



Г.Г. Куц  
Ж.М. Соколова  
Л.И. Шангина

# ПРИБОРЫ И УСТРОЙСТВА ОПТИЧЕСКОГО И СВЧ-ДИАПАЗОНОВ



ТОМСКИЙ МЕЖВУЗОВСКИЙ ЦЕНТР ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ

**Г.Г. Куц, Ж.М. Соколова, Л.И. Шангина**

**ПРИБОРЫ И УСТРОЙСТВА  
ОПТИЧЕСКОГО И СВЧ  
ДИАПАЗОНОВ**

**Учебное пособие**

Рекомендовано  
Сибирским региональным отделением  
учебно-методического объединения высших учебных заведений РФ  
по образованию в области радиотехники, электроники,  
биомедицинской техники и автоматизации  
для межвузовского использования в качестве учебного пособия  
для студентов, обучающихся  
по направлению подготовки 654200 «Радиотехника »  
и 654400 «Телекоммуникации»

**2012**

Министерство образования Российской Федерации

**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ  
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)**

**Г.Г. Куц, Ж.М. Соколова, Л.И. Шангина**

# **ПРИБОРЫ И УСТРОЙСТВА ОПТИЧЕСКОГО И СВЧ ДИАПАЗОНОВ**

**Учебное пособие**

Томск 2012

УДК 621.371(075.8) + 537.8(075.8)

Рецензенты:

**А.А. Тихомиров** институт оптического мониторинга Томского СОРАН,  
зам. директора, профессор д-р. тех. наук

**А.П. Коханенко** д-р физ.-мат. наук, проф. каф. квантовой  
электроники и оптоинформатики Томск. гос. ун-та:

**А.Е. Мудров**, проф., зав. каф. электроники Томск. гос. ун-та:

**Куц Г.Г., Соколова Ж.М., Шангина Л.И.**

Приборы и устройства оптического и СВЧ диапазонов Учебное пособие. –  
Томск: Томск. Гос. ун-т систем управления и радиоэлектроники, 2012. - 414 с.

В учебном пособии рассматриваются принципы работы, схемы устройств электровакуумных, полупроводниковых, квантовых приборов СВЧ и оптического диапазонов, общие вопросы оптоволоконных и диэлектрических волноводов. Значительное внимание в книге уделено: физике рассматриваемых явлений в оптоэлектронных и СВЧ устройствах; изучению зависимости и оптимизации параметров и характеристик приборов и устройств от геометрии и режима работы.

Учебное пособие предназначено для студентов, обучающихся по направлению подготовки 654200 (210300) «Радиотехника» и 654400 (210400) «Телекоммуникации» по дисциплинам «Оптоэлектронные и квантовые приборы и устройства», «Устройства СВЧ и антенны».

© Куц Г.Г., Соколова Ж.М., Шангина Л.И., 2012

© Томск. гос. ун-т систем управления  
и радиоэлектроники, 2012

## ОГЛАВЛЕНИЕ

Предисловие.....	10
<b>РАЗДЕЛ ПЕРВЫЙ</b>	
<b>Электронные приборы СВЧ</b>	
<b>Глава 1. Физические основы электронных приборов СВЧ.....</b>	<b>11</b>
1.1. Характеристики и особенности электровакуумных приборов ...	12
1.2. Основные уравнения и параметры электронных приборов СВЧ.....	12
1.3. Электродинамические системы приборов СВЧ.....	20
1.3.1. Резонаторы .....	21
1.3.2.Замедляющие системы, их параметры и характеристики... ..	26
1.3.3. Разновидности замедляющих систем.....	33
1.4. Принципы модуляции электронного потока.....	39
1.5. Отбор энергии от электронного потока .....	44
1.6. Волны в электронных потоках.....	46
1.7. Фазовые и амплитудные условия усиления и генерации сигналов.....	48
<b>Глава 2. Электровакуумные приборы СВЧ.....</b>	<b>51</b>
2.1.Пролетные усилительные клистроны .....	51
2.2. Отражательные клистроны .....	63
2.3. Лампы бегущей и обратной волны О – типа.....	64
2.3.1.Элементы линейной теории ЛБВ .....	65
2.3.2 Лампы бегущей волны типа О .....	67
2.3.3 Лампы обратной волны типа О .....	68
2.4. Электронные приборы со скрещенными полями .....	71
2.4.1 Движение электронов в скрещенных полях .....	71
2.4.2 Многорезонаторные магнетроны .....	73
2.4.3 Лампы бегущей и обратной волны типа М.....	90
<b>Глава 3. Полупроводниковые приборы СВЧ .....</b>	<b>93</b>
3.1.Элементы зонной теории твердого тела.....	93
3.2. Лавинно-пролетные диоды .....	104
3.3. Диоды Ганна .....	115
<b>Глава 4. Применение и регистрация СВЧ энергии .....</b>	<b>124</b>
4.1. Нагрев СВЧ энергией .....	124
4.1.1. Поведение диэлектриков в электромагнитных полях.....	124
4.1.2. Электродинамические системы нагрева диэлектриков полями бегущих волн.....	128
4.1.3.Нагрев в поле стоячей волны.....	135
4.2. Измерение СВЧ энергии .....	146
4.2.1. Калориметрические методы.....	146
4.2.2. Боллометрические методы .....	148
4.2.3. Термоэлектрические методы .....	152

## РАЗДЕЛ ВТОРОЙ

### Квантовые приборы СВЧ и оптического диапазона

Введение.....	157
<b>Глава 5. Взаимодействие электромагнитного излучения с веществом.....</b>	<b>159</b>
5.1. Принцип квантового усиления .....	159
5.2. Типы квантовых переходов. Коэффициенты Эйнштейна.....	162
5.2.1. Спонтанные переходы и их вероятность.....	163
5.2.2. Индуцированные переходы и их вероятность.....	165
5.2.3. Безызлучательные переходы и их вероятность.....	172
5.3. Уширение спектральных линий.....	174
5.3.1. Естественная ширина спектральной линии.....	174
5.3.2. Ширина полосы пропускания резонатора Фабри-Перо....	174
5.3.3. Причины уширения спектральной линии.....	175
5.4. Общие уравнения баланса .....	176
5.4.1. Взаимодействие монохроматического поля с двухуровневой квантовой системой.....	178
5.4.2. Мощность поглощения как функция частоты и амплитуды поля. Эффект насыщения .....	179
<b>Глава 6. Усиление и генерация электромагнитного излучения.....</b>	<b>181</b>
6.1. Принцип работы квантовых усилителей и генераторов и пороговые условия генерации.....	183
6.2. Методы получения инверсной населенности в квантовых генераторах и усилителях.....	183
6.2.1. Двухуровневая схема.....	184
6.2.2. Трехуровневая схема.....	185
6.2.3. Четырехуровневая схема.....	187
<b>Глава 7. Двухуровневые квантовые генераторы и усилители.....</b>	<b>190</b>
7.1. Квантовые стандарты частоты.....	190
7.1.1. Принцип действия молекулярного квантового генератора..	190
7.1.2. Условие самовозбуждения генератора.....	192
7.2. Квантовые парамагнитные усилители СВЧ (КПУ).....	196
7.2.1. Парамагнетизм в кристаллах.....	196
7.2.2. Энергетические уровни парамагнитных ионов.....	197
7.2.3. Электронный парамагнитный резонанс (ЭПР).....	198
7.2.4. Устройство квантового парамагнитного усилителя.....	200
7.2.5. Особенности квантового парамагнитного усилителя и области применения.....	204
	206
<b>Глава 8. Общие вопросы построения оптических квантовых генераторов и усилителей.....</b>	<b>207</b>
8.1. Диапазон оптических волн и его особенности.....	207
8.2. Классификация, используемых в оптико-электронных приборах	209

лазеров.....	209
8.3. Структурная схема оптико-электронного квантового прибора....	209
8.4. Оптические резонаторы.....	211
8.5. Свойства лазерного излучения. Характеристики излучения лазеров.....	225
<b>Глава 9. Оптические квантовые генераторы на твердом теле.....</b>	<b>233</b>
9.1. Устройство твердотельных генераторов.....	233
9.2. Генераторы на рубине.....	236
9.3. Неодимовые лазеры .....	237
9.4. Режимы работы твердотельных лазеров. Основные уравнения.	238
9.4.1. Режим непрерывной генерации твердотельного лазера....	240
9.4.2. Пички и режим свободной генерации.....	241
9.4.3. Лазер с модуляцией добротности.....	245
9.4.4. Синхронизация мод в твердотельном лазере и генерация ультракоротких импульсов (УКИ).....	247
9.4.5. Спектроскопические и угловые характеристики твердотельных лазеров.....	249
<b>Глава 10. Газовые лазеры.....</b>	<b>251</b>
10.1. Особенности газов как активного вещества для лазеров.....	251
10.2. Механизмы возбуждения.....	251
10.3. Гелий-неоновый лазер.....	253
10.4. Ионный аргоновый лазер.....	256
10.5. Молекулярный лазер на углекислом газе.....	258
<b>Глава 11. Жидкостные лазеры.....</b>	<b>263</b>
11.1. Общая характеристика и особенности жидкостных лазеров....	263
11.2. Жидкостные лазеры на органических красителях.....	264
11.2.1. Способы накачки лазеров на основе красителей.....	267
11.3. Спектральные характеристики и управление спектром лазера на основе красителей.....	272
<b>Глава 12. Полупроводниковые излучатели света.....</b>	<b>275</b>
12.1. Полупроводниковые лазеры.....	275
12.1.1. Рекомбинация электронов и дырок.....	278
12.1.2. Условия и методы получения инверсии населенности в полупроводниках .....	279
12.1.3. Инжекционные полупроводниковые лазеры.....	283
12.1.4. Оптические и электрические свойства полупроводниковых лазеров.....	291
12.2. Светоизлучающие диоды.....	297
12.2.1. Устройство СИД на гомо- и гетеропереходах.....	298
12.2.2. Типы конструкций СИД и их особенности.....	300
<b>Глава 13. Приборы управления лазерным излучением.....</b>	<b>303</b>

13.1. Методы модуляции. Классификация.....	304
13.2. Электрооптические модуляторы света.....	305
13.2.1. Электрооптический эффект Поккельса.....	307
13.2.2. Параметры и характеристики модуляторов света.....	314
13.3. Акустооптические методы управления параметрами лазерного излучения.....	314
13.3.1. Дифракция Рамана—Ната.....	316
13.3.2. Брегговская дифракция.....	317

## РАЗДЕЛ ТРЕТИЙ

### Прием и передача излучения в оптоэлек- тронике

<b>Глава 14. Приемники оптического излучения .....</b>	<b>322</b>
14.1. Физические принципы приема оптического излучения.....	322
14.1.1. Поглощение света в твердых телах.....	323
14.1.2. Явление фотоэффекта .....	325
14.1.3. Фотомагнитоэлектрический и фотогальванический эффект.....	327
14.2. Параметры фотоприемников .....	331
14.3. Типы фотоприемников.....	333
14.3.1. Полупроводниковые фотоприемники.....	340
14.3.2. Фотоэлектронные приборы.....	347
<b>Глава 15. Средства передачи оптического излучения .....</b>	<b>352</b>
15.1. Диэлектрические световоды.....	352
15.1.1. Распространение волн в световодах.....	352
15.1.2. Типы световодов, их характеристики.....	354
15.1.3. Элементы связи.....	357
15.2. Волоконно – оптические световоды.....	360
15.2.1. Свойства оптического волокна.....	362
15.2.2. Параметры характеристики оптических волокон .....	364
15.2.3. Ввод и вывод излучения.....	369

## РАЗДЕЛ ЧЕТВЕРТЫЙ

### Опто-электронные системы и их применения

<b>Глава 16. Системы обработки и хранения информации.....</b>	<b>371</b>
16.1. Структура системы оптической обработки сигналов.....	371
16.1.1. Принцип аналоговой оптической обработки информации с помощью когерентных световых полей.....	374
16.2. Способы выполнения математических операций оптическими системами.....	376
16.2.1. Вычисление корреляционных функций.....	379



16.2.2. Преобразование Фурье в оптической системе.....	379
16.2.3. Пространственная фильтрация.....	384
16.2.4. Согласованная фильтрация.....	387
16.2.5. Распознавание образов.....	392

## **Глава 17. Различные применения оптико-электронных**

<i>квантовых приборов и систем</i> .....	393
17.1. Применение твердотельных лазеров.....	394
17.2. Применение газовых лазеров.....	395
17.3. Области применения полупроводниковых источников излучения.....	398
17.4 Акусто – оптические системы обработки радиосигналов.....	399
17.4.1. Классический акустооптический спектроанализатор.....	400
17.4.2. Акустооптические процессоры .....	404
17. 4.3.Перспективы развития акустооптических систем.....	406
Список осокращений и обозначений параметров.....	408
Список рекомендованной литературы.....	411

## ПРЕДИСЛОВИЕ

Характерной чертой непрерывного прогресса в области электронных и квантовых приборов является создание более мощных генераторов, более чувствительных приемников и освоение более короткого диапазона волн. С освоением коротковолновой части СВЧ - диапазона значительно расширилась информационная емкость каналов связи, а непрерывный рост объема передаваемой информации стимулирует освоение более высоких частот. Основная цель данного методического пособия – дать представление о фундаментальных физических процессах, лежащих в основе современной электроники и оптоэлектроники и, тем самым, определить её возможности и пути её дальнейшего развития, рассмотреть принципы действия, особенности конструкций, технических характеристик СВЧ приборов и устройств оптоэлектроники, подготовить специалистов к технически грамотному их применению.

Настоящее пособие делится на четыре раздела. В первом разделе дается изложение общих вопросов, лежащих в основе электронных приборов СВЧ – диапазона (вакуумных и твердотельных). При изложении материала, обращено внимание на работу наиболее распространенных приборов, таких как: клистроны, лампы бегущей и обратной волн, магнетроны, генераторы и усилители на диодах Ганна и на лавинно-пролетных диодах. Приведены их основные характеристики и параметры. Проводится сравнительный анализ приборов и делается заключение об их применении в том или ином частотном диапазоне, или величине передаваемой мощности с точки зрения применения их в бытовой радиоэлектронной аппаратуре.

Во втором разделе кратко излагаются вопросы, лежащие в основе квантовых устройств. Описываются конкретные квантовые устройства.

Расширению областей применения приборов оптической электроники способствовало создание волоконно-оптических элементов, на основе которых совершенствуются системы ближней и дальней связи. Поэтому, в третьем разделе описываются приемники и средства передачи оптического излучения.

В четвертом разделе рассмотрены основы обработки и хранения информации и различные применения оптико-электронных и квантовых приборов и систем.

Пособие включает лишь самый необходимый материал для дальнейшего изучения конкретных приборов. При этом рассмотрение строится так, чтобы эти вопросы можно было легко понять. Некоторая часть теоретических вопросов опущена и для углубленного изучения курса приводится список рекомендуемой литературы.

Введение, раздел 1 написаны Шангиной Л.И., Соколовой Ж.М.; раздел 2 и 4– Куц Г.Г и Шангиной Л.И.; раздел 3- Шангиной Л.И.

## РАЗДЕЛ ПЕРВЫЙ

### ЭЛЕКТРОННЫЕ ПРИБОРЫ СВЧ

#### Введение

**Назначение и применение приборов СВЧ.** Электронными приборами СВЧ называются приборы, использующие принцип отдачи электронным потоком энергии высокочастотному (ВЧ) полю. Длительное время пролёта электронов является необходимым условием их работы.

Электроника сверхвысоких частот изучает вопросы взаимодействия электронных потоков с переменными электромагнитными полями в системах, где время пролета электронов через пространство взаимодействия оказывается большим или сравнимым с периодом колебаний. К приборам СВЧ относятся электровакуумные приборы: клистроны, магнетроны, лампы бегущей и обратной волны (ЛБВ и ЛОВ), гироприборы и твердотельные полупроводниковые приборы. Из всех СВЧ полупроводниковых приборов рассмотрены только генераторы на ЛПД и диодах Ганна.

“ СВЧ диапазон ” перекрывает область частот 300 МГц - 300 ГГц. Приборы СВЧ (ПСВЧ) широко применяются в аппаратуре для телевидения, для радиосвязи, радиолокации, радионавигации, кибернетики, для различных технологических процессов, в многоканальных радиопередатчиках и т.д.

#### Глава 1

#### Физические основы электровакуумных приборов СВЧ

**Классификация СВЧ приборов.** Электронные приборы классифицируются по способу управления электронным потоком с помощью электромагнитных полей, по характеру применения статических полей.

**Управление током.** Изменение в пространстве плотности электронного потока носит название управления током. Существует два принципа управления электронным потоком:

- Электростатическое управление;
- Динамическое управление.

**Использование постоянных электрических и магнитных полей.** Приборы СВЧ по характеру использования статических магнитных полей подразделяются на приборы типа О и типа М.

В приборах типа О статическое магнитное поле или вовсе не используется, или используется для фокусировки электронного потока. К этим приборам относятся клистроны, ЛБВ О и ЛОВ О.

