ПРИОРИТЕТНЫЙ НАЦИОНАЛЬНЫЙ ПРОЕКТ «ОБРАЗОВАНИЕ»

РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ

Л.А. СЕВАСТЬЯНОВ, К.П. ЛОВЕЦКИЙ,
А.П. ГОРОБЕЦ, О.Н. БИКЕЕВ

МЕТОДЫ И АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧ В МОДЕЛЯХ ОПТИЧЕСКИХ ПОКРЫТИЙ

Учебное пособие

Москва

2008

Инновационная образовательная программа
Российского университета дружбы народов

«Создание комплекса инновационных образовательных программ и формирование инновационной образовательной среды, позволяющих эффективно реализовывать государственные интересы РФ через систему экспорта образовательных услуг»

Экспертное заключение –

доктор физико-математических наук, профессор *А.В. Крянев*

Севастьянов Л.А., Ловецкий К.П., Горобец А.П., Бикеев О.Н.

Методы и алгоритмы решения задач в моделях оптических покрытий: Учеб. пособие. – М.: РУДН, 2008. – 171 с.

Учебное пособие является вводным в предмет моделирования и методы моделирования взаимодействия света с тонкопленочными оптическими наноструктурами. Предмет моделирования излагается на нескольких уровнях описания. Вначале даются физические основы векторной природы электромагнитного излучения оптического диапазона в вакууме и в бездисперсных однородных средах и средах с дисперсией. Далее рассматривается распространение монохроматических плоских волн в стратифицированной изотропной среде, а также распространение монохроматических плоских волн в анизотропных средах. Заканчивается курс изложением основ отражения и преломления монохроматических плоских волн на периодических структурах.

Для магистров и аспирантов, обучающихся по направлению «Прикладная математика и информатика».

Учебное пособие выполнено в рамках инновационной образовательной программы Российского университета дружбы народов, направление «Комплекс экспортоориентированных инновационных образовательных программ по приоритетным направлениям науки и технологий», и входит в состав учебно-методического комплекса, включающего описание курса, программу и электронный учебник.

Содержание

Общее описание курса.

Учебное пособие открывает цикл учебников, посвященных основам и приложениям методов математического моделирования к предметной области «тонкие пленки сложного оптического состава, характерные размеры неоднородностей которых меньше длины волны падающего света». При этом математическое моделирование понимается авторами в смысле, объединяющем методы теоретической и математической физики и прикладные методы проектирования физических приборов. Предметом математического моделирования являются такие оптические покрытия со сложной геометрической и диэлектрической структурой, взаимодействие с которыми электромагнитного излучения оптического диапазона требует использования модели, базирующейся на уравнениях Максвелла, когда квантовыми эффектами еще можно пренебречь. А именно, учет квантового и нелинейного характера взаимодействия электромагнитного излучения с веществом оптического покрытия вносится в используемые нами математические модели в виде «эффективных параметров», которые получаются после усреднения по небольшим объемам малых квантовых и нелинейных поправок.

Пособие является вводным в предмет моделирования и методы моделирования взаимодействия света с тонкопленочными оптическими наноструктурами. Предмет моделирования излагается на нескольких уровнях описания:

- популярный обзор в первой теме;
- физические основы векторной природы электромагнитного излучения оптического диапазона в вакууме и в бездисперсных однородных средах и средах с дисперсией;
- распространение монохроматических плоских волн в стратифицированной изотропной среде;
- распространение монохроматических плоских волн в анизотропных

средах;

- отражение и преломление монохроматических плоских волн на периодических структурах и т. п.

Актуальные методы моделирования, излагаемые в последующих учебных пособиях цикла «Оптика наноструктур», базируются на конструкциях, которые описываются со все усложняющимися деталями теоретических моделей по мере усложнения предмета исследований.

Предлагаемый курс, как и вся магистерская программа, посвящены, в основном, многослойным оптическим покрытиям, которые являются важной составной частью современных устройств в оптике и оптоэлектронике и широко используются при создании систем телекоммуникаций, электронных дисплеев, оптических запоминающих устройств и полупроводниковых приборов. Постоянно растет потребность в синтезе новых типов оптических покрытий для улучшения параметров подобных устройств. В последнее время сделаны значительные шаги в области развития методов синтеза многослойных оптических покрытий, применяемых при создании жидкокристаллических дисплеев. Большинство известных численных методов синтеза основано на минимизации функции, оценивающей разность между требуемой и реальной спектральными характеристиками покрытия. В рассматриваемой работе основными целями оптимизации являются:

- Моделирование и конструирование ретардеров (фазовых пластинок) для компенсации отрицательного влияния жидкого кристалла и поляроидов на контраст изображения.
- Моделирование многослойных оптических систем для создания широкополосных отражающих покрытий (зеркал).

При расчете параметров оптимальных конструкций необходимо уметь решать две задачи. Первая (прямая задача) заключается в том, чтобы по известным характеристикам материала (диэлектрическая проницаемость и толщина покрытия) определить такие свойства системы, как контрастность (при отклонении от нормального угла зрения), эффективность, цветность и т.п. Вторая (обратная задача) подразумевает определение оптических и/или геометрических (толщины) характеристик отдельного материала по измеренным данным. В нашем случае такими данными служат интенсивности пропускания и отражения при разных углах падения света (получаемые с помощью спектрофотометра) и толщина (в качестве начального приближения) нанесенного на подложку материала (измеряемая профилометром).

ОСНОВНЫЕ СВОЙСТВА ЭЛЕКТРОМАГНИТНОГО ПОЛЯ

Светом следует называть только видимое излучение, от которого нам светло. И на заре оптических исследований возможность визуальной регистрации излучения была действительно принципиальной. Сейчас же техника регистрации значительно усовершенствовалась, так что «видимость» излучения с λ = 400…750 нм не дает особых преимуществ этой области. Так что удобное слово свет может быть отнесено не только к этой области спектра.

На шкале электромагнитных волн (рисунок 1) принято выделять области ядерного, рентгеновского, ультрафиолетового (УФ-), видимого, инфракрасного (ИК-) и радиоизлучения. Это историческое деление, Правильно же будет выделять какую-либо область на основании физического смысла. При этом существенны два критерия – природа генерации и (или) поглощения излучения, либо – особенности и принцип действия технических устройств, преимущественно используемых для работы с ним.

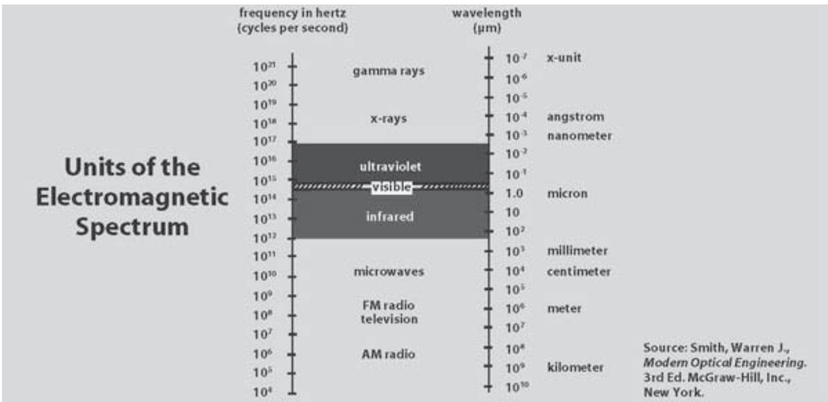

Рис. 1. Шкала электромагнитных волн.

К *оптическому* относят только излучение в области УФ-, видимого и ИК- диапазонов. Генерация и поглощение в этой области возникает при изменении состояния *валентных* электронов и *колебательных* или *вращательных* возбуждений в атомных ансамблях – молекулах и твердых телах. Эти процессы и будут являться для нас основным предметом рассмотрения.

Коротковолновый край оптической области – *рентгеновское* излучение. В понятии *рентгеновского* излучения или *рентгеновской* области спектра заложено предположение о том, что оно связано с возбуждением *остовных*, *внутренних* электронов атомов. Но какого-либо универсального определения того, какие именно электронные состояния нужно отнести к *остовным*, а какие – к *валентным*, дать невозможно. Тем более, невозможно провести четкую энергетическую границу между ними. Поэтому и границу между рентгеновским и УФ- излучением можно провести лишь условно, где-то в области 25÷100 эВ (длина волны $\lambda \approx$ 12÷50 нм). Обычно считают к тому же, что если источник излучения газоразрядный, то это – УФ- излучение.

Длинноволновую границу оптической области еще труднее определить и обосновать. При $h\nu \leq 0{,}5$ эВ ($\lambda > 2{,}5$ мкм) начинается область частот собственных колебаний атомных ионных остовов. Чем тяжелее ион и чем слабее его связь с соседними атомами, тем меньше частота и энергия колебательных квантов. Частоты колебаний атомов в молекулах и кристаллах могут быть сколь угодно малы, практически – до нуля. Энергии электронных переходов снизу тоже не ограничены, так что длинноволновый край "оптической" области можно определить только основываясь на традиционных представлениях или, что лучше, на типах экспериментальных устройств, применяемых для генерации или регистрации излучения. Можно считать, что она простирается до ~ 1 мм (10-3 эВ), далее – радиоволны, с радиоэлектронными методами их генерации и регистрации, использующими законы движения свободных электронов.

Всю оптическую область 50…10–3 эВ (от 25 нм до 1 мм) иногда все же условно делят на "электронную" и "колебательную", с границей около 0,5 эВ. Условность заключается не только в том, что при малых энергиях возможны и электронные переходы, но и в том, что при больших энергиях могут наблюдаться обертоны колебательных возмущений.

При описании принципов работы спектральных приборов обычно ограничиваются рассмотрением "УФ-видимой" области, так что интересующий нас диапазон составит λ = 190…800 нм. Освоив работу в этом узком, но наиболее удобном и доступном для измерений диапазоне, можно освоить и более широкую область.

Тема 1. Электромагнитное поле в недиспергирующей среде

Изложение следует широко известной книге Борна и Вольфа [1]

1.1. Уравнения Максвелла и материальные уравнения связи

При изменении электрических зарядов в пространстве устанавливается возбужденное состояние, которое называют электромагнитным полем. Основными векторами поля считают векторы $\mathbf{E}$ и $\mathbf{H}$, а посредством векторов $\mathbf{D}$ и $\mathbf{B}$ описывают влияние среды.

Замечание. Однако в специальной теории относительности при изучении электродинамики движущихся сред, обязательно иное представление. Уравнения Максвелла (1.1)—(1.4) (см. ниже) можно разделить на две группы. Одну, содержащую $\mathbf{E}$ и $\mathbf{B}$, состоящую из двух однородных уравнений (с правой частью, равной нулю), и другую, содержащую $\mathbf{D}$ и $\mathbf{H}$, — из двух неоднородных уравнений (с правой частью, отличной от нуля). При релятивистском преобразовании пространства и времени (преобразование Лоренца) уравнения каждой группы преобразуются совместно. При этом они остаются инвариантными, если величины $\mathbf{j}/c$ и ρ преобразуются как четырехмерный вектор, а каждая из пар $\mathbf{E}$, $\mathbf{B}$ и $\mathbf{D}$, $\mathbf{H}$ — как антисимметричный тензор второго ранга. Поскольку в группу, состоящую из неоднородных уравнений, входят заряды и токи (которые представляют влияние среды), мы должны считать, что соответствующая пара ($\mathbf{D}$, $\mathbf{H}$) отражает влияние среды.

Пространственные и временные производные характеристик электромагнитного поля связаны уравнениями Максвелла. Эти уравнения имеют вид:

$$rot\,\mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c}\mathbf{j} \tag{1.1}$$

$$rot\,\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} \tag{1.2}$$

$$div\,\mathbf{D} = 4\pi\rho \qquad (1.3)$$

$$div\,\mathbf{B} = 0 \qquad (1.4)$$

Уравнения (1.1) – (1.4) записаны в так называемой гауссовой системе единиц, то есть электрические величины ($\mathbf{E}$, $\mathbf{D}$, $\mathbf{j}$ и ρ) измеряются в электростатических единицах, а магнитные величины ($\mathbf{H}$ и $\mathbf{B}$) в электромагнитных единицах. Постоянная c в (1.1) и (1.2) связывает единицы заряда в обеих системах; она равна скорости света в вакууме и составляет приблизительно $3\cdot 10^{10}\,см/сек$.

Можно считать уравнение (1.3) определением плотности электрического заряда ρ, а соотношение (1.4) означает, что не существует свободных магнитных зарядов. Из уравнения (1.1) следует (поскольку $div\,rot(*) \equiv 0$), что $div\,\mathbf{j} = -\frac{1}{4\pi} div\,\mathbf{D}$, или, используя (1.3),

$$\frac{\partial\rho}{\partial t} + div\,\mathbf{j} = 0 \qquad (1.5)$$

По аналогии с соотношением, встречающимся в гидродинамике, уравнение (1.5) называется уравнением непрерывности. Оно отражает свойство сохранения заряда в окрестности любой точки. Действительно, если мы проинтегрируем (1.5) по некоторой области пространства, то получим, используя теорему Гаусса,

$$\frac{d}{dt}\int \rho dV + \int \mathbf{j}\cdot\mathbf{n} dS = 0, \qquad (1.6)$$

причем первый интеграл берется по объему этой области, а второй по поверхности, ограничивающей ее. Здесь $\mathbf{n}$ — единичный вектор внешней нормали, S - поверхность, охватывающая рассматриваемую область. Из этого уравнения следует, что полный заряд

$$e = \int \rho dV, \qquad (1.7)$$

заключенный внутри рассматриваемой области, может увеличиться только при наличии электрического тока, протекающего через поверхность S

$$J = \int \mathbf{j} \cdot \mathbf{n} dS . \tag{1.8}$$

Если все величины, характеризующие поле, не зависят от времени и отсутствуют токи ($\mathbf{j} = \mathbf{0}$), то такое электрическое поле называют статическим. Если эти величины не зависят от времени, но имеются токи ($\mathbf{j} \neq \mathbf{0}$), то мы говорим о стационарном поле. Векторы электромагнитных полей в области оптических частот очень быстро изменяются во времени, но источники полей обычно таковы, что при рассмотрении не мгновенных значений величин, а усредненных по любому макроскопическому интервалу времени, свойства поля оказываются не зависящими от времени. Мы будем употреблять термин «стационарный» в более широком смысле для описания поля указанного типа.

Уравнения Максвелла (1.1)—(1.4) связывают шесть основных величин $\mathbf{E}, \mathbf{H}, \mathbf{B}, \mathbf{D}, \mathbf{j}$ и ρ. Для того чтобы при заданном распределении зарядов и токов уравнения допускали единственное решение для векторов поля, к этим уравнениям необходимо добавить соотношения, описывающие поведение веществ под влиянием поля. Такие соотношения называются материальными уравнениями. В общем случае они довольно сложны, но для тел, находящихся в покое друг относительно друга (или в состоянии очень медленного движения) и состоящих из изотропных веществ (то есть веществ, физические свойства которых в каждой точке не зависят от направления), эти уравнения принимают относительно простую форму.

В недиспергирующих средах выполняются материальные уравнения связи:

$$\mathbf{j} = \sigma \mathbf{E} \tag{1.9}$$

$$\mathbf{D} = \varepsilon \mathbf{E} \tag{1.10}$$

$$\mathbf{B} = \mu \mathbf{H} \qquad (1.11)$$

Уравнение (1.9) является дифференциальной формой закона Ома. Вещества, для которых $\sigma \neq 0$ (вернее, не равно пренебрежимо малой величине; точный смысл этого неравенства здесь обсуждать невозможно), называются проводниками.

Замечание. Очень хорошими проводниками служат металлы, но существуют и другие классы хорошо проводящих веществ, такие, как ионные растворы в жидкостях, а также в твердых телах. Проводимость металлов уменьшается с повышением температуры. Для других классов веществ, называемых полупроводниками (например, германий), проводимость сильно увеличивается с ростом температуры.

Вещества, для которых σ пренебрежимо мала, называются изоляторами или диэлектриками. Их электрические и магнитные свойства полностью определяются величинами ε и μ. Для большинства веществ магнитная проницаемость μ практически равна единице. Если это не так, то есть если μ заметно отличается от единицы, то мы называем такое вещество магнетиком. В частности, если $\mu > 1$, то вещество называют парамагнетиком (например, платина, кислород, азот), если же $\mu < 1$,— то диамагнетиком (например, висмут, медь, водород, вода).

Для чрезвычайно сильных полей, которые получаются, например, при фокусировке света, генерируемого лазером, к правым частям материальных уравнений, возможно, придется добавить члены, содержащие компоненты векторов поля в степени, большей единицы.

Во многих случаях величины σ, ε и μ не зависят от напряженности полей, однако часто поведение вещества невозможно описать таким простым способом. Например, ток в газе свободных ионов, который определяется средней скоростью ионов, в любой момент времени зависит не от мгновенного значения $\mathbf{E}$, а от всех его предыдущих значений. В так называемых ферромагнитных веществах (вещества, являющиеся очень сильными магнетиками, например железо, кобальт и никель) магнитная индукция $\mathbf{B}$ определяется предысторией поля $\mathbf{H}$, а не его мгновенным значением. В этом случае говорят, что вещество проявляет гистерезис. В некоторых диэлектриках наблюдается подобная же зависимость от предыстории для электрического смещения. Гистерезисные эффекты для высокочастотных полей, встречающихся в оптике, как правило, незначительны.

В основной части настоящего учебного пособия мы будем рассматривать распространение света в таких средах, в которых не происходит заметного его затухания (например, воздух, стекло). Подобные среды называются оптически прозрачными. Они должны быть непроводящими ($\sigma = 0$), поскольку наличие проводимости приводит к выделению джоулева тепла и, следовательно, к диссипации электромагнитной энергии.

1.2. Граничные условия на поверхностях раздела и сохранение энергии электромагнитного поля.

В оптике часто встречается ситуация, когда свойства среды (величины σ, ε и μ) резко меняются на поверхностях разделов двух сред. Выведем соотношения, описывающие этот переход.

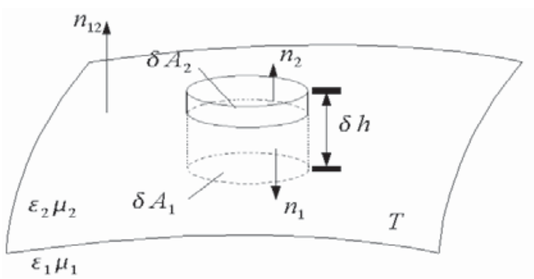

Рис. 2. К выводу граничных условий для нормальных компонент $\mathbf{B}$ и $\mathbf{D}$.

Заменим поверхность резкого раздела T тонким, переходным слоем, внутри которого ε и μ быстро, но непрерывно меняются от значений около T с одной стороны слоя до их значений вблизи T с другой его стороны. Внутри этого слоя построим небольшой квазицилиндр, ограниченный с боков «частоколом» нормалей к T.

Основаниями цилиндра на каждой стороне T служат небольшие площадки δA_1 и δA_2, параллельные поверхности T (рис. 2). Поскольку во всем цилиндре вектор $\mathbf{B}$ и его производные непрерывны, мы можем применить теорему Гаусса к интегралу от $div\,\mathbf{B}$, взятому по объему цилиндра. Тогда, согласно (1.4), получим

$$\int div\,\mathbf{B}dV = \int \mathbf{B}\cdot\mathbf{n}dS = 0 \qquad (1.12)$$

Здесь $\mathbf{n}$ — единичный вектор внешней нормали; второй интеграл берется по поверхности цилиндра.

Так как площадки δA_1 и δA_2 предполагаются малыми, можно считать, что на них $\mathbf{B}$ принимает постоянные значения $\mathbf{B}^{(1)}$ и $\mathbf{B}^{(2)}$. Тогда выражение (1.12) можно заменить следующим:

$$\mathbf{B}^{(1)}\cdot\mathbf{n}_1\delta A_1 + \mathbf{B}^{(2)}\cdot\mathbf{n}_2\delta A_2 + \text{вклад от стенок} = 0. \qquad (1.13)$$

Если высота цилиндра δh стремится к нулю, переходный слой переходит в поверхность, а вклад от стенок цилиндра исчезает при условии, что отсутствует поверхностный поток магнитной индукции. Такой поток никогда не наблюдается и, следовательно, в пределе

$$(\mathbf{B}^{(1)}\cdot\mathbf{n}_1 + \mathbf{B}^{(2)}\cdot\mathbf{n}_2)\delta A = 0 \qquad (1.14)$$

где δA — площадь пересечения нашего цилиндра с поверхностью T. Если $\mathbf{n}_{12}$ — единичный вектор нормали, направленный из первой среды во вторую, то $\mathbf{n}_1 = -\mathbf{n}_{12}$, $\mathbf{n}_2 = \mathbf{n}_{12}$, и из соотношения (1.14) получим

$$\mathbf{n}_{12}\cdot(\mathbf{B}^{(2)} - \mathbf{B}^{(1)}) = 0 \qquad (1.15)$$

то есть нормальная компонента вектора магнитной индукции остается непрерывной при переходе через границу раздела.

Подобным же образом можно рассмотреть электрическое смещение $\mathbf{D}$, но в этом случае при наличии зарядов появится дополнительный член. Тогда из уравнения (1.3) вместо соотношения (1.12) мы получим

$$\int div\, \mathbf{D} dV = \int \mathbf{D} \cdot \mathbf{n} dS = 4\pi \int \rho dV \tag{1.16}$$

При сближении площадок δA_1 и δA_2 полный заряд остается конечным и, следовательно, объемная плотность становится бесконечной. При этом вместо объемной плотности заряда ρ необходимо ввести поверхностную плотность заряда $\hat{\rho}$, определяемую соотношением:

$$\lim_{\delta h \to 0} \int \rho dV = \int \hat{\rho} dA \tag{1.17}$$

Позже нам понадобится также понятие поверхностной плотности тока $\hat{\mathbf{j}}$, которая определяется аналогичным образом, а именно

$$\lim_{\delta h \to 0} \int \mathbf{j} dV = \int \hat{\mathbf{j}} dA \tag{1.18}$$

Если площадку δA и высоту δh выбрать достаточно малыми, то из уравнения (1.16) получим $\mathbf{D}^{(1)} \cdot \mathbf{n}_1 \delta A_1 + \mathbf{D}^{(2)} \cdot \mathbf{n}_2 \delta A+$ вклад от стенок$=4\pi\rho\delta A$.

Вклад от стенок стремится к нулю с уменьшением δh, и поэтому в пределе при $\delta h \to 0$ получим

$$\mathbf{n}_{12} \cdot \left(\mathbf{D}^{(2)} - \mathbf{D}^{(1)} \right) = 4\pi\hat{\rho}, \tag{1.19}$$

то есть при наличии на поверхности раздела слоя с поверхностной плотностью заряда $\hat{\rho}$, нормальная компонента вектора электрического смещения при переходе через эту поверхность испытывает скачок, равный $4\pi\hat{\rho}$.

Исследуем поведение тангенциальных компонент. Заменим поверхность резкого раздела переходным слоем, а цилиндр, показанный на рис. 2, «прямоугольной» площадкой, стороны которой параллельны и перпендикулярны поверхности T (рис.3).

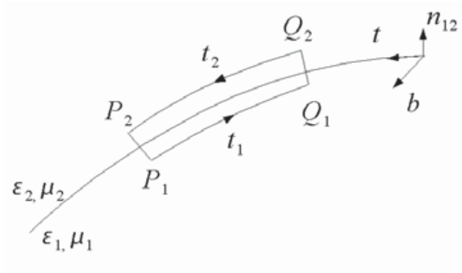

Рис. 3. К выводу граничных условий для тангенциальных компонент E и H

Пусть $\mathbf{b}$ — единичный вектор, перпендикулярный плоскости рассматриваемого прямоугольника. Тогда на основании теоремы Стокса получим из (1.2)

$$\int rot\,\mathbf{E}\cdot\mathbf{b}dS = \int \mathbf{E}\cdot dr = -\frac{1}{c}\int \mathbf{B}\cdot\mathbf{b}dS. \qquad (1.20)$$

Здесь первый и третий интегралы берутся по площади прямоугольника, а второй — вдоль его границ. Если длины $P_1Q_1\left(=\delta s_1\right)$ и $P_2Q_{21}\left(=\delta s_2\right)$ малы, то на каждой из этих сторон вектор E можно заменить постоянными значениями $\mathbf{E}^{(1)}$ и $\mathbf{E}^{(2)}$. Подобным же образом постоянным значением можно заменить и вектор $\mathbf{B}$. Тогда из (20) найдем

$$\mathbf{E}^{(1)}\cdot t_1\delta s_1 + \mathbf{E}^{(2)}\cdot t_2\delta s_2 + \text{вклад от концов} = -\frac{1}{c}\mathbf{B}\cdot\mathbf{b}\,\delta s\delta h, \qquad (1.21)$$

где δs — линейный элемент, по которому прямоугольник пересекается с поверхностью раздела. Если теперь постепенно уменьшать высоту прямоугольника, то вклад от концов P_1P_2 и Q_1Q_2 будет стремиться к нулю при условии, что $\mathbf{E}$ в пределе не имеет достаточно резких особенностей (такую возможность мы исключаем). Предположим также, что остается конечным и $\mathbf{B}$; тогда в пределе при $\delta h \to 0$ получим

$$\left(\mathbf{E}^{(1)} \cdot t_2 + \mathbf{E}^{(2)} \cdot t_2\right)\delta s = 0 . \tag{1.22}$$

Если $\mathbf{t}$ — единичный вектор касательной к поверхности (см. рис. 1.2), то $\mathbf{t}_1 = -\mathbf{t} = -\left[\mathbf{b}\times\mathbf{n}_{12}\right]$, $\mathbf{t}_2 = \mathbf{t} = \left[\mathbf{b}\times\mathbf{n}_{12}\right]$, и из (1.22) следует $\mathbf{b}\cdot\left[\mathbf{n}_{12}\times\left(\mathbf{E}^2 - \mathbf{E}^{(1)}\right)\right] = 0$.

Так как ориентация прямоугольника, а, следовательно, и единичного вектора $\mathbf{b}$ произвольна, ясно, что

$$\left[\mathbf{n}_{12}\times\left(\mathbf{E}^2 - \mathbf{E}^{(1)}\right)\right] = 0 , \tag{1.23}$$

то есть тангенциальная компонента электрического вектора непрерывна на поверхности раздела.

Наконец, рассмотрим поведение тангенциальной компоненты магнитного вектора. Анализ проводится аналогичным образом, но при наличии токов появится дополнительный член. Вместо условия (1.21) в этом случае получим

$$\begin{aligned} &\mathbf{H}^{(1)} \cdot t_1\delta s_1 + \mathbf{H}^{(2)} \cdot t_2\delta s_2 + \text{вклад от концов} \\ &\qquad = \frac{1}{c}\mathbf{D}\cdot\mathbf{b}\,\delta s\,\delta h + \frac{4\pi}{c}\hat{\mathbf{j}}\cdot\mathbf{b}\,\delta s \end{aligned} \tag{1.24}$$

Переходя, как и раньше, к пределу $\delta h \to 0$, находим

$$\left[\mathbf{n}_{12}\times\left(\mathbf{H}^2 - \mathbf{H}^1\right)\right] = \frac{4\pi}{c}\hat{\mathbf{j}} \tag{1.25}$$

Из условия (1.25) следует, что при наличии тока с поверхностной плотностью $\hat{\mathbf{j}}$ тангенциальная компонента магнитного вектора (рассматриваемая как вектор) испытывает скачок, равный $\frac{4\pi}{c}\left[\hat{\mathbf{j}}\times\mathbf{n}_{12}\right]$.

В рамках электромагнитной теории интенсивность света интерпретируется как плотность мощности электромагнитного поля. Из уравнений Максвелла (1.1) и (1.2) следует, что

$$\mathbf{E}\cdot rot\,\mathbf{H}-\mathbf{H}\cdot rot\,\mathbf{E}=\frac{4\pi}{c}j\cdot\mathbf{E}+\frac{1}{c}\mathbf{E}\cdot\mathbf{D}+\frac{1}{c}\mathbf{H}\cdot\mathbf{B} \qquad (1.26)$$

Члены, стоящие слева, с помощью хорошо известного векторного тождества можно выразить через дивергенцию векторного произведения $\mathbf{H}$ и $\mathbf{E}$, то есть

$$\mathbf{E}\cdot rot\,\mathbf{H}-\mathbf{H}\cdot rot\,\mathbf{E}=-div(\mathbf{E}\times\mathbf{H}) \qquad (1.27)$$

Из (1.26) и (1.27) получим

$$\frac{1}{c}(\mathbf{E}\cdot\mathbf{D}+\mathbf{H}\cdot\mathbf{B})+\frac{4\pi}{c}\mathbf{j}\cdot\mathbf{E}+div(\mathbf{E}\times\mathbf{H})=0 \qquad (1.28)$$

Умножив это равенство на $c/4\pi$, проинтегрировав по произвольному объему и использовав теорему Гаусса, найдем

$$\frac{1}{4\pi}\int(\mathbf{E}\times\mathbf{D}+\mathbf{H}\times\mathbf{B})dV+\int\mathbf{j}\cdot\mathbf{E}dV+\frac{c}{4\pi}\int(\mathbf{E}\times\mathbf{H})\cdot\mathbf{n}dS=0 \qquad (1.29)$$

Здесь последний интеграл берется по границе объема, а $\mathbf{n}$ — единичный вектор внешней нормали.

Соотношение (1.29) непосредственно вытекает из уравнения Максвелла и поэтому выполняется независимо от справедливости материальных уравнений (1.9)—(1.11). Оно выражает закон сохранения энергии для электромагнитного поля. Рассмотрим его для случая, когда удовлетворяются материальные уравнения (1.9)—(1.11). Позже будет проведено обобщение этого закона для случая анизотропных сред, где материальные уравнения принимают более сложную форму.

Используя материальные уравнения, найдем

$$\left.\begin{aligned}\frac{1}{4\pi}(\mathbf{E}\times\mathbf{D})=\frac{1}{4\pi}\mathbf{E}\cdot\frac{\partial}{\partial t}(\varepsilon\mathbf{E})=\frac{1}{8\pi}\frac{\partial}{\partial t}(\varepsilon\mathbf{E}^2)=\frac{1}{8\pi}\frac{\partial}{\partial t}(\mathbf{E}\times\mathbf{D}),\\ \frac{1}{4\pi}(\mathbf{H}\times\mathbf{B})=\frac{1}{4\pi}\mathbf{H}\cdot\frac{\partial}{\partial t}(\mu\mathbf{H})=\frac{1}{8\pi}\frac{\partial}{\partial t}(\mu\mathbf{H}^2)=\frac{1}{8\pi}\frac{\partial}{\partial t}(\mathbf{H}\times\mathbf{B}).\end{aligned}\right\}. \qquad (1.30)$$

Полагая

$$\omega_e=\frac{1}{8\pi}\mathbf{E}\times\mathbf{D},...,\omega_m=\frac{1}{8\pi}\mathbf{H}\times\mathbf{B} \qquad (1.31)$$

и

$$W = \int \left(\omega_e + \omega_m \right) dV\,, \tag{1.32}$$

преобразуем соотношение (1.29) к виду

$$\frac{dW}{dt} + \int j \cdot \mathbf{E} dV + \frac{c}{4\pi} \int \left(\mathbf{E} \times \mathbf{H} \right) \cdot \mathbf{n} dS = 0 \tag{1.33}$$

Величина W представляет полную энергию, заключенную внутри объема, и, следовательно, ω_e можно отождествить с плотностью электрической энергии, а ω_m — с плотностью магнитной энергии поля. Вектор

$$\mathbf{S} = \frac{c}{4\pi} \left[\mathbf{E} \times \mathbf{H} \right] \tag{1.34}$$

называется вектором Умова-Пойнтинга и описывает плотность потока энергии электромагнитного поля.

Замечание. В общем случае плотности электрической и магнитной энергии определяются выражениями $\omega_e = \frac{1}{4\pi} \int \mathbf{E} \cdot d\mathbf{D}$, …, $\omega_m = \frac{1}{4\pi} \int \mathbf{H} \cdot d\mathbf{B}$. Они переходят в (1.31), когда связь между $\mathbf{E}$ и $\mathbf{D}$, а также между $\mathbf{H}$ и $\mathbf{B}$ линейна, как предполагается здесь.

Тема 2. Электромагнитные волны.

2.1. Волновое уравнение.

Уравнения Максвелла связывают между собой векторы поля системой дифференциальных уравнений. Дифференциальные уравнения, которым должен удовлетворять каждый из векторов в отдельности, можно получить путем исключения остальных векторов. Мы ограничимся рассмотрением области поля, не содержащей ни зарядов, ни токов.

Подставим выражение для **B** из материального уравнения (1.11) во второе уравнение Максвелла (1.2), разделим обе его части на μ и применим операцию rot. Это дает

$$rot\left(\frac{1}{\mu} rot\,\mathbf{E}\right) + \frac{1}{c} rot\,\dot{\mathbf{H}} = 0\,. \tag{2.1}$$

Продифференцируем затем первое уравнение Максвелла (1.1) по времени, используем уравнение для **D** (1.10) и исключим $rot\,\mathbf{H}$ из системы двух уравнений, содержащей получающееся уравнение и уравнение (2.1). Тогда получим

$$rot\left(\frac{1}{\mu} rot\,\mathbf{E}\right) + \frac{\varepsilon}{c^2}\ddot{\mathbf{E}} = 0\,. \tag{2.2}$$

Если использовать тождества $rot\,uv \equiv u\,rot\,v + \left(grad\,u\right)\times v$ и $rot\,rot \equiv grad\,div - V^2$, то (2.2) примет вид

$$\nabla^2\mathbf{E} - \frac{\varepsilon\mu}{c^2}\dot{\mathbf{E}} + \left[grad\left(\ln\mu\right)\right]\times rot\,\mathbf{E} - grad\,div\mathbf{E} = 0 \tag{2.3}$$

Используя снова материальное уравнение для **D** и применяя тождество $div\,uv \equiv u\,div\,v + v\,grad\,u$, найдем из (1.3) соотношение

$$\varepsilon\,div\,\mathbf{E} + \mathbf{E}\,grad\,\varepsilon = 0\,. \tag{2.4}$$

Следовательно, уравнение (2.3) можно записать в виде

$$\nabla^2\mathbf{E}-\frac{\varepsilon\mu}{c^2}\ddot{\mathbf{E}}+\left[grad\left(\ln\mu\right)\right]\times rot\,\mathbf{E}+grad\left[\mathbf{E}\cdot grad\left(\ln\varepsilon\right)\right]=0 \qquad (2.5)$$

Подобным же образом получается уравнение для $\mathbf{H}$

$$\nabla^2\mathbf{H}-\frac{\varepsilon\mu}{c^2}\ddot{\mathbf{H}}+\left[grad\left(\ln\varepsilon\right)\right]\times rot\,\mathbf{H}+grad\left[\mathbf{H}\cdot grad\left(\ln\mu\right)\right]=0 \qquad (2.6)$$

В частности, если среда однородна, то $grad\left(\ln\varepsilon\right)=grad\left(\ln\mu\right)=0$, и соотношения (5) и (6) принимают вид

$$\nabla^2\mathbf{E}-\frac{\varepsilon\mu}{c^2}\ddot{\mathbf{E}}=0,\quad \nabla^2\mathbf{H}-\frac{\varepsilon\mu}{c^2}\ddot{\mathbf{H}}=0 \qquad (2.7)$$

Это обычные уравнения волнового движения. Они означают, что существуют электромагнитные волны, распространяющиеся со скоростью

$$v=\frac{c}{\sqrt{\varepsilon\mu}}. \qquad (2.8)$$

В любой волновой теории света элементарным процессом считают гармоническую волну в пространстве — времени. Если ее частота лежит в интервале от $4\cdot10^{14}сек^{-1}$ до $7{,}5\cdot10^{14}сек^{-1}$ (приблизительно), то она вызывает у человека физиологическое ощущение определенного цвета. Действительная связь между цветом и частотой сложна и будет затронута в пятом учебном пособии магистерской программы «Оптика наноструктур».

Перед тем, как рассматривать решения векторных волновых уравнений (2.7), опишем несколько характерных решений скалярного волнового уравнения. Эти понятия и величины имеют не только отвлеченный математический смысл, но и могут быть использованы в качестве простейших моделей в оптике.

2.2. Скалярные волны

В однородной среде в областях, свободных от токов и зарядов, каждая из декартовых компонент $V(\mathbf{r},t)$ векторов поля удовлетворяет, согласно (2.7), однородному волновому уравнению

$$\nabla^2 V - \frac{1}{\upsilon^2}\frac{\partial^2 V}{\partial t^2} = 0 \qquad (2.9)$$

Рассмотрим простейшие решения этого уравнения.

Плоские волны.

Пусть $\mathbf{r}(x,y,z)$ — радиус-вектор точки P, а $\mathbf{s}(s_x,s_y,s_z)$ — единичный вектор с фиксированным направлением. Говорят, что любое решение уравнения (2.9) вида

$$V = V(\mathbf{r}\cdot\mathbf{s},t) \qquad (2.10)$$

представляет собой плоскую волну, так как в каждый момент времени величина V постоянна в плоскостях $\mathbf{r}\cdot\mathbf{s} = const$, которые перпендикулярны к единичному вектору $\mathbf{s}$.

