$$R(\theta) = 2 \cos \frac{\pi}{\lambda} (\beta x + \delta),$$

что дает для интенсивности знакомое выражение типа $\cos^2$

$$R^2(\theta) = 4 \cos^2 \frac{\pi}{\lambda} (\beta x + \delta), \quad (6.05)$$

или

$$R^2(\theta) = 2 \left[1 + \cos \frac{2\pi}{\lambda} (\beta x + \delta) \right] \quad (6.06)$$

Ввиду некогерентности источника каждая полоска дает ряд отдельных интерференционных полос. Поскольку полосы имеют одинаковую яр-

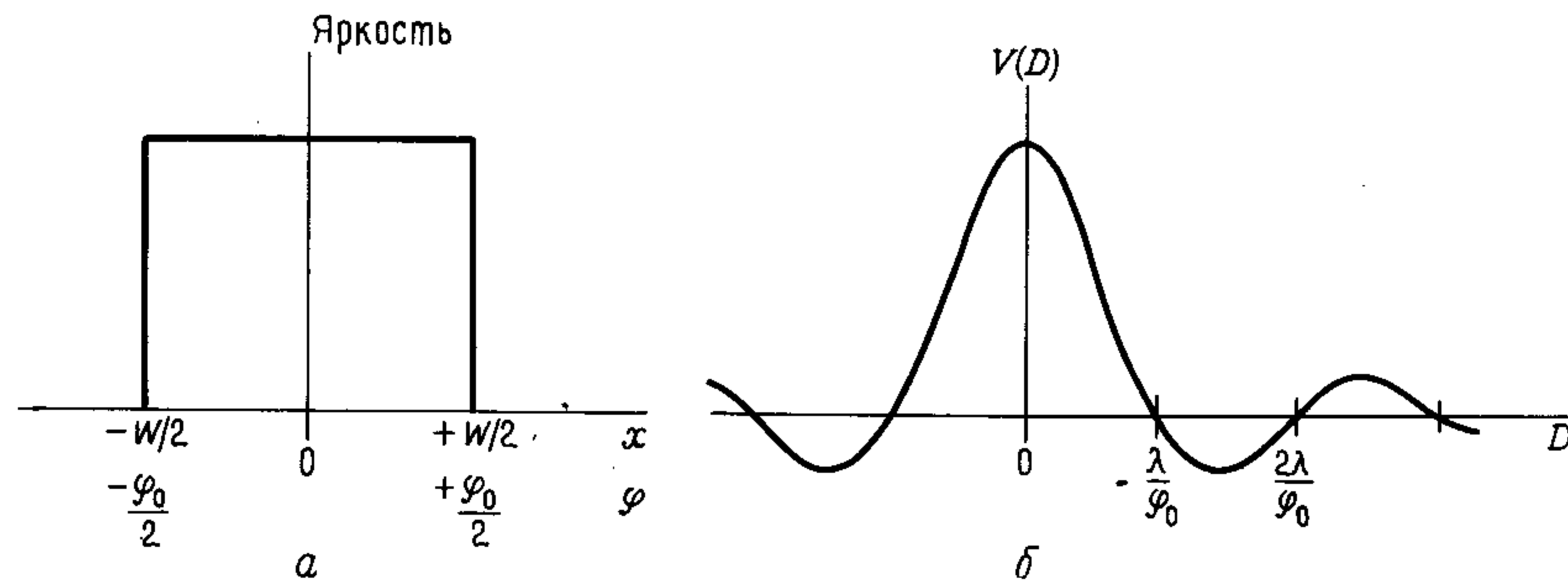


Рис. 6.4. а – распределение яркости источника; б – кривая видности.

кость, мы можем умножить приведенное выше выражение на площадь полоски $f(x)dx$, чтобы представить взвешенный вклад от этой полоски в полную картину интерференционных полос. В пределе, когда ширина полоски стремится к нулю, результирующая интенсивность $I(\theta)$ в направлении θ находится интегрированием уравнения (6.06):

$$I(\theta) = \int_{-W/2}^{+W/2} f(x) \left[1 + \cos \frac{2\pi}{\lambda} (\beta x + \delta) \right] dx, \quad (6.07)$$

где опущен множитель 2.

Для вычисления видности полос, согласно ее обычному определению [уравнение (1.05)], необходимо получить максимальное и минимальное значения интенсивности на результирующей картине. В качестве одного из примеров Майкельсон рассматривал прямоугольный источник равномерной яркости (рис. 6.4, а), для которого была задана его ширина W . Распределение источника $f(x)$, ориентированное параллельно щелям, теперь постоянно и может быть для простоты принято равным единице. В результате интегрирования уравнение (6.07) дает

$$I(\theta) = W + \frac{\lambda}{\pi\beta} \sin \frac{\pi\beta W}{\lambda} \cos \frac{2\pi\delta}{\lambda} \quad (6.08)$$

Максимальное значение I достигается при $\delta = n\lambda$, где n – ноль или целое, а минимум – при $\delta = (n \pm 1/2)\lambda$. Учитывая, что масштаб на рис. 6.3 сильно искажен, мы можем записать $r\phi_0 = W$ и $r\beta = D$, откуда $\beta W = \phi_0 D$. В результате получаются следующие выражения для $I_{\max}$ и $I_{\min}$:

$$I_{\max} = W + \frac{\lambda}{\pi\beta} \sin \frac{\pi\phi_0 D}{\lambda}, \quad (6.09)$$

$$I_{\min} = W - \frac{\lambda}{\pi\beta} \sin \frac{\pi\phi_0 D}{\lambda} \quad (6.10)$$

Следовательно, функция видности $V(D)$ определяется выражением

$$V(D) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{\sin \frac{\pi\phi_0 D}{\lambda}}{\frac{\pi\phi_0 D}{\lambda}} \quad (6.11)$$

(Обозначение $V(D)$ подчеркивает наше желание найти V как функцию D для заданной λ и фиксированного неизвестного значения ϕ_0 .)

Это выражение представляет собой sinc-функцию, известную нам из предыдущих глав как дифракционная картина, и фурье-преобразование апертурной функции, имеющей такую же «форму», как и распределение яркости на рис. 6.4, а. Сходство этих двух совершенно различных примеров не является случайным совпадением, но подробнее об этом поговорим ниже (разд. 6.4.1). Кривая видности на рис. 6.4, б спадает до нуля при $D = D' = \lambda/\phi_0$, затем повторно при $D = D'' = 2\lambda/\phi_0$ и т.д. Это ее поведение согласуется с интерпретацией, приведенной в предыдущем разделе.

Теперь нужно учесть неравномерность распределения яркости. Если положим, согласно Майкельсону [37], что полная интенсивность от элементарной полоски S на расстоянии x (рис. 6.3) равна $B(x)$, то этим исключается дополнительная трудность, связанная с формой источника. Фактически это сводит задачу к определению одномерного распределения интенсивности. Более того, если рассмотреть частный случай четной функции $B(x)$ (распределение яркости, которое симметрично относительно оси на рисунке) и если сделать замену переменной x на угловую переменную ϕ (изменяющуюся от $-\phi_0/2$ до $+\phi_0/2$), то легко получается следующее соотношение для видности:

$$V(D) = \frac{\int B(\phi) \cos(2\pi D\phi/\lambda) d\phi}{\int B(\phi) d\phi} \quad (6.12)$$

Для одномерного случая в общем виде, когда $B(\phi)$ не является четной функцией, но по-прежнему принимается то, что Майкельсон называл «полной интенсивностью полоски шириной dx ...», получается выражение следующего вида:

$$[V(D)]^2 = \frac{[\int B(\phi) \cos(2\pi D\phi/\lambda) d\phi]^2 + [\int B(\phi) \sin(2\pi D\phi/\lambda) d\phi]^2}{[\int B(\phi) d\phi]^2} \quad (6.13)$$

Майкельсон пророчески сделал вывод, что из измерений видности «можно найти распределение интенсивности излучения шаровых масс, которое должно послужить ценным ключом для нахождения распределения температуры и плотности в газовых туманностях».

Аналогичная ситуация возникла с теоретической работой Майкель-

сона, посвященной зависимости видности интерференционных полос от спектрального распределения (разд. 6.3.2). Именно в этой связи лорд Рэлей [50] обратил внимание на фурье-соотношение между видностью полос и (спектральным) распределением интенсивности. Такое соотношение применимо и здесь, и, завершая этот раздел, будет интересно и поучительно несколько глубже рассмотреть вопрос фурье-преобразования с современной точки зрения.

Числитель в уравнении (6.12) представляет собой косинус-преобразование Фурье от $B(\phi)$, а знаменатель выполняет роль простого масштабного коэффициента. Нетрудно заметить, что для каждого значения длины базы D видность дает информацию об одном конкретном фурье-компоненте распределения яркости. Это легко выясняется с помощью теоремы о свертке (разд. 4.5). Выразив наблюдаемую интерференционную картину при данном D в виде свертки $B(\phi)$ с инструментальным откликом, мы получаем из теоремы о свертке, что фурье-преобразование этой свертки является произведением отдельных преобразований. Но преобразование инструментального отклика представляет собой набор полос вида $\cos^2$, у которых имеется единственная пространственная частота, определяемая значением D . Поэтому оказывается, что преобразование от наблюдаемой дифракционной картины лепестков при данном D содержит информацию лишь об одной гармонике в распределении яркости источника.

Связь между видностью полос и когерентностью рассматривается в разд. 6.4, но сначала мы обсудим проявление видности лепестков как функции спектрального, а не пространственного распределения источника излучения, воспользовавшись на этот раз описанием спектрального интерферометра Майкельсона.

6.3. СПЕКТРАЛЬНЫЙ ИНТЕРФЕРОМЕТР МАЙКЕЛЬСОНА

6.3.1. Введение

В отличие от звездного интерферометра спектральный интерферометр основан на явлении интерференции при делении амплитуд (разд. 1.4). Основы его конструкции разработаны Майкельсоном в 1881 г. в связи с экспериментом по проверке возможности движения Земли относительно эфира. С этой целью он совместно с И.В. Морли (исторический опыт Майкельсона–Морли) намеревался создать прибор большого размера. Но основные схемные решения были использованы для измерения спектральных длин волн (позднее для эталонирования метра в единицах длины волны красной линии кадмия) и изучения тонкой структуры спектра. Именно эти спектроскопические приложения сохраняют свое значение и даже становятся все более важными в наши дни.

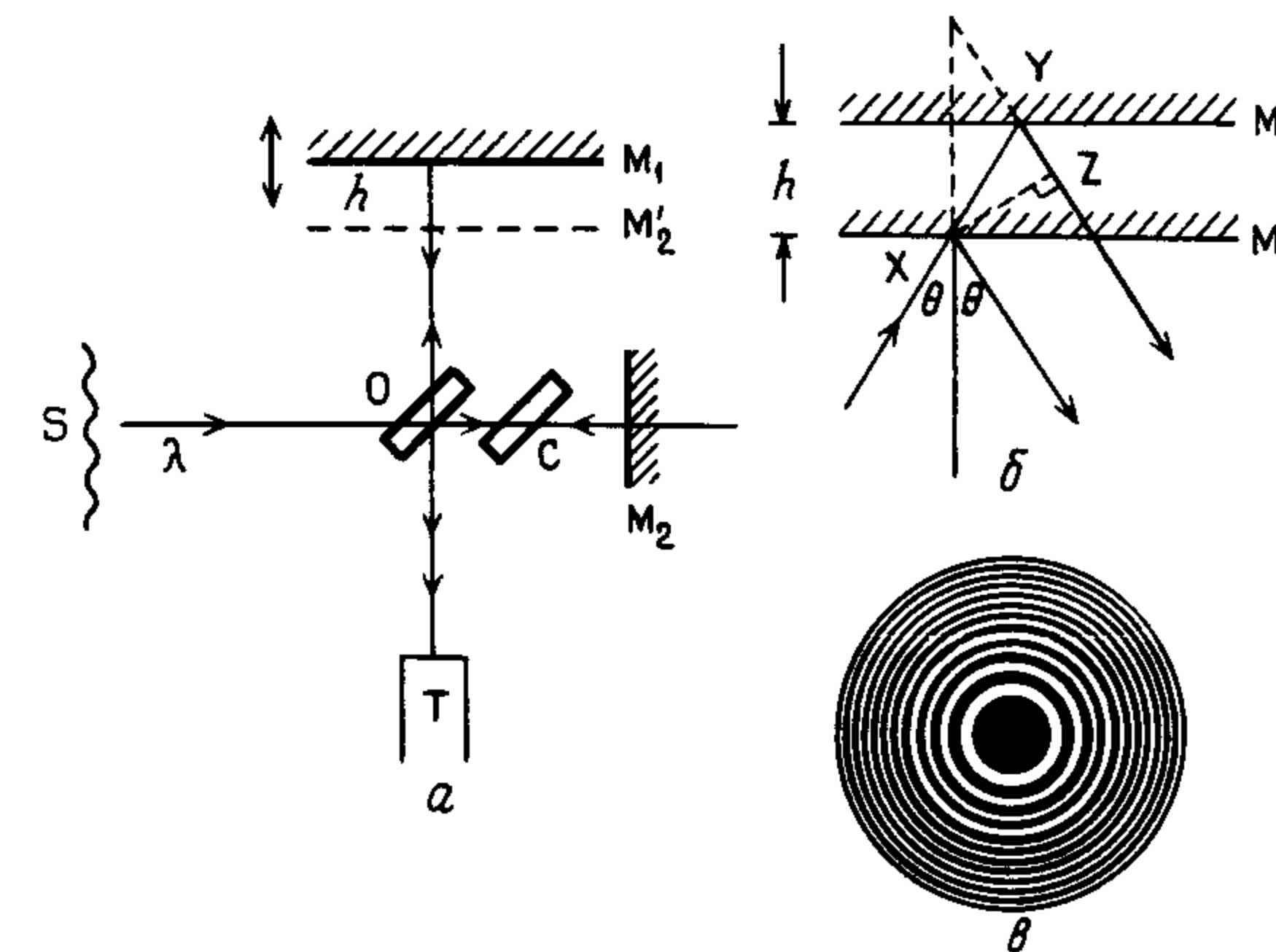


Рис. 6.5. Спектральный интерферометр Майкельсона. *a* – общий вид схемы (отражение на стеклянных пластинках O и C не показано); b – разность путей между отраженными лучами $XY + YZ = WZ = 2h \cos \theta$; δ – вид интерференционных полос для квазимонохроматического света.

На рис. 6.5, *a* схематично показано устройство одного из первых вариантов интерферометра. Свет от источника S (обычно протяженного) делится по амплитуде задней поверхностью стеклянной пластинки O с полупрозрачным серебряным покрытием на два пучка, один из которых отражается, а другой пропускается. Отраженный пучок достигает зеркала M_1 и после этого возвращается, частично проходя через O в телескоп T . В то же время другой пучок, который вначале прошел через делитель пучка, поступает на зеркало M_2 и также возвращается к O , откуда он частично отражается к телескопу. Поскольку идущий к M_1 пучок проходит через пластинку O в общей сложности три раза по сравнению с одним разом для пучка, идущего к M_2 , то обычно в точку C помещается компенсирующая пластинка той же толщины и из того же материала, что и O . В общем случае M_1 и M_2 находятся на различных расстояниях от O и между двумя лучами преднамеренно вводится разность хода (компенсирующая пластинка предназначена только для выравнивания дисперсионного хода через стекло). Соединившись вместе, два пучка создают интерференцию, результат которой определяется разностью хода между ними.

Зеркала M_1 и M_2 размещаются взаимно перпендикулярно друг другу, а делитель пучка – под углом 45° к ним. При наблюдении в телескоп изображение M_2 , формируемое O , располагается параллельно M_1 (или совпадает с ним) в M_2 . Поэтому наблюдаемая в телескоп интерференционная картина похожа на картину с одной пластинкой на рис. 1.8, хотя в представленном примере она получается при отражении от воображаемой «воздушной пластинки». Лучи от протяженного источника с длиной волны λ поступают на систему в широком диапазоне углов, и потому образуются яркие концентрические кольца (рис. 6.5, *b*) (ср. с рис. 1.8, *b*).

Окружности соответствуют направлениям с углами θ , для которых возникает усиление при сложении пар волновых цугов. Это условие определяется выражением

$$2h \cos \theta = m\lambda, \quad (6.14)$$

где m — целое или нуль, а h — расстояние между зеркалами (рис. 6.5, б). При этом предполагается, что два интерферирующих пучка одинаковым образом меняют фазу на делителе пучка. Если это условие не выполняется, то к разности фаз, связанной с разностью хода, должна быть добавлена постоянная величина. Соответственно смещаются и все интерференционные полосы.

Одно из зеркал (M_1 на рисунке) может поступательно перемещаться в указанном направлении. Изменение h приводит к расширению или сжатию картины колец; при увеличении h кольца расходятся от своего центра, как будто бы они там и возникают, а при уменьшении h они сжимаются к центру.

Выражение для радиального распределения интенсивности в направлении от центра дифракционной картины при заданных значениях h и длины волны λ легко получить путем известного нам метода векторных диаграмм. Если, например, амплитуды излучения, поступающего в телескоп двумя путями, сделаны равными, скажем, A , то результирующая интенсивность I_1 в направлении θ системы колец определяется выражением

$$I_1 = 4A^2 \cos^2(\psi/2) \quad (6.15)$$

с разностью фаз

$$\psi = \frac{2\pi}{\lambda} 2h \cos \theta \quad (6.16)$$

В результате получаем

$$I_1 = 4A^2 \cos^2\left(\frac{\pi}{\lambda} 2h \cos \theta\right) \quad (6.17)$$

Поэтому для идеального монохроматического излучения интерференционные полосы имеют вид $\cos^2$, как показано на рис. 6.6, а. Кроме того, из упомянутой выше зависимости картины колец от изменения h следует, что при постепенном уменьшении или увеличении h детектирующее устройство в любой точке картины (оно может располагаться на оси, т. е. $\theta = 0$) будет регистрировать синусоидальное изменение интенсивности. Если бы излучение было полностью монохроматическим, то цуги волн имели бы бесконечную длину (разд. 4.6) и синусоидальная картина функции видности не зависела бы от влияния разности хода, обусловленной интерферирующими пучками света. Если бы такая

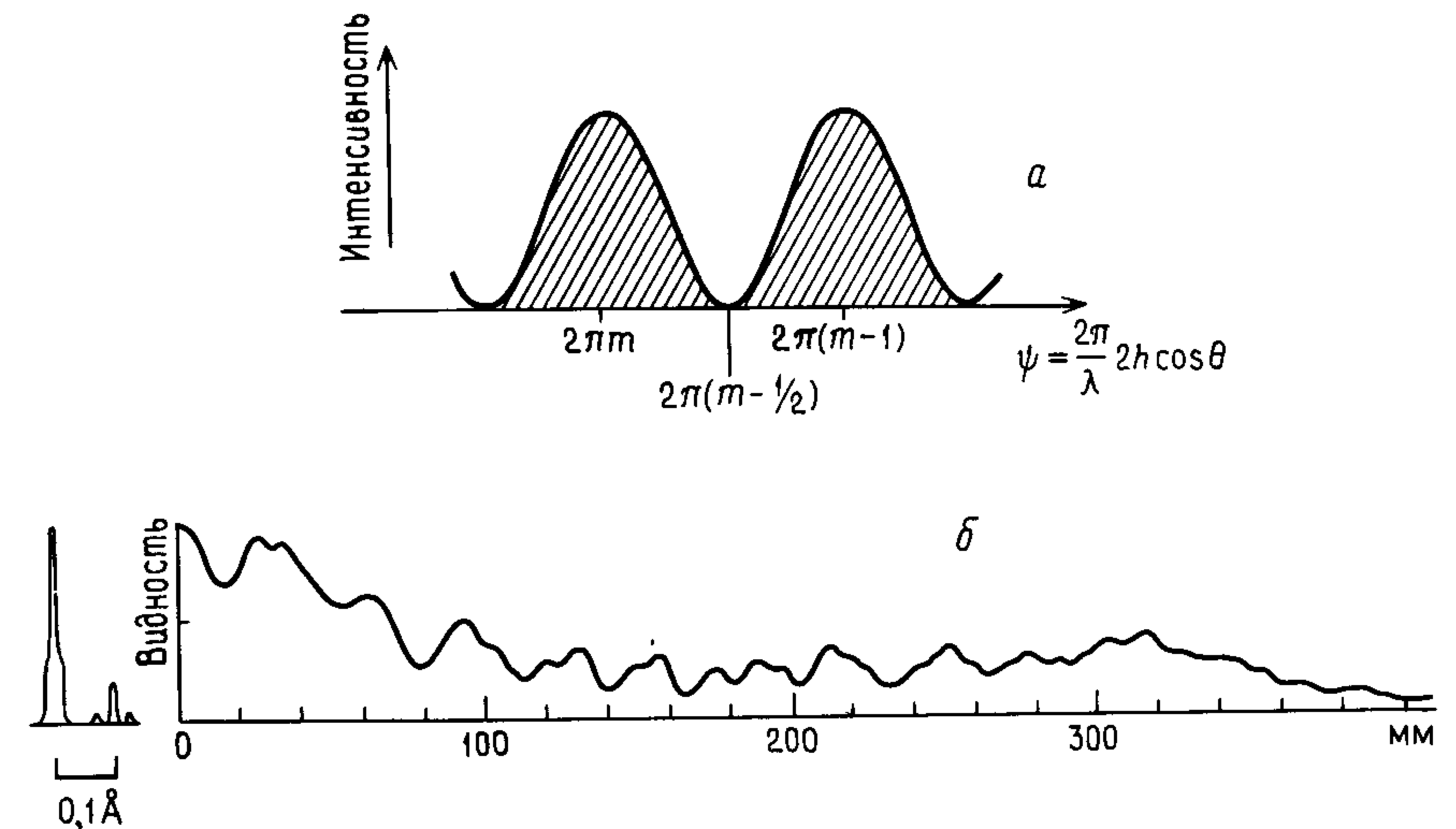


Рис. 6.6. а — интерференционные полосы типа $\cos^2$; б — результат Майкельсона для линии Hg ($\lambda = 5461 \text{ \AA}$) [43].

картина наблюдалась на самом деле, то можно было бы заключить, что излучение полностью монохроматично. Если же, наоборот, функция видности от другого источника излучения падает до нуля всякий раз, когда вносится разность хода, то можно считать, что излучение источника имеет широкий спектр, поскольку волновые цуги должны быть коротки (разд. 4.6). Именно такой количественный подход к анализу оптических спектров и является основой для использования интерферометрического метода.

Рассмотрим другой гипотетический пример. Предположим, что исследуемое излучение представляет собой комбинацию двух полностью монохроматических излучений с близкими длинами волн λ_1 и λ_2 . В этом случае изменяющаяся картина интенсивности, регистрируемая нашим детектором, сложнее, чем в приведенном выше примере монохроматического излучения на одной длине волны. Для заданного положения детектора θ найдутся такие значения h , при которых кольца двух систем почти или совершенно точно совпадают и детектор регистрирует более сильный сигнал. Это происходит, например, при h , равном такому h_1 , что

$$2h_1 \cos \theta = p\lambda_1 \approx q\lambda_2,$$

где p и q — целые. (На практике, если разность $\lambda_1 - \lambda_2$ мала, две системы колец при таком значении h будут полностью совпадать в достаточно широком интервале углов.)

Увеличение (или уменьшение) h вновь вызывает разделение двух

групп колец, хотя и незначительное, и детектор регистрирует последовательное прохождение максимума меньшей интенсивности и ненулевого минимума. Характер изменения сигнала будет определяться разностью двух длин волн, их относительной интенсивностью излучения, а также в конкретных примерах формой линии и ее тонкой структурой. Поскольку две системы колец движутся от (или в направлении) центра картины в различном темпе [см. уравнение (6.14)], то достигается значение $h = h_2$, при котором вновь происходит «совпадение» и сигнал на детекторе опять возрастает. В этом случае одна из последовательностей колец опережает другую на один целый интервал между интерференционными полосами. Это условие можно выразить в виде

$$2h_2 \cos \theta = (p + k)\lambda_1 \approx (q + k + 1)\lambda_2,$$

где k – некоторое число.

Из приведенных выше уравнений вычитанием находим, что

$$2(h_2 - h_1) \cos \theta = k\lambda_1 = (k + 1)\lambda_2.$$

Этот результат можно использовать для различных целей. Предположим, например, что λ_1 известно и наблюдения выполняются на оси или вблизи нее, т.е. $\cos \theta = 1$. Положим $2(h_2 - h_1) = H$. Тогда

$$\lambda_2 = H\lambda_1 / (H + \lambda_1)$$

Следовательно, измерение H позволяет вычислить λ_2 .

Высокая точность этого метода при определении λ_2 иллюстрируется приближенным расчетом, который показывает, что разность длин волн между линиями дублета натрия ($\lambda = 589,0$ и $589,6$ нм) требует для получения двух последовательных совпадений перемещения зеркала на расстояние около $0,3$ мм. (Мы не будем вдаваться в детали вопроса о разрешающей способности как таковой, поскольку на самом деле нас интересует метод Фурье, описанный в разд. 6.5.)