В приборах типа М движение электронов происходит в скрещенных постоянных  $\vec{E}$  и  $\vec{H}$  полях. К этим приборам относятся: магнетроны, ЛБВ М и ЛОВ М, платинотроны и др.

## 1.1. Характеристики и особенности электровакуумных приборов

Одной из центральных задач современной электроники является наращивание выходной мощности СВЧ генераторов и усилителей. Необходимость повышения *мощности* диктуется в основном развитием новейших средств ракетной техники, противоракетной обороны и дальней связи. Максимальные значения выходной мощности при работе в импульсном режиме составляют 30 МВт, магнетронов – 10 МВт, усилительных ЛБВ – 8 МВт (в дециметровом диапазоне).

Второй важной характеристикой является *перестройка частоты*. Наиболее перспективными считаются генераторы с электронной перестройкой частоты (ЛОВ типа “М” и “О”), усилительные ЛБВ.

Для увеличения *дальности действия* радиотехнических средств большое значение имеет уменьшение уровня собственных шумов приёмных усилительных приборов. В данное время малошумящими приборами в диапазоне СВЧ являются детекторные и смесительные диоды и ЛБВ.

Важной проблемой для мощных СВЧ генераторов и усилителей является проблема *повышения КПД*. Наиболее эффективными приборами следует считать клистроны и приборы магнетронного типа, КПД которых достигает в сантиметровом диапазоне длин волн 60%.

*Повышение* требований к *надёжности* действующего радиоэлектронного оборудования ставит перед СВЧ электронными приборами задачу повышения срока службы приборов.

**Особенности электровакуумных приборов СВЧ.** Электровакуумные приборы, предназначенные для работы в диапазоне СВЧ ( $10^8 \div 10^{12}$ ) Гц, занимают особое положение среди современных электронных и газоразрядных приборов. Вместе с тем существуют обстоятельства, принципиально отличающие электронику СВЧ от “обычной” (низкочастотной) электроники. Рассмотрим их подробнее.

1. Первым важным фактором, характеризующим электронику СВЧ, является *соизмеримость периода колебаний со временем пролёта электронов* между электродами “лампы”.

2. Второе основное затруднение обусловлено: *соизмеримостью рабочей длины волны и линейных размеров цепи*, в которую включается электровакуумный прибор. Длина волны становится соизмеримой также с геометрическими размерами электродов, вводов и другими конструктивными элементами лампы.

3. В - третьих, цепи с сосредоточенными параметрами на СВЧ имеют столь серьёзные недостатки, что применение их в сантиметровом и более коротковолновом диапазоне практически исключено. В диапазоне СВЧ *обязательными элементами цепей являются волноводы и полые резонаторы*, т.е. цепи с распределёнными или полураспределёнными параметрами.

## 1.2. Основные уравнения и параметры электронных приборов СВЧ

Исходя из особенностей электро- вакуумных приборов, определим их исходные уравнения и основные параметры. Материал этой главы излагается на основе учебников, рекомендуемых вузами радиотехнических специальностей [1-14].

Наиболее общий подход к явлениям электродинамики обеспечивается применением знакомых из теории электромагнитного поля уравнений Максвелла [1,9]:

$$\operatorname{rot} \bar{H} = \bar{J}_{\text{полн}} = \rho \mathbf{v} + \frac{\partial \bar{D}}{\partial t}, \quad \operatorname{rot} \bar{E} = -\frac{\partial \bar{B}}{\partial t}, \quad \operatorname{div} \bar{D} = \rho, \quad \operatorname{div} \bar{B} = 0. \quad (1.1)$$

В этих уравнениях  $\mathbf{D}$  и  $\mathbf{B}$  – векторы электрического смещения и магнитной индукции, которые через диэлектрические и магнитные проницаемости связаны с векторами напряженностей электрического и магнитного полей:

$$\bar{D} = \epsilon \epsilon_0 \bar{E}, \quad \bar{B} = \mu \mu_0 \bar{H},$$

$$\text{где } \epsilon_0 = [1/(36\pi)] \cdot 10^{-9} \text{ Ф/м}, \quad \mu_0 = 4 \cdot \pi \cdot 10^{-7} \text{ Гн/м}. \quad (1.1,а)$$

Плотность полного тока определяется плотностями конвекционного тока и тока смещения:

$$\bar{J}_{\text{полн}} = \bar{J}_{\text{конв}} + \bar{J}_{\text{см}} = \rho \bar{V} + \frac{\partial \bar{D}}{\partial t}, \quad (1.2)$$

где  $\rho$  - плотность объемного заряда,  $\bar{V}$  – скорость движения электронного потока.

Из 1-го и 3-го уравнений системы (1.1) следует уравнение непрерывности:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \bar{J}_{\text{конв}} = 0. \quad (1.3)$$

Уравнения Максвелла можно дополнить уравнением движения (уравнение Лоренца) - зависимость скорости электронов  $\mathbf{v}$  от величины электрического и магнитного полей:

$$\frac{d(m\mathbf{v})}{dt} = \bar{F} = q \cdot \{\bar{E} + [\bar{V}, \bar{B}]\} = \bar{F}_E + \bar{F}_H, \quad (1.4)$$

где  $m = m_0 / \sqrt{1 - \frac{V^2}{c^2}}$  – масса электрона,  $m_0$  – масса покоящегося электрона ( $m_0 = 9,108 \cdot 10^{-31}$  кг),  $\bar{F}_E$  – сила, обусловленная действием электрического поля;  $\bar{F}_H$  – сила, обусловленная действием магнитного поля. Если ( $|v| \ll c$ ), то релятивистской поправкой можно пренебречь. В дальнейшем будем считать, что  $m = m_0$ .

При замене в (1.4)  $q = -e$ , где  $e$  свободные электроны ( $e \cong 1,6 \cdot 10^{-19}$  Кл), получим уравнение движения одиночного электрона:

$$\frac{d(m\mathbf{v})}{dt} = \bar{F} = -e \cdot \{\bar{E} + [\bar{V}, \bar{B}]\}. \quad (1.5)$$

Можно легко показать, что в большинстве практически важных случаев достаточно учитывать в уравнении (1.4), кроме постоянных составляющих  $\bar{E}$  и

$\bar{V}$ , лишь переменную составляющую электрического поля. Это возможно вследствие того, что

$$\frac{|\bar{F}_H|}{|\bar{F}_E|} \leq \frac{|\epsilon\mu_0[\bar{V}; \bar{H}]|}{|c\bar{E}|} \leq \frac{\bar{V}}{c}, \text{ при } V \ll c, \text{ где } c - \text{ скорость света } (3 \cdot 10^8 \text{ м/с}) \text{ в вакууме. Без}$$

учета переменных полей при  $U_0 = \text{const}$  и  $V \ll c$  уравнение движения в этом случае приводится к виду:

$$m \frac{d^2x}{dt^2} = e \left( -\frac{U_0}{d} \right). \quad (1.6)$$

Далее, пренебрегается квантовым характером взаимодействия между зарядами и полем, так как учёт квантовых процессов необходим лишь на значительно более высоких частотах.

Нужно помнить, что в электронике СВЧ главным является взаимодействие электронов с полем СВЧ, поэтому важно знать основные параметры, характеризующие электронный поток и СВЧ поле.

**Скорость электрона**, приобретённую в потенциальном электрическом поле, находят исходя из закона сохранения энергии  $W_k = W_{\text{п}}$ :

$$eU = \frac{m V^2}{2}, \text{ откуда } V = \sqrt{\frac{2eU_0}{m}}, \quad (1.7)$$

где  $U_0$  - разность потенциалов между рассматриваемой точкой пространства и точкой, где скорость электрона равна нулю. Подставляя в (1.7) величины  $e$  и  $m$ , получаем расчетное выражение:

$$V \approx 5,95 \cdot 10^5 \sqrt{U_0}. \quad (1.8)$$

Заметим, что в (1.7) не учитывают возможного изменения  $U$  за время движения частицы и строго применяют только для статического поля. Считаем, что время пролета частиц  $\tau$  много меньше периода сигнала ( $T$ ).

Система рассмотренных основных уравнений СВЧ требует для своего решения граничных условий. Такими являются, кроме обычных (известных из теории поля) – начальные скорости частиц на фиксированных поверхностях и в фиксированные моменты времени.

**Время пролета электрона  $\tau$ .** Найдем  $\tau$  для простейшего плоского вакуумного диода, электроды которого образованы двумя бесконечными параллельными плоскостями 1 и 2, находящимися на расстоянии  $d$  друг от друга (рис. 1.1).

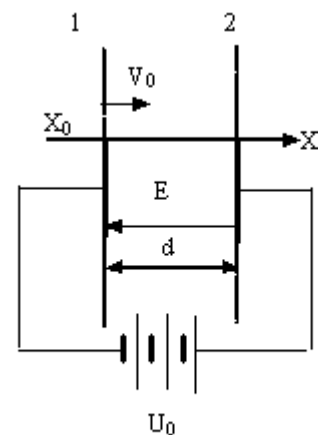


Рис. 1.1. Схема плоского диода

При использовании уравнения движения (1.6) и граничных условий ( $v = v_0$ ,  $t = t_0$ ,  $x = x_0$ ), конечное выражение для скорости электронного потока выглядит следующим образом:

$$v = v_0 + \frac{eU_0}{md} (t - t_0), \quad (1.9)$$

где время пролета равно  $\tau = t - t_0$ .

Траектория движения электрона по оси  $x$  тогда запишется:

$$x = x_0 + v_0(t - t_0) + \frac{eU_0}{md} \frac{(t - t_0)^2}{2}. \quad (1.10)$$

Подставляя в уравнение (1.10) координату второго электрода  $x = x_0 + d$  и считая, что  $t - t_0 = \tau$ , то при  $U \neq 0$ ,  $v_0 = 0$ , получаем время пролета электронами межэлектродного промежутка  $d$  в режиме насыщения:

$$\tau = d \sqrt{\frac{2m}{eU_0}}. \quad (1.11)$$

**Уравнение движения с учетом переменного поля.** Если между электродами плоского зазора приложено переменное напряжение

$U = U_m \sin \omega t$ , то двойное интегрирование уравнения движения (1.6) дает:

$$x = x_0 + \frac{eU_m}{\omega^2 md} [(\omega t - \omega t_0) \cdot \cos \omega t_0 - \sin \omega t + \sin \omega t_0]. \quad (1.12)$$

Отсюда следует, что время пролета теперь имеет различную величину для электронов, вошедших в зазор в разные моменты времени  $t_0$ . В этом случае невозможно ввести "истинное" время пролета электрона.

Если на электроды приложить одновременно постоянное и переменное напряжения, т.е.  $U = U_0 + U_m \sin \omega t$ , то при соизмеримых величинах  $U_0$  и  $U_m$   $\tau$  сильно различается в зависимости от начального момента времени  $t_0$ . Однако, при  $U_m/U_0 \ll 1$  можно пренебречь малыми изменениями времени пролета, обусловленными переменной составляющей напряжения, и вычислять невозмущенное  $\tau$ .

**Угол пролета.** Абсолютная величина  $\tau$  сама по себе недостаточно полно характеризует влияние инерции электронов на работу прибора. Поведение электромагнитного поля в значительной степени зависит от соизмеримости  $\tau$  с периодом колебаний  $T$ . Поэтому вводится понятие угла пролета электронов  $\theta$ :

$$\theta = 2\pi \frac{\tau}{T} \text{ (рад)} \quad \text{или} \quad \theta = 360^\circ \frac{\tau}{T} \text{ (град)}.$$

Используя соотношение  $\omega = 2\pi/T$ , определение угла пролета можно привести к виду:

$$\theta = \omega \tau. \quad (1.13)$$

С физической точки зрения угол пролета показывает изменение фазы напряжения, приложенного к электродам, за время движения электрона между электродами. При  $U_m \ll U_0$ , угол пролета с учетом формулы (1.11) равен:

$$\theta = \omega d \sqrt{\frac{2m}{eU_0}}.$$

Если же на электроды наложено только малое переменное поле и электроны поступают в зазор с большой скоростью, то угол пролета описывается уравнением:  $\theta = \omega d / V_0$ .

Последнее уравнение можно применить при условии  $U_m \ll \frac{mV_0^2}{2e}$ .

**Пространственно-временные диаграммы (ПВД).** Движение электронов между электродами можно наглядно иллюстрировать графиками зависимости координаты

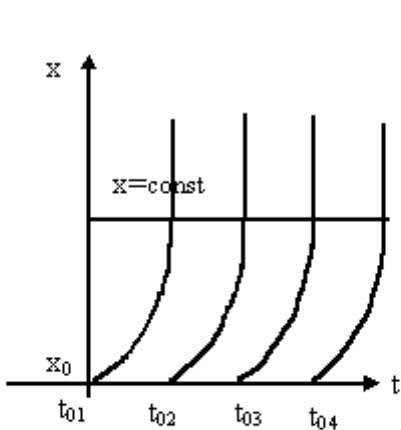


Рис. 1.2. Движение электронов в плоском диоде

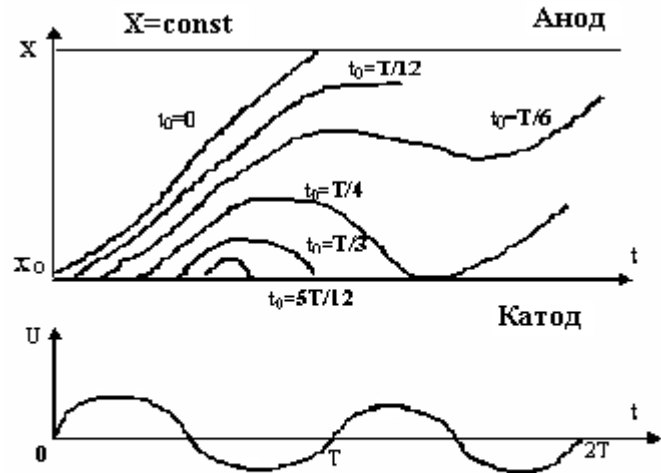


Рис. 1.3. Движение электронов под действием переменного поля

ординаты электрона от времени - так называемыми пространственно-временными диаграммами. На рис. 1.2 приведено семейство парабол, отражающих движение электронов в плоском диоде в отсутствие пространственного заряда при нулевых начальных скоростях  $V_0$ . На электроды диода наложено постоянное напряжение. Электроны, вышедшие из катода через любые равные интервалы времени, достигают анода (любой плоскости  $x=\text{const}$ ) через равные интервалы.

На рис. 1.3 приведен вид диаграммы, построенной по уравнению (1.12). Она описывает движение электронов под действием переменного напряжения при равенстве нулю постоянной составляющей. ( $v_0=0$ ; пространственный заряд не учтен). Как видно не все электроны, имитированные катодом, достигают анода.

При большом междуэлектродном расстоянии  $d$  до половины вышедших электронов возвращаются на катод, рассеивая на нем свою кинетическую энергию. Некоторые электроны достигают анода лишь после нескольких “качаний” в междуэлектродном пространстве. ПВД позволяет судить по густоте кривых, пересекающих линию  $x=\text{const}$  о мгновенной плотности конвекционного тока на фиксированном расстоянии от катода, например, на аноде. Касательная к кривой на ПВД определяет мгновенную скорость и, следовательно, кинетическую энергию электрона в каждой точке пространства в любой момент времени. Одновременно



могут быть сделаны выводы о мгновенной и средней мощности, рассеиваемой электронами на аноде и катоде.