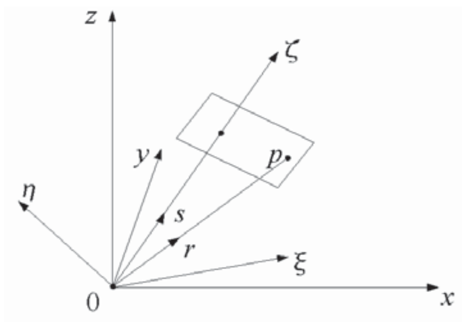

Рис. 4. Распространение плоской волны.

Удобно выбрать новое положение декартовых осей $O\xi$, $O\eta$, $O\zeta$ так, чтобы ось $O\varsigma$ была направлена по $\mathbf{s}$. Тогда (смотри рисунок 4)

$\mathbf{r} \cdot \mathbf{s} = \zeta$ и $\dfrac{\partial}{\partial x} = s_x \dfrac{\partial}{\partial \zeta}$, $\quad \dfrac{\partial}{\partial y} = s_y \dfrac{\partial}{\partial \zeta}$, $\quad \dfrac{\partial}{\partial z} = s_z \dfrac{\partial}{\partial \zeta}$.

Отсюда легко получить, что $\nabla^2 V = \dfrac{\partial^2 V}{\partial \zeta^2}$ и, следовательно, волновое уравнение (2.9) запишется в виде

$$\frac{\partial^2 V}{\partial \zeta^2} - \frac{1}{\upsilon^2} \frac{\partial^2 V}{\partial t^2} = 0 \tag{2.11}$$

Если положить

$\zeta - \upsilon t = p, \quad \zeta + \upsilon t = q,$, то (2.11) примет вид $\dfrac{\partial^2 V}{\partial p \partial q} = 0$.

Общим решением этого уравнения служит

$V = V_1(p) + V_2(q) = V_1(r \cdot s - \upsilon t) + V_2(r \cdot s + \upsilon t)$,

где V_1 и V_2 — произвольные функции.

Мы видим, что аргумент функции V_1 не изменяется при замене аргументов (ζ, t) на $(\zeta + \upsilon\tau, t + \tau)$, где τ произвольно. Следовательно, V_1 представляет возмущение, которое распространяется со скоростью υ в положительном направлении оси ζ. Аналогично $V_2(\zeta + \upsilon t)$ — это возмущение, распространяющееся со скоростью υ в отрицательном направлении оси ζ.

Сферические волны.

Теперь рассмотрим решения, зависящие от расстояния между точкой наблюдения и источником волны, т. е. решения вида $V = V(r, t)$, где $r = |r| = \sqrt{x^2 + y^2 + z^2}$. Используя соотношения $\dfrac{\partial}{\partial x} = \dfrac{\partial r}{\partial x}\dfrac{\partial}{\partial r} = \dfrac{x}{r}\dfrac{\partial}{\partial r}$ и так далее, непосредственным расчетом найдем

$\nabla^2 V = \dfrac{1}{r}\dfrac{\partial^2}{\partial r^2}(rV)$

и, следовательно, волновое уравнение (2.1) примет вид

$$\frac{\partial^2}{\partial r^2}(rV) - \frac{1}{\upsilon^2}\frac{\partial^2}{\partial t^2}(rV) = 0 \qquad (2.12)$$

Общим решением этого уравнения служит

$$V = \frac{V_1(r - \upsilon t)}{r} + \frac{V_2(r + \upsilon t)}{r} \qquad (2.13)$$

где V_1 и V_2 — по-прежнему произвольные функции. В правой части равенства (12) первый член представляет сферическую волну, расходящуюся от начала координат, второй — сферическую волну, сходящуюся к началу координат, причем скорость распространения обоих волн равна υ.

Гармонические волны.

В точке r_0 пространства возмущение, вызванное волной, зависит только от времени, то есть

$$V(r_0, t) = F(t) \qquad (2.14)$$

Особый интерес представляет периодическая функция F, поэтому рассмотрим случай, когда F имеет вид:

$$F(t) - a\cos(\omega t + \delta) \qquad (2.15)$$

Величина $a(>0)$ называется амплитудой, а аргумент косинуса $(\omega t + \delta)$ — фазой. Величина

$$\nu = \frac{\omega}{2\pi} = \frac{1}{T} \qquad (2.16)$$

называется частотой и представляет число колебаний в секунду. Величина ω называется угловой (или циклической) частотой и дает число колебаний в 2π секунд. При замене t на $t + T$ значение функции F остается неизменным, поэтому T является периодом колебаний. Волновые функции (т. е. решения волнового уравнения) в форме (2.15) называют гармоническими относительно времени.

Рассмотрим вначале волновую функцию, которая представляет гармоническую плоскую волну, распространяющуюся в направлении, заданном единичным вектором s. Она получается при замене в формуле (2.15) t на $t-(r\cdot s)/\upsilon$, т. е.

$$V(r,t)=a\cos\left[\omega\left(t-\frac{rs}{\upsilon}\right)+\delta\right] \tag{2.17}$$

Уравнение (2.17) не изменится, если $r\cdot s$ заменить на $s\cdot r+\lambda$, где

$$\lambda=\upsilon\frac{2\pi}{\omega}=\upsilon T \tag{2.18}$$

Величина λ называется длиной волны. Полезно также определить длину волны в вакууме λ_0

$$\lambda_0=cT=n\lambda \tag{2.19}$$

где $n=\sqrt{\varepsilon\mu}$ - так называемый показатель преломления среды, λ_0- это длина волны, распространяющейся в вакууме с той же частотой.

Удобно также ввести векторы $\mathbf{k}_0$ и $\mathbf{k}$, направления которых совпадают с направлением распространения $\mathbf{s}$, а длины соответственно равны

$$k_0=2\pi/\lambda_0=\omega/c \tag{2.20}$$

и

$$k=nk_0=2\pi/\lambda=n\omega/c=\omega/\upsilon \tag{2.21}$$

Вектор $\mathbf{k}=k\mathbf{s}$ называется волновым вектором или вектором распространения в среде, а $\mathbf{k}_0=k_0\mathbf{s}$ — соответствующим вектором для вакуума.

Фазовая скорость.

Рассмотрим теперь гармонические волны более сложной формы. В общем случае вещественную гармоническую скалярную волну с частотой ω можно определить как вещественное решение волнового уравнения вида

$$\frac{ds}{dt} = \frac{\omega}{\mathbf{q}\, grad\ g} \tag{2.22}$$

где $a(>0)$ и g — вещественные скалярные функции положения. Поверхность

$$g(r) = const \tag{2.23}$$

называют поверхностью постоянной фазы, или фазовым фронтом. В отличие от предыдущего случая, поверхности постоянной амплитуды волны (2.22), вообще говоря, не совпадают с поверхностями постоянной фазы. Говорят, что такая волна неоднородна.

Расчеты, связанные с гармоническими волнами, упрощаются, если использовать экспоненциальные функции вместо тригонометрических.

Уравнение (2.22) можно записать в виде

$$V(r,t) = \mathrm{Re}\{U(r)\exp[-i\omega t]\} \tag{2.24}$$

где

$$U(r) = a(r)\exp[ig(r)], \tag{2.25}$$

а символ Re означает, что берется вещественная часть. Подставляя (2.25) в волновое уравнение (2.1), мы найдем, что U должно удовлетворять уравнению

$$\nabla^2 U + n^2 k_o^2 U = 0 \tag{2.26}$$

Величину U называют комплексной амплитудой волны. В частности, для плоской волны имеем

$$g(r) = \omega\left(\frac{rs}{\upsilon}\right) - \delta = k(r \cdot s) - \delta = \kappa \cdot r - \delta \tag{2.27}$$

Замечание. Если операции, производимые над V, линейны, то в выражении (2.24) можно опустить символ Re и оперировать прямо с комплексной функцией. При этом вещественная часть окончательного выражения будет представлять изучаемою физическую величину. Однако если приходится иметь дело с нелинейными операциями, такими, как возведение в квадрат и т. д. (например, при расчетах плотности электрической или магнитной энергии), то, вообще говоря, необходимо взять действительные части и оперировать только с ними.

В отличие от плоской гармонической волны, волна более общего вида (2.24) не периодична в пространстве. Однако фаза $\omega t - g(r)$` одинакова для (r,t) и $(r+dt, t+dt)$ при условии, что

$$\omega dt - (grad\ g)\cdot dr = 0 \tag{2.28}$$

Обозначив через $\mathbf{q}$ единичный вектор в направлении dr и написав $dr = \mathbf{q}ds$, найдем отсюда

$$\frac{ds}{dt} = \frac{\omega}{\mathbf{q}\ grad\ g} \tag{2.29}$$

Эта величина минимальна, когда вектор $\mathbf{q}$ перпендикулярен к поверхности постоянной фазы, т. е. когда $\mathbf{q} = (grad\ g)/|grad\ g|$, причем тогда значение (2.29) будет равно

$$\upsilon^{(p)}(r) = \frac{\omega}{|grad\ g|} \tag{2.30}$$

Величина $\upsilon^{(p)}(r)$ называется фазовой скоростью и равна скорости, с которой распространяется каждая поверхность постоянной фазы. Для плоской электромагнитной волны из (2.27) найдем $grad\ g = \mathbf{k}$ и, учитывая (2.20), получим $\upsilon^{(p)} = \omega / k = c/\sqrt{\varepsilon\mu}$. Для волн более сложной формы, фазовая скорость $\upsilon^{(p)}$ в общем случае отличается от $c/\sqrt{\varepsilon\mu}$ и меняется от точки к точке даже в однородной среде.

Замечание. При достаточно большой частоте фазовая скорость приблизительно равна отношению $c/\sqrt{\varepsilon\mu}$ даже для волн, у которых поверхности постоянной фазы не являются плоскими.

Замечание. Необходимо отметить, что выражение (2.29) для ds/dt не является компонентой фазовой скорости в направлении q. С другой стороны, величина, обратная ей, т. е. величина

$$\frac{dt}{ds}=\frac{q\,grad\,g}{\omega} \tag{2.31}$$

как видно из этого выражения, есть компонента вектора $(grad\,g)/\omega$ в направлении q.

2.3. Волновые пакеты. Групповая скорость.

Монохроматические волны, рассматриваемые в предыдущем разделе,— это идеализация, никогда строго не реализующаяся на практике. Из теоремы Фурье следует, что волну $V(r,t)$ (если она удовлетворяет определенным, очень общим условиям [2]) можно рассматривать как суперпозицию монохроматических волн разных частот, а именно

$$V(r,t)=\int_0^\infty a_\omega(r)\cos\left[\omega t-g_\omega(r)\right]d\omega \tag{2.32}$$

Здесь снова удобно воспользоваться комплексным представлением, в котором V отождествляется с вещественной частью соответствующей комплексной волны

$$V(r,t)=\mathrm{Re}\int_0^\infty a_\omega(r)\exp\left\{-i\left[\omega t-g_\omega(r)\right]\right\}d\omega \tag{2.33}$$

Если Фурье-амплитуды a_ω заметно отличаются от нуля лишь внутри узкого интервала $\bar{\omega}-\frac{1}{2}\Delta\omega\le\omega\le\bar{\omega}+\frac{1}{2}\Delta\omega$, $\quad(\Delta\omega/\bar{\omega})<<1$ вблизи средней частоты $\bar{\omega}$, то волну можно назвать «почти монохроматической». В этом случае обычно говорят о волновой группе, или волновом пакете.

Замечание. Строго говоря, для того чтобы функция V обладала свойствами, которые обычно приписывают волновому пакету, необходимо также предположить, что во всем эффективном интервале частот фазовую функцию g_ω можно аппроксимировать линейной функцией ω.

Для иллюстрации некоторых основных свойств волнового пакета рассмотрим вначале волну, которая распространяется вдоль оси $\left.\begin{aligned}\frac{E_x}{a_1}\sin\delta_2 - \frac{E_y}{a_2}\sin\delta_1 = \cos\tau\sin(\delta_2-\delta_1),\\ \frac{E_x}{a_1}\cos\delta_2 - \frac{E_y}{a_2}\cos\delta_1 = \sin\tau\sin(\delta_2-\delta_1).\end{aligned}\right\}$ и образована в результате суперпозиции двух плоских монохроматических волн с одинаковыми амплитудами и слегка различными частотами и волновыми числами.

$$V(z,t) = a\exp\left[-i(\omega t - kz)\right] + a\exp\left\{-i\left[(\omega+\delta\omega)t - (k+\delta k)z\right]\right\} \quad (2.34)$$

В соответствии с принятым ранее соглашением символ Re здесь опущен. Уравнение (2.34) можно записать в форме

$$\begin{aligned}V(z,t) - a\left\{\exp\left[\frac{1}{2}i(t\delta\omega - z\delta k)\right] + \exp\left[-\frac{1}{2}i(t\delta\omega - z\delta k)\right]\right\}\\ \times\exp\left[-i(\bar{\omega}t - \bar{k}z)\right] = 2a\cos\left[\frac{1}{2}(t\delta\omega - z\delta k)\right]\exp\left[-i(\bar{\omega}t - \bar{k}z)\right]\end{aligned} \quad (2.35)$$

где

$$\bar{\omega} = \omega + \frac{1}{2}\delta\omega, \quad \bar{k} = k + \frac{1}{2}\delta k \quad (2.36)$$

— средняя частота и среднее волновое число соответственно. Можно считать что выражение (2.35) описывает плоскую волну с частотой $\bar{\omega}$ и длиной волны $2\pi / \bar{k}$, распространяющуюся в направлении оси z. Однако амплитуда этой волны не постоянна, а изменяется во времени и пространстве от нуля до значения $2a$. (рис. 5), что вызывает хорошо известное явление биений.

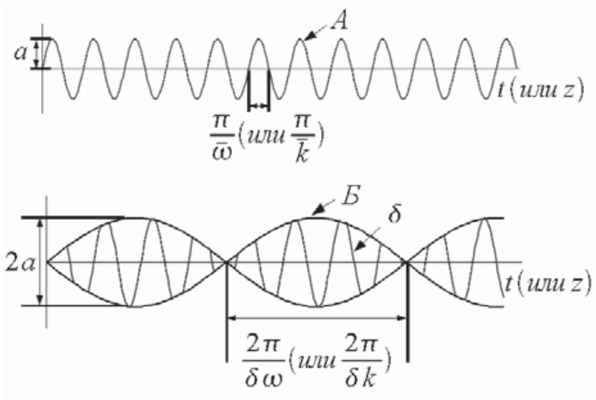

Рис.5. A **- волна** $a\cos(\bar{\omega}t-\bar{k}z)$**;** $Б$ **- волна** $2a\cos\left[\frac{1}{2}(t\delta\omega-z\delta k)\right]$**;** $В$ **- волновая группа** $2a\cos\left[\frac{1}{2}(t\delta\omega-z\delta k)\right]\times\cos(\bar{\omega}t-\bar{k}z)$**.**

На рис. 5 абсцисса представляет одну из двух независимых переменных (t или z), тогда как другая сохраняется постоянной.

Соседние максимумы функции, описывающей амплитуду, находятся на расстояниях

$$\left.\begin{array}{l}\delta t=\dfrac{4\pi}{\delta\omega}\quad \text{при фиксированном } z\\ \text{или}\\ \delta z=\dfrac{4\pi}{\delta k}\quad \text{при фиксированном } t,\end{array}\right\}\tag{2.37}$$

а максимумы функции, связанной с фазой,— на расстояниях

$$\left.\begin{array}{l}\delta t=\dfrac{2\pi}{\delta\bar{\omega}}\quad \text{при фиксированном } z\\ \text{или}\\ \delta z=\dfrac{2\pi}{\delta\bar{k}}\quad \text{при фиксированном } t.\end{array}\right\}\tag{2.38}$$

Следовательно, поскольку считается, что $\delta\omega / \bar{\omega}$ и $\delta k / \bar{k}$ малы по сравнению с единицей, амплитуда будет меняться медленно по сравнению с изменением другого члена.

Из (2.35) вытекает, что плоскости постоянной амплитуды и, в частности, максимумы амплитуды распространяются со скоростью

$$\upsilon^{(g)} = \frac{\delta\omega}{\delta k}, \tag{2.39}$$

тогда как плоскости постоянной фазы распространяются со скоростью

$$\upsilon^{(p)} = \bar{\omega} / \bar{k}\ . \tag{2.40}$$

Величина $\upsilon^{(g)}$ называется групповой скоростью волны. Поскольку V удовлетворяет волновому уравнению, частота ω и волновое число k связаны (см. (2.21)) друг с другом (в среде с показателем преломления n) соотношением

$$k = n(\omega)\omega / c\,; \tag{2.41}$$

здесь показатель преломления n зависит от ω. Уравнение (41) выражает дисперсию волны. В недиспергирующей изотропной среде n не зависит от ω; в такой среде и фазовая скорость $\upsilon^{(p)}$, и групповая скорость $\upsilon^{(g)}$ равны c / n. Однако в диспергирующей среде эти две скорости в общем случае различны.

Так как, но предположению, $\delta\omega$ мало, $\delta\omega / \delta k$ можно заменить дифференциальным соотношением $\delta\omega / \delta k$ тогда выражение для групповой скорости запишется в виде

$$\upsilon^{(g)} = \frac{d\omega}{dk}. \tag{2.42}$$

Мы покажем, что фактически это соотношение выполняется при более общих условиях. Рассмотрим одномерную волновую группу.

$$V(z,t) = \int\limits_{(\Delta\omega)} a_{\omega} \exp[-i(\omega t - kz)]\, d\omega, \tag{2.43}$$

где $\Delta\omega$ означает небольшой интервал вблизи средней частоты $\overline{\omega}\,(\Delta\omega/\overline{\omega} \quad 1)$, в котором a_ω заметно отличается от нуля. Пусть $\overline{k} = n(\overline{\omega})\overline{\omega}/c$ - соответствующее волновое число. Тогда последнее соотношение можно переписать в форме

$$V(z,t) = A(z,t)\exp[-i(\overline{\omega}t - \overline{k}z)], \tag{2.44}$$

где

$$A(z,t) = \int\limits_{(\Delta\omega)} a_\omega \exp\{-i[(\omega-\overline{\omega})t - (k-\overline{k})z]\}\,d\omega$$

$$\int\limits_{(\Delta\omega)} a_\omega \exp\left\{-i\left[(\omega-\overline{\omega})\left\{t-\left(\frac{dk}{d\omega}\right)_{\overline{\omega}} z\right\}\right]\right\}d\omega, \tag{2.45}$$

если $\Delta\omega$ достаточно мало. Снова V можно интерпретировать как плоскую волну с переменной амплитудой, частотой $\overline{\omega}$ и волновым числом $\overline{k}$, распространяющуюся в направлении z. Амплитуда $A(z,t)$ представляет суперпозицию гармонических волн с частотами $\omega-\overline{\omega}$. Так как $\Delta\omega/\overline{\omega}$ мало по сравнению с единицей, $A(z,t)$ медленно меняется по сравнению с изменением второго члена. В общем случае $A(z,t)$ комплексно и дает вклад $\arg(A)$ в фазу $(\overline{\omega}t - \overline{k}z)$. Как мы видим, поверхности

$$t = \left(\frac{dk}{d\omega}\right)_{\overline{\omega}} z, \tag{2.46}$$

играют особую роль: на каждой такой поверхности $A(z,t)$ постоянно. Следовательно, скорость перемещения какого-либо значения A, а также максимума $|A|$ определяется, как и раньше, *групповой скоростью*

$$\upsilon^{(g)} = \left(\frac{d\omega}{dk}\right)_{\overline{k}}. \tag{2.47}$$

Легко показать справедливость следующих соотношений между групповой и фазовой скоростями:

$$\upsilon^{(g)} = \frac{d}{dk}(v^{(p)}k) = v^{(p)} + k\frac{dv^{(p)}}{dk} = v^{(p)} - \lambda\frac{dv^{(p)}}{d\lambda}, \tag{2.48}$$

причем все величины здесь относятся к средней частоте $\overline{\omega}$.

Наконец, рассмотрим трехмерный волновой пакет общего вида

$$V(r,t) = \mathrm{Re}\int_{\Delta\omega} a_\omega(r)\exp\left\{-i\left[\omega t - g_\omega(r)\right]\right\}d\omega. \tag{2.49}$$

По аналогии с (2.43) выделим член, соответствующий средней частоте $\overline{\omega}$, и для достаточно малых $\Delta\omega$ напишем

$$V(r,t) = A(r,t)\exp\left\{-i\left[\overline{\omega}t - g_{\overline{\omega}}(r)\right]\right\}, \tag{2.50}$$

где

$$\begin{gathered} A(r,t) = \int_{(\Delta\omega)} a_\omega(r)\exp\left(-i\left\{(\omega-\overline{\omega})t - \left[g_\omega(r) - g_{\overline{\omega}}(r)\right]\right\}\right) \\ \int_{(\Delta\omega)} a_\omega(r)\exp\left(-i\left\{(\omega-\overline{\omega})\left[t - \left(\frac{\partial g(r)}{\partial\omega}\right)_{\overline{\omega}}\right]\right\}\right)d\omega. \end{gathered} \tag{2.51}$$

Выражение (2.50) представляет волну с частотой $\overline{\omega}$, амплитуда которой $A(r,t)$ (обычно комплексная) меняется и в пространстве, и во времени, причем ее изменение так же медленно по сравнению с изменением второго члена. Можно ожидать, что поверхность

$$t = \left[\frac{\partial g(r)}{\partial\omega}\right]_{\overline{\omega}} \tag{2.52}$$

будет играть особую роль. Однако теперь амплитудная функция A не обязательно постоянна на каждой такой поверхности, так как здесь Фурье-амплитуды a_ω зависят не только от частоты, но и от положения. Значение поверхности, описываемой (2.52), станет ясным, если мы будем рассматривать абсолютную величину амплитуды $M = |A|$. Тогда

$$\begin{aligned} M^2(r,t) &= A(r,t)\cdot A^*(r,t) = \\ &= \int_{(\Delta\omega)}\int_{(\Delta\omega')} a_\omega(r)\, a_\omega(r)\exp\left\{-i(\omega-\omega')\left[t - \left(\frac{\partial g(r)}{\partial\omega}\right)_{\overline{\omega}}\right]\right\}d\omega d\omega'. \end{aligned} \tag{2.53}$$

Очевидно, что мнимая часть двойного интеграла равна нулю, так как величина M^2 вещественна. (Формально в этом легко убедиться, если поменять местами независимые переменные ω и ω' и заметить, что при этом мнимая часть подынтегрального выражения меняет знак.) Следовательно,

$$M^2(r,t) = \int\limits_{(\Delta\omega)}\int\limits_{(\Delta\omega')} a_\omega(r)\,a_\omega(r)\cos\left\{(\omega-\omega')\left[t-\left(\frac{\partial g(r)}{\partial\omega}\right)_{\bar{\omega}}\right]\right\}d\omega d\omega'. \quad (2.54)$$

Рассматривая какую-нибудь определенную точку $r = r_0$ и вспоминая, что a_ω либо положительно, либо равно нулю, мы видим, что $M^2(r_0,t)$ достигает максимального значения, когда аргумент косинуса равен нулю, т. е. когда $t = \left(\frac{\partial g(r)}{\partial\omega}\right)_{\bar{\omega}}$. Таким образом, соотношение (2.52) представляет поверхности, на которых в момент времени t абсолютное значение амплитуды максимально в указанном выше смысле. Поэтому в общем случае разумно определить групповую скорость трехмерной волновой группы как скорость, с которой перемещаются эти поверхности. Рассмотрим малое смешение $dr = ds$, где q — единичный вектор в направлении нормали к поверхности. Согласно (2.52) соответствующее изменение δt определяется выражением

$$\delta t = \delta s\left|grad\left(\frac{\partial g(r)}{\partial\omega}\right)_{\bar{\omega}}\right|, \quad (2.55)$$

и, следовательно, в общем случае групповая скорость трехмерной группы равна

$$\upsilon^{(g)} = \frac{1}{\left|grad\left(\frac{\partial g}{\partial\omega}\right)_{\bar{\omega}}\right|}. \quad (2.56)$$

Это выражение нужно сравнить с выражением для фазовой скорости гармонической волны общего вида (2.32), т. е. с выражением

$$\upsilon^{(p)} = \frac{1}{\left| grad \frac{g}{\omega} \right|}. \qquad (2.57)$$

В частном случае распространения группы плоских волн в направлении z имеем $g_\omega = kz$, и (2.56) переходит в (2.47).

Замечание. Эффективный интервал частот $\Delta\omega$ представляет собой важный параметр, относящийся к волновой группе; по существу эта величина определяет скорость изменения амплитуды и фазы. Если дисперсия среды невелика, волновая группа проходит значительное расстояние без заметного «размывания». При таких обстоятельствах групповая скорость, которую можно считать скоростью распространения группы как целого, является также скоростью распространения энергии. Однако в общем случае это неверно. В частности, в области аномальной дисперсии групповая скорость может превысить скорость света или стать отрицательной.

2.4. Векторные волны

В задачах взаимодействия электромагнитных волн с материальным объектом, характерные размеры которого меньше длины волны, теории скалярных волн недостаточно для его описания.

Электромагнитная плоская волна общего вида.

Простейшим электромагнитным полем является поле плоской волны. В этом случае каждая из компонент векторов поля, а, следовательно, и векторы E и H зависят лишь от переменной $u = r \cdot s - \upsilon t$, т. е.

$$E = E(r \cdot s - \upsilon t), \qquad H = H(r \cdot s - \upsilon t) \qquad (2.58)$$

здесь s, как и раньше, единичный вектор в направлении распространения.

Обозначим точкой дифференцирование по t и штрихом дифференцирование по переменной u. Тогда

$$\dot{E} = -\upsilon E', \quad (rotE)_x = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = E'_z s_y - E'_y s_z = (s \times E')_x \tag{2.59}$$

Подставляя эти выражения в уравнения Максвелла (1.1) и (1.2) при условии $j = 0$ и, используя материальные уравнения (1.10) и (1.11), получим

$$s \times H' + \frac{\varepsilon \upsilon}{c} E' = 0, \quad s \times E' - \frac{\mu \upsilon}{c} H' = 0 \tag{2.60}$$

После интегрирования соотношения (2.60) (считаем при этом, что $\upsilon / c = 1/\sqrt{\varepsilon\mu}$ и что постоянная интегрирования равна нулю, т. е. учитываем граничные условия на бесконечности) получим

$$E = -\sqrt{\frac{\mu}{\varepsilon}}(s \times H), \quad H = \sqrt{\frac{\varepsilon}{\mu}}(s \times E). \tag{2.61}$$

Скалярное умножение на $\mathbf{s}$ дает

$$\mathbf{E} \cdot \mathbf{s} = \mathbf{H} \cdot \mathbf{s} = 0 \tag{2.62}$$

Это соотношение выражает «поперечность» плоских электромагнитных волн, т. е. оно показывает что электрический и магнитный векторы лежат в плоскостях, перпендикулярных к направлению распространения. Из соотношений (2.61) и (2.62) вытекает, что E, H и s образуют правую ортогональную систему векторов. Из равенств (2.61) следует также, что

$$\sqrt{\mu} H = \sqrt{\varepsilon} E, \tag{2.63}$$

где $E = |\mathbf{E}|$, $H = |\mathbf{H}|$.

Рассмотрим количество энергии, которое протекает в единицу времени через элемент площади, перпендикулярный направлению распространения. Вообразим цилиндр, ось которого параллельна s, а площадь поперечного сечения равна единице. Тогда количество энергии, которое протекает через основание цилиндра в единицу времени, равно энергии, содержащейся в части цилиндра длиной υ. Следовательно, поток энергии равен $\upsilon\omega$, где ω — плотность энергии. Согласно (2.63), а также (1.31) плотность энергии определяется выражением

$$\omega = \frac{\varepsilon}{4\pi} E^2 = \frac{\mu}{4\pi} H^2. \tag{2.64}$$

С другой стороны, вектор Пойнтинга, в соответствии с (1.34), равен

$$S = \frac{c}{4\pi} EH\mathbf{s} = \frac{c}{4\pi}\sqrt{\frac{\varepsilon}{\mu}} E^2\mathbf{s} = \frac{c}{4\pi}\sqrt{\frac{\varepsilon}{\mu}} H^2\mathbf{s}. \tag{2.65}$$

Мы видим, что, в согласии с пунктом 1.2, вектор Пойнтинга представляет собой поток энергии и по величине, и по направлению распространения.

Гармоническая электромагнитная плоская волна.

Особый интерес представляет случай плоской волны гармонической во времени, т. е. случай, когда каждая из декартовых компонент векторов **E** и **H** имеет вид

$$a\cos(\tau + \delta) = \operatorname{Re}\left\{a\exp\left[-i(\tau + \delta)\right]\right\} \qquad (a > 0). \tag{2.66}$$

Здесь τ обозначает переменную часть фазового множителя, т. е.

$$\tau = \omega\left(t - \frac{r \cdot s}{\upsilon}\right) = \omega t - \kappa \cdot r. \tag{2.67}$$

Выберем ось z в направлении s. Тогда отличными от нуля будут лишь x-и y-компоненты E и H, поскольку, в соответствии с (2.62), поле поперечно.

Рассмотрим характер кривой, которую описывает вектор напряженности электрического поля в плоскости волнового фронта. Эта кривая является геометрическим местом точек, координаты которых (E_x, E_y) равны

$$E_x = a_1\cos(\tau + \delta_1), \qquad E_y = a_2\cos(\tau + \delta_2), \qquad E_z = 0. \tag{2.68}$$

2.5. Эллиптическая поляризация.

Для того чтобы исключить τ из первых двух уравнений (2.68), перепишем их в виде

$$E_x / a \; = \cos\tau\cos\delta_1 - \sin\delta_1, \qquad E_y / a_2 = \cos\tau\cos\delta_2 - \sin\tau\sin\delta_2. \quad (2.69)$$

Следовательно,

$$\left.\begin{aligned} \frac{E_x}{a_1}\sin\delta_2 - \frac{E_y}{a_2}\sin\delta_1 = \cos\tau\sin\left(\delta_2 - \delta_1\right), \\ \frac{E_x}{a_1}\cos\delta_2 - \frac{E_y}{a_2}\cos\delta_1 = \sin\tau\sin\left(\delta_2 - \delta_1\right). \end{aligned}\right\} \quad (2.70)$$

Возводя в квадрат и складывая, получим

$$\left(\frac{E_x}{a_1}\right)^2 + \left(\frac{E_y}{a_2}\right)^2 - 2\frac{E_x}{a_1}\frac{E_y}{a_2}\cos\delta = \sin^2\delta, \quad (2.71)$$

где

$$\delta = \delta_2 - \delta_1. \quad (2.72)$$

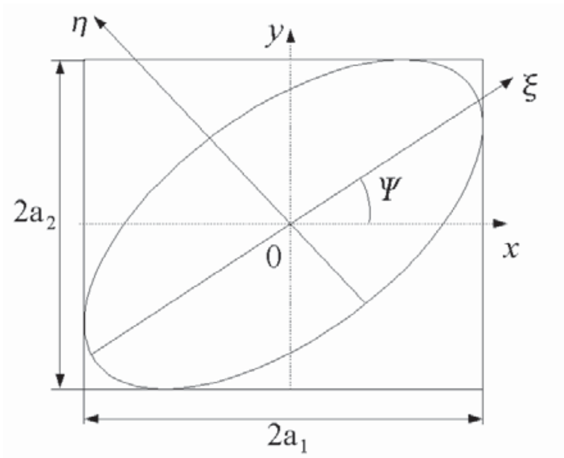

Рис 6. Эллиптически поляризованная волна. Эллипс, соответствующий колебанию электрического вектора.

Соотношение (2.71) является уравнением конического сечения. Оно имеет форму эллипса, так как соответствующий детерминант неотрицателен, т. е.

$$\begin{vmatrix} \dfrac{1}{a_1^2} & -\dfrac{\cos\delta}{a_1 a_2} \\ -\dfrac{\cos\delta}{a_1 a_2} & \dfrac{1}{a_2^2} \end{vmatrix} = \frac{1}{a_1^2 a_2^2}\left(1-\cos^2\delta\right) = \frac{\sin^2\delta}{a_1^2 a_2^2} \geq 0.$$

Этот эллипс вписан в прямоугольник, стороны которого параллельны осям координат и имеют длины $2a_1$ и $2a_2$ (рис. 6.). Эллипс касается сторон прямоугольника в точках $\left(\pm a_1, \pm a_2\cos\delta\right)$ и $\left(\pm a_1\cos\delta, \pm a_2\right)$. В этом случае говорят, что волна, описываемая (2.68), эллиптически поляризована. Легко видеть, что волна, связанная с магнитным вектором, также поляризована эллиптически. Из (2.61) и (2.68) следует

$$\left.\begin{aligned} &H_x = -\sqrt{\varepsilon/\mu}E_y = -\sqrt{\varepsilon/\mu}a_2\cos\left(\tau+\delta_2\right), \\ &H_y = \sqrt{\varepsilon/\mu}E_x = \sqrt{\varepsilon/\mu}a_1\cos\left(\tau+\delta_1\right), \\ &H_z = 0. \end{aligned}\right\} \tag{2.73}$$

Конец магнитного вектора описывает эллипс, который вписан в прямоугольник со сторонами длиной $2\sqrt{\varepsilon/\mu}a_2$ и $2\sqrt{\varepsilon/\mu}a_1$, параллельными осями x и y.

В общем случае оси эллипса не параллельны осям O_x и O_y. Пусть O_ξ и O_η — новые оси, направленные по осям эллипса, а $\psi\left(0 \leq \psi < \pi\right)$ — угол между O_x и направлением главной оси O_ξ (см. рис. 1.6). Тогда компоненты E_ξ и E_η будут связаны с E_x и E_y соотношениями

$$E_\xi = E_x\cos\psi + E_y\sin\psi, \quad E_\eta = -E_x\sin\psi + E_y\cos\psi. \tag{2.74}$$

Если $2a$ и $2b$ $\left(a \geq b\right)$ — длины осей эллипса, то уравнение эллипса относительно осей O_ξ, O_η будет иметь вид

$$E_\xi = a\cos(\tau + \delta_0), \qquad E_\eta = \pm b\sin(\tau + \delta_0). \tag{2.75}$$

Наличие двух знаков указывает на возможность двух направлений движения конца электрического вектора, описывающего эллипс.

Чтобы определить a и b, сравним (2.74) и (2.75), и используем соотношение (2.69), тогда

$$a(\cos\tau\cos\delta_0 - \sin\tau\sin\delta_0) =$$
$$= a_1(\cos\tau\cos\delta_1 - \sin\tau\sin\delta_1)\cos\psi + a_2(\cos\tau\cos\delta_2 - \sin\tau\sin\delta_2)\sin\psi.$$
$$\pm b(\sin\tau\cos\delta_0 + \cos\tau\sin\delta_0) =$$
$$= -a_1(\cos\tau\cos\delta_1 - \sin\tau\sin\delta_1)\sin\psi + a_2(\cos\tau\cos\delta_2 - \sin\tau\sin\delta_2)\cos\psi.$$

Приравнивая коэффициенты при $\cos\tau$ и $\sin\tau$, получим

$$\begin{aligned} a\cos\delta_0 &= a_1\cos\delta_1\cos\psi + a_2\cos\delta_2\sin\psi, \\ a\sin\delta_0 &= a_1\sin\delta_1\cos\psi + a_2\sin\delta_2\sin\psi, \end{aligned} \tag{2.76}$$

$$\begin{aligned} \pm b\cos\delta_0 &= a_1\sin\delta_1\sin\psi - a_2\sin\delta_2\cos\psi, \\ \pm b\sin\delta_0 &= -a_1\cos\delta_1\sin\psi + a_2\cos\delta_2\cos\psi. \end{aligned} \tag{2.77}$$

Возведя в квадрат и складывая уравнения (2.76) и (2.77) с использованием соотношения (2.72), находим

$$\left.\begin{aligned} a^2 &= a_1^2\cos^2\psi + a_2^2\sin^2\psi + 2a_1a_2\cos\psi\cos\delta, \\ b^2 &= a_1^2\sin^2\psi + a_2^2\cos^2\psi - 2a_1a_2\cos\psi\sin\psi\cos\delta. \end{aligned}\right\} \tag{2.78}$$

Следовательно,

$$a^2 + b^2 = a_1^2 + a_2^2. \tag{2.79}$$

Умножим теперь (20а) на (21а), (20б) на (21б) и сложим. Это даст

$$\pm ab = a_1a_2\sin\delta. \tag{2.80}$$

Деля (21а) на (20а) и (21б) на (20б), получим

$$\pm\frac{b}{a} = \frac{a_1\sin\delta_1\sin\psi - a_2\sin\delta_2\cos\psi}{a_1\cos\delta_1\cos\psi + a_2\cos\delta_2\sin\psi} = \frac{-a_1\cos\delta_1\sin\psi + a_2\cos\delta_2\cos\psi}{a_1\sin\delta_1\cos\psi + a_2\sin\delta_2\sin\psi}.$$

Отсюда находим следующее уравнение относительно ψ

$$\left(a_1^2 - a_2^2\right)\sin 2\psi = 2a_1a_2\cos\delta\cos 2\psi \tag{2.81}$$

Удобно ввести такой вспомогательный угол $\alpha(0 \le \alpha \le \pi/2)$, чтобы

$$a_2 / a_1 = tg\alpha. \qquad (2.82)$$

Тогда предыдущее уравнение (2.81) примет вид $tg2\psi = \frac{2a_1a_2}{a_1^2 - a_2^2}\cos\delta = \frac{2tg\alpha}{1 - tg^2\alpha}\cos\delta.$

Его можно упростить:

$$tg2\psi = (tg2\alpha)\cos\delta. \qquad (2.83)$$

Из (2.80) и (2.81) мы найдем также

$$\pm\frac{2ab}{a^2 + b^2} = \frac{2a_1a_2}{a_1^2 + a_2^2}\sin\delta = (\sin 2\alpha)\sin\delta. \qquad (2.84)$$

Пусть $\chi(-\pi/4 \le \chi \le \pi/4)$ — другой вспомогательный угол, такой что

$$\pm b / a = tg\chi \qquad (2.85)$$

Численное значение $tg\chi$ определяет величину отношения осей эллипса, а знак при χ характеризует два варианта, которые можно использовать при описании эллипса. Перепишем уравнение (2.84) в виде

$$\sin 2\chi = (\sin 2\alpha)\sin\delta. \qquad (2.86)$$

Полезно кратко просуммировать результаты. Если заданы a_1, a_2 и разность фаз δ, относящиеся к произвольному положению оси, и если $\alpha(0 \le \alpha \le \pi/2)$ — угол, определяемый соотношением

$$tg\alpha = a_2 / a_1, \qquad (2.87)$$

то главные полуоси эллипса a и b и угол $\psi(0 \le \psi < \pi)$, который большая ось эллипса образует с осью O_x, находятся из формул

$$\begin{aligned} &a^2 + b^2 = a_1^2 + a_2^2, \\ &tg2\psi = (tg2\alpha)\cos\delta, \\ &\sin 2\chi = (\sin 2\alpha)\sin\delta, \end{aligned} \qquad (2.88)$$

где $\chi\left(-\pi/4 \le \chi \le \pi/4\right)$ — вспомогательный угол, определяющий форму и ориентацию эллипса колебаний, а именно

$$tg\chi = \pm b/a. \tag{2.89}$$

Наоборот, если известны длины осей a и b и ориентация эллипса (т. е. заданы a, b и ψ), то эти формулы позволяют найти амплитуды a_1, a_2 и разность фаз δ.