Этот способ использования интерферометра аналогичен более ранним наблюдениям Физо [19], обнаружившего в опыте с кольцами Ньютона, что кольца 500-го порядка от натриевого источника почти полностью исчезают (т.е. видность равна нулю), но снова обретают свою четкость на 1000-м порядке. Он заключил, что излучение натрия представлено дублетом, для которого кольцо 1000-го порядка на большей длине волны совпадает с кольцом 1001-го порядка на меньшей длине волны, и поэтому разность длин волн двух линий составляет около $1/1000$ от их среднего значения.

Однако Майкельсон понимал, что при таком методе анализа теряется много информации. Он сделал визуальные оценки (выраженные в количественном масштабе с помощью отдельного изошренного калибровочного эксперимента) видности интерференционных полос в зависимости от перемещения зеркала. Он осознавал, что «кривая видности» содержит очень детальную информацию о спектре источника света.

Уже в 1887 г. Майкельсоном на основе тщательных наблюдений было показано, что «красная линия водорода является очень тесным дублетом; то же самое относится к зеленой линии таллия».

Проведенное им математическое исследование этих вопросов наряду с важным вкладом, внесенным опубликованной вскоре после этого работой Рэлея, рассматриваются в следующем разделе, поскольку они служат отправной точкой для введения в основы метода преобразования Фурье.

6.3.2. Видность лепестков и спектральное распределение

Работа Майкельсона [40] в контексте нашего обсуждения является настолько, можно сказать, проясняющей и важной в применении к современным проблемам, что полезно иметь четкое представление о ее содержании. Важное вводное утверждение по сути вопроса состояло в следующем: «Общая формула для видности полос, обусловленных интерференцией двух пучков света от неоднородного источника с переменной разностью путей такая же, как и для источника конечного размера с переменным «параллаксом». («Однородный» следует понимать как монохроматичный.)

Иными словами, было высказано предположение, что взаимосвязь между видностью и спектральным распределением является такой же, как связь между видностью лепестков и пространственным распределением источника.

Майкельсон исходил из следующего выражения для интенсивности света с длиной волны λ в направлении θ картины полос, полученной на спектральном интерферометре:

$$I_1 = 4 \cos^2 \left(\frac{\pi l}{\lambda} \right),$$

где l – разность хода [$2h \cos \theta$ в уравнении (6.17)].

В спектроскопии принято пользоваться волновыми числами ($\sigma = 1/\lambda$), а не длинами волн. Тогда приведенное выше уравнение записывается как

$$I_1 = 4 \cos^2 \pi \sigma l. \quad (6.18)$$

Для полихроматического света интенсивности полос в направлении θ , обусловленные различными волновыми числами, являются аддитивными, поскольку когерентность между ними отсутствует. Если распределение интенсивности спектра $B(\sigma)$ простирается от σ_1 до σ_2 , то полная интенсивность в направлении θ определяется выражением

$$I = \int_{\sigma_1}^{\sigma_2} I_1 d\sigma = 4 \int_{\sigma_1}^{\sigma_2} B(\sigma) \cos^2 \pi \sigma l d\sigma \quad (6.19)$$

Заменим σ на $\bar{\sigma} + \sigma'$, где $\bar{\sigma}$ – среднее волновое число. Тогда замена $B(\sigma)$ на $f(\sigma')$ дает (без учета числового множителя)

$$I = \int_{-a/2}^{+a/2} f(\sigma') \cos^2 \pi l (\bar{\sigma} + \sigma') d\sigma', \quad (6.20)$$

где спектральный интервал от σ_1 до σ_2 определен теперь в пределах от $\sigma' = -a/2$ до $+a/2$ около среднего значения.

Разложение члена $\cos^2$ и использование обозначений Майкельсона

$$\left. \begin{aligned} \int f(\sigma') d\sigma' &= P, & \int f(\sigma') \cos 2\pi l \sigma' d\sigma' &= C, \\ 2\pi l \bar{\sigma} &= \theta, & \int f(\sigma') \sin 2\pi l \sigma' d\sigma' &= S \end{aligned} \right\} \quad (6.21)$$

дает

$$I = P + C \cos \theta - S \sin \theta$$

(в обозначения величин внесены некоторые изменения).

Если a достаточно мало, то можно пренебречь изменением C и S с изменением θ и максимумы и минимумы возникают, когда

$$\operatorname{tg} \theta = -S/C, \quad (6.22)$$

откуда

$$I = P \pm \sqrt{C^2 + S^2}$$

и поэтому

$$V^2 = (C^2 + S^2)/P^2 \quad (6.23)$$

Майкельсон обратил внимание, что это выражение имеет в точности такую же форму, как и выражение для видности, относящейся к пространственно протяженному источнику. Уравнение (6.13) относится к распределению яркости $B(\phi)$, в то время как здесь мы имеем распределение интенсивности в спектре $f(\sigma')$. Именно в связи с уравнением (6.21) Рэлей [50] указал на присутствие соотношения Фурье. Одновременно он обратил внимание на трудность его использования при нахождении спектрального распределения из измерений видности ввиду отсутствия информации о фазах гармоник распределения: «...сама по себе кривая видности дает не C и S , а только $C^2 + S^2$; из этого мы должны заключить, что в общем случае с кривой видности согласуется бесконечное множество структур...»

Несмотря на это ограничение, Майкельсон изыскал возможность нахождения структур некоторых простых спектральных линий, и они были с убедительностью подтверждены в последующих исследованиях.

Выражения для C и S в уравнениях (6.21) являются коэффициентами

в фурье-представлении для $f(\sigma')$, которое Майкельсон записал в виде

$$f(\sigma') = \int C \cos 2\pi l \sigma' dl + \int S \sin 2\pi l \sigma' dl$$

Принимая во внимание из (6.22) и (6.23), что $C = PV \cos \theta$ и $S = PV \sin \theta$, получаем

$$f(\sigma') = \int V \cos (2\pi l \sigma' + \theta) dl \quad (6.24)$$

Если функция $f(\sigma')$ симметрична (напомним, что σ' – это волновое число относительно среднего значения $\bar{\sigma}$), то $S = 0$ и

$$f(\sigma') = \int V \cos 2\pi l \sigma' dl,$$

что позволяет найти $f(\sigma')$ интегрированием полученной экспериментально кривой видности V (или $V(D)$, как мы обозначали ее раньше). В случае асимметричной $f(\sigma')$ требовалось знать V и θ . Последняя величина (которая является также функцией l) была названа Майкельсоном «фазовой кривой». В ряде примеров он оценил θ из сравнения с «почти однородным симметричным источником на почти такой же длине волны» (интересно отметить, что для этого Майкельсон сам разработал синусное и косинусное механическое аналоговое вычислительное устройство, которое он назвал «анализатор гармоник»). На рис. 6.6, б показан пример его исследований. Это кривая видности для ртути ($\lambda = 5461 \text{ \AA}$) и полученное из нее спектральное распределение, в котором по оценкам две главные компоненты отстоят друг от друга на $0,1 \text{ \AA}$ ($1 \text{ \AA} = 0,1 \text{ нм}$).

Для полного использования потенциальных возможностей этого метода оставалось ждать изобретения современной цифровой вычислительной машины. В приложении к измерениям длин волн двухлучевой интерферометр был заменен многолучевым методом, использованным в интерферометре Фабри – Перо. Затем в 50-х годах началось возрождение метода, послужившее основой современной фурье-спектроскопии (разд. 6.5).

6.4. ЧАСТИЧНАЯ КОГЕРЕНТНОСТЬ, КОРРЕЛЯЦИЯ И ВИДНОСТЬ

Последующее изложение понятия частичной когерентности в основном соответствует упрощенному подходу Борна и Вольфа в их книге «Принципы оптики» и базируется главным образом на работах Цернике [64] и Гопкинса [28]. В отличие от предыдущих работ оно более тесно связано с экспериментом.

На рис. 6.7 рассматривается поле освещенности, создаваемое в плоскости S источником W , который является одновременно протяженным и полихроматичным. На этом поле взяты две точки $C1$ и $C2$. Размещая в каждой из них экран с точечными отверстиями, можно экспериментально измерять видность интерференционных полос, созданных све-

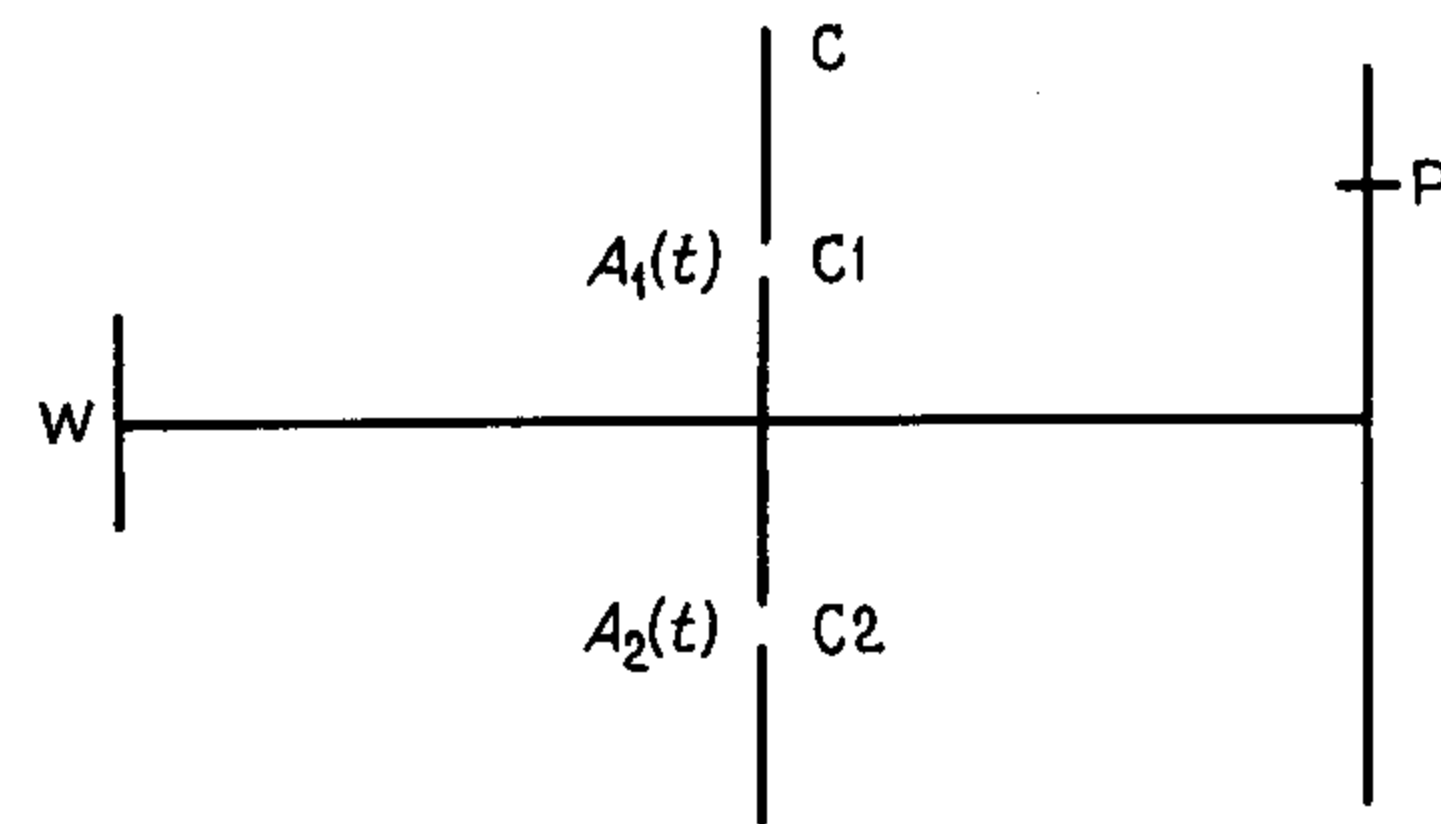


Рис. 6.7.

том от этих конкретных областей поля. Нас интересует зависимость, которая позволяет вычислить по таким измерениям временную и пространственную когерентность освещенности. Видность полос фактически является не чем иным, как убедительным физическим проявлением когерентности, и мы постараемся аналитически связать ее с более фундаментальным свойством.

Рассмотрим возникновение интерференции в некоторой произвольной точке Р (вне оси). Между интервалами прохождения расстояний С1–Р и С2–Р существует разностная задержка во времени, пусть она равна τ . Без потери общности эту разность можно отнести (исключительно с целью упрощения записываемых нами уравнений) к одному пути. Если комплексные амплитуды в С1 и С2 равны соответственно $A_1(t)$ и $A_2(t)$, то интенсивность в Р можно записать в виде

$$I_P = \langle [A_1(t + \tau) + A_2(t)] [A_1^*(t + \tau) + A_2^*(t)] \rangle, \quad (6.25)$$

где звездочкой обозначено комплексное сопряжение. (В целях упрощения этого представления были опущены коэффициенты передачи комплексной амплитуды при прохождении света через апертуры («функции распространения»), которые позволяют учитывать размер и расстояние апертур от Р.)

Угловые скобки $\langle \rangle$ используются в этом уравнении для обозначения того, что измерение интенсивности является (обычно) усредненным во времени процессом. Предполагается, что усреднение выполняется за интервал времени, который больше времени когерентности света.

Перемножение дает

$$I_P = \langle A_1(t + \tau) A_1^*(t + \tau) \rangle + \langle A_2(t) A_2^*(t) \rangle + \langle A_1(t + \tau) A_2^*(t) \rangle + \langle A_1^*(t + \tau) A_2(t) \rangle \quad (6.26)$$

Первые два члена представляют собой интенсивности в С1 и С2 соответственно, которые можно обозначить I_P^{C1} и I_P^{C2} . (В первом члене τ только смещает начальный момент усреднения и это обычно не влияет на величину среднего: система предполагается «стационарной».)

Сумма третьего и четвертого членов является суммой комплексного

числа с соответствующим комплексно сопряженным и потому равна лишь его удвоенной действительной части. Поэтому уравнение можно записать в виде

$$I_P = I_P^{C1} + I_P^{C2} + 2\Re[\Gamma_{12}(\tau)], \quad (6.27)$$

где $\Re[\Gamma_{12}(\tau)]$ представляет собой действительную часть величины

$$\Gamma_{12}(\tau) = \langle A_1(t + \tau) A_2^*(t) \rangle \quad (6.28)$$

В этом определении Γ_{12} усреднение во времени можно выразить через интегрирование, и тогда со ссылкой на разд. 4.7 становится очевидным, что $\Gamma_{12}(\tau)$ может быть описана как *комплексная функция кросс-корреляции* между освещенностями поля в С1 и С2, для которой колебания в С1 рассматриваются на время τ позже, чем в С2. В современном контексте физической оптики $\Gamma_{12}(\tau)$ также часто называется *комплексной функцией взаимной когерентности* поля освещенности в этих точках. Тогда $\Re[\Gamma_{12}(\tau)]$ называется *взаимной когерентностью*, и из сравнения с нашим анализом в разд. 1.1 дифракции на двойной апертуре ее роль в уравнении (6.27) четко соответствует отмеченному в уравнении (1.07) «интерференционному члену».

Если С2 совпадает с С1, то уравнение (6.28) принимает вид

$$\Gamma_{11}(\tau) = \langle A_1(t + \tau) A_1^*(t) \rangle, \quad (6.29)$$

что является *комплексной функцией самокогерентности* (или *автокорреляции*) поля в С1. При $\tau = 0$ Γ_{11} переходит в чистую интенсивность С1, и в этом упрощенном случае мы можем записать

$$\left. \begin{aligned} I_P^{C1} &= \Gamma_{11}(0), \\ I_P^{C2} &= \Gamma_{22}(0). \end{aligned} \right\} \quad (6.30)$$

Прежде чем двинуться дальше, полезно выполнить нормировку выражений типа уравнения (6.28), чтобы когерентность не зависела от модулей амплитуд (последние фактически не имеют отношения к когерентностям). Нормированная комплексная функция когерентности определяется как *комплексная степень взаимной когерентности* (или *кросс-корреляции*) $\gamma_{12}(\tau)$ в соответствии со следующим выражением:

$$\gamma_{12}(\tau) = \frac{\langle A_1(t + \tau) A_2^*(t) \rangle}{[\langle |A_1|^2 \rangle \langle |A_2|^2 \rangle]^{1/2}} = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0) \Gamma_{22}(0)}} \quad (6.31)$$

Ее называют также *фазокогерентным коэффициентом*.

С помощью неравенства Шварца в математике можно показать, что

$$0 \leq |\gamma_{12}(\tau)| \leq 1.$$

Тогда уравнение (6.27) принимает вид

$$I_P = I_P^{C1} + I_P^{C2} + 2\sqrt{I_P^{C1}}\sqrt{I_P^{C2}} \mathcal{R}(\gamma_{12}(\tau)), \quad (6.32)$$

где $\mathcal{R}(\gamma_{12}(\tau))$ обозначает действительную часть от $\gamma_{12}(\tau)$.

Это уравнение известно как *общий закон интерференции стационарных оптических полей*, или как *общий закон интерференции для частично когерентного света*. Важно, что он позволяет вычислить $\mathcal{R}(\gamma_{12}(\tau))$ по экспериментально измеренным значениям. Для представления выбранной пары точек поля освещенности используется, как и на рис. 6.7, экран с расположенными на нем определенным образом точечными отверстиями. Интенсивность I_P измеряется в разных положениях P , соответствующих определенному диапазону τ . Также измеряется интенсивность в P от каждого из точечных отверстий в отдельности I_P^{C1} и I_P^{C2} . Эти три величины дают $\mathcal{R}(\gamma_{12}(\tau))$ для каждого значения τ :

$$\mathcal{R}(\gamma_{12}(\tau)) = \frac{I_P - I_P^{C1} - I_P^{C2}}{2\sqrt{I_P^{C1}}\sqrt{I_P^{C2}}} \quad (6.33)$$

Точный смысл $\mathcal{R}(\gamma_{12}(\tau))$ становится более ясным, если записать $\gamma_{12}(\tau)$ в виде

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)| \exp(i\phi_{12}(\tau)), \quad (6.34)$$

где $\phi_{12}(\tau)$ рассматривается как обобщенный фазовый угол между полями в P . Тогда мы можем записать:

$$\phi_{12}(\tau) = \alpha_{12}(\tau) + \delta, \quad (6.35)$$

где δ – фазовый угол, обусловленный только внутренней разностью хода между $C1 \rightarrow P$ и $C2 \rightarrow P$.

Действительную часть $\gamma_{12}(\tau)$ в уравнении (6.34) можно теперь выразить как

$$\mathcal{R}(\gamma_{12}(\tau)) = |\gamma_{12}(\tau)| \cos[\alpha_{12}(\tau) + \delta], \quad (6.36)$$

и уравнение (6.32) принимает вид

$$I_P = I_P^{C1} + I_P^{C2} + 2\sqrt{I_P^{C1}}\sqrt{I_P^{C2}} |\gamma_{12}(\tau)| \cos[\alpha_{12}(\tau) + \delta]. \quad (6.37)$$

Его сходство с уравнением (1.01) теперь еще больше, и путем сравнения этих уравнений становится очевидным, что если величина $|\gamma_{12}(\tau)|$ равнялась бы своему предельному значению, единице, то интенсивность в P была бы такой же, как в случае идеального монохроматического света с разностью фаз $\alpha_{12}(\tau)$ между волнами в $C1$ и $C2$. При таких условиях поле излучения в $C1$ и $C2$ должно быть когерентно. Если величина $|\gamma_{12}(\tau)|$ равна другому своему предельному значению, нулю, то последний член в уравнении (6.37) отсутствует, интерференция отсутствует,

поле излучения в $C1$ и $C2$ некогерентно и наблюдаемая интенсивность является простой суммой независимых интенсивностей от $C1$ и $C2$. Что наиболее важно, промежуточные условия, т.е. *частичная когерентность*, могут быть количественно охарактеризованы с помощью величины $|\gamma_{12}(\tau)|$, называемой *степенью взаимной когерентности*.

Это представление приводит к удобной эквивалентной трактовке частичной когерентности как смеси когерентного и некогерентного света с отношением интенсивностей

$$\frac{I_{\text{ког}}}{I_{\text{неког}}} = \frac{|\gamma_{12}(\tau)|}{1 - |\gamma_{12}(\tau)|} \quad (6.38)$$

Однако нужно хорошо помнить о статистической в своей основе природе физического явления, которое моделируется этим подходом в целом.

Возвращаясь к уравнению (6.37), отметим, что мы до сих пор еще не видели, каким образом можно получить модуль и аргумент γ_{12} из экспериментальных измерений; у нас два неизвестных и только одно уравнение. Оценим вновь наше положение. Вначале для получения общей картины был постулирован источник, являющийся протяженным как в пространстве, так и по спектру. Все наши рассуждения до сих пор учитывали это, и в результате различные уравнения относительно γ_{12} не имеют ограничений по отношению к когерентности освещенности. Теперь вернемся к рис. 6.7 и проведем сравнение различных точек $C1$ и $C2$ в выборочной плоскости. Ясно, что эта схема в особенности чувствительна к пространственной (поперечной) когерентности. Для получения связи γ_{12} с наблюдаемыми величинами разумно рассмотреть случай, когда временная когерентность не вносит искажений (разд. 6.4.1). Функция $\Gamma_{11}(\tau)$ особенно удобна для изучения временной когерентности, поскольку она характеризует степень сохранения фазовых соотношений для отдельных волновых углов.

6.4.1. Пространственная когерентность

Предположим, что протяженный источник на рис. 6.7 является квазимонохроматичным в том смысле, что время когерентности много больше связанного с P максимального значения для области положений $C1$ и $C2$, в которых мы хотим изучить пространственную когерентность поля излучения. При таких условиях максимальное и минимальное значения I_P получаются из уравнения (6.37) при $\cos[\alpha_{12}(\tau) + \delta] = \pm 1$, так как при сделанных выше предположениях уровень интенсивности фона (и другие факторы, с которыми приходится иметь дело при детальном анализе) будут очень медленно меняться по интерференционной картине. Кроме того, если само τ очень мало, то можно показать, что степень взаимной когерентности $|\gamma_{12}(\tau)|$ лишь незначительно отличается от $|\gamma_{12}(0)|$, т.е. V и γ по существу не зависят от τ . Все это вместе взятое

дает следующее выражение для видности:

$$V_{12} = 2 |\gamma_{12}| \frac{\sqrt{I_P^{C1}} \sqrt{I_P^{C2}}}{I_P^{C1} + I_P^{C2}}. \quad (6.39)$$

Это и есть искомое соотношение, и если связанные с C1 и C2 отдельные интенсивности равны (как это зачастую бывает), то *видность полос равна степени взаимной когерентности (кросс-корреляции)*, т.е.

$$V_{12} = |\gamma_{12}|. \quad (6.40)$$

В отношении уравнения (6.36) следует заметить, что фазовый член α_{12} взаимной когерентности (т.е. любая разность фаз между полями в C1 и C2) определяется величиной смещения центрального максимума полос относительно оси, поскольку положение на оси соответствует $\delta = 0$.

Поэтому экспериментальные измерения видности и положения интерференционных полос непосредственно дают информацию как о модуле, так и об аргументе комплексной степени пространственной когерентности для поля излучения протяженного источника.

Помня об этом соотношении между видностью полос и корреляцией, мы вернемся к сходству между парами Фурье, упомянутому в разд. 6.2.2, а именно парой видность полос – распределение яркости на рис. 6.4 и парой дифракционная картина – апертурная функция, хорошо знакомой нам из предыдущих глав. Как было указано в свое время, это сходство не является случайным или присущим лишь конкретному примеру. Можно показать, что так называемая «картина» комплексной степени когерентности (кросс-корреляция) в плоскости, освещаемой протяженным источником, совершенно аналогична картине комплексных амплитуд дифракции от апертуры того же размера и формы, что и данный источник. Формально это выражается теоремой ван Циттерта–Цернике, которую можно найти в более специальных пособиях.

Помимо своего теоретического значения, указанная теорема особенно важна с вычислительной точки зрения, поскольку, как правило, расчет когерентности оказывается более трудным, чем расчет соответствующей дифракционной картины.

Наиболее важно то, что распределение яркости можно вычислить на основе преобразования Фурье от кросс-корреляционной функции, полученной с помощью данных о фазе и амплитуде видности полос. Из нашего анализа спектрального интерферометра следует, что аналогичная связь существует между автокорреляционной функцией и спектральным распределением. Этот вопрос рассматривается в следующем разделе.

6.4.2. Временная когерентность

Здесь мы покажем, что фурье-преобразование комплексной функции *самокогерентности* (или *автокорреляционной функции*) $\Gamma_{11}(\tau)$ является распределением интенсивности спектра (*спектром мощности*). Начнем

с записи определения среднего во времени от $\Gamma_{11}(\tau)$ в уравнении (6.29) в виде интеграла:

$$\Gamma_{11}(\tau) = \int A(t + \tau) A^*(t) dt, \quad (6.41)$$

где использовано предыдущее определение A и A^* . Тогда по определению преобразованием $\Gamma_{11}(\tau)$ является

$$T[\Gamma_{11}(\tau)] = \int \Gamma_{11}(\tau) \exp(2\pi i \nu \tau) d\tau, \quad (6.42)$$

где ν и τ – сопряженная пара частота – время.