**Уравнение наведенного тока. Наведение тока при движении свободных зарядов в плоском зазоре.** Вопрос о связи между движением электронов в лампе и током, протекающим во внешней цепи лампы, лежит в основе всей современной электроники. Рассмотрим для простоты плоский вакуумный зазор (рис. 1.4). Во внешней цепи зазора включен источник постоянного или переменного напряжения  $U$ , имеющий нулевое внутреннее сопротивление. Если свободные заряды отсутствуют, то на электродах зазора имеются поверхностные заряды

$$+Q_0 \text{ и } -Q_0,$$

где  $Q_0 = \frac{\epsilon_0 S U}{d} = \epsilon_0 E_0 S$ ,  $E_0 = \frac{U}{d}$ ;  $S$  - поверхность электродов,  $d$  - расстояние между пластинами. Внесение в зазор заряда  $+q$  наводит за счет электростатической индукции поверхностные заряды  $-q_1$  и  $-q_2$ , удовлетворяющие уравнению сохранения заряда:  $q - q_1 - q_2 = 0$ . Эти заряды вызывают возникновение поля, накладывающегося на поле  $E_0$ . Напряженность поля слева от слоя и справа (рис. 1.4, в) будет равна:

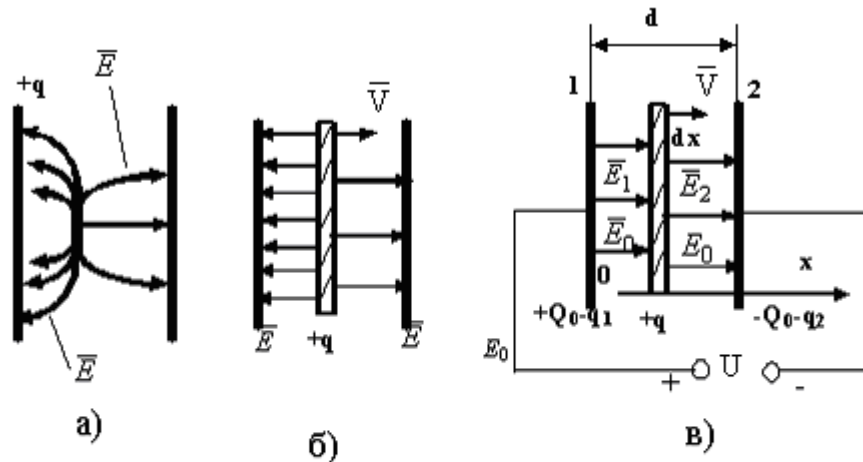


Рис. 1.4. Движение электронов в плоском зазоре.  
а - для точечного заряда, б - для заряда в виде слоя  
в - напряженность поля слева и справа от слоя

$$E_1 = + \frac{+Q_0 - q_1}{\epsilon_0 S}, \quad E_2 = - \frac{-Q_0 - q_2}{\epsilon_0 S}. \quad (1.14)$$

С учетом того, что к зазору приложено напряжение  $U$ , можно записать:

$$E_1(x) + E_2(d - x) - U = E_0 \cdot d,$$

где  $x$  - текущая координата заряженного слоя. Подставляем в последнее выражение значения  $E_1$  и  $E_2$  из (1.14). После несложных арифметических выкладок и с учетом общего количества зарядов ( $q_1 + q_2 = q$ ), значения  $q_1$  и  $q_2$  определяются следующими выражениями:

$$q_2 = q \frac{x}{d}; \quad q_1 = q \left( 1 - \frac{x}{d} \right).$$

Ток, регистрируемый во внешней цепи, будет равен:

$$i_{\text{полн}} = + \frac{dQ_1}{dt} = - \frac{dQ_2}{dt} = \frac{dQ_0}{dt} + \frac{q}{d} \frac{dx}{dt} = \frac{dQ_0}{dt} + \frac{qV}{d}. \quad (1.15)$$

Итак, ток во внешней цепи складывается из тока смещения  $\frac{dQ_0}{dt}$  и тока, создаваемого движущимся зарядом. Физический смысл второй компоненты полного тока легко понять на основе рассмотрения электрических силовых линий, выходящих из точечного заряда и оканчивающихся на одном из электродов. По мере движения заряда количество линий меняется и, в результате, электрические заряды переходят с одного электрода на другой через внешнюю цепь. Ток, создаваемый во внешней цепи движущимся электрическим зарядом, называют наведенным током. Для плоского зазора выражение мгновенного наведенного тока:

$$i_{\text{навед}} = \frac{qV}{d}. \quad (1.16)$$

**Форма импульсов наведенного тока.** Найдем форму импульсов тока, протекающего во внешней цепи плоского диода при движении в нем точечных зарядов электронов. Для этого подставим в (1.16) скорость электрона в плоском зазоре из уравнения (1.9), тогда заменяя  $q = -e$ , получим следующее выражение:

$$i_{\text{навед}} = \frac{qV}{d} = - \frac{e}{d} \left[ V_0 + \frac{eU_0}{md} (t - t_0) \right] = \left[ \frac{eV_0}{d} + \frac{e^2 U_0}{md^2} (t - t_0) \right] \quad (1.16a)$$

Форма импульсов  $I_{\text{навед}}$  показана на рис. 1.5, на котором (а) соответствует импульсу наведенного тока, например, в пространстве сетка-анод ( $V_0 \neq 0$ ). При  $t = t_0$  скачок тока на  $\frac{eV_0}{d}$ , затем линейное изменение, и в момент попадания электрона на второй электрод ток падает скачком до нуля. Если же первый электрод является катодом  $V_0 \approx 0$  (рис. 1.5,б), то импульсы наведенного тока имеют форму треугольника. Площадь импульса равна заряду электрона  $e$ .

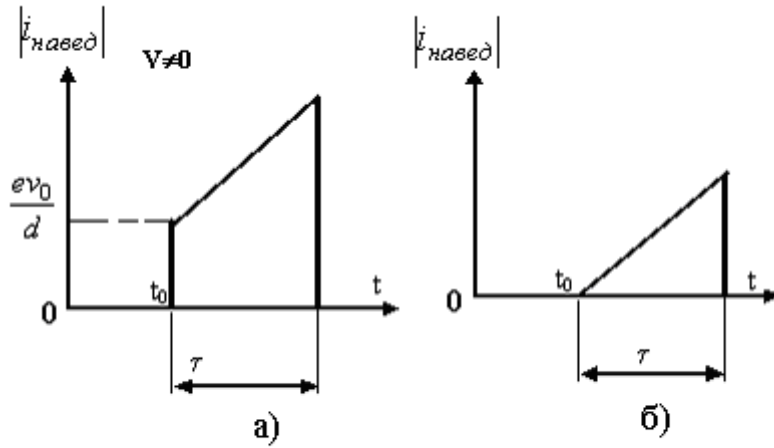


Рис. 1.5. Форма импульса наведенного тока

Выводы: Длительность импульса тока от отдельных электронов равна времени пролета  $\tau$  в данном зазоре. Ток прекращается, а не возникает в момент попадания электрона на электрод. Полный ток, протекающий во внешней цепи любого электронного прибора, имеет две основные составляющие: наведенный ток и емкостной ток. Так как полный ток во всех участках последовательной цепи одинаков, то внутри зазора, где заряды отсутствуют, полный ток представлен током смещения, а там где есть заряды – током переноса. Наведенный ток создается электронами и направлен навстречу движению электронов в зазоре.

**Наведение тока в плоском зазоре при прохождении промодулированного по плотности электронного потока.** Будем рассматривать движение в двух сеточном зазоре непрерывного потока с переменной объемной плотностью  $\rho$  при постоянной скорости электронов  $v_0$  (рис. 1.6).

Пусть плотность пучка, поступающего через входную сетку, меняется по гармоническому закону:

$$\rho = \rho_0 + \rho_1 \sin \omega t,$$

где  $\rho_0$  - постоянная составляющая,  $\rho_1$  - амплитуда переменной составляющей плотности заряда.

В фиксированной плоскости конвекционный ток определяется выражением:  $i_{\text{конв}} = \rho S V_0 = I_0 + I_m \sin \omega t$ ,

где  $I_0 = \rho_0 S V_0$ ,  $I_m = \rho_1 S V_0$ ,  $S$  - площадь сечения тока.

Наведенный ток  $di_{\text{навед}}$ , созданный зарядом  $dq$  равен:

$$di_{\text{навед}} = dq \frac{V_0}{d} = i_{\text{конв}} \frac{dx}{d}, \quad (1.17)$$

где  $i_{\text{навед}}$  - наведенный ток, протекающий во внешней цепи зазора, когда сетки соединены накоротко ( $E=0$ ).

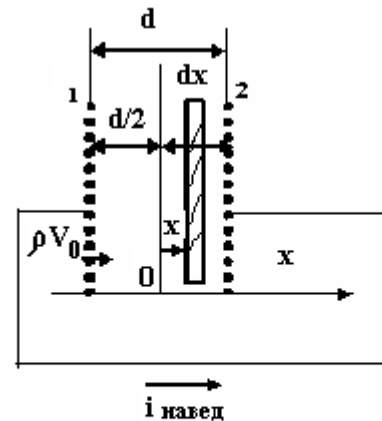


Рис. 1.6. Движение в сеточном зазоре непрерывного потока электронов с переменной объемной плотностью

Предполагаем, что  $V_0 = \text{const}$  внутри зазора (т.к.  $E=0$ ) и что в момент  $t_0$  слой проходит через центр зазора. Суммарный наведенный ток, получаем интегрированием по зазору от  $-d/2$  до  $+d/2$ :

$$i_{\text{навед}} = \int_{-d/2}^{d/2} \left\{ I_0 + I_m \sin \left( \omega t + \frac{\omega x}{2V_0} \right) \right\} \frac{dx}{d} = I_0 + I_m \sin \omega t_0 \frac{\sin \frac{\omega t}{2V_0}}{\frac{\omega t}{2V_0}}. \quad (1.17a)$$

Учитывая, что угол пролета электронов равен  $\theta = \omega d / V_0$ , введем обозначение:

$$\gamma = \sin \frac{\theta}{2} / \frac{\theta}{2}, \quad (1.18)$$

где  $\gamma$  - коэффициент взаимодействия, определяющий эффективность поведения тока в зазоре.

Тогда уравнение тока, наведенного модулированным электронным потоком, приобретает вид:

$$i_{\text{навед}} = I_0 + \gamma I_m \sin \omega t_0. \quad (1.19)$$

Множитель  $\gamma$  обычно называют коэффициентом взаимодействия электронного потока с электрическим полем в зазоре.

Зависимость  $\gamma$  от  $\theta$  имеет вид (рис. 1.7). Из анализа приведенных уравнений заключаем, что если угол пролета мал, то  $\gamma \rightarrow 1$  и  $i_{\text{навед}} \approx i_{\text{конв}}$ . С увеличением угла пролета  $\theta$ , происходит уменьшение амплитуды наведенного тока. При  $\theta = 2\pi$   $I_{\text{навед}} \approx 0$ , хотя конвекционный ток остается неизменным. Практически коэффициент эффективности равен  $\gamma = 0,6-0,9$ .

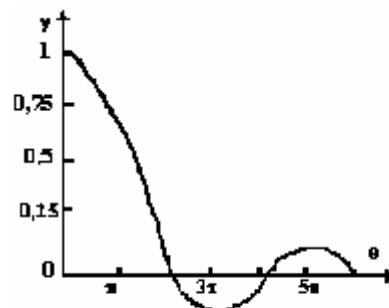


Рис. 1.7. Зависимость коэффициента эффективности от угла пролета

### 1.3. Электродинамические системы приборов СВЧ

В основе работы электровакуумных приборов с электростатическим управление лежит изменение плотности электронного потока с помощью сетки, находящейся под отрицательным постоянным потенциалом по отношению к катоду, при поступлении переменного сигнала, создающей модуляцию конвекционного тока.

Наиболее ценной и характерной особенностью электростатического управления при низких частотах является практически нулевое потребление мощности от источника напряжения. Однако, на СВЧ, как показывает опыт, управление с

помощью отрицательной сетки начинает требовать затрат мощности, быстро возрастающих с ростом частоты. Поэтому на высоких частотах применяется динамическое управление электронным потоком.

В приборах СВЧ электромагнитные поля резонаторов или замедляющих электродинамических систем задействованы в принципе работы. Переменные электрические поля, вводимые или наводимые в электродинамических системах приборов СВЧ, взаимодействуют с пропускаемыми электронными потоками. В результате такого взаимодействия может происходить усиление сигнала или генерация мощности на выбранной частоте, которые определяются параметрами и характеристиками электродинамических систем.

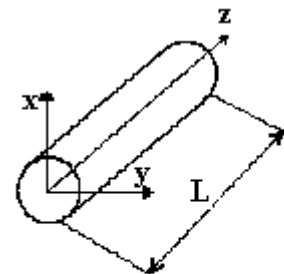


Рис. 1.8. Объемный резонатор

### 1.3.1. Резонаторы

Объемные резонаторы используются в приборах СВЧ резонансного типа. Резонаторами называют полые металлические тела, внутри которых концентрируется переменное электромагнитное поле. Область существования этого поля  $V$  можно отделить от остального пространства условной границей  $S$ , излучение энергии через которую отсутствует или незначительно. Количество энергии, запасаемой в электромагнитном поле объемного резонатора  $W_{\text{зап}}$ , зависит от резонансной частоты  $\omega_0$ . По мере приближения к частоте  $\omega_0$  количество энергии резко возрастает.

Простейшим объемным резонатором может служить отрезок волновода (рис. 1.8), закрытый с обеих сторон металлическими крышками.

Резонанс в системе наступает тогда, когда выполняется фазовое условие резонанса, имеющее следующий вид:

$$\beta L + \varphi_1 + \beta L + \varphi_2 = 2\pi n, \quad (1.20)$$

где  $\beta = \frac{2\pi}{\lambda_0}$  — фазовая постоянная волны в линии, имеющей длину волны  $\lambda_0$ ,

называемую в резонаторах резонансной;

$n=1, 2, \dots$  — целое число;

$\varphi_1, \varphi_2$  — набег фазы волны при ее отражении от торцевых крышек. Фазовый набег волны при отражении от металлической крышки равен  $\pi$ , в отсутствие крышки — нулю.

Резонансная длина волны (частота) объемного резонатора определяется типом волны, числом полуволн, укладывающихся на длине  $L$ , и геометрией резонатора.

В приборах СВЧ применяются [3,4] следующие типы резонаторов: магнетронные (рис. 1.9), тороидальные (рис. 1.10,а), коаксиальные с сосредоточенной емкостью (рис. 1.10,б), призматические (на базе прямоугольного волновода) с укорачивающей емкостью (рис. 1.10,в), многоззорные (рис. 1.10,г) и различные их модификации.

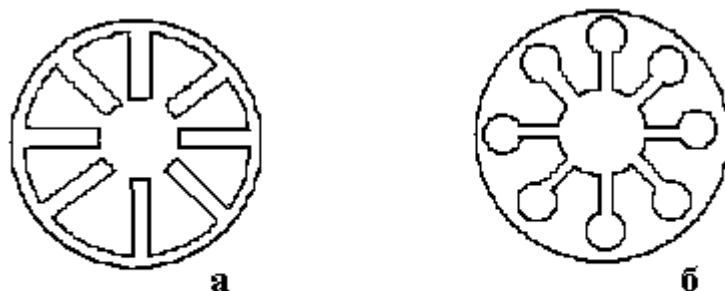


Рис. 1.9. Магнетронные резонаторы: а – секторообразные, б – щель – отверстие

Основными (собственными) параметрами резонаторов являются те, которые легко измеряются. В диапазоне СВЧ такими параметрами приняты:

- **резонансная длина волны  $\lambda_0$**  (резонансная частота  $f_0$  или круговая частота  $\omega_0$ ), которая определяется типом волны и геометрией резонатора;
- **активная проводимость  $G_0$** , являющаяся мерой активных потерь в резонаторе;
- **собственная или ненагруженная добротность  $Q_0$** , которая определяется отношением активной и реактивной проводимостей резонатора вблизи резонансной частоты.

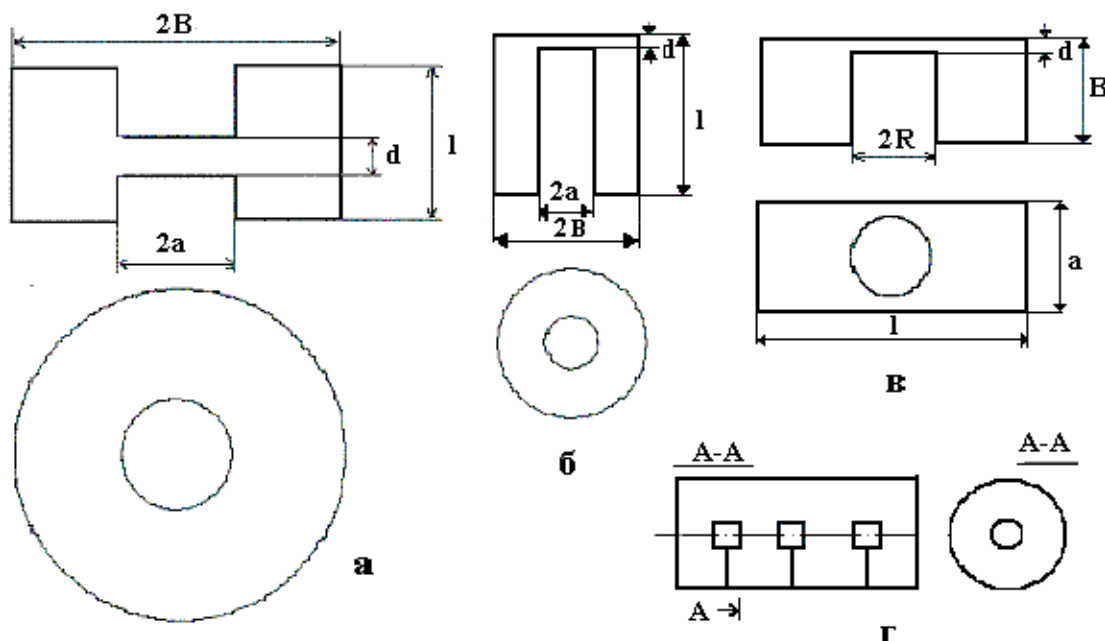


Рис. 1.10. Типы резонаторов СВЧ приборов

Параметры  $\lambda_0$ ,  $G_0$ ,  $Q_0$  могут полностью заменить собой основные параметры контуров низкочастотной радиотехники: индуктивность- $L_0$ , емкость-  $C_0$ , сопротивление-  $R$ . При определении собственных параметров резонатор предполагается изолированным от внешних нагрузок. Влияние внешних нагрузок учитывается дополнительно в понятиях внешней  $Q_{вн}$  и нагруженной  $Q_{н}$  добротностях:

$$Q_0 = \omega_0 \frac{W_{зап}}{P_{пот.р}} , \quad Q_{вн} = \omega_0 \frac{W_{зап}}{P_{пот.нагр}} , \quad (1.21,а)$$

$$Q_{н} = \omega_0 \frac{W_{зап}}{P_{пот.рез} + P_{пот.нагр}} . \quad (1.21,б)$$

Добротности связаны уравнением

$$\frac{1}{Q_{н}} = \frac{1}{Q_0} + \frac{1}{Q_{вн}} , \quad (1.22)$$

где  $W_{зап}$  – запасенная энергия в резонаторе;

$P_{пот.р}$  – мощность потерь в резонаторе (обычно в стенках);

$P_{пот.нагр.}$  – мощность потерь в нагрузке;

$P_{пот.}$  – мощность полных потерь в резонаторе и в нагрузке.