2.6. Линейная и круговая поляризации.

Наиболее важны два специальных случая, когда эллипс поляризации вырождается либо в прямую линию, либо в окружность. Согласно формулам (2.69), (2.70) эллипс перейдет в прямую линию при

$$\delta = \delta_2 - \delta_1 = m\pi \qquad (m = 0, \pm 1, \pm 2, ...). \tag{2.90}$$

Тогда

$$\frac{E_y}{E_x} = (-1)^m \frac{a_2}{a_1}, \tag{2.91}$$

и мы говорим о линейной поляризации $\mathbf{E}$. Одну из координатных осей, например, x, можно выбрать вдоль этой прямой. Тогда остается лишь одна компонента, а именно, E_x. Более того, поскольку векторы напряженностей электрического и магнитного полей ортогональны и лежат в плоскости, перпендикулярной к направлению z, компонента H_x тоже исчезает, и, следовательно, вектор $\mathbf{H}$ линейно поляризован в направлении y.

Другой важный специальный случай - это случай круговой поляризации волны, когда эллипс вырождается в круг. Ясно, что необходимым условием этого является превращение в квадрат описанного выше прямоугольника, то есть выполнение соотношения $a_1 = a_2 = a$.

Кроме того, одна из компонент вектора $\mathbf{E}$ должна равняться нулю, когда другая достигает своего максимума. Отсюда следует, что

$$\delta = \delta_2 - \delta_1 = \frac{m\pi}{2} \qquad (m = \pm 1, \pm 3, \pm 5, ...) . \tag{2.92}$$

и уравнение (15) переходит в уравнение окружности

$$E_x^2 + E_y^2 = a^2 . \tag{2.93}$$

В случае *правой* поляризации $\sin\delta > 0$, так что

$$\delta = \frac{\pi}{2} + 2m\pi \qquad (m = 0, \pm 1, \pm 2, ...) , \tag{2.94}$$

$$E_x = a\cos(\tau + \delta_1),\ E_y = a\cos(\tau + \delta_1 + \frac{\pi}{2}) = -a\sin(\tau + \delta_1) . \tag{2.95}$$

Для *левой* поляризации $\sin\delta < 0$, так что

$$\delta = -\frac{\pi}{2} + 2m\pi \qquad (m = 0, \pm 1, \pm 2, ...) , \tag{2.96}$$

$$E_x = a\cos(\tau + \delta_1),\ E_y = a\cos(\tau + \delta_1 - \frac{\pi}{2}) = a\sin(\tau + \delta_1) . \tag{2.97}$$

Если вместо вещественного представления используется комплексное (то есть если вместо косинусов в (2.68) написаны экспоненциальные функции), то

$$\frac{E_y}{E_x} = \frac{a_2}{a_1}\exp\left[i(\delta_2 - \delta_1)\right] = \frac{a_2}{a_1}e^{i\delta} , \tag{2.98}$$

и из значения этого отношения можно сразу определить характер поляризации.

Линейная поляризация электрической волны ($\delta = m\pi,\, m = 0, \pm 1, \pm 2, ...$)

$$\frac{E_y}{E_x} = (-1)^m \frac{a_2}{a_1} .$$

Правая круговая поляризация электрической волны ($a_1 = a_2,\, \delta = \pi/2$)

$$\frac{E_y}{E_x} = \exp\left[i\pi/2)\right] = i .$$

Левая круговая поляризация электрической волны ($a_1 = a_2,\, \delta = -\pi/2$)

$$\frac{E_y}{E_x} = \exp[-i\pi / 2)] = -i .$$

В более общем случае можно показать, что для правой эллиптической поляризации мнимая часть отношения E_y / E_x положительна, тогда как для левой эллиптической поляризации она отрицательна.

На рисунке 7 показаны эллипсы поляризации при разных значениях δ.

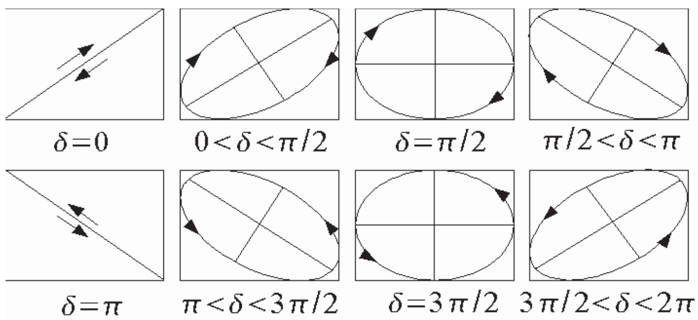

Рис. 7. Эллиптическая поляризация при различных значениях разности фаз δ.

2.7. Характеристика состояния поляризации с помощью параметров Стокса.

Для характеристики эллипса поляризации необходимы три независимые величины, например амплитуды a_1 и a_2 и разность фаз δ или малая и большая оси a,b и угол ψ, характеризующий ориентацию эллипса. Для практических целей состояние поляризации удобно охарактеризовать некоторыми параметрами, обладающими одинаковой физической размерностью; они были введены Стоксом в 1852 г. при его исследованиях частично поляризованного света.

Параметрами Стокса для плоской монохроматической волны служат четыре величины:

$$s_0 = a_1^2 + a_2^2,\ s_1 = a_1^2 - a_2^2,\ s_2 = 2a_1a_2\cos\delta,\ s_3 = 2a_1a_2\sin\delta \tag{2.99}$$

Лишь три из них независимы, так как справедливо тождество

$$s_0^2 = s_1^2 + s_2^2 + s_3^2 \tag{2.100}$$

Очевидно, что параметр s_0 пропорционален интенсивности волны. Параметры s_1, s_2 и s_3 простым образом связаны с углом $\psi\,(0 \le \psi < \pi)$, характеризующим ориентацию эллипса, и углом $\chi\,(-\pi/4 \le \chi < \pi/4)$, характеризующим эллиптичность и направление вращения. Выполняются следующие соотношения:

$$s_1 = s_0 \cos 2\chi \cos 2\psi \tag{2.101}$$

$$s_2 = s_0 \cos 2\chi \sin 2\psi \tag{2.102}$$

$$s_3 = s_0 \sin 2\chi\,. \tag{2.103}$$

Выражение (2.103) следует из (2.82) и (2.86). Для вывода двух других соотношений заметим, что, согласно уравнению, предшествующему (2.83),

$$s_2 = s_1 tg 2\psi\,. \tag{2.104}$$

Соотношение (2.101) получается, если выражения (2.104) и (2.103) подставить в (2.100). Наконец, (2.102) получается при подстановке (2.101) в (2.104).

Тема 3. Отражение и преломление плоской волны

Поскольку электромагнитное поле должно быть определено во всем пространстве, для его единственности кроме условий на границах раздела сред нужно поставить определенные условия на бесконечности. Условия на бесконечности должны удовлетворять следующим требованиям: физическому – они должны выделять волны, уходящие в бесконечность, и исключать возможность появления волн, приходящих из бесконечности; математическому – они должны выделять единственное решение системы уравнений Максвелла. Первым сформулировал условия на бесконечности А.Зоммерфельд. Затем для более сложных геометрий были предложены и другие условия на бесконечности. Наиболее полно они изложены в научных работах Тихонова, Самарского, Свешникова и их учеников.

В пункте 1.2 были получены соотношения, которым должны удовлетворять векторы поля на поверхностях, где физические свойства среды претерпевают разрыв. Применим теперь эти формулы к исследованию распространения плоской волны, падающей на плоскую границу, разделяющую две однородные изотропные среды.

3.1. Законы отражения и преломления Снеллиуса.

Если на границу двух однородных сред с разными оптическими свойствами падает плоская волна, она разделяется на две волны: проходящую во вторую среду и отраженную. Из предположения, что прошедшая и отраженная волны являются плоскими и монохроматичными, следует существование частного решения. Из граничного условия на бесконечности следует единственность решения, т. е. невозможность других решений. Из предположения, что эти волны являются плоскими, выведем выражения для их амплитуд и направлений распространения.

Плоская волна, распространяющаяся в направлении единичного вектора $s^{(i)}$, полностью определена, если известно поведение возмущения во времени $F(t)$ в одной точке пространства. Если $F(t)$ представляет зависимость возмущения от времени в какой-то одной точке, то эта зависимость в другой точке, отстоящей от первой на r, будет $F\left[t-(r\cdot s)/\upsilon\right]$. На границе двух сред вторичные поля будут так же изменяться во времени, как и первичное поле падающей волны. Следовательно, если $s^{(r)}$ и $s^{(t)}$ — единичные векторы в направлении распространения отраженной и прошедшей волн, то, приравнивая аргументы трех волновых функций в точке r на граничной плоскости $z=0$,получим граничное условие: $E_i\exp+E_r\exp=E_t\exp$ (индексы i, r и t относятся соответственно к падающей, отраженной и проходящей (переломленной) волнам). Оно выполняется в любой момент времени, следовательно, совпадают $\exp\{*\}$ у всех трех вкладов и, соответственно, совпадают их фазы:

$$t-\frac{r\cdot s^{(i)}}{\upsilon_1}=t-\frac{r\cdot s^{(r)}}{\upsilon_1}=t-\frac{r\cdot s^{(t)}}{\upsilon_2}, \tag{3.1}$$

где υ_1 и υ_2 — скорости распространения в одной и другой средах. Учитывая, что $r=\{x,y,0\}$, находим из (3.1)

$$\frac{xs_x^{(i)}+ys_y^{(i)}}{\upsilon_1}=\frac{xs_x^{(r)}+ys_y^{(r)}}{\upsilon_1}=\frac{xs_x^{(t)}+ys_y^{(t)}}{\upsilon_2}. \tag{3.2}$$

Равенства (3.2) должны выполняться для любых значений x и y на границе, и поэтому

$$\frac{s_x^{(i)}}{\upsilon_1}=\frac{s_x^{(r)}}{\upsilon_1}=\frac{s_x^{(t)}}{\upsilon_2},\qquad \frac{s_y^{(i)}}{\upsilon_1}=\frac{s_y^{(r)}}{\upsilon_1}=\frac{s_y^{(t)}}{\upsilon_2}. \tag{3.3}$$

Плоскость, определяемая вектором $s^{(i)}$ и нормалью к границе, называется плоскостью падения. Соотношения (3.3) показывают, что и $s^{(r)}$, и $s^{(t)}$ лежат в этой плоскости.

Считая плоскость xz плоскостью падения и обозначая через θ_i, θ_r и θ_t углы, которые $s^{(i)}, s^{(r)}$ и $s^{(t)}$ образуют с осью O_z, получим (рис. 8)

$$\left.\begin{aligned} s_x^{(i)} &= \sin\theta_i, \quad s_y^{(i)} = 0, \quad s_z^{(i)} = \cos\theta_i, \\ s_x^{(r)} &= \sin\theta_r, \quad s_y^{(r)} = 0, \quad s_z^{(r)} = \cos\theta_r, \\ s_x^{(t)} &= \sin\theta_t, \quad s_y^{(t)} = 0, \quad s_z^{(t)} = \cos\theta_t. \end{aligned}\right\} \tag{3.4}$$

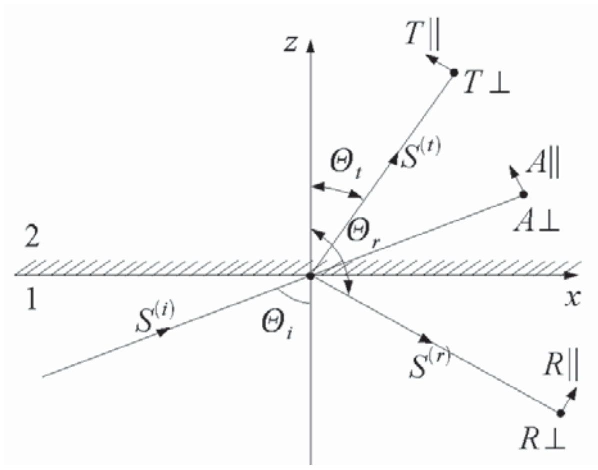

Рис. 8. Преломление и отражение плоской волны. Плоскость падения.

Если волна распространяется из первой среды во вторую, компонента вектора s вдоль оси z положительна; если волна распространяется в противоположном направлении, эта компонента отрицательна, т. е.

$$\begin{aligned} &s_z^{(i)} = \cos\theta_i \geq 0, \quad s_z^{(r)} = \cos\theta_r \leq 0, \\ &s_z^{(t)} = \cos\theta_t \geq 0. \end{aligned} \tag{3.5}$$

Подставляя значения (3.4) в первую систему равенств (3.3), получим

$$\frac{\sin\theta_i}{\upsilon_1} = \frac{\sin\theta_r}{\upsilon_1} = \frac{\sin\theta_t}{\upsilon_2}. \tag{3.6}$$

Следовательно, $\sin\theta_r = \sin\theta_t$. Используя (3.5), мы находим, что $\cos\theta_r = -\cos\theta_i$, поэтому

$$\theta_r = \pi - \theta_i \tag{3.7}$$

Это соотношение вместе с утверждением, что нормаль $s^{(r)}$ к отраженной волне лежит в плоскости падения, составляет закон отражения.

Используя соотношение $n = \sqrt{\varepsilon\mu}$, из (3.6) найдем также

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{\upsilon_1}{\upsilon_2} = \sqrt{\frac{\varepsilon_2\mu_2}{\varepsilon_1\mu_1}} = \frac{n_2}{n} = n_{12}. \tag{3.8}$$

Соотношение $\sin\theta_i / \sin\theta_t = n_2 / n_1$ вместе с утверждением, что нормаль $s^{(t)}$ к преломленной волне лежит в плоскости падения, составляет закон преломления (или закон Снеллиуса).

Если $n_2 > n_1$, то $n_{12} > 1$, и мы говорим, что оптическая плотность второй среды больше, чем первой. В этом случае, учитывая (3.6), имеем

$$\sin\theta_t = \frac{1}{n_{12}}\sin\theta_i < \sin\theta_i \tag{3.9}$$

так что для каждого угла падения существует вещественный угол преломления θ_t. Однако если вторая среда оптически менее плотна, чем первая (т. е. если $n_{12} < 1$), то вещественное значение θ_t мы получим лишь для таких углов падения θ_i, для которых $\sin\theta_i \le n_{12}$. Для больших значений θ_i; имеет место так называемое полное внутреннее отражение.

3.2. Формулы Френеля.

Здесь мы рассмотрим амплитуды отраженной и преломленной волн. Предположим, что обе среды (однородные и изотропные) обладают нулевой проводимостью и, следовательно, совершенно прозрачны; их магнитные проницаемости фактически будут отличаться от единицы на пренебрежимо малые величины, и поэтому мы положим $\mu_1 = \mu_2 = 1$.

Пусть A — амплитуда электрического вектора поля падающей волны, будем считать ее комплексной величиной с фазой, равной постоянной части аргумента волновой функции. Переменная ее часть имеет вид

$$\tau_i = \omega\left(t - \frac{r \cdot s^{(t)}}{\upsilon_1}\right) = \omega\left(t - \frac{x\sin\theta_i + z\cos\theta_i}{\upsilon_1}\right). \tag{3.10}$$

Разложим каждый вектор на две компоненты - параллельную (снабдим ее индексом) и перпендикулярную (индекс $\perp$) плоскости падения. Выбор положительных направлений для параллельных компонент указан на рис. 8. Перпендикулярные компоненты располагаются перпендикулярно к плоскости рисунка.

Тогда компоненты электрического вектора поля падающей волны запишутся в виде

$$\left.\begin{aligned} &E_x^{(i)} = -A\,\cos\theta_i \exp(-i\tau_i), \quad E_y^{(i)} = A_\perp \exp(-i\tau_i),\\ &E_z^{(i)} = -A\,\sin\theta_i \exp(-i\tau_i). \end{aligned}\right\} \tag{3.11}$$

Компоненты магнитного вектора сразу же получаются (при $\mu = 1$) из соотношения (2.61)

$$H = \sqrt{\varepsilon} s \times E. \tag{3.12}$$

Отсюда

$$\left.\begin{aligned} &H_x^{(i)} = -A_\perp \cos\theta_i \sqrt{\varepsilon_1} \exp(-i\tau_i),\\ &H_y^{(i)} = -A\,\sqrt{\varepsilon_1} \exp(-i\tau_i),\\ &H_z^{(i)} = A_\perp \sin\theta_i \sqrt{\varepsilon_1} \exp(-i\tau_i). \end{aligned}\right\} \tag{3.13}$$

Аналогично, если T и R — комплексные амплитуды прошедшей и отраженной волн, то соответствующие компоненты электрического и магнитного векторов равны следующим величинам.

Поле прошедшей волны

$$\left.\begin{aligned}
&E_x^{(t)} = -T\cos\theta_t \exp(-i\tau_t); && \mathrm{H}_x^{(t)} = -T_\perp \cos\theta_t \sqrt{\varepsilon_2}\exp(-i\tau_t);\\
&E_y^{(t)} = T_\perp \exp(-i\tau_t); && \mathrm{H}_y^{(t)} = -T\sqrt{\varepsilon_2}\exp(-i\tau_t);\\
&E_z^{(t)} = T\sin\theta_t \exp(-i\tau_t); && \mathrm{H}_z^{(t)} = T_\perp \cos\theta_t \sqrt{\varepsilon_2}\exp(-i\tau_t),
\end{aligned}\right\} \tag{3.14}$$

где

$$\tau_t = \omega\left(t - \frac{r\cdot s^{(t)}}{\upsilon_2}\right) = \omega\left(t - \frac{x\sin\theta_t + z\cos\theta_t}{\upsilon_2}\right). \tag{3.15}$$

Поле отраженной волны

$$\left.\begin{aligned}
&E_x^{(r)} = -R\cos\theta_r \exp(-i\tau_r); && \mathrm{H}_x^{(r)} = -R_\perp \cos\theta_r \sqrt{\varepsilon_1}\exp(-i\tau_r);\\
&E_y^{(r)} = R_\perp \exp(-i\tau_r); && \mathrm{H}_y^{(r)} = -R\sqrt{\varepsilon_1}\exp(-i\tau_r);\\
&E_z^{(r)} = R\sin\theta_r \exp(-i\tau_r); && \mathrm{H}_z^{(r)} = R_\perp \cos\theta_r \sqrt{\varepsilon_1}\exp(-i\tau_r),
\end{aligned}\right\} \tag{3.16}$$

Граничные условия (1.23) и (1.25) требуют, чтобы на границе тангенциальные составляющие векторов E и H были непрерывны. Следовательно, должны выполняться соотношения

$$\left.\begin{aligned}
E_x^{(i)} + E_x^{(r)} = E_x^{(t)}, \quad E_y^{(i)} + E_y^{(r)} = E_y^{(t)},\\
H_x^{(i)} + H_x^{(r)} = H_x^{(t)}, \quad H_y^{(i)} + H_y^{(r)} = H_y^{(t)},
\end{aligned}\right\} \tag{3.17}$$

при этом условия для нормальных компонент **B** и **D** (1.15) и (1.19) будут удовлетворяться автоматически. Подставляя в (18) значения всех компонент и используя тот факт, что $\theta_r = \cos(\pi - \theta_i) = -\cos\theta_i$, получим четыре соотношения

$$\left.\begin{aligned}
&\cos\theta_i (A - R) = \cos\theta_t T, && A_\perp + R_\perp = T_\perp,\\
&\sqrt{\varepsilon_1}\cos\theta_i (A_\perp - R_\perp) = \sqrt{\varepsilon_2}\cos\theta_t T_\perp, && \sqrt{\varepsilon_1}(A + R) = \sqrt{\varepsilon_2} T.
\end{aligned}\right\} \tag{3.18}$$

Заметим, что уравнения (3.18) делятся на две группы, одна из которых содержит лишь компоненты, параллельные плоскости падения, а другая — только компоненты, перпендикулярные ей.

Следовательно, волны этих двух типов независимы друг от друга.

Можно решить уравнения (3.18) относительно компонент отраженной и прошедшей волн, выразив их через компоненты падающей волны. Вновь используя соотношение $n = \sqrt{\varepsilon}$,получим

$$T = \frac{2n_1 \cos\theta_i}{n_2 \cos\theta_i + n_1 \cos\theta_t} A, \quad T_\perp = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t} A_\perp, \tag{3.19}$$

$$R = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} A, \quad \mathrm{R}_\perp = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} A_\perp \tag{3.20}$$

Уравнения (3.19) и (3.20) называются формулами Френеля. Впервые они были выведены Френелем в несколько менее общем виде в 1823 г. на основе его теории, рассматривавшей свет как колебания упругой среды. Эти соотношения пишутся обычно в другой форме, которую можно получить из (3.19) и (3.20), используя закон преломления (3.8), а именно в форме

$$T = \frac{2\sin\theta_t \cos\theta_i}{\sin(\theta_i - \theta_t)\cos(\theta_i - \theta_t)} A, \quad T_\perp = \frac{2\sin\theta_t \cos\theta_i}{\sin(\theta_i + \theta_t)} A_\perp, \tag{3.21}$$

$$R = \frac{tg(\theta_i - \theta_t)}{tg(\theta_i + \theta_t)} A, \qquad \mathrm{R}_\perp = -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} A_\perp. \tag{3.22}$$

Так как θ_i и θ_t вещественны (случай полного внутреннего отражения пока исключаем), то тригонометрические функции, стоящие в правой части уравнений (3.21) и (3.22), также вещественны. Следовательно, фаза каждой компоненты отраженной и прошедшей волн либо равна по фазе соответствующей компоненты падающей волны, либо отличается от нее на π. Так как знаки T и $T_\perp$ совпадают со знаками A и $A_\perp$, фаза прошедшей волны равна фазе падающей. В случае же отраженной волны фаза будет зависеть от относительных значений θ_i и θ_t. Если оптическая плотность второй среды больше, чем первой $(\varepsilon_2 < \varepsilon_1)$, то $\theta_t < \theta_i$; и поэтому, согласно (3.20), знаки $R_\perp$ и $A_\perp$ различны и фазы отличаются друг от друга на π. При этом значение $tg(\theta_i - \theta_t)$ положительно, но знаменатель $tg(\theta_i + \theta_t)$

может стать отрицательным при $\theta_i + \theta_t > \pi / 2$, и в том случае фазы R и A отличаются друг от друга на π. Аналогичное рассмотрение можно провести для случая, когда вторая среда оптически менее плотна, чем первая.

Для нормального падения $\theta_i = 0$ и, следовательно, $\theta_t = 0$; тогда соотношения (3.19) и (3.20) примут вид

$$T = \frac{2}{n+1} A, \quad T_{\perp} = \frac{2}{n+1} A_{\perp}, \tag{3.23}$$

$$R = \frac{n-1}{n+1} A, \quad \mathrm{R}_{\perp} = -\frac{n-1}{n+1} A_{\perp}. \tag{3.24}$$

3.3. Отражательная и пропускательная способности; поляризация при отражении и преломлении.

Рассмотрим теперь, как энергия поля падающей волны распределяется между двумя вторичными полями.

Интенсивность света (снова считаем $\mu = 1$), согласно (2.65), равна

$$S = \frac{c}{4\pi} \sqrt{\varepsilon} E^2 = \frac{cn}{4\pi} E^2. \tag{3.25}$$

Поэтому количество энергии в первичной волне, которое попадает на единицу площади поверхности раздела за 1 сек, будет равно

$$J^{(r)} = S^{(i)} \cos\theta_i = \frac{cn_1}{4\pi} |A|^2 \cos\theta_i. \tag{3.26}$$

Для отраженной и преломленной волн энергия, покидающая единицу площади поверхности раздела за 1 сек, определяется подобными же выражениями, а именно:

$$J^{(r)} = S^{(r)} \cos\theta_i = \frac{cn_1}{4\pi} |R|^2 \cos\theta_i, \quad J^{(t)} = S^{(t)} \cos\theta_t = \frac{cn_2}{4\pi} |T|^2 \cos\theta_t \tag{3.27}$$

Отношения

$$\Re = \frac{J^{(r)}}{J^{(i)}} = \frac{|R|^2}{|A|^2} \quad \text{и} \quad \Im = \frac{J^{(t)}}{J^{(i)}} = \frac{n_2}{n_1}\frac{\cos\theta_t}{\cos\theta_i}\frac{|T|^2}{|A|^2} \tag{3.28}$$

называют соответственно отражательной и пропускательной способностью. Легко проверить, что в соответствии с законом сохранения энергии

$$\Re + \Im = 1 \tag{3.29}$$

Как видно из (2.65), если $\mu \neq 1$, множитель n_2 / n_1 в выражении для $\Im$ следует заменить величиной $\frac{\sqrt{\varepsilon_2 / \mu_2}}{\sqrt{\varepsilon_1 / \mu_1}}$.

Отражательная и пропускательная способности зависят от поляризации падающей волны. Их можно выразить через соответствующие отражательную и пропускательную способности для света, поляризованного параллельно и перпендикулярно плоскости падения.

Пусть вектор E падающей волны образует с плоскостью падения угол α_i Тогда

$$A = A\cos\alpha_i, \quad A_{\perp} = A\sin\alpha_i. \tag{3.30}$$

Пусть, далее,

$$\begin{aligned} J^{(i)} &= \frac{cn_1}{4\pi}\left|A\right|^2 \cos\theta_i = J^{(i)}\cos^2\alpha_i, \\ J_{\perp}^{(i)} &= \frac{cn_1}{4\pi}\left|A_{\perp}\right|^2 \cos\theta_i = J^{(i)}\sin^2\alpha_i, \end{aligned} \tag{3.31}$$

(30)

и

$$\begin{aligned} J^{(r)} &= \frac{cn_1}{4\pi}\left|R\right|^2 \cos\theta_i \\ J_{\perp}^{(r)} &= \frac{cn_1}{4\pi}\left|R_{\perp}\right|^2 \cos\theta_i \end{aligned} \tag{3.32}$$

(31)

Тогда

$$\begin{aligned}\Re &= \frac{J^{(r)}}{J^{(i)}} = \frac{J^{(r)} + J_{\perp}^{(r)}}{J^{(i)}} = \\ &= \frac{J^{(r)}}{J^{(i)}}\cos^2\alpha_i + \frac{J_{\perp}^{(r)}}{J_{\perp}^{(i)}}\sin^2\alpha_i = \\ &= \Re\ \cos^2\alpha_i + \Re_{\perp}\sin^2\alpha_i,\end{aligned} \tag{3.33}$$

где

$$\left.\begin{aligned}\Re\ &= \frac{J^{(r)}}{J^{(i)}} = \frac{\left|R\ \right|^2}{\left|A\ \right|^2} = \frac{tg^2\left(\theta_i - \theta_t\right)}{tg^2\left(\theta_i + \theta_t\right)}, \\ \Re_{\perp} &= \frac{J_{\perp}^{(r)}}{J_{\perp}^{(i)}} = \frac{\left|R_{\perp}\right|^2}{\left|A_{\perp}\right|^2} = \frac{\sin^2\left(\theta_i - \theta_t\right)}{\sin^2\left(\theta_i + \theta_t\right)}.\end{aligned}\right\} \tag{3.34}$$

Подобным же образом получим

$$\Im = \frac{J^{(t)}}{J^{(i)}} = \Im\ \cos^2\alpha_i + \Im_{\perp}\sin^2\alpha_i, \tag{3.35}$$

где

$$\left.\begin{aligned}\Im\ &= \frac{J^{(t)}}{J^{(i)}} = \frac{n_2}{n_1}\frac{\cos\theta_t}{\cos\theta_i}\frac{\left|T\ \right|^2}{\left|A\ \right|^2} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2\left(\theta_i + \theta_t\right)\cos^2\left(\theta_i - \theta_t\right)}, \\ \Im_{\perp} &= \frac{J_{\perp}^{(t)}}{J_{\perp}^{(i)}} = \frac{n_2}{n_1}\frac{\cos\theta_t}{\cos\theta_i}\frac{\left|T_{\perp}\right|^2}{\left|A_{\perp}\right|^2} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2\left(\theta_i + \theta_t\right)}.\end{aligned}\right\} \tag{3.36}$$

Снова можно показать, что

$$\Re\ + \Im\ = 1, \quad \Re_{\perp} + \Im_{\perp} = 1. \tag{3.37}$$

Для нормального падения различие между параллельной и перпендикулярной компонентами исчезает, и из (3.23), (3.24) и (3.28) находим

$$\Re = \left(\frac{n-1}{n+1} \right)^2, \quad \Im = \frac{4\pi}{(n+1)^2}. \tag{3.38}$$

Отсюда следует, что

$$\lim_{n \to 1} \Re = 0, \quad \lim_{n \to 1} \Im = 1. \tag{3.39}$$

Аналогичные результаты получаются также для предельных значений $\Re$, $\Im$ и $\Re_{\perp}, \Im_{\perp}$. Это легко увидеть из (3.34) и (3.36), если учесть, что, согласно закону преломления, $\theta_t \to \theta_i$ при $n \to 1$. Следовательно, чем меньше различие в оптической плотности обеих сред, тем меньше энергии уносится отраженной волной.

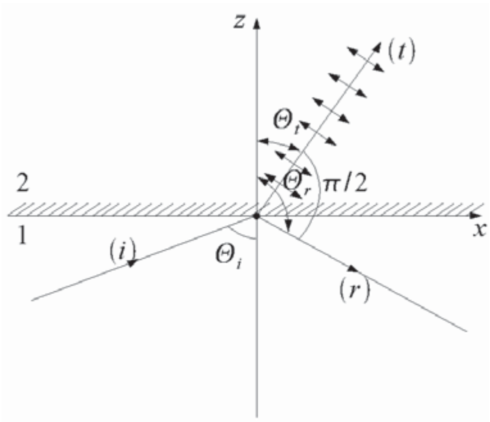

Рис 9. К определению угла полной поляризации (угол Брюстера).

Знаменатели в (3.34) и (3.36) конечны, за исключением случая $\theta_i + \theta_t = \pi / 2$. Тогда $tg(\theta_i + \theta_t) = \infty$ и, следовательно, $\Re = 0$. В этом случае (рис. 9) отраженный и преломленный лучи перпендикулярны друг другу, а из закона преломления следует (так как теперь $\sin\theta_t = \sin\{\pi / 2 - \theta_i\} = \cos\theta_i$), что

$$tg\theta_i = n. \tag{3.40}$$

Угол θ_i, определяемый этим выражением, называется углом полной поляризации или углом Брюстера.

3.4. Полное внутреннее отражение.

До сих пор мы исключали случай, когда закон преломления

$$\sin\theta_t = \frac{\sin\theta_i}{n_{12}} \tag{3.41}$$

не дает вещественного значения для угла преломления θ_t. Сейчас мы исследуем этот случай. Он реализуется при распространении света из оптически более плотной среды в среду с меньшей оптической плотностью, т. е. когда $n_{12} = n_2 / n_1 = \sqrt{\varepsilon_2\mu_2 / \varepsilon_1\mu_1} < 1$, при условии, что угол падения θ_i, превосходит критическое значение θ_i, определяющееся выражением

$$\sin\overline{\theta}_i = n_{12} \tag{3.42}$$

Если $\theta_i = \overline{\theta}_i$, то $\sin\theta_t = 1$, т. е. $\theta_t = 90^0$, так что направление распространения света касательно к поверхности раздела. Если θ_i превышает предельное значение в $\overline{\theta}_i$, свет не входит во вторую среду. Весь падающий свет отражается обратно в первую среду, и мы говорим о полном внутреннем отражении.

Тем не менее, электромагнитное поле во второй среде не равно нулю, отсутствует лишь поток энергии через границу. Действительно, если в фазовом множителе (3.15) прошедшей волны мы положим

$$\sin\theta_t = \frac{\sin\theta_i}{n}, \quad \cos\theta_t = \pm i\sqrt{\frac{\sin^2\theta_i}{n^2}} - 1 \tag{3.43}$$

(нижний индекс $_{12}$ у n_{12} опущен), то получим

$$\exp\left(-i\tau_t\right) = \exp\left[-i\omega\left(t\frac{x\sin\theta_i}{n\upsilon_2}\right)\right]\exp\left[\pm\frac{\omega z}{\upsilon^2}\sqrt{\frac{\sin^2\theta_i}{n^2} - 1}\right]. \tag{3.44}$$

Выражение (3.44) описывает неоднородную волну, которая распространяется вдоль поверхности раздела в плоскости падения (т. е. в направлении x) и меняется экспоненциально с изменением расстояния z от этой поверхности. Конечно, физический смысл имеет лишь отрицательный знак перед квадратным корнем в выражении (3.44), в противном случае при увеличении расстояния амплитуда росла бы неограниченно. Как мы видим, амплитуда очень быстро уменьшается с глубиной проникновения z, причем эффективная глубина проникновения равна $\upsilon_2 / \omega = \lambda_2 / 2\pi$, т. е. сравнима с длиной волны.

Эта волна не является поперечной, поскольку, как будет показано ниже, компонента электрического вектора в направлении распространения не равна нулю.

Чтобы применить формулы Френеля (3.20), (3.22) к случаю полного внутреннего отражения, перепишем их в виде

$$\delta z_1, \delta z_2, \delta z_3, \ldots, \delta z_n \tag{3.45}$$

Подставив в эти выражения значения величин (3.43) и помня, что перед квадратным корнем необходимо брать верхний знак, получим

$$\left.\begin{aligned} R &= \frac{n^2\cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{n^2\cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}} A, \\ R_\perp &= \frac{\cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{\cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}} A_\perp. \end{aligned}\right\} \tag{3.46}$$

Следовательно,

$$|R| = |A|, \quad |R_\perp| = |A_\perp|, \tag{3.47}$$

т. е. для каждой компоненты интенсивность света, отраженного при полном внутреннем отражении, равна интенсивности падающего света.

Замечание. Хотя во второй среде поле отлично от нуля, легко видеть, что поток энергии через поверхность отсутствует. Точнее, можно показать, что хотя в общем случае компонента вектора Пойнтинга в направлении, нормальном к границе, конечна, ее значение, усредненное по времени, равно нулю. Это означает, что не существует постоянного потока во вторую среду, а энергия течет туда и обратно. Запишем для $z=0$ компоненты поля прошедшей волны вдоль осей x и y и используем (3.43). (Тут необходимо брать вещественные выражения для E и H, так как поток энергии является квадратичной функцией компонент.) Отмечая сопряженную комплексную величину звездочкой, получим из (3.14)

$$E_x^{(t)} = -\frac{1}{2}\left(T\cos\theta_t \exp\left(-i\tau_t^0\right) + T^*\cos^*\theta_t \exp\left(+i\tau_t^0\right)\right) =$$

$$= -\frac{1}{2}\sqrt{\frac{\sin^2\theta_i}{n^2} - 1}\cdot\left(T\exp\left(-i\tau_t^0\right) - T^*\exp\left(+i\tau_t^0\right)\right),$$

$$E_y^{(t)} = \frac{1}{2}\left(T_\perp \exp\left(-i\tau_t^0\right) + T_\perp^*\exp\left(+i\tau_t^0\right)\right),$$

$$H_x^{(t)} = -\frac{1}{2}\left(T_\perp\cos\theta_t\sqrt{\varepsilon_2}\exp\left(-i\tau_t^0\right) + T_\perp^*\cos\theta_t\sqrt{\varepsilon_2}\exp\left(+i\tau_t^0\right)\right) =$$

$$= -\frac{1}{2}\sqrt{\varepsilon_2}\sqrt{\frac{\sin^2\theta_i}{n^2} - 1}\left(T_\perp\exp\left(-i\tau_t^0\right) - T_\perp^*\exp\left(+i\tau_t^0\right)\right),$$

$$H_y^{(t)} = -\frac{1}{2}\sqrt{\varepsilon_2}\left(T\exp\left(-i\tau_t^0\right) + T^*\exp\left(+i\tau_t^0\right)\right),$$

где $\tau_t^0 = \omega\left(t - \dfrac{x\sin\theta_i}{n\upsilon_2}\right)$. Если мы рассмотрим среднее по времени значение величины $S_z^{(t)} = \dfrac{c}{4\pi}\left(E_x^{(t)}H_y^{(t)} - E_y^{(t)}H_x^{(t)}\right)$ в интервале $-t' \le t \le t''$, где t' велико по сравнению с периодом $T = 2\pi/\omega$, то увидим, что оба члена исчезают при $z=0$. Действительно, один из них содержит множитель

$$\frac{1}{2t'}\int_{-t'}^{t'}\left(T^2\exp\left(-2i\tau_t^0\right) - \left(T^*\right)^2\exp\left(+2i\tau_t^0\right)\right)dt =$$

$$= \left(T^2\exp\left(+\frac{2i\omega x\sin\theta_i}{n\upsilon^2}\right) - \left(T^*\right)^2\exp\left(-\frac{2i\omega x\sin\theta_i}{n\upsilon^2}\right)\right)O\left(\frac{T}{t'}\right),$$

который пренебрежимо мал при $t' \quad T$; другой содержит подобный же множитель с $T_\perp$ вместо T

Вместе с тем расчет показывает, что среднее по времени значение двух других компонент вектора $S^{(t)}$ для $z=0$, а именно $S_x^{(t)}$ и $S_y^{(t)}$, в общем случае оказываются конечными. Поэтому энергия не проникает во вторую среду, а течет вдоль поверхности раздела в плоскости падения.

