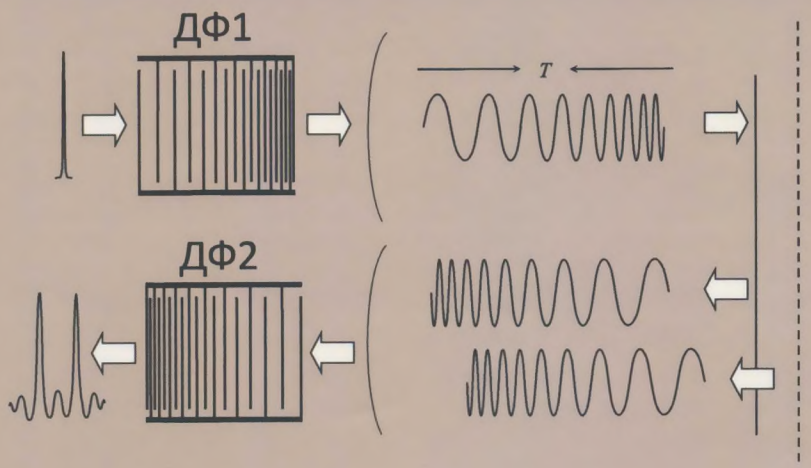
А.В. Войцеховский,
А.А. Скрыльников

ФУНКЦИОНАЛЬНАЯ АКУСТОЭЛЕКТРОНИКА



УДК
ББК
В 65

Рецензенты:
Зав. каф. электронных приборов ТУСУР,
д-р физ.-мат. наук, проф. *С.М. Шандаров,*
декан факультета фотоники и оптоинформатики СПбГУ ИТМО,
д-р физ.-мат. наук, проф. *С.А. Козлов*

Войцеховский А.В., Скрыльников А.А.
В 65 Функциональная акустоэлектроника: Учебное пособие. –
Томск: Томский государственный университет, 2013. – 362 с.

ISBN 978-5-7511-2151-8

Даны введение в функциональную электронику, физические основы работы устройств функциональной акустоэлектроники, представлены элементы теории упругости и нелинейных взаимодействий волн и электромагнитных полей, рассмотрены вопросы возбуждения и распространения объемных и поверхностных акустических волн в различных средах и физические принципы работы устройств на поверхностных акустических волнах. Предложены практические задания по расчёту и оптимизации параметров некоторых простейших устройств.

Для студентов и аспирантов физических, физико-технических, радиофизических и оптотехнических специальностей, а также специалистов в области функциональной электроники.

УДК
ББК

ISBN 978-5-7511-2151-8

© А.В. Войцеховский, А.А. Скрыльников, 2013
© Томский государственный университет, 2013

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

А.В. Войцеховский, А.А. Скрыльников

ФУНКЦИОНАЛЬНАЯ АКУСТОЭЛЕКТРОНИКА

Учебное пособие
Издание 2-е, исправленное и дополненное

Рекомендовано УМО по образованию в области приборостроения и оптотехники в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлению подготовки 200600 «Фотоника и оптоинформатика»

Издательство Томского университета
2013

ПРЕДИСЛОВИЕ

Современная функциональная электроника включает в себя ряд разделов, связанных с акустическими явлениями в твердых телах. К таким разделам относятся: акустоэлектроника, акустооптика и магнитоакустика, в которых принцип действия приборов определяется взаимодействием акустических волн с волнами объёмного заряда (электроны), оптическими волнами и волнами магнитной природы. Среды, в которых распространяются акустические волны, должны обладать, соответственно, электропроводностью, спонтанной поляризацией и намагниченностью, а также быть оптически активными.

Цель данного пособия – изложение физических основ построения устройств функциональной акустоэлектроники. Рассмотрены вопросы возбуждения и распространения акустических волн в материальных средах, обладающих упругими свойствами, электронной проводимостью и спонтанной поляризацией. Обсуждаются нелинейные свойства твердых тел и явления, возникающие при взаимодействии акустических волн и переменных электрических полей, включая трёхволновые взаимодействия. Изложены физические принципы функционирования устройств обработки информации на поверхностных акустических волнах.

Во введении даны начала функциональной электроники. Основное внимание уделено тенденциям развития современной интегральной электроники, а также физическим и схемотехническим ограничениям микроэлектроники. Здесь же приведены физические основы устройств функциональной электроники.

В первом разделе представлены элементы теории упругости, необходимые для дальнейшего изложения материала пособия. Вводятся понятия деформации, упругих сил, упругой энергии, тензоров деформаций, упругих напряжений, модулей упругости. При этом рассмотрены деформации в изотропной среде, кристаллах и пьезоэлектриках.

Второй раздел посвящен акустическим волнам в различных средах. Рассмотрены объёмные и поверхностные акустические волны в упругой, изотропной, кристаллической, пьезоэлектрической и полупроводниковой средах. Подробно проанализированы характеристики поверхностных акустических волн Рэлея, Лява, Стоунли, Лэмба, а также волн Гуляева-Блюстейна. Важное внимание уделено взаимодействию акустических волн с электронами в полупроводниках.

В третьем разделе представлены нелинейные взаимодействия акустических волн, в том числе взаимодействие волн с высокочастотными электрическими полями в пьезоэлектрических средах. Представлен анализ возможных трёхволновых взаимодействий в пьезоэлектрических кристаллах, являющийся основой для создания приборов акустоэлектроники. В одном из параграфов описаны нелинейные свойства твердых тел: решёточная нелинейность монокристаллов, концентрационная нелинейность полупроводников. В заключительном параграфе дана классификация нелинейных акустоэлектронных устройств обработки информации, основанных на взаимодействии акустических волн с учетом нелинейных свойств твердых тел.

Четвёртый раздел является центральным в учебном пособии. В нём представлены устройства на акустических волнах. В начале раздела рассмотрены устройства на объёмных акустических волнах: резонаторы, линии задержки, преобразователи электрических сигналов в акустические. Основное внимание уделено устройствам на поверхностных акустических волнах. Рассмотрены структуры для возбуждения и приёма акустических волн. При этом наиболее подробно проанализированы свойства встречно-штыревых преобразователей – основных элементов всех устройств акустоэлектроники и акустооптики. Проведено рассмотрение характеристик направляющих элементов для поверхностных акустических волн – многополосковых ответвителей и отражательных решёток. В ряде разделов описываются устройства на основе встречных штыревых преобразователей и многополосковых ответвителей (отражательных решёток): линии задержки, резонаторы и полосовые фильтры. Большое внимание среди нелинейных акустоэлектронных приборов уделено усилителям, автогенераторам и синтезаторам частоты. В заключительных параграфах раздела представлены дисперсионные фильтры, акустоэлектронные конвольверы и корреляторы, акустические запоминающие устройства и устройства спектрального анализа сигналов, включая Фурье-процессоры.

Каждый из разделов заканчивается контрольными вопросами.

Второе издание пособия дополнено компьютерным практикумом. По структуре практикум представляет собой ряд задач по расчёту и оптимизации простейших ПАВ устройств. Каждая задача содержит теоретическое описание устройства, дополняющее и расширяющее сведения, представленные в основной части пособия, контрольные вопросы и расчётные задания. Выполнение расчётных за-

даний ориентировано на использование пакетов прикладных программ типа MathCad или MathLab. Выбранные для численного исследования характеристики позволяют проанализировать основные и наиболее общие физические факторы и закономерности, определяющие свойства различных ПАВ устройств.

Задача пособия – сформировать современное представление о наиболее динамично развивающемся разделе функциональной электроники: акустоэлектронике, рассматривающем вопросы взаимодействия акустических волн, а также взаимодействие их с волнами другой природы, включая электромагнитные поля, в различных средах, в том числе с нелинейными свойствами. Для освоения материала не требуется предварительных знаний в области акустики, так как все необходимые сведения имеются в первом и втором разделах настоящей книги.

Необходимость издания данного пособия вызвана тем, что в России по данной тематике имеется недостаточно учебной и технической литературы. Изданная в девяностые годы прошлого века монографическая литература по акустоэлектронным приборам не может в полной мере использоваться в качестве учебных пособий. Учебные пособия, изданные в последние годы, а именно Кравченко А. Ф. «Физические основы функциональной электроники» (2000 г.), А. А. Щука «Функциональная электроника» (1998 г.), содержат лишь краткие разделы по акустоэлектронике среди других разделов функциональной электроники. Данное учебное пособие, на наш взгляд, восполняет пробел в учебной литературе по этому разделу функциональной электроники.

Учебное пособие написано по материалам лекций, прочитанных автором А.В. Войцеховским в Томском государственном университете. В список литературы включены монографии, посвященные различным аспектам затронутых проблем, что позволит более детально изучить их заинтересованным читателям.

ВВЕДЕНИЕ
В ФУНКЦИОНАЛЬНУЮ ЭЛЕКТРОНИКУ

Тенденция развития интегральной электроники. Физические и схемотехнические ограничения. Развитие электроники началось с момента изобретения вакуумной лампы. С самого начала своего развития разработки в области электроники были необходимы для создания радиотехнических систем и устройств беспроводной связи. Разрабатывались радиотехнические схемы устройств для обработки, хранения и передачи информации, состоящие из отдельных электронных приборов. До изобретения транзистора этот путь развития основывался на использовании в качестве активных элементов электронных ламп. С появлением транзистора многие принципы схемотехники претерпели серьезные изменения. Особенности твердотельных приборов заставили искать новые схемотехнические решения, оптимальные для нового класса электронных приборов – твердотельных. К концу 50-х и началу 60-х годов, когда были решены основные принципы конструирования электронных схем на твердотельных (полупроводниковых) приборах, появилась идея интеграции – изготовления в одном кристалле необходимых электронных приборов и элементов, а также соединений между ними. Так появились интегральные схемы. Дальнейшее развитие электроники шло по пути микроминиатюризации. Это дало возможность во много раз снизить габариты, массу, потребляемую мощность, стоимость и значительно повысить надежность и функциональные возможности электронной аппаратуры. С развитием технологии повышалась степень интеграции, т. е. число электронных элементов, приходящихся на один кристалл (или на 1 мм^2 площади кристалла), расширялся выпуск и номенклатура изделий микроэлектроники. Появились большие и сверхбольшие интегральные схемы (БИС и СБИС), в вычислительной технике однокристальный микропроцессор и однокристальная микро-ЭВМ.

Надо сказать, что при всей прогрессивности, при всех достоинствах идея интегральной электроники не несла в себе ничего принципиально нового. Это был все тот же схемотехнический путь развития электроники, но уже в интегральном варианте. Интегральные схемы воспроизводили на кристалле кремния известные схемные решения, полученные на дискретных полупроводниковых приборах (транзисторах, диодах и т. д.). Однако развитие интегральной электроники

накладывало отпечаток и на схемотехнику. Появились новые типы приборов, например, МДП-транзисторы, многоэмиттерные и многоколлекторные транзисторы, схемы с инжекционным питанием, способствующие появлению новых схемных решений радиоэлектронных и других устройств. Однако сама схемотехническая концепция развития электроники оставалась незыблемой.

Схемотехнический путь развития интегральной электроники связан с изготовлением на одном кристалле огромного количества электронных приборов, соединений между ними, изолирующих слоев. Каждый электронный прибор состоит из микроскопических областей полупроводника с вполне определенными свойствами, изолирующих слоев, омических контактов и т. п. Все эти микроскопические области создаются в ходе технологических процессов. С этой точки зрения схемотехническая интеграция является технологической интеграцией. За все время хранения и эксплуатации интегральных микросхем свойства этих областей не должны изменяться. Всякое изменение свойств этих областей представляет собой деградацию. Такие области можно рассматривать как статические неоднородности в кристалле, созданные с помощью технологических процессов. Обработка информации осуществляется продвижением информационного сигнала из области одной статической неоднородности в другую, при этом происходит непрерывное преобразование физических величин из одной в другую. Другими словами, схемотехническая интеграция – это чисто технологическая интеграция, интеграция электронных приборов, состоящих из статических неоднородностей.

Размеры этих областей – статических неоднородностей – весьма малы. С возрастанием уровня интеграции наблюдается систематическая тенденция к их уменьшению. Уже сейчас на отдельных образцах интегральных схем минимальные размеры элементов составляют менее 1 мкм. Расчеты показывают, что с уменьшением размеров статических неоднородностей неизбежно придется столкнуться с целым рядом физических ограничений. Так, с уменьшением размеров при определенных величинах статических неоднородностей теряют смысл статистические законы распределения примесей, и ставится под сомнение справедливость используемых моделей транзисторов. Кроме того, возникают проблемы с достижением значительных плотностей тока, необходимых для обеспечения быстродействия, с уменьшением рабочих потенциалов и получением заданной надежности, с отводом рассеиваемой в малых объемах тепловой мощности и др. Пределом уменьшения раз-

меров элемента топологии (ширина линий, зазоров между ними и т. д.) является величина 0,05 мкм.

В то же время есть все основания предполагать, что еще до достижения физических пределов интеграции мы столкнемся с весьма существенными технологическими ограничениями. По мере уменьшения размеров элемента топологии существенно снижается воспроизводимость технологических процессов, повышается вероятность появления случайных дефектов. Большую сложность представляет переход к субмикронным линиям нанесения рисунка схемы (литография). Фотолитография уже не может удовлетворить технологов, поскольку длина волны излучения в видимом свете составляет 0,4–0,7 мкм. Возникла необходимость перехода к рентгено-, а затем и к электронной литографии. Термические процессы и процессы ионного легирования также связаны с возникновением дефектов различной природы, их влияние на технологические процессы резко возрастает при переходе к субмикронной технологии.

Еще одним серьезным ограничением в схемотехнической интегральной электронике является проблема межсоединений. С ростом уровня интеграции возрастает и доля площади кристалла, занимаемая соединениями между схемотехническими элементами (транзисторами, диодами, образуемыми ими ячейками и т. д.). Толщина металлических и особенно диэлектрических слоев не должна быть меньше 0,05 мкм. Проблема дефектов обостряется, в частности, наличие пор в диэлектрике приводит к коротким замыканиям и обрывам узких тонких токоведущих дорожек.

Другими словами, «тирания количеств» в схемотехнической интегральной электронике, опирающейся на технологические принципы интеграции, оборачивается «тиранией дефектов». Ясно, что беспредельно увеличивать уровень интеграции нельзя.

В последние годы в твердотельной электронике ведутся работы по исследованию ряда новых физических принципов и эффектов с целью создания принципиально новых электронных устройств. Отличительной чертой таких устройств является то, что в основе их работы не используются схемотехнические принципы. Функции схемотехники выполняют в них непосредственно те или иные физические процессы.

Характерной чертой этих физических процессов является наличие и использование для обработки или хранения информации динамических неоднородностей в однородном объеме твердого тела.

Примером таких динамических неоднородностей могут быть ганновские электрические домены, цилиндрические магнитные домены, пакеты зарядов в приборах с зарядовой связью, поверхностные и объемные акустические волны, спиновые волны и др. Представляется, что использование в интегральной электронике комбинаций различных физических явлений даст определенные преимущества и позволит повысить степень интеграции по сравнению с технологическими методами интеграции.

Следует особо подчеркнуть, что основной принципиальной чертой интеграции физических явлений служит использование динамических неоднородностей для выполнения определенных функций. С этой точки зрения данное направление в интегральной электронике называют функциональной электроникой.

В настоящее время существует несколько направлений исследований, основанных на непосредственном использовании физических явлений, потенциально пригодных для создания функциональных устройств. Это – акусто- и магнитоэлектроника, акусто- и магнитооптика и т.д.

Перспективы и преимущества развития функциональной электроники несомненны. Однако, несмотря на определенные успехи, достигнутые функциональной электроникой к настоящему времени, она находится еще на начальном этапе своего развития. Многие устройства, предложенные на принципах функциональной электроники, еще не вышли из стадии разработки, а многие идеи пока еще вообще не реализованы. Одним словом, здесь еще много нерешенных проблем. К числу таких проблем, например, относятся поиск новых сред и новых физических процессов, позволяющих решать задачи хранения и обработки информации, изучение возможности использования взаимодействия динамических неоднородностей различной природы в одной многослойной структуре и многие другие. Следует особо подчеркнуть, что развитие функциональной электроники не исключает развития традиционной схемотехнической микроэлектроники, а дополняет и расширяет ее границы и области применения.

Современная электроника в значительной степени является интегральной электроникой, она реализует в одном кристалле функциональную схему, будь то определенную логику, запоминающее или вычислительное устройство в одном кристалле. Сегодня микроэлектроника является схемотехнической и решает задачи создания определенных схем устройств, выполняющих заданные функции. Важ-

ным следствием появления ИС стало уменьшение размеров, энергоемкости и, самое главное, повышение надежности устройств. К другим преимуществам относятся простота серийного производства и высокие эксплуатационные показатели, что позволяет широко применять изделия микроэлектроники.

Темпы развития интегральной электроники впечатляющие. Главными направлениями исследований в микроэлектронике продолжают оставаться поиск путей уменьшения размеров элементов ИС, а значит, повышения степени интеграции, и, соответственно, скорости вычислений, разработка эффективных технологий, а также применение новых материалов.

За счет чего должен продолжаться рост степени интеграции? В принципе здесь возможны три пути: уменьшение топологических размеров элементов интегральных схем; увеличение площади кристалла; переход к расположению элементов прибора в объеме кристалла, так называемой трехмерной электронике. Каждый из этих путей характеризуется весьма существенными сложностями и не может рассматриваться как единственно перспективный путь решения проблемы.

Наибольшие сложности в этом отношении представляет переход к трехмерной микроэлектронике. Этот путь связан с очень серьезными технологическими трудностями, хотя поиски в этом направлении ведутся весьма интенсивно.

Наиболее простым путем на первый взгляд представляется увеличение размера кристалла. Однако практика промышленного производства ИС показывает, что с увеличением площади кристалла резко возрастает вероятность возникновения дефектов. К числу таких дефектов могут относиться и внутренние дефекты исходного полупроводникового материала, и дефекты, вносимые в ходе выполнения технологических операций изготовления ИС.

Естественно, что чем большее количество последовательных прецизионных технологических операций включает в себя технологический процесс, тем более вероятным становится возникновение в кристалле тех или иных дефектов. В результате процент выхода годных интегральных схем с увеличением площади кристалла падает соответственно по экспоненциальному закону. Поэтому кристалл должен быть достаточно малым, чтобы вероятность возникновения дефектов была минимальна. В результате усовершенствования производства кремниевых подложек и технологии изготовления ИС плотность дефектов постоянно уменьшается, что позволяет увеличивать размеры кристаллов. Одна-

ко нет оснований считать, что увеличение размеров кристалла существенно снизит число дефектов на единицу площади. По-видимому, это увеличение будет медленным, так как существенные изменения размеров кристаллов потребуют значительных капиталовложений в создание нового технологического оборудования.

Микроминиатюризация элементов и рост уровня интеграции являются доминирующим направлением совершенствования изделий интегральной схемотехнической электроники. Предполагается, что кремний в силу своих уникальных свойств и прогресса кремниевой технологии останется основным материалом, несмотря на поиск и внедрение новых материалов.

Анализ традиционных путей развития интегральной электроники показывает, что в настоящее время достигнут настолько высокий уровень интеграции, что при его дальнейшем повышении приходится считаться с рядом физических и технологических ограничений. Дело в том, что нельзя беспредельно уменьшать размеры приборов не нарушая их работоспособности в соответствии с заложенными в них физическими принципами и не вызывая резкого удорожания их производства из-за необходимости значительных капиталовложений в создание более совершенного технологического и измерительного оборудования.

В настоящее время наиболее распространенным методом уменьшения элементов и кристаллов ИС является масштабирование. Масштабный принцип подхода к уменьшению размеров предполагает, что все размеры элемента, например МДП-транзистора, а также напряжение питания уменьшаются в одно и то же число раз. Это позволяет, не изменяя технологических процессов и топологии ИС, улучшить параметры ИС пропорционально масштабу этого уменьшения.

Масштабирование удачно применялось при уменьшении размеров элементов от 10 мкм до единиц микрометра. Однако при переходе к длинам менее 1 мкм оно действует не столь эффективно, и необходимо учитывать ограничения, связанные с физическими эффектами, возникающими при малых геометрических размерах. Одно из ограничений связано с тем, что рабочее напряжение не может быть ниже определенного значения. Так, например, для нормальной работы транзисторов и надежного выполнения ими пороговых функций (при работе в режиме электронного ключа) напряжение питания кремниевых ИС на МОП-транзисторах не может быть меньше 1 В. Вследствие этого, начиная с некоторого момента, напряжение уже

не может быть уменьшено пропорционально размерам приборов. Если напряжение питания задано, то с уменьшением размеров электрические поля в диэлектрике и в обедненных областях растут. Особое значение в этом случае приобретают связанные с высокими электрическими полями такие явления, как пробой *p-n* переходов и диэлектрика, появление «горячих» электронов и туннелирование их в диэлектрический слой. Поэтому при заданном напряжении существует минимальная толщина диэлектрика или обедненного слоя, при которой электрическое поле еще не превышает пробивного значения или не является причиной появления других нежелательных эффектов. Эта минимальная толщина определяет минимум всех других размеров прибора и, таким образом, ставит предел миниатюризации приборов такого типа.

При приближении к предельным размерам транзисторов начинают проявляться и другие нежелательные физические эффекты. Наиболее существенным из них является эффект предпорогового включения транзистора, приводящий к постепенному возрастанию проводимости канала. Другой эффект, ограничивающий уменьшение размеров, связан с туннелированием «горячих» электронов, образующихся в канале под действием высоких электрических полей, через слой подзатворного окисла. Вырываясь из канала в слой SiO_2, они захватываются имеющимися в нем ловушками и заряжают их, в результате чего изменяется пороговое напряжение включения (или выключения) МОП-транзистора.

Нельзя не учитывать и ограничение предельных размеров элементов ИС, обусловленное флуктуациями числа атомов легирующих примесей. Малые размеры элементов современных ИС и малые концентрации примесей ведут к тому, что рабочая область прибора (область объемного заряда *p-n* перехода, канал МОП-транзистора и др.) может содержать довольно малое количество примесных атомов. Поскольку последние распределены в решетке полупроводника случайным образом, их число будет флуктуировать от прибора к прибору относительно некоторого среднего значения. Эти отклонения приводят к разбросу от прибора к прибору его важнейших параметров, таких, как пороговое напряжение (для МОП-транзисторов), ток и др. По мере повышения уровня миниатюризации влияние флуктуаций становится все более существенным. Заметим, что разброс; обусловленный флуктуациями, не уменьшается с понижением температуры. Эти ограничения, пожалуй, являются наиболее серьезны-

ми и, по имеющимся в литературе сведениям, приводят к четырехкратному возрастанию минимальных размеров элементов.

К сказанному следует добавить, что при малых размерах, о которых шла речь выше, теряют смысл представления о концентрации примесей и законе ее распределения. В этом случае, видимо, становится сомнительной и справедливость используемых математических моделей приборов, без которых невозможно машинное проектирование.

Еще одним существенным фактором, препятствующим уменьшению размеров транзистора, является неоднородность пленок, используемых при формировании его структуры, и трудность в связи с этим точного воспроизведения его размеров. Для МОП-транзистора, например, трудно воспроизвести длину канала с приемлемой точностью при размерах менее 1 мкм. Пока длина канала превышала 1 мкм, его воспроизводимость определялась «качеством» процессов получения и переноса изображения и последующих технологических операций. При меньших размерах начинает сказываться структура (неоднородность) материалов, определяющих конечный размер канала.

Следует иметь в виду, что все приведенные выше рассуждения велись без учета дефектности материалов. Известно, что в бездислокационном кремнии, который используется для биполярных и полевых БИС, содержатся разнообразные дефекты: точечные дефекты и их скопления (так называемые кластеры), дефекты упаковки, включения второй фазы с линейными размерами от десятых долей микрометра до нескольких микрометров, слоистые неоднородности в распределении примесей, связанные с условиями вытягивания монокристаллов и имеющие размеры в несколько десятков микрометров и др. Учет дефектов материала приведет к еще более жестким ограничениям.

Говоря о проблемах развития схемотехнической микроэлектроники, невозможно обойти вниманием проблему межсоединений. По мере увеличения числа вентилей на кристалле существенно возрастает доля площади кристалла, занимаемая межсоединениями. Вследствие этого дальнейшее повышение уровня интеграции, особенно в логических схемах, весьма проблематично. В настоящее время при площади кристалла порядка 1 $см^2$ уже 80% будут занимать межсоединения и лишь 20% – логические схемы. Таким образом, проблема межсоединений вырастает в одну из серьезнейших среди проблем создания ИС с высоким уровнем интеграции (БИС и СБИС). Один из путей уменьшения занимаемой межсоединениями площади кристал-

ла это просто введение дополнительных уровней межсоединений. Расчеты показывают, что увеличение числа уровней разводки со стандартных трех (внутриячеечные соединения, межячеечные соединения, питание) даст возможность снизить площадь, занимаемую межсоединениями, с 80 до 50%. В то же время многоуровневая разводка сама по себе представляет весьма сложную проблему, особенно при малых топологических размерах.

Однако проблема межсоединений не сводится только к проценту занимаемой ими площади кристалла. Уменьшение топологических размеров существенно уменьшает сечение токоведущих дорожек и увеличивает отношение их длины к сечению.

Следствием этого являются увеличение сопротивления соединительных проводников и другие нежелательные явления. Обеспечение быстродействия сводится в той или иной степени к скорости перезарядки емкостей, т. е. к протеканию по токоведущим дорожкам токов относительно высокой плотности. Здесь нельзя не считаться с явлением электромиграции – диффузионного процесса перемещения атомов в твердом теле при высоких плотностях тока, разрушающего токов едущие дорожки. Так, для алюминиевых токоведущих дорожек безопасной может считаться плотность тока, не превышающая 10^5 А/см^2. При ширине токоведущей дорожки 0,1 мкм и ее толщине, равной 0,25 ширины, допустимая величина тока составит уже $4 \cdot 10^5$ А/см^2, т. е. существенно превысит безопасную. Приходится считаться и с падением напряжения на токоведущих дорожках от выхода, например, одной логической ячейки до входа следующей. Минимальное входное напряжение определяется порогом надежного срабатывания. Выходное напряжение предыдущей логической ячейки в этом случае должно быть больше входного напряжения последующей ячейки на величину падения напряжения на токоведущей дорожке. При малой ширине токоведущей дорожки мала будет и ее толщина, и толщина диэлектрических слоев, и расстояние между токоведущими дорожками в плане. В этом случае возрастают напряженности электрического поля, и приходится считаться с возможностью электрического пробоя через объем или по поверхности диэлектрика.

Все перечисленные выше факторы ограничивают уменьшение размеров токоведущих дорожек.

Повышение уровня интеграции не является единственной проблемой схемотехнической микроэлектроники. Серьезной проблемой

является также и повышение быстродействия. Эти две проблемы входят в известное противоречие. С одной стороны, для повышения информационных скоростей и тактовых частот нужно уменьшать размеры ИС и повышать плотность упаковки. С другой стороны, повышение мощности в целях ускорения переключения, как правило, требует увеличения расстояния между отдельными элементами схемы по причинам, связанным с тепловым режимом, что приводит к уменьшению плотности упаковки и увеличению задержки на распространение сигнала по соединительным линиям. Поэтому из-за энергетических ограничений предельные значения уровня интеграции и быстродействия не могут быть реализованы одновременно.

Мощность, рассеиваемая в кристалле в виде тепла, должна быть меньше тепловой мощности, которая может быть отведена. От твердой поверхности при равномерном распределении на ней источников тепла в воздух может быть отведена максимальная удельная мощность около 2 Вт/см^2, а в охлаждающую жидкость – примерно 20 Вт/см^2. При работе же СБИС с предельным быстродействием плотность выделяемой мощности может достичь 10^4 Вт/см^2, что на три порядка превышает удельную плотность мощности, которая может быть отведена известными способами. Поэтому в настоящее время можно говорить либо о СБИС с умеренным быстродействием, либо о ССИС с умеренным уровнем интеграции. Тепловые ограничения можно ослабить путем снижения рабочей температуры устройства. Так как мощность и температура связаны соотношением $P \sim T^2$, то при низких температурах должна уменьшиться мощность рассеяния. Грубые оценки показывают, что при снижении температуры от 300 до 77 К потери уменьшаются в 16 раз. Снижение температуры вызывает и другие полезные эффекты. Один из них – смягчение ограничений, связанных с сопротивлением соединений. Пониженное сопротивление металлов позволяет сделать более узкими соединительные линии, вследствие этого снижается площадь, занимаемая межсоединениями, сильно уменьшается электромиграция и др. Очевидно, что при низких температурах можно достичь более высоких плотностей тока в устройствах без снижения их быстродействия. Однако система охлаждения требует дополнительных затрат, увеличивает массу аппаратуры и не всегда удобна в использовании.

Одной из тенденций в повышении быстродействия схем является переход от использования кремния к использованию полупроводниковых материалов с более высокими значениями подвижности, на-

пример, арсенида галлия или соединения германий-кремний. Тогда без повышения мощности и уменьшения плотности упаковки можно было бы достичь высокого быстродействия. Однако здесь встает проблема отработки на арсениде галлия или германий-кремнии технологии, достаточно управляемой и воспроизводимой, отвечающей требованиям промышленного производства БИС и СБИС. Выхода технологии ИС на новых материалах на уровень технологии кремния можно будет ожидать не ранее чем через 10–15 лет. Можно предположить, что на первом этапе высокое быстродействие будет достигнуто за счет снижения уровня интеграции кремниевых ИС.

Другим путем повышения быстродействия может явиться параллельная обработка информации несколькими устройствами. Имеется точка зрения, что повышения производительности на этом пути можно будет достичь значительно быстрее и проще, чем за счет увеличения только скорости обработки.

Остановимся еще на одном очень важном моменте. ИС с высокой степенью интеграции представляют собой ансамбль огромного числа (10^6–10^8) элементов. Это делает необходимым проводить анализ их работы на основе статистических закономерностей. При высоком уровне интеграции и большом сроке службы в схеме происходит столь большое число событий, что могут иметь место и маловероятные процессы. Нельзя, например, не учитывать случайные сбои в работе БИС и СБИС, обусловленные тепловыми флуктуациями, которые приобретают особенно существенное значение при снижении напряжения питания.

Серьезным является и вопрос, связанный с надежностью работы СБИС в условиях воздействия различного рода ионизирующих излучений. С уменьшением размеров элементов схемы работают все с меньшим рассеянием энергии, а функционирование элементов обеспечивается все меньшим количеством свободных электронов и дырок. Энергия, затрачиваемая, например, на операции в схемах логики, близка по значению к энергии частиц, испускаемых в процессе естественной радиоактивности (до 10 МэВ), и к энергии частиц космического излучения.

В элементах памяти, о которых идет речь, информация записывается в виде наличия или отсутствия заряда на конденсаторе. Когда высокоэнергетическая частица порождает вблизи прибора пары электрон – дырка, часть добавочных носителей заряда собирается на конденсаторе, изменяя его заряд, и если этот дополнительный заряд достаточно велик, он искажает содержащуюся информацию.

Серьезной проблемой является также проблема проектирования при высоких уровнях интеграции. В настоящее время кристаллы БИС и СБИС уже не могут традиционно рассматриваться как компоненты или элементная база. Проектирование таких кристаллов выводит конструкторов уже на уровень подсистем. В дальнейшем можно ожидать и выхода на уровень систем. Стоимость и сроки проектирования непрерывно возрастают. При малых объемах производства стоимость проектирования может составлять для БИС и СБИС до 50% их себестоимости.

Создается также опасность, что СБИС может оказаться морально устаревшей не только к моменту начала производственного выпуска, но и к моменту окончания проектирования. Решить эту проблему может только широкое использование средств автоматизации при проектировании. Только на этом пути возможно минимизировать сроки проектирования СБИС.

Одну из серьезных проблем представляет контроль работоспособности изготовленных схем с высоким уровнем интеграции. Одним из направлений решения этой проблемы является введение в них специальных средств обнаружения и коррекции ошибок, диагностики отказов и самовосстановления. Все эти средства должны являться неотъемлемой частью этих схем, что делает их еще более сложными.

Подводя итог сказанному, можно сделать вывод, что уменьшение размеров элементов ИС и повышение уровня интеграции связаны с серьезными физическими ограничениями. В то же время есть все основания предполагать, что еще до достижения физических пределов интеграции мы столкнемся с весьма существенными технологическими ограничениями. К их числу следует отнести усложнение и удорожание технологических процессов и оборудования, повышение требований к контрольно-измерительной аппаратуре при уменьшении размеров элементов, ужесточение требований по чистоте технологических сред и реактивов и по точности совмещения при резко возросшем числе совмещаемых элементов, существенное снижение процента выхода годных приборов с увеличением площади кристалла за счет повышения вероятности возникновения дефектов в кристалле.

Изготовление бездефектных кристаллов, разработка технологии, не создающей дефектов, являются весьма сложной технической проблемой. Проблема эта усложняется и тем, что при малых размерах элемента существенное значение приобретают дефекты размером 1 мкм и менее.

Таким образом, условия массового производства, требования воспроизводимости технологических процессов вряд ли дадут возможность подойти «вплотную» к барьерам физических ограничений, а продвижение к ним будет происходить все труднее и медленнее.

Предшествующий материал позволяет сделать следующие заключения:

1. Традиционный путь развития ИС памяти – схемотехническое направление, направление активных и пассивных элементов, соединенных с помощью межсоединений (токоведущих дорожек) в электрическую схему, направление, в котором носителем информации является некоторая совокупность довольно большого количества СН (включая в это число и внутриячеечные соединения).

2. Основным способом обработки и хранения информации для схемотехнических решений ЗУ является последовательный метод обработки, основанный на Булевой алгебре.

3. Статические неоднородности создаются в ходе технологических процессов. Именно в этих процессах возникают как явные, так и скрытые дефекты. Таким образом, уменьшение числа статических неоднородностей и облегчение требований к ним повышает экономические показатели, надежность схем памяти.

4. В ячейке полупроводниковой схемы памяти с носителем единичной информации происходит постоянное преобразование от одного вида к другому: «потенциал-заряд-ток-потенциал»... Эти многочисленные преобразования могут являться причиной сбоев схемы.

5. Наиболее уязвимым местом в схеме памяти является такой вид СН, как межсоединения. Именно они могут являться основной причиной ненадежности, а также ограничить продвижение устройств и приборов схемотехнической электроники в направлении повышения уровня интеграции.

6. Межсоединения в схемотехнической электронике образуют жесткие связи, которые практически невозможно изменить в ходе эксплуатации системы.

7. Увеличение степени интеграции за счет увеличения площади кристалла при наиболее употребимых топологических размерах встречает трудности из-за резкого возрастания дефектообразования (по экспоненте) с ростом площади кристалла.

8. Проблема многоуровневых межсоединений является крайне сложной. Тем более трудно пока говорить о трехмерных системах памяти.

Ближайший прогноз (2010 г.) на достижение критических значений параметров ИС при использовании современной технологии:

– число элементов – 10^7-10^8;

– размер транзисторов – 30 нм;

– длина затвора – 13 нм;

– рабочее напряжение – 0,6-0,75 В;

– тактовая частота – 2,6 ТГц.

Таким образом, одной из серьезнейших проблем современной микроэлектроники, в том числе схем памяти, является проблема существенного уменьшения количества статических неоднородностей вообще и межсоединений в частности. Альтернативной сугубо схемотехническому пути развития микроэлектроники является электроника динамических неоднородностей или функциональная электроника.

Физические основы работы устройств функциональной электроники. Выход необходимо искать в направлении использования динамических неоднородностей для целей обработки и хранения информации. Существует большое число физических явлений, позволяющих возбуждать динамические неоднородности в различных континуальных средах, а также управлять ими с помощью статических неоднородностей.

Несмотря на многообразие используемых физических явлений и эффектов, носителей информационного сигнала, сред распространения сигнала и механизмов его детектирования, можно выделить и классифицировать общие элементы и свойства приборов функциональной электроники.

Для создания прибора функциональной электроники, прежде всего, необходимо выбрать среду, в которой можно создать как статические, так и динамические неоднородности. Среда может находиться в любом агрегатном состоянии, важно, чтобы она имела распределенные параметры по направлению распространения информационного сигнала.

Физические свойства тонкопленочных структур значительно отличаются от свойств объемного образца того же материала. Значительное влияние оказывают дислокации, вакансии, границы зерен, двойники, а также стехиометрия состава, механические напряжения и деформации.

Направленное распространение возбуждения, создаваемого в элементарной ячейке среды, осуществляется благодаря физическим свя-

зям, свойствам среды распространения. Такое возбуждение представляет собой информационный сигнал, механизм возникновения, хранения или распространения которого определяется как свойствами среды, так и физическими явлениями или их совокупностью. Возбуждение может проявляться, например, в избыточной или недостаточной концентрации свободных электронов или дырок, во взаимодействии акустических волн высокой частоты с электронами твердого тела, в заданной ориентации спинов связанных электронов и т. п.

Возбуждения, возникающие и распространяющиеся в среде, являются динамическими неоднородностями. Динамическая неоднородность представляет собой локальный объем в однородной среде, создаваемый в результате обратимых физических процессов, не имеющий внутри себя статических неоднородностей и способный перемещаться по объему континуальной среды, в результате взаимодействия со статическими неоднородностями или динамическими неоднородностями другого или такого же вида. Динамическая неоднородность способна переносить информационный сигнал и ее не следует смешивать с неравновесным состоянием статической неоднородности. Длительность существования ДН может быть различной: кратковременной (например, в ПЗС) или долговременной (в ЦМД). ДН может исчезать при снятии внешнего возбуждающего фактора (ПАВ) или сохраняться неограниченно долго (электрические и магнитные домены). Такие неоднородности создаются физическими методами и в зависимости от свойств среды имеют различные времена жизни, определяющие длительность хранения информации или ее распространения от одного элементарного объема к другому. Время жизни определяет также и степень деградации такой неоднородности.

Динамические неоднородности возникают в результате неравновесных физических процессов и обладают следующими особенностями:

организуются физическими средствами в процессе возникновения, распространения или хранения сигнала;

возникновение, исчезновение или деградация определяются совокупностью обратимых физических процессов;

обладают способностью управляемого переноса неоднородности по объему;

осуществляют перенос информационного сигнала в аналоговой или дискретной форме.

Генерация, преобразование или усиление информационного сигнала в среде в форме динамической неоднородности осуществляются с помощью статических неоднородностей.

Статическая неоднородность представляет собой локальную область на поверхности или в объеме среды, созданную и сохраняемую в данном месте в результате заданного технологического процесса. При воздействии внешних факторов (подключение напряжения, освещение и т. п.) В них могут возникнуть неравновесные процессы. Всякое неконтролируемое изменение свойств СН рассматривается как деградация.

Статические неоднородности используются как в схемотехнической, так и в несхемотехнической электронике, однако при этом они выполняют различные функции. В первом случае носителем информации является состояние совокупности СН, во втором – СН используется как средство генерирования, управления и детектирования носителей информации – динамических неоднородностей. Примерами СН могут служить *p–n*-переходы, контакты, межсоединения, резисторные, емкостные элементы и т. п.

Именно статические неоднородности определяют, основные принципы схемотехнической электроники. Добиться уменьшения их количества можно только путем упрощения схемотехники. При этом можно рассчитывать на весьма несущественное уменьшение количества статических неоднородностей за счет снижения функциональных возможностей устройства. Более того, схемотехнические упрощения могут в гораздо большей степени повысить вероятность сбоев. Остается предположить, что снижение числа статических неоднородностей может быть достигнуто только путем отказа от традиционных схемотехнических принципов обработки и хранения информации и перехода к иным, несхемотехническим принципам.

В статических неоднородностях под воздействием внешних факторов возникают неравновесные физические явления, приводящие к генерации динамических неоднородностей. Статические неоднородности предназначены для ввода, вывода, управления, распространения или хранения информационного сигнала в среде. Функции элементарных ячеек среды, находящихся в начале тракта передачи информационного сигнала, сводятся к генерации и передаче сигнала в следующую ячейку. Для ячеек, находящихся в тракте передачи сигнала, характерна функция управления этим сигналом, ввод в него нужной информации, ее направленный перенос, регули-

рование скорости передачи либо управление хранением. Ячейки, находящиеся в конце тракта передачи информационного сигнала, обеспечивают его детектирование, вывод из среды в адекватной, удобной для обработки форме.

Переход от классических принципов интегральной электроники, основанных на схемотехнических решениях, статических неоднородностях и технологической интеграции, к несхемотехническим принципам функциональной электроники, основанным на использовании физических принципов интеграции и динамических неоднородностей, позволяет рассчитывать, что снижение количества статических неоднородностей приведет к снижению вероятности появления дефектов, возникающих при технологических методах создания статических неоднородностей (создание динамических неоднородностей не должно сопровождаться дефектообразованием). Однако при этом значительно возрастают требования к качеству и совершенству используемого материала (кристаллической решетки, поверхности, границы раздела и т.д.), так как материал становится той средой, где развиваются процессы.

Таким образом, использование принципов несхемотехнической функциональной электроники и физической интеграции позволяет рассчитывать на возможность увеличения рабочих площадей при тех же топологических размерах, что и в случае классического схемотехнического решения. Преимущества от использования физических методов интеграции не исчерпываются только этим эффектом, дальнейшее их выявление связано с проведением существенного объема работ по исследованию возможностей и сравнительных характеристик этих двух методов.

Основные направления развития функциональной электроники в сравнении с традиционной интегральной электроникой показаны в таблице. Останавливаясь на основных направлениях работ в области функциональной электроники, следует обратить внимание на необходимость более глубокого изучения физических и технологических ограничений в этих двух направлениях, и, прежде всего, в схемотехнической электронике как в более проработанном и традиционном направлении.

Эта проблема уже изучалась, однако целый ряд вопросов физики, конструирования и технологии нельзя считать полностью решенными. Более того, для ряда направлений функциональной электроники аналогичные вопросы еще просто не рассматривались, что не дает возможно-

сти в ряде случаев провести обоснованное сравнение и сделать выводы о перспективах и возможностях того или иного направления.

Таблица

Направления развития функциональной электроники и физические явления, используемые в функциональной электронике

Функциональная электроника	
Физические явления	**Направления**
Оптические явления	Оптоэлектроника
Акустические явления	Акустоэлектроника
Магнитные явления	Магнитоэлектроника
Явления в полупроводниках	Полупроводниковая электроника
Криоэлектронные явления (сверхпроводимость)	Криоэлектроника
Квантово-размерные явления	Квантовая электроника
Явления в диэлектриках	Диэлектрическая электроника
Тепловые явления	Теплоэлектроника
Явления в молекулярных системах	Молекулярная электроника
Пьезоэлектрические явления	Пьезоэлектроника
Электрохимические явления	Хемоэлектроника
Автоволновые явления	Автоволновая электроника
Явления самоорганизации	Синергетика
Явления живой природы	Бионика

К числу фундаментальных проблем функциональной электроники следует отнести необходимость изучения минимально возможных размеров динамических неоднородностей различного вида, скорости их образования и исчезновения, энергетических характеристик этих процессов, подвижности динамических неоднородностей, длительности их существования, причин деградации, влияния внешних факторов на их устойчивость (например, на устойчивость к воздействию альфа-частиц) и т.п.

Эта проблема выдвигает в свою очередь необходимость решения проблемы метрологии основных параметров динамических неоднородностей: размерных, временных, энергетических, надежностных и т.д. на более высоком по сравнению с существующим уровне.

Уже в настоящее время устройства функциональной электроники находят широкое применение. Нет, пожалуй, ни одной системы беспроводной радиосвязи, где бы сегодня не использовались устройства на поверхностных акустических волнах (ПАВ). Они используются и в маленьких и простых пейджерах, и в сложных

приёмниках для сотовой связи, телевизорах и оптоволоконных системах передачи данных.

Столь широкое применение объясняется малыми габаритами, незначительным весом, отсутствием энергопотребления, технологией изготовления совместимой с производством ИС, возможностью использования корпусов для автоматизированного поверхностного монтажа.

Большинство преимуществ ПАВ-устройств обусловлено непосредственно их физической структурой: малым весом и габаритами; линейной (или определяемой требованиями) фазой; фактором формы, приближающимся к единице (очень высокая прямоугольность); исключительным внеполосным подавлением; температурной стабильностью. Поскольку центральная частота и форма частотной характеристики определяются топологией, они не требуют сложной настройки в аппаратуре и не могут расстроиться в процессе эксплуатации. Технология изготовления, совместимая с полупроводниковым производством, позволяет выпускать их в большом объёме с высокой воспроизводимостью.

Одной из наиболее интересных задач функциональной электроники является исследование возможности интеграции различных физических эффектов в одном устройстве. Наряду с развитием устройств функциональной электроники, использующих какой-либо один физический эффект (ПЗС, ЦМД, ПАВ и т.п.) и динамические неоднородности одного вида, необходимо прорабатывать принципы создания многослойных структур с соприкасающимися распределенными средами разной природы. Возникает необходимость изучения процессов взаимодействия динамических неоднородностей разной природы, которые происходят вблизи границы раздела двух сред, представляющей собой в общем случае гетеропереход.

К интересным явлениям относится возможность параллельного переноса всего массива информации из одной матрицы в другую, при этом особую важность будут представлять техника ПЗС и оптоэлектронные процессы. Перспективность использования принципов ПЗС в устройствах функциональной электроники определяется также и тем, что информация в них может обрабатываться в цифровой и аналоговой форме и вводиться световой инжекцией или инжекцией с электронно-дырочных переходов. В многослойных устройствах с обменом информацией между слоями возможно получение трехмерных интегральных устройств, что представляется весьма перспективным направлением функциональной электроники.

Немаловажное значение имеет поиск и синтез новых материалов, пригодных для получения в них динамических неоднородностей, к которым относится группа таких «искусственных» материалов, как сверхрешеточные структуры. В этой области изучаются технологические проблемы получения таких структур, возможности создания на их основе приборов с падающим участком характеристики (получение в них динамических неоднородностей той или иной природы до настоящего времени не рассматривались).

Представляют интерес также возможности использования в устройствах функциональной электроники тонких монокристаллических сегнетоэлектрических пленок, а также динамических неоднородностей, связанных с локальным изменением структуры вещества, как это имеет место, например, в халькогенидах. Перспективными представляются и динамические неоднородности типа магнитных вихрей, возникающих в магнитных материалах при сверхнизких температурах в области сверхпроводимости. Но главное – это получение совершенных, бездефектных материалов.

В области прикладных исследований важное место занимает проблема созданий устройств долговременной энергонезависимой перепрограммируемой памяти на принципах функциональной электроники с высокой плотностью записи информации, достаточно малым временем выборки и высокой надежностью. Такие устройства, обладающие преимуществами по массо-габаритным и энергетическим показателям, по быстродействию и механической устойчивости, заменят устройства внешней памяти, например, на магнитных дисках.

Другой важной проблемой в области прикладных исследований являются поиски различных путей построения логических устройств на принципах функциональной электроники.

Очень перспективным представляется создание электронных моделей биологических систем на базе устройств функциональной электроники. Можно предположить, что несхемотехническая электроника динамических неоднородностей, свободная в известной мере от «тирании межсоединений», является лучшим материалом для создания таких моделей, что, в свою очередь, может стать составной частью решения проблемы создания искусственного интеллекта.

Функциональная несхемотехническая электроника не противопоставляется схемотехнической электронике. Эти два направления в интегральной электронике должны и будут развиваться одновремен-

но и параллельно, взаимно дополняя друг друга, давая возможность создавать сложные функционально интегрированные системы.

Анализ работы существующих приборов и устройств функциональной электроники (ФЭ) показывает, что они содержат следующие основные элементы.

1. Континуальную среду, представляющую собой непрерывную совокупность вещества с определенной временной, пространственной или пространственно-временной структурой в макроскопических масштабах. В ней формируются носители информационного сигнала и осуществляется их направленный перенос или хранение.

2. Динамическую неоднородность в виде локального объема на поверхности или в объеме среды с отличными от окружения характеристиками, которая создается физическими процессами. Эта неоднородность способна перемещаться по объему или поверхности континуальной среды в качестве носителя информационного сигнала. Динамическая неоднородность служит носителем информационного сигнала.

3. Устройство генерации и управления динамической неоднородностью в виде статической или динамической неоднородности, отличной от носителя информационного сигнала, позволяющее осуществить направленный перенос, обработку или хранение носителей информационного сигнала.

4. Детектор для считывания информационного сигнала и вывода его из устройства.

В настоящее время известно большое число разновидностей основных элементов устройств функциональной электроники (УФЭ). Например, в качестве континуальной среды может использоваться твердое тело с различными физическими свойствами, классифицированными по различным признакам, например, по проводимости. Могут применяться среды с другим агрегатным состоянием, что значительно расширяет круг УФЭ.

Итак, классифицируем элементы, введя иерархическую нумерацию:

1. Континуальная среда:
1.1 – твердотельная;
1.1.1 – сверхпроводящая;
1.1.2 – проводящая;
1.1.3 – полупроводниковая;
1.1.4 – диэлектрическая;

1.1.5 – магнитная;
1.2 – жидкостная;
1.3 – плазменная;
1.3.1 – газовая;
1.3.2 – твердотельная;
1.4 – газовая среда;
1.5 – автоволновая и т.п.

Приведенный перечень не претендует на полноту. Классификацию можно продолжить на следующем уровне. Например, для полупроводниковой среды можно записать:

1.1.3.1 – кремний;
1.1.3.2 – германий;
1.1.3.3 – соединения $A^{II}B^{IV}$;
1.1.3.4 – соединения $A^{III}B^{V}$ и т. д.

В качестве носителя информационного сигнала – динамических неоднородностей – используются фазовые состояния области вещества информационного тракта, например, магнитный или зарядовый домен, заряд группы подвижных носителей (зарядовый пакет), набор квантов магнитного потока (вихрей), переносимых в сверхпроводниках волной магнитного поля и т. д. Анализ возможного использования различных типов динамических неоднородностей позволяет классифицировать их следующим образом:

2. Динамическая неоднородность:
2.1 – частицы;
2.1.1 – электроны;
2.1.2 – ионы;
2.1.3 – магнитные ядра;
2.2 – зарядовые пакеты;
2.3 – квазичастицы;
2.3.1 – дырки;
2.3.2 – куперовские пары;
2.3.3 – фононы;
2.3.4 – поляроны и т. п.
2.4 – токовые шнуры;
2.5 – домены;
2.5.1 – тепловые;
2.5.2 – электрические;
2.5.3 – магнитные;
2.6 – магнитные вихри;

2.7 – волны;
2.7.1 – акустические;
2.7.2 – тепловые;
2.7.3 – электростатические;
2.7.4 – объемные зарядовые;
2.7.5 – магнитостатические;
2.7.6 – рекомбинационные;
2.7.7 – автоволны;
2.7.8 – уединенные (солитоны);
2.8 – лазерные и т. п.

Приведенный перечень также не претендует на полноту, но его структура такова, что позволяет включать новые научные сведения без нарушения структуры и смысла квалификации.

Весьма разнообразны варианты реализации устройств управления динамическими неоднородностями, роль которых могут выполнять статические неоднородности. Они могут быть классифицированы по разным независимым признакам: физическим явлениям, принципам управления движением динамическими неоднородностями, конструктивными особенностями:

3. Устройства управления:
3.1 – механическое;
3.1.1 – встречно-штыревой преобразователь;
3.2 – тепловое;
3.3 – электрическое;
3.3.1 – полевые электроды;
3.3.2 – *p*–*n*-переходы;
3.3.3 – инжектирующий диод и т. п.
3.4 – магнитное;
3.4.1 – внешние поля;
3.4.2 – аппликации;
3.5 – оптическое;
3.6 – комбинированное;
3.7 – конструктивное;
3.8 – технологическое;
3.8.1 – объемное;
3.8.2 – поверхностное.

Следует подчеркнуть, что особенностью УФЭ является интеграция эффектов в различных континуальных средах. Поэтому при разработке УФЭ в поисках оптимальных вариантов необходимо анали-

зировать различные комбинации эффектов, выбирать соответствующие конструктивно-технологические решения.

Целесообразно применять динамические неоднородности одного типа в разных средах, например, зарядовые пакеты в полупроводнике и заряды переполяризации в сегнетоэлектриках.

Детекторы информационных сигналов могут различаться по ряду независимых признаков, таких как физические явления (п. 4.1), служащие для детектирования информационного сигнала и его вывода из информационного тракта, конструктивные решения (п. 4.2), зависящие от определенного физического явления, положенного в основу работы детектора, а также технологическое исполнение (п. 4.3) детектирующего устройства. В процессе детектирования могут использоваться обратные (дополнительные) явления по отношению к явлениям, служащим для генерации динамических неоднородностей. Таким образом, существует объективно большое число вариантов элементов, образующих информационный массив, из которого разработаны реально существующие и могут быть сконструированы гипотетические УФЭ.

В данном учебном пособии рассмотрены физические основы построения устройств функциональной акустоэлектроники. Поэтому приведено изложение элементов теории упругости и вопросы распространения упругих волн в различных средах.

Контрольные вопросы

1. Каковы физические и схемотехнические ограничения интегральной электроники?
2. В чём принципиальное отличие интегральной электроники и функциональной электроники?
3. Какова взаимосвязь функциональной электроники и интегральной электроники?
4. Какие Вы знаете основные элементы устройств функциональной электроники?
5. Перечислите виды динамических неоднородностей?
6. Назовите физические процессы и явления в твердых телах, применяющиеся в приборах функциональной электроники.
7. Каковы отличительные признаки статических и динамических неоднородностей?

1. ЭЛЕМЕНТЫ ТЕОРИИ УПРУГОСТИ

1.1. Понятие деформации. Тензор деформаций. Виды деформации

Под действием различных внешних сил твердое тело может изменять свою форму и объем, то есть деформироваться. Если после снятия внешнего воздействия деформация полностью исчезает, ее называют упругой. В противном случае говорят о пластической деформации. В дальнейшем мы всегда будем иметь в виду только упругую деформацию.

В макроскопической теории сплошных сред твердое тело рассматривается как непрерывное множество бесконечно малых частиц – материальных точек. Очевидно, деформация тела наступает в результате движения его точек. Однако изменение положения точек может быть связано не только с деформированием, а также с простым перемещением и поворотом всего тела как единого целого. Но при деформировании в отличие от любого движения всего тела расстояние между его точками изменяется (рис. 1.1).

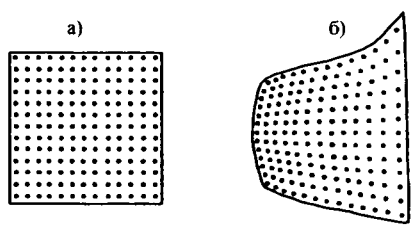


Рис. 1.1. При деформировании тела расстояние между его точками изменяется: а – расположение точек до деформации, б – после деформации

Будем определять положение каждой точки тела радиус-вектором в декартовой прямоугольной системе координат x_i ($i = 1,2,3$) . Пусть точки тела изменили свое положение. Если обозначить начальное положение некоторой точки как $\vec{r}$, а после ее перемещения – $\vec{r}'$, то можно записать

$$\vec{r}' = \vec{r} + \vec{u} ,$$

где $\vec{u}$ – вектор смещения точки. Изменение расстояния между точками в разных частях тела может отличаться, поэтому выделим в теле бесконечно малую область и рассмотрим в этой области какие-либо две точки. Вектор, соединяющий их начальные положения, обозначим как $d\vec{r}$, а после их смещения – $d\vec{r}'$. Связь между ними определится соотношением

$$d\vec{r}' = d\vec{r} + d\vec{u}, \tag{1.1}$$

где $d\vec{u}$ – разница смещений выбранных точек. В покомпонентной записи (1.1) представляется в виде

$$dx_i' = dx_i + du_i, \quad (i = 1,2,3). \tag{1.2}$$

Возведем соотношение (1.1) в квадрат

$$\left(d\vec{r}'\right)^2 = \left(d\vec{r}\right)^2 + 2d\vec{r}d\vec{u} + \left(d\vec{u}\right)^2.$$

Начальное расстояние между точками $\left|d\vec{r}\right|$ обозначим как $d\ell$, а расстояние после их смещения $\left|d\vec{r}'\right|$ – как $d\ell'$. Тогда изменение квадрата расстояния между точками

$$\left(d\ell'\right)^2 - \left(d\ell\right)^2 = 2d\vec{r}d\vec{u} + \left(d\vec{u}\right)^2. \tag{1.3}$$

Далее будем использовать известное правило сокращенного обозначения сумм. Оно заключается в следующем. Если в какой-либо индексированной величине или произведении таких величин дважды встречается парный индекс (обозначенный одной и той же буквой), это неявно подразумевает суммирование по всем возможным значениям этого индекса. Индекс суммирования называется немым. Он может быть заменен другим символом, но так, чтобы его новое обозначение не совпадало с обозначением иных возможно присутствующих в выражении непарных свободных индексов. Например, если a_{ik} – компоненты тензора второго ранга, а b_k – компоненты вектора $(i,k = 1,2,3)$, то запись $a_{ik}b_k$ означает $a_{i1}b_1 + a_{i2}b_2 + a_{i3}b_3$, что эквивалентно записи $a_{im}b_m$ $(i,m = 1,2,3)$. Отметим, что при сокра-

щенной записи символы взятия производной по координатам $\partial/\partial x_i$ понимаются как компоненты псевдовектора $\nabla = (\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3)$. Таким образом, выражение $\partial b_k/\partial x_k$ рассматривается как символическое произведение $\nabla_k b_k$ и означает сумму $\partial b_1/\partial x_1 + \partial b_2/\partial x_2 + \partial b_3/\partial x_3$.

С помощью сформулированного выше правила обозначения сумм можно записать полный дифференциал каждой компоненты вектора смещения как

$$du_i = \frac{\partial u_i}{\partial x_1} dx_1 + \frac{\partial u_i}{\partial x_2} dx_2 + \frac{\partial u_i}{\partial x_3} dx_3 = \frac{\partial u_i}{\partial x_k} dx_k = \frac{\partial u_i}{\partial x_m} dx_m . \qquad (1.4)$$

Здесь $i, k, m = 1, 2, 3$.
Заметим также, что:

$$(d\vec{u})^2 = du_1 du_1 + du_2 du_2 + du_3 du_3 = du_i du_i,$$
$$d\vec{r} d\vec{u} = dx_1 du_1 + dx_2 du_2 + dx_3 du_3 = dx_i du_i.$$

Теперь (1.3) можно придать следующий вид:

$$(d\ell')^2 - (d\ell)^2 = \frac{\partial u_i}{\partial x_k} dx_i dx_k + \frac{\partial u_i}{\partial x_k} dx_i dx_k + \frac{\partial u_i}{\partial x_k} dx_k \frac{\partial u_i}{\partial x_m} dx_m .$$

Во втором члене правой части поменяем немые индексы $i \leftrightarrow k$, а в третьем – $i \leftrightarrow m$. В результате получим

$$(d\ell')^2 - (d\ell)^2 = 2U_{ik} dx_i dx_k , \qquad (1.5)$$

где введено обозначение

$$U_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_m}{\partial x_i} \frac{\partial u_m}{\partial x_k} \right). \qquad (1.6)$$

Как правило, при упругой деформации относительное смещение точек значительно меньше расстояния между ними – $|d\vec{u}| << |d\vec{r}|$, по-

этому квадратичный член $(d\bar{u})^2$ в правой части выражения (1.2) обычно отбрасывают. В этом случае

$$U_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right). \qquad (1.7)$$

Выражениями (1.5) – (1.7) определяется изменение расстояния между бесконечно близкими точками тела при его деформировании. При любых преобразованиях координат, не изменяющих длины линейного элемента, выражение (1.5) является скалярным инвариантом. В дальнейшем мы всегда будем иметь в виду только такие преобразования координат. Так как (1.5) определяет скаляр, а дифференциалы координат dx_i – компоненты вектора, то при упомянутых преобразованиях набор девяти величин U_{ik} ведет себя как тензор второго ранга. Его называют тензором деформаций.

В общем случае деформация в различных частях тела отличается, и компоненты тензора деформаций зависят от координат. Если компоненты U_{ik} одинаковы во всем теле, то деформацию называют однородной.

Совокупность компонент тензора деформаций можно записать в виде квадратной матрицы:

$$U_{ik} = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{pmatrix}.$$

Из определений (1.6) и (1.7) легко видеть, что члены матрицы, расположенные симметрично относительно главной диагонали, попарно равны друг другу:

$$U_{ik} = U_{ki},$$

то есть тензор деформаций – симметричный тензор, и лишь шесть из девяти его компонент могут быть заданы независимо. Следовательно, любая деформация вполне описывается шестью независимыми

величинами, например, тремя диагональными компонентами U_{11}, U_{22}, U_{33} и тремя недиагональными – U_{12}, U_{13}, U_{23}.

Как и всякий симметричный тензор в некоторой бесконечно малой области тела, тензор деформаций всегда можно привести к диагональному виду. То есть в данной области тела, можно выбрать такую систему координат, в которой отличны от нуля только три диагональные компоненты – U_{11}, U_{22} и U_{33}. Эти координаты называются главными осями тензора. Поскольку деформация тела может быть неоднородной, направления главных осей в разных частях тела, вообще говоря, неодинаковы. Тот факт, что в системе главных осей недиагональные компоненты отсутствуют, не означает уменьшения числа независимых величин, требуемых для описания деформации, так как при этом добавляются три угла, определяющих положение главных осей.

Определим смысл компонент тензора деформаций.

Выделим в теле бесконечно малую прямоугольную площадку, образованную векторами $\vec{a}$ и $\vec{b}$, параллельными соответственно осям x_1, x_2 и ортогональными оси x_3 (рис. 1.2). В результате малой деформации площадки компоненты векторов $\vec{a}$ и $\vec{b}$ получат некоторые малые приращения. Для определенности на рисунке эти приращения изображены положительными.

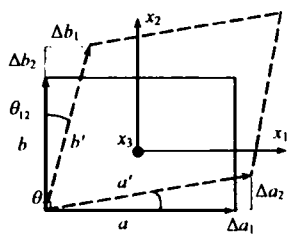


Рис. 1.2. Деформация бесконечно малой прямоугольной площадки, ортогональной оси x_3

Изменения компонент любого вектора при деформировании определяются соотношениями (1.2) и (1.4):

$$dx_i' - dx_i = \frac{\partial u_i}{\partial x_k} dx_k .$$

Учитывая, что вектор $\vec{a}$ направлен вдоль оси x_1, а вектор $\vec{b}$ – вдоль оси x_2, для приращений компонент этих векторов вдоль их первоначальных направлений получаем

$$\Delta a_1 = \frac{\partial u_1}{\partial x_1} a = U_{11} a, \quad \Delta b_2 = \frac{\partial u_2}{\partial x_2} b = U_{22} b .$$

Откуда можно заключить, что компоненты тензора деформаций U_{11} и U_{22} описывают относительное изменение расстояния вдоль соответствующих осей координат:

$$\Delta a_1 / a = U_{11}, \quad \Delta b_2 / b = U_{22} .$$

При положительных приращениях Δa_1 и Δb_2 диагональные компоненты также положительны и вдоль осей x_1, x_2 среда испытывает растяжение. Если же какое-нибудь из этих приращений (или оба) меньше нуля, то соответствующая диагональная компонента станет отрицательной и будет описывать сжатие среды.

Рассмотрим теперь изменения компонент векторов $\vec{a}$ и $\vec{b}$ в направлениях, перпендикулярных исходным:

$$\Delta a_2 = \frac{\partial u_2}{\partial x_1} a, \quad \Delta b_1 = \frac{\partial u_1}{\partial x_2} b .$$

Они приведут к повороту векторов вокруг оси x_3 в плоскости площадки на углы:

$$\theta_{21} \cong \frac{\Delta a_2}{a + \Delta a_1} \cong \frac{\Delta a_2}{a} = \frac{\partial u_2}{\partial x_1}, \quad \theta_{12} \cong \frac{\Delta b_1}{b + \Delta b_2} \cong \frac{\Delta b_1}{b} = \frac{\partial u_1}{\partial x_2} .$$

В результате чего угол между векторами $\vec{a}$ и $\vec{b}$ изменится на величину

$$\theta_{21} + \theta_{12} = \frac{\partial u_2}{\partial x_1} + \frac{\partial u_1}{\partial x_2} = 2U_{21} = 2U_{12}.$$

Если $U_{12} > 0$, то угол между векторами уменьшится, если же $U_{12} < 0$ – увеличится. Таким образом, компоненты U_{12} и U_{21} описывают изменение угла в координатной плоскости (x_1, x_2) между векторами, первоначально направленными по осям x_1, x_2.

Рассматривая деформацию элементарных площадок, ортогональных двум другим осям, можно сформулировать следующие выводы: диагональные компоненты U_{11}, U_{22}, U_{33} описывают относительные растяжения (или сжатия) параллельно осям x_1, x_2 и x_3, соответственно. Недиагональные компоненты U_{ik} $(i \neq k)$ описывают изменение угла между векторами, первоначально направленными по осям x_i и x_k, в плоскости (x_i, x_k) на величину $\theta_{ik} = 2U_{ik}$.

Здесь необходимо сделать одно пояснение. Если диагональные компоненты тензора деформаций равны нулю, то растяжения по осям координат также равны нулю. Но это вовсе не означает, что растяжение среды в других направлениях отсутствует, что можно увидеть хотя бы из рис. 1.2. Даже если приращения $\Delta a_1 = \Delta b_2 = 0$, а вместе с ними и компоненты $U_{11} = U_{22} = 0$, но $2U_{12} = \theta_{12} + \theta_{21} \neq 0$, то из-за встречного поворота векторов $\vec{a}$ и $\vec{b}$ длина диагоналей площадки изменяется. Можно найти относительное растяжение среды в произвольном направлении. В случае малой деформации

$$(d\ell')^2 - (d\ell')^2 = (d\ell' + d\ell)(d\ell' - d\ell) \cong 2d\ell(d\ell' - d\ell).$$

Подставляя это выражение в (1.5), получаем:

$$\frac{d\ell' - d\ell}{d\ell} = U_{ik} \nu_i \nu_k,$$

где $\nu_i = dx_i / d\ell$ – компоненты единичного вектора в выбранном направлении. Если, например, лишь U_{12} и U_{21} отличны от нуля, то во всех направлениях, не совпадающих с осями координат, когда ν_1 и ν_2 одновременно не равны нулю, среда испытывает растяжение. Таким образом, недиагональные компоненты тензора деформаций также описывают некоторые растяжения.

Часто бывает удобным в качестве шести независимых величин, описывающих деформацию в некоторой системе координат, использовать так называемые технические деформации. Их совокупность записывают в виде матрицы, называемой матрицей технических деформаций:

$$\begin{pmatrix} \gamma_1 & \tau_{12} & \tau_{13} \\ \tau_{12} & \gamma_2 & \tau_{23} \\ \tau_{13} & \tau_{23} & \gamma_3 \end{pmatrix} = \begin{pmatrix} U_{11} & 2U_{12} & 2U_{13} \\ 2U_{12} & U_{22} & 2U_{23} \\ 2U_{13} & 2U_{23} & U_{33} \end{pmatrix}. \qquad (1.8)$$

Для элементарного прямоугольного параллелепипеда, ребра которого направлены по координатным осям, диагональные элементы матрицы технических деформаций γ_i описывают относительное изменение длины ребер вдоль соответствующей оси. Их называют растяжениями. Недиагональные элементы τ_{ik} описывают изменение углов между ребрами. Вследствие изменения углов между ребрами противоположные грани параллелепипеда смещаются навстречу друг другу параллельно координатным плоскостям (рис. 1.3). В связи с этим недиагональные элементы τ_{ik} называют сдвигами.

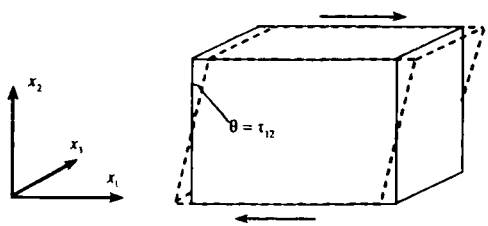


Рис. 1.3. Сдвиг граней параллелепипеда при изменении угла между его ребрами

Следует иметь в виду, что технические деформации не образуют тензора. Для их определения в какой-либо системе координат необходимо использовать компоненты тензора деформаций, соответствующие данной системе координат.

Как уже указывалось, в некоторой бесконечно малой области тела путем приведения к главным осям тензору деформаций можно придать диагональный вид. Это означает, что деформация любой бесконечно малой области тела является совокупностью трех растяжений (или сжатий) по трем, взаимно перпендикулярным направлениям.

Найдем изменение объема, обусловленное малой деформацией. Пусть ребра элементарного прямоугольного параллелепипеда направлены вдоль главных осей x_1, x_2, x_3, так что недиагональные компоненты тензора деформаций равны нулю. Если длина ребра, параллельного главной оси x_1, до деформации равна dx_1, а после деформации $-dx_1'$, то, как мы установили,

$$\left(dx_1' - dx_1\right) = U_{11}dx_1 .$$

Аналогично, для двух других ребер

$$\left(dx_2' - dx_2\right) = U_{22}dx_2 , \quad \left(dx_3' - dx_3\right) = U_{33}dx_3 .$$

Откуда:

$$dx_1' = \left(1 + U_{11}\right)dx_1 , \quad dx_2' = \left(1 + U_{22}\right)dx_2 , \quad dx_3' = \left(1 + U_{33}\right)dx_3 .$$

Перемножая эти соотношения и ограничиваясь членами, линейными по величине деформации, получаем

$$dV' = \left(1 + U_{11}\right)\left(1 + U_{22}\right)\left(1 + U_{33}\right)dV \cong \left(1 + U_{11} + U_{22} + U_{33}\right)dV ,$$

где $dV = dx_1dx_2dx_3$ – объем параллелепипеда до деформации, а $dV' = dx_1'dx_2'dx_3'$ – после деформации. Относительное изменение объема при малой деформации

$$\frac{dV' - dV}{dV} = U_{11} + U_{22} + U_{33}.$$

То есть сумма диагональных элементов тензора деформаций равна относительному изменению объема. Поскольку преобразования координат не изменяют объема, данное утверждение остается справедливым и в системе координат, в которой тензор деформаций недиагонален. Это согласуется с известным утверждением, что сумма диагональных элементов тензора является скалярным инвариантом.

Если сумма диагональных элементов во всех точках тела равна нулю, то деформация приводит к изменению формы тела без изменения его объема. Такую деформацию называют сдвигом. Если же форма тела остается подобной самой себе, меняется лишь его объем, деформацию называют всесторонним растяжением (или сжатием).

Полезно заметить, что сумма диагональных элементов тензора деформаций есть дивергенция вектора смещения:

$$\operatorname{div}\vec{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = U_{mm}.$$

Поэтому при деформации сдвига во всех точках тела $\operatorname{div}\vec{u} = 0$. Если же объем изменяется, то $\operatorname{div}\vec{u} \neq 0$.

Всякую деформацию в произвольной системе координат одновременно для всех точек тела можно представить как сумму чистого сдвига и всестороннего растяжения:

$$U_{ik} = \left(U_{ik} - \frac{1}{3}\delta_{ik}U_{mm}\right) + \frac{1}{3}\delta_{ik}U_{mm}, \qquad (1.9)$$

где δ_{ik} – единичный инвариантный тензор:

$$\delta_{ik} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}.$$

Какова бы ни была зависимость компонент тензора деформаций от координат, первый член в правой части (1.9) представляет собой

чистый сдвиг, так как сумма его диагональных элементов тождественно равна нулю, в чем легко убедиться, если учесть, что $\delta_{kk} = 3$. Второй член описывает всестороннее растяжение (или сжатие). Диагональные компоненты тензора, описывающего всестороннее растяжение, равны между собой, а недиагональные компоненты всегда отсутствуют. При этом растяжения во всех направлениях одинаковы и пропорции тела сохраняются.

1.2. Упругие силы. Тензор упругих напряжений

Из опыта известно, что, когда под влиянием внешних сил твердое тело деформируется, в нем возникают внутренние силы, противостоящие внешнему воздействию. Причину этого легко объяснить исходя из атомного строения вещества. В недеформированном твердом теле расположение атомов соответствует состоянию некоторого внутреннего равновесия. Деформация приводит к смещению атомов относительно друг друга, равновесие нарушается, и появляются межатомные силы, стремящиеся восстановить их первоначальное расположение. Эти силы называют упругими, а порождаемое ими объемное механическое напряжение – упругим напряжением. В недеформированном теле упругое напряжение отсутствует.

Существенным для теории упругости является тот факт, что межатомные силы являются близкодействующими. Каждый атом тела взаимодействует только со своим ближайшим окружением. В макроскопической теории, пренебрегающей дискретным строением вещества и рассматривающей материальную среду как непрерывную совокупность точек, радиус действия этих сил следует считать равным нулю. Проще говоря, упругое напряжение передается от каждой точки тела только к ее непосредственному окружению. Поэтому взаимодействие двух смежных частей деформированного твердого тела может быть описано распределением сил на разделяющей эти части поверхности.

Поверхностная плотность сил внутри твердого тела, находящегося в состоянии упругого напряжения, напоминает давление в покоящейся жидкости, но имеются существенные отличия. Механическое напряжение покоящейся жидкости вполне описывается скалярным давлением, так как сила давления в каждой точке во всех направле-

ниях одинакова и на любую поверхность действует в перпендикулярном направлении. Упругая реакция возникает лишь на изменение объема, но отсутствует при сдвиге. Иногда говорят, что жидкость обладает упругостью объема, но не имеет упругости формы. В твердом же теле механическое напряжение появляется не только при изменении объема, но и при деформации сдвига. Упругая сила, действующая на границе смежных частей твердого тела, в общем случае к этой границе не ортогональна, и имеет касательные компоненты. Кроме того, деформации в разных направлениях могут отличаться. Следовательно, в одной и той же точке для разных поверхностей действующие на них силы, вообще говоря, неодинаковы. Для того чтобы полностью описать механическое напряжение в некоторой точке деформированного твердого тела, необходимо найти компоненты сил сразу для всех поверхностей, содержащих данную точку.

Выделим в деформированном теле некоторую точку P и рассмотрим в этой точке элементарную площадку s. Со стороны прилегающих к площадке частей тела на нее действуют две силы (рис. 1.4).

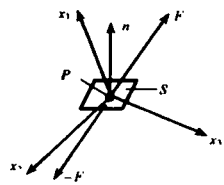


Рис. 1.4. Упругие силы, действующие на малую площадку внутри деформированного тела

По принципу равенства действия и противодействия эти силы равны по величине, но противоположны по направлению. Чтобы избежать неоднозначности и определить, о какой из них идет речь, введем единичный вектор нормали к площадке $\vec{n}$. Далее будем иметь в виду силу, действующую на площадку со стороны той части тела, для которой направление единичной нормали является внутренним.

Определим вектор площадки $\vec{s} = \vec{n}s$ с компонентами $s_k = n_k s$, где n_k – компоненты единичного вектора нормали $\vec{n}$. Размер и ориентация площадки полностью определяются тремя величинами

s_k $(k=1,2,3)$. При заданном напряжении в выбранной точке действующая на площадку сила будет зависеть только от переменных s_k:

$$\vec{F} = \vec{F}(s_1, s_2, s_3)$$

Поверхностная плотность силы на площадке запишется в виде

$$\vec{T} = \frac{d\vec{F}}{ds} = \frac{\partial \vec{F}}{\partial s_1}\frac{ds_1}{ds} + \frac{\partial \vec{F}}{\partial s_2}\frac{ds_2}{ds} + \frac{\partial \vec{F}}{\partial s_3}\frac{ds_3}{ds} = \frac{\partial \vec{F}}{\partial s_k} n_k .$$

В покомпонентной записи

$$T_i = T_{ik} n_k , \qquad (1.10)$$

где введено обозначение

$$T_{ik} = \frac{\partial F_i}{\partial s_k} .$$

На основании (1.10) можно сделать вывод, что компоненты поверхностной плотности силы в некоторой точке для произвольно ориентированной элементарной площадки можно найти, если в этой точке заданы девять величин T_{ik}. Так как T_i и n_k – компоненты векторов, совокупность девяти величин T_{ik} образует тензор второго ранга, который называют тензором упругих напряжений или просто тензором напряжений:

$$T_{ik} = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}.$$

По определению, T_{ik} есть i-тая компонента поверхностной плотности силы на элементарной площадке, перпендикулярной k-му направлению. Например, на единицу площади поверхности, ортогональной оси x_1, действует одна нормальная сила T_{11} и две тангенци-

альные силы: T_{21} (вдоль оси x_2) и T_{31} (вдоль x_3). По определению компонента T_{ik} будет положительной, если при положительной нормали вдоль оси x_k упругая сила, описываемая этой компонентой, направлена в положительном направлении оси x_i.

В общем случае упругое напряжение в различных точках тела может отличаться, и компоненты T_{ik} зависят от координат. Если компоненты тензора T_{ik} одинаковы во всех точках тела, напряжение называют однородным.

Покажем, что тензор напряжений – симметричный тензор. Выделим в деформированном теле небольшой кубик. Пусть ребра кубика имеют длину h и направлены по осям координат (рис. 1.5). На каждую грань кубика со стороны его окружения действуют одна нормальная и две сдвиговые силы. Для определенности все они выбраны положительными. Если длина ребра кубика достаточно мала, то изменением компонент тензора T_{ik} внутри него можно пренебречь. Тогда упругие силы на противоположных гранях кубика одинаковы по величине, но имеют встречное направление.

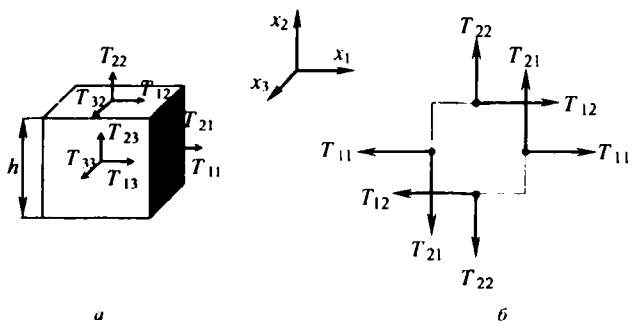


Рис. 1.5. Компоненты поверхностной плотности упругих сил, действующих на грани кубического объема (*а*) и на его сечение, ортогональное оси x_3 (*б*)

На каждое сечение кубика действует момент сил, создаваемый сдвиговыми компонентами тензора напряжений. Например, на сечение, ортогональное оси x_3, действует момент сил относительно центра кубика $h^3(T_{12}-T_{21})$. Это должно приводить к вращению кубика

вокруг оси x_3 с угловым ускорением, которое по законам механики равно отношению момента сил к моменту инерции. Момент инерции пропорционален произведению массы кубика на квадрат его ребра. Поскольку масса кубика пропорциональна третьей степени ребра, то момент инерции будет величиной порядка h^5, а угловое ускорение, таким образом, величиной порядка $1/h^2$. При стремлении размеров кубика к нулю угловое ускорение неограниченно возрастает. Поскольку по физическим причинам это невозможно, момент сил должен равняться нулю, откуда $T_{12} = T_{21}$. Повторяя такие же рассуждения для сечений, ортогональных двум другим осям координат, заключаем, что тензор напряжений является симметричным:

$$T_{ik} = T_{ki}.$$

Приведенные выше рассуждения, конечно же, не означают, что момент сил, создаваемый упругим напряжением, всегда равен нулю. Соотношения симметрии выражают свойство компонент T_{ik}, взятых в одной и той же точке. При неоднородном напряжении и конечных размерах кубика из-за различия компонент тензора T_{ik} на противоположных гранях момент сил может быть не равен нулю.

Ввиду установленных свойств симметрии из девяти компонент тензора напряжений только шесть могут быть заданы независимо. Следовательно, любое упругое напряжение вполне описывается шестью независимыми величинами. В системе главных осей сдвиговые компоненты исчезают, и тензор напряжений приобретает диагональный вид. Так как упругое напряжение может быть неоднородным, положение главных осей в разных частях тела, вообще говоря, различно.

1.3. Закон Гука. Тензор модулей упругости. Сокращенная запись упругих модулей

Как уже говорилось, деформирование твердого тела приводит к появлению в нем упругого напряжения. Следовательно, между компонентами тензоров напряжений T_{ik} и деформаций U_{ik} должна существовать некоторая связь.

Из многочисленных наблюдений известно, что при малой деформации твердых тел силы упругой реакции пропорциональны величине деформации. Так, при небольшом продольном растяжении упругого изотропного стержня возникающее осевое напряжение T пропорционально относительному удлинению $\Delta\ell/\ell$. Этот эмпирический факт известен как закон Гука:

$$T = C\Delta\ell/\ell\,. \qquad (1.11)$$

Коэффициент пропорциональности C характеризует упругость материала стержня. Обратную ему величину $S = 1/C$ называют упругой податливостью материала.

Рассматривая малую неоднородную деформацию твердого тела сложной формы, логично предположить, что в любой бесконечно малой области тела закон Гука выполняется. Это значит, что если в этой области имеется некоторое малое растяжение, связанное с какой-либо компонентой тензора деформаций U_{ik}, то возникающие упругие силы T_{ik} должны быть пропорциональны этой компоненте. Ввиду того, что упругость твердой среды в разных направлениях часто бывает неодинаковой, коэффициенты пропорциональности для разных растяжений могут отличаться. Если имеется нескольких ненулевых U_{ik}, полное напряжение будет складываться из напряжений, создаваемых каждой из компонент U_{ik}. Другими словами, следует предположить, что в любой точке тела каждая компонента тензора напряжений должна быть линейной комбинацией всех компонент тензора деформаций:

$$T_{ik} = C_{iklm}U_{lm}, \quad (i,k,l,m = 1,2,3)\,. \qquad (1.12)$$

По существу, девять уравнений (1.12) являются естественным обобщением соотношения (1.11) на случай малой деформации произвольной упругой среды. В связи с этим (1.12) иногда называют обобщенным законом Гука, а чаще просто законом Гука. Коэффициенты C_{iklm} характеризуют упругость данного материала на тот или иной вид деформации, их называют упругими модулями.

Так как T_{ik} и U_{lk} – компоненты тензоров второго ранга, совокупность модулей C_{iklm} образует тензор четвертого ранга, называемый тензором модулей упругости. Иногда его также называют просто тензором упругости. Из уравнений закона Гука (1.12) компоненты тензора модулей упругости C_{iklm} могут быть найдены дифференцированием компонент тензора напряжений по компонентам тензора деформаций:

$$C_{iklm} = \frac{\partial T_{ik}}{\partial U_{lm}}. \qquad (1.13)$$

Как всякий тензор четвертого ранга, тензор модулей упругости имеет 3^4=81 компоненту. Однако ввиду установленных ранее свойств симметрии компонент T_{ik} и U_{lm} упругие модули C_{iklm} должны не изменяться при перестановках $i \leftrightarrow k$ и $l \leftrightarrow m$:

$$C_{iklm} = C_{kilm} = C_{ikml} = C_{ikml}. \qquad (1.14)$$

Следовательно, из 81 упругого модуля различными могут быть не более 36. Основываясь на этом, совокупность упругих модулей записывают в виде квадратной матрицы $\mathbf{C}_{\alpha\beta}$ $(\alpha,\beta = 1,2,3,4,5,6)$, называемой матрицей упругих постоянных:

$$\mathbf{C}_{\alpha\beta} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} & \mathbf{C}_{14} & \mathbf{C}_{15} & \mathbf{C}_{16} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{C}_{23} & \mathbf{C}_{24} & \mathbf{C}_{25} & \mathbf{C}_{26} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{C}_{33} & \mathbf{C}_{34} & \mathbf{C}_{35} & \mathbf{C}_{36} \\ \mathbf{C}_{41} & \mathbf{C}_{42} & \mathbf{C}_{43} & \mathbf{C}_{44} & \mathbf{C}_{45} & \mathbf{C}_{46} \\ \mathbf{C}_{51} & \mathbf{C}_{52} & \mathbf{C}_{53} & \mathbf{C}_{54} & \mathbf{C}_{55} & \mathbf{C}_{56} \\ \mathbf{C}_{61} & \mathbf{C}_{62} & \mathbf{C}_{63} & \mathbf{C}_{64} & \mathbf{C}_{65} & \mathbf{C}_{66} \end{pmatrix}. \qquad (1.15)$$

Ее члены получаются путем замены в компонентах тензора упругости C_{iklm} каждой пары перестанавливаемых индексов, пробегающих значения от 1 до 3, единственным индексом, изменяющимся от 1 до 6, по правилу:

$ik \to \alpha,\ \ lm \to \beta$

$11 \to 1\ \ 22 \to 2\ \ 33 \to 3\ \ 23,32 \to 4\ \ 13,31 \to 5\ \ 12,21 \to 6.$

Каждая упругая постоянная $\mathbf{C}_{\alpha\beta}$ в зависимости от ее индексов равна одному, двум или четырем неразличимым модулям:

$$\mathbf{C}_{\alpha\beta} = C_{iklm}, \qquad \text{когда } i = k,\ l = m,\ \ \alpha,\beta = 1,2,3;$$

$$\mathbf{C}_{\alpha\beta} = C_{iklm} = C_{ikml}, \qquad \text{когда } i = k,\ l \neq m,\ \ \alpha = 1,2,3,\ \beta = 4,5,6;$$

$$\mathbf{C}_{\alpha\beta} = C_{iklm} = C_{kilm}, \qquad \text{когда } i \neq k,\ l = m,\ \ \alpha = 4,5,6,\ \beta = 1,2,3;$$

$$\mathbf{C}_{\alpha\beta} = C_{iklm} = C_{kilm} = C_{ikml} = C_{kiml}, \qquad \text{когда } i \neq k,\ l \neq m,\ \ \alpha,\beta = 4,5,6.$$

Например:

$$\mathbf{C}_{11} = C_{1111},$$
$$\mathbf{C}_{24} = C_{2223} = C_{2232},$$
$$\mathbf{C}_{65} = C_{1213} = C_{1231} = C_{2113} = C_{2131}.$$

Заметим, что вследствие симметрии тензора напряжений из девяти уравнений закона Гука (1.12) только шесть являются различными. Матричная запись упругих модулей позволяет заменить девять уравнений закона Гука (1.12) шестью независимыми соотношениями и придать им сокращенный вид. Это можно сделать следующим образом.