Подстановка Γ_{11} из уравнения (6.41) дает

$$T[\Gamma_{11}(\tau)] = \iint_{\tau t} A(t + \tau) A^*(t) \exp(2\pi i \nu \tau) dt d\tau = \quad (6.43)$$

$$= \int_t A^*(t) \exp(2\pi i \nu t) dt \int_{\tau} A(t + \tau) \exp[-2\pi i \nu (t + \tau)] d\tau = T[A^*(t)] T[A(t)]. \quad (6.44)$$

Второй интеграл в уравнении (6.43) взят по τ и не зависит от t , поскольку система предполагается стационарной – см. объяснения в связи с уравнением (6.26).

Полагая $F^*(\nu) = T[A^*(t)]$ и $F(\nu) = T[A(t)]$, получаем

$$T[\Gamma_{11}(\tau)] = F^*(\nu) F(\nu) = |F(\nu)|^2. \quad (6.45)$$

Полученное равенство преобразования Фурье от автокорреляционной функции со спектром интенсивности (мощности) является формулировкой *теоремы Винера–Хинчина* (см. разд. 4.7.1).

Это выражение совпадает с результатом подхода Майкельсона–Рэлея в разд. 6.3.2. Величину $|\Gamma_{11}(\tau)|$, или ее нормированный аналог $|\gamma_{11}(\tau)|$, можно связать с видностью таким же образом, как было сделано для $|\gamma_{12}|$ в уравнении (6.39). Напомним, что распределение интенсивности в спектре $|F(\nu)|^2$ было обозначено в разд. 6.3.2 как $f(\sigma')$. Кривые видности, полученные с помощью двухлучевого спектрального интерферометра, можно интерпретировать как представляющие $|\gamma_{12}(\tau)|$, функцию разницы длины пути (соответствующей времени τ), которая была введена с целью сравнения волнового пакета с самим собой.

Теперь элегантность и симметрия двух пар фурье-преобразований стала для нас поразительно очевидной. Кривая видности в спектроскопии определена во временном пространстве, т.е. она является функцией временной задержки, внесенной в два оптических пути спектрального интерферометра, в котором волновой пакет сопоставляется сам с собой (автокорреляция); здесь преобразование представляет собой интенсивность (мощность) спектра источника. В звездном (пространственном) интерферометре кривая видности является функцией расстояния между двумя точками поля освещенности, которые сравниваются (кросс-корреляция); ее преобразование представляет собой пространственное угловое распределение яркости источника.

6.5. ФУРЬЕ-СПЕКТРОСКОПИЯ

Мы видели в разд. 6.3, что в спектральном интерферометре Майкельсона каждый волновой пакет света делится по амплитуде на два волновых пакета, которые после введения между ними разности пути сводятся вместе. Было показано, что характер изменения видности интерференционных полос с изменением разности пути (путем перемещения зеркала M_1 на рис. 6.5) определяется спектральным составом света. Однако систематическое практическое применение этой возможности задерживалось до появления современных вычислительных машин. Сейчас главным образом благодаря прогрессу за последнее время в создании небольших специализированных ЭВМ и микропроцессоров, ставших неотъемлемой частью лабораторного оборудования, преимущества фурье-спектропии используются все шире.

В этом разделе мы более детально рассмотрим метод Фурье в контексте некоторых современных методов и приложений.

Схема, используемая в некоторых типах современных фурье-спектрометров, показана на рис. 6.8. Она отличается от схемы на рис. 6.5 одной главной особенностью: свет от источника сводится в пучок (коллимируется) зеркалом C до деления амплитуд делителем пучка B . Это вариант Тваймана–Грина для интерферометра Майкельсона. Коллимирование позволяет сделать все поперечное сечение поля освещенности в инструменте соответствующим осевому ($\theta = 0$) направлению на рис. 6.5. Поэтому кольцевые полосы отсутствуют и все поле имеет равномерную яркость. Возникающие при перемещении зеркала M_1 изменения интенсивности измеряются с помощью показанной на рисунке системы зеркала и детектора. Таким образом, для рассматриваемого нами гипотетического случая монохроматического света детектор снова должен регистрировать синусоидальный характер изменения интенсивности излучения. Если волновое число равно σ_1 и слагаемые пучки имеют равные амплитуды A_1 , то интенсивность в зависимости от l дается выражением

$$I_{\sigma_1}(l) = 4A_1^2 \cos^2(\pi\sigma_1 l),$$

что соответствует уравнению (6.17) при разности хода $l = 2h$, поскольку $\theta = 0$. Это выражение можно переписать в виде

$$I_{\sigma_1}(l) = 2A_1^2 (1 + \cos 2\pi\sigma_1 l) \quad (6.46)$$

Таким образом, измеренная детектором интенсивность света со спектральным распределением $B(\sigma)$ определяется без учета числового множителя выражением

$$I(l) = \int_0^\infty B(\sigma) (1 + \cos 2\pi\sigma l) d\sigma \quad (6.47)$$

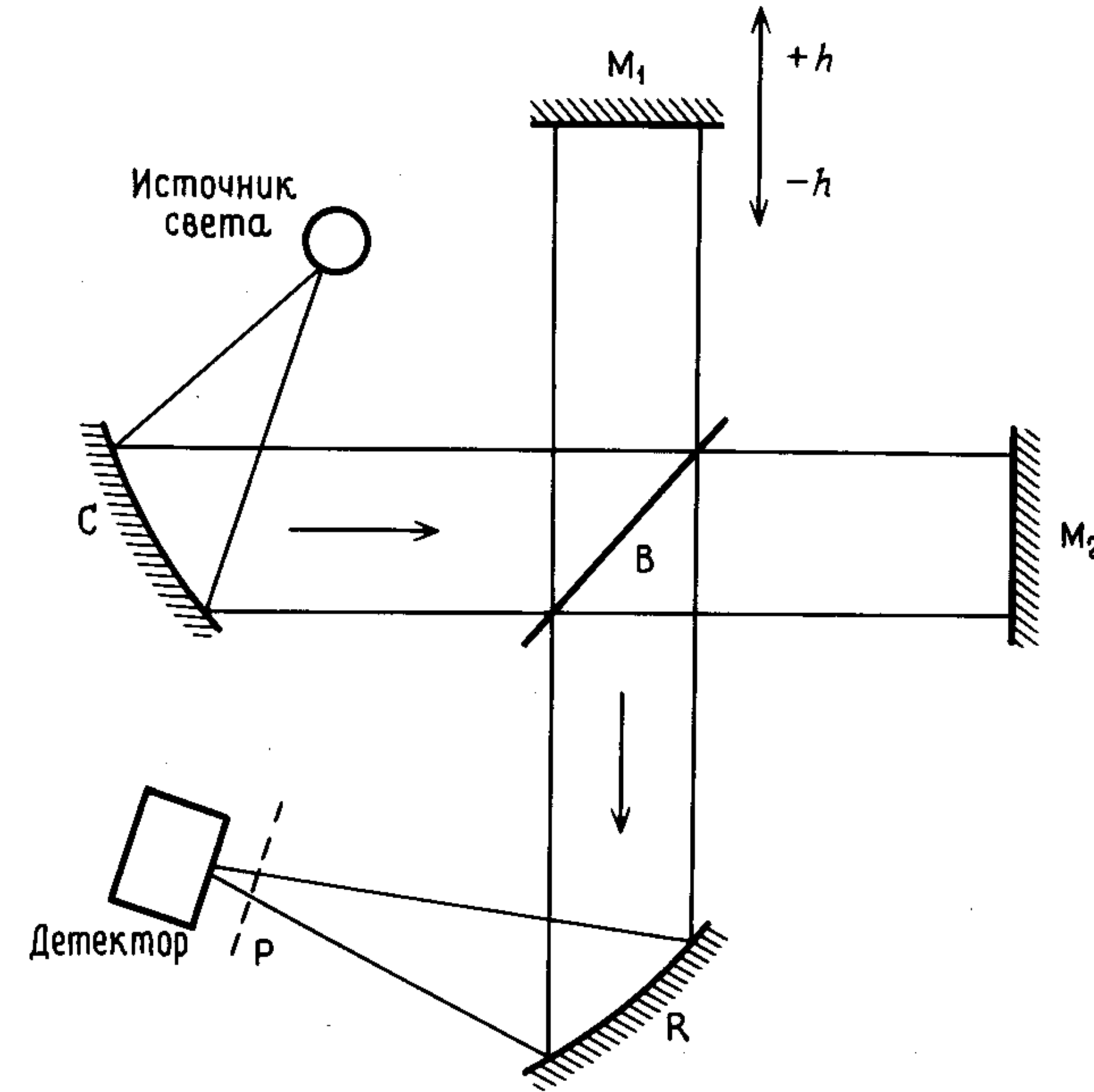


Рис. 6.8. Фурье-спектропия [разность пути $2(BM_1 - BM_2) = 2h = l$].

Этот интеграл можно записать в виде

$$I(l) = \int_0^\infty B(\sigma) d\sigma + \int_0^\infty B(\sigma) \cos(2\pi\sigma l) d\sigma. \quad (6.48)$$

Когда два зеркала находятся на равном оптическом расстоянии от делителя пучка, то $l = 0$ и $I(l) = I(0) = 2 \int_0^\infty B(\sigma) d\sigma$. Поэтому мы имеем

$$I(l) - \frac{1}{2} I(0) = \int_0^\infty B(\sigma) \cos(2\pi\sigma l) d\sigma \quad (6.49)$$

Это выражение для косинусного преобразования Фурье, и мы можем записать

$$B(\sigma) = \int_0^\infty \left[I(l) - \frac{1}{2} I(0) \right] \cos(2\pi\sigma l) dl. \quad (6.50)$$

Уравнение (6.50), получаемое более строго в специальной литературе, часто называется основным уравнением фурье-спектропии. Оно дает возможность вычислить для каждого конкретного σ интенсивность спектра $B(\sigma)$ путем интегрирования выражения в правой части.

Посмотрим, что сюда входит. Величина $I(l)$ представляет собой из-

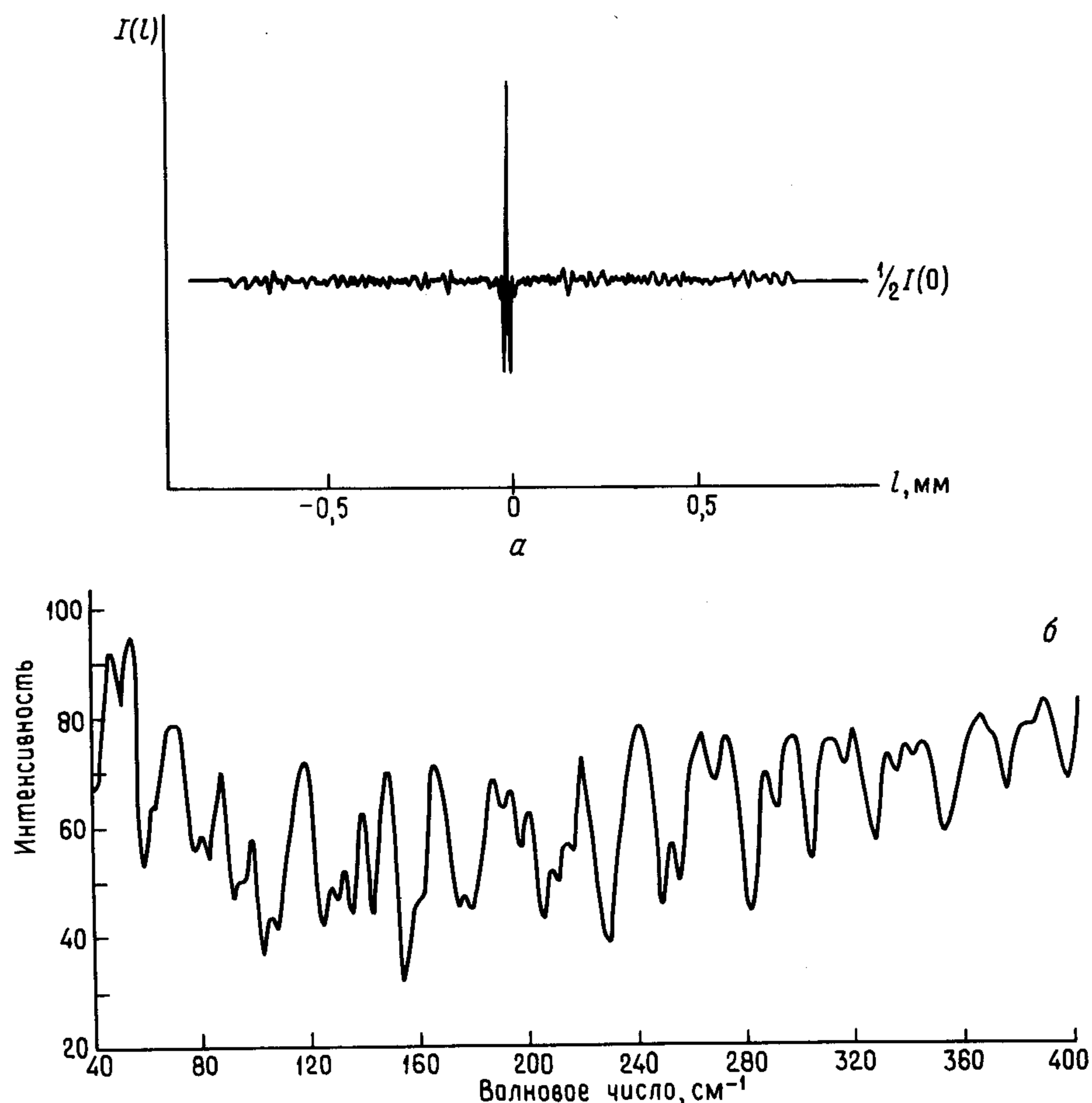


Рис. 6.9. а – интерферограмма; б – вычисленный по ней спектр. (Белл, 1972.)

меренную как функция от l интенсивность. Изображенная графически, она называется «интерферограммой». Ее пример показан на рис. 6.9, а. (Будем считать, что имеются и отрицательные значения l .) Далее, в любой интерферограмме при $l=0$ отсутствует какая-либо разность хода между слагаемыми волновыми фронтами, амплитуды складываются и $I = I(0) = \sum (2A)^2 = 4\sum A^2$. С другой стороны, при больших l , когда интерференции уже больше не происходит (т.е. l превышает длину когерентности для излучения), то складываются отдельные интенсивности, что дает $2\sum A^2 = I(0)/2$. Для излучения с такими свойствами из рис. 6.9, а следует, что кривой видности будет соответствовать зависимость $[I(l) - \frac{1}{2}I(0)]/\frac{1}{2}I(0)$ (видности в определенном раньше смысле). Поэтому мы

можем сказать, что уравнение (6.50) связывает $B(\sigma)$ с преобразованием Фурье от кривой видности, полученной из интерферограммы. На рис. 6.9, б показан результат, соответствующий рис. 6.9, а.

Заметим, что вычисление $B(\sigma)$ в уравнении (6.50) для каждого конкретного волнового числа, например σ_1 , сводится к умножению всей кривой $[I(l) - \frac{1}{2}I(0)]$ на $\cos(2\pi\sigma_1 l)$, т.е. вычисление площади под кривой произведения дает конкретное значение $B(\sigma_1)$. Так же был получен и график на рис. 6.9, б, правда, с помощью ЭВМ. Эта процедура напоминает получение коэффициентов Фурье в гл. 3 (см. рис. 3.2), за тем исключением, что теперь мы имеем дело с нерегулярными картинками, и потому используется интеграл, а не ряды. Кроме того, расчеты выполняются в пространстве частота–время, где характеристиками являются σ и l (l соответствует задержке во времени).

Экспериментальные измерения для отрицательных значений l выполнить так же легко, как для положительных, хотя интерферограмма и должна быть симметричной. Если по какой-либо причине она не оказывается симметричной, то косинусное преобразование Фурье использовать нельзя, поскольку оно явно предполагает, что $I(l)$ четная функция. По причинам, которые в данной книге не рассматриваются, некоторые приборы в действительности делаются асимметричными. В ряде из них образец помещается в одно из плеч прибора, например перед зеркалом M_2 . Их использование относится к области «амплитудной спектроскопии», включающей измерение фазы наряду с амплитудой и требующей вычисления комплексного преобразования Фурье.

Относительно уравнения (6.50) следует заметить, что в нем предполагается возможность точного определения положения с $l=0$, а также возможность интегрирования от $l=0$ до $l=\infty$. Первое из этих условий осуществить трудно, а второе вообще невозможно (рис. 6.9). Более сложные методы, упомянутые выше, конечно, помогают преодолеть первую трудность, а математическая аподизация используется в качестве средства минимизации ошибок, обусловленных несоответствием пределов в интегрировании фурье-преобразования [ср. использование аподизации для исключения вторичных колец вокруг диска Эри (разд. 2.3)]. Однако разрешающая способность прибора, хотя и ограниченная, может быть, очевидно, очень большой, поскольку она определяется по существу экспериментально достижимым верхним пределом l .

Интерпретацию сути метода можно также описать на языке автокорреляции. Мы только что отметили, что введение разности хода l приводит к временной задержке – задержке между моментами прихода волновых цугов, которые сводятся вместе для интерференции. Это полностью соответствует сути автокорреляции (разд. 4.7), поскольку каждая пара волновых цугов до прихода к делителю пучка относится к одному волновому цугу и по существу каждый первичный волновой цуг сравнивается сам с собой со сдвигом по длине. Временная задержка,

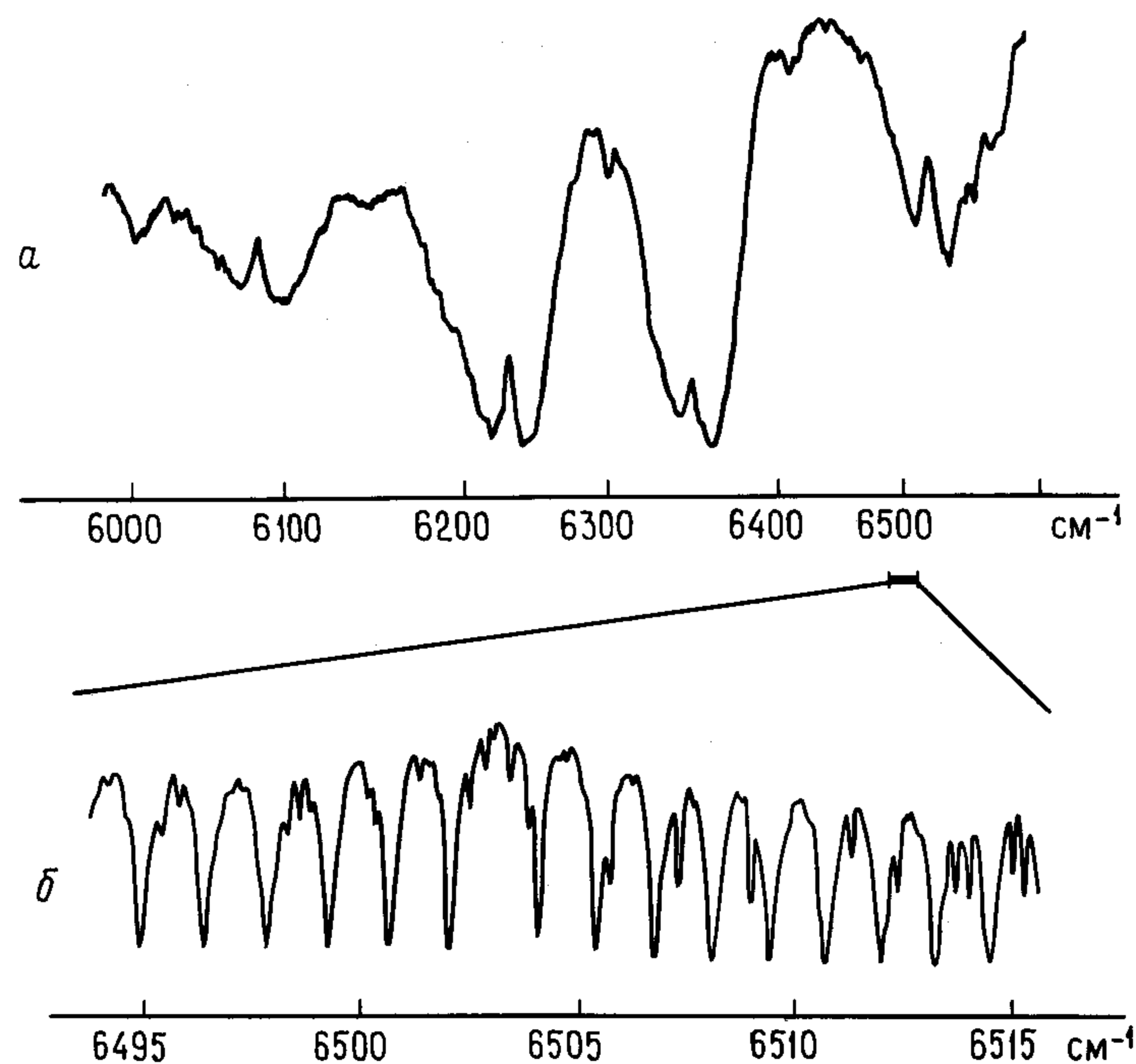


Рис. 6.10. Спектр излучения Венеры. *a* – обычная дифракционная спектроскопия; *б* – результат фурье-спектроскопии, показывающий детали вращательной структуры CO_2 в области вблизи 6505 см^{-1} [21].

τ , равна l/c , где c – скорость света. Поэтому амплитуды каждой пары слагаемых волновых цугов могут быть записаны как $A(t)$ и $A(t + \tau)$. Тогда при каждом значении l измеряется не что иное, как $\langle A(t) A(t + \tau) \rangle$. Это автокорреляционная функция для света, и из разд. 6.4 мы знаем, что ее преобразование Фурье равно спектру мощности излучения $|F(\nu)|^2$ (теорема Винера – Хинчина).

Вернемся к измерению l . Изменение l достигается при перемещении зеркала, и механическая точность, с которой его можно контролировать, ставит предел точности прибора в целом. В результате рассматриваемый метод применим главным образом для работы на длинных волнах и широко использовался на протяжении многих лет в инфракрасной области. Однако в настоящее время благодаря всестороннему техническому совершенствованию область длин волн расширена в видимый и даже ультрафиолетовый диапазоны.

В особенности важны исследования поглощения веществом в инфракрасном диапазоне ввиду характера получаемой при этом информации. В этом диапазоне вещества поглощают излучение по причинам, обусловленным целым рядом явлений. К ним относятся колебания (внутри-

и межмолекулярные) и вращения молекул относительно определенных химических связей.

Помимо использования в фундаментальных исследованиях различных материалов, инфракрасные спектры являются мощным средством анализа. На рис. 6.10 приведен пример, который интересен с нескольких точек зрения. На рис. 6.10, *б* мы имеем спектр излучения Венеры, полученный с помощью фурье-спектрометра (см. следующий раздел). Он свидетельствует о поглощении света молекулами двуокиси углерода в атмосфере планеты: все детали спектра можно отождествить с колебательными модами CO_2 . На рис. 6.10, *а* показан тот же спектр, полученный на обычном спектрометре с дифракционной решеткой.

Здесь снова можно указать, что группы атомов, прикрепленные к цепочкам из атомов углерода в органических молекулах, являются, как правило, центрами химической активности молекулы. Эти «функциональные группы» имеют характерные частоты колебаний и с аналитической точки зрения обеспечивают простые, быстрые и надежные свидетельства, позволяющие отнести сложное вещество к соответствующему химическому классу.

Этот метод используется в дополнение к другим методам неорганической химии, в том числе для изучения физических свойств материалов в твердом состоянии. Он применим также в исследовании плазменных процессов, явлений релаксации и т.п.

Для проведения лабораторных исследований поглощающих свойств материалов источник света подбирается в соответствии с изучаемой спектральной областью, а материал, спектр поглощения которого исследуется, помещается в ячейку для образцов *P* на рис. 6.8. В качестве делителя пучка для далекой инфракрасной области (волновые числа меньше 400 см^{-1} , см. приложение Г) можно использовать майларовую или полиэтиленовую пленку. Для измерений в области от 200 до 4000 см^{-1} разработаны устройства, где применяется германиевая пленка, нанесенная на различные подложки (например, KBr , CsI). В области 2000 – $16\,000 \text{ см}^{-1}$ применяются пленки из окиси железа Fe_2O_3 на подложке CaF_2 . При этом возможно использование не показанной на рисунке компенсирующей пластинки.