Наконец, вычислим изменения фаз компонент отраженной и падающей волн. На основании (3.47) мы можем положить

$$\frac{R}{A}=e^{i\delta} \quad , \qquad \frac{R_\perp}{A_\perp}=e^{i\delta} \quad . \tag{3.48}$$

Согласно (57) каждая из величин $R\ /\ A$ и $R_\perp\ /\ A_\perp$ имеет форму $z\left(z^*\right)^{-1}$. Следовательно, если α — аргумент z (т. е. $z=ae^{i\alpha}$, где a и α вещественны), то

$$e^{i\delta}=z\left(z^*\right)^{-1}=e^{2i\alpha}, \text{ т.е. } tg\frac{\delta}{2}=tg\alpha,$$

и поэтому

$$tg\frac{\delta}{2}=-\frac{\sqrt{\sin^2\theta_i-n^2}}{n^2\cos\theta_i}, \qquad tg\frac{\delta_\perp}{2}=-\frac{\sqrt{\sin^2\theta_i-n^2}}{\cos\theta_i}. \tag{3.49}$$

Отсюда видно, что обе компоненты испытывают скачки фаз разной величины. Вследствие этого линейно поляризованный свет при полном внутреннем отражении окажется поляризованным эллиптически.

Тема 4. Распространение волн в слоистой среде.

4.1. Теория диэлектрических пленок

Среда, свойства которой постоянны на каждой плоскости, перпендикулярной к фиксированному направлению, называется слоистой средой. Если считать это специальное направление осью z декартовой системы координат, то

$$\varepsilon = \varepsilon(z), \quad \mu = \mu(z). \tag{4.1}$$

Рассмотрим распространение плоской гармонической электромагнитной волны через такую среду. Это естественное обобщение простого случая, рассмотренного выше.

Теория слоистых сред приобретает важное значение в оптике в связи с многослойными системами, т. е. с системами тонких плоскопараллельных пленок. Такие пленки можно изготовлять методом напыления в высоком вакууме, а их толщину можно контролировать с очень большой точностью. Они находят множество полезных приложений. Например, их можно применять в качестве просветляющих пленок, т. е. в качестве покрытий, которые уменьшают отражение от данной поверхности. Вместе с тем тонкие пленки при соответствующих условиях будут увеличивать отражение. Нанесенные на поверхность стекла пленки можно использовать для разделения пучка; такие устройства применяются в интерферометрии для разделения падающего пучка на две части. При определенных условиях многослойная система может служить фильтром, который пропускает (или отражает) лишь выделенные участки спектра, Многослойные системы употребляются также в качестве поляризаторов.

Вопрос о диэлектрических и металлических пленках очень широко обсуждался в научной литературе. Было предложено много схем для расчета оптических свойств многослойных систем. Мы следуем схеме, изложенной в монографии [1]

Здесь мы будем заниматься лишь диэлектрической слоистой средой.

4.2. Основные дифференциальные уравнения.

Рассмотрим плоскую гармоническую электромагнитную волну, распространяющуюся через слоистую среду. В частном случае, когда волна поляризована линейно и ее электрический вектор перпендикулярен к плоскости падения, мы будем говорить о поперечной электрической волне (обозначаемой TE); если она поляризована линейно и ее магнитный вектор перпендикулярен к плоскости падения, мы будем говорить о поперечной магнитной волне (обозначаемой TM). Любую произвольно поляризованную плоскую волну можно разложить на две волны, одна из которых является волной TE-типа, а другая — TM-типа. Так как граничные условия двух поляризаций на поверхности раздела не зависят друг от друга (смотрите тему 3), то эти две волны также будут взаимно независимы. Более того, если поменять местами E и H и одновременно ε и $-\mu$, то уравнения .Максвелла не изменятся. Поэтому любую теорему, относящуюся к TM-волнам, сразу же можно вывести из соответствующего результата для TE-волн с помощью такой замены. Таким образом, достаточно изучить подробно лишь TE-волны.

Возьмем в качестве плоскости падения плоскость yz, причем z — направление поперек слоев. Для волны TE-типа $E_y = E_z = 0$, и уравнения Максвелла переходят в следующие шесть скалярных уравнений (зависимость от времени предполагается в виде $\exp(-i\omega t)$):

$$\frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial z} + \frac{i\varepsilon\omega}{c} E_x = 0, \tag{4.2}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = 0, \tag{4.3}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = 0, \tag{4.4}$$

$$\frac{i\omega\mu}{c}H_x = 0, \tag{4.5}$$

$$\frac{\partial E_x}{\partial z} - \frac{i\omega\mu}{c}H_y = 0, \tag{4.6}$$

$$\frac{\partial E_x}{\partial y} + \frac{i\omega\mu}{c}H_z = 0. \tag{4.7}$$

Эти уравнения показывают, что H_y, H_z и E_x зависят только от y и z. Исключая H_y и H_z из (4.2), (4.6), и (4.7), найдем

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} + n^2 k_0{}^2 E_x = \frac{d(\ln\mu)}{dz}\frac{\partial E_x}{dz}, \tag{4.8}$$

где

$$n^2 = \varepsilon\mu, \quad k_0 = \omega / c = 2\pi / \lambda_0. \tag{4.9}$$

Будем искать решение (4.8) в виде произведения двух функций, одна из которых зависит от y, а другая только от z:

$$E_x(y,z) = Y(y)U(z). \tag{4.10}$$

Тогда уравнение (4.8) примет вид

$$\frac{1}{Y}\frac{d^2Y}{dy^2} = -\frac{1}{U}\frac{d^2U}{dz^2} - n^2k_0^2 + \frac{d(\ln\mu)}{dz}\frac{1}{U}\frac{dU}{dz} \tag{4.11}$$

Левая его часть зависит лишь от y, тогда как правая — лишь от z. Следовательно, (4.11) может выполняться лишь в том случае, если каждая его часть равна постоянной (скажем, $-K^2$), т. е.

$$\frac{1}{Y}\frac{d^2Y}{dy^2} = -K^2, \tag{4.12}$$

$$\frac{d^2U}{dz^2} - \frac{d(\ln\mu)}{dz}\frac{dU}{dz} + n^2k_0{}^2U = K^2U. \tag{4.13}$$

Удобно положить

$$K^2 = k_0{}^2\alpha^2. \tag{4.14}$$

Тогда уравнение (4.12) дает $Y = const \cdot \exp(ik_0 \alpha y)$, и, следовательно, E_x имеет вид

$$E_x = U(z)\exp\left[i(k_0 \alpha y - \omega t)\right], \tag{4.15}$$

где $U(z)$ — функция z (возможно, комплексная). Из (4.6) и (4.7) мы видим, что выражения для H_y и H_z имеют такую же форму, т. е.

$$H_y = V(z)\exp\left[i(k_0 \alpha y - \omega t)\right], \tag{4.16}$$

$$H_z = W(z)\exp\left[i(k_0 \alpha y - \omega t)\right]. \tag{4.17}$$

Амплитудные функции U, V и W на основании (4.2), (4.6) и (4.7) связаны следующими уравнениями:

$$V' = ik_0(\alpha W + \varepsilon U), \tag{4.18}$$

$$U' = ik_0 \mu V, \tag{4.19}$$

$$\alpha U + \mu W = 0; \tag{4.20}$$

здесь штрих означает дифференцирование по z. Подставляя W из (4.20) в (4.18), мы получим вместе с (4.19) систему из двух дифференциальных уравнений первого порядка относительно U и V

$$\begin{aligned} U' &= ik_0 \mu V, \\ V' &= ik_0\left(\varepsilon - \frac{\alpha^2}{\mu}\right)U. \end{aligned} \tag{4.21}$$

Переходя к уравнениям, содержащим лишь по одной неизвестной функции, окончательно получим следующие линейные дифференциальные уравнения второго порядка относительно U и V:

$$\frac{d^2U}{dz^2} - \frac{d(\ln \mu)}{dz}\frac{dU}{dz} + k_0^{\,2}(n^2 - \alpha^2)U = 0, \tag{4.22}$$

$$\frac{d^2V}{dz^2} - \frac{d\left[\ln\left(\varepsilon - \frac{\alpha^2}{\mu}\right)\right]}{dz}\frac{dV}{dz} + k_0^{\,2}(n^2 - \alpha^2)V = 0. \tag{4.23}$$

В соответствии с правилом, которое является следствием симметрии уравнений Максвелла, сразу же можно написать, что для волны TM-типа $\left(H_y = H_z = 0\right)$ неисчезающие компоненты векторов поля имеют вид

$$H_x = U(z)\exp\left[i\left(k_0\alpha y - \omega t\right)\right], \tag{4.24}$$

$$E_y = -V(z)\exp\left[i\left(k_0\alpha y - \omega t\right)\right], \tag{4.25}$$

$$E_z = -W(z)\exp\left[i\left(k_0\alpha y - \omega t\right)\right], \tag{4.26}$$

причем

$$U' = ik_0\varepsilon V, \quad V' = ik_0\left(\mu - \frac{\alpha^2}{\varepsilon}\right)U, \tag{4.27}$$

а W и U связаны соотношением

$$\alpha U + \varepsilon W = 0. \tag{4.28}$$

Функции W и V удовлетворяют следующим линейным дифференциальным уравнениям второго порядка:

$$\frac{d^2U}{dz^2} - \frac{d(\ln\varepsilon)}{dz}\frac{dU}{dz} + k_0^{\,2}\left(n^2 - \alpha^2\right)U = 0, \tag{4.29}$$

$$\frac{d^2V}{dz^2} - \frac{d\left[\ln\left(\mu - \frac{\alpha^2}{\varepsilon}\right)\right]}{dz}\frac{dV}{dz} + k_0^{\,2}\left(n^2 - \alpha^2\right)V = 0. \tag{4.30}$$

В общем случае U, V и W являются комплексными функциями z. Поверхности постоянной амплитуды E_x определяются из уравнения

$$|U(z)| = const,$$

поверхности постоянной фазы — из уравнения

$$\varphi(z) + k_0\alpha y = const,$$

где $\varphi(z)$ — фаза U. Оба семейства поверхностей в общем случае не совпадают, так что E_x (и аналогично H_y и H_z) будет неоднородной волной. Для небольшого смещения (dy, dz) вдоль поверхности постоянной фазы имеем

$$\varphi'(z)dz + k_0\alpha dy = 0;$$

следовательно, если через θ обозначить угол между нормалью к поверхности постоянной фазы и O_z, то получим

$$tg\theta = -\frac{dz}{dy} = \frac{k_0\alpha}{\varphi'(z)}.$$

В специальном случае однородной плоской волны имеем

$$\varphi(z) = k_0 nz\cos\theta, \quad \alpha = n\sin\theta. \tag{4.31}$$

Следовательно, соотношение $\alpha = const$ с условием (4.14) можно рассматривать как обобщение закона преломления Снеллиуса для слоистых сред.

4.3. Характеристическая матрица для слоистой среды.

Решения только что выведенных дифференциальных уравнений, подчиняющиеся соответствующим граничным условиям, и различные теоремы, относящиеся к слоистым средам, удобнее представлять в матричной форме.

Поскольку функции $U(z)$ и $\mathrm{V}(z)$ в пункте 4.2. удовлетворяют линейным дифференциальным уравнениям второго порядка (типа (4.22) и (4.23)), каждую из них можно выразить в виде линейной комбинации двух частных решений из фундаментальной системы, например, U_1, U_2 и V_1, V_2. Эти частные решения не могут быть произвольными; они должны быть независимыми. Это означает, что определитель Вронского этих двух решений должен быть отличен от нуля [3]. В частности они могут быть связаны дифференциальными уравнениями первого порядка (4.21), а именно

$$\left.\begin{aligned} U_1' &= ik_0\mu V_1, \\ V_1' &= ik_0\left(\varepsilon - \frac{\alpha^2}{\mu}\right)U_1. \end{aligned}\right\} \qquad \left.\begin{aligned} U_2' &= ik_0\mu V_2, \\ V_2' &= ik_0\left(\varepsilon - \frac{\alpha^2}{\mu}\right)U_2. \end{aligned}\right\} \tag{4.32}$$

Отсюда следует, что

$$V_1U_2' - U_1'V_2 = 0, \qquad U_1V_2' - V_1U_2 = 0,$$

так что

$$\frac{d}{dz}\left(U_1V_2 - U_2V_1\right) = 0.$$

Это соотношение означает, что определитель

$$D = \begin{vmatrix} U_1 & V_1 \\ U_2 & V_2 \end{vmatrix}, \tag{4.33}$$

соответствующий любым двум произвольным решениям системы (4.21), постоянен, т. е. что D является инвариантом нашей системы уравнений.

Для наших целей наиболее удобно выбрать частные решения:

$$U_1 = f(z), \quad U_2 = F(z), \quad V_1 = g(z), \quad V_2 = G(z), \tag{4.34}$$

удовлетворяющие граничным условиям

$$f(0) = G(0) = 0 \quad \text{и} \quad \mathrm{F}(0) = g(0) = 1. \tag{4.35}$$

Тогда решение с начальными условиями вида

$$U(0) = U_0, \quad V(0) = V_0 \tag{4.36}$$

можно выразить через решения (4.34) в виде

$$U = FU_0 + fV_0, \qquad V = GU_0 + gV_0,$$

или в матричной форме

$$Q = NQ_0, \tag{4.37}$$

где

$$Q = \begin{bmatrix} U(z) \\ V(z) \end{bmatrix}, \quad Q_0 = \begin{bmatrix} U_0 \\ V_0 \end{bmatrix}, \quad \mathrm{N} = \begin{bmatrix} \mathrm{F}(z) & f(z) \\ G(z) & g(z) \end{bmatrix}. \tag{4.38}$$

Из соотношения $D = const$ следует, что определитель квадратной матрицы N постоянен. Значение этой постоянной сразу же можно получить, полагая $z = 0$, что дает

$$|N| = Fg - fG = 1.$$

Обычно удобнее выражать U_0 и V_0 через функции $U(z)$ и $V(z)$. Разрешая относительно U_0 и V_0 получим

$$Q_0 = MQ, \tag{4.39}$$

где

$$M = \begin{bmatrix} g(z) & -f(z) \\ -G(z) & F(z) \end{bmatrix}. \tag{4.40}$$

Эта матрица также унимодулярная, т. е.

$$|M| = 1. \tag{4.41}$$

Смысл матрицы M ясен; она связывает x- и y-компоненты электрического (или магнитного) векторов на плоскости $z = 0$ с этими компонентами на произвольной плоскости $z = const$. Таким образом, мы видим, что для полного определения поля достаточно знать U и V. Следовательно, для того чтобы узнать, как распространяется плоская монохроматическая волна через слоистую среду, последнюю необходимо охарактеризовать лишь соответствующей унимодулярной 2×2-матрицей. По этой причине мы будем называть M характеристической матрицей слоистой среды.

Ниже рассматривается форма характеристической матрицы для случаев, представляющих особый интерес.

Однородная диэлектрическая пленка.

В этом случае величины ε, μ и $n = \sqrt{\varepsilon\mu}$ постоянны. Если θ — угол между нормалью к волне и осью z, то, согласно (4.31), имеем

$$\alpha = n\sin\theta.$$

Для волны TE-типа получим в соответствии с (4.22) и (4.23)

$$\frac{d^2U}{dz^2}+\left(k_0^{\,2}n^2\cos^2\theta\right)U=0,\qquad \frac{d^2V}{dz^2}+\left(k_0^{\,2}n^2\cos^2\theta\right)V=0.$$

Легко видеть, что решения этих уравнений, удовлетворяющие соотношениям (4.21), имеют вид

$$\left.\begin{aligned}&U(z)=A\cos\left(k_0nz\cos\theta\right)+B\sin\left(k_0nz\cos\theta\right),\\&V(z)=\frac{1}{i}\sqrt{\frac{\varepsilon}{\mu}}\cos\theta\left\{B\cos\left(k_0nz\cos\theta\right)-A\sin\left(k_0nz\cos\theta\right)\right\}.\end{aligned}\right\}\tag{4.42}$$

Следовательно, частное решение (4.34), удовлетворяющее граничным условиям (4.35), запишется следующим образом:

$$\left.\begin{aligned}&U_1=f(z)=\frac{i}{\cos\theta}\sqrt{\mu/\varepsilon}\sin\left(k_0nz\cos\theta\right),\\&V_1=g(z)=\cos\left(k_0nz\cos\theta\right),\\&U_2=F(z)=\cos\left(k_0nz\cos\theta\right),\\&V_2=G(z)=i\sqrt{\mu/\varepsilon}\cos\theta\sin\left(k_0nz\cos\theta\right).\end{aligned}\right\}\tag{4.43}$$

Если мы положим

$$p=\sqrt{\varepsilon/\mu}\cos\theta,\tag{4.44}$$

то получим характеристическую матрицу в виде

$$M(z)=\begin{bmatrix}\cos\left(k_0nz\cos\theta\right) & -\dfrac{i}{p}\sin\left(k_0nz\cos\theta\right)\\ -ip\sin\left(k_0nz\cos\theta\right) & \cos\left(k_0nz\cos\theta\right)\end{bmatrix}.\tag{4.45}$$

Те же уравнения будут справедливы для волны TM-типа, если p заменить на

$$q=\sqrt{\frac{\mu}{\varepsilon}}\cos\theta\tag{4.46}$$

Слоистая среда, состоящая из тонких однородных пленок.

Рассмотрим две смежные слоистые среды, первая из которых занимает пространство от $z = 0$ до $z = z_1$, а вторая — от $z = z_1$ до $z = z_2$. Если $M_1(z)$ и $\mathrm{M}_2(z)$ — характеристические матрицы двух сред, то

$$Q_0 = M_1(z)Q(z_1), \quad Q(z_1) = M_2(z_2 - z_1)Q(z_2),$$

так что $Q_0 = M(z_2)Q(z_2)$, где $\mathrm{M}(z_2) = M_1(z_1)M_2(z_2 - z_1)$.

Этот результат немедленно можно обобщить на случай непрерывного ряда слоистых сред, расположенных в областях

$0 \le z \le z_1, \ z_1 \le z \le z_2, ..., z_{N-1} \le z \le z_N$.

Если $M_1, M_2, ..., M_N$ — характеристические матрицы сред, то

$$\begin{aligned} Q_0 &= M(z_N)Q(z_N), \\ M(z_N) &= M_1(z_1)M_2(z_2 - z_1)\cdot ... \cdot M_N(z_N - z_{N-1}) \end{aligned} \tag{4.47}$$

С помощью последнего соотношения легко вывести соответствующее выражение для характеристической матрицы любой слоистой среды: мы разбиваем эту среду на очень большое число тонких пленок толщиной $\delta z_1, \delta z_2, \delta z_3, ..., \delta z_n$. Если их максимальная толщина достаточно мала, можно считать, что ε, μ и n постоянны в каждой пленке. Из (4.45) видно, что в этом случае характеристическая матрица j-й пленки приближенно равна

$$M_j = \begin{bmatrix} 1 & -\dfrac{1}{p_j}k_0 n_j \delta z_j \cos\theta_j \\ -ipk_0 n_j \delta z_j \cos\theta_j & 1 \end{bmatrix}.$$

Следовательно, характеристическая матрица всей среды, рассматриваемой как совокупность тонких пленок, приблизительно равна

$$M = \prod_{j=1}^{N} M_j = \begin{bmatrix} 1 & -ik_0 B \\ -ik_0 A & 1 \end{bmatrix}, \tag{4.48}$$

где

$$A = \sum_{j=1}^{N} p_j n_j \delta z_j \cos\theta_j = \sum_{j=1}^{N} \left(\varepsilon_j - \frac{\alpha^2}{\mu_j} \right) \delta z_j,$$

$$B = \sum_{j=1}^{N} \frac{n_j}{p_j} \delta z_j \cos\theta_j = \sum_{j=1}^{N} \mu_j \delta z_j.$$

Здесь также оставлены лишь члены до первого порядка по δz включительно. Переходя к пределу при $N \to \infty$ таким образом, чтобы максимум $|\delta z_j|$ стремился к нулю, получим

$$M = \begin{bmatrix} 1 & -ik_0\mathcal{B} \\ -ik_0\mathcal{A} & 1 \end{bmatrix}, \tag{4.49}$$

где

$$\mathcal{A} = \int \left(\varepsilon - \frac{\alpha^2}{\mu} \right) dz, \quad \mathcal{B} = \int \mu dz, \tag{4.50}$$

а интегрирование проводится по всему интервалу изменения z. Выражение (4.49) дает первое приближение для характеристической матрицы произвольной слоистой среды. Оставляя в разложении величин $\cos(k_0 n\delta z\cos\theta)$ и $\sin(k_0 n\delta z\cos\theta)$ и в произведении (4.48) члены более высоких порядков, можно получить следующие приближения.

Поскольку для непоглощающей среды ε и μ вещественны, то видно также, что характеристическая матрица непоглощающей слоистой среды имеет вид

$$M = \begin{bmatrix} a & ib \\ ic & d \end{bmatrix}, \tag{4.51}$$

где a, b, c и d вещественны.

4.4. Коэффициенты отражения и пропускания.

Рассмотрим плоскую волну, падающую на слоистую среду, которая занимает область от $z = 0$ до $z = z_1$. И с обеих сторон граничит с однородными полубесконечными средами. Выведем выражения для амплитуд и интенсивностей отраженной и прошедшей волн.

Пусть A, R и T обозначают, как и раньше, амплитуды (возможно, комплексные) векторов напряженности электрического поля для волн TE-типа (и векторов напряженности магнитного поля для волн TM-типа) падающей, отраженной и преломленной волн. Далее пусть ε_1, μ_1 и ε_i, μ_i — диэлектрические и магнитные проницаемости первой и последней сред, а θ_1 и θ_i — углы между нормалями к падающей и прошедшей волнам и направлением оси z (направлением стратификации).

Граничные условия, приведенные в пункте 1.2, требуют, чтобы тангенциальные компоненты векторов E и H были непрерывны на каждой из двух поверхностей раздела слоистой среды. Вместе с соотношением $H = \sqrt{\varepsilon / \mu} s \times E$, это приводит к следующим соотношениям для волны TE-типа:

$$\left.\begin{aligned} &U_0 = A + R, \qquad U(z_1) = T, \\ &V_0 = p_1(A - R), \quad V(z_1) = p_i T, \end{aligned}\right\} \tag{4.52}$$

где

$$p_1 = \sqrt{\varepsilon_1 / \mu_1} \cos\theta_1, \qquad p_i = \sqrt{\varepsilon_i / \mu_i} \cos\theta_i. \tag{4.53}$$

Четыре величины U_0, V_0, U и V, определяемые равенствами (4.52), связаны основным соотношением (4.39); следовательно,

$$A + R = (m'_{11} + m'_{12} p_i) T, \qquad p_1 (A - R) = (m'_{21} + m'_{22} p_i) T, \tag{4.54}$$

где m_{ij} — элементы характеристической матрицы седы при $z = z_1$.

Из (4.54) мы найдем коэффициенты отражения и пропускания пленки

$$r = \frac{R}{A} = \frac{(m'_{11} + m'_{12} p_i) p_1 - (m'_{21} + m'_{22} p_i)}{(m'_{11} + m'_{12} p_i) p_1 + (m'_{21} + m'_{22} p_i)}, \tag{4.55}$$

$$t = \frac{T}{A} = \frac{2 p_1}{(m'_{11} + m'_{12} p_i) p_1 + (m'_{21} + m'_{22} p_i)}. \tag{4.56}$$

Отражательная и пропускательная способности, выраженные через r и t имеют вид

$$\Re = |r|^2, \quad \Im = \frac{p_i}{p_1}|t|^2. \tag{4.57}$$

Фазу δ_r величины r можно назвать изменением фазы при отражении, а фазу δ_t величины t — изменением фазы при пропускании. Изменение фазы δ_r отнесено к первой поверхности раздела, тогда как δ_t — к плоской границе между слоистой средой и последней полубесконечной средой.

Соответствующие формулы для волны TM -типа сразу же получаются из (4.55) — (4.57) путем замены величины p_1 и p_i на величины

$$q_1 = \sqrt{\mu_1 / \varepsilon_1}\cos\theta_1, \qquad q_i = \sqrt{\mu_i / \varepsilon_i}\cos\theta_i. \tag{4.58}$$

В этом случае r и t являются отношениями амплитуд магнитных, а не электрических векторов.

Тема 5. Электромагнитное поле в анизотропной среде

5.1. Определение анизотропной среды

Будем считать среду анизотропной, если связь между индукцией и напряженностью по отношению к электромагнитным волнам определяется тензорами $\hat{\varepsilon}_{ik}$ и $\hat{\mu}_{ik}$. Поэтому для анизотропной среды материальные уравнения приобретают вид:

$$D_i = \hat{\varepsilon}_{ik} E_k \tag{5.1}$$

$$B_i = \hat{\mu}_{ik} H_k \tag{5.2}$$

Таким образом, анизотропия среды характеризуется различной по разным направлениям способностью среды реагировать на действие падающей на неё электромагнитной волны. Реакция эта состоит в смещении электрических зарядов под действием электрического поля волны. В случае изотропных сред это смещение происходило параллельно вектору напряженности электрического поля волны; в анизотропных средах это не так – смещение зарядов происходит под некоторым углом по отношению к вызывающему его вектору напряженности. Если же анизотропная среда обладает магнитными свойствами, то можно привести аналогичные рассуждения, имея в виду не смещение зарядов, а возникновение магнитного момента, не совпадающего по направлению с напряженностью магнитного поля волны.

Конкретные значения составляющих тензоров диэлектрической и магнитной проницаемостей зависят от конкретной анизотропной среды (например – кристалла), однако общими признаками их является зависимость от частоты и симметричность. Последнее обусловлено термодинамикой кристаллических сред, более подробно с этим вопросом можно ознакомиться в [4].

5.2. Уравнения Максвелла для анизотропной среды

В силу ряда обстоятельств при анализе оптических свойств анизотропных сред удобнее пользоваться системой единиц Гаусса. В этой системе единиц уравнения Максвелла имеют вид:

$$rot\vec{H} = \frac{1}{c}\cdot\frac{\partial}{\partial t}\vec{D} \qquad (5.3)$$

$$rot\vec{E} = -\frac{1}{c}\frac{\partial}{\partial t}\vec{B}, \qquad (5.4)$$

где с – скорость света в вакууме.

$$\begin{aligned} div\vec{D} &= 0, \\ div\vec{B} &= 0. \end{aligned} \qquad (5.5)$$

Совместно с (5.1) и (5.2) эти уравнения полностью описывают процесс распространения электромагнитной волны в непроводящей и несегнетоэлектрической (т.е. не имеющей спонтанной поляризации) анизотропной среде.

В дальнейшем под термином электромагнитные волны будем понимать только плоские монохроматические волны, зависимость полевых векторов которых от пространственных координат $\vec{r}$ и времени t определяется формулами:

$$\begin{aligned} \vec{E}(\vec{r},t) &= \vec{E}_0 \exp[-i(\omega t - \vec{k}\cdot\vec{r})] \\ \vec{H}(\vec{r},t) &= \vec{H}_0 \exp[-i(\omega t - \vec{k}\cdot\vec{r})] \\ \vec{D}(\vec{r},t) &= \vec{D}_0 \exp[-i(\omega t - \vec{k}\cdot\vec{r})] \\ \vec{B}(\vec{r},t) &= \vec{B}_0 \exp[-i(\omega t - \vec{k}\cdot\vec{r})] \quad , \end{aligned} \tag{5.6}$$

где ω- круговая частота, $\vec{k}\cdot\vec{r}$ - скалярное произведение волнового вектора и радиус-вектора точки наблюдения.

Поскольку мы ограничили класс электромагнитных волн только плоскими монохроматическими волнами, мы можем подставить выражения для полевых векторов в уравнения Максвелла. Нетрудно заметить, что действия на экспоненциальные векторные функции вида (5.6) операторов rot, div, $\frac{\partial}{\partial t}$ сводятся к векторному и скалярному умножению на $i\vec{k}$ и умножению на $-i\omega$ соответственно. Таким образом уравнения (5.3) – (5.5) приобретают вид:

$$\left[\vec{k}\times\vec{H}\right] = -k_0\vec{D} \ , \qquad \left[\vec{k}\times\vec{E}\right] = k_0\vec{B} \tag{5.7}$$

$$\vec{k}\cdot\vec{D} = 0 \ , \qquad \vec{k}\cdot\vec{B} = 0 \tag{5.8}$$

где через $k_0 = \frac{\omega}{c}$ мы обозначили волновое число вакуума.

Из уравнений (5.8) видно, что в анизотропной среде поперечными векторами электромагнитной волны являются только векторы индукции $\vec{D}$ и $\vec{B}$, которые лежат в плоскости волнового фронта. Вспомним, что в изотропных средах поперечными векторами были также и векторы напряженностей. Кроме того, из уравнений (5.7) и (5.8) видно, что, по-прежнему, векторы напряженностей и векторы индукции попарно ортогональны друг другу, но отсутствует коллинеарность между напряженностями и индукциями полей (см. (5.1) и (5.2)).

В изотропной среде направление фазовой скорости электромагнитной волны и направление переноса энергии этой волной совпадали. Направление фазовой скорости волны в анизотропной среде мы уже определили; это направление задается волновым вектором. Определим теперь направление переноса энергии электромагнитной волны в анизотропной среде, которое, в общем случае, задается вектором плотности потока энергии (или вектором Умова-Пойнтинга):

$$\vec{P} = \frac{c}{4\pi}\left[\vec{E} \times \vec{H}\right] \tag{5.9}$$

По аналогии с (5.8) можно записать:

$$\vec{P} \cdot \vec{E} = 0 \,, \qquad \vec{P} \cdot \vec{H} = 0 \tag{5.10}$$

Так как векторы $\vec{D}$ и $\vec{E}$ не коллинеарны, так же, впрочем, как $\vec{B}$ и $\vec{H}$, то и направление фазовой скорости электромагнитной волны в анизотропной среде не совпадает с направлением потока энергии. Следовательно, для полного описания процесса распространения электромагнитных волн в анизотропной среде (в отличие от среды изотропной) необходимо учитывать этот факт.

На практике, за редким исключением, мы имеем дело с немагнитными, оптически прозрачными средами. Это означает, что тензор магнитной восприимчивости можно считать единичным $\hat{\mu}_{ij} = 1$. В этом случае вектор индукции магнитного поля волны $\vec{B}$ в наших выкладках мы можем заменить вектором напряженности $\vec{H}$. Взаимная ориентация всех векторов, характеризующих процесс распространения электромагнитной волны в анизотропной среде, показана на рис. 10.

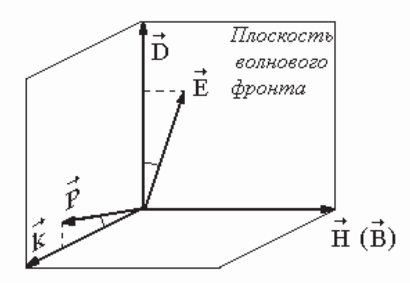

Рис 10. Ориентация полевых векторов плоской электромагнитной волны, распространяющейся в немагнитной анизотропной среде.

Итак, подводя промежуточный итог, можно сказать, что электромагнитная волна в анизотропной среде по-прежнему поперечная, но в плоскости волнового фронта $\vec{k} \cdot \vec{r} = const$ теперь лежат векторы индукции, а векторы напряженности могут выходить из этой плоскости.

5.3. Волновое и дисперсионное уравнения

Для немагнитной анизотропной среды можно получить волновое уравнение точно так же, как и в случае изотропных сред. Для этого исключим из уравнений Максвелла (5.7) вектор напряженности магнитного поля:

$$\left[\vec{k} \times \left[\vec{k} \times \vec{E}\right]\right] + k_0^2 D = 0$$

Воспользовавшись материальным уравнением (5.1), можно записать

$$\left[\vec{k} \times \left[\vec{k} \times \vec{E}\right]\right] + k_0^2 \hat{\varepsilon} \vec{E} = 0 \qquad (5.11)$$

Это векторное уравнение мы будем называть волновым уравнением для плоских электромагнитных волн в анизотропной среде. Раскрывая двойное векторное произведение, получим:

$$\vec{k}\left(\vec{k}\cdot\vec{E}\right)-k^2\vec{E}+k_0^2\hat{\varepsilon}\vec{E}=0 \ .$$

Поскольку в анизотропной среде, как это было показано выше, векторы $\vec{k}$ и $\vec{E}$ не ортогональны (в отличие от изотропных сред), то скалярное произведение уже нельзя полагать равным нулю. Для дальнейшего анализа полученное нами уравнение удобнее представить в виде однородной системы трех линейных уравнений относительно проекций волнового вектора и вектора напряженности электрического поля волны на оси координат. В матричной форме эта система уравнений будет иметь вид:

$$\begin{vmatrix} \left(\tau^2-K^2+\varepsilon_{11}\right) & \left(\beta\tau+\varepsilon_{12}\right) & \left(\gamma\tau+\varepsilon_{13}\right) \\ \left(\beta\tau+\varepsilon_{12}\right) & \left(\beta^2-K^2+\varepsilon_{22}\right) & \left(\beta\gamma+\varepsilon_{23}\right) \\ \left(\gamma\tau+\varepsilon_{13}\right) & \left(\beta\gamma+\varepsilon_{23}\right) & \left(\gamma^2-K^2+\varepsilon_{33}\right) \end{vmatrix} \cdot \begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix} = 0 \qquad (5.12)$$

где $\beta=\dfrac{k_y}{k_0}$, $\gamma=\dfrac{k_z}{k_0}$, $\tau=\dfrac{k_x}{k_0}$ нормированные на волновое число вакуума проекции волнового вектора и $K^2=\beta^2+\gamma^2+\tau^2$.

Здесь также учтена симметрия тензора диэлектрической проницаемости анизотропной среды.

Как известно, система уравнений (5.12) будет иметь нетривиальное решение, если её определитель будет равен нулю

$$\det\begin{vmatrix} \left(\tau^2-K^2+\varepsilon_{11}\right) & \left(\beta\tau+\varepsilon_{12}\right) & \left(\gamma\tau+\varepsilon_{13}\right) \\ \left(\beta\tau+\varepsilon_{12}\right) & \left(\beta^2-K^2+\varepsilon_{22}\right) & \left(\beta\gamma+\varepsilon_{23}\right) \\ \left(\gamma\tau+\varepsilon_{13}\right) & \left(\beta\gamma+\varepsilon_{23}\right) & \left(\gamma^2-K^2+\varepsilon_{33}\right) \end{vmatrix} = 0 \qquad (5.13)$$

Полученное уравнение (5.13) в дальнейшем мы будем называть дисперсионным уравнением для плоской электромагнитной волны в анизотропной среде. Дисперсионным это уравнение названо потому, что оно определяет зависимость длины волнового вектора волны от её частоты (напомним, что $k_0 = \frac{\omega}{c}$). Иногда это уравнение называют уравнением Френеля.

Как уже отмечалось, тензор диэлектрической проницаемости анизотропной среды обладает свойством симметрии, т.е. $\varepsilon_{ij} = \varepsilon_{ji}$. Следовательно, мы всегда можем выбрать такую систему координат, в которой этот тензор будет иметь диагональный вид:

$$\hat{\varepsilon} = \begin{vmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{vmatrix}.$$

Такую систему координат назовем кристаллографической системой координат. В ней дисперсионное уравнение несколько упрощается и принимает вид:

$$\det \begin{vmatrix} \left(\varepsilon_1 + \tau^2 - K^2\right) & \beta\tau & \gamma\tau \\ \beta\tau & \left(\varepsilon_2 + \beta^2 - K^2\right) & \beta\gamma \\ \gamma\tau & \beta\gamma & \left(\varepsilon_3 + \gamma^2 - K^2\right) \end{vmatrix} = 0.$$

Раскрывая определитель, получим

$$\tau^2 \frac{\varepsilon_2 + \varepsilon_3 - K^2}{\varepsilon_2 \varepsilon_3} + \beta^2 \frac{\varepsilon_1 + \varepsilon_3 - K^2}{\varepsilon_1 \varepsilon_3} + \gamma^2 \frac{\varepsilon_1 + \varepsilon_2 - K^2}{\varepsilon_1 \varepsilon_2} = 1 \qquad (5.14)$$

Как видно из полученного выражения (5.14), дисперсионное уравнение описывает некоторую поверхность четвертого порядка в кристаллографической системе координат, которую описывает конец волнового вектора, будучи отложенным от начала координат во всех возможных направлениях. Назовем эту поверхность поверхностью волновых векторов. Чтобы ясно представить себе её вид, применим следующий приём.