Заметим, что с помощью понятий добротностей  $Q_0$ ,  $Q_{вн}$ ,  $Q_{н}$  можно определить коэффициент полезного действия резонатора  $\eta_{рез}$  в виде

$$\eta_{рез} = \frac{Q_{н}}{Q_{вн}} = 1 - \frac{Q_{н}}{Q_0} . \quad (1.23)$$

Величина  $\eta_{рез}$  для большинства случаев резонаторов составляет 80-90%.

Если известны выражения электрических и магнитных полей в резонаторе, то собственную добротность можно рассчитать по формуле

$$Q_0 = \frac{2}{\delta} \frac{\int_V |H|^2 dV}{\oint_S |H_t|^2 dS} ,$$

где  $\delta$  – глубина проникновения поля в стенки резонатора;

$|H|$  - величина магнитного поля в объеме резонатора;

$|H_t|$  - величина тангенциального магнитного поля на стенках резонатора;

$V$  – объем резонатора;

$S$  – площадь поверхности внутренних стенок.

Для определения величины  $Q_0$  с помощью метода измерительной линии рекомендуется [4] использовать соотношение

$$Q_0 = \frac{\omega_0}{2G_0} \left( \frac{dB}{d\omega} \right)_{\omega \rightarrow \omega_0} . \quad (1.24)$$

Аналогично для измерения  $Q_{вн}$  и  $Q_{н}$  рекомендуется использовать выражения

$$Q_{\text{BH}} = \frac{\omega_0}{2G_{\text{H}}} \left( \frac{dB}{d\omega} \right)_{\omega \rightarrow \omega_0}, \quad (1.25, \text{a})$$

$$Q_{\text{H}} = \frac{\omega_0}{2(G_0 + G_{\text{H}})} \left( \frac{dB}{d\omega} \right)_{\omega \rightarrow \omega_0}, \quad (1.25, \text{б})$$

где  $G_{\text{H}}$  – активная проводимость нагрузки резонатора;

$\left( \frac{dB}{d\omega} \right)_{\omega \rightarrow \omega_0}$  – скорость изменения реактивной проводимости резонатора

вблизи резонансной частоты.

Резонатор может быть представлен параллельной или последовательной эквивалентной схемой (рис. 1.11). В диапазоне СВЧ при анализе работы генераторов и усилителей чаще используется параллельная схема (рис. 1.11,в), причем

$B_0 = \omega C_0 - \frac{1}{\omega L_0}$ . Активная проводимость резонатора определяется выражением

$$G_0 = \frac{2P_{\text{пот.р.}}}{U_{\text{m}}^2}, \quad \text{где } U_{\text{m}} = -\int_a^b \bar{E} d\bar{l} \text{ – амплитуда СВЧ напряжения в резонаторе,}$$

измеряемая между заданными точками (обычно это зажимы ввода – вывода энергии); напряженность электрического поля  $E$  берется между теми же точками резонатора.

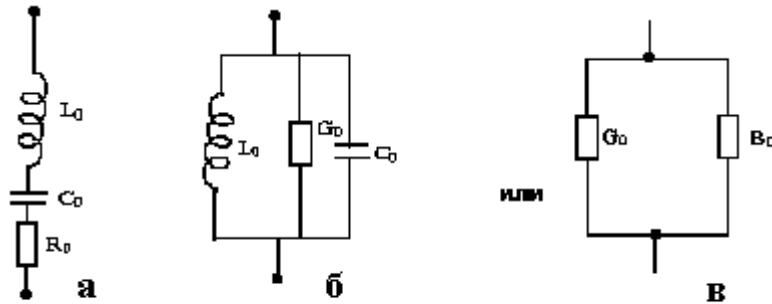


Рис. 1.11. Эквивалентные схемы полого резонатора: а-последовательная; б, в – параллельные.

Заменяя  $P_{\text{пот.р.}}$  и  $U_{\text{m}}$  выражениями через характеристики поля, получим

$$G_0 = \sqrt{\frac{\omega \mu_{\text{ст}} \oint |H_t|^2 dS}{2\sigma_{\text{ст}} \left\{ \int_a^b \bar{E} d\bar{l} \right\}^2}}, \quad (1.26)$$

здесь  $\mu_{\text{ст}}$  – абсолютная магнитная проницаемость стенок резонатора;

$\sigma_{\text{ст}}$  – удельная проводимость материала стенок резонатора.

В случае электронных СВЧ приборов  $G_0$  определяется по отношению к тем точкам, в которых резонатор пронизывается электронным потоком, и составляет  $10^{-4} - 10^{-5} \text{ 1/Ом}$ .



Ниже приведены некоторые формулы для определения резонансной частоты или длины волны резонаторов [1,4], показанных на рис. 1.10.

Для резонатора рис. 1.10,а

$$\lambda_0 = \pi\alpha \sqrt{\frac{2h}{d} \left( 1 + \frac{4d}{\pi\alpha} \ln \frac{h}{d} \right) \ln \frac{b}{a}} .$$

Для резонатора рис. 1.10,б частота определяется путем решения трансцендентного уравнения:

$$\omega_0 \frac{\epsilon\pi\alpha^2}{d} - \frac{1}{Z_C} \operatorname{ctg} \frac{\omega_0 L}{c} = 0 ,$$

где  $Z_C = 60 \ln \frac{b}{a}$  - волновое сопротивление коаксиальной линии.

В случае резонатора (рис. 1.10,в) резонансная длина волны определяется также трансцендентным соотношением:

$$\frac{1}{Z_C} \operatorname{ctg} \left\{ \frac{2\pi(1-R)}{\lambda_B} \right\} = \frac{2\pi^2 R^2 \epsilon_0 c}{d\lambda_0} ,$$

где  $c$  – скорость света;

$Z_C = \frac{377\epsilon}{\sqrt{a^2 - \left(\frac{\lambda_0}{2}\right)^2}}$  – волновое сопротивление волновода на волне типа  $H_{10}$ ;

$$\lambda_\epsilon = \frac{\lambda_0}{\sqrt{1 - \left(\frac{\lambda_0}{2a}\right)^2}} \quad - \text{длина волны в волноводе};$$

$a$ - размер широкой стенки волновода.

В случае резонатора магнетронного типа (рис. 1.9,б) собственная частота отдельного резонатора (ячейки) определяется [3] формулой  $\omega_0 = \frac{1}{R} \sqrt{\frac{d}{\epsilon\mu\pi\alpha}}$ , так как емкость и индуктивность резонатора- ячейки равны соответственно

$$C_0 = \frac{\epsilon\alpha h}{d}, \quad L_0 = \frac{\mu\pi R^2}{h}, \quad \text{размеры видны на рис. 1.12.}$$

Высокочастотное электрическое поле в квазистатических (тороидальных) резонаторах сосредоточено в узких зазорах (рис. 1.13), через которые пропускается электронный поток (вдоль электрического поля) с целью взаимодействия и изменения скорости.

Перестройка частоты объемных резонаторов (рис.1.10) осуществляется путем изменения одного из размеров: высоты резонатора  $L$  (поршнем) или высоты зазора  $d$  (перемещением внутреннего выступа), или посредством изменения диаметра (путем симметричного погружения нескольких металлических стержней в область с магнитным полем), если возможно.

Связь резонатора с нагрузкой осуществляется различными способами в зависимости от типа линии. Если вывод энергии коаксиального типа, то используется петлевой диполь, если выход в торец волновода (рис. 1.14), то в резонаторе выполняется отверстие связи.

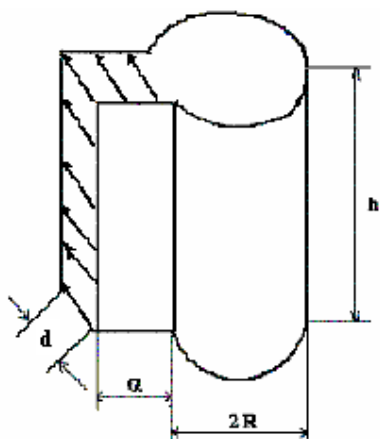


Рис. 1.12. Резонатор типа щель-отверстие

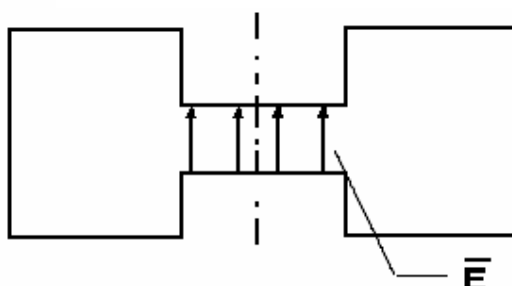


Рис. 1.13. Торондальный резонатор

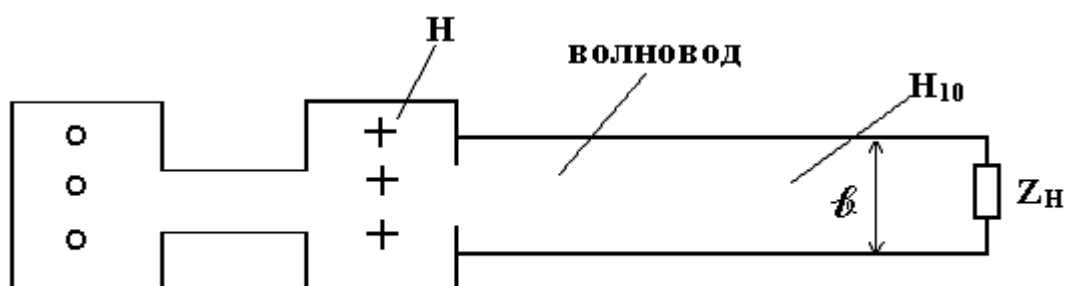


Рис. 1.14

### 1.3.2. Замедляющие системы, их параметры и характеристики

Некоторые типы ЗС [3,5] приведены на рис. 1.15.

Чаще всего одна из металлических плоскостей ( $x$ -const или  $y$ -const) замедляющей структуры имеет гладкую поверхность, а другая – состоит из ряда периодически повторяющихся неоднородностей. Иногда обе стенки имеют периодическую структуру (рис. 1.15.г). Расстояние между соседними идентичными элементами ЗС называется шагом структуры  $L$ , который может не совпадать с пространственным периодом  $D$  (рис. 1.15). Период  $D$  системы превышает шаг  $L$  в многоступенчатых структурах в несколько раз. Так, на рис. 1.15,а – в показаны одноступенчатые ЗС, когда  $L=D$ , а на рис. 1.15,г – двухступенчатая ( $L = D/2$ ).

Анализ решения волнового уравнения для передающих линий типа ЗС [4] показывает, что замедление волны возможно только в линиях, где поверхностное сопротивление одной из плоскостей имеет реактивный характер. Замедляющие структуры, приведенные на рис. 1.15, удовлетворяют этому требованию.

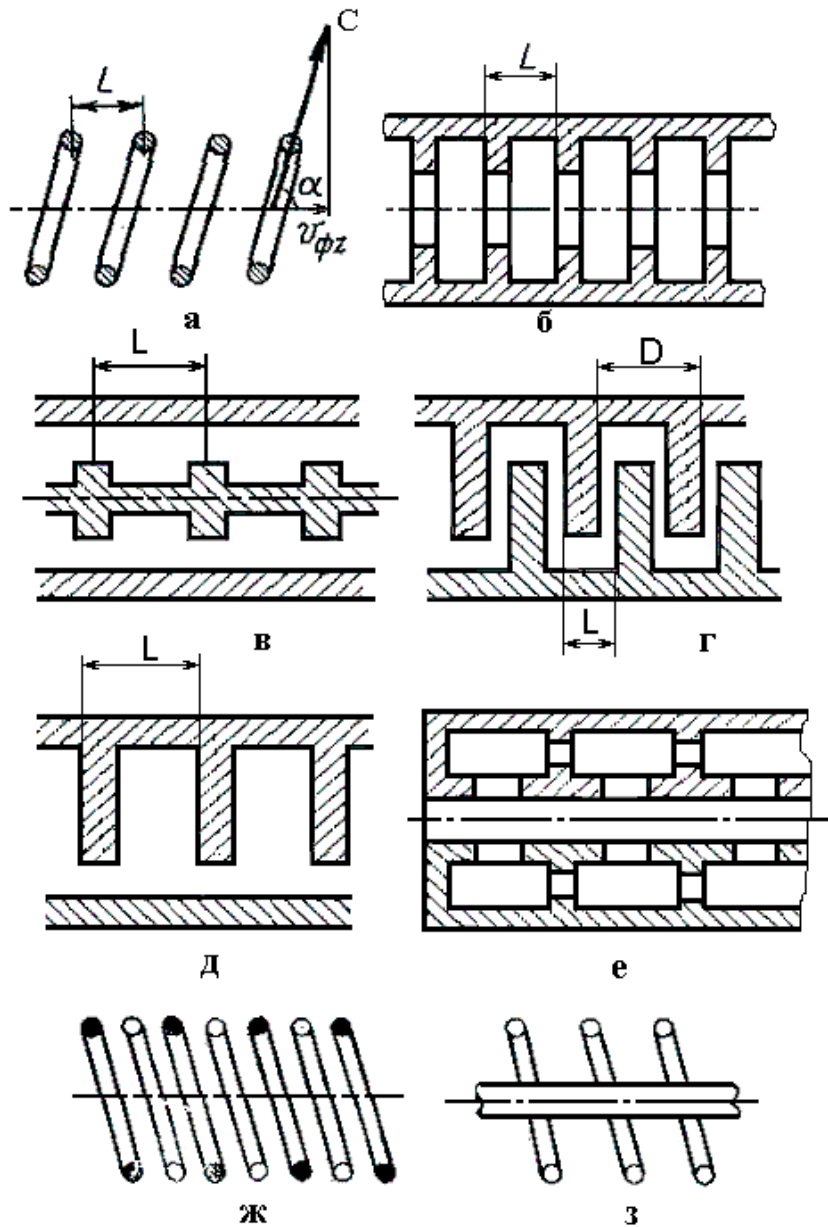


Рис. 1.15. Виды замедляющих систем: а - спиральная; б - цепочка связанных резонаторов; в - коаксиальная линия с диафрагмами во внутреннем проводнике; г - встречные штыри; д - гребенчатая; е - цепочка связанных резонаторов с магнитной связью; ж - двойная спираль; з - спираль с центральным металлическим стержнем.

Принцип замедления волны легко увидеть на примере спиральной ЗС (рис. 1.15,а), представляющей собой коаксиальную линию передачи со спиральным внутренним проводником. При возбуждении такой линии генератором СВЧ колебаний волна распространяется по спиральному проводнику со скоростью, близкой к скорости света.

Тогда, за время движения волны вдоль одного витка спирали длиной  $2\pi\alpha$ , где  $\alpha$  – радиус спирали, волна переместится вдоль оси системы на величину шага спирали  $L$ . Отношение длины витка к шагу спирали составляет величину замедления волны

$$K_3 = \frac{2\pi\alpha}{L},$$

где  $K_3$  – называется коэффициентом замедления.

**Электрическое поле в ЗС.** При распространении электромагнитной волны по идентичным ячейкам любой ЗС фаза колебаний в соседних ячейках отличается на некоторую величину  $\varphi$ .

$$E(z + L) = E(z)e^{j\varphi}. \quad (1.27)$$

Картина электрического поля на отрезке гребенчатой ЗС при  $\varphi < \pi$  и фиксированном моменте времени  $t$  показана на рис. 1.16.

Электрическое поле короткозамкнутых щелевых резонаторов сосредоточено вблизи щелей (рис. 1.16,а). Поэтому зависимость продольной составляющей поля  $E_z$  от координаты  $z$  (рис 1.16,б) представляет собой серию импульсов, имеющих синусоидальную огибающую [2].