Выбор детектора также зависит от спектральной области. Как правило, используются детектор Голея, пирозлектрический болометр или детектор на полупроводниках с фотопроводимостью. Для уменьшения поглощения парами воды воздух из системы откачивается до давления 10^{-2} – 10^{-3} мм рт. ст. Конечно, даже и в этих условиях необходимо получать интерферограммы с образцом и без него, чтобы внести поправку на характерную инструментальную интерферограмму.

Все спектроскопические исследования в инфракрасном диапазоне имеют одно общее свойство: низкие энергии, обусловленные большими длинами волн. В далеком инфракрасном диапазоне, переходящем в радиодиапазон СВЧ, энергия излучения может оказаться в тысячи раз меньше, чем в видимом спектре (см. приложение Г). Это те области, где,

как мы видели, метод интерферометрии выигрывает по сравнению с классическими методами спектроскопии, в которых используются призмы и спектрометры с дифракционными решетками. Это обусловлено двумя причинами. Во-первых, как показал П. Жакино, все интерферометры с круговой симметрией, как, например, интерферометр Майкельсона, принимают больший поток излучения, чем щелевые спектрометры равной разрешающей способности [29, 30]. Во-вторых, метод интерферометрии «мультиплексирует» спектр, т.е. дает информацию для всего спектра одновременно. Это было количественно показано Фелжетом в 1951 г. [17] (см. также [18]). При перемещении подвижного зеркала все поступающее на интерферометр излучение используется для построения интерферограммы, и она содержит всю информацию о спектре излучения. В результате получается значительное увеличение в отношении сигнал/шум по сравнению с ожидаемым для диспергирующего инструмента, например спектрометра с дифракционной решеткой, где спектр просматривается последовательно по длинам волн. (Для некоторых типов детекторов отношение сигнал/шум можно еще более улучшить с помощью усреднения сигнала, и это используется сейчас повсеместно.)

Кроме спектроскопических исследований в инфракрасной, а также в упомянутых ранее видимой и ультрафиолетовой областях, фурье-методы применяются в настоящее время и в других видах спектроскопии, включая спектроскопию ядерного магнитного резонанса (ЯМР), масс-спектроскопию и ее модификацию, известную как спектроскопия ион-циклотронного резонанса (ИЦР).

6.6. ПРИЛОЖЕНИЯ В АСТРОНОМИИ

Взглянув на выражение

$$\theta = 1,22\lambda/a \quad (2.07)$$

для предела углового разрешения θ телескопа с апертурой диаметром a , можно сделать вывод, что в радиоспектре с длинами волн в диапазоне $10\text{ см} - 10^5\text{ см}$ (приложение Г) потребуется телескоп-рефлектор с диаметром в $2 \cdot 10^5 - 2 \cdot 10^9$ раз большим, чем для эквивалентного оптического телескопа, работающего на волне около 500 нм . Кроме того, в то время как оптический телескоп сразу образует в своем поле зрения изображения звезд, радиотелескоп дает только один сигнал в виде напряжения, обусловленного излучением, принимаемым от всех направлений, определяемых его разрешающей способностью. Неудивительно поэтому, что в течение нескольких лет после того, как в 1931 г. К. Дж. Янским были открыты космические или галактические шумы, все еще не было ясным, откуда они приходят.

В последующие годы Янский и другие исследователи выполнили различные наблюдения. В 1945 г. И. В. Эпплтон сообщил, что в диапазо-

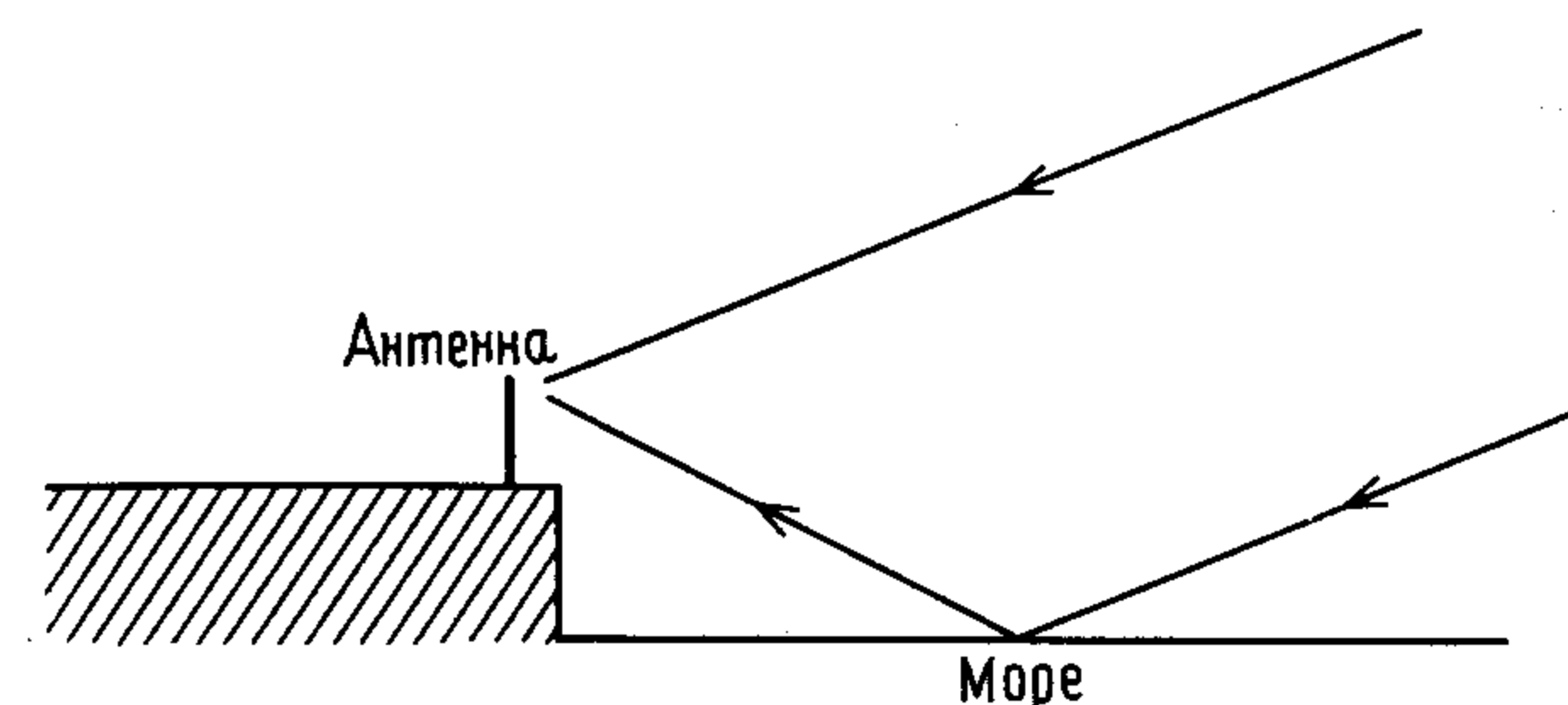


Рис. 6.11. Морской интерферометр.

не 60 МГц (волна 5 м) определенная часть излучения связана с сильной активностью солнечных пятен. В 1946 г. Д. С. Хей с сотрудниками открыл вариации интенсивности галактического шума в направлении созвездия Лебедь, высказав предположение, что это конкретное излучение возникает в дискретном источнике за пределами Солнечной системы. Так началось исследование радионеба.

Конечно, в этой небольшой книге недостаточно места для изложения удивительной и занимательной истории радиоастрономии. Тем не менее здесь целесообразно упомянуть о двух важных достижениях, в которых используются методы Фурье.

6.6.1. Апертурный синтез

В 1946 г. антенны радиолокаторов имели на волне $1,5\text{ м}$ ширину луча около 10° , чего было совершенно недостаточно для выделения, например, областей на поверхности Солнца из общего галактического фона. Интерферометрия представляла возможность разрешения этой трудности, и потому Д. Л. Поззи и его коллеги в Австралии провели такие же наблюдения, как и Эпплтон, используя антенну, расположенную высоко на выступающем над морем утесе в Сиднее. Установка (рис. 6.11), которая служила интерферометром, по своей схеме была аналогична опыту Ллойда с зеркалом в оптике. Интерференция возникала между прямыми радиосигналами и сигналами, отраженными от поверхности моря. По существу, метод был тем же самым, что и в звездном интерферометре Майкельсона, но с недостатком, состоящим в наличии фиксированной базовой линии. С помощью той же антенны Болтон и Стэнли [4] успешно зарегистрировали лепестки от источника в Лебеде-созвездии, которое поднималось лишь незначительно над горизонтом в Сиднее. Австралийские исследователи нашли также другие источники, включая небольшой интенсивный источник в созвездии Тельца. Этот объект наряду с другими первыми «радиозвездами» быстро был отождествлен с Крабовидной туманностью.

Одновременно Мартин Райл (ныне сэр) и Д. Д. Вонберг в Кембридже также исследовали излучение Солнца. Они пользовались интерферометром, который был прямым аналогом интерферометра Майкельсона

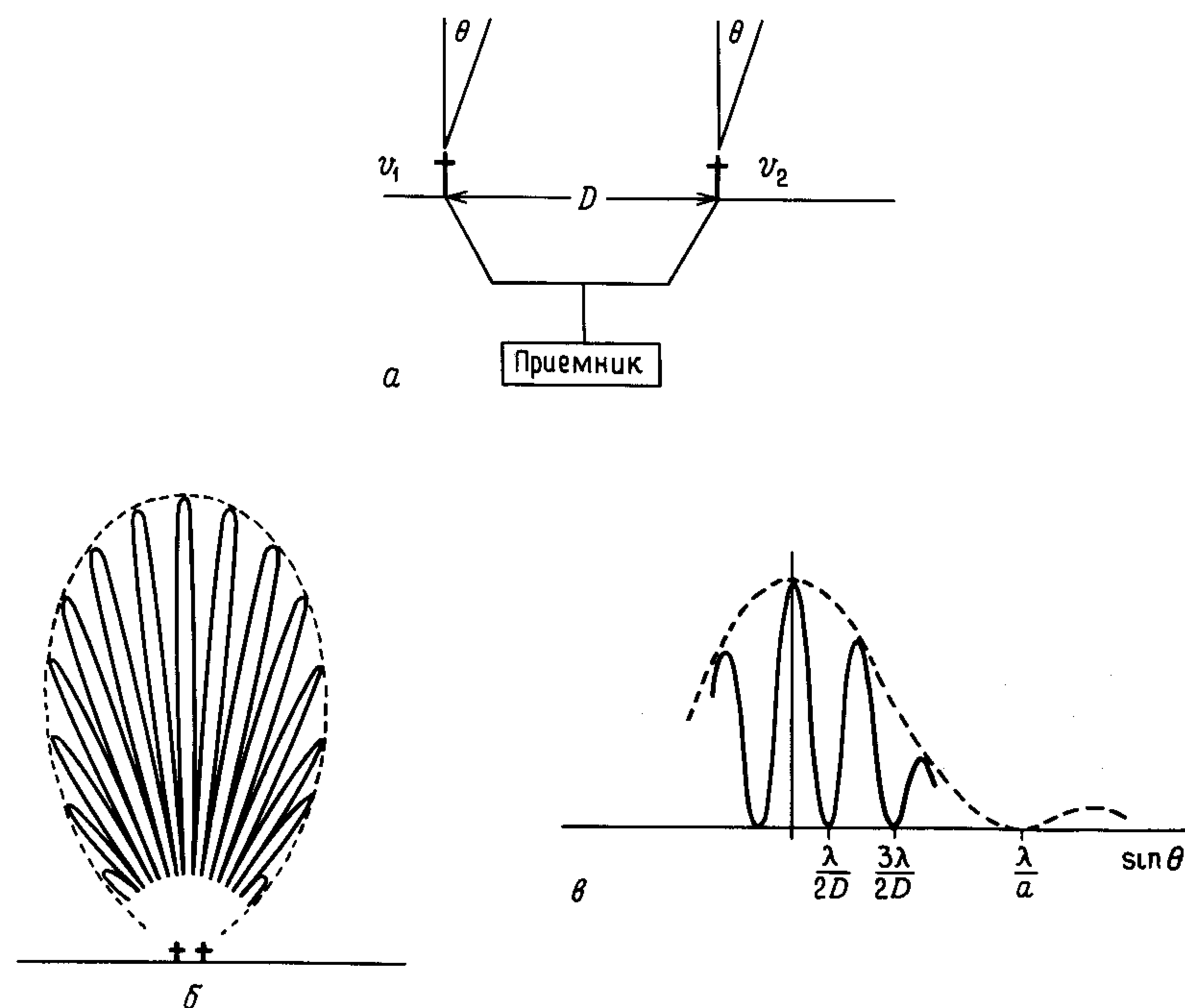


Рис. 6.12. Аналог звездного интерферометра Майкельсона в радиодиапазоне.

(рис. 6.12, а). Он имел две системы антенн (дипольного типа), разнесенных по горизонтали в направлении восток–запад, и принятые сигналы подавались на приемное устройство. На рис. 6.12, б показан вид «приемной полярной диаграммы» такого интерферометра, огибающая которой представляет собой полярную диаграмму одиночной антенны. Целесообразность применения такой схемы была связана с тем, что если расстояние между последовательными минимумами в этой «двухпертурной» схеме сделать достаточно большим по сравнению с угловым размером Солнца, то при прохождении полярной диаграммой через Солнце, вызванном вращением Земли, Солнце в отличие от галактического фона должно регистрироваться в виде осциллирующей кривой.

Точность в этом методе оценивается с учетом рис. 6.12, в, из которого мы находим, что угловой размер лепестков на рис. 6.12, б равен полуширине центрального максимума для параболической антенны с диаметром, равным расстоянию между двумя антеннами, образующими интерферометр. Иными словами, по способности локализовать источники и измерять их угловые диаметры интерферометр можно сравнить с обычным параболическим радиотелескопом, если его диаметр равен

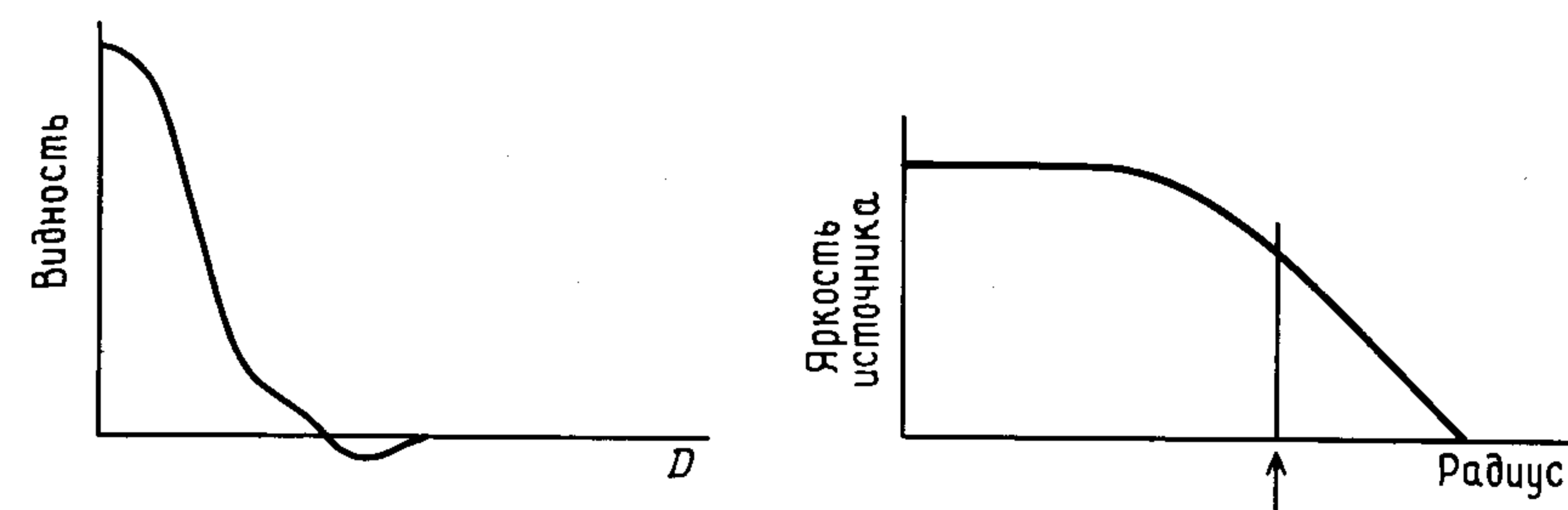


Рис. 6.13. Кривая видности и радиальное распределение радиояркости по диску Солнца (стрелкой отмечен край Солнца в оптике) [50].

удвоенному расстоянию между двумя антеннами интерферометра¹⁾.

Во время появления в 1946 г. [54] большого солнечного пятна, когда излучение Солнца существенно возросло, Райл и Вонберг воспользовались своим прибором для определения углового диаметра радиоисточника на Солнце. Для различных расстояний между антеннами они измерили отношение максимума к минимуму лепестков, образующих интерференционную кривую. На основе этих результатов они заключили, что угловой диаметр источника составляет $10''$. Так как это значение существенно не превышало диаметр визуально наблюдаемого солнечного пятна, они заключили, что радиоисточник относится к визуальному пятну или по крайней мере связан с ним.

С помощью двухантенного интерферометра Райл и Смит [55] открыли самый яркий источник северного неба в Кассиопее, но не смогли его отождествить. Они оценили, что его угловой диаметр меньше $6''$.

Принципы Фурье в интерферометрии с переменной базой, позволяющие получить фактическую структуру радиоисточника, были заложены Поззи с коллегами в вышеупомянутой работе. Стэйни [59] в Кембридже использовал для проверки теории, разработанные в конце 40-х годов, согласно которым излучение Солнца в отсутствие солнечных пятен было необычайно сильным в направлении лимба на волнах около 60 см. По существу так же, как это было описано для интерферометра Майкельсона (разд. 6.2.2), видность лепестков была измерена для расстояний между антеннами вплоть до 365 длин волн. Поскольку ориентация антенной системы была фиксированной, вычисления должны были исходить из предположения о круговой симметрии источника. Фурье-преобразование кривой видности давало радиальное распределение интенсивности. (Строго говоря, здесь должно иметь место преобразование Фурье–Бесселя.) На рис. 6.13 показан общий вид результатов с отсутствием указаний на уярчение к краю, чего ожидали некоторые исследователи.

¹⁾ Отметим, что точность локализации источников определяется отношением сигнал/шум. В современных системах она может достигать $\sim 1\%$ от ширины интерференционного лепестка.—Прим. ред.

Корреляционный принцип методов интерферометрии играет очень важную роль. К нему мы теперь и перейдем.

На рис. 6.12, а v_1 и v_2 представляют напряжения, принимаемые двумя антеннами и подаваемые на приемник, выходной сигнал которого пропорционален среднему от их произведения¹⁾. Кроме того, v_1 и v_2 содержат «шумовой» компонент, так как отдельные антенны имеют широкую диаграмму направленности. Из-за своей различной природы этот шум между антеннами не коррелирует²⁾. Поэтому мы можем записать:

$$\langle v_1 v_2 \rangle = \langle (v_{u1} + v_{c1})(v_{u2} + v_{c2}) \rangle = \quad (6.51)$$

$$= \langle v_{u1} v_{u2} \rangle + \langle v_{u1} v_{c2} \rangle + \langle v_{c1} v_{u2} \rangle + \langle v_{c1} v_{c2} \rangle,$$

где индексы u и c относятся к некоррелирующим и коррелирующим сигналам соответственно.

Если время усреднения достаточно большое, то все члены, за исключением последнего, равны нулю, так как они представляют собой произведения некоррелирующих напряжений. Усредненное во времени произведение v_1 и v_2 просто должно быть равно

$$\langle v_1 v_2 \rangle = \langle v_{c1} v_{c2} \rangle. \quad (6.52)$$

Однако время усреднения должно быть конечным, шумовые флуктуации будут сохраняться и требуемый сигнал может оказаться малым по сравнению с ненужными шумами. Для снижения этого эффекта Райл [53] предложил метод переключения фазы, при котором одна из антенн поочередно подключалась в фазе и противофазе с помощью отрезка линии передачи длиной $\lambda/2$ (рис. 6.12, а).³⁾ Измерялась разность между двумя выходными сигналами, что позволяло получить

$$\langle (v_1 + v_2)^2 \rangle - \langle (v_1 - v_2)^2 \rangle = 4 \langle v_1 v_2 \rangle \quad (6.53)$$

Впоследствии были разработаны более совершенные методы, но принцип остается тем же самым.

Интерферометр в таком виде измеряет кросс-корреляцию между сигналами на двух антеннах как функцию расстояния между ними; ее преобразование Фурье представляет пространственное распределение яркости источника (разд. 6.4.1).

¹⁾ Указанный случай относится к корреляционному радиоинтерферометру. — Прим. ред.

²⁾ Этот протяженный фон имеет ту же природу, что и излучение радиоисточника, но корреляция сигналов отсутствует из-за большой протяженности области излучения — малого значения функции видности (хотя этот сигнал создает общую «засветку» интерферограммы). С другой стороны, на вход приемника поступают шумы, излучаемые самими высокочастотными кабелями, и эти шумы некогерентны. — Прим. ред.

³⁾ Таким способом М. Райл исключил влияние «засветки». Полностью исключить шумовые флуктуации невозможно, но указанный способ позволяет снизить флуктуации, порождаемые нестабильной работой приемника. — Прим. ред.

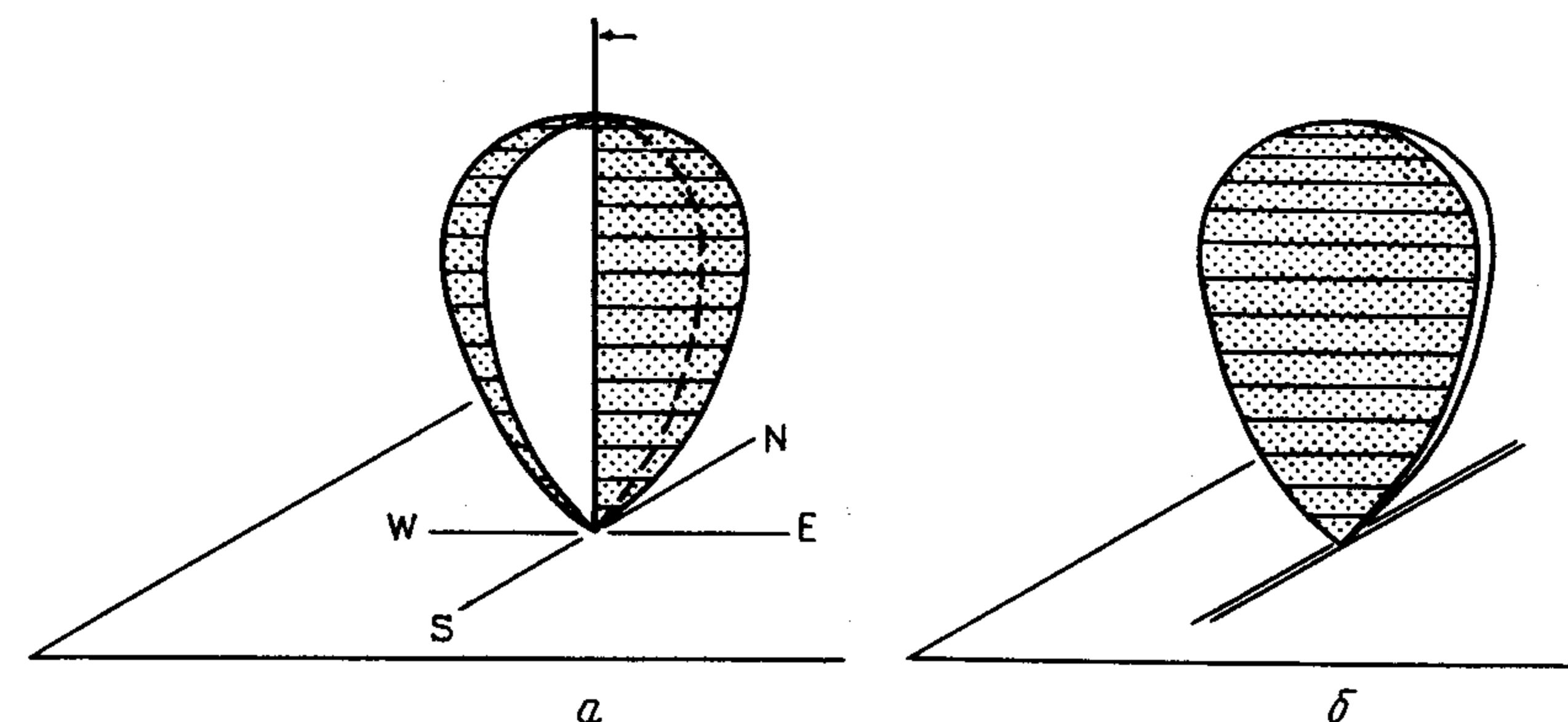


Рис. 6.14. Крест Миллса. а — «карандашная» диаграмма направленности (показана стрелкой); б — «веерная» диаграмма.