Пользуясь симметрией тензора напряжений, с помощью сформулированного выше правила замены перестанавливаемых индексов определим шесть независимых компонент напряжения как последовательность величин $\mathbf{T}_\alpha \ \ (\alpha = 1,2,3,4,5,6)$:

$$\mathbf{T}_1 = T_{11},\ \mathbf{T}_2 = T_{22},\ \mathbf{T}_3 = T_{33},$$
$$\mathbf{T}_4 = T_{23} = T_{32},\ \mathbf{T}_5 = T_{13} = T_{31},\ \mathbf{T}_6 = T_{12} = T_{21}. \qquad (1.16)$$

Компоненты $\mathbf{T}_1$, $\mathbf{T}_2$ и $\mathbf{T}_3$ описывают растягивающие напряжения по осям координат, а $\mathbf{T}_4$, $\mathbf{T}_5$ и $\mathbf{T}_6$ – независимые сдвиговые напряжения на координатных плоскостях.

Так как равные друг другу недиагональные компоненты тензора деформаций U_{lm} $(l \neq m)$ в правой части (1.12) повторяются дважды с одинаковыми коэффициентами, удобно в качестве независимых компонент деформации $\mathbf{U}_\alpha$ $(\alpha = 1,2,3,4,5,6)$ использовать технические деформации (1.8):

$$\begin{aligned} &\mathbf{U}_1 = U_{11},\ \mathbf{U}_2 = U_{22},\ \mathbf{U}_3 = U_{33}, \\ &\mathbf{U}_4 = 2U_{23},\ \mathbf{U}_5 = 2U_{13},\ \mathbf{U}_6 = 2U_{12}. \end{aligned} \tag{1.17}$$

Величины $\mathbf{U}_1$, $\mathbf{U}_2$ и $\mathbf{U}_3$ определяют растяжения по координатным осям, а $\mathbf{U}_4$, $\mathbf{U}_5$ и $\mathbf{U}_6$ – встречный сдвиг слоев, параллельных координатным плоскостям. Тогда девять уравнений закона Гука заменяется шестью независимыми соотношениями:

$$\mathbf{T}_\alpha = \mathbf{C}_{\alpha\beta}\mathbf{U}_\beta, \quad (\alpha,\beta = 1,2,3,4,5,6). \tag{1.18}$$

Закон Гука для произвольной упругой среды можно также представить в виде зависимости компонент тензора деформаций от величины приложенных напряжений:

$$U_{ik} = S_{iklm}T_{lm}, \tag{1.19}$$

где S_{iklm} – называют тензором модулей упругой податливости. Очевидно, компоненты S_{iklm} должны обладать теми же свойствами симметрии, что и C_{iklm}, и лишь 36 из 81 модуля упругой податливости могут быть различными. Поэтому их также можно записать в виде квадратной матрицы $\mathbf{S}_{\alpha\beta}$ $(\alpha,\beta = 1,2,3,4,5,6)$. Однако для сокращенной записи (1.19) способ определения элементов $\mathbf{S}_{\alpha\beta}$ будет другим. Во-первых, в каждом уравнении (1.19) сдвиговые компоненты тензора напряжений T_{lm} в силу симметрии повторяются дважды с одинаковыми коэффициентами. Значит, при сокращенной записи, когда пара сдвиговых сил $T_{lm}(l \neq m)$ заменяется одной компонентой $\mathbf{T}_\beta$ $(\beta = 4,5,6)$, коэффициент при ней должен быть равен удвоенному значению соответствующего модуля. Во-вторых, при замене в левой

части (1.19) недиагональных компонент $U_{ik}\,(i \neq k)$ на технические деформации $\mathbf{U}_\alpha\ (\alpha = 4,5,6)$ правая часть должна быть увеличена вдвое. На основании этих рассуждений, чтобы при сокращенной записи соотношений (1.19) избежать появления дополнительных численных коэффициентов, элементы матрицы $\mathbf{S}_{\alpha\beta}\ (\alpha,\beta = 1,2,3,4,5,6)$ определяют по следующим правилам:

$$\begin{aligned}
&\mathbf{S}_{\alpha\beta} = S_{iklm}, && \text{если } i = k,\ l = m,\ \alpha,\beta = 1,2,3;\\
&\mathbf{S}_{\alpha\beta} = 2S_{iklm}, && \text{если } i = k,\ l \neq m,\ \alpha = 1,2,3,\ \beta = 4,5,6;\\
&\mathbf{S}_{\alpha\beta} = 2S_{iklm}, && \text{если } i \neq k,\ l = m,\ \alpha = 4,5,6,\ \beta = 1,2,3;\\
&\mathbf{S}_{\alpha\beta} = 4S_{iklm}, && \text{если } i \neq k,\ l \neq m,\ \alpha,\beta = 4,5,6.
\end{aligned}$$

Тогда девять соотношений (1.19) заменятся шестью независимыми уравнениями:

$$\mathbf{U}_\alpha = \mathbf{S}_{\alpha\beta}\mathbf{T}_\beta,\quad \alpha,\beta = 1,2,3,4,5,6\,.$$

Используя сокращенную запись, необходимо всегда иметь в виду, что матрицы $\mathbf{C}_{\alpha\beta}$ и $\mathbf{S}_{\alpha\beta}$ не образуют тензоров, а компоненты $\mathbf{T}_\alpha$ и $\mathbf{U}_\alpha$ – векторов. Для их определения в некоторой системе координат следует использовать компоненты тензоров C_{iklm}, S_{iklm}, T_{ik} и U_{ik}, соответствующие данной системе координат. Во избежание путаницы в дальнейшем для обозначения компонент векторов и тензоров будем использовать латинские индексы, изменяющиеся от 1 до 3, а для обозначения независимых компонент напряжения и деформации, а также членов матриц, описывающих упругие свойства, – греческие, изменяющиеся от 1 до 6.

1.4. Упругая энергия. Свойства матрицы упругих постоянных

Упругой энергией называют потенциальную энергию деформированного тела, связанную с наличием в его объеме упругих напряжений. По сути, это дополнительная потенциальная энергия меж-

атомных сил, возникающая из-за смещения атомов из положения равновесия.

Выделим в деформированном теле кубик, ребра которого имеют длину h и направлены вдоль осей координат (рис. 1.6). Размеры кубика будем считать настолько малыми, чтобы считать деформацию и напряжение однородными.

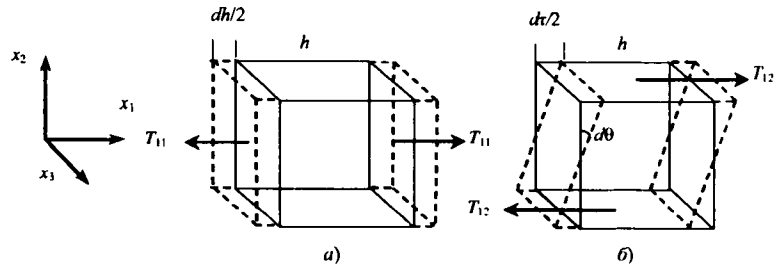


Рис. 1.6. Однородное растяжение (*а*) и сдвиг (*б*) малого кубического объема

Как указывалось ранее, любая деформация малой области тела представляет собой совокупность шести независимых деформаций $\mathbf{U}_\alpha$ (1.17). Компоненты $\mathbf{U}_1$, $\mathbf{U}_2$ и $\mathbf{U}_3$ описывают растяжения по координатным осям, а $\mathbf{U}_4$, $\mathbf{U}_5$ и $\mathbf{U}_6$ – относительный сдвиг слоев, параллельных координатным плоскостям. Если каждая из величин $\mathbf{U}_\alpha$ получает малое приращение, потенциальная энергия кубика изменяется. Это изменение будет равно работе сил, действующих на грани кубика со стороны его окружения.

Пусть длина ребер вдоль оси x_1 увеличилась на малую величину dh (рис. 1.6, *а*). Обе грани кубика, перпендикулярные оси x_1, получат смещение $dh/2$. Одна из граней сместится в положительном направлении оси x_1, а другая – в отрицательном. В результате этого компонента $\mathbf{U}_1$ получит приращение:

$$d\mathbf{U}_1 = dh/h .$$

Остальные компоненты $\mathbf{U}_\alpha$ не изменят своего значения.

Работа равна произведению перемещения на составляющую силы вдоль направления перемещения. Компоненты сил на гранях кубика определяются соотношением (1.11). Рассматриваются силы, действующие на кубик со стороны внешнего окружения. Поэтому для грани, смещающейся в положительном направлении, вектор единичной нормали имеет компоненты $(1,0,0)$, а на нее в направлении перемещения действует сила $T_{11}h^2$. Для противоположной грани компоненты единичной нормали (–1,0,0), а действующая на нее сила равна $\left(-T_{11}h^2\right)$. Сдвиговые напряжения на обеих гранях перпендикулярны направлению перемещения и работы не совершают. Остальные напряжения описывают силы, приложенные к остальным четырем граням. Но центры этих граней не испытывают перемещения, изменяется лишь их площадь, и, значит, работа данных сил также равна нулю. Следовательно, работа, связанная с приращением $\mathbf{U}_1$, может быть представлена в виде

$$dA_1 = T_{11}h^2\, dh/2 - T_{11}h^2\left(-dh/2\right) = T_{11}h^3 d\mathbf{U}_1 .$$

Если использовать независимые компоненты напряжения $\mathbf{T}_\alpha$, определенные соотношениями (1.16), то эту работу можно также записать как

$$dA_1 = h^3 \mathbf{T}_1 d\mathbf{U}_1 .$$

Выполняя аналогичные рассуждения для граней, ортогональных двум другим осям, получим:

$$dA_2 = h^3 \mathbf{T}_2 d\mathbf{U}_2 ,$$
$$dA_3 = h^3 \mathbf{T}_3 d\mathbf{U}_3 .$$

Рассмотрим, далее, сдвиг граней, перпендикулярных оси x_2, во встречных направлениях параллельно x_1 на величину $d\tau/2$ (рис. 1.6, *б*). Этот сдвиг уменьшит угол между ребрами кубика, первоначально направленными вдоль осей x_1, x_2. Уменьшение угла определяется компонентой $\mathbf{U}_6$:

$$d\mathbf{U}_6 = d\theta = \frac{d\tau}{2}.$$

Остальные величины $\mathbf{U}_\alpha$ останутся неизменными.

Для верхней грани вектор единичной внешней нормали имеет компоненты (0,1,0), а для нижней – (0,-1,0). В направлении сдвига верхней грани, таким образом, действует сила $T_{12}h^2$, а в направлении сдвига нижней грани – сила $\left(-T_{12}h^2\right)$. Суммарная работа этих сил

$$dA_6 = T_{12}h^2\left(d\tau/2\right) - T_{12}h^2\left(-d\tau/2\right) = h^3\mathbf{T}_6 d\mathbf{U}_6.$$

Компоненты T_{22} и T_{32}, также описывающие напряжения на рассматриваемых гранях, работы не совершают, так как они ортогональны перемещению. Остальные напряжения приложены к другим граням, центры которых не испытывают перемещения, поэтому работа этих напряжений равна нулю. Аналогичные сдвиги, перпендикулярные двум другим осям, приводят к соотношениям:

$$dA_5 = h^3\mathbf{T}_5 d\mathbf{U}_5,$$
$$dA_4 = h^3\mathbf{T}_4 d\mathbf{U}_4.$$

Суммируя все рассмотренные составляющие работы, получаем выражение, описывающее изменения упругой энергии кубика при любой деформации:

$$dA_1 + dA_2 + \ldots + dA_6 = h^3\mathbf{T}_\alpha d\mathbf{U}_\alpha.$$

Если обозначить объемную плотность упругой энергии тела как Φ, то ее изменение представляется в виде

$$d\Phi = \mathbf{T}_\alpha d\mathbf{U}_\alpha. \qquad (1.20)$$

В тензорной записи:

$$d\Phi = T_{ik}\, dU_{ik} \,. \qquad (1.21)$$

При упругой обратимой деформации энергия определяется только текущим значением компонент деформации и не зависит от того, как они изменялись, поэтому выражение (1.20) представляет собой полный дифференциал функции Φ, как функции независимых переменных $\mathbf{U}_\alpha$. Компоненты напряжения $\mathbf{T}_\alpha$ находятся дифференцированием объемной плотности упругой энергии Φ по компонентам деформации:

$$\mathbf{T}_\alpha = \frac{\partial \Phi}{\partial \mathbf{U}_\alpha} \,, \qquad (1.22)$$

или в тензорной записи:

$$T_{ik} = \frac{\partial \Phi}{\partial U_{ik}} \,. \qquad (1.23)$$

Если закон Гука выполняется, то из (1.18) и (1.22) получаем

$$\mathbf{C}_{\alpha\beta}\mathbf{U}_\beta = \frac{\partial \Phi}{\partial \mathbf{U}_\alpha} \,.$$

Непосредственной проверкой можно убедиться, что дифференцирование этого выражения по $\mathbf{U}_\beta$ приводит к соотношению

$$\mathbf{C}_{\alpha\beta} = \frac{\partial^2 \Phi}{\partial \mathbf{U}_\beta \partial \mathbf{U}_\alpha} \,.$$

Так как упругая энергия зависит только от текущих значений $\mathbf{U}_\alpha$, а не от того, как они изменялись, порядок дифференцирования можно изменить. Откуда вытекает, что матрица упругих постоянных – симметричная матрица

$$\mathbf{C}_{\alpha\beta} = \mathbf{C}_{\beta\alpha} \,,$$

а компоненты тензора модулей упругости должны удовлетворять следующим соотношениям симметрии:

$$C_{iklm} = C_{lmik} . \tag{1.24}$$

Вследствие симметрии матрицы упругих постоянных только 21 из 36 ее членов могут быть различными. Таким образом, для описания упругих свойств любой среды требуется не более 21 упругой постоянной.

Найдем явное выражение объемной плотности упругой энергии Φ через компоненты $\mathbf{U}_\alpha$ для деформации, подчиняющейся закону Гука. Подставим (1.18) в (1.20):

$$d\Phi = \mathbf{C}_{\alpha\beta}\mathbf{U}_\beta d\mathbf{U}_\alpha .$$

Проинтегрируем это выражение:

$$\Phi = \int \mathbf{C}_{\alpha\beta}\mathbf{U}_\alpha d\mathbf{U}_\beta = \frac{1}{2}\int \mathbf{C}_{\alpha\beta}\left(\mathbf{U}_\alpha d\mathbf{U}_\beta + \mathbf{U}_\beta d\mathbf{U}_\alpha\right) = \frac{1}{2}\int \mathbf{C}_{\alpha\beta} d\left(\mathbf{U}_\alpha \mathbf{U}_\beta\right).$$

Окончательно получаем

$$\Phi = \frac{1}{2}\mathbf{C}_{\alpha\beta}\mathbf{U}_\alpha\mathbf{U}_\beta . \tag{1.25}$$

Произвольная константа интегрирования отсутствует, так как упругая энергия недеформированной среды должна быть равна нулю. Данное выражение по существу обобщает известную из элементарной физики формулу, выражающую энергию растянутой пружины W через ее упругость k и величину растяжения $\Delta\ell$:

$$W = \frac{1}{2}k\left(\Delta\ell\right)^2 .$$

Объемная плотность упругой энергии при любой деформации должна быть положительной, в противном случае среда становилась бы неустойчивой. Это накладывает определенные ограничения на возможные значения упругих постоянных. Данные ограничения вы-

текают из условий положительной определенности симметричной квадратичной формы (1.25). Как известно из линейной алгебры, для того чтобы симметричная квадратичная форма (1.25) была положительно определена, ее коэффициенты должны удовлетворять условиям:

$$\mathbf{C}_{11} > 0,\ \begin{vmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12} & \mathbf{C}_{22} \end{vmatrix} > 0,\ \begin{vmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{12} & \mathbf{C}_{22} & \mathbf{C}_{23} \\ \mathbf{C}_{13} & \mathbf{C}_{23} & \mathbf{C}_{33} \end{vmatrix} > 0,\ \ldots,\ \det\left[\mathbf{C}_{\alpha\beta}\right] > 0\,. \quad (1.26)$$

Из (1.26), в частности, следует, что диагональные члены матрицы упругих постоянных всегда положительны.

Необходимо отметить, что соотношение (1.25) нуждаются в одном замечании. Записывая приращение объемной плотности упругой энергии (1.20), мы не учитывали изменение объема при деформировании. Строго говоря, формула (1.25) определяет упругую энергию, приходящуюся на единицу объема среды в недеформированном состоянии. Выражение, описывающее объемную плотность энергии деформированной среды, должно содержать поправку на изменение объема. Но так как изменение объема равно сумме диагональных элементов тензора деформаций, то упомянутая поправка будет пропорциональна третьей степени деформаций. В рассматриваемом приближении малых деформаций этой поправкой можно пренебречь.

1.5. Упругие свойства кубического кристалла и изотропной среды

Выражение (1.25), определяющее упругую энергию единицы объема произвольной слабодеформированной среды, в тензорной записи имеет вид

$$\Phi = \frac{1}{2} C_{iklm} U_{ik} U_{lm}\,, \quad (1.27)$$

где упругие модули удовлетворяют соотношениям симметрии (1.14) и (1.24):

$$C_{iklm} = C_{kilm} = C_{kiml} = C_{ikml} \text{ и } C_{iklm} = C_{lmik} . \qquad (1.28)$$

Упругая энергия является скалярным инвариантом, то есть не изменяет своего значения при любых преобразованиях координат. Если среда обладает некоторой внутренней симметрией, существуют соответствующие этой симметрии преобразования координат, при которых свойства среды не изменяются. Это значит, что при таких преобразованиях могут изменяться только компоненты тензора деформаций U_{ik}, а упругие модули C_{iklm} должны оставаться теми же самыми.

Рассмотрим кубический кристалл. Выберем систему координат, оси которой x_1, x_2, x_3 совпадают с ребрами элементарной ячейки кристалла. Симметрия кубического кристалла такова, что его свойства не изменяются при зеркальном отражении относительно любой из координатных плоскостей, а также повороте кристалла на 90° вокруг каждой из осей.

Однократное зеркальное отражение относительно координатной плоскости означает инверсию координатной оси, ортогональной данной плоскости. Инверсия оси приводит к смене знаков некоторых компонент тензора деформаций. Например, при отражении относительно плоскости (x_2, x_3) компоненты U_{12} и U_{13} сменят знак, а компонента U_{11} и другие компоненты, не имеющие среди индексов единиц, останутся неизменными. Произведения $U_{ik}U_{lm}$, имеющие в наборе индексов нечетное число единиц, сменят знак. Напротив, знак произведений $U_{ik}U_{lm}$, имеющих среди индексов четное число единиц, не изменится. Для того чтобы при зеркальном отражении относительно плоскости (x_2, x_3) энергия (1.27) сохранялась, необходимо потребовать, чтобы все модули C_{iklm}, имеющие в наборе индексов нечетное число единиц, равнялись нулю.

Выполняя аналогичные рассуждения для зеркального отражения относительно двух других плоскостей, приходим к выводу, что все компоненты тензора C_{iklm}, имеющие нечетное число одинаковых индексов, равны нулю. В результате нулевыми будут только 21 мо-

дуль. Среди них в силу общих свойств симметрии (1.28) только 9 являются независимыми. В качестве независимых модулей можно выбрать, например,

$$C_{1111}, C_{2222}, C_{3333}, C_{1122}, C_{1133}, C_{2233}, C_{1212}, C_{1313}, C_{2323}. \qquad (1.29)$$

Нетрудно убедиться, что матрица упругих постоянных кубического кристалла имеет лишь 12 ненулевых элементов.

Рассмотрим далее последовательность преобразований, состоящую из поворота на 90° вокруг x_3 и последующего отражения относительно плоскости (x_1, x_3):

$$(x_1, x_2, x_3) \to (-x_2, x_1, x_3) \to (x_2, x_1, x_3).$$

Данная последовательность преобразований, очевидно, сводится к замене $x_1 \leftrightarrow x_2$. Для сохранения величины упругой энергии необходимо потребовать выполнения равенств:

$$C_{1111} = C_{2222}, \; C_{1133} = C_{2233}, \; C_{1313} = C_{2323}.$$

Аналогично, замена $x_2 \leftrightarrow x_3$ приводит к дополнительным равенствам:

$$C_{2222} = C_{3333}, \; C_{1133} = C_{1122}, \; C_{1212} = C_{1313}.$$

Замена $x_1 \leftrightarrow x_3$ не порождает каких-либо новых соотношений между упругими модулями. В результате, лишь 3 модуля из последовательности (1.29) оказываются независимыми.

Таким образом, упругие свойства кубического кристалла вполне описываются тремя независимыми постоянными. В качестве независимых постоянных принято использовать $\mathbf{C}_{11} = C_{1111}$, $\mathbf{C}_{44} = C_{2323}$ и $\mathbf{C}_{12} = C_{1122}$. Матрица упругих постоянных кубического кристалла имеет вид

$$\begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{C}_{12} & 0 & 0 & 0 \\ \mathbf{C}_{12} & \mathbf{C}_{11} & \mathbf{C}_{12} & 0 & 0 & 0 \\ \mathbf{C}_{12} & \mathbf{C}_{12} & \mathbf{C}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{C}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{C}_{44} \end{pmatrix},$$

а объемная плотность упругой энергии:

$$\Phi = \frac{1}{2}\Big[\mathbf{C}_{11}\left(\mathbf{U}_1^2 + \mathbf{U}_2^2 + \mathbf{U}_3^2\right) + \mathbf{C}_{44}\left(\mathbf{U}_4^2 + \mathbf{U}_5^2 + \mathbf{U}_6^2\right) + \\ + 2\mathbf{C}_{12}\left(\mathbf{U}_1\mathbf{U}_2 + \mathbf{U}_1\mathbf{U}_3 + \mathbf{U}_2\mathbf{U}_3\right)\Big].$$

Условия положительной определенности упругой энергии (1.26) для кубического кристалла приводят к соотношениям:

$$\mathbf{C}_{11} > |\mathbf{C}_{12}|, \quad \mathbf{C}_{11} + 2\mathbf{C}_{12} > 0, \quad \mathbf{C}_{44} > 0.$$

Рассмотрим теперь упругие свойства изотропной среды. В изотропной упругой среде выбор направления осей координат совершенно произволен. Любые преобразования координат не должны требовать изменения упругих модулей. Это значит, что объемная плотность упругой энергии должна представлять собой линейную комбинацию скалярных инвариантов второй степени относительно компонент тензора деформаций. Из компонент тензора можно составить только два линейно независимых инварианта второй степени. Один из них – квадрат суммы диагональных элементов $(U_{mm})^2$, другой – сумма квадратов всех элементов $U_{ik}U_{ik}$. Поэтому для изотропной среды можно записать

$$\Phi = \frac{\lambda}{2}(U_{mm})^2 + \mu U_{ik}U_{ik}. \qquad (1.30)$$

Из (1.30) следует, что для описания упругих свойств изотропной среды достаточно всего двух независимых постоянных λ и μ. Их называют коэффициентами Лямэ.

Если использовать компоненты напряжения $\mathbf{U}_\alpha$, то (1.30) можно придать следующий вид:

$$\Phi = \frac{1}{2}\Big[(2\lambda+\mu)\left(\mathbf{U}_1^2+\mathbf{U}_2^2+\mathbf{U}_3^2\right)+\mu\left(\mathbf{U}_4^2+\mathbf{U}_5^2+\mathbf{U}_6^2\right)+$$
$$+2\lambda\left(\mathbf{U}_1\mathbf{U}_2+\mathbf{U}_1\mathbf{U}_3+\mathbf{U}_2\mathbf{U}_3\right)\Big].$$

Сопоставляя это выражение с выражением для энергии кубического кристалла, получаем:

$$\mathbf{C}_{11}=2\mu+\lambda,\ \mathbf{C}_{12}=\lambda,\ \mathbf{C}_{44}=\left(\mathbf{C}_{11}-\mathbf{C}_{12}\right)/2=\mu.$$

Нетрудно видеть, что матрица упругих постоянных для изотропной среды имеет те же ненулевые элементы, что и для кубического кристалла, но постоянная $\mathbf{C}_{44}$ уже не является независимой. В связи с этим величину $\eta = 2\mathbf{C}_{44}/\left(\mathbf{C}_{11}-\mathbf{C}_{12}\right)$ часто используют как меру анизотропии кубического кристалла. Для изотропной среды $\eta = 1$, а в кубических кристаллах η может быть как больше, так и меньше единицы.

Дифференцируя (1.30) по компонентам тензора деформаций, на основании (1.23) получаем закон Гука для изотропной среды:

$$T_{ik}=\lambda\delta_{ik}U_{mm}+2\mu U_{ik}. \qquad (1.31)$$

В разделе 1.1 мы отмечали, что всякая деформация может быть представлена в виде чистого сдвига и всестороннего растяжения (1.9). Используя это, соотношению (1.30) можно придать следующий вид:

$$\Phi=\frac{K_V}{2}U_{ik}U_{ik}+\mu\left(U_{ik}-\frac{1}{3}\delta_{ik}U_{mm}\right)^2, \qquad (1.32)$$

где введено обозначение

$$K_V = \frac{3\lambda + 2\mu}{3}.$$

Так как упругая энергия положительная величина при любой деформации, то из (1.32) заключаем, что постоянные μ и K_V – положительные величины. Знак коэффициента λ не может быть определен из энергетических представлений.

Дифференцируя (1.32) по U_{ik}, получаем закон Гука для изотропной среды в виде

$$T_{ik} = K_V U_{mm}\delta_{ik} + 2\mu\left(U_{ik} - \frac{1}{3}\delta_{ik}U_{mm}\right). \qquad (1.33)$$

При чистом сдвиге сумма диагональных элементов тензора деформаций равна нулю, и возникающие напряжения пропорциональны коэффициенту μ, который на этом основании называется модулем сдвига. При объемном расширении изотропной среды без сдвига второй член в правой части (1.33) равен нулю. Постоянную K_V в связи с этим называют модулем объемной упругости. Таким образом, модуль K_V описывает упругость объема, а модуль сдвига μ – упругость формы.

В принципе, для описания упругих свойств изотропной среды достаточно любой пары независимых констант, например, членов матрицы упругих постоянных, либо коэффициентов Лямэ, либо модулей сдвига и объемной упругости. Выбор определяется традицией или удобством, вытекающим из практического содержания решаемой задачи. Во многих приложениях упругие свойства изотропных материалов описывают модулем Юнга и коэффициентом Пуассона. Их определяют следующим образом.

Рассмотрим брусок, сделанный из упругого изотропного материала. Пусть он подвергнут малому однородному растяжению вдоль своей оси с помощью сил F, действующих на торцы бруска с площадью A, а на его боковой поверхности внешние силы отсутствуют. При этом поперечные размеры бруска, вообще говоря, изменяются (рис. 1.7).

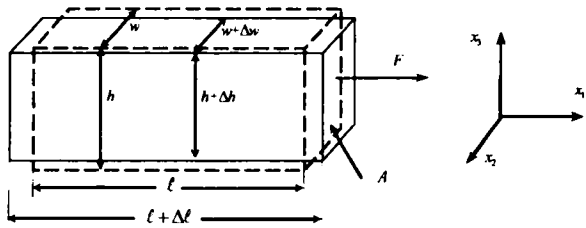


Рис. 1.7. Однородное растяжение изотропного прямоугольного бруска. Пунктирной линией показан брусок в недеформированном состоянии

Поскольку деформация однородна, то все компоненты U_{ik}, а вместе с ними и компоненты тензора напряжений T_{ik} постоянны вдоль бруска, поэтому их можно найти из граничных условий. На боковой поверхности внешние силы отсутствуют, откуда $T_{ik}n_k = 0$. Так как на боковой поверхности вектор единичной нормали не имеет компоненты вдоль оси x_1, можно заключить, что все компоненты тензора деформаций за исключением T_{11} равны нулю. Значение компоненты T_{11} определится по величине растягивающей силы на торцах бруска:

$$T_{11} = F/A\,.$$

Компоненты тензора деформаций запишутся как

$$U_{11} = \frac{\Delta\ell}{\ell},\quad U_{22} = \frac{\Delta w}{w},\quad U_{11} = \frac{\Delta h}{h}\,.$$

Все остальные U_{ik} равны нулю. Уравнения закона Гука (1.31) при этом имеют вид:

$$\lambda\left(\frac{\Delta\ell}{\ell}+\frac{\Delta w}{w}+\frac{\Delta h}{h}\right)+2\mu\frac{\Delta\ell}{\ell}=\frac{F}{A},$$
$$\lambda\left(\frac{\Delta\ell}{\ell}+\frac{\Delta w}{w}+\frac{\Delta h}{h}\right)+2\mu\frac{\Delta w}{w}=0,$$
$$\lambda\left(\frac{\Delta\ell}{\ell}+\frac{\Delta w}{w}+\frac{\Delta h}{h}\right)+2\mu\frac{\Delta h}{h}=0.$$

Из них нетрудно установить, что нормальное напряжение в поперечном сечении бруска пропорционально его относительному продольному растяжению:

$$\frac{F}{A}=\mathrm{E}\frac{\Delta\ell}{\ell},$$

где введено обозначение

$$\mathrm{E}=\frac{(3\lambda+2\mu)\mu}{\lambda+\mu}=\frac{9K_V\mu}{3K_V+\mu}.$$

Коэффициент пропорциональности E между напряжением и относительным растяжением в направлении действия напряжения называют модулем Юнга. Так как $K_V,\mu>0$, то модуль E – положительная величина. Следует сказать, что модуль Юнга иногда используют также для описания упругих свойств анизотропных материалов. Однако в этом случае он зависит от выбранного направления в материале.

Из уравнений закона Гука также следует, что относительное изменение поперечных размеров одинаково как для высоты, так и для ширины бруска и пропорционально его относительному удлинению:

$$\frac{\Delta h}{h}=\frac{\Delta w}{w}=-\sigma\frac{\Delta\ell}{\ell}.$$

Характеристика материала σ называется коэффициентом (или отношением) Пуассона:

$$\sigma = \frac{\lambda}{2(\lambda+\mu)} = \frac{3K_V - 2\mu}{2(3K_V + \mu)}.$$

Ввиду того, что модули K_V и μ – положительные величины, коэффициент Пуассона σ может изменяться только от -1 при $K_V << \mu$ до $1/2$ при $K_V >> \mu$. Фактически же в реальных материалах коэффициент Пуассона всегда лежит в пределах $0 < \sigma < 1/2$, так как не известны тела, которые бы утолщались при продольном растяжении, хотя это и не противоречит энергетическим представлениям. Когда $\sigma = 0$ среду называют абсолютно сжимаемой, а при $\sigma = 1/2$ – абсолютно несжимаемой. Если $\sigma > 0$, то коэффициент Лямэ λ также является положительным.

1.6. Деформация пьезоэлектриков

Известно, что под действием электрического поля диэлектрические тела поляризуются. Механизм электрической поляризации кроется во внутреннем строении диэлектрической среды. Каждый ее атом состоит из положительно заряженного ядра и отрицательно заряженной электронной оболочки. Сами атомы в целом также часто не являются электрически нейтральными, а представляют собой положительные и отрицательные ионы. Электрическое поле смещает положительные заряды относительно отрицательных, и среда приобретает объемный дипольный момент.

В некоторых кристаллах поляризация появляется не только под действием электрического поля, а также при их деформировании. Такие кристаллы называют пьезоактивными кристаллами или пьезоэлектриками. В пьезоэлектриках в силу определенных особенностей строения их кристаллической решетки при деформировании происходит смещение положительных ионов относительно отрицательных, и помимо механического напряжения появляется электрическая поляризация. Вследствие этого на поверхности деформированного пьезоэлектрического тела возникают поляризационные заряды, которые создают электрическое поле. Когда деформация невелика, создаваемое телом электрическое поле пропорционально деформации. Это явление называют прямым пьезоэффектом. Имеет место и обратный пьезоэффект – воздействие слабого электрического поля

на пьезоэлектрик приводит не только к его электрической поляризации, но также вызывает пропорциональное полю механическое напряжение.

Обратный пьезоэффект не следует путать с электрострикцией – изменением объема тела пропорционально квадрату электрического поля. По сравнению с пьезоэффектом электрострикция является эффектом более высокого порядка малости по величине электрического поля. В дальнейшем мы ограничимся слабыми электрическими полями и электрострикцией будем пренебрегать.

Необходимо отметить, что в пиро- и сегнетоэлектриках возможна спонтанная поляризация даже в отсутствие каких-либо внешних воздействий. По этой причине, говоря о поляризации, мы будем иметь в виду только ее изменение, связанное с электрическим полем и произведенной деформацией.

Рассмотрим сначала диэлектрическую среду, не обладающую пьезоэлектрическими свойствами. Пусть под влиянием электрического поля с напряженностью $\vec{E}$ в среде создана поляризация, которую будем описывать объемной плотностью дипольного момента $\vec{P}$. Когда электрическое поле достаточно мало, компоненты векторов $\vec{P}$ и $\vec{E}$ связывает соотношение

$$P_i = \varepsilon_0 \chi_{ik} E_k , \tag{1.34}$$

где ε_0 – диэлектрическая проницаемость вакуума, χ_{ik} – тензор второго ранга, называемый тензором поляризуемости диэлектрика. Тензорный характер поляризуемости объясняется тем, что направление смещения зарядов может не совпадать с направлением электрического поля, и векторы $\vec{P}$ и $\vec{E}$, в общем случае, не параллельны. Компоненты тензора поляризуемости определяются из (1.34) дифференцированием компонент вектора поляризации по компонентам напряженности электрического поля:

$$\chi_{ik} = \frac{1}{\varepsilon_0} \frac{\partial P_i}{\partial E_k} . \tag{1.35}$$

Рассмотрим энергию поляризации диэлектрика. Ее можно найти как работу, совершаемую электрическим полем при установлении поляризации. Пусть в единице объема среды содержится N положи-

тельных зарядов q, которые под действием электрического поля E_1 получили вдоль оси x_1 смещение dx_1. Совершаемая при этом работа равна $E_1 qNdx_1$. Но $qNdx_1$ представляет собой приращение компоненты дипольного момента P_1, то есть dP_1. Значит, работа по изменению поляризации вдоль x_1 равна $E_1 dP_1$. Аналогично, при смещении зарядов вдоль двух других осей получим $E_2 dP_2$ и $E_3 dP_3$. Суммарная работа по поляризации среды – $E_i dP_i$. Если обозначить объемную плотность энергии поляризации как W, то ее приращение

$$dW = E_i dP_i . \tag{1.36}$$

Подставим сюда (1.34):

$$dW = \varepsilon_0 \chi_{ik} E_i dE_k . \tag{1.37}$$

Энергия поляризации определяется только текущим значением компонент электрического поля, но не тем, как они изменялись, поэтому (1.36) представляет собой полный дифференциал W как функции независимых переменных E_i. Значит,

$$\varepsilon_0 \chi_{ik} E_i = \frac{\partial W}{\partial E_k} .$$

Дифференцируя по E_i, имеем

$$\varepsilon_0 \chi_{ik} = \frac{\partial^2 W}{\partial E_i \partial E_k} .$$

Порядок дифференцирования можно поменять, следовательно, тензор поляризуемости χ_{ik} – симметричный тензор.

Интегрируя (1.37), получаем энергию поляризованного диэлектрика:

$$W = \frac{1}{2}\varepsilon_0 \chi_{ik} E_i E_k .$$

Так как энергия поляризации должна быть положительной при любом направлении электрического поля, диагональные компоненты тензора χ_{ik} также положительны.

Вектор электрической индукции в среде определяется выражением

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} \tag{1.38}$$

В слабых электрических полях компоненты векторов $\vec{D}$ и $\vec{E}$ связаны соотношениями:

$$D_i = \varepsilon_{ik} E_k ,$$

где тензор второго ранга

$$\varepsilon_{ik} = \varepsilon_0 (\delta_{ik} + \chi_{ik}) \tag{1.39}$$

называется тензором диэлектрической проницаемости. Очевидно, тензор диэлектрической проницаемости ε_{ik} является симметричным и его диагональные компоненты – положительные величины.

Полезно ввести понятие диэлектрической проницаемости для некоторого направления. Пусть $\vec{\nu}$ – единичный вектор в направлении действия электрического поля. Его компоненты представляют собой косинусы углов вектора $\vec{E}$ с осями координат. Очевидно, можно записать

$$E_k = E\nu_k , \tag{1.40}$$

где $E = |\vec{E}|$. Обозначим через D проекцию вектора индукции на направление электрического поля. Она найдется как скалярное произведение $D_i \nu_i$:

$$D = D_i \nu_i = \varepsilon_{ik} \nu_i E_k .$$

Подставляя выражение (1.40), получаем

$$D = (\varepsilon_{ik} \nu_i \nu_k) E .$$

Величина $\varepsilon = (\varepsilon_{ik} \nu_i \nu_k)$ может быть названа диэлектрической проницаемостью в направлении вектора $\vec{\nu}$. Для любого фиксированного направления она является скалярным инвариантом и может быть вычислена в любой системе координат. В главных осях тензора ε_{ik} величина ε равна

$$\varepsilon = \varepsilon_{11} \nu_1^2 + \varepsilon_{22} \nu_2^2 + \varepsilon_{33} \nu_3^2 ,$$

откуда следует, что диэлектрическая проницаемость для любого направления – всегда положительная величина.

Рассмотрим теперь пьезоактивную среду. Ясно, что в пьезоактивной среде три уравнения (1.34) должны быть заменены соотношениями, которые учитывали бы прямой пьезоэффект – вклад деформации в электрическую поляризацию. В свою очередь, уравнения закона Гука (1.12), устанавливающие зависимость механических напряжений от компонент тензора деформаций, должны быть заменены некоторыми другими соотношениями, которые бы учитывали вклад в механическое напряжение электрического поля, то есть обратный пьезоэффект. Таким образом, в пьезоэлектрике компоненты вектора поляризации P_i и компоненты тензора напряжений T_{ik} должны рассматриваться как функции компонент U_{ik} и E_i:

$$P_i = P_i (E_1, E_2, ..., U_{11}, U_{12}, ...),$$
$$T_{ik} = T_{ik} (E_1, E_2, ..., U_{11}, U_{12}, ...).$$

Имея в виду малую деформацию и слабое электрическое поле, данные выражения можно разложить в ряд, ограничившись линейными членами:

$$P_m = P_m (0, ...) + \left(\frac{\partial P_m}{\partial E_i} \right) E_i + \left(\frac{\partial P_m}{\partial U_{ik}} \right) U_{ik} , \qquad (1.41)$$

2. АКУСТИЧЕСКИЕ ВОЛНЫ В РАЗЛИЧНЫХ СРЕДАХ

2.1. Движения в упругом теле

Выделим в деформированном упругом теле какой-нибудь объем V, ограниченный поверхностью S, и найдем действующую на него силу $\vec{F}$. В общем случае она будет состоять из двух частей. Одна из них является суммой поверхностных сил, действующих на объем со стороны окружающего материала. Другая часть создается внешними силами, действующими на данный объем на расстоянии (например, силой тяжести):

$$\vec{F} = \int_V \vec{g} dV + \oint_S \vec{T} ds ,$$

где $\vec{g}$ – объемная плотность внешних дальнодействующих сил, а $\vec{T}$ – поверхностная плотность сил на S. В покомпонентном виде:

$$F_i = \int_V g_i dV + \oint_S T_i ds , \qquad (2.1)$$

Учитывая (1.10), можно написать

$$\oint_S T_i ds = \oint_S T_{ik} n_k ds , \qquad (2.2)$$

где T_{ik} – компоненты тензора напряжений, а n_k – компоненты единичной внешней нормали к поверхности S. С помощью теоремы Остроградского–Гаусса поверхностный интеграл в правой части (2.2) преобразуем в объемный:

$$\oint_S T_{ik} n_k ds = \int_V \frac{\partial T_{ik}}{\partial x_k} dV .$$

Тогда (2.1) приобретает вид

$$T_{ik} = T_{ik}(0,...) + \left(\frac{\partial T_{ik}}{\partial E_m}\right) E_m + \left(\frac{\partial T_{ik}}{\partial U_{lm}}\right) U_{lm}. \qquad (1.42)$$

Частные производные по каждой из переменных берутся при постоянном значении всех остальных переменных.

Так как рассматриваются механическое напряжение и поляризация, связанные только с электрическим полем и деформацией, первые члены в (1.41) и (1.42) будем опускать. С учетом сказанного запишем соотношения (1.41) и (1.42) в следующем виде:

$$P_m = \varepsilon_0 \chi^U_{mi} E_i + e^E_{mik} U_{ik}, \qquad (1.43)$$

$$T_{ik} = C^E_{iklm} U_{lm} - e^U_{mik} E_m, \qquad (1.44)$$

где использованы обозначения:

$$\chi^U_{mi} = \frac{1}{\varepsilon_0}\frac{\partial P_m}{\partial E_i}, \quad e^E_{mik} = \frac{\partial P_m}{\partial U_{ik}}, \qquad (1.45)$$

$$C^E_{iklm} = \left(\frac{\partial T_{ik}}{\partial U_{lm}}\right), \quad e^U_{mik} = -\left(\frac{\partial T_{ik}}{\partial E_m}\right). \qquad (1.46)$$

Первый член правой части (1.43) описывает изменение поляризации среды под действием электрического поля при фиксированной деформации. Набор коэффициентов χ^U_{mi} образуют тензор второго ранга, который по аналогии с (1.35) можно назвать тензором поляризуемости пьезоэлектрика. Его компоненты определяются при постоянном значении компонент тензора деформаций. Второй член представляет собой вклад в поляризацию малой деформации, то есть прямой пьезоэффект. Прямой пьезоэффект описывается тензором третьего ранга e^E_{mik}, компоненты которого определяются при постоянном электрическом поле. Совокупность величин C^E_{iklm} образует тензор четвертого ранга и описывает упругую реакцию пьезоэлектрика на деформацию при постоянном электрическом поле. По аналогии с (1.13) величины C^E_{iklm} можно

назвать упругими модулями пьезоэлектрика. Наконец, набор коэффициентов e^{U}_{mik} описывает механическое напряжение, порождаемое электрическим полем – обратный пьезоэффект. Данный набор, очевидно, образует тензор третьего ранга, компоненты которого определяются при фиксированной деформации. Свойства тензоров, введенных соотношениями (1.45) и (1.46), можно найти из энергетических представлений.

Рассмотрим деформированное пьезоэлектрическое тело, находящееся во внешнем электрическом поле. Пусть сначала компоненты тензора деформаций получили некоторые малые изменения, а электрическое поле поддерживается постоянным. Приращение упругой энергии единицы объема тела при малом изменении компонент тензора деформаций определяется соотношением (1.21). К нему надо прибавить работу, связанную с изменением поляризации тела (1.36). Изменение поляризации при фиксированном электрическом поле, как следует из (1.43):

$$dP^{E}_{m} = e^{E}_{mik} dU_{ik} .$$

Таким образом, при малом изменении компонент тензора деформаций и постоянном электрическом поле суммарное приращение энергии на единицу объема равно

$$d\Phi^{E} = T_{ik} dU_{ik} + e^{E}_{mik} E_{m} dU_{ik} .$$

С учетом (1.44) получаем

$$d\Phi^{E} = \left(C^{E}_{iklm} U_{lm} + e^{E}_{mik} E_{m} - e^{U}_{mik} E_{m} \right) dU_{ik} . \qquad (1.46)$$

Пусть теперь компоненты тензора деформации остаются постоянными, а компоненты электрического поля получат малое изменение. Связанное с этим приращение энергии вновь определяется соотношением (1.36), но изменение поляризации теперь равно:

$$dP^{U}_{m} = \varepsilon_{0} \chi^{U}_{ml} dE_{l} .$$

Таким образом, приращение энергии при малом изменении электрического поля и фиксированной деформации записывается как

$$W^U = \varepsilon_0 \chi^U_{mi} E_m dE_i = \varepsilon_0 \chi^U_{im} E_i dE_m . \qquad (1.48)$$

Складывая выражения (1.47) и (1.48), получаем приращение энергии на единицу объема пьезоэлектрика при независимом изменении компонент тензора деформаций и электрического поля:

$$d\Psi = \left(C^E_{iklm} U_{lm} + e^E_{mik} E_m - e^U_{mik} E_m \right) dU_{ik} + \varepsilon_0 \chi^U_{im} E_i dE_m . \qquad (1.49)$$

Выражение (1.49) представляет собой полный дифференциал Ψ как функции независимых переменных U_{ik} и E_m. Следовательно:

$$\begin{aligned} \frac{\partial \Psi}{\partial U_{ik}} &= \left(C^E_{iklm} U_{lm} + e^E_{mik} E_m - e^U_{mik} E_m \right), \\ \frac{\partial \Psi}{\partial E_m} &= \varepsilon_0 \chi^U_{mi} E_i . \end{aligned} \qquad (1.50)$$

Продифференцируем первое из соотношений (1.50) по E_m, а второе – по U_{ik} и поменяем порядок дифференцирования. В результате получим

$$\frac{\partial^2 \Psi}{\partial U_{ik} \partial E_m} = e^E_{mik} - e^U_{mik} = 0 .$$

Откуда следует важный вывод: прямой и обратный пьезоэффекты описываются одним и тем же тензором. Далее его будем называть пьезоэлектрическим тензором и обозначать как e_{mik}. Ввиду симметрии компонент T_{ik} и U_{ik} пьезоэлектрический тензор симметричен относительно перестановки второго и третьего индексов. В связи с этим из 27 его компонент лишь 18 могут быть независимыми, и для его записи можно также использовать ранее рассмотренный матричный метод:

$$e_{mik} \to \mathbf{e}_{m\alpha} \quad (m = 1,2,3, \ \alpha = 1,2,3,4,5,6) .$$

В наиболее общем случае прямоугольная 3×6 матрица $\mathbf{e}_{i\alpha}$ имеет 18 различных ненулевых членов. Обычно по причине различных внутренних симметрий число ненулевых членов значительно меньше.

Беря повторные производные выражений (1.50), получаем

$$\frac{\partial^2 \Psi}{\partial U_{ik} \partial U_{lm}} = C^E_{iklm}, \quad \frac{\partial^2 \Psi}{\partial E_m \partial E_i} = \varepsilon_0 \chi^U_{mi} .$$

Так как порядок дифференцирования может быть изменен, то тензор поляризуемости χ^U_{mi} симметричен, а упругие модули C^E_{iklm} обладают теми же свойствами симметрии, которые были ранее установлены для упругих модулей непьезоактивной среды.

Интегрирование (1.49) дает

$$\Psi = \frac{1}{2} C^E_{iklm} U_{lm} U_{ik} + \frac{1}{2} \varepsilon_0 \chi^U_{im} E_i E_m ,$$

откуда следует, что диагональные компоненты тензора χ^U_{im} – положительные величины.

Компоненты вектора электрической индукции в пьезоэлектрике определятся из (1.38) и (1.43):

$$D_m = \varepsilon^U_{mi} E_i + e_{mik} U_{ik} , \qquad (1.51)$$

где ε^U_{mi} – симметричный тензор диэлектрической проницаемости пьезоэлектрика, определенный при фиксированной деформации

$$\varepsilon^U_{ik} = \varepsilon_0 \left(\delta_{ik} + \chi^U_{ik} \right) .$$

Надстрочные символы в обозначении тензора диэлектрической проницаемости и тензора модулей упругости, указывающие на способ их определения, для краткости обычно опускают, а соотношения прямого и обратного пьезоэффектов записывают как

$$D_m = \varepsilon_{im} E_i + e_{mik} U_{ik},$$
$$T_{ik} = C_{iklm} U_{lm} - e_{mik} E_m. \qquad (1.52)$$

Эти соотношения также называют уравнениями состояния пьезоэлектрика.

В принципе, вид соотношений, описывающих пьезоэффект, достаточно условен и определяется тем, какие переменные считаются независимыми. Иногда рассматривают состояние пьезоэлектрика при заданных величинах электрического поля и механического напряжения, а деформацию и индукцию считают их функциями. В этом случае соотношения для пьезоэффекта принимают вид

$$D_m = \varepsilon_{im}^T E_i + d_{mik} T_{ik},$$
$$U_{ik} = S_{iklm}^E T_{lm} + d_{mik} E_m.$$

где компоненты тензора диэлектрической проницаемости ε_{im}^T определяются при постоянной величине напряжения, а упругие податливости S_{iklm}^E – при постоянном электрическом поле. Пьезоэлектрический тензор d_{mik} описывает вклад в индукцию и деформацию, соответственно, напряжения и электрического поля. Между тензорами $\varepsilon_{im}^T, S_{iklm}^E, d_{mik}$ и тензорами $\varepsilon_{im}^U, C_{iklm}^E, e_{mik}$ существует достаточно сложная связь, которая здесь не рассматривается.

До сих пор, говоря об упругой деформации различных тел, здесь и в разделе 1.4, мы не учитывали того факта, что в процессе деформирования многие тела изменяют свою температуру. Кроме того, нагревание или охлаждение тела может происходить по внешним причинам. Среди различных (в термодинамическом смысле) типов деформации можно выделить изотермическую и адиабатическую деформации. Если процесс деформирования протекает изотермически и обратимо, то изменение энергии тела, связанное с деформированием, равно приращению его свободной энергии при постоянной температуре. Соответственно, компоненты материальных тензоров ε_{im}, C_{iklm}, e_{mik} должны рассматриваться как изотермические величины. При быстропеременной деформации (например, при распространении высокочастотных упругих волн) теплообмена между различными частями тела и окружающим пространством практически не происходит. В этом случае изменение энергии, связанное с деформированием, будет равно приращению внутренней энергии тела

при постоянной его энтропии, и компоненты тензоров ε_{im}, C_{iklm}, e_{mik} следует рассматривать как адиабатические величины.

В заключение отметим, что макроскопическая теория деформации определяет только некоторые свойства материальных констант, но не задает их численных значений. Эти значения могут быть получены только на основе микроскопической квантовой теории, учитывающей природу сил межатомного взаимодействия и особенности внутреннего строения вещества, либо из экспериментальных исследований.

Контрольные вопросы

1. Каков физический смысл диагональных и недиагональных элементов тензора деформаций?
2. Докажите, что тензор напряжений является симметричным тензором на примере деформации кубика малого объема.
3. Назовите число независимых упругих модулей тензора модулей упругости для случаев изотропной среды, кристалла кубической симметрии и в общем случае.
4. Приведите формулу для объёмной плотности упругой энергии при деформации и проанализируйте её для изотропной среды.
5. Приведите соотношения, характеризующие связь коэффициентов Лямэ с модулем объёмной упругости, коэффициентом Пуассона и модулем Юнга.
6. Проанализируйте соотношения для прямого и обратного пьезоэффектов при наличии и без внешних электрического и механического полей.

$$F_i = \int_V g_i dV + \int_V \frac{\partial T_{ik}}{\partial x_k} dV .$$

Поскольку данное соотношение справедливо для произвольного объема, то величины

$$f_i = \frac{\partial T_{ik}}{\partial x_k}$$

можно рассматривать как компоненты объемной плотности силы, создаваемой внутренними напряжениями.

Рассмотрим в теле какую-нибудь бесконечно малую частицу с объемом dV. Компоненты действующей на нее силы будут равны

$$\left(f_i + g_i\right) dV = \left(\frac{\partial T_{ik}}{\partial x_k} + g_i\right) dV .$$

Когда тело находится в равновесии, сумма сил, действующих на любую его частицу, должна равняться нулю. Если же равновесие по какой-нибудь причине нарушилось, то согласно второму закону Ньютона эту сумму следует приравнять произведению массы частицы на ее ускорение

$$\rho dV \frac{\partial^2 u_i}{\partial t^2} = \left(\frac{\partial T_{ik}}{\partial x_k} + g_i\right) dV ,$$

где u_i – компоненты вектора смещения частицы, ρ – плотность материала. Сокращая на dV, получаем уравнение движения упругой среды

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial T_{ik}}{\partial x_k} + g_i .$$

Если внешние дальнодействующие силы отсутствуют,

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial T_{ik}}{\partial x_k}. \qquad (2.3)$$

Для поиска решений уравнения движения необходима связь компонент тензора напряжений с компонентами вектора смещения. Данная связь будет различной для разного рода сред. В непьезоактивных средах эта связь определяется законом Гука, а в пьезоэлектриках – уравнениями пьезоэффекта.

При решении уравнений движения в ограниченных объемах и на границах раздела сред требуются краевые условия для компонент смещения и напряжения, которые обычно формулируются из простых физических соображений.

Рассмотрим, например, границу двух упругих сред. Все величины, относящиеся к одной среде, будем обозначать надстрочным индексом 1, а к другой – 2. Если границы сред жестко связаны, то, очевидно, при переходе через границу вектор смещения должен быть непрерывен:

$$u_i^{(1)} = u_i^{(2)}.$$

Кроме того, в соответствии с принципом равенства действия и противодействия должны быть равны и противоположны компоненты силы, действующей на границу раздела со стороны прилегающих к ней сред. Обозначив вектор нормали к границе $\vec{n}$, можно записать

$$T_{ik}^{(1)} n_k = T_{ik}^{(2)} n_k.$$

Часто упругая среда граничит с неупругой (например, вакуум), или средой, чьей упругостью можно пренебречь (воздух). Для этого случая граничное условие для компонент напряжения приобретает вид

$$T_{ik} n_k = 0.$$

Движение пьезоактивной среды может индуцировать переменное электрическое поле, которое, в свою очередь, будет создавать переменное магнитное поле. Поэтому для полного описание движения пьезоэлектриков наряду с уравнением движения (2.3) необходимо

использовать, вообще говоря, полную систему уравнений Максвелла. В присутствии границ помимо краевых условий, налагаемых на смещения и напряжения, требуются также стандартные граничные условия для компонент электромагнитного поля.

2.2. Объемные упругие волны в изотропной среде

Рассмотрим безграничную однородную изотропную среду. На основании закона Гука для изотропной среды (1.31) и определения тензора деформаций (1.7) находим

$$\frac{\partial T_{ik}}{\partial x_k} = (\lambda + \mu)\frac{\partial}{\partial x_i}\left(\frac{\partial u_k}{\partial x_k}\right) + \mu\left(\frac{\partial^2 u_i}{\partial x_k^{\,2}}\right),$$

и уравнение движения (2.3) приобретает вид

$$\rho\frac{\partial^2 u_i}{\partial t^2} - (\lambda + \mu)\frac{\partial}{\partial x_i}\left(\frac{\partial u_k}{\partial x_k}\right) - \mu\left(\frac{\partial^2 u_i}{\partial x_k^{\,2}}\right) = 0.$$

В векторной форме

$$\rho\frac{\partial^2 \vec{u}}{\partial t^2} - (\lambda + \mu)\,\mathrm{graddiv}\,\vec{u} - \mu\nabla^2\vec{u} = 0. \qquad (2.4)$$

Распределение смещений на множестве точек среды $\vec{u}(r)$ образует векторное поле, имеющее непрерывные производные. Из векторного анализа известно, что такое поле на всем множестве точек всегда может быть представлено в виде суммы соленоидального (вихревого) поля $\vec{u}_T$ и безвихревого поля $\vec{u}_L$:

$$\vec{u} = \vec{u}_T + \vec{u}_L \;\; \vec{u} = \vec{u}_T + \vec{u}_L, \qquad (2.5)$$

которые удовлетворяют условиям:

$$\mathrm{div}\,\vec{u}_T = 0, \qquad (2.6)$$

$$\operatorname{rot} \vec{u}_L = 0 . \qquad (2.7)$$

Известно также, что если дивергенция и ротор некоторого векторного поля исчезают во всем пространстве, то это поле тождественно равно нулю.

Подставим (2.5) в (2.4) с учетом (2.6):

$$\rho \frac{\partial^2 \left(\vec{u}_T + \vec{u}_L\right)}{\partial t^2} - \left(\lambda + \mu\right) \operatorname{graddiv} \vec{u}_L - \mu \nabla^2 \left(\vec{u}_T + \vec{u}_L\right) = 0 .$$

Далее, используя операторное тождество $\operatorname{graddiv} \equiv \operatorname{rotrot} + \nabla^2$ и принимая во внимание (2.7), получаем

$$\rho \frac{\partial^2 \left(\vec{u}_T + \vec{u}_L\right)}{\partial t^2} - \left(\lambda + 2\mu\right) \nabla^2 \vec{u}_L - \mu \nabla^2 \vec{u}_T = 0 . \qquad (2.8)$$

Возьмем от (2.8) дивергенцию

$$\operatorname{div} \left[\rho \frac{\partial^2 \vec{u}_L}{\partial t^2} - \left(\lambda + 2\mu\right) \nabla^2 \vec{u}_L \right] = 0 .$$

Ротор векторного выражения в квадратных скобках ввиду условия (2.7) равен нулю. Из одновременного равенства нулю ротора и дивергенции следует, что это выражение тождественно равно нулю:

$$\frac{\rho}{\lambda + 2\mu} \frac{\partial^2 \vec{u}_L}{\partial t^2} - \nabla^2 \vec{u}_L = 0 . \qquad (2.9)$$

Аналогично, беря ротор от уравнения (2.8), имеем

$$\operatorname{rot} \left[\rho \frac{\partial^2 \vec{u}_T}{\partial t^2} - \mu \nabla^2 \vec{u}_T \right] = 0 .$$

Дивергенция векторного выражения в квадратных скобках в силу (2.6) равна нулю, следовательно,

$$\frac{\rho}{\mu}\frac{\partial^2 \vec{u}_T}{\partial t^2} - \nabla^2 \vec{u}_T = 0 . \qquad (2.10)$$

Таким образом, векторы смещения $\vec{u}_L$ и $\vec{u}_T$ удовлетворяют независимым волновым уравнениям (2.9) и (2.10). Поэтому в безграничной изотропной среде соленоидальное и безвихревое движение частиц среды происходит независимо и возможно существование двух независимых видов объемных волн. Исследуем их подробнее.

Рассмотрим волновое уравнение

$$\frac{1}{v^2}\frac{\partial^2 \vec{u}}{\partial t^2} - \nabla^2 \vec{u} = 0 . \qquad (2.11)$$

Будем искать поле смещений в виде плоской однородной монохроматической волны:

$$\vec{u}(\vec{r},t) = \vec{u}_0 \exp\left[j\left(\omega t - \kappa_i x_i\right)\right], \qquad (2.12)$$

где $\vec{u}_0$ – постоянный вектор, определяющий амплитуду и направление движения точек среды (поляризацию волны), ω – циклическая частота, κ_i – компоненты волнового вектора $\vec{\kappa}$. Модуль волнового вектора $\kappa = |\vec{\kappa}|$ называют постоянной распространения, или волновым числом. Точки с фиксированным значением фазы $\omega t - \kappa_i x_i = const$ образуют плоский волновой фронт. Он ортогонален волновому вектору и перемещается в пространстве в направлении $\vec{\kappa}$ со скоростью $v_\phi = \omega/\kappa$, называемой фазовой скоростью волны. В однородной волне для всех точек волнового фронта смещение одинаково.

Под истинным вектором смещения, конечно же, следует понимать вещественную часть выражения (2.12), однако ввиду линейности волнового уравнения комплексная форма решения возможна и часто оказывается более удобной, чем запись через вещественные тригонометрические функции, так как существенно упрощает математические преобразования. В частности, дифференцирование (2.12) по времени эквивалентно умножению на $j\omega$, а по x_i – на $(-j\kappa_i)$. Отметим также, что умножение на

мнимую единицу в силу соотношения $\pm j = \exp(\pm j\pi/2)$ означает дополнительный фазовый сдвиг на $\pi/2$.

Подставляя решение (2.12) в (2.11), получаем

$$\omega^2 = \kappa^2 v^2. \qquad (2.13)$$

Для любой монохроматической волны уравнение, устанавливающее связь частоты с волновым вектором $\omega = \omega(\vec{\kappa})$, называется дисперсионным уравнением. Оно позволяет найти важные характеристики волны. В частности, из (2.13) вытекает, что фазовая скорость упругих волн в безграничной изотропной среде $v_\phi = v$ и не зависит от направления распространения. В принципе, любую объемную волну можно представить как волновой пакет, составленный из решений вида (2.12), имеющих различные фазы и амплитуды. Как известно, скорость и направление распространения огибающей волнового пакета определяет вектор групповой скорости:

$$\vec{v}_{гр} = \frac{\partial \omega}{\partial \vec{\kappa}}.$$

Из дисперсионного уравнения (2.13) находим

$$\vec{v}_{гр} = v \frac{\partial}{\partial \vec{\kappa}} \sqrt{\vec{\kappa}\vec{\kappa}} = v \frac{\vec{\kappa}}{|\vec{\kappa}|}.$$

В изотропной среде направление групповой скорости совпадает с направлением волнового вектора, а ее величина равна фазовой скорости.

Волновое движение частиц среды будет порождать волны деформации. Для соленоидальной волны дивергенция (2.11) должна равняться нулю:

$$\operatorname{div} \vec{u} = \frac{\partial u_i}{\partial x_i} = -j u_{0i} \kappa_i \exp\left[j\left(\omega t - \kappa_i x_i\right)\right] = -j(\vec{u}, \vec{\kappa}) = 0.$$

Отсюда следует, во-первых, что направление движения частиц среды ортогонально волновому вектору. Во-вторых, что создаваемая волной деформация не связана с изменением объема и является чистой деформацией сдвига (см. раздел 1.1). Таким образом, соленоидальная волна является поперечной волной сдвига (рис. 2.1, а). Из сравнения (2.10) и (2.11) находим, что фазовая и групповая скорости этой волны:

$$v_T = \sqrt{\mu/\rho}\,. \qquad (2.14)$$

Для безвихревой волны ротор (2.12) должен быть равен нулю, откуда вытекает условие:

$$\left[\vec{u}, \vec{\kappa}\right] = 0\,.$$

Это означает, что движение частиц параллельно направлению распространения. Так как $\vec{u} \| \vec{\kappa}$, то $\mathrm{div}\vec{u} \neq 0$, и безвихревая волна представляет собой продольную волну чередующихся сжатий и растяжений (рис. 2.1, б). Ее фазовая и групповая скорости равны:

$$v_L = \sqrt{(2\mu + \lambda)/\rho}\,. \qquad (2.15)$$

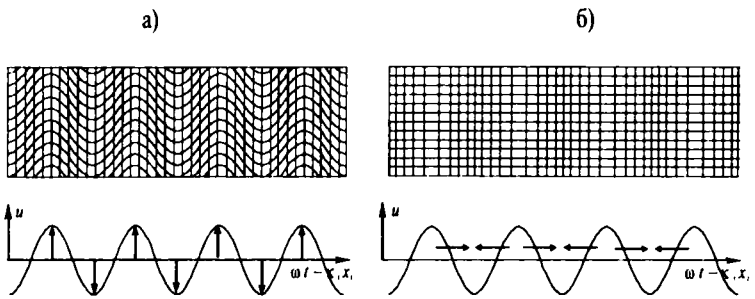


Рис. 2.1. Распределение деформаций и смещений в упругих объемных волнах: а – поперечная волна сдвига; б – продольная волна сжатия-разряжения

Отметим, что скорости распространения обоих типов волн определяются только упругими свойствами среды и не зависят от часто-

ты. Значения v_T и v_L иногда называют соответственно поперечной и продольной скоростью звука. Так как λ и μ – положительные величины, то из (2.14) и (2.15) вытекает неравенство $v_L > v_T\sqrt{2}$. То есть в безграничной изотропной среде скорость продольной волны всегда больше скорости поперечной.

2.3. Объемные упругие волны в кристаллах в отсутствие пьезоэффекта

Рассмотрим однородный непьезоактивный кристалл. В таком кристалле зависимость напряжений от деформаций определяет общая форма закона Гука (1.12). Продифференцируем (1.12) по координатам и подставим выражение для тензора деформаций (1.7):

$$\frac{\partial T_{ik}}{\partial x_k} = \frac{1}{2} C_{iklm}\left(\frac{\partial^2 u_m}{\partial x_k \partial x_l} + \frac{\partial^2 u_l}{\partial x_k \partial x_m}\right).$$

Принимая во внимание общие свойства симметрии упругих модулей (1.28), произведем во втором члене правой части замену немых индексов $l \leftrightarrow m$. В результате находим

$$\frac{\partial T_{ik}}{\partial x_k} = C_{iklm}\frac{\partial^2 u_m}{\partial x_k \partial x_l},$$

и уравнение движения (2.3) для непьезоактивного однородного кристалла приобретает следующий вид:

$$\rho\frac{\partial^2 u_i}{\partial t^2} - C_{iklm}\frac{\partial^2 u_m}{\partial x_k \partial x_l} = 0. \qquad (2.16)$$

Исследуем решения (2.16) в форме плоской однородной монохроматической волны (2.12). Подставляя (2.12) в (2.16), имеем

$$\rho\omega^2 u_i - \kappa^2 C_{iklm} v_k v_l u_m = 0,$$

где ν_i – компоненты единичного вектора в направлении распространения волны:

$$\vec{\nu} = \vec{\kappa}/\kappa .$$

Используя очевидное равенство $u_i = \delta_{im} u_m$, получаем однородную систему трех линейных уравнений относительно компонент смещения u_m:

$$\left(\rho\omega^2\delta_{im} - \kappa^2 C_{iklm}\nu_k\nu_l\right)u_m = 0 . \qquad (2.17)$$

Однородная система линейных уравнений имеет нетривиальные решения лишь при равенстве нулю ее главного определителя. Приравнивая главный определитель нулю:

$$\det\left[\rho\omega^2\delta_{im} - \kappa^2 C_{iklm}\nu_k\nu_l\right] = 0 ,$$

получаем дисперсионное уравнение третьей степени относительно ω^2.

При любом направлении распространения в кристалле дисперсионное уравнение имеет, в общем случае, три разных корня $\omega^2 = \omega^2(\vec{\kappa})$ – три ветки закона дисперсии. Каждая из веток описывает некоторый тип упругой волны или, как говорят, моду, которая может существовать в данном кристалле. Подставляя корни дисперсионного уравнения в (2.17) и решая полученную систему уравнений, можно найти поляризацию каждой из трех мод. В изотропной среде этими модами были $\omega = v_L\kappa$ (продольная волна) и две поперечные моды с совпадающими законами дисперсии $\omega = v_T\kappa$ и отвечающие двум независимым взаимно ортогональным направлениям поляризации.

Чтобы исследовать моды в кристалле, рассмотрим тензор $\Gamma_{im} = C_{iklm}\nu_k\nu_l$. Легко установить, что вследствие общих свойств модулей упругости (1.28) он симметричен. Поэтому для любого заданного направления распространения $\vec{\nu}$ в каждом кристалле можно найти систему координат – систему главных осей, в ко-

торой Γ_{im} диагонален, и дисперсионное уравнение принимает вид

$$\left(\rho\omega^2-\kappa^2\Gamma_1\right)\left(\rho\omega^2-\kappa^2\Gamma_2\right)\left(\rho\omega^2-\kappa^2\Gamma_3\right)=0,$$

где Γ_1, Γ_2, Γ_3 – главные значения тензора Γ_{im}. При этом система уравнений (2.17) распадается на три независимых уравнения:

$$\left(\rho\omega^2-\kappa^2\Gamma_1\right)u_1=0,$$
$$\left(\rho\omega^2-\kappa^2\Gamma_2\right)u_2=0,$$
$$\left(\rho\omega^2-\kappa^2\Gamma_3\right)u_3=0,$$

где u_i – компоненты смещений по главным осям. Корню дисперсионного уравнения $\omega^2=\kappa^2\Gamma_1/\rho$ соответствует мода с поляризацией $\vec{u}_1=(u_1,0,0)$, корню $\omega^2=\kappa^2\Gamma_2/\rho$ – мода $\vec{u}_2=(0,u_2,0)$ и, наконец, корню $\omega^2=\kappa^2\Gamma_3/\rho$ – мода $\vec{u}_3=(0,0,u_3)$. Поляризации мод взаимно ортогональны, а их фазовые скорости равны

$$v_1=\sqrt{\Gamma_1/\rho},\ v_2=\sqrt{\Gamma_2/\rho},\quad v_3=\sqrt{\Gamma_3/\rho}$$

и не зависят от частоты, однако зависят от направления распространения. Если компоненты C_{iklm} и ν_i выражены в главных осях, то

$$v_1=\sqrt{\frac{C_{1k1l}\nu_k\nu_l}{\rho}},\ v_2=\sqrt{\frac{C_{2k2l}\nu_k\nu_l}{\rho}},\quad v_3=\sqrt{\frac{C_{3k3l}\nu_k\nu_l}{\rho}}. \qquad (2.18)$$

Можно показать, что в силу требования положительной определенности упругой энергии все главные значения тензора Γ_{im} положительны и фазовые скорости всех трех мод – вещественные величины.

В общем случае направление распространения не совпадает ни с одной из главных осей тензора Γ_{im} и волны нельзя отнести ни к продольным, ни к поперечным. Если в кристалле существует какое-либо направление распространения, совпадающее с одной из главных

осей тензора Γ_{im}, мода с поляризацией вдоль этой оси становится чисто продольной, а две другие – чисто поперечными. Такое направление в кристалле называют направлением чистой моды или акустической осью кристалла. В любом случае направление распространения составляет с одной из главных осей тензора Γ_{im} наименьший угол, и в кристалле распространяются одна квазипродольная и две квазипоперечные волны.

Помимо фазовой скорости, дисперсионное уравнение позволяет найти величину скорости и направление распространения огибающей волнового пакета, составленного из монохроматических волн, то есть групповую скорость. В изотропной среде, как мы видели в разделе 2.2, зависимость $\omega(\vec{\kappa})$ сводилась к простой пропорциональности частоты волновому числу. Коэффициент пропорциональности не зависел от направления распространения, а вектор групповой скорости был направлен вдоль волнового вектора. В кристалле же, как следует из дисперсионных соотношений, это не так, и векторы $\vec{\kappa}$ и $\vec{v}_{гр}$, вообще говоря, не коллинеарны. В частном случае распространения вдоль акустической оси направления волнового вектора и групповой скорости совпадают. Величина групповой скорости, так же как и фазовой, не зависит от частоты.

В кристаллах, имеющих различные внутренние типы симметрии при определенных направлениях волнового вектора, выражения для фазовых скоростей распространяющихся мод можно упростить. В качестве примера рассмотрим кубический кристалл. Направим оси системы координат (x_1, x_2, x_3) вдоль рёбер элементарной ячейки кристалла. Исследуем волну, распространяющуюся вдоль оси x_1. Принимая во внимание свойства симметрии кубического кристалла, рассмотренные в разделе 1.5, находим, что в этом случае тензор Γ_{im} имеет всего три диагональных ненулевых элемента, два из которых равны между собой:

$$\Gamma_{11} = C_{1111} = \mathbf{C}_{11}, \quad \Gamma_{22} = \Gamma_{33} = C_{1212} = \mathbf{C}_{44}.$$

То есть рёбра элементарной ячейки кубического кристалла являются его акустическими осями. Вдоль ребра элементарной ячейки кубического кристалла могут распространяться одна продольная мода с

фазовой скоростью $v_L = \sqrt{C_{11}/\rho}$ и две поперечные взаимно ортогональные вырожденные моды с одинаковыми скоростями $v_T = \sqrt{C_{44}/\rho}$.

2.4. Объемные упругие волны в пьезоэлектриках

Рассмотрим однородный пьезоактивный кристалл. В пьезоэлектриках зависимость напряжений от деформаций определяют соотношения пезоэффекта (1.52). Дифференцируя второе из соотношений (1.52) по координатам, получаем

$$\frac{\partial T_{ik}}{\partial x_k} = C_{iklm}\frac{\partial U_{lm}}{\partial x_k} - e_{mik}\frac{\partial E_m}{\partial x_k}. \qquad (2.19)$$

Первый член правой части выразим через компоненты смещения подобно тому, как это делалось для непьезоактивного кристалла:

$$C_{iklm}\frac{\partial U_{lm}}{\partial x_k} = C_{iklm}\frac{\partial^2 u_m}{\partial x_l \partial x_k}. \qquad (2.20)$$

Подставляя (2.19) и (2.20) в (2.3), получаем уравнение движения однородного пьезоактивного кристалла:

$$\rho\frac{\partial^2 u_i}{\partial t^2} - C_{iklm}\frac{\partial^2 u_m}{\partial x_k \partial x_l} = -e_{mik}\frac{\partial E_m}{\partial x_k}. \qquad (2.21)$$

Рассматривая распространение упругой волны в пьезоэлектрике, можно утверждать, что порождаемые ее деформации будут приводить к появлению переменного электрического поля, которое, в свою очередь, будет индуцировать переменное магнитное поле. Поэтому для совместного описания распространения механических и электромагнитных колебаний, помимо уравнения (2.21), следует использовать, вообще говоря, полную систему уравнений Максвелла:

$$\operatorname{rot}\vec{E} = -\mu_0 \frac{\partial \vec{H}}{\partial t}, \quad \operatorname{rot}\vec{H} = \frac{\partial \vec{D}}{\partial t}, \qquad (2.22)$$
$$\operatorname{div}\vec{D} = 0, \quad \operatorname{div}\vec{H} = 0,$$

где μ_0 – магнитная проницаемость, $\vec{H}$ – вектор напряженности переменного магнитного поля. Однако из-за значительной разницы в скорости распространения звука и электромагнитной волны электрическое поле для упругой волны является, по существу, квазистационарным и для его описания можно использовать более простые уравнения электростатики. Покажем это.

Пусть плоская однородная монохроматическая волна распространяется по оси x_1:

$$\vec{u}(\vec{r},t) = \vec{u}_0 \exp\left[j\left(\omega t - \kappa x_1\right)\right].$$

Имея в виду линейность уравнения движения (2.21), логично считать, что такого же вида функция описывает распределение электрического и магнитного полей, порожденных данной волной. Представим электрическое поле волны в виде суммы вихревого и потенциального полей:

$$\vec{E} = \vec{E}^{rot} + \vec{E}^{\varphi}, \text{ где } \operatorname{rot}\vec{E}^{\varphi} = 0, \ \operatorname{div}\vec{E}^{rot} = 0.$$

С точностью до порядка величин из первого уравнения (2.22) получаем $\kappa\left|\vec{E}^{rot}\right| \sim \mu_0\omega\left|\vec{H}\right|$, а из второго $-\kappa\left|\vec{H}\right| \sim \omega\left|\vec{D}\right| \sim \omega\varepsilon\left|\vec{E}\right|$, где ε – диэлектрическая проницаемость. Отношение вихревой части электрического поля к полному электрическому полю

$$\frac{\left|\vec{E}^{rot}\right|}{\left|\vec{E}\right|} \sim \frac{v^2}{c^2},$$

где $c = \left(\varepsilon\mu_0\right)^{-\frac{1}{2}}$ – скорость света в среде, $v = \omega/\kappa$ – фазовая скорость упругой волны (скорость звука). Так как скорость звука примерно на пять порядков меньше скорости света, то это отношение весьма мало и вихревой частью электрического поля можно пренебречь. Следо-

вательно, для описания электрического поля, сопровождающего упругую волну в пьезоэлектрике, можно использовать скалярный потенциал $\vec{E} = -\mathrm{grad}\varphi$

Из третьего уравнения (2.22) и первого из соотношений пьезоэффекта (1.52) следует, что электрический потенциал должен удовлетворять анизотропному уравнению Лапласа:

$$\frac{\partial}{\partial x_i}\left(-\varepsilon_{im}\frac{\partial\varphi}{\partial x_m} + e_{ilm}U_{lm}\right) = 0 .$$

Подставляя сюда выражение для тензора деформаций (1.7) и учитывая свойство симметрии пьезоэлектрического тензора по второму и третьему индексам, получаем

$$e_{ilm}\frac{\partial^2 u_m}{\partial x_i \partial x_l} = \varepsilon_{im}\frac{\partial^2\varphi}{\partial x_i \partial x_m}, \tag{2.23}$$

а уравнение движения (2.21) приобретает вид

$$\rho\frac{\partial^2 u_i}{\partial t^2} - C_{iklm}\frac{\partial^2 u_m}{\partial x_k \partial x_l} = e_{mik}\frac{\partial^2\varphi}{\partial x_k \partial x_m} . \tag{2.24}$$

Система уравнений (2.23) и (2.24) является исходной для поиска решений, описывающих распространения упругих возмущений в пьезоактивной среде.

Рассмотрим решение в виде плоской однородной монохроматической волны:

$$u_i(\vec{r},t) = u_{0i}\exp\left[j(\omega t - \kappa_k x_k)\right],$$
$$\varphi(\vec{r},t) = \varphi_0\exp\left[j(\omega t - \kappa_k x_k)\right]$$

Подставляя это решение в (2.23), (2.24), приходим к следующей системе уравнений:

$$\begin{aligned} &e_{ilm}u_m\nu_i\nu_l = \varphi\varepsilon, \\ &\rho\omega^2 u_i - \kappa^2 C_{iklm}\nu_k\nu_l u_m = \kappa^2\varphi e_{mik}\nu_k\nu_m, \end{aligned} \tag{2.25}$$

где ν_i – компоненты единичного вектора в направлении распространения, $\varepsilon = \varepsilon_{im}\nu_i\nu_m$ – диэлектрическая проницаемость в этом направлении. Из первого уравнения (2.25) выразим потенциал через компоненты смещений:

$$\varphi = \frac{\left(e_{ilm}\nu_i\nu_l\right)}{\varepsilon}u_m,$$

и подставим во второе уравнение (2.25). В результате получаем однородную систему трех линейных уравнений:

$$\left(\rho\omega^2\delta_{im} - \kappa^2\Gamma^e_{im}\right)u_m = 0,$$

где компоненты симметричного тензора Γ^e_{im} имеют вид

$$\Gamma^e_{im} = C_{iklm}\nu_k\nu_l + \frac{\left(e_{pqi}\nu_p\nu_q\right)\left(e_{pqm}\nu_p\nu_q\right)}{\varepsilon}.$$

Если компоненты материальных тензоров C_{iklm}, e_{ilm} и компоненты единичного вектора в направлении распространения ν_i приведены к главным осям тензора Γ^e_{im}, то его главные значения запишутся как:

$$\Gamma^e_1 = C_{1kl1}\nu_k\nu_l + \frac{\left(e_{pq1}\nu_p\nu_q\right)^2}{\varepsilon},$$

$$\Gamma^e_2 = C_{2kl2}\nu_k\nu_l + \frac{\left(e_{pq2}\nu_p\nu_q\right)^2}{\varepsilon},$$

$$\Gamma^e_3 = C_{3kl3}\nu_k\nu_l + \frac{\left(e_{pq3}\nu_p\nu_q\right)^2}{\varepsilon}.$$

Они определяют три независимые моды взаимно ортогональной поляризации по главным осям. Фазовые скорости мод равны:

$$v_1^e = \sqrt{\frac{C_{1kl1}\nu_k\nu_l}{\rho}} + \left(1 + K_1^2\right),$$

$$v_2^e = \sqrt{\frac{C_{2kl2}\nu_k\nu_l}{\rho}} + \left(1 + K_2^2\right),$$

$$v_3^e = \sqrt{\frac{C_{3kl3}\nu_k\nu_l}{\rho}} + \left(1 + K_3^2\right),$$

где введено обозначение:

$$K_1^2 = \frac{\left(e_{pq1}\nu_p\nu_q\right)^2}{\varepsilon C_{1km1}\nu_k\nu_m}, \quad K_2^2 = \frac{\left(e_{pq2}\nu_p\nu_q\right)^2}{\varepsilon C_{2km2}\nu_k\nu_m}, \quad K_3^2 = \frac{\left(e_{pq3}\nu_p\nu_q\right)^2}{\varepsilon C_{3km3}\nu_k\nu_m}.$$

Так как K_i^2 – положительные величины, то K_i являются вещественными. Их называют коэффициентами электромеханической связи. Они различны для разных мод и зависят от направления распространения.

Сравнивая выражения для фазовых скоростей распространяющихся объемных мод в пьезоэлектрике и соответствующие им выражения для непьезоактивного кристалла (2.18), делаем вывод, что пьезоэффект приводит к некоторому увеличению фазовых скоростей объемных волн по сравнению с соответствующими волнами в непьезоактивном кристалле. Можно сказать, что пьезоэлектричество увеличивает эффективные упругие константы в $\left(1 + K^2\right)$ раз, то есть делает среду для волн более «жесткой». Однако на практике даже для самых сильных пьезоэлектриков коэффициенты электромеханической связи значительно меньше единицы, и относительное увеличение фазовых скоростей объемных волн за счет пьезоэффекта

$$\frac{v^e - v}{v} \cong \frac{K^2}{2} \tag{2.26}$$

не превышает нескольких процентов. Таким образом, объемная упругая волна в пьезоэлектриках имеет преимущественно механическую природу, а пьезоэффект представляет собой лишь малое возмущение.

2.5. Поверхностные акустические волны в изотропном полупространстве

Во многих практических случаях приходится иметь дело с распространением упругих возмущений в структурах, содержащих плоские границы раздела. Среди возможных решений уравнения движения в таких структурах значительный практический интерес представляют так называемые поверхностные волны, распространяющиеся вдоль границы раздела сред и быстро затухающие при удалении от нее. В зависимости от свойств граничащих сред, ориентации вектора смещения по отношению к границе раздела и направлению распространения выделяют различные виды поверхностных волн. Число их достаточно велико. В дальнейшем мы ограничимся лишь некоторыми часто встречающимися на практике и наиболее изученными видами поверхностных волн.

Пусть упругая однородная изотропная среда занимает полупространство $x_3 < 0$ и граничит с вакуумом или какой-либо другой средой, чьей упругостью можно пренебречь, например воздухом. Не нарушая общности, можно считать, что поверхностная волна распространяется в положительном направлении оси x_1. Плоскость, образованную нормалью к поверхности и направлением распространения, называют сагиттальной плоскостью (рис. 2.2). Исследуем поверхностную волну с вектором смещения в сагиттальной плоскости. Ее амплитуда постоянна на границе полупространства и убывает вглубь среды. Данная волна получила название волны Рэлея.

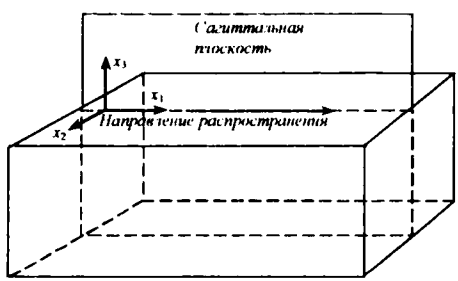


Рис. 2.2. Система координат, используемая при анализе волны Рэлея

Для анализа волны Рэлея вновь запишем волновое уравнение:

$$\frac{1}{\mathrm{v}^2}\frac{\partial^2 u}{\partial t^2}-\nabla^2 u=0\,, \tag{2.27}$$

где u – одна из компонент векторов $\vec{u}_T$ или $\vec{u}_L$, v – соответствующая фазовая скорость. В волне Рэлея присутствуют только компоненты u_1 и u_3. Будем искать решение в виде плоской неоднородной монохроматической волны, распространяющейся по оси x_1:

$$u=f\left(x_3\right)\exp\left[j\left(\omega t-\kappa x_1\right)\right]. \tag{2.28}$$

В отличие от рассмотренных ранее однородных волн ее амплитуда зависит от x_3, но по-прежнему не зависит от x_2. Подставляя (2.28) в (2.27), получаем уравнение

$$\frac{\partial^2 f}{\partial x_3^2}-\left(\kappa^2-\omega^2/\mathrm{v}^2\right)f=0\,. \tag{2.29}$$

Если $\kappa^2<\omega^2/\mathrm{v}^2$, то $f\left(x_3\right)$ – периодическая функция и выражение (2.28) описывает волну, не исчезающую во всем объеме среды и не связанную с поверхностью. Если же $\kappa^2>\omega^2/\mathrm{v}^2$, то решениями уравнения (2.29) являются экспоненциальные функции $\exp\left(\pm x_3\beta\right)$, где $\beta=\sqrt{\kappa^2-\omega^2/\mathrm{v}^2}$ – вещественное число. Поверхностной волне соответствует та из них, которая описывает смещения, экспоненциально убывающие по мере удаления от границы полупространства вглубь среды. Таким образом, мы приходим к следующему решению волнового уравнения, описывающему поверхностную волну:

$$\begin{aligned} u&=C\exp\left(x_3\beta\right)\exp\left[j\left(\omega t-\kappa x_1\right)\right],\\ \beta&=\sqrt{\kappa^2-\omega^2/\mathrm{v}^2}\,, \end{aligned} \tag{2.30}$$

где произвольная константа C не зависит от времени и координат.

На границе полупространства силы, действующие со стороны вакуума, равны нулю. Так как в выбранной системе координат вектор нормали к границе направлен по оси x_3, данное граничное условие сводится к требованию $T_{13} = T_{23} = T_{33} = 0$. Отсюда с учетом закона Гука для изотропной среды (1.31) и определения тензора деформаций следует:

$$\left.\begin{aligned} &\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} = 0 \\ &\lambda \frac{\partial u_1}{\partial x_1} + (\lambda + 2\mu)\frac{\partial u_3}{\partial x_3} = 0 \end{aligned}\right\} \text{при } x_3 = 0. \qquad (2.31)$$

Как ранее указывалось, в общем случае вектор смещения, описывающий движение частиц среды, является суммой векторов $\vec{u}_L$ и $\vec{u}_T$, компоненты каждого из которых удовлетворяют волновому уравнению (2.27) со скоростями $v = v_L$ для компонент $\vec{u}_L$ и $v = v_T$ для компонент $\vec{u}_T$. В неограниченной среде эти две части представляют собой две независимо распространяющиеся волны. В случае же поверхностных волн вследствие требований (2.31), налагаемых на суммарные компоненты смещения на границе, это не так. Поверхностная волна должна быть некоторой вполне определенной линейной комбинацией безвихревой и соленоидальной волн. Кроме того, для того чтобы выполнялись граничные условия одновременно для всех точек поверхности, необходимо, чтобы безвихревая и соленоидальная части волны являлись одинаковыми функциями x_1. Следовательно, постоянная распространения κ для обеих составляющих должна быть одинакова. При этом постоянные β, описывающие убывание амплитуды каждой из волн при удалении от границы, из-за различия скоростей v_L и v_T будут отличаться:

$$\begin{aligned} \beta_L &= \sqrt{\kappa^2 - \omega^2/v_L^2}, \\ \beta_T &= \sqrt{\kappa^2 - \omega^2/v_T^2}. \end{aligned} \qquad (2.32)$$

Соленоидальная часть волны должна удовлетворять условию $\operatorname{div} u_T = 0$, или

$$\frac{\partial u_{T1}}{\partial x_1} + \frac{\partial u_{T3}}{\partial x_3} = 0 .$$

Ввиду соотношений (2.30) это условие приводит к равенству

$$-j\kappa u_{T1} + \beta_T u_{T3} = 0 .$$

Чтобы данное равенство выполнялось, следует положить:

$$\begin{aligned} u_{T1} &= \beta_T A \exp(\beta_T x_3) \exp\left[j(\omega t - \kappa x_1) \right], \\ u_{T3} &= j\kappa A \exp(\beta_T x_3) \exp\left[j(\omega t - \kappa x_1) \right]. \end{aligned} \qquad (2.33)$$

где A – произвольная константа.