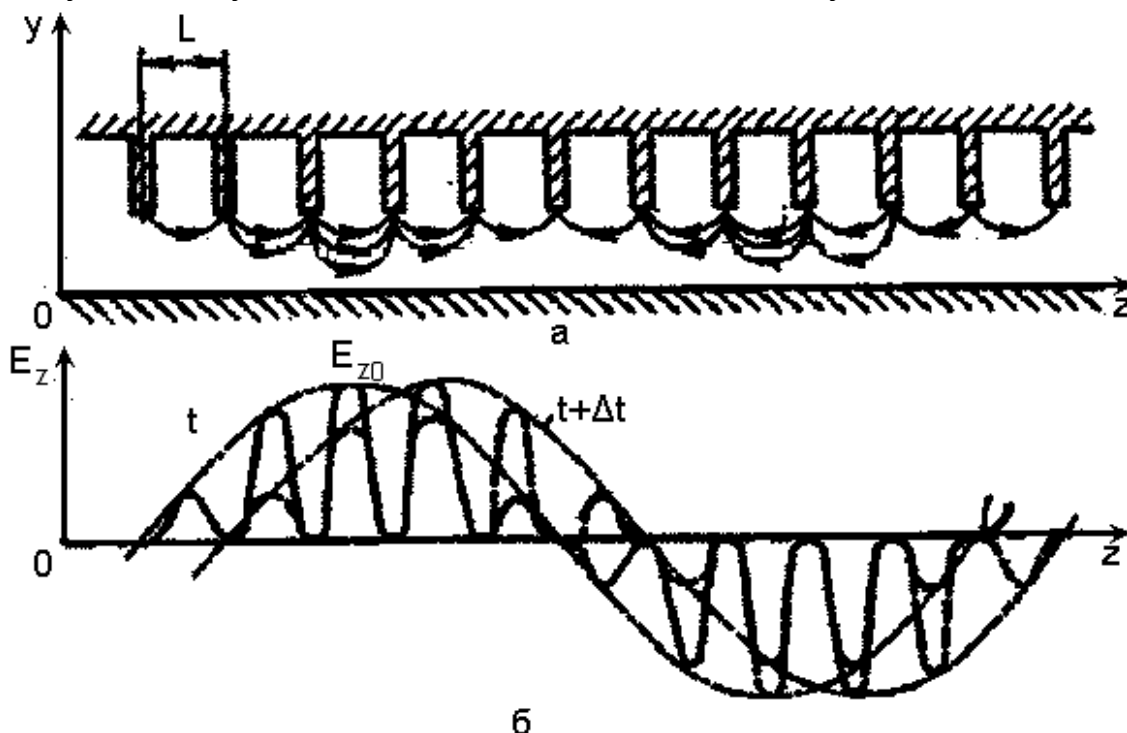


Рис.1 16. Распределение электрического поля вдоль длины ЗС

Через интервал времени  $\Delta t$  волна переместится на величину  $\Delta z$  и максимум поля  $E_z$  окажется вблизи следующих ячеек. Через половину периода фаза колебаний в каждой из ячеек ЗС изменится на противоположную величину. Таким образом, поле  $E_z$  будет меняться в такт с частотой колебаний генератора  $\omega$ . В любой момент времени функция  $E_z(z)$  представляет собой последовательность импульсов с

периодом следования, равным периоду структуры  $L$ . Такая пространственно-периодическая функция может быть представлена, подобно временной последовательности импульсов, комплексным периодическим рядом Фурье в виде

$$E_z(x, y, z, t) = \sum_{m=-\infty}^{m=+\infty} E_{zm}(x, y, z, m) e^{j\left[\omega t - \frac{z}{L}(\varphi + 2\pi m)\right]}, \quad (1.28)$$

где  $m$  – любое целое число.

Из (1.28) видно, что электрическое поле в периодической системе выражается суммой бесконечного ряда прямых (при  $m > 0$ ) и обратных волн ( $m < 0$ ), называемых пространственными гармониками. Математически пространственные гармоники появляются в результате разложения сложного (периодического, несинусоидального) поля (рис. 1.16,б) в ряд Фурье, каждое слагаемое которого является гармонической функцией, характеризуется своей амплитудой, фазовой скоростью  $V_{\varphi m}$  и длиной волны  $\lambda_{zm}$ . Пространственные гармоники не существуют отдельно сами по себе. Они определяют общий волновой процесс в замедляющей системе на одной частоте, связаны определенными соотношениями между амплитудами и фазовыми скоростями.

Величина поля  $E_z$  в поперечной плоскости  $z = \text{const}$  уменьшается при удалении от поверхности ЗС. Это уменьшение тем больше, чем больше коэффициент замедления и чем выше номер пространственной гармоники.

**Параметры волны в ЗС.** Введем соотношения для основных параметров волны, распространяющейся в ЗС. Фазовый набег волны на периоде ЗС  $L$  для гармоники  $m$  из (1.16) имеет вид

$$\varphi_m = \beta_m Z = \frac{Z}{L} \left( \varphi + \frac{2\pi m}{q} \right), \quad (1.29)$$

где  $q$  – число ступеней в ЗС, в случае одноступенчатых ЗС (рис.1.15)  $q=1$ ;  $\beta_m$  – фазовая постоянная волны  $m$ -ой гармоники в ЗС равняется

$$\beta_m = \frac{\varphi + 2\pi m / q}{L}. \quad (1.30)$$

Величина фазовой скорости в ЗС определяется соотношением

$$V_{\varphi m} = \frac{\omega}{\beta_m} = \frac{\omega L}{\varphi + 2\pi m / q}, \quad (1.31)$$

а длина замедленной волны

$$\lambda_{zm} = \frac{2\pi}{\beta_m} = \frac{V_{\varphi m}}{f} = \frac{2\pi L}{\varphi + 2\pi m / q}. \quad (1.32)$$

Отметим, что электромагнитное поле в неоднородной ( $D \geq \lambda_3$ ) периодической ЗС нельзя описать одной пространственной гармоникой, т.к. не выполняются граничные условия для одной отдельно взятой гармоники. Граничным условиям удовлетворяет лишь суперпозиция бесконечного числа полей пространственных гармоник, бегущих с разной фазовой скоростью, а, следовательно, имеющих разные амплитуды и фазовые набег на одной частоте колебаний.

Пространственные гармоники существуют все одновременно и на одной круговой частоте  $\omega$ , которая задается генератором.

Перемещение некоторой точки огибающей амплитуды поля, полученной в результате суммирования пространственных гармоник, характеризуется групповой скоростью

$$V_r = \frac{d\omega}{d\beta_m} = \frac{V_{\phi m}}{1 - \frac{\omega}{V_{\phi m}} \frac{\partial V_{\phi m}}{\partial \omega}}. \quad (1.33)$$

Так как знак постоянной  $m$  может быть любой, то из (1.31) следует, что фазовая скорость может быть положительной и отрицательной, групповая скорость всегда положительная величина. В первом случае гармоники называются прямыми, во втором, при  $V_{\phi m} < 0$ , – обратными. Положительная величина  $V_{\phi m}$  означает совпадение направлений фазовой и групповой скоростей, отрицательный знак перед фазовой скоростью  $V_{\phi m}$  говорит о том, что групповая скорость гармоники противоположна по направлению вектору фазовой скорости.

**Коэффициент замедления.** Отношение скорости света к фазовой скорости волны в замедляющей системе называется коэффициентом замедления

$$K_3 = \frac{c}{V_\phi}. \quad (1.34)$$

Для различных видов ЗС величина  $K_3 = 3 \div 40$ . Чем меньше коэффициент замедления ЗС, тем в более мощных приборах СВЧ используется такая замедляющая система.

С учетом (1.31) и (1.32) коэффициент замедления (1.34) гармоники  $m$  может быть представлен в другом виде

$$K_{3m} = \frac{\lambda}{\lambda_{3m}} = c \left( \frac{\varphi + 2\pi m / q}{2\pi L} \right). \quad (1.35)$$

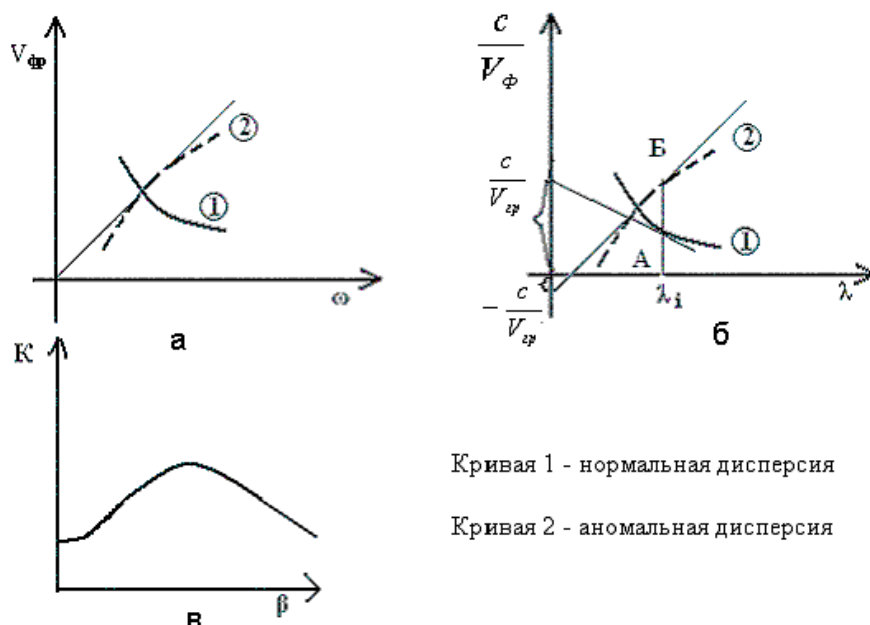
При  $m = 0$  (нулевая гармоника),  $K_{3m} = K_3$ , тогда для нулевой гармоники

$$K_3 = \frac{\lambda}{\lambda_3} = \frac{c\varphi}{2\pi L}. \quad (1.36)$$

### Дисперсионная характеристика.

Наиболее существенные свойства ЗС отражаются ее дисперсионной характеристикой. Зависимость фазовой скорости от частоты, представленная графически, называется дисперсионной характеристикой.

Дисперсионную характеристику (рис. 1.17) можно представлять кроме графика функции



Кривая 1 - нормальная дисперсия

Кривая 2 - аномальная дисперсия

Рис.1. 17. Виды дисперсионных характеристик

$$V_{\phi m} = \varphi(\omega); \quad \text{также в виде [4]} \quad K_{3m} = \varphi(\lambda) \quad \text{или} \quad K = f(\beta_m), \quad \text{где}$$

$$K = \frac{2\pi}{\lambda_0}.$$

Если абсолютная величина фазовой скорости уменьшается с ростом частоты, то дисперсия называется нормальной (рис. 1.17,а, б - кривая 1), в противном случае – аномальной (рис. 1.17,а, б - кривая 2). Помимо этого различают положительную (прямую) и отрицательную (обратную) дисперсии в зависимости от знака  $V_{\phi}$ . Для прямых гармоник ( $m>0$ ) дисперсия положительная ( $V_{\phi}$ ,  $V_{\Gamma}$  одного направления), а для обратных гармоник ( $m<0$ ) – дисперсия отрицательная (фазовая и групповая скорости противоположно направлены). В лампах бегущей волны применяются ЗС с положительной дисперсией, в лампах обратной волны – с отрицательной дисперсией.

Графики дисперсионных зависимостей обладают интересными свойствами. Если провести касательные к точкам кривой (А, Б на рис. 1.17,б) до пересечения с осью  $c/V_{\phi}$ , то в точке пересечения оси ординат в случае фиксированной  $\lambda_1$  можно найти величину  $c/V_{\Gamma}$ . Величина  $c/V_{\Gamma}$  в случае кривой 1 положительная (дисперсия положительная), в случае кривой 2 – отрицательная (дисперсия отрицательная).

Таким образом, каждая точка на дисперсионной характеристике несет информацию о величине  $V_{\phi}$ ,  $K_3$ ,  $\beta$ ,  $V_{\Gamma}$  для заданной рабочей частоты, а также информацию о величине и направлении движения скорости энергии волны  $v_{\Gamma}$ . Изучение характера дисперсионной зависимости ЗС необходимо для конструирования СВЧ приборов.

Выражение (1.35) используется для построения дисперсионных характеристик ЗС в виде зависимостей  $K_{3m}(\lambda)$  для разных гармоник  $r$ . При этом должна быть известна зависимость  $\varphi(\lambda)$  или  $\varphi(\omega)$ , которая определяется геометрией ЗС. Если  $\varphi = \text{const}$ , то график зависимости  $K_3(\lambda)$  будет линейным. Прямые, выходящие из начала системы координат ( $K_3$  и  $\lambda$ ), являются линиями постоянных  $\varphi$ .

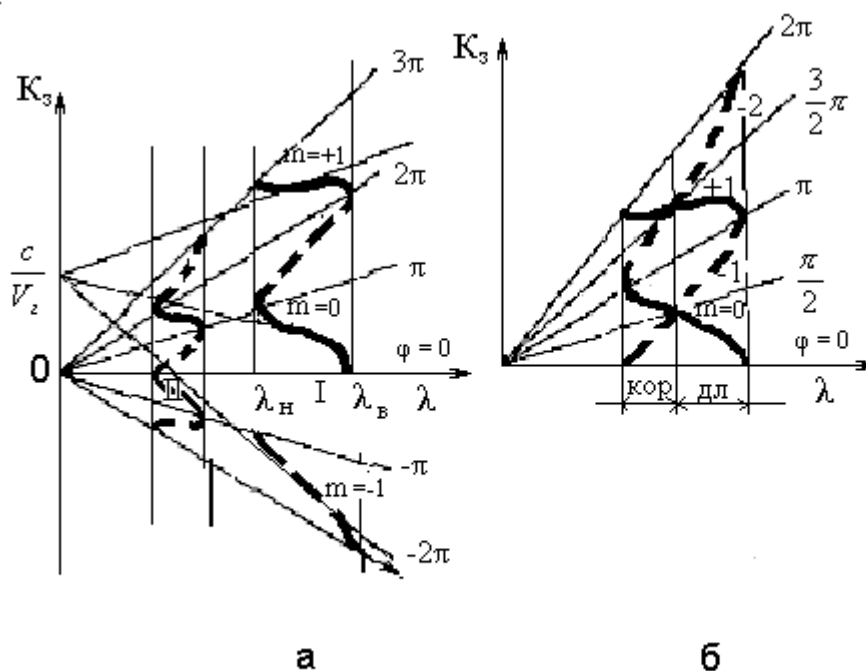


Рис. 1.18 Дисперсионные характеристики замедляющих систем.

Характерной особенностью периодических структур является наличие у распространяющегося типа волны двух длин волн (частот) отсечки – нижней и верхней, соответствующих значениям фазового сдвига на ячейку:  $\varphi = 0$  и  $\pi$ . При фазовых сдвигах  $\varphi = 0$  и  $\pi$  распространение волн в ЗС невозможно, так как  $c/V_{\Gamma}$  стремится к бесконечности.

На рис 1.18 изображены линии  $\varphi = \pm\pi, \pm 2\pi$ . Линия  $\varphi=0$  совпадает с осью абсцисс. Вертикальными линиями отмечены границы полосы пропускания системы. Дисперсионная характеристика нулевой гармоники заключена между осью абсцисс и прямой  $\varphi=\pi$ ; первой гармоники, при  $m=1$ , - между линиями  $\varphi=2\pi$  и  $\varphi=3\pi$ ; а при  $m=-1$ , т.е. характеристика первой обратной гармоники, находится между прямыми  $\varphi=-\pi$  и  $\varphi=-2\pi$  и т.д. Обычно дисперсионные кривые обратных гармоник переносят в верхнюю (положительную) полуплоскость чертежа и изображают пунктиром (рис. 1.18,б).

По дисперсионным характеристикам, зная рабочую длину волны в свободном пространстве, легко определить коэффициент замедления, фазовую скорость любой пространственной гармоники, проектируя соответствующую точку на ось ординат. Линия, касательная к дисперсионной характеристике в точке, заданной длиной волны  $\lambda_0$ , определяет отсчет на оси ординат, соответствующий величине  $c/V_{\Gamma}$ . Групповые скорости всех гармоник при заданной длине волны одинаковы.

**Электрическое поле пространственных гармоник.** Во всех электронных СВЧ приборах продольное электрическое поле пространственной гармоники играет важную роль во взаимодействии электронов с волной. Обычно выбирается основная пространственная гармоника, у которой фазовая скорость  $V_{\Phi m}$  максимальная. При этом коэффициент замедления (1.34) минимальная величина.

С ростом номера гармоники  $m$  уменьшается величина поля  $E_{zm}$  на поверхности ЗС и увеличивается спад поля в поперечном сечении по мере удаления от поверхности (рис. 1.19).