Предположим, что волновой фронт электромагнитной волны расположен параллельно оси Ох. В этом случае никаких изменений фазы волны вдоль оси Ох не будет, следовательно, мы можем считать $\tau = 0$. Тогда для тех волн, волновой вектор которых лежит в плоскости yOz, дисперсионное уравнение (5.14) принимает вид:

$$\beta^2 \frac{\varepsilon_1 + \varepsilon_3 - K^2}{\varepsilon_1 \varepsilon_3} + \gamma^2 \frac{\varepsilon_1 + \varepsilon_2 - K^2}{\varepsilon_1 \varepsilon_2} = 1, \text{ где } K^2 = \beta^2 + \gamma^2.$$

После элементарных алгебраических преобразований получаем

$$\left(\varepsilon_1 - \beta^2 - \gamma^2\right) \cdot \left(\varepsilon_2 \beta^2 + \varepsilon_3 \gamma^2 - \varepsilon_2 \varepsilon_3\right) = 0 \qquad (5.15)$$

Таким же образом можно вычислить, что для волнового вектора, лежащего в плоскости XOZ ($\beta = 0$)

$$\left(\varepsilon_2 - \gamma^2 - \tau^2\right) \cdot \left(\varepsilon_1 \tau^2 + \varepsilon_3 \gamma^2 - \varepsilon_1 \varepsilon_3\right) = 0 \qquad (5.16)$$

И, наконец, для плоскости XOY, где $\gamma = 0$

$$\left(\varepsilon_3 - \beta^2 - \tau^2\right) \cdot \left(\varepsilon_2 \beta^2 + \varepsilon_1 \tau^2 - \varepsilon_1 \varepsilon_2\right) = 0 \qquad (5.17)$$

Из полученных выражений (5.15) – (5.17) видно, что след от сечения поверхности волновых векторов координатными плоскостями во всех случаях представляет собой две замкнутые кривые. Графически это представлено на рис. 11. Следовательно, для волны, распространяющейся в каком-либо заданном направлении, существует два различных значения волнового вектора, при которых дисперсионное уравнение (5.13) выполняется. Из рис. 11 также видно, что эти два значения волнового вектора совпадают, если волна распространяется вдоль направления **O.O**. Это направление называется *оптической осью* анизотропной среды. В рассмотренном нами случае имеются две оптические оси, расположенные симметрично в плоскости XOZ, но в общем случае их положение зависит от конкретных значений компонентов тензора диэлектрической проницаемости.

Поскольку два решения дисперсионного уравнения являются независимыми (т.е. собственными числами дисперсионного уравнения), то можно сказать, что вдоль выбранного направления в анизотропной среде могут распространяться две независимых электромагнитных волны, отличающиеся друг от друга значением фазовой скорости. Для определённости назовем эти две волны быстрой и медленной. Вдоль оптических осей среды скорости быстрой и медленной волн становятся одинаковыми и, поэтому, различия между ними исчезают.

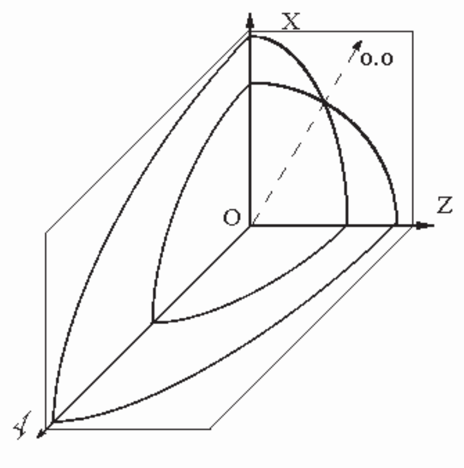

Рис. 11. Сечение поверхности волновых векторов координатными плоскостями.

Система координат XYZ совпадает с кристаллографической системой координат анизотропной среды.

Определив значения собственных волновых чисел для быстрой и медленной волн, распространяющихся вдоль произвольного направления в анизотропной среде, необходимо выяснить и их поляризации. Ранее мы выяснили, что поперечными векторами для электромагнитной волны в анизотропной среде являются $\vec{D}$ и $\vec{B}(\vec{H})$. Условимся, что под направлением поляризации волны будем понимать направление, определяемое вектором $\vec{D}$. Для того, чтобы определить собственные значения вектора индукции, необходимо подставить в волновое уравнение (5.12) собственные значения волнового вектора и решить это уравнение относительно проекций вектора напряженности электрического поля $\vec{E}$. Затем из уравнений Максвелла мы сможем вычислить искомые значения векторов $\vec{D}$ для быстрой и медленной волн, тем самым определив их поляризацию. Проделаем все эти операции в системе координат, совпадающей с кристаллографической. При этом пусть обе собственные волны имеют волновые векторы лежащие, например, в плоскости XOZ, то есть $\beta = 0$.

Волновое уравнение (5.12) для этого случая будет выглядеть следующим образом:

$$\begin{vmatrix} (\varepsilon_1 - \gamma^2) & 0 & \gamma\tau \\ 0 & (\varepsilon_2 - \gamma^2 - \tau^2) & 0 \\ \gamma\tau & 0 & (\varepsilon_3 - \tau^2) \end{vmatrix} \cdot \begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix} = 0 \qquad (5.18)$$

В общем случае система трех однородных линейных уравнений вида $\sum_{i=1}^{3} a_{ij} E_j = 0$ имеет решения в виде $E_j = cA_{kj}$, где A_{kj} - алгебраическое дополнение элемента a_{kj} [5], c – произвольная постоянная. Поскольку k изменяется от 1 до 3, мы имеем три различных набора проекций вектора $\vec{E}$ на оси координат. Однако все эти три набора описывают параллельные векторы и, поэтому, мы можем выбирать любой из них. На практике удобно пользоваться следующей таблицей, в которой выписаны все три решения для проекций вектора напряженности электрического поля волны:

Таблица 1.

	i = 1	i = 2	i = 3
$E_x =$	$a_{22}a_{33}-(a_{23})^2$	$a_{13}a_{23}-a_{12}a_{33}$	$a_{12}a_{23}-a_{13}a_{22}$
$E_y =$	$a_{13}a_{23}-a_{12}a_{33}$	$a_{11}a_{33}-(a_{13})^2$	$a_{12}a_{13}-a_{11}a_{23}$
$E_z =$	$a_{12}a_{23}-a_{13}a_{22}$	$a_{12}a_{13}-a_{11}a_{23}$	$a_{11}a_{22}-(a_{12})^2$

Учитывая конкретный вид детерминанта системы (5.18), получим следующие значения для проекций вектора $\vec{E}$:

$$\vec{E} = \left\{ \begin{matrix} \left(\varepsilon_2 - \gamma^2 - \tau^2\right)\cdot\left(\varepsilon_3 - \tau^2\right) & 0 & -\gamma\tau\left(\varepsilon_2 - \gamma^2 - \tau^2\right) \\ 0 & \varepsilon_1\varepsilon_3 - \varepsilon_3\gamma^2 - \varepsilon_1\tau^2 & 0 \\ -\gamma\tau\left(\varepsilon_2 - \gamma^2 - \tau^2\right) & 0 & \left(\varepsilon_1 - \gamma^2\right)\cdot\left(\varepsilon_2 - \gamma^2 - \tau^2\right) \end{matrix} \right\} \quad (5.19)$$

Собственные значения для постоянных распространения обеих волн мы уже получили ранее – это есть решения уравнения (5.16): $\tau_1^2 = \varepsilon_2 - \gamma_1^2$ и $\tau_2^2 = \dfrac{\varepsilon_3\left(\varepsilon_1 - \gamma_2^2\right)}{\varepsilon 1}$. Подставляя эти решения в (5.19), выясним, что собственные волны в данном случае имеют следующие проекции вектора $\vec{E}$:

$$\vec{E}_1 = \left\{ \begin{matrix} 0, & 0, & 0 \\ 0, & \varepsilon_1\left(\varepsilon_3 - \varepsilon_2\right) - \gamma_1^2\left(\varepsilon_3 - \varepsilon_1\right), & 0 \\ 0, & 0, & 0 \end{matrix} \right\} \begin{matrix} (i=1) \\ (i=2) \\ (i=3) \end{matrix}$$

Обозначим $\overline{\varepsilon} = \varepsilon_1\left(\varepsilon_3 - \varepsilon_2\right) - \gamma_2^2\left(\varepsilon_3 - \varepsilon_1\right)$, тогда

$$\vec{E}_2 = \begin{Bmatrix} -\dfrac{\varepsilon_3\gamma_2^2}{\varepsilon_1}\cdot\dfrac{\overline{\varepsilon}}{\varepsilon_1}, & 0, & \gamma_2^2\tau_2\dfrac{\overline{\varepsilon}}{\varepsilon_1} \\ 0, & 0, & 0 \\ \gamma_2^2\tau_2\dfrac{\overline{\varepsilon}}{\varepsilon_1}, & 0, & -\left(\varepsilon_1-\gamma_2^2\right)\cdot\dfrac{\overline{\varepsilon}}{\varepsilon_1} \end{Bmatrix} \quad \begin{matrix} (i=1) \\ (i=2) \\ (i=3) \end{matrix}$$

Итак, мы выяснили, что компоненты одной из собственных волн $\vec{E}$ равны $\{0, E_y, 0\}$, в то время как компоненты другой равны $\{E_x, 0, E_z\}$.

Исключая, как и ранее, из уравнений Максвелла (5.7)(1.7) векторы магнитного поля, можно вычислить вектор индукции $\vec{D}$ через известный вектор напряженности $\vec{E}$:

$$\left[\vec{K}\times\left[\vec{K}\times\vec{E}\right]\right] = -\vec{D},$$

или, раскрывая двойное векторное произведение,

$$\begin{matrix} D_x & \gamma(\gamma E_x - \tau E_z) - \beta(\tau E_y - \beta E_x) \\ D_y = & \tau(\tau E_y - \beta E_x) - \gamma(\beta E_z - \gamma E_y). \\ D_z & \beta(\beta E_z - \gamma E_y) - \tau(\gamma E_x - \tau E_z) \end{matrix}$$

Учитывая, что в рассматриваемом нами случае $\beta = 0$, получим компоненты вектора индукции:

$$D_x = \gamma(\gamma E_x - \tau E_z)$$
$$D_y = K^2 E_y$$
$$D_Z = \tau(\tau E_z - \gamma E_x)$$

Таким образом, у собственной волны с индексом 1 вектор индукции имеет компоненты $\vec{D}_1 = \{0, D_y, 0\}$, а у волны с индексом 2 $\vec{D}_2 = \{D_x, 0, D_y\}$. Отсюда видно, что эти векторы ортогональны. Кроме того, поскольку проекции волнового вектора β, γ и τ в наших рассуждениях есть величины действительные, то и поляризации обеих волн являются линейными.

Аналогичные результаты можно получить, рассмотрев и два оставшихся варианта направлений распространения собственных волн: $\gamma = 0$ и $\tau = 0$.

Подводя итог, сформулируем весьма важное свойство электромагнитных волн, распространяющихся в анизотропной среде – в каждом заданном направлении в анизотропной среде могут распространять две электромагнитных волны с различными фазовыми скоростями, обладающие ортогональными линейными поляризациями.

Из этого вытекает ещё одно существенное различие в процессе распространения плоских волн в изотропных и анизотропных средах. Явление преломления плоской волны, падающей на поверхность анизотропной среды, принципиально отличается от аналогичного случая на границе двух изотропных сред. Действительно, законы преломления по-прежнему получаются из условия равенства тангенциальной составляющей волнового вектора волны на границе раздела, но в анизотропной среде у волны есть два различных значения волнового вектора, для каждого из них необходимо выполнение этого равенства. Удовлетворить этому требованию можно лишь в том случае, если в анизотропной среде будет две преломленных волны, распространяющихся в различных направлениях. Это явление получило название двойного лучепреломления.

5.4. Электромагнитные волны в одноосной анизотропной среде

До сих пор мы рассматривали анизотропную среду, тензор диэлектрической восприимчивости которой в кристаллографической системе координат тремя главными числами ε_1, ε_2 и ε_3. Однако довольно широкий класс реально существующих анизотропных сред обладает тензором диэлектрической проницаемости, в котором $\varepsilon_1 = \varepsilon_2 = \varepsilon_o$ а $\varepsilon_3 = \varepsilon_e$:

$$\hat{\varepsilon} = \begin{vmatrix} \varepsilon_o & 0 & 0 \\ 0 & \varepsilon_o & 0 \\ 0 & 0 & \varepsilon_e \end{vmatrix} \qquad (5.20)$$

(смысл употребляемых здесь индексов будет пояснен немного позднее). Такие среды принято называть одноосными в отличие от тех сред, которые мы рассматривали выше. Заранее можно сделать вывод, что оптическая ось в таких средах единственная; ниже это будет доказано более строго. Вид тензора диэлектрической проницаемости, определяемый всего двумя числами, позволяет существенно упростить приведенные выше выкладки.

И так, начнем по порядку. Волновое уравнение для одноосной среды, что естественно, останется прежним (5.11) либо (5.12), а дисперсионное уравнение (5.14) упростится:

$$\left(K^2 - \varepsilon_o\right)\cdot\left(\varepsilon_o\beta^2 + \varepsilon_e\gamma^2 + \varepsilon_o\tau^2 - \varepsilon_o\varepsilon_e\right) = 0 \qquad (5.21)$$

Теперь мы, уже не прибегая к методу сечений, сразу можем сказать, что поверхность волновых векторов для одноосной среды представляет собой для одной из волн шар радиуса $\sqrt{\varepsilon_o}$, а для другой – эллипсоид вращения с полуосями $\sqrt{\varepsilon_o}$ и $\sqrt{\varepsilon_e}$. Заметим, что фазовая скорость первой волны не зависит от направления распространения, поэтому такую волну принято называть обыкновенной. Фазовая скорость второй волны по-прежнему зависит от направления распространения, эту волну обычно называют необыкновенной. Любая из этих волн может быть быстрой либо медленной, это определяется конкретными значениями ε_o и ε_e. В зависимости от соотношения между этими величинами одноосные анизотропные среды принято подразделять на отрицательные ($\varepsilon_o > \varepsilon_e$) и положительные ($\varepsilon_e > \varepsilon_o$). В первом случае обыкновенная волна будет медленной, а необыкновенная – быстрой. Во втором случае всё наоборот. Сечение поверхности волновых векторов плоскостью XOZ показано на рис. 12.

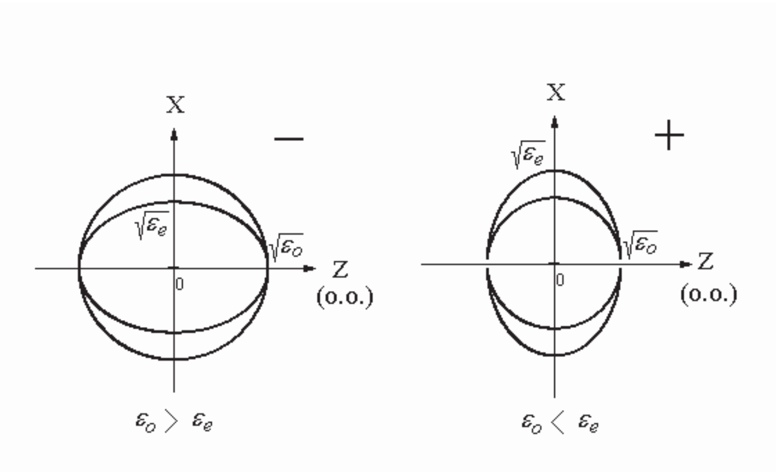

Рис. 12. Отрицательная и положительная анизотропная среда.

Оптическая ось, как это видно из рисунка, теперь только одна и совпадает с одной из кристаллографических осей анизотропной среды (в данном случае с OZ). Все сечения поверхности волновых векторов, перпендикулярные оптической оси представляют собой окружности.

Определим поляризации обыкновенной и необыкновенной волн. Для этого подставим в волновое уравнение (5.12) значения собственных волновых векторов этих волн. Для обыкновенной волны $K^2 = \varepsilon_o$ и (5.12) приобретает вид:

$$\begin{vmatrix} \tau_o^2 & \beta_o\tau_o & \gamma_o\tau_o \\ \beta_o\tau_o & \beta_o^2 & \beta_o\gamma_o \\ \gamma_o\tau_o & \beta_o\gamma_o & \left(\gamma_o^2 + \Delta\varepsilon\right) \end{vmatrix} \cdot \begin{vmatrix} E_{ox} \\ E_{oy} \\ E_{oZ} \end{vmatrix} = 0 \tag{5.22}$$

Пользуясь таблицей 1, запишем

$$\vec{E}_o = \left\{ \begin{matrix} \beta_o^2\Delta\varepsilon & -\beta_o\tau_o\Delta\varepsilon & 0 \\ -\beta_o\tau_o\Delta\varepsilon & \tau_o^2\Delta\varepsilon & 0 \\ 0 & 0 & 0 \end{matrix} \right\} \begin{matrix} i=1 \\ i=2 \\ i=3 \end{matrix}.$$

Мы получили, что вектор напряженности электрического поля обыкновенной волны имеет проекции $\vec{E}_o = \left\{E_{ox},\ E_{oy},\ 0\right\}$. Проекции вектора индукции обыкновенной волны можно определить из известного нам уравнения $\left[\vec{K}\times\left[\vec{K}\times\vec{E}\right]\right] = -\vec{D}$, откуда следует, что

$$\vec{D}_o = \begin{Bmatrix} \beta_o^2\Delta\varepsilon\varepsilon_o, & -\beta_o\tau_o\Delta\varepsilon\varepsilon_o, & 0 \\ -\beta_o\tau_o\Delta\varepsilon\varepsilon_o, & \tau_o^2\Delta\varepsilon\varepsilon_o, & 0 \\ 0, & 0, & 0 \end{Bmatrix} \begin{matrix} i=1 \\ i=2 \\ i=3 \end{matrix} \qquad (5.23)$$

Отсюда можно заключить, что поляризация обыкновенной воны в одноосной среде всегда перпендикулярна оптической оси этой среды.

Волновое уравнение для необыкновенной волны получается аналогичным образом:

$$\begin{vmatrix} \dfrac{\varepsilon_o\tau_e^2 + \Delta\varepsilon\left(\gamma_e^2 - \varepsilon_o\right)}{\varepsilon_o} & \tau_e\beta_e & \tau_e\gamma_e \\ \tau_e\beta_e & \dfrac{\varepsilon_o\beta_e^2 + \Delta\varepsilon\left(\gamma_e^2 - \varepsilon_o\right)}{\varepsilon_o} & \beta_e\gamma_e \\ \tau_e\gamma_e & \beta_e\gamma_e & \varepsilon_e\dfrac{\gamma_e^2}{\varepsilon_o} \end{vmatrix} \cdot \begin{vmatrix} E_{ex} \\ E_{ey} \\ E_{ez} \end{vmatrix} = 0 \qquad (5.24)$$

Это уравнение имеет следующие решения для компонент вектора напряженности электрического поля необыкновенной волны:

$$\vec{E}_e = \begin{Bmatrix} -\gamma_e^2\tau_e^2\dfrac{\Delta\varepsilon}{\varepsilon_o}, & -\beta_e\gamma_e^2\tau_e\dfrac{\Delta\varepsilon}{\varepsilon_o}, & \gamma_e\tau_e\left(\varepsilon_o - \gamma_e^2\right)\dfrac{\Delta\varepsilon}{\varepsilon_o} \\ -\beta_e\gamma_e^2\tau_e\dfrac{\Delta\varepsilon}{\varepsilon_o}, & -\beta_e^2\gamma_e^2\dfrac{\Delta\varepsilon}{\varepsilon_o}, & \beta_e\gamma_e\left(\varepsilon_o - \gamma_e^2\right)\dfrac{\Delta\varepsilon}{\varepsilon_o} \\ \gamma_e\tau_e\left(\varepsilon_o - \gamma_e^2\right)\dfrac{\Delta\varepsilon}{\varepsilon_o}, & \beta_e\gamma_e\left(\varepsilon_o - \gamma_e^2\right)\dfrac{\Delta\varepsilon}{\varepsilon_o}, & -\left(\varepsilon_o - \gamma_e^2\right)^2\dfrac{\Delta\varepsilon}{\varepsilon_o} \end{Bmatrix} \begin{matrix} i=1 \\ i=2 \\ i=3 \end{matrix} \qquad (5.25)$$

Соответственно, компоненты вектора индукции будут иметь вид

$$\vec{D}_e = \left\{ \begin{matrix} -\gamma_e^2\tau_e^2\Delta\varepsilon, & -\beta_e\gamma_e^2\tau_e\Delta\varepsilon, & \gamma_e\tau_e\left(\beta_e^2+\tau_e^2\right)\Delta\varepsilon \\ -\beta_e\gamma_e^2\tau_e\Delta\varepsilon, & -\beta_e^2\gamma_e^2\Delta\varepsilon, & \beta_e\gamma_e\left(\beta_e^2+\tau_e^2\right)\Delta\varepsilon \\ \gamma_e\tau_e\left(\varepsilon_o-\gamma_e^2\right)\Delta\varepsilon, & \gamma_e\beta_e\left(\varepsilon_o-\gamma_e^2\right)\Delta\varepsilon, & -\left(\varepsilon_o-\gamma_e^2\right)\left(\beta_e^2+\tau_e^2\right)\Delta\varepsilon \end{matrix} \right\} \begin{matrix} i=1 \\ i=2 \\ i=3 \end{matrix} \quad (5.26)$$

Ранее мы выдвигали утверждение о том, что собственные волны анизотропной среды имеют ортогональные поляризации. Проверим, когда выполняется это условие для одноосной анизотропной среды. Если векторы $\vec{D}_e$ и $\vec{D}_o$ ортогональны, то их скалярное произведение должно быть равным нулю. Зная проекции этих векторов на оси кристаллографической системы координат (5.23) и (5.26), можем записать:

$$\vec{D}_e \cdot \vec{D}_o = \begin{matrix} -\beta_o\gamma_e^2\tau_e\Delta\varepsilon^2\varepsilon_o\left(\beta_o\tau_e-\beta_e\tau_o\right) & i=1 \\ \beta_e\gamma_e^2\tau_o\Delta\varepsilon^2\varepsilon_o\left(\beta_o\tau_e-\beta_e\tau_o\right) & i=2 \end{matrix} \quad (5.27)$$

Из (5.27) видно, что поляризации собственных волн ортогональны когда, в частности, выполняется $\beta_o\tau_e=\beta_e\tau_o$ или $\dfrac{\beta_e}{\beta_o}=\dfrac{\tau_e}{\tau_o}$.

Это равенство выполняется в единственном случае, когда направления распространения этих волн совпадают. Кроме этого, ортогональность поляризаций имеет место при распространении волн в одной из кристаллографических плоскостей, где одна из проекций волнового вектора $\beta_{o(e)},\gamma_e,\tau_{o(e)}$ равна нулю.

5.5. Произвольная ориентация кристаллографических осей анизотропной среды

Все предыдущие рассуждения мы проводили, предполагая, что система координат совпадает с кристаллографической. Однако для многих практических задач часто оказывается удобной не кристаллографическая, а какая-то другая декартова система, которую назовем лабораторной системой координат. Поскольку обе системы ортогональные, то в общем случае лабораторная система оказывается повернутой на произвольный угол относительно кристаллографической системы. Очевидно, что тензор диэлектрической проницаемости анизотропной среды в лабораторной системе уже не будет диагональным. Попытаемся вычислить его компоненты, для чего введем три единичных вектора $\vec{a}^0$, $\vec{b}^0$, $\vec{c}^0$, которые направлены по осям кристаллографической системы координат. Другие три единичных вектора $\vec{x}^0, \vec{y}^0, \vec{z}^0$ направим по осям лабораторной системы координат. Тогда эти две системы координат будут связаны между собой как:

$$\begin{aligned} \vec{x}^0 &= \alpha_1\vec{a}^0 + \alpha_2\vec{b}^0 + \alpha_3\vec{c}^0 \\ \vec{y}^0 &= \beta_1\vec{a}^0 + \beta_2\vec{b}^0 + \beta_3\vec{c}^0 \\ \vec{z}^0 &= \gamma_1\vec{a}^0 + \gamma_2\vec{b}^0 + \gamma_3\vec{c}^0 \end{aligned} \tag{5.28}$$

где $\alpha_1 = \vec{x}^0 \cdot \vec{a}^0$ - косинус угла между осями OX в лабораторной и кристаллографической системах координат, $\alpha_2 = \vec{x}^0 \cdot \vec{b}^0$ - косинус угла между осью OX в лабораторной системе и осью OY в кристаллографической системе координат и т.д. Девять направляющих косинусов связаны между собой условием ортогональности:

$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$, $\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 = 0$ и т.д.

Матрица, составленная из направляющих косинусов, называется матрицей поворота

$$\alpha_{ij} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix} \tag{5.29}$$

Произвольный поворот одной декартовой системы координат относительно другой, так же декартовой, всегда можно представить как три последовательных поворота вокруг каждой из осей первой системы координат. Поэтому всегда полезно знать матрицы поворота вокруг каждой из осей. Для примера приведем вид матрицы поворота на угол ϕ

вокруг оси OX (смотри рис. 13.), причем поворот осуществляется по часовой стрелке (если смотреть в направлении оси OX):

$$\alpha_{ij}^{OX} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \tag{5.30}$$

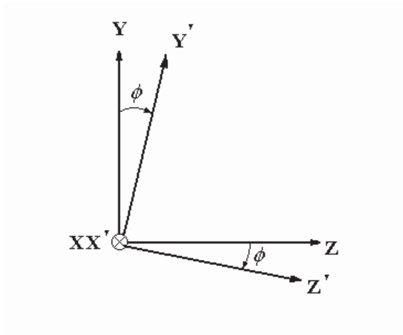

Рис. 13. Поворот по часовой стрелке системы координат X'Y'Z' вокруг оси OX системы координат XYZ на угол ϕ.

Матрицы поворота вокруг двух других координатных осей записываются аналогично.

Хорошо известно [5], что компоненты произвольного вектора $\vec{V}$ при переходе из одной системы координат в другую преобразуются по закону:

$V_i = \alpha_{ij} V_j$.

С другой стороны комбинация девяти чисел составляет тензор только в том случае, если эти числа при поворотах системы координат преобразовываются так же, как тензорное произведение двух векторов [5]. Например, элемент ε_{ab} должен преобразовываться по тому же правилу, что и произведение компонент векторов $u_a v_b$. Следовательно, тензор при повороте системы координат должен преобразовываться по закону:

$$\varepsilon_{ij} = \sum_k \sum_l \alpha_{ik} \alpha_{jl} \varepsilon_{kl}, \quad \text{или} \quad \varepsilon_{ij} = \sum_k \sum_l \alpha_{ik} \varepsilon_{kl} \alpha_{lj}^T \tag{5.31}$$

где значок «Т» означает транспонирование.

Поясним эти преобразования на примере. Описанный выше поворот системы координат вокруг оси ОХ на угол ϕ (5.30) преобразует диагональный тензор к виду:

$$\varepsilon_{ij} = \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \lambda_3 \end{vmatrix} \cdot \begin{vmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{vmatrix} \cdot \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{vmatrix} \cdot \begin{vmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{vmatrix} =$$

$$= \begin{vmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2\cos^2\phi + \varepsilon_3\sin^2\phi & (\varepsilon_2 - \varepsilon_3)\sin\phi\cos\phi \\ 0 & (\varepsilon_2 - \varepsilon_3)\sin\phi\cos\phi & \varepsilon_2\sin^2\phi + \varepsilon_3\cos^2\phi \end{vmatrix}$$

Таким образом, последовательно применяя преобразования поворота, можно выразить компоненты тензора диэлектрической проницаемости в системе координат, повернутой на произвольный угол относительно кристаллографической системы координат.

Тема 6. Дифракция электромагнитных волн

Использованное ранее понятие плоской электромагнитной волны является в известном смысле идеализацией. В реальной ситуации, когда фазовый фронт электромагнитной волны чем-либо ограничен, возникает явление ***дифракции*** волны на препятствиях, обеспечивающих это ограничение фазового фронта. В качестве примера можно привести отверстие конечных размеров в бесконечном непрозрачном для волны экране. Вблизи краев этого отверстия и будет возникать дифракция электромагнитной волны, суть которой проявляется в том, что часть энергии электромагнитной волны будет проникать в область геометрической тени (см. рис 14).

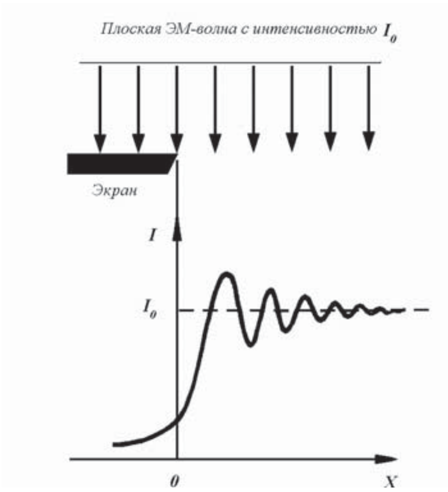

Рис.14. Поперечное распределение интенсивности плоской электромагнитной волны после прохождение через отверстие в непрозрачном экране.

Если под электромагнитной волной понимать световую электромагнитную волну, то в понятие дифракции включают все случаи отклонения направления распространения света от прямолинейного, не обусловленные его преломлением или отражением.

6.1. Дифракционный интеграл Кирхгофа

Поляризация световой волны играет важную роль при отражении и преломлении световых волн; при рассмотрении дифракции ее роль можно не учитывать, поскольку длина световой волны предполагается много меньшей любого линейного размера препятствия, на которое эта волна падает. Исходя из этого, при описании процесса дифракции можно пользоваться скалярным волновым уравнением в форме:

$$\nabla^2\psi = \varepsilon\mu\frac{\partial^2\psi}{\partial t^2} \tag{6.1}$$

где $\frac{1}{\sqrt{\varepsilon\mu}}$ - скорость света в среде, а ψ - любая из компонент электромагнитного поля волны.

Напомним, что оператор ∇ является вектором с координатами $\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}$.

Ограничим рассмотрение случаем монохроматической электромагнитной волны с зависимостью от времени вида $\exp(i\omega t)$, тогда, после дифференцирования по времени, скалярное волновое уравнение примет вид:

$$\nabla^2\psi + k^2\psi = 0 \tag{6.2}$$

где $k = \frac{2\pi}{\lambda} = \omega\sqrt{\varepsilon\mu}$ есть модуль волнового вектора или волновое число в среде с проницаемостями ε и μ.

Поскольку в дальнейшем мы будем рассматривать в основном немагнитные среды, можно считать $\mu = 1$.

Уравнение (6.2) известно как уравнение Гельмгольца. Его решения существенно упрощается при использовании δ-функции Дирака. Предполагается, что читатель знаком с ее свойствами. Поэтому рассмотрим, наряду с (6.2), уравнение вида

$$\nabla^2 G + k^2 G = \delta(\vec{r} - \vec{r}') \tag{6.3}$$

Дельта-функция векторного аргумента может быть представлена в виде

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z') \tag{6.4}$$

Функция G, известная как функция Грина, зависит от двух наборов переменных:

$$G = G(x, y, z, x', y', z') \tag{6.5}$$

Решение уравнения Гельмгольца можно получить следующим образом [6]. Умножим обе его части на G, а обе части (6.3) на ψ и вычтем одно уравнение из другого. После интегрирования по некоторому объему V получим:

$$\psi(x', y', z') = \int_V \left(\psi \nabla^2 G - G \nabla^2 \psi\right) dV \tag{6.6}$$

Правая часть получена в результате интегрирования произведения $\delta(\vec{r} - \vec{r}')\psi$. Подынтегральное выражение можно переписать в виде

$$\psi \nabla^2 G - G \nabla^2 \psi = \nabla \cdot (\psi \nabla G - G \nabla \psi) \tag{6.7}$$

Используя известную теорему Стокса, преобразующую интеграл по объему в интеграл по поверхности, искомое решение уравнения Гельмгольца может быть записано в виде:

$$\psi(x', y', z') = \int_S \left(\psi \frac{\partial G}{\partial n} - G \frac{\partial \psi}{\partial n}\right) dS \tag{6.8}$$

Интегрирование проводится по замкнутой поверхности S, для которой n – направление внешней нормали. Точка с координатами (x', y', z'), в которой вычисляется функция ψ, расположена внутри замкнутой поверхности S (см. рис. 15). Выражение (6.8) принято называть *дифракционным интегралом Кирхгофа*.

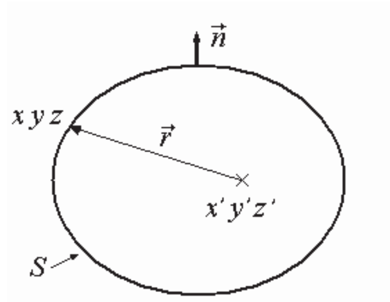

Рис.15 Геометрическая иллюстрация представления интеграла (6.8). Поверхность *S* окружает точку наблюдения $x'y'z'$. Единичный вектор $\vec{n}$ нормален к поверхности *S*. Точка *xyz* принадлежит поверхности *S*.

Таким образом, для вычисления поля в точке наблюдения необходимо знать значение функции ψ на некоторой поверхности S, охватывающей эту точку наблюдения. Кроме того, необходимо знать функцию Грина. Для вычисления этой функции обратимся к уравнению (6.3), одним из важных решений которого является *функция Грина для свободного пространства*, которую иногда называют *функцией точечного источника*. Эта функция фактически описывает волну единичной амплитуды со сферическим фазовым фронтом:

$$G = -\frac{1}{4\pi}\frac{\exp(-ikr)}{r} \qquad (6.9)$$

где $r = \sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2}$.

Подставляя функцию Грина для свободного пространства в дифракционный интеграл, имеем:

$$\psi(x',y',z')=\frac{1}{4\pi}\int_S\left[\frac{\partial\psi}{\partial n}\frac{\exp(-ikr)}{r}-\psi\frac{\partial}{\partial n}\left(\frac{\exp(-ikr)}{r}\right)\right]dS \qquad (6.10)$$

Для практических расчетов часто полезно использовать некоторые приближения дифракционного интеграла (6.10). Первое приближение, которое почти всегда справедливо для оптических задач, основано на малости длины волны света по сравнению с расстоянием до точки наблюдения r. Это приближение математически можно выразить следующим образом:

$$k \gg \frac{1}{r} \qquad (6.11)$$

Тогда можно пренебречь производной от $1/r$ по сравнению с производной от $\exp(-ikr)$ и записать приближенное значение дифракционного интеграла:

$$\psi(x',y',z')\approx\frac{1}{4\pi}\int_S\left(\frac{\partial\psi}{\partial n}+ik\frac{\partial r}{\partial n}\psi\right)\frac{\exp(-ikr)}{r}dS \qquad (6.12)$$

Следующее приближение состоит в том, чтобы получить более простое для r путем разложения его в ряд Тейлора. Это возможно в тех случаях, когда размеры дифракционного препятствия x, y малы по сравнения с расстоянием до точки наблюдения z, т.е. $|x'-x| \ll |z'-z|$ и $|y'-y| \ll |z'-z|$. Тогда

$$r = z'-z+\frac{1}{2}\frac{(x'-x)+(y'-y)}{z'-z}+\ldots \qquad (6.13)$$

Поясним сделанные приближения на простейшем примере дифракционной задачи, а именно – рассмотрим дифракцию света на щели в бесконечном, плоском непрозрачном экране, геометрия которой представлена на рис. 16. Пусть на этот экран слева падает под углом γ плоская световая волна. В качестве поверхности S выберем плоскость экрана.

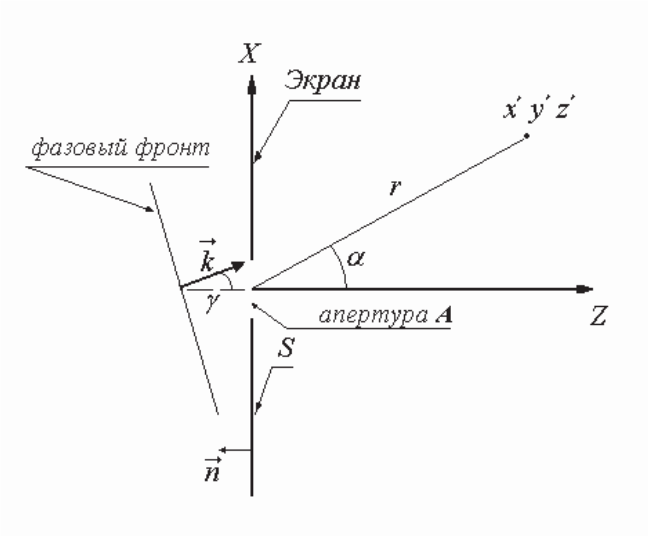

Рис. 16. Дифракция плоской волны на щели в бесконечном непрозрачном экране.