Безвихревая часть удовлетворяет условию $\operatorname{rot} \vec{u}_L = 0$, или

$$\frac{\partial u_{L1}}{\partial x_3} - \frac{\partial u_{L3}}{\partial x_1} = 0 .$$

Откуда

$$\beta_L u_{L1} + j\kappa u_{L3} = 0 .$$

Это равенство выполняется, если:

$$\begin{aligned} u_{L1} &= \kappa B \exp(\beta_L x_3) \exp\left[j(\omega t - \kappa x_1) \right], \\ u_{L3} &= j\beta_L B \exp(\beta_L x_3) \exp\left[j(\omega t - \kappa x_1) \right], \end{aligned} \qquad (2.34)$$

где B – произвольная константа.

Заметим, что в поверхностной волне в силу ее неоднородности соленоидальная и безвихревая составляющие уже не имеют простого наглядного смысла поперечных и продольных волн. Каждая из этих

составляющих имеет как параллельные, так и перпендикулярные по отношению к направлению распространения компоненты смещения.

Компоненты суммарного смещения определятся как

$$u_1 = u_{L1} + u_{T1},$$
$$u_3 = u_{L3} + u_{T3}.$$

Подставляя эти выражения в (2.31), получаем систему двух уравнений относительно неизвестных постоянных A и B:

$$\left(\kappa^2 + \beta_T^2\right) A + 2\kappa\beta_L B = 0,$$
$$2\kappa\beta_T A + \left[\frac{\lambda}{\mu}\left(\beta_L^2 - \kappa^2\right) + 2\beta_L^2\right] B = 0.$$

Из соотношений (2.32) и выражений для скоростей соленоидальных и безвихревых волн (2.14), (2.15) можно получить:

$$\frac{\lambda}{\mu} = \frac{\beta_T^2 - \kappa^2}{\beta_L^2 - \kappa^2} - 2.$$

Теперь система уравнений для определения неизвестных постоянных A и B приобретает следующий вид:

$$\left(\kappa^2 + \beta_T^2\right) A + 2\kappa\beta_L B = 0,$$
$$2\beta_T \kappa A + \left(\kappa^2 + \beta_T^2\right) B = 0. \qquad (2.35)$$

Для существования волны Рэлея требуется, чтобы система однородных линейных уравнений (2.35) имела ненулевые решения. Приравнивая главный определитель системы к нулю, получаем дисперсионное уравнение, устанавливающее связь частоты и постоянной распространения:

$$\left(\kappa^2 + \beta_T^2\right)^2 = 4\kappa^2\beta_T\beta_L. \qquad (2.36)$$

Оно называется уравнением Рэлея.

Для поиска явной зависимости $\omega = \omega(\kappa)$ возведем (2.36) в квадрат и подставим в него выражения (2.32):

$$\left(2\kappa^2 - \frac{\omega^2}{v_T^2}\right)^4 = 16\kappa^4\left(\kappa^2 - \frac{\omega^2}{v_T^2}\right)\left(\kappa^2 - \frac{\omega^2}{v_L^2}\right). \qquad (2.37)$$

Надо иметь в виду, что уравнение (2.37) получено путем возведения в квадрат и может иметь лишние корни по сравнению с исходным уравнением Рэлея. Представим искомую зависимость в виде

$$\omega = \xi v_T \kappa, \qquad (2.38)$$

где ξ – некоторая неизвестная функция. Поскольку ξ заранее не определена, такая запись всегда возможна. Подставляя (2.38) в (2.37), сокращая на κ^8 и раскрывая скобки, получаем:

$$\xi^6 - 8\xi^4 + 8\xi^2\left(3 - 2\frac{v_T^2}{v_L^2}\right) - 16\left(1 - \frac{v_T^2}{v_L^2}\right) = 0. \qquad (2.39)$$

Отсюда видно, что ξ зависит только от отношения скоростей v_T^2/v_L^2, которое является определенной характеристикой данной упругой среды. Отношение скоростей можно выразить через коэффициент Пуассона σ:

$$\frac{v_T^2}{v_L^2} = \frac{1-2\sigma}{2(1-\sigma)}, \quad \sigma = \frac{\lambda}{2(\lambda+\mu)}.$$

Таким образом, для волны Рэлея частота пропорциональна постоянной распространения. Поэтому фазовая скорость не зависит от частоты, а определяется только упругими свойствами среды:

$$v_R = \frac{\omega}{\kappa} = v_T \xi.$$

Волна Рэлея, так же как и рассмотренные ранее объемные волны, не имеет дисперсии фазовой скорости.

Величина ξ должна быть, разумеется, вещественной положительной, причем $\xi < 1$, так как β_T и β_L должны быть вещественными. Исследование показывает, что уравнение (2.39) имеет единственный корень, удовлетворяющий данным условиям и являющийся одновременно корнем исходного уравнения (2.36). Приближенное значение этого корня:

$$\xi_R = \frac{0.87 + 1.12\sigma}{1 + \sigma}.$$

При изменении коэффициента Пуассона σ от 0 до 1/2 ξ_R изменяется от 0.87 до 0.96. Следовательно, скорость рэлеевской волны v_R всегда несколько меньше (в среднем примерно на 10%) скорости поперечной объемной волны и, конечно же, меньше скорости продольной объемной волны, так как $v_L > \sqrt{2} v_T$.

Длина волны Рэлея Λ_R – пространственный период вдоль направления распространения определится через постоянную распространения:

$$\Lambda_R = \frac{2\pi v_R}{\omega} = \frac{2\pi v_T \xi_R}{\omega} = \Lambda_T \xi_R.$$

Так как $\xi_R < 1$, то пространственный период деформаций и напряжений в рэлеевской волне всегда меньше, чем период деформаций и напряжений вдоль поверхности, которые могла бы создать любая объемная волна, движущаяся под произвольным углом. Значит, рэлеевская волна на поверхности изотропного полупространства, несмотря на присутствие в ней продольных и поперечных смещений, не может совпадать по фазе ни с одной из объемных мод и не может привести к их возбуждению.

Из (2.35) можно выразить отношение неизвестных A и B:

$$\frac{B}{A} = -\frac{\kappa^2 + \beta_T^2}{2\kappa\beta_L} = -\frac{2\kappa\beta_T}{\kappa^2 + \beta_T^2}.$$

Компоненты смещения в рэлеевской волне определятся соотношениями:

$$u_1 = \hat{u}_1 \exp\left[j\left(\omega t - \kappa x_1\right)\right], \quad u_3 = j\hat{u}_3 \exp\left[j\left(\omega t - \kappa x_1\right)\right],$$

где амплитуды смещения по осям x_1 и x_3 равны:

$$\hat{u}_1 = \beta_T A\left[\exp\left(\beta_T x_3\right) - \frac{2\kappa^2}{\kappa^2 + \beta_T^2} \exp\left(\beta_L x_3\right)\right],$$

$$\hat{u}_3 = \kappa A\left[\exp\left(\beta_T x_3\right) - \frac{2\beta_T \beta_L}{\kappa^2 + \beta_T^2} \exp\left(\beta_L x_3\right)\right].$$

Зная компоненты смещения, нетрудно найти напряжения и деформации. На рис. 2.3 показана мгновенная картина смещений и деформаций в волне Рэлея. Точками обозначены положения частиц среды, которые в отсутствие волны расположены на равных расстояниях по вертикали и горизонтали одна от другой.

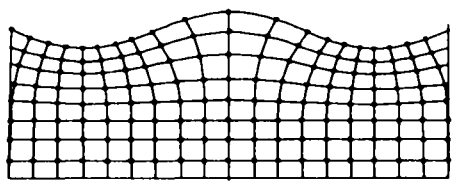


Рис. 2.3. Мгновенная картина смещений и деформаций в волне Рэлея

На рис. 2.4 представлены зависимости амплитуд смещений и амплитуд напряжений в рэлеевской волне от глубины. Кривые даны в безразмерной форме: амплитуды смещений отнесены к амплитуде нормального смещения на поверхности, а амплитуды напряжений – к поверхностному значению амплитуды компоненты T_{11}. Глубина отложена в долях длины рэлеевской волны. Смещение, нормальное к поверхности, сначала возрастает, а затем монотонно убывает. Продольное смещение меняет знак на глубине примерно $0.2\Lambda_R$. Имея нулевое значение на поверхности, компоненты нормального напряжения сначала возрастают, на глуби-

не около $0.3\Lambda_R$, достигают максимума, а затем быстро убывают. Продольное напряжение в точке максимума нормального напряжения проходит через нуль и меняет знак.

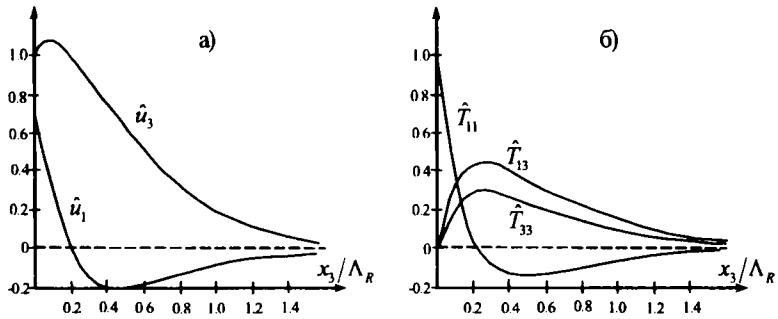


Рис. 2.4. Зависимость нормированных амплитуд смещений (а) и напряжений (б) в волне Рэлея от относительной глубины. Коэффициент Пуассона $\sigma = 0{,}25$

Так как поперечная компонента смещения сдвинута по фазе по отношению к продольной компоненте на $\pi/2$, траекториями движения частиц среды являются эллипсы. Вращение по эллипсу вблизи поверхности происходит против часовой стрелки, а на глубине $0.2\Lambda_R$, когда продольное смещение меняет знак, направление вращения меняется на противоположное. Отметим также, что на глубине более $1.5\Lambda_R$ движение и напряжение практически отсутствуют, то есть можно считать, что глубина проникновения волны Рэлея $\sim 1.5\Lambda_R$.

Рэлеевская волна характеризуется продольными и поперечными смещениями в сагиттальной плоскости и имеет, таким образом, вертикальную поляризацию. Исследуем теперь возможность существования в изотропном полупространстве поверхностной волны с вектором смещения, ортогональным сагиттальной плоскости. Такая волна будет представлять собой поперечную волну горизонтальной поляризации. Вектор смещения этой волны имеет единственную компоненту u_2.

Рассмотрим решение волнового уравнения в виде плоской поперечной поверхностной волны:

$$u_2 = A\exp(\beta_T x_3)\exp\left[j(\omega t - \kappa x_1)\right].$$

С помощью закона Гука (1.31) легко определить, что механические напряжения T_{13} и T_{33}, создаваемые этой волной на поверхности тождественно обращаются в нуль, в то время как $T_{23} = -j\mu\beta_T u_2$. Следовательно, если $\beta_T \neq 0$, то нельзя удовлетворить граничному условию $T_{23} = 0$ при $x_3 = 0$. Соответственно, не существует решения вида волны, связанной с поверхностью. Однако, если $\beta_T = 0$, то все компоненты нормального напряжения T_{i3} на границе полупространства равны нулю. Такое решение описывает просто скользящую вдоль поверхности однородную поперечную волну сдвига, которую называют SH-волной. Фазовая скорость этой волны, очевидно, равна v_T и не зависит от частоты.

Распространяющаяся вдоль поверхности однородная волна горизонтальной поляризации, как будет видно в дальнейшем, неустойчива в том смысле, что небольшое изменение граничных условий или свойств среды превращает ее в поверхностную. Примером этой неустойчивости являются волны Лява – второй основной тип поверхностных волн, возникающих при добавлении к полупространству твердого слоя, являющегося механической нагрузкой для полупространства.

2.6. Акустические волны в полупространстве со слоем

Пусть изотропный слой толщины h расположен на поверхности изотропной подложки, занимающей полупространство $x_3 < 0$. Рассмотрим плоскую поперечную волну горизонтальной поляризации, распространяющуюся вдоль границы подложки и слоя в положительном направлении оси x_1 (рис. 2.5). Будем считать, что соприкасающиеся поверхности слоя и подложки жестко связаны, так что смещения в обеих средах на границе раздела одинаковы.

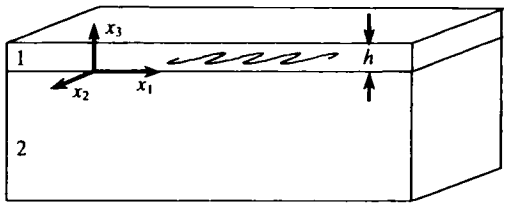


Рис. 2.5. Горизонтальная волна сдвига на границе полупространства и твердого слоя

Как в слое (среда 1), так и в подложке (среда 2) единственная ненулевая компонента смещения u_2 должна удовлетворять волновому уравнению со значениями скоростей поперечных объемных волн в каждой из сред. В среде 2 выберем решение в виде поверхностной волны, затухающей вглубь полупространства:

$$u_2^{(2)} = A\exp(\beta_2 x_3)\exp\left[j(\omega t - \kappa x_1)\right],$$

где $\beta_2 = \sqrt{\kappa^2 - \omega^2/v_{T2}^2}$ – вещественная положительная величина, A – произвольная константа. В среде 1 ищем решение в виде неоднородной волны:

$$u_2^{(1)} = f(x_3)\exp\left[j(\omega t - \kappa x_1)\right].$$

Для функции $f(x_3)$ получаем уравнение

$$\frac{d^2 f}{dx_3^2} + \beta_1^2 f = 0,$$
$$\beta_1 = \sqrt{\omega^2/v_{T1}^2 - \kappa^2}.$$

Далее будет установлено, что величина β_1 должна быть вещественной, поэтому

$$f(x_3) = B\sin(\beta_1 x_3) + C\cos(\beta_1 x_3),$$

где B, C – некоторые константы. Таким образом, решение для слоя приобретает вид

$$u_2^{(1)} = \left[B \sin\left(\beta_1 x_3\right) + C \cos\left(\beta_1 x_3\right) \right] \exp\left[j\left(\omega t - \kappa x_1\right) \right].$$

На свободной границе слоя $x_3 = h$ должно выполняться условие отсутствия нормального напряжения $T_{32}^{(1)} = 0$, то есть

$$\frac{\partial u_2^{(1)}}{\partial x_3} = 0 \quad \text{при } x_3 = h. \tag{2.40}$$

На плоскости $x_3 = 0$ должны выполняться условия непрерывности нормальных компонент напряжения и компонент смещения:

$$\mu_1 \frac{\partial u_2^{(1)}}{\partial x_3} = \mu_2 \frac{\partial u_2^{(2)}}{\partial x_3}, \quad u_2^{(1)} = u_2^{(2)} \quad \text{при } x_3 = 0, \tag{2.41}$$

где μ_1, μ_2 – модули сдвига каждой из сред.

Подстановка выражений для смещения в каждой из сред в краевые условия (2.40) и (2.41) после необходимых вычислений приводит к системе трех однородных линейных уравнений относительно неизвестных A, B и C:

$$\begin{aligned} & B\cos\left(\beta_1 h\right) - C\sin\left(\beta_1 h\right) = 0, \\ & \mu_2 \beta_2 A - \mu_1 \beta_1 B = 0, \\ & A - C = 0. \end{aligned} \tag{2.42}$$

Приравнивая определитель системы к нулю, получаем дисперсионное уравнение:

$$\operatorname{tg}\left(\beta_1 h\right) = \frac{\mu_2 \beta_2}{\mu_1 \beta_1}, \tag{2.43}$$

устанавливающее в неявном виде зависимость $\omega(\kappa)$. Проанализируем это уравнение.

Величина $\beta_2 = \sqrt{\kappa^2 - \omega^2/v_{T2}^2}$ вследствие требования поверхностного характера волны должна быть вещественной и положительной, поэтому $\kappa > \omega/v_{T2}$. Рассмотрим теперь величину $\beta_1 = \sqrt{\omega^2/v_{T1}^2 - \kappa^2}$. Она может быть вещественной или чисто мнимой. Предположим, что она чисто мнимая. Обозначим $\beta_1 = i\alpha$, где α – вещественная величина. Тогда уравнение (2.43) приобретает вид

$$-\mathrm{th}(\alpha h) = \frac{\mu_2 \beta_2}{\mu_1 \alpha}.$$

При любом α знаки левой и правой частей уравнения противоположны. Следовательно, для мнимых β_1 дисперсионное уравнение не имеет решений. Этим подтверждается сделанное ранее предположение о вещественности величины β_1. Поэтому должно быть выполнено неравенство $\kappa < \omega/v_{T1}$. Объединяя полученные неравенства, приходим к необходимому условию существования решения дисперсионного уравнения:

$$\omega/v_{T2} < \kappa < \omega/v_{T1}.$$

Оно может быть выполнено, если только $v_{T1} < v_{T2}$ (замедляющий слой). При этом фазовая скорость волны $v = \omega/\kappa$ будет лежать в пределах

$$v_{T1} < v < v_{T2}.$$

Нетрудно показать, что при выполнении этого условия решение дисперсионного равнения всегда существует.

Таким образом, если на поверхности однородной изотропной подложки присутствует замедляющий слой, поверхностная волна горизонтальной поляризации всегда существует. Ее называют волной Лява. Фазовая скорость волны Лява всегда меньше скорости поперечной объемной волны в подложке, но больше скорости этой волн в замедляющем слое. При малой толщине слоя и низкой частоте дисперсионное уравнение имеет единственный корень. По мере увеличения толщины слоя и частоты число корней увеличивается.

Эти корни соответствуют различным распространяющимся модам – волнам Лява разного порядка.

Фазовые и групповые скорости волн Лява обладают дисперсией. На рис. 2.6 представлены зависимости фазовых и групповых скоростей трех первых волн Лява от безразмерного параметра $\omega h/v_{T1}$. При высоких частотах слоя фазовая и групповая скорости волн приближаются к скорости поперечной волны в слое v_{T1}. С понижением частоты фазовые скорости монотонно возрастают и постепенно стремятся к скорости волны в подложке v_{T2}. Групповые скорости имеют минимум в области наибольшей дисперсии фазовой скорости. Для каждой волны Лява, кроме первой, при заданных параметрах слоя и подложки существует некоторая минимальная частота отсечки, ниже которой распространение этой волны становится невозможным.

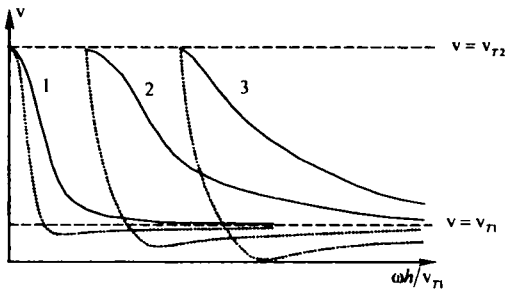


Рис. 2.6. Дисперсия фазовых (сплошные линии) и групповых (штриховые) скоростей трех первых волн Лява

Из системы уравнений (2.42) можно выразить неизвестные константы B и C через единственную произвольную константу A. Если обозначить $u_0 = A/\cos(\beta_1 h)$, тогда смещения приобретают вид

$$u_2 = \hat{u}_2\left[j\left(\omega t - \kappa x_1\right)\right],$$

где амплитуды в каждой из сред определяются выражениями:

$$\hat{u}_2^{(1)} = u_0 \cos\left[\beta_1 (x_3 - h)\right],$$
$$\hat{u}_2^{(2)} = u_0 \cos(\beta_1 h) \exp(\beta_2 x_3).$$

На рис. 2.7 представлены распределения амплитуды смещений по глубине структуры для волн Лява первых трех порядков. Амплитуда смещений в волне первого порядка монотонно убывает при удалении от верхней (свободной) границы слоя, а во второй и третьей волнах Лява в слое имеются узлы и пучности смещений. С ростом толщины слоя волна Лява все меньше проникает в подложку и почти целиком локализуется в слое.

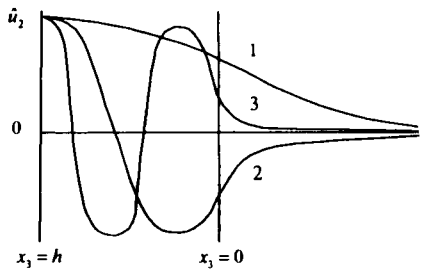


Рис. 2.7. Распределение амплитуды смещений в волнах Лява первых трех порядков

Часто под волной Лява понимают только волну первого порядка, существующую при любой толщине замедляющего слоя. При очень малых толщинах $(\omega h / v_{T1} << 1)$ смещения в слое постоянны по толщине и медленно убывают в подложке. Фазовая скорость лишь немного меньше, чем v_{T2}.

В полупространстве со слоем возможны также волны вертикальной поляризации с вектором смещения в сагиттальной плоскости. Решение для этих волн можно получить, суммируя в каждой из сред парциальные волны, подобно тому, как это делалось в разделе 2.5. Однако вычисления становятся существенно сложнее.

В подложке (среда 2) смещение для рэлеевской волны будет попрежнему являться суммой соленоидальной и безвихревой волн вида (2.30):

$$u^{(2)} = C\exp(x_3\beta)\exp\left[j(\omega t - \kappa x_1)\right].$$

Однако для среды 1 ввиду конечной толщины слоя уже нет оснований отбрасывать решения, пропорциональные $\exp(-x_3\beta)$, и вместо (2.30) общее решение волнового уравнения в слое надо записывать в виде:

$$u^{(1)} = \left[A\exp(x_3\beta) + B\exp(-x_3\beta)\right]\exp\left[j(\omega t - \kappa x_1)\right].$$

В результате суммарные компоненты смещения в слое $u_{1,3}^{(1)} = u_{L1,3}^{(1)} + u_{T1,3}^{(1)}$ будут содержать четыре произвольные константы. Граничные условия требуют, чтобы механические напряжения T_{i3} обращались в нуль на свободной поверхности $x_3 = h$. В тоже время на границе раздела слоя и подложки $(x_3 = 0)$ напряжения и смещения должны быть непрерывны. Из данных требований получается система шести линейных уравнений, связывающих шесть неизвестных констант. Условие наличия ненулевых решений системы приводит к дисперсионному уравнению.

В общем виде дисперсионное уравнение достаточно громоздко и полный его анализ может быть выполнен только численными методами. При $\omega h \to 0$ дисперсионное уравнение имеет единственный корень, отвечающий поверхностной рэлеевской волне. При увеличении толщины и частоты появляются новые корни, описывающие различные моды рэлеевской волны в системе полупространство – твердый слой. Скорость первой моды является наименьшей. С ростом номера моды фазовая скорость возрастает. Высшие моды вертикальной поляризации в полупространстве с замедляющим слоем иногда именуют волнами Сезавы.

Если скорость волн в материале слоя меньше, чем в подложке – $v_{R1} < v_{R2}$ (замедляющий слой), то на низких частотах толщина слоя намного меньше длины волны, и скорость основной моды стремится к скорости рэлеевских волн в материале подложки v_{R2}. На высоких частотах волна концентрируется у свободной поверхности слоя, и скорость ее распространения стремится к скорости рэлеевской волны в материале слоя v_{R1}. Если скорости волн в материале слоя оказываются большими, чем в подложке – $v_{R1} > v_{R2}$, скорость основной

моды изменяется от значения v_{R2} на нулевой частоте до значения v_{T2}, при котором происходит частотная отсечка.

Возможность управления скоростью распространения упругих поверхностных волн в твердом теле с помощью слоя на его поверхности может быть использована для создания акустических поверхностных волноводов. Если на поверхности твердого тела нанесена полоска из соответствующим образом подобранного материала, то вследствие уменьшения скорости поверхностных волн под этой полоской по сравнению со свободной поверхностью может возникать волноводный эффект. При этом поверхностная волна локализуется под полоской и образуется акустический волновод.

2.7. Волны Лява в полупространстве со слабой поверхностной неоднородностью

Важным с практической точки зрения обобщением волн Лява являются поперечные поверхностные волны в полупространстве с небольшой поверхностной неоднородностью. Такая неоднородность появляется, например, при химической и механической обработке поверхности, либо при лазерном воздействии, а также при ионной имплантации.

Рассмотрим простейший случай упругого изотропного материала, занимающего полупространство $x_3 < 0$. Предположим, что плотность ρ и модуль сдвига μ изменяются по следующим законам:

$$\rho(x_3) = \rho_0 + \Delta\rho \exp(x_3/h), \quad \mu(x_3) = \mu_0 - \Delta\mu \exp(x_3/h),$$

где h – характерная глубина неоднородности, ρ_0 и μ_0 – значения плотности и модуля сдвига в глубине материала, $\Delta\rho$ и $\Delta\mu$ – их отклонения на поверхности, причем $|\Delta\rho| << \rho_0$ и $|\Delta\mu| << \mu_0$.

Будем искать решение уравнения движения (2.3) в виде поперечной плоской неоднородной волны, распространяющейся вдоль поверхности в положительном направлении оси x_1:

$$u_2 = AF(x_3)\exp(-\beta x_3)\exp\left[j(\omega t - \kappa x_1)\right], \qquad (2.44)$$

где A – произвольная постоянная, $F(x_3)$ – неизвестная функция, $\beta = \sqrt{\kappa^2 - \omega^2/v_T^2}$, κ – неизвестная постоянная распространения, $v_T = \sqrt{\mu_0/\rho_0}$ – скорость распространения поперечной объемной волны в однородной среде с параметрами ρ_0, μ_0.

В уравнении движения (2.3) тензор напряжений будет иметь тот же вид, что и в случае однородной среды, но при дифференцировании его компонент следует учитывать, что входящие в него упругие модули зависят от координат. В рассматриваемом нами случае горизонтальной волны (2.44) будут присутствовать только две независимые компоненты тензора деформаций:

$$U_{21} = \frac{1}{2}\frac{\partial u_2}{\partial x_1} \quad \text{и} \quad U_{23} = \frac{1}{2}\frac{\partial u_2}{\partial x_3}.$$

Подставляя их в закон Гука, имеем

$$T_{21} = \mu\frac{\partial u_2}{\partial x_1}, \quad T_{23} = \mu\frac{\partial u_2}{\partial x_3}.$$

Тогда уравнение движения (2.3) приобретает следующий вид:

$$\rho\frac{\partial^2 u_i}{\partial t^2} - \mu\left(\frac{\partial^2 u_2}{\partial x_1^2} + \frac{\partial^2 u_2}{\partial x_3^2}\right) + \frac{d\mu}{dx_3}\frac{\partial u_2}{\partial x_3} = 0.$$

Поскольку производная $d\mu/dx_3$ имеет порядок величины $\Delta\mu$, то последний член в левой части уравнения можно отбросить:

$$\rho\frac{\partial^2 u_i}{\partial t^2} - \mu\left(\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial x_3^2}\right) = 0. \qquad (2.45)$$

Подставим (2.44) в укороченное уравнение (2.45). Делая необходимые преобразования, с учетом малости величин $\Delta\rho/\rho_0$ и $\Delta\mu/\mu_0$ получаем дифференциальное уравнение для неизвестной функции $F(x_3)$:

$$\frac{d^2F}{dx_3^2}+2\beta\frac{\partial F}{\partial x_3}+\frac{\omega^2}{v_T^2}\xi F=0, \qquad (2.46)$$

где обозначено:

$$\xi=\delta\exp(x_3/h),$$
$$\delta=(\Delta\rho/\rho_0+\Delta\mu/\mu_0).$$

Очевидно, функция F должна зависеть от ξ, а при $\xi=0$ тождественно обращаться в единицу. Поэтому, представим ее в виде разложения по малому параметру ξ:

$$F(x_3)=1+\alpha_1\xi+\alpha_2\xi^2+\ldots.$$

Подставляя это разложение в уравнение (2.46) и удерживая только линейные по ξ члены, получаем

$$F(x_3)=1-\frac{(\omega h/v_T)^2}{(1+2\beta h)}\xi. \qquad (2.47)$$

На поверхности полупространства должна отсутствовать компонента нормального напряжения T_{32}:

$$\mu\frac{\partial u_2}{\partial x_3}=0 \text{ при } x_3=0.$$

Подставляя в данное граничное условие решение (2.44) с учетом (2.47), получаем дисперсионное уравнение

$$\frac{\omega^2 h^2}{v_T^2}\frac{\delta}{\beta h}=\frac{1+2\beta h}{1+\beta h}.$$

Из него находим:

$$\beta=\frac{\omega^2 h}{v_T^2}\delta,\quad \kappa=\frac{\omega}{v_T}\left(1+\frac{\omega^2 h^2}{v_T^2}\frac{\delta^2}{2}\right).$$

Поверхностная поперечная волна (2.44) будет существовать, если неоднородность $\delta>0$, что соответствует уменьшению скорости поперечной объемной волны вблизи поверхности.

Таким образом, при замедляющей неоднородности на поверхности упругого полупространства наряду с рэлеевской может распространяться дополнительная поверхностная поперечная волна, являющаяся некоторым обобщением волны Лява. Она локализована в слое толщиной $z_Л=1/\beta$:

$$z_Л=\frac{1}{\delta h}\frac{v_T^2}{\omega^2}=\frac{1}{\delta h}\frac{\Lambda_T^2}{4\pi^2}.$$

Ее фазовая скорость и длина волны

$$v_Л=v_T\left(1-\frac{\omega^2 h^2}{v_T^2}\frac{\delta^2}{2}\right)=v_T\left(1-\frac{2\pi^2 h^2}{\Lambda_T^2}\delta^2\right),$$

$$\Lambda_Л=\Lambda_T\left(1-\frac{2\pi^2 h^2}{\Lambda_T^2}\delta^2\right)$$

лишь незначительно меньше фазовой скорости и длины объемной поперечной волны.

Сделаем численные оценки. Пусть характерная глубина неоднородности равна половине длины волны: $h=\Lambda_T/2$. То при величине поверхностной неоднородности $\delta=5\%$ волна Лява локализована в слое $z_Л\cong\Lambda_Л$, а ее фазовая скорость примерно на 1% меньше скорости поперечной объемной волны.

2.8. Волны Стоунли

Волны Стоунли возникают на границе раздела двух упругих полуограниченных сред. Волны Стоунли подразделяют на два вида – волны горизонтальной и вертикальной поляризаций (рис. 2.8).

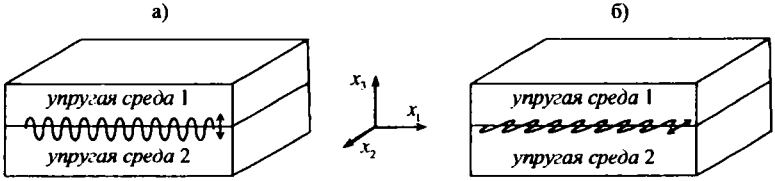


Рис. 2.8. Волны Стоунли на границе двух сред:
а – вертикальная поляризация; б – горизонтальная поляризация

Пусть граница жестко связанных изотропных сред совпадает с плоскостью $x_3 = 0$. Рассмотрим распространение плоской поверхностной волны с вертикальной поляризацией в положительном направлении оси x_1. Повторяя рассуждения, сделанные в разделе 2.5, будем считать, что волна в каждом из полупространств является суммой соленоидальной и безвихревой составляющих. Их компоненты удовлетворяют волновому уравнению (2.27) со скоростями v_L и v_T в каждой из сред. Обозначим компоненты смещения при $x_3 > 0$ как $u_i^{(1)}$, а при $x_3 < 0$ как $u_i^{(2)}$. Для их записи можно воспользоваться выражениями (2.32) – (2.34), делая для среды 1 замену $\beta_T \to -\beta_T$ и $\beta_L \to -\beta_L$:

$$
\begin{aligned}
u_1^{(1)} &= \left[-\beta_{T1}A_1 \exp\left(-\beta_{T1}x_3\right) + \kappa B_1 \exp\left(-\beta_{L1}x_3\right)\right]\exp\left[j\left(\omega t - \kappa x_1\right)\right],\\
u_3^{(1)} &= \left[j\kappa A_1 \exp\left(-\beta_{T1}x_3\right) - j\beta_{L1}B_1 \exp\left(-\beta_{L1}x_3\right)\right]\exp\left[j\left(\omega t - \kappa x_1\right)\right],\\
u_1^{(2)} &= \left[\beta_{T2}A_2 \exp\left(\beta_{T2}x_3\right) + \kappa B_2 \exp\left(\beta_{L2}x_3\right)\right]\exp\left[j\left(\omega t - \kappa x_1\right)\right],\\
u_3^{(2)} &= \left[j\kappa A_2 \exp\left(\beta_{T2}x_3\right) + j\beta_{L2}B_2 \exp\left(\beta_{L2}x_3\right)\right]\exp\left[j\left(\omega t - \kappa x_1\right)\right],\\
\beta_{T1,2} &= \sqrt{\kappa^2 - \omega^2/v_{T1,2}^2}, \quad \beta_{L1,2} = \sqrt{\kappa^2 - \omega^2/v_{L1,2}^2}.
\end{aligned}
\qquad (2.48)
$$

Компоненты тензора напряжений T_{13} и T_{33} в средах 1 и 2 определяются законом Гука (1.31):

$$T_{13} = \mu\left(\frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}\right), \quad T_{33} = \lambda\left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_3}{\partial x_3}\right) + 2\mu\frac{\partial u_3}{\partial x_3}. \tag{2.49}$$

Так как поверхности сред жестко связаны, на границе $x_3 = 0$ должны выполняться условия непрерывности напряжений и смещений:

$$\left.\begin{array}{l} T_{13}^{(1)} = T_{13}^{(2)}, \quad T_{33}^{(1)} = T_{33}^{(2)} \\ u_1^{(1)} = u_1^{(2)}, \quad u_3^{(1)} = u_3^{(2)} \end{array}\right\} \text{ при } x_3 = 0. \tag{2.50}$$

Подставляя (2.48) и (2.49) в (2.50), получаем однородную систему четырех линейных уравнений относительно неизвестных амплитуд $A_{1,2}$, $B_{1,2}$. Приравнивая нулю ее главный определитель, получаем дисперсионное уравнение

$$\begin{vmatrix} 2\mu_1\dfrac{\beta_{L1}}{\kappa} & 2\mu_2\dfrac{\beta_{L2}}{\kappa} & \mu_1\left(1+\dfrac{\beta_{T1}^2}{\kappa^2}\right) & \mu_2\left(1+\dfrac{\beta_{T2}^2}{\kappa^2}\right) \\ -\mu_1\left(1+\dfrac{\beta_{T1}^2}{\kappa^2}\right) & \mu_2\left(1+\dfrac{\beta_{T2}^2}{\kappa^2}\right) & -2\mu_1\dfrac{\beta_{T1}}{\kappa} & 2\mu_2\dfrac{\beta_{T2}}{\kappa} \\ 1 & -1 & \dfrac{\beta_{T1}}{\kappa} & -\dfrac{\beta_{T2}}{\kappa} \\ \dfrac{\beta_{L1}}{\kappa} & \dfrac{\beta_{L2}}{\kappa} & 1 & 1 \end{vmatrix} = 0. \tag{2.51}$$

Искомой поверхностной волне соответствует вещественный корень, удовлетворяющий условию:

$$\kappa > \omega^2/v_{T1}^2, \omega^2/v_{T2}^2. \tag{2.52}$$

Только в этом случае β_{T1}, β_{T2}, β_{L1}, β_{L2} – вещественные величины, и соотношения (2.48) описывают волну, локализованную вблизи границы сред.

Дисперсионное уравнение (2.51) достаточно сложно, и в общем случае произвольных параметров сред его исследование возможно только численными методами. Анализ показывает, что дисперсионное уравнение не всегда имеет корень, удовлетворяющий условию (2.52), то есть волна Стоунли с вертикальной поляризацией в отличие от волны Рэлея существует только при определенных соотношениях между параметрами граничащих сред. Приведем два примера. Прежде всего отметим, что если $\rho_2/\rho_1 \to 0$, то уравнение (2.52) переходит в уравнение Рэлея (2.36), а волна Стоунли вертикальной поляризации переходит в волну Рэлея. Это легко проверить, если в первых двух строках определителя (2.51) сделать подстановки $\mu_1 = \rho_1 v_{T1}^2$ и $\mu_2 = \rho_2 v_{T2}^2$. Можно также показать, что при достаточно близких скоростях упругих волн в средах ($v_{L1} \cong v_{L2}$ и $v_{T1} \cong v_{T2}$), но разных плотностях граничащих сред $(\rho_1 \neq \rho_2)$ волны Стоунли вертикальной поляризации всегда существуют.

Волны Стоунли, как и волны Рэлея, не обладают дисперсией. Их фазовая скорость всегда меньше скоростей продольных и поперечных волн в граничащих средах. Толщина слоев локализации волн Стоунли обычно близка к ее длине волны.

В отличие от волн вертикальной поляризации горизонтальные волны Стоунли не могут распространяться на границе изотропных сред. Это следует из того простого факта, что для горизонтальной волны на границе сред должна быть непрерывна компонента напряжения $T_{23} = -j\mu\beta_T u_2$. В сочетании с требованием непрерывности u_2 это возможно, только если свойства сред полностью совпадают. Однако горизонтальные волны Стоунли могут существовать при наличии промежуточного замедляющего слоя, в котором скорость поперечных волн меньше, чем в обеих средах. Их еще называют обобщенными волнами Лява. Кроме того, волны Стоунли горизонтальной поляризации могут существовать на границах анизотропных сред.

2.9. Нормальные волны в тонких пластинах

В предыдущих разделах мы предполагали, что `подложка, несущая поверхностную волну, является полубесконечной. Такое допущение вполне оправдано, когда глубина проникновения поверхно-

стной волны мала по сравнению с толщиной подложки и нет излучения энергии от поверхности вглубь материала. Если же это не выполняется, необходимо учитывать также граничные условия на нижней поверхности. Влияние обеих поверхностей существенно сказывается в тонких плоскопараллельных пластинах, где упругие волны распространяются в виде так называемых нормальных волн. Их обычно также относят к поверхностным, хотя движение в этих волнах происходит не только у поверхности, но и в толще пластины.

Нормальные волны бывают двух поляризаций: вертикальной поляризации – волны Лэмба, и горизонтальной поляризации – поперечные нормальные волны (рис. 2.9). Они имеют либо симметричную, либо антисимметричную картину смещений точек среды относительно центральной линии по толщине пластины. Соответственно, волны обоих видов поляризации делят на симметричные и антисимметричные.

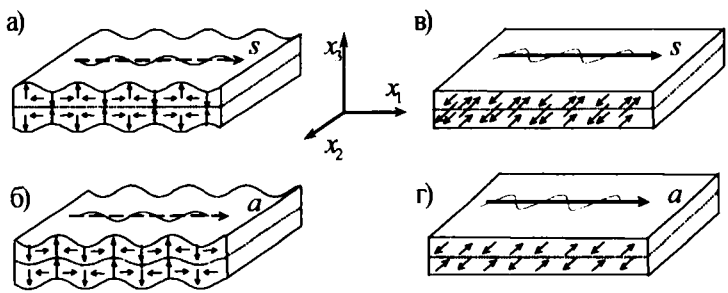


Рис. 2.9. Упругие волны в тонкой пластине: а – симметричная, б – антисимметричная волны Лэмба; в, г – симметричная и антисимметричная поперечные нормальные волны

Волны Лэмба характеризуются смещениями в сагиттальной плоскости (x_1, x_3). В симметричных волнах знаки компоненты u_1 одинаковы по обе стороны от центральной линии, а знаки компоненты u_3 – противоположны. Таким образом, по своей природе это волны растяжения пластины. В антисимметричных волнах Лэмба при переходе через центральную линию компонента u_1 изменяет свой знак, а компонента u_3 сохраняет знак, так что движение пластины для этих волн является изгибным

Рассмотрим пластину из некоторого изотропного материала, ограниченную плоскостями $x_3 = \pm h/2$. Поскольку смещение точек среды в волне Лэмба происходит в сагиттальной плоскости (x_1, x_3), парциальные волны, составляющие волну Лэмба, имеют такой же вид, что и в случае рэлеевских волн с той, однако, разницей, что, записывая выражение для каждой из парциальных волн, необходимо также учитывать решение волнового уравнения, пропорциональное $\exp(-\beta_{L,T} x_3)$. В волне Лэмба парциальные волны имеют вид:

$$\begin{aligned}
u_{T1} &= \left[\beta_T A \exp(\beta_T x_3) - \beta_T B \exp(-\beta_T x_3)\right] \exp\left[j(\omega t - k x_1)\right], \\
u_{T3} &= \left[j\kappa A \exp(\beta_T x_3) + j\kappa B \exp(-\beta_T x_3)\right] \exp\left[j(\omega t - k x_1)\right], \\
u_{L1} &= \left[\kappa C \exp(\beta_L x_3) + \kappa D \exp(-\beta_L x_3)\right] \exp\left[j(\omega t - k x_1)\right], \\
u_{L3} &= \left[j\beta_L C \exp(\beta_L x_3) - j\beta_L D \exp(-\beta_L x_3)\right] \exp\left[j(\omega t - k x_1)\right],
\end{aligned} \tag{2.53}$$

где β_T, β_L выражаются через постоянную распространения и фазовые скорости поперечных и продольных объемных волн соотношениями (2.32). Неизвестные постоянные A, B, C и D найдутся из требований равенства нулю компонент тензора напряжений T_{13} и T_{33} на обеих поверхностях пластины $x_3 = \pm h/2$. В результате получается однородная система четырех линейных уравнений:

$$\begin{aligned}
&\left(\beta_T^2 + \kappa^2\right) \mathrm{ch}\left(\beta_T h/2\right)(A + B) + 2\kappa\beta_L \,\mathrm{ch}\left(\beta_L h/2\right)(C - D) = 0, \\
&2\kappa\beta_T \,\mathrm{sh}\left(\beta_T h/2\right)(A + B) + \left(\beta_T^2 + \kappa^2\right) \mathrm{sh}\left(\beta_L h/2\right)(C - D) = 0, \\
&\left(\beta_T^2 + \kappa^2\right) \mathrm{sh}\left(\beta_T h/2\right)(A - B) + 2\kappa\beta_L \,\mathrm{sh}\left(\beta_L h/2\right)(C + D) = 0, \\
&2\kappa\beta_T \,\mathrm{ch}\left(\beta_T h/2\right)(A - B) + \left(\beta_T^2 + \kappa^2\right) \mathrm{ch}\left(\beta_L h/2\right)(C + D) = 0.
\end{aligned} \tag{2.54}$$

Простые рассуждения показывают, что решение системы (2.54) существует, если выполняется одно из двух условий. Первое из них

$$\frac{\mathrm{th}\left(\beta_L h/2\right)}{\mathrm{th}\left(\beta_T h/2\right)} = \frac{4\beta_L \beta_T \kappa^2}{\left(\beta_T^2 + \kappa^2\right)^2}, \tag{2.55}$$

$$A = B,\ \ C = -D\,.$$

При этом компонента смещения $u_1 = u_{T1} + u_{L1}$, как видно из (2.53), является нечетной функцией x_3, а компонента $u_3 = u_{T3} + u_{L3}$ – четной. Следовательно, соотношение (2.55) является дисперсионным уравнением для антисимметричных волн Лэмба. Второе условие:

$$\frac{\mathrm{th}\left(\beta_T h/2\right)}{\mathrm{th}\left(\beta_L h/2\right)} = \frac{4\beta_L\beta_T\kappa^2}{\left(\beta_T^2 + \kappa^2\right)^2}, \qquad (2.56)$$

$$A = -B,\ \ C = D\,.$$

В этом случае компонента u_1 – четная функция x_3, а u_3 – нечетная, и дисперсионное уравнение (2.56), таким образом, описывает симметричные волны Лэмба.

В зависимости от толщины пластины, частоты и упругих параметров среды уравнения (2.55) и (2.56) имеют различное число корней, и волны Лэмба могут распространяться в форме одной или нескольких мод, каждая из которых характеризуется своим распределением смещений и напряжений. Фазовые скорости мод отличаются и зависят от частоты.

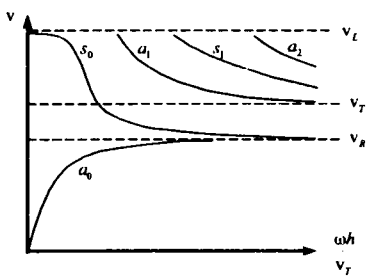


Рис. 2.10. Дисперсионные кривые для симметричных (s) и антисимметричных (a) мод изотропной пластины толщиной h; v_R, v_T, v_L – скорости рэлеевской и объемных поперечной и продольной волн соответственно

На рис. 2.10 показаны дисперсионные кривые для нескольких первых мод Лэмба. При малых значениях параметра $\omega h/v_T$ распро-

страняются только моды нулевого порядка s_0 и a_0, которые представляют собой соответственно продольную и изгибную волны в пластине. По мере увеличения толщины пластины и частоты число распространяющихся мод возрастает.

При малых $\omega h/\mathrm{v}_T$ симметричная мода s_0 весьма похожа на продольную волну в неограниченной среде: в ней преобладает продольная компонента смещения, и только потому, что обе поверхности пластины свободны, появляется небольшое поперечное смещение. Из-за уменьшения эффективной продольной упругости пластины за счет податливости свободных поверхностей фазовая скорость немного меньше фазовой скорости объемной продольной волны.

С ростом параметра $\omega h/\mathrm{v}_T$ моды s_0 и a_0 все более становятся похожими друг на друга, а их фазовые скорости стремятся к скорости рэлеевской волны v_R. Смещения все более локализуются у поверхностей пластины, и каждая из мод s_0 и a_0 превращается в две рэлеевские волны на обеих поверхностях пластины. Это дает некоторое представление о свойствах рэлеевской волны, возбуждаемой на одной из поверхностей пластины. Такую волну можно рассматривать как сумму симметричной и антисимметричной мод Лэмба с одинаковыми амплитудами и фазами. Ввиду свойств симметрии смещений эти моды суммируются на верхней поверхности и компенсируют друг друга на нижней. Однако при конечных значениях толщины пластины моды имеют различающиеся фазовые скорости v_a и v_s, а значит, и разные волновые числа $\kappa_a = \omega/\mathrm{v}_a$ и $\kappa_s = \omega/\mathrm{v}_s$. Вследствие этого по мере распространения между ними будет накапливаться фазовый сдвиг. После прохождения волной расстояния $x_c = \pi/(\kappa_a - \kappa_s)$ фазовый сдвиг станет равным π, и на верхней поверхности моды будут гасить друг друга, в то время как на нижней поверхности складываться. Сущность данного явления заключается в том, что происходит перекачка энергии рэлеевской волны с одной поверхности на другую. Поэтому расстояние x_c называют длиной связи. На расстоянии $2x_c$ амплитуда волны на верхней поверхности снова достигает максимума, таким образом, рэлеевская волна претерпевает обратную перекачку. Расчеты показывают, что длина связи более чем на три порядка превосходит длину волны даже в том случае, когда толщина пластины равна всего нескольким длинам волн. По существу, это означает, что при распространении рэлеев-

ских волн по поверхности пластин во многих практических случаях влиянием нижней поверхности можно пренебречь.

Моды Лэмба порядков выше нулевого появляются, начиная с некоторых критических для каждой моды значений $\omega h/\mathrm{v}_T$. Их можно найти, положив в дисперсионных уравнениях (2.55) и (2.56) $\kappa = 0$. При докритических $\omega h/\mathrm{v}_T$ постоянная κ – мнимая, и распространение соответствующей моды невозможно. В критической точке появляющаяся мода Лэмба представляет стоячую волну в поперечном сечении пластины, а ее фазовая скорость равна бесконечности и групповая – нулю. Когда значение параметра $\omega h/\mathrm{v}_T$ превышает критическое, фазовая скорость волны Лэмба становится конечной, а групповая – отличной от нуля.

При достаточно больших толщинах и высоких частотах у всех мод Лэмба, кроме s_0 и a_0, остается только одна поперечная компонента смещения u_3. Ее амплитуда на поверхности пластины мала по сравнению с амплитудой в толще пластины. Для мод s_0 и a_0, как ранее отмечалось, напротив, наблюдается своеобразный скин-эффект. Фазовые и групповые скорости всех высших мод при $\omega h/\mathrm{v}_T \to \infty$ стремятся к v_T.

Рассмотрим далее поперечные нормальные волны. Эти волны имеют единственную компоненту смещения u_2, параллельную поверхности пластины и ортогональную направлению распространения. Таким образом, она является горизонтальной поперечной волной сдвига. Общее решение волнового уравнения запишем в следующем виде:

$$u_2 = \left[A\cos(\beta x_3) + B\sin(\beta x_3)\right]\exp\left[j(\omega t - \kappa x_1)\right],$$
$$\beta = \sqrt{\omega^2/\mathrm{v}_T^2 - \kappa^2}. \qquad (2.57)$$

Как и волны Лэмба, поперечные нормальные волны делятся на две группы – симметричные и антисимметричные. Очевидно, симметричным волнам соответствует решение (2.57) с коэффициентом $B = 0$, для антисимметричных волн $A = 0$.

На обеих поверхностях пластины $x_3 = \pm h/2$ единственная компонента нормального напряжения T_{32} должна равняться нулю. Из этого

условия получаем выражения для смещения в каждом из видов волн. В симметричной волне:

$$u_2^{(s)} = A\cos\left(\frac{m_s\pi}{h}x_3\right)\exp\left[j\left(\omega t - \kappa_s x_1\right)\right],$$
$$\kappa_s = \sqrt{\left(\frac{\omega}{v_T}\right)^2 - \left(\frac{m_s\pi}{h}\right)^2}, \quad m_s = 0,\ 1,\ 2... \tag{2.58}$$

В антисимметричной волне:

$$u_2^{(a)} = A\sin\left(\frac{m_a\pi}{h}x_3\right)\exp\left[j\left(\omega t - \kappa_a x_1\right)\right],$$
$$\kappa_a = \sqrt{\left(\frac{\omega}{v_T}\right)^2 - \left(\frac{m_a\pi}{h}\right)^2}, \quad m_a = \frac{1}{2},\ \frac{3}{2},\ \frac{5}{2}... \tag{2.59}$$

Из выражений (2.58), (2.59) нетрудно установить, что при заданных значениях ω и h в пластине может распространяться только определенное число волн, для которых постоянная распространения является вещественной. Данное число тем больше, чем больше отношение $2h/\Lambda_T = \omega h/\pi v_T$. Симметричные волны порядков выше нулевого $m_s > 0$ и все антисимметричные волны появляются только при некоторых критических значения $\omega h/v_T$:

$$\left(\omega h/v_T\right)_{кр} = m_{s,a}\pi .$$

Фазовые и групповые скорости поперечных нормальных волн общем случае зависят от частоты:

$$v_{s,a} = \frac{v_T}{\sqrt{1-\left(m_s\pi v_T/\omega\right)^2}}, \quad v_{s,a}^{гр} = v_T\sqrt{1-\left(m_s\pi v_T/\omega\right)^2}.$$

Если $2h/\Lambda_T < 1/2$, то распространяется только одна нулевая симметричная мода. Для нее $m_s = 0$, амплитуда смещения одинакова во

всех точках, фазовая и групповая скорости не зависят от частоты и равны скорости поперечных объемных волн v_T. Иначе говоря, данная мода является объемной SH-волной, о которой уже говорилось в разделе 2.6.

2.10. Волны на поверхности непьезоактивных кристаллов

Для изотропных сред свойства поверхностных волн не зависят от выбора той или иной плоскости в качестве свободной поверхности или от направления распространения волны на этой плоскости. Совсем иная ситуация имеет место в случае различного рода кристаллов, где, как мы уже видели в разделе 2.3 на примере объемных волн, все свойства – фазовая скорость, групповая скорость и тип смещения зависят от направления распространения.

Пусть кристалл занимает полупространство $x_3 < 0$ и граничит с вакуумом (рис. 2.11).

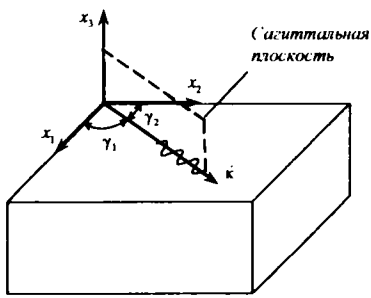


Рис. 2.11. Система координат, используемая для анализа упругих волн на поверхности кристалла

Будем искать решение уравнения движения кристалла

$$\rho \frac{\partial^2 u_i}{\partial t^2} - C_{iklm} \frac{\partial^2 u_m}{\partial x_k \partial x_l} = 0 \qquad (2.60)$$

в виде плоской неоднородной волны, распространяющейся вдоль поверхности $x_3 = 0$:

$$u_i = A_i \exp(\beta x_3) \exp\left[j(\omega t - \kappa \alpha_1 x_1 - \kappa \alpha_2 x_2) \right], \qquad (2.61)$$

где $\alpha_1 = \cos\gamma_1$ и $\alpha_2 = \cos\gamma_2$ – направляющие косинусы волнового вектора на плоскости $x_3 = 0$. В изотропной среде свойства волн не зависят от ориентации волнового вектора и оси координат выбираются произвольно. В кристалле необходимо учитывать направление распространения. Поэтому обычно положение осей x_1, x_2 выбирается по возможности совпадающим с направлениями высокой симметрии кристалла, а ориентация волнового вектора определяется по отношению к этим осям.

Выражение (2.61) описывает волну, связанную с поверхностью, если $\mathrm{Re}(\beta) > 0$. Для изотропной среды величина β была вещественной, а для кристаллов, как будет видно в дальнейшем, в общем случае ее следует считать комплексной. Введем новую переменную $\alpha_3 = j\beta$. Тогда решение приобретает следующий вид:

$$u_i = A_i \exp\left[j(\omega t - \kappa \alpha_i x_i) \right]. \qquad (2.62)$$

По форме оно соответствует волне, распространяющейся в неограниченной анизотропной среде, где все три величины α_i являются направляющими косинусами волнового вектора, а выражение (2.62) описывает плоскую однородную волну, не исчезающую во всем пространстве. Для неоднородных волн, связанных с поверхностью, величина α_3 комплексная и имеет положительную мнимую часть. Подстановка решения (2.62) в (2.60) приводит к системе трех однородных линейных уравнений для неизвестных коэффициентов A_i:

$$\left(\rho\omega^2 \delta_{im} - \kappa^2 C_{iklm} \alpha_k \alpha_l \right) A_i = 0. \qquad (2.63)$$

Приравнивая нулю ее главный определитель, получаем дисперсионное уравнение:

$$\det\left[\rho\omega^2\delta_{im} - \kappa^2 C_{iklm}\alpha_k\alpha_l\right] = 0\,. \qquad (2.64)$$

В разделе 2.3 для случая объемных волн при заданном направлении волнового вектора (то есть при заданных значениях направляющих косинусов α_i) соотношение (2.64) рассматривалось как кубическое уравнение относительно квадрата частоты и имело три вещественных корня, которые соответствовали трем независимым модам взаимно ортогональной поляризации. В случае же поверхностных волн направление волнового вектора определяют только косинусы α_1 и α_2, а α_3 – неизвестный комплексный параметр. Поэтому с точки зрения удобства анализа поверхностных волн дисперсионное соотношение (2.64) следует трактовать как уравнение шестой степени относительно α_3 с действительными значениями частоты и волнового числа, которые присутствуют в уравнении в качестве некоторых параметров.

Поскольку все коэффициенты при степенях α_3 в дисперсионном уравнении – действительные величины, в общем случае оно имеет три пары комплексно сопряженных корней. Три корня удовлетворяют условию $\mathrm{Im}(\alpha_3) > 0$ и описывают волны, локализованные у поверхности. Три другие корня с отрицательной мнимой частью $\mathrm{Im}(\alpha_3) < 0$ связаны с волнами, неограниченно возрастающими при $x_3 \to \infty$, и для полуограниченной среды должны быть отброшены. В частном случае какой-либо из корней может оказаться чисто вещественным. Этот корень будет описывать волну, двигающуюся вдоль поверхности, но не исчезающую во всем объеме среды.

Обозначим корни дисперсионного уравнения с положительной мнимой частью как $\alpha_3^{(1)}, \alpha_3^{(2)}$ и $\alpha_3^{(3)}$. Зная эти корни, для каждого из них с помощью системы уравнений (2.63) можно выразить две из трех постоянных $A_i^{(n)}$ через третью, например $A_1^{(n)}$. Таким образом находим частные решения уравнения движения, имеющие одинаковые значения частоты и волнового числа. Общее решение будет линейной комбинацией частных решений:

$$u_i = \sum_{n=1}^{3} C_r A_i^{(n)} \exp\left[j\left(\omega t - \kappa\alpha_1 x_1 - \kappa\alpha_2 x_2 - \kappa\alpha_3^{(n)} x_3\right)\right], \qquad (2.65)$$

где C_n – неизвестные коэффициенты. Они могут быть найдены из граничных условий на плоскости $x_3 = 0$, которые заключаются в требовании равенства нулю компонент тензора напряжений T_{i3}:

$$T_{i3} = C_{i3lm} \frac{\partial u_m}{\partial x_l} = 0 \text{ при } x_3 = 0. \qquad (2.66)$$

Подставляя решение (2.65) в (2.66), получаем систему трех линейных уравнений для неизвестных C_n. Главный определитель этой системы – определитель граничных условий состоит из членов вида

$$d_{in} = C_{i3lm} A_l^{(n)} \alpha_m^{(n)}, \text{ где } \alpha_1^{(n)} = \alpha_1, \ \alpha_2^{(n)} = \alpha_2. \qquad (2.67)$$

Ненулевое решение для поверхностной волны (2.65) будет существовать, если найденные из уравнений (2.63), (2.64) величины $A_i^{(n)}$ и $\alpha_3^{(n)}$ обращают определитель граничных условий в ноль. В отличие от рэлеевской волны для изотропной среды волна на поверхности кристалла состоит, вообще говоря, не из двух, а из трех связанных между собой через граничные условия парциальных волн. В ней присутствуют все три компоненты смещения. Таким образом, задача поиска характеристик поверхностной волны, распространяющейся в некотором заданном направлении, сводится к подбору такого соотношения между частотой и волновым числом (то есть к подбору фазовой скорости), при котором определитель граничных условий равен нулю.

Приведенная выше схема расчета характеристик упругой волны на поверхности кристалла показывает, что данная задача (за исключением простейших случаев) может быть решена только численными методами. К настоящему времени сделаны расчеты для многих кристаллов при различных ориентациях поверхности и направлениях распространения. Далее для иллюстрации некоторых новых свойств поверхностных волн, связанных с анизотропией, остановимся лишь на случае простейшей ориентации кубического кристалла.

Пусть координатные оси направлены вдоль ребер элементарной ячейки кубического кристалла, а распространение волны происходит параллельно плоскости $x_3 = 0$ (рис. 2.12).

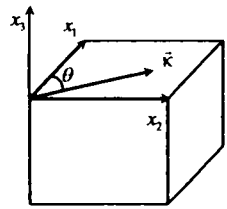


Рис. 2.12. Упругая волна на поверхности кубического кристалла

Поведение волн зависит от направления распространения и величины фактора анизотропии кубического кристалла $\eta = 2\mathbf{C}_{44}/(\mathbf{C}_{11} - \mathbf{C}_{12})$.

Прежде всего, следует сказать, что при $\eta \approx 1$ свойства поверхностных волн аналогичны случаю изотропной среды, их фазовая скорость слабо зависит от направления и может быть приближенно найдена по формулам, определяющим скорость рэлеевской волны в изотропной среде.

Для кристаллов с $\eta > 1$, в отличие от изотропных материалов, существует две поверхностные волны. Волна 1 при $\theta = 0°$ имеет смещения только в сагиттальной плоскости, однако распределение смещений по глубине имеет вид затухающей осциллирующей функции (рис. 2.13). В связи с этим ее иногда называют обобщенной рэлеевской волной.

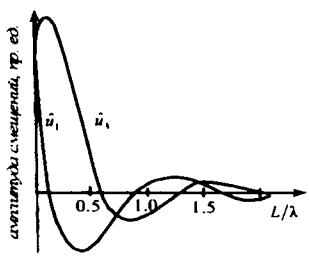


Рис. 2.13. Распределение смещений по глубине кристалла в обобщенной волне Рэлея $(\eta > 1)$

При увеличении угла θ от 0° до 45° в волне 1 появляется компонента смещения, ортогональная сагиттальной плоскости. При этом глубина проникновения возрастает, а в направлении, близком к диагональному $(\theta = 45°)$, волна плавно переходит в чисто поперечную объемную волну горизонтальной поляризации. Скорость волны 1 зависит от направления распространения.

Волна 2 строго существует только в случае $\theta = 45°$. При отклонении от этого угла она уже не удовлетворяет граничным условиям ни при каких действительных волновых числах. Однако если предположить, что волновое число имеет небольшую мнимую часть, то граничные условия удовлетворяются. Это значит, что при распространении вдоль поверхности волна 2 затухает. Причина затухания связана с тем, что фазовая скорость волны v превосходит фазовую скорость объемной поперечной волны горизонтальной поляризации v_T^{zop}. Смещения в волне 2 оказываются в фазе с объемной волной, распространяющейся под углом к поверхности и уносящей энергию

от поверхности вглубь среды. Такого вида волны получили название вытекающих волн. На свободной поверхности изотропной среды вытекающие волны невозможны, так как скорость поверхностной волны всегда меньше любой из скоростей объемных волн. Но в кристалле вследствие анизотропии упругих свойств скорости поперечных волн могут сильно различаться и соотношение $v > v_T^{zop}$ можно реализовать. Излучение энергии возрастает при увеличении отклонения от диагонального, а при $\theta \cong 20°$ эта волна переходит в объемную.

Для кубических кристаллов с фактором анизотропии $\eta < 1$ поверхностная волна ни при каких направлениях распространения не вырождается в объемную. Несмотря на анизотропию, ее фазовая скорость практически не зависит от направления. Распределение смещений по глубине среды не содержит осциллирующей части и аналогично соответствующему распределению в рэлеевских волнах изотропной среды. В случае углов распространения $\theta = 0°$ и $\theta = 45°$ волна имеет смещения только в сагиттальной плоскости, а при других углах присутствует также компонента, параллельная поверхности и ортогональная волновому вектору.

2.11. Волны на поверхности пьезоэлектриков

Как было показано в разделе 2.4, в пьезоактивной среде упругие волны сопровождаются квазистационарным электрическим полем, которое может быть описано скалярным электростатическим потенциалом φ. Распределение потенциала и компонент смещения, создаваемых упругой волной в пьезоэлектрике, описывает связанная система уравнений (2.23) и (2.24):

$$\begin{aligned} e_{ilm}\frac{\partial^2 u_m}{\partial x_i \partial x_l} &= \varepsilon_{im}\frac{\partial^2 \varphi}{\partial x_i \partial x_m}, \\ \rho\frac{\partial^2 u_i}{\partial t^2} - C_{iklm}\frac{\partial^2 u_m}{\partial x_k \partial x_l} &= e_{mik}\frac{\partial^2 \varphi}{\partial x_k \partial x_m}. \end{aligned} \tag{2.68}$$

В безграничной среде однородные плоские волны имеют по преимуществу механическую природу, а пьезоэффект проявляется только в виде небольшой поправки к эффективным упругим постоянным. Исследуем теперь влияние пьезоэлектрического эффекта на распространение поверхностных волн.

Пусть, как и ранее, среда занимает полупространство $x_3 < 0$ (рис. 2.11). Рассмотрим поверхностную волну:

$$\begin{aligned} u_i &= A_i \exp\left[j(\omega t - \kappa\alpha_k x_k)\right], \\ \varphi &= A_4 \exp\left[j(\omega t - \kappa\alpha_k x_k)\right], \end{aligned} \tag{2.69}$$

где α_1, α_2 – направляющие косинусы волнового вектора в плоскости $x_3 = 0$, α_3 – неизвестная комплексная величина, которая для поверхностной волны удовлетворяет условию $\mathrm{Im}(\alpha_3) > 0$, A_4 – постоянная, определяющая амплитуду потенциала на поверхности. Подставляя (2.69) в (2.68), получаем однородную систему четырех линейных уравнений относительно неизвестных постоянных A_r $(r = 1,2,3,4)$:

$$\left(\rho\omega^2\delta'_{rs} - \kappa^2\Gamma_{rs}\right)A_r = 0, \quad (r,s = 1,2,3,4), \tag{2.70}$$

где

$$\delta'_{rs} = \begin{cases} 1, & r = s \neq 4 \\ 0, & r \neq s, \quad r = s = 4 \end{cases},$$

$$\Gamma_{im} = C_{iklm}\alpha_k\alpha_l,$$

$$\Gamma_{i4} = \Gamma_{4i} = e_{kli}\alpha_k\alpha_l,$$

$$\Gamma_{44} = -\varepsilon_{kl}\alpha_k\alpha_l.$$

$$(i,k,l,m = 1,2,3).$$

Для заданного направления распространения, приравнивая главный определитель системы уравнений (2.70) к нулю, получаем алгебраическое уравнение восьмой степени с действительными коэффициентами относительно неизвестной величины α_3:

$$\det\left[\rho\omega^2\delta'_{rs} - \kappa^2\Gamma_{rs}\right] = 0. \tag{2.71}$$

Данное дисперсионное уравнение в общем случае имеет четыре пары комплексно сопряженных корней, среди которых следует выбрать корни $\alpha_3^{(n)}$ $(n = 1,2,3,4)$ с положительной мнимой частью. Зная корни $\alpha_3^{(n)}$, из системы уравнений (2.70) можно найти частные решения вида (2.69). Общее решение определится как линейная комбинация частных решений:

$$\begin{aligned} u_i &= \sum_{n=1}^{4} C_n A_i^{(n)} \exp\left[j\left(\omega t - \kappa\alpha_1 x_1 - \kappa\alpha_2 x_2 - \kappa\alpha_3^{(n)} x_3\right)\right], \\ \varphi &= \sum_{n=1}^{4} C_n A_4^{(n)} \exp\left[j\left(\omega t - \kappa\alpha_1 x_1 - \kappa\alpha_2 x_2 - \kappa\alpha_3^{(n)} x_3\right)\right] \end{aligned} \tag{2.72}$$

Упругая волна на поверхности пьезоэлектрика в общем случае состоит из четырех парциальных волн. Неизвестные коэффициенты C_n, описывающие вклад каждой из парциальных волн, должны быть найдены из граничных условий на поверхности $x_3 = 0$.

Граничными условиями для механических величин является отсутствие компонент напряжения T_{3i} на поверхности. С учетом пьезоэффекта эти условия запишутся как

$$T_{3i} = C_{3ikl}\frac{\partial u_k}{\partial x_l} + e_{k3i}\frac{\partial \varphi}{\partial x_k} = 0 \quad \text{при } x_3 = 0\,. \tag{2.73}$$

Ввиду наличия в волне электрического поля, помимо механических граничных условий, необходимо также использовать электрические граничные условия. Обычно рассматривают два случая. В первом из них пространство над поверхностью среды – вакуум, а проводники, и, следовательно, свободные заряды на границе отсутствуют. Такую поверхность называют электрически свободной. Во втором случае предполагается, что поверхность покрыта тонким слоем металла, который обеспечивает короткое замыкание горизонтальной составляющей электрического поля, но не изменяет механических граничных условий. Данную поверхность называют металлизированной или электрически закороченной. Заметим, что нормальная к электрически закороченной поверхности компонента электрического поля может быть отличной от нуля. То есть закоротка поверхности не устраняет полного электрического поля в кристалле. Два описанных случая, вообще говоря, приводят к разным скоростям поверхностных волн. Различие между скоростями отражает степень связи волны с электрическим возмущением поверхности.

В случае электрически свободной поверхности потенциал в вакууме, который обозначим как φ', удовлетворяет уравнению Лапласа:

$$\nabla^2\varphi' = 0\,.$$

На границе среды и вакуума потенциал должен быть непрерывен. Для того чтобы это условие выполнялось одновременно для всех точек поверхности, выражение для потенциала в вакууме должно иметь следующий вид:

$$\varphi' = \sum_{n=1}^{4} C_n A_4^{(n)} \exp\left[j\left(\omega t - \kappa\alpha_1 x_1 - \kappa\alpha_2 x_2\right)\right] f\left(x_3\right), \tag{2.74}$$

где $f\left(x_3\right)$ – неизвестная функция, удовлетворяющая условию $f\left(0\right) = 1$. Нетрудно видеть, что на границе $x_3 = 0$ потенциалы, опре-

деляемые выражениями (2.74) и (2.72), совпадают. Подставляя выражение (2.74) в уравнение Лапласа, получаем следующее уравнение для функции $f(x_3)$:

$$\frac{\partial^2 f}{\partial x_3^2} - \kappa^2\left(\alpha_1^2 + \alpha_2^2\right) f = 0 .$$

Сумма квадратов направляющих косинусов равна единице: $\alpha_1^2 + \alpha_2^2 = 1$, поэтому $f(x_3) = \exp(\pm\kappa x_3)$. Решение со знаком плюс в экспоненте следует отбросить, так как оно неограниченно возрастает при удалении от поверхности. Таким образом, распределение потенциала в вакууме принимает вид

$$\varphi' = \sum_{n=1}^{4} C_n A_4^{(n)} \exp(-\kappa x_3)\exp\left[j\left(\omega t - \kappa\alpha_1 x_1 - \kappa\alpha_2 x_2\right)\right]. \qquad (2.75)$$

Запишем нормальные компоненты электрической индукции по обе стороны от поверхности:

$$D_3 = -\varepsilon_0 \frac{\partial\varphi'}{\partial x_3} = \kappa\varepsilon_0\varphi' \quad \text{при } x_3 > 0 ,$$

$$D_3 = e_{3kl}\frac{\partial u_k}{\partial x_l} - \varepsilon_{3k}\frac{\partial\varphi}{\partial x_k} \quad \text{при } x_3 < 0 .$$

Как известно, в отсутствие свободных зарядов нормальная компонента индукции на поверхности непрерывна. Откуда для электрически свободной поверхности получаем следующее граничное условие:

$$e_{3kl}\frac{\partial u_k}{\partial x_l} - \varepsilon_{3k}\frac{\partial\varphi}{\partial x_k} = \kappa\varepsilon_0\varphi \quad \text{при } x_3 = 0 . \qquad (2.76)$$

В случае металлизированной поверхности потенциал на поверхности равен нулю, то есть электрическое граничное условие имеет вид

$$\varphi = 0 \text{ при } x_3 = 0. \qquad (2.77)$$

Подставляя решение (2.72) в граничные условия (2.73) и (2.76) для электрически свободной поверхности, либо в граничные условия (2.73) и (2.77) для закороченной, получаем систему четырех однородных линейных уравнений для неизвестных коэффициентов C_n. Главный определитель этой системы – определитель граничных условий будет состоять из членов

$$d_{in} = C_{3ikl} A_k^{(n)} \alpha_l^{(n)} + e_{k3i} A_4^{(n)} \alpha_k^{(n)},$$
$$d_{4n} = e_{3kl} A_k^{(n)} \alpha_l^{(n)} - \left(\varepsilon_{3k} \alpha_k^{(n)} + j\varepsilon_0\right) A_4^{(n)}$$

для свободной поверхности и

$$d_{in} = C_{3ikl} A_k^{(n)} \alpha_l^{(n)} + e_{k3i} A_4^{(n)} \alpha_k^{(n)},$$
$$d_{4n} = A_4^{(n)}$$

для металлизированной. Так же как и в непьезоактивном кристалле, ненулевое решение для поверхностной волны (2.72) будет существовать, если найденные из уравнений (2.70), (2.71) величины $A_r^{(n)}$, $\alpha_3^{(n)}$ $(r, n = 1,2,3,4)$ обращают определитель граничных условий в ноль.

В случае произвольного направления распространения вопрос о существовании того или иного вида поверхностной волны в данном пьезоэлектрическом кристалле и ее свойствах может быть решен только численными методами. Однако в некоторых специальных случаях, когда направление распространения совпадает с направлением чистой моды (акустической осью кристалла), расчеты существенно упрощаются. Рассмотрим подробнее два таких случая.

Во-первых, если сагиттальная плоскость является плоскостью симметрии кристалла, то поверхностная волна имеет рэлеевскую структуру, а ее компоненты смещения связаны с потенциалом. Такую пьезоактивную волну называют жесткой рэлеевской волной. При этом ортогональная сагиттальной плоскости компонента смещения не связана с остальными тремя величинами -- смещениями в сагиттальной плоскости и потенциалом.

Во-вторых, если сагиттальная плоскость перпендикулярна оси симметрии второго или более высоких четных порядков в кристалле, компоненты смещения в сагиттальной плоскости связаны между собой, но не связаны с поперечной компонентой и потенциалом. При этом рэлеевская волна распространяется так, как будто все пьезоэлектрические константы равны нулю. Другими словами, данная мода является непьезоактивной.

Поперечная компонента смещения и потенциал в некоторых кристаллах оказываются связанными. При этом возникает новый тип поверхностной волны – волна Гуляева–Блюстейна, распространяющаяся только в пьезоэлектриках. В качестве примера рассмотрим кристалл ниобата лития ($LiNbO_3$), часто используемый в устройствах акустоэлектроники. В данном кристалле, как и многих других пьезоэлектриках, выделяют три взаимно ортогональных оси симметрии, которые обозначают X, Y и Z. Ориентацию поверхности и направление распространения обозначают по отношению к этим осям. Такие материальные характеристики кристалла, как тензор модулей упругости, тензор, диэлектрической проницаемости и пьезоэлектрический тензор, также определяют по отношению к осям X, Y и Z. Первый из рассмотренных выше случаев реализуется в ниобате лития при распространении в направлении Z в плоскости среза кристалла Y (YZ-геометрия). В YZ-геометрии могут распространяться жесткие рэлеевские волны. Так как смещения в данной волне связаны с потенциалом, скорость распространения зависит от электрических граничных условий.

В разделе 2.4 было показано, что для каждой объемной волны может быть введен коэффициент электромеханической связи K, определяющий увеличение скорости этой волны в $\left(1+K^2/2\right)$ раз. Для рэлеевской волны в силу ее сложного компонентного состава не существует единственного электромеханического коэффициента. Тем не менее для описания электромеханической связи в поверхностной волне по аналогии с (2.26) вводят эффективный коэффициент K, квадрат которого определяется как удвоенное относительное изменение скорости рэлеевской волны, вызываемое электрическим закорачиванием поверхности:

$$\frac{K^2}{2}=\frac{\mathrm{v}-\mathrm{v}_0}{\mathrm{v}}, \qquad (2.78)$$

где v – скорость на свободной поверхности, а v_0 – на электрически закороченной. Для *YZ*-геометрии $LiNbO_3$ эффективный коэффициент электромеханической связи составляет примерно 0.21. Это одно из наибольших значений для подложек, пригодных для высокочастотных устройств акустоэлектроники, использующих поверхностные волны.

Интересно рассмотреть теперь такую геометрию распространения, при которой потенциал связан только с поперечной компонентой смещения, когда наряду с непьезоактивной рэлеевской волной может распространяться горизонтально поляризованная поперечная волна Гуляева–Блюстейна, в отсутствие пьезоэффекта вырождающаяся в объемную волну сдвига SH. Например, волна Гуляева–Блюстейна может распространяться вдоль оси *X* на *Y* срезе пьезоэлектрического кристалла CdS. Скорость объемной волны сдвига SH, распространяющейся при такой ориентации в неограниченном кристалле, равна $v_T = \sqrt{C_{44}\left(1+K^2\right)/\rho}$, где $K^2 = e_{15}^2/\left(\varepsilon_{11}C_{44}\right)$. Можно показать, что объемная волна в данной геометрии не удовлетворяет граничным условиям, а поверхностная им удовлетворяет.

Пусть ось x_1 совпадает с направлением *X*, а кристалл занимает полупространство $x_3 < 0$, так что *Y* срез расположен в плоскости $x_3 = 0$. Если поверхность электрически закорочена, то распределение единственной компоненты смещения u_2 и потенциала φ по глубине кристалла в волне Гуляева–Блюстейна описывается следующими выражениями:

$$u_2 = A\exp(\kappa b x_3)\exp\left[j(\omega t - \kappa x_1)\right],$$

$$\varphi = A\frac{e_{15}}{\varepsilon_{11}}\left[\exp(\kappa b x_3) - \exp(\kappa x_3)\right]\exp\left[j(\omega t - \kappa x_1)\right],$$

где $b = K^2/\left(1+K^2\right)$. Фазовая скорость v_0 и глубина проникновения волны h_0 на закороченной поверхности равны:

$$v_0 = v_T\sqrt{1-b^2}, \quad h_0 = \frac{\Lambda_T}{2\pi K^2}.$$

Для электрически свободной поверхности:

$$u_2 = A\exp(\kappa b\varepsilon_r x_3)\exp[j(\omega t - \kappa x_1)],$$

$$\varphi = A\frac{e_{15}}{\varepsilon_{11}}[\exp(\kappa b\varepsilon_r x_3) - \varepsilon_r \exp(\kappa x_3)]\exp[j(\omega t - \kappa x_1)],$$

где $\varepsilon_r = \varepsilon_0/(\varepsilon_0 + \varepsilon_{11})$, а фазовая скорость и глубина проникновения:

$$v = v_T\sqrt{1-\varepsilon_r^2 b^2}, \quad h = \frac{\Lambda_T}{2\pi K^2 \varepsilon_r}.$$

Поскольку K^2 – малая величина, волна Гуляева–Блюстейна по сравнению с волной Рэлея относительно глубоко проникает в среду. Для кристалла CdS глубина проникновения на закороченной поверхности составляет примерно $4\Lambda_T$, а на свободной возрастает до $\approx 50\Lambda_T$. Кроме того, уменьшение амплитуды смещения с глубиной в отличие от волны Рэлея описывается простой экспонентой. Поскольку $b^2 << 1$, фазовая скорость волны Гуляева–Блюстейна очень близка к скорости объемной сдвиговой волны SH, особенно на сводной поверхности, так как $\varepsilon_r^2 << 1$. Относительное уменьшение скорости волны, связанное с металлизацией поверхности имеет порядок K^4.

2.12. Распространение акустических волн в пьезоэлектрических полупроводниках

До сих пор мы рассматривали распространение упругих волн только в диэлектрических средах. В данном разделе будут исследованы явления, которые возникают в кристалле, обладающем как пьезоэлектрическими свойствами, так и проводимостью. Распространение пьезоактивной упругой волны в таком кристалле сопровождается переменным электрическим полем, которое будет взаимодействовать со свободными носителями заряда.

Динамические свойств ансамбля свободных носителей заряда в проводящей среде, как известно, характеризует время максвелло-

ской релаксации $\tau_M = \varepsilon/\sigma$ (где σ – проводимость кристалла). При появлении в среде локального электрического поля свободные носители заряда перераспределяются и за время τ_M экранируют это поле. Для анализа взаимодействия свободных носителей с переменными полями удобно ввести частоту релаксации проводимости ω_c, которая определяется как величина, обратная времени максвелловской релаксации $\omega_c = 1/\tau_M$. Если временной период изменения электрического поля значительно больше времени максвелловской релаксации, то есть частота волны много меньше частоты релаксации проводимости $(\omega << \omega_c)$, то свободные носители достаточно быстро перераспределяются в кристалле и компенсируют пьезоэлектрическое поле. В этом случае поведение кристалла будет точно таким же, как в отсутствие пьезоэффекта. Напротив, если частота упругой волны порядка или больше частоты ω_c $(\omega \geq \omega_c)$, то свободные носители не успевают перераспределиться таким образом, чтобы заметно уменьшить пьезоэлектрическое поле, которое в этом случае будет иметь фактически такую же величину, как в диэлектрическом кристалле. При этом проявляется ряд эффектов взаимодействия упругой волны со свободными носителями заряда. Данные эффекты включают в себя затухание и дополнительную дисперсию волны, а также сопутствующие им акустоэлектрические явления. Кроме того, постоянное электрическое поле, вызывающее дрейф носителей заряда в направлении распространения волны, при некоторых условиях может приводить к ее усилению.

Среди полупроводниковых материалов пьезоэлектрическими свойствами обладают многие бинарные соединения группы $A_{III}B_V$ (GaAs, InSb и др.), а также соединения группы $A_{II}B_{VI}$ (CdS, ZnO и др.). Частоты релаксации проводимости этих материалов ω_c лежат в звуковом, радио и СВЧ диапазонах, а эффекты взаимодействия упругих волн со свободными носителями достаточно велики и могут иметь определенный практический интерес.

Стоит сказать, что в непьезоактивных полупроводниках также наблюдается взаимодействие упругих волн со свободными носителями заряда. Это связано с тем, что деформации, создаваемые волной приводят к локальному изменению ширины запрещенной зоны и возмущению равновесного распределения носителей, что также приводит к дисперсии и затуханию волн, а также к акустоэлектрическим

явлениям. Но величина этих эффектов на несколько порядков меньше, чем в пьезоэлектриках.

Рассмотрим взаимодействие плоской объемной упругой волны со свободными носителями полупроводникового образца. Упругая волна, как мы видели ранее, в общем случае содержит поперечные и продольные смещения частиц среды. Эти смещения в пьезоэлектрике приводят к поляризации кристаллической решетки и созданию квазистационарных электрических полей. С точки зрения взаимодействия с носителями заряда с плоскими объемными волнами, однако, существенным оказывается лишь продольное по отношению к волне электрическое поле. Это поясняет рис. 2.14.

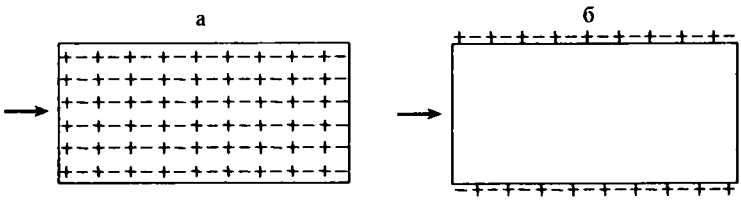


Рис. 2.14. Схематическое изображение распределения поляризационных зарядов, создаваемых плоской объемной волной в пьезоэлектрике: а – для продольной поляризации, б – для поперечной. Стрелками обозначено направление распространения волны

Предполагая существование объемной волны, мы должны потребовать, чтобы ее длина была много меньше размеров образца. При этом поперечные компоненты пьезоэлектрической поляризации вызывают появление зарядов только на гранях образца (рис. 2.14, б). Однако продольная компонента пьезоэлектрической поляризации создает заряды и в объеме образца, которые находятся на расстоянии половины длины упругой волны (рис. 2.14, а). Поэтому создаваемые волной продольные поля будут значительно превосходить поперечные, и действие продольных полей на свободные носители намного сильнее действия любых поперечных полей, обусловленных поляризационными зарядами на границах образца. Следовательно, распространение плоских объемных упругих волн будет сопровождаться значительными электрическими полями и взаимодействие волны со свободными носителями будет достаточно сильным, только когда возникает продольная компонента пьезоэлектрической поляризации.

Далее будем использовать простую модель невырожденного полупроводника с одним типом носителей (для определенности, электронов). Пусть в направлении x распространяется чисто продольная или чисто сдвиговая волна. Создаваемая ею деформация U равна $\partial u/\partial x$, где u – смещение частиц, параллельное оси x в продольной волне или ортогональное этой оси для сдвиговой волны. Выше было показано, что существенными могут быть только продольные пьезоэлектрические поля. Таким образом, по существу мы имеем дело с одномерной задачей. Поэтому соотношения пьезоэффекта (1.51) можно записать в следующем виде:

$$\begin{aligned} D &= \varepsilon E + eU, \\ T &= cU - eE. \end{aligned} \tag{2.78}$$

Здесь c и e – модуль упругости и пьезоэлектрическая константа, соответствующая выбранным направлению и поляризации волны. Другие уравнения, описывающие движение среды и свободных носителей, имеют вид:

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial T}{\partial x} \text{ (уравнение движения среды)}, \tag{2.79}$$

$$\frac{\partial D}{\partial x} = -qn \text{ (уравнение Пуассона)}, \tag{2.80}$$

$$\frac{\partial J}{\partial x} = q \frac{\partial n}{\partial t} \text{ (уравнение непрерывности)}. \tag{2.81}$$

В дальнейшем ограничимся только частотами, существенно меньшими частоты релаксации импульса электронов, что допустимо при $\omega < 10^{11}$ Гц. Это позволяет использовать низкочастотные характеристики полупроводника. Будем также предполагать, что длина волны значительно больше длины свободного пробега электронов, так что плотность тока не зависит от производных электрического поля:

$$J = q\left(n_0 + n\right)\mu E + q\mathcal{D}\frac{\partial n}{\partial x} \text{ (уравнение для полного тока)}. \tag{2.82}$$

В выражениях (2.79) – (2.82) n_0 – равновесная концентрация свободных электронов, n – отклонение от нее, вызванное волной, q, μ и D – соответственно заряд, подвижность и коэффициент диффузии электронов.

Из (2.78) – (2.79) вытекает следующее волновое уравнение:

$$\rho\frac{\partial^2 u}{\partial t^2} = c\frac{\partial^2 u}{\partial x^2} - e\frac{\partial E}{\partial x}. \tag{2.83}$$

В отсутствие пьезоэффекта $(e=0)$ решением (2.83) является плоская упругая волна, распространяющаяся с фазовой скоростью $v_0 = \sqrt{c/\rho}$. Считая, что пространственно-временная зависимость всех переменных соответствует плоской однородной волне:

$$\begin{aligned} E &= E_0 + E_1 \exp\left[j(\omega t - \kappa x)\right], \\ D &= D_0 + D_1 \exp\left[j(\omega t - \kappa x)\right], \end{aligned} \tag{2.84}$$

уравнения (2.80) – (2.82) можно свести к алгебраическим соотношениям. В выражении (2.84) E_0 представляет собой внешнее постоянное электрическое поле, приложенное к кристаллу вдоль оси x для создания в нем дрейфа электронов, D_0 – электрическая индукция, связанная с полем E_0. В линейном приближении в получающихся алгебраических соотношениях можно пренебречь произведением малых амплитуд nE_1. Данное приближение правомерно, когда деформация настолько мала, что $n << n_0$. Из упомянутых алгебраических соотношений и уравнений пьезоэффекта (2.78) можно выразить электрическое поле через производные от смещения. В результате волновое уравнение (2.83) приобретает следующий вид:

$$\rho\frac{\partial^2 u}{\partial t^2} = c'\frac{\partial^2 u}{\partial x^2}, \tag{2.85}$$

где обозначено

$$c' = c\left[1 + \frac{e^2/c\varepsilon}{1 - j\omega_c/\omega\left[1 + (\mu\kappa/\omega)E_0 - j(\kappa^2/\omega)\mathcal{D}\right]^{-1}}\right]. \qquad (2.86)$$

Все эффекты, связанные с влиянием пьезоэлектричества и проводимости кристалла на распространение упругой волны включены в новый эффективный модуль упругости c'.