Для улучшения углового разрешения при изучении распределения галактического излучения был разработан ряд различных систем с незаполненной апертурой. Они обладали способностью измерять только кросс-корреляцию (т.е. $\langle v_{c1} v_{c2} \rangle$), благодаря чему удавалось «перемножить» полярные диаграммы двух различных антенн. Принцип, согласно которому эффективная площадь для получения двумерного углового разрешения не обязательно должна быть такой же, как площадь для достижения данной чувствительности, лежит в основе определения «незаполненной апертуры».

Райл [53] описал преимущества длинных антенных решеток и различных схем переключения фазы для улучшения параметров систем. Успешным примером реализации такого подхода стал «крест Миллса» в Австралии, названный по имени его конструктора Б.Н. Миллса. Различные виды этого интерферометра были впоследствии построены в ряде стран. Первоначальный вариант [44] состоял из двух взаимно перпендикулярных узких «апертур» в виде линейных решеток в направлениях С–Ю и В–З (рис. 6.14, а). Каждая решетка состояла из 250 диполей (на длину волны 3,5 м), расположенных вдоль плеча длиной 1500 футов (около 457 м), и, следовательно, имела полярную диаграмму, напоминающую веер, как показано на рис. 6.14, б. При соединении двух «вееров» вместе в области их пересечения образуется узкая «карандашная» диаграмма направленности. Изящный способ переключения позволил измерить только те сигналы, которые принимались в пределах этой узкой диаграммы. Для измерения положения угла места диаграммы направленности между всеми элементами решетки С–Ю подключались кабели соответствующей длины — вводилась разность фаз.

До сих пор мы в самых общих чертах описывали некоторые основные оптические принципы получения высокой разрешающей силы на радиоволнах, а также то, каким образом методы с «незаполненной

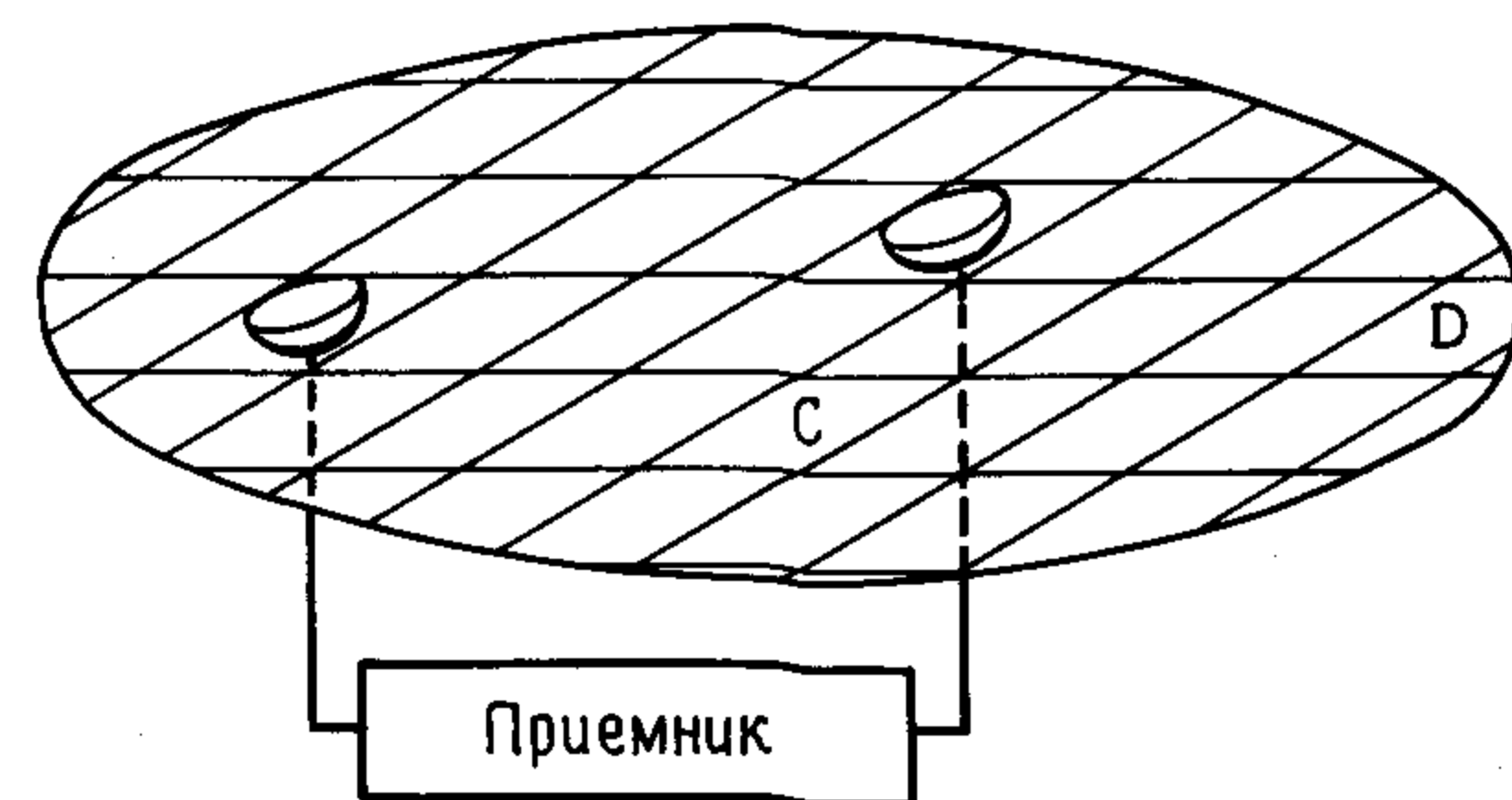


Рис. 6.15. Апертурный синтез.

апертурой» помогают достичь этой цели. Но собирающая площадь этих устройств сравнительно мала, и это делает их непригодными для исследований, требующих высокой чувствительности, там, где критическим фактором является минимальный поток радиоизлучения, который система может зарегистрировать. Для решения сложной проблемы проведения обзоров удаленных «радиозвезд», связанной с локализацией слабых сигналов, группа Райла из Маллардской радиоастрономической обсерватории при Кавендишской лаборатории в Кембридже первой применила интерферометр с переменной базой для «апертурного синтеза». Интерферометр был использован так, чтобы не только достичь углового разрешения, свойственного очень большой одиночной антенне, но и приблизиться к ней по чувствительности. Метод, детально разработанный в работе Райла и Хьюиша в 1960 г. [56], можно понять, представив себе одиночную большую апертуру, которая разделена на отдельные элементарные площадки (рис. 6.15). Если можно было бы измерить волновые фронты, принимаемые этими площадками, с помощью размещаемого в них поочередно небольшого радиотелескопа соответствующей площади и сложить полученные сигналы точно по фазе, то результат был бы таким же, как и для всей апертуры. Фазы между отдельно измеряемыми сигналами можно определить из парных измерений на интерферометре с переменной базой. Как можно понять из рисунка, нет необходимости проводить измерения для всех пар, некоторые из них (например, CD) являются избыточными. Впервые этот тип синтеза был реализован в 1957 г. на волне 7,9 м с эффективным угловым разрешением 1° коллегой Райла Д.Х. Блитом.

Поскольку два элемента интерферометра в апертурном синтезе не обязательно должны иметь равные размеры, создавались различные системы, например большая неподвижная антенная решетка с малой поворотной параболической антенной.

В 1962 г. Райл, используя вращение Земли, положил начало методу «суперсинтеза». Принцип его показан на рис. 6.16, где две антенны интерферометра A и B разнесены на расстояние D в направлении В–З. База вместе с Землей совершает оборот каждые 24 ч и каждая антенна за один оборот проходит все положения на суточном кольце с радиусом D относительно другой антенны. При разных D полностью синтези-

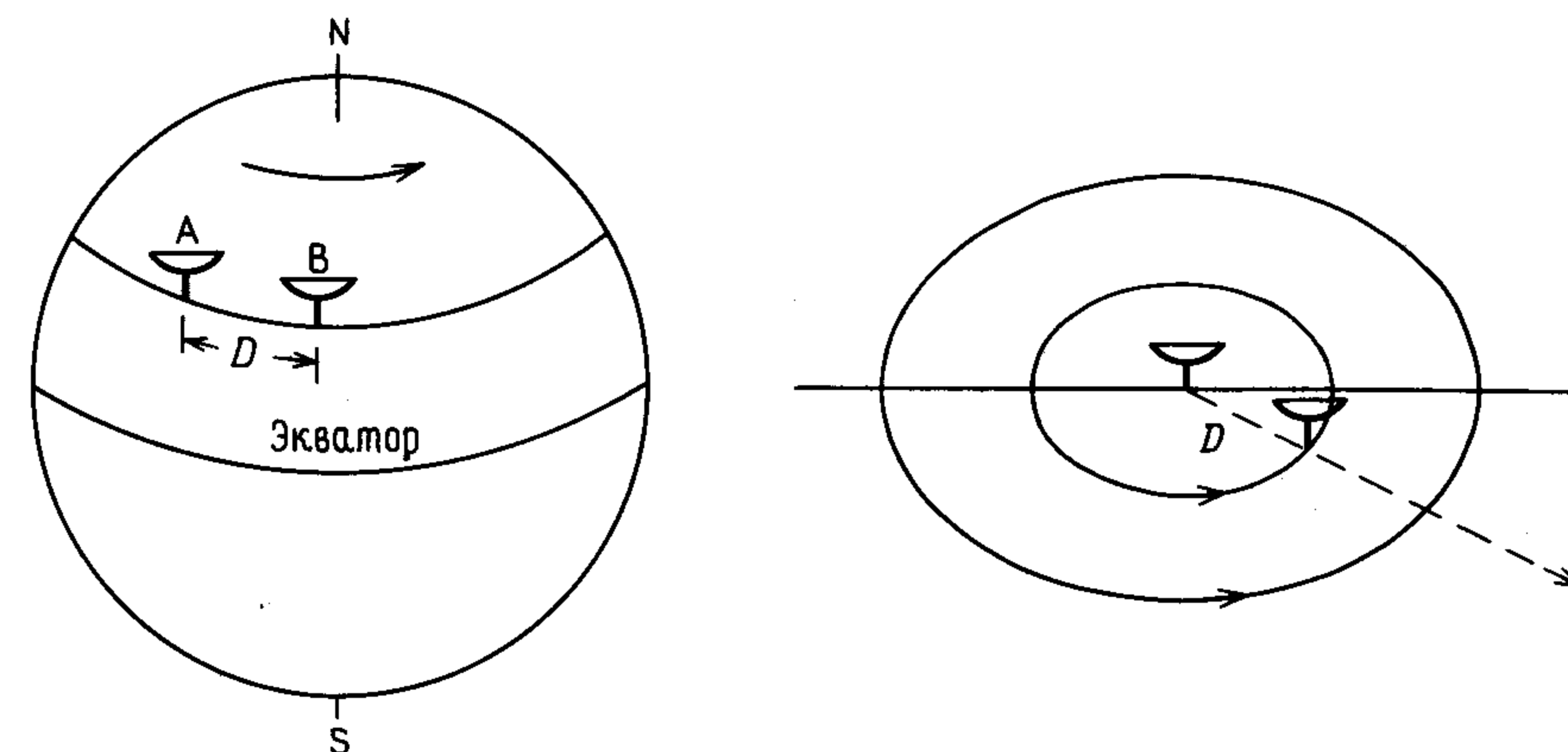


Рис. 6.16. Суперсинтез.

руется круглая площадка с радиусом, равным максимальному расстоянию между двумя антеннами. (Строго говоря, эти площадки являются эллиптическими, кроме как в направлении на Полярную звезду.)

В этих методах синтеза схема приемника, показанная упрощенно на рис. 6.17, позволяет при каждом значении D и угле проекции базы измерять косинусную $C(D)$ и синусную $S(D)$ компоненты на выходе интерферометрического приемника. Учитывая уравнения (6.12) и (6.13), можно показать, что в общем случае с использованием экспоненциального представления можно записать

$$C(D) + iS(D) = \int_{\phi} B(\phi) \exp(2\pi i D \phi / \lambda) d\phi. \quad (6.54)$$

Тем самым с помощью фурье-преобразования можно вычислить распределение яркости $B(\phi)$. Конечно, на практике это выглядит и теоретически, и технически более сложным, чем следует из приведенного уравнения. Дело не только в том, что задача является трехмерной, но также в том, что антенны имеют диаграмму, которая искажает отклик типа $\cos^2$, предполагаемый при получении указанного уравнения, и т.п. Наше изложение имеет целью передать только основные принципы.

Для получения углового разрешения $0,6''$ на волне 2 см использовался 5-км телескоп в Кембридже.

В 1974 г. Райл и Хьюиш совместно были удостоены Нобелевской премии по физике за свои пионерские исследования в радиоастрофизике. Райл за свои наблюдения и изобретения, в частности за метод апертурного синтеза, а Хьюиш за решающую роль, которую он сыграл в открытии пульсаров (пульсары были открыты с помощью антенны другого типа, крупной решетки из 2048 диполей, установленных на площади 18 км^2 , и специально построенной для изучения быстро мерцаю-

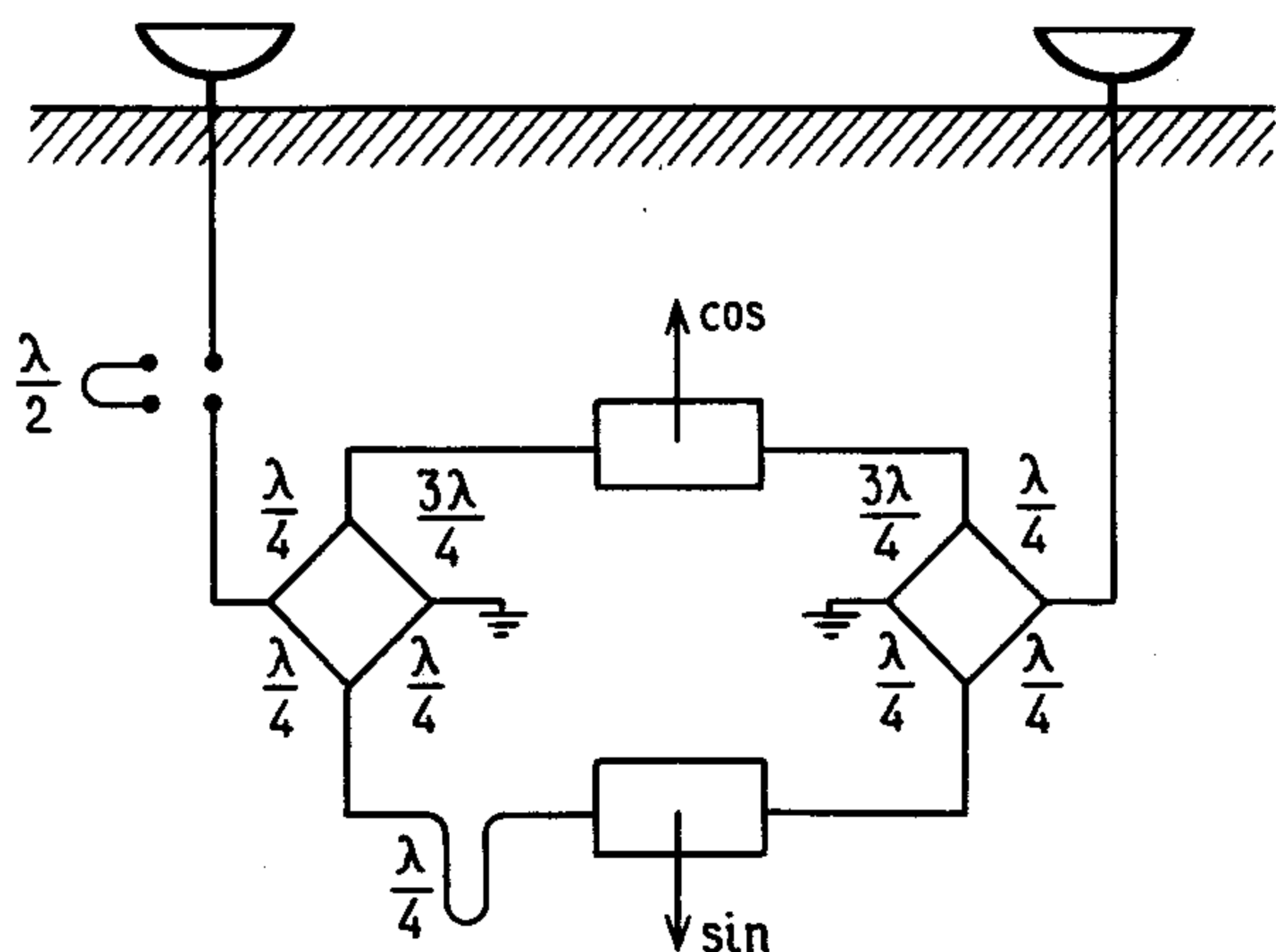


Рис. 6.17. Синусная и косинусная компоненты, обеспечивающие данные о фазе.

ших радиоисточников, для которых апертурный синтез непригоден по своей природе).

Для наблюдения в тончайших деталях радиогалактик и квазаров в апертурном синтезе наших дней используются базы, простирающиеся через континенты и океаны. Хотя радиоканалы оказались неприемлемыми для сведения сигналов из-за трудностей с фазовой стабильностью, изобретение точных атомных часов в конце 60-х годов означало, что сигналы можно записывать отдельно на магнитную ленту. После этого записи объединяются и воспроизводятся для получения требуемого интерферометрического выхода. Однако и в этом случае возникают фазовые флуктуации, и их действие эквивалентно тому, что положения максимумов лепестков становятся расплывчатыми. Способы преодоления этой трудности применимы в основном лишь в исследованиях источников очень малых размеров или более протяженных источников, включающих яркий точечный источник, который может служить в качестве фазового репера. Связанные с этим ограничения теперь в значительной мере преодолены с помощью нового метода. Его принцип был предложен Р.К. Дженисоном в радиообсерватории Джодрелл-Бэнк, однако с того времени он пребывал в забвении. В 1974 г. этот метод был возрожден и применен в интерферометрии со сверхдлинными базами А.Е. Роджерсом в США. Указанный метод замыкания фазы основан на том, что при измерении разности фаз между парами антенн, которые образуют замкнутый контур (например, три пары в треугольнике), и последующем сложении этих разностей ложные эффекты подавляются. При наличии антенн, расположенных по всему миру, стало реальным создание радиотелескопа размером с Землю, дающего «изображения», свободные от атмосферных и других эффектов и позволяющие достигнуть разрешения в $0,0001''$. Обрисовывая перспективные возможности, связанные с таким разрешением, можно заметить, что оно примерно соответствует угловому диаметру горячих, наблюдаемых в визуальной

области звезд. Их изображения, полученные даже с самым лучшим обычным оптическим телескопом, сильно замыкаются турбулентностью земной атмосферы.

Более полное описание этих методов и подробные результаты, которые были получены с их помощью (но не упоминались нами), можно найти в дополнительной литературе, указанной в конце книги.

Корреляционные принципы интерферометрических методов привели к другому ценному достижению в астрономии (радио и оптической), к которому мы должны теперь перейти. Речь пойдет об инструменте, на первый взгляд противоречащем условиям наблюдения интерференции: в нем складываются интенсивности, а не амплитуды.

6.6.2. Интерферометр интенсивностей

Принцип интерферометра интенсивностей принадлежит Р. Ханбери Брауну, который в 1949 г. рассматривал в Джодрелл-Бэнк две проблемы, с которыми столкнулись радиоастрономы: как освоить по-настоящему длинные базы и как преодолеть влияние атмосферной турбулентности. Напоминание о Джодрелл-Бэнк неизбежно ассоциируется с фотографией крупного полноповоротного радиотелескопа диаметром 250 футов (около 76 м). Он был построен для продолжения программы работ А.К. Б. Ловелла (ныне сэра Бернарда Ловелла) и его сотрудников по изучению радиолокационных отражений от метеоров и ливней космических лучей. Помимо указанных исследований эта крупная чашеобразная антенна (каких теперь, различных по форме и размерам, много в мире) сыграла важную роль в развитии различных аспектов интерферометрии со сверхдлинными базами.

В своей книге «Интерферометр интенсивностей» Ханбери Браун вспоминает о том, как в 1949 г. стоял вопрос об измерении угловых размеров двух самых известных внеземных радиоисточников, Лебедя А и Кассиопеи А.

«В то время мы знали лишь то, что их угловые размеры не превышают существенно несколько минут дуги, но не располагали никакими данными о том, насколько они могут быть малы. Если, как полагали некоторые, они оказались бы такими же малыми, как видимые звезды, то для метровых волн потребовалось бы разместить приемные станции на противоположных концах Земли. Можно ли создать радиоинтерферометр с базой, которая при необходимости могла бы простирается на десятки, сотни или даже тысячи километров? Непосредственная техническая трудность в реализации существующих решений была связана с обеспечением когерентности гетеродинов в двух удаленных точках, и я начал размышлять, действительно ли это условие необходимо. Может быть, принимаемые в двух точках волны удастся сравнивать каким-либо другим способом? Для примера я вообразил простой детектор, который демодулирует волны от источника и отображает их в виде обыч-

ного шума, который наблюдается на катодно-лучевом осциллографе. Если сфотографировать одновременно шум на двух станциях, будут ли две фотографии одинаковы? Этот вопрос непосредственно привел меня к идее корреляции флуктуаций интенсивности и принципу интерферометра интенсивности».

Таким образом, идея состояла в том, что если флуктуации интенсивности на двух близких антеннах коррелировали, то уменьшение корреляции (отсюда корреляционный интерферометр) с увеличением базы позволяло бы определять угловой размер источника (это был бы аналог метода Майкельсона, использующий интенсивности для измерения диаметров оптически видимых звезд). Тогда трудность, связанная с взаимной нестабильностью далеко разнесенных гетеродинов, была бы преодолена. (В то время не были разработаны атомные часы, которые сейчас используются в интерферометрии с длинными базами.)

Эта идея была успешно опробована в Джодрелл-Бэнк на интерферометре, работающем на частоте 125 мГц, для измерения углового размера Солнца [22]. Годом позже он был использован для измерения упоминавшихся двух радиоисточников, которые вызывали споры. По воле случая они оказались больше, чем ожидалось, и потребовалась база размером всего в несколько километров. Создалось впечатление, будто кузнечный молот был использован только для того, чтобы расколоть орех. Тем не менее Ханбери Браун и его коллеги обнаружили, что когда радиоисточники сильно мерцали из-за ионосферных неоднородностей, данные измерения корреляции незначительно искажались, поэтому была попутно решена и другая проблема. Теория интерферометра интенсивностей была сформулирована на математической основе Ханбери Брауном и Р.К. Твиссом в 1954 г. [23]. Хотя в радиоастрономии этот метод получил ограниченное применение из-за того, что в нем измеряемый сигнал должен быть больше по сравнению с уровнем шума в приемнике, в оптической астрономии он имел большой практический выход. Решение проблемы атмосферной турбулентности и явилось стимулом к исследованию вопроса о том, можно ли использовать интерферометр интенсивностей в оптической астрономии. Напомним, что атмосферная турбулентность¹⁾ была серьезной трудностью в методе Майкельсона для измерения диаметров звезд.

В 1956 г. вариант рассматриваемого инструмента (рис. 6.18) из двух параболических зеркал (M_1 и M_2) диаметром 1,56 м и с переменной базой до 14 м был использован впервые для измерения углового диаметра Сириуса. Каждое из зеркал в отдельности давало размытое

¹⁾ В случае оптической астрономии трудности, связанные с влиянием атмосферы, существенно уменьшаются при применении космического телескопа. Для зеркала диаметром 2,4 м на орбите высотой 500 км максимальное угловое разрешение будет составлять 0,1" по сравнению с наземным телескопом, для которого оно не может быть лучше 1".