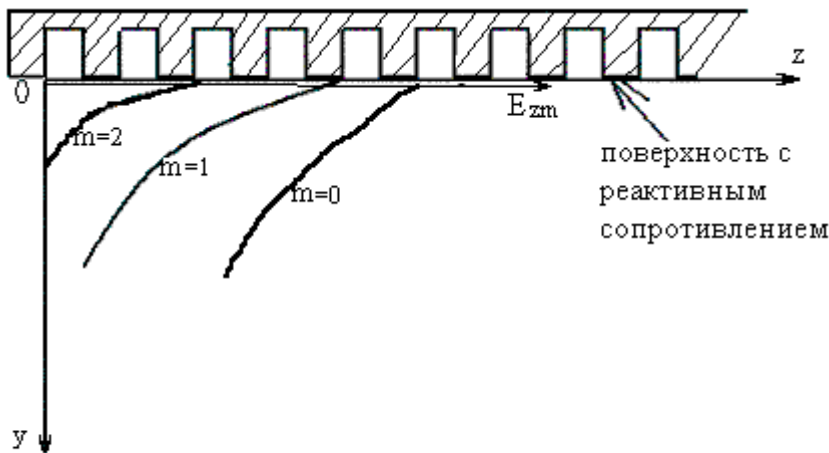


Рис.1.19. Затухание амплитуды электрического поля гармоник при удалении от поверхности ЗС

При использовании высоких номеров гармоник (рис. 1.19) уменьшается эффективность взаимодействия электронного потока, проходящего вдоль ЗС при  $y=\text{const}$ , с электрическим полем волны, т.к. происходит ослабление поля.

**Сопротивление связи.** В приборах СВЧ обычно рассчитывают и измеряют мощность электромагнитной волны  $P$ , которая зависит от величины напряженно-



сти продольного электрического поля. Для оценки этого взаимодействия вводят параметр ЗС, называемый сопротивлением связи, который для  $m$ -ой гармоники определяется выражением:

$$R_{св.m} = \frac{|E_{zm}|^2}{(2\beta_m^2 P)}, \quad (1.36)$$

где  $P$  - мощность, передаваемая через поперечное сечение ЗС. Сопротивление связи является функцией поперечных координат. Разные пространственные гармоники имеют разные величины сопротивлений связи. С увеличением номера гармоники  $m$  сопротивление связи падает, так как падает величина напряженности поля-  $E$ . На границах полосы пропускания передаваемая мощность равна нулю, а сопротивление связи в этих точках обращается в бесконечность. Поэтому в качестве рабочих пространственных гармоник выбирается обычно нулевая ( $m=0$ ) или для коротких волн первая ( $m=1$ ). Сопротивление связи для разных ЗС находится в пределах  $R_{св}=10 - n100$  Ом.

Для ввода и вывода энергии требуется ЗС согласовать с передающими линиями. Для этих целей следует применять параметр - волнового сопротивления ЗС, который имеет вид:

$$W_{ЗС} = \frac{E_{\perp}}{H_{\perp}} = \frac{\left| \int_a^b E_{\perp} dl \right|^2}{2P}.$$

Криволинейный интеграл вычисляется между точками на проводящих поверхностях по линии, лежащей в плоскости поперечного сечения системы. Волновое сопротивление определяется поперечными компонентами суммарного электрического поля  $E_{\perp}$ , а не полем одной используемой пространственной гармоники. Волновое сопротивление ЗС не зависит от номера гармоник, но может значительно отличаться от волновых сопротивлений передающих линий. Для их согласования в приборах СВЧ используются специальные трансформирующие устройства.

### 1.3.3. Разновидности замедляющих систем

**Спиральная замедляющая система** (рис 1.20) наиболее используемая в СВЧ приборах как самая широкополосная система. Отношение максимальной и минимальной частот для спирали в большинстве случаев равно двум, т.е. порядка октавы.

Дисперсионные характеристики спирали изображены на рис 1.21.

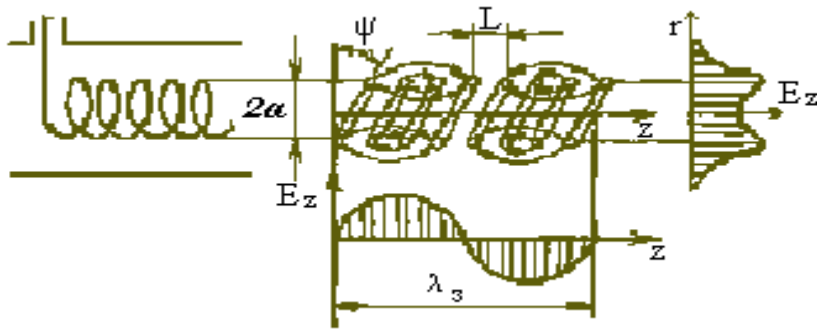


Рис. 1.20. Поле  $E_z$  в спиральной замедляющей системе

Коэффициент замедления определяется соотношением

$$K_s = \frac{\sqrt{(2\pi a)^2 + L^2}}{L^2} = \frac{1}{\sin \psi}, \quad (1.37)$$

где  $\Psi$  - угол намотки, и, очевидно, может регулироваться изменением геометрии  $\alpha$ ,  $L$ . Формула (1.37) соответствует случаю нулевой гармоники. Чем меньше угол намотки спирали  $\Psi$  (рис 1.20), т.е. меньше фазовый сдвиг  $\phi$ , тем более пологой становится дисперсионная характеристика на основной волне, а в области больших длин волн она вообще не имеет отсечки..

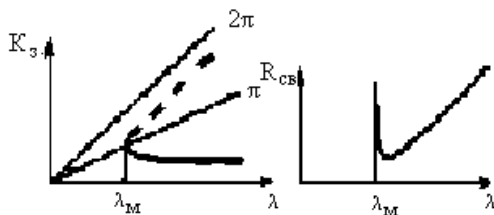


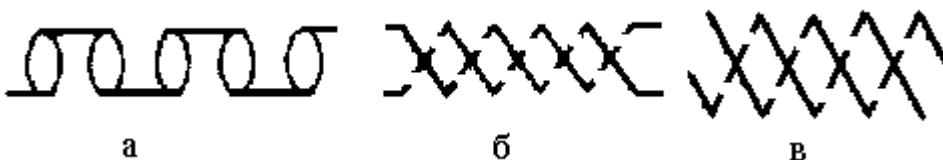
Рис. 1.21. Дисперсионные характеристики спирали

Сопротивление связи спирали растет с ростом длины волны. В СВЧ приборах коротковолнового диапазона спиральные ЗС не используются, так как размеры их становятся малыми и трудновыполнимыми.

К спиральным ЗС можно отнести следующие модификации:

- система типа кольцо – стержень (рис.1.22,а), которая предназначается для применения в мощных ЛБВ;
- система встречно-намотанных спиралей (рис.1.22, б), предназначается для использования в ЛБВ;
- двухзаходная спираль (рис.1.22,в), используется в ЛОВО;
- плоская спираль (меандр) и спираль с прямоугольными витками, успешно используются в усилителях прямой волны М – типа (рис.1.22,г, д).

Для ввода – вывода энергии в ЗС применяются следующие способы: *кондуктивный способ* или гальванический, при котором провод спирали напрямую соединяется с внутренним проводником наружной коаксиальной линии.



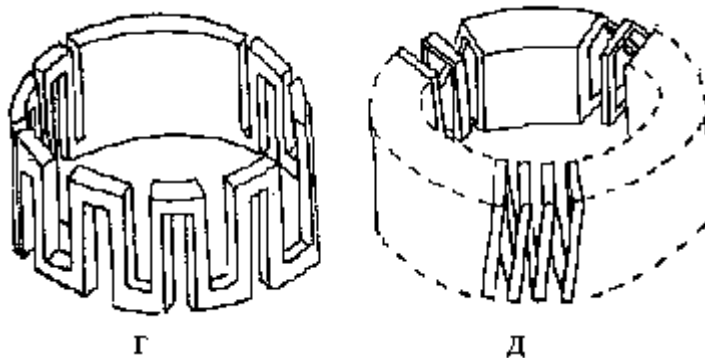


Рис. 1.22. Виды спиральных ЗС

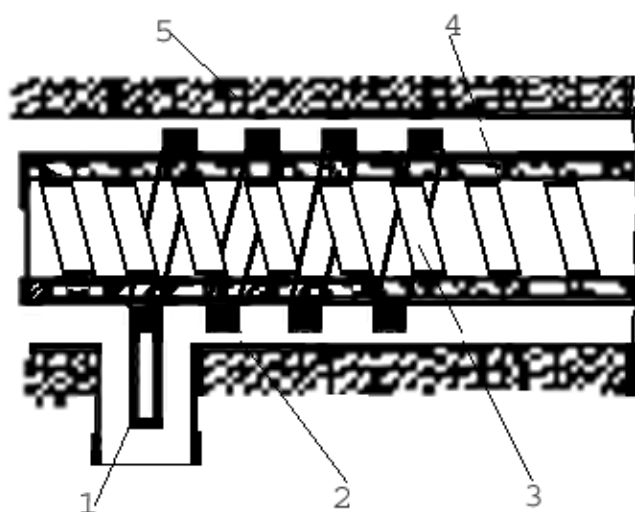


Рис. 1.23. Вывод энергии на связанных спиралах.

Вывод энергии типа связанные спиралы [3] показан на рис.1.23. Внутренний проводник коаксиальной линии (1) переходит во внешнюю спираль (2), надеваемую на колбу лампы (4), в которой закреплена спиральная замедляющая система (3). Длина области связи спиралей подбирается кратной числу замедленных полуволн. Для наилучшей передачи энергии фазовые скорости в спиралах должны быть равны. Это выполняется подбором диаметров и периодов спиралей: отношение диаметров должно равняться отношению периодов системы связанных спиралей.

Волноводно – коаксиальный вывод энергии [5] широко используется в коротковолновом диапазоне. Замедляющая система заканчивается полыми цилиндриками, которые являются штыревыми возбудителями электромагнитного поля в прямоугольном волноводе (рис. 1.24).

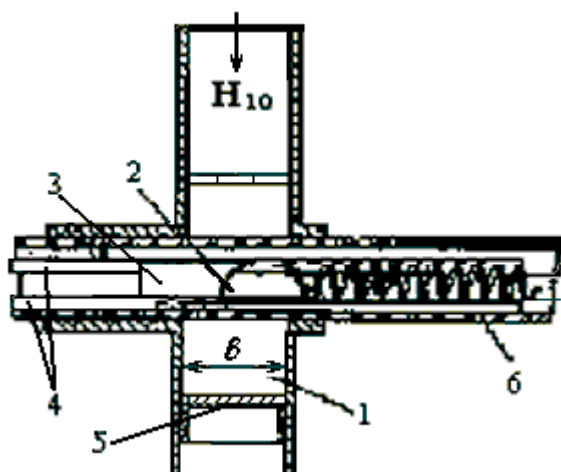


Рис. 1.24. Волноводно-коаксиальный вывод энергии: 1 – волновод; 2 – цилиндрический штырь, в который переходит спиральная ЗС; 3 – цилиндр; 4 – стеклянный стержень, фиксирующий спираль; 5 – поршень; 6 – колба.

**Система встречно штыревого типа.** Система встречных штырей состоит из двух гребенок, вставленных одна в другую (рис. 1.25). При этом между гребенчатыми структурами образуется петляющий волновод, по которому со скоростью света распространяется волна типа -Т. Вдоль оси системы фазовая скорость этой волны замедляется до величины

$$\frac{V_{\phi m}}{c} = \frac{L}{(2\ell + L) + \lambda(2m + 1)}, \quad (1.38)$$

где  $\ell$  - высота штырей.

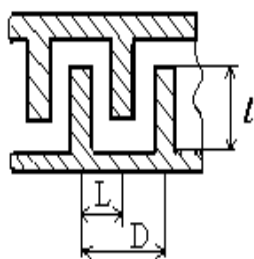


Рис. 1.25. ЗС типа встречные штыри

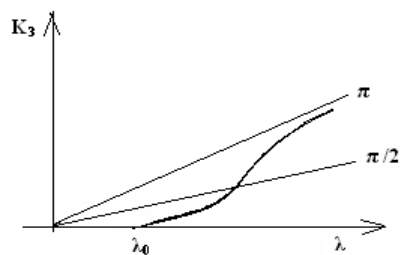
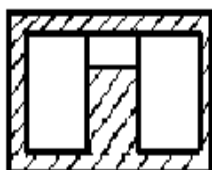


Рис. 1.26 Дисперсионная характеристика двухступенчатой ЗС типа встречные штыри

При номере гармоники  $m = -1$  фазовая скорость  $V_{\phi(-1)}$  (1.38) становится максимальной, а гармоника, при условии  $2\ell + L < \lambda$ ;  $L \ll \lambda$  и  $\ell < \lambda/4$ , становится основной. Так как при этих условиях фазовая скорость равна

$$V_{\phi(-1)} = -\frac{L \cdot c}{\lambda - (2\ell + L)} \quad (1.39)$$

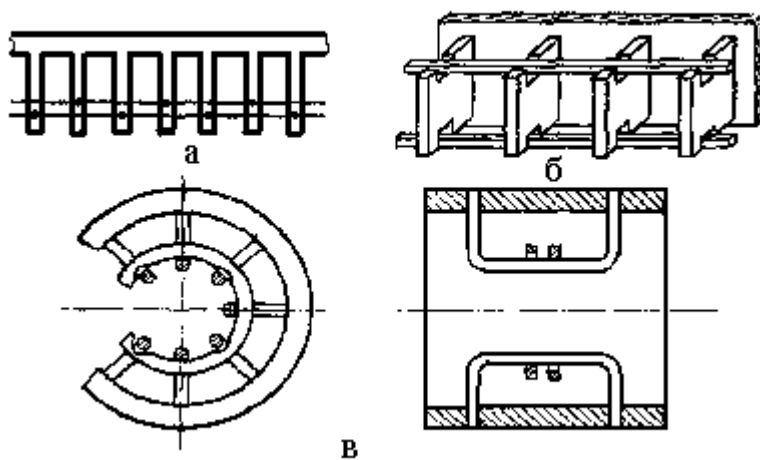


Рис. 1.27. Ламельные замедляющие системы со связками: а- с двойными двухсторонними; б- с одинарными двухсторонними; в- бугельная системы с внутренними связками.

и имеет знак минус, то это говорит об обратном направлении движения групповой скорости в ЗС типа встречные штыри и о возможности использовании этих ЗС в лампах обратной волны.

Дисперсионная характеристика встречно – штыревой ЗС изображена на рис. 1.26.

Разновидностью системы встречно штыревого типа являются [3,5] стержневые или ламельные системы со связками, показанные на рис. 1.27,а, 1.27,б, используемые в магнетронах и ЛОВМ, и бугельная система со связками (рис.1.27,в), применяемая в амплитронах – усилителях обратной волны типа М.

**Замедляющие системы типа цепочка связанных резонаторов (ЦСР)** – состоит из цилиндрических резонаторов, возбуждающихся на колебании вида  $E_{010}$  (рис.1.28) с разными фазовыми сдвигами  $\varphi$  поля в соседних резонаторах [4]. Система имеет еще название диафрагмированный волновод.

На рис. 1.28 показаны поля для крайних случаев:  $\varphi = 0$  и  $\varphi = \pi$ . Связь между резонаторами происходит через центральные отверстия в торцевых стенках, за счет этого создается центральный канал с интенсивным электрическим полем и емкостной связью. При  $\varphi = 0$  связь максимальная, при  $\varphi = \pi$  -связь ослабляется, так как силовые линии в соседних резонаторах имеют противоположное направление, и емкость связи уменьшается, рабочая длина волны  $\lambda_{\pi}$  уменьшается по отношению к  $\lambda_0$  ( $\lambda_0 = 2,62 \cdot R$  - резонансная длина волны для колебания  $E_{010}$  в резонаторе).

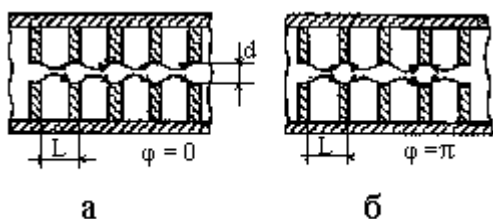


Рис. 1.28. Электрическое поле в системе типа цепочка связанных резонаторов при  $\varphi = 0$  (а) и  $\varphi = \pi$  (б).

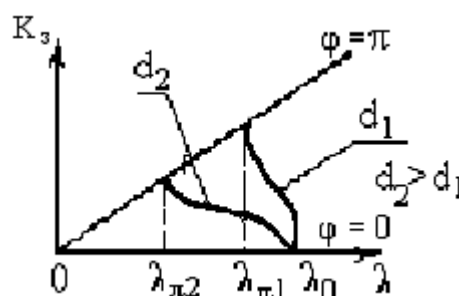


Рис. 1.29. Дисперсионная характеристика цепочки связанных резонаторов

Дисперсионная характеристика системы типа цепочка связанных резонаторов приведена на рис. 1.29, из которой следует, что дисперсия довольно крутая, но при увеличении диаметра отверстий длина  $\lambda_{\pi}$  укорачивается и характеристика становится положе.

Для широкополосных приборов СВЧ диаметр отверстия  $d$  почти приближается к диаметру цилиндрического волновода, хотя при этом сопротивление связи также уменьшается. Цепочка связанных резонаторов применяется в усилителях прямой волны.