Решение можно сильно упростить, если считать, что взаимодействие между электронами и упругой волной мало. В этом приближении постоянную распространения можно записать как

$$\kappa = \frac{\omega}{v} - j\alpha, \qquad (2.87)$$

где ее мнимая часть удовлетворяет условию

$$\alpha << \frac{\omega}{v}. \qquad (2.88)$$

Волновое уравнение (2.85) можно решить теперь с помощью подстановки

$$u = u_1 \exp\left[j(\omega t - \kappa x)\right]. \qquad (2.89)$$

Скорость волны v и коэффициент затухания α на единицу длины тогда определяются выражениями:

$$v = \rho^{-1/2}\,\mathrm{Re}\left[(c')^{1/2}\right], \qquad (2.90)$$

$$\alpha = \omega\rho^{1/2}\,\mathrm{Im}\left[(c')^{-1/2}\right]. \qquad (2.91)$$

Слабое взаимодействие означает также, что дисперсия мала, то есть v отличается от v_0 незначительно. Поэтому в выражении (2.86) можно κ заменить на ω/v_0. Подставляя (2.86) в (2.90) и (2.91), получаем:

$$v = v_0\left\{1+\frac{K^2}{2}\frac{1+\left(\omega_c^2/\gamma^2\omega^2\right)\left(\omega^2/\omega_c\omega_D\right)\left[1+\left(\omega^2/\omega_c\omega_D\right)\right]}{1+\left(\omega_c^2/\gamma^2\omega^2\right)\left[1+\left(\omega^2/\omega_c\omega_D\right)\right]^2}\right\}, \qquad (2.92)$$

$$\alpha = \frac{\omega}{v_0}\frac{K^2}{2}\left\{\frac{\omega_c/\gamma\omega}{1+\left(\omega_c/\gamma\omega\right)^2\left[1+\left(\omega^2/\omega_c\omega_D\right)\right]^2}\right\}. \qquad (2.93)$$

В формулах (2.92) и (2.93) использованы следующие обозначения:

$$\omega_c = \frac{qn_0\mu}{\varepsilon} \text{ – частота релаксации проводимости;} \qquad (2.94)$$

$$\omega_D = \frac{v_0^2}{\mathcal{D}} \text{ – частота диффузии электронов;} \qquad (2.95)$$

$$\gamma = 1+\frac{\mu E_0}{v_0} = 1-\frac{v_d}{v_0} \text{ – параметр дрейфа;} \qquad (2.96)$$

$$K = \sqrt{e^2/c\varepsilon} \text{ – коэффициент электромеханической связи.} \qquad (2.97)$$

Так как для типичных полупроводников $K^2 < 10^{-2}$, а абсолютная величина члена в фигурных скобках в формуле (2.93) не превышает 0.5, приближение слабого взаимодействия (2.88) вполне оправдано.

На рисунках 2.15 и 2.16 показана зависимость скорости волны v и коэффициента затухания α от величины $\gamma\omega/\omega_c$. Смысл и значение параметра $\omega^2/\omega_c\omega_D$ определим позднее. Рассмотрим кривую A на рис. 2.15. При $\gamma\omega/\omega_c = 0$ скорость упругой волны минимальна и приблизительно равна v_0, а при $|\gamma\omega/\omega_c| >> 1$ она стремится к несколько большему значению $v_0\left[1+\left(K^2/2\right)\right]$, отвечающему непроводящему кристаллу. Такую зависимость легко объяснить, если перейти в систему отсчета, движущуюся со скоростью дрейфа электронов. В данной системе отсчета $\gamma\omega$ есть эффективная частота волны (с

учетом эффекта Доплера), а $\gamma\omega/\omega_c$ – отношение эффективной частоты к частоте проводимости. Когда $\gamma\omega/\omega_c = 0$, эффективная частота упругой волны равна нулю и электрическое поле для электронов является электростатическим. В результате электроны скапливаются в потенциальных ямах, образуя сгустки. Поле этих сгустков почти полностью компенсирует продольное пьезоэлектрическое поле волны. С ростом отношения $|\gamma\omega/\omega_c|$ условия группировки ухудшаются. При $|\gamma\omega/\omega_c| >> 1$ эффективная частота значительно больше частоты проводимости, и, находясь в быстропеременном поле, электроны практически не группируются, продольное пьезоэлектрическое поле проявляются в полную силу, и материал для волн становится более жестким.

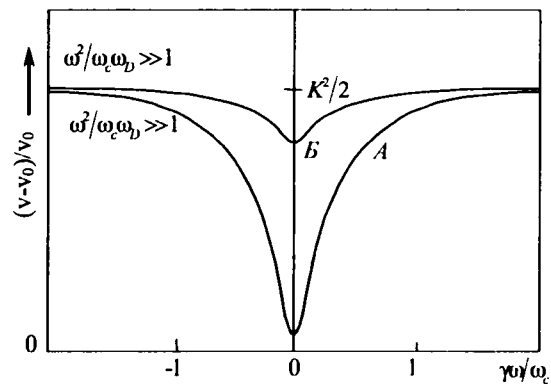


Рис. 2.15. Зависимость скорости упругой волны от параметра дрейфа

Зависимость коэффициента затухания α от $\gamma\omega/\omega_c$ (рис. 2.16) также можно рассмотреть, исходя из эффективной частоты упругой волны $\gamma\omega$. При равенстве нулю эффективной частоты между волной и электронами устанавливается полное равновесие. Поэтому при $\gamma = 0$ средняя сила, действующая на упругую волну со стороны электронов, равна нулю. Если $\gamma > 0$ $(v_d < v_0)$, электронные сгустки формируются в ускоряющем электрическом поле упругой волны, и она затухает, отдавая энергию электронам. Когда $\gamma < 0$ $(v_d > v_0)$, электроны скапливаются в тормозящем поле и затухание меняется

на усиление. При $|\gamma\omega/\omega_c| >> 1$ сгустки практически не формируются, и затухание (усиление) стремится к нулю.

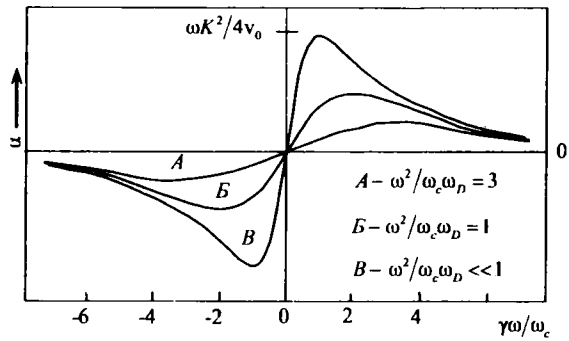


Рис. 2.16. Зависимость коэффициента затухания от параметра дрейфа

В выражениях (2.92) и (2.93) важную роль играет величина $\omega^2/\omega_c\omega_D$. Ее смысл можно раскрыть, если использовать соотношение Эйнштейна

$$\mathcal{D} = \mu kT/q$$

в сочетании с выражениями (2.94), (2.95). Здесь обозначено: T – абсолютная температура, k – постоянная Больцмана. Нетрудно получить

$$\frac{\omega^2}{\omega_c\omega_D} = \kappa^2 L_D^2, \qquad (2.98)$$

где L_D – радиус экранирования Дебая:

$$L_D = \sqrt{\varepsilon kT/q^2 n_0}\,. \qquad (2.99)$$

При размерах электронного сгустка, меньших L_D, границы электронного сгустка размываются тепловым движением электронов.

Кривая B на рис. 2.15 соответствует случаю $\kappa^2 L_D^2 >> 1$, когда дебаевский радиус значительно превосходит длину упругой волны. Поэтому образование сгустков затруднено и скорость волны практически не зависит от γ. В обратном случае $\kappa^2 L_D^2 << 1$ диффузия не препятствует образованию сгустков, но относительное изменение скорости волны тем не менее не превышает $K^2/2$. При слабой электромеханической связи этим изменением также можно пренебречь.

Значительно сильнее диффузия сказывается, как видно из кривых на рис. 2.16, на затухании и усилении акустической волны. Используя (2.98), запишем (2.93) в следующем виде:

$$\alpha = \frac{K^2}{2v_0} \frac{\omega_c/\gamma}{1 + (\omega_c/\gamma)^2 1/\omega^2 \left(1 + \kappa^2 L_D^2\right)^2}. \qquad (2.100)$$

Максимальное значение $|\alpha|$ достигается при факторе дрейфа

$$|\gamma|_{max} = \frac{\omega_c}{\omega}\left(1 + \kappa^2 L_D^2\right) \qquad (2.101)$$

и равно

$$|\alpha|_{max} = \frac{K^2}{4v_0} \frac{\omega}{1 + \kappa^2 L_D^2}. \qquad (2.102)$$

По мере увеличения дебаевского радиуса величина $|\alpha|_{max}$ уменьшается от $\omega K^2/4v_0$ (при $\kappa L_D << 1$) до сколь угодно малых значений, когда дебаевский радиус значительно превосходит длину упругой волны $(\kappa L_D >> 1)$.

Диффузионное размывание проявляется также на частотной зависимости модуля коэффициента затухания:

$$|\alpha| = \frac{K^2}{2} \frac{|\omega_c/\gamma|}{1 + (\omega_c/\gamma)^2 1/\omega^2 \left(1 + \omega^2 L_D^2/v_0^2\right)^2}. \qquad (2.103)$$

Величина $|\alpha|$ как функция ω имеет максимум, когда $\omega = v_0/L_D$. Затухание или усиление волны возрастает с увеличением частоты до тех пор, пока $\omega < v_0/L_D$, то есть $\Lambda > 2\pi L_D$. На более высоких частотах усиление (или затухание) уменьшается.

До сих пор мы рассматривали взаимодействие упругой пьезоактивной волны со свободными носителями заряда в линейном приближении, когда нелинейный член $q\mu n E_1$ в уравнениях отбрасывался. Если амплитуда волны возрастает настолько, что неравновесная концентрация n становится сравнимой с равновесной n_0, этот член уже отбрасывать нельзя. Включение нелинейного члена означает, что чисто гармоническая волна не может распространяться без искажений. В этом случае она должна выражаться в виде рядов Фурье:

$$n = \sum_m n_m \cos(m\omega t - \kappa_m x + \phi_m),$$
$$E = E_0 + \sum_m E_m \cos(m\omega t - \kappa_m x + \theta_m). \qquad (2.104)$$

Аналогичные ряды запишутся и для других величин (u, U и т. д.). Вследствие нелинейности каждая частотная компонента будет взаимодействовать со всеми остальными компонентами. Поэтому при достаточно большой амплитуде волны будут наблюдаться заметная генерация гармоник и параметрическое взаимодействие упругих волн.

Нелинейный член $q\mu n E_1$ в выражении для плотности тока (2.82) можно вычислить достаточно простым способом, если использовать результаты линейной теории для n и E_1. Это будет эквивалентно учету только членов с $m = 1$ в разложениях (2.104). Подстановка значений n и E_1, полученных из линейной теории, в выражение для плотности тока приводит к перемножению двух косинусов. Среднее за период значение произведения косинусов в общем случае отлично от нуля, что означает появление постоянного тока, создаваемого упругой волной, – акустоэлектрического тока. Линейная теория дает следующие выражения для n и E_1:

$$n=\frac{\omega_c e}{q v_0}\frac{u_1}{\gamma-j\left[\left(\omega_c/\omega\right)+\left(\omega/\omega_D\right)\right]},$$
$$E_1=-\frac{e}{\varepsilon}\frac{\gamma-j\left(\omega/\omega_D\right)}{\gamma-j\left[\left(\omega_c/\omega\right)+\left(\omega/\omega_D\right)\right]}u_1. \qquad (2.105)$$

Среднее значение величины $q\mu nE_1$ определится выражением

$$J_{ae}=qn\mu\frac{1}{2}\mathrm{Re}\left(nE_1^*\right).$$

Если подставить сюда (2.105) и учесть выражение для коэффициента поглощения (2.93), то получаем выражение для акустоэлектронного тока в следующем виде:

$$J_{ae}=-\mu c u_1^2\alpha=-\frac{2\alpha\mu W}{v_0}, \qquad (2.106)$$

где W – поток энергии волны. Из этого выражения видно, что при $\alpha>0$ (затухание) J_{ae} представляет собой поток электронов в направлении распространения волны, а при $\alpha<0$ (усиление) ток J_{ae} направлен навстречу волне. Такое простое соотношение между акустоэлектрическим током и коэффициентом затухания представляется естественным. Если упругую волну и совокупность всех свободных электронов рассматривать как две раздельные взаимодействующие системы, то энергия и импульс, теряемые упругой волной, передаются электронам. Поэтому в режиме затухания непрерывные потери импульса волной приводят к появлению постоянного тока в направлении движения волны. Напротив, в режиме усиления волна увеличивает свой импульс за счет электронов, а реакция электронов сводится к образованию потока в направлении, противоположном движению волны.

Контрольные вопросы

1. Какими методами возбуждаются поверхностные и объемные акустические волны?

2. Какие типы поверхностных акустических волн возбуждаются в различных средах?

3. Охарактеризуйте виды объемных акустических волн, распространяющихся в изотропной упругой среде и в кристаллах.

4. Назовите отличительные свойства объемных акустических волн в пьезоэлектриках.

5. Приведите распределение амплитуд смещений и напряжений в рэлеевской волне по глубине как характеристику поверхностного типа волны.

6. При каких условиях реализуются волны Лява в упругой среде?

7. Объясните, при каких параметрах соприкасающихся сред волны Стоунли переходят в волны Рэлея или в обобщенные волны Лява.

8. Проанализируйте характер распространения акустических волн в пластинах.

9. Приведите схему расчёта характеристик акустической волны, распространяющейся на поверхности кристалла.

10. Проанализируйте условия распространения поверхностных волн в пьезоэлектриках для случаев электрически свободной поверхности и поверхности, покрытой металлическим слоем.

11. В чём заключается эффект усиления акустических волн в пьезополупроводниках, и каковы условия его реализации (параметры среды, внешнего поля, акустической волны)?

3. НЕЛИНЕЙНЫЕ ЭФФЕКТЫ И ВЗАИМОДЕЙСТВИЕ ВОЛН

3.1. Нелинейные взаимодействия акустических волн

Взаимодействие объемных акустических волн. Если по твердому телу распространяется акустическая волна, то уравнение, описывающее смещение частиц тела от положений равновесия, имеет вид

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial T_{ik}}{\partial x_k}, \tag{3.1}$$

где ρ – плотность тела; u_i – компонент вектора смещения частиц в волне ($\vec{u}$); T_{ik} – компонент тензора механических напряжений; x_k – пространственная координата.

Смещения частиц вызывают деформацию тела, которая в линейном приближении описывается тензором деформаций:

$$u_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right). \tag{3.2}$$

Если амплитуды смещений частиц малы, то колебания частиц в волне можно считать гармоническими, так как потенциальная энергия атомов твердого тела при небольших отклонениях от положений равновесия пропорциональна квадрату отклонения. В этом случае мала и деформация, и в разложении свободной энергии твердого тела F по компонентам тензора деформаций достаточно учесть члены вида $u_{ik}u_{lm}$:

$$F = F_0 + \frac{1}{2} c_{iklm} u_{ik} u_{lm}, \tag{3.3}$$

где F_0 – свободная энергия тела без учета воздействия акустической волны; c_{iklm} – компоненты тензора линейной упругости.

Тогда тензоры механических напряжений и деформаций связаны линейным законом Гука:

$$T_{ik} = \frac{\partial F}{\partial u_{ik}} = c_{iklm} u_{lm}, \qquad (3.4)$$

а решения уравнения (3.1) могут быть записаны в комплексной форме:

$$\vec{u} = \vec{u}_0 \exp\left[j\left(\omega t - \vec{\kappa}\vec{r}\right)\right],$$

где $\vec{u}_0$ – амплитуда смещения; ω – частота колебаний (волны); $\vec{r}$ – радиус-вектор; $\vec{\kappa}$ – волновой вектор, т. е. вектор, ориентированный в направлении распространения волны и имеющий модуль $|\vec{\kappa}| = \omega/V = 2\pi/\Lambda$ (V – скорость распространения волны; Λ – длина волны). Реальные смещения частиц описываются действительной частью функции $\vec{u}$.

В общем случае в анизотропном твердом теле (кристалле) в данном направлении $\vec{r}$ могут распространяться три объемные акустические волны (ОАВ): квазипродольная L и две квазипоперечные (FS и SS), одна из которых имеет большую скорость по сравнению с другой.

Если по твердому телу распространяются две волны, характеризующиеся тензорами деформаций $u^{(1)}$ и $u^{(2)}$, то механические напряжения, вызванные ими, складываются:

$$T_{ik} = c_{iklm}\left(u_{lm}^{(1)} + u_{lm}^{(2)}\right),$$

т. е. происходит простое наложение волн без их взаимодействия.

Рассмотрим теперь акустические волны такой интенсивности, что амплитуды смещений частиц в волне достаточно велики. Тогда необходимо учесть ангармонический характер колебаний атомов в решетке твердого тела, связанный с непараболичностью потенциальной ямы атома в решетке и сказывающийся в том, что при сближении атомов сила отталкивания растет быстрее, чем сила притяжения при увеличении расстояния между атомами. Это значит, что при

достаточно больших деформациях разложение (3.3) следует дополнить членом, пропорциональным произведению $u_{ij}u_{kl}u_{mn}$:

$$F = F_0 + \frac{1}{2}c_{iklm}u_{ik}u_{lm} + \frac{1}{6}C_{iklmjn}u_{ik}u_{lm}u_{jn}, \tag{3.5}$$

где C_{iklmjn} – компоненты тензора нелинейной упругости.

Нелинейный закон Гука в этом случае имеет вид

$$T_{ik} = c_{iklm}u_{lm} + \frac{1}{2}C_{iklmjn}u_{lm}u_{jn}. \tag{3.6}$$

Решение уравнения (3.1) с учетом соотношения (3.6) уже не будет, строго говоря, представлять собой синусоидальные колебания, и форма волны по мере распространения изменяется. В пределе синусоидальная волна может выродиться в пилообразную. Однако этими эффектами обычно пренебрегают, считая, что по твердому телу попрежнему распространяются три волны: *L*, *FS* и *SS*.

Теперь одновременное распространение двух акустических волн в твердом теле, а следовательно, участие частиц тела одновременно в двух колебательных движениях, приводит не только к сложению колебаний, но и к их перемножению, т. е. происходит нарушение принципа суперпозиции волн:

$$\begin{aligned} T_{ik} &= c_{iklm}\left(u_{lm}^{(1)} + u_{lm}^{(2)}\right) + \frac{1}{2}C_{iklmjn}\left(u_{lm}^{(1)}u_{jn}^{(1)} + u_{lm}^{(2)}u_{jn}^{(2)}\right) + \\ &+ \frac{1}{2}C_{iklmjn}\left(u_{lm}^{(1)}u_{jn}^{(2)} + u_{jn}^{(1)}u_{lm}^{(2)}\right). \end{aligned} \tag{3.7}$$

Второй член в (3.7) описывает эффект генерации второй гармоники распространяющихся акустических волн (эффект самовоздействия волн), а третий обусловливает возможность взаимодействия волн в нелинейной среде.

Для того чтобы колебания частиц твердого тела усиливались и приводили к макроскопическому эффекту взаимодействия двух волн с частотами ω_1 и ω_2 и волновыми векторами $\vec{\kappa}_1$ и $\vec{\kappa}_2$, необходимо, чтобы возбуждения в отдельных точках области взаимодействия возникали в соответствующих фазах, т. е. чтобы имелось согласова-

ние колебаний отдельных частиц во времени и пространстве – синхронизм.

В квантовой акустике колебание частицы тела принято интерпретировать как существование в некоторой точке тела квазичастицы – фонона, а взаимодействие акустических волн как фонон-фононное взаимодействие. Как всегда при взаимодействии частиц, должны выполняться законы сохранения их энергий и импульсов. Следовательно, если при фонон-фононном взаимодействии возникает новая частица, т. е. генерируется волна с частотой ω_3 и волновым вектором $\vec{\kappa}_3$, то должны быть выполнены условия

$$\hbar\omega_1 \pm \hbar\omega_2 = \hbar\omega_3,$$
$$\hbar\vec{\kappa}_1 \pm \hbar\vec{\kappa}_2 = \hbar\vec{\kappa}_3,$$

где $\hbar$ – постоянная Планка.
Отсюда

$$\left.\begin{aligned}\omega_1 \pm \omega_2 &= \omega_3\\ \vec{\kappa}_1 \pm \vec{\kappa}_2 &= \vec{\kappa}_3\end{aligned}\right\}. \qquad (3.8)$$

Эти соотношения называются условиями синхронизма взаимодействующих волн. Из них следует, что при взаимодействии двух волн порождаются колебания комбинационных (суммарной и разностной) частот с соответствующими волновыми векторами.

Для анализа взаимодействия акустических волн удобно использовать метод дисперсионных диаграмм. Для этого в трехмерной системе координат (ω, κ_x, κ_y) – для неколлинеарного взаимодействия волн или в двумерной системе (ω, κ) – для коллинеарного взаимодействия строятся дисперсионные характеристики волн, т. е. зависимости $\omega(\kappa)$. При коллинеарном взаимодействии это некоторые кривые (или при отсутствии дисперсии – прямые), а в случае неколлинеарного взаимодействия – некоторые конические поверхности.

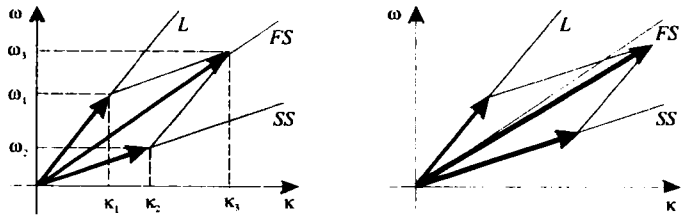


Рис. 3.1. Дисперсионные диаграммы синхронного (*а*) и несинхронного (*б*) коллинеарных взаимодействий акустических волн $L(\omega_1;\vec{\kappa}_1)$ и $SS(\omega_2;\vec{\kappa}_2)$

Пусть, например, исследуется попутное взаимодействие двух бездисперсионных волн – продольной с частотой ω_1 и волновым вектором $\vec{\kappa}_1$ – $L(\omega_1;\vec{\kappa}_1)$ и поперечной $SS(\omega_2;\vec{\kappa}_2)$, причем $\vec{\kappa}_1 \parallel \vec{\kappa}_2$, т.е. взаимодействие коллинеарное. На дисперсионной диаграмме этим волнам будут сопоставлены векторы $\{\omega_1;\vec{\kappa}_1\}$ и $\{\omega_2;\vec{\kappa}_2\}$, концы которых лежат на соответствующих дисперсионных прямых (рис. 3.1). Суммой указанных векторов явится вектор $\{\omega_3;\kappa_3\}=\{\omega_1+\omega_2;\ \kappa_1+\kappa_2\}$. Если его конец попадет на дисперсионную прямую для волны *FS* (рис. 3.1, *а*), то произойдет синхронное взаимодействие, порождающее акустическую волну:

$$L(\omega_1)+SS(\omega_2)=FS(\omega_1+\omega_2).$$

Если конец полученного при сложении вектора не попадет на дисперсионную прямую (рис. 3.1, *б*), то взаимодействие будет несинхронным и произойдет генерация не синусоидальной волны, а волны с пространственными или временными биениями.

Условия синхронизма накладывают ограничения и на вид взаимодействующих волн. Так, в изотропной среде (когда скорости двух поперечных волн не различаются) возможны лишь следующие синхронные взаимодействия:
коллинеарные

$$L(\omega_1)+L(\omega_2)=L(\omega_1\pm\omega_2),$$
$$S(\omega_1)+S(\omega_2)=S(\omega_1\pm\omega_2),$$

неколлинеарные

$$L(\omega_1)+L(\omega_2)=S(\omega_1-\omega_2),$$
$$S(\omega_1)+S(\omega_2)=L(\omega_1+\omega_2),$$
$$L(\omega_1)+S(\omega_2)=L(\omega_1\pm\omega_2),$$
$$L(\omega_1)+S(\omega_2)=S(\omega_1-\omega_2).$$

В анизотропных средах возможен 21 тип различных синхронных взаимодействий акустических волн.

Взаимодействие возможно и лишь при вполне определенных соотношениях частот перемножаемых колебаний. Например, для процесса, проиллюстрированного на дисперсионной диаграмме рис. 3.1, *а*,

$$\kappa_3=\kappa_1+\kappa_2,$$

здесь κ_1, κ_2, κ_3 – проекции соответствующих волновых векторов на направление распространения волн. Тогда

$$\frac{\omega_3}{\upsilon_{FS}}=\frac{\omega_1+\omega_2}{\upsilon_{FS}}=\frac{\omega_1}{\upsilon_L}+\frac{\omega_2}{\upsilon_{SS}},$$

где υ_{FS}, υ_L и υ_{SS} – скорости соответствующих волн в направлении $\vec{r}$. Отсюда соотношение частот должно удовлетворять условию

$$\frac{\omega_2}{\omega_1}=\frac{1-\upsilon_{FS}/\upsilon_L}{\upsilon_{FS}/\upsilon_{SS}-1}.$$

В случае неколлинеарного взаимодействия уравнения для определения ω_2/ω_1 более сложны. Пусть, например, при взаимодействии двух продольных волн $L(\omega_1;\vec{\kappa}_1)$ и $L(\omega_2;\vec{\kappa}_2)$, распространяющихся под углом β друг к другу, образуется поперечная волна $S(\omega_3;\vec{\kappa}_3)$ на разностной частоте $\omega_3=\omega_1-\omega_2$. Тогда, как видно из рис. 3.2,

$$\left(\frac{\omega_1-\omega_2}{\upsilon_S}\right)^2=\left(\frac{\omega_1}{\upsilon_L}\right)^2+\left(\frac{\omega_2}{\upsilon_L}\right)^2-\frac{2\omega_1\omega_2\cos\beta}{\upsilon_L^2}.$$

Откуда можно получить выражение для ω_2/ω_1 или при заданном соотношении частот ω_2 и ω_1 вычислить угол, при котором такой тип взаимодействий возможен.

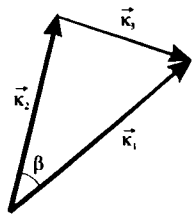


Рис. 3.2. Диаграмма неколлинеарного взаимодействия с образованием волны на разностной частоте

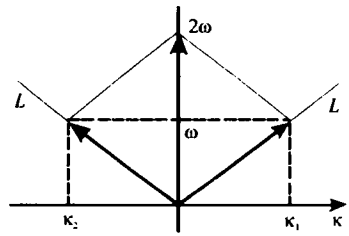


Рис. 3.3. Дисперсионная диаграмма вырожденного взаимодействия акустических волн и вырожденного электроакустического взаимодействия

Простейший макет для наблюдения нелинейных взаимодействий ОАВ состоит из звукопровода, входных преобразователей и выходного электрода. Пьезоэлектрические преобразователи, прикрепленные к торцам звукопровода, выполняемого обычно также из пьезоэлектрического материала, возбуждают в нем однотипные волны одинаковой частоты. Волны распространяются навстречу друг другу и, следовательно, $\vec{\kappa}_1=-\vec{\kappa}_2$. Как видно из дисперсионной диаграммы (рис. 3.3), при таком взаимодействии, называемом вырожденным, так как $\omega_1=\omega_2=\omega$, образуется однородное в пространстве и имеющее частоту $\omega_3=2\omega$ поле механических напряжений. Поскольку $\vec{\kappa}_3=0$, волна не распространяется, и взаимодействие поэтому является несинхронным, однако поле механических напряжений за счет линейного пьезоэффекта вызывает появление однородного электрического поля той же частоты 2ω. Электрическое напряжение снимается со сплошного электрода, называемого обычно параметрическим.

В твердом теле наряду с взаимодействием акустических волн друг с другом могут происходить и взаимодействия акустических волн с полями иной (немеханической) природы. Наиболее важным является электроакустическое взаимодействие, т. е. взаимодействие акустической волны с переменным электрическим полем. Условия синхронизма (3.8) при таких взаимодействиях остаются неизменными, а для анализа взаимодействий, для выяснения вопроса о том, будет ли в результате него генерироваться бегущая акустическая волна, можно пользоваться обычными дисперсионными диаграммами. Например, диаграмма на рис. 3.3 иллюстрирует вырожденное электроакустическое взаимодействие. Если входной преобразователь возбудит в звукопроводе акустическую волну с частотой ω и волновым вектором $\vec{\kappa}_1$, а на параметрический электрод в момент прохождения под ним волны будет подано однородное в пространстве ($\vec{\kappa}_3 = 0$) электрическое поле на частоте 2ω, то в результате электроакустического взаимодействия возникнет «обращенная» волна, т. е. акустическая волна того же типа, что и «прямая», с той же частотой ω, но с волновым вектором $\vec{\kappa}_2 = -\vec{\kappa}_1$.

Если преобразователи возбуждают в звукопроводе волны различных поляризаций (L и S), то можно осуществить следующие невырожденные электроакустические взаимодействия:

$$S(\omega_1;\vec{\kappa}) + E(\omega_2 - \omega_1;0) = L(\omega_2;\vec{\kappa});$$
$$L(\omega_1;\vec{\kappa}) + E(\omega_1 - \omega_2;0) = S(\omega_2;\vec{\kappa});$$
$$S(\omega_1;-\vec{\kappa}) + E(\omega_1 + \omega_2;0) = L(\omega_2;\vec{\kappa});$$
$$L(\omega_1;-\vec{\kappa}) + E(\omega_1 + \omega_2;0) = S(\omega_2;-\vec{\kappa}).$$

Электроакустическое взаимодействие играет важную роль и при образовании фононного (или электроакустического) эха в пьезоэлектриках.

Различают эффекты двух- и трехимпульсного эха. Для наблюдения двухимпульсного фононного эха на ОАВ кристалл помещают в двухрезонаторную систему, причем один резонатор настроен на частоту ω, а другой – на частоту 2ω. В момент времени t=0 электрический компонент электромагнитного импульса, поданного на первый резонатор, возбуждает в кристалле за счет линейного пьезоэффекта упругие волны частоты ω. При распространении этих волн по кри-

сталлу происходит их рассогласование по фазе из-за различия скоростей волн различной поляризации, небольших различий в направлениях волновых векторов и упругого рассеяния волн на дефектах кристалла. Если в момент времени $t = \tau$ подать на второй резонатор электромагнитный импульс с частотой 2ω, то при условии, что акустические волны будут находиться в той части кристалла, которая помещена во второй резонатор, произойдет электроакустическое взаимодействие, порождающее «обращенные» волны с теми же частотами и фазовыми соотношениями, что и «прямые» волны, но с противоположным направлением волнового вектора.

Поскольку упругие волны рассогласовывались по фазе в течение времени τ, то при распространении в обратном направлении они снова окажутся в фазе через такой же интервал времени τ. В результате в момент $t = 2\tau$ сфазированные ОАВ создадут за счет обратного пьезоэффекта принимаемый первым резонатором электрический сигнал частоты ω, который и представляет собой сигнал двухимпульсного фононного эха (рис. 3.4, *а*).

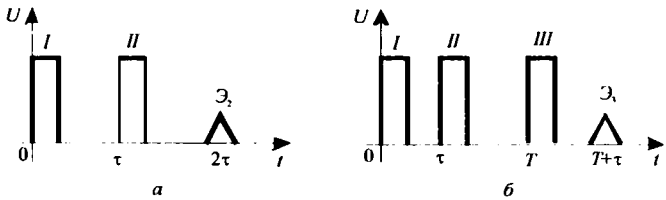


Рис. 3.4. Последовательность импульсов при наблюдении двух- (*а*) и трехимпульсного (*б*) фононного эха: *I*, *II*, *III* – зондирующие импульсы; $Э_2$ – сигнал двухимпульсного эха; $Э_3$ – сигнал трехимпульсного эха

Амплитуда сигнала двухимпульсного эха зависит от времени задержки между импульсами τ по закону

$$A_{Э_2} \sim \exp\left(-\frac{2\tau}{T_2}\right),$$

где T_2 – время релаксации, обусловленное процессами затухания акустических волн в кристалле.

Для наблюдения эффекта трехимпульсного фононного эха кристалл обычно помещают между пластинами конденсатора, на

которые подают последовательно три электрических импульса с частотой ω в моменты времени 0, τ и T ($T > \tau$). Тогда при $t = T + \tau$ на обкладках конденсатора появляется сигнал эха (рис. 3.4, *б*).

Можно выделить следующие нелинейные взаимодействия, приводящие к появлению трехимпульсного фононного эха. Акустические волны $B_{\text{П}}$, генерируемые первым импульсом и распространяющиеся по кристаллу (при этом, как и при образовании двухимпульсного эха, происходит их рассогласование по фазе), взаимодействуют с электрическим полем второго импульса E_{II}. Возникает не зависящее от времени (так как разностная частота равна нулю) неоднородное в пространстве поле упругих механических деформаций (рис. 3.5, *а*) и соответствующее ему за счет линейного пьезоэффекта электрическое поле. Под действием этих полей в кристалле происходит неоднородное перераспределение свободных зарядов (или, возможно, дефектов в неидеальных кристаллах), которое может сохраниться в течение длительного времени после затухания создавших поля колебаний и поддерживать стационарную неоднородную деформацию. Происходит так называемая запись акустической голограммы Γ, а сам эффект носит название «акустоэлектронной памяти».

Если третий импульс E_{III} подан на кристалл через время T, меньшее, чем время разрушения голограммы, то произойдет ее считывание – образование «обращенных» акустических волн B_{O} с частотой ω в результате взаимодействия поля деформаций с электрическим полем третьего импульса (рис. 3.5, *б*). Волны сфазируются через время τ после подачи третьего импульса, что и вызовет появление в момент $t = T + \tau$ сигнала трехимпульсного эха с амплитудой:

$$A_{\text{Э}} \sim \exp\left(-\frac{2\tau}{T_2} - \frac{T-\tau}{T_1}\right),$$

где T_1 – время сохранения акустической голограммы.

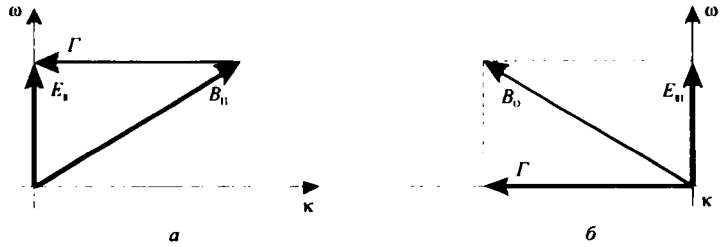


Рис. 3.5. Дисперсионные диаграммы процессов записи (*а*) и считывания (*б*) акустической программы: $B_{П}$, $B_{О}$ – прямая и отраженные волны; E_{II} – электрическое поле второго импульса; E_{III} – электрическое поле третьего импульса; $Г$ – акустическая голограмма

Возможна и иная последовательность нелинейных взаимодействий, приводящих к образованию трехимпульсного фононного эха, например запись голограммы при взаимодействии двух акустических волн и считывание ее бегущей акустической волной и др.

3.2. Взаимодействие акустических волн с высокочастотными электрическими полями

Задачу о нелинейном взаимодействии волн в кристаллах можно распространить и на взаимодействие с электромагнитными волнами. Так как $\upsilon/c << 1$, то это означает, что можно рассматривать взаимодействие с квазистатическими переменными электрическими полями. Геометрия таких взаимодействий будет определяться на дисперсионных диаграммах так же, как и раньше, с учетом того, что электрическому полю будет соответствовать либо ось частот (в случае однородных полей $\kappa = 0$), либо точка (ω_0, κ_0), где ω_0 соответствует частоте генератора (приемника) электрического поля, а κ_0 – неоднородности электрического поля, задаваемой системой электродов. Дисперсионные диаграммы возможных коллинеарных взаимодействий для этого случая показаны в табл. 3.1. Механизмы такого взаимодействия могут быть различными. В пьезоэлектрических кристаллах определяющими будут нелинейность пьезоэффекта и электрострикция. Среди эффектов коллинеарного взаимодействия, рассмотренных в табл 3 1, большое значение как для понимания физики такого взаимодействия, так и с точки зрения практического приме-

нения в устройствах акустоэлектроники (устройствах свертки и корреляции) имеет встречное взаимодействие акустических волн одинаковых частот – вырожденная акустическая свертка. Как видно из дисперсионной диаграммы (см. табл. 3.1), в этом случае вектор результирующего процесса замыкается на ось частот, т.е. результатом встречного взаимодействия может быть однородное электромагнитное поле с частотой 2ω, так как

$$(\omega,\kappa)+(\omega,-\kappa)\to(2\omega,0). \tag{3.9}$$

Будем рассматривать взаимодействие волны $A_+(U_+,E_+)$, распространяющейся в положительном направлении оси X, и встречной волны $A_-(U_-,E_-)$. Тогда результирующий процесс будет определяться уравнением

$$\frac{\partial A_3}{\partial x}=\frac{\kappa_+\kappa_-}{8}\Gamma_3 A_+ A_- \langle 2\omega,0\rangle, \tag{3.10}$$

где A_3 – однородные электрическое поле и деформация в области взаимодействия, определяемые интегралом

$$A_3=\frac{\kappa_+\kappa_-}{8}\Gamma_3\langle 2\omega,0\rangle\int_0^L A_+ A_- dx. \tag{3.11}$$

Здесь используем обозначения $e^{-i(\omega t-\kappa x)}\equiv\langle\omega,\kappa\rangle$,

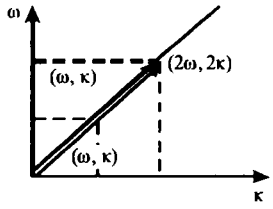
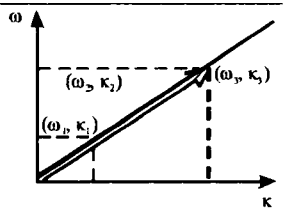
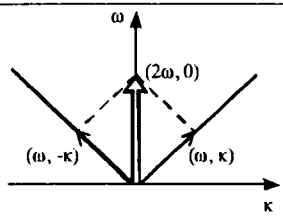
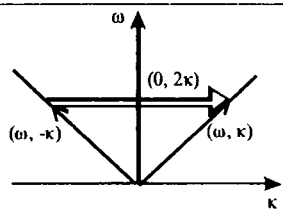
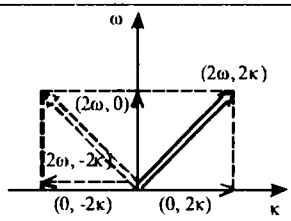
$$e^{-2i\omega t}\equiv\langle 2\omega,0\rangle;\quad e^{-2i\kappa x}\equiv\langle 0,2\kappa\rangle,\ \text{и т. п.}$$

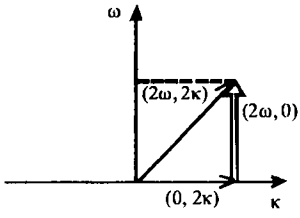
Параметр взаимодействия Γ_3 для этого случая можно определить, воспользовавшись нелинейным уравнением состояния пьезоэлектрического кристалла:

$$D_m=e_{mkl}u_{kl}+\frac{1}{2}e_{mkqr}u_{kl}u_{qr}+f_{mnkl}u_{kl}E_n+\varepsilon_{mn}E_n+\frac{1}{2}r_{mnp}E_nE_p. \tag{3.12}$$

Таблица 3.1

Коллинеарные взаимодействия акустических волн одинаковых поляризаций

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega + \omega = 2\omega$ $\kappa + \kappa = 2\kappa$	Генерация акустической гармоники
	$\omega_1 + \omega_2 = \omega_3$ $\kappa_1 + \kappa_2 = \kappa_3$ $\omega_1 = \omega_3 - \omega_2$ $\kappa_1 = \kappa_3 - \kappa_2$	Сигнал суммарной частоты Сигнал разностной частоты
	$\omega + \omega = 2\omega$ $\kappa - \kappa = 0$	Вырожденная свертка
	$\omega - \omega = 0$ $\kappa - (-\kappa) = 2\kappa$	Запись акустического сигнала в устройствах памяти
	$2\omega + 0 = 2\omega$ $0 \pm 2\kappa = \pm 2\kappa$	Считывание записанного сигнала электрическим импульсом

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$0 + 2\omega = 2\omega$ $2\kappa - 2\kappa = 0$	Считывание записанного сигнала акустической волной

В области взаимодействия электрическое поле и вектор смещения будут определяться суммой соответствующих векторов встречных волн:

$$E_m = E_m^+ + E_m^-, \quad u_i = u_i^+ + u_i^-. \tag{3.13}$$

Вводя обозначения

$$\begin{aligned} u_i^+ &= U^+ p_i^+ \langle \omega, \kappa \rangle, \quad u_i^- = U^- p_i^- \langle \omega, -\kappa \rangle, \\ E_m^+ &= E^+ q_m^+ \langle \omega, \kappa \rangle, \quad E_m^- = E^- q_m^- \langle \omega, -\kappa \rangle, \end{aligned} \tag{3.14}$$

и учитывая связь между вектором смещения и пьезоэлектрическим полем для волны основной частоты:

$$\varphi_{,n} = -\frac{E_k}{\varepsilon} u_{k,n}, \quad Eq_m = -i\kappa \frac{E_k p_k}{\varepsilon} U_0 q_m,$$

индукцию в пьезоэлектрическом кристалле можем записать в виде

$$D_m = D_m^{+\omega} + D_m^{-\omega} + D_m^{+2\omega} + D_m^{-2\omega} + D_m^e + D_m^n, \tag{3.15}$$

где

$$D_m^{+\omega} = i\kappa U^+ \left(e_{mkl} p_k p_l - \varepsilon_{mn} q_n \frac{E_k p_k}{\varepsilon} \right) \langle \omega, \kappa \rangle \tag{3.16}$$

– поляризация, сопровождающая акустическую волну на основной частоте ω. При $\mathsf{E}_k = 0$ это будет поперечная поляризация, соответствующая поперечной пьезоактивности;

$$D_m^{+2\omega} = -\kappa^2 \left(U^+\right)^2 \left[e_{mkl} p_k n_l + e_{mklqr} n_l n_r p_k p_q - \right.$$
$$\left. - f_{mnkl} n_l p_k q_n \frac{\mathsf{E}_r p_r}{\varepsilon} + \varepsilon_{mnp} q_m q_p \left(\frac{\mathsf{E}_r p_r}{\varepsilon}\right)^2 \right] \langle 2\omega, 0 \rangle \tag{3.17}$$

– поляризация на удвоенной частоте – гармоника поляризации. Аналогичные соотношения получим для обратных волн, т. е. для $D_m^{-\omega}$ и $D_m^{-2\omega}$. Эти эффекты соответствуют обычному линейному распространению звука и генерации акустической гармоники. Наибольший интерес представляют члены D_m^c и D_m^n, соответствующие эффектам свертки и памяти на встречных волнах. Эти эффекты позволяют объяснить некоторые физические явления, например электроакустическое или фононное эхо.

$$\left.\begin{matrix} D_m^c \\ D_m^n \end{matrix}\right\} = \frac{1}{2}\kappa^2 \left(U^+ U^-\right) \left[2 e_{mkl} n_k n_l p_n^+ p_n^- + \right.$$
$$+ e_{mklqr} n_l n_r \left(p_k^+ p_q^+ + p_q^+ p_k^-\right) -$$
$$- f_{mnkl} n_l \left(p_k^+ q_n^- + p_n^+ q_k^-\right) \frac{\mathsf{E}_r p_r}{\varepsilon} + \tag{3.18}$$
$$\left. + r_{mnp} \left(q_n^+ q_p^- + q_n^- q_p^+\right) \left(\frac{\mathsf{E}_r p_r}{\varepsilon}\right)^2 \right] \begin{Bmatrix} \langle 2\omega, 0 \rangle \\ \langle 0, 2\kappa \rangle \end{Bmatrix}.$$

При взаимодействии встречных волн одинаковой поляризации – случай, часто осуществляемый на практике, – выражение (3.18) можно упростить, так как $p_k^+ = p_k^-$, $q_n^+ = q_n^-$. Тогда индукцию свертки можно представить в виде

$$D_m^c = \kappa^2 \left(U^+ U^-\right) e_{mklqr}^+ n_l n_r p_k p_q , \tag{3.19}$$

где e'_{mklqr} – эффективный коэффициент нелинейности пьезоэффекта с учетом геометрической нелинейности. Пусть встречное взаимодействие объемных волн наблюдается в структуре, показанной на рис. 3.6. Тогда напряжение свертки на электродах будет

$$V_c = \int_0^d E dx = \int_0^d \frac{D}{\varepsilon} dx = \frac{D^c}{\varepsilon} d, \tag{3.20}$$

а ток свертки

$$I_c = \frac{dD}{dt} S = 2i\omega D^c S. \tag{3.21}$$

Здесь d – толщина кристалла (расстояние между электродами), S – площадь электрода кристалла. Выражение для выходного импеданса устройства свертки можно представить в виде

$$Z = \frac{V_c}{I_c} = \frac{d}{S 2i\omega} = \frac{1}{2i\omega C}, \tag{3.22}$$

где C – емкость между электродами кристалла.

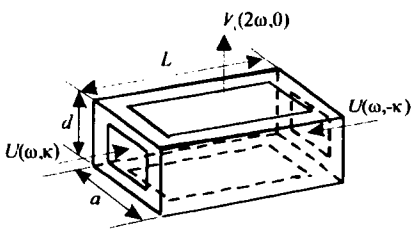


Рис. 3.6. Схема наблюдения эффекта свертки на объемных акустических волнах

Явление акустической свертки принято характеризовать фактором M, определяемым выражением

$$M = \frac{V_c}{(P_+ P_-)^{1/2}}, \tag{3.23}$$

где P_+ и P_- – акустические мощности встречных волн. Удобно определить этот фактор при $P_+ = P_-$, причем

$$P_+ = \omega\kappa\rho\upsilon^2\left(U^+\right)^2 ad\,. \tag{3.24}$$

Подставляя (3.24) и (3.20) в (3.23) и учитывая (3.19), получим

$$M = \frac{e^+_{mkqr} n_l n_r p_k p_q p_m}{\varepsilon\rho\upsilon^2} = \frac{\tilde{E}}{\varepsilon}\left(\rho\upsilon^3\right)^{-1}, \tag{3.25}$$

где

$$\tilde{E} = e^+_{mkqr} n_l n_r p_k p_q p_m\,. \tag{3.26}$$

Особенности взаимодействия поверхностных акустических волн друг с другом и с объемными волнами. Более эффективно, чем на ОАВ, нелинейные взаимодействия удается наблюдать на поверхностных акустических волнах (ПАВ) – волнах, распространяющихся вдоль свободной поверхности твердого тела или вдоль границы твердого тела с другими средами. Различают волны Рэлея, распространяющиеся вдоль границы твердого тела с вакуумом или достаточно разреженной газовой средой; волны на границе «твердое тело – жидкость»; волны Стоунли на границе раздела двух твердых сред; волны Лява на границе «твердое полупространство – твердый слой» и др.

Особенность ПАВ состоит в том, что колебания частиц в волне сильно затухают при удалении от границы вглубь твердого тела (на расстоянии порядка длины волны) и, следовательно, энергия ПАВ сосредоточивается в узком приграничном слое звукопровода. Возбуждение ПАВ в пьезоэлектрических средах можно осуществить встречно-штыревым преобразователем (ВШП), подав на него электрический сигнал соответствующей частоты. Для возбуждения ПАВ достаточно создать поле механических напряжений лишь вблизи поверхности звукопровода, поэтому даже при сравнительно небольших электрических напряжениях на ВШП удается получить амплитуду звуковой волны, необходимую для наблюдения нелинейных взаимодействий.

Условия взаимодействия ПАВ в основном такие же, как и ОАВ. В частности, должны выполняться условия синхронизма для частот и волновых векторов. Однако взаимодействия ПАВ обладают рядом особенностей. Например, на ПАВ легко наблюдать несинхронные невырожденные взаимодействия ($\omega_1 \neq \omega_2$), при которых в области пересечения пучков волн результирующее возмущение будет существовать в виде электрического поля с частотой ω_3 и пространственной неоднородностью, определяемой волновым вектором $\vec{\kappa}_3$. Если в области взаимодействия поместить решетку электродов с той же периодичностью, т.е. такую, у которой расстояние между штырями $d_{\text{ВШП}}$ выбрано из соотношения

$$d_{\text{ВШП}} = \frac{\pi \upsilon_{\text{ПАВ}}}{|\omega_1 - \omega_2|},$$

то результирующее поле можно будет выделить приемником, настроенным на частоту $\omega_1 + \omega_2$.

Особый случай представляют взаимодействия ОАВ с ПАВ, когда перемножение колебаний происходит в тонком слое, пропорциональном длине ПАВ. При таких взаимодействиях условия синхронизма выполняются не для самих волновых векторов $\vec{\kappa}$, а для их проекций на границу кристалла и возможны следующие типы взаимодействий:

$$\text{ПАВ}(\omega_1) + \text{ПАВ}(\omega_2) = \text{ОАВ}(\omega_1 + \omega_2);$$
$$\text{ПАВ}(\omega_1) + \text{ОАВ}(\omega_2) = \text{ОАВ}(\omega_2 \pm \omega_1);$$
$$\text{ПАВ}(\omega_1) + \text{ОАВ}(\omega_2) = \text{ПАВ}(\omega_1 - \omega_2).$$

Возможно также наблюдение эффекта двухимпульсного фононного эха при электроакустическом взаимодействие ПАВ. «Прямая» поверхностная волна с частотой ω_1 взаимодействовала с электрическим полем частоты ω_2, поданным не на сплошной, а распределенный параметрический электрод, расстояние между штырями которого рассчитывалось так, чтобы при взаимодействии образовывались «обращенные» волны на разностной частоте $\omega_3 = \omega_2 - \omega_1$. «Обращен-

ные» ПАВ принимались несколькими ВШП, расположенными на пути распространения волн. Причем оказалось, что амплитуда «обращенных» ПАВ растет по мере их удаления от района взаимодействия. Это означает, что при распространении «обращенных» ПАВ по звукопроводу происходит их сфазирование, т. е. эффект, характерный для двухимпульсного фононного эха.

Данное электроакустическое взаимодействие ПАВ являлись коллинеарными, так как вектор, характеризующий пространственную периодичность электрического поля, был коллинеарен волновому вектору ПАВ. Однако возможны и неколлинеарные взаимодействия ПАВ, если преобразователь, возбуждавший входную ПАВ, располагается перпендикулярно (в общем случае угол может быть произвольным) распределенному параметрическому электроду, на который подается электрическое поле. Генерируемая в результате взаимодействия волна принималась преобразователем, расположенным под некоторым углом к направлению распространения входной волны, который рассчитывался из условий синхронизма для неколлинеарного электроакустического взаимодействия.

3.3. Анализ возможных трехволновых взаимодействий в пьезоэлектрических кристаллах

Как было показано выше, квадратичная нелинейность, которая имеет место в кристаллах, открывает возможность трехволновых взаимодействий в кристаллах, т. е. взаимодействий двух волн с образованием третьей волны комбинационной частоты. Нелинейность пьезоэффекта, квадратичная по своему характеру, позволяет включить в эти взаимодействия высокочастотные электрические поля. Таким образом, даже в коллинеарном случае возрастает число возможных видов взаимодействий, которые могут приводить к интересным физическим эффектам. Так, в табл. 3.1 уже были рассмотрены эффекты генерации гармоник:

$$(\omega, \kappa) + (\omega, \kappa) = (2\omega, 2\kappa) \tag{3.27}$$

и коллинеарного взаимодействия волн:

$$(\omega_1, \kappa_1) \pm (\omega_2, \kappa_2) = (\omega_1 \pm \omega_2, \kappa_1 \pm \kappa_2). \tag{3.28}$$

В этих случаях результирующий вектор замыкается на дисперсионную характеристику той же волны.

Встречное взаимодействие волн одинаковых поляризаций (фазовых скоростей) и одинаковых частот дает эффект акустической свертки:

$$(\omega,\kappa)+(\omega,-\kappa)=(2\omega,0), \tag{3.29}$$

в котором результирующий вектор замыкается на ось частот (электромагнитную ось), и в пространстве взаимодействия образуется однородное электрическое поле. Это же взаимодействие приводит к образованию постоянного во времени неоднородного электрического поля (эффект акустической памяти), табл. 3.1. Неоднородное поле может быть записано в кристалле и потом считано электрическим импульсом или акустической волной.

В табл. 3.2 рассмотрены эффекты невырожденной свертки (волны одинаковых поляризаций, но разных частот) на суммарной

$$(\omega_1,\kappa_1)+(\omega_2,\kappa_2)=(\omega_1+\omega_2,\kappa_1-\kappa_2) \tag{3.30}$$

и разностной

$$(\omega_1,\kappa_1)-(\omega_2,\kappa_2)=(\omega_1-\omega_2,\kappa_1+\kappa_2) \tag{3.31}$$

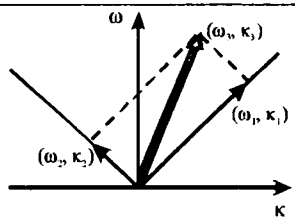
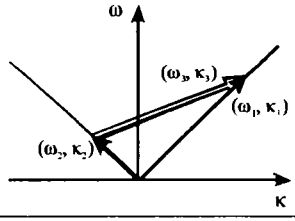
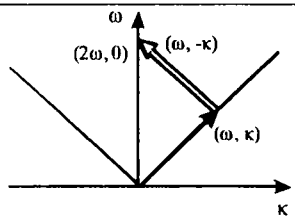
частотах. В этих случаях результирующий вектор не попадает на дисперсионную характеристику естественного волнового процесса. Такой процесс на частоте $\omega_3=\omega_1\pm\omega_2$ может быть принят искусственной системой – решеткой с периодом

$$\Lambda_3=\frac{\lambda_1\lambda_2}{\lambda_2\pm\lambda_1} \tag{3.32}$$

(знак – соответствует суммарной частоте, λ_1 и λ_2 – длины взаимодействующих волн).

Таблица 3.2

Коллинеарные взаимодействия акустических волн одинаковых поляризаций

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega_3 = \omega_1 + \omega_2$ $\kappa_3 = \kappa_1 - \kappa_2$	Невырожденная свертка на суммарной частоте $\Lambda_3 = \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1}$
	$\omega_1 - \omega_2 = \omega_3$ $\kappa_1 + \kappa_2 = \kappa_3$	Невырожденная свертка на разностной частоте $\Lambda_3 = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2}$
	$2\omega - \omega = \omega$ $0 - \kappa = -\kappa$	Генерация обратной волны вырожденное параметрическое усиление прямой волны

Возможны и обратные процессы. Так, взаимодействие электрического поля и акустической волны

$$(2\omega, 0) + (\omega, \kappa) = (\omega, -\kappa) \qquad (3.33)$$

создает волну, распространяющуюся в обратном направлении относительно первичной акустической волны. Такая генерация обратной волны может рассматриваться как параметрический процесс при накачке на удвоенной частоте. В случае взаимодействия волн разных поляризаций также можно наблюдать генерацию обратной волны. При этом, как показано в табл. 3.3, взаимодействие однородного электрического поля частоты ω и «прямой» продольной волны приводит к генерации «обратной» сдвиговой волны

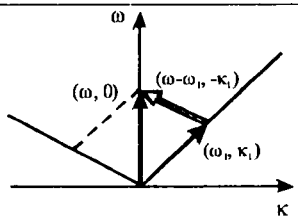
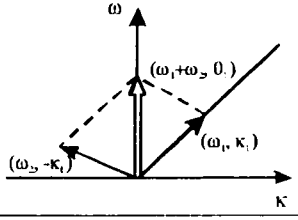
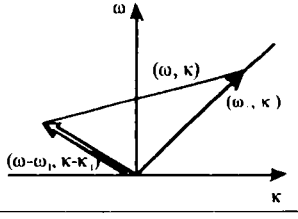
$$(\omega,0)+(\omega_1,\kappa_1)=(\omega-\omega_1,-\kappa), \tag{3.34}$$

частота которой равна

$$\omega_3=\omega-\omega_1=\omega_1\frac{\upsilon_s}{\upsilon_L}. \tag{3.35}$$

Таблица 3.3

Коллинеарные взаимодействия акустических волн разных поляризаций и разных частот

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega-\omega_1=\omega_3$ $0-\kappa_1=\kappa_3$	Генерация обратной волны новой поляризации
	$\omega_1+\omega_2=\omega_3$ $\kappa_1-\kappa_1=0$	Квазивырожденная свертка для волн разной поляризации
	$\omega-\omega_1=\omega_3$ $\kappa-\kappa_1=\kappa_3$ $\kappa=\frac{\pi}{L}$ $L=\frac{\pi V_L V_S}{\omega V_L-\omega_1(V_L-V_S)}$	Генерация обратной волны параметрической решеткой

При взаимодействии встречных волн разных поляризаций можно наблюдать эффект квазивырожденной свертки:

$$(\omega_1,\kappa_1)+(\omega_2,\kappa_2)=(\omega_1+\omega_2,0), \tag{3.36}$$

т.е. образование однородного электрического поля в области взаимодействия. Такой эффект интересен для создания устройства свертки в сверхвысокочастотном диапазоне, так как позволяет развязать входы по поляризации и частоте.

При взаимодействии волн разных поляризаций также можно использовать параметрические решетки. В этом случае взаимодействие одной из волн (продольной) с полем решетки, на которую подан сигнал с частотой ω, будет приводить к генерации другой объемной (сдвиговой) волны:

$$(\omega,\kappa)+(\omega_1,\kappa_1)=(\omega-\omega_1,\kappa_2), \tag{3.37}$$

как указано на дисперсионной диаграмме табл. 3.3.

В табл. 3.4 рассмотрены эффекты, возможные при взаимодействии волн разных поляризаций, но одинаковых частот. В этом случае в области взаимодействия может образоваться постоянное во времени (за время взаимодействия) неоднородное электрическое поле:

$$(\omega_1,\kappa_1)+(\omega_2,\kappa_2)=(0,\kappa_1+\kappa_2). \tag{3.38}$$

Это поле может приводить к записи акустической решетки на каких-либо примесных центрах кристалла (эффекты акустической памяти и возникновение трехимпульсного электроакустического эха). В этом варианте взаимодействия можно наблюдать также квазивырожденную свертку:

$$(\omega,\kappa_1)+(\omega,-\kappa_2)=(2\omega,\kappa_1-\kappa_2). \tag{3.39}$$

Наконец, в этом варианте можно наблюдать отражение акустической волны с изменением поляризации от параметрической решетки, на которую подано постоянное (по времени) электрическое поле, создающее периодическое изменение фазовой скорости:

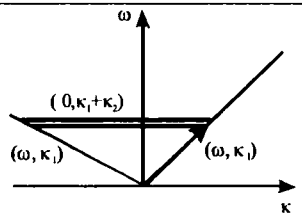
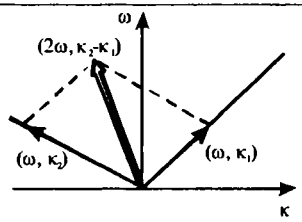
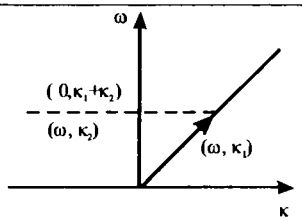
$$(0,\kappa)+(\omega_1,\kappa_1)=(\omega_1,-\kappa_2), \tag{3.40}$$

где

$$\kappa = \frac{\pi}{L}; \quad L = \frac{\pi \upsilon_L \upsilon_S}{\omega\left(\upsilon_L + \upsilon_S\right)}. \tag{3.41}$$

Таблица 3.4

Коллинеарные взаимодействия акустических волн различных поляризаций и одинаковых частот

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega_3 = \omega - \omega = 0$ $\kappa_3 = \kappa_1 + \kappa_2$	Запись сигнала при встречном взаимодействии волн разных поляризаций
	$\omega_3 = \omega + \omega = 2\omega$ $\kappa_3 = \kappa_2 - \kappa_1$ $L = \frac{\pi V_L V_S}{\omega\left(V_L - V_S\right)}$	Квазивырожденная свертка с регистрацией параметрической решеткой
	$\omega_3 = \omega - \omega = 0$ $\kappa_3 = \kappa_1 + \kappa_2$	Генерация обратной волны при взаимодействии с параметрической решеткой

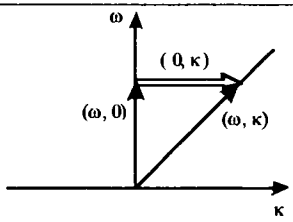
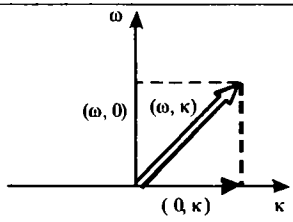
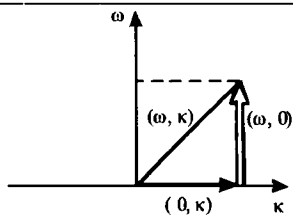
Взаимодействие бегущей акустической волны с однородным электрическим полем также может быть использовано для записи и считывания акустических сигналов, как показано в табл. 3.5.

Запись и считывание сигналов можно осуществлять, используя взаимодействие поля решетки электродов с акустической волной. Как показано в табл. 3.6, запись сигнала, переносимого акустической волной с частотой ω, будет происходить, если на решетку электродов подать импульс с частотой ω. Считывание можно проводить как акустической волной, так и электрическим импульсом. Следует отметить, что осуществление взаимодействия объемных акустических

волн с электрическим полем решетки затрудняется тем, что поле решетки неглубоко проникает в кристалл. Однако этот эффект может играть существенную роль при взаимодействии неоднородных волн, возникающих при косом падении акустических волн на границу раздела.

Таблица 3.5

Взаимодействия акустических волн с однородным электрическим полем

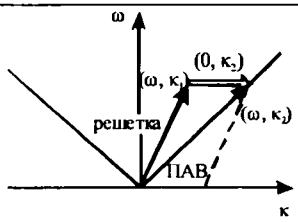
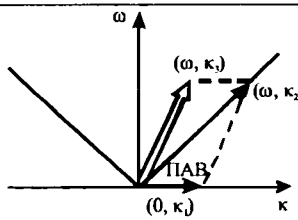
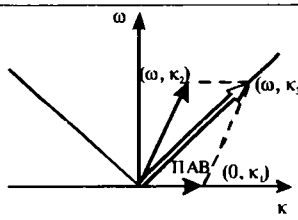
Дисперсионная диаграмма	Условия синхронизма	Примечание
	ω – ω = 0 κ – 0 = κ	Запись сигнала электрическим импульсом
	ω + 0 = ω 0 + κ = κ	Считывание записанного сигнала электрическим импульсом
	ω – 0 = ω κ – κ = 0	Считывание записанного сигнала акустической волной

Наконец, представляют интерес эффекты параметрического взаимодействия полей двух периодических решеток (табл. 3.7). В этом случае возможна параметрическая генерация волны. Если на две решетки электродов с периодами L_1 и L_2 подать сигналы с частотами ω_1 и ω_2 соответственно, то можно возбудить акустическую волну с частотой

$$\omega = \omega_1 + \omega_2 = \frac{\pi\upsilon(L_1 + L_2)}{L_1 L_2}. \qquad (3.42)$$

Таблица 3.6

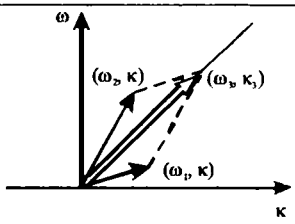
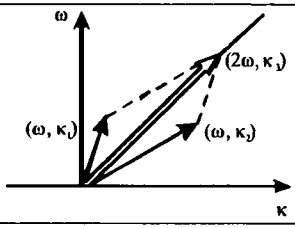
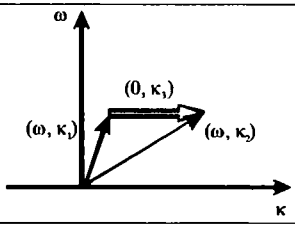
Коллинеарные взаимодействия акустических волн одинаковых поляризаций

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega - \omega = 0$ $\kappa_1 - \kappa_2 = \kappa_3$	Невырожденная запись сигнала $\Lambda_3 = \frac{\lambda_1\lambda_2}{\lambda_1 - \lambda_2}$
	$\omega - 0 = \omega$ $-\kappa_1 + \kappa_2 = \kappa_3$	Невырожденное считывание записанного сигнала акустической волной
	$0 + \omega = \omega$ $\kappa_1 + \kappa_2 = \kappa_3$	Невырожденное считывание записанного сигнала электрическим импульсом

Такое взаимодействие можно использовать для генерации поверхностных волн в диапазоне сверхвысоких частот с использованием решеток большего периода, чем это требуется на основной частоте. Частным случаем такого взаимодействия является генерация второй гармоники. Наконец, параметрическое взаимодействие электрических полей двух решеток можно использовать для записи сигналов.

Таблица 3.7

Параметрическое взаимодействие полей периодических решеток

Дисперсионная диаграмма	Условия синхронизма	Примечание
	$\omega_3 = \omega_1 + \omega_2$ $\kappa_3 = \kappa + \kappa$	Параметрическое возбуждение акустической волны решеткой
	$\omega + \omega = 2\omega$ $\kappa_1 + \kappa_2 = \kappa_3$	Генерация второй гармоники двухпериодической решеткой
	$\omega - \omega = 0$ $\kappa_1 - \kappa_2 = \kappa_3$	Запись сигнала двухпериодической решеткой

В случае однородного электрического поля в направлении распространения акустической волны амплитуда акустической волны будет определяться уравнением

$$\frac{\partial A_2}{\partial x} = \frac{1}{8}\Gamma_2 U_1 E_2, \tag{3.43}$$

а параметр взаимодействия

$$\Gamma_2 = \frac{\kappa_1^2 e^*_{mslkj} n_t n_l p_k^{(1)} q_m p_s^{(2)}}{\rho \kappa_2 v_2^2}. \tag{3.44}$$

Если рассматривать процесс генерации обратной волны, то Γ_2 будет параметром взаимодействия этого процесса. Воспользовавшись (3.26), Γ_2 можем представить в виде

$$\Gamma_2 = \frac{\kappa_1^2 \tilde{E}}{\kappa_2 \rho v_2^2}. \qquad (3.45)$$

Метод расчета этого параметра совпадает с методом расчета фактора М. Следовательно, генерация обратной волны возможна в тех же направлениях, что и акустическая свертка. Амплитуда сигнала обратной волны при длине области взаимодействия L будет

$$A_2 = \frac{\kappa_1^2 \tilde{E}}{8\kappa_2 \rho v_2^2} L. \qquad (3.46)$$

В случае равных частот акустических волн (см. (3.25))

$$\Gamma_2 = \omega \varepsilon M. \qquad (3.47)$$

Таким образом, параметр взаимодействия для генерации обратной волны Γ_2 совпадает с точностью до $\omega\tau$ с фактором свертки М. Следовательно, анализ анизотропии фактора М позволяет судить и о других процессах трехволнового взаимодействия акустических волн с электрическим полем в кристаллах.

3.4. Нелинейные свойства твердых тел

Решеточная нелинейность монокристаллов. Ранее было показано, что для взаимодействия акустических волн в твердом теле среда, по которой волна распространяется, должна обладать нелинейными свойствами, и рассмотрено взаимодействие волн при наличии упругой нелинейности кристалла. Строго говоря, любое деформируемое твердое тело обладает нелинейными свойствами уже в силу нелинейной связи компонент тензора механических деформаций с производными от смещения, т. е. при достаточно больших смещениях тензор деформаций должен быть записан в виде (1.6)

$$u_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_k}{\partial x_i} \cdot \frac{\partial u_k}{\partial x_j} \right).$$

Эта особенность конечных деформаций не зависит от физических свойств деформируемого тела и обычно называется геометрической нелинейностью. Существование геометрической нелинейности приводит к тому, что упругая нелинейность кристаллов описывается не обычным тензором c_{iklm}, а «эффективным» тензором нелинейной упругости

$$C^{*}_{ijklqr} = c_{ijrl}\delta_{qk} + c_{jlqr}\delta_{ik} + c_{jklr}\delta_{iq} + C_{ijklqr},$$

где δ – символ Кронекера, который для пары индексов *ij* определяется как

$$\delta_{ij} = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Однако упругая нелинейность лишь частный случай решеточной нелинейности, т. е. нелинейности, обусловленной свойствами кристаллической решетки. Например, поскольку в пьезоэлектрических кристаллах распространение акустической волны (т. е. волны механических деформаций) сопровождается распространением и волны электрического поля, то при достаточно большой амплитуде возбуждения, когда поляризуемость атомов становится нелинейной, проявляется еще один механизм взаимодействия – нелинейный пьезоэффект.

Для описания нелинейных взаимодействий акустических волн в пьезоэлектрике воспользуемся разложением электрической энтальпии кристалла $H_{ЭЛ}$ по степеням компонент вектора напряженности электрического поля $\vec{E}$ и тензора деформаций, учитывая члены до третьего порядка:

$$H_{ЭЛ} = H_{ЭЛ0} - \frac{1}{2}\chi_{ij}E_iE_j - \frac{1}{6}\tilde{\chi}_{ijk}E_iE_jE_k +$$
$$+\frac{1}{2}c_{ijkl}u_{ij}u_{kl} + \frac{1}{6}C_{ijklmn}u_{ij}u_{kl}u_{mn} - e_{ijk}E_iu_{jk} -$$
$$-\frac{1}{2}\tilde{e}_{ijklm}E_iu_{jk}u_{lm} - \frac{1}{2}Q_{ijkl}E_iE_ju_{kl},$$

где χ_{ij} и $\tilde{\chi}_{ijk}$ – компоненты тензоров линейной и нелинейной диэлектрических проницаемостей; e_{ijk} и $\tilde{e}_{ijklm}$ – компоненты тензоров линейного и нелинейного пьезоэффектов; Q_{ijkl} – компонент тензора электрострикции.

В этом случае свойства кристалла описываются не законом Гука, а более общими уравнениями состояния:

$$\left.\begin{aligned} T_{ij} &= \frac{\partial H_{ЭЛ}}{\partial u_{ij}} = c_{ijkl}u_{kl} + \frac{1}{2}C_{ijklmn}u_{kl}u_{mn} - e_{kij}E_k - \frac{1}{2}Q_{klij}E_kE_l - \tilde{e}_{kijlm}E_ku_{lm};\\ D_i &= \frac{\partial H_{ЭЛ}}{\partial E_i} = \chi_{ij}E_j + \frac{1}{2}\tilde{\chi}_{ijk}E_jE_k + e_{ijk}u_{jk} + Q_{ijkl}E_ju_{kl} + \frac{1}{2}\tilde{e}_{ijklm}u_{jk}u_{lm}, \end{aligned}\right\} \quad (3.48)$$

где D_i – компонент вектора электрической индукции.

Если в (3.48) исключить линейные члены, то для нелинейных составляющих механического напряжения и электрической индукции получим

$$\left.\begin{aligned} T_{ij}^{НЛ} &= \frac{1}{2}C_{ijklmn}u_{kl}u_{mn} - \frac{1}{2}Q_{klij}E_kE_l - \tilde{e}_{kijlm}E_ku_{lm};\\ D_i^{НЛ} &= \frac{1}{2}\tilde{\chi}_{ijk}E_jE_k + Q_{ijkl}E_ju_{kl} + \frac{1}{2}\tilde{e}_{ijklm}u_{jk}u_{lm}. \end{aligned}\right\} \quad (3.49)$$

Из соотношений (3.49) видно, что появление акустической и электроакустической волн на комбинационных частотах может быть вызвано взаимодействиями типа «деформация – деформация», «электрическое поле – электрическое поле» и «деформация – электрическое поле». При этом решеточную нелинейность кристалла можно подразделить на упругую (тензор C_{ijklmn}), пьезоэлектрическую ($\tilde{e}_{iklm}$), электрострикционную (Q_{ijkl}) и диэлектрическую ($\tilde{\chi}_{ijk}$).

Концентрационная нелинейность полупроводников. Особый случай представляют взаимодействия акустических волн, распространяющихся в полупроводниковом материале. Смещение атомов решетки при прохождении акустической волны приводит к изменению внутрикристаллических полей, и тем самым к воздействию волны на свободные электроны полупроводника. В зависимости от типа кристалла это воздействие может быть чисто ионным, пьезоэлектрическим, электро- или магнитострикционным или носить потенциал-деформационный характер за счет изменения ширины запрещенной зоны полупроводника под действием деформации кристалла. Наиболее эффективным на частотах до 10^{11} Гц является пьезоэлектрический механизм воздействия акустической волны на электроны проводимости в пьезополупроводниковых средах.

Нелинейные эффекты, проявляющиеся при взаимодействии акустических волн с электронами проводимости, связаны с нелинейной зависимостью кинетических коэффициентов от концентрации электронов вследствие связи времени релаксации импульса электронов с их энергией, с «рекомбинационной» нелинейностью, обусловленной наличием произведений переменных концентраций электронов в зоне проводимости и на ловушках в выражении для скорости захвата и др. Доминирующим типом нелинейности является концентрационная нелинейность, связанная с перераспределением электронов проводимости в присутствии акустической волны и проявляющаяся в нелинейности уравнений для электрического тока, потока энергии и уравнения баланса энергии.

При распространении двух акустических волн, сопровождаемых волнами электрического поля:

$$E_1 = E_1^0 \exp\left[i\left(\omega_1 t - \vec{\kappa}_1 \vec{r}_1\right)\right],$$
$$E_2 = E_2^0 \exp\left[i\left(\omega_2 t - \vec{\kappa}_2 \vec{r}_2\right)\right],$$

где E_i^0 – амплитуда соответствующей волны, в полупроводнике возникнут и волны концентрации свободных электронов:

$$N_1 = N_1^0 \exp\left[i\left(\omega_1 t - \vec{\kappa}_1 \vec{r}_1\right)\right],$$
$$N_2 = N_2^0 \exp\left[i\left(\omega_2 t - \vec{\kappa}_2 \vec{r}_2\right)\right].$$

При этом электрическое поле одной акустической волны может взаимодействовать с носителями, сопровождающими другую волну, что приводит к образованию нелинейной индукции $D^{НЛ}$ и нелинейной составляющей электрического тока:

$$I^{НЛ} \sim \frac{\partial D^{НЛ}}{\partial t}.$$

Без учета процессов диффузии ток $I^{НЛ}$ связан с параметрами взаимодействующих волн как

$$I^{НЛ} = q\mu\left(N_1 E_2 + N_2 E_1\right),$$

где q и μ – заряд и подвижность носителей соответственно.

Концентрационная нелинейность начинает сказываться, как правило, при меньших амплитудах акустических волн, чем другие типы нелинейности, а эффективность взаимодействия при этом оказывается на несколько порядков выше, чем, например, при взаимодействии за счет упругой нелинейности. Эффективность взаимодействий, обусловленных концентрационной нелинейностью, легко поддается управлению с помощью изменения числа носителей (например, при подсветке фотополупроводников) или при приложении смещающего электрического поля. Кроме того, важным свойством полупроводников (в отличие от обычных диэлектриков) является существование эффекта акустоэлектронной памяти.

Применение полупроводниковых материалов позволяет использовать для осуществления нелинейного взаимодействия акустических волн еще один физический механизм. Если над звукопроводом, по которому распространяются ПАВ, поместить полупроводниковую пластинку с созданной на ее поверхности матрицей диодов (например, *p-n* диодов или диодов Шотки), то перемножение электрических полей, сопровождающих ПАВ, и запоминание неоднородного распределения зарядов будет происходить за счет нелинейности вольт-амперной характеристики диодов.

3.5. Интегральные преобразования сигналов, получаемые в результате нелинейных взаимодействий акустических волн

Устройства свертки используются для целого ряда преобразований радиосигналов. Если акустические волны в твердом теле возбуждаются радиочастотными сигналами, то с помощью нелинейных взаимодействий можно осуществить различные интегральные преобразования входных сигналов.

Пусть один из входных сигналов имеет вид $F = F(t)\exp\left[i\left(\omega_1 t - \vec{\kappa}_1\vec{r}_1\right)\right]$, а второй – $G = G(t)\exp\left[i\left(\omega_2 t - \vec{\kappa}_2\vec{r}_2\right)\right]$. В каждой точке кристалла, в которой существуют акустические волны, вызванные обоими сигналами, с амплитудами u_1, зависящей от $F(t)$, и u_2, зависящей от $G(t)$, происходит их перемножение.

Тогда выходной сигнал, соответствующий напряжению, снимаемому с распределенного параметрического электрода, находящегося в области взаимодействия акустических волн, можно представить в виде

$$U(t) \sim \int_{-L/2}^{L/2} F\left(t - \frac{z}{\upsilon}\right) G\left(t + \frac{z}{\upsilon}\right) dz ,$$

где L – длина области взаимодействия (параметрического электрода); υ – скорость акустической волны.

Если длительность сигналов не превышает длины области взаимодействия, то выходной сигнал представляет собой акустическую свертку входных сигналов:

$$U(t) \sim \int_{-\infty}^{\infty} F(\tau) G(2t - \tau) d\tau . \qquad (3.50)$$

Если сигнал F подается на кристалл в виде электрического поля, а сигнал G возбуждает в нем акустическую волну, то выходной сигнал принимает вид функции акустической корреляции:

$$U(t) \sim \int_{-\infty}^{\infty} F(\tau) G(2\tau - t)\, d\tau . \qquad (3.51)$$

Следует отметить, что выражения (3.50) и (3.51) отличаются от принятых определений свертки и корреляции двух функций, имеющих вид

$$U_{св}(t) \sim \int_{-\infty}^{\infty} F(\tau) G(t - \tau)\, d\tau , \qquad (3.52)$$

$$U_{кор}(t) \sim \int_{-\infty}^{\infty} F(\tau) G(\tau + t)\, d\tau , \qquad (3.53)$$

поэтому для получения функций (3.52) и (3.53) необходимы либо дополнительные преобразования входных сигналов (обращение во времени, сжатие или растяжение одного из сигналов), либо использование двух последовательных нелинейных взаимодействий (четырехфононные взаимодействия).

Рассмотрим некоторые частные случаи интегральных преобразований, которые могут быть реализованы на практике.

1. Когда в выражении (3.50) сигнал F является δ-функцией, на выходе имеем сигнал G, сжатый во времени в два раза:

$$U(t) \sim \int_{-\infty}^{\infty} \delta(\tau) G(2t - \tau)\, d\tau = G(2t) . \qquad (3.54)$$

2. Когда в выражении (3.51) сигнал G имеет вид δ-функции, на выходе имеем сигнал F, растянутый во времени в два раза:

$$U(t) \sim \int_{-\infty}^{\infty} F(\tau) \delta(2\tau - t)\, d\tau = F(t/2) . \qquad (3.55)$$

3. Когда в выражении (3.51) сигнал F имеет вид δ-функции, то происходит обращение сигнала G во времени:

$$U(t) \sim \int_{-\infty}^{\infty} \delta(\tau) G(2\tau - t)\, d\tau = G(-t). \qquad (3.56)$$

Рассмотренные интегральные преобразования входных сигналов соответствуют взаимодействию волн одной природы: ПАВ с ПАВ или ОАВ с ОАВ. Если же использовать взаимодействия объемной волны с поверхностной или взаимодействие ОАВ с электрическим полем с образованием ПАВ, то можно получить дополнительные виды интегральных преобразований: акустическую свертку сигналов, сжатую во времени, функцию корреляции, растянутую во времени и др. В табл. 3.8 приведены интегральные преобразования, осуществляемые в случае, если волновой вектор ПАВ $\vec{\kappa}_1$ и тангенциальная составляющая κ_2' волнового вектора ОАВ $\vec{\kappa}_2$ коллинеарны, а угол между $\vec{\kappa}_1$ и $\vec{\kappa}_2$ равен β. Изменяя β и соотношение скоростей взаимодействующих волн, можно управлять временным масштабом, в котором образуются функции типа свертки и корреляции.

Таблица 3.8

Интегральные преобразования сигналов при взаимодействии бегущих ПАВ и ОАВ

Тип взаимодействия	Вид выходного сигнала $U(t)$	η	$U(t)$ при $G = \delta$	$U(t)$ при $F = \delta$
$\text{ПАВ}(\omega_1) + \text{ОАВ}(\omega_0) = E(\omega_1 + \omega_0)$	$\int G(\tau) F[(\eta+1)t - \eta\tau]\, d\tau$	$-\frac{v_{\text{ПАВ}}}{v_{\text{ОАВ}}}\cos\beta$	$F[(\eta+1)t]$	$G\left(\frac{\eta+1}{\eta}t\right)$
$\text{ОАВ}(\omega_0) + E(\omega_2) = \text{ПАВ}(\omega_2 - \omega_0)$	$\int G(\tau) F[\eta\tau - (\eta-1)t]\, d\tau$	$\frac{v_{\text{ОАВ}} - v_{\text{ПАВ}}\cos\beta}{v_{\text{ОАВ}}}$	$F[(1-\eta)t]$	$G\left(\frac{\eta-1}{\eta}t\right)$
$\text{ПАВ}(\omega_1) + \text{ОАВ}(\omega_0) = \text{ПАВ}(\omega_0 - \omega_1)$	$\int G(\tau) F[\eta\tau - (\eta-1)t]\, d\tau$	$\frac{v_{\text{ПАВ}}\cos\beta - v_{\text{ОАВ}}}{2v_{\text{ОАВ}}}$	$F[(1-\eta)t]$	$G\left(\frac{\eta-1}{\eta}t\right)$

Все перечисленные виды интегральных преобразований сигналов можно осуществить с помощью нелинейных акустоэлектронных процессоров. В зависимости от режима работы и вида входных сигналов процессоры подразделяются на устройства свертки, корреляторы, устройства обращения сигналов во времени и др. (см. табл. 3.9).

Основные требования, предъявляемые к конструкции, параметрам и выбору материалов процессоров, работающих в различных режимах, достаточно близки. Поэтому рассмотрим вначале наиболее простой режим работы нелинейного процессора – режим акустической свертки, а сами устройства будем называть устройствами свертки (УС).

3.6. Классификация нелинейных акустоэлектронных устройств обработки информации

Рассмотренные взаимодействия акустических волн, основанные на нелинейных свойствах твердых тел, позволяют создавать целый ряд устройств обработки сигналов, которые можно выделить в отдельный класс – нелинейные акустоэлектронные устройства. Если в основу классификации таких устройств положить характер распространяющихся упругих волн и механизм нелинейного взаимодействия волн между собой и с другими излучениями, то все нелинейные акустоэлектронные устройства можно подразделить на три группы (табл. 3.9).

К первой группе относят устройства, основанные на нелинейности, связанной с ангармоничностью колебаний кристаллической решетки твердого тела. В нее входят устройства на основе упругой, пьезоэлектрической и других решеточных нелинейностей.

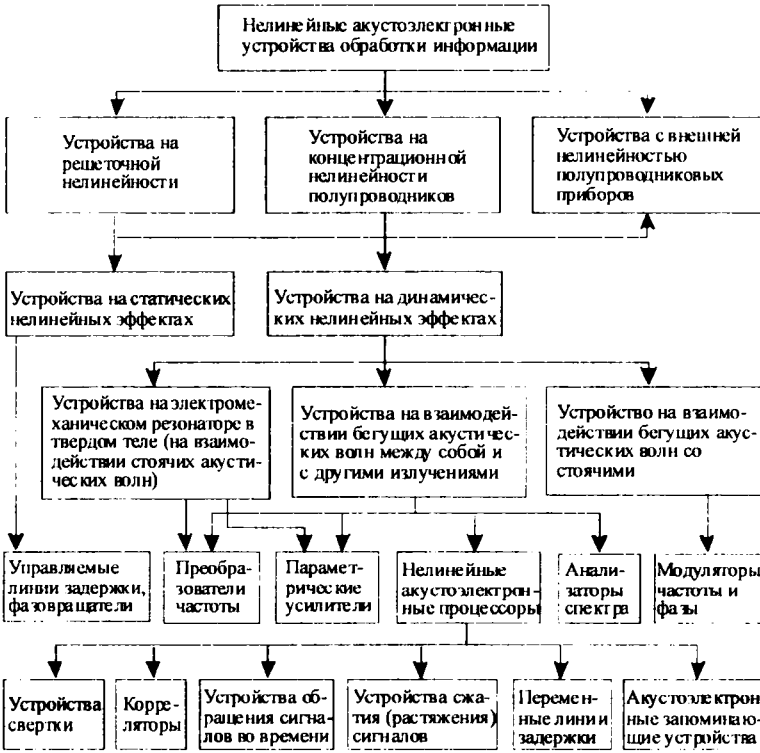
Вторую группу составляют устройства, основанные на концентрационной нелинейности полупроводников, обусловленной взаимодействием акустических волн с электронами проводимости. Такой тип взаимодействий имеет место в пьезополупроводниках, а также в слоистых структурах «полупроводник – пьезоэлектрик».

В третью группу включены устройства с внешней нелинейностью полупроводниковых приборов. В этом случае акустическая часть устройств является линейной, а нелинейное взаимодействие электрических сигналов происходит на внешних нелинейных элементах, например на полупроводниковых диодах.

Вне зависимости от механизма взаимодействия нелинейные эффекты в твердых телах можно разделить на статические и динамические.

Таблица 3.9

Классификация нелинейных акустоэлектронных устройств



Статические эффекты заключаются в изменении параметров распространения упругих волн в кристаллах под действием приложенных к звукопроводу постоянных или медленно меняющихся сил. К числу таких взаимодействий можно отнести гидростатическое давление, электрическое и магнитное поля, температуру и т. п. Статические нелинейные эффекты лежат в основе работы управляемых акустоэлектронных приборов, управляемых линий задержки, фазовращателей и т. д.

К динамическим относятся эффекты самовоздействия или взаимодействия акустических волн между собой, а также с волнами другой природы. При этом образуются акустические гармоники или волны на комбинационных частотах. На основе динамических эффектов в настоящее время создан обширный круг

устройств обработки сигнальной информации. Среди этих устройств можно выделить устройства, основанные на размерном резонансе твердых тел (на стоячих упругих волнах). Взаимодействие стоячих волн между собой используется для создания преобразователей частоты и параметрических усилителей. Параметрическое усиление можно получить и при использовании взаимодействия стоячих волн с внешним радиочастотным электрическим полем.