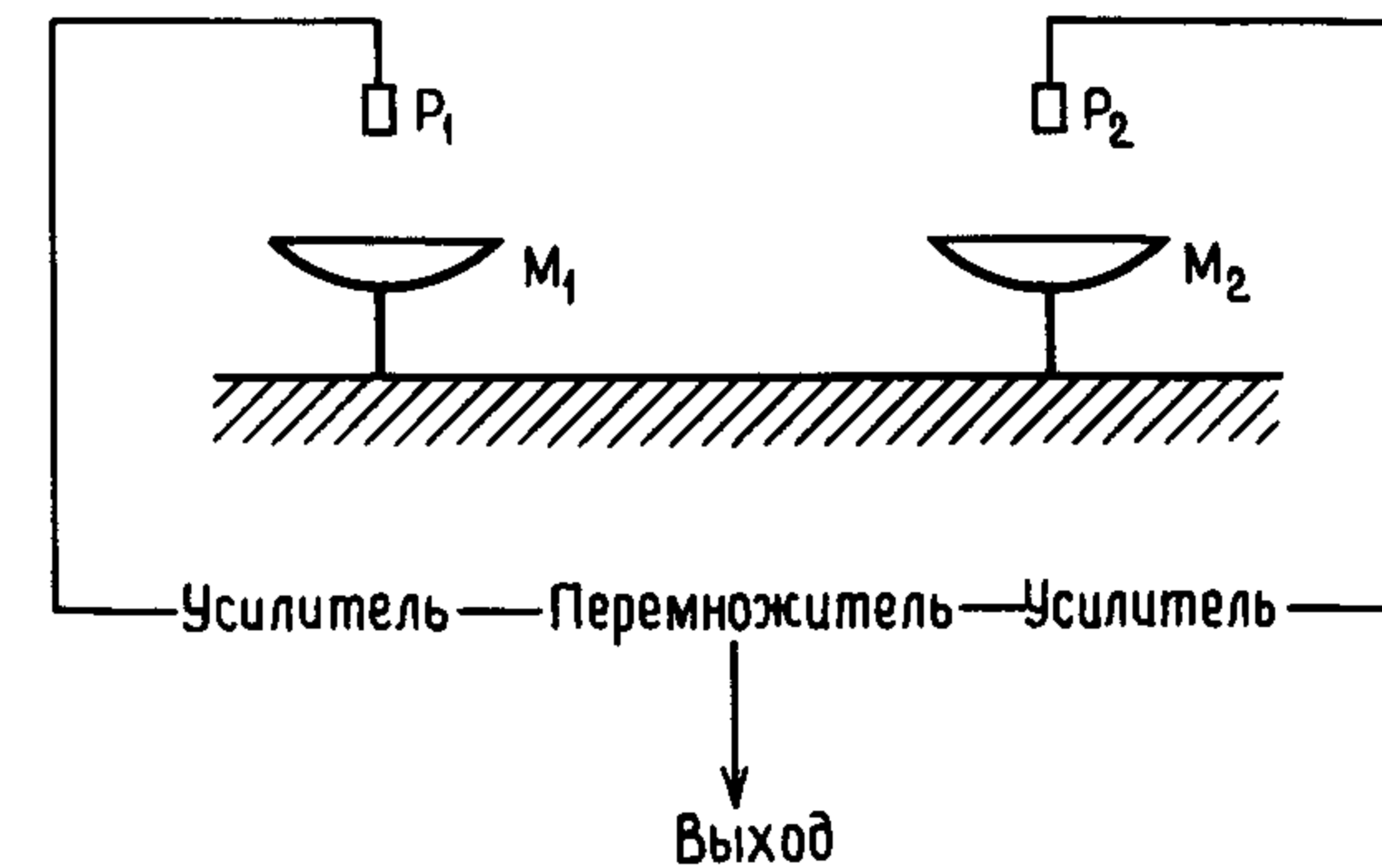


Рис. 6.18. Интерферометр интенсивностей (корреляционный).

изображение размером 8 мм, которое направлялось на катод фотоумножителя (P_1 и P_2). После усиления сигналы перемножались. Среднее за несколько часов значение произведения являлось непосредственной мерой кросс-корреляции между флуктуациями интенсивности света на двух зеркалах. Угловой диаметр Сириуса, вычисленный по падению корреляции с увеличением базы (т.е. падению видности в методе Майкельсона) оказался равным $0,0068'' \pm 0,0005''$ в хорошем согласии со значением $0,0063'' \pm 0,0006''$, предсказанным теоретической астрофизикой на совсем другой основе.

Теория оптического интерферометра интенсивности осложняется из-за квантовой природы фотоэлектрического эффекта [25].

Пусть в точках M_1 и M_2 на рис. 6.18 мгновенные интенсивности равны $I_1(t)$ и $I_2(t)$. Флуктуации $I_1(t)$ можно записать в виде $\Delta I_1(t) = I_1(t) - \langle I_1(t) \rangle$ и аналогично для $I_2(t)$. В предположении, что выходные токи фотоумножителей пропорциональны интенсивности падающего на них света, среднее по времени произведение этих токов пропорционально $\langle \Delta I_1 \Delta I_2 \rangle$. Статические расчеты показывают, что последнее выражение пропорционально квадрату модуля кросс-корреляции между M_1 и M_2 . Поэтому фурье-преобразование дает пространственное распределение яркости источника, при условии (которое выполняется при измерении диаметров звезд), что источник симметричен.

Сравнительно малочувствительный к качеству зеркал и нестабильностям базы и менее чем его предшественники подверженный атмосферным возмущениям этот метод стал важным средством оптической астрономии и активно использовался в самые последние годы Ханбери Брауном и его коллегами в Австралии.

Приложение А. Скалярное описание электромагнитных волн

А.1. ПРЕДСТАВЛЕНИЕ СКАЛЯРНОЙ ВОЛНЫ

Если точка Р на рис. А.1 вращается вокруг точки С с равномерной скоростью, то ее проекция Р' на ось Оу совершает простые гармонические колебания относительно О. Если Р' прикреплена к концу струны, как показано на рис. А.1,б, то передаваемое по струне волнообразное движение аналогично модели скалярной волны для света (без учета поляризации). Чтобы описать это движение аналитически, мы поступим следующим образом. Вначале заметим, что смещение точки Р' относительно О при колебаниях определяется выражением

$$\left. \begin{aligned} y &= A \cos \frac{2\pi t}{T} \\ &= A \cos 2\pi \nu t \\ &= A \cos \omega t \end{aligned} \right\} \quad (\text{А.01})$$

где A – амплитуда колебаний, T – период колебаний (с), ν – частота (колебаний/с), ω – угловая частота (радиан/с).

Поэтому выражение для волнового движения, распространяющегося вдоль струны, можно записать в виде:

$$y = A \cos \left[\omega t - \frac{2\pi x}{\lambda} \right]. \quad (\text{А.02})$$

За каждое колебание волна проходит расстояние в одну длину волны (λ), и потому скорость волны равна

$$v = \lambda/T,$$

откуда следует известное соотношение

$$v = \lambda \nu.$$

(Читатель может убедиться, что отрицательный знак в уравнении (А.02) относится к волне, распространяющейся в положительном направлении x и наоборот.)

Часто уравнение (А.02) записывают иначе в виде

$$y = A \cos (\omega t - kx), \quad (\text{А.03})$$

где $k = 2\pi/\lambda$ и поэтому $v = \omega/k$.

Величина k называется по-разному: *угловое волновое число, постоян-*

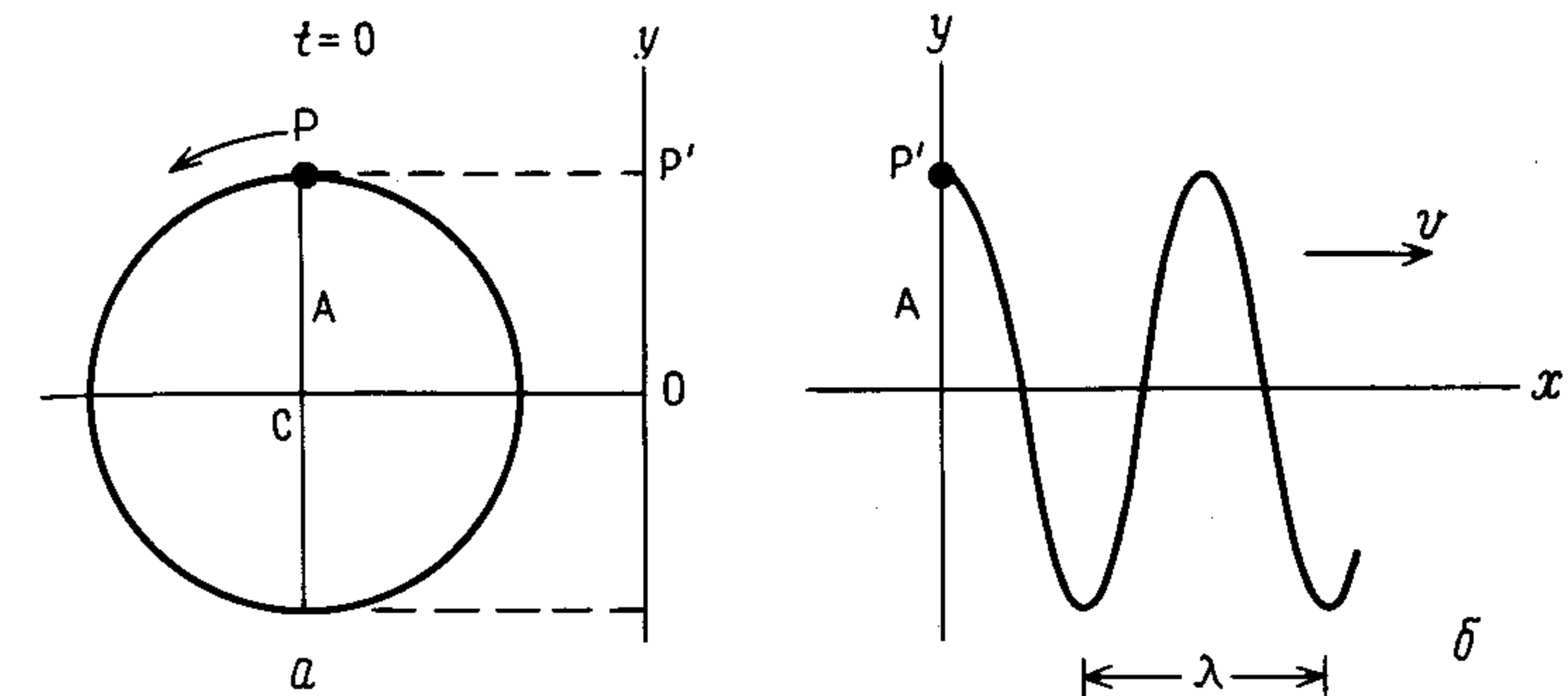


Рис. А.1.

ная длины волны, волновая постоянная и постоянная распространения. Часто эта величина используется просто для обозначения $1/\lambda$ и тогда зовется (спектроскопическим) волновым числом (более широко принятое обозначение σ использовалось в разд. 6.3.2).

До сих пор для рассматриваемой волны $y = A$ при $x = 0$ в момент $t = 0$. Аналогичная волна изображена на рис. А.2,б, но она задержана относительно только что рассмотренной на расстояние Δx . Поэтому ее уравнение записывается в виде

$$y = A \cos \left[\omega t - \frac{2\pi}{\lambda} (x + \Delta x) \right].$$

Из рис. А.2 очевидно, что $\Delta x/\lambda = \alpha/2\pi$, и поэтому

$$y = A \cos \left[\left(\omega t - \frac{2\pi x}{\lambda} \right) - \alpha \right] \quad (\text{А.04})$$

$$= A \cos (\omega t - kx - \alpha), \quad (\text{А.05})$$

где $\alpha = \frac{2\pi}{\lambda} \Delta x$ называется *фазой, фазовым углом или фазовой постоянной*.

В случае волны, описываемой уравнением (А.03), α называется *разностью фаз*, а Δx – *разностью пути*. Таким образом, мы получаем часто используемое соотношение

$$(\text{Разность фаз}) = \frac{2\pi}{\lambda} \times (\text{Разность пути}).$$

А.2. КОМБИНАЦИЯ ВОЛН

А.2.1. Принцип суперпозиции: интерференция

В этой книге часто приходится находить результирующую освещенность, когда несколько световых волн, различающихся только амплитудой и фазой, накладываются в некоторой точке, скажем на экране.

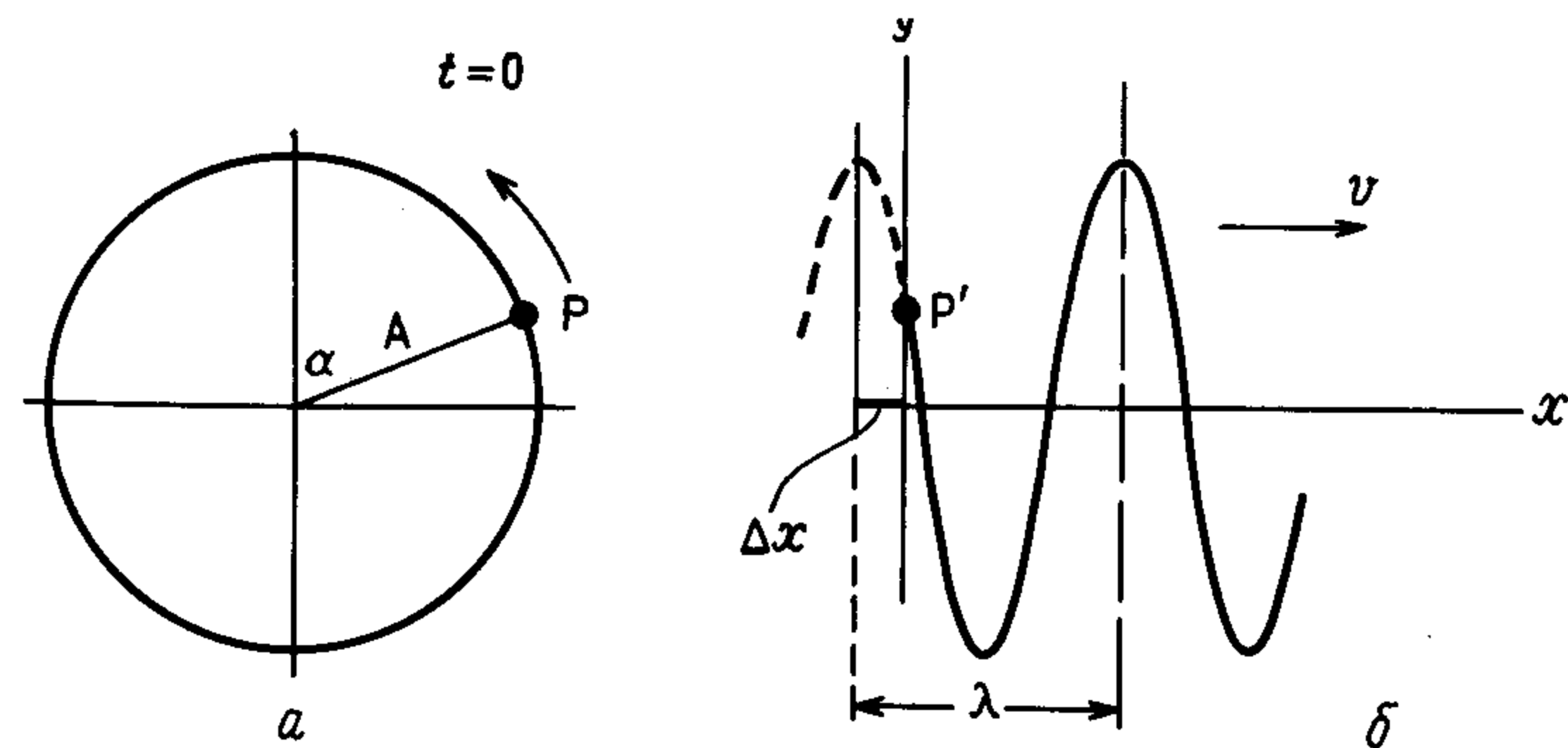


Рис. А.2.

В соответствии с принципом суперпозиции результирующее возмущение в этой точке представляет собой сумму отдельных возмущений.

Рассмотрим распространение двух волн

$$y_1 = A \cos(\omega t - kx),$$

$$y_2 = A \cos(\omega t - kx - \alpha),$$

где α — разность фаз между ними. Тогда

$$y_1 + y_2 = 2A \cos \frac{\alpha}{2} \cos \left(\omega t - kx - \frac{\alpha}{2} \right),$$

или, в общем виде,

$$Y = R \cos(\omega t - kx - \theta). \quad (\text{A.06})$$

Этот результат представляет собой точно такую же световую волну, как отдельные падающие волны, с неизменной частотой и скоростью распространения. Однако ее амплитуда не равна простой сумме амплитуд отдельных приходящих волн. Между ними имеется *интерференция*, степень которой зависит от α . Если разность фаз α между двумя приходящими волнами определяется выражением $\alpha = n 2\pi$ (где n — нуль или целое), что соответствует нулевой разности пути или разности пути в целое число длин волн, то амплитуда результирующей волны действительно равна $2A$. Однако если $\alpha = (n + 1/2)2\pi$, то две волны полностью не совпадают и результирующая амплитуда равна нулю. Разность фаз между этими двумя крайними случаями дает результат с промежуточными значениями амплитуд.

Аналогичный результат получается, если амплитуды двух падающих волн не равны, хотя здесь, конечно, не может быть полного подавления.

Заметим, что если указанные волны вновь разделить после прохождения ими места, где они интерферируют, то оказывается, что интерференция не оказала на них никакого влияния. Заметим также, что интерференция может возникать только между волнами, имеющими одну и ту же скорость.

Для удобства место расчета интерференции можно взять при $x = 0$.

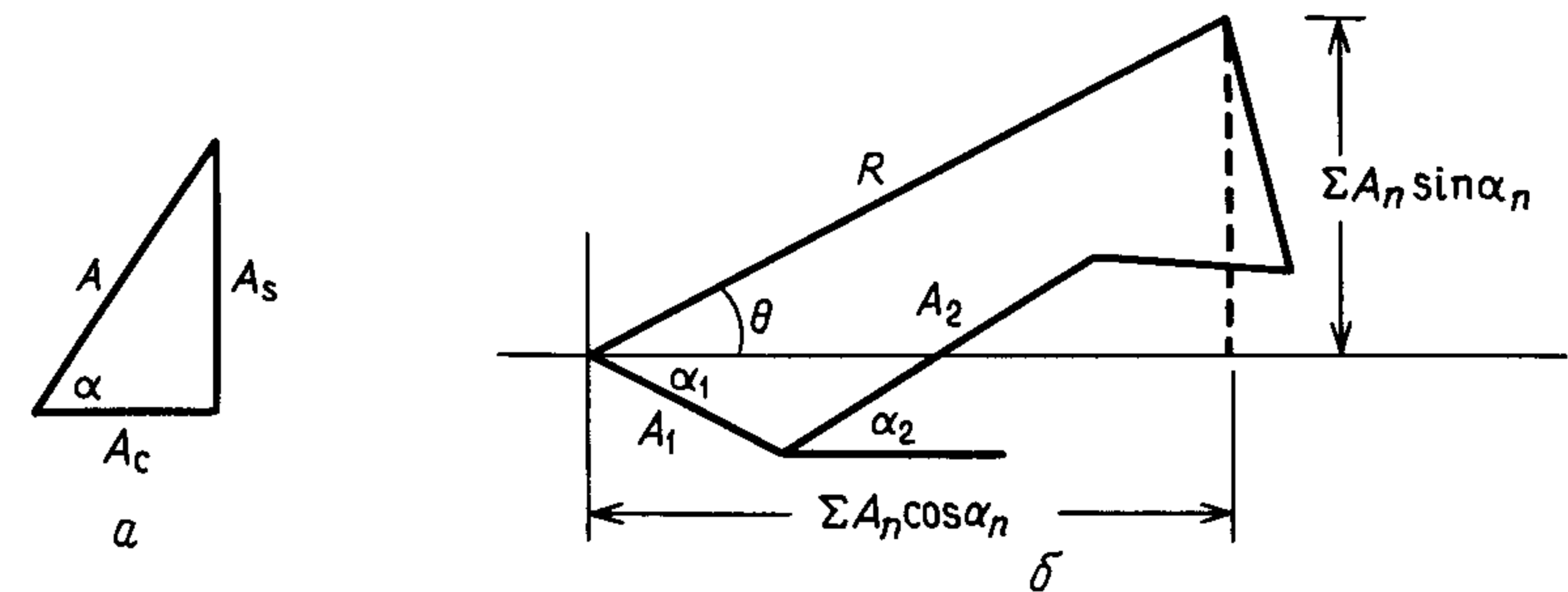


Рис. А.3.

Уравнение (A.04) принимает вид смещения

$$y = A \cos(\omega t - \alpha). \quad (\text{A.07})$$

Результирующее смещение от волн с различными амплитудами и фазами можно записать в виде суммы

$$Y = \sum_n A_n \cos(\omega t - \alpha_n), \quad (\text{A.08})$$

которая сводится к выражению

$$Y = R \cos(\omega t - \theta). \quad (\text{A.09})$$

Иногда суммирование легче всего выполнить с помощью векторной диаграммы, описанной в следующем разделе, а иногда алгебраически. Наиболее простым для алгебраического суммирования является использование экспоненциального обозначения (разд. 2.3).

А.2.2. Векторные диаграммы

Выражение (A.07) можно разложить

$$y = A \cos \alpha \cos \omega t + A \sin \alpha \sin \omega t = A_c \cos \omega t + A_s \sin \omega t,$$

где $A_c = A \cos \alpha$, $A_s = A \sin \alpha$. A_c и A_s тогда содержат информацию об амплитуде и фазе волны, как показано на рис. А.3,а. При этом

$$A = (A_c^2 + A_s^2)^{1/2},$$

$$\alpha = \arctg(A_s/A_c).$$

A_c и A_s напоминают компоненты двухмерного вектора, но поскольку углы представляют здесь собой фазу, а не направление, то для того, чтобы подчеркнуть различие, используется термин *фазовый вектор*. Поэтому суммирование уравнения (A.08) можно выполнить графически, и рис. А.3,б показывает, что результирующая амплитуда [R в уравнении (A.09)] представляется выражением

$$R = [(\sum A_n \sin \alpha_n)^2 + (\sum A_n \cos \alpha_n)^2]^{1/2}, \quad (\text{A.10})$$

а фаза

$$\theta = \arctg \left[\frac{\sum A_n \sin \alpha_n}{\sum A_n \cos \alpha_n} \right].$$

А.2.3. Экспоненциальное представление

Другой общеизвестный метод основан на выражении

$$y = A \cos(\omega t - \alpha)$$

в виде

$$y = \Re A \exp[i(\omega t - \alpha)], \quad (\text{A.11})$$

где $\Re$ указывает на то, что имеется в виду только действительная часть выражения; обычно для краткости $\Re$ опускается. Величина $Ae^{-i\alpha}$ называется комплексной амплитудой.

Сумма [уравнение (A.08)]

$$Y = \sum_n A_n \cos(\omega t - \alpha_n)$$

переписывается в виде

$$Y = \sum A_n \exp[i(\omega t - \alpha_n)],$$

где $\Re$ подразумевается.

Как и раньше, результат суммирования имеет вид

$$Y = R \cos(\omega t - \theta),$$

поэтому

$$R \exp[i(\omega t - \theta)] = \sum_n A_n \exp[i(\omega t - \alpha_n)],$$

т.е.

$$R \exp(-i\theta) = \sum_n A_n \exp(-i\alpha_n) = Z. \quad (\text{A.12})$$

Таким образом, комплексная амплитуда $Re^{-i\theta}$ суммы равна сумме комплексных амплитуд.

Суммирование выполняется алгебраически, и тогда умножение полученного выражения на соответствующее комплексно сопряженное дает R , так как

$$R^2 = ZZ^* = |Z|^2.$$

Этот метод напоминает метод фазовых векторов, но с суммированием в комплексной плоскости. Таким образом, разложение уравнения (A.12)

$$R \cos - i R \sin \theta = \sum A_n \cos \alpha_n - i \sum A_n \sin \alpha_n$$

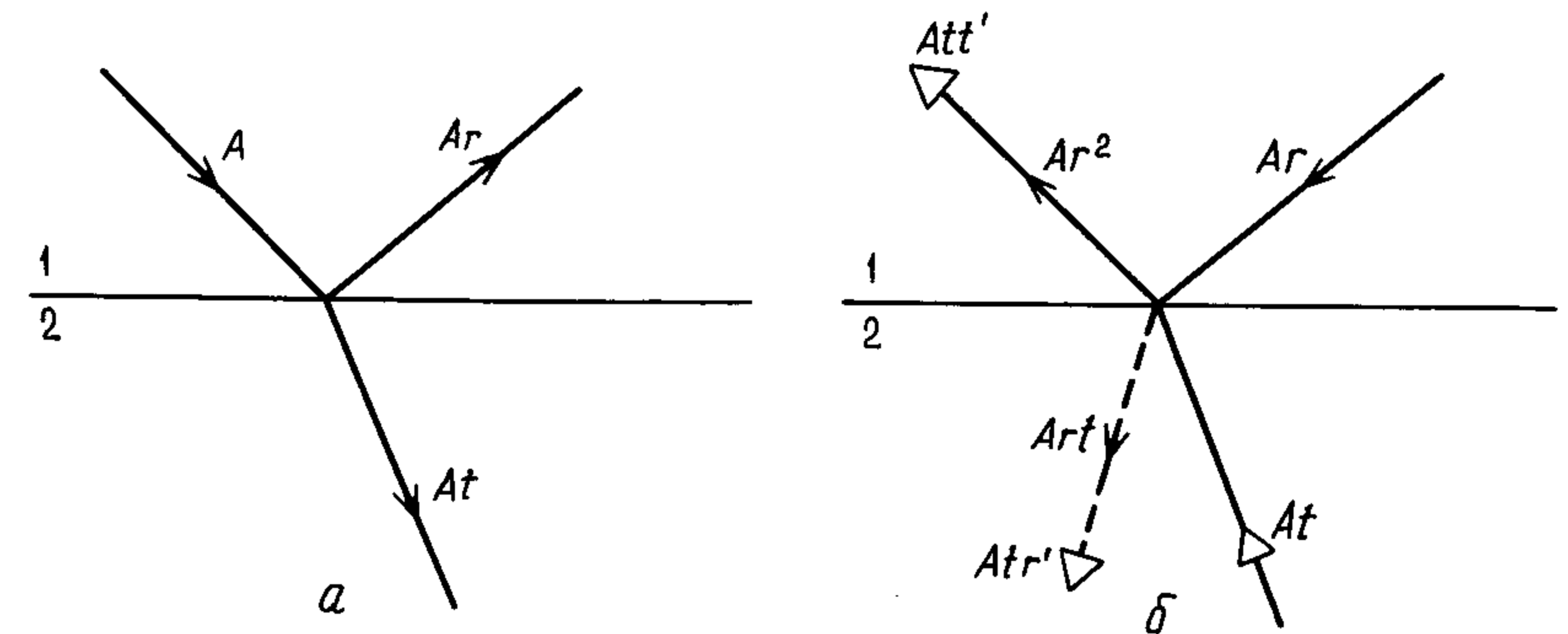


Рис. Б.1.