Разновидностью рассмотренной системы является система ЦСР с магнитной связью между резонаторами. Конструктивно кроме центрального отверстия в диафрагме делаются щели в области магнитных силовых линий (рис. 1.30). Щели нарушают радиальную симметрию электромагнитного поля относительно оси ЗС.

Чтобы увидеть какие изменения произойдут в дисперсионной характеристике, следует рассмотреть поле в щелях при разных фазовых сдвигах на ячейке ЗС.

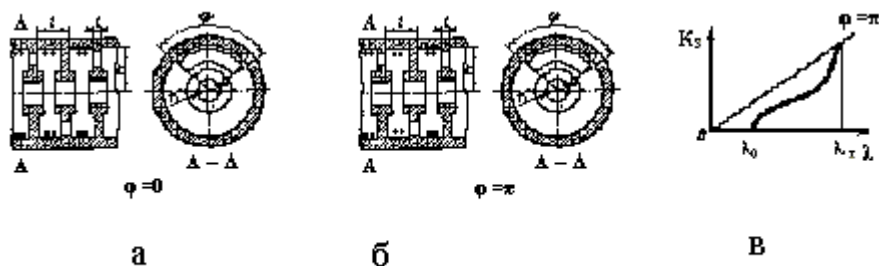


Рис. 1.30. Цепочки связанных резонаторов с магнитной связью

При  $\varphi = 0$  (рис. 1.30, а) линии магнитного поля во всех резонаторах в области щелей направлены в одну сторону, стенка может быть удалена и ничего не изменится. Длина волны  $\lambda_0$  останется такой же, как у одного изолированного резонатора на волне типа  $E_{010}$ . ( $\lambda_0 = 2,62 \cdot R$ , где  $R$  – радиус резонатора).

При сдвиге на периоде системы  $\varphi = \pi$  магнитные линии в щелях будут противоположно направлены. Возникающие в стенках диафрагмы токи  $J_C$ , по правилу буравчика имеют одинаковое направление (рис 1.30, б). Так как линии полного тока непрерывны, то в щелях связи возбуждается электрическое поле, создающее ток смещения, равный суммарному току в стенках диафрагмы в пределах щели. Возникновение электрического поля щели и связанного с ним магнитного поля эквивалентно появлению последовательной дополнительной индуктивности, что уменьшает частоту резонатора со щелью. Длина  $\lambda_\pi$  становится больше  $\lambda_0$ , в результате дисперсионная характеристика цепочки связанных резонаторов со щелями связи обладает отрицательной дисперсией (рис. 1.30,в). Изменение размеров щелей и их местоположения в диафрагме приводит к изменению наклона дисперсионной кривой.

Возбуждение ЗС типа цепочки связанных резонаторов осуществляется через отверстие в боковой стенке или с торца волновода (рис 1.31).

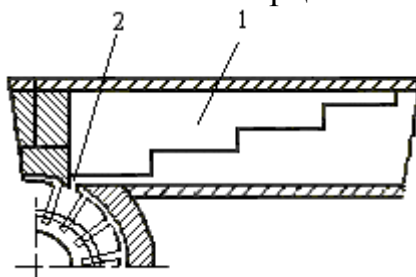


Рис. 1.31. Связь волновода с ЗС типа цепочки связанных резонаторов.

Различные способы ввода или вывода энергии СВЧ колебаний [14] приведены на рис.1.32. При разработке вывода энергии ЗС вопросу его широкополосности уделяется особое внимание. Для расширения рабочей полосы частот ис-

пользуются плавные или многоступенчатые переходы. В случае плавных переходов изменяется форма или размеры проводников ЗС или передающей линии. Так, например, в плавных переходах со стандартного прямоугольного волновода к системе гребенчатого типа (в,г) изменяется глубина щелей ЗС, а в системе встречные штыри (б) – узкий размер волновода изменяется по экспоненте. В переходах от стандартной коаксиальной линии к системе лестничного типа (е) изменяется плавно диаметр и форма проводников линии. Многоступенчатый трансформатор (а) использован в системах встречно-штыревого типа для дециметрового диапазона.

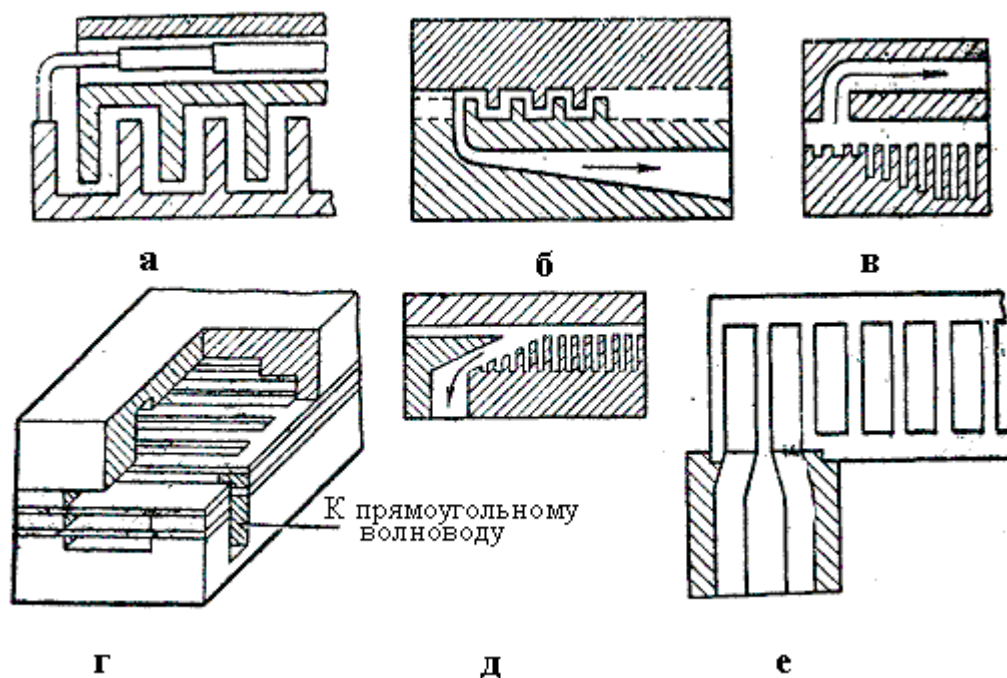


Рис. 1.32. ЗС типа встречные штыри — коаксиальная линия (а) с использованием двухступенчатого трансформатора сопротивлений; система встречные штыри — волновод с плавным изменением узкого размера волновода (б); гребенчатая система с переменной высотой гребней — волновод (в,г); лестничная система — волновод (д); лестничная система — коаксиальная линия с плавным изменением размеров диаметров (е).

Качество согласования проверяется экспериментально, посредством измерения диапазонной зависимости коэффициента стоячей волны, который не должен быть более 1,3.

#### 1.4. Принципы модуляции электронного потока

От управляющего устройства требуется, чтобы в некоторой плоскости, соответствующей выходному устройству, конвекционный ток менялся по закону

$$i_{\text{конв}} = I_0 + i(t),$$

где  $i(t)$  — периодическая функция времени,  $I_0$  — постоянная составляющая конвекционного тока.

Переменная составляющая конвекционного тока  $i(t)$  не обязательно должна создаваться в самом управляющем устройстве, поэтому целесообразно применять

такое воздействие на электроны, чтобы они, двигаясь в пространстве между управляющим и выходным устройствами, постепенно группировались, образуя сгустки как раз при входе в выходное устройство. Это можно осуществлять путем периодического изменения скоростей электронов.

Периодическое изменение скорости электронов, производимое в специальном управляющем устройстве (например, в резонаторе или замедляющей системе), получило название **модуляции электронного потока по скорости**.

Сначала рассмотрим *резонансные устройства* для создания скоростной модуляции.

Для того чтобы получить наиболее сильное изменение скоростей, электронный поток следует пропускать через отверстие в той части резонатора, где сосредоточено преимущественно электрическое поле. В случае тороидального резонатора, таким участком является емкостной зазор, который выполняется часто в виде двух плоских параллельных сеток. Направление начальной скорости электронов  $v_0$  должно по возможности совпадать с направлением электрических силовых линий в зазоре. Скоростная модуляция, происходящая при параллельности вектора скорости электронов и вектора напряженности электрического поля, имеет продольный характер. От поперечной модуляции избавляются наложением на систему постоянного по времени продольного магнитного поля с бесконечно большой напряженностью.

**Модуляция электронного потока по скорости.** Рассмотрим процесс изменения скорости гармоническим напряжением  $U = U_m \sin \omega t$ , наложенным на электроды плоского зазора. Начало координат выберем в центре зазора. Энергия, приобретаемая одиночным электроном при прохождении пути  $dx$  внутри зазора, равна:

$$dW = \frac{eU_m}{d} \sin \omega t dx.$$

Будем считать напряжение  $U_m$  положительным, если оно ускоряет электроны, двигающиеся в направлении оси  $x$ .

Полное приращение  $\Delta W$  найдем, интегрированием по ширине зазора, предполагая, что зазор находится посередине электрода:

$$\Delta W = \int_{-\frac{d}{2}}^{\frac{d}{2}} \frac{eU_m}{d} \sin\left(\omega t_0 + \frac{\omega x}{V_0}\right) dx = eU_m \sin \omega t_0 \frac{\sin \frac{\theta}{2}}{\frac{\theta}{2}} = e \gamma U_m \sin \omega t_0,$$

где мы выразили  $t = t_0 + \frac{x}{V_0}$ , пренебрегая малым изменением скорости электрона внутри зазора.

Исходя из полной кинетической энергии электрона

$$\frac{mV^2}{2} = eU_0 + e \cdot \gamma U_m \sin \omega t_0,$$

найдем его скорость на выходе из зазора:



$$V = \sqrt{\frac{2e}{m} U_0 \left(1 + \frac{\gamma U_m}{U_0} \sin \omega t_0\right)}.$$

Полагая  $U_m \ll U_0$ , разложим это выражение в ряд и, отбросив все члены выше первой степени, получим:

$$V \approx V_0 + \frac{\gamma U_m}{2U_0} \sin \omega t + \dots$$

Обозначив за  $V_1$  коэффициент, стоящий перед синусом ( $V_1 = \frac{\gamma \cdot U_m}{2U_0} V_0$ ),

получим: 
$$V = V_0 + V_1 \sin \omega t. \quad (1.40)$$

**Модуляция электронного потока по плотности.** Это преобразование должно происходить автоматически за пределами управляющего устройства в пространстве между модулирующим и выходным зазорами. В современной электронике СВЧ существует 2 метода: дрейфа и тормозящего поля. Схемы устройств, обеспечивающих оба вида преобразования, показаны на рис. 1.33 и 1.34.

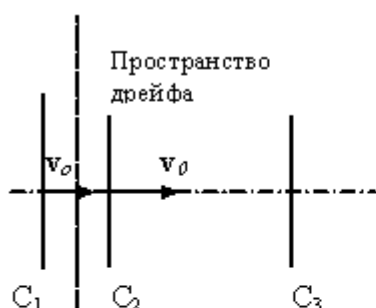


Рис. 1.33. Схема модуляции электронного потока в пространстве дрейфа



Рис. 1.34. Схема модуляции электронного потока в пространстве группирования

Первый метод (см. рис. 1.33, где  $C_1$ - $C_2$  – пространство между сетками резонатора,  $C_2$ - $C_3$  – пространство дрейфа,  $C_3$ -плоскость выходного устройства) основывается на движении модулированных электронов в пространстве, свободном от каких-либо внешних постоянных или переменных полей (пространство дрейфа), где происходит их группирование. Второй метод (см. рис. 1.34, где  $C_1$ - $C_2$  – пространство между сетками резонатора,  $C_3$ - отражатель) базируется на движении электронов в пространстве  $C_2$ - $C_3$  с постоянным во времени тормозящим электрическим полем.

На рис. 1.35 построена ПВД движения электронов в системе, изображенной на рис. 1.33, где в сеточном зазоре  $C_1$ -  $C_2$ , происходит модуляция электронов по скорости, а в промежутке  $C_2$  -  $C_3$  модуляция по плотности. Все пространство  $C_1$  –  $C_3$  находится под по-

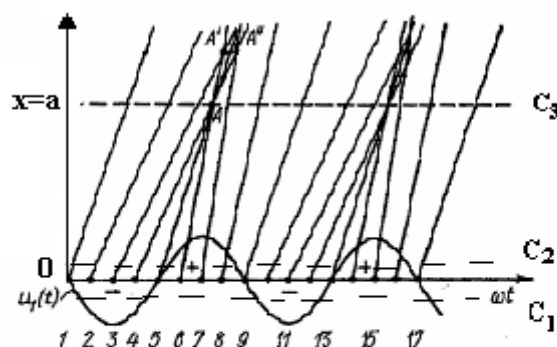


Рис. 1.35. Пространственно-временная диаграмма для модуляции электронного потока в пространстве дрейфа

стоянным потенциалом. Пространство между модулирующим зазором и выходным устройством, свободное от полей, называют **пространством дрейфа** или **пространством группировки**.

Рассмотрим движение электронов, прошедших модулирующий зазор в различные моменты времени  $t_0$ . Электроны типов 1,5,9,13 прошедшие зазор в моменты нулевого ВЧ поля, не изменяют своих скоростей. Электроны типа 7,15 скачком увеличивают свою скорость, в то время как электроны типа 3,11, проходя зазор в момент максимального тормозящего поля, несколько замедляются в сравнении с электронами типов 5 и 13. В результате, в некоторой плоскости  $X=\text{const}=a$ , где может быть расположен выходной зазор, электронный поток оказывается сгруппированным, т.е. приобретает модуляцию по плотности. Глубина этой модуляции зависит от расстояния между управляющим и выходным устройствами, т.е. от времени пролета в пространстве дрейфа. При фиксированной частоте и заданных значениях напряжений  $U_0$  и  $U_m$  существует оптимальная длина пространства дрейфа, обеспечивающая наибольшую переменную составляющую конвекционного тока в выходном устройстве. Значит, можно сказать, что центрами сгустков, приходящих в выходное устройство при преобразовании методом дрейфа, являются электроны, прошедшие середину управляющего зазора в момент нулевого ВЧ при переходе поля от тормозящего к ускоряющему. Период следования сгустков равен периоду модулирующего напряжения.

Рассмотрим теперь схему преобразования модуляции по скорости в модуляцию электронного потока по плотности методом тормозящего поля рис. 1.34. В пространстве позади управляющего зазора в отличие от метода дрейфа имеется продольное постоянное во времени тормозящее электрическое поле. Оно создается специальным электродом - отражателем, находящемся под отрицательным потенциалом по отношению к катоду и ускоряющему электроду.

Пространство между управляющим зазором и отражателем называется пространством группировки. Рассмотрим ПВД для этого случая (см. рис. 1.36). Здесь в пространстве группировки имеется лишь постоянное тормозящее поле. Если потенциал отражателя близок к потенциалу катода, то происходит модуляция возвращающегося электронного потока по плотности за счет оседания части электронов на отражателе. Более эффективная модуляция, когда все электроны, не достигая отражателя, двигаются обратно по направлению к управляющему зазору. Как видно из рис. 1.36 электроны типов 3 и 7, прошедшие управляющий зазор соответственно при максимально ускоряющем и максимально тормозящем полях, группируются относительно электрона типа 5.

Таким образом, в случае преобразования методом тормозящего поля, сгустки образуются носительно электронов, прошедших центр зазора в момент левого высокочастотного поля (ВЧ) поля при переходе от ускоряющего к тормозящему. В этом различие методов дрейфа и тор-

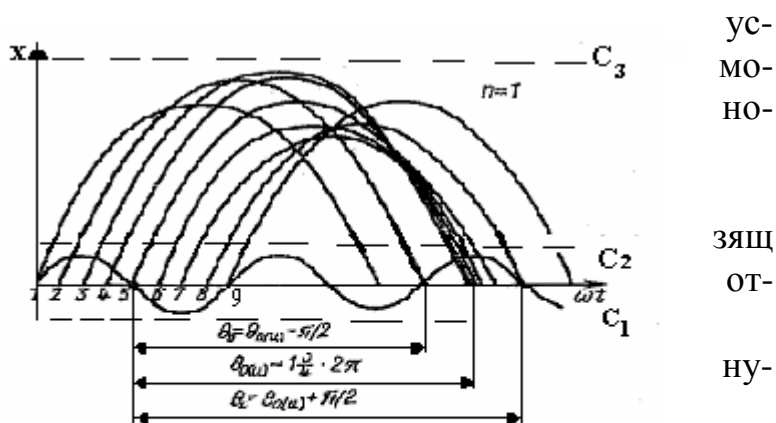


Рис. 1.36. Пространственно-временная диаграмма для модуляции электронного потока в пространстве группирования

зющего поля.

Существует другая особенность рассмотренных методов. В методе дрейфа может быть использована модуляция по плотности в выходном устройстве, находящимся позади зазора при прямом однократном пролете электронов. Этот способ лежит в основе современных пролетных двух резонаторных и многорезонаторных клистронов.