Наибольшими потенциальными возможностями обладают устройства, основанные на взаимодействии бегущих акустических волн между собой и с волнами другой природы. Сюда можно отнести устройства на основе взаимодействия ОАВ или ПАВ между собой, устройства на взаимодействиях ОАВ с ПАВ, устройства, основанные на взаимодействии бегущих волн с радиочастотными электрическим или магнитным полями. Причем для создания таких устройств можно использовать как коллинеарные, так и неколлинеарные взаимодействия. К этой же группе относятся и устройства, основанные на взаимодействии электрических полей между собой с образованием бегущих акустических волн.

Как видно из табл. 3.9, на основе взаимодействий бегущих волн можно создавать целый ряд устройств обработки сигнальной информации: преобразователи частоты, нелинейные акустоэлектронные процессоры – устройства, предназначенные для выполнения интегральных преобразований типа свертки и корреляции, анализаторы спектра и др.

Отдельно можно выделить устройства, основанные на взаимодействии бегущих акустических волн со стоячими. Результатом такого взаимодействия является фазовая модуляция бегущей волны. Это явление лежит в основе работы акустоэлектронных модуляторов частоты и фазы.

Классификация нелинейных акустоэлектронных устройств по характеру распространения акустических волн и по механизму, обусловливающему нелинейные свойства твердого тела, позволяет выбирать вид устройства в зависимости от требуемой обработки сигнальной информации, а также определять возможность разработки новых нелинейных акустоэлектронных устройств. Вместе с тем из табл. 3.9 видно, что целый ряд одних и тех же устройств обработки сигналов можно реализовать, используя различные механизмы нелинейности.

Контрольные вопросы

1. Какова интерпретация взаимодействия акустических волн в квантовой акустике?
2. В чём суть метода дисперсионных диаграмм при анализе взаимодействия акустических волн?
3. Какие дополнительные нелинейные члены в законе Гука учитывают одновременное распространение двух акустических волн?
4. Изобразите дисперсионную диаграмму для коллинеарного и неколлинеарного взаимодействий акустических волн.
5. Нарисуйте дисперсионную диаграмму для взаимодействия электрического поля и акустической волны в пьезоэлектрике.
6. Запишите условия синхронизма взаимодействия двух акустических волн.
7. Перечислите виды взаимодействующих волн в изотропной среде при выполнении условия синхронизма для коллинеарного и неколлинеарного случаев.
8. Каковы виды взаимодействия электрического поля и акустической волны в пьезоэлектрике.
9. В чём суть явления фононного эха при распространении акустических волн в кристалле?
10. Каковы особенности взаимодействия поверхностных акустических волн друг с другом и с объемными волнами?
11. Какие физические эффекты реализуются для трёхволновых взаимодействий в пьезоэлектрических кристаллах?
12. При каких условиях взаимодействия акустических волн реализуются:
    - функция свертки на суммарной и разностной частотах;
    - генерация обратной волны;
    - запись сигнала;
    - считывание записанного сигнала
13. Виды нелинейности свойств кристаллов.
14. Напишите основные формулы интегральных преобразований входных сигналов при взаимодействии акустических волн для случаев двух поверхностных и поверхностной и объёмной волн.
15. Назовите основные типы акустоэлектронных устройств обработки информации.

4. УСТРОЙСТВА ОБРАБОТКИ СИГНАЛОВ НА АКУСТИЧЕСКИХ ВОЛНАХ

4.1. Устройства на объемных акустических волнах

Исторически первые акустоэлектронные устройства использовали объемные акустические волны. Простейшим устройством на основе объемных акустических волн является линия задержки (рис. 4.1, а). Она состоит из звукопровода и двух пьезоэлектрических преобразователей на его торцах. В качестве преобразователей используются пьезокристаллические плоскопараллельные пластинки, на гранях которых напылены металлические электроды. На входном преобразователе переменное напряжение, приложенное к электродам, благодаря обратному пьезоэффекту вызывает вибрацию пластинки и возбуждение акустической волны в звукопроводе. Работа приемного преобразователя основана на прямом пьезоэффекте: вибрации пластинки, порождаемые упругой волной, приводят к появлению электрического сигнала. Выходное напряжение представляет собой задержанную копию входного сигнала. Время задержки зависит от длины пути акустической волны и от скорости ее распространения. Низкая скорость волн позволяет получать достаточно большие времена задержки при малых габаритных размерах, как правило, несколько микросекунд на каждый сантиметр пути акустической волны. В тех случаях, когда требуются большие задержки, используются звукопроводы сложной формы, обеспечивающие многократное отражение волны от плоских граней, увеличивающее путь акустической волны.

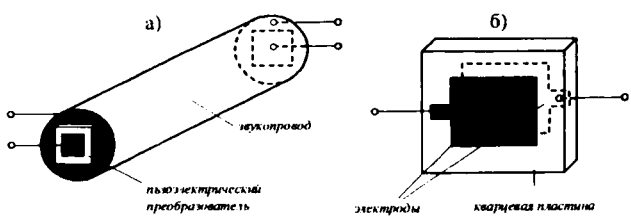


Рис. 4.1. Устройства на объемных акустических волнах:
а – простейшая линия задержки; б – резонатор

Другое широко распространенное акустическое устройство – кристаллический резонатор (рис. 4.1, б). Резонатор представляет собой плоскопараллельную пластинку кристаллического кварца с металлическими электродами на обеих сторонах. Если длина и ширина пластинки намного превышают ее толщину, то пластинка резонирует на частотах, при которых на ее толщине укладывается целое число полуволн акустической волны. Так как кристалл кварца является пьезоэлектриком, акустические колебания пластинки можно возбуждать электрическим способом. Однако если толщина пластинки равна четному числу полуволн, электрические поля, связанные с упругими колебаниями, будут иметь симметрию относительно центра резонатора, потенциалы на электродах будут одинаковы и связь между напряжением на электродах резонатора и механическими колебаниями отсутствует. Поэтому используются нечетные моды колебаний, когда на толщине пластинки укладывается нечетное число полуволн. Добротность кварцевого резонатора достигает очень высоких для радиоэлектронных элементов значений (до 10^6) при высокой температурной стабильности. Это позволяет использовать кварцевые резонаторы для стабилизации частоты генераторов и в качестве узкополосных фильтров.

Для расчета характеристик преобразователя и резонатора необходимо установить связь амплитуды упругих колебаний пластинки с электрическими потенциалами на электродах.

Пусть плоская однородная упругая волна (продольная или поперечная) распространяется вдоль оси x, а пьезоэффект отсутствует. Зависимость смещения u и напряжения T в волне от координаты x в направлении распространения и времени определяет одномерная форма уравнения движения и закона Гука:

$$\frac{\partial T}{\partial x} = \rho \frac{\partial^2 u}{\partial t^2}, \quad T = c \frac{\partial u}{\partial x}, \tag{4.1}$$

где ρ – плотность материала, c – соответствующий данной волне упругий модуль. Продифференцируем второе из уравнений по времени и введем колебательную скорость движения частиц среды $v = \partial u / \partial t$. В результате уравнения (4.1) приобретают следующий вид:

$$\frac{\partial T}{\partial x}=\rho\frac{\partial v}{\partial t},\quad \frac{\partial v}{\partial x}=\frac{1}{c}\frac{\partial T}{\partial t} \tag{4.2}$$

Соотношения (4.2) имеют форму стандартных уравнений линии передачи и могут применяться аналогичным образом. Чтобы убедиться в этом, сравним их с телеграфными уравнениями для тока I и напряжения V, описывающими распространение электромагнитных колебаний в линиях передачи:

$$\frac{\partial V}{\partial x}=-L\frac{\partial I}{\partial t},\quad \frac{\partial I}{\partial x}=-C\frac{\partial V}{\partial t}, \tag{4.3}$$

где L и C – индуктивность и емкость на единицу длины линии. Распространяющаяся в линии монохроматическая волна с частотой ω имеет постоянную распространения $\kappa=\omega\sqrt{LC}$ и характеризуется волновым импедансом

$$Z_0=\sqrt{L/C}\,.$$

Для волн в положительном направлении оси x $V,I\sim\exp\left[j\left(\omega t-\kappa x\right)\right]$, и отношение напряжения и тока в линии равно

$$\frac{V}{I}=Z_0\,.$$

Для волн в обратном направлении $V,I\sim\exp\left[j\left(\omega t+\kappa x\right)\right]$:

$$\frac{V}{I}=-Z_0\,.$$

Произведем в (4.2) замену T на $-V$ и v на I. Тогда уравнения (4.2) будут эквивалентны телеграфным уравнениям (4.3), где роль индуктивности L играет плотность материала ρ, а роль емкости C – упругая податливость среды $1/c$ для данного вида волны. При

этом аналогом импеданса электромагнитной теории является величина

$$Z_0 = -\frac{T}{v} = \sqrt{\rho c}\,, \tag{4.4}$$

которую называют удельным акустическим импедансом упругой среды.

Рассмотренная выше аналогия позволяет использовать соотношения, известные из электромагнитной теории, для расчета амплитуд прошедшей и отраженной волны при нормальном падении акустической волны на границу раздела сред (рис. 4.2). Коэффициенты отражения $\mathcal{R}$ и прохождения $\mathcal{T}$ по амплитуде, связанные со «скачком» импеданса на границе сред определятся выражениями:

$$\mathcal{R} = \frac{Z_{02} - Z_{01}}{Z_{02} + Z_{01}}, \quad \mathcal{T} = \frac{2Z_{02}}{Z_{02} + Z_{01}}, \tag{4.5}$$

а коэффициент прохождения по мощности $\mathcal{T}_P$:

$$\mathcal{T}_P = 1 - |\mathcal{R}|^2. \tag{4.6}$$

Так как объемные акустические волны не имеют частотной дисперсии, акустический импеданс, а вместе с ним коэффициенты прохождения и отражения не зависят от частоты и могут использоваться для сигналов произвольного спектрального состава.

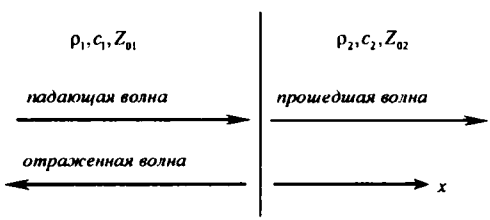


Рис. 4.2. Отражение объемной упругой волны при нормальном падении на границу раздела сред

Очевидно, что для эффективной передачи упругих колебаний из одной среды в другую необходимо, чтобы параметр $\mathcal{R}$ принимал по возможности наименьшее значение. Согласование сред наступает при условии $Z_{01} = Z_{02}$. Для минимизации отражения от границы раздела сред, имеющих значительно отличающиеся импедансы, иногда используют промежуточные четвертьволновые согласующие слои (рис. 4.3). Их действие вполне аналогично действию «просветляющих» слоев, используемых в оптике для уменьшения отражения света от поверхности прозрачных сред.

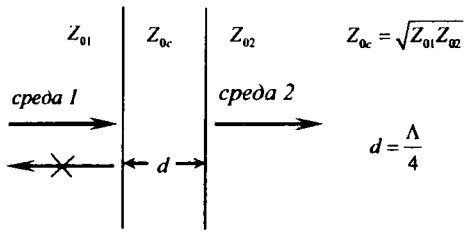


Рис. 4.3. Согласование акустических импедансов сред с помощью $\Lambda/4$ -слоя

Входной импеданс слоя с толщиной d, размещенного на поверхности полуограниченного полупространства, определяется выражением

$$Z_{in} = Z_{0c} \frac{Z_{02} + jZ_{0c} \operatorname{tg} \kappa_c d}{Z_{0c} + jZ_{02} \operatorname{tg} \kappa_c d}, \qquad (4.7)$$

где Z_{02} – импеданс полупространства, Z_{0c} – импеданс материала слоя, κ_c – постоянная распространения акустической волны в слое. Когда толщина слоя равна четверти длины акустической волны $d = \Lambda/4$, то есть $\kappa_c d = \pi/2$, то

$$Z_{in} = \frac{Z_{0c}^2}{Z_{02}}.$$

Если выбрать $Z_{0c} = \sqrt{Z_{01} Z_{02}}$, то $Z_{in} = Z_{01}$ и отражение на границе среды 1 и слоя будет отсутствовать. Таким образом, выбирая материал слоя и его толщину, можно обеспечить приблизительное согласование двух сред с весьма сильно различающимися импедансами, хотя полное согласование возможно только на одной частоте. Нетрудно показать, что чем больше отличие импедансов сред 1 и 2, тем уже полоса согласования. Для расширения полосы частот существуют методы согласования с использованием нескольких четвертьволновых слоев.

Рассмотрим теперь физическую модель преобразователя или резонатора, представляющую собой плоскопараллельную пьезоэлектрическую пластину с тонкими металлическими электродами на обеих поверхностях (рис. 4.4). Пусть пластина ортогональна оси x, а ее поперечные размеры значительно больше толщины и длины акустической волны. В этом случае можно считать, что все величины, описывающие упругую волну вдоль оси x – механическое напряжение T, смещение u, электрическое поле E и индукция D, не зависят от поперечных координат. Очевидно также, что следует рассматривать только волны, порождающие продольное электрическое поле, так как поперечная компонента электрического поля на поверхности электродов равна нулю и пьезоэлектрическая связь между напряжением на электродах и поперечным электрическим полем акустической волны будет отсутствовать.

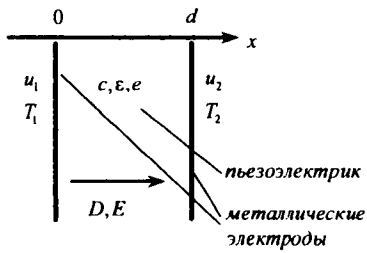


Рис. 4.4. Физическая модель преобразователя

Одномерные уравнения, описывающие волны в пьезоэлектрике, имеют вид

$$\frac{\partial T}{\partial x} = \rho \frac{\partial^2 u}{\partial t^2} \quad \text{(уравнение движения)}, \tag{4.8}$$

$$\begin{aligned} T &= c\frac{\partial u}{\partial x} - eE \\ D &= \varepsilon E + e\frac{\partial u}{\partial x} \end{aligned} \quad \text{(уравнения пьезоэффекта)}, \tag{4.9}$$

$$\frac{\partial D}{\partial x} = 0 \quad \text{(отсутствие в среде свободных зарядов)}. \tag{4.10}$$

Здесь ρ – плотность материала, c, e и ε – соответственно модуль упругости, пьезоэлектрическая константа и диэлектрическая проницаемость пьезоэлектрика для выбранного типа волны. Из второго уравнения (4.9) выразим электрическое поле:

$$E = \frac{D}{\varepsilon} - \frac{e}{\varepsilon}\frac{\partial u}{\partial x}. \tag{4.11}$$

Подставим это выражение в первое из уравнений (4.9):

$$T = c'\frac{\partial u}{\partial x} - \frac{e}{\varepsilon}D, \tag{4.12}$$

где $c' = c\left(1 + K^2\right)$ – «ужесточенный» модуль упругости, $K^2 = e^2/\varepsilon c$ – квадрат коэффициента электромеханической связи. Дифференцируя (4.12) по x с учетом (4.10) и подставляя в уравнение движения (4.8), получаем волновое уравнение

$$\frac{\partial^2 u}{\partial x^2} = \frac{\rho}{c'}\frac{\partial^2 u}{\partial t^2}. \tag{4.13}$$

Далее будем рассматривать монохроматические волны, для которых T, u, E и D пропорциональны множителю $\exp(j\omega t)$. Общим решением уравнения (4.13) будет линейная комбинация волн, распространяющихся в противоположных направлениях:

$$u = a\exp(j\kappa x) + b\exp(-j\kappa x), \tag{4.14}$$

где $\kappa = \omega\sqrt{\rho/c'}$ – постоянная распространения в пьезоэлектрической среде. Здесь и далее для сокращения записи мы опускаем общий множитель $\exp(j\omega t)$. Неизвестные константы a и b найдутся из граничных условий на поверхности преобразователя:

$$u(0) = u_1 \text{ и } u(d) = u_2, \tag{4.15}$$

где u_1, u_2 – смещения при $x = 0$ и $x = d$. Подставляя решение (4.14) в граничные условия (4.15), находим неизвестные a и b. В результате выражение для смещения приобретает следующий вид:

$$u = u_1(\cos\kappa x - \operatorname{ctg}\kappa d\sin\kappa x) + u_2\csc\kappa d\sin\kappa x. \tag{4.16}$$

Напряжение на электродах и плотность тока через преобразователь определятся как

$$V = \int_0^d E dx, \quad J = \frac{\partial D}{\partial t} = j\omega D. \tag{4.17}$$

Подставляя (4.16) в (4.11), (4.12) и (4.17) с учетом (4.10) и обозначая механические напряжения при $x = 0$ и $x = d$ как T_1 и T_2, получаем следующую систему уравнений:

$$\begin{aligned} T_1 &= -u_1\omega Z_{0c}\operatorname{ctg}\kappa d + u_2\omega Z_{0c}\csc\kappa d - \frac{e}{j\omega\varepsilon}J, \\ T_2 &= -u_1\omega Z_{0c}\csc\kappa d + u_2\omega Z_{0c}\operatorname{ctg}\kappa d - \frac{e}{j\omega\varepsilon}J, \\ V &= \frac{e}{\varepsilon}u_1 - \frac{e}{\varepsilon}u_2 + \frac{d}{j\omega\varepsilon}J, \end{aligned} \tag{4.18}$$

где $Z_{0c} = \sqrt{\rho c'}$ – акустический импеданс пьезоэлектрика.

Для анализа преобразователя удобно представить его в виде эквивалентного шестиполюсника, имеющего два акустических входа и один электрический (рис. 4.5).

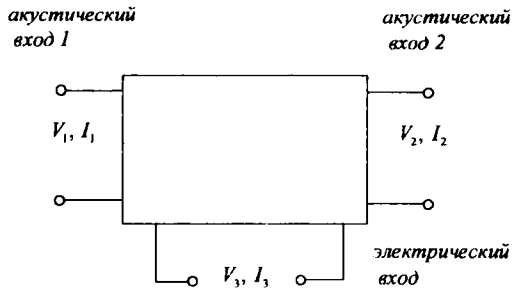


Рис. 4.5. Пьезоэлектрический преобразователь как шестиполюсник

На акустических входах 1 и 2 действуют акустические напряжения $V_1 = -T_1$ и $V_2 = -T_2$, а соответствующие акустические токи равны $I_1 = v_1 = j\omega u_1$ и $I_2 = v_2 = j\omega u_2$. На входе 3 электрическое напряжение $V_3 = V$, а ток $I_3 = JS$, где S – площадь преобразователя. Теперь (4.18) приобретает следующий вид:

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} -jZ_{0c}\,\mathrm{ctg}\,\kappa d & jZ_{0c}\,\mathrm{csc}\,\kappa d & e/j\omega dC_0 \\ -jZ_{0c}\,\mathrm{csc}\,\kappa d & jZ_{0c}\,\mathrm{ctg}\,\kappa d & e/j\omega dC_0 \\ e/j\omega\varepsilon & -e/j\omega\varepsilon & 1/j\omega C_0 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix}, \qquad (4.19)$$

где $C_0 = \varepsilon S/d$ – электрическая емкость преобразователя при деформации, равной нулю.

Система уравнений (4.19), в принципе, позволяет рассчитать все свойства преобразователя в любом режиме работы. Пусть, например, преобразователь работает в режиме приема акустического сигнала, поступающего на вход 1. Зададим импеданс нагрузки на электрическом входе $V_3/I_3 = Z_3$. Если считать, что электроды не создают механической нагрузки, акустический импеданс на входе 2 для волн, прошедших через преобразователь и уходящих от него, равен

$Z_2 = V_2/I_2 = Z_{02}$. Тогда из (4.19) можно выразить входной импеданс $Z_1 = V_1/I_1$ через Z_2, Z_3 и определить коэффициент отражения волн от входа 1. Далее, задавая амплитуду падающей волны, можно найти амплитуду отраженной волны, напряжение и ток в электрической нагрузке и амплитуду акустической волны, прошедшей через преобразователь в среду 2.

Найдем электрический импеданс преобразователя в режиме излучения (рис. 4.6).

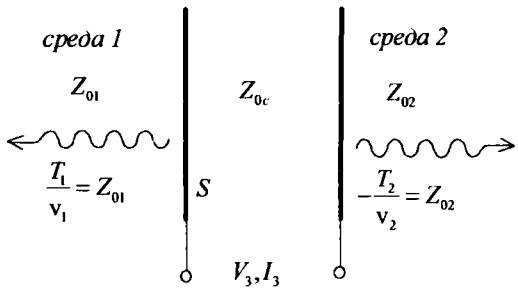


Рис. 4.6. Преобразователь в режиме излучения

Предположим, что прилегающие к преобразователю среды имеют акустические импедансы Z_{01} и Z_{02}, а металлические электроды достаточно тонкие и не создают механической нагрузки. Тогда для волны, возбуждаемой в среде 2,

$$V_2 = Z_{02} I_2 . \tag{4.20}$$

В среде 1 волна распространяется в обратном направлении, поэтому

$$V_1 = -Z_{01} I_1 . \tag{4.21}$$

Подставляя (4.20) и (4.21) в (4.19), после необходимых преобразований находим электрический импеданс преобразователя:

$$Z_3 = \frac{V_3}{I_3} = \frac{1}{j\omega C_0}\Bigg[1 + \\ + \frac{K^2}{\left(1+K^2\right)}\frac{1}{\kappa d}\frac{j\left(Z_{01}+Z_{02}\right)Z_{0c}\sin\kappa d - 2Z_{0c}^2\left(1-\cos\kappa d\right)}{\left(Z_{0c}^2 + Z_{01}Z_{02}\right)\sin\kappa d - j\left(Z_{01}+Z_{02}\right)Z_{0c}\cos\kappa d}\Bigg]. \quad (4.22)$$

Данное выражение можно также записать в следующем виде:

$$Z_3 = \frac{1}{j\omega C_0} + Z_a, \\ Z_a = \frac{1}{j\omega C_0}\frac{K^2}{\left(1+K^2\right)}\frac{1}{\kappa d}\frac{j\left(Z_{01}+Z_{02}\right)Z_{0c}\sin\kappa d - 2Z_{0c}^2\left(1-\cos\kappa d\right)}{\left(Z_{0c}^2 + Z_{01}Z_{02}\right)\sin\kappa d - j\left(Z_{01}+Z_{02}\right)Z_{0c}\cos\kappa d}. \quad (4.23)$$

Величина Z_a описывает акустический вклад в электрический импеданс. Вещественную часть Z_a называют сопротивлением излучения $R_a = \mathrm{Re}\, Z_a$.

Полное количественное исследование зависимости электрического импеданса преобразователя от параметров прилегающих сред и частоты возбуждения может быть выполнено только численными методами. Однако для многих практически важных случаев возможен качественный анализ, основанный на эквивалентных схемах преобразователя, которые строятся на основе различных приближений и допущений.

Рассмотрим пьезоэлектрический резонатор.

В этом случае акустическая нагрузка со стороны окружающей среды практически отсутствует. Полагая $Z_{01} = Z_{02} = 0$, из (4.22) получаем

$$Z_3 = \frac{1}{j\omega C_0}\left[1 - \frac{K^2}{1+K^2}\frac{\mathrm{tg}\left(\kappa d/2\right)}{\left(\kappa d/2\right)}\right]. \quad (4.24)$$

На частотах, при которых на толщине резонатора укладывается нечетное число полуволн (то есть при $\kappa d = \left(2m+1\right)\pi$, $m = 0,1,\ldots$), элек-

трический импеданс становится бесконечно большим. Соответствующие резонансные частоты определятся выражением

$$\omega_{0m} = \frac{\pi v}{d}(2m+1) = \omega_0(2m+1),\ (m = 0,\ 1,\ \ldots), \tag{4.25}$$

где $v = \sqrt{c(1+K^2)/\rho}$ – скорость акустической волны, ω_0 – частота основного резонанса $(m=0)$. Механические резонансы пластины могут иметь место и на частотах, при которых на ее толщине укладывается четное число полуволн. Но электрические поля, связанные с четными модами, будут иметь симметричное распределение относительно центра резонатора, и разность потенциалов между электродами на этих частотах равна нулю, то есть электрическая связь с четными модами отсутствует. Если подставить $\kappa d = 2m\pi$ в (4.23), то акустический вклад в импеданс $Z_a = 0$.

Используя соотношение (4.25), можно записать:

$$\kappa d = \frac{\omega d}{v} = \omega\frac{(2m+1)\pi}{\omega_{0m}} = \frac{\pi\omega}{\omega_0},$$

$$Z_3 = \frac{1}{j\omega C_0}\left[1 - \frac{K^2}{1+K^2}\frac{\mathrm{tg}(\pi\omega/2\omega_0)}{(\pi\omega/2\omega_0)}\right]. \tag{4.26}$$

С помощью известного разложения

$$\mathrm{tg}\,x = \sum_{m=0}^{\infty}\frac{2x}{\left[(2m+1)\pi/2\right]^2 - x^2}$$

выражению (4.26) можно придать следующий вид:

$$Z_3 = \frac{1}{j\omega C_0}\left(1 - \sum_{m=0}^{\infty}\frac{K_m^2\omega_{0m}^2}{\omega_{0m}^2 - \omega^2}\right), \tag{4.27}$$

где введен эффективный коэффициент электромеханической связи m-й моды K_m:

$$K_m^2 = \frac{K_{eff}^2}{(1+2m)^2},$$
$$K_{eff}^2 = K_0^2 = \frac{8}{\pi^2}\frac{K^2}{1+K^2}. \qquad (4.28)$$

С ростом номера моды эффективный коэффициент электромеханической связи K_m уменьшается. Выражение (4.27) преобразуем следующим образом:

$$Z_3 = \frac{1}{j\omega C_0}\left[1 - \sum_{m=0}^{\infty} K_m^2 + \sum_{m=0}^{\infty} K_m^2\left(1 - \frac{\omega_{0m}^2}{\omega_{0m}^2 - \omega^2}\right)\right] =$$
$$= \frac{1}{j\omega C_0\left(1+K^2\right)} + \sum_{m=0}^{\infty}\frac{j\omega K_m^2}{C_0\left(\omega_{0m}^2 - \omega^2\right)}. \qquad (4.29)$$

Считая электромеханическую связь слабой, и рассматривая частоты вблизи m-го механического резонанса, можно записать:

$$Z_3 \cong \frac{1}{j\omega C_0} + \frac{j\omega K_m^2}{C_0\left(\omega_{0m}^2 - \omega^2\right)}. \qquad (4.30)$$

В такой форме импеданс точно соответствует эквивалентной схеме на рис. 4.7, а.

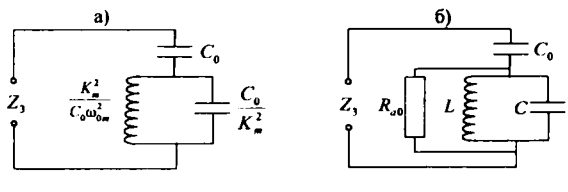


Рис. 4.7. Эквивалентная схема резонатора: а – в отсутствии потерь ($Z_{01}, Z_{02} = 0$); б – при ненулевых значениях акустических нагрузок ($Z_{01}, Z_{02} \neq 0$)

Параллельный колебательный контур с эффективными значениями емкости $C = C_0/K_m^2$ и индуктивности $L = K_m^2/\left(C_0\omega_{0m}^2\right)$ соединен последовательно с емкостью C_0. В данной эквивалентной схеме имеется электрический параллельный резонанс на частоте ω_{0m}, когда импеданс резонатора становится бесконечно большим, и последовательный резонанс на частоте ω_{1m}, при котором импеданс резонатора обращается в ноль. Отношение этих частот выражается соотношением

$$\omega_{1m}/\omega_0 = \left(1 + K_m^2\right)^{-1/2}.$$

Оценим влияние акустических нагрузок Z_{01} и Z_{02} на частоте параллельного резонанса. Из (4.23) для m-го резонанса $\left(\kappa d = (2m+1)\pi\right)$ получаем

$$R_{a0} = \frac{\pi K_{eff}^2}{2\omega_{0m} C_0 (2m+1)} \frac{Z_{0c}}{Z_{01} + Z_{02}}. \qquad (4.31)$$

На эквивалентной схеме, учитывающей ненулевые значения акустических нагрузок (рис. 4.7, б), резистор R_{a0} подключен последовательно с емкостью C_0 и параллельно колебательному контуру. Добротность резонатора

$$Q_a = \omega_{0m} R_{a0} C = \frac{\pi(2m+1)}{2} \frac{Z_{0c}}{Z_{01} + Z_{02}} \qquad (4.32)$$

с увеличением номера моды возрастает, но импеданс резонатора на резонансной частоте уменьшается $\left(R_a \sim (2m+1)^{-2}\right)$.

Использование пьезоэлектрических резонаторов ограничивается в основном частотами до 50 МГц. На частоте 50 МГц толщина кварцевой пластинки для основной моды должна составлять ~ 35 мкм. Изготовление столь тонких плоскопараллельных пластинок – слож-

ная технологическая задача. Поэтому на частотах более 1 МГц, как правило, используют более толстые резонаторы, работающие на третьей или пятой модах колебаний. Резонатор также не может быть слишком толстым, а номер моды слишком большим, так как из-за падения импеданса резонатора его резонансные свойства проявляются слабее.

Рассмотрим далее преобразователь для широкополосного возбуждения объемных акустических волн.

Обычно для улучшения эффективности преобразования электрического сигнала в акустический преобразователь подключается к источнику электрических колебаний через последовательную подстроечную индуктивность L_0, компенсирующую статическую емкость C_0 на центральная частоте (рис. 4.8). Наибольшая мощность, рассеиваемая преобразователем в акустической нагрузке на центральной частоте, достигается при выполнении условия равенства импеданса преобразователя R_{a0} внутреннему сопротивлению генератора R_g.

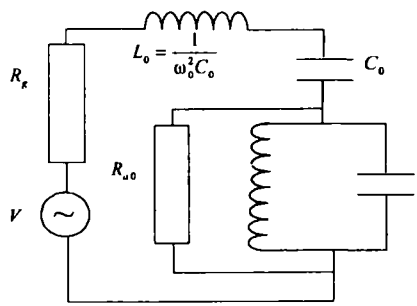


Рис. 4.8. Эквивалентная схема нагруженного преобразователя, подключенного к генератору через согласующую индуктивность: R_g – внутреннее сопротивление генератора, L_0 – согласующая индуктивность

Полоса частот преобразования ограничивается двумя факторами. Во-первых, тем, что сопротивление излучения R_a изменяется с частотой вследствие акустических свойств преобразователя. Связанная с этим фактором полоса хаоактеризуется акустической добротностью преобразователя Q_a, которая определяется как $Q_a = \omega_0/\Delta\omega_a$,

где ω_0 – центральная частота, $\Delta\omega_a$ – ширина акустической частотной характеристики по уровню 3 дБ. Для оценки акустической добротности и полосы можно воспользоваться соотношением (4.32). Вторая причина связана с тем, что схема компенсации емкости преобразователя C_0 имеет эффективную добротность, называемую электрической добротностью Q_e. Она определяет электрическую полосу преобразователя $\Delta\omega_e/\omega_0 = 1/Q_e$. Вблизи частоты ω_0 реактивное сопротивление параллельного контура велико, и величина Q_e определится как добротность последовательного колебательного контура из емкости C_0, компенсирующей индуктивности L_0, сопротивления излучения R_{a0} и внутреннего сопротивления генератора R_g:

$$Q_e = \frac{1}{\omega_0 C_0 \left(R_{a0} + R_g\right)} = \frac{2}{\pi K_{eff}^2} \frac{R_{a0}}{R_{a0} + R_g} \frac{Z_{01} + Z_{02}}{Z_{0c}}. \qquad (4.33)$$

Эффективная полоса преобразователя будет определяться большей из величин Q_a и Q_e. Из соотношений (4.32) и (4.33) легко видеть, что при малых Z_{01} и Z_{02} эффективную полосу преобразования будет ограничивать акустическая добротность, а при больших Z_{01} и Z_{02} – электрическая. Следовательно, оптимальная полоса преобразования достигается при условии $Q_e = Q_a$. Данное условие выполняется, если

$$\sqrt{\frac{R_{a0} + R_g}{R_{a0}}} = \frac{2}{\pi K_{eff}} \frac{Z_{01} + Z_{02}}{Z_{0c}}. \qquad (4.34)$$

При этом полоса преобразования по уровню 3 дБ равна:

$$\frac{\Delta\omega}{\omega_0} = \frac{2}{\pi} \frac{Z_{01} + Z_{02}}{Z_{0c}}. \qquad (4.35)$$

Рассмотрим преобразователь с воздушной тыльной нагрузкой $\left(Z_{01} = 0\right)$, возбуждающий волну в звукопроводе с импедансом Z_{02}

(рис. 4.9, а). Найдем оптимальное значение отношения Z_{02}/Z_{0c} для преобразователя, электрически согласованного с генератором $\left(R_{a0} = R_g\right)$. Из (4.34) получаем

$$Z_{02}/Z_{0c} = \pi K_{eff}/\sqrt{2}\,.$$

Для ниобата лития $K_{eff} \approx 0.22$, откуда $Z_{02}/Z_{0c} \approx 0.5$, а полоса преобразования, определяемая соотношением (3.35), составляет примерно 32 %. Применение пьезокерамик с высоким коэффициентом электромеханической связи позволяет увеличить Z_{02}/Z_{0c} и довести полосу преобразования до октавы (≈ 66 %) при сохранении электрического согласования с генератором. Если импеданс материала звукопровода Z_{02} не соответствует оптимальному значению, можно использовать четвертьволновые согласующие слои. Однако, как уже указывалось, эти слои обеспечивают согласование только в ограниченной полосе частот.

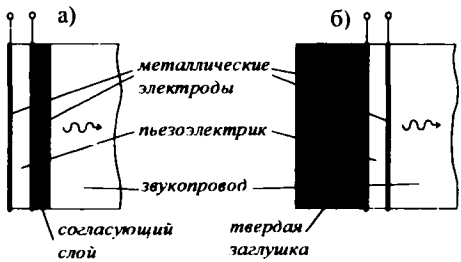


Рис. 4.9. Преобразователь с воздушной тыльной нагрузкой и согласующим слоем (а), с жесткой тыльной нагрузкой (б)

Существенного расширения полосы преобразования можно добиться, применяя на тыльной поверхности жесткую нагрузку с импедансом $Z_{01} > Z_{0c}, Z_{02}$ (рис. 4.9, б). Однако при этом преобразователь уже не будет электрически согласован с генератором и эффективность преобразования резко снижается. Пусть, например, $Z_{01} + Z_{02} = 2Z_{0c}$. В этом случае полоса преобразования составляет более двух октав. Однако требуемое отношение внутреннего сопро-

тивления генератора к сопротивлению преобразователя при $K_{eff} \approx 0.22$, как следует из (4.34), должно быть более 30.

4.2. Возбуждение и прием поверхностных акустических волн

Для возбуждения поверхностных акустических волн (ПАВ) необходимо создать на поверхности упругого твердого тела периодическое во времени и пространстве механическое воздействие, структура которого соответствовала бы выбранному типу волны. В настоящее время известно большое число разнообразных методов возбуждения ПАВ. Наиболее распространенными являются два типа: механические методы, использующие пьезоэлектрические преобразователи специальной формы, и электрические – на основе металлических тонкопленочных структур, размещенных на поверхности пьезоэлектрической подложки.

Для возбуждения ПАВ первым из названных методов используют в основном два вида преобразователей – клиновидный и гребенчатый.

В клиновидном преобразователе (рис. 4.10, а) колебания пьезоэлектрической пластинки под действием приложенного переменного напряжения возбуждают в клине упругую продольную волну с длиной Λ_L, которая, достигая подложки, создает на ее поверхности периодическое во времени и по координате механическое воздействие с пространственным периодом $\Lambda_L/\sin\theta$. Угол θ определяется соотношением между скоростью продольной волны в клине v_L и скоростью рэлеевской волны в подложке v_R. Переход в волну Рэлея будет эффективным, если угол θ выбран таким, чтобы пространственный период возмущения вдоль поверхности подложки был равен длине рэлеевской волны, то есть $\sin\theta = \Lambda_L/\Lambda_R = v_L/v_R$. Откуда следует, что скорость v_L в клине должна быть меньше скорости v_R в подложке. Поскольку скорости объемных волн в изотропных телах всегда больше скорости рэлеевской волны, объемные волны на границе раздела испытывают полное внутреннее отражение и не проникают в подложку, а неод-

нородная волна, распространяющаяся параллельно поверхности и убывающая вглубь подложки, возбуждает волну Рэлея.

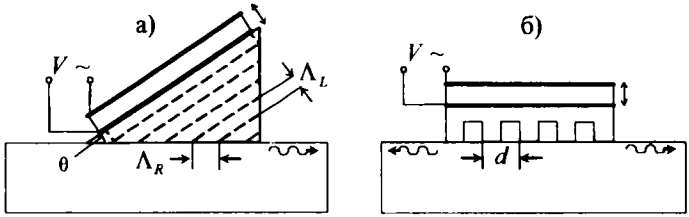


Рис. 4.10. Механические преобразователи волн Рэлея: а – возбуждение продольной волной в клине; б – возбуждение с помощью металлической гребенки

Преимущества клиновидного преобразователя состоят в том, что при всех частотах угол клина θ зависит только от скоростей продольных волн в клине и волны Рэлея в подложке, а также в том, что поверхностные волны излучаются только в одном направлении. Недостаток связан с требованием $v_L < v_R$. Это приводит к необходимости изготовления клина из материалов с низкой скоростью объемных волн, например из пластмасс, в которых скорость продольного звука достаточно низка. Но в таких материалах велико поглощение ультразвуковых волн, особенно на высоких частотах.

Второй возможный вид преобразователя для механического возбуждения ПАВ представлен на рис. 4.10, б. Металлическая гребенчатая структура с периодически расположенными параллельными выступами прижимается к поверхности подложки. На ее противоположной стороне размещается пьезоэлектрическая пластинка. Равномерное движение системы выступов под действием колебаний пластинки приводит к возбуждению волны Рэлея, распространяющейся в обе стороны от преобразователя. Наилучшее возбуждение рэлеевской волны будет происходить при условии равенства шага гребенки d длине волны Λ_R. С увеличением числа выступов гребенки эффективность преобразования увеличивается, но полоса преобразователя уменьшается. В дополнение к рэлеевской волне заметная доля энергии излучается в виде объемной продольной волны, уходящей от поверхности.

Максимальная частота, на которой волны Рэлея успешно возбуждаются с помощью рассмотренных механических методов, не превышает 30–50 МГц. На более высоких частотах клиновидный преоб-

разователь нельзя использовать из-за сильного поглощения упругих продольных волн в самом клине. Применение гребенчатого преобразователя ограничено, главным образом, технологическими трудностями изготовления гребенки с шагом менее 0.1 мм.

Механическое напряжение, соответствующее структуре ПАВ, можно также создать путем воздействия пространственно периодического электрического поля высокой частоты с помощью металлических электродов, размещенных на поверхности соответствующим образом ориентированного пьезокристалла. В настоящее время такой способ является наиболее используемым на высоких частотах (вплоть до 3 ГГц).

Существуют две основные конфигурации электродов (рис. 4.11). В первой из них все электроды на верхней поверхности кристалла имеют одинаковый электрический потенциал. Высокочастотный генератор подсоединяется между этими электродами и сплошным электродом на нижней поверхности. Данный преобразователь получил название однофазного (рис. 4.11, а). Преобразователь другой конфигурации состоит из двух вложенных одна в другую и изолированных друг от друга гребенок электродов, к которым присоединяется возбуждающий генератор, так что соседние электроды имеют противоположные потенциалы (рис. 4.11, б). Данную конфигурацию называют двухфазным или встречно-штыревым преобразователем (ВШП). Оба преобразователя с прямой пьезоэлектрической связью можно использовать и для обратного преобразования акустического сигнала в электрический.

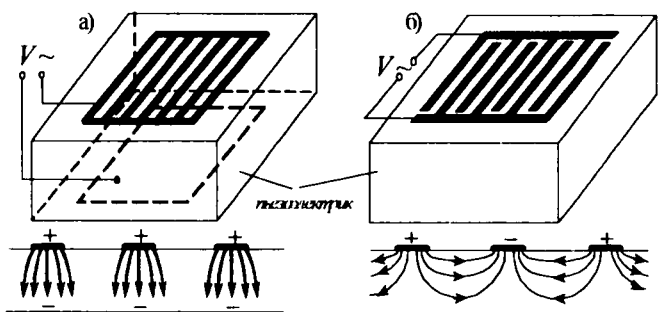


Рис. 4.11. Возбуждение ПАВ с помощью прямой пьезоэлектрической связи с электрическим полем волны: а – однофазная система; б – двухфазная система (ВШП)

Очевидно, для однофазного преобразователя максимальная эффективность связи электрического напряжения на электродах с электрическим полем поверхностной волны будет достигаться при равенстве расстояния между электродами длине ПАВ, а для ВШП – при равенстве расстояния между соседними электродами половине длины ПАВ. Частота, для которой данное условие выполняется, называется частотой акустического синхронизма. Ясно, что при одинаковом шаге электродов частота акустического синхронизма для однофазной системы в два раз выше, чем для ВШП. Это делает применение однофазной системы на высоких частотах более предпочтительным. При малом шаге возможное короткое замыкание между электродами в однофазной системе не влияет на работу преобразователя, а ВШП делает неработоспособным. Кроме того, при значительной интенсивности сигнала в двухфазной системе в отличие от однофазной существует угроза электрического пробоя между электродами.

Тем не менее в акустоэлектронных устройствах на поверхностных волнах, работающих на частотах до 1 ГГц, чаще всего используются встречно-штыревые преобразователи. Применение однофазной системы существенно ограничивается необходимостью использования тыльного электрода на обратной стороне подложки. При заданном потенциале электродов напряженность поля, создаваемая однофазным преобразователем в подложке, обратно пропорциональна ее толщине, в то время как для ВШП напряженность поля определяется расстоянием между соседними электродами преобразователя. Минимальная толщина подложки обычно определяется технологическими причинами и составляет не менее десяти длин волн ПАВ. Так как при заданной частоте акустического синхронизма шаг ВШП в два раза меньше, чем у однофазного преобразователя, можно сделать вывод, что эффективность однофазного преобразователя, по крайней мере, в 20 раз хуже, чем встречно-штыревого. Помимо сказанного, требование доступа к обратной стороне подложки усложняет конструкцию акустоэлектронного устройства.

Электродные преобразователи могут быть также использованы для связи с поверхностными акустическими волнами в непьезоактивных подложках (рис. 4.12). Возможны две конфигурации взаимодействия. В первой из них между электродами и подложкой размещается тонкая пьезоэлектрическая пленка (рис. 4.12, а), во второй – электроды размещены на границе подложки и пьезоэлектрика

(рис. 4.12, б). Соприкасающиеся поверхности подложки и пьезоэлектрика должны быть жестко связаны. Этим обеспечивается передача упругих деформаций между пьезоэлектриком и подложкой.

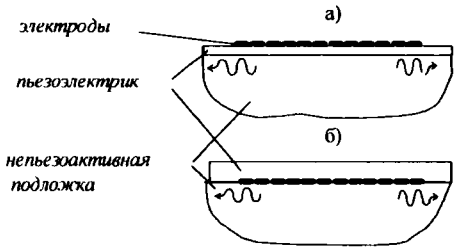


Рис. 4.12. Возбуждение ПАВ металлическими электродами в упругой непьезоактивной среде: а - пьезоэлектрическая пленка между электродами и непьезоактивной средой; б - электроды на границе пьезоэлектрика и непьезоактивной среды

В обеих конфигурациях связь электрического потенциала на электродах с упругими возмущениями в подложке происходит через электрическое поле пьезоэлектрика. Следует иметь в виду, что на частотную характеристику преобразователя будет оказывать влияние дисперсия поверхностных волн, распространяющихся в слоистых средах.

4.3. Встречно-штыревой преобразователь

Как отмечалось в предыдущем разделе, встречно-штыревые преобразователи (ВШП) являются основой большинства акустоэлектронных устройств. Поэтому подробное описание свойств ВШП весьма важно как для анализа этих устройств, так и для их проектирования.

Рассмотрим ВШП на поверхности пьезоэлектрической подложки (рис. 4.13). Почти всегда материал и ориентацию подложки выбирают таким образом, чтобы вдоль преобразователя (ортогонально штырям) мог распространяться только один тип упругой поверхностной волны. Чаще всего единственной распространяющейся модой является жесткая рэлеевская волна. Волны Гуляева–Блюстейна используются значительно реже, так как из-за большой глубины проникновения требуются толстые подложки для исключения связи с нижней поверхностью.

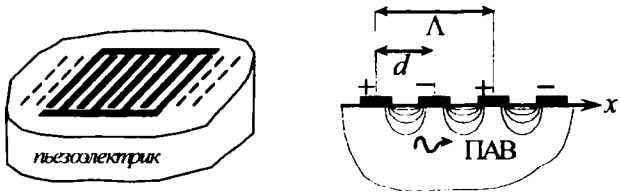


Рис. 4.13. Взаимодействие электродов ВШП с акустическим возмущением в пьезоэлектрической подложке

Электрическое напряжение на всех штыревых электродах, принадлежащих каждой из вложенных решеток, находится в одной и той же фазе, в то время как фронт акустической волны последовательно проходит под каждым электродом с некоторой задержкой, зависящей от скорости волны и расстояния между электродами. Таким образом, штыревые электроды решеток оказываются соединенными электрически параллельно, а акустически – последовательно. На частоте акустического синхронизма $f_0 = v/2d$, когда расстояние между соседними штырями равно половине длины ПАВ $(d = \Lambda/2)$, между электрическим напряжением на штыревых электродах и электрическим полем волны существует достаточно сильная связь, которая позволяет использовать ВШП как источник и приемник ПАВ в различных акустоэлектронных устройствах.

Основными разновидностями ВШП являются: однородный преобразователь с одинаковым расстоянием между штырями и постоянной их длиной; неэквидистантный преобразователь с переменным расстоянием между штырями; аподизованный преобразователь с переменной длиной перекрытия штырей (рис. 4.14). Несмотря на простую конструктивную схему таких преобразователей, даже в случае простейшего однородного преобразователя строгий математический анализ его характеристик довольно сложен. Это связано, во-первых, со сложностью задачи нахождения электрических полей в произвольной системе поверхностных электродов и, во-вторых, с трудностями строгого учета электроакустических взаимодействий в электродной системе преобразователя. Однако для многих практических случаев получены различные приемлемые приближения.

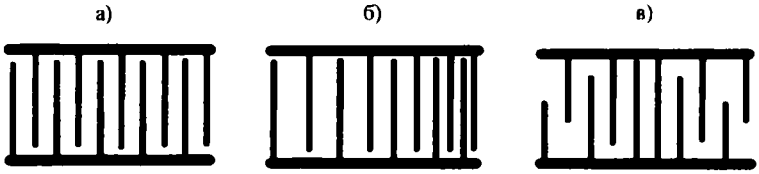


Рис. 4.14. Различные виды ВШП: а – неаподизованный эквидистантный преобразователь; б – неэквидистантный преобразователь; в – эквидистантный аподизованный преобразователь

При рассмотрении возбуждения ПАВ с помощью поверхностных электродов обычно принимают, что электроды имеют нулевую толщину и массу и не влияют на упругие свойства поверхности подложки. Кроме того, полагается, что длина электродов значительно больше расстояния между ними, а электромеханическая связь является слабой. Установлено, что электрический потенциал на поверхности пьезоэлектрического полупространства φ_s связан с поверхностной плотностью заряда $q(x)$ выражениями:

$$\varphi_s(\omega)=\frac{K^2}{2(\varepsilon_p+\varepsilon_0)}q(\omega),$$
$$q(\omega)=\int_{-\infty}^{\infty}q(x)\exp\left(-j\frac{\omega}{v}x\right)dx. \qquad (4.36)$$

Здесь K – эффективный коэффициент электромеханической связи для поверхностной волны (2.78), v – скорость акустической волны в системе электродов на поверхности пьезоэлектрика, $\varepsilon_p=\sqrt{\varepsilon_{33}\varepsilon_{11}-\varepsilon_{13}^2}$ – эффективная диэлектрическая проницаемость, ε_{ik} – компоненты тензора диэлектрической проницаемости пьезоэлектрика, ε_0 – диэлектрическая проницаемость вакуума. Из (4.36) следует, что в отсутствие дисперсии ПАВ частотная зависимость потенциалов электродов $\varphi_s(\omega)$, или, что то же самое, частотная характеристика ВШП $H(\omega)=\varphi_s(\omega)$ определяется исключительно геометрией электродов, определяющей поверхностное распределение зарядов.

Так как электромеханическая связь предполагается слабой, в качестве первого приближения можно принять, что распределение зарядов на электродах такое же, как в отсутствие волны. Тогда неизвестная поверхностная плотность заряда может быть найдена из решения электростатической задачи о распределении электрического поля, создаваемого последовательностью бесконечно длинных тонких проводящих электродов, размещенных на границе диэлектрического полупространства. Эта классическая задача имеет строгое решение. Распределение зарядов на электродах определяет формула Келдыша-Седова:

$$q(x)=\sum_{m=0}^{N-2}\gamma_m x^m \Big/ \sqrt{\prod_{k=1}^{N}(x-a_k)(x-b_k)}\,, \qquad (4.37)$$

где a_k и b_k – координаты начала и конца k–того электрода, N – число электродов, γ_m – некоторые коэффициенты, выбираемые так, чтобы при $a_k < x < b_k$ потенциал был равен φ_k – заданному потенциалу k–того электрода.

Соотношения (4.36) – (4.37) позволяют в принципе найти зависимость потенциалов электродов от параметров подложки и частоты ПАВ и, таким образом, полностью исследовать любой неаподизованный преобразователь с произвольным расположением штырей. Но даже в простейшем случае двухэлектродного преобразователя эти соотношения пригодны только для численных расчетов, а получить простые аналитические выражения для частотной характеристики преобразователя не удается.

Для качественного анализа частотной характеристики используют различные модели распределения зарядов на электродах. Наиболее просто результат получается в модели с одной или двумя дельта-функциями. В упрощенном виде суть данной модели заключается в следующем (рис. 4.15). Реальное распределение зарядов для узких электродов заменяется последовательностью дельта-функций:

$$q(x)=\sum_{n=1}^{N} q_k \delta(x-x_k)\,, \qquad (4.38)$$

где q_k и x_k – заряд и координата середины k–того электрода. В тех случаях, когда нельзя пренебречь шириной электродов, используют модель двух дельта-функций. При этом реальное распределение зарядов на каждом электроде заменяется суммой дельта-функций, расположенных на его краях:

$$q(x)=\frac{1}{2}\sum_{n=1}^{N} q_k\left[\delta(x-a_k)+\delta(x-b_k)\right]. \qquad (4.39)$$

Неизвестные коэффициенты, входящие в (4.38) и (4.39), находятся по потенциалам электродов и их емкости, которую определяют из решения соответствующей электростатической задачи. Несмотря на довольно грубую аппроксимацию распределения зарядов на электродах, модель дельта-функций во многих практических случаях дает приемлемое совпадение с результатами строгой теории.

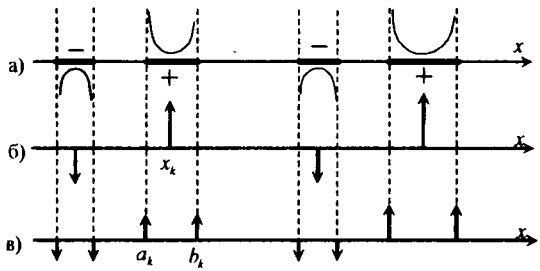


Рис. 4.15. Распределение зарядов на электродах ВШП: а – реальное распределение; б - в модели с одной дельта-функцией; в - в модели с двумя дельта-функциями

Частотная характеристика ВШП в модели с одной дельта-функцией имеет простой вид:

$$H(\omega)=H_0\sum_{k=1}^{N} q_k \exp\left(-j\frac{\omega}{v}x_k\right), \qquad (4.40)$$

где H_0 зависит только от параметров пьезоэлектрика. Более строгий анализ показывает, что H_0 зависит, хотя и довольно слабо от частоты. В дальнейшем для простоты мы этой зависимостью будем пре-

небрегать, полагая $H_0 = 1$. Выражение (4.40) по существу означает, что амплитуда генерируемой ПАВ является суммой амплитуд N независимых элементарных источников волн с частотой приложенного к преобразователю напряжения. Модель с дельта-функциями весьма удобна еще и потому, что с помощью (4.40) позволяет анализировать частотную характеристику аподизованного преобразователя. В этом случае заряды на электродах q_k пропорциональны межэлектродной емкости, то есть степени перекрытия апертур соседних штырей.

Рассмотрим случай возбуждения поверхностной волны с помощью однородного неаподизованного ВШП. Будем считать, что ширина электродов значительно меньше периода решеток и можно использовать модель с одной дельта-функцией. Пусть центр первого электрода преобразователя имеет координату $x_1 = 0$, а расстояние между центрами соседних электродов $-d$, то координата k-того электрода равна $x_k = (k-1)d$. Если пренебречь краевыми эффектами, заряды на соседних электродах однородного преобразователя равны по величине и противоположны по знаку. В результате частотная характеристика N-электродного преобразователя (4.40) с точностью до множителя, не зависящего от частоты, приобретает вид

$$H(\omega) = \sum_{k-1}^{N} c_k \exp\left[-j\frac{\omega}{v}(k-1)d\right], \qquad (4.41)$$

где $c_k = (-1)^k$. Вычисляя сумму геометрической прогрессии (4.41) и отбрасывая фазовый множитель, находим абсолютное значение частотной характеристики – амплитудно-частотную характеристику:

$$|H(f)| = \left|\frac{\sin\left(\frac{N\pi}{2}\frac{f-f_0}{f_0}\right)}{\sin\left(\frac{\pi}{2}\frac{f-f_0}{f_0}\right)}\right|, \qquad (4.42)$$

где $f_0 = 2d/\mathrm{v}$ – частота акустического синхронизма. Из этого соотношения видно, что данная характеристика имеет основной максимум на частоте акустического синхронизма f_0 и побочные максимумы на нечетных гармониках $(2m+1)f_0$, $(m=1,2,...)$ (рис. 4.16). Ограничиваясь областью вблизи основного максимума, можно видеть, что нули $H(f)$, ближайшие к основному максимуму, расположены на частотах $f_0 \pm 2f_0/N$. То есть ВШП работает эффективно в полосе частот, ограниченной значениями $f_0 \pm f_0/N$, при которых амплитуда снижается на 4 дБ по отношению к максимальному значению. Поэтому можно считать, что ширина акустической полосы пропускания

$$\Delta f = 2f_0/N. \tag{4.43}$$

Она характеризуется акустической добротностью преобразователя

$$Q_a = \frac{f_0}{\Delta f} = N_2, \tag{4.44}$$

где $N_2 = 2N$ – число пар штырей.

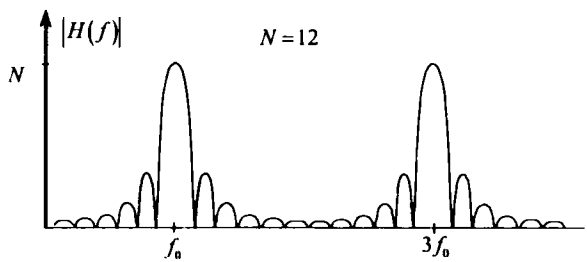


Рис. 4.16. Амплитудно-частотная характеристика однородного 12-электродного ВШП

На частотные свойства преобразователя влияют также его электрические характеристики. Они формируют электрическую полосу преобразователя. Расчет электрических характеристик проще всего

вести на основе последовательной или параллельной эквивалентных схем. Установлено, что наиболее точной является параллельная эквивалентная схема (рис. 4.17). Она содержит статическую емкость преобразователя – C_0, активную $G_a(f)$ и реактивную $B_a(f)$ составляющие акустической проводимости:

$$C_0 = N_2 C_2 W\ ,\ G_a(f) = G_{a0}\frac{\sin p}{p},$$

$$B_a = G_{a0}\frac{\sin 2p - 2p}{2p^2}, \qquad (4.45)$$

$$G_{a0} = 8K^2 f_0 C_2 W N_2^2. \qquad (4.46)$$

В соотношениях (4.45), (4.46) K – коэффициент электромеханической связи, f_0 – частота акустического синхронизма, C_2 – емкость пары штырей на единицу их длины, W – апертура преобразователя, N_2 – число пар штырей, $p = \left[(f - f_0)/f_0\right]\pi N_2$.

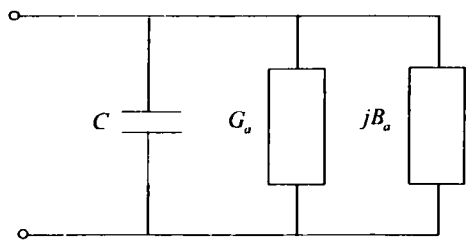


Рис. 4.17. Параллельная эквивалентная схема однородного ВШП

Электрическую полосу ВШП можно рассчитать по эквивалентной схеме включения преобразователя во внешнюю цепь (рис. 4.18). Внешняя цепь содержит генератор с внутренней проводимостью G_0 и согласующую индуктивность L. Согласующая индуктивность предназначена для компенсации статической емкости преобразователя. Если выбрать $G_0 = G_{a0}$, $f_0 L = 1/f_0 C_0$, то на частоте акустического синхронизма преобразователь будет согласован с генератором. Следовательно, условие согласования можно удовлетворить, выби-

рая соответствующим образом количество пар штырей и апертуру преобразователя ПАВ. Добротность электрической схемы, обусловливающая ширину электрической полосы, определяется выражением

$$Q_e = \frac{2\pi f_0 C}{2G_{a0}} = \frac{\pi}{8K^2 N_2}, \tag{4.47}$$

откуда электрическая полоса

$$\Delta f / f_0 = 1/Q_э = \frac{8K^2 N_2}{\pi}. \tag{4.48}$$

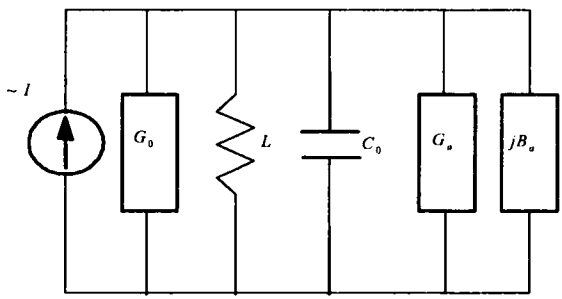


Рис. 4.18. Электрическая эквивалентная схема однородного ВШП в режиме возбуждения: G_0 – внутренняя проводимость генератора, L – согласующая индуктивность, C_0 – статическая емкость преобразователя, G_a, B_a – действительная и мнимая части акустической проводимости

Сопоставляя (4.43), (4.44) и (4.47), (4.48), нетрудно сделать вывод, что полоса пропускания ВШП при большом количестве штырей ограничена акустической полосой, а при малом – его электрической полосой. Максимальная полоса преобразования достигается при равенстве акустической и электрической добротностей, откуда определяется оптимальное число пар электродов преобразователя:

$$N_2 = \sqrt{\pi / 8K^2}. \tag{4.49}$$

Относительная ширина полосы преобразования при оптимальном числе штырей:

$$\frac{\Delta f}{f_0} = \sqrt{\frac{8K^2}{\pi}}.$$

На рис. 4.19 показана зависимость акустической и электрической добротностей от числа штырей преобразователя N для некоторых пьезоэлектрических материалов. Оптимальное число электродов в случае сильных пьезоэлектриков меньше, чем в случае более слабых. Соответственно, для широкополосного возбуждения ПАВ следует выбирать подложки по возможности с большим коэффициентом электромеханической связи. В случае ниобата лития оптимальное число штырей равно шести, а относительная ширина полосы пропускания составляет около 30 %.

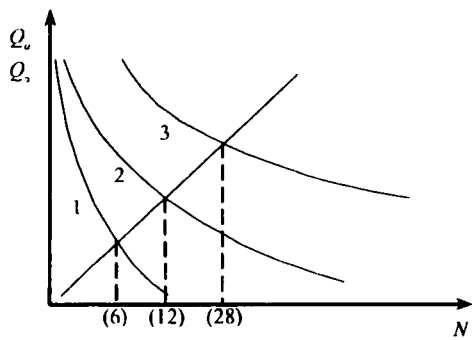


Рис. 4.19. Зависимость акустической и электрической добротности от числа штырей N для различных материалов: 1 – ниобат лития; 2 – германат висмута; 3 – кварц

Рассмотренные выше преобразователи обладают очевидным свойством двунаправленности. При электрическом возбуждении генерируемые поверхностные акустические волны распространяются в обе стороны вдоль оси преобразователя. Если одна из волн является нежелательной, на пути ее распространения размещают поглотитель. При этом энергетические потери на одно преобразование составляют не менее 3 дБ, а при наличии в устройстве 2 преобразователей – не менее 6 дБ. Однако известны конструкции однонаправленных преобразователей. Их работа основана на принципе взаимной компенсации волн, излученных с различными фазами. Примером однона-

правленного преобразователя является преобразователь, показанный на (рис. 4.20). Он состоит из двух одинаковых ВШП, расположенных в общем акустическом потоке на расстоянии $\Lambda(n+1/4)$, где n – произвольное целое число. На частотах вблизи акустического резонанса волны в одном направлении складываются в фазе, в другом – взаимно гасятся. Требуемые фазовые соотношения обеспечиваются при сдвиге фазы между ВШП на 90°. Существенным ограничением применения однонаправленных преобразователей является узкая полоса пропускания, а также усложнение конструкции за счет необходимости включения электромагнитного фазовращателя. Известны также и другие конструкции однонаправленных преобразователей, которые будут рассмотрены ниже.

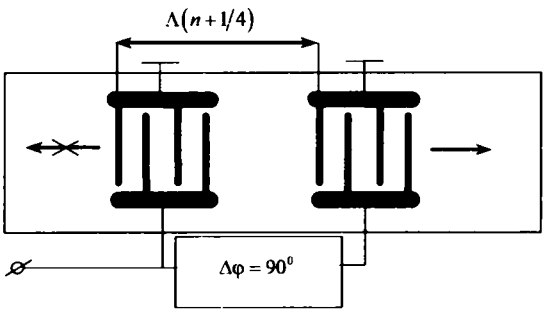


Рис. 4.20. Конструкция двухфазного однонаправленного ВШП

В заключение отметим, что встречно-штыревые преобразователи, помимо поверхностных волн, возбуждают и детектируют также объемные волны, которые в ряде случаев следует учитывать при конструировании акустоэлектронных устройств.

4.4. Многополосковый ответвитель и его применение

Многополосковый ответвитель (МПО) представляет собой систему электрически изолированных друг от друга тонких проводящих электродов, нанесенных на поверхность пьезоэлектрического звукопровода перпендикулярно направлению распространения поверхностной акустической волны. Расстояние между электродами МПО

всегда меньше половины длины поверхностной акустической волны. Так как электрические сигналы распространяются практически мгновенно, МПО позволяет осуществлять перенос акустической энергии в поперечном направлении. Благодаря этому свойству МПО нашли широкое применение при создании различных акустоэлектронных компонентов.

Пусть на поверхности звукопровода имеется два параллельных акустических канала равной ширины, в одном из которых возбуждается поверхностная волна $A_1 \exp(-j\kappa x)$, падающая на МПО, а на входе другого возбуждение отсутствует (рис. 4.21). Поверхностная волна, проходя под электродами МПО, вследствие пьезоэффекта наводит на них электрические потенциалы, которые постоянны вдоль каждого электрода. При этом сами электроды МПО будут служить источниками вторичных акустических волн, излучаемых по всей апертуре МПО. Ввиду того, что в МПО на длину поверхностной волны приходится более двух электродов $(d < \Lambda/2)$, их суммарное излучение является однонаправленным. Это вытекает из того факта, что распределение фаз напряжения на электродах соответствует фазовой последовательности первичной возбуждающей волны, распространяющейся в положительном направлении оси x. В результате на выходе МПО, в общем случае, в каждом из каналов будет присутствовать прошедшая волна.

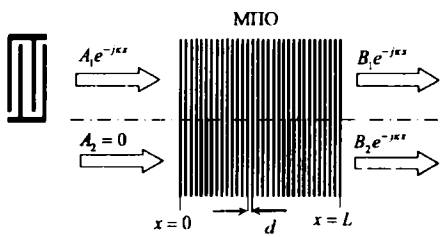


Рис. 4.21. Многополосковый ответвитель для связи между двумя акустическими каналами

Так как амплитуда падающей волны в верхнем канале A_1, а в нижнем канале падающая волна отсутствует $(A_2 = 0)$, входную волну можно представить как результат интерференции двух мод с рав-

ными амплитудами $A_1/2$. Одна из них симметричная, а другая – антисимметричная (рис. 4.22). Из-за влияния электродов, как мы увидим далее, эти две моды имеют несколько различающиеся фазовые скорости. Следовательно, в процессе распространения в МПО между модами появится некоторой сдвиг фаз, зависящий от длины L, и суммарное волновое поле на выходе МПО будет отличаться от суммарного поля на входе. При некоторой длине ответвителя L_T, называемой длиной взаимодействия, разность фаз между модами на выходе МПО равна π. В этом случае суммарное распределение волнового поля на выходе МПО будет давать нулевое значение амплитуды в верхнем канале $B_1 = 0$ и ненулевое значение B_2 – в нижнем канале. Если пренебречь потерями, то очевидно $|B_2| = |A_1|$.

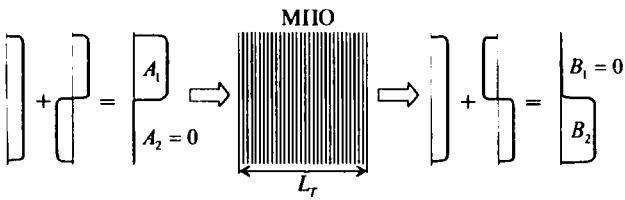


Рис. 4.22. Процесс полной передачи акустической энергии в параллельный канал при длине МПО, равной L_T

Длина взаимодействия, при которой происходит полная передача энергии в параллельный канал, определится из соотношения

$$\left|\frac{\omega}{v_s} - \frac{\omega}{v_a}\right| L_T = \pi , \qquad (4.50)$$

где v_s и v_a – соответственно фазовые скорости симметричной и антисимметричной моды. Обозначая $v_s - v_a = \Delta v$ и предполагая, что относительное различие скоростей мод мало $(\Delta v/v << 1)$, из (4.50) находим

$$L_T = \Lambda\left(\frac{v}{2\Delta v}\right). \qquad (4.51)$$

Разность фазовых скоростей Δv легко оценить в предельном случае, когда пространственный период электродов МПО много меньше длины акустической волны Λ. При этом решетка электродов по существу эквивалентна неоднородно проводящему слою с нулевой проводимостью в направлении распространения волны x и бесконечной проводимостью вдоль электродов. Как мы ранее видели, однородный проводящий слой уменьшает скорость поверхностной волны на величину, определяемую эффективным коэффициентом электромеханической связи:

$$\frac{\Delta v}{v} \cong \frac{K^2}{2}. \qquad (4.52)$$

Этот эффект обусловлен закорачиванием касательного электрического поля на поверхности пьезоэлектрика. Закорачивание электрического поля можно рассматривать как результат перераспределения зарядов, индуцируемых волной в проводнике. В симметричной моде распределение зарядов по длине электрода является однородным. Данной моде соответствуют электрические поля в направлении распространения x и в направлении нормали к поверхности, а вдоль электродов поле отсутствует (если пренебречь краевыми эффектами). Таким образом, в хорошем приближении можно считать, что электроды МПО не влияют на распространение симметричной моды. В случае же антисимметричной моды заряды на электродах в верхнем канале равны и противоположны по знаку зарядам в нижнем канале. Следовательно, распространение антисимметричной моды связано с перераспределением зарядов на электродах. Поэтому можно считать, что скорость антисимметричной моды близка к скорости акустической волны на металлизированной поверхности. На основании данных рассуждений и соотношений (4.51) и (4.52) можно написать следующую оценку длины взаимодействия:

$$L_T \sim \frac{\Lambda}{K^2}.$$

Данная оценка найдена из модели неоднородно проводящего слоя и требует дальнейших уточнений, связанных с конечной шириной электродов и наличием зазоров между ними. Если, как часто бывает на практике, ширина электродов равна половине шага решетки d, то связь одного канала с другим осуществляется в течение

только половины времени распространения мод в МПО. Откуда сразу очевидна необходимость удвоения длины взаимодействия L_T. Если ввести коэффициент металлизации $p = a/d$, где a – ширина электрода, то длину взаимодействия следует увеличить в $1/(1-p) = d/(d-a)$ раз. Кроме того, следует учесть и другой важный эффект, связанный с конечной шириной электродов. Отдельный электрод имеет электрическое напряжение, соответствующее среднему значению амплитуды волны на его ширине. Изменение амплитуды волны по ширине электрода описывается выражением $A(x) = A\exp\left[-j(2\pi/\Lambda)x\right]$. Поэтому средняя амплитуда определится как

$$A(\theta) = \frac{1}{a}\int_{-a/2}^{a/2} A\exp\left[-j(2\pi/\Lambda)x\right]dx = A\frac{\sin\theta}{\theta},$$

где $\theta = \pi a/\Lambda$. В случае бесконечно узких электродов коэффициент $\sin\theta/\theta$ равен единице. При конечной ширине он меньше единицы. Это приводит к некоторому уменьшению эффективной электромеханической связи в процессе преобразования акустического сигнала в электрический в верхнем канале, а также электрического сигнала в акустический в нижнем. В результате длина взаимодействия L_T увеличивается на квадрат этого коэффициента. Данный эффект можно также интерпретировать как частичное закорачивание продольного электрического поля симметричной моды на ширине электрода, приводящее к уменьшению разницы скоростей симметричной и антисимметричной мод. Таким образом, приходим к следующему выражению для длины взаимодействия:

$$L_T = \frac{\Lambda}{K^2}\frac{d}{d-a}\left[\frac{\theta}{\sin\theta}\right]^2. \qquad (4.53)$$

Для ниобата лития значение K^2 приближенно составляет 0.05, откуда при малой ширине электродов МПО получается нижняя оценка длины взаимодействия $L_T \sim 20\Lambda$. Если шаг решетки равен $\Lambda/4$, то требуется 80 электродов. Для материла с низким коэффици-

ентом электромеханической связи, такого как кварц, длина взаимодействия составляет $\sim 500\Lambda$, откуда ясно, что применение МПО оправдано только в случае достаточно сильных пьезоэлектриков.

Исследуем теперь общий случай возбуждения МПО упругой волной, распространяющейся в обоих каналах. Обозначим произвольные амплитуды симметричной и антисимметричной мод как A_s и A_a, а их постоянные распространения $-\kappa_s$ и κ_a. Для амплитуд волн в каждом канале можно записать:

$$A_1(x) = A_s \exp(-j\kappa_s x) + A_a \exp(-j\kappa_a x),$$
$$A_2(x) = A_s \exp(-j\kappa_s x) - A_a \exp(-j\kappa_a x).$$

С помощью несложных преобразований эти соотношения можно привести к следующему виду:

$$A_1(x) = \left[(A_s + A_a)\cos\left(\frac{\kappa_s - \kappa_a}{2} x\right) - \right.$$
$$\left. - j(A_s - A_a)\sin\left(\frac{\kappa_s - \kappa_a}{2} x\right)\right] \exp\left(-j\frac{\kappa_s + \kappa_a}{2} x\right),$$
$$A_2(x) = \left[(A_s - A_a)\cos\left(\frac{\kappa_s - \kappa_a}{2} x\right) - \right.$$
$$\left. - j(A_s + A_a)\sin\left(\frac{\kappa_s - \kappa_a}{2} x\right)\right] \exp\left(-j\frac{\kappa_s + \kappa_a}{2} x\right).$$

Поскольку представляют интерес только амплитуды сигналов в каждом из каналов, а также относительный сдвиг фаз между ними, то одинаковый для обоих каналов фазовый сдвиг $\exp\left[-jx(\kappa_s + \kappa_a)/2\right]$ можно отбросить.

Амплитуды сигнала в каналах на входе ответвителя $(x = 0)$:

$$A_1 = A_1(0) = A_s + A_a,$$
$$A_2 = A_2(0) = A_s - A_a,$$

а на выходе $(x = L)$:

$$B_1 = A_1(L) = A_1 \cos\phi - jA_2 \sin\phi,$$
$$B_2 = A_2(L) = A_2 \cos\phi - jA_1 \sin\phi, \tag{4.54}$$

где

$$\phi = L(\kappa_s - \kappa_a)/2 = \pi L/(2L_T).$$

Соотношения (4.54) можно переписать в матричной форме:

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \tag{4.55}$$

где коэффициенты S_{ik} образуют матрицу рассеяния МПО:

$$[S] = \begin{pmatrix} \cos\phi & -j\sin\phi \\ -j\sin\phi & \cos\phi \end{pmatrix} \tag{4.56}$$

В рассмотренном выше случае наличия входного сигнала только в верхнем канале $(A_2 = 0)$ получаем:

$$B_1 = A_1 \cos\phi,$$
$$B_2 = -jA_1 \sin\phi. \tag{4.57}$$

На рис. 4.23 представлена зависимость выходных амплитуд в каждом из каналов от нормированной длины МПО. Данная зависимость имеет циклический характер. При малых L с ростом длины МПО амплитуда волны в нижнем канале B_2 нарастает, а в верхнем B_1 уменьшается, так как часть энергии тратится на возбуждения нижнего канала. При достижении значения $L = L_T$ $(\phi = \pi/2)$ происходит полная передача энергии из верхнего канала в нижний. В данном случае распределение амплитуды волны на выходе вполне аналогично распределению на входе ответвителя с той, однако, разницей, что теперь возбуждающая волна

присутствует только в нижнем канале, а в верхнем равна нулю. Поэтому по мере дальнейшего увеличения длины амплитуда B_2 теперь будет постепенно уменьшаться, а B_1 – нарастать.

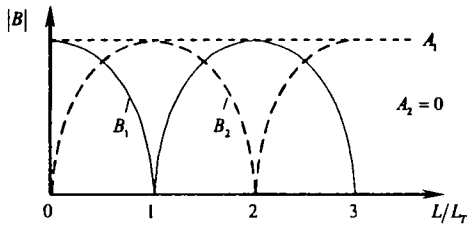


Рис. 4.23. Зависимость амплитуд на выходе каждого из каналов от длины МПО

Проведенный выше анализ не учитывает дифракции волн, которая приводит к некоторой расходимости акустических пучков. В результате между каналами появляется дополнительная паразитная связь, и полной перекачки энергии между каналами не происходит. Для улучшения акустической развязки каналов применяют МПО, конструкция которого показана на рис. 4.24.

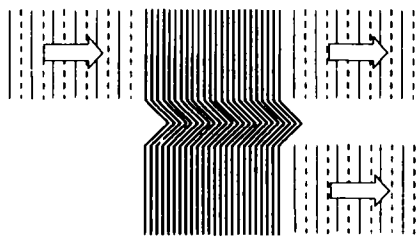


Рис. 4.24. МПО с улучшенной развязкой между акустическими каналами

Здесь, кроме двух основных каналов, имеется центральный относительно узкий канал, в котором электроды расположены под углом таким образом, что шаг, измеренный перпендикулярно электродам, оказывается другим. При этом распределение фаз напряжения на электродах в центральном канале не соответствует фазовой последовательности поверхностной волны с данной частотой, и связь электродов с ПАВ в центральном канале отсутствует.

Большое практическое значение имеет многополосковый ответвитель с равным делением энергии в каналах (3 дБ МПО), когда $\phi = \pi/4$, $L = L_T/2$ и

$$B_1 = A_1/\sqrt{2}, \quad B_2 = -j\,A_1/\sqrt{2}. \tag{4.58}$$

На основе 3 дБ МПО создан ряд устройств, которые рассмотрены ниже.

На рис. 4.25, а показана конструкция U-образного отражателя. По сути, данный отражатель представляет собой 3дБ МПО, но изогнутый. Если длина изогнутых участков ответвителя мала по сравнению с длиной прямых участков, то устройство эквивалентно обычному МПО на рис. 4.25, б.

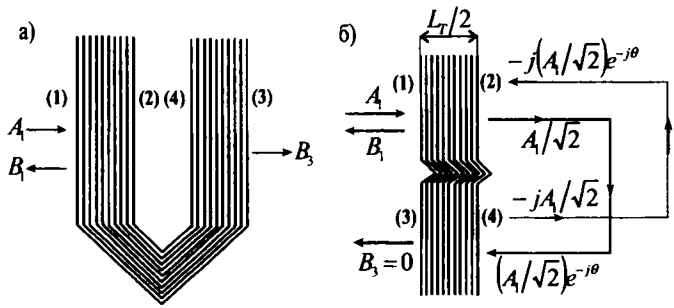


Рис. 4.25. Многополосковый U-образный отражатель (а) и эквивалентная ему структура (б), демонстрирующая прохождение сигналов

Предположим, что входной сигнал A_1 подается только на ввод (1). Выходной сигнал ввода (2), равный $A_1/\sqrt{2}$, в силу изогнутой геометрии поступает на ввод (4) с некоторой фазовой задержкой θ. Аналогично, волна $-j\,A_1/\sqrt{2}$, выходящая из ввода (4), подвергается такой же задержке перед попаданием на ввод (2). Если теперь для выходных сигналов вводов (1) и (3) записать соотношения (4.58), то при наличии этих вторичных сигналов получаем:

$$B_1 = \frac{-j}{\sqrt{2}}\left(e^{-j\theta}\frac{A_1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}\left(-je^{-j\theta}\frac{A_1}{\sqrt{2}}\right) = -je^{-j\theta}A_1,$$
$$B_3 = \frac{1}{\sqrt{2}}\left(e^{-j\theta}\frac{A_1}{\sqrt{2}}\right) - \frac{j}{\sqrt{2}}\left(-j\frac{A_1 e^{-j\theta}}{\sqrt{2}}\right) = 0. \qquad (4.59)$$

Так как $|B_1| = |A_1|$, U-образный МПО действует как идеальный отражатель ПАВ. Однако следует иметь в виду, что данное устройство, строго говоря, обладает качествами идеального отражателя только на одной фиксированной частоте из-за частотной зависимости параметра L_T. Однако в достаточно широкой полосе частот коэффициент отражения близок единице.

U-образный отражатель возвращает энергию в тот же самый канал. Комбинация двух отражателей с 3дБ МПО позволяет создать реверсивный акустический элемент, отражающий волну в параллельный канал. Его устройство показано на рис. 4.26.

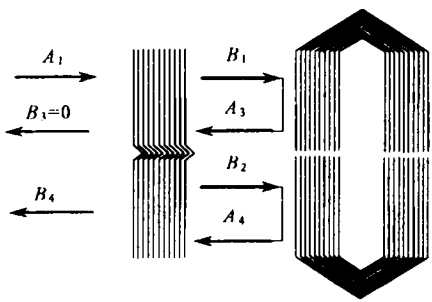


Рис. 4.26. МПО отражатель ПАВ с переводом энергии в параллельный канал

Падающая в верхнем канале волна A_1 с помощью 3 дБ МПО разделяется на две волны с амплитудами B_1 и B_2, которые определяются соотношением (4.58). С помощью (4.59) находим волны, отраженные от U-образных элементов:

$$A_3 = -je^{-j\theta}B_1 = -je^{-j\theta}A_1/\sqrt{2},$$
$$A_4 = -je^{-j\theta}B_2 = -e^{-j\theta}A_1/\sqrt{2}.$$

Теперь, применяя общие соотношения для МПО с входным сигналом в обоих каналах (4.54) при $\phi = \pi/4$, получаем:

$$B_3 = \frac{1}{\sqrt{2}}\left(A_3 - jA_4\right) = \frac{A_1}{2}\left(-je^{-j\theta} + je^{-j\theta}\right) = 0,$$

$$B_4 = \frac{1}{\sqrt{2}}\left(A_4 - jA_3\right) = \frac{A_1}{2}\left(-e^{-j\theta} - e^{-j\theta}\right) = -A_1 e^{-j\theta}.$$

Таким образом, волна полностью отражается, но распространяется по соседнему каналу. Отражающие структуры рассмотренного типа находят применение в линиях длительной задержки электрических сигналов на поверхностных акустических волнах.

U-образный МПО с равным делением энергии может быть также использован для создания однонаправленного преобразователя электрической энергии в акустическую. Для этого ВШП размещается между двумя плечами МПО, как показано на рис. 4.27, а. Если A – амплитуда возбуждаемой ВШП волны, то падающие на вводы (2) и (4) волны равны по амплитуде, но подвергаются разным фазовым задержкам, зависящим от положения преобразователя внутри ответвителя. В эквивалентной структуре на рис. 4.27, б волны, выходящие из вводов (1) и (3), определяются выражениями:

$$B_1 = \frac{A}{\sqrt{2}}\left(e^{-j\theta_1} - je^{-j\theta_2}\right), \quad B_3 = \frac{A}{\sqrt{2}}\left(e^{-j\theta_2} - je^{-j\theta_1}\right).$$

Если выполняется условие $\theta_1 - \theta_2 = \pi/2$, то $B_1 = A\sqrt{2}e^{-j\theta_1}$, $B_3 = 0$, и тем самым достигается однонаправленное излучение. То факт, что при этом выходящая волна имеет амплитуду, в $\sqrt{2}$ раза большую, чем создаваемая ВШП, обусловлен сложением двух волн, первоначально распространяющихся в противоположных направлениях. Полная энергия пропорциональна $\left(A\sqrt{2}\right)^2 = 2A^2$, что получается суммированием двух волн, излучаемых обычным преобразователем в противоположных направлениях.

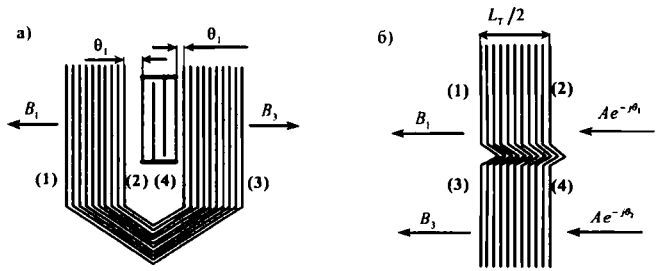


Рис. 4.27. Однонаправленный преобразователь (а) и эквивалентная ему структура (б)

Еще одна модификация основной топологии МПО может быть использована для связи двух акустических каналов разной ширины. Если апертура выходного пучка меньше апертуры входного, амплитуда акустической волны после прохождения МПО возрастает. Ответвители такой структуры используются в конвольверах на ПАВ, где повышением объемной плотности энергии волны за счет сужения пучка повышается эффективность устройства. Асимметричный МПО со сжатием пучка показан на рис. 4.28. Широкий канал с периодом d_1 соединен с узким каналом, в котором период электродов d_2 несколько меньше. Расчет показывает, что для полной передачи энергии из верхнего канала в нижний соотношение периодов решеток в каналах должно удовлетворять условию

$$d_1 - d_2 = \frac{\Lambda(d_1 + d_2)}{4L_T}\frac{W_1 - W_2}{W_1 + W_2}, \qquad (4.60)$$

Если это условие выполнено, то волна полностью перекачивается из одного канала в другой, когда длина МПО равна:

$$L_T' = L_T \frac{W_1 + W_2}{2\sqrt{W_1 W_2}}. \qquad (4.61)$$

В соотношениях (4.60), (4.61) W_1, W_2 – соответственно апертуры входного и выходного пучков, L_T – длина МПО, необходимая для полной перекачки энергии, при равной апертуре пучков. При коэф-

фициенте сжатия $W_1/W_2 = 10$ отличие шагов решеток составляет около 1 %, а требуемая длина МПО примерно равна $1.5L_T$.

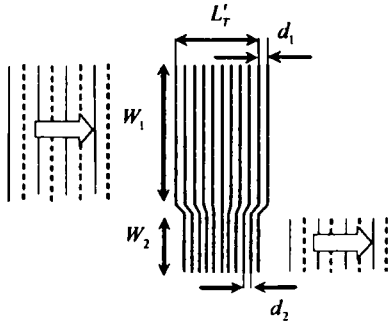


Рис. 4.28. Асимметричный МПО для сжатия акустического пучка

Выше всюду предполагалось, что число электродов МПО на длине волны больше двух, что не приводит к сколько-нибудь заметному отражению ПАВ. Однако когда шаг электродов приближается к половине длины акустической волны $\Lambda/2$, небольшие отражения от каждого электрода складываются в фазе, и при большом их числе появляется полоса заграждения (полоса режекции), которая используется при создании отражательных решеток (ОР) ПАВ. В МПО эта область частот исключается. Обычно устройства, использующие МПО, работают на частотах $(0.3 \div 0.9) f_0$, где $f_0 = \mathrm{v}/2d$ – центральная частота полосы заграждения.

4.5. Отражательные решетки

Отражательная решетка (ОР) представляет систему протяженных периодических неоднородностей, размещенных на поверхности звукопровода перпендикулярно направлению распространения поверхностной волны. Возможные примеры таких неоднородностей представлены на рис. 4.29. Каждая неоднородность в общем случае создает как механическую, так и электрическую нагрузку поверхности, что приводит к некоторому отражению акустической волны. Если

период структуры близок к целому числу полуволн ПАВ, то отраженные волны складываются в фазе и при достаточно большом числе неоднородностей суммарный коэффициент отражения решетки стремится к единице. В дальнейшем мы будем говорить о металлических электродах, хотя получаемые результаты не зависят от конкретного механизма рассеяния отдельной неоднородности и могут быть в полной мере перенесены на другие типы периодических структур.

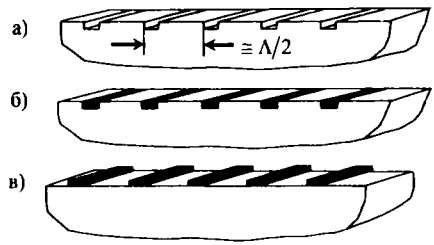


Рис. 4.29. Периодические неоднородности на поверхности звукопровода: а – прямоугольные пазы; б – внедренные металлические электроды; в – наложенные металлические электроды

Рассмотрим бесконечную решетку, состоящую из одинаковых изолированных электродов с шагом d, показанную на рис. 4.30.