и приравнивание действительной и мнимой частей дает

$$R \cos \theta = \sum A_n \cos \alpha_n,$$

$$R \sin \theta = \sum A_n \sin \alpha_n$$

и, следовательно,

$$R = [(\sum A_n \cos \alpha_n)^2 + (\sum A_n \sin \alpha_n)^2]^{1/2},$$

$$\theta = \arctg \left[\frac{\sum A_n \sin \alpha_n}{\sum A_n \cos \alpha_n} \right],$$

как в уравнении (A.10).

В связи с интерпретацией, согласно которой векторная диаграмма может рассматриваться как сложение комплексных амплитуд в комплексной плоскости, необходимо отметить, что вращение фазового вектора на угол α соответствует умножению комплексной амплитуды на $e^{i\alpha}$. Именно потому чаще всего и используется экспоненциальное обозначение, поскольку последний тип операции зачастую является более легким, чем приведение тригонометрических формул.

Приложение Б. Представление Стокса для отражения и преломления

Классический подход выглядит следующим образом. На рис. Б.1,а луч света с амплитудой A дает отраженный луч с амплитудой Ar и преломленный луч с амплитудой At на линии раздела двух диэлектрических сред 1 и 2 с различными показателями преломления. Величины r и t определяют долю отраженной и преломленной (пропущенной) амплитуд соответственно.

Из принципа обратимости световых лучей следует, что если обратить отраженный и преломленный лучи, то они должны дать падающий луч с амплитудой A . Это схематично показано на рис. Б.1,б. Обра-

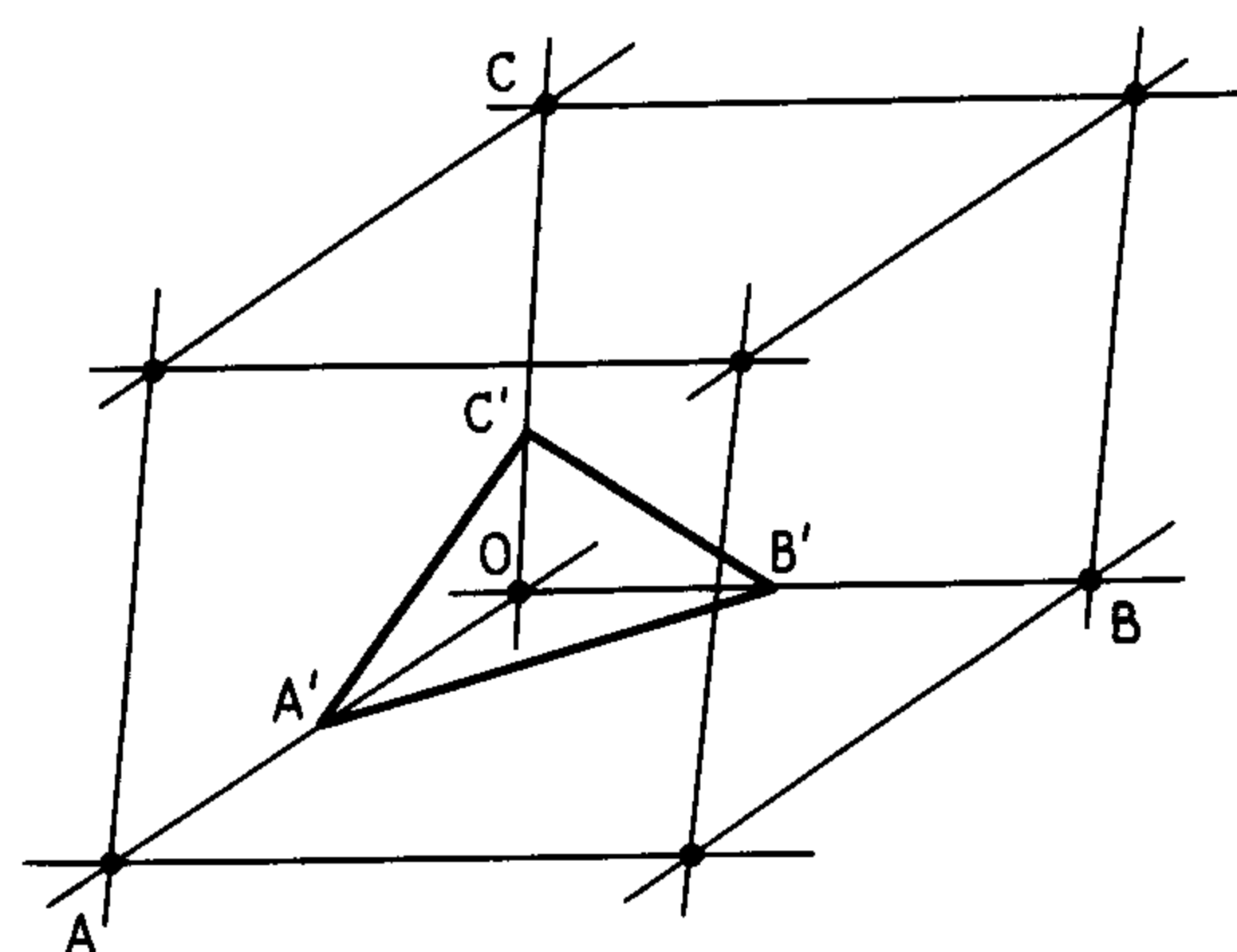


Рис. В.1.

щение луча с амплитудой Ar дает отражение назад в первой среде с амплитудой Ar^2 и преломленный луч с амплитудой Art . Аналогично обращение луча с амплитудой At дает отражение назад во вторую среду с амплитудой Atr' и преломленный в первую среду луч с амплитудой Att' , где r' и t' — доли для событий с переходом 2–1 в противоположность переходу 1–2.

Поэтому для воспроизведения рис. В.1,а при обращении мы должны приравнять следующие выражения:

$$Att' + Ar^2 = A,$$

т.е.

$$tt' = 1 - r^2,$$

и

$$Art + Atr' = 0,$$

т.е. $r = -r'$.

Различные знаки амплитуд в этом последнем равенстве указывают на изменение фазы на π при отражении в процессах 1/2 или 2/1. Опыт показывает, что обычно это происходит, когда свет падает на границу с той стороны, на которой скорость выше (меньше показатель преломления).

Приложение В. Дифракция рентгеновских лучей на кристаллах. Эквивалентность условия Лауэ и концепции отражения Брэгга

Прямоугольный параллелепипед на рис. В.1 изображает ячейку кристаллической решетки, в которой каждый узел решетки представлен идентичной группой атомов. В разд. 2.7 отмечалось, что дифракционные рентгеновские максимумы возникают в данном направлении от кри-

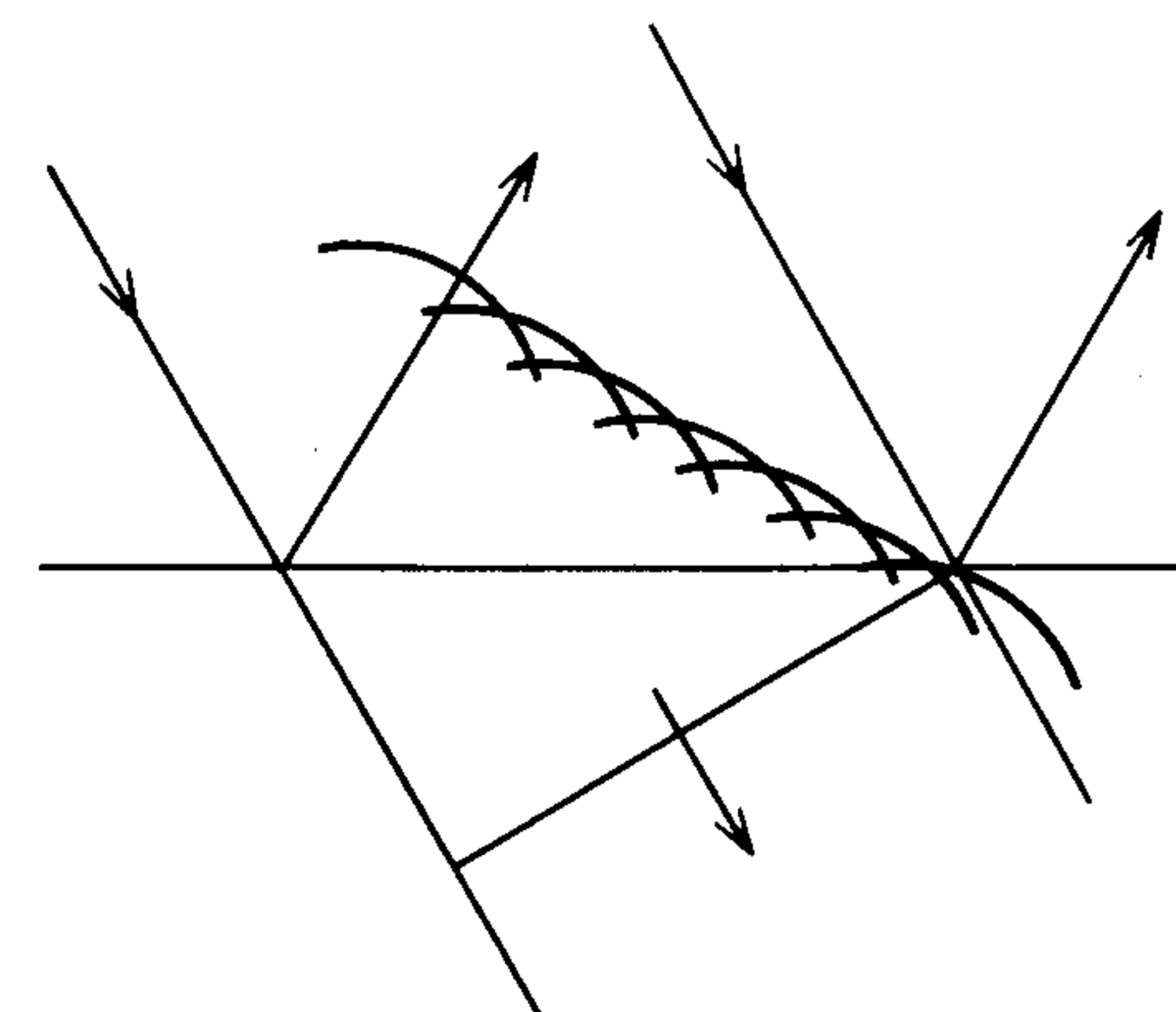


Рис. В.2. Отражение волнового фронта рассеивающими центрами, лежащими в одной плоскости.

сталла с такой решеткой, если рентгеновские лучи, рассеянные в этом направлении группами атомов, связанных с каждым из узлов решетки A , B и C , совпадают по фазе с излучением, рассеянным группой, связанной с узлом O . Условия Лауэ [уравнение (2.18)] формулируют это требование, а именно

$$\left. \begin{aligned} (s - s_0) \cdot a h \lambda \\ (s - s_0) \cdot b k \lambda \\ (s - s_0) \cdot c l \lambda \end{aligned} \right\} , \quad (B.01)$$

где s_0 , s — единичные векторы, задающие направления падающих и рассеянных рентгеновских лучей; a , b , c — векторы смещения; h , k , l — целое или нуль, а λ — длина волны рентгеновских лучей.

Объясняя то же самое явление отражением, Брэгг отмечал, что требование совпадения по фазе рассеянных в A , B и C лучей аналогично оптическому отражению в зеркале с точками A , B , C на его поверхности: при любом угле падения отражение происходит под углом, равным углу падения (рис. В.2). Как объяснял Брэгг, эта связь отраженной и падающей волн приводит к тому, что волны, рассеянные всеми точками в двух пространственных направлениях относительно плоскости решетки, совпадают одна с другой по фазе. Однако в противоположность отражению в оптике, при падении рентгеновских лучей на плоскость кристаллической решетки амплитуда отраженного пучка составляет очень малую долю от амплитуды падающего пучка. Большая часть излучения проходит через кристалл. Кроме того, лучи, отраженные от последующих плоскостей решетки, параллельных первой, в общем случае не будут совпадать по фазе один с другим. Усиление, однако, можно получить путем подбора угла падения. Как показано на рис. В.3, для этого требуется, чтобы разность пути $X'Y'Z - XYZ$ равнялась целому числу длин волн. Поскольку $Y'Y = Y'W$, то это эквивалентно требованию

$$VW = n\lambda,$$

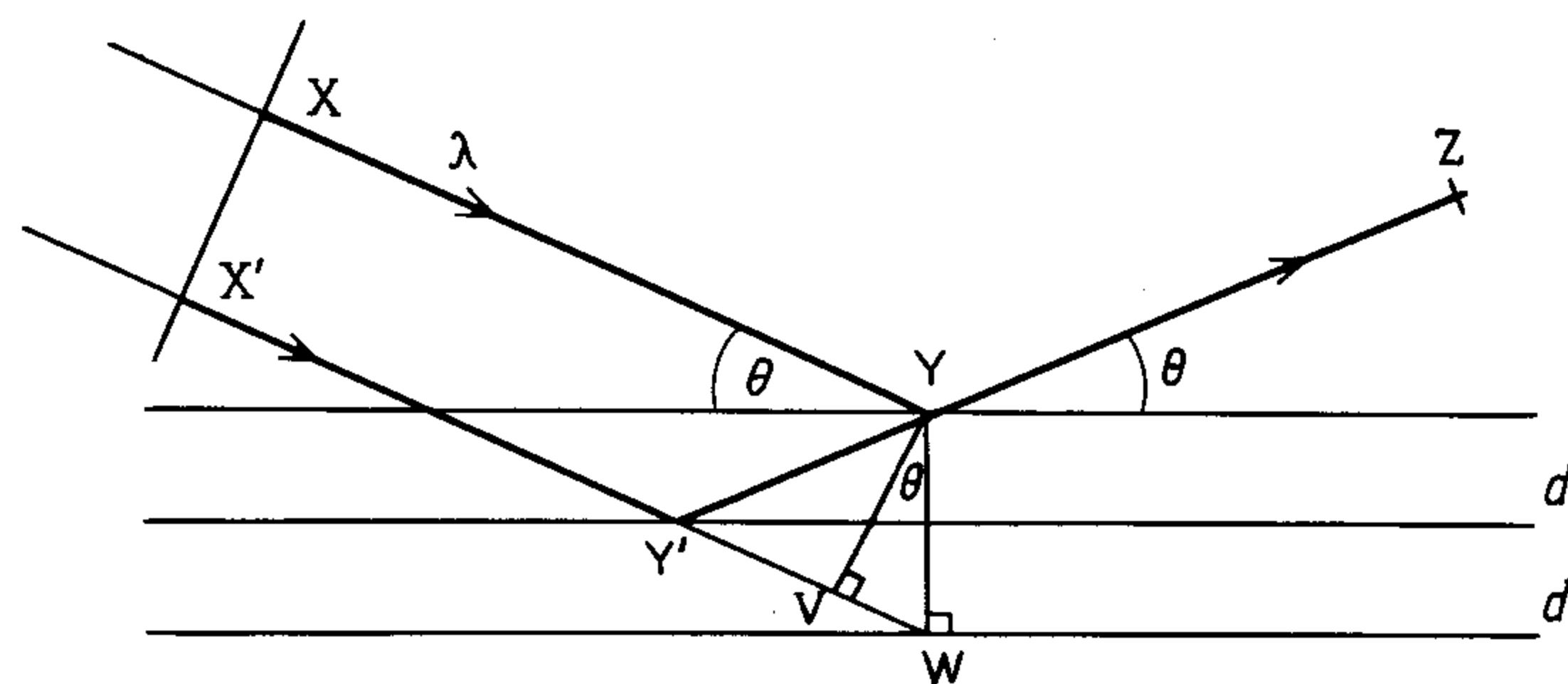


Рис. В.3. Отражение волнового пакета последовательными плоскостями кристаллической решетки, разделенными расстоянием d .

т.е.

$$2d \sin \theta = n\lambda, \quad (\text{В.02})$$

где d – расстояние между рассматриваемыми плоскостями решетки, а целое число n – «порядок» отражения.

Это знакомое нам уравнение Брэгга, а особые значения углов скольжения θ называют «углами Брэгга». При выполнении условия, выраженного указанным уравнением, рентгеновские волны от всех узлов кристаллической решетки усиливают друг друга; при других углах падения интерференция приводит к снижению интенсивности. Как со свойственной ему проницательностью отмечал Брэгг, это уравнение «представляет собой разновидность знакомого соотношения в оптике, которое определяет цвета при отражении от тонких пленок» (Брэгг, 1975).

Для установления эквивалентности интерпретаций Лауэ и Брэгга и нахождения отражающих плоскостей, которые соответствуют конкретным значениям h, k, l в уравнении В.01, мы поступим следующим образом.

Построим на рис. В.1 плоскость $A'B'C'$, такую, что $OA' = OA/h$, $OB' = OB/k$, $OC = OC/l$, где h, k, l – целые величины в уравнениях Лауэ. Так как h, k, l – целые, то отсюда следует, что A', B', C' лежат в плоскостях решетки, которые проходят через узлы решетки повсюду в той же самой структуре (для которой мы изобразили только одну единичную ячейку). Поскольку при выводе уравнений Лауэ отмечалось, что для получения максимума необходимо, чтобы излучение, рассеянное в узлах A, B, C , опережало (или отставало) излучение от узла O на h, k, l длин волн, то излучение от A', B', C' должно отличаться только на одну длину волны.

Более удачный вид $A'B'C'$ для очередного шага рассуждений показан на рис. В.4,а, где показана также часть следующей плоскости решетки,

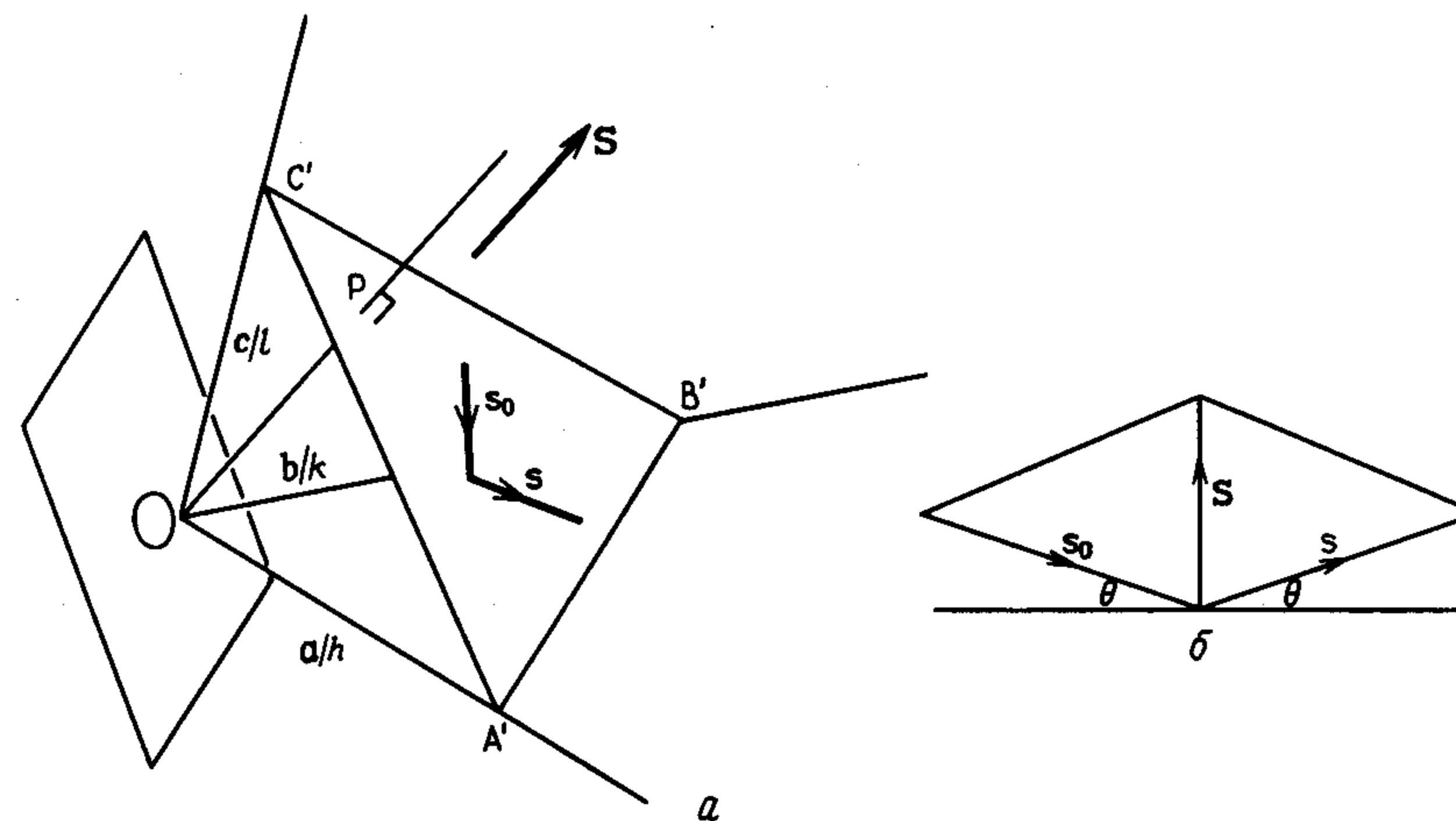


Рис. В.4.

параллельной $A'B'C'$ и проходящей через O . Теперь можно снова обратиться к уравнениям Лауэ.

При подстановке $s - s_0 = S$ уравнение (В.01) принимает вид

$$\left. \begin{aligned} \frac{S}{\lambda} \cdot \frac{a}{h} &= 1, \\ \frac{S}{\lambda} \cdot \frac{b}{k} &= 1, \\ \frac{S}{\lambda} \cdot \frac{c}{l} &= 1. \end{aligned} \right\} \quad (\text{В.03})$$

Вычитание второго из первого дает

$$\frac{S}{\lambda} \cdot \left(\frac{a}{h} - \frac{b}{k} \right) = 0,$$

т.е. направление S перпендикулярно $\overrightarrow{B'A'}$.

Вычитание других пар в уравнении (В.03) показывает, что S также перпендикулярно $\overrightarrow{A'C'}$ и $\overrightarrow{C'B'}$. Поэтому S перпендикулярно плоскости $A'B'C'$. На рис. В.4,б мы видим теперь, что S пересекает s и s_0 . Поэтому отсюда следует, что s и s_0 имеют одинаковый наклон к $A'B'C'$, в соответствии с представлением об отражении от плоскости решетки $A'B'C'$.

Поскольку узлы в плоскости $A'B'C'$ дают при рассеянии разность пути в одну длину волны по сравнению с узлами в параллельной плоскости, проходящей через O , то мы также получаем картину Брэгга с усилением излучения, рассеянного на последовательных плоскостях решетки. Если s_0 и s расположены под равными углами скольжения

к $A'B'C'$, которые Брэгг обозначил θ , то мы можем показать, что расстояние OP между этими плоскостями равно d в уравнении Брэгга при $n = 1$. Из рисунка имеем

$$OP = \frac{a}{h} \cdot \frac{S}{|S|}$$

С учетом первого из уравнений Лауэ это дает

$$OP = \frac{\lambda}{|S|}$$

Согласно рис. В.4,б, $|S| = 2 \sin \theta$, и поэтому

$$OP = \frac{\lambda}{2 \sin \theta}$$

равно d в уравнении Брэгга при $n = 1$.

Таким образом, мы показали следующее: во-первых, рассеянные рентгеновские лучи, дающие в соответствии с уравнениями Лауэ дифракционный максимум, испытывают также и отражение, аналогичное оптическому, причем отражаются они в плоскостях структуры решетки, определяемых величинами h, k, l в уравнениях Лауэ; во-вторых, расстояния между этими плоскостями решетки и угол отражения на них определяются уравнением Брэгга.