В соответствии с выбранной геометрией

$$\frac{\partial}{\partial n} = -\frac{\partial}{\partial z},$$

$$\frac{\partial \psi}{\partial n} = -\frac{\partial \psi}{\partial z} = ik_z\psi = ik\psi\cos\gamma,.$$

$$\frac{\partial r}{\partial n} = -\frac{\partial r}{\partial z} = \frac{z'-z}{r} = \cos\alpha.$$

Дифракционный интеграл в форме (6.12) в этом случае будет выглядеть как:

$$\psi(x',y',z') = \frac{ik}{4\pi}\int_A (\cos\gamma + \cos\alpha)\psi(x,y,z)\frac{\exp(-ikr)}{r}dS \qquad (6.14)$$

Поверхность S, по которой производится интегрирование, включает в себя экран, показанный на рис. 16 и бесконечную полусферу, ограничивающую пространство справа от экрана. Поскольку поле на бесконечной полусфере в силу условия излучения равно нулю, то эту часть поверхности S можно исключить из области интегрирования. Значения функции ψ и ее нормальной производной на экране очень малы и могут быть измерены только в непосредственной близости от краев щели, поэтому их вкладом в интеграл можно пренебречь в первом приближении. Исходя из этих положений, область интегрирования можно ограничить лишь апертурой А.

Во многих случаях апертуру щели в экране можно считать малой по сравнению с расстоянием до точки наблюдения. Если же при этом последняя настолько близка к оси Z, что $\cos\alpha \approx 1$, то, при нормальном падении плоской волны на экран ($\cos\gamma = 1$), дифракционный интеграл можно считать равным

$$\psi\left(x', y', z'\right) = \frac{i}{\lambda}\int\limits_A \psi\left(x, y, z\right)\frac{\exp\left(-ikr\right)}{r}dS \qquad (6.15)$$

Физический смысл интеграла (6.15) достаточно ясен. Каждая точка отверстия в экране может рассматриваться как источник сферической волны (смотрите (6.9)). Амплитуда источника определяется амплитудой падающей волны в этой точке отверстия. Поле в точке наблюдения $x'y'z'$ образуется в результате суперпозиции всех сферических волн, достигающих этой точки. Данная трактовка известна как *принцип Гюйгенса*.

Возвращаясь к (6.14), используем указанное ранее приближение (6.13) Здесь следует заметить, что, прежде чем пользоваться этим приближением, необходимо удостовериться, что произведение k на член третьего порядка в (6.13) действительно много меньше единицы. При выполнении этого условия дифракционный интеграл можно приближенно представить в виде

$$\begin{aligned}\psi(x',y',z') &= \frac{i}{2\lambda}\exp(-ikz')\times \\ &\times\int_A(\cos\gamma+\cos\alpha)\frac{\psi}{z'-z}\exp\left[ik\left(z-\frac{(x'-x)^2+(y'-y)^2}{2(z'-z)}\right)\right]dS\end{aligned} \tag{6.16}$$

Дифракционный интеграл в таком виде наиболее часто используется на практике. Обычно различают два случая.

При вычислении поля достаточно близко к отверстию, так близко, что член x^2+y^2 в показателе степени экспоненты следует принимать во внимание, имеет место *дифракция Френеля*.

Если же этот член пренебрежимо мал, то следует говорить о *дифракции Фраунгофера*.

Таким образом, можно сделать следующий вывод. Дифракция Фраунгофера имеет место вдали от самого предмета, вызывающего дифракцию Дифракция Френеля проявляется также в областях, достаточно удаленных от источника дифракции, таких, чтобы приближение (6.16) выполнялось, но все же не настолько далеких, чтобы можно было говорить о дифракции Фраунгофера.

6.2. Двумерная дифракция.

Рассмотрение задачи о дифракции иногда значительно упрощается, если вместо трехмерной задачи искать решение только в двухмерном случае. Исследования дифракционных интегралов, проведенные выше, и в частности важная аппроксимация (6.16), основаны на разложении поля падающей волны на непрерывный набор источников сферических волн. Сферические волны своим появлением обязаны именно трехмерной постановке задачи. Для двухмерных задач эквивалентом сферических волн являются волны с цилиндрическим фазовым фронтом. Для двухмерных задач характерно, что все полевые переменные, так же как и геометрия пространства, не зависят от одной из пространственных координат. Символически это можно выразить с помощью соотношения $\frac{\partial}{\partial y}=0$.

В двухмерном случае дифракционный интеграл (6.8) можно преобразовать к виду:

$$\psi(x',y')=\int_C\left(\psi\frac{\partial G}{\partial n}-G\frac{\partial\psi}{\partial n}\right)dC \tag{6.17}$$

Здесь была использована двумерная интегральная теорема о преобразовании интеграла по поверхности в интеграл по замкнутой кривой, охватывающей эту поверхность. По-прежнему, направление n совпадает с направлением внешней нормали к кривой C. Функция Грина для свободного двумерного пространства также хорошо известна и представляется с помощью функции Ханкеля второго рода:

$$G(x,x',z,z')=\frac{i}{4}H_0^{(2)}(kr) \tag{6.18}$$

Характерной особенностью функции Ханкеля является то, что при малых длинах волн λ ее аргумент kr всегда большой. Это позволяет использовать асимптотическое выражение для функции Ханкеля при большом аргументе

$$H_0^{(2)}(kr)=\sqrt{\frac{2}{\pi kr}}\exp\left[-i\left(kr-\frac{\pi}{4}\right)\right] \tag{6.19}$$

Используя эту асимптотику, можно дифракционный интеграл (6.10) переписать для двухмерной геометрии

$$\psi(x',y')=\frac{1}{\sqrt{8\pi k}}\exp(-i\frac{\pi}{4})\int_C\left[\frac{\partial\psi}{\partial n}\frac{\exp(-ikr)}{\sqrt{r}}-\psi\frac{\partial}{\partial n}\left(\frac{\exp(-ikr)}{\sqrt{r}}\right)\right]dC \tag{6.20}$$

Здесь $r=\sqrt{(z'-z)^2+(x'-x)^2}$.

Из сравнения (6.10) и (6.20) видно, что амплитуда цилиндрических волн уменьшается с расстоянием как $r^{-1/2}$ вместо r^{-1}, что имело место в случае сферических волн. И, наконец, важное для нас приближенное значение дифракционного интеграла (6.16) в случае двумерной геометрии приобретает более простой вид::

$$\psi(x',z') = \frac{1}{2\sqrt{\lambda}}\exp[-i(kz' - \frac{\pi}{4})]\times \\ \times\int_C \frac{\cos\gamma + \cos\alpha}{\sqrt{z'-z}}\psi(x,z)\exp[ik(z - \frac{(x'-x)^2}{2(z'-z)})]dC \qquad (6.21)$$

Посмотрим, как в двумерном случае будет выглядеть полученное нами ранее решение задачи о дифракции на щели в непрозрачном экране. Предполагая, что амплитуда падающей волны на щели везде одинакова (ψ_0), используя приближения $\cos\gamma = 1$(нормальное падение волны) и $\cos\alpha \approx 1$ в зоне Фраунгофера можно записать

$$\psi(x',z') = \frac{\exp[-i(kz' - \frac{\pi}{4})]}{\sqrt{\lambda}}\frac{\psi_0}{\sqrt{z'}}\exp(-i\frac{kx'^2}{2z'})\int_{-\frac{d}{2}}^{\frac{d}{2}}\exp\left[ik\left(\frac{x'}{z'}\right)x\right]dx \qquad (6.22)$$

Так как мы уже предположили, что $\cos\alpha \approx 1$,или, с другой стороны, $x' << z'$, то можно заменить z' на r везде, кроме показателя экспоненты, куда входит произведение малой величины $\frac{x'}{z'}$ и большого числа k. Но, тем не менее, в этом показателе экспоненты, стоящей перед интегралом, мы вправе использовать приближение $r = z' + \frac{x'^2}{2z'}$. В показателе экспоненты, стоящей под интегралом, в рамки сделанных приближений укладывается замена

$$\alpha \approx tg\alpha = \frac{x'}{z'} \approx \frac{x'}{r} \qquad (6.23)$$

Тогда, проводя интегрирование в (1.22), получим

$$\psi(x',z') = \psi(r,\alpha) = \psi_0 \frac{2\exp\left[-i(kr - \frac{\pi}{4})\right]}{k\sqrt{\lambda r}} \frac{\sin(\pi \frac{d}{\lambda}\alpha)}{\alpha} \qquad (6.24)$$

Дифракционное поле (6.24) имеет форму цилиндрической волны, спадающей с расстоянием как $r^{-\frac{1}{2}}$. Эта волна промодулирована по углу множителем $\frac{\sin(\pi \frac{d}{\lambda}\alpha)}{\alpha}$, имеющим вид функции $\frac{\sin x}{x}$. Другими словами, проходя через узкую щель, свет рассеивается в стороны тем больше, чем уже щель или чем больше длина волны.

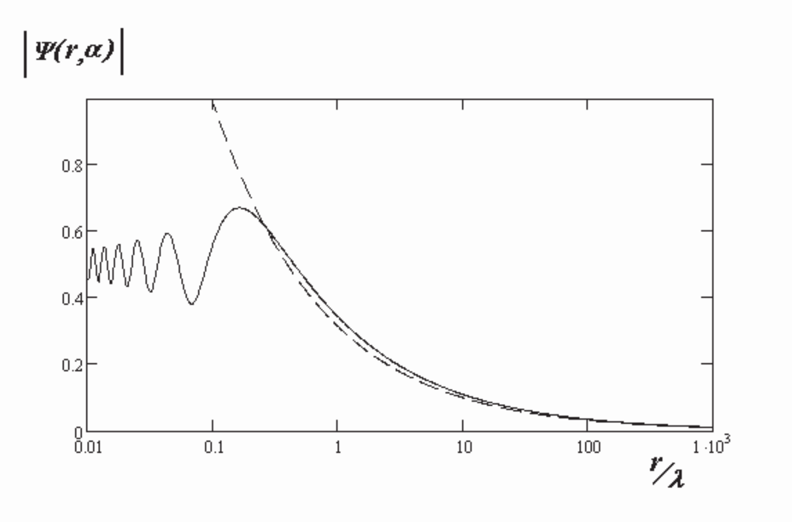

Рис. 17. Зависимость модуля напряженности поля дифракции от расстояния между экраном и точкой наблюдения, расположенной на оси Z. Сплошная кривая – вычисления по формуле (6.21), пунктир – по формуле (6.24). Ширина щели равна 0,7 длины волны.

Проиллюстрируем полученные результаты численными расчетами. На рис 17 показаны зависимости модуля поля, вычисленного по (6.21) (сплошная кривая) и по (6.24) (пунктирная кривая). Поля вычислялись на оси Z при ширине щели d = 0.7λ. Из рисунка видно, что зона дифракции Фраунгофера имеет место, по крайней мере, начиная с расстояния в сотню длин волн от плоскости экрана. В зоне дифракции Френеля зависимость модуля напряженности поля носит осциллирующий характер.

6.3. Одномерная дифракционная решетка

Дифракционные решетки являются устройствами, применяемыми при анализе спектра немонохроматического света. Простой пример такой решетки приведен на рис. 18. Непрозрачный экран имеет большое число (N) щелей шириной d, расположенных параллельно друг другу, причем расположены они строго периодически на расстоянии a друг от друга. Векторы $\mathbf{k}$, показанные на рисунке суть волновые векторы, причем волновой вектор падающей волны k_i нормален к плоскости дифракционной решетки.

Вектор $\mathbf{k}_0$ есть волновой вектор волны, продолжающей за экраном распространятся в первоначальном направлении. Векторы $\mathbf{k}_1$ и $\mathbf{k}_2$ соответствуют волнам, которые испытали дифракцию на решетке.

Задачу будем решать в приближении (иногда говорят «в зоне») Фраунгофера. Для этого заменим расстояние от центра каждой щели до точки наблюдения выражением:

$$r_n = z' + \frac{\left(x' - x_n\right)^2}{2z'} = r_0 - \frac{x' x_n}{z'} + \frac{x_n^2}{2z'}, \quad \text{где}\, r_0 = z' + \frac{x'^2}{2z'} \tag{6.25}$$

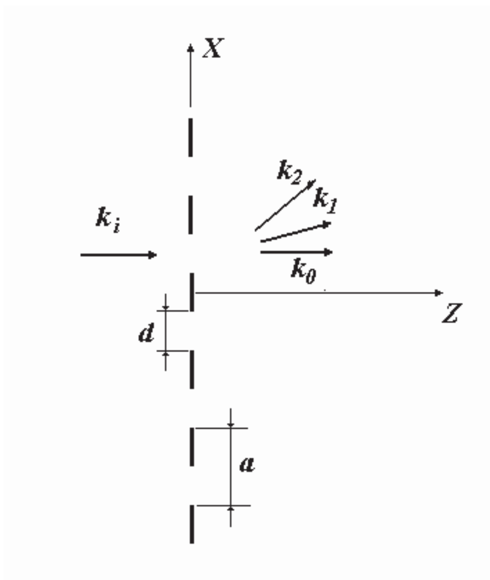

Рис. 18. Дифракционная решетка, состоящая из системы периодических щелей в непрозрачном экране

Поскольку мы хотим найти решение в зоне Фраунгофера, то величиной порядка x_n^2 можно смело пренебречь. Координата $x_n = na$ определяет центр n-ой щели дифракционной решетки. Предположим, что, хотя решетка содержит очень много щелей, ее полный размер мал по сравнению с расстоянием, на котором наблюдается дифракционное поле. Исходя из (6.24) и (6.25), запишем поле на больших расстояниях от дифракционной решетки в приближении Фраунгофера:

$$\psi(r_0, a) = \frac{2\psi_0}{k\sqrt{r_0\lambda}} \exp\left[-i\left(kr_0 - \frac{\pi}{4}\right)\right] \frac{\sin(\frac{\pi d}{\lambda}\alpha)}{\alpha} \sum_{n=-N/2}^{N/2} \exp(ik\alpha na) \qquad (6.26)$$

Здесь было использовано соотношение (6.23), следовательно, полученное выражение справедливо только для малого угла α, под которым точка наблюдения видна из центра дифракционной решетки.

Сумму, входящую в выражение (6.26), можно вычислить в явном виде

$$\sum_{n=-N/2}^{N/2}\exp(ik\alpha na)=\exp(-ik\alpha\frac{N}{2}a)\frac{\exp\left[ik\alpha(N+1)a\right]-1}{\exp(ik\alpha a)-1}=$$
$$=\frac{\sin\frac{1}{2}k\alpha(N+1)a}{\sin\frac{1}{2}k\alpha a} \qquad (6.27)$$

Подставляя вычисленное значение суммы в (6.26) и учитывая, что $k=\frac{2\pi}{\lambda}$, получим

$$\psi(r_0,\alpha)=\psi_0\frac{\sqrt{\lambda}}{\pi\sqrt{r_0}}\exp[-i(kr_0-\pi/4)]\frac{\sin(\frac{\pi d}{\lambda}\alpha)}{\alpha}\frac{\sin(\frac{\pi D}{\lambda}\alpha)}{\sin(\frac{\pi a}{\lambda}\alpha)} \qquad (6.28)$$

Здесь $D=a(N+1)$ - часть экрана, занимаемая только щелями.

Выражение для поля на больших расстояниях от решетки содержит два сомножителя, один из которых описывает дифракционное поле от одной, отдельно взятой щели, тогда как второй член учитывает суммарное влияние всех щелей. Результаты численного расчета по (6.28) (1.29) приведены на рис. 19а. При расчетах использовались следующие параметры: Число щелей N =300, период решетки a = 30λ, ширина щели d = 20λ. Точка наблюдения находилась на расстоянии 105λ от решетки.

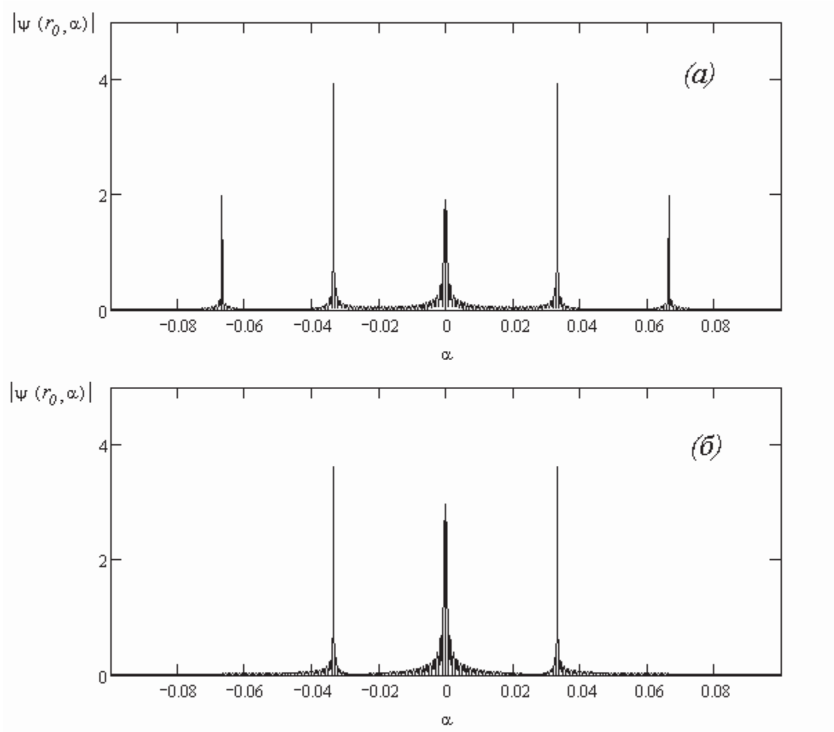

Рис. 19. Угловая зависимость модуля напряженности поля в зоне дифракции Фраунгофера для щелевой дифракционной решетки (а) и решетки с синусоидальной прозрачностью (б).

На практике дифракционные решетки малого периода часто изготавливают, используя голографический процесс. Если не вдаваться в тонкости этого процесса, можно сказать, что в результате получается периодическая структура, функция прозрачности которой меняется по гармоническому закону. Все наши рассуждения, приведенные выше, в этом случае по-прежнему справедливы. Необходимо лишь описанный выше экран, составленный участками с абсолютно прозрачными щелями, заменить аналогичным, функция пропускания которого имеет вид:

$$T(x) = 1 + t\cos(2\pi \frac{x}{a}). \qquad (6.29)$$

Математически это обстоятельство можно учесть, если распределение напряженности поля на экране $\psi(x,z)$, ранее считавшееся константой ψ_0, представить в виде:

$$\psi(x,z) = T(x)\psi_0. \qquad (6.30)$$

Возвращаясь к (6.21) и используя обычные приближения, применяемые к случаю дифракции Фраунгофера, получим

$$\psi(x',z') = \frac{\psi_0}{\sqrt{r_0\lambda}}\exp\left[-i(kr_0 - \frac{\pi}{4})\right]\int_{-aN/2}^{aN/2}\left[1 + t\cos(2\pi\frac{x}{a})\right]\exp(ik\alpha x)dx \qquad (6.31)$$

Здесь под числом N следует понимать число периодов функции пропускания экрана.

Интеграл в (6.31) можно вычислить, выразив косинус через экспоненты. В результате интегрирования получим

$$\psi(x',z') = \frac{\psi_0}{\sqrt{r_0\lambda}}\exp[-i(kr_0 - \frac{\pi}{4})]\times$$
$$\times[2\frac{\sin(ak\alpha\frac{N}{2})}{k\alpha} + t\frac{\sin[(ak\alpha + 2\pi)\frac{N}{2}]}{k\alpha + \frac{2\pi}{a}} + t\frac{\sin[(ak\alpha - 2\pi)\frac{N}{2}]}{k\alpha - \frac{2\pi}{a}}] \qquad (6.32)$$

Действие решетки с гармонической функцией прозрачности существенно отличается от действия щелевой решетки. В то время как щелевая решетка обеспечивает появление в зоне дифракции Фраунгофера множества максимумов напряженности поля, гармоническая решетка имеет только три максимума. Часто эти максимумы называют порядками дифракции. Нулевой порядок дифракции описывается первым слагаемым в квадратных скобках; это просто дифракционная картина экрана как некоего целого объекта, Этот порядок имеет максимальное значение при

$\alpha = 0$. Каждое из двух остальных слагаемых в скобках соответствует одному боковому лепестку. Угловая зависимость дифракционного поля для решетки с синусоидальным профилем функции прозрачности представлена на рис. 19б. Расчет проводился по (6.32) при глубине модуляции t =0,5 с использованием тех же параметров, что и для щелевой решетки.

Завершим рассмотрение дифракционных одномерных решеток анализом фазовых решеток. Две рассмотренные выше решетки модулировали амплитуду падающей волны по закону, заданному периодической функцией прозрачности экрана. Можно получить дифракционную решетку, модулируя не амплитуду, а фазу падающей плоской волны. Это произойдет, если на пути волны поместить совершенно прозрачный экран переменной толщины. Действительно, фаза прошедшей волны зависит от длины оптического пути света, проходящего через прозрачный экран. Если же сделать прозрачный экран с периодически меняющейся толщиной, то его функция прозрачности может быть записана в виде:

$$T(x) = \exp[i(1 + t\cos\frac{2\pi x}{a})] \qquad (6.33)$$

Подставляя это выражение в (6.21), в зоне дифракции Фраунгофера получим

$$\psi(x', z') = \frac{\psi_0}{\sqrt{r_0\lambda}} \exp[-i(kr_0 - \frac{\pi}{4})] \int\limits_{-D/2}^{D/2} \exp[i(1 + t\cos\left(2\pi\frac{x}{a}\right) + k\alpha x)]dx \quad (6.34)$$

Интеграл в этом выражении оценить гораздо труднее, чем интегралы, вычисленные ранее.

Поэтому приведем лишь результаты численных расчетов по (6.34) с параметрами, которые использовались в случае амплитудной модуляции функции прозрачности экрана (см. рис. 20).

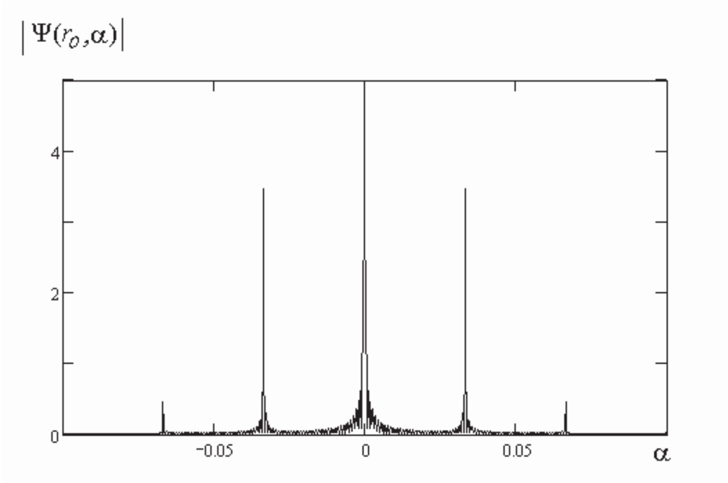

Рис. 20. Угловая зависимость модуля напряженности поля для фазовой дифракционной решетки.

Итак, нами рассмотрены все возможные типы тонких одномерных дифракционных решеток. Нетрудно заметить, что дифракционный интеграл Фраунгофера (6.32) (6.34), описывающий изменения, которые претерпевает плоская электромагнитная волна, проходящая через эти решетки, имеет вид интеграла Фурье. Это совпадение не случайно. Существует альтернативный способ отыскания дифракционного поля решетки, основанный на разложение этого поля в спектр плоских волн, источником которых является рассматриваемая решетка. Подробнее этот метод будет описан в другом разделе курса.

Все приведенные выше рассуждения проводились в предположении, что толщина дифракционной решетки пренебрежимо мала по сравнению с ее периодом и, уж тем более, с расстоянием до точки наблюдения. Однако существует широкий класс периодических структур, для которых это предположение не имеет смысла. В качестве примера можно привести периодическую структуру, образованную системой чередующихся тонких слоев различных диэлектриков. Для анализа электродинамических свойств подобных структур требуется иной математический аппарат, отличающийся от применяемого нами в этом разделе. Подробнее с этим аппаратом читатель может познакомиться в следующем разделе.

6.4. Волны в периодических структурах

В этом разделе рассматриваются задачи распространения волн в структурах, свойства которых изменяются в пространстве по периодическому закону. Примерами таких структур могут служить многослойные оптические покрытия, искусственно созданные оптические среды (т.н. метаматериалы) и т.д. Распространение волн в средах с периодически изменяющимися параметрами сопровождается появлением новых качественных особенностей, наиболее заметных в тех случаях, когда длина волны становится сравнимой с характерным пространственным периодом данной структуры. Ниже мы на простейшем примере такой структуры рассмотрим эти особенности и продемонстрируем математические приемы расчета, применяемые для анализа подобных периодических структур [7].

Рассмотрим среду, в которой диэлектрическая проницаемость является периодической функцией какой-либо пространственной координаты, например z:

$$\varepsilon(z) = \varepsilon_1\left[1 + m\cos(2Kz)\right] \tag{6.35}$$

где $K = \dfrac{2\pi}{\Lambda}$,Λ - пространственный период структуры и ε_1 есть среднее значение диэлектрической проницаемости.

Ограничимся рассмотрением лишь немагнитных сред, для которых $\mu = 1$. Тогда волновое уравнение (2.7) можно записать в виде

$$\nabla^2\vec{E} - \frac{\varepsilon}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} - grad\ div\vec{E} = 0 \tag{6.36}$$

Из материального уравнения следует, что

$\varepsilon \, div\vec{E} + \vec{E} \cdot \, grad\varepsilon = 0$.

Подставляя это в (6.36) получим следующее волновое уравнение для волны с гармонической зависимостью от времени вида $\exp(-i\omega t)$:

$$\nabla^2 \vec{E} + \frac{\omega^2}{c^2}\varepsilon(z)\vec{E} + grad(\vec{E} \cdot \frac{grad\ \varepsilon(z)}{\varepsilon(z)}) = 0 \quad (6.37)$$

Рассмотрим плоскую электромагнитную волну типа TE, распространяющуюся вдоль оси z. Такая волна имеет только одну проекцию вектора напряженности электрического поля, а именно - E_x. Скалярное произведение, стоящее под знаком градиента, в этом случае будет равно нулю. Для такой волны волновое уравнение упрощается и принимает вид:

$$\frac{\partial^2 E_x(z)}{\partial z^2} + k_0^2 \varepsilon(z) E_x(z) = 0 \quad (6.38)$$

Здесь k_0 - волновое число вакуума $= \frac{\omega}{c}$.

Подставляя сюда зависимость диэлектрической проницаемости от координаты (6.35), окончательно получим

$$\frac{\partial^2 E_x(z)}{\partial z^2} + k^2 \left[1 + m\cos(2Kz)\right] E_x(z) = 0 \quad (6.39)$$

где $k^2 = k_0^2 \varepsilon_1$ - среднее волновое число среды.

Введем новые переменные

$Kz = \xi, \quad \frac{k^2}{K^2} = \eta, \quad m\frac{k^2}{K^2} = \gamma.$

Тогда уравнение (6.39) примет вид

$$\frac{\partial^2 \Psi}{\partial \xi^2} + \left[\eta + \gamma\cos(2\xi)\right]\Psi = 0 \quad (6.40)$$

Уравнение типа (6.40) впервые получено ***Матье*** и носит его имя. Как установил ***Флоке***, общее решение уравнения (6.40) записывается в форме:

$$\Psi = C_1 F_1(\xi)\exp(\nu\xi) + C_2 F_2(\xi)\exp(-\nu\xi) \qquad (6.41)$$

Здесь C_1, C_2 - произвольные константы, F_1, F_2 - периодические функции. В общем случае (6.41) представляет собой суперпозицию двух волн, распространяющихся навстречу друг другу, причем случай действительного или комплексного волнового числа ν соответствует волнам с экспоненциально уменьшающейся амплитудой.

Основные результаты, получаемые из решения уравнения Матье, проиллюстрированы на рис. 21.

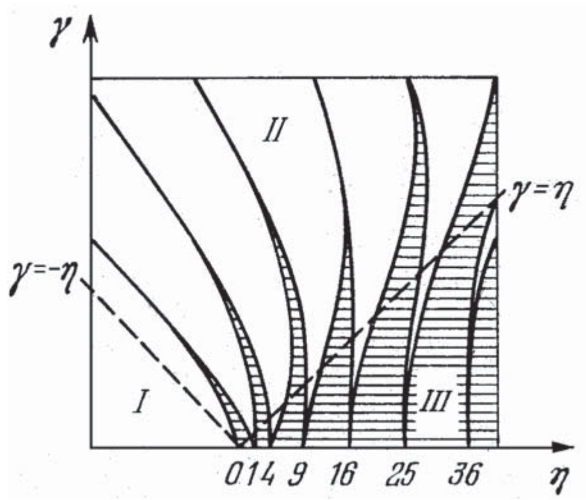

Рис. 21. Основные свойства решений уравнения Матье

Рисунок 21 показывает связь переменных η и γ. В незаштрихованных областях величина ν комплексная или действительная, что означает экспоненциальное уменьшение амплитуды этих волн по мере их распространения. В заштрихованных областях ν чисто мнимая величина, следовательно, здесь волны распространяются вдоль структуры без уменьшения амплитуды. На границах графика действительная часть ν обращается в нуль.

Всю плоскость η, γ удобно разбить на три части с помощью двух прямых линий $\gamma=\eta$ и $\gamma=-\eta$, изображенных на рис. 21 пунктирными линиями. В области I , где $\gamma<-\eta$, распространяющиеся волны существовать не могут. В области II, где $-\eta<\gamma<\eta$, заштрихованные области (или окна прозрачности структуры) сужаются по мере увеличения γ и в пределе становятся прямыми, параллельными направлению $\gamma=-\eta$. В области III, где $\gamma<\eta$, области прозрачности становятся широкими, а полосы непропускания – узкими. Полуось $\eta>0$ соответствует случаю распространения волн в однородной среде ($m=0$). Вблизи оси η, где $\gamma<<\eta$, мы имеем дело со случаем слабо неоднородной среды.

Решения уравнения (6.40) выражаются через специальные *функции Матье*, которые хорошо изучены и табулированы.

Если же изучаемая среда модулирована с помощью произвольной периодической функции, отличной от синусоидальной, то в этом случае мы имеем дело с ***уравнением Хилла***:

$$\frac{\partial^2\Psi}{\partial z^2}+k^2\left[1+mf(z)\right]\Psi=0 \qquad (6.42)$$

где $f(z)$ - произвольная периодическая функция.

С помощью (6.42) можно решить практически важную задачу о распространении волн в слоистой периодической структуре, что мы и проделаем на примере структуры, образованной слоями с различными значениями диэлектрической проницаемости. Функция модуляции такой структуры показана на рис. 22.

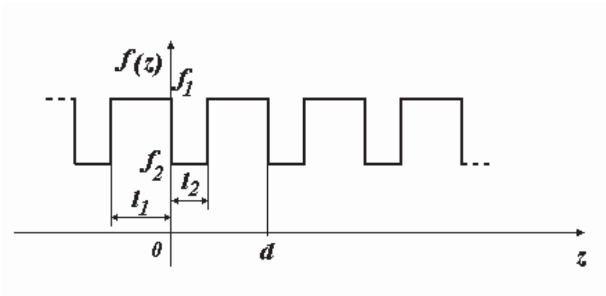

Рис. 22. Периодическая функция модуляции, входящая в (6.42)

Нашей задачей является определение дисперсии структуры с модуляцией диэлектрической проницаемости вида $\varepsilon(z)=\varepsilon_1(1+mf(z))$. Напомним, что термином дисперсия мы обозначаем зависимость волнового числа структуры $\tilde{k}$ от частоты ω.

Введем обозначения:

$k^2(1+mf_1)=k_1^2, \quad k^2(1+mf_2)=k_2^2$.

Решение уравнения Хилла имеет вид

$$\Psi(z)=\begin{cases} C_1\exp(ik_1z)+D_1\exp(-ik_1z), & -l_1<z<0 \\ C_2\exp(ik_2z)+D_2\exp(-ik_2z), & 0<z<l_2 \end{cases} \tag{6.43}$$

В силу теоремы Флоке для одной из встречных волн (2.7) можно записать

$$\Psi(z)=F(z)\exp(i\tilde{k}z) \tag{6.44}$$

где $F(z)$ - функция, обладающая периодом d.

Отсюда следует, что в интервале $l_2<z<d$ решение запишется в виде

$$\Psi(z)=C_1\exp(i\tilde{k}d)\exp[ik_1(z-d)]+D_1\exp(i\tilde{k}d)\exp[-ik_1(z-d)] \tag{6.45}$$

Выражения (6.43) и (6.45) описывают поле волны в трех областях при $-l_1<z<d$, разделенных двумя границами $z=0$ и $z=l_2$. Воспользуемся граничными условиями для полей на этих границах и потребуем равенства

на границах самой функции $\Psi(z)$ и ее производной. В результате получим четыре уравнения:

$$C_1 + D_1 = C_2 + D_2, \qquad C_1 k_1 - D_1 k_1 = C_2 k_2 - D_2 k_2,$$
$$C_1 \exp(i\tilde{k}d)\exp(-ik_1 l_1) + D_1 \exp(i\tilde{k}d)\exp(ik_1 l_1) =$$
$$= C_2 \exp(ik_2 l_2) + D_2 \exp(-ik_2 l_2) \qquad (6.46)$$
$$C_1 \exp(i\tilde{k}d)\exp(-ik_1 l_1) - D_1 \exp(i\tilde{k}d)\exp(ik_1 l_1) =$$
$$= C_2 \exp(ik_2 l_2) k_2 - D_2 \exp(-ik_2 l_2) k_2$$

Условием совместности этой системы линейных уравнений относительно четырех коэффициентов C_1, C_2, D_1, D_2 является равенство нулю ее определителя:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ k_1 & -k_1 & k_2 & -k_2 \\ e^{(i\tilde{k}d)}e^{(-ik_1 l_1)} & e^{(i\tilde{k}d)}e^{(ik_1 l_1)} & e^{(ik_2 l_2)} & e^{(-ik_2 l_2)} \\ k_1 e^{(i\tilde{k}d)}e^{(-ik_1 l_1)} & -k_1 e^{(i\tilde{k}d)}e^{(ik_1 l_1)} & k_2 e^{(ik_2 l_2)} & -k_2 e^{(-ik_2 l_2)} \end{vmatrix} = 0.$$

Вычисляя этот определитель, получим уравнение:

$$e^{(i2\tilde{k}d)} + \left(\frac{k_1^2 + k_2^2}{k_1 k_2} \sin(k_1 l_1)\sin(k_2 l_2) - 2\cos(k_1 l_1)\cos(k_2 l_2) \right) e^{(i\tilde{k}d)} + 1 = 0 \quad (6.47)$$

Решая это уравнение относительно искомого $\tilde{k}$, получим дисперсионное уравнение:

$$2\cos(\tilde{k}d) + \left(\frac{k_1^2 + k_2^2}{k_1 k_2} \sin(k_1 l_1)\sin(k_2 l_2) - 2\cos(k_1 l_1)\cos(k_2 l_2) \right) = 0 \qquad (6.48)$$

Анализировать полученное дисперсионное уравнение (6.48) удобнее, если положить одинаковыми оптические толщины слоев, т.е. $k_1 l_1 = k_2 l_2$. Это означает, что при распространении в каждом слое фаза волны изменяется на одну и ту же величину. В этом случае дисперсионное уравнение упрощается и приобретает вид:

$$\cos(\tilde{k}d) = 1 - \frac{(k_1 + k_2)^2}{2k_1k_2}\sin^2(k_1l_1) \qquad (6.49)$$

Правая часть дисперсионного уравнения (6.49), вычисленная как функция от величины k_1l_1, показана на рис.23.

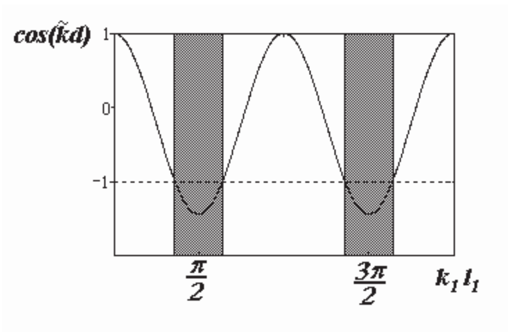

Рис.23. Правая часть дисперсионного уравнения (6.49)

Легко заметить, что можно подобрать такое значение величины k_1l_1, при котором решение дисперсионного уравнения возможно лишь при чисто мнимом значении $\tilde{k}$. Очевидно, что эта ситуация соответствует полосе непрозрачности структуры, другими словами, амплитуда одной из встречных волн (6.44) будет экспоненциально убывать по мере проникновения в структуру, а другой, распространяющейся в противоположном направлении – возрастать. При достаточно широкой полосе непрозрачности волна может полностью от нее отразиться и такую структуру можно рассматривать как некий аналог зеркала для электромагнитной волны, причем зеркала, изготовленного только из диэлектрических слоев.