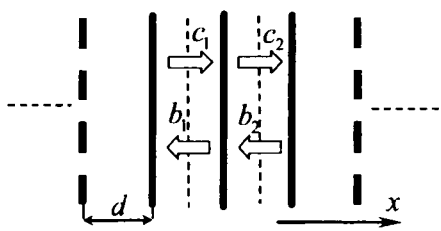


Рис. 4.30. Распространение ПАВ в бесконечной регулярной решетке электродов

Найдем зависимость свойств решетки от характеристик рассеяния отдельного электрода. Будем считать, что на некоторой заданной частоте ω акустическое возмущение в зазорах между электродами содержит множители $\exp(\pm j\kappa x)$, описывающие распространяющиеся волны, где κ – волновое число ПАВ на свободной по-

верхности. Рассматривая отдельный электрод, обозначим амплитуды волн слева от него символами c_1 и b_1, а амплитуды волн справа – символами c_2 и b_2. Указанные амплитуды относятся к точкам, отстоящим на расстоянии $d/2$ от центра электрода.

Уходящие от электрода волны b_1 и c_2 могут быть выражены через падающие волны c_1 и b_2 с помощью матрицы рассеяния S_{ik}:

$$\begin{pmatrix} b_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ b_2 \end{pmatrix}. \tag{4.62}$$

Из симметрии задачи следует, что $S_{22} = S_{11}$ и $S_{12} = S_{21}$. В явном виде

$$\begin{aligned} b_1 &= S_{11}c_1 + S_{12}b_2, \\ c_2 &= S_{12}c_1 + S_{22}b_2. \end{aligned} \tag{4.63}$$

Пусть присутствует только падающая слева волна, то, приравнивая b_2 нулю, из (4.63) получаем

$$b_1 = S_{11}c_1 \text{ и } c_2 = S_{12}c_1. \tag{4.64}$$

Таким образом, S_{11} определяет амплитуду отраженной волны, а S_{12} – амплитуду прошедшей. Пусть r – коэффициент отражения электрода, отнесенный к его центру, а t – его коэффициент прохождения. Учитывая, что сечения, в которых определены амплитуды, отстоят от центра электрода на $d/2$, можно записать:

$$\begin{aligned} S_{11} &= r\exp(-j\kappa d), \\ S_{12} &= t\exp(-j\kappa d). \end{aligned} \tag{4.65}$$

Найдем связь коэффициентов отражения и прохождения. Если считать, что потери отсутствуют, то энергия на входе равна энергии на выходе.

$$|c_1|^2 + |b_2|^2 = |c_2|^2 + |b_1|^2 . \qquad (4.66)$$

Взяв $b_2 = 0$ и подставляя (4.64) – (4.65) в (4.66), получаем

$$|r|^2 + |t|^2 = 1 . \qquad (4.67)$$

Если все четыре амплитуды b_1, b_2, c_1, c_2 не равны нулю, то из (4.63) и (4.66) следует

$$|c_1|^2 + |b_2|^2 = |c_1|^2 + |b_2|^2 + 2|c_1||b_2|\left(r^* t + t^* r\right), \qquad (4.68)$$

Откуда

$$r^* t + t^* r = 0 . \qquad (4.69)$$

Объединяя равенства (4.67) и (4.69), получаем

$$r^2 = t^2 - t/t^* . \qquad 4.70)$$

Вновь используя (4.67), в итоге находим

$$r/t = \pm j|r/t| . \qquad (4.71)$$

Если не учитывать неоднозначности фазы, кратной π, то с помощью данного соотношения коэффициент отражения может быть найден исходя из коэффициента прохождения.

Для дальнейших целей удобно выразить волны, распространяющиеся справа от электрода, через волны, распространяющиеся слева от него, с помощью волновой матрицы $[T_{ik}]$:

$$\begin{pmatrix} c_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ b_1 \end{pmatrix} . \qquad (4.72)$$

Элементы волновой матрицы T_{ik} можно найти, выразив из соотношений (4.63) амплитуды b_2, c_2 через амплитуды b_1, c_1:

$$c_2 = \left(S_{12} - S_{11}^2/S_{12}\right)c_1 + \left(S_{11}/S_{12}\right)b_1,$$
$$b_2 = \left(-S_{11}/S_{12}\right)c_1 + \left(1/S_{12}\right)b_1. \qquad (4.73)$$

Используя соотношения (4.70) и (4.71), находим

$$[T_{ik}] = \begin{pmatrix} 1/\tau^* & r/t \\ -r/t & 1/\tau \end{pmatrix}, \qquad (4.74)$$

где введен параметр $\tau = S_{12}$ – коэффициент прохождения для одного периода структуры:

$$\tau = t\exp\left(-j\kappa d\right). \qquad (4.75)$$

Полученные соотношения применимы к каждому периоду регулярной решетки. Амплитуды волн c_2 и b_2 справа от одного электрода можно отождествить с амплитудами волн c_1 и b_1, существующими слева от последующего электрода. В случае бесконечной решетки решение описывает распространяющийся волновой процесс, если для каждого электрода амплитуды c_2 и b_2 совпадают с амплитудами c_1 и b_1 с точностью до некоторого фазового сдвига. Пусть этот фазовый сдвиг равен $-\gamma d$, так что величину γ можно интерпретировать как волновое число. Поэтому рассмотрим решения, в которых

$$c_2 = c_1\exp\left(-j\gamma d\right),$$
$$b_2 = b_1\exp\left(-j\gamma d\right), \qquad (4.76)$$

что соответствует распространяющейся волне, если волновое число γ – действительная величина. Используя волновую матрицу (соотношения (4.72) и (4.74)), с помощью (4.70) находим дисперсионное уравнение относительно γ:

$$2\cos\left(\gamma d\right) = 1/\tau + 1/\tau^*. \qquad (4.77)$$

Это уравнение можно переписать в более удобном виде, если представить коэффициент прохождения в экспоненциальной форме:

$$t = |t| \exp(j\theta_t),$$

где $|t|$ и θ_t – соответственно модуль и фаза коэффициента прохождения электрода. Тогда $\tau = |t| \exp\left[j(\theta_t - \kappa d)\right]$, и дисперсионное уравнение (4.77) приобретает следующий вид:

$$\cos(\gamma d) = \cos\left(\frac{\omega}{v} d - \theta_t\right) \Big/ |t|, \qquad (4.78)$$

где $v = \omega/\kappa$ – фазовая скорость ПАВ на свободной поверхности.

Обычно значение θ_t мало, а $|t| \to 1$, поэтому при большинстве значений ω, существует действительное решение $\gamma \cong \omega/\kappa$. Тем не менее поскольку $|t| < 1$, то правая часть уравнения (4.78) больше $+1$ или меньше -1, когда значение $\omega d/v - \theta_t$ близко к величине, кратной π. Тогда корень γ становится комплексным и волна экспоненциально затухает, что указывает на наличие полосы режекции. Полагая $\omega d/v - \theta_t = \pi$, из (4.78) находим $\mathrm{Im}\,\gamma \cong |r|/d$. Таким образом, в центре полосы режекции на длине

$$L_r \cong d/|r| \qquad (4.79)$$

амплитуда волны уменьшается в e раз.

На краях полосы режекции правая часть уравнения (4.78) равна ± 1. Используя (4.67), это условие можно записать в виде

$$\omega d/\kappa - \theta_t = m\pi \pm \arcsin(|r|), \quad m = 0, \pm 1, \pm 2, \ldots$$

Обычно θ_t и $|r|$ слабо зависят от частоты, а значение r мало. Поэтому ширина полосы режекции на оси частот задается приближенным выражением

$$\Delta\omega \cong 2v|r|/d\,. \qquad (4.80)$$

Рассмотрим теперь решетку конечной длины из N электродов (рис. 4.31).

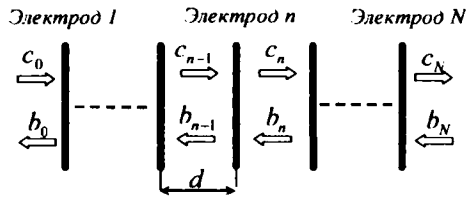


Рис. 4.31. Распространение волны в конечной регулярной решетке электродов

Найдем амплитуду отраженной волны при условии, что ПАВ падает на один из краев. Будем использовать волновую матрицу единичного электрода. Предположим, что краевыми эффектами можно пренебречь, и, следовательно, для всех электродов можно пользоваться одной и той же матрицей. Тогда амплитуды волн по обе стороны n-го электрода связаны уравнением

$$\begin{pmatrix} c_n \\ b_n \end{pmatrix} = [T_{ik}] \begin{pmatrix} c_{n-1} \\ b_{n-1} \end{pmatrix}, \qquad (4.81)$$

где матрица $[T_{ik}]$ задается формулой (4.74), $n = 1...N$. Последовательно используя это уравнение применительно к электродам с номерами от 1 до N, получаем соотношение

$$\begin{pmatrix} c_N \\ b_N \end{pmatrix} = [T_{ik}]^N \begin{pmatrix} c_0 \\ b_0 \end{pmatrix}, \qquad (4.82)$$

которое связывает амплитуды волн справа от N-го электрода с амплитудами волн на левом краю решетки. Если имеется только одна падающая волна $(b_N = 0)$, то из (4.82) при условии $\gamma \approx \omega/v$ вытекает следующая формула для модуля коэффициента отражения по амплитуде:

$$|R| = \left|\frac{b_0}{c_0}\right| = \left[1 + \left|\frac{t}{r}\right|^2 \frac{\sin^2(\omega d/\mathrm{v})}{\sin^2(N\omega d/\mathrm{v})}\right]^{-2}. \qquad (4.83)$$

На рис. 4.32 показан график зависимости модуля коэффициента отражения $|R|$ от нормированной частоты $2fd/\mathrm{v}$.

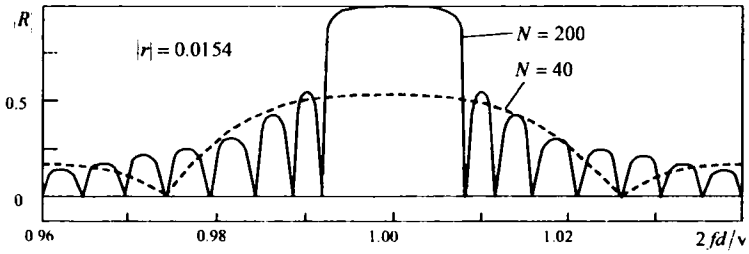


Рис. 4.32. Зависимость модуля коэффициента отражения конечной решетки металлических электродов от частоты при различном их числе

Значение $2fd/\mathrm{v} = 1$ соответствует условию равенства шага решетки половине длины поверхностной акустической волны. Предполагается, что отношение $|t/r|$ не зависит от частоты, что обычно является хорошим приближением. Для достаточно большого числа электродов $N|r| >> 1$ коэффициент отражения близок к единице при условии, что частота падающей волны находится в полосе режекции бесконечной решетки.

4.6. Линии задержки на поверхностных акустических волнах

Одним из важнейших и широко применяемых в электронике устройств являются различные линии задержки (ЛЗ). Так как поверхностные волны являются наиболее медленными упругими волнами, можно получать большие времена задержки при сравнительно малых размерах устройства.

Простейшая линия задержки состоит из двух однородных ВШП, имеющих одинаковый шаг между электродами и размещенных на поверхности пьезоэлектрического звукопровода параллельно друг другу (рис. 4.33). Для устранения нежелательных отражений ПАВ от торцов звукопровода, искажающих сигнал, используются поглотители. Как правило, для уменьшения потерь за счет отражения сигналов от ЛЗ оба преобразователя электрически согласованы с источником и нагрузкой, а число электродов выбирается из условия максимальной полосы пропускания, чтобы обеспечить задержку импульсных сигналов.

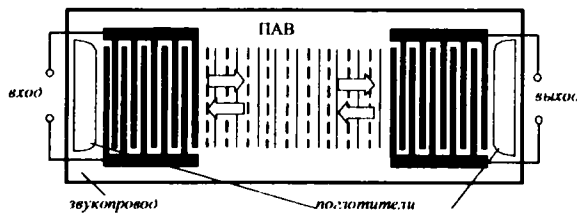


Рис. 4.33. Простейшая линия задержки на поверхностных волнах с двумя однородными ВШП

Недостатком данной конструкции ЛЗ является наличие отражений ПАВ от преобразователей, что приводит к появлению ложных эхо-импульсов, соответствующих нечетному числу проходов ПАВ между преобразователями. Обычно существенным является лишь первый эхо-импульс, связанный с тройным проходом акустического сигнала. Его амплитуда даже в случае согласованных преобразователей остается достаточно большой. При длительном входном сигнале передний фронт трехпролетного сигнала может появиться на выходе линии раньше заднего фронта основного сигнала, что приводит к искажениям.

Анализ отражений и потерь в ЛЗ можно провести с помощью матрицы рассеяния преобразователя. Каждый преобразователь рассматривается как шестиполюсник с тремя входами, в котором вход *3* является электрическим, а входы *1* и *2* – акустическими (рис. 4.34).

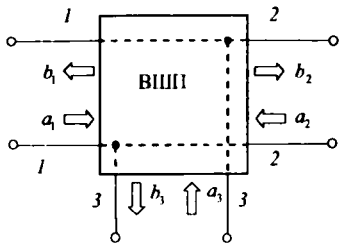


Рис. 4.34. Преобразователь как шестиполюсник с одним электрическим и двумя симметричными акустическими входами

Матрица рассеяния S_{ik} связывает амплитуды сигналов, приходящих на вход a_k, с амплитудами выходящих сигналов b_i:

$$b_i = S_{ik} a_k . \tag{4.84}$$

Ввиду симметрии акустических входов при одновременной замене $a_1 \leftrightarrow a_2$ и $b_1 \leftrightarrow b_2$ коэффициенты при a_1 и a_2, а также амплитуда b_3 не должны изменяться. Следовательно, $S_{11} = S_{22}$, $S_{12} = S_{21}$, $S_{13} = S_{23}$ и $S_{31} = S_{32}$. Введем обозначения: $S_1 = S_{11}$, $S_2 = S_{12}$, $S_{13} = S_3$, $S_{31} = S_4$. Пусть преобразователь электрически согласован. При возбуждении со стороны входа *3* отраженный сигнал $b_3 = 0$, а на акустических входах *1* и *2* присутствуют только выходящие волны, то есть $a_1 = a_2 = 0$, откуда $S_{33} = 0$. Таким образом, матрица рассеяния приобретает следующий вид:

$$[S_{ik}] = \begin{pmatrix} S_1 & S_2 & S_3 \\ S_2 & S_1 & S_3 \\ S_4 & S_4 & 0 \end{pmatrix}. \tag{4.85}$$

Поскольку идеальный преобразователь является пассивным устройством, не поглощающим и не генерирующим энергии, должен выполняться закон сохранения энергии:

$$a_i a_i^* = b_i b_i^* ,$$

откуда с учетом (4.84)

$$S_{ik} S_{im}^* = \delta_{im} . \tag{4.86}$$

Из (4.86) следует:

$$\begin{aligned} &S_2 = -S_1, \quad S_4 = \pm S_3, \\ &|S_3|^2 = |S_4|^2 = \frac{1}{2}, \quad |S_1|^2 = |S_2|^2 = \frac{1}{4}. \end{aligned} \tag{4.87}$$

Полученные соотношения позволяют определить потери на преобразование в электрически согласованной линии задержки и уровень трехпролетного сигнала. Пусть мощность сигнала на электрическом входе равна A^2. После преобразования половина акустической мощности поступит на вход *1* и будет поглощена. Вторая половина $A^2/2$ достигнет входа 1 второго преобразователя, где одна четверть от нее отразится ($A^2/8$), одна четверть пройдет на вход 2 и поглотится, а половина ($A^2/4$) будет преобразована в электрический сигнал. Таким образом, потери на преобразования в ЛЗ составляют не менее 6 дБ – по 3 дБ на каждый преобразователь. Уровень трехпролетного импульса только на 12 дБ ниже основного прошедшего сигнала, так как данный импульс претерпевает два дополнительных отражения, которые приводят к потерям по 6 дБ на каждое отражение.

Простейшим способом подавления трехпролетного сигнала является рассогласование преобразователей. В этом случае подавление этого сигнала достигается за счет увеличения вносимых потерь. При увеличении потерь на преобразование от 6 до 12 дБ уровень ложного сигнала снижается на 33 дБ.

Значительное уменьшение вносимых потерь при одновременном подавлении ложных сигналов можно достичь в ЛЗ на рис. 4.35. Два

приемных преобразователя R и R' симметрично расположены относительно передающего преобразователя T. Выходные сигналы преобразователей R и R' объединяются, и передающий преобразователь T, таким образом, не имеет потерь на 3 дБ, обусловленных двунаправленностью излучения. Весьма важно, что T теперь является в принципе идеальным поглотителем акустических сигналов, отраженных от R и R', и ложный эхо-импульс тройного прохождения исчезает. Однако полное подавление происходит только при идеальном согласовании преобразователя T, что может быть достигнуто лишь на центральной частоте. Кроме того, требуется идеальная симметричность преобразователя T и его симметричное расположение, а также полная идентичность преобразователей R и R'. При изготовлении ЛЗ эти условия могут быть выполнены лишь приближенно. Тем не менее получаемое таким образом дополнительное подавление трехпролетного сигнала достигает 20 дБ.

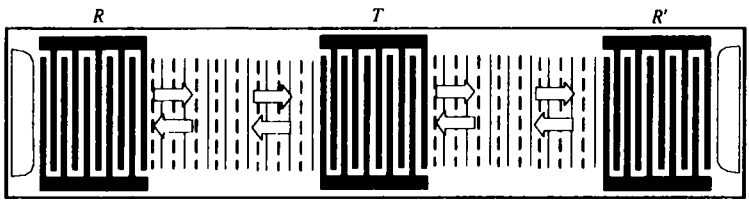


Рис. 4.35. Линия задержки с двумя симметричными приемными преобразователями

Другой возможный метод уменьшения потерь и подавления ложного сигнала заключается в использовании различных видов однонаправленных преобразователей, которые мы рассматривали в предыдущих разделах. Этот метод заметно усложняет конструкцию, но на практике оказывается более эффективным, чем рассмотренные выше.

Кроме сигнала тройного прохождения, имеется и другой источник ложных сигналов – объемные волны, неизбежно возбуждаемые преобразователем. Подавления объемных волн добиваются путем придания шероховатости обратной стороне подложки, либо нанесением поглощающего покрытия. Эффективным способом устранения влияния объемных волн является применение многополоскового ответвителя с полной передачей энергии в параллельный канал

(рис. 4.36). Так как объемные волны не переносятся МПО в параллельный канал, они не достигают выходного преобразователя.

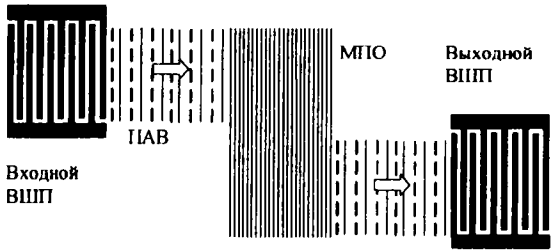


Рис. 4.36. Линия задержки с МПО для подавления связи с объемными волнами

Рассмотренные линейные ЛЗ обеспечивают задержку не более 50–100 мкс. При большей задержке требуемая длина звукопровода становится слишком большой. В то же время часто требуются времена задержки в одну и более миллисекунд. Одним из путей решения данной задачи является объединение нескольких относительно коротких ЛЗ с разделительными усилителями (рис. 4.37). Периодическое применение усилителей делает допустимыми относительно большие потери в каждой из линий. При этом вносимые потери улучшают подавление трехпролетного сигнала.

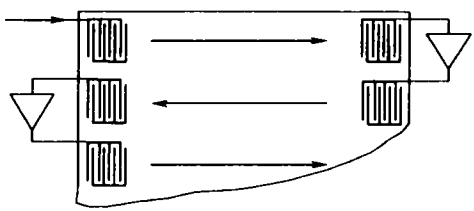


Рис. 4.37. Линия длительной задержки с разделительными усилителями

В линиях длительной задержки могут использоваться многополосковые ответвители и отражательные структуры на их основе. В ЛЗ на рис. 4.38 входной сигнал с помощью МПО переносится в параллельный поток, дважды отражается и вновь через МПО поступает на выход. Преимущество данной ЛЗ заключается в многократном использовании длины звукопровода.

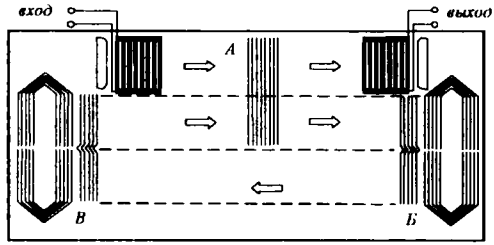


Рис. 4.38. Линия задержки на структурах МПО. Траектория волны: вх. ВШП–А–Б–В–А–вых. ВШП

Рассмотренные варианты ЛЗ рассчитаны на получение фиксированной величины задержки. В то же время часто требуется изменяемая величина задержки. Устройства с изменяемой величиной задержки часто реализуются в виде многоотводных линий задержки (МЛЗ). Для построения МЛЗ можно также использовать многополосковые ответвители. Пример такого устройства показан на рис. 4.39.

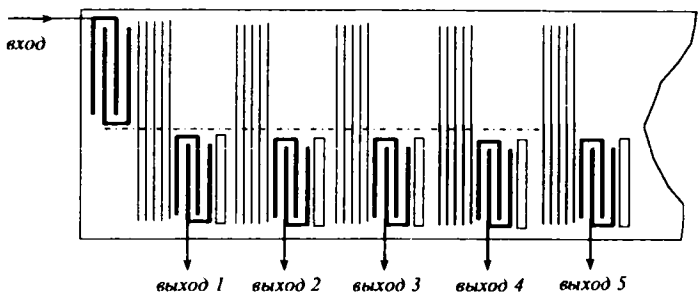


Рис. 4.39. Многоотводная линия задержки на МПО

Поверхностная волна, возбуждаемая встречно-штыревым входным преобразователем, находящимся в верхнем канале, последовательно с некоторым временным интервалом проходит через ряд коротких МПО, которые производят частичный отбор энергии в нижний канал. Волна, возбуждаемая МПО в нижнем канале, принимается соответствующим выходным преобразователем. С тыльной стороны выходных ВШП нанесены поглотители, которые демпфируют волну, прошедшую через преобразователь. Это предотвращает даль-

нейшее распространение ПАВ в нижнем канале и ее частичное переизлучение в верхний канал, что приводило бы к искажению характеристик МЛЗ.

4.7. Резонаторы на ПАВ

Одним из важнейших и широко применяемых акустоэлектронных устройств являются резонаторы на объемных акустических волнах, которые мы рассматривали в разделе 4.1. Их работа основана на размерном механическом резонансе плоскопараллельной пьезокристаллической пластины при многократном отражении акустической волны от ее граней. Электрический резонанс наступает при равенстве толщины пластины нечетному числу полуволн. На низких частотах используется основная мода колебаний, а на высоких – обычно третья гармоника. Несмотря на простоту конструкции и хорошие электрические характеристики применение резонаторов на объемных волнах ограничено частотами до 50–100 МГц. На более высоких частотах пластина становится слишком тонкой и хрупкой и ее изготовление все более сложным, а использование высших мод колебаний ухудшает электрические характеристики резонатора.

Значительно более высоких рабочих частот можно достичь в резонаторах на поверхностных акустически волнах (ПАВ-резонаторах). По принципу действия ПАВ-резонаторы схожи с резонаторами на объемных волнах. Но если резонаторная полость последних ограничена гранями пластины, то в ПАВ-резонаторах ее ограничивают распределенные отражатели – отражательные решетки (ОР), расположенные на упругой подложке в общем акустическом потоке (рис. 4.40).

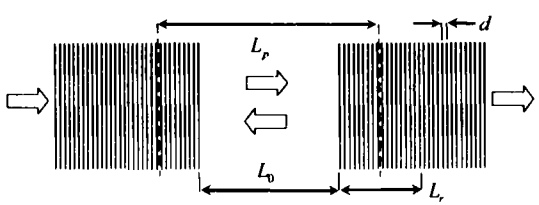


Рис. 4.40. Проходной ПАВ-резонатор на основе отражательных решеток (ОР). Пунктиром показаны положения эффективных дискретных отражателей

Заметим, что использование дискретных отражателей для создания ПАВ-резонаторов, например границ подложки, не может быть эффективным. Это следует из того простого факта, что пучности продольного и поперечного смещений в стоячей волне Рэлея сдвинуты относительно друг друга на четверть длины волны. Таким образом, граничное условие не может быть удовлетворено, и дискретный идеальный отражатель волн Рэлея принципиально не реализуем. При падении на край подложки поверхностная волна частично отражается, но значительная доля энергии преобразуется в объемные волны, поэтому добротность резонатора будет низкой.

Элементы ОР обычно представляют собой металлические полосковые электроды или канавки, ортогональные направлению распространения ПАВ. Каждый из этих элементов вызывает незначительное возмущение поверхности, которое приводит к малому отражению ПАВ и возбуждению слабой объемной волны. Если шаг решетки равен половине длины ПАВ, то отраженные от отдельных элементов поверхностные волны складываются в фазе, и при достаточно большом числе элементов суммарный коэффициент отражения ПАВ от решетки стремится к единице. Теоретическую зависимость коэффициента отражения от частоты, полученную в пренебрежении генерацией объемных волн, выражает соотношение (4.83). Так как скорости объемных волн больше скорости ПАВ, то на рабочей частоте, соответствующей максимуму отражения ПАВ, объемные волны от отдельных элементов не будут находиться в фазе, и генерация решеткой объемных волн весьма незначительна. Таким образом, отражательные решетки обеспечивают близкий к единице коэффициент отражения ПАВ при незначительных потерях на генерацию объемных волн.

Проходной ПАВ-резонатор аналогичен оптическому резонатору Фабри-Перо. Но в резонаторе Фабри-Перо толщина отражающих зеркал пренебрежимо мала по сравнению с его длиной, и резонансные свойства определяются расстоянием между зеркалами. В случае же ПАВ-резонатора волна проникает внутрь отражателей, экспоненциально затухая по глубине. Эффективная глубина проникновения L_r определяется как такое расстояние в отражателе, на котором амплитуда волны уменьшается в e раз. Значение L_r в центре полосы режекции определяется соотношением (4.79). Заметим, что эффективная глубина проникновения будет зависеть от частоты, она минимальна в центре полосы режекции решетки и стремится к беско-

нечности на ее краях. В ПАВ резонаторах величина L_r может быть близка, а часто даже больше расстояния между решетками L_0. Поэтому эффективная длина резонатора L_p больше чем L_0.

Эффективную длину ПАВ-резонатора L_p определим следующим образом. Заменим распределенные отражатели некоторыми эффективными дискретными отражателями, размещенными в глубине решеток, и найдем расстояние между ними. Энергия, запасенная в объеме резонатора, пропорциональна среднему значению квадрата амплитуды волны A_0 и расстоянию между эффективными отражателями:

$$W \sim L_p A_0^2 . \tag{4.88}$$

Внутри решетки зависимость амплитуды ПАВ от глубины при достаточно большом числе отражающих элементов имеет вид

$$A(x) = A_0 \exp(-x/L_r),$$

Энергия внутри одной решетки

$$W_p \sim A_0^2 \int_0^\infty \exp(-2x/L_r)\,dx = A_0^2 L_r/2 ,$$

а в зазоре между решетками

$$W_0 \sim A_0^2 L_0 .$$

Полная энергия в резонаторе

$$W \sim A_0^2 L_0 + A_0^2 L_r .$$

Сравнивая это выражение с формулой (4.88), получаем

$$L_p = L_0 + L_r . \tag{4.89}$$

Другое важное отличие ПАВ-резонатора заключается в том, что в ПАВ-резонаторе существует два механизма частотной селекции: зависимость коэффициента отражения решеток от частоты и собственно резонансные свойства промежутка между эффективными дискретными отражателями. В оптическом резонаторе Фабри-Перо зеркала обычно являются широкополосными с коэффициентом отражения, близким к единице. В результате на частотной зависимости коэффициента пропускания наблюдаются периодические пики, чередующиеся с полосами заграждения. В случае ПАВ-резонатора, если на частоте максимального отражения решетки расстояние $L_0 = n\Lambda/2$ $(n = 1,2,3,...)$, характеристика пропускания будет иметь такой же вид, как и для резонатора Фабри-Перо (рис. 4.41, а). Если же на этой частоте $L_0 = (\Lambda/4)(2n-1)$, то характеристика пропускания имеет провал в середине каждого пика, как это показано на рис. 4.41, б.

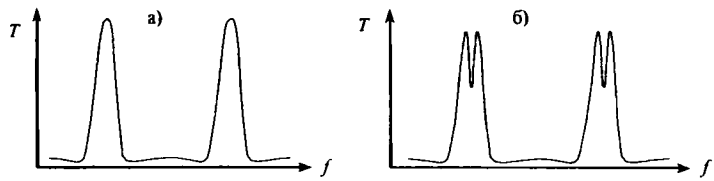


Рис. 4.41 Характеристика пропускания проходного резонатора. Эффективное расстояние между отражателями при резонансе решетки равно: а – четному числу четвертей длины волны; б – нечетному числу четвертей волны

Рассмотренные свойства проходного ПАВ-резонатора могут быть основой при конструировании различных устройств. Однако применение проходного резонатора связано с некоторыми технологическими трудностями. Коэффициенты отражения решеток не могут быть слишком близки к единице, а сами решетки должны быть идентичны. Если данное условие нарушается, пики пропускания значительно уменьшаются. Это накладывает достаточно жесткие требования на точность изготовления решеток. Указанных трудностей по большей части можно избежать, помещая в резонатор один или несколько ВШП (рис. 4.42). В таких резонаторах избирательные свойства непосредственно переводятся в электрические характеристики. При этом отражательные решетки можно сделать достаточно

протяженными, чтобы избежать излучения акустической энергии через отражатели.

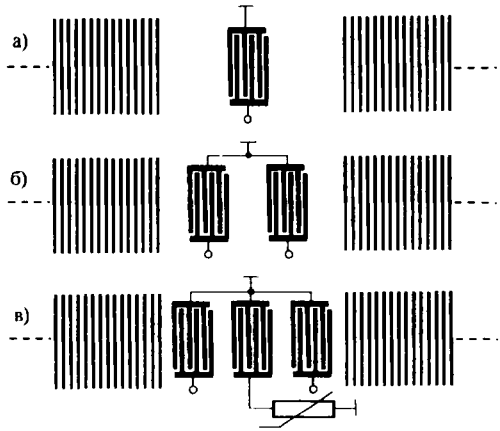


Рис. 4.42. ПАВ-резонаторы, нагруженные встречно-штыревыми преобразователями: а) одновходовой резонатор; б) двухвходовой резонатор; в) трехвходовой резонатор с подстройкой резонансной частоты

Рассмотрим ПАВ-резонатор с одним ВШП (рис. 4.42, а). Его работу можно достаточно просто проанализировать, если учесть, что резонатор представляет собой весьма узкополосное устройство и с хорошим приближением можно принять, что в пределах рабочей полосы резонатора ВШП работает вблизи своей синхронной частоты f_0. Это значит, что можно пренебречь реактивной проводимостью преобразователя, проявляющей себя на частотах, далеких от синхронной частоты.

Электрическая эквивалентная схема одновходового ПАВ-резонатора представлена на рис. 4.43. Она аналогична схемам, используемым для описания акустических резонаторов на объемных волнах. Здесь C_0 – статическая емкость преобразователя, C_1 и L_1 – динамические реактивные сопротивления, R_1 – сопротивление излучения при акустическом резонансе. Выражения для элементов эквивалентной схемы при N парах штырей ВШП, емкости одной пары C_2 и коэффициенте отражения решеток R имеют следующий вид:

$$C_0 = NC_2,\ \ R_0 = \left(8N^2 f_0 K^2 C_2\right)^{-1},$$
$$R_1 = R_0 \frac{1-|R|}{1+|R|},\ \ L_1 = \frac{N_p}{4}\frac{R_0}{f_0},\ \ C_1 = \frac{1}{4\pi f_0^2 L_1}, \tag{4.90}$$

где $N_p = L_p/\Lambda$ – расстояние между эффективными дискретными отражателями резонатора, выраженное в длинах волн ПАВ. Импеданс устройства будет минимален на частоте последовательного резонанса f_r и достигает максимума на частоте параллельного резонанса f_a. Для резонаторов с высокой добротностью (при малых R_1) из эквивалентной схемы следует, что

$$f_a - f_r = \frac{f_0}{2}\frac{C_1}{C_0} = f_0 \frac{4K^2}{\pi^2}\frac{N}{N_p}. \tag{4.91}$$

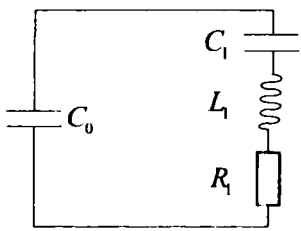


Рис. 4.43. Электрическая эквивалентная схема одновходового ПАВ-резонатора

Величина $f_0/(f_a - f_r)$ может рассматриваться как параметр качества ПАВ-резонатора. Однако ее не следует отождествлять с добротностью, которая имеет вид

$$Q = \frac{\pi}{2} N_p. \tag{4.92}$$

Ясно, что добротность может быть повышена за счет увеличения зазора между отражательными решетками. Однако увеличение этого зазора может привести к появлению дифракционных по-

терь. Кроме того, расстояние между соседними модами резонатора будет уменьшаться.

Для проектирования резонаторов с одночастотным откликом необходимо сузить полосу пропускания ВШП. При совпадении синхронной частоты ВШП с частотой требуемого резонанса соседние отклики резонатора совпадают с первыми нулями частотной характеристики преобразователя, только если $N = N_p$. Однако, даже если преобразователь заполняет весь зазор, N_p больше N по крайней мере на величину $N_r = L_r/\Lambda$. Но если сделать $N >> N_r$, можно обеспечить близость нулей характеристики ВШП и соседних откликов резонатора.

Двухвходовой резонатор (рис. 4.42, б) аналогичен обычной линии задержки. Однако на резонансной частоте коэффициент передачи этой линии резко возрастает. Двухвходовой резонатор может быть использован в качестве узкополосного фильтра, обеспечивая ту же полосу пропускания, что и фильтр обычного типа, при достаточно малых размерах.

Одним из способов подстройки частоты ПАВ-резонатора является введение дополнительного ВШП (рис. 4.42, в). Изменение нагрузочного импеданса этого преобразователя влияет на время распространения ПАВ между отражателями. Таким способом удается осуществлять относительную перестройку частоты в пределах до 0.5 %.

В области частот выше 100 МГц ПАВ-резонаторы обладают несомненными преимуществами перед своими аналогами, работающими на объемных волнах. Независимо от рабочей частоты размеры звукопровода всегда могут быть достаточно большими. Современная технология позволяет изготовить отражатели и преобразователи с субмикронным шагом. При этом верхняя граница частотного диапазона примерно на порядок превышает соответствующую величину для объемно-волновых аналогов. Добротность ПАВ-резонаторов составляет 10^4 и более.

4.8. Акустоэлектронный усилитель

Ранее было установлено, что при акустоэлектронном взаимодействии возможен эффект усиления волны. Продольная компонента

пьезоэлектрического поля волны взаимодействует со свободными носителями заряда, дрейфующими в направлении ее распространения, и при определенных условиях наблюдается усиление волны за счет энергии тянущего поля. Данный эффект может быть использован для создания акустоэлектронных усилителей.

Возможны следующие разновидности акустоэлектронных усилителей: монолитный усилитель на объемных или поверхностных волнах в пьезополупроводниковом кристалле; пленочный усилитель с полупроводниковой пленкой на пьезокристаллической подложке или с пьезокристаллической пленкой на полупроводящей подложке; усилитель с воздушным зазором между пьезоэлектриком и полупроводниковой пленкой. В последних двух видах усилителей используются поверхностные акустические волны.

В технологическом отношении наиболее прост, очевидно, монолитный усилитель. Однако совмещение в одном кристалле пьезо- и полупроводящих свойств существенно ограничивает возможности управления параметрами усилителя. Наибольшей привлекательностью для акустоэлектронных устройств обработки сигналов обладает усилитель с полупроводниковой пленкой. Процессы акустоэлектронного взаимодействия в этом усилителе происходят в тонкой полупроводниковой пленке, поэтому электрическое сопротивление пленки может быть довольно большим, что уменьшает рассеиваемую тепловую мощность. Пленочный усилитель также может быть частью интегрального акустоэлектронного устройства, уменьшая потери, вносимые различными компонентами. Так как эффект усиления является однонаправленным, то есть при распространении ПАВ против направления дрейфа волна испытывает дополнительное поглощение, уменьшается трехпролетный сигнал, возникающий из-за отражения от преобразователей.

Динамический диапазон пленочного усилителя сверху ограничивается наступлением концентрационной нелинейности, заключающейся в том, что все свободные носители группируются в сгустки и дальнейший рост усиления невозможен. Снизу динамический диапазон ограничивается собственными шумами полупроводниковой пленки.

К сожалению, применение пленочных усилителей существенно ограничивается высокими требуемыми напряжениями питания (несколько сот и даже тысяч вольт) и большими собственными шумами. По своим эксплуатационным характеристикам они значительно уступают транзисторным усилителям.

Конструктивная схема пленочного усилителя представлена на рис. 4.44.

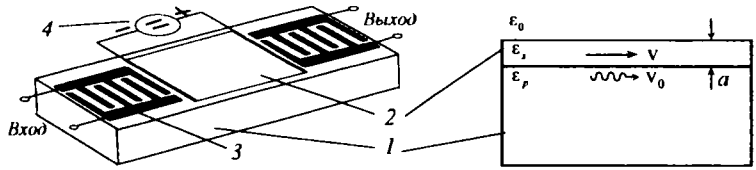


Рис. 4.44. Пленочный усилитель: 1 – пьезоэлектрическая подложка, 2 – полупроводниковая пленка, 3 – ВШП, 4 – источник постоянного смещения

Пленочный усилитель (ПУ) состоит из пьезоэлектрической подложки с нанесенными на ее поверхности входным и выходным встречно-штыревыми преобразователями и полупроводниковой пленкой, на которую подается постоянное смещение от внешнего источника. Поверхностная упругая волна распространяется вдоль границы раздела пленки и подложки. Электрическое поле волны проникает в полупроводниковую пленку, где взаимодействует с дрейфующими свободными носителями полупроводника. Коэффициент усиления ПУ на единицу длины G [дБ/м] при толщине пленки a, значительно меньшей длины волны ПАВ и дебаевского радиуса экранирования $(a << \Lambda, L_D)$, равен

$$G = 8.68\frac{K^2}{2}\frac{\omega}{v_0}\frac{-\gamma\varepsilon_r a\omega_c/v_0}{(\omega/\omega_D + \varepsilon_r a\omega_c/v_0)^2 + \gamma^2}. \qquad (4.93)$$

В этом выражении: K – эффективный коэффициент электромеханической связи, $\omega_c = \sigma/\varepsilon_s$ – частота проводимости, σ – удельная проводимость пленки, $\gamma = 1 - v/v_0$ – фактор дрейфа, v – дрейфовая скорость свободных носителей, v_0 – скорость ПАВ, $\varepsilon_r = \varepsilon_s/(\varepsilon_p + \varepsilon_0)$, ε_s – диэлектрическая проницаемость пленки, ε_p – диэлектрическая проницаемость подложки, $\omega_D = v_0^2/\mathcal{D}$ – диффузионная частота, $\mathcal{D}$ – коэффициент диффузии носителей.

Из соотношения (4.93) следует, что для усиления необходимо, чтобы скорость дрейфа свободных носителей в полупроводнике v

превышала скорость волны v_0 $(\gamma < 0)$. Максимальное усиление наступает, если

$$\gamma = \gamma_{max} = -\left(\omega/\omega_D + \varepsilon_r a \omega_c / v_0\right). \qquad (4.94)$$

Коэффициент усиления при заданных ω и γ по мере роста толщины пленки сначала увеличивается, достигая максимального значения, когда выполняется условие

$$a = a_{max} = \sqrt{\gamma^2 + (\omega/\omega_D)^2} \Big/ \left(\varepsilon_r \omega_c / v_0\right), \qquad (4.95)$$

а затем уменьшается. Такая зависимость связана с тем, что с ростом толщины увеличивается область взаимодействия, но одновременно падает электрическое сопротивление пленки. При малых a сопротивление пленки велико, и коэффициент усиления с ростом толщины увеличивается. Когда толщина становится больше a_{max}, из-за малого электрического сопротивления пленка начинает шунтировать пьезоэлектрическое поле волны, акустоэлектронное взаимодействие уменьшается, и коэффициент усиления снижается.

На величину коэффициента усиления влияет диффузия свободных носителей заряда в пленке. Как видно из (4.93), коэффициент усиления линейно растет с увеличением частоты, пока $\omega << \omega_D$. При $\omega \sim \omega_D$ рост уменьшается, и на высоких частотах $\omega >> \omega_D$ коэффициент усиления начинает уменьшаться. Данное явление называют диффузионным ограничением. Если диффузия пренебрежимо мала $\left(\omega/\omega_D \to 0\right)$, максимальное усиление при оптимальном значении фактора дрейфа равно:

$$G_{0\,max} = 8.68 \frac{K^2}{4} \frac{\omega}{v_0}. \qquad (4.96)$$

При наличии диффузии максимальное значение коэффициента усиления определится следующим выражением:

$$G_{max} = 8.68\frac{K^2}{4}\frac{\omega}{v_0}\frac{\varepsilon_r a\omega_c / v_0}{(\omega/\omega_D + \varepsilon_r a\omega_c / v_0)}. \qquad (4.97)$$

Очевидно, что $G_{max}/G_{0max} < 1$.

В качестве критерия влияния диффузионного ограничения можно принять величину отношения G_{max}/G_{0max}. Обычно считается, что диффузия несущественна, если $G_{max}/G_{0max} > 0.5$. Этот критерий позволяет сформулировать требования к электрофизическим параметрам полупроводниковой пленки для усилителя, рассчитанного на заданную рабочую частоту ω. Отношение G_{max}/G_{0max} будет больше 0.5, если

$$\omega/\omega_D < \varepsilon_r a\omega_c / v_0 . \qquad 4.98)$$

После несложных преобразований из (4.98) получаем

$$an > \frac{\omega\varepsilon_p}{v_0}\frac{kT}{q^2}, \qquad (4.99)$$

где n – концентрация свободных носителей в пленке, k – постоянная Больцмана, T – абсолютная температура, q – заряд электрона. Численная оценка показывает, что, например, при рабочей частоте 100 МГц и толщине пленки $a = 0.1$ мкм на подложке из ниобата лития концентрация электронов n должна быть больше 10^{12} см$^{-3}$, что легко достигается в большинстве полупроводниковых материалов. Таким образом, диффузионное ограничение в ПУ оказывается несущественным, и в расчетах им, как правило, пренебрегают.

При конструировании усилителя необходимо учитывать величину тепловой мощности, рассеиваемую пленкой. В расчете на единицу площади величина рассеиваемой тепловой мощности определится соотношением

$$P^T = a\sigma E^2 = qn\mu a E^2, \qquad (4.100)$$

где μ – подвижность носителей, E – напряженность постоянного электрического поля, приложенного к пленке. Так как усиление име-

ет место, только если скорость носителей $v = \mu E$ превосходит скорость волны v_0, то минимальная (пороговая) мощность, рассеиваемая пленкой, при которой начинается усиление, определится из условия $\mu E = v_0$ $(\gamma = 0)$ и будет равна:

$$P_{min}^{T} = \frac{qnav_0^2}{\mu}. \qquad (4.101)$$

Для того чтобы снизить пороговую мощность и уменьшить нагрев усилителя, следует выбирать материалы с большой подвижностью носителей заряда. Для пленки антимонида индия InSb с толщиной 0.1 мкм при комнатной температуре $P_{min}^{T} \cong 10$ мВт/см^2. В то же время можно взять материал с меньшей подвижностью, но и с существенно меньшей концентрацией носителей и получить более низкое значение пороговой мощности.

При $\mu E > v_0$ $(\gamma < 0)$ рассеиваемая мощность равна

$$P^{T} = P_{min}^{T}(1-\gamma)^2. \qquad (4.102)$$

Оптимальное значение γ для пленки InSb с упомянутыми выше параметрами равно $|\gamma|_{max} \cong 10^2$. Получаемое при этом значение $P^T = 100$ Вт/см^2 практически недостижимо из-за теплового разрушения усилителя. Предельная величина P_{max}^{T}, не приводящая к тепловому разрушению, зависит от параметров полупроводника и пьезоэлектрика, степенью согласованности температурных коэффициентов расширения пленки и подложки.

Практическая реализация пленочного усилителя осложняется большими технологическими трудностями получения пленки, связанными как с подбором пар материалов с близкими значениями периодов кристаллических решеток и температурных коэффициентов расширения, так и со сложностью самого процесса выращивания на пьезоэлектрической подложке полупроводниковой пленки с заданными электрофизическими свойствами.

Указанные технологические трудности можно преодолеть в пленочном усилителе с воздушным зазором (ПУВЗ), у которого в отличие от ПУ полупроводниковая пленка наносится на диэлектриче-

скую подложку и приподнята над пьезоэлектриком, так что между ними образуется небольшой воздушный зазор (рис. 4.45). Поскольку выбор материалов полупроводниковой пленки и пьезоэлектрика совершенно произволен, появляется возможность создания усилителя с оптимальными параметрами.

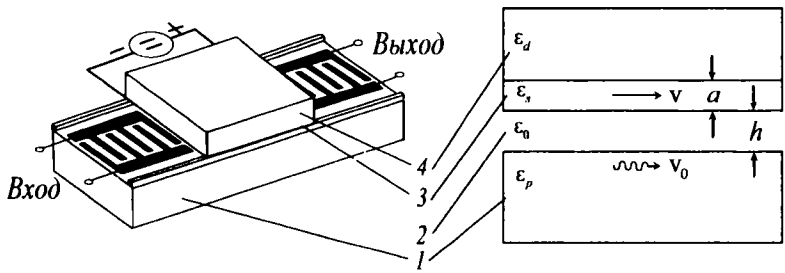


Рис. 4.45. Конструктивная схема усилителя с воздушным зазором между полупроводниковой пленкой и пьезоэлектрической подложкой: *1* – пьезоэлектрическая подложка, *2* – воздушный зазор, *3* – полупроводниковая пленка, *4* – диэлектрик

Коэффициент усиления ПУВЗ определяется следующим выражением:

$$G = 8.68\frac{K^2}{2}\frac{\omega}{v_0}\frac{-\gamma\varepsilon_r a\,\omega_c/v_0}{\left[\omega/\omega_D + \left(1+\kappa h\,\varepsilon_p/\varepsilon_0\right)\varepsilon_r a\omega_c/v_0\right]^2+\gamma^2}. \qquad (4.103)$$

Предполагается, что толщина пленки и воздушного зазора удовлетворяет следующим условиям: $a \ll \Lambda, L_D$ и $\kappa h \ll 1$, где h – величина зазора. При заданной частоте и факторе дрейфа оптимальная толщина пленки равна:

$$a_{max} = \frac{\sqrt{\gamma^2+(\omega/\omega_D)^2}}{\left(\varepsilon_r\,\omega_c/v_0\right)\left(1+\kappa h\,\varepsilon_p/\varepsilon_0\right)}. \qquad (4.104)$$

Поскольку напряженность электрического поля волны на поверхности полупроводника за счет зазора меньше, чем на поверхности пьезоэлектрической подложки, оптимальное значение толщины несколько меньше, чем в усилителе без зазора.

Ограничение роста усиления ПУВЗ с увеличением частоты связано главным образом с наличием зазора, а не с процессом диффузии, как в ранее рассмотренном ПУ. Коэффициент усиления достигает своего максимального значения уже на частоте

$$\omega = (v/h)(\varepsilon_0/\varepsilon_p), \tag{4.105}$$

которая обычно значительно меньше ω_D. Таким образом, наличие зазора заметно сужает полосу усиливаемых частот. Оценки показывают, что коэффициент усиления усилителя с зазором на частоте 100 МГц при зазоре $h = 0.1$ мкм и оптимальных параметрах пленки не превышает 60 дБ/см, тогда как для усилителя без зазора он может достигать 100 дБ/см и более.

4.9. ПАВ-автогенераторы и синтезаторы частоты

Типичный автогенератор на ПАВ содержит усилитель, охваченный цепью положительной обратной связи, в которую включен стабилизирующий частотно-задающий элемент – ПАВ-резонатор, либо линия задержки (рис. 4.46).

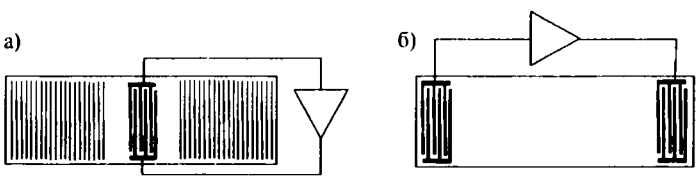


Рис. 4.46. Автогенераторы на поверхностных акустических волнах: а – генератор с ПАВ-резонатором, б – генератор с линией задержки

Важнейшими параметрами любого автогенератора являются рабочая частота и стабильность. По своей стабильности ПАВ-генераторы приближаются к своим аналогам, использующим кварцевые резонаторы на объемных волнах. Однако рабочие частоты последних не превышают 50–100 МГц, причем на частотах выше 20 МГц используются высшие гармоники. Для получения более высоких частот приходится использовать умножители. В то же время ПАВ-генераторы позволяют непосредственно генерировать колеба-

ния с частотами от 20 МГц до 3 ГГц, что позволяет избавиться от умножителей и делает их более компактными и дешевыми. Малогабаритные ПАВ-элементы, применяемые в генераторах, хорошо сочетаются с микроэлектронной планарной технологией. На одной подложке можно разместить несколько линий задержки или резонаторов на различные рабочие частоты вместе с электронным обеспечением (усилители, ключи). Это значительно расширяет функциональные возможности устройства.

Структурная схема генератора с ПАВ-резонатором представлена на рис. 4.46, а. В цепь положительной обратной связи усилителя включен одновходовой ПАВ-резонатор, состоящий из ВШП, размещенного между двумя отражательными решетками, находящимися на поверхности пьезоэлектрической подложки. На резонансной частоте коэффициент передачи резонатора резко возрастает и начинается генерация. Высокая добротность ПАВ-резонаторов $\left(Q > 10^4\right)$ позволяет создавать высокостабильные генераторы, перекрывающие диапазон от 20 МГц до 3 ГГц. Однако это несомненное достоинство не совместимо с возможностью широкой перестройки частоты.

Большими функциональными возможностями обладает генератор с линией задержки в цепи обратной связи (рис. 4.46, б). По существу, такой генератор представляет собой кольцевой резонатор с компенсацией полных потерь. Этот метод хорошо известен для электромагнитных волн оптического и СВЧ диапазонов. Резонансные частоты определяются условием, при котором полный фазовый сдвиг в цепи обратной связи кратен 2π:

$$\frac{\omega L}{\mathrm{v}} + \phi = 2\pi m \quad (m = 1,2...), \tag{4.106}$$

где L – длина звукопровода, v – скорость ПАВ, ϕ – полный набег фазы в преобразователях и усилителе. Протяженность звукопровода линии задержки обычно составляет сотни и даже тысячи длин волн ПАВ, поэтому величиной ϕ можно пренебречь. Тогда из (4.106) получаем

$$f_m = m(\mathrm{v}/L). \tag{4.107}$$

Таким образом, возможные частоты генерации образуют эквидистантный спектр с интервалом между соседними модами

$$\Delta f = \mathrm{v}/L\,. \qquad 4.108)$$

Так как длина линии задержки велика, обычно этот интервал меньше полосы пропускания усилителя, и генерация может быть получена на нескольких различных модах раздельно или даже одновременно. На практике обычно стремятся получить одночастотный режим, хотя в некоторых применениях требуется одновременное наличие нескольких частот. В любом случае желателен эффективный контроль возникающих мод.

Для получения одночастотного режима работы необходимо, чтобы коэффициент усиления в петле обратной связи был меньше единицы для всех ненужных частот, то есть надо ввести дополнительную избирательность для ослабления всех мод, кроме одной. Это задачу можно решить с помощью внешних цепей. Существует, однако, способ дополнительной селекции внутри самого генератора. Для этого один из преобразователей делают достаточно протяженным и, следовательно, имеющим узкую полосу пропускания (рис. 4.47, а). Второй преобразователь существенно короче и не оказывает влияния на избирательные свойства. В частности, если длину протяженного преобразователя выбрать такой, что все частоты спектра, кроме выбранной, совмещаются с нулями частотной характеристики ВШП, будет возбуждаться единственная мода (рис. 4.47, б). В разделе 4.3 мы видели, что ближайшие к резонансной частоте нули преобразователя расположены на расстоянии f_0/N_2 от центральной частоты f_0 (N_2 – число пар штырей). Используя соотношение (4.108), получаем, что для требуемого совпадения нулей характеристики ВШП с подавляемыми частотами необходимо, чтобы эффективная длина линии задержки (расстояние между серединами преобразователей) равнялась длине протяженного преобразователя $L = \Lambda N_2$, где Λ – длина волны ПАВ.

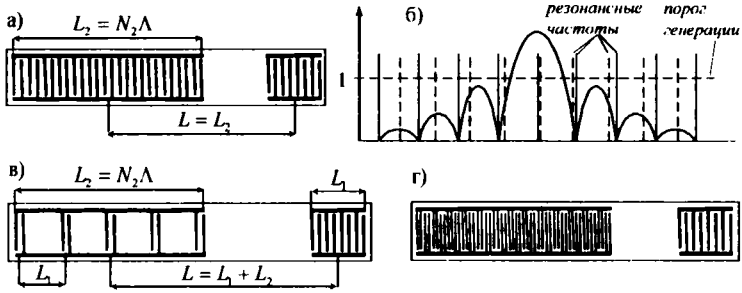


Рис. 4.47. Выбор геометрии преобразователя для одночастотного режима: а – применение протяженного преобразователя; б – частотная характеристика коэффициента передачи в петле обратной связи и положение резонансных частот; в – применение прореженного преобразователя; г – применение преобразователя с расщепленными штырями

Отметим, что сформулированное выше условие не является жестким. Даже если подавляемые частоты не совпадают строго с нулями преобразователя, для одночастотного режима генерации вполне достаточно, чтобы общий коэффициент передачи в петле положительной обратной связи для этих частот с учетом частотной характеристики протяженного преобразователя был меньше единицы. На рис. 4.47, б вертикальными линиями обозначено положение резонансных частот при их совпадении с нулями АЧХ ВШП (сплошные линии) и для того случая, когда они смещены относительно упомянутых нулей (пунктир). Горизонтальной чертой изображен порог генерации, соответствующий полному коэффициенту передачи в петле обратной связи, равному единице. В обоих случаях для всех мод за пределами главного лепестка АЧХ преобразователя условие генерации не выполняется.

Протяженные однородные ВШП с большим числом штырей имеют один серьезный недостаток. Так как на центральной частоте период решетки преобразователя равен половине длины волны, в случае большого числа электродов существенно возрастает отражение от преобразователя, что ухудшает характеристики генератора.

Одной из возможностей преодоления этого недостатка протяженного ВШП является его периодическое прореживание – удаление части штырей (рис. 4.47, в). Вследствие уменьшения общего числа штырей отражение снижается, но и АЧХ ВШП также изменяется. Анализ показывает, что в автогенераторе на линии задержки с про-

реженным преобразователем для подавления всех частот, кроме основной, необходимо выполнение двух условий. Во-первых, длина короткого преобразователя должна быть равна расстоянию между группами электродов прореженного ВШП. Во-вторых, эффективная длина линии задержки должна равняться сумме длин протяженного и короткого ВШП.

Другой возможностью является «расщепление» электродов протяженного ВШП (рис. 4.47, г). Каждый одиночный электрод заменяется двумя или более электродами, присоединенными к одной и той же общей шине. В этом случае на длину поверхностной акустической волны приходится более двух электродов и отражение от преобразователя также снижается. Недостатком такой конструкции является усложнение технологического процесса нанесения ВШП с малым межэлектродным расстоянием.

Как мы уже говорили, рабочая частота ПАВ-генератора в основном определяется длиной линии задержки, а также, в некоторой степени, от фазового сдвига в цепи обратной связи. Очевидно, что небольшой подстройки частоты можно добиться, введя в петлю обратной связи фазосдвигающее устройство. Для того чтобы в одномодовом режиме такая перестройка не привела бы к скачкообразной смене моды, изменение частоты не должно быть больше ширины основного лепестка частотной характеристики линии задержки. Следовательно, полное дополнительное изменение фазы не должно превышать $\pm\pi/2$. Из уравнения (4.106) следует, что изменение частоты δf, связанное со сдвигом фазы $\delta\phi$, равно:

$$\frac{\delta f}{f} = -\frac{\delta\phi}{2\pi}\frac{\Lambda}{L}. \qquad (4.109)$$

При $L = 100\Lambda$ диапазон перестройки частоты составляет $\delta f / f = \pm 0.25\ \%$. Если фазосдвигающей устройство сделать электрически управляемым, то ПАВ-генератор может работать в режиме ЧМ-модуляции.

Другая возможность небольшой перестройки частоты реализуется в ПАВ-генераторе, показанном на рис. 4.48. В общем акустическом потоке преобразователя *1* размещены два преобразователя *2* и *3*, сдвинутые на четверть длины акустической волны. Потенциометр,

включенный между преобразователями *2* и *3*, позволяет изменять частоту на величину $\Lambda/4L$.

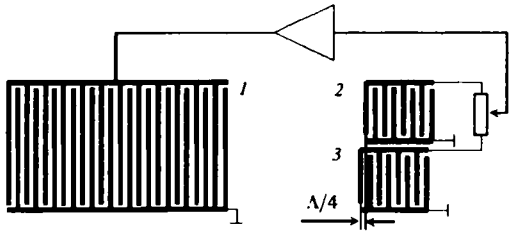


Рис. 4.48. Схема ПАВ-генератора с двойным преобразователем

В том случае, когда линия задержки ПАВ-генератора является широкополосной, существует возможность одновременной генерации нескольких мод. В этих условиях весьма важным становится нелинейное поведение усилителя. Если одна из мод превысила порог генерации, то коэффициент усиления для других мод уменьшается. Подобное явление наблюдается в лазерах, где линия излучения вещества существенно шире расстояния между модами резонатора, и носит название конкуренции мод. Когда коэффициент усиления выбран на много больше, чем порог генерации одной моды, то остальные моды не возбуждаются, то есть имеется тенденция к одномодовому режиму работы. При включении такого генератора амплитуда моды с наименьшими потерями растет быстрее всех других и при приближении к насыщению подавляет конкурирующие моды ниже порога их возбуждения. Начальная генерация возникает из широкополосного шума, одинакового для всех мод. Однако возбуждение одной из мод может происходить не из шума, а под действием более мощного внешнего сигнала на частоте, соответствующей выбранной моде. Данная мода первой достигнет насыщения и будет подавлять все другие конкурирующие моды, для которых усиление теперь станет очень малым.

Описанный выше эффект используется в ПАВ-генераторах с широкополосной перестройки частоты. Такие устройства называют также программируемыми ПАВ-генераторами или ПАВ-синтезаторами. На рис. 4.49 представлена структурная схема одного из вариантов ПАВ-синтезатора.

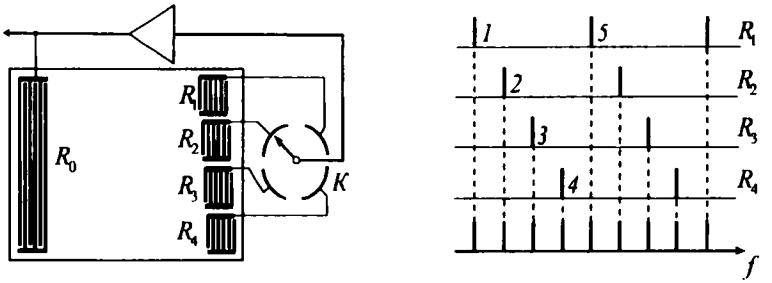


Рис. 4.49. ПАВ-синтезатор: структурная схема и принцип формирования спектра

В акустическом потоке входного преобразователя R_0 параллельно расположены четыре выходных преобразователя R_1 – R_4, сдвинутые относительно друг друга на $\Lambda/4$ вдоль направления распространения акустической волны. Роторный реверсивный безразрывный коммутатор K позволяет циклически чередовать каналы $R_1 - R_4$, включаемые в цепь обратной связи. Коммутатор может быть электронным и управляться, например, микроконтроллером. Все преобразователи являются широкополосными, то есть условие одномодового режима заведомо не выполняется. Пусть в момент включения роторный коммутатор подключен к преобразователю R_1 и установилась частота *1*. При переключении на канал преобразователя R_2 в генераторе возбудится ближайшая к частоте *1* частота *2* из спектра канала R_2. Изменение частоты будет равно $\Lambda/4L$. Последовательно подключая каналы R_3, R_4 и снова R_1 в прямом и обратном направлениях, можно получить последовательное дискретное изменение частоты в пределах полосы пропускании преобразователей. Более тонкое или даже плавное изменение частоты можно осуществить посредством замены простого коммутатора потенциометрическими цепями, включенными между преобразователями.

4.10. Полосовые фильтры на поверхностных акустических волнах

Фильтрами называют стационарные пассивные линейные системы, преобразующие спектра сигнала. Они находят весьма широкое применение в радиоэлектронике. Например, в качестве узкополосных фильтров часто применяются различные резонаторы на объемных и поверхностных акустических волнах. В данном разделе мы рассмотрим принципы построения ПАВ-фильтров на основе аподизованных ВШП. Доступность волнового поля ПАВ, малая длина волны и низкая скорость, а также простота управления свойствами ВШП позволяют создавать компактные, высокостабильные и дешевые фильтры с большим разнообразием характеристик.

Электрический фильтр, как и любую пассивную линейную систему, можно адекватно описать в частотной и временной области. Требования к большинству типов электрических фильтров задаются в частотной области, поэтому на практике обычно оперируют с частотными характеристиками.

Если на вход фильтра поступает монохроматический сигнал единичной амплитуды $U_{вх}(\omega) = \exp(j\omega t)$, то на выходе будет также присутствовать монохроматический сигнал той же частоты, но с другими значениями амплитуды и фазы, которые зависят от частоты входного сигнала:

$$U_{вых}(\omega) = A(\omega)\exp\left[j\omega t - j\Phi(\omega)\right].$$

Коэффициент передачи фильтра $H(\omega)$ равен:

$$H(\omega) = \frac{U_{вых}}{U_{вх}} = A(\omega)\exp\left[-j\Phi(\omega)\right]. \qquad (4.110)$$

Зависимость коэффициента передачи фильтра от частоты $H(\omega)$ называют передаточной характеристикой, а иногда также частотной характеристикой. Вещественные положительные функции $A(\omega) = |H(\omega)|$ и $\Phi(\omega)$ называют, соответственно, амплитудной (АЧХ) и фазовой (ФЧХ) частотными характеристиками фильтра.

Заметим, что если в спектральной полосе сигнала $A(\omega)$ является константой, а фазовый сдвиг линейно зависит от частоты, то сигнал передается без искажения, изменяется лишь его амплитуда и появляется некоторая временная задержка, одинаковая для всех спектральных составляющих сигнала. Например, если $\Phi(\omega) = t_0\omega$, то из (4.110) следует

$$U_{вых}(\omega) = A(\omega)\exp\left[j\omega(t - t_0)\right].$$

При синтезе структуры и топологии некоторых видов фильтров, в частности ПАВ-фильтров, весьма полезным является описание их свойств во временной области. Так как фильтр – линейное устройство, то при подаче на его вход сигнала произвольного спектрального состава $u_{вх}(t)$ сигнал на выходе $u_{вых}(t)$ должен линейно зависеть от $u_{вх}(t)$. Наиболее общей формой такой зависимости является свертка входного сигнала с некоторой функцией $h(t)$:

$$u_{вых}(t) = \int_{-\infty}^{\infty} h(t - t')u_{вх}(t')dt'. \qquad (4.111)$$

Очевидно, функция $h(t)$ должна удовлетворять условию $h(t) = 0$ при $t < 0$, так как сигнал на выходе не может появиться раньше входного сигнала. Если входной сигнал представляет собой бесконечно короткий дельта-импульс в момент $t' = 0$, то из (4.111) получаем

$$u_{вых}(t) = h(t) \text{ при } u_{вх}(t') = \delta(t'). \qquad (4.112)$$

На этом основании функцию $h(t)$ называют импульсным откликом или импульсной характеристикой фильтра.

Используя соотношения (4.110) и (4.111), нетрудно убедиться, что импульсный отклик и передаточная характеристика связаны парой преобразований Фурье:

$$H(\omega)=\int_{-\infty}^{\infty}h(t)\exp(-j\omega t)dt\,. \qquad (4.113)$$

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(\omega)\exp(j\omega t)d\omega\,. \qquad (4.114)$$

Таким образом, для полного описания фильтра достаточно определить его импульсный отклик, либо передаточную характеристику.

Рассмотрим фильтр, структура которого показана на рис. 4.50. Входной сигнал поступает на идеальную многоотводную линию задержки с равномерным расположением выводов. С выводов снимаются задержанные копии входного сигнала, которые затем подвергаются амплитудному взвешиванию и суммируются, образуя выходной сигнал. Фильтры данного вида получили название трансверсальных.

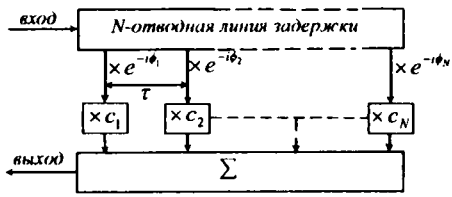


Рис. 4.50. Трансверсальный фильтр на основе многоотводной линии задержки

Если на входе трансверсального фильтра действует гармонический сигнал единичной амплитуды, то выходной сигнал есть передаточная характеристика фильтра:

$$H(\omega)=\sum_{k=1}^{N}c_k e^{-j\phi_k}\,, \qquad (4.115)$$

где ϕ_k – фазовый сдвиг сигнала, снимаемого с k-того отвода, N – число отводов. Предполагается, что взвешивание не приводит к дополнительной фазовой задержке, и, следовательно, все коэффициенты c_k – вещественные числа. Фазовый сдвиг ϕ_k может быть записан как $\phi_k=\omega x_k/\mathrm{v}$, где x_k – координата k-того отвода, v – скорость рас-

пространения сигнала в линии задержки. Тогда передаточная характеристика приобретает следующий вид:

$$H(\omega)=\sum_{k=1}^{N} c_k \exp\left(-j\frac{\omega}{v}x_k\right). \qquad (4.116)$$

Координата k-того отвода может быть выражена через временную задержку сигнала между двумя соседними отводами τ:

$$x_k = kv\tau .$$

Принимая во внимание это выражение, находим импульсный отклик трансверсального фильтра как обратное преобразование Фурье (4.114) передаточной характеристики (4.116):

$$h(t)=\sum_{k=1}^{N} c_k \delta(t-k\tau). \qquad (4.117)$$

При вычислении учитывается, что

$$\delta(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(j\omega t)d\omega . \qquad (4.118)$$

Введем гладкую вещественную функцию $u(t)$, значения которой в моменты времени $t_k = k\tau$ равняются коэффициентам c_k. Пусть, кроме того, $u(t)=0$ для $t<0$ и $t>t_N$. При сделанных предположениях (4.117) запишется как

$$h(t)=\sum_{k=-\infty}^{\infty} u(k\tau)\delta(t-k\tau)=u(t)\sum_{k=-\infty}^{\infty}\delta(t-k\tau). \qquad (4.119)$$

Это выражение представляет собой дискретную модель сигнала $u(t)$. Величину задержки между двумя соседними отводами линии τ называют интервалом дискретизации. Соответствующая частота $\omega_\tau = 2\pi/\tau$ называется частотой дискретизации. Так как

$$g(t) = \sum_{k=-\infty}^{\infty} \delta(t - k\tau)$$

– периодическая функция времени с периодом τ, ее можно представить рядом Фурье:

$$g(t) = \sum_{m=-\infty}^{\infty} g_m \exp(jm\omega_\tau t),$$

где коэффициенты разложение равны:

$$g_m = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} g(t) \exp(-j\omega_\tau mt)\, dt = \frac{\omega_\tau}{2\pi}.$$

Таким образом:

$$h(t) = u(t) \frac{\omega_\tau}{2\pi} \sum_{m=-\infty}^{\infty} \exp(jm\omega_\tau t).$$

Осуществляя преобразование Фурье, в результате находим

$$H(\omega) = \frac{\omega_\tau}{2\pi} \sum_{m=-\infty}^{\infty} U(\omega - m\omega_\tau), \qquad (3.120)$$

где $U(\omega)$ – преобразование Фурье от $u(t)$.

Амплитудно-частотная характеристика трансверсального фильтра представлена на рис. 4.51. Она представляет собой бесконечную сумму копий амплитудных спектров непрерывного сигнала $u(t)$. Так как $u(t)$ – вещественная функция, ее спектр симметричен относительно точки $\omega = 0$ (рис. 4.51, а). Предполагается, что $U(\omega)$ – функция полосового типа, то есть ее значения пренебрежимо малы вне области частот от ω_1 до ω_2. Если частота дискретизации превышает $2\omega_2$, полосы частот каждого из слагаемых в сумме (4.120) не перекрываются (рис. 4.51, б). При этом в области частот $|\omega| \le \omega_\tau/2$

$$H(\omega)=\frac{\omega_\tau}{2\pi}U(\omega). \qquad (4.121)$$

То есть при $\omega_\tau \geq 2\omega_2$ в полосе частот $|\omega| \leq \omega_\tau/2$ спектры непрерывного и дискретного сигналов с точностью до константы совпадают. Минимальное значение частоты дискретизации $\omega_{\tau\min}=2\omega_2$, при котором полосы частот отдельных слагаемых в сумме (4.120) не перекрываются, носит название частоты Найквиста. Если $\omega_\tau < 2\omega_2$, наблюдается явление наложения спектров, и частотная характеристика искажается (рис. 4.51, в).

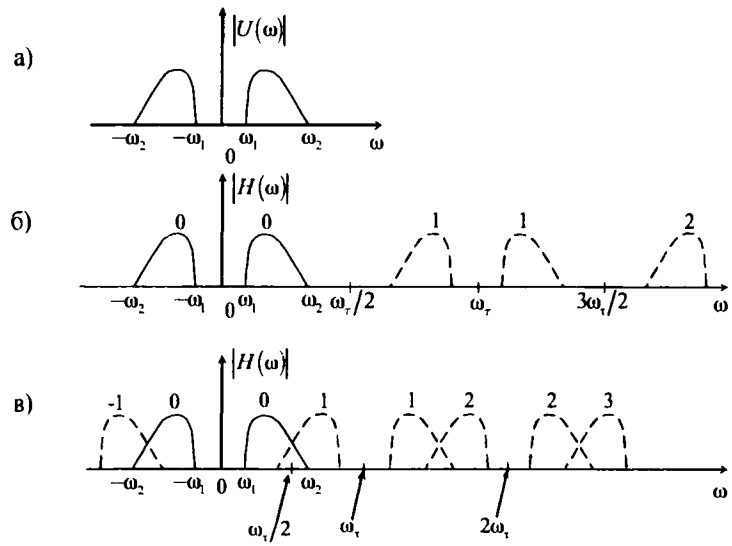


Рис. 4.51. Амплитудно-частотная характеристика трансверсального фильтра: а амплитудный спектр непрерывного сигнала; б) АЧХ фильтра при частоте дискретизации большей частоты Найквиста; в) наложение спектров при частоте дискретизации меньшей частоты Найквиста

На проведенном выше анализе основывается процедура создания трансверсального фильтра с требуемыми свойствами. Пусть задана некоторая частотная характеристика вплоть до частоты ω_2, выше

которой она обращается в ноль. Вначале выбирается сигнал $u(t)$ конечной длительности T, такой, что его преобразование Фурье хорошо аппроксимирует заданный коэффициент передачи на частотах $|\omega| \le \omega_2$ и незначительно отличается от нуля при $\omega > \omega_2$. Затем производится дискретизация сигнала $u(t)$ с периодом $\tau < 4\pi/\omega_2$, в результате чего находится последовательность действительных весовых коэффициентов $c_k = u(k\tau)$, количество отводов линии задержки $N = T/\tau + 1$ и шаг расположения отводов $d = \tau v$. Следует отметить, что в случае синтеза частотных характеристик с резкими переходами от полосы пропускания к полосе заграждения при выборе функции $u(t)$ возникают определенные трудности, которые будут рассмотрены ниже.

Помимо требуемой частотной характеристики трансверсальный фильтр имеет «зеркальные» характеристики на частотах выше $\omega_\tau/2$. Обычно их подавляют с помощью дополнительного фильтра нижних частот. Так как частота дискретизации выбирается выше частоты Найквиста, это позволяет снизить требования к крутизне спада АЧХ фильтра нижних частот.

Принципы построения ПАВ-фильтров на основе ВШП во многом базируются на теории рассмотренного выше трансверсального фильтра. Основанием к этому является эквивалентность частотной характеристики ВПШ, получаемой в модели дельта-источников (4.40) и частотной характеристики трансверсального фильтра (4.116). Пьезоэлектрическая подложка с нанесенными на ее поверхность электродами играет роль линии задержки, сами электроды являются отводами, а общие шины суммируют сигналы электродов. Таким образом, создание ВШП с заданной передаточной характеристикой требует взвешивания сигналов, снимаемых с электродов.

Существует два способа взвешивания. Один из них называют внешним взвешиванием, а другой – непосредственным. При внешнем взвешивании к одиночным изолированным электродам или секциям из нескольких электродов подводятся различные по фазе и напряжению сигналы. Это выполняется с помощью внешних емкостных или резистивных делителей, подключаемых к электродам преобразователя. Однако на практике используется редко, так как усложняет конструкцию преобразователя. При непосредственном взвешивании, называемом аподизацией, преобразователь имеет

только две общие шины для всех электродов. Необходимый импульсный отклик формируется путем изменения топологии преобразователя (рис. 4.52).

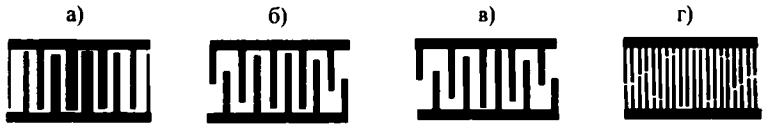


Рис. 4.52. Аподизованные ВШП: а – переменная ширина электродов; б – переменная апертура перекрытия штырей; в) переменная апертура перекрытия с добавлением пассивных электродов (изображены серыми); г) переменная апертура перекрытия с расщепленными штырями

Поскольку межэлектродная емкость пропорциональна ширине электродов, то изменением ширины (рис. 4.52, а) в принципе можно достичь требуемой функции аподизации и обеспечить заданный импульсный отклик преобразователя. Недостатком данного метода является весьма жесткое требование к точности получения электродов заданной ширины. Это обстоятельство является существенным на высоких частотах, когда ширина электродов составляет единицы микрон и менее.

Другая возможность заключается в изменении степени перекрытия электродов (рис. 4.50, б, в, г). Метод аподизации на рис. 4.52, б имеет два серьезных недостатка. Во-первых, на участках с малым перекрытием штырей наблюдается дифракция акустической волны. Во-вторых, скорость распространения поверхностной волны на металлизированной поверхности несколько ниже, чем на свободной. В преобразователе с переменной величиной перекрытия штырей степень металлизации неоднородна по апертуре, поэтому распространение ПАВ в преобразователе будет сопровождаться нарушением прямолинейности волнового фронта. Данные недостатки проявляются в виде значительного искажения импульсного отклика преобразователя.

С целью их преодоления в преобразователь вводят дополнительные пассивные электроды, которые не принимают участия в генерации сигнала (рис. 4.52, в). При этом дифракция и искажение фронта волны существенно уменьшаются, хотя полного устранения этих явлений не происходит, так как активные и пассивные электроды по-разному влияют на распространение ПАВ.

В настоящее время этот метод является самым распространенным инструментом аподизации.

Так как для точного воспроизведения требуемого импульсного отклика необходимы протяженные преобразователи с большим числом штырей, коэффициент отражения от них будет довольно значительным. Для уменьшения отражения применяют аподизованные преобразователи с расщепленными штырями (рис. 4.52, г). Однако в этом случае также серьезно возрастают требования к точности нанесения штырей.

Типичный ПАВ фильтр содержит два преобразователя, один из которых является неаподизованным широкополосным, а второй аподизованным – формирует заданную частотную характеристику (рис. 4.53, а). Иногда оба преобразователя выполняются аподизованными. Но в этом случае для устранения влияния искажений волнового фронта связь между преобразователя осуществляется через многополосковый ответвитель (рис. 4.53, б). Применение МПО, как уже говорилось, оправдывает себя только на подложках из сильных пьезоэлектриков. Передаточная характеристика ПАВ-фильтра с двумя преобразователями определяется выражением

$$H(\omega)=\sum_{k=1}^{N}\sum_{m=1}^{M}c_k c_m \exp\left[-j\frac{\omega}{v}(x_k-x_m)\right], \qquad (4.122)$$

где N и M – число электродов входного и выходного ВШП, c_k и c_m – дискретные выборки импульсных откликов, x_k и x_m – координаты расположения штырей. Поскольку в данном случае оба преобразователя влияют на передаточную характеристику, возможности синтеза фильтра с заданными параметрами расширяются.

Рассмотрим процедуру аподизации ВШП с целью создания полосового фильтра с равномерным пропусканием в полосе частот $[\omega_1,\omega_2]$, максимально возможным подавлением за пределами этой полосы и линейной ФЧХ. Фильтры данного типа находят широкое применение при конструировании различной радиоэлектронной аппаратуры, например телевизионных приемников.

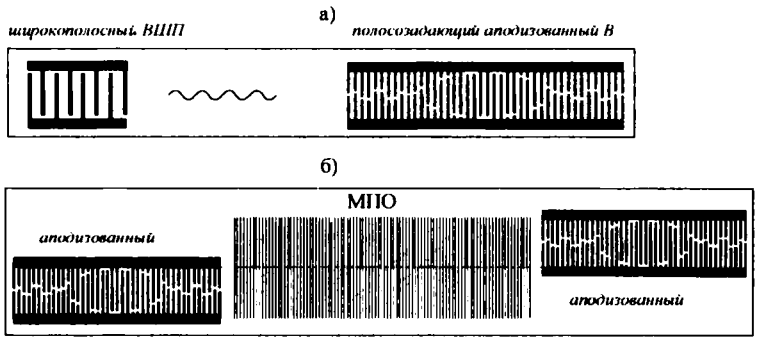


Рис. 4.53. ПАВ-фильтры: а – фильтр с широкополосным неаподизованным и полосозадающим аподизованным ВШП; б) фильтр с двумя аподизованными ВШП, связанными через МПО

Передаточная характеристика идеального полосового фильтра имеет вид

$$H(\omega)=\begin{cases} A_0 \exp(-j\omega_0 t), & \omega_1 \le \omega \le \omega_2, \\ 0, & \omega < \omega_1, \omega > \omega_2, \end{cases} \qquad (4.123)$$

где ω_0 – центральная частота полосы пропускания. АЧХ этого фильтра и его импульсный отклик представлены на рис. 4.54 (а и б). Так как преобразование Фурье прямоугольной функции имеет бесконечную протяженность, то в результате дискретизации импульсного отклика получается бесконечная последовательность отсчетов. Очевидно, ВШП с бесконечным числом штырей не может быть реализован. Импульсный отклик ВШП всегда конечен, то есть он является фильтром с конечной импульсной характеристикой (КИХ-фильтр). Поэтому при синтезе ВШП используют некоторое ограниченное число отсчетов (рис. 4.54, в). В результате АЧХ, получаемая с помощью преобразования Фурье от ограниченного дискретного набора отсчетов идеального импульсного отклика, приобретает вид, показанный на рис. 4.54, г. Наблюдаются значительные пульсации в полосе пропускания, плохое подавление сигнала за пределами этой полосы и невысокая крутизна срезов на ее границах. В большинстве случаев данная АЧХ является неприемлемой. Увеличение количества отсчетов импульсного отклика увеличивает крутизну срезов, но уровень пульсаций АЧХ в полосе про-

пускания и за ее пределами не уменьшается. Такое поведение АЧХ получило название явления Гиббса.

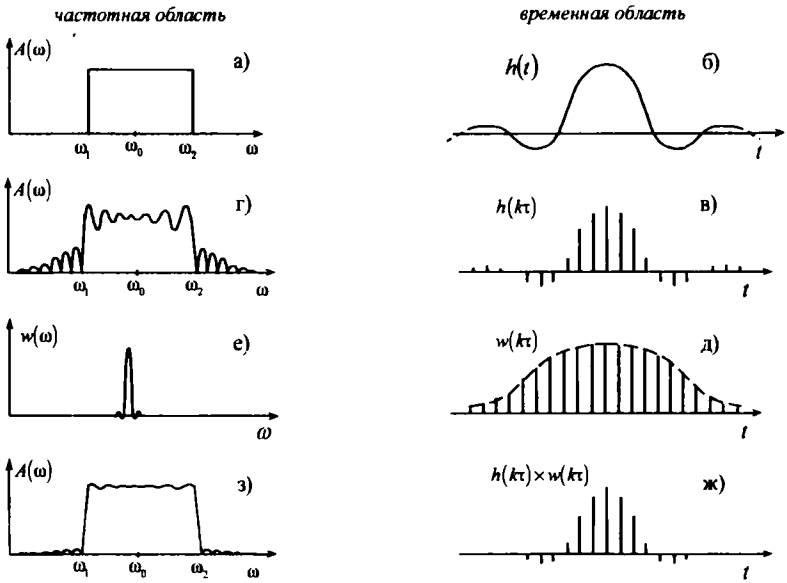


Рис. 4.54. Дискретизация импульсного отклика ВШП при синтезе полосового фильтра: а – АХЧ идеального полосового фильтра; б – его импульсный отклик; в – дискретные отсчеты импульсного отклика; г – АЧХ, получаемая в результате дискретизации и ограничения количества отсчетов; д – оконная функция; е – спектр оконной функции; ж – отсчеты импульсного отклика, взвешенные оконной функцией; з – АЧХ взвешенного преобразователя

Широко распространенным и достаточно эффективным способом борьбы с явлением Гиббса служит использование так называемых оконных функций. Оконная функция – это некоторая гладкая медленная функция, спектр которой имеет достаточно узкий основной лепесток и малый уровень боковых лепестков (рис. 4.54, д, е). Дискретные отсчеты импульсного отклика взвешиваются путем умножения на последовательность значений оконной функции и используются для аподизации ВШП (рис. 4.54, ж). При правильном подборе оконной функции АЧХ фильтра в полосе пропускания заметно выравнивается, а подавление за ее пределами увеличивается (рис. 4.54, з). Отметим, что простое усечение бесконечной последо-

вательности отсчетов импульсного отклика эквивалентно применению прямоугольной оконной функции.

Действие оконной функции легко понять, если учесть, что умножение во временной области эквивалентно свертке в частотной области. Поэтому АЧХ, получаемая в результате взвешивания отсчетов импульсного отклика, равна свертке идеальной АЧХ со спектром оконной функции. Таким образом, уровень пульсаций в полосе пропускания и подавление за ее пределами будет зависеть от амплитуды боковых лепестков оконной функции в частотной области, а крутизна среза от ширины основного лепестка. К сожалению, требование низкого уровня боковых лепестков и узости основного лепестка являются взаимно противоречивыми. Однако для практики найдено большое количество видов оконных функций, позволяющих получать достаточно хорошие результаты. Наиболее известны функции Хэмминга и Кайзера. Вид этих и других применяемых функций можно найти в литературе. Прямоугольная функция окна при достаточно большом числе отсчетов имеет узкий основной лепесток, но значительный уровень боковых лепестков, который с ростом числа отсчетов не уменьшается. Поэтому ее использование приводит к неудовлетворительным результатам.

Рассмотренные выше частотные и временные параметры фильтров определялись только их собственными свойствами. Такие параметры реализуются лишь при полном согласовании фильтра с внешними цепями и называются характеристическими. Параметры же, реализуемые в конкретных условиях, вообще говоря, отличаются от характеристических и называются рабочими.

4.11. Дисперсионные ПАВ-фильтры

Дисперсионный фильтр представляет собой устройство, предназначенное для получения задержки, существенно изменяющейся с частотой. Их часто также называют дисперсионными линиями задержки.

Основные применения дисперсионных фильтров связаны с использованием частотно-модулированных сигналов особого вида, несущая частота которых линейно изменяется со временем (ЛЧМ-сигналы). ЛЧМ-сигнал с прямоугольной огибающей имеет вид

$$a(t)=\begin{cases} a_0 \exp\left[j2\pi\left(f_0 t+\frac{\alpha}{2}t^2 \right)\right], & 0\le t\le T, \\ 0, & t<0, t>T, \end{cases} \qquad (4.124)$$

где a_0 – амплитуда, T – длительность сигнала, f_0 – частота при $t=0$, α – параметр, определяющий скорость изменения фазы. Мгновенная частота сигнала

$$f=\frac{d}{dt}\left(f_0 t+\frac{\alpha}{2}t^2 \right)=f_0+\alpha t \qquad (4.125)$$

линейно зависит от времени. Если $\alpha>0$ частота возрастает, а при $\alpha<0$ – убывает. Мгновенная частота изменяется в диапазоне $f_0\le f\le f_0+B$, где $B=|\alpha|T$ – ширина спектра. Спектр ЛЧМ-сигнала выражается через интегралы Френеля, однако при $BT>>1$ амплитудный спектр ЛЧМ сигнала может быть приближенно аппроксимирован прямоугольной функцией, а фазовый – параболой. Обращает на себя внимание тот факт, что ширина спектра ЛЧМ-сигнала при увеличении длительности растет, тогда как у обычного прямоугольного импульса с неизменной несущей спектр при увеличении длительности сужается. ЛЧМ-сигнал может создаваться дисперсионным фильтром, у которого задержка монохроматического сигнала линейно зависит от его частоты.