Приложение Г. Электромагнитный спектр

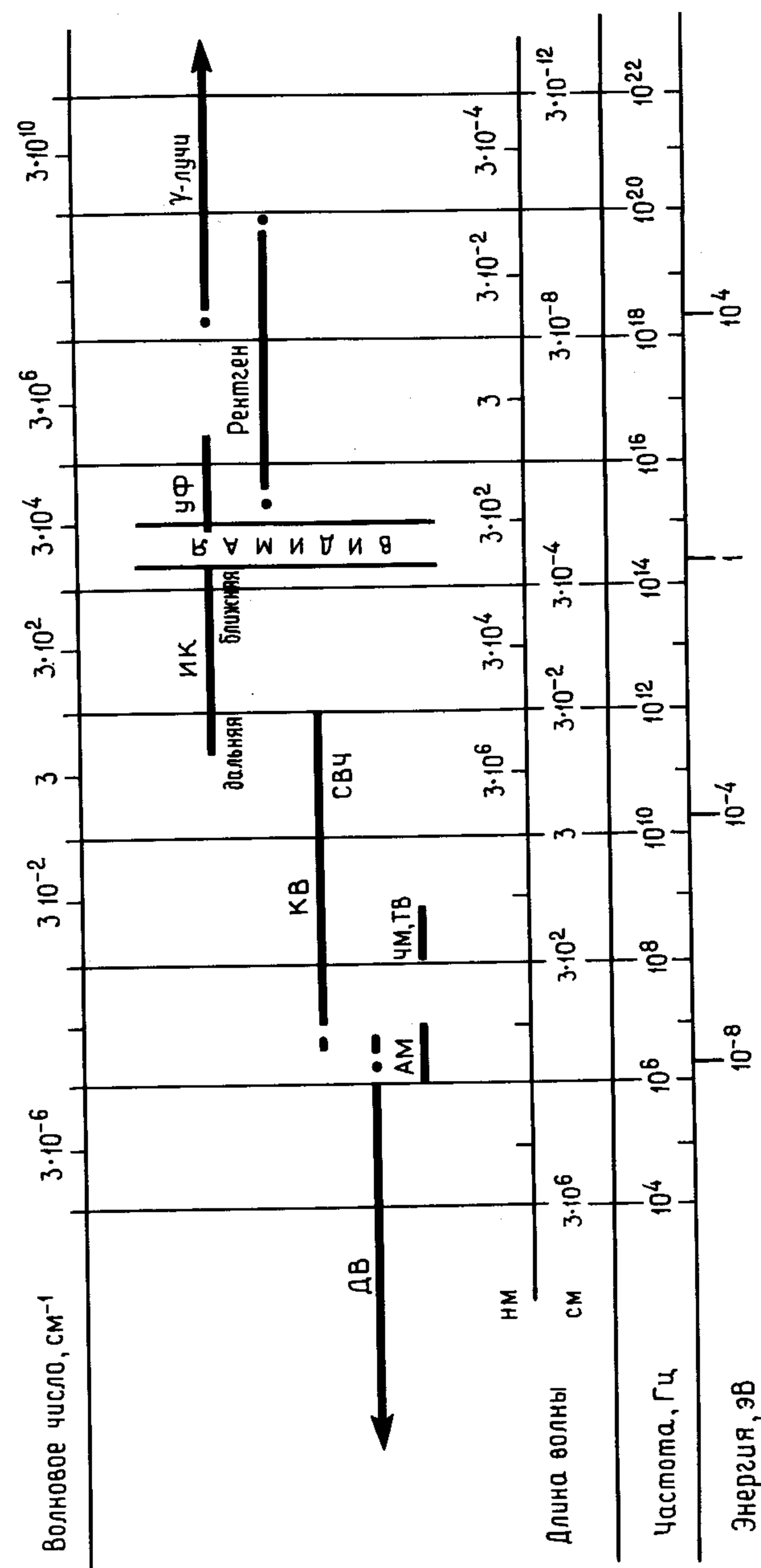


Рис Г 1 Примерные границы основных областей электромагнитного спектра (Скорость принята равной $3 \cdot 10^{10}$ см/с; 1 эВ соответствует частоте $2,42 \cdot 10^{14}$ Гц, длине волны 1240 нм)

Приложение Д. Некоторые полезные формулы

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B,$$

$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B,$$

$$\sin 2\theta = 2 \sin \theta \cos \theta,$$

$$\cos 2\theta = \cos^2 \theta - \sin^2 \theta = 2 \cos^2 \theta - 1 = 1 - 2 \sin^2 \theta,$$

$$\sin A + \sin B = 2 \sin \frac{1}{2}(A + B) \cos \frac{1}{2}(A - B),$$

$$\sin A - \sin B = 2 \cos \frac{1}{2}(A + B) \sin \frac{1}{2}(A - B),$$

$$\cos A + \cos B = 2 \cos \frac{1}{2}(A + B) \cos \frac{1}{2}(A - B),$$

$$\cos A - \cos B = -2 \sin \frac{1}{2}(A + B) \sin \frac{1}{2}(A - B),$$

$$e^{i\theta} = \cos \theta + i \sin \theta \text{ (формула Эйлера)}^1,$$

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2},$$

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

¹⁾ Охарактеризована Ричардом Фейнманом как «восхитительная драгоценность, самая замечательная формула математики».

Литература

1. Abbe E. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. Archiv. f. Mikroskop. Anat., **9**, 413–468. (Collected Works (1904), Vol. 1), 1873.
2. Barer R. Phase contrast microscopy. Research, **8**, 341–343, 1955.
3. Blythe J.H. A new type of pencil beam aerial for radio astronomy. Roy. Ast. Soc. Monthly Notices, **117**, 644–651, 1957.
4. Bolton J.G. and Stanley G.J. Variable source of radio frequency radiation in the constellation of Cygnus. Nature, **161**, 312–313, 1948.
5. Bragg W.L. The determination of parameters in crystal structures by means of Fourier analysis. Proc. Roy. Soc. A, **123**, 537–559, 1929.
6. Bragg W.L. An optical method of representing the results of X-ray analysis. Z. f. Krist., **70**, 475–492, 1929.
7. Bragg W.L. The crystalline state, Vol. I. Bell, London, 1933.
8. Bragg W.L. A new type of 'X-ray microscope'. Nature, **143**, 678, 1939.
9. Bragg W.L. The X-ray microscope. Nature, **149**, 470–471, 1942.
10. Bragg W.L. Lightning calculations with light. Nature, **154**, 69–72, 1944.
11. Buerger M.J. Generalised microscopy and the two-wavelength microscope. J. Appl. Phys., **21**, 909–917, 1950.
12. Connes P. Advances in Fourier spectroscopy. Phys. Bull., **22**, 26–28, 1971.
13. DeRosier D.J. and Klug A. Reconstruction of three-dimensional structures from electron micrographs. Nature, **217**, 130–134, 1968.
14. DeRosier D.J. and Klug A. Structure of the tubular variants of the head of bacteriophage T4 (polyheads). J. Mol. Biol., **65**, 469–488, 1972.
15. Elias P. Optics and communication theory. Jour. Opt. Soc. Amer., **43**, 229–232, 1953.
16. Elias P., Grey D.S. and Robinson D.Z. Fourier treatment of optical processes. Jour. Opt. Soc. Amer., **42**, 127–134, 1952.
17. Fellgett P. PhD thesis, University of Cambridge, UK, 1951.
18. Fellgett P. A propos de la théorie du spectrometre interférentiel multiplex. J. Phys. Radium, **19**, 187, 1958.
19. Fizeau H. Ann. Chim. Phys., **66**, 429, 1862.
20. Fizeau H. C.R. Acad. Sci. Paris, **66**, 934, 1868.
21. Gabor D. A new microscope principle. Nature, **161**, 777–778, 1948.
22. Hanbury Brown R. and Jennison, R.C. Unpublished, 1950.
23. Hanbury Brown R. and Twiss, R.Q. A new type of interferometer for use in radio astronomy. Phil. Mag., **45**, 663–682, 1954.
24. Hanbury Brown R. and Twiss, R.Q. A test of a new type of stellar interferometer on Sirius. Nature, **178**, 1046–1048, 1956.
25. Hanbury Brown R. and Twiss R.Q. Interferometry of the intensity fluctuations

- in light—III. Applications to astronomy. *Proc. Roy. Soc. A*, **248**, 199–221, 1958.
26. *Hanson A.W. and Lipson H.* A simplified fly's eye procedure. *Acta Cryst.*, **5**, 145, 1952.
 27. *Harburn G., Taylor C.A. and Welberry T.R.* Atlas of optical transforms. Bell, London, 1975.
 28. *Hopkins H.H.* The concept of partial coherence in optics. *Proc. Roy. Soc. A*, **208**, 263–277, 1951.
 29. *Jacquinot P.* New developments in interference spectroscopy. *Rep. Progr. Phys.*, **23**, 267–312, 1960.
 30. *Jacquinot P. and Dufour C.J.* Conditions optiques d'emploi des cellules photo-électriques dans les spectrographes et les interféromètres. *J. Rech. C.N.R.S.*, **6**, 91–103, 1948.
 31. *Klug A. and Berger J.E.* An optical method for the analysis of periodicities in electron micrographs, and some observations on the mechanism of negative staining. *J. Mol. Biol.*, **10**, 565–569, 1964.
 32. *Klug A. and Finch J.T.* Structure of viruses of the papilloma-polyoma type. *J. Mol. Biol.*, **31**, 1–12, 1968.
 33. *Leith E.N. and Upatnieks J.* Reconstructed wavefronts and communication theory. *J. Opt. Soc. Amer.*, **52**, 1123–1130, 1962.
 34. *Leith E.N. and Upatnieks J.* Wavefront reconstruction with continuous-tone objects. *J. Opt. Soc. Amer.*, **53**, 1377–1381, 1963.
 35. *Leith E.N. and Upatnieks J.* Wavefront reconstruction with diffused illumination and 3-dimensional objects. *J. Opt. Soc. Amer.*, **54**, 1295–1301.
 36. *Leith E.N. and Upatnieks J.* Wavefront reconstruction photography. *Physics Today*, **18**, No. 8, 26–32, 1965.
 37. *Michelson A.A.* On the application of interference methods to astronomical measurements. *Phil. Mag.*, Ser 5, **30**, 1–21, 1890.
 38. *Michelson A.A.* Measurement by light-waves. *Am. Jnl. Sci.*, **39**, 115–121, 1890.
 39. *Michelson A.A.* Visibility of interference-fringes in the focus of a telescope. *Phil. Mag.*, Ser. 5, **31**, 256–259, 1891.
 40. *Michelson A.A.* On the application of interference-methods to spectroscopic measurements—I. *Phil. Mag.*, Ser 5, **31**, 338–346, 1891.
 41. *Michelson A.A.* On the application of interference-methods to spectroscopic measurements—II. *Phil. Mag.*, Ser 5, **34**, 280–299.
 42. *Michelson A.A.* On the application of interference methods to astronomical measurements. *Astrophys. Jnl*, **51**, 257–262, 1920.
 43. *Michelson A.A.* Studies in optics. University of Chicago Press, Chicago (Phoenix Science Series edn. 1962), 1927.
 44. *Mills B.Y. and Little A.G.* A high resolution aerial system of a new type. *Australian J. Phys.*, **6**, 272–278, 1953.
 45. *Padawer J.* The Nomarski interference-contrast microscope. An experimental basis for image interpretation. *J. Roy. Mic. Soc.*, **88**, 305–349.
 46. *Pawsey J.L., Payne-Scott R. and McCready L.L.* Radiofrequency energy from the Sun. *Nature*, **157**, 158–159, 1946.
 47. *Porter A.B.* On the diffraction theory of microscopic vision. *Phil. Mag.*, Ser 6, **11**, 154–166, 1906.
 48. *Pykett I.L.* NMR imaging in medicine. *Sci. Amer.*, **246**, 54–64, 1982.
 49. *Rayleigh Lord.* On the manufacture and theory of diffraction-gratings. *Phil. Mag.*, **47**, 81–93, 193–205 (Scientific Papers, Vol. I, 199–221, CUP, Cambridge), 1874.

50. *Rayleigh Lord.* On the interference bands of approximately homogeneous light: in a letter to Prof. A. Michelson. *Phil. Mag. Ser 5*, **34**, 407–411 (Scientific Papers, Vol. IV, 15–18, CUP, Cambridge), 1892.
51. *Rayleigh Lord.* On the theory of optical images, with special reference to the microscope. *Phil. Mag.*, **42**, 167–195 (Scientific Papers, Vol. IV, 235–260, CUP, Cambridge), 1896.
52. *Rayleigh Lord.* On the theory of optical images, with special reference to the microscope. *Journ. R. Micr. Soc.*, 474–482 (Scientific Papers, Vol. V, 118–125, CUP, Cambridge), 1903.
53. *Ryle M.* A new radio interferometer and its application to the observation of weak radio stars. *Proc. Roy. Soc. A*, **211**, 351–375, 1952.
54. *Ryle M. and Vonberg D.D.* Solar radiation on 175 Mc/s. *Nature*, **158**, 339–340, 1946.
55. *Ryle M. and Smith F.G.* A new intense source of radio-frequency radiation in the constellation of Cassiopeia. *Nature*, **162**, 462–463.
56. *Ryle M. and Hewish A.* The synthesis of large radio telescopes. *Roy. Ast. Soc. Monthly Notices*, **120**, 220–230, 1960.
57. *Schade O.H.* Electro-optical characteristics of television systems. *RCA Review*, **9**, 5–37, 245–286, 490–530, 653–686.
58. *Shannon C.* The mathematical theory of communication. University of Illinois, Urbana, 1949.
59. *Stanier H.M.* Distribution of radiation from the undisturbed Sun at a wavelength of 60 cm. *Nature*, **165**, 354–355, 1950.
60. *Taylor C.A. and Lipson H.* Optical transforms. Bell, London, 1964.
61. *Vander Lugt A.B.* Signal detection by complex spatial filtering. Radar Lab., Report No. 4594–22–T, University of Michigan, Ann Arbor.
62. *Vander Lugt A.B.* Signal detection by complex spatial filtering. *IEEE Trans. Inform. Theory*, IT-10:2, 1964.
63. *Wiener N.* The extrapolation, interpolation and smoothing of stationary time series. Technology Press and Wiley, New York, 1949.
64. *Zernike F.* The concept of degree of coherence and its applications to optical problems. *Physica*, **5**, 785–895, 1938.
65. *Zernike F.* Phase contrast, a new method for the microscopic observation of transparent objects. *Physica*, **9**, 686–698, 974–986, 1942.

Дополнительная литература

- Bahcall J.N. and Spitzer L.* The space telescope. *Sci. Amer.*, **247**, 38–49, 1982.
- Barnes K.R.* The optical transfer function. Hilger, London, 1971.
- Bell R.J.* Introduction to Fourier transform spectroscopy. Academic Press, New York, 1972.
- Born M and Wolf E.* Principles of optics, 6th edn. Pergamon, Oxford, 1980.
- Bracewell R.N.* The Fourier transform and its applications, 2nd edn. McGraw-Hill, New York, 1979.
- Bragg W.L.* (Phillips D.C. and Lipson H., eds.) The development of X-ray analysis. Bell, London, 1975.
- Brown E.B.* Modern optics. Reinhold, New York, 1965.
- Casasent D. (ed.)* Optical data processing, Springer, Berlin, 1978.
- Christiansen W.N. and Hogbom J.A.* Radiotelescopes. CUP, London, 1969.
- Clarke R.H. and Brown J.* Diffraction theory and antennas. Ellis Horwood, Chichester, UK, 1980.
- Culling C.F.* A modern microscopy: elementary theory and practice. Butterworth, London, 1974.
- Cutrona L.J., Leith E.N., Palermo C.J. and Porcello L.J.* Optical data processing and filtering systems. *IRE Trans. Inform. Theory*, IT-6, 386–400, 1960.
- Ditchburn R.W.* Light, 3rd edn. Blackie, London, 1976.
- Duffieux P.M.* The Fourier transform and its application to optics, 2nd edn. Wiley, New York, 1983.
- Duffieux P.M.* L'intégrale de Fourier et ses applications à l'optique. Published privately: Rennes, France, 1946.
- Fowles G.R.* Introduction to modern optics, 2nd edn. Holt, Rinehart and Winston, New York, 1975.
- Françon M.* Modern applications of physical optics, English edn. (transl. Scripta Technica). Wiley, New York, 1963.
- Françon M.* Diffraction: coherence in optics, English edn. (transl. Jeffrey B and edited Sanders J.H.). Pergamon, London, 1966.
- Françon M.* Optical interferometry, English edn. (transl. Wilmanns I.). Academic Press, New York, 1966.
- Françon M.* Holography, English edn. (transl. Spruch G.M.). Academic Press, New York, 1974.
- Françon M.* Optical image formation and processing, English edn. (transl. Jaffe B.M.). Academic Press, New York, 1979.
- Gabor D.* (Edwards Memorial Lecture) Holography: 1948–1971. The City University, London, 1972.
- Gaskill J.D.* Linear systems, Fourier transforms and optics, Wiley, New York, 1978.
- Chatak A.* Optics. Tata McGraw-Hill, New Delhi, 1977.

- Goodman J.W.* Introduction to Fourier optics. McGraw-Hill, New York, 1968.
- Graham N.* Does the brain perform a Fourier analysis of the visual scene? *TINS*, **2**, 207–208, 1979.
- Griffiths P.R. (ed.)* Transform techniques in chemistry. Heyden, London, 1978.
- Hanbury Brown R.* The intensity interferometer. Taylor and Francis, London, 1974.
- Harburn G., Taylor C.A. and Welberry T.R.* Atlas of optical transforms. Bell, London, 1975.
- Hawkes P.W. (ed.)* Computer processing of electron microscope images. Springer, Berlin, 1980.
- Hecht E. and Zajac A.* Optics. Addison-Wesley, Reading, Mass., 1974.
- Hewish A.* The synthesis of giant radio telescopes. *Sci. Progr.*, **53**, 355–368, 1965.
- Hey J.S.* The radio universe. Pergamon, Oxford, 1971.
- Hey J.S.* The evolution of radio astronomy. Elek Science, London, 1973.
- Hopkins H.H.* 21st Thomas Young Oration. The application of frequency response techniques in optics. *Proc. Phys. Soc.* **79**, 889–919, 1962.
- Jenkins F.A. and White H.E.* Fundamentals of optics. 4th edn. McGraw-Hill, Kogakusha, Tokyo, 1976.
- Jennison R.C.* Introduction to radio astronomy. Newnes, London, 1966.
- Kihlborq L. (ed.)* Nobel Symposium 47. Direct imaging of atoms in crystals and molecules. The Royal Swedish Academy of Sciences, Stockholm, 1979.
- Klein M.W.* Optics, Wiley, New York, 1970.
- Kraus J.D.* Radio astronomy. McGraw-Hill, New York, 1966.
- Lee S.H. (ed.)* Optical information processing—fundamentals. Springer, Berlin, 1981.
- Lipson H.S.* Crystals and X-rays. Wykeham Publications, London, 1970.
- Lipson H. (ed.)* Optical transforms. Academic Press, London, 1972.
- Lipson H. and Taylor C.A.* Fourier transforms and X-ray diffraction. Bell, London, 1958.
- Lipson S.G. and Lipson H.* Optical physics, 2nd edn. CUP, London, 1981.
- Longhurst R.S.* Geometrical and physical optics, 3rd edn. Longman, London, 1973.
- McLachlan D.* The role of optics in applying correlation functions to pattern recognition. *J. Opt. Soc. Amer.*, **52**, 454–459.
- Marathay A.S.* Elements of optical coherence theory. Wiley, New York, 1982.
- Michelson A.A.* Studies in optics. University of Chicago Press, Chicago (Phoenix Science Series edn., 1962), 1927.
- Mertz L.* Transformations in optics. Wiley, New York, 1965.
- O'Neill E.L.* Spatial filtering in optics. *IRE Trans. Inform. Theory*, IT-2, 56–65, 1956.
- O'Neill E.L.* Introduction to statistical optics. Addison-Wesley, Reading, Mass., 1963.
- Readhead A.C.S.* Radio astronomy by very-long-baseline interferometry. *Sci. Amer.*, **247**, 39–47, 1982.
- Rogers G.L.* The Abbe theory of microscopic vision and the Gibbs phenomenon. *Amer. Jnl. Phys.*, **22**, 384–389, 1954.
- Schawlow A.L. (ed.)* Lasers and light (Readings from Scientific American). Freeman, San Francisco, 1969.
- Shannon C.* The mathematical theory of communication. University of Illinois, Urbana, 1949.
- Shulman R.A.* Optical data processing. Wiley, New York, 1970.
- Smith H.M.* Principles of holography, 2nd edn. Wiley, New York, 1975.
- Smith F.G.* Radio astronomy, 3rd edn. Penguin Books, London, 1966.
- Smith F.G. and Thomson J.H.* Optics, Wiley, London, 1971.
- Stark H. (ed.)* Applications of optical Fourier transforms. Academic Press, New York, 1982.

- Stokes A.R.* Three-dimensional diffraction theory of microscope image formation. Proc. Roy. Soc. A, **212**, 264–274, 1952.
- Stone J.M.* Radiation and optics. McGraw-Hill, New York, 1963.
- Stroke G.W.* An introduction to coherent optics and holography, 2nd edn. Academic Press, New York, 1969.
- Taylor C.A.* Images. Wykeham Publications, London, 1978.
- Taylor C.A. and Lipson H.* Optical transforms. Bell. London, 1964.
- Welford W.T.* Optics, 2nd edn. OUP, Oxford, 1981.
- Wiener N.* The Fourier integral and certain of its applications, CUP, London, 1933.
- Wiener N.* The extrapolation, interpolation and smoothing of stationary time series. Technology Press and Wiley, New York, 1949.
- Williams E.W.* Applications of interferometry. Methuen, London, 1941.

Содержание

От редактора перевода	5
Предисловие	7
Глава 1. Вступление	9
1.1. Введение	9
1.2. Когерентность и источники света	14
1.3. Формирование оптического изображения	18
1.4. Интерференция при делении амплитуд	25
Глава 2. Дифракция Фраунгофера	27
2.1. Введение	27
2.2. Картина дифракции на одиночной щели	27
2.3. Круглая апертура	31
2.4. Двойная апертура	35
2.5. Решетка из N щелей	39
2.6. Двухмерные решетки	42
2.7. Кристаллы в качестве трехмерных решеток	43
Глава 3. Ряды Фурье и периодические структуры	49
3.1. Введение	49
3.2. Ряды Фурье	50
3.3. Нахождение коэффициентов Фурье: четные функции	52
3.4. Оптические и кристаллические дифракционные решетки: физическая интерпретация членов Фурье	54
3.5. Ряды Фурье: общие определения	57
Глава 4. Преобразование Фурье, свертка и корреляция	62
4.1. Введение	62
4.2. Преобразование Фурье и дифракция на одиночной щели	62
4.3. Дифракционная картина решетки как произведение преобразований	68
4.4. Свертка	71
4.5. Теорема о свертке и дифракция	75
4.6. Преобразования Фурье и световые волны	77
4.7. Корреляция	79
Глава 5. Формирование и обработка оптического изображения	84
5.1. Введение	84
5.2. Формирование некогерентного оптического изображения	88
5.3. Формирование когерентного оптического изображения	92

5.4. Голография	104
5.5. Оптическая обработка	109
Глава 6. Интерферометрия и источники излучения	122
6.1. Введение	122
6.2. Звездный интерферометр Майкельсона	123
6.3. Спектральный интерферометр Майкельсона	130
6.4. Частичная когерентность, корреляция и видность	137
6.5. Фурье-спектроскопия	144
6.6. Приложения в астрономии	150
Приложения	
Приложение А. Скалярное описание электромагнитных волн	162
Приложение Б. Представление Стокса для отражения и преломления	167
Приложение В. Дифракция рентгеновских лучей на кристаллах. Эквивалентность условий Лауэ и концепции отражения Брэгга	
Приложение Г. Электромагнитный спектр	173
Приложение Д. Некоторые полезные формулы	174
Литература	175
Дополнительная литература	178

Уважаемый читатель!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и др. просим присылать по адресу: 129820, Москва, И-110, 1-й Рижский пер., д. 2, Изд-во «Мир».

Стюард И. Г.

ВВЕДЕНИЕ В ФУРЬЕ-ОПТИКУ

Научный редактор М. Ф. Путов
Мл. редактор Н. В. Корниенко
Художник В. П. Груздев
Художественный редактор М. Н. Кузьмина
Технический редактор З. И. Резник
Корректор Н. А. Гиря

ИБ № 5259

Сдано в набор 08.05.84.
Подписано к печати 10.10.84.
Формат 60 × 90/16. Объем 5,75 б. л.
Бумага офсетная № 1.
Гарнитура таймс. Печать офсетная.
Усл. печ. л. 11,50. Усл. кр.-отг. 23,51.
Уч.-изд. л. 11,04. Изд. № 27/3424.
Тираж 4.200 экз. Зак. 458. Цена 1 р. 70 к.

ИЗДАТЕЛЬСТВО «МИР»
Москва, 1-й Рижский пер., 2.

Можайский полиграфкомбинат Союзполиграфпрома при
Государственном комитете СССР по делам издательств,
полиграфии и книжной торговли. 143200, г. Можайск,
ул. Мира, 93.