Подводя итог, можно заключить, что характерной особенностью распространения электромагнитной волны в диэлектрической среде с поперечной стратификацией является возникновения полос непрозрачности, в которых волна не может проникнуть глубоко в такую среду. В этих обстоятельствах среда ведет себя как зеркало, что приводит к отражению электромагнитной волны от такой слоистой структуры. В силу того, что поглощение света в диэлектрических слоях, применяемых для изготовления подобных структур, может быть весьма малым, то появляется возможность создавать оптические зеркала с коэффициентом отражения (по интенсивности) близким к единице. Достижение таких значений коэффициента отражения в случае использования металлических пленок в качестве отражающей поверхности в большинстве случаев невозможно. Наиболее часто диэлектрические зеркала применяются для изготовления резонаторов лазеров, поскольку, в силу малых потерь на поглощение, могут выдерживать воздействие весьма интенсивных световых пучков.

Тема 7. Краткие сведения из векторного анализа

В этом разделе векторный анализ изложен в объеме, необходимом для понимания некоторых математических выкладок в данном пособии; ни к полноте, ни к математической строгости мы не стремились [8].

7.1. Векторная алгебра

Векторная алгебра предполагается читателю известной, здесь мы лишь напомним некоторые основные ее определения и формулы.

Скалярное произведение векторов

$$\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}, \qquad \vec{b} = b_x\vec{i} + b_y\vec{j} + b_z\vec{k},$$

где $\vec{i}, \vec{j}$ и $\vec{k}$ суть единичные векторы, направленные по осям x, y, z, равно

$$\vec{a}\vec{b} = (\vec{a}\cdot\vec{b}) = \vec{b}\vec{a} = ab\cos(a,b) = a_x b_x + a_y b_y + a_z b_z.$$

Векторное произведение $\left[\vec{a}\times\vec{b}\right]$ векторов $\vec{a}$ и $\vec{b}$ является вектором, перпендикулярным к $\vec{a}$ и $\vec{b}$ и по абсолютной величине равным площади параллелограмма, построенного на этих векторах

$$\left|\left[\vec{a}\times\vec{b}\right]\right| = ab\sin(a,b),$$

$$\left[\vec{a}\times\vec{b}\right] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = \vec{i}\left(a_y b_z - a_z b_y\right) + \vec{j}\left(a_z b_x - a_x b_z\right) + \vec{k}\left(a_x b_y - a_y b_x\right);$$

$$\left[\vec{a}\times\vec{b}\right] = -\left[\vec{b}\times\vec{a}\right].$$

Направление вектора $\left[\vec{a}\times\vec{b}\right]$ определяется из требования, чтобы векторы $\vec{a}$, $\vec{b}$ и $\left[\vec{a}\times\vec{b}\right]$ образовывали *правовинтовую* систему (см. рис. 24)

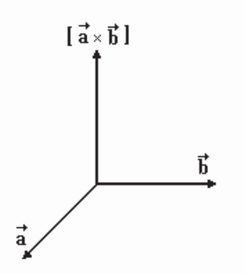

Рис. 24. Векторное произведение

Смешанное или векторно-скалярное, произведение трех векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$ является скаляром и численно равно объему параллелепипеда, построенного на этих векторах:

$$\vec{a}\left[\vec{b}\times\vec{c}\right]=\vec{b}\left[\vec{c}\times\vec{a}\right]=\vec{c}\left[\vec{a}\times\vec{b}\right]=-\vec{b}\left[\vec{a}\times\vec{c}\right]=-\vec{a}\left[\vec{c}\times\vec{b}\right]=\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.$$

Двойное векторное произведение векторов $\vec{a}$, $\vec{b}$ и $\vec{c}$ равно

$$\left[\vec{a}\times\left[\vec{b}\times\vec{c}\right]\right]=\vec{b}\left(\vec{a}\cdot\vec{c}\right)-\vec{c}\left(\vec{a}\cdot\vec{b}\right)=-\left[\left[\vec{b}\times\vec{c}\right]\times\vec{a}\right]$$

Если векторы являются функциями некоторой скалярной переменной ***t***, то при соблюдении обычных условий можно дифференцировать векторы по этой переменной. При этом имеют место соотношения

$$\frac{d}{dt}(\vec{a}+\vec{b})=\frac{d\vec{a}}{dt}+\frac{d\vec{b}}{dt}, \qquad \frac{d}{dt}(\phi\vec{a})=\phi\frac{d\vec{a}}{dt}+\frac{d\phi}{dt}\vec{a},$$

$$\frac{d(\vec{a}\cdot\vec{b})}{dt}=\left(\frac{d\vec{a}}{dt}\cdot\vec{b}\right)+\left(\vec{a}\cdot\frac{d\vec{b}}{dt}\right)$$

и т.д.

7.2. Векторные и скалярные поля. Градиент

Рассмотрим скалярное поле функции $\phi(\vec{R}) = \phi(x, y, z)$. Таким полем является, например, поле температуры неравномерно нагретого тела $(\phi = T)$, поле плотности неоднородного тела $(\phi = \tau)$, поле электростатического потенциала и т.п.

Пусть скаляр ϕ имеет в точке P_0 значение ϕ_0, и пусть при перемещении $d\vec{S}$ по направлению вектора $\vec{S}$ мы переходим из точки P_0 в точку P, где скаляр ϕ имеет значение ϕ_S. Приращение ϕ при этом перемещении равно $d\phi = \phi_S - \phi_0$. Предел отношения этого приращения $d\phi$ к числовому значению перемещения ds обозначается через $\frac{\partial \phi}{\partial s}$ и называется производной скаляра ϕ в точке P_0 по направлению $\vec{S}$:

$$\frac{\partial \phi}{\partial s} = \lim_{ds \to 0} \frac{\phi_S - \phi_0}{ds} \tag{7.1}$$

Очевидно, что значение этой производной существенно зависит от выбора направления $\vec{S}$ и что ее ни в коем случае нельзя смешивать с обыкновенной частной производной по скалярному параметру S.

Для изучения зависимости производной $\frac{\partial \phi}{\partial s}$ от направления дифференцирования $\vec{S}$ рассмотрим те точки поля, в которых ϕ имеет одинаковое значение, например ϕ_0. Совокупность этих точек, вообще говоря, образует собой поверхность, которая называется *поверхностью уровня*, или *эквипотенциальной поверхностью*. Аналитически поверхность эта характеризуется уравнением

$\phi(x, y, z) = \phi_0$.

Обозначим через $\vec{n}$ нормаль к поверхности уровня $\phi = \phi_0$, направленную в сторону возрастания ϕ, и покажем, что, зная производную

$\frac{\partial \phi}{\partial n}$ по направлению этой нормали, можно определить значение производной скаляра ϕ по любому направлению $\vec{S}$.

Пусть поверхность уровня, проходящая через лежащую в направлении $\vec{S}$ точку P_S, пересекает нормаль $\vec{n}$ (или ее продолжение в обратном направлении) в точке P_n (рис. 25). Значение ϕ в точке P_n равно значению ϕ в точке P_S ($\phi_n = \phi_S$) и $P_0P_S = \frac{P_0P_n}{\cos(\vec{S},\vec{n})}$.

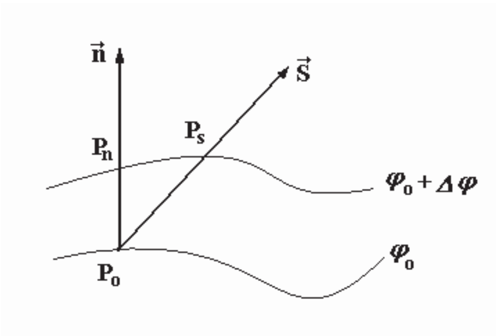

Рис. 25. Нормаль к поверхности уровня

Поэтому

$$\left(\frac{\partial \phi}{\partial s}\right)_0 = \lim_{P_0P_S \to 0} \frac{\phi_S - \phi_0}{P_0P_S} = \cos(\vec{S},\vec{n}) \qquad \lim_{P_0P_n \to 0} \frac{\phi_n - \phi_0}{P_0P_S} = \left(\frac{\partial \phi}{\partial n}\right)_0 \cos(\vec{S},\vec{n}).$$

Таким образом,

$$\frac{\partial \phi}{\partial s} = \frac{\partial \phi}{\partial n}\cos(\vec{S},\vec{n}) \tag{7.2}$$

Вектор, численно равный $\frac{\partial \phi}{\partial n}$ и направленный по нормали к поверхности уровня в сторону возрастания ϕ, носит название *градиента скаляра* ϕ:

$$grad\phi = \frac{\partial\phi}{\partial n}\vec{n} \tag{7.3}$$

Поэтому уравнение (7.2) может быть записано так:

$$\frac{\partial\phi}{\partial s} = \left|grad\phi\right|\cos(\vec{S},\vec{n}) = grad_S\phi \tag{7.4}$$

Стало быть, производная $\frac{\partial\phi}{\partial s}$ равна проекции вектора градиента ϕ на направление $\vec{S}$. Если, в частности, ввести систему декартовых координат x, y, z, оси которой направлены параллельно единичным векторам $\vec{i}, \vec{j}$, $\vec{k}$ то, согласно уравнению (7.4), получим

$$grad_x\phi = \frac{\partial\phi}{\partial x}, \quad grad_y\phi = \frac{\partial\phi}{\partial y}, \quad grad_z\phi = \frac{\partial\phi}{\partial z} \tag{7.5}$$

т.е.

$$\begin{gathered} grad\phi = \vec{i}\frac{\partial\phi}{\partial x} + \vec{j}\frac{\partial\phi}{\partial y} + \vec{k}\frac{\partial\phi}{\partial z}, \\ \left|grad\phi\right| = \sqrt{\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2} \end{gathered} \tag{7.6}$$

Из уравнения (7.4) следует, как это, впрочем, и непосредственно явствует из рис. 25, что направление градиента $\vec{n}$ есть направление наиболее быстрого возрастания скаляра ϕ, а направление ($-\vec{n}$) есть направление наиболее быстрого убывания ϕ. В направлениях же, перпендикулярных к $\vec{n}$, т.е. касательных к поверхности уровня, значение ϕ вовсе не изменяется ($\frac{\partial\phi}{\partial s} = 0$).

Итак, если известно поле скаляра ϕ, то в каждой точке этого поля можно определить вектор $grad\phi$, перпендикулярный поверхностям уровня этого скаляра. Если провести систему *ортогональных траекторий* поверхностей уровня, т.е. систему линий, перпендикулярных этим поверхностям, то в каждой точке поля *направление* градиента будет совпадать с направлением этих линий. Поэтому ортогональные траектории поверхностей уровня носят название *линий градиента*.

7.3. Поток вектора через поверхность

Если задано поле произвольного, но дифференцируемого скаляра $\phi(\vec{R})$, то тем самым задано и поле производных этого скаляра по произвольному направлению. Инвариантной, т.е. не зависящей от выбора системы координат характеристикой этого поля производных является, как мя видели, поле вектора $grad\phi$. Нам предстоит теперь определить инвариантные характеристики поля пространственных производных произвольного векторного поля $\vec{a}(\vec{R})$ в окрестности точки, где это поле задано. К этим характеристикам, естественно приводит рассмотрение поверхностных и криволинейных интегралов вектора $\vec{a}$. Мы начнем с исследования поверхностных интегралов

В векторном поле выделим мысленно бесконечно малую плоскую площадку dS, т.е. площадку столь малую, что во всех ее точках векторное поле с заданной степенью точности остается постоянным по величине и направлению. Проведем нормаль к этой площадке и условимся одно из направлений этой нормали $\vec{n}$ считать положительным, или внешним, а другое – отрицательным, или внутренним. Если задано направление обхода контура площадки, то направление положительной нормали мы будем выбирать так, чтобы нормаль образовывала вместе с контуром *правовинтовую* систему. Обратно, если задано направление внешней нормали, то мы будем соответственным образом выбирать направление положительного обхода контура площадки.

Наконец, если направление обхода контура и направление нормали к его плоскости заданы независимо друг от друга, то мы будем для краткости говорить, что направление обхода и направление нормали составляют *правовинтовую* систему, если они удовлетворяют упомянутому условию, и *левовинтовую* систему, если они ему не удовлетворяют.

Направление нормали мы будем характеризовать совпадающим с ней единичным вектором $\vec{n}$.

Потоком вектора $\vec{a}$ через бесконечно малую площадку dS называется величина

$$dN = \vec{a} \cdot \vec{n} dS = a\cos(\vec{a},\vec{n})dS = a_n dS \quad (7.7)$$

где a - значение вектора на площадке dS, а a_n - проекция этого вектора на направление $\vec{n}$. Площадка dS выбрана нами бесконечно малой именно для того, чтобы вектор $\vec{a}$ имел на этой площадке одно определенное значение (см. рис. 26).

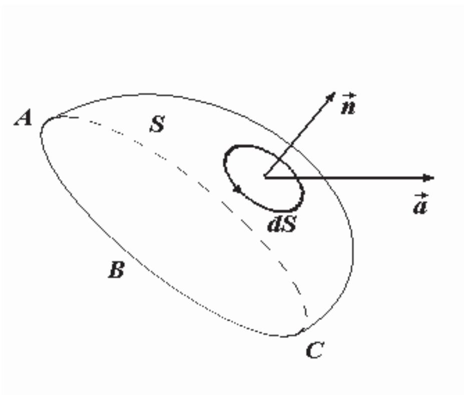

Рис. 26. Поток вектора

Чтобы определить поток вектора $\vec{a}$ через поверхность конечных размеров, нужно разбить ее на бесконечно малые площадки dS так, чтобы не только вектор $\vec{a}$ оставался постоянным на каждой площадке, но чтобы и самые площадки можно было бы считать плоскими (рис. 26). Одну из сторон поверхности S назовем внутренней, а другую – внешней и выберем соответственным образом направление внешних нормалей к каждому из элементов dS. Потоком N вектора $\vec{a}$ через поверхность S называется алгебраическая сумма потоков $a_n dS$ через отдельные элементы этой поверхности. Это суммирование тождественно с операцией нахождения определенного интеграла:

$$N = \iint_S a_n dS$$

и называется *интегрированием по поверхности* S. Оно обозначается двойным интегралом потому, что поверхность имеет два измерения. Однако для упрощения записи мы в этом курсе обозначили двукратные интегралы, как и интегралы однократные, одним единственным знаком интеграла:

$$N = \int_S a_n dS \tag{7.8}$$

Напомним, что во всех поверхностных (и только в поверхностных) интегралах мы обозначили элемент интегрирования через dS.

Часто приходится вычислять поток векторного поля через замкнутые поверхности (поверхности шара, куба и т.д.). Мы будем отмечать это обстоятельство кружком на знаке интеграла следующим образом $N = \int_S a_n dS$.

7.4. Теорема Гаусса. Дивергенция

Поверхностный интеграл $\int_S a_n dS$ можно преобразовать в объемный; в этом заключается содержание одной из важнейших теорем векторного анализа – *теоремы Гаусса.*

Рассмотрим сначала поток dN произвольного, но дифференцируемого вектора $\vec{a}$ через поверхность бесконечно малого параллелепипеда и выберем для удобства вычислений направление осей координат x, y, z так, чтобы они совпадали с ребрами этого параллелепипеда dx, dy, dz (см. рис. 27).

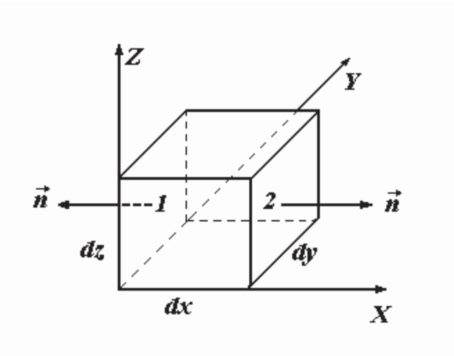

Рис. 27. Геометрический смысл дивергенции

Интеграл $dN = \int\limits_S a_n ds$ сводится в этом случае к сумме шести интегралов по каждой из граней параллелепипеда. Воспользовавшись известной из интегрального исчисления теоремой о среднем, можно каждый из этих шести интегралов представить как произведение площади грани на некоторое среднее значение нормальной слагающей вектора $\vec{a}$ на данной грани.

Рассмотрим сначала поток вектора $\vec{a}$ через две параллельные грани 1 и 2, перпендикулярные оси ***X***. Поток через переднюю грань 2 равен

$$a_{2n}(x + dx, \overline{y}, \overline{z})dS = a_{2x}(x + dx, \overline{y}, \overline{z})dxdz$$

где $\overline{y}$ и $\overline{z}$ - некоторые средние значения координат y и z на грани 2 и a_2 - значение вектора $\vec{a}$ на грани 2; поток через заднюю грань 1 равен

$$a_{1n}ds = -a_{1x}dydz \qquad \left(a_{1n} = a_{1n}(x, \overline{y}, \overline{z}\right),$$

где a_1 - значение вектора $\vec{a}$ на грани 1, ибо внешняя нормаль к этой грани направлена противоположно оси ***X***. Стало быть, общий поток через грани 1 и 2 равен $\left(a_{2x} - a_{1x}\right)dydz$.

Разность $a_{2x} - a_{1x}$ есть приращение проекции вектора a_x при изменении координаты ***x*** на расстояние ***dx*** между гранями 1 и 2. С точностью до бесконечно малых второго порядка приращение это равно

$$a_{2x}(x+dx, y, z) - a_{1x}(x, y, z) = \frac{\partial a_x}{\partial x} dx,$$

где ввиду бесконечной малости параллелепипеда под $\frac{\partial a_x}{\partial x}$ можно понимать значение этой производной в любой точке параллелепипеда. Таким образом, общий поток через обе грани, перпендикулярные к оси ***X***, равен $\frac{\partial a_x}{\partial x} dxdydz$.

Для потоков через пары граней, перпендикулярных осям ***Y*** и ***Z***, получим аналогично $\frac{\partial a_y}{\partial y} dxdydz$ и $\frac{\partial a_z}{\partial z} dxdydz$.

Складывая полученные выражения, получим общий поток вектора $\vec{a}$ через все шесть граней элементарного параллелепипеда

$$dN = \int_S a_n dS = \left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) dxdydz \qquad (7.9)$$

Слоящую в скобках сумму производных вектора $\vec{a}$ по координатам принято для краткости обозначить символом $div\vec{a}$:

$$div\vec{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \qquad (7.10)$$

Если, кроме того, ввести для бесконечно малого элемента объема обозначение dV: $dV = dxdydz$, то выражение для потока dN примет вид

$$dN = div\vec{a}dV \qquad (7.11)$$

Эту формулу, выражающую поток вектора $\vec{a}$ через поверхность бесконечно малого параллелепипеда, нетрудно обобщить для поверхности произвольной формы и размеров. Рассмотрим произвольную замкнутую поверхность S. Разобьем ограниченный ею объем V системой взаимно перпендикулярных плоскостей на совокупность бесконечно малых кубических элементов. Конечно, крайние, смежные с поверхностью S, элементы объема, вообще говоря, не будут иметь кубической формы; однако путем дальнейшего их дробления можно достигнуть того, что грани крайних кубиков с любой степенью точности будут совпадать с заданной поверхностью S. Вычислим с помощью уравнения (7.11) поток вектора $\vec{a}$ через поверхность каждого кубика, лежащего внутри S, и сложим полученные выражения:

$$\sum dN = \sum div\vec{a}dV = \iiint_V div\vec{a}dV .$$

В этом уравнении тройной интеграл означает, что суммирование подынтегрального выражения должно быть произведено по всем элементам трехмерного объема V, заключенного внутри поверхности S. Однако мы договорились, что мы обозначаем интегралы любой кратности одним единственным интегралом, различие же интегралов разной кратности достигается различным обозначением элементов интегрирования:

- элемент объема (трехкратного интеграла) обозначается через dV,
- элемент поверхности (двукратного интеграла) через dS,
- элемент линии (одинарного интеграла) через ds.

Грани всех элементарных кубиков, составляющих в совокупности объем V, могут быть разделены на два класса – грани внешние, совпадающие с элементами поверхности S, и грани внутренние, ограничивающие смежные кубики друг от друга. Очевидно, что в сумму $\sum dN$ поток вектора $\vec{a}$ через каждую внутреннюю грань войдет дважды: при подсчете потока через поверхность кубика, лежащего по одну сторону от этой грани, и при подсчете потока через поверхность кубика, лежащего по другую сторону от нее. Так как нормаль к грани, внешняя по отношению к первому кубику, противоположна нормали к той же грани, внешней по отношению ко второму кубику, то оба потока через эту грань будут иметь противоположные знаки. Следовательно, все члены суммы $\sum dN$, относящиеся к внутренним граням, сократятся, и сумма эта сведется к сумме потоков вектора $\vec{a}$ чрез одни лишь внешние грани кубиков, совпадающие с элементами поверхности S. Таким образом, $\sum dN$ оказывается равной потоку N вектора $\vec{a}$ через заданную поверхность S, и, стало быть,

$$N = \int_S a_n dS = \int_V div\vec{a}dV \qquad (7.12)$$

Это выражение представляет собой *теорему Гаусса*: поток вектора $\vec{a}$, являющегося непрерывной функцией точки, через произвольную замкнутую поверхность S равен интегралу от дивергенции этого вектора по объему V, ограниченному этой поверхностью.

Если поверхность S столь мала, что во всех лежащих внутри нее точках $div\vec{a}$ можно считать величиной постоянной, то в уравнении (7.12) $div\vec{a}$ можно вынести за знак интеграла. Стало быть, поток dN через бесконечно малую замкнутую поверхность S произвольной формы выражается той же формулой (7.11):

$$dN = \int_S a_n dS = div\vec{a}dV\ ,$$

как и поток через поверхность элементарного параллелепипеда. Так как эта формула справедлива лишь в предельном случае бесконечно малой поверхности, то ее правильнее записать в следующей форме:

$$div\vec{a} = \lim_{\Delta V \to 0}\left(\frac{\int_S a_n ds}{\Delta V}\right) \qquad (7.13)$$

Правильнее всего считать эту формулу определением *понятия дивергенции*: дивергенция вектора $\vec{a}$ в данной точке поля есть предел, к которому стремится отношение потока вектора $\vec{a}$ через произвольную, окружающую эту точку, поверхность к ограниченному этой поверхностью объему ΔV (при $\Delta V \to 0$). Из этого определения дивергенции следует, что значение ее вовсе не зависит от выбора системы координат, т.е. дивергенция вектора есть истинный скаляр.

Отметим в заключение, что в гидродинамике дивергенция скорости жидкости $\vec{v}$ имеет непосредственное физическое значение. Действительно, в каждой точке жидкости $div\vec{v} = \lim\limits_{\Delta V \to 0} \dfrac{\int v_n dS}{dV}$ равна рассчитанному на единицу объема количеству жидкости, вытекающей из элемента объема dV, окружающего рассматриваемую точку. Название «дивергенция», что в переводе с латыни значит расхождение или расходимость, было избрано для этой величины именно потому, что жидкость растекается или расходится из тех и только из тех точек или участков занимаемого ею пространства, в которых $div\vec{v} > 0$. Очевидно, что в этих точках должны быть расположены источники жидкости. По аналогии, те точки поля произвольного вектора $\vec{a}$, в которых $div\vec{a} \neq 0$, принято называть истоками этого поля. Числовое же значение $div\vec{a}$ называется силой или обильностью истоков поля; в зависимости от знака дивергенции сила истоков может быть как положительной, так и отрицательной. Иногда отрицательным истокам поля дают название стоков поля. Векторные поля, у которых $div\vec{a} = 0$, называют свободными от источников, или *соленоидальными*.

7.5. Циркуляция вектора. Ротор вектора. Теорема Стокса

Преобразование интеграла вектора по замкнутой поверхности в интеграл по объему привело нас к понятию дивергенции вектора. Рассмотрим теперь интеграл от вектора по замкнутой кривой.

Пусть в векторном поле $\vec{a}(\vec{R})$ задана некоторая кривая $\boldsymbol{L}$ и вместе с тем задано, какое из двух возможных направлений движения вдоль этой кривой считается положительным. Разбиваем кривую $\boldsymbol{L}$ на бесконечно малые элементы $d\vec{s}$, направление которых совпадает с направлением положительного движения вдоль линии, и умножаем каждый элемент $d\vec{s}$ скалярно на значение вектора $\vec{a}$ в соответствующей точке поля. Предел суммы этих произведений $\vec{a}\cdot d\vec{s} = a_s ds$ при $ds \to 0$, распространенный на все элементы кривой, называется линейным интегралом вектора $\vec{a}$ вдоль кривой $\boldsymbol{L}$:

$$\int_L \vec{a} d\vec{s} = \int_L a_s ds .$$

Если кривая $\boldsymbol{L}$ замкнута, то линейный интеграл вектора $\vec{a}$ вдоль нее называется циркуляцией $\vec{a}$ вдоль $\boldsymbol{L}$:

$$C(\vec{a}) = \oint_L \vec{a} d\vec{s} = \oint_L a_s ds \qquad (7.14)$$

Предположим, что контур $\boldsymbol{L}$ представляет собой контур плоского прямоугольника $\boldsymbol{ABCD}$, и выберем оси $\boldsymbol{X}$ и $\boldsymbol{Y}$ декартовых координат так, чтобы они были параллельны сторонам этого прямоугольника и пересекались в его центре (см. рис. 28). Пусть стороны этого прямоугольника равны соответственно Δx и Δy. Если выбрать направление положительного обхода контура так, чтобы соответствующая положительная нормаль к площади прямоугольника была направлена по оси $\boldsymbol{Z}$ (рис. 28), то

$$C = \oint_S a_s ds = \int_A^B a_x dx + \int_B^C a_y dy + \int_C^D a_x dx + \int_D^A a_y dy .$$

Воспользовавшись известной из интегрального исчисления теоремой о среднем значении, получим (при $\vec{n} \| OZ$)

$$C = \oint a_s ds = a'_x \Delta x + a''_y \Delta y - a'''_x \Delta x - a''''_y \Delta y ,$$

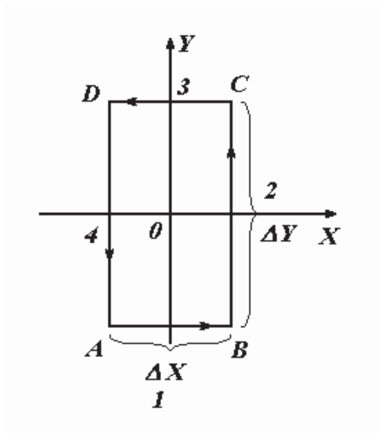

Рис. 28. Циркуляция вектора

где a_x', a_y'' и т.д. суть средние значения a_x и a_y на первой, второй и т.д. сторонах прямоугольника; отрицательный знак, например, последнего члена суммы объясняется тем, что интегрирование по стороне ***AD*** производится в направлении убывания координаты ***y***.

Будем стремить теперь длину сторон прямоугольника к нулю. Тогда с точностью до величин второго порядка малости среднее значение составляющей a_y на отрезке ***BC***, отстоящем от отрезка ***AD*** на расстояние Δx по направлению оси ***OX***, будет отличаться от значения a_y на отрезке ***AD*** на величину $\frac{\partial a_y}{\partial x}\Delta x$: $a_y'' = a_y'''' + \frac{\partial a_y}{\partial x}\Delta x$.

Соответственно $a_x''' = a_x' + \frac{\partial a_x}{\partial y}\Delta y$, ибо ***CD*** отстоит от ***AB*** на расстояние Δy по направлению оси ***OY***. При этом в пределе при бесконечно малых размерах прямоугольника мы можем понимать под $\frac{\partial a_y}{\partial x}$ и $\frac{\partial a_x}{\partial y}$ значения этих величин в центре ***0*** прямоугольника. Внося эти выражения в предшествующее равенство, получим (при $\vec{n} \| OZ$):

$$dC = \int a_s ds = \left(a''_y - a''''_y\right) \Delta y - \left(a''_x - a'_x\right) \Delta x = \frac{\partial a_y}{\partial x} \Delta x \Delta y - \frac{\partial a_x}{\partial y} \Delta x \Delta y,$$

где мы заменили ***C*** через dC, чтобы отметить, что соотношение это справедливо лишь для бесконечно малого прямоугольника. Обозначая, наконец, площадь прямоугольника $\Delta x \Delta y$ через dS, получим окончательно при $\vec{n} \| OZ$:

$$dC = \int a_s ds = \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) dS \qquad (7.15)$$

Так как оси OX, OY, OZ образуют правовинтовую систему, то, совершив круговую перестановку индексов x, y, z, мы получим, очевидно, циркуляцию вектора $\vec{a}$ по контуру бесконечно малого прямоугольника, положительная нормаль к которому направлена по оси OX или по оси OY:

при $\vec{n} \| OX$

$$dC = \int a_s ds = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) dS,$$

при $\vec{n} \| OY$

$$dC = \int a_s ds = \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) dS. \qquad (7.16)$$

Фигурирующие в формулах (7.15) и (7.16) комбинации производных проекций вектора $\vec{a}$ являются, как мы покажем, проекциями некоторого вектора, который принято обозначать через $rot\vec{a}$:

$$rot_x \vec{a} = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right), \quad rot_y \vec{a} = \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right), \quad rot_z \vec{a} = \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \qquad (7.17)$$

Вектор $rot\vec{a}$ может быть назван векторной пространственной производной вектора $\vec{a}$ (в отличие от скалярной пространственной производной $div\vec{a}$). Заметим, что $rot\vec{a}$ может быть представлен в форме символического определителя:

$$rot\vec{a} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix},$$

где $\vec{i}, \vec{j}, \vec{k}$ суть единичные векторы, направленные по осям x, y, z.

С помощью обозначений (7.17) выражения (7.15) и (7.16) могут быть записаны следующим образом:

$$dC = \int_L a_s ds = rot_n\vec{a}\ dS \tag{7.18}$$

причем под $\vec{n}$ нужно понимать положительную нормаль к площадке dS, составляющую правовинтовую систему с направлением положительного обхода контура этой площадки.

Перейдем теперь к рассмотрению циркуляции вектора по контуру произвольной формы и размера. Проведем поверхность S так, чтобы она опиралась на контур L, т.е. чтобы этот контур являлся пограничным контуром поверхности S. Разобьем затем эту поверхность двумя взаимно перпендикулярными системами параллельных линий на совокупность бесконечно малых элементов (см. рис. 29), которые, благодаря своей малости, могут считаться плоскими. Применив к каждому из этих элементов уравнение (7.18) и сложив полученные выражения, найдем:

$$\sum dC = \sum \int a_s ds = \sum rot_n\vec{a}\ dS = \int_S rot_n\vec{a}\ dS,$$

где $\vec{n}$ есть внешняя нормаль к dS, причем внешняя сторона поверхности S должна быть выбрана в соответствии с направлением положительного обхода ее контура (правовинтовая система).

При интегрировании по контурам элементарных площадок каждая граница ***AB*** двух смежных площадок пройдется два раза и притом в противоположных направлениях; поэтому в сумме $\sum\int a_s ds$ встретятся оба члена $\int_A^B a_s ds$ и $\int_B^A a_s ds$, в совокупности дающие ноль.

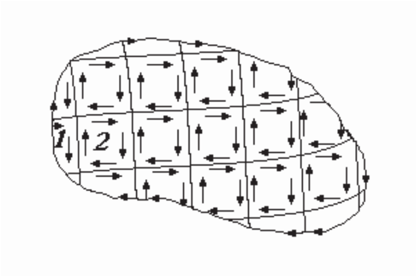

Рис. 29. Циркуляция вектора по контуру произвольной формы

Таким образом, $\sum\int a_s ds$ сведется к сумме членов, относящихся к одним лишь наружным границам площадок, т.е. к интегралу вектора $\vec{a}$ по внешнему контуру L площади S, откуда

$$\sum dc = \sum\int a_s ds = \int_L a_s ds = C,$$

где ***C*** означает циркуляцию вектора $\vec{a}$ по контуру L. Внося это выражение в предшествующее уравнение, получим

$$C = \int_L a_s ds = \int_S rot_n \vec{a}\, dS \tag{7.19}$$

При выводе этой формулы мы не приняли во внимание, что наружные (прилегающие к контуру L) элементарные площадки, вообще говоря, не будут иметь прямоугольной формы, тогда как справедливость уравнения (7.18) доказана нами лишь для площадок прямоугольных. Однако при неограниченном уменьшении размера прямоугольников ломаная линия, составленная из наружных сторон крайних прямоугольников, сколь угодно точно совпадает с контуром L площади S. Основываясь на этом, можно придать выводу уравнения (7.19) совершенно точную форму.

Таким образом, единственное условие справедливости уравнения (7.19)состоит в требовании непрерывности и дифференцируемости вектора $\vec{a}$ во всех точках поверхности S.

Уравнение это выражает собой так называемую теорему Стокса, которая гласит: *циркуляция произвольного вектора $\vec{a}$ по замкнутой кривой L равна потоку ротора этого вектора через поверхность S, опирающуюся на кривую L.*

Форма поверхности S при этом остается совершенно неопределенной. Стало быть, через любые две поверхности S_1 и S_2, если только они обладают одним и тем же контуром L, проходит одинаковый поток ротора вектора $\vec{a}$, равный циркуляции этого вектора по общему контуру этих поверхностей.

Из уравнения (7.19), между прочим, сразу следует, что

$$\int rot_n\vec{a}\ dS = 0 \qquad (7.20)$$

так как в случае замкнутой поверхности S контур L стягивается в точку, следовательно, $\boldsymbol{C} = 0$.

Переходя от уравнения (7.19) обратно к столь малому элементу поверхности dS, что его можно считать плоской площадкой, во всех точках которой $rot\vec{a}$ сохраняет постоянное значение, мы сможем вынести $rot\vec{a}$ за знак интеграла и написать: $dC = \int a_s ds = rot_n\vec{a}\ dS$, что совпадает с уравнением (7.18). Поскольку уравнение (7.19) применимо к поверхности любой формы, постольку и формула (7.18) так же применима к бесконечно малым площадкам любой формы. Так как эта формула справедлива лишь в предельном случае бесконечно малой поверхности, то правильнее записать ее следующим образом:

$$rot_n\vec{a} = \lim_{dS\to 0}\frac{\int a_s ds}{dS} \qquad (7.21)$$

Таким образом, проекция вектора $rot\vec{a}$ в данной точке поля $\boldsymbol{P}$ по данному направлению $\vec{n}$ равна пределу отношения циркуляции вектора $\vec{a}$ по контуру произвольной площадке dS, проходящей через $\boldsymbol{P}$ и перпендикулярной к $\vec{n}$, к поверхности этой площадки dS.

Отсюда следует, что значение проекции $rot\,\vec{a}$ вовсе не зависит от выбора системы координат, т.е. $rot\vec{a}$ действительно является истинным вектором.

7.6. Оператор набла. Вторые производные. Производные от произведений

Выше мы ознакомились с рядом дифференциальных операций над векторами и скалярами: образование градиента скаляра (7.6), дивергенции вектора (7.13), ротора вектора (7.21) и т.д. При применении векторного анализа приходится встречаться с целым рядом других дифференциальных выражений.

Оперирование этими выражениями может быть упрощено и уложено в простую и стройную схему введением в рассмотрение *символического дифференциального оператора Гамильтона*. Оператор этот обозначается знаком ∇; в декартовой системе координат он имеет вид:

$$\nabla = \vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} \tag{7.22}$$

где $\vec{i}$, $\vec{j}$, $\vec{k}$ - единичные векторы по осям OX, OY, OZ. Иными словами, ∇ есть векторный оператор, проекции которого на оси координат равны

$$\nabla_x = \frac{\partial}{\partial x}, \quad \nabla_y = \frac{\partial}{\partial y}, \quad \nabla_z = \frac{\partial}{\partial z} \tag{7.23}$$

Так, например, действие ∇ на скаляр ϕ нужно положить равным

$$\nabla\phi = \vec{i}\,\frac{\partial\phi}{\partial x} + \vec{j}\,\frac{\partial\phi}{\partial y} + \vec{k}\,\frac{\partial\kappa}{\partial z}.$$

Стало быть, согласно (7.6),

$$\nabla\phi = grad\phi. \qquad (7.24)$$

Таким образом, $\nabla\phi$ действительно может быть названа пространственной производной от ϕ, ибо вектор $grad\phi$ вполне характеризует изменения, испытываемые скаляром ϕ при перемещении «точки наблюдения» (т.е. при изменении координат x, y, z). Подобно этому, и другие выражения, включающие в себя оператор ∇, тоже характеризуют собой те или иные соотношения между значениями скалярных и векторных функций в смежных точках пространства.

С известными ограничениями, о которых будет сказано ниже, можно образовывать произведения ∇ с другими векторами и скалярами так, как если бы ∇ был истинным, а не символическим вектором.

Так, например, скалярное произведение символического вектора ∇ на произвольный вектор $\vec{a}$ равно

$$\nabla \cdot \vec{a} = \nabla_x a_x + \nabla_y a_y + \nabla_z a_z = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z},$$

т.е., согласно (7.10),

$$\nabla\vec{a} = div\vec{a} \qquad (7.25)$$

Помимо скалярного произведения символического вектора ∇ на вектор $\vec{a}$, можно образовать и векторное произведение этих векторов, которое, как легко заметить, представляет собой ротор вектора $\vec{a}$

$$[\nabla \times \vec{a}] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix} = rot\vec{a} \qquad (7.26)$$

Применение оператора ∇ весьма упрощает нахождение вторых и старших производных от скалярных и векторных величин. Так, например, квадрат вектора ∇ равен

$$\nabla^2 = \nabla_x^2 + \nabla_y^2 + \nabla_z^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Оператор ∇^2 часто называют лапласианом. Действие лапласиана на скаляр сводится к следующему

$$\nabla^2\phi = \nabla(\nabla\phi) = \mathit{div}\ \mathit{grad}\ \phi\ =\ \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} \qquad (7.27)$$

Совершенно иной смысл имеет выражение $\mathit{graddiv}\vec{a}$:

$$\mathit{grad}\ \mathit{div}\ \vec{a} = \nabla(\nabla\vec{a}) = \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z}\right)\left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}\right).$$

Оно вовсе не равно $\nabla^2\vec{a}$, подобно тому, как при оперировании с обычными векторами $\vec{b}\left(\vec{b}\cdot\vec{a}\right) \neq b^2\vec{a}$.