Дисперсионные фильтры широко применяются в радиолокационных станциях (РЛС). В обычной РЛС расстояние определяется по времени задержки электромагнитного импульса, отраженного от объекта. Ясно, что для обнаружения удаленных объектов и получения высокого разрешения по дальности излучаемый импульс должен быть одновременно коротким и иметь большую энергию. Это требует создания передатчиков с высоким значением пиковой мощности, что связано с определенными техническими трудностями. Применение дисперсионных фильтров существенно облегчает решение данной задачи. На рис. 4.55 показана схема обработки сигналов в РЛС с применением дисперсионных фильтров. С помощью фильтра ДФ1 короткий импульс с широким спектром трансформируется в ЛЧМ-цуг. ЛЧМ-сигнал можно сделать достаточно продолжительным, поэтому при той же мощности передатчика энергия ЛЧМ-импульса

будет значительно больше энергии короткого импульса, излучаемого обычной РЛС. После отражения от объекта сигнал поступает в приемник, где обрабатывается с помощью дисперсионного фильтра ДФ2. Импульсный отклик этого фильтра является обращенной во времени копией отраженного сигнала. Такой фильтр называют согласованным. В результате принятый ЛЧМ-цуг вновь трансформируется в короткий импульс, который имеет форму острого корреляционного пика с шириной примерно $1/B$. Для фильтра с $BT >> 1$ эта ширина будет намного меньше длительности входного сигнала, поэтому согласованные фильтры также часто называют фильтрами сжатия импульсов. Величина BT называется коэффициентом сжатия. По существу, сжатие импульса эквивалентно повышению пиковой мощности передатчика в BT раз. Если в обычной РЛС перекрытие отраженных импульсов ограничивает разрешение по дальности, то в РЛС со сжатием импульса, цуги ЛЧМ-сигналов, отраженных от близких объектов, могут значительно перекрываться.

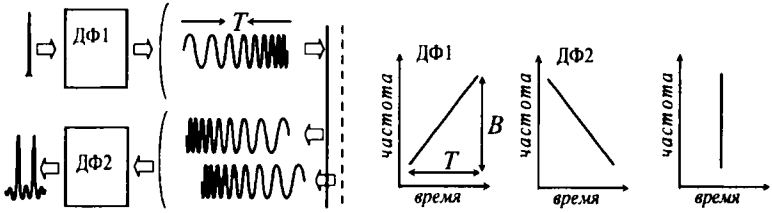


Рис. 4.55. Использование дисперсионных фильтров в РЛС для разрешения сигналов удаленных целей

Важное для практических целей свойство ЛЧМ-сигнала состоит в том, что форма его огибающей при амплитудной модуляции повторяет форму зависимости его спектральной плотности от частоты. На этом свойстве ЛЧМ-сигналов основана работа устройств, в которых с помощью дисперсионных фильтров выполняется преобразование Фурье и выходной сигнал оказывается пропорциональным спектру входного сигнала.

Дисперсионные фильтры успешно реализуются в виде компактных устройств на поверхностных акустических волнах. Для получения задержки, зависящей от частоты, можно использовать естественную дисперсию некоторых видов волн (волны Лява, Лэмба), соз-

давая на поверхности звукопровода соответствующие покрытия. Однако большее распространение получили методы, использующие недиспергирующие поверхностные волны (волны Рэлея), а необходимая дисперсионная характеристика получается за счет применения неэквидистантных ВШП либо неоднородных отражательных решеток. В этом случае появляются большие возможности управления характером и величиной дисперсии. В неэквидистантном ВШП шаг между штырями плавно изменяется вдоль преобразователя. При соответствующем подборе шага короткий импульс на входе преобразователя трансформируется на его выходе в цуг ЛЧМ-колебания с длительностью, определяемой длиной преобразователя (рис. 4.56). Шаг между штырями на одном из концов преобразователя соответствует наинизшей синхронной частоте, а на противоположном конце – наивысшей. При достаточно большой протяженности ВШП можно перекрыть большой интервал частот, для которых электроды преобразователя будут эффективно взаимодействовать с акустической волной.

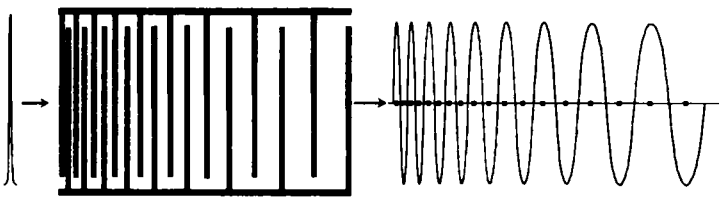


Рис. 4.56. Неэквидистантный ВШП – генератор ЛЧМ-сигналов

Зависимость фазы от координаты x вдоль преобразователя, предназначенного для генерации ЛЧМ-импульса длительностью T и полосой частот B, на основании (4.124) имеет вид

$$\Phi(x)=\frac{2\pi f_0}{V}x+\frac{\pi B}{V^2T}x^2, \qquad (4.126)$$

где V – скорость ПАВ. Набег фазы волны при похождении от первого до k-го штыря должен равняться $(k-1)\pi$. Полагая, что координата первого штыря $x_1=0$, получаем уравнение для определения положения остальных штырей:

$$\frac{2f_0}{V}x_k + \frac{B}{V^2T}x_k^2 = k - 1. \qquad (4.127)$$

Так как длина преобразователя $x_N = VT$, то из (3.127) находим общее число штырей:

$$N = 1 + 2f_0T + BT. \qquad (4.128)$$

Обычно $B << f_0$ и $BT >> 1$, поэтому $N \approx 2f_0T$. Из (4.126) можно также найти зависимость расстояния между соседними штырями от координаты по длине преобразователя:

$$d(x) = \frac{1}{2}\frac{V}{f_0 + xB/(VT)}. \qquad (4.129)$$

Используя широкополосные свойства неэквидистантного ВШП, можно создать линию задержки для видеоимпульсов (рис. 4.57).

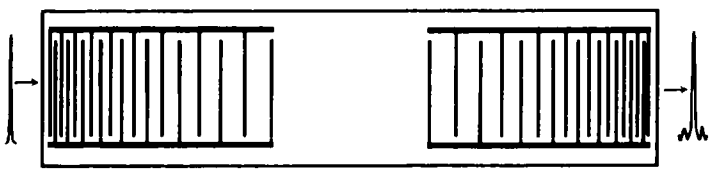


Рис. 4.57. Широкополосная линия задержки для коротких импульсов на снове двух неэквидистатнтных ВШП

Входной ВШП преобразует подаваемый на вход видеоимпульс в ЛЧМ-цуг. Закон расположения штырей в выходном преобразователе выбирается обратным по отношению к входному, то есть выходной преобразователь является согласованным фильтром. В результате на выходе линии задержки формируется задержанный импульс, который по своим параметрам близок к исходному. Импульсной отклик выходного согласованного ВШП, как обращенная во времени копия ЛЧМ-сигнала (4.124), определится выражением

$$h(t) = \exp\left[j2\pi\left(f_0t + Bt - \frac{B}{2T}t^2 \right)\right], \quad (0 \le t \le T), \qquad (4.130)$$

где для простоты полагается $a_0 = 1$. Тогда сигнал на выходе равен:

$$u_{вых} = \int_0^T a(t)h(t-t')dt . \qquad (4.131)$$

Выполняя необходимые вычисления и опуская осциллирующий фазовый множитель, получаем следующую форму огибающей сигнала:

$$u_{вых} = \frac{\sin[\pi B(t-T)]}{\pi B(t-T)} . \qquad (4.132)$$

График этой функции показан на рис. 4.58. Огибающая достигает максимума при $t = T$, первые нули расположены в точках $t = T \pm 1/B$, а ширина основного лепестка по уровню 4дБ $\Delta t \cong 1/B$. Кроме главного максимума, имеется ряд побочных, ближайшие из которых только на 13 дБ меньшей амплитуды основного, что часто является неприемлемым для практики.

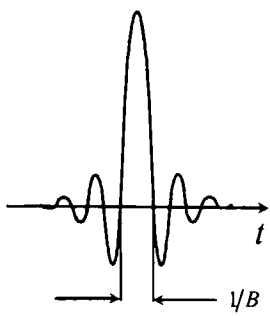


Рис. 4.58. Сигнал на выходе линии задержки импульсов с двумя дисперсионными преобразователями, один из которых обращен во времени

Появление побочных максимумов можно объяснить тем, что частотная характеристика фильтра ограничена прямоугольным окном. Следовательно, для подавления побочных лепестков могут быть использованы оконные функции, применение которых обсуждалось в предыдущем разделе. Однако в данном случае

процедура взвешивания выполняется не во временной, а в частотной области.

Если неэквидистантный преобразователь возбуждается непрерывным монохроматическим сигналом, то энергия излучается главным образом той областью ВШП, где расстояние между электродами близко к половине длины волны. Эффективное число электродов зависит от дисперсионной характеристики и может быть легко оценено исходя из того, что полная фазовая расстройка в пределах рабочего диапазона частот не должна превышать π:

$$N_e/N = 2/\sqrt{BT}\,. \tag{4.133}$$

Изменение частоты Δf, требуемое для перехода от одного ряда действующих электродов к соседнему ряду, определится как $B(N_e/N)$, то есть

$$\Delta f = 2\sqrt{B/T}\,. \tag{4.134}$$

До сих пор предполагалось, что на возбуждаемую волну не влияют остальные части преобразователя, не находящиеся в синхронизме. Однако в длинных преобразователях проявляются нежелательные эффекты в виде отражения волны, дифракции, а также некоторое ее затухание, связанное с отводом энергии штырями при распространении вдоль звукопровода. Все эти явления при большой величине BT сильно искажают частотную характеристику фильтра. Данные проблемы могут быть решены с помощью наклонных преобразователей (рис. 4.59). Если угол наклона θ выбран таким, что число электродов на каждом участке $W\cos\theta$ (W – апертура ВШП) близко к N_e, то различные спектральные компоненты распространяются в параллельных каналах, не пересекая других частей преобразователя.

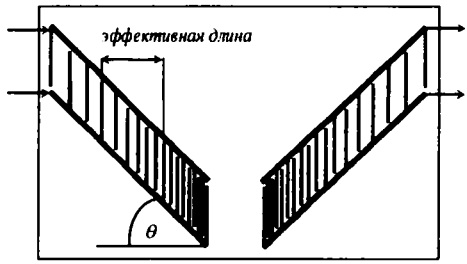


Рис. 4.59. ПАВ-фильтр с двумя наклонными преобразователями. Длина взаимодействия с акустической волной много меньше протяженности дисперсионных преобразователей

При использовании наклонных ВШП преодолеваются также проблемы, связанные с генерацией объемных волн. Обычный короткий преобразователь не согласован по фазе с объемными волнами, так как скорость волны Рэлея всегда меньше скоростей объемных мод. Но в протяженном дисперсионном ВШП участки с большим периодом могут эффективно взаимодействовать с объемными волнами на верхних частотах рабочего диапазона. Ввиду того, что, как уже говорилось, различные частотные составляющие попадают на различные участки, возбуждение объемных волн не происходит.

Недостатком наклонной конструкции является неэффективное использование звукопровода, которое особенно сказывается при больших значениях BT.

Помимо неэквидистантных ВШП, для создания дисперсионных фильтров могут использоваться неоднородные отражательные решетки. Одна из возможных структур дисперсионного фильтра на основе неоднородных отражательных решеток представлена на рис. 4.60. Поверхностные волны, излучаемые входным широкополосным преобразователем, испытывают двукратное отражение и достигают выходного преобразователя. Отражательные решетки могут представлять собой параллельные разомкнутые металлические полоски, либо канавки на поверхности звукопровода с глубиной, существенно меньшей длины акустической волны. Так как шаг решетки является переменным, длина U-образной траектории для каждой спектральной компоненты различна.

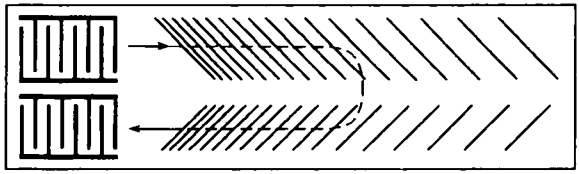


Рис. 4.60. Дисперсионный фильтр на основе неоднородных отражательных структур типа «елочка»

Применение отражательных решеток вместо протяженных наклонных преобразователей имеет ряд серьезных преимуществ. Отражательные решетки мало чувствительны к дефектам изготовления, тогда как замыкание одной пары ВШП делает весь преобразователь неработоспособным. Распространение ПАВ определяется двумя отражениями. Объемные волны и другие ложные сигналы практически не достигают приемного преобразователя с фронтом, параллельным электродам. Таким образом, ложные отклики эффективно подавляются. Поскольку частотные компоненты сигнала распространяются исключительно по U-образным траекториям, легко вводить дополнительную фазовую коррекцию, необходимую для компенсации дефектов нанесения решетки, с помощью тонкой пленки, наносимой между решетками. Кроме того, при полной заданной дисперсии требуется вдвое меньшая длина подложки по сравнению с обычной конструкцией с линейным расположением преобразователей.

4.12. Акустические конвольверы, корреляторы и запоминающие устройства

Как было показано в разделе 3, при достаточно сильном механическом или электрическом воздействии материальные среды становятся нелинейными. Вследствие этого распространение акустических волн большой амплитуды сопровождается нелинейными эффектами, которые не наблюдаются, когда амплитуды волн малы. Например, при распространении волны с частотой ω появляется вторая гармоника 2ω, а при значительной интенсивности волны – и высшие гармоники. В случае нелинейного взаимодействия волн

с частотами ω_1 и ω_2 возникают гармоники с комбинационными частотами $k\omega_1 \pm m\omega_2 \quad (k, m = 1, 2, 3, ...)$.

Генерация гармоник происходит при коллинеарном (попутном или встречном) распространении, а также когда волновые векторы взаимодействующих волн образуют некоторый угол. В акустоэлектронике чаще всего используется встречное распространение, что позволяет осуществлять ряд полезных для практических целей функциональных преобразований сигналов. Соответствующие устройства называются акустическими конвольверами.

В акустических конвольверах могут быть использованы как объемные, так и поверхностные акустические волны. Очевидно, что эффективность нелинейного взаимодействия будет тем выше, чем больше объемная плотность энергии взаимодействующих волн. Тот факт, что определяющую роль для проявления нелинейного взаимодействия играет именно плотность энергии, а не сама энергия, делает использование поверхностных волн более предпочтительным. ПАВ распространяются в поверхностном слое толщиной порядка длины волны Λ. Если в некотором устройстве для исключения дифракции требуется акустический пучок шириной $N\Lambda$, то при заданной мощности взаимодействующих волн в объемном устройстве, требующем пучок с обоими поперечными размерами $N\Lambda$, плотность энергии в N раз меньше, чем в аналогичном устройстве на ПАВ. Поскольку типичное значение $N \sim 50$, это представляет собой весьма ощутимое преимущество.

Простейший конвольвер на поверхностных акустических волнах состоит из пьезоэлектрической подложки, на противоположных концах которой размещены встречно-штыревые преобразователи, а между ними – металлический поверхностный электрод, называемый параметрическим (рис. 4.61). В вырожденных ПАВ-конвольверах параметрический электрод является сплошным и используется совместно с симметричным электродом на тыльной стороне подложки. В невырожденных конвольверах он образован двумя вложенными гребенками электродов и аналогичен по структуре ВШП. Обычно входные сигналы $u_1(t)$ и $u_2(t)$ поступают на преобразователи, которые возбуждают встречные акустические волны, а результат их нелинейного взаимодействия в виде выходного сигнала $u_3(t)$ считывается параметрическим электродом. При некоторых видах обра-

ботки сигналов один из преобразователей используется как выход конвольвера, а параметрический электрод – как вход.

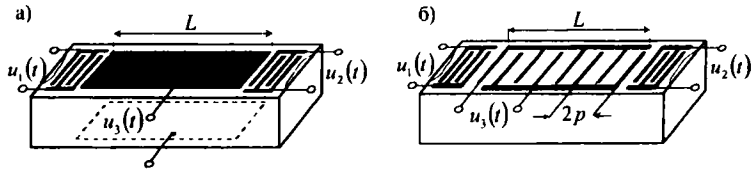


Рис. 4.61. ПАВ-конвольверы: а – вырожденный конвольвер, б – невырожденный ПАВ-конвольвер

Для взаимодействия акустических волн можно использовать нелинейные свойства самого пьезоэлектрического звукопровода, однако в ПАВ-конвольверах между параметрическим электродом и пьезоэлектрической подложкой обычно размещают полупроводниковую пленку. Так как концентрационная нелинейность полупроводника, как правило, наступает раньше, чем другие виды нелинейности, это позволяет получать генерацию комбинационных частот при меньшей мощности входных сигналов. В дальнейшем при анализе работы конвольвера будем считать, что нелинейность достаточно слаба. Это, во-первых, означает, что генерацией гармоник выше второй можно пренебречь. Во-вторых, позволяет считать, что нелинейное взаимодействие волн не вызывает их затухания по мере распространения вдоль звукопровода.

Пусть на ВШП с обеих сторон конвольвера поступают гармонические сигналы с частотами ω_1 и ω_2. Тогда вдоль звукопровода навстречу друг другу будут распространяться две акустические волны:

$$\begin{aligned} f(x,t) &= a_1 \cos(\omega_1 t - \kappa_1 x), \\ g(x,t) &= a_2 \cos(\omega_2 t + \kappa_2 x), \end{aligned} \qquad (4.135)$$

где a_1, a_2 – амплитуды, $\kappa_1 = \omega_1/v$ и $\kappa_2 = \omega_2/v$ – волновые числа, v – скорость ПАВ. Вследствие нелинейности, помимо волн (4.135), на поверхности звукопровода будут присутствовать возмущения вида:

$$f^2(x,t)=a_1^2\cos^2(\omega_1 t-\kappa_1 x)=\frac{1}{2}a_1^2+\frac{1}{2}a_1^2\cos\left[2(\omega_1 t-\kappa_1 x)\right],$$
$$g^2(x,t)=a_2^2\cos^2(\omega_2 t+\kappa_2 x)=\frac{1}{2}a_2^2+\frac{1}{2}a_2^2\cos\left[2(\omega_2 t+\kappa_2 x)\right], \quad (4.136)$$

а также возмущение, обусловленное взаимодействием волн:

$$f(x,t)g(x,t)=\frac{a_1 a_2}{2}\left\{\cos\left[(\omega_1+\omega_2)t+(\kappa_1-\kappa_2)x\right]+\right.$$
$$\left.+\cos\left[(\omega_1-\omega_2)t+(\kappa_1+\kappa_2)x\right]\right\}. \quad (4.137)$$

Генерацию комбинационных частот $\omega_1\pm\omega_2$ можно трактовать как появление третьего фонона с частотой $\omega_3=\omega_1\pm\omega_2$ в результате столкновения двух фононов с частотами ω_1 и ω_2. Условия данного взаимодействия представляют собой законы сохранения энергии и квазиимпульса фононов, которые при встречном распространении имеют следующий вид:

$$\hbar\omega_3=\hbar\omega_1\pm\hbar\omega_2,$$
$$\kappa_3\hbar=\kappa_1\hbar\mp\kappa_2\hbar. \quad (4.138)$$

Это соотношения удобно представить в виде дисперсионных диаграмм на плоскости (κ,ω) (рис. 4.62). В отсутствие дисперсии фазовой скорости каждому из двух сталкивающихся фононов соответствует некоторый вектор. Углы, образуемые векторами с осью частот, равны. Сумма и разность этих векторов будет давать частоту и пространственный период получаемого колебания. При столкновении фононов (ω_1,κ_1) и (ω_2,κ_2) появляются колебания с частотами $\omega_1\pm\omega_2$ и пространственными периодами $2\pi/(\kappa_1\mp\kappa_2)$. В вырожденном случае, когда частоты и постоянные распространения встречных волн равны ($\omega_1=\omega_2=\omega$ и $\kappa_1=\kappa_2=\kappa$), возникает постоянное поле, периодическое в пространстве: $\omega_3=0$, $\kappa_3=2\kappa$, и переменное колебание удвоенной частоты, однородное по всей длине взаимодействия: $\omega_3=2\omega$, $\kappa_3=0$. Следует отметить, что колебания комбинационных частот $\omega_1\pm\omega_2$ не являются поверхност-

ными волнами, так как отношения $(\omega_1+\omega_2)/(\kappa_1-\kappa_2)$ и $(\omega_1-\omega_2)/(\kappa_1+\kappa_2)$ не равны фазовой скорости ПАВ. Следовательно, данные колебания существуют только в области взаимодействия встречных волн.

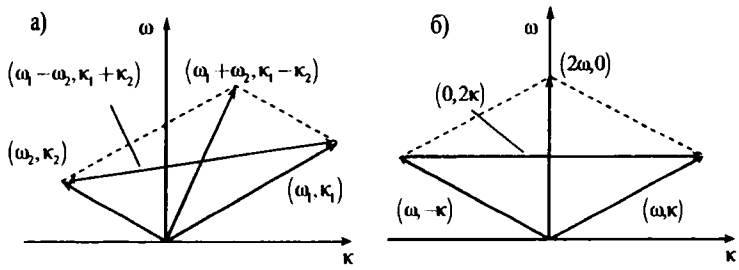


Рис. 4.62. Дисперсионная диаграмма нелинейного взаимодействия встречных волн в недиспергирующей среде: а – частоты и постоянные распространения встречных волн отличаются; б – частоты и постоянные встречных волн равны

Каждому из указанных выше колебаний соответствует некоторое электрическое поле в пьезоэлектрическом материале. Параметрический электрод конвольвера позволяет селективно считывать электрическое поле, порождаемое взаимодействующими волнами.

В вырожденном конвольвере сплошной параметрический электрод интегрирует электрическое поле по всей длине взаимодействия, считывая лишь те составляющие, которые не зависит от координаты x. Чтобы показать это, рассмотрим интеграл

$$\int_{-L/2}^{L/2}\cos(\kappa x-\omega t)\,dx=2\pi\cos\omega t\,\frac{L}{2\pi}\,\frac{\sin(\kappa L/2)}{(\kappa L/2)}. \qquad (4.139)$$

При достаточно большой длине параметрического электрода правая часть отличается от нуля только в малой окрестности точки $\kappa=0$. Если $L\to\infty$, то

$$\int_{-\infty}^{\infty}\cos(\kappa x-\omega t)\,dx=2\pi\cos\omega t\,\delta(\kappa)=2\pi v\cos\omega t\,\delta(\omega). \qquad (4.140)$$

На основании этого соотношения можно сделать вывод, что вклад в выходной сигнал будут давать только члены, описывающие пространственно независимое колебание. Если частоты входных сигналов ω_1 и ω_2 равны, именно такой оказывается составляющая с суммарной частотой в выражении (4.137). Что касается постоянных полей, присутствующих в (4.136), то ввиду того, что постоянный ток через параметрический электрод не проходит, их вклад в выходной сигнал равен нулю.

Таким образом, при достаточно большой протяженности параметрического электрода напряжение на выходе вырожденного конвольвера возникает только при равенстве частот входных гармонических сигналов. Частота этого напряжения равна удвоенной частоте сигналов, а амплитуда пропорциональна произведению амплитуд сигналов. Такие устройства получили название билинейных. Можно показать, что

$$V_3 = \sqrt{P_1 P_2}\, M/W\,, \qquad (4.141)$$

где V_3 – амплитуда выходного напряжения холостого хода на частоте 2ω, P_1 и P_2 – мощности входных сигналов, W – апертура параметрического электрода, которая равна апертурам преобразователей, M – параметр нелинейности. Считается, что явления дифракции, дисперсии и затухания волн несущественны. Параметр нелинейности M зависит только от свойств и ориентации подложки. В принципе, он может быть рассчитан, но обычно определяется из эксперимента. Для $LiNbO_3$ Y,Z -среза $M = 1.2 \cdot 10^{-4}$ В×м/Вт.

На практике эффективность конвольвера определяют с помощью коэффициента билинейности C, связывающего мощности сигналов, подводимых к входам, с мощностью выходного сигнала, рассеиваемого в нагрузке. Поскольку устройство является билинейным, то мощность на выходе P_0 пропорциональна произведению входных мощностей $P_1 P_2$, что дает возможность ввести коэффициент билинейности [дБ]:

$$C = 10 \lg\left(P_0 / P_1 P_2\right), \qquad (4.142)$$

где все мощности выражают в милливаттах. Коэффициент билинейности учитывает все виды потерь в устройстве, в том числе и потери, обусловленные затуханием и дифракцией волн.

Если частоты входных сигналов различаются, то пространственный период составляющей с суммарной частотой составляет $2\pi/|\kappa_2 - \kappa_1|$. Эту составляющую можно выделить с помощью встречно-штыревой структуры, подобной ВШП (рис. 4.61, б). Для этого необходимо, чтобы шаг структуры удовлетворял условию $p = \pi/|\kappa_1 - \kappa_2|$. Конвольвер с такой структурой параметрического электрода называется невырожденным. Серьезным недостатком невырожденного конвольвера является интенсивное взаимодействие параметрического электрода с объемными волнами, что существенно ограничивает его применение. Далее мы ограничимся исследованием методов обработки сигналов с помощью вырожденного конвольвера.

Пусть на встречно-штыревые преобразователи вырожденного конвольвера поступают некоторые сложные сигналы, описываемые действительными функциями $u_1(t)$ и $u_2(t)$. На вид этих сигналов не накладывается никаких ограничений, кроме требования конечной длительности. В остальном функции $u_1(t)$ и $u_2(t)$ остаются произвольными.

Если считать полосы пропускания входных ВШП достаточными для неискаженной передачи сигналов, дисперсия отсутствует, а уровни мощности сигналов малы, то мгновенные значения поля ПАВ в некоторой точке звукопровода x описываются функциями $u_1(t - L/2\mathrm{v} - x/\mathrm{v})$ и $u_2(t - L/2\mathrm{v} + x/\mathrm{v})$. Далее для краткости будем отбрасывать постоянную временную задержку сигналов в конвольвере $L/2\mathrm{v}$ и записывать мгновенное распределение поля ПАВ в виде $u_1(t - x/\mathrm{v})$ и $u_2(t + x/\mathrm{v})$.

Вследствие нелинейности подложки в каждой точке звукопровода помимо волн $u_1(t - x/\mathrm{v})$ и $u_2(t + x/\mathrm{v})$, генерируются колебания $u_1^2(t - x/\mathrm{v})$, $u_2^2(t + x/\mathrm{v})$, а также $u_1(t - x/\mathrm{v})u_2(t + x/\mathrm{v})$. Представим каждую из волн интегралом Фурье:

$$u_1(t-x/\mathrm{v})=\int_0^{\infty}\Phi_1(\omega)\cos(\omega t-\kappa x+\varphi_1(\omega))d\omega,$$
$$u_2(t-x/\mathrm{v})=\int_0^{\infty}\Phi_2(\omega)\cos(\omega t-\kappa x+\varphi_2(\omega))d\omega \qquad (4.143)$$

где $\Phi_1(\omega)$, $\Phi_2(\omega)$ и $\varphi_1(\omega)$, $\varphi_2(\omega)$ – соответственно модули и фазы спектров сигналов $u_1(t)$ и $u_2(t)$. Рассмотрим сигнал, создаваемый на параметрическом электроде исходными волнами $u_1(t-x/\mathrm{v})$, $u_2(t+x/\mathrm{v})$ и возникающими из-за нелинейности подложки колебаниями $u_1^2(t-x/\mathrm{v})$, $u_2^2(t+x/\mathrm{v})$ и $u_1(t-x/\mathrm{v})u_2(t+x/\mathrm{v})$. Если протяженность цугов волн $u_1(t-x/\mathrm{v})$ меньше длины параметрического электрода, то интегрирование по длине параметрического электрода можно заменить интегрированием в бесконечных пределах.

Вклад в суммарное напряжение на параметрическом электроде будут давать следующие интегралы:

$$I_1(t)=\int_{-\infty}^{\infty}u_1(t-x/\mathrm{v})dx,\quad I_2(t)=\int_{-\infty}^{\infty}u_2(t+x/\mathrm{v})dx,$$
$$I_3(t)=\int_{-\infty}^{\infty}u_1^2(t-x/\mathrm{v})dx,\quad I_4(t)=\int_{-\infty}^{\infty}u_2^2(t+x/\mathrm{v})dx,$$
$$I_5(t)=\int_{-\infty}^{\infty}u_1(t-x/\mathrm{v})u_2(t+x/\mathrm{v})dx.$$

Подставляя в эти интегралы соотношения (4.143) и применяя формулу (4.140), получаем:

$$I_1(t)=\pi\mathrm{v}\Phi_1(0)\cos\varphi_1(0),\quad I_2(t)=\pi\mathrm{v}\Phi_1(0)\cos\varphi_1(0),$$
$$I_3(t)=\pi\mathrm{v}\int_0^{\infty}\Phi_1^2(\omega)d\omega,\qquad I_4(t)=\pi\mathrm{v}\int_0^{\infty}\Phi_2^2(\omega)d\omega,$$
$$I_5=\pi\mathrm{v}\int_0^{\infty}\Phi_1(\omega)\Phi_2(\omega)\cos[2\omega t+\varphi_1(\omega)+\varphi_2(\omega)]d\omega.$$

Так как параметрический электрод не пропускает постоянного тока, сигнал на выходе конвольвера будет определяться только интегралом I_5:

$$u_3(t) \sim \int_{-\infty}^{\infty} u_1(t - x/\mathrm{v}) u_2(t + x/\mathrm{v}) dx . \qquad (4.144)$$

Введя переменную $\tau = t - x/\mathrm{v}$, преобразуем (4.144) к следующему виду:

$$u_3(t) \sim \int_{-\infty}^{\infty} u_1(\tau) u_2(2t - \tau) d\tau . \qquad (4.145)$$

Это соотношение выражает важнейшее свойство акустического конвольвера: сигнал на параметрическом электроде является сверткой входных сигналов, сжатой во времени в два раза. Эффект сжатия сигнала во времени физически объясняется тем, что здесь относительная скорость взаимодействующих волн равна $2\mathrm{v}$, а не v, как в устройствах на ПАВ, где акустическая волна взаимодействует со статическими неоднородностями.

Рассмотрим различные применения конвольвера для обработки сигналов.

Если не учитывать сжатия во времени, выходной сигнал конвольвера формально совпадает с выходным сигналом фильтра, несмотря на то, что последний является линейным устройством. Выходной сигнал фильтра представляет собой свертку входного сигнала с импульсной характеристикой фильтра. В акустическом конвольвере роль импульсной характеристики выполняет один из входных сигналов, который в этом случае называют опорным. Изменяя вид опорного сигнала, можно произвольно изменять импульсный отклик устройства. Таким образом, акустический конвольвер может применяться в качестве линейного фильтра с возможностью динамической перестройки его характеристик.

Если один из входных сигналов инвертирован во времени, то на параметрическом электроде выделяется функция корреляции сигналов. Пусть для определенности сигнал u_1 инвертирован. Тогда из выражения (4.145), имеем

$$u_3(t) \sim \int_{-\infty}^{\infty} u_2(-\tau) u_2(2t-\tau)\, d\tau = \int_{-\infty}^{\infty} u_1(\tau) u_2(2t+\tau)\, d\tau . \qquad (4.146)$$

То есть выходной сигнал равен функции корреляции сигналов $u_1(t)$ и $u_2(t)$, также сжатой во времени в два раза.

Обращение сигнала во времени может быть осуществлено с помощью устройства на рис. 4.63, а. Принцип его работы можно объяснить следующим образом. Когда на одном из входов конвольвера действует некоторый сигнал $u(t)$, а на противоположном – его обращенная копия $u(-t)$, то, как было ранее установлено, на параметрическом электроде формируется сигнал корреляции, который в данном случае будет иметь вид острого корреляционного пика. На основании принципа взаимности можно сразу заключить, что если теперь на один из входов подать сигнал $u(t)$, а на параметрический электрод дельта-импульс, то должна возбуждаться волна, распространяющаяся в направлении, противоположном исходной волне, которую возбуждает сигнал $u(t)$. В результате на том же входе формируется обращенный во времени сигнал $u(-t)$. Его можно отделить от исходного сигнала $u(t)$ с помощью циркулятора. Важно подчеркнуть, что инвертированный сигнал не сжат и не растянут во времени.

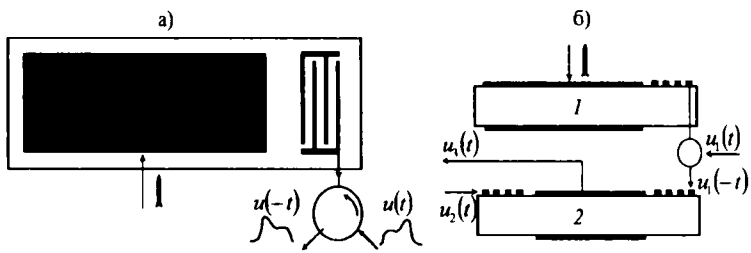


Рис. 4.63. Нелинейное устройство обращения сигналов во времени (а) и коррелятор на его основе (б). Задержанный и обращенный во времени сигнал отделяется от входного посредством циркулятора

На рис. 4.63, б показана схема коррелятора на основе эффекта обращения времени. Коррелятор состоит из двух конвольверов.

Конвольвер *1* выполняет функцию обращения во времени входного сигнала $u_1(t)$, который затем поступает на один из входов конвольвера *2*, а на второй его вход поступает сигнал $u_2(t)$. Функция корреляции сигналов $u_1(t)$ и $u_2(t)$, сжатая в два раза по времени, выделяется на параметрическом электроде конвольвера *2*.

Конвольвер на поверхностных волнах может быть использован как запоминающее устройство. Для этого между параметрическим электродом и поверхностью пьезоэлектрического звукопровода размещается матрица диодов Шотки (рис. 4.64). Запоминаемый электрический сигнал $u_1(t)$ поступает на преобразователь, расположенный у левого края подложки. Когда волновой пакет целиком находится в области взаимодействия, на параметрический электрод подается короткий отрицательный импульс, смещающий диоды в прямом направлении. Время переключения диодов Шотки составляет доли наносекунд. Диоды быстро накапливают заряды, пропорциональные потенциалу полного электрического поля. Это поле представляет собой сумму пространственно независимого поля смещающего импульса и поля, обусловленного поверхностной акустической волной. После окончания действия импульса записи, накопленные в диодах заряды, создают обратное смещение и запирают диоды. Постоянная времени разряда закрытых диодов определяется токами утечки и может достигать 100 мс. Таким образом, мгновенное распределение поля ПАВ фиксируется в диодной матрице.

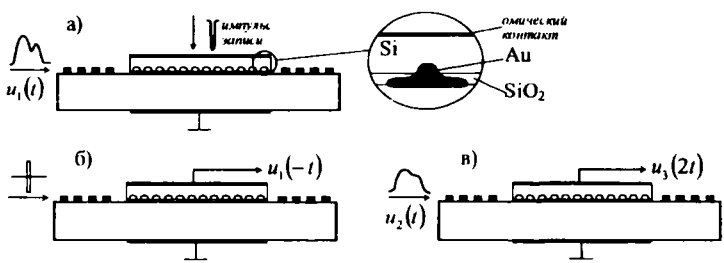


Рис. 4.64. Запоминающий конвольвер с диодами Шотки: а – запись сигнала; б – считывание, в – корреляционная обработка с использованием запомненного сигнала

Для того чтобы распределение, записанное в виде накопленных зарядов, точно описывало поданный сигнал, расстояние между дио-

дами в соответствии с принципом дискретизации должно быть меньше половины длины ПАВ, а длительность импульса смещения должна быть меньше полупериода несущей частоты сигнала.

Записанный сигнал может быть воспроизведен через некоторое время. Для этого на преобразователь, через который производилась запись, подается короткий импульс, возбуждающий широкополосную ПАВ (рис. 4.64, б). Поле этой ПАВ нелинейно взаимодействует с постоянным полем, обусловленным накопленным в диодах зарядом. В результате на параметрическом электроде возникает сигнал, являющийся обращенной во времени копией записанного сигнала $u_1(-t)$. Если импульс считывания подать на противоположный преобразователь, то можно получить сигнал без обращения во времени. Необходимо отметить, что процедура считывания является неразрушающей. По существу, при считывании лишь немного изменяется ширина обедненной зоны диода, а переноса заряда через эту зону не происходит. Поэтому записанный сигнал допускает многократное считывание без ощутимого искажения.

В общем случае, если сигнал $u_2(t)$ подан на входной преобразователь, через который производилась запись сигнала $u_1(t)$ (рис. 4.64, в), то выходной сигнал на параметрическом электроде $u_3(t)$ является корреляционной функцией $u_1(t)$ и $u_2(t)$:

$$u_3(t) \sim \int_{-\infty}^{\infty} u_1(\tau) u_2(t+\tau)\, d\tau . \qquad (4.147)$$

В отличие от обычного конвольвера здесь нет необходимости в предварительном обращении одного из сигналов, а получаемая функция корреляции не сжата во времени. Очевидно, если входной сигнал $u_2(t)$ подать на противоположный преобразователь, выходной сигнал будет сверткой сигналов $u_1(t)$ и $u_2(t)$, которая также не будет сжата во времени. Причина отсутствия сжатия во времени заключается в том, что в отличие от обычного конвольвера, в запоминающем конвольвере один из взаимодействующих сигналов является не бегущей волной, а статическим неоднородным распределением электрического поля.

4.13. Устройства спектрального анализа сигналов. Фурье-процессоры

В разделе 4.11 было показано, что дисперсионные фильтры играют ключевую роль в РЛС со сжатием импульса. В этом параграфе обсуждаются их применения, связанные со спектральным анализом. Речь будет идти о фильтрах, импульсная характеристика которых представляет собой ЛЧМ-сигнал. Общие принципы спектральной обработки не зависят от технической реализации. Однако использование устройств на ПАВ, в особенности фильтров сжатия на ОР, может дать существенные технические преимущества.

Одной из возможностей спектральной обработки сигналов является система измерения частоты, известная под названием приемника со сжатием импульса, который содержит два дисперсионных фильтра – фильтр расширения и фильтр сжатия. Частота непрерывного гармонического сигнала определяется по времени задержки выходного импульса, которое линейно зависит от частоты. Поскольку система является линейной, то с ее помощью можно анализировать сигнал, состоящий из суммы нескольких гармонических составляющих; при этом каждой из этих составляющих соответствует свой выходной импульс. Отсюда можно заключить, что форма выходного сигнала характеризует его амплитудный спектр.

Преобразование Фурье некоторого сигнала можно получить с помощью Фурье-процессоров. Они подобны приемнику со сжатием импульса, но содержат не два, а три дисперсионных фильтра. Как в приемнике со сжатием импульса, так и в Фурье-процессорах время выполнения преобразования обычно порядка $1/\Delta f$, где Δf – разрешающая способность по частоте. Таким образом, преобразование осуществляется значительно быстрее, чем обычным анализатором спектра, в котором сигнал сначала смешивается с колебанием гетеродина качающейся частоты, а затем поступает на выходной узкополосный фильтр.

Структурная схема приемника со сжатием импульса показана на рис. 4.65. Входной гармонический сигнал, имеющий угловую частоту Ω, умножается на ЛЧМ-сигнал, получаемый в результате подачи короткого импульса на вход фильтра расширения. Результат перемножения поступает на фильтр сжатия, на выходе которого формируется короткий импульс. Нижняя часть рисунка поясняет принцип действия устройства с помощью частотно-временных характеристик.

Предполагается, что огибающие импульсных характеристик обоих фильтров, входящих в систему, прямоугольные. Характеристики, соответствующие расширенному импульсу и сигналу на входе фильтра сжатия, сдвинуты по частоте на величину Ω. Полоса частот, занимаемая сигналом на входе фильтра сжатия, полностью перекрывает диапазон его частотной характеристики. Для этого центральные частоты обоих фильтров должны различаться, а полоса пропускания фильтра расширения должна быть шире полосы пропускания фильтра сжатия. Кроме того, значение Ω не может выходить за пределы определенного частотного интервала, равного разности полос пропускания обоих устройств. Характеристики, показанные на рисунке, соответствуют двум крайним значениям частоты Ω, обозначенным Ω_1 и Ω_2. При некотором значении частоты сигнала на входе выходной сигнал фильтра сжатия имеет огибающую, приближенно описываемую выражением $\sin(\pi B_c t)$, где B_c – полоса пропускания фильтра сжатия. Эта функция напоминает функцию, описывающую выходной сигнал системы со сжатием ЛЧМ-импульса без взвешивания.

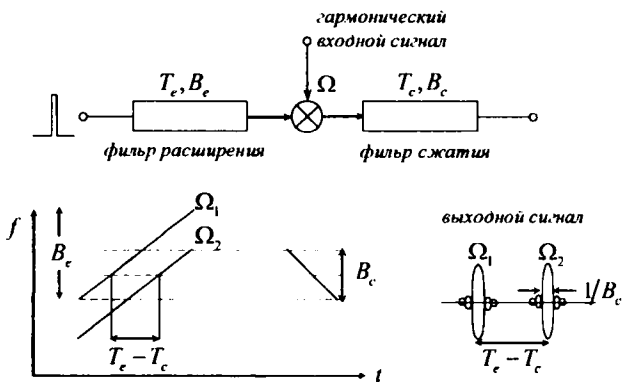


Рис. 4.65. Структурная схема приемника со сжатием импульса и частотно-временные характеристики, поясняющие принцип действия

Как видно из рис. 4.65, спектр расширенного импульса попадает в полосу пропускания фильтра сжатия в моменты времени, изменяющиеся в зависимости от Ω. Таким образом, время появления выходного импульса зависит от частоты входного сигнала, причем

эта зависимость линейная, так как в системе используются ЛЧМ-фильтры. Значение частоты входного сигнала можно определить, детектируя выходной импульс и измеряя его задержку относительно импульса, возбуждающего фильтр расширения.

Важными параметрами являются разрешающая способность по частоте и число разрешаемых точек. Если известен параметр μ – наклон дисперсионной характеристики фильтра расширения, то время задержки выходного сигнала τ находится из соотношения $\Omega = -2\pi\mu\tau + const$. Длительность выходного импульса приближенно равна $1/B_c$, следовательно, разрешающая способность $\Delta f = |\mu| B_c = 1/T_c$, где T_c – длительность импульсной характеристики фильтра сжатия. Диапазон входных частот равен $(B_e - B_c)$, где B_e – полоса пропускания фильтра расширения. Поэтому число разрешаемых точек определится как $(B_e - B_c)T_c$.

По технологическим причинам желательно, чтобы значение произведения $T_e B_e$ фильтра расширения было минимально допустимым. Но уменьшение этого произведения ухудшает разрешение по частоте. Минимально допустимое значение $T_e B_e$ можно определить следующим образом. Пусть $B_c = kB_e$, тогда число разрешаемых точек равно $(k - k^2)T_e B_e$. Дифференцируя это выражение, убеждаемся, что при заданном числе точек минимальное значение $T_e B_e$ получается, если $k = 1/2$. Таким образом, полоса пропускания фильтра расширения должна быть в два раза шире полосы пропускания фильтра сжатия. Соответствующее число разрешаемых точек равно $T_e B_e/4$.

Если перемножитель линеен по отношению к сигналу, то система получается линейной в том смысле, что входному сигналу, состоящему из нескольких гармонических колебаний, соответствует последовательность выходных импульсов, появляющихся в разные моменты времени. Однако выходной сигнал, как будет далее показано, не является преобразованием Фурье от входного колебания, а соответствует функции, называемой скользящим преобразованием Фурье. На практике это различие обычно не принципиально.

На практике в фильтре сжатия часто применяют амплитудное взвешивание, что позволяет снизить уровень боковых лепестков вы-

ходного импульса, как это делается в РЛС со сжатием импульса. При этом удается снизить уровень боковых лепестков выходного импульса и, следовательно, разрешать пару спектральных составляющих, имеющих близкие частоты, но различный уровень мощности. По сравнению с системой без амплитудного взвешивания здесь несколько ухудшается разрешающая способность и, соответственно, уменьшается число разрешаемых точек, так как ширина основного лепестка выходного импульса несколько увеличивается. Однако этот недостаток можно скомпенсировать, расширив полосы пропускания фильтров.

В основе работы приемника со сжатием импульса и Фурье процессоров лежит глубокая связь между теорией преобразования Фурье и теорией сигналов с внутриимпульсной ЧМ. Рассмотрим сигнал $f(t)$, имеющий преобразование Фурье $F(\omega)$. Заменив t на τ и ω на $2\pi\mu t$, получаем

$$F(2\pi\mu t) = \int_{-\infty}^{\infty} f(\tau)\exp(-j2\pi\mu t\tau)\,d\tau. \qquad (4.148)$$

С помощью тождества $2t\tau = t^2 + \tau^2 - (t-\tau)^2$ это равенство преобразуется к следующему виду:

$$F(2\pi\mu t) = \exp(-j2\pi\mu t^2)\int_{-\infty}^{\infty} f(\tau)\exp(-j2\pi\mu\tau^2)\exp\left[j\pi\mu(t-\tau)^2\right]d\tau. \qquad (4.149)$$

Эта форма записи преобразования Фурье называется ЛЧМ-преобразованием (ЛЧМП). Эквивалентная форма для дискретных сигналов, применяемая при анализе цифровых устройств и приборов с зарядовой связью, известна как ЛЧМZ-преобразование (ЛЧМZП). Преобразование вида (4.149) представляет собой умножение на комплексный сигнал с внутриимпульсной ЧМ, свертку с подобным сигналом и, наконец, умножение на третий аналогичный сигнал. Соответственно, эту процедуру называют П–С–П схемой. Если ввести функции

$$h_e(t) = h_m(t) = \exp\left(-j\pi\mu t^2\right),$$
$$h_c(t) = \exp\left(j\pi\mu t^2\right), \tag{4.150}$$

то в символической форме выражение (4.150) можно записать в следующем виде:

$$F(2\pi\mu t) = h_m(t)\left[\left\{f(t)h_e(t)\right\} * h_c(t)\right]. \tag{4.151}$$

На рис. 4.66 показана реализация этого процесса с помощью дисперсионных фильтров.

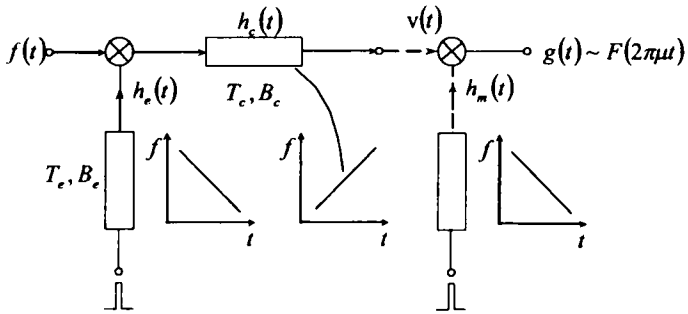


Рис. 4.66. Фурье-процессор типа П-С-П

Сигналы $h_e(t)$ и $h_m(t)$ получаются путем ударного возбуждения дисперсионных фильтров с импульсными характеристиками $h_e(t)$ и $h_m(t)$. Свертка с $h_c(t)$ осуществляется путем пропускания сигнала через фильтр с импульсной характеристикой $h_c(t)$. Наконец, умножение на $h_m(t)$ изменяет фазу выходного сигнала, но не его амплитуду. Следовательно, если требуется определить лишь амплитудный спектр сигнала, то последнее перемножение можно опустить. В этом случае система, по сути дела, становится аналогичной приемнику со сжатием импульса. Структурная схема, приведенная на рис. 4.66, интерпретирует выражение (4.151) довольно упрощенно, так как на практике импульсные характеристики ре-

альных устройств являются действительными и имеют конечную длительность. Более подробный анализ, учитывающий эти факторы, проводится далее.

Дуальная форма рассмотренной выше системы изображена на рис. 4.67. Она осуществляет свертку, последующее умножение и, наконец, еще одну свертку и носит название С–П–С-схемы.

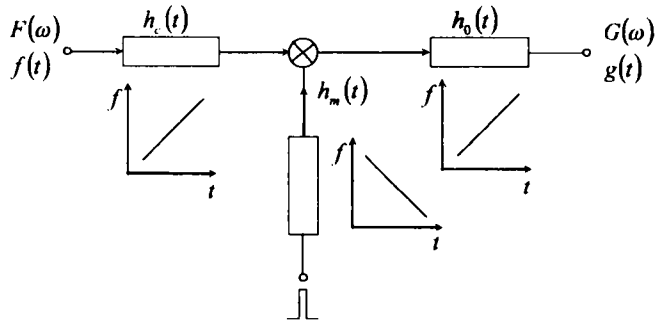


Рис. 4.67. Фурье-процессор типа С–П–С

Входной сигнал $f(t)$ свертывается с $h_c(t)$, умножается на $h_m(t)$ и вновь свертывается с $h_0(t)$. Показать, что в результате получается преобразование Фурье, можно, рассмотрев данные сигналы в частотной области. Обозначим символами $H_c(\omega)$, $H_m(\omega)$ и $H_0(\omega)$ частотные коэффициенты передачи фильтров, являющиеся преобразованием Фурье от $h_c(t)$, $h_m(t)$ и $h_0(t)$ соответственно. Умножение на $h_m(t)$ соответствует свертке с $H_m(\omega)$ в частотной области согласно теореме о свертке. Если $F(\omega)$ – спектр входного сигнала, то спектр $G(\omega)$ выходного сигнала имеет вид

$$G(\omega)=\frac{1}{2\pi}H_0(\omega)\left[\left\{F(\omega)H_c(\omega)\right\}*H_m(\omega)\right], \quad (4.152)$$

что аналогично выражению (4.151). Убедимся, что (4.151) определяет преобразование Фурье входного сигнала. Для этого заметим, что

преобразование Фурье от $\exp\left(j\pi\mu t^2\right)$ пропорционально $\exp\left(-j\alpha\omega^2\right)$, где $\alpha = 1/(4\pi\mu)$. Поэтому, опуская несущественные мультипликативные коэффициенты, можно принять $H_c(\omega) = H_0(\omega) = \exp\left(-j\alpha\omega^2\right)$ и $H_m(\omega) = \exp\left(j\alpha\omega^2\right)$. Подстановка этих соотношений в выражение (4.152) дает

$$G(\omega) = f(-\omega/2\pi\mu), \tag{4.153}$$

где $f(t)$ – исходный входной сигнал. Тогда из определения интеграла Фурье следует, что выходной сигнал описывается зависимостью $g(t) = \mu F(2\pi\mu t)$, в сущности тождественной результату, получаемому по схеме П–С–П.

Для более строгого описания функционирования схемы Фурье-процессора, как уже говорилось выше, необходимо принять во внимание тот факт, что длительность импульсных характеристик устройств конечна. Для краткости ограничимся только схемой П–С–П. Далее для простоты будем считать, что импульсные характеристики фильтров центрированы относительно момента времени $t = 0$. Конечно же импульсные характеристики реальных устройств, подчиняющиеся принципу причинности, центрированы относительно некоторых положительных значений t, таких, чтобы сигнал обращался в нуль при $t < 0$. Однако такое допущение связано с постоянной задержкой и не влияет на форму результирующего выходного сигнала.

В П–С–П системе, изображенной на рис. 4.66, выходной сигнал фильтра сжатия имеет вид

$$v(t) = \left[f(t)h_e(t)\right] * h_c(t), \tag{4.154}$$

что одновременно является выходным сигналом приемника со сжатием импульса. Пусть ω_e – центральная частота импульсной характеристики фильтра расширения, μ – скорость изменения частоты и T_e – длительность импульса. Тогда

$$h_e(t) = \text{rect}(t/T_e)\cos(\omega_e t - \pi\mu t^2), \qquad (4.155)$$

где $\text{rect}(x) = 1$, если $|x| \le 1$ и 0, если $|x| > 1$. Определяя аналогичным образом величины ω_c и T_c для фильтра сжатия, получаем

$$h_c(t) = \text{rect}(t/T_c)\cos(\omega_c t - \pi\mu t^2), \qquad (4.156)$$

Подставив эти соотношения в выражение (4.154), находим:

$$\begin{aligned} v(t) = &\int_{-\infty}^{\infty} f(\tau)\,\text{rect}(\tau/T_e)\cos(\omega_e\tau - \pi\mu\tau^2)\times \\ &\times rect\left[(t-\tau)/T_c\right]\cos\left[\omega_c(t-\tau) + \pi\mu(t-\tau)^2\right]d\tau. \end{aligned} \qquad (4.157)$$

В этом выражении произведение косинусов дает косинусы суммы и разности аргументов. Слагаемое, содержащее разностный аргумент, пренебрежимо мало, так как сигналы можно считать узкополосными, а этот член соответствует отрицательным частотам для $h_e(t)$ и положительным для $h_c(t)$ или наоборот. Поэтому

$$v(t) = \int_{-\infty}^{\infty} f(\tau)\,\text{rect}(\tau/T_e)\,rect\left[(t-\tau)/T_c\right]\cos(\omega_c t + \pi\mu t^2 - \Omega\tau)\,d\tau, \qquad (4.158)$$

где Ω линейно зависит от t в соответствии с формулой

$$\Omega(t) = \omega_c - \omega_e + 2\pi\mu t. \qquad (4.159)$$

Рассмотрим теперь частные случаи, когда длительность импульсной характеристики фильтра расширения мала $(T_e < T_c)$ и когда она велика $(T_e > T_c)$. Как выше отмечалось, последний случай относится к приемнику со сжатием импульса.

Вначале нужно выяснить роль функций rect( ) в выражении (4.158). Если все остальные сомножители в подынтегральном выра-

жении равны единице, то при $T_e < T_c$ сигнал $v(t)$ принимает постоянное значение для $|t| < (T_c - T_e)/2$ и имеет линейные фронт и срез. Таким образом, результирующая длительность трапецеидального импульса равна $T_c + T_e$. Область, в которой значение сигнала постоянно, соответствует интервалу времени, в течение которого расширенный импульс «целиком» находится внутри фильтра сжатия. Ясно, что система способна создавать на выходе преобразование Фурье только в моменты времени, соответствующие области постоянства сигнала. Поэтому проводимое далее рассмотрение относится только к этой области.

Оказывается, что в этой области, то есть для $|t| < (T_c - T_e)/2$, в выражении (4.158) достаточно учесть лишь первую функцию $\mathrm{rect}(\ )$, в то время как второй функцией $\mathrm{rect}(\ )$ можно пренебречь. Тогда

$$\bar{v}(t) = \frac{1}{4}\exp\left(j\omega_c t + j\pi\mu t^2\right)\int_{-\infty}^{\infty} f_1(\tau)\exp(-j\Omega\tau)\,d\tau + \qquad (4.160)$$

+комплексно сопряженная величина.

Для удобства введена функция

$$f_1(t) = f(t)\,rect(t/T_c), \qquad (4.161)$$

описывающая отрезок входного сигнала длительностью T_c. Ясно, что интеграл в выражении (4.160) равен $F_1(\Omega)$, где $F_1(\omega)$ – преобразование Фурье от $f_1(t)$. Если $\Phi_1(\omega)$ – фаза $F_1(\omega)$, то выходной сигнал фильтра сжатия представляется в виде

$$v(t) = \frac{1}{2}\left|F_1(\Omega)\right|\cos\left[\omega_c t + \pi\mu t^2 + \Phi_1(\Omega)\right]$$
$$\text{при } |t| < \frac{1}{2}(T_c - T_e). \qquad (4.162)$$

Отсюда следует, что огибающая выходного сигнала пропорциональна амплитудному спектру сигнала $f_1(t)$, описываемого выражением

(4.161). Частотная переменная Ω линейно связана с временной переменной для выходного сигнала, как видно из формулы (4.159). Так как в системе преобразуется отрезок входного сигнала длительностью T_c, то разрешающая способность системы приближенно равна $1/T_c$. Кроме того, выходной сигнал описывает спектр лишь в ограниченной полосе частот. Эта полоса шириной $\Delta\Omega$ определяется длительностью интервала времени, в течение которого справедливо выражение (4.162), и получается подстановкой в (4.159), откуда $\Delta\Omega = 2\pi(B_e - B_c)$. Скорость изменения частоты μ в обоих выражениях может быть положительной или отрицательной; в последнем случае выходной сигнал представляет собой не прямое, а обратное преобразование Фурье.

Если необходимо определить не только амплитудный, но и фазовый спектр сигнала, то следует провести дополнительное умножение на сигнал $h_m(t)$ с внутриимпульсной ЧМ, как показано на рис. 4.66. Это исключает квадратичный фазовый член $\pi\mu t^2$ в выходном сигнале фильтра сжатия, который описывается выражением (4.162). Скорость изменения частоты сигнала $h_m(t)$ равна $(-\mu)$, а его длительность должна, по крайней мере, превышать $T_c - T_e$. Обозначая центральную частоту сигнала $h_m(t)$ символом ω_m, получаем выходной сигнал:

$$g(t) = \frac{1}{4}\left|F_1(\Omega)\right|\cos\left[(\omega_c + \omega_m)t + \Phi_1(\Omega)\right]$$
$$\text{при } |t| < \frac{1}{2}(T_c - T_e). \qquad (4.163)$$

Здесь предполагается, что член с разностной частотой подавлен вследствие соответствующей фильтрации.

Рассмотрим далее случай $T_e > T_c$, что имеет место в приемнике со сжатием импульса. Выходной сигнал $v(t)$, как и ранее, описывается выражением (4.158). Функции rect( ) сами по себе дают постоянный результат при $|t| < (T_e - T_c)/2$. Полная длительность выходного

сигнала, как и ранее, равна $T_e + T_c$. В интервале $|t| < (T_e - T_c)/2$ первая из функций rect( ) в выражении (4.158) может быть опущена, тогда

$$v(t) = \frac{1}{4}\exp\left(j\omega_c t + j\pi\mu t^2\right)\int_{-\infty}^{\infty} f_2(\tau, t)\exp(-j\Omega\tau)\,d\tau + \quad (4.164)$$

+комплексно сопряженная величина,

где

$$f_2(\tau, t) = f(\tau)\,rect\left[(t-\tau)/T_c\right], \quad (4.165)$$

а $\Omega(t)$ по-прежнему определяется выражением (4.159). Обозначим символом $F_2(\Omega, t)$ интеграл в выражении (4.164). Он представляет собой преобразование функции $f_2(\tau, t)$ из пространства τ в пространство Ω. Если функция $F_2(\Omega, t)$ является аргументом $F_2(\Omega, t)$, то

$$v(t) = \frac{1}{2}\left|F_2(\Omega, t)\right|\cos\left[\omega_c t + \pi\mu t^2 + \Phi_2(\Omega, t)\right] \quad (4.166)$$

при $|t| < \frac{1}{2}(T_e - T_c)$.

Этот результат сходен с найденным для короткого фильтра расширения, но полученное здесь преобразование не вполне совпадает с преобразованием Фурье от входного сигнала. Можно показать, что

$$F_2(\Omega, t) = \int_{t-T_c/2}^{t+T_c/2} f(\tau)\exp(-j\Omega\tau)\,d\tau. \quad (4.167)$$

Отсюда следует, что для любых значений t выходной сигнал соответствует одной точке спектра отрезка входного сигнала $f(t)$. В то же время положение этого отрезка зависит от текущего значения t,

поэтому выражение (4.167) называют скользящим преобразованием Фурье. В общем случае получить компоненты преобразования Фурье входного сигнала из скользящего преобразования невозможно. Однако если сигнал представляет собой конечную сумму гармонических колебаний, то абсолютное значение выходного сигнала при скользящем и при обычном преобразованиях оказывается одним и тем же. Имеющиеся различия, если не требуется знание фазового спектра, несущественны.

Контрольные вопросы

1. Приведите эквивалентную схему преобразователя электрического сигнала в акустический (для объёмных волн) и проанализируйте оптимальные значения добротности и полосы преобразования.
2. Каковы ограничения применения однофазной системы решетки металлических электродов по сравнению с двухфазной системой (ВШП) для генерации поверхностных акустических волн?
3. Чем определяется оптимальное число штырей встречно-штыревого преобразователя и при каких условиях оно достигается?
4. Физический смысл акустической и электрической добротности встречно-штыревого преобразователя.
5. Приведите эквивалентную схему встречно-штыревого преобразователя и охарактеризуйте его акустические добротность и полосу пропускания.
6. В чём отличие явлений генерации и детектирования поверхностных акустических волн встречно-штыревым преобразователем.
7. Какие устройства для управления распространением поверхностных акустических волн реализованы на многополосковых ответвителях?
8. В чём принципиальное отличие функционирования устройств на многополосковых ответвителях и отражательных решётках? Каковы их рабочие частоты?
9. Каков физический смысл элементов волновой матрицы, описывающей распространение поверхностных акустических волн в отражательной решётке?
10. В каких устройствах функциональной акустоэлектроники применяется отражательная решётка?
11. В каких устройствах на основе встречно-штыревых преобразователей достигается однонаправленный режим генерации и распространения поверхностных акустических волн?
12. Приведите основные типы линий задержки поверхностных акустических волн.
13. Каким образом топология встречно-штыревых преобразователей определяет их частотные свойства?
14. В чём отличие резонаторов оптических и поверхностных акустических волн?
15. Приведите и проанализируйте формулы для описания част-

ной характеристики и добротности акустического резонатора.

16. Опишите конструкцию усилителя поверхностных акустических волн.

17. Проанализируйте особенности характеристик автогенераторов поверхностных акустических волн с акустическим резонатором и на основе линии задержки.

18. Каковы принципы построения трансверсального фильтра поверхностных акустических волн на основе многоотводной линии задержки?

19. Проведите анализ характеристик полосковых фильтров поверхностных акустических волн на основе анодизированных встречно-штыревых преобразователей.

20. Приведите основные конструкции дисперсионных фильтров поверхностных акустических волн и проанализируйте их характеристики.

21. Проиллюстрируйте дисперсионными диаграммами принцип действия конвольверов, корреляторов, запоминающих устройств.

22. Приведите сравнительный анализ конструкций корреляторов, конвольверов и запоминающих устройств на поверхностных акустических волнах.

23. Проанализируйте принципы функционирования Фурье-процессоров на основе дисперсионных фильтров по схемам П-С-П и С-П-С (С – свёртка, П – перемножение).

Расчёт и оптимизация параметров устройств на поверхностных акустических волнах. Компьютерный практикум

Задача 1. Возбуждение и приём поверхностных акустических волн. Встречно-штыревой преобразователь

Основным устройством для возбуждения и приёма ПАВ является встречно-штыревой преобразователь (ВШП). Его важной отличитель-

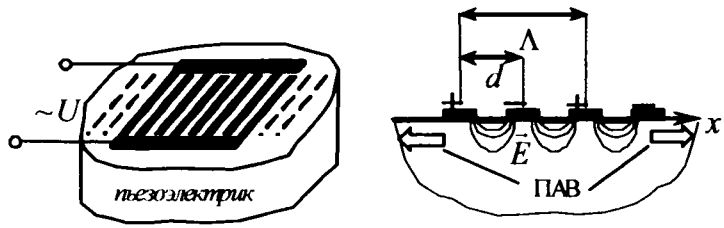


Рис П1.1. Возбуждение ПАВ с помощью ВШП

ной особенностью является возможность совмещения функций возбуждения (или приёма) ПАВ с одновременной обработкой сигнала.

ВШП представляет собой пару плоских металлических гребёнок электродов, вложенных одна в другую и размещённых на пьезоэлектрической подложке. В простейшем случае пространственный период обеих гребёнок одинаков и постоянен по всей длине преобразователя, а штыревые электроды гребёнок имеют одинаковую длину и ширину (рис. П1.1). Такой преобразователь получил название эквидистантного неаподизованного или однородного. Для возбуждения ПАВ к общим шинам гребёнок подводится электрический сигнал. В результате потенциалы электрического поля на соседних штыревых электродах, принадлежащих разным гребёнкам, оказываются противоположными, и в приповерхностной области пьезоэлектрической подложки создаётся периодическое переменное электрическое поле. Это поле благодаря пьезоэффекту возбуждает в подложке две поверхностные упругие вол-

ны, распространяющиеся в обе стороны от преобразователя и сопровождающиеся переменным электрическим полем.

Очевидно, что для эффективного преобразования электрической энергии в акустическую пространственный период изменения электрического поля гребёнок должен соответствовать пространственному периоду электрического поля поверхностной акустической волны. В этом случае элементарные волны, излучаемые каждым из участков ВШП, будут складываться. В противном случае интерференция элементарных волн, генерируемых разными участками ВШП, будет частично, либо полностью гасить суммарную ПАВ, возбуждаемую преобразователем. Частота сигнала, удовлетворяющая условию наилучшего возбуждения, называется частотой акустического синхронизма f_0. Она определится из условия $2d = \Lambda$, где d – расстояние между соседними электродами, принадлежащими разным гребёнкам, $\Lambda = v/f_0$ – длина волны ПАВ, v – фазовая скорость ПАВ:

$$f_0 = v/2d . \qquad (П1.1)$$

В режиме приёма ПАВ ввиду обратимости пьезоэффекта поверхностная акустическая волна, проходящая под штыревыми электродами преобразователя, наводит на них напряжение, которое, суммируясь на общих шинах гребёнок, приводит к появлению на выходе электрического сигнала. Амплитуда сигнала максимальна также при условии соответствия периода электрического поля ПАВ периоду расположения штыревых электродов, то есть при равенстве частоты ПАВ и частоты акустического синхронизма. На практике для осуществления различных функций обработки сигнала используются преобразователи, в которых шаг электродов, их длина или ширина являются переменными. В соответствии с принятой к настоящему времени терминологией, если длина или ширина электродов являются переменными, ВШП называют аподизованным. Когда переменным является период преобразователя, его именуют неэквидистантным (рис. П1.2).

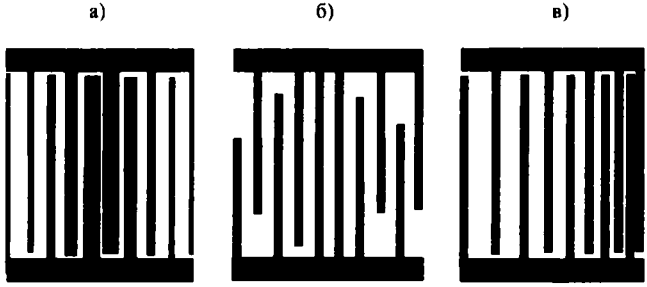


Рис. П1.2. Различные виды ВШП:
а – эквидистантный аподизованный с переменной шириной электродов;
б – эквидистантный аподизованный с переменной длиной электродов;
в – неэквидистантный неаподизованный

Для описания процесса возбуждения акустической волны системой параллельных планарных электродов на поверхности пьезоэлектрического полупространства предположим, что длина электродов значительно больше расстояния между ними. Так как на практике длина штыревых электродов ВШП обычно составляет от нескольких десятков до нескольких сотен длин волн ПАВ, такое допущение вполне возможно. В этом случае влиянием концов электродов можно пренебречь и считать распределение электрического заряда вдоль длины электрода равномерным, в то время как в поперечном электродам направлении распределение заряда на каждом электроде не известно.

Пусть $q(\omega, x) = q(x)\exp(j\omega t)$ – поверхностная плотность заряда на электродах, возбуждаемых гармоническим внешним источником (рис. П1.3). Осциллирующий заряд создаёт переменное электрическое поле, которое, в свою очередь, возбуждает две поверхностные

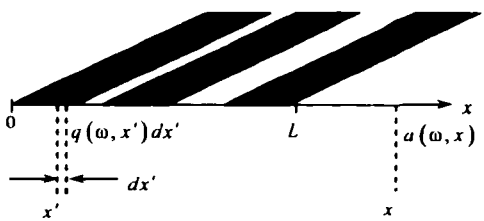


Рис. П1.3. Возбуждение пьезоэлектрического полупространства планарными металлическими электродами на его поверхности

волны вида $a_0 \exp\left[j(\omega t - \kappa x)\right]$ и $a_0 \exp\left[j(\omega t + \kappa x)\right]$, распространяющиеся в противоположных направлениях перпендикулярно электродам вдоль поверхности пьезоэлектрика, где $\kappa = \omega / v$ – волновое число ПАВ. В силу линейности системы поверхностная волна $da(x, x', \omega)$ в некоторой точке с координатой x справа от преобразователя, создаваемая элементарным зарядом $q(\omega, x')dx'$, имеющим координату x', будет пропорциональна величине этого заряда:

$$da(x, x', \omega) = \alpha q(\omega, x') \exp\left[-j\kappa(x - x')\right]dx', \qquad (П1.2)$$

где коэффициент α определяет связь между зарядом и акустическим возбуждением и не зависит от топологии электродов. Если длина области подложки, занятой электродами, равна L, то суммарная амплитуда ПАВ в некоторой точке x запишется как

$$a(\omega, x) = \alpha \int_0^L q(\omega, x') \exp\left[-j\kappa(x - x')\right]dx'. \qquad (П1.3)$$

В точках x', лежащих вне преобразователя, $q(x') = 0$, поэтому в выражении (П1.3) пределы интегрирования можно заменить на бесконечные:

$$a(\omega, x) = \alpha \int_{-\infty}^{\infty} q(\omega, x') \exp\left[-j\kappa(x - x')\right]dx'. \qquad (П1.4)$$

Переходя к амплитудам, окончательно можно записать:

$$a_0(\omega) = \alpha q(\omega),$$
$$q(\omega) = \int_{-\infty}^{\infty} q(x') \exp\left(j\frac{\omega}{v}x'\right)dx'. \qquad (П1.5)$$

Так как коэффициент α не зависит от геометрии преобразователя, а используемые в устройствах на ПАВ акустические волны, как правило, не имеют дисперсии (обычно это волны Рэлея), последние соотношения позволяют сделать важнейший для теории ВШП вывод: зависимость амплитуды генерируемой ПАВ от частоты (то есть частотная характеристика ВШП) представляет собой Фурье-образ функции пространственного распределения поверхностной плотности заряда на штыревых электродах ВШП. Ввиду того, что частотная характеристика преобразователя равна Фурье-образу его импульсной характеристики, можно также сделать вывод, что импульсный отклик преобразователя совпадает с

пространственным распределением плотности поверхностного заряда на его электродах.

Задача расчёта распределения заряда на электродах и определения коэффициента связи α требует совместного решения волновых уравнений для пьезоэлектрического полупространства с учётом всех возможных видов упругих волн при наличии электродов на поверхности и уравнений электростатики для распределения электрического поля электродов. Такое решение приводит к довольно сложным интегральным уравнениям, анализ которых возможен только численными методами. Однако для многих практических случаев найдены достаточно хорошие приближения.

Будем считать, что электромеханическая связь является слабой. Это позволяет предположить, что электрическое поле ПАВ мало и не влияет на распределение поверхностного заряда электродов, то есть вторичное акустоэлектронное преобразование отсутствует. Предположим также, что электроды являются идеально проводящими и бесконечно тонкими. При данных допущениях распределение заряда может быть найдено из решения соответствующей электростатической задачи. Данное решение приводит к распределению заряда, описываемого известной формулой Келдыша–Седова:

$$q(x)=\sum_{m=0}^{N-1}\gamma_m x^m\left(\prod_{n=1}^{N}(x-a_n)(b_n-x)\right)^{-1/2},\quad x\in[a_n,b_n]. \quad (\text{П1.6})$$

Здесь a_n и b_n – координаты краёв n-го электрода, N – число электродов, γ_m – набор коэффициентов, выбираемых таким образом, чтобы потенциал n-того электрода был равен заданному значению φ_n. В промежутке между электродами $q(x)=0$. Потенциалы электродов φ_n найдутся из соотношений:

$$\begin{aligned}\varphi_n &= \int_L q(x')G_e(x-x')\,dx',\ x\in[a_n,b_n],\\ G_e(x-x') &= \frac{1}{\pi\varepsilon_0(1+\varepsilon_p)}\ln|x-x'|,\end{aligned} \quad (\text{П1.7})$$

где $G_e(x-x')$ – электростатическая функция Грина для диэлектрического полупространства, ε_0 – абсолютная диэлектрическая проницаемость вакуума, $\varepsilon_p=\sqrt{\varepsilon_{33}\varepsilon_{11}-\varepsilon_{13}}$ – относительная эффективная

диэлектрическая проницаемость подложки, ε_{ik} – компоненты тензора диэлектрической проницаемости подложки. Однако даже в случае небольшого числа электродов формулы (П1.6) – (П1.7) всё ещё довольно сложны для анализа влияния топологии преобразователя на

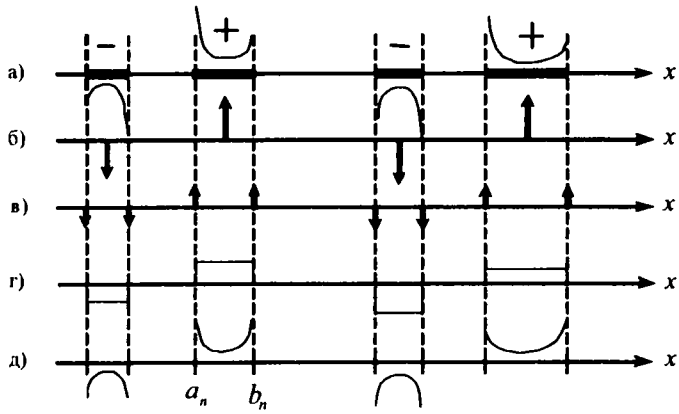


Рис. П1.4. Распределение зарядов 4-электродного преобразователя: а – реальное; б – в модели с одной дельта-функцией; в – в модели с двумя дельта-функциями; г – в модели равномерного распределения, д – в модели уединённого электрода

его частотную характеристику. Поэтому для функции распределения зарядов применяются различные приближённые модели. Наиболее простые из них – модель с одной или двумя дельта-функциями, модель равномерного распределения и модель уединённого электрода (рис. П1.4).

В модели с одной дельта-функцией заряд каждого электрода считается сосредоточенным в его центре, и плотность распределения заряда описывается соотношением

$$q(x) = \sum_{k=1}^{N} \frac{q_n}{W} \delta(x - x_n), \qquad (П1.8)$$

где q_n и x_n – соответственно заряд и координата середины n-го электрода, N – число электродов, W – апертура преобразователя. Очевидно, данная модель может быть использована в случае достаточно узких электродов. В тех случаях, когда нельзя пренебречь их шириной, можно использовать модель с двумя дельта-функциями. При этом реальное

распределение зарядов на каждом электроде заменяется суммой дельта-функций, расположенных на его краях:

$$q(x)=\frac{1}{2}\sum_{n=1}^{N}\frac{q_n}{W}\left[\delta(x-a_n)+\delta(x-b_n)\right]. \qquad (П1.9)$$

Довольно часто используется модель равномерного распределения, в которой заряд каждого электрода считается равномерно распределённым по его ширине:

$$q(x)=\begin{cases}\dfrac{q_n}{W|b_n-a_n|} & \text{при } x\in[a_n,b_n],\\ 0 & \text{при } x\notin[a_n,b_n].\end{cases} \qquad (П1.10)$$

Наконец, в модели уединённого электрода в качестве функции распределение заряда принимается точное распределение заряда на каждом электроде, вычисляемое в отсутствие других электродов. Это распределение вытекает из формулы Келдыша–Седова для одиночного электрода $(N=1)$:

$$q(x)=\begin{cases}\dfrac{q_n}{W}\dfrac{1}{\pi\sqrt{(x-a_n)(b_n-x)}} & \text{при } x\in[a_n,b_n],\\ 0 & \text{при } x\notin[a_n,b_n].\end{cases} \qquad (П1.11)$$

Заряды электродов q_n определяются по их потенциалам и межэлектродной ёмкости, которые находятся из решения соответствующей электростатической задачи.

Несмотря на довольно грубую аппроксимацию распределения зарядов, рассмотренные модели во многих практических случаях дают приемлемое совпадение с результатами строгой теории. Преимущество этих моделей заключается в возможности получения простых выражений для частотных характеристик различных преобразователей, в том числе аподизованных и неэквидистантных. Однако такой подход даёт только общий вид частотной характеристики ВШП в относительных единицах и не позволяет найти эквивалентные электрические параметры ВШП, а также определить эффективность преобразования.

Рассмотрим частотную характеристику однородного ВШП с одинаковой шириной и длиной электродов и постоянным шагом между ними. Используя модель с одной дельта-функцией, из соотношений (П1.5) и (П1.8) получаем

$$a_0(\omega) = \alpha \frac{1}{W} \sum_{n=1}^{N} q_n \exp\left(j \frac{\omega}{\mathrm{v}} x_n \right). \qquad (П1.12)$$

Пусть центр первого электрода преобразователя имеет координату $x_1 = 0$, а расстояние между центрами соседних электродов – d, то координата центра n-го электрода равна $x_n = (n-1)d$. Если число положительно заряженных электродов равно числу отрицательно заряженных, то заряды на соседних электродах равны по величине и противоположны по знаку – $q_n = q_0 (-1)^{n-1}$. Конечно, данное предположение не вполне точно, так как заряды электродов вблизи краёв преобразователя должны быть, очевидно, несколько меньше зарядов внутренних электродов. Однако для упрощённого анализа можно принять, что преобразователь достаточно длинный и влиянием его краёв можно пренебречь. В этом случае соотношение (П1.12) приобретает следующий вид:

$$a_0(\omega) = \alpha \frac{q_0}{W} \sum_{n=1}^{N} \exp\left[-j(n-1)\left(\frac{\omega}{\mathrm{v}} d - \pi \right) \right]. \qquad (П1.13)$$

Оно представляет собой сумму геометрической прогрессии, вычисляя которую, находим:

$$a_0(\omega) = \frac{\alpha q_0}{W} \exp\left[j \frac{(N-1)}{2} \left(\pi - \frac{\omega}{\mathrm{v}} d \right) \right] \frac{\sin\left[\frac{N}{2} \left(\frac{\omega}{\mathrm{v}} d - \pi \right) \right]}{\sin\left[\frac{1}{2} \left(\frac{\omega}{\mathrm{v}} d - \pi \right) \right]}. \qquad (П1.14)$$

Если преобразователь, имеющий N_2 пар штыревых электродов $(N_2 = N/2)$, возбуждается напряжением единичной амплитуды $U = \exp(j\omega t)$, то абсолютный заряд каждого электрода равен

$$q_0 = \frac{C}{N_2} = \frac{2C}{N}, \qquad (П1.15)$$

где C – электростатическая ёмкость преобразователя. Принимая во внимание, что

$$\frac{\omega}{\mathrm{v}} d - \pi = \frac{2\pi f}{\mathrm{v}} d - \pi = \pi \frac{f - f_0}{f_0}, \qquad (П1.16)$$

где $f_0 = \mathrm{v}/(2d)$ – частота акустического синхронизма, и отбрасывая несущественный фазовый множитель, линейно зависящий от часто-

ты и описывающий групповую задержку сигнала, получаем выражение для АЧХ однородного преобразователя:

$$A(f)=|a_0|=\alpha\frac{2C}{W}\left|\frac{1}{N}\frac{\sin\left(\frac{N\pi}{2}\cdot\frac{f-f_0}{f_0}\right)}{\sin\left(\frac{\pi}{2}\cdot\frac{f-f_0}{f_0}\right)}\right|. \qquad (П1.17)$$

Из соотношения (П1.17) следует, что АЧХ однородного ВШП имеет максимумы на частоте акустического синхронизма f_0 и её нечётных гармониках $(2n+1)f_0$, $n=1,2,3...$ Ограничиваясь частотами вблизи основного максимума, легко видеть, что нули $A(f)$ расположены на частотах $f_0 \pm 2f_0/N$. Таким образом, можно приближённо принять, что ВШП работает эффективно в полосе частот $f_0 \pm f_0/N$, на краях которой амплитуда снижается на ≈ 4 дБ по отношению к максимальному значению. Более точная величина полосы преобразования, определённая по уровню 3 дБ, равна:

$$\Delta f \simeq 0{,}89(2f_0/N) \simeq 0{,}89\, f_0/N_2\,. \qquad (П1.18)$$

Рассмотренная выше полоса, определяемая геометрией преобразователя, носит название акустической полосы. На частотные свойства ВШП влияют также его эквивалентные электрические параметры. Для анализа влияния электрических параметров можно использовать параллельную, или последовательную эквивалентную схему (рис. П1.5). Оба представления полностью взаимозаменяемы. Элементы последовательной схемы могут быть пересчитаны в элементы параллельной схемы с помощью стандартных методов теории цепей

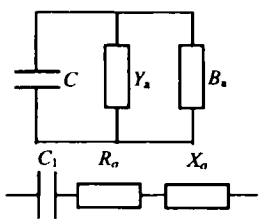


Рис. П1.5. Параллельная и последовательная электрические эквивалентные схемы ВШП: C – параллельная статическая ёмкость, Y_a и B_a – активная и реактивная части проводимости излучения, C_1 – последовательная статическая ёмкость, R_a и X_a – активная и реактивная части сопротивления излучения

В параллельной эквивалентной схеме ВШП имеет статическую межэлектродную емкость C, активную и реактивную составляющие

проводимости излучения $Y_a(f)+jB_a(f)$. Активная часть проводимости излучения описывает затраты электрической энергии на генерацию ПАВ. Вблизи частоты акустического синхронизма реактивная составляющая проводимости B_a мала по сравнению со статической ёмкостью C, и ею можно пренебречь. Эквивалентные элементы последовательной схемы выразятся соотношениями:

$$R_a = \frac{Y_a}{Y_a^2 + (\omega C)^2}, \quad \frac{1}{\omega C_1} = \frac{\omega C}{Y_a^2 + (\omega C)^2}. \qquad (П1.19)$$

Как будет показано далее, активная проводимость излучения однородного преобразователя, имеющего N_2 пар электродов, на частоте акустического синхронизма определяется выражением

$$Y_a = \tilde{y} K^2 f_0 N_2 C, \qquad (П1.20)$$

где K – коэффициент электромеханической связи, $\tilde{y}$ – численный коэффициент, зависящий от отношения ширины электродов к шагу ВШП. При типичных значениях этого отношения $\tilde{y} \approx 8$. Вслучае слабой электромеханической связи проводимость преобразователя должна, очевидно, носить, в основном ёмкостной характер:

$$\omega C >> Y_a. \qquad (П1.21)$$

Откуда с учётом (П1.20) вытекает количественное выражение условия слабой электромеханической связи:

$$\frac{4}{\pi} K^2 N_2 << 1. \qquad (П1.22)$$

При выполнении данного условия выражения (П1.19) приобретают следующий вид:

$$R_a \approx \frac{Y_a}{(\omega C)^2}, \quad C_1 \approx C. \qquad (П1.23)$$

Электрическая добротность параллельной эквивалентной схемы ВШП равна

$$Q_e = \frac{2\pi f_0 C}{Y_a}. \qquad (П1.24)$$

Из (П1.24) и (П1.20) при $\tilde{y} = 8$ получаем

$$Q_e = \frac{\pi}{4K^2 N_2}. \qquad (П1.25)$$

Электрическая добротность определяет электрическую полосу ВШП:

$$\Delta f = \frac{f_0}{Q_e} = f_0 \frac{4K^2 N_2}{\pi}. \qquad (П1.26)$$

Меньшая из двух полос – акустическая или электрическая – будет ограничивать эффективную полосу преобразования, которую называют собственной или характеристической полосой ВШП. Как видно из выражений (П1.18) и (П1.26), акустическая полоса с ростом числа пар штырей убывает как $1/N_2$, в то время как электрическая полоса пропорциональна N_2. При большом числе штырей собственная полоса ограничена акустическими свойствами преобразователя, а при малом – электрическими. Максимально широкая характеристическая полоса преобразования будет достигаться при условии равенства электрической и акустической полос:

$$f_0 \frac{4K^2 N_2}{\pi} \simeq 0{,}89 \frac{f_0}{N_2}. \qquad (П1.27)$$

Отсюда можно определить число пар штыревых электродов, при котором полоса преобразования однородного ВШП имеет наибольшую величину:

$$N_2 \simeq \frac{\sqrt{0{,}89\pi}}{2K} \simeq \frac{0{,}83}{K}. \qquad (П1.28)$$

Сама же характеристическая относительная полоса преобразования при оптимальном числе электродов равна:

$$\left(\Delta f / f_0\right)_{\max} \simeq 1{,}07K. \qquad (П1.29)$$

Следует отметить, что максимальная относительная ширина полосы преобразования зависит исключительно от пьезоэлектрических свойств подложки. Для сильных пьезоэлектриков ширина полосы больше, чем для слабых. На практике число электродов ВШП часто не соответствует оптимальному, а выбирается исходя из требования реализации заданной частотной характеристики.

Рассчитаем теперь АЧХ однородного преобразователя с использованием модели двух дельта-функций (рис. П1.6). Пусть расстояние между центрами соседних электродов равно d, левый край первого электрода преобразователя имеет координату $a_1 = 0$,

а правый – $b_1 = a_1 + p$, где p – ширина электродов. Тогда координата левого края n-го электрода равна $a_n = (n-1)d$, а правого края –

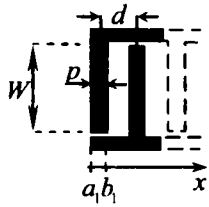


Рис. П1.6. К расчёту АЧХ однородного ВШП:
W– апертура преобразователя,
p – ширина штыревых электродов,
d – расстояние между серединами соседних электродов,
p/d– коэффициент металлизации

$b_n = (n-1)d + p$. Подставляя эти значения a_n и b_n в соотношения (П1.5) и (П1.9), с учётом (П1.15) и (П1.16) получаем

$$A(f) = \alpha \frac{2C}{W} \left| \cos\left(\frac{\pi f}{2 f_0} \frac{p}{d} \right) \right| \cdot \left| \frac{1}{N} \frac{\sin\left(\frac{N\pi}{2} \frac{f - f_0}{f_0} \right)}{\sin\left(\frac{\pi}{2} \frac{f - f_0}{f_0} \right)} \right|. \quad \text{(П1.30)}$$

Аналогичный расчёт АЧХ ВШП в модели равномерного распределения заряда на электродах приводит к следующему выражению:

$$A(f) = \alpha \frac{2C}{W} \left| \frac{\sin\left(\frac{\pi f}{2 f_0} \frac{p}{d} \right)}{\left(\frac{\pi f}{2 f_0} \frac{p}{d} \right)} \right| \cdot \left| \frac{1}{N} \frac{\sin\left(\frac{N\pi}{2} \frac{f - f_0}{f_0} \right)}{\sin\left(\frac{\pi}{2} \frac{f - f_0}{f_0} \right)} \right|. \quad \text{(П1.31)}$$

Модель уединённого электрода даёт

$$A(f) = \alpha \frac{2C}{W} \left| J_0\left(\frac{\pi f}{2 f_0} \frac{p}{d} \right) \right| \cdot \left| \frac{1}{N} \frac{\sin\left(\frac{N\pi}{2} \frac{f - f_0}{f_0} \right)}{\sin\left(\frac{\pi}{2} \frac{f - f_0}{f_0} \right)} \right|, \quad \text{(П1.32)}$$

где $J_0(x)$ – функция Бесселя первого рода нулевого порядка.

Нетрудно видеть, что для достаточно узких электродов $(p/d \to 0)$ все четыре рассмотренные модели дают одинаковый вид АЧХ однородного ВШП.

Обобщая на случай произвольного распределения заряда, для однородного ВШП можно записать:

$$A(f) = \alpha \frac{2C}{W} |e(f)| \cdot |H(f)|. \quad \text{(П1.33)}$$

В этом выражении функция $e(f)$ представляет собой Фурье-образ распределения заряда на отдельном электроде, её называют множителем элемента. В зависимости от характера распределения заряда на электроде множитель элемента имеет вид:

$$e(f)=\begin{cases}1 & \text{одна дельта-функция,}\\ \cos\left(\dfrac{\pi f}{2f_0}\dfrac{p}{d}\right) & \text{две дельта-функции,}\\ \sin\left(\dfrac{\pi f}{2f_0}\dfrac{p}{d}\right)\Big/\left(\dfrac{\pi f}{2f_0}\dfrac{p}{d}\right) & \text{равномерное распределение,}\\ \mathrm{J}_0\left(\dfrac{\pi f}{2f_0}\dfrac{p}{d}\right) & \text{уединённый электрод.}\end{cases} \qquad (П1.34)$$

Функция $H(f)$ – есть Фурье-образ распределения заряда в преобразователе с бесконечно узкими электродами, её называют множителем системы. Множитель системы не зависит от характера распределения заряда на отдельном электроде, а определяется исключительно числом электродов и шагом ВШП, задающим значение частоты акустического синхронизма:

$$H(f)=\frac{1}{N}\frac{\sin\left(\dfrac{N\pi}{2}\dfrac{f-f_0}{f_0}\right)}{\sin\left(\dfrac{\pi}{2}\dfrac{f-f_0}{f_0}\right)}. \qquad (П1.35)$$

В окрестности частоты акустического синхронизма зависимостью $e(f)$ по сравнению с зависимостью $H(f)$ во многих случаях, особенно при большом числе электродов, можно пренебречь и использовать для анализа АЧХ модель с одной дельта-функцией. Однако на других частотах влияние множителя элемента может быть существенным.

Рассмотрим далее процедуру расчёта проводимости излучения ВШП.

Известны различные методы расчёта проводимости излучения ВШП. Среди этих методов наиболее общим и строгим является физический метод, основанный на методе функции Грина для пьезоэлектрического полупространства.

Пусть ПАВ возбуждается ВШП длиной L и с апертурой W, распределение зарядов на электродах которого $q(x)$ произвольно, но известно. Если преобразователь питается напряжением единичной амплитуды $U = e^{j\omega t}$, а $I_0(x)e^{j\omega t} = j\omega q(x)e^{j\omega t}$ – плотность тока преобразователя, то комплексная электрическая мощность, поступающая в подложку, равна:

$$P = \frac{1}{2} W \int_L \varphi(x) I_0^*(x)\,dx = -\frac{1}{2} j\omega W \int_L \varphi(x) q(x)\,dx, \quad (\text{П1.36})$$

где φ – электрический потенциал на поверхности. Эту мощность также можно записать как

$$P = \frac{1}{2} Y |U|^2 = \frac{1}{2} Y, \quad (\text{П1.37})$$

где Y – полная входная проводимость преобразователя. Вещественная часть Y определяет проводимость излучения, а мнимая – статическую ёмкость преобразователя:

$$\begin{aligned} Y_a &= \operatorname{Re} Y, \\ \omega C &= -\operatorname{Im} Y. \end{aligned} \quad (\text{П1.38})$$

Из (П1.36) и (П1.37) получаем

$$Y = -j\omega W \int_L \varphi(x) q(x)\,dx. \quad (\text{П1.39})$$

Таким образом, для расчёта проводимости и статической ёмкости ВШП необходимо знать связь поверхностного потенциала и поверхностной плотности заряда.

Потенциал $\varphi(x)$ на поверхности пьезоэлектрика связан с поверхностным зарядом соотношением

$$\varphi(x) = \int_{-\infty}^{\infty} G(x - x') q(x')\,dx', \quad (\text{П1.40})$$

где $G(x - x')$ – функция Грина пьезоэлектрического полупространства. Она определяет потенциал на поверхности пьезоэлектрика, создаваемый бесконечно тонкой нитью заряда. Расчёт функции Грина в общем случае может быть выполнен только численными методами. Но в практически важном случае рэлеевских волн найдено аналитическое выражение, позволяющее получить простое решение, которое имеет следующий вид:

$$G = G_e + G_R + G_V, \quad (\text{П1.41})$$

где G_e – электростатическая часть, ответственная за статическую ёмкость электродов, G_R – соответствует возбуждению ПАВ, G_V – описывает возбуждение объёмных волн, уходящих от преобразователя вглубь пьезоэлектрика. Возбуждение объёмной волны в большинстве случаев является нежелательным явлением в устройствах на ПАВ. Поэтому принимаются специальные меры для их устранения. Отметим, что с ростом числа электродов относительный вклад объёмных волн в излучение ВШП быстро уменьшается. В дальнейшем будем считать, что объёмные волны отсутствуют, а проводимость излучения связана лишь с возбуждением ПАВ.

Электростатическая функция Грина определяется соотношением (П1.7), а функция Грина для рэлеевских волн имеет следующий вид:

$$G_R(x-x') = j\frac{K^2}{2\varepsilon_0(1+\varepsilon_p)}\exp\left(j\kappa|x-x'|\right), \qquad (П1.42)$$

где $K^2 = 2\Delta v/v$ – квадрат эффективного коэффициента электромеханической связи для рэлеевской волны, равный удвоенной относительной разнице скоростей волн Рэлея на свободной и покрытой идеально проводящей плёнкой поверхности.

Используя соотношения (П1.38) – (П1.42), нетрудно получить следующее выражение:

$$Y_a = \frac{\omega K^2 W}{2\varepsilon_0(1+\varepsilon_p)}\left|q(\omega)\right|^2. \qquad (П1.43)$$

Статическую ёмкость преобразователя можно найти, используя электростатическую функцию Грина (П1.7):

$$C = \frac{W}{\pi\varepsilon_0(1+\varepsilon_p)}\int_L dx\int_L dx' q(x)q(x')\ln|x-x'|. \qquad (П1.44)$$

Распределение заряда на электродах $q(x)$ в (П1.43) и его Фурье-образ $q(\omega)$ в (П1.44) берутся при единичной амплитуде напряжения на ВШП.

Помимо значения ёмкости и проводимости ВШП, физический метод позволяет найти коэффициент α, определяющий связь между поверхностной плотностью заряда и амплитудой возбуждаемой ПАВ. Электрическая мощность, поступающая в преобразователь, равна

$$P = \frac{1}{2} Y_a = \frac{\omega K^2 W}{4\varepsilon_0 \left(1 + \varepsilon_p\right)} \left|q(\omega)\right|^2 . \qquad (П1.45)$$

В силу закона сохранения энергии эта мощность должна равняться сумме мощностей двух возбуждаемых ПАВ, распространяющихся в противоположных направлениях от ВШП. Считая ширину пучка возбуждаемых ПАВ равной апертуре ВШП, суммарную мощность обеих ПАВ можно записать как

$$P = \frac{W|A|^2}{2} + \frac{W|A|^2}{2} = \alpha^2 W \left|q(\omega)\right|^2 . \qquad (П1.46)$$

Из сравнения (П1.45) и (П1.46) находим

$$\alpha = \frac{K}{2} \sqrt{\frac{\omega}{\varepsilon_0 \left(1 + \varepsilon_p\right)}} . \qquad (П1.47)$$

Отсюда следует, что для эффективного акустоэлектронного преобразования следует выбирать подложки с возможно большим коэффициентом электромеханической связи и возможно меньшей диэлектрической проницаемостью. Влияние коэффициента электромеханической связи на эффективность преобразования вполне очевидно. Влияние диэлектрической проницаемости также легко понять, если учесть, что при заданном заряде электродов (токе электродов) напряжённость электрического поля в подложке тем больше, чем меньше её диэлектрическая проницаемость.

Отметим, что в соответствии с рассмотренной теорией коэффициент α зависит, хотя и слабо, от частоты. Так что, строго говоря, эта зависимость также влияет на АЧХ преобразователя. Однако при обычной ширине полосы преобразования < 20 % главным фактором, определяющим АЧХ, является топология электродов, а зависимостью $\alpha(\omega)$ можно пренебречь.

Найдём ёмкость и проводимость излучения однородного преобразователя на частоте акустического синхронизма. Если пренебречь влиянием краёв преобразователя и в качестве функции распределения заряда на электродах принять известное из электростатики распределение заряда в бесконечно длинной регулярной решётке, то ёмкость ВШП можно записать в следующем виде:

$$C = N_2 C_2 W , \qquad (П1.48)$$

где

$$C_2 = \varepsilon_0 \left(1 + \varepsilon_p\right) \tilde{c}\left(p/d\right) \qquad (П1.49)$$

и представляет собой ёмкость пары электродов на единицу их длины (погонную ёмкость). Коэффициент $\tilde{c}(p/d)$ зависит от степени металлизации p/d:

$$\tilde{c}\left(\frac{p}{d}\right)=\frac{\mathrm{K}\left(\sin^2\frac{\pi p}{2d}\right)}{\mathrm{K}\left(\cos^2\frac{\pi p}{2d}\right)}, \qquad (\Pi 1.50)$$

где $\mathrm{K}(x)$ – полный эллиптический интеграл первого рода.

Используя выражения (П1.5), (П1.33), (П1.43), (П1.48) и (П1.49), проводимость излучения вблизи частоты акустического синхронизма можно представить в следующем виде:

$$Y_a=\tilde{y}f_0K^2CN_2, \qquad (\Pi 1.51)$$

где

$$\tilde{y}(p/d)=4\pi\tilde{c}^2(p/d)e^2(f_0), \qquad (\Pi 1.52)$$

$e(f_0)$ – значение множителя элемента на частоте акустического синхронизма.

Зависимость коэффициентов $\tilde{c}$ и $\tilde{y}$ от степени металлизации p/d, рассчитанная на основании распределения заряда в бесконечной регулярной решётке, показана на рис. П1.7.

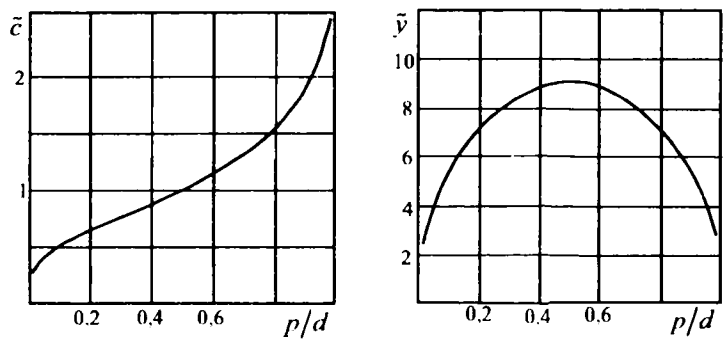


Рис. П1.7. Зависимость коэффициентов $\tilde{c}$ и $\tilde{y}$ от степени металлизации p/d

Электрическая добротность ВШП определяется выражением:

$$Q_e=\frac{2\pi f_0C}{\mathrm{Y}_a}=\frac{2\pi}{\tilde{y}K^2N_2}. \qquad (\Pi 1.53)$$

Из приведённой зависимости $\tilde{y}(p/d)$ следует, что минимальное значение электрической добротности и, следовательно, максимальная ширина собственной полосы преобразования достигаются при коэффициенте металлизации $p/d = 0{,}5$. При изменении степени металлизации в диапазоне от $p/d = 0{,}2$ до $p/d = 0{,}8$ среднее значение коэффициента $\tilde{y} \approx 8$.

Необходимо отметить, что периодические неоднородности, создаваемые электродами преобразователя, приводят к отражению от них акустических волн. Коэффициент отражения от отдельного электрода незначителен. Однако если расстояние между соседними электродами равно половине длины акустической волны, отражённые волны будут складываться в фазе. Вследствие этого при большом числе электродов вблизи частоты акустического синхронизма может наблюдаться сильное отражение ПАВ от преобразователя. Надо иметь в виду, что в рассмотренной выше теории отражение от электродов не учитывается. Поэтому полученные для проводимости выражения пригодны лишь для расчёта устройств, в которых отражённые волны не участвуют в формировании АХЧ или заметно на неё не влияют.

Контрольные вопросы

1. Перечислите основные виды ВШП.
2. Что такое частота акустического синхронизма?
3. Объясните суть простейших моделей для расчёта АЧХ преобразователей.
4. Сформулируйте факторы, влияющие на собственную полосу преобразования однородного ВШП, и объясните, как добиться максимально широкой полосы преобразования.
5. Сформулируйте кратко суть физической модели расчёта ёмкости и проводимости излучения ВШП и определите границы её применимости.
6. Какие требования предъявляются к пьезоэлектрической подложке?
7. Как влияет коэффициент металлизации на собственную полосу преобразования ВШП?

Расчётные задания

1. Используя данные Таблицы П1, определите оптимальное число электродов однородного преобразователя для следующих пьезоэлектрических подложек: а) кварц ST, X; б) танталат лития Z, Y; в) ниобат лития Z, Y. Для каждой из подложек найдите максимальную относительную полосу преобразования.

2. Для центральной частоты $f_0 = 100$ МГц найдите шаг преобразователя и постройте нормированную АЧХ 8-электродного преобразователя:

$$A(f) = \sqrt{\frac{f}{f_0}}\left|e(f)\right| \cdot \left|H(f)\right|$$

с использованием модели одной дельта-функции, двух дельта-функций, модели однородного распределения и модели уединённого электрода при различных значениях коэффициента металлизации $p/d = 0{,}2;\ 0{,}33;\ 0{,}5;\ 0.7$. Расчёт АЧХ выполнить в диапазоне частот $f = (50 \div 350)$ МГц. Сравните полученные зависимости и сформулируйте сходства и различия.

Для расчёта значений функции Бесселя использовать следующие аппроксимирующие многочлены:

$x \le 3$:

$$J_0(x) = 1 - 2{,}2499997(x/3)^2 + 1{,}2656208(x/3)^4 - 0{,}3163866(x/3)^6$$
$$+0{,}0444479(x/3)^8 - 0{,}0039444(x/3)^{10} + 0{,}00021(x/3)^{12}.$$

$x \ge 3$:

$$J_0(x) = x^{-1/2}\xi(x)\cos\theta(x),$$

$$\xi(x) = 0{,}79788456 - 0{,}00000077(3/x) - 0{,}0055274(3/x)^2 -$$
$$-0{,}00009512(3/x)^3 + 0{,}00137237(3/x)^4 - 0{,}00072805(3/x)^5 +$$
$$+0{,}00014476(3/x)^6,$$

$$\theta = x - 0{,}78539816 - 0{,}04166397(3/x) - 0{,}00003954(3/x)^2 +$$
$$+0{,}00262573(3/x)^3 - 0{,}00054125(3/x)^4 - 0{,}00029333(3/x)^5 +$$
$$+0{,}00013558(3/x)^6.$$

Задача 2. Потери в устройствах на ПАВ. Линия задержки на поверхностных акустических волнах

В электронике достаточно широко применяются различные линии задержки (ЛЗ). Поверхностные акустические волны являются довольно медленными волнами. Скорость их распространения на пять порядков меньше скорости света, благодаря чему на их основе можно получать значительные времена задержки при небольших габаритах устройства.

Простейшая линия задержки на ПАВ состоит из двух однородных ВШП – входного и выходного, которые размещены на поверхности пьезоэлектрического звукопровода (рис. П2.1). Для того чтобы ПАВ, достигающие торцов подложки, не отражались и не вносили искажений в сигнал, на эти торцы обычно наносится акустический поглотитель. Если необходимо задерживать сигналы сложного спектрального состава, полоса пропускания обоих преобразователей должна совпадать и быть по возможности максимально широкой.

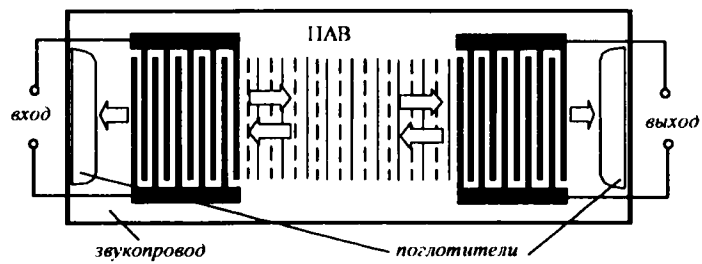


Рис.П2.1. Линия задержки с двумя ВШП на общей подложке

Одним из важнейших параметров любого ПАВ устройства и, в частности, линии задержки является величина вносимых потерь, которые приводят к ослаблению обрабатываемого сигнала.

Основными причинами потерь, вносимых ПАВ устройствами, являются:

– двунаправленность входного и выходного ВШП;
– рассогласование ВШП с внешними электрическими цепями;
– затухание сигнала в электродах преобразователя;
– затухание ПАВ в подложке;
– дифракционная расходимость акустического пучка.

Анализ потерь, связанных с двунаправленностью ВШП, можно провести с помощью матрицы рассеяния ВШП. Каждый ВШП рассматривается как шестиполюсник с тремя портами, в котором порт *3* является электрическим, а порты *1* и *2* – акустическими (рис. П2.2).

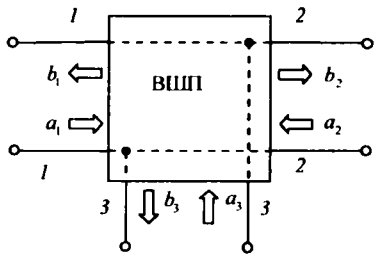


Рис. П2.2. Преобразователь как шестиполюсник с двумя акустическим портами *1* и *2* и одним электрическим – *3*

Связь между амплитудами сигналов a_k, приходящих на порты, с амплитудами выходящих сигналов b_i описывается матрицей рассеяния S_{ik}:

$$b_i = S_{ik} a_k . \qquad (П2.1)$$

Пусть преобразователь согласован со стороны электрического порта и потери отсутствуют. Тогда из соображений симметрии и закона сохранения энергии можно показать, что матрица рассеяния имеет следующий вид:

$$[S_{ik}] = \begin{pmatrix} S_1 & S_2 & S_3 \\ S_2 & S_1 & S_3 \\ S_4 & S_4 & 0 \end{pmatrix}, \qquad (П2.2)$$

где

$$S_2 = -S_1, \quad S_4 = \pm S_3,$$
$$|S_3|^2 = |S_4|^2 = \frac{1}{2}, \quad |S_1|^2 = |S_2|^2 = \frac{1}{4}. \qquad (П2.3)$$

Если мощность сигнала на электрическом порте первого ВШП равна A^2, то после преобразования половина этой мощности $A^2/2$ попадёт на акустический порт 1 и будет поглощена. Вторая половина достигнет порта 1 второго ВШП, где одна четверть от неё $A^2/8$ отразится, одна четверть пройдёт на порт 2 и поглотится, а половина $A^2/4$ преобразуется в сигнал электрического порта второго ВШП. Таким образом, результатом двунаправленного преобразования яв-

ляется ослабление мощности сигнала в 4 раза, или на 6 дБ – по 3 дБ на каждый преобразователь.

Следует отметить, что существуют несколько видов ВШП, обеспечивающих однонаправленное излучение ПАВ. Однако однонаправленный характер излучения в них, строго говоря, имеет место только в весьма узкой полосе частот. Кроме того, конструкции однонаправленных ВШП значительно сложнее.

Наличие отражения от акустических портов ВШП приводит к появлению на выходе ЛЗ, помимо основного сигнала, ряда дополнительных эхо-сигналов, соответствующих нечётному числу проходов ПАВ между преобразователями. Обычно существенным является только первый из них, связанный с тройным прохождением сигнала между входным и выходным ВШП. Его амплитуда только на 12 дБ меньше основного сигнала, что обусловлено двумя дополнительными преобразованиями. При длительном входном сигнале передний фронт трёхпролётного сигнала может появиться на выходе ЛЗ раньше заднего фронта основного сигнала, что приводит к искажениям.

Если преобразователи не согласованы с внешними электрическими цепями, уровень трёхпролётного сигнала по сравнению с уровнем основного сигнала уменьшается, но вносимые потери возрастают.

Потери сигнала, обусловленные рассогласованием ВШП с внешними цепями и поглощением мощности в электродной системе ВШП, могут быть найдены с помощью эквивалентной схемы возбу-

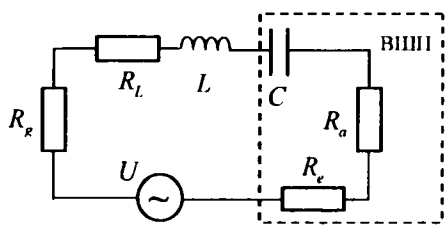


Рис. П2.3. Эквивалентная схема ВШП, подключённого к внешнему источнику сигнала

ждающей цепи, изображённой на рис. П2.3.

Преобразователь имеет статическую межэлектродную емкость C, сопротивление излучения R_a и сопротивление штыревых электродов R_e. Внешняя цепь содержит генератор с внутренним сопро-

тивлением R_g, согласующую индуктивность L, предназначенную для компенсации статической ёмкости ВШП, и последовательное активное сопротивление R_L, позволяющее согласовать генератор с преобразователем.

Если ширина электродов в два раза меньше расстояния между ними (то есть коэффициент металлизации равен 0,5), элементы эквивалентной схемы ВШП выражаются следующим образом:

$$C = \varepsilon_0 \left(1 + \varepsilon_p\right) W N_2,$$

$$R_a = \frac{2K^2}{\pi^2 v \varepsilon_0 \left(1 + \varepsilon_p\right)} \left(\frac{W}{\Lambda}\right)^{-1}, \qquad \text{(П2.4)}$$

$$R_e = \frac{4\rho}{N_2} \frac{W}{\Lambda}$$

Здесь ε_0 – абсолютная диэлектрическая проницаемость вакуума, ε_p – относительная эффективная диэлектрическая проницаемость подложки, K – коэффициент электромеханической связи, v – скорость ПАВ, Λ – длина волны ПАВ, W – апертура ВШП, N_2 – число пар электродов, ρ – поверхностное сопротивление электродов [Ом/□]. Диэлектрическая проницаемость, коэффициент электромеханической связи и скорость ПАВ определяются материалом подложки, ориентацией её поверхности и направлением распространения ПАВ.

Эквивалентная схема цепи приёмного ВШП вполне аналогична за исключением того, что источником сигнала является сам преобразователь, а сопротивление генератора должно быть заменено на сопротивление нагрузки. В общем случае под величиной $Z = R_L + j\omega L$ следует понимать полный импеданс внешней электрической цепи, согласующей преобразователь с источником сигнала либо с нагрузкой.

Если эдс источника сигнала равна U, то ток в цепи I выразится как отношение эдс источника к полному сопротивлению всей схемы:

$$I = \frac{U}{\left(R_g + R_a + R_e + R_L + j\omega L + \dfrac{1}{j\omega C}\right)}. \qquad \text{(П2.5)}$$

Напряжение на активном сопротивлении R_a равно

$$U_a = IR_a. \qquad (П2.6)$$

Тогда мощность излучения преобразователя

$$P_a = \frac{1}{2}U_a I^* = \frac{1}{2}\frac{|U|^2 R_a}{\left[\left(R_a + R_к + R_e + R_L\right)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2\right]}. \qquad (П2.7)$$

При заданной величине эдс генератора мощность излучения достигает своего максимального значения, если ёмкость и сопротивление излучения ВШП на частоте акустического синхронизма ω_0 удовлетворяют условиям:

$$\begin{aligned} R_a &= R_к + R_e + R_L, \\ \omega_0 C &= \frac{1}{\omega_0 L}. \end{aligned} \qquad (П2.8)$$

Максимальная электрическая мощность, которую генератор может отдавать в согласованную нагрузку,

$$P_0 = \frac{|U|^2}{8R_к}. \qquad (П2.9)$$

Потери, связанные с рассогласованием и поглощением энергии в электродной системе, определятся как отношение мощности, которую источник сигнала способен отдавать в согласованную нагрузку, к мощности, рассеиваемой сопротивлением излучения:

$$T = \frac{P_0}{P_a} = \frac{\left(R_a + R_к + R_e + R_L\right)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}{4R_Г R_A}. \qquad (П2.10)$$

Общепринято эту величину выражать в децибелах:

$$T[\text{дБ}] = 10\lg\frac{\left(R_a + R_к + R_e + R_L\right)^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}{4R_к R_a}. \qquad (П2.11)$$

Если $R_L = 0$ и $R_e = 0$, то при

$$R_a = R_к \text{ и } \omega_0 C = \frac{1}{\omega_0 L} \qquad (П2.12)$$

рассматриваемые потери отсутствуют $(T = 0\ \text{дБ})$.

Для уменьшения поглощения мощности в электродах необходимо, чтобы их поверхностное сопротивление было малым. Поэтому для нанесения преобразователя необходимо выбирать материалы с высокой проводимостью, а сами электроды делать не слишком тонкими. Однако по мере увеличения толщины возрастает механическая нагрузка на поверхности подложки, что приводит к росту коэффициента отражения ПАВ от преобразователя. Это вызывает дополнительные потери и искажения. Установлено, что компромисс между величиной отражения и потерями в случае часто используемых электродов из алюминия достигается при величине поверхностного сопротивления $\rho \approx 0,5 \div 0,7$ Ом/□.

Кроме величины вносимых потерь параметры схемы согласования ВШП с источником сигнала или нагрузкой также влияют на максимально достижимую полосу преобразования. Эта полоса определяется двумя факторами: акустическими свойствами преобразователя и добротностью электрической схемы согласования ВШП с источником или приёмником сигнала.

Электрическую полосу преобразования можно найти из эквивалентной схемы на рис. П2.3. Добротность этой схемы равна:

$$Q_e = \frac{2\pi f_0 C}{R_a + R_g + R_e + R_L} = \frac{2\pi f_0 C}{R_a(1+\eta)} = Q_{e0}(1+\eta)^{-1}, \quad (П2.13)$$

где Q_{e0} – собственная добротность ВШП,

$$\eta = \left(R_g + R_L + R_e\right)/R_a . \quad (П2.14)$$

Величина Q_e определяет относительную электрическую полосу схемы:

$$\left(\Delta f/f_0\right)_e = (1+\eta)/Q_{e0} . \quad (П2.15)$$

Собственная добротность ВШП равна:

$$Q_{e0} = \frac{\pi}{4K^2 N_2}, \quad (П2.16)$$

Следовательно,

$$\left(\Delta f/f_0\right)_e = \frac{4K^2 N_2}{\pi}(1+\eta). \quad (П2.17)$$

Относительная акустическая полоса преобразователя определяется выражением

$$\left(\Delta f/f_0\right)_a = 0,89/N_2 . \quad (П2.18)$$

Так как электрическая полоса пропорциональна числу пар электродов N_2, а акустическая полоса обратно пропорциональна N_2, то максимальная полоса преобразования достигается при условии равенства акустической и электрической полос, когда число пар электродов равно:

$$N_2 = \sqrt{\frac{0{,}89\pi}{4K^2(1+\eta)}}. \qquad (П2.19)$$

Сама же полоса преобразования при этом равна:

$$(\Delta f/f_0)_{max} = (\Delta f/f_0)_0\sqrt{1+\eta}, \qquad (П2.20)$$

где $(\Delta f/f_0)_0$ – собственная относительная полоса ВШП. Как видно из последнего выражения, наличие внешней цепи увеличивает максимально достижимую ширину полосы преобразования. С помощью соотношений (П2.11) и (П2.20) нетрудно показать, что увеличение полосы преобразования в два раза всегда приводит к увеличению вносимого затухания на 6 дБ, а с учётом наличия двух преобразователей – на 12 дБ.

На практике, например в ПАВ-фильтрах, используются ВШП с большим числом электродов, у которых акустическая полоса ВШП уже, чем электрическая полоса схемы согласования, так что АЧХ фильтра определяется исключительно топологией преобразователя. Для этого обычно достаточно, чтобы электрическая полоса схемы согласования превосходила в 2 раза акустическую полосу ВШП:

$$\frac{4K^2N_2}{\pi}(1+\eta) = 2(0{,}89/N_2), \qquad (П2.21)$$

откуда

$$N_2 = \sqrt{\frac{0{,}89\pi}{4K^2}\cdot\frac{2}{(1+\eta)}}. \qquad (П2.22)$$

Соответствующая этому количеству пар электродов полоса преобразования будет равна:

$$(\Delta f/f_0) = (\Delta f/f_0)_0\sqrt{(1+\eta)/2}. \qquad (П2.23)$$

Как следует из (П2.8) и (П2.11), минимальные потери рассогласования наступают при $\eta = 1$. В этом случае оптимальное число пар электродов равно:

$$N_2 = \sqrt{0{,}89\pi/(4K^2)}. \qquad (П2.24)$$

Формулы (П2.4) и (П2.11) позволяют найти апертуру ВШП, при которой потери рассогласования являются минимальными. Пусть сопротивление электродов ВШП пренебрежимо мало, тогда условие минимума потерь запишется как

$$R_a = R, \qquad (П2.25)$$

где $R = R_g + R_L$ – активное сопротивление внешней цепи. Используя (П2.4), получаем:

$$\frac{W}{\Lambda} = \frac{2K^2}{\pi^2 \nu \varepsilon_0 \left(1+\varepsilon_p\right) \mathrm{v}} \cdot \frac{1}{R}. \qquad (П2.26)$$

Если сопротивлением электродов пренебречь нельзя, то

$$\frac{W}{\Lambda} = \frac{RN_2}{8\rho}\left(\sqrt{1+\frac{32K^2\rho}{\pi^2 \nu \varepsilon_0 \left(1+\varepsilon_p\right) N_2 R^2}} - 1\right). \qquad (П2.27)$$

Рассмотренные выше потери, обусловленные двунаправленностью ВШП и его рассогласованием с внешними цепями, связаны с преобразованием акустической энергии в электрическую и обратно. Их называют потерями преобразования. Помимо потерь преобразования, в устройствах на ПАВ присутствуют потери распространения. Они обусловлены поглощением ПАВ в пьезоэлектрике и дифракционной расходимостью акустического пучка. В линии задержки ввиду возможно большого расстояния между преобразователями потери распространения могут быть существенны.

Поглощение ПАВ в пьезоэлектриках связано с их рассеянием на тепловых колебаниях решетки и давлением на поверхность подложки окружающей газовой среды. Установлено, что затухание ПАВ с достаточной для практических расчётов точностью можно описать формулой

$$B_M = \alpha_M f + \beta_M f^2, \qquad (П2.28)$$

где B_M – коэффициент поглощения, имеющий размерность дБ/см; f – частота в гигагерцах. Коэффициент α_M описывает поглощение, связанное с давлением окружающей газовой среды, а коэффициент β_M – с рассеянием на тепловых колебаниях. Как видно из выражения (П2.28), поглощение существенно возрастает при повышении рабочей частоты. Современные материалы, используемые в ПАВ устройствах, обеспечивают приемлемую величину B_M вплоть до частот в несколько гигагерц. Значения коэффициентов α_M и β_M,

а также другие параметры некоторых часто применяемых подложек представлены в табл. П1.

Полные потери, учитывающие преобразование и распространение ПАВ в ЛЗ, определятся соотношением

$$T_{сум}[\text{дБ}] = 6\ \text{дБ} + T_1[\text{дБ}] + T_2[\text{дБ}] + B_M(f)[\text{дБ}] \cdot L, \quad (\text{П2.29})$$

где T_1 и T_2 – потери рассогласования входного и выходного ВШП, L – длина линии задержки. Дополнительное слагаемое 6 дБ связано с двунаправленностью обоих ВШП. Уровень трёхпролётного сигнала по отношению к прошедшему сигналу будет определяться удвоенным значением $T_{сум}$.

Помимо указанных выше потерь, в длинных ЛЗ могут иметь место дополнительные потери, связанные с дифракционным расширением фронта ПАВ, которое возрастает по мере удаления от ВШП. Количественный учёт дифракционных потерь довольно сложен. Но можно сформулировать простые критерии, позволяющие оценить их влияние.

Поле ПАВ, генерируемое однородным преобразователем, имеет много общего с полем, возникающим при дифракции однородного светового потока на щели. В частности, имеется ближняя зона излучения или зона дифракции Френеля – область, где ширина пучка практически равна апертуре ВШП. В дальней зоне, или зоне Фраунгофера, пучок расходится под углом, зависящим от апертуры преобразователя. Используя оптическую аналогию, дифракционное расширение акустического луча можно оценить по величине параметра:

$$F = \frac{4\Lambda L}{W^2}, \quad (\text{П2.30})$$

где Λ – длина акустической волны, L – расстояние от точки наблюдения до преобразователя, W – апертура преобразователя. Значение $F < 1$ соответствует зоне Френеля (ближней зоне). В этой зоне энергия акустической волны практически полностью сосредоточена в границах пространственной полосы, ширина которой определяется апертурой ВШП, и дифракционными потерями можно пренебречь. Значения $F > 1$ соответствуют зоне Фраунгофера (дальней зоне), в которой дифракционная расходимость становится существенной.

Основное отличие дифракции ПАВ от оптической дифракции связано с сильной анизотропией пьезоэлектрических материалов,

используемых в качестве подложек в ПАВ устройствах. Вследствие данной анизотропии фазовая скорость волны существенно зависит от направления распространения, и в общем случае это направление не совпадает с направлением распространения энергии волны, которое определяется вектором групповой скорости. При этом картина дифракции значительно усложняется. В случае распространения вдоль направления чистой моды подложки фазовая и групповая скорости совпадают. Изменение фазовой скорости ПАВ при изменении направления распространения приводит к уменьшению или увеличению расходимости акустического фронта по сравнению с изотропной средой. Уменьшение расходимости акустического пучка называют автоколлимацией. В большинстве случаев этот эффект является желательным в ПАВ устройствах. Автоколлимация весьма полезна в случае применения протяжённых многоэлектродных преобразователей, а также в случае большого расстояния между преобразователями. В протяженных преобразователях автоколлимация позволяет избежать искажения частотной характеристики, а в длинных линиях задержки – дополнительных дифракционных потерь.

В общем случае учёт дифракционной расходимости в случае произвольного направления в анизотропной подложке является сложной задачей, однако для практической оценки можно использовать аналогичный изотропному случаю параметр

$$F = \frac{4\Lambda L}{W^2}|1+\gamma|, \qquad (П2.31)$$

где γ – параметр анизотропии подложки. Значения γ для некоторых направлений в ряде часто используемых подложек приведены в табл. П1. При $\gamma > 0$ дифракционные потери больше, чем в изотропной среде, а при $\gamma < 0$ – меньше. Идеальное значение $\gamma = -1$, при котором граница зоны Френеля стремится к бесконечности, то есть дифракционная расходимость полностью подавлена эффектом автоколлимации. Для того чтобы вся акустическая энергия, излученная входным преобразователем, была принята выходным, преобразователи должны располагаться в ближней зоне относительно друг друга.

Обычно величина задержки в ЛЗ на ПАВ не превышает нескольких микросекунд. При больших задержках общие потери могут оказаться неприемлемыми. Для получения больших времён задержки применяют каскадное включение нескольких ЛЗ, размещённых на

общей подложке, с промежуточными усилителями, компенсирующими потери каждой из ЛЗ.

Контрольные вопросы

1. Как устроена простейшая линия задержки на ПАВ?
2. Что такое трёхпролётный сигнал?
3. Какие виды потерь присутствуют в линии задержки на ПАВ?
4. Из чего складываются потери на преобразование?
5. Как влияют внешние согласующие цепи на полосу преобразования ВШП?
6. Каково условие отсутствия дифракционных потерь?

Расчётные задания

1. Используя табл. 1 Приложения, найти оптимальное число пар электродов N_2, апертуру W, сопротивление излучения R_a, ёмкость C и сопротивление электродов однородного ВШП, согласованного с активной нагрузкой $R = 50$ Ом для 3 различных подложек: а) ниобат лития Z,Y-ориентация; б) кварц ST,X-ориентация; в) германат висмута [001],[100]-ориентация.

Расчёт провести для двух значений центральной частоты:

$f_0 = 200$ МГц,

$f_0 = 1000$ МГц.

Поверхностное сопротивление электродов $\rho = 0{,}5$ Ом/□.

2. Используя результаты расчётов, полученные в п. 1, для ЛЗ длиной $L = 5$ мм и $L = 10$ мм рассчитать время задержки, потери распространения и потери преобразования. Для каждого случая найти уровень трёхпролётного сигнала и оценить влияние дифракционной расходимости.

Задача 3. ПАВ-фильтры на основе встречно-штыревых преобразователей

Фильтрами называют линейные системы, преобразующие спектр сигнала. Если на вход фильтра поступает гармонический сигнал $U_{вх}(\omega) = U_0 e^{j\omega t}$, то на выходе будет присутствовать монохроматический сигнал той же частоты, но с другим значением амплитуды и фазы:

$$U_{вых}(\omega) = U_0 A(\omega) \exp\left[j\omega t - j\Phi(\omega) \right]. \qquad (П3.1)$$

Комплексный коэффициент передачи фильтра

$$H(\omega) = \frac{U_{вых}}{U_{вх}} = A(\omega) \exp\left[-j\Phi(\omega) \right] \qquad (П3.2)$$

называют передаточной функцией фильтра. Модуль и фазу $H(\omega)$ – положительные вещественные функции $A(\omega)$ и $\Phi(\omega)$ называют соответственно амплитудно-частотной и фазовой характеристиками фильтра. Если в полосе частот сигнала сложного спектрального состава $A(\omega)$ является константой, а $\Phi(\omega)$ линейно зависит от частоты, то данный сигнал передаётся без искажений, испытывая только временную задержку. Фурье-образ передаточной функции фильтра называют импульсным откликом:

$$h(t) = \int_{-\infty}^{\infty} H(\omega) \exp(j\omega t) d\omega. \qquad (П3.3)$$

Функцию фильтра с заданными характеристиками может выполнять встречно-штыревой преобразователь. Это утверждение основано на том факте, что передаточная характеристика ВШП представляет собой Фурье-образ функции пространственного распределения зарядов на электродах, а импульсный отклик ВШП повторяет эту функцию. Поскольку с помощью аподизации и изменения расстояния между электродами можно весьма гибко управлять распределением зарядов на электродах, на основе ВШП можно получать большое разнообразие импульсных откликов и соответствующих этим откликам передаточных характеристик.

В практических приложениях достаточно часто встречается следующая задача: из входного широкополосного сигнала выделить некоторую область частот, внутри которой спектральные компонен-

ты сигнала передаются без искажений, а за пределами данной полосы полностью подавляются. Фильтры такого типа называются полосовыми. Как было показано в разделе 1, однородный ВШП по существу представляет собой полосовой фильтр, эффективно работающий в диапазоне частот $f_0 \pm f_0/N$ (f_0 – частота акустического синхронизма, N – число штырей). Фазовая характеристика ВШП является линейной функцией частоты (см. формулу **Ошибка! Источник ссылки не найден.**). Изменяя число штырей и шаг между ними, можно обеспечить преобразование сигнала в заданной полосе частот. Однако в большинстве случаев однородный ВШП не пригоден в качестве полосового фильтра. Это обусловлено несколькими причинами: АЧХ однородного ВШП имеет большую неравномерность в полосе пропускания; неполное подавление в области непропускания – наличие боковых лепестков значительной амплитуды; низкая крутизна среза АЧХ на границе полосы пропускания. Рассмотрим далее принципы построения полосового ПАВ фильтра, свободного от вышеперечисленных недостатков.

В общем случае произвольных зарядов на электродах эквидистантного преобразователя его передаточная характеристика в модели одиночных дельта-функций определяется выражением

$$H(f) = \sum_{n=1}^{N} q_n (-1)^{n-1} \exp\left(j \frac{2\pi f}{\mathrm{v}} x_n \right), \qquad (\text{П3.4})$$

где q_n – абсолютное значение заряда n-го электрода, $x_n = (n-1)d$ – координата центра этого электрода, d – расстояние между центрами соседних электродов, v – скорость ПАВ. Импульсный отклик находится с помощью преобразования Фурье от (П3.4):

$$h(t) = \sum_{n=1}^{N} q_n (-1)^{n-1} \delta\left(t - \frac{x_n}{\mathrm{v}} \right). \qquad (\text{П3.5})$$

Как видно из (П3.5), импульсный отклик ВШП представляет собой последовательность N дельта-импульсов, момент появления которых определяется положением электродов, а их амплитуда и полярность – величиной и знаком заряда электрода. Длительность импульсного отклика равна L/v (L – длина преобразователя). Это означает, что ВШП, по существу, представляет собой трансверсальный фильтр, в котором роль линии задержки выполняет пьезоэлектрическая подложка, штыревые электроды выполняют функцию отводов, а общие шины – функцию сумматора (рис. П3.1). Коэффици-

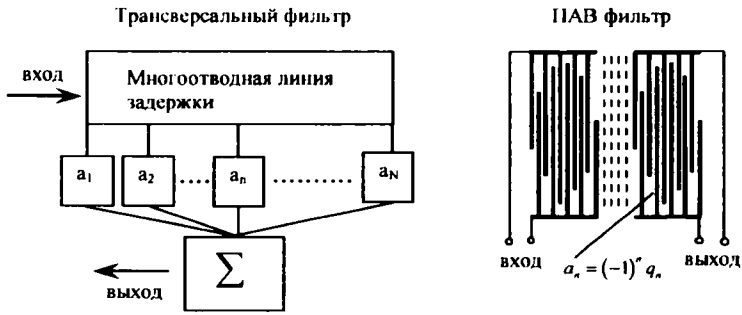


Рис. П3.1. Аналогия между трансверсальным фильтром и ПАВ фильром на основе ВШП

енты взвешивания отводов a_n могут изменяться путём изменения апертуры электродов. Но имеются некоторые отличия. Ввиду того что преобразователь имеет встречно-штыревую структуру, полярности соседних отводов противоположны в отличие от трансверсального фильтра, где полярности всех отводов могут быть одинаковыми. Отсюда следует, что на основе ВШП не может быть создано классического фильтра нижних частот, так как отклик ВШП на нулевой частоте отсутствует. Передаточная характеристика ВШП всегда является функцией полосового типа. Поскольку ПАВ фильтр содержит как минимум два преобразователя, то есть является комбинацией двух трансверсальных фильтров, появляется дополнительная гибкость при синтезе заданной передаточной характеристики.

ВШП, как и трансверсальный фильтр, всегда имеет конечную длину, поэтому его импульсный отклик ограничен во времени. Но для реализации полосовых передаточных характеристик требуются импульсные отклики бесконечной протяжённости. Кроме того, импульсный отклик трансверсального фильтра представляет взвешенную дискретную совокупность выборок непрерывного сигнала.

Поясним это на примере. Рассмотрим полосовой фильтр, имеющий прямоугольную АЧХ с полосой пропускания в интервале частот Δf и центром на частоте f_0:

$$A(f) = \begin{cases} A_0, & f_0 - \Delta f/2 < f < f_0 + \Delta f/2, \\ 0, & f < f_0 - \Delta f/2,\ f > f_0 + \Delta f/2. \end{cases} \qquad (П3.6)$$

Импульсный отклик данного фильтра с точностью до несущественного множителя имеет вид

$$g(t)=\frac{\sin \pi \Delta f t}{\pi \Delta f t}\cos\left(2\pi f_0 t\right). \qquad (П3.7)$$

Так как обычно $\Delta f << f_0$, второй множитель в правой части (П3.7) изменяется значительно медленнее первого, поэтому импульсный отклик представляет собой высокочастотное колебание с частотой f_0, модулированное медленно меняющимся множителем, и содержит бесконечную последовательность лепестков (рис. П3.2). А зна-

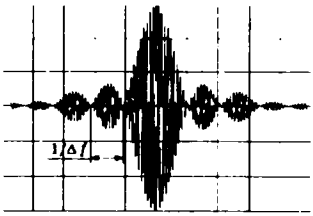


Рис. П3.2. Импульсный отклик идеального полосового фильтра. Ширина основного лепестка $2/\Delta f$, а боковых $-1/\Delta f$

чит, всегда конечный импульсный отклик ВШП ни при каком распределении зарядов не будет в точности соответствовать бесконечному импульсному отклику идеального фильтра. Однако, как будет показано далее, путём соответствующего выбора топологии ВШП можно получить АЧХ, максимально соответствующую заданной.

В дальнейшем удобно записать выражение (П3.5) в другом виде:

$$h(t)=\sum_{n=-N/2+1}^{N/2} q_n(-1)^{n-1}\delta\left(t-\frac{x_n}{\mathrm{v}}\right). \qquad (П3.8)$$

Оно отличается от выражения (П3.5) только смещением центра преобразователя из точки $x = Nd/2$ в точку $x = 0$. Пусть $\tilde{q}(t)$ – гладкая медленная функция, принимающая в моменты времени $t_n = x_n/\mathrm{v}$ значения, равные q_n. Нетрудно видеть, что $h(t)$ можно рассматривать как результат дискретизации некоторой непрерывной функции

$$\tilde{h}(t)=\tilde{q}(t)\cos\left(2\pi f_0 t\right) \qquad (П3.9)$$

с последующим ограничением числа дискретных отсчётов. Сигнал $\tilde{h}(t)$ представляет собой амплитудно-модулированное синусоидальное колебание с максимумами и минимумами в моменты времени

$t_n = x_n/\mathrm{v}$. Значение $\tilde{h}(t)$ в этих точках равно абсолютному заряду соответствующего электрода q_n. Предположим, что функция $\tilde{q}(t)$ достаточно медленная, так что спектр $\tilde{h}(t)$ ограничен частотами $f < 2f_0$. Так как период отсчётов $h(t)$ равен $t_n - t_{n-1} = 1/(2f_0)$, то в соответствии с теорией дискретизации при бесконечном числе отсчётов сигнала $h(t)$ на частотах $f < 2f_0$ его спектр будет совпадать со спектром непрерывного сигнала $\tilde{h}(t)$. Необходимо отметить, что спектр дискретного сигнала содержит также зеркальные характеристики, совпадающие с основной на нечётных гармониках f_0. Предполагается, что отклик преобразователя на этих гармониках подавлен.

Из приведённых выше рассуждений вытекает простой рецепт построения преобразователя, АЧХ которого соответствует заданной. Для этого необходимо найти импульсный отклик преобразователя, разместить электроды, подключенные к положительной общей шине в точках максимума импульсного отклика, а к отрицательной общей шине – в точках минимума. Затем обеспечить распределение зарядов на электродах в соответствии с огибающей импульсного отклика.

Наиболее удобным и самым распространённым приёмом управления распределением зарядов на электродах является аподизация – изменение ширины электродов либо их апертуры. При этом можно считать, что величина заряда на каждом электроде будет пропорциональна его ширине в случае аподизации ширины либо степени его перекрытия с соседними электродами, имеющими противоположную полярность, в случае аподизации апертуры (рис. П3.3).

Недостатком преобразователя с аподизацией ширины электродов (рис. П3.3, а) являются весьма жёсткие требования к точности их нанесения. Это обстоятельство является особенно существенным на высоких частотах, когда ширина электродов составляет единицы микрон и менее.

Другая возможность – аподизация апертуры (рис. П3.3, б). Этот метод имеет два серьёзных недостатка. Во-первых, на участках с малым перекрытием штырей наблюдается дифракция акустической волны. Во-вторых, скорость распространения поверхностной волны на металлизированной поверхности несколько ниже, чем на свободной. Вследствие этого в преобразователе с переменной величиной

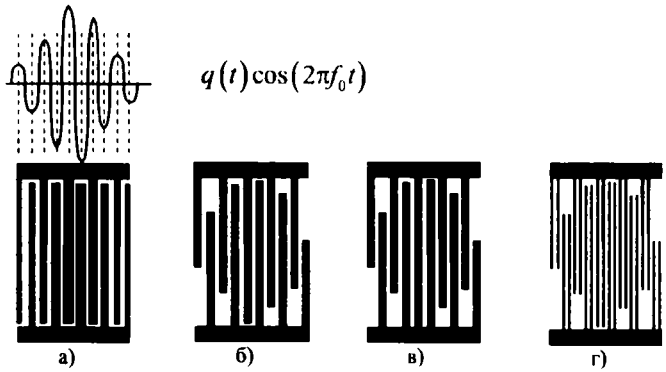


Рис. П3.3. Импульсный отклик и его реализация в аподизованных ВШП: а) ВШП с переменной шириной электродов; б) ВШП с переменной апертурой; в) ВШП с переменной апертурой и ложными электродами; г) ВШП с переменной апертурой и расщеплёнными штырями

перекрытия штырей степень металлизации неоднородна по апертуре, поэтому распространение ПАВ под преобразователем будет сопровождаться нарушением прямолинейности волнового фронта. Данные недостатки проявляются в виде значительного искажения импульсного отклика преобразователя. С целью их преодоления в преобразователь вводят дополнительные пассивные электроды, которые не принимают участия в генерации сигнала (рис. П3.3, в). При этом дифракция и искажение фронта волны существенно уменьшаются, хотя полного устранения этих явлений не происходит, так как активные и пассивные электроды по-разному влияют на распространение ПАВ. В настоящее время этот метод является самым распространенным методом аподизации.

Для точного воспроизведения требуемого импульсного отклика необходимы протяженные преобразователи с большим числом штырей. Расстояние между штырями ВШП равно половине длины ПАВ. Хотя коэффициент отражения ПАВ от одиночного штыря достаточно мал, отражаемые от штырей элементарные волны будут складываться, и коэффициент отражения от протяжённого ВШП может быть довольно значительным. Для уменьшения отражения применяют преобразователи с расщепленными штырями (рис. П3.3, г). Так как в этом случае расстояние между штырями составляет четверть длины волны, отражённые от штырей ПАВ

будут складываться в противофазе. При этом, однако, также ужесточаются требования к точности нанесения штырей.

Точное воспроизведение заданной АЧХ, как выше говорилось, требует бесконечно большого числа отсчётов импульсного отклика и, следовательно, бесконечно большого числа электродов. На практике число электродов ограничено. В результате АЧХ преобразователя искажается. Это проявляется в виде значительных пульсаций в полосе пропускания и наличии боковых лепестков в полосе заграждения. Данное поведение АЧХ преобразователя с конечным числом электродов получило название явления Гиббса. Математически оно обусловлено плохой аппроксимацией функции рядом Фурье с конечным числом членов, если эта функция имеет точки с большой производной. В полосовых фильтрах такими точками являются границы полосы пропускания.

Для борьбы с явлением Гиббса применяют так называемые оконные функции. Оконная функция представляет собой последовательность коэффициентов, на которые при аподизации домножаются дискретные отсчёты импульсного отклика ВШП. Поскольку умножение во временной области соответствует свёртке в частотной области, частотная характеристика получаемого фильтра является сверткой исходной частотной характеристики и частотной характеристики оконной функции. Следовательно, «хорошая» функция окна должна иметь частотную характеристику с узким центральным лепестком и малой амплитудой боковых лепестков, так как ширина центрального лепестка определяет крутизну среза АЧХ на границах полосы пропускания, а уровень боковых лепестков – пульсации АЧХ вблизи этих границ. К сожалению, данное требование является противоречивым. Но для практических приложений найден ряд оконных функций, позволяющих получать неплохие результаты.

Очевидно, что простое ограничение числа членов разложения Фурье эквивалентно использованию прямоугольной функции окна. Спектр прямоугольной функции имеет узкий центральный лепесток, но значительные боковые лепестки, что, как указывалось, приводит к неудовлетворительным результатам. Не останавливаясь на принципах построения оконных функций, отметим, что во многих случаях удовлетворительные результаты дает «окно Хэмминга», определяемое соотношением

$$w_n = 0.52 + 0.48\cos\frac{\pi n}{N}, \quad -\frac{N}{2}+1 \le n \le \frac{N}{2}. \qquad (П3.10)$$

Применение «окна Хэмминга» позволяет снизить пульсации АЧХ в полосе пропускании и значительно уменьшить уровень боковых лепестков в полосе заграждения. Однако крутизна среза АЧХ на границах полосы пропускания несколько хуже, чем в случае прямоугольного окна.

Контрольные вопросы

1. Дайте определение передаточной функции и импульсного отклика фильтра.
2. При каких условиях импульсный отклик ВШП порождает полосовую АЧХ?
3. Какие способы аподизации могут быть использованы для синтеза полосового фильтра? Сформулируйте их достоинства и недостатки.
4. Что такое оконная функция и для чего она применяется?

Расчётные задания

1. Для центральной частоты и полосы пропускания, задаваемых преподавателем, синтезируйте импульсный отклик полосового фильтра.
2. Ограничившись 3 лепестками импульсного отклика, рассчитайте и постройте функцию аподизации ВШП, определите число электродов и шаг между ними. Рассчитайте и постройте АЧХ аподизованного преобразователя, размещённого на подложке из ниобата лития *Y,Z*-ориентации.
3. Рассчитайте и постройте АЧХ аподизованного преобразователя с помощью оконной функции Хэмминга.

Задача 4. Взаимодействие акустических волн с потоком заряженных частиц. Плёночный усилитель

Известно, что электромагнитная волна, распространяющаяся в направлении потока заряженных частиц, может достаточно сильно с ними взаимодействовать. Для этого необходимо, во-первых, чтобы скорости движения частиц и волны были близки (условие синхронизма), во-вторых, волна должна содержать в своём составе компоненту электрического поля вдоль направления распространения. При определённых условиях такое взаимодействие приводит к усилению волны за счёт кинетической энергии частиц. Это явление получило успешное применение в приборах вакуумной электроники – лампе бегущей волны (ЛБВ) и лампе обратной волны (ЛОВ). Так как скорость движения свободной электромагнитной волны равна скорости света, а в своём составе она содержит только поперечное электрическое поле, в ЛБВ и ЛОВ для её замедления и придания ей необходимого компонентного состава используются специальные замедляющие структуры.

Распространение упругих волн в пьезоактивных средах, как объёмных (ОАВ), так и поверхностных (ПАВ), сопровождается переменным электрическим полем. Эти волны содержат в своём составе продольную компоненту пьезоэлектрического поля, а скорость их распространения близка к скорости дрейфового движения свободных носителей заряда в полупроводниковых материалах под действием постоянного электрического поля. Например, при подвижности носителей заряда $\mu = 10^4\ \text{см}^2/(\text{В}\cdot\text{с})$ и напряжённости тянущего поля $E = 30\ \text{В/см}$ их дрейфовая скорость составляет $3000\ \text{м/с}$, что примерно соответствует скорости звука в типичных пьезоэлектриках. Это служит основанием для создания твердотельных аналогов ЛБВ и ЛОВ и исследования их свойств.

Возможны различные варианты взаимодействия акустических волн со свободными носителями заряда (рис. П4.1).

Взаимодействие с объёмными волнами (рис. П4.1, а) можно получить в полупроводниковых материалах, обладающих пьезоэлектрическими свойствами. К числу таких материалов относятся многие бинарные соединения $A^{III}B^{V}$ и $A^{II}B^{VI}$, например, арсенид галлия (GaAs) и сульфид кадмия (CdS). Для достижения эффективного

взаимодействия ОАВ со свободными носителями заряда и управления параметрами усилителя необходима возможность независимого выбора пьезоэлектрических и электрофизических параметров кристалла. Это существенно ограничивает применение объёмных волн.

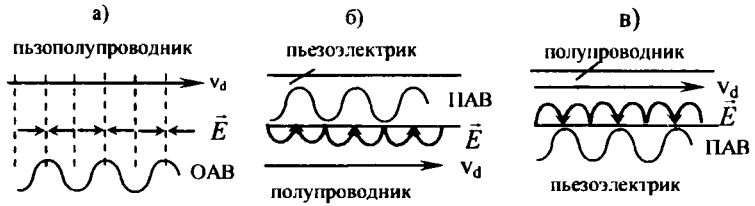


Рис П4.1. Варианты акустоэлектронного взаимодействия свободных носителей заряда с пьезоэлектрическим полем акустической волны: а) взаимодействие ОАВ со свободными носителями в пьезополупроводнике; б) взаимодействие ПАВ и свободных носителей в структуре пьезоэлектрическая плёнка на полупроводниковой подложке; в) взаимодействие ПАВ и свободных носителей в структуре полупроводниковая плёнка на пьезоэлектрической подложке

В структурах, в которых пьезоэлектрическая плёнка размещена на полупроводниковой подложке (рис. П4.1, б), либо полупроводниковая плёнка – на пьезоэлектрической подложке (рис. П4.1, в), используются поверхностные акустические волны, электрическое поле которых проникает в объем полупроводника. В этом случае параметры пьезоэлектрика и полупроводника могут подбираться независимо.

Наиболее перспективна структура с полупроводниковой плёнкой на пьезоэлектрической подложке. К этому имеются следующие основания. Во-первых, нет смысла использовать толстый полупроводниковый кристалл, так как глубина проникновения электрического поля ПАВ в полупроводник равна длине экранирования Дебая. Поэтому в акустоэлектронном взаимодействии участвует только тонкий приповерхностный слой полупроводника. Остальная часть кристалла будет только поглощать мощность внешнего источника, создающего тянущее поле. Во-вторых, данная структура хорошо согласуется с конструкциями других устройств на ПАВ и может быть частью большой интегральной акустоэлектронной схемы, размещённой на общей пьезоэлектрической подложке. И в-третьих, акустическая волна в тонкой пьезоэлектрической плёнке подвержена дисперсии, что может вызвать нежелательную дополнительную зависимость коэффициента усиления от частоты.

Рассмотрим взаимодействие акустических волн со свободными носителями заряда на примере объёмных продольных волн. Для этого случая имеется простое аналитическое решение, позволяющее установить основные особенности взаимодействия акустических волн с дрейфующими свободными носителями заряда.

Для невырожденного полупроводника n-типа расчёт приводит к следующему выражению для постоянной распространения акустической волны κ :

$$\kappa = \frac{\omega}{v} - j\alpha, \qquad \alpha = \frac{\omega}{v}\frac{K^2}{2}\frac{\dfrac{\omega_c}{\gamma\omega}}{1+\dfrac{\omega_c^2}{\gamma^2\omega^2}\left(1+\dfrac{\omega^2}{\omega_c\omega_d}\right)^2}. \qquad (П4.1)$$

В формулах (П4.1) использованы следующие обозначения:

$\omega_c = qn\mu/(\varepsilon\varepsilon_0)$ – частота проводимости, q – заряд электрона, μ – подвижность электронов, n – их концентрация, ε_0 – абсолютная диэлектрическая проницаемость вакуума, ε – относительная диэлектрическая проницаемость полупроводника;

$\omega_d = v^2/D$ – диффузионная частота электронов, v – фазовая скорость волны, D – коэффициент диффузии электронов;

$\gamma = 1 - \mu E_0/v = (v - v_d)/v$ – фактор дрейфа, v_d – скорость дрейфа электронов, E_0 – напряжённость постоянного электрического поля, создающего дрейф носителей;

K – коэффициент электромеханической связи.

Зависимость коэффициента затухания от параметра дрейфа, рассчитанная по формуле (П4.1), представлена на рис. П4.2.

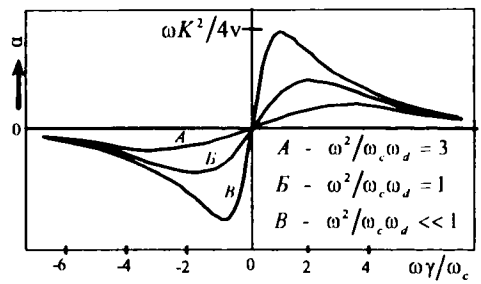


Рис. П4.2. Зависимость коэффициента затухания от параметра дрейфа

Смысл параметра $\omega^2/\omega_c\omega_d$ определим несколько позже, а сейчас предположим, что $\omega^2/\omega_c\omega_d << 1$, и объясним общий ход зависимости. Для этого удобно перейти в систему отсчёта, движущуюся со скоростью дрейфа электронов. В данной системе отсчёта $\omega\gamma$ есть эффективная частота волны (с учётом эффекта Доплера), а $\gamma\omega/\omega_c$ – отношение эффективной частоты к частоте проводимости.

Частота проводимости является величиной, обратной времени максвелловской релаксации $\omega_c = 1/\tau_m$. Как известно из электродинамики, τ_m – время, необходимое для свободных зарядов в проводящей среде, чтобы перегруппироваться в соответствии с распределением электрического поля.

Таким образом, если частота изменения гармонического электрического поля меньше частоты проводимости $\omega < \omega_c$, то свободные заряды успевают перегруппироваться. В обратном случае $\omega > \omega_c$ группировка ослабляется, и в предельном случае $\omega >> \omega_c$ полностью исчезает.

Когда скорость электронов близка к фазовой скорости волны ($|\gamma| << 1$), эффективная частота упругой волны в движущейся системе отсчёта мала. Если эта частота меньше частоты проводимости, электроны скапливаются в потенциальных ямах, образованных периодическим пространственным распределением продольного электрического поля, образуя сгустки (рис. П4.3).

Причём в случае $\gamma = 0$ электронные сгустки располагаются в точках с нулевым значением поля, и волна не испытывает влияние электронов ($\alpha = 0$). Если $\gamma > 0$, то есть $v_d < v$, то волна опережает электроны, а сгустки смещаются в область ускоряющего электрического поля. Электроны, испытывая ускоряющее действие, увлекаются волной, их кинетическая энергия возрастает, а волна, теряя энергию, постепенно затухает. Коэффициент затухания $\alpha > 0$. В противоположном случае $\gamma < 0$ ($v_d < v$) электронные сгустки, находясь в тормозящем поле, теряют свою кинетическую энергию, которая передаётся волне, что приводит к её усилению. Коэффициент затухания $\alpha < 0$.

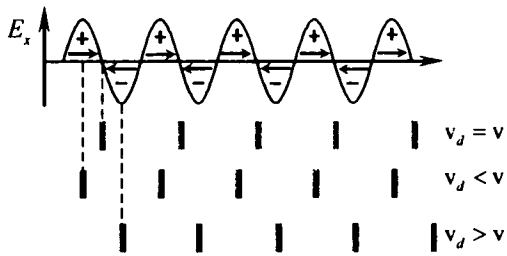


Рис. П4.3. Положение электронных сгустков в зависимости от соотношения скоростей дрейфа электронов и фазовой скорости упругой волны. Знаками «+» и «–» обозначены соответственно ускоряющие и замедляющие полупериоды волны, стрелками – направление силы, действующей на электроны со стороны волны

Когда эффективная частота в движущейся системе отсчёта становится значительно больше частоты проводимости $|\gamma\omega/\omega_c| >> 1$, электроны находятся в быстропеременном поле и не скапливаются в сгустки. Взаимодействие электронов с волной ослабляется, и коэффициент затухания стремится к нулю.

На величину постоянной затухания α влияет также параметр $\omega^2/\omega_c\omega_d$. Смысл этого параметра можно раскрыть, используя известное соотношение Эйнштейна:

$$D = \frac{\mu kT}{q}. \qquad (П4.2)$$

Здесь k – постоянная Больцмана, T – абсолютная температура. С помощью несложных преобразований можно получить следующее выражение:

$$\frac{\omega^2}{\omega_c\omega_d} = \left(2\pi\frac{L_D}{\Lambda}\right)^2, \qquad (П4.3)$$

где $\Lambda = \omega/(2\pi v)$ – длина упругой волны, L_D – длина экранирования Дебая:

$$L_D = \sqrt{\varepsilon kT/q^2 n} \qquad (П4.4)$$

Она определяет минимальный размер электронного сгустка, ограниченный тепловой диффузией электронов. Если длина волны Λ велика по сравнению с L_D, то есть $\omega^2/\omega_c\omega_d << 1$, то при усилении все

электроны скапливаются в области максимального тормозящего поля, испытывая максимальное тормозящее воздействие (рис. П4.4).

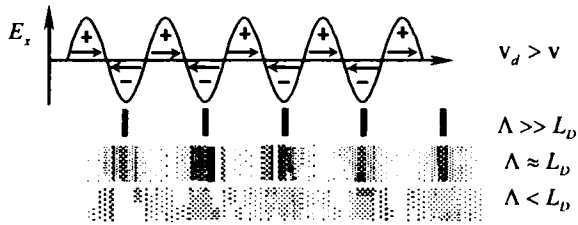


Рис. П4.4. Зависимость размеров электронных сгустков от соотношения длины акустической волны и длины экранирования Дебая

По мере увеличения частоты размеры электронного сгустка постепенно сравниваются с протяжённостью области тормозящего поля, эффективность взаимодействия уменьшается, и усиление падает. Наконец, при $L_D >> \Lambda$ диффузия практически полностью размывает электронные сгустки, и усиление практически отсутствует.

Уменьшение коэффициента усиления при возрастании частоты, связанное с диффузионным размыванием сгустков, получило название диффузионного ограничения.

Заметим, что в устройствах вакуумной электроники – ЛБВ и ЛОВ частотная зависимость коэффициента усиления обусловлена исключительно дисперсионными свойствами структуры, замедляющей электромагнитные волны. Тогда как акустические волны не имеют дисперсии, а зависимость коэффициента усиления от частоты при акустоэлектронном взаимодействии определяется параметрами электронной плазмы.

Как уже говорилось, наиболее перспективной следует считать схему взаимодействия, при которой полупроводниковая плёнка размещается на пьезоэлектрической подложке. Усилитель, в котором используется данный тип взаимодействия, называют плёночным. Его конструкция представлена на рис. П4.5. Усилитель состоит из пьезоэлектрической подложки с нанесенными на нее входным и выходным встречно-штыревыми преобразователями и полупроводниковой пленкой, на которую подается постоянное смещение от внешнего источника. Дополнительно усилитель может быть снабжен системой охлаждения (на рисунке не показана).

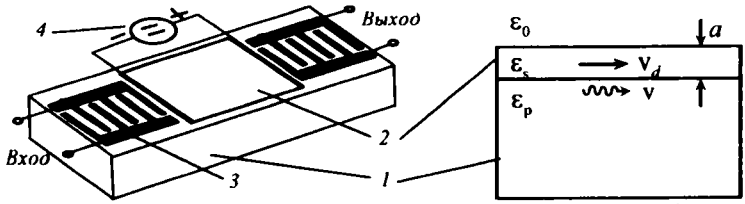


Рис. П4.5. Пленочный усилитель:
1 – пьезоэлектрическая подложка, 2 – полупроводниковая пленка,
3 – ВШП, 4 – источник постоянного смещения

Процесс взаимодействия электрического поле ПАВ и свободных носителей заряда происходит в полупроводниковой плёнке, толщина которой a выбирается из условия $a < L_D$, так что весь объём плёнки участвует в процессе усиления. Удельный коэффициент усиления плёночного усилителя в случае невырожденного полупроводника с одним типом носителей – электронами равен

$$G = -8{,}68\frac{K^2}{2}\frac{\omega}{v}\frac{\gamma\varepsilon_r a\omega_c/v}{\left(\omega/\omega_D + \varepsilon_r a\omega_c/v\right)^2 + \gamma^2}\ [дБ/м], \qquad (П4.5)$$

где $\varepsilon_r = \varepsilon_s/\left(1+\varepsilon_p\right)$, ε_s – относительная диэлектрическая проницаемость полупроводниковой плёнки, ε_p – относительная диэлектрическая проницаемость пьезоэлектрической подложки. Как и при взаимодействии с объёмной волной, для усиления необходимо, чтобы скорость электронов превосходила скорость ПАВ ($\gamma < 0$). Максимальное усиление наступает при факторе дрейфа:

$$\gamma_{max} = -\left(\omega/\omega_d + a\varepsilon_r\omega_c/v\right). \qquad (П4.6)$$

Коэффициент усиления при заданных величинах ω и γ по мере роста толщины плёнки сначала увеличивается, достигая максимального значения, когда выполняется условие

$$a_{max} = \sqrt{\gamma^2 + \left(\omega/\omega_D\right)^2}\Big/\left(\varepsilon_r\omega_c/v\right), \qquad (П4.7)$$

а затем уменьшается. Такое поведение коэффициента усиления связано с тем, что с ростом толщины (вплоть до L_D) увеличивается область взаимодействия, но одновременно падает электрическое сопротивление плёнки. При малых a сопротивление плёнки велико, и с увеличением толщины коэффициент усиления растёт. Когда толщина становится больше a_{max}, из-за малого электрического со-

противления плёнка начинает шунтировать пьезоэлектрическое поле ПАВ, акустоэлектронное взаимодействие ослабляется и коэффициент усиления снижается.

На величину усиления влияет диффузия носителей заряда в плёнке. Как видно из (П4.5), коэффициент усиления при оптимальном факторе дрейфа (П4.6) линейно растёт с увеличением частоты пока $\omega/\omega_c\omega_d << a\varepsilon_r/\mathrm{v}$. С дальнейшим ростом частоты он достигает максимума, когда $\omega/\omega_c\omega_d = a\varepsilon_r/\mathrm{v}$, и затем уменьшается вследствие диффузионного ограничения.

При расчёте усилителя необходимо учитывать величину тепловой мощности, рассеиваемой плёнкой. В расчёте на единицу площади это мощность равна

$$P^T = qn\mu aE_0^2 . \qquad (П4.8)$$

Предельная величина рассеиваемой пленкой тепловой мощности, ограниченная её тепловым разрушением, составляет $\approx 0{,}1$ Вт/см^2. Применяя принудительное охлаждение, можно довести это величину до $0{,}3 \div 0{,}4$ Вт/см^2.

Так как усиление возможно только когда скорость электронов $\mathrm{v}_d = \mu E$ превосходит скорость ПАВ, то минимальная (пороговая) мощность, рассеиваемая плёнкой, при которой начинается усиление, определится из условия $\gamma = 0$ и будет равна:

$$P^T_{\min} = \frac{qna\mathrm{v}^2}{\mu} . \qquad (П4.9)$$

При прочих равных условиях для того чтобы снизить пороговую мощность и уменьшить нагрев усилителя, следует выбирать материалы с большой подвижностью свободных носителей заряда. Для плёнки антимонида индия InSb толщиной 0,1 мкм при комнатной температуре $P^T_{\min} \cong 10$ мВт/см^2. В то же время можно взять материал с небольшой подвижностью, но с существенно меньшей концентрацией носителей и получить меньшее значение пороговой мощности. Однако при малой подвижности для выполнения условия $\mathrm{v}_d = \mu E > \mathrm{v}$ требуется более высокое напряжение питания.

Динамический диапазон плёночного усилителя ограничивается сверху наступлением концентрационной нелинейности. Она наступает при плотности энергии ПАВ в несколько ватт на сантиметр в

квадрате. Снизу динамический диапазон ограничивается собственными шумами усилителя уровнем $\approx 0{,}1$ мВт/см^2.

Практическая реализация плёночного усилителя осложняется большими технологическим трудностями получения плёнки, связанными как с подбором пар материалов с близкими значениями периодов кристаллических решёток и температурных коэффициентов расширения, так и со сложностью самого процесса выращивания на пьезоэлектрической подложке полупроводниковой плёнки с заданными электрофизическими параметрами.

Указанные технологические трудности можно преодолеть в плёночном усилителе с воздушным зазором, у которого в отличие от рассмотренной выше конструкции плёнка наносится на диэлектрическую подложку и приподнята над подложкой, так что между ними образуется небольшой воздушный зазор (рис. П4.6).

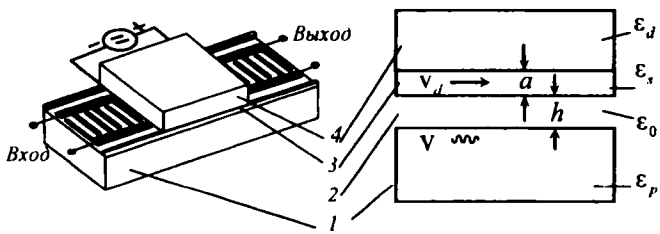


Рис. П4.6. Конструктивная схема усилителя с воздушным зазором между полупроводниковой пленкой и пьезоэлектрической подложкой: *1* – пьезоэлектрическая подложка, *2* – воздушный зазор, *3* – полупроводниковая пленка, *4* – диэлектрик

Коэффициент усиления данного усилителя определяется следующим выражением:

$$C = -8{,}68\frac{K^2\omega}{2v}\frac{\gamma\varepsilon_r a\omega_c/v}{\left[\omega/\omega_D+\left(1+\omega h\varepsilon_p/v\right)\varepsilon_r a\omega_c/v\right]^2+\gamma^2}\left[\frac{\text{дБ}}{\text{м}}\right]. \quad (\text{П4.10})$$

Предполагается, что толщина пленки и воздушного зазора удовлетворяет следующим условиям: $a < L_D$ и $h\omega/v << 1$, где h – величина зазора. При заданной частоте и факторе дрейфа оптимальная толщина пленки равна:

$$a_{max} = \frac{\sqrt{\gamma^2+\left(\omega/\omega_D\right)^2}}{\left(\varepsilon_r\,\omega_c/v\right)\left(1+\omega h\varepsilon_p/v\right)}. \quad (\text{П4.11})$$

Поскольку напряженность электрического поля волны на поверхности полупроводника за счет зазора меньше, чем на поверхности пьезоэлектрической подложки, оптимальное значение толщины несколько меньше, чем в усилителе без зазора.

Ограничение роста усиления с увеличением частоты в этом усилителе связано главным образом с наличием зазора, а не с процессом диффузии. Коэффициент усиления достигает своего максимального значения уже на частоте

$$\omega = (v/h)(1/\varepsilon_p), \qquad (П4.12)$$

которая обычно значительно меньше соответствующего значения для усилителя без зазора. Таким образом, наличие зазора заметно сужает полосу усиливаемых частот. Оценки показывают, что коэффициент усиления усилителя с зазором на частоте 100 МГц при зазоре $h = 0{,}1$ мкм и оптимальных параметрах пленки не превышает 60 дБ/см, тогда как для усилителя без зазора он может достигать 100 дБ/см и более.

Контрольные вопросы

1. Каковы условия эффективного взаимодействия электромагнитной волны и потока заряженных частиц?
2. Перечислите возможные схемы акустоэлектронного взаимодействия свободных носителей заряда полупроводника и упругой волны.
3. При каких условиях наблюдается усиление в акустоэлектронном усилителе? В чём заключается диффузионное ограничение и каковы его причины?
4. Приведите варианты конструкции плёночных усилителей. Сформулируйте достоинства и недостатки каждой конструкции.
5. Чем ограничивается динамический диапазон плёночного усилителя?

Расчётные задания

Рассчитайте и постройте частотную зависимость коэффициента усиления плёночного усилителя. Подберите оптимальное напряжение питания и концентрацию носителей, обеспечивающие максимально возможное усиление. Определите полосу усиления. Найдите пороговое значение мощности, рассеиваемое плёнкой.

Выполнить задание для двух типов плёночных усилителей – без зазора и с зазором 0,5 мкм, и двух полупроводниковых материалов – кремния (p-Si) и антимонита индия n-InSb при комнатной температуре. Сравните полученные результаты.

Данные для расчёта:

– длина усилителя 5 мм, толщина полупроводниковой плёнки a =0,05; 0,1; 0,2 мкм;

– максимальная рассеиваемая мощность 0,1 Вт/см^2;

– диэлектрическая проницаемость подложки $\varepsilon_p = 52$;

– коэффициент электромеханической связи $K = 0,22$;

Электрофизические параметры полупроводниковой плёнки:

p-Si: $\varepsilon_s = 11,7$; $\mu = 130$ см$^2/(\text{В}\cdot\text{с})$

n-InSb: $\varepsilon_s = 16$; $\mu = 68000$ см$^2/(\text{В}\cdot\text{с})$

Таблица III

Параметры некоторых пьезоэлектрических подложек, применяемых в устройствах на ПАВ

Материал подложки	Плоскость среза, направление распространения	Скорость ПАВ, м/с	Коэфф. поглощения ПАВ		Коэфф. э/м связи K	Отн. диэл. проницаемость, ε_p	Парам. анизотр., γ
			α_m	β_m			
Кварц SiO_2	Y,X ST,X	3159 3158	6,80 8,30	1,42 1,49	0,042 0,035	4,52 4,55	0,653 0,378
Тантал-лат лития $LiTaO_3$	Z,Y	3329	2,31	0,96	0,11	47,9	1,24
Германат висмута $Bi_{12}GeO_{20}$	[001],[100] [111],[100]	1681 1708	8,63 8,49	1,13 1,11	0,12 0,13	43,6	-0,3 0,35
Ниобат лития $LiNbO_3$	Z,Y 41,5°,X	3488 4000	2,52 1,88	0,54 0,75	0,22 0,24	50,2	-1,08 -0,45

ЛИТЕРАТУРА

1. *Кравченко А.Ф.* Физические основы функциональной электроники: Учебное пособие. Новосибирск: Изд-во Новосиб. ун-та, 2000. 444 с.
2. *Щука А.А.* Функциональная электроника: Учебник для вузов. М.: МИРЭА, 1998. 261 с.
3. *Щука А.А.* Электроника: Учебное пособие / Под ред. проф. А.С. Сигова. СПб.: БХВ-Петербург, 2005. 800 с.
4. *Степанченко И.П.* Основы микроэлектроники: Учебное пособие для вузов. М.: Лаборатория базовых знаний, 2004. 488 с.
5. Микроэлектроника и полупроводниковые приборы / Сборник статей под общей редакцией А.А. Васенкова и Я.А. Федотова. Вып. 10. М.: Радио и связь, 1989. 192 с.
6. *Федотов Я.А., Старостин В.В., Щука А.А.* Электроника четвертого поколения (о функциональной электронике). М.: Знание, 1985. 64 с.
7. *Васенков А.А., Федотов Я.А.* Функциональная электроника. Основные направления работ // Электронная промышленность. М.: ЦНИИ «Электроника», 1983. Вып. 8. С. 3-5.
8. *Зюбрин А.И., Бурак Я.В., Савицкий И.В.* Акустоэлектроника: Учебное пособие. Львов: Львов. ун-т, 1980. 100 с.
9. *Сугано Т., Икома Т., Такэиси Е.* Введение в микроэлектронику. М.: Мир, 1988. 320 с.
10. *Кравченко А.Ф., Талипов Н.Х., Шамирзиев Т.С.* Задачи по физическим основам информационных технологий. Новосибирск: Изд-во Новосиб. ун-та, 2000. 24 с.
11. *Ландау Л.Д., Лифшиц Е.М.* Теория упругости. М.: Наука, ГРФМЛ, 1965. 202 с.
12. *Най Дж.* Физические свойства кристаллов. М.: Мир, 1967. 386 с.
13. *Киттель Ч.* Введение в физику твердого тела. М.: ГИФМЛ, 1965. 696 с.
14. *Викторов И.А.* Звуковые поверхностные волны в твердых телах. М.: Наука, 1981. 288 с.
15. *Олинер А., Фарнел Дж., Джерард Г., Эш Э., Слободник Э., Смит Г.* Поверхностные акустические волны / Под ред. А. Олинера. М.: Мир, 1981. 390 с.
16. Физическая акустика. Принципы и методы / Под ред. У. Мэзона, Р. Терстона. Т. 7. М.: Мир, 1974. 430 с.
17. *Руденко О.В., Солуян С.И.* Теоретические основы нелинейной акустики. М.: Наука, ГРФМЛ, 1975. 288 с.
18. *Зарембо Л.К., Красильников В.А.* Введение в нелинейную акустику. М.: Наука, ГРФМЛ, 1966. 520 с.
19. *Лямов В.Е.* Поляризационные эффекты и анизотропия взаимодействия акустических волн в кристаллах. М.: МГУ, 1983. 224 с.
20. *Бондаренко В.С., Бочков Б.Г., Громашевский В.Л., СоболевБ.В.* Нелинейные акустоэлектронные устройства и их применение. М.: Радио и связь, 1985. 160 с.
21. *Зарембо Л.К., Тимошенко В.И.* Нелинейная акустика. М.: МГУ, 1984. 104 с.
22. *Скучик Е.* Основы акустики. Т. 2. М.: Мир, 1976. 542 с.

23. Физическая акустика. Принципы и методы / Под ред. У. Мэзона, Р. Терстона. Т. 6. М.: Мир, 1973. 432 с.
24. *Лепендин Л.Ф.* Акустика: Учебное пособие для вузов. М.: Высшая школа, 1978. 448 с.
25. *Виноградова М.Б., Руденко О.В., Сухоруков А.П.* Теория волн. М.: Наука, ГРФМЛ, 1990. 432 с.
26. *Кулаков А.В., Румянцев А.А.* Введение в физику нелинейных процессов. М.: Наука, 1988. 160 с.
27. *Бондаренко В.С., Бочков Б.Г., Громашевский В.Л., Соболев Б.В.* Нелинейные акустоэлектронные устройства и их применение / Под ред. В.С. Бондаренко. М.: Радио и связь, 1985. 160 с.
28. *Каринский С.С.* Устройства обработки сигналов на ультразвуковых поверхностных волнах. М.: Советское радио, 1975. 176 с.
29. *Речицкий В.И.* Акустоэлектронные радиокомпоненты: элементы и устройства на поверхностных акустических волнах. М.: Советское радио, 1980. 264 с.
30. *Морган Д.* Устройства обработки сигналов на поверхностных акустических волнах. М.: Радио и связь, 1990. 416 с.
31. *Кочемасов В.Н., Долбня Е.В., Соболь Н.В.* Акустоэлектронные Фурье-процессоры. М.: Радио и связь, 1987. 168 с.
32. *Кайно Г.* Акустические волны: устройства, визуализация и аналоговая обработка сигналов. М.: Мир, 1990. 656 с.
33. *Гуляев Ю.В., Медведь А.В.* Некоторые перспективы совершенствования акустоэлектронных устройств для обработки информации // Проблемы современной радиотехники и электроники. М.: Наука, 1987. С. 187-195.
34. *Гуляев Ю.В.* Акустоэлектронные устройства для систем связи и обработки информации / Сборник статей под ред. акад. В.А. Котельникова Проблемы современной радиотехники и электроники. М.: Наука, 1980. С. 297-319.
35. Фильтры на поверхностных акустических волнах (расчёт, технология и применение): Пер. с англ. / Под ред. Г. Мэтьюза. М.: Радио и связь, 1981. 472 с.
36. *Балышева О. Л.* Материалы для акустоэлектронных устройств: Учебное пособие. СПб.: ГУАП, 2005. 50 с.
37. *Дмитриев В. Ф.* Устройства интегральной электроники: Акустоэлектроника. Основы теории, расчёта и проектирования: Учебное пособие. СПб.: ГУАП, 2006. 169 с.

СОДЕРЖАНИЕ

Предисловие ... 3

Введение в функциональную электронику ... 6
Контрольные вопросы ... 29

1. Элементы теории упругости ... 30
1.1. Понятие деформации. Тензор деформаций. Виды деформации ... 30
1.2. Упругие силы. Тензор упругих напряжений ... 40
1.3. Закон Гука. Тензор модулей упругости. Сокращенная запись упругих модулей ... 44
1.4. Упругая энергия. Свойства матрицы упругих постоянных ... 49
1.5. Упругие свойства кубического кристалла и изотропной среды ... 55
1.6. Деформация пьезоэлектриков ... 63
Контрольные вопросы ... 73

2. Акустические волны в различных средах ... 74
2.1. Движения в упругом теле ... 74
2.2. Объемные упругие волны в изотропной среде ... 77
2.3. Упругие объемные волны в кристаллах в отсутствие Пьезоэффекта ... 82
2.4. Объемные упругие волны пьезоэлектриках ... 86
2.5. Поверхностные акустические волны в изотропном Полупространстве ... 91
2.6. Акустические волны в полупространстве со слоем ... 100
2.7. Волны Лява в полупространстве со слабой поверхностной неоднородностью ... 107
2.8. Волны Стоунли ... 111
2.9. Нормальные волны в тонких пластинах ... 113
2.10. Волны на поверхности непьезоактивных кристаллов ... 120
2.11. Волны на поверхности пьезоэлектриков ... 126
2.12 Распространение акустических волн в пьезоэлектрических Полупроводниках ... 133
Контрольные вопросы ... 145

3. Нелинейные эффекты взаимодействия акустических волн ... 146
3.1. Нелинейные взаимодействия акустических волн ... 146
3.2. Взаимодействие акустических волн с высокочастотными электрическими полями ... 156
3.3. Анализ возможных трехволновых взаимодействий в пьезоэлектрических кристаллах ... 164

3.4. Нелинейные свойства твердых тел 173
3.5. Интегральные преобразования сигналов, получаемые в результате нелинейных взаимодействий акустических волн 178
3.6. Классификация нелинейных акустоэлектронных устройств обработки информации 181
Контрольные вопросы 184

4. Устройства обработки информации на акустических волнах 185
4.1. Устройства на объемных акустических волнах 185
4.2. Возбуждение и прием поверхностных акустических волн 202
4.3. Встречно-штыревой преобразователь 206
4.4. Многополосковый ответвитель и его применение 216
4.5. Отражательные решетки 228
4.6. Линии задержки на поверхностных акустических волнах 236
4.7. Резонаторы на поверхностных акустических волнах 242
4.8. Акустоэлектронный усилитель 248
4.9. Автогенераторы и синтезаторы частоты на поверхностных акустических волнах 255
4.10. Полосовые фильтры на поверхностных акустических волнах 262
4.11. Дисперсионные фильтры на поверхностных акустических волнах 273
4.12. Акустические конвольверы и корреляторы, запоминающие устройства 281
4.13. Устройства спектрального анализа сигналов. Фурье-процессоры 293
Контрольные вопросы 304

Приложение. Расчёт и оптимизация параметров устройств на поверхностных акустических волнах. Компьютерный практикум 307
Задача 1. Возбуждение и приём поверхностных акустических волн 307
Контрольные вопросы 324
Расчётные задания 325
Задача 2. Потери в устройствах на ПАВ. Линии задержки 326
Контрольные вопросы 336
Расчётные задания 336
Задача 3. ПАВ-фильтры на основе встречно-штыревых преобразователей 337
Контрольные вопросы 344
Расчётные задания 344
Задача 4. Взаимодействие акустических волн с потоком заряженных частиц. Плёночный усилитель 345
Контрольные вопросы 354
Расчётные задания 354
Параметры пьезоэлектрических подложек для устройств на ПАВ 356

Литература 358

Учебное издание

Александр Васильевич ВОЙЦЕХОВСКИЙ
Александр Аркадьевич СКРЫЛЬНИКОВ

ФУНКЦИОНАЛЬНАЯ АКУСТОЭЛЕКТРОНИКА

Учебное пособие
Издание 2-е, исправленное и дополненное

Редактор В.С. Сумарокова
Компьютерная верстка Е.В. Дьяковой

Подписано в печать 24.12.2012.
Формат 60x84 1/16. Бумага офсетная № 1. Печать офсетная.
Печ. л. 22,5; усл. печ. л. 20,9 ; уч.-изд. л. 20,0. Тираж 300. Заказ 2.

ООО «Издательство ТГУ», 634029, г. Томск, ул. Никитина, 4
ООО «Интегральный переплет», 634040, г. Томск, ул. Высоцкого, 28, стр. 1