Выражение же $\nabla^2\vec{a}$ имеет, очевидно, следующий смысл:

$$\nabla^2\vec{a} = \left(\nabla\cdot\nabla\right)\vec{a} = \frac{\partial^2\vec{a}}{\partial x^2} + \frac{\partial^2\vec{a}}{\partial y^2} + \frac{\partial^2\vec{a}}{\partial z^2} \qquad (7.28)$$

то есть представляет собой вектор, Проекция которого на ось *X*, например, равна

$$\left(\nabla^2\vec{a}\right)_x = \nabla^2 a_x = \frac{\partial^2 a_x}{\partial x^2} + \frac{\partial^2 a_x}{\partial y^2} + \frac{\partial^2 a_x}{\partial z^2} \qquad (7.29)$$

Конечно, $\nabla^2\phi$ и $\nabla^2\vec{a}$ нельзя смешивать с $\left(\nabla\phi\right)^2$ и $\left(\nabla\vec{a}\right)^2$; так, например,

$$\left(\nabla\phi\right)^2 = \left(\mathit{grad}\ \phi\right)^2 = \left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2.$$

Известные формулы векторной алгебры

$$\left[\vec{b}\times\left(\vec{b}\phi\right)\right] = 0, \quad \vec{b}\left[\vec{b}\times\vec{a}\right] = 0, \quad \left[\vec{b}\times\left[\vec{b}\times\vec{a}\right]\right] = \vec{b}\left(\vec{b}\cdot\vec{a}\right) - \left(\vec{b}\cdot\vec{b}\right)\vec{a}$$

остаются справедливыми и при замене вектора $\vec{b}$ символическим вектором ∇ (при любых $\vec{a}$ и ϕ):

$$\left[\nabla\left(\nabla\phi\right)\right]=\left[\nabla grad\phi\right]=rot\ grad\phi=0,$$

$$\nabla\left[\nabla\vec{a}\right]=\nabla rot\ \vec{a}=div\ rot\ \vec{a}=0, \qquad (7.30)$$

$$\left[\nabla\left[\nabla\times\vec{a}\right]\right]=\nabla\left(\nabla\vec{a}\right)-\nabla^2\vec{a}, \text{ или } rot\ rot\ \vec{a}=grad\ div\ \vec{a}-\nabla^2\vec{a}$$

В справедливости этих выражений легко убедиться непосредственным вычислением в декартовых координатах. Так, например,

$$div\ rot\ \vec{a}=\frac{\partial rot_x\vec{a}}{\partial x}+\frac{\partial rot_y\vec{a}}{\partial y}+\frac{\partial rot_z\vec{a}}{\partial z}=$$

$$=\frac{\partial}{\partial x}\left(\frac{\partial a_z}{\partial y}-\frac{\partial a_y}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial a_x}{\partial z}-\frac{\partial a_z}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial a_y}{\partial x}-\frac{\partial a_x}{\partial y}\right)=0.$$

Итак, поскольку оператор ∇ входит сомножителем в произведения, содержащие в себе один-единственный истинный скаляр или вектор, постольку произведения эти можно преобразовывать по обычным правилам векторной алгебры. Однако, если в произведение входят два или несколько истинных скаляров или векторов, то правила эти становятся неприменимыми и нуждаются в видоизменениях. Подобно тому, как в обычном анализе действие дифференциала на произведение дифференцируемых функций выражается в виде $d(\phi\psi)=\psi\cdot d\phi+\phi\cdot d\psi$, так и в случае действия ∇ на произведение скаляров или векторов операция дифференцирования должна быть выполнена над каждым из сомножителей в отдельности. Так, например, при дифференцировании произведения двух скаляров или скаляра и вектора получаем

$$\begin{aligned}
\nabla(\phi\psi)&=\psi(\nabla\phi)+\phi(\nabla\psi) &&\Leftrightarrow\quad \mathrm{grad}(\phi\psi)=\psi\,\mathrm{grad}\,\phi+\phi\,\mathrm{grad}\,\psi\\
\nabla(\phi\vec{a})&=\phi(\nabla\vec{a})+\vec{a}(\nabla\phi) &&\Leftrightarrow\quad \mathrm{div}(\phi\vec{a})=\phi\,\mathrm{div}\,\vec{a}+\vec{a}\,\mathrm{grad}\,\phi \qquad (7.31)\\
\left[\nabla\times(\phi\vec{a})\right]&=\phi\left[\nabla\times\vec{a}\right]+\left[(\nabla\phi)\times\vec{a}\right] &&\Leftrightarrow\quad \mathrm{rot}(\phi\vec{a})=\phi\,\mathrm{rot}\,\vec{a}+\left[\mathrm{grad}\,\phi\times\vec{a}\right]
\end{aligned}$$

Несколько сложнее обстоит дело при скалярном дифференцировании произведения двух векторов.

Отсылая за доказательствами к курсам векторного анализа, приведем соответствующие формулы:

$$\nabla\left[\vec{a}\times\vec{b}\right]=\vec{b}\left[\nabla\times\vec{a}\right]-\vec{a}\left[\nabla\times\vec{b}\right]$$

$$\operatorname{div}\left[\vec{a}\times\vec{b}\right]=\vec{b}\cdot\operatorname{rot}\vec{a}-\vec{a}\cdot\operatorname{rot}\vec{b}\quad .$$

$$\operatorname{grad}(\vec{a}\cdot\vec{b})=\left(\vec{b}\cdot\nabla\right)\vec{a}+\left(\vec{a}\cdot\nabla\right)\vec{b}+\left[\vec{b}\times\operatorname{rot}\vec{a}\right]+\left[\vec{a}\times\operatorname{rot}\vec{b}\right]$$

$$\operatorname{rot}\left[\vec{a}\times\vec{b}\right]=\left(\vec{b}\cdot\nabla\right)\vec{a}-\left(\vec{a}\cdot\nabla\right)\vec{b}+\vec{a}\operatorname{div}\vec{b}-\vec{b}\operatorname{div}\vec{a} \tag{7.32}$$

Элементарные операции пространственного дифференцирования сводятся к образованию градиента, дивергенции и ротора. Все эти операции имеют определенный геометрический смысл и потому инвариантны по отношению к преобразованию системы координат. Иными словами, значение выражений $\operatorname{grad}\phi$, $\operatorname{div}\vec{a}$, $\operatorname{rot}\vec{a}$ не зависит от выбора системы координат. Все соотношения между выражениями, выведенные нами выше, тоже носят инвариантный характер, хотя при доказательстве их мы всякий раз пользовались декартовой системой координат. Однако во все соотношения входят лишь инвариантные выражения для градиента, дивергенции, ротора и т.д. Стало быть, форма этих соотношений не может измениться при переходе к иным системам координат.

7.7. Интегральные соотношения. Теорема Грина

Формулы Гаусса (7.12) и Стокса (7.19) представляют собой основные интегральные соотношения векторного анализа; исходя из них, можно получить и ряд других важных соотношений между пространственными (объемными, поверхностными и линейными) интегралами скалярных и векторных величин.

Формула Гаусса (7.12) позволяет без труда доказать важную для векторного анализа и его приложений *теорему Грина*. Для этой цели в формуле Гаусса

$$\int \operatorname{div}\vec{a}\, dV = \int a_n dS$$

положим: $\vec{a} = \psi \cdot \operatorname{grad}\phi = \psi \nabla \phi$, где ϕ и ψ - два произвольных скалярных поля. Согласно (7.31) и (7.27)

$$div\vec{a} = \psi \cdot div\ grad\phi + grad\psi \cdot grad\phi = \psi \nabla^2 \phi + (\nabla \phi)(\nabla \psi).$$

Далее, $a_n = \psi \cdot grad_n \phi = \psi \dfrac{\partial \psi}{\partial n}$. Поэтому из (7.12) следует, что

$$\int \left\{ \psi\ \nabla^2 \phi + (\nabla \phi)(\nabla \psi) \right\} dV = \int \psi\ \frac{\partial \phi}{\partial n} dS \qquad (7.33)$$

где интеграл правой части должен быть взят по замкнутой поверхности S, ограничивающей область интегрирования V. Эта формула и выражает собой *теорему Грина*. Для некоторых целей удобно преобразовать формулу (7.33), заменив в ней ψ на ϕ, и обратно; вычтя полученное таким образом уравнение из (7.33), получаем

$$\int (\psi\ \nabla^2 \phi - \phi\ \nabla^2 \psi) dV = \int (\psi \frac{\partial \phi}{\partial n} - \phi \frac{\partial \psi}{\partial n}) dS \qquad (7.34)$$

Как указывалось, применение теоремы Гаусса ограничено требованием непрерывности вектора $\vec{a}$ и конечности его первых производных в области интегрирования V. Поэтому теорема Грина непосредственно применима лишь к конечным и непрерывным скалярным функциям точки ϕ и ψ, обладающими в области интегрирования V производными первого и второго порядков.

7.8. Важнейшие формулы векторного анализа

$\dfrac{\partial \phi}{\partial s} = \lim\limits_{ds \to 0} \dfrac{\phi_S - \phi_0}{ds}$ (производная скаляра ϕ по направлению $\vec{s}$)

$$grad\phi = \frac{\partial \phi}{\partial n} \vec{n}$$

$$\nabla \phi = grad\phi = \vec{i}\, \frac{\partial \phi}{\partial x} + \vec{j}\, \frac{\partial \phi}{\partial y} + \vec{k}\, \frac{\partial \phi}{\partial z}$$

$$grad_s\phi = \frac{\partial\phi}{\partial s}$$

$$grad\phi(\psi) = \frac{\partial\phi}{\partial\psi} grad\psi$$

$$\int_S a_n dS = \int_V div\vec{a} dV \qquad \text{(теорема Гаусса)}$$

$$div\vec{a} = \lim_{\Delta V \to 0} \frac{\int a_n dS}{\Delta V}$$

$$\nabla\vec{a} = div\vec{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}$$

$$\int_L a_s ds = \int_S rot_n\vec{a} dS \qquad \text{(теорема Стокса)}$$

$$rot_n\vec{a} = \lim_{dS \to 0} \frac{\int a_s ds}{dS}$$

$$[\nabla \times \vec{a}] = rot\vec{a} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix}$$

$$\int_S rot_n\vec{a} dS = 0$$

Вторые производные

$$div\ grad\phi = \nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}$$

$$\nabla^2\vec{a} = \frac{\partial^2\vec{a}}{\partial x^2} + \frac{\partial^2\vec{a}}{\partial y^2} + \frac{\partial^2\vec{a}}{\partial z^2}$$

$$rot\ grad\phi = 0$$

$$div\ rot\ \vec{a} = 0$$

$$rot\ rot\ \vec{a} = grad\ div\ \vec{a} - \nabla^2\vec{a}$$

Производные от произведений

$$grad(\phi\psi) = \psi\ grad\phi + \phi\ grad\psi$$

$$div(\phi\vec{a}) = \phi\ div\ \vec{a} + \vec{a}\ grad\ \phi$$

$$rot(\phi\vec{a}) = \phi\ rot\ \vec{a} + \left[grad\phi \times \vec{a}\right]$$

$$div\left[\vec{a} \times \vec{b}\right] = \vec{b}\ rot\vec{a} - \vec{a}\ rot\vec{b}$$

$$grad(\vec{a} \cdot \vec{b}) = (\vec{b}\nabla)\vec{a} + (\vec{a}\nabla)\vec{b} + \left[\vec{b} \times rot\vec{a}\right] + \left[\vec{a} \times rot\vec{b}\right]$$

$$rot\left[a \times b\right] = (\vec{b}\nabla)\vec{a} - (\vec{a}\nabla)\vec{b} + \vec{a}\ div\vec{b} - \vec{b}\ div\vec{a}$$

Теорема Грина

$$\int\left\{\ \psi\ \nabla^2\phi + (\nabla\phi)(\nabla\psi)\right\}\ dV = \int\psi\frac{\partial\phi}{\partial n}dS$$

ЛИТЕРАТУРА

1. М.Борн, Э.Вольф. Основы оптики. –М.: Наука, 1973, 721 с.
2. В.А.Ильин, Э.Г.Поздняк. Основы математического анализа. Часть 2. – М.: Наука, 2000, 447 с.
3. А.Н.Тихонов, А.Б.Васильева, А.Г.Свешников. Дифференциальные уравнения. –М.: Наука, 1998, 231 с.
4. Л.Д.Ландау, Е.М.Лифшиц. Электродинамика сплошных сред. –М.: Наука, 1992, 661 с.
5. Воеводин В.В. Линейная алгебра (2-е изд.). Наука, 1980. - 400 стр.
6. Д. Маркузе «Оптические волноводы» Под ред. В.В. Шевченко. М., «Мир» 1974.
7. Виноградова М.Б., Руденко О.В., Сухоруков А.П. Теория волн. – М., «Наука» 1979.
8. Тамм И.Е. Основы теории электричества. М., «Наука» 1989.

Цели и задачи магистерской программы «Оптика наноструктур»

Целью учебно-методического комплекса магистерской программы «Оптика наноструктур» является формирование у студентов четкого представления об основных принципах функционирования современных дифракционных оптических элементов и устройств, тонкопленочных многослойных покрытий; о законах взаимодействия электромагнитного излучения видимого диапазона с материалом. Целью является также изучение способов и возможностей математического синтеза и компьютерного проектирования дифракционных оптических покрытий. Полученные знания закрепляются в оптической лаборатории и дисплейном классе на примерах изучения конкретных дифракционных оптических элементов и многослойных покрытий со сложной геометрией.

Задачей учебно-методического комплекса магистерской программы «Оптика наноструктур» является обучение студентов навыкам самостоятельного анализа технических заданий на проектирование дифракционных оптических элементов и устройств. Они должны научиться выбирать из имеющихся в наличии алгоритмов и программ математического синтеза или разрабатывать их самостоятельно. В результате обучения обретут навыки ориентации в научной и бизнес информации с целью выбора нужной функции или нужного инструмента для реализации известной функции в области проектирования и создания дифракционных оптических наноструктур.

Цели и задача курса
«Методы и алгоритмы решения задач в моделях оптических покрытий»

Курс «Методы и алгоритмы решения задач в моделях оптических покрытий» является составной частью магистерской программы «Оптика наноструктур». Магистерская программа «Оптика наноструктур» реализуется в рамках направления «Прикладная математика и информатика» и направления «Прикладная математика и физика», а возможно и других направлений. В составе магистерской программы «Оптика наноструктур» курс «Методы и алгоритмы решения задач в моделях оптических покрытий» является обязательным, привязанным к семестру. Для других магистерских программ этот курс может быть курсом по выбору без привязки к семестру или факультативным на усмотрение методической комиссии программы. Курс носит теоретический и практический характер.

Целью курса является подробное ознакомление студентов с устойчивыми современными методами численного решения математических задач, возникающих при изучении взаимодействия электромагнитного излучения в области светового диапазона с веществом, в особенности с наноструктурами. Эта область знаний особенно быстро развивается в последние годы в связи с широким применением наноэлементов и тонких (менее одного микрометра толщиной) пленок, используемых в производстве жидко-кристаллических дисплеев, солнечных батарей на основе диэлектриков, фотоэмиссионных диодов, просветляющих покрытий, поляризаторов, миниатюрных лазеров, управляемых оптических элементов. Задачи оптики наноструктур практически не поддаются аналитическому решению, поэтому важным является не только освоение теоретического материала, но и изучение эффективных численных методов, используемых при решении данного класса задач, приобретение навыков создания программного обеспечения для численного моделирования различных оптических наноструктур.

Задачей курса «Методы и алгоритмы решения задач в моделях оптических покрытий» является формирование у студентов навыков работы на современной измерительной аппаратуре. Задачей курса является также обучение студентов использованию строгих методов связанных волн при решении задач моделирования современных оптических устройств на основе тонкопленочных покрытий и дифракционных оптических элементов. Это позволит им при необходимости разрабатывать новое программное обеспечение. Безусловной задачей курса является также освоение существующего программного обеспечения, ориентированного на расчет и проектирование оптических покрытий. В результате обучения они получат умение и навыки правильно оценить сложность научно-исследовательских и конструкторских заданий на разработку дифракционных оптических элементов и устройств, аргументированно выбирать метод решения конструкторской задачи, а затем экономично и эффективно выполнять компьютерный дизайн требуемого дифракционного оптического покрытия или устройства.

Трудоемкость курса составляет 3 кредита; 2 часа лекций и 2 часа лабораторных занятий в дисплейном классе в неделю.

Инновационность курса.

Курс является инновационным по содержанию и по литературе, он включает в себя последние научные достижения в области решения задач дифракционной оптики, когда характерные размеры исследуемых объектов не превышают либо сравнимы с длиной волны оптического излучения. Эта область знаний интенсивно развивалась в последнее время, но лишь недавно были созданы устойчивые алгоритмы и разработаны численные методы решения задач для многослойных решеток. Следует отметить, что для оптических однослойных и многослойных решеток с характерными размерами больше длины волны оптического излучения устойчивые методы решения известны с середины прошлого века. Сейчас алгоритмы решения оптических задач в субволновой области распространяются на объекты со сложной геометрией, такие как двумерные решетки с произвольным профилем, трехмерные решетки (фотонные кристаллы) и на анизотропные материалы. Они востребованы, поскольку позволяют создавать математические модели взаимодействия излучения с веществом в наномасштабах, а затем с их помощью проектировать новые эффективные устройства в высокотехнологичных областях медицины, энергетики, инфокоммуникаций и приборостроения.

В ходе проведения занятий по этому курсу разработчики предполагают использование традиционных методик преподавания, принятой в странах болонской системы образования, то есть с использованием кредитной системы оценки знаний.

Наряду с традиционными элементами преподавания математических методов решения прикладных задач, разработчики курса предполагают воспользоваться хорошо зарекомендовавшим себя опытом МФТИ и подобных вузов. Для этого в рамках подпрограммы «Оптика наноструктур» осуществляется закупка уникального измерительного и аналитического оборудования для выполнения измерений разнообразных характеристик оптических наноустройств с целью использования этого оборудования в учебном процессе и для проведения научно-исследовательских работ преподавателями, аспирантами и студентами.

По окончании магистратуры по направлению «Оптика наноструктур» выпускники Российского университета дружбы народов станут конкурентно-способными специалистами в области проектирования современных оптических устройств, которые не будут испытывать затруднений при последующем трудоустройстве.

Данное направление научно-практических разработок сформировалось лишь в последние 10 – 15 лет. Поэтому наблюдается сильный дефицит учебно-методической литературы не только в России, но и во всем мире. Разрабатываемые в рамках инновационной программы «Оптика наноструктур» учебные пособия восполнят в некоторой степени этот пробел и составят основной список литературы для слушателей курсов. Вместе с ними следует использовать несколько учебников и монографий, вышедших в свет к настоящему времени и перечисленные в списке литературы. Курс базируется на публикациях научных статей мировых лидеров исследований в данной области в научной периодике, диссертационных работах их учеников, включающих работы по непосредственному моделированию, дизайну и последующему изготовлению лабораторных образцов оптических элементов и устройств. В список дополнительной и рекомендуемой литературы включены все научно-исследовательские публикации, положенные в основу предлагаемого курса.

В качестве практических заданий, курсовых работ и тем рефератов слушателям магистерской программы будут предложены актуальные проблемы и задачи, решение которых востребовано современным уровнем развития высокотехнологичных отраслей промышленности и научно-исследовательских лабораторий.

Структура курса «Методы и алгоритмы решения задач в моделях оптических покрытий» (с указанием количества часов аудиторных /самостоятельной работы на темы)

Темы лекций

Тема 1. Основы моделирования оптических структур. Геометрическая оптика, скалярная волновая оптика, векторная оптика уравнений Максвелла. Примеры оптических изображений и приборов. Последние достижения в моделировании оптических устройств, впечатляющие успехи в технологии создания оптических структур (LCD, солнечные батареи, LED – фотоэмиссионные диоды, оптические покрытия со специальными свойствами – просветляющие покрытия, фильтры, поляризаторы, управляемые оптические структуры), бабочка Morpho, хвост павлина. (1 пара)

Тема 2. Необходимые формулы векторной алгебры и векторного анализа.

Система уравнений Максвелла, материальные уравнения связи. Интегральная формулировка уравнений Максвелла, граничные условия. Волновые уравнения для напряженностей электрического и магнитного полей. Решения волнового уравнения, монохроматическое поле и его характеристики. Формализм комплексных амплитуд. Уравнение Гельмгольца. Наблюдаемые величины при сложении полей. (1 пара)

Тема 3. Электромагнитная плоская волна общего вида. Вектор Пойнтинга и закон сохранения энергии электромагнитного поля. Поперечность плоских монохроматических электромагнитных волн в однородной среде, волновой вектор, показатель преломления.

Поляризация плоских монохроматических электромагнитных волн. Линейная и круговая поляризация. Представление комплексного параметра. Вектор Джонса. (2 пары)

<u>Тема 4.</u> Прохождение света через границу раздела двух сред: закон преломления, закон отражения, полное внутреннее отражение.

Положительный и отрицательный индекс преломления, метаматериалы.

Формулы Френеля, соотношения между амплитудами падающих, преломленных и отраженных волн. Нормальное падение, угол Брюстера, просветление оптики. (2 пары)

<u>Тема 5.</u> Распространение света в анизотропных средах. Диэлектрический тензор анизотропной среды. Распространение плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде.

Эллипсоид показателей преломления. Фазовая скорость, групповая скорость, скорость переноса энергии. Классификация анизотропных сред. Двойное лучепреломление плоских монохроматических электромагнитных волн на границе анизотропного кристалла. (2 пары)

<u>Тема 6.</u> Распространение света в двуосных кристаллах. Оптическая активность. Фарадеево вращение. Анализ распространения плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде методом связанных мод. Уравнения движения для состояния поляризации.(2 пары)

<u>Тема 7.</u> Прохождение плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие. Вектор Джонса исследуемого образца. Описание расширенного метода Джонса 4х4. Матрицы Мюллера, вектор Стокса. Метод Берремана описания прохождения плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие. Обобщения и модификации метода Берремана по результатам работ Палто и Шуберта. (4 пары)

Темы семинарских и практических занятий

Лабораторные работы в оптической лаборатории и дисплейном классе.

1. Произведение комплексных матриц и комплексной матрицы на комплексный вектор
2. Метод LU-разложения для решения комплексных систем линейных алгебраических уравнений
3. Метод Якоби для решения комплексных систем линейных алгебраических уравнений
4. Метод Якоби для отыскания собственных векторов и собственных значений комплексных матриц.

Темы коллоквиумов и контрольных работ

- Алгоритм LU-разложения для решения комплекснозначной СЛАУ.
- Метод Якоби для решения СЛАУ с комплексными коэффициентами.
- Метод Якоби для отыскания собственных векторов и собственных значений комплексной матрицы.
- Примеры физических постановок задач, приводящих к решению СЛАУ с комплексными матрицами: схема Берремана $\hat{\varepsilon} = A\varepsilon_{diag}A^{-1}$ и решение ОДУ для определения пропускания и отражения с вычислением экспоненты от матрицы.

Описание системы контроля знаний:

Общие правила выполнения контрольных заданий;

Требования к оформлению работы

Постановка задачи.

1. Краткая формулировка задачи.
2. Развернутая постановка задачи с указанием основных режимов работы и их сценариев.

Алгоритм решения.

1. Математическое описание алгоритма.
2. Структура алгоритма ядра программы (укрупненная блок-схема).

Тестирование.

1. Описание основных режимов тестирования алгоритма и программы и результатов работы программы.
2. Список возможных ошибок и аномалий, описание реакции программы на них.

Заключение.

Содержит общие комментарии и замечания исполнителя о выполненной работе.

Приложение.

Приложение должно содержать текст программы (полная распечатка или распечатка алгоритма ядра программы).

Работа должна быть представлена в виде распечатанного текста и на дискете (Word + Delphi и/или C++).

Рекомендации к составлению отчета

Оформление.

отчет по работе должен быть оформлен в форме Word-файла.

Содержание отчета.

Каждый пункт задания вычислительного эксперимента должен найти свое отражение в отчете.

Каждый раздел отчета должен содержать:

формулировку цели эксперимента

описание исходных данных - приближаемая функция, интервал и порядок приближения, метод приближения и т.п.

результаты эксперимента, представленные в форме таблиц, гистограмм и графиков

иллюстрационный материал в виде копий экрана с графиками зависимостей погрешности приближения, вида приближаемой функции и т.п.

выводы, следующие из результатов эксперимента в контексте его цели.

Примерные типы письменных работ и форм устного контроля

Образец оформления реферата на тему «Оптимизация многослойных оптических кристаллических компонент» приведен в приложении 1.

Шкала оценок, итоговые оценки (методика выставления)

Бально-рейтинговая методика оценки уровня знаний по обязательной дисциплине «Методы и алгоритмы решения задач в моделях оптических покрытий», привязанной к семестру

Порядок начисления баллов за семестр.

Контрольная работа № 1: 0 – 40 баллов

Теоретические вопросы: 0 – 10 баллов

Практические задания: 0 – 30 баллов

Контрольная работа № 2: 0 – 40 баллов

Теоретические вопросы: 0 – 10 баллов

Практические задания: 0 – 30 баллов

Контрольная работа № 3: 0 – 20 баллов

Теоретические вопросы: 0 – 20 баллов

Шкала бально-рейтинговой системы.

Баллы за семестр	Автоматическая оценка		Баллы за итоговый контроль знаний	Общая сумма баллов	Итоговая оценка
	Итоговая оценка	Дополнительные баллы			
78 – 80	зачет	по 5 баллов за каждый свыше 76**	0 – 20*	86 – 100	зачет
41 – 77	Нет	Нет	0 – 20 0 – 20	51 – 97 41 – 50	зачет незачет
< 41	незачет	Нет	Нет	Нет	незачет

* студент имеет право не проходить итоговый контроль знаний.

** дополнительные баллы начисляются автоматически:

за 86 баллов, набранных в семестре, начисляется дополнительно 6 баллов (общая сумма баллов – 92);

за 87 баллов – 12 баллов (99);

за 88 баллов – 18 баллов (106);

за 89 баллов – 24 балла (113);

за 90 баллов – 30 баллов (120).

Академическая этика, соблюдение авторских прав.

Все имеющиеся в тексте сноски тщательно выверены и снабжены «адресами». Авторы не включали в свою работу выдержки из работ других авторов без указания на это, не пересказывали чужих работ близко к тексту без отсылки к ним. Авторы также не использовали чужих идей без указания первоисточников. Это касается и источников, найденных в интернете. В необходимых случаях указан полный адрес сайта.

Программа курса УМК

<u>Аннотированное содержание курса.</u>

Первый модуль трудоемкостью в 1 кредит составляют:

- теоретический материал следующего содержания.

Основы моделирования оптических структур. Геометрическая оптика, скалярная волновая оптика, векторная оптика уравнений Максвелла. Примеры оптических изображений и приборов. Последние достижения в моделировании оптических устройств, впечатляющие успехи в технологии создания оптических структур. Оптические наноструктуры в природе - бабочка Morpho, хвост павлина.

Необходимые формулы векторной алгебры и векторного анализа. Система уравнений Максвелла, материальные уравнения связи. Интегральная формулировка уравнений Максвелла, граничные условия. Волновые уравнения для напряженностей электрического и магнитного полей. Решения волнового уравнения, монохроматическое поле и его характеристики. Формализм комплексных амплитуд. Уравнение Гельмгольца. Наблюдаемые величины при сложении полей.

Электромагнитная плоская волна общего вида. Вектор Пойнтинга и закон сохранения энергии электромагнитного поля. Поперечность плоских монохроматических электромагнитных волн в однородной среде, волновой вектор, показатель преломления.

Поляризация плоских монохроматических электромагнитных волн. Линейная и круговая поляризация. Представление комплексного параметра. Вектор Джонса.

Прохождение света через границу раздела двух сред: закон преломления, закон отражения, полное внутреннее отражение.

Положительный и отрицательный индекс преломления, метаматериалы.

Формулы Френеля, соотношения между амплитудами падающих, преломленных и отраженных волн. Нормальное падение, угол Брюстера, просветление оптики.

- практическими занятиями в оптической лаборатории в течение 14 академических часов,
- самостоятельные занятия по написанию рефератов и выполнению курсовых работ.

В конце этого модуля проводится промежуточный контроль знаний.

Второй модуль трудоемкостью в 1 кредит составляют:

- теоретический материал на темы:

Распространение света в анизотропных средах. Диэлектрический тензор анизотропной среды. Распространение плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде.

Эллипсоид показателей преломления. Фазовая скорость, групповая скорость, скорость переноса энергии. Классификация анизотропных сред. Двойное лучепреломление плоских монохроматических электромагнитных волн на границе анизотропного кристалла.

Распространение света в двуосных кристаллах. Оптическая активность. Фарадеево вращение. Анализ распространения плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде методом связанных мод. Уравнения движения для состояния поляризации.

- практическими занятиями в дисплейном классе в течение 8 академических часов,
- самостоятельные занятия по написанию рефератов и выполнению курсовых работ.

В конце этого модуля проводится промежуточный контроль знаний.

Третий модуль трудоемкостью в 1 кредит составляют:

- теоретический материал, излагаемый в седьмой теме:

Прохождение плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие. Вектор Джонса исследуемого образца. Описание расширенного метода Джонса 4x4. Матрицы Мюллера, вектор Стокса.

Метод Берремана описания прохождения плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие.

Обобщения и модификации метода Берремана по результатам работ Палто и Шуберта.

- практическими занятиями в дисплейном классе в течение 8 академических часов,

- самостоятельные занятия по написанию рефератов и выполнению курсовых работ.

В конце этого модуля проводится итоговый контроль знаний.

<u>Список обязательной и дополнительной литературы с указанием соответствия разделов источника (постранично) разделам читаемого курса</u>

Список обязательной литературы.

1. Методы и алгоритмы решения задач в моделях оптических покрытий/Под ред. Л.А Севастьянова: Учебное пособие. - М.: Изд-во РУДН (готовится к печати).
2. Родионов С.А. Основы оптики. Конспект лекций.– СПб: СПб ГИТМО (ТУ), 2000. - 167 с.
3. Ильинский А.С., Кравцов В.В., Свешников А.Г. Математические модели электродинамики. М.: Высшая школа. 1991. 224с.
4. Ярив А., Юх П. Оптические волны в кристаллах. М. Мир.1983.
5. Терлецкий Я.П., Рыбаков Ю.П. Электродинамика. М.: Высш. шк., 1990.
6. D.W. Berreman. Numerical m0odelling of twisted nematic devices. // Phil. Trans. R. Soc. Lond. A, 1983, v.309, p203-216.
7. С.П. Палто. Алгоритм решения оптической задачи для слоистых анизотропных сред. // ЖЭТФ, 2001, т.119, с.638-648.

8. M. Schubert **Theory and Application of Generalized Ellipsometry.** In Handbook of Ellipsometry. Harland Tompkins and Eugene Irene (Editors) William Andrews Publications, Norwich, NY 2005, 875 p.

Список дополнительной литературы и источников в интернете.

9. Аззам Р., Башара Н. Эллипсометрия и поляризованный свет. М. Мир.1981.
10. D.W. Berreman. Dynamical theory of X-ray diffraction in flat, focusing and distorted crystals by Abeles's matrix method. // Phys. Rev. B, 1976, v.14, p.4313-4317.
11. М.Борн,Э.Вольф. Основы оптики. – М.: Наука, 1973, 721 с.
12. Джеррард А., Берч Дж.М.. Введение в матричную оптику. М. Мир.1978.
13. Д.В. Сивухин. Общий курс физики. Том IV. Оптика. - М.: Физматлит, 1985, 792 с.

Темы рефератов, курсовых работ, эссе

Темы рефератов.

1. Оптические дифракционные решетки (1 и 2-х мерные, 3-х мерные).
2. Жидко-кристаллические ячейки (хиральные ячейки).
3. Ячейки для солнечных батарей.
4. Фотоизлучающие диоды.
5. Зеркала – многослойные структуры.
6. Физическая постановка, приводящая к задаче диагонализации тензора диэлектрической проницаемости, метод Берремана.
7. Соотношения для коэффициентов прохождения T и коэффициентов отражения R.

8. Оптические методы измерения компонентов тензора диэлектрической проницаемости анизотропных покрытий. Метод, основанный на измерении коэффициентов отражения и пропускания (Берриман и др. – спектрофотометр.
9. Оптические методы измерения компонентов тензора диэлектрической проницаемости анизотропных покрытий. Метод, основанный на измерении фазовой скорости направляемой волны, распространяющейся вдоль анизотропного покрытия (Бикеев, Горобец, лабораторные)

<u>Учебный тематический план курса УМК (календарный план, структурированный по видам учебных занятий)</u>

Календарный план (20 недель) учебных занятий по обязательной дисциплине «Методы и алгоритмы решения задач в моделях оптических покрытий», привязанный к первому семестру магистратуры.

Виды и содержание учебных занятий				
Неделя	Лекции	Число часов	Семинарские занятия	Число часов
1	Основы моделирования оптических структур. Геометрическая оптика, скалярная волновая	2	Знакомство с приборами и оборудованием	2

	оптика, векторная оптика уравнений Максвелла. Примеры оптических изображений и приборов. Последние достижения в моделировании оптических устройств, впечатляющие успехи в технологии создания оптических структур (LCD, солнечные батареи, LED – фотоэмиссионные диоды, оптические покрытия со специальными свойствами – просветляющие покрытия, фильтры, поляризаторы, управляемые оптические структуры), бабочка Morpho, хвост павлина.			
2	Необходимые формулы векторной алгебры и векторного анализа. Система уравнений Максвелла, материальные уравнения связи.	2	Изучение работы спектрофотометра.	2

	Интегральная формулировка уравнений Максвелла, граничные условия. Волновые уравнения для напряженностей электрического и магнитного полей.			
3	Решения волнового уравнения, монохроматическое поле и его характеристики. Формализм комплексных амплитуд. Уравнение Гельмгольца. Наблюдаемые величины при сложении полей.	2	Изучение работы поляриметра	2
4	Электромагнитная плоская волна общего вида. Вектор Пойнтинга и закон сохранения энергии электромагнитного поля. Поперечность плоских монохроматических электромагнитных волн в однородной среде, волновой вектор, показатель преломления.	2	Изучение работы профилометра	2

5	Поляризация плоских монохроматических электромагнитных волн. Линейная и круговая поляризация. Представление комплексного параметра. Вектор Джонса.	2	Изучение работы поляризационного микроскопа	2
6	Прохождение света через границу раздела двух сред: закон преломления, закон отражения, полное внутреннее отражение. Положительный и отрицательный индекс преломления, метаматериалы.	2	Знакомство с программным обеспечением оптометрической аппаратуры.	2
7	Формулы Френеля, соотношения между амплитудами падающих, преломленных и отраженных волн. Нормальное падение, угол Брюстера, просветление оптики.	2	Обзор численных методов решения (прямых) оптических задач.	2
8	Промежуточный контроль знаний (Контрольная работа №1)			2

9	Распространение света в анизотропных средах. Диэлектрический тензор анизотропной среды. Распространение плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде.	2	Методы сопряжения программного обеспечения с приборами. Форматы и структура измеренных данных	2
10	Эллипсоид показателей преломления. Фазовая скорость, групповая скорость, скорость переноса энергии. Классификация анизотропных сред. Двойное лучепреломление плоских монохроматических электромагнитных волн на границе анизотропного кристалла.	2	Комплексная арифметика компилятора Delphi.	2
11	Распространение света в двуосных кристаллах. Оптическая активность. Фарадеево вращение	2	Методы решения СЛАУ над полем комплексных чисел	2

12	Анализ распространения плоских монохроматических электромагнитных волн в диэлектричеки анизотропной среде методом связанных мод. Уравнения движения для состояния поляризации.	2	Определение собственных значений и собственных векторов комплекснозначных матриц	2
13		2	Визуализация результатов измерений и расчетных характеристик.	2
14	Промежуточный контроль знаний (Контрольная работа №2)			2
15	Прохождение плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие.	2	Произведение комплексных матриц и комплексной матрицы на комплексный вектор	2
16	Вектор Джонса исследуемого образца. Описание расширенного	2	Метод LU-разложения для решения	2

	метода Джонса 4х4. Матрицы Мюллера, вектор Стокса.		комплексных систем линейных алгебраических уравнений	
17	Метод Берремана описания прохождения плоской монохроматической поляризованной электромагнитной волны через однослойное или многослойное изотропное или анизотропное оптическое покрытие.	2	Метод Якоби для решения комплексных систем линейных алгебраических уравнений	2
18	Обобщения и модификации метода Берремана по результатам работ Палто и Шуберта	2	Метод Якоби для отыскания собственных векторов и собственных значений комплексных матриц.	2
19	Заключительный обзор курса. Консультации по подготовке к итоговому контролю знаний.	2	Заключительный обзор курса. Консультации.	2
20	Итоговый контроль знаний			2