Во втором же случае удобнее реализовать модуляцию по плотности в плоскости, совпадающей с плоскостью модулирующего зазора. Последнее позволяет совместить управляющее и выходное устройство в одном зазоре. Этот принцип лежит в основе отражательных клистронов.

**Нерезонансные устройства для создания скоростной модуляции.** Рассмотренные выше управляющие устройства являются узкополосными, т.к. имеющийся в них СВЧ зазор входит в состав полого резонатора. Для создания скоростной модуляции в широкой полосе частот можно перейти к нерезонансным системам с длительным взаимодействием электронного пучка с полем бегущей волны. В качестве модулирующего устройства может быть применена замедляющая система - устройство с замедленной электромагнитной волной. Рассмотрим отрезок замедляющей системы, например, спирали, на вход которой подан СВЧ сигнал (рис. 1.37). В эту систему поступает электронный пучок. Для качественного объяснения скоростной модуляции рассмотрим случай синхронизма, когда  $V_0 = V_\phi$ , где  $V_\phi$  - фазовая скорость замедленной электромагнитной волны.

График распределения продольной составляющей бегущей волны в системе координат  $Z'$ , двигающихся со скоростью  $V_\phi$ , показан на рис. 1.38.

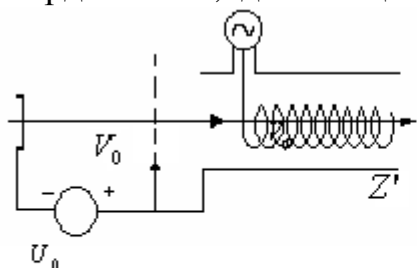


Рис. 1.37. Отрезок спиральной замедляющей системы

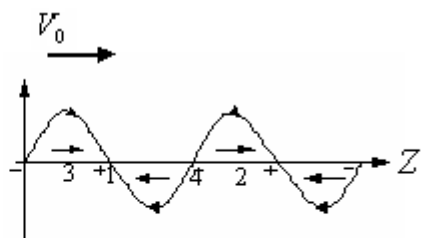


Рис. 1.38. График распределения продольной составляющей бегущей волны в ЗС

При  $V_0 = V_\phi$  электроны типов 1 и 4 остаются неподвижными, так как поле для них равно нулю. Электроны 3 ускоряются, а 2 - замедляются, поэтому они собираются в сгустки вокруг электрона типа 1, расположенных в областях перехода СВЧ поля от ускоряющей к тормозящей полуволне.

Напряженность СВЧ поля в замедляющей системе значительно меньше, чем в резонаторе. Однако угол пролёта электронов здесь ничем не ограничен, и путь, на протяжении которого поле модулирует электронный поток, гораздо больше, чем в резонаторе. В результате коэффициент усиления может быть достигнут такого же порядка, что и в высокодобротных системах с кратковременным взаимодействием.

## 1.5. Отбор энергии от электронного потока

Конечной целью работы подавляющего большинства электровакуумных приборов СВЧ, том числе генераторных и усилительных, является преобразование энергии постоянного тока в ВЧ энергию. Поэтому большой интерес представляет механизм отбора ВЧ энергии от электронного потока и условия, при которых может быть достигнуто наиболее полное превращение энергии.

**Отбор энергии от электронного потока с помощью резонансных колебательных систем.**

Обратимся к плоскому 2-х сеточному зазору (рис.1.13), в который поступает модулированный по плотности электронный поток. Если между электродами зазора включено активное сопротивление  $R$  (рис. 1.39), то в нем рассеивается мощность, обусловленная протеканием наведенного тока. Наведённый ток, проходя по внешнему сопротивлению  $R$ , создаёт на нём падение напряжения, с полярностью, обозначенной на рис. 1.39. Тогда электрод, к которому двигаются электроны, оказывается под отрицательным потенциалом, т.е. в момент прохождения электронов в зазоре автоматически создаётся тормозящее электрическое поле (см. рис. 1.16). В результате взаимодействия электронного потока, с электрическим полем волны, кинетическая энергия электрона уменьшается на величину энергии, отданной полю волны, следовательно, во внешнюю цепь.

**Вывод:** Энергия передаётся во внешнюю цепь в процессе движения электрона в продольном тормозящем электрическом поле. В диапазоне СВЧ используется объёмный резонатор, т.к. резонансных частотах он обладает активным сопротивлением.

Электроны пролетают через ёмкостную часть резонатора, где сосредоточено преимущественно электрическое поле СВЧ (рис. 1.13).

Если через зазор резонатора (рис. 1.15) пропустить один электронный сгусток, то колебания, возникающие в резонаторе СВЧ, будут затухающими. Для поддержания колебаний надо в нужные моменты снова пропускать через зазор сгустки - очевидно, в момент, когда поле будет тормозящим. Отсюда следует, что наибольший отбор энергии от электронного потока будет в том случае, когда резонансная частота резонатора равна частоте доведения сгустков.

На основании этого идеальная форма конвекционного течения будет следующей (рис. 1.40), чтобы  $\Delta t \ll T/2$ .

Для эффективного взаимодействия электронного потока с ВЧ полем

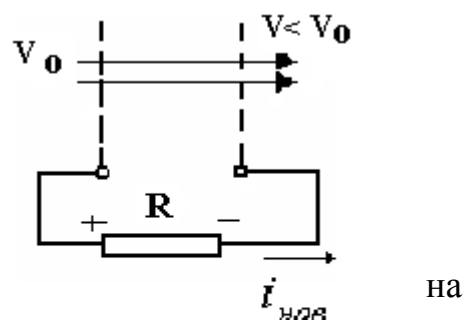


Рис. 1.39. Схема двухсеточного зазора

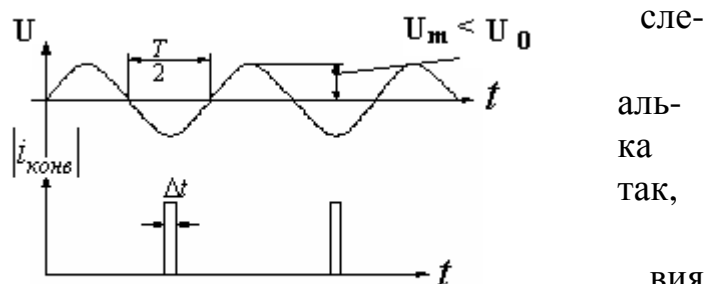


Рис. 1.40. Форма конвекционного тока

пролета должен быть меньше  $\pi$  ( $\theta < \pi$ ).

По этой причине большинство электронных приборов, использующих объёмные резонаторы, принято относить к классу устройств с кратковременным взаимодействием электронов с электромагнитным полем. Недостаток таких устройств - их узкополосность.

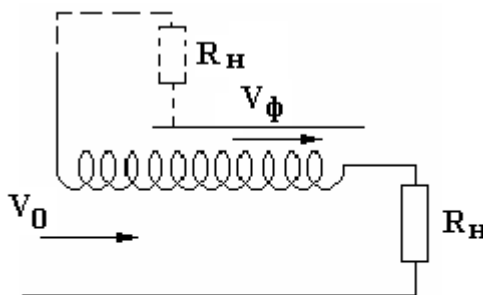
### ***Отбор энергии от электронного потока с помощью нерезонансных колебательных систем.***

Одним из путей расширения полосы частот, рассмотренных устройств является применение замедляющих систем. В СВЧ диапазоне используют для этих целей спиральные, встречно-штыревые замедляющие системы, типа гребёнка, т.д.

Ширина полосы частот рассматриваемого устройства определяется дисперсионной характеристикой замедляющей системы, т.е.  $V_\phi = F(\omega)$ . Чем слабее дисперсия, тем больше полоса частот, в которой при заданной  $V_0$  обеспечивается синхронизм скоростей ( $V_0 = V_\phi$ ) между электронами и волной.

Для СЗС -  $\Delta f \approx 50 \div 100\%$  от  $f_0$ . Поэтому в широкополосных усилителях на ЛБВ используются спиральные замедляющие системы. Резонатор же, даже при  $\theta_H = 100$  даёт  $\Delta f = 1\%$  от  $f_0$ .

Направления  $V_\phi$  и  $V_{гр}$  совпадают и  $V_0 = (V_\phi)_0$ , вывод энергии находится на конце системы (рис. 1.41).



**Рис. 1.41. Эквивалентная схема вывода энергии из замедляющей системы**

При использовании пространственных гармоник ЗС с отрицательной дисперсией, когда направление фазовой скорости  $V_\phi$  противоположно групповой скорости  $V_{гр}$  и  $V_0 = (V_\phi)_{-1}$ , энергия должна отбираться в нагрузку с конца системы, обращённого в сторону источника электронного потока. Так как передача энергии электронов полю бегущей волны может происходить на большом протяжении ЗС, то электронные приборы, использующие замедляющие системы, принято относить к классу приборов с длительным взаимодействием электронов с электромагнитным полем. Это ЛОВ, ЛБВ, магнетроны и т.д.

## **1.6. Волны в электронных потоках**

Следуя методике [4], рассмотрим электронный поток - как некоторую среду, в которой распространяются электромагнитные волны. Возьмем электронный поток, движущийся в пространстве дрейфа, где отсутствуют внешние электрические поля. Предположим, что движение электронов происходит только вдоль оси  $z$ . Наложим на этот поток электромагнитные возмущения, например, с помощью простейшего 2-х сеточного зазора. Будем полагать, что все величины, характеризующие поле и электронный поток  $E, H, V, A, \rho$ , изменяются по гармоническому закону с запаздыванием вдоль оси  $z$ . Эту зако-

номерность можно выписать в ком-плексной форме множителем  $e^{j\omega t - \Gamma z}$ , где  $\Gamma$  – постоянная распространения рассматриваемой волны. В любом сечении волновода, при распространении волны типа E ( $H_z=0$ ) скорость электронов, плотность объемного заряда, продольную составляющую вектора-потенциала можно записать:

$$\left. \begin{aligned} V &= V_0 + V_1 \cdot e^{j\omega t - \Gamma z}, \\ \rho &= \rho_0 + \rho_1 \cdot e^{j\omega t - \Gamma z}, \\ A &= A_0 + A_1 \cdot e^{j\omega t - \Gamma z}, \end{aligned} \right\} \quad (1.41)$$

где  $V_1, \rho_1, A_1$  – амплитуды переменных составляющих скорости, плотности заряда, вектора-потенциала соответственно.

Общее уравнение, описывающее распространение электромагнитных волн в произвольной, однородной, передающей линии в присутствии электронного потока имеет вид:

$$\nabla_{xy}^2 \cdot A_1 + A_1 \cdot (k^2 + \Gamma^2) \cdot \left( 1 - \frac{\omega_n^2}{(\omega + j \cdot \Gamma \cdot V_0)^2} \right) = 0, \quad (1.42)$$

где  $\omega_n$  – является круговой частотой собственных колебаний электронного потока.

$$\begin{aligned} \text{Обозначим:} \quad & (k^2 + \Gamma^2) \cdot \left( 1 - \frac{\omega_n^2}{(\omega + j \cdot \Gamma \cdot V_0)^2} \right) = P^2 \\ & \omega_n = 2\pi \cdot f_n, \quad f_n = 8980 \cdot \sqrt{N} \text{ [Гц]}, \end{aligned} \quad (1.43)$$

где  $N$  – концентрация электронов в пучке.

Тогда, рассматриваемое волновое уравнение, записывается в виде:

$$\nabla_{xy}^2 \cdot A_1 + A_1 \cdot P^2 = 0. \quad (1.44)$$

Это уравнение позволяет в принципе, зная граничные условия, найти структуру поля, параметры, характеризующие рассматриваемый волновой процесс.

Рассмотрим случай, когда отсутствуют вариации поля плотности зарядов и всех прочих величин по переменному сечению линии и электронного пучка, т.е.  $\nabla_{xy}^2 \cdot A_1 = 0$ .

Тогда уравнение (1.42) даёт:

$$A_1 \cdot (k^2 + \Gamma^2) \cdot \left( 1 - \frac{\omega_n^2}{(\omega + j \cdot \Gamma \cdot V_0)^2} \right) = 0. \quad (1.45)$$

Случай  $A_1 = 0$  интереса не представляет.

Решение  $k^2 + \Gamma^2 = 0$  также является тривиальным и соответствует волне типа ТЕМ, распространяющейся в отсутствии электронного потока в свободном пространстве.

Рассмотрим остающееся условие, зависящее от присутствия электронного луча:

$$1 - \frac{\omega_n^2}{(\omega + j \cdot \Gamma \cdot V_0)^2} = 0, \quad \frac{\omega_n}{\omega + j \cdot \Gamma \cdot V_0} = \pm 1.$$

Наибольший интерес представляет постоянная распространения  $\Gamma$ , которая оказывается равной:

$$\Gamma = j \cdot \frac{\omega \pm \omega_n}{V_0}. \quad (1.46)$$

Таким образом,  $\Gamma$  получилась чисто мнимой и имеет два значения, соответствующие различным знакам. Обозначим  $\Gamma = j\beta$ , где  $\beta$  - фазовая постоянная.

В общем случае фазовая постоянная определяется через круговую частоту  $\omega$  и фазовую скорость волн  $V_\phi$  в виде:

$$\beta = \omega / V_\phi. \quad (1.47)$$

Отсюда фазовая скорость волн, распространяющихся вдоль электронного потока, равна:

$$V_\phi = V_0 \cdot \frac{\omega}{\omega \pm \omega_n}. \quad (1.48)$$

Это уравнение показывает, что вдоль электронного луча могут распространяться две волны, скорости которых соответственно несколько больше и несколько меньше  $V_0$ . Как следует из этого уравнения, полученные волны имеют дисперсию и уже по этой причине не могут быть отнесены к классу волн ТЕМ. Волновое число  $p$  для этих волн равно 0.

Поперечные координаты являются независимыми константами, поэтому все пять  $E_x, E_y, H_x, H_y, H_z$  составляющих поля равны нулю, а  $E_z$  составляющая равна:

$$E_z = -j \cdot D \cdot \frac{k^2 - \beta^2}{\omega \cdot \epsilon_0} \cdot e^{j\omega t - \Gamma z}. \quad (1.49)$$

Следовательно, рассматриваемые волны можно отнести к особому типу электрических волн. В отсутствии электронного потока такие волны существовать не могут. Физическую сущность волн легко понять, рассматривая изменение плотности заряда  $\rho$  вдоль луча. Поскольку можно показать, что  $\rho_1=0$ , то вдоль луча периодически происходят сжатие и уплотнение электронного газа. Продольное электрическое поле  $E_z$  обусловлено именно этими уплотнениями.

Такие волны могут быть названы волнами сжатия, волнами пространственного заряда или электронными волнами.

В общем случае всякое возмущение в электронном луче приводит к появлению парных волн, скорость которых зависит от скорости движения электронов и граничных условий за пределами луча. Первая волна пространственного заряда, фазовая скорость  $V_\phi$  которой несколько меньше средней скорости электронов называется медленной волной. Вторая волна,

скорость которой больше  $V_0$  называется быстрой волной пространственного заряда.

Как видно из (1.46), постоянная распространения  $\Gamma$  является мнимой величиной, поэтому нарастания или затухания волн пространственного заряда не происходит; эффект усиления или генерирования колебаний отсутствует. Однако при существовании электронных потоков, двигающихся с разными скоростями и взаимодействующих с волнами, может происходить нарастание амплитуды. На этом принципе основаны усилительные лампы, которые будут рассмотрены далее.

### 1.7. Фазовые и амплитудные условия самовозбуждения

Автоколебания в электродинамических системах возникают и поддерживаются при выполнении фазовых и амплитудных условий самовозбуждения.

Усиление электромагнитных колебаний возможно в том случае, когда между входом и выходом усилителя создана активная среда, обеспечивающая превращение энергии, подводимой извне, в энергию электромагнитных колебаний, совпадающих по частоте с входным сигналом. Например, входной сигнал создаёт в усилителе электромагнитное поле, взаимодействующее с электронным потоком. В режиме усиления результатом взаимодействия является преобразование части потенциальной или кинетической энергии электронов в энергию электромагнитных колебаний. Иными словами, в результате взаимодействия вырабатывается СВЧ - мощность  $P_э$ , называемая **электронной мощностью**, которую электронный поток передаёт СВЧ – полю. Баланс активных мощностей усилителя выражается равенством:

$$P_{вх} + P_э = P_{кон} + P_{вых}, \quad (1.50)$$

где  $P_{вх}$  - мощность входного сигнала,  $P_э$  - электронная мощность,

$P_{кон}$  - мощность потерь в колебательной системе (контурная мощность),

$P_{вых}$  - мощность выходного сигнала.

Если мощность, потребляемую от источника питания, обозначить за  $P_0$ , то КПД усилителя определяется соотношением:

$$\eta = \frac{P_{вых} - P_{вх}}{P_0} \quad (1.51)$$

Электронный КПД по определению равен:

$$\eta_э = \frac{P_э}{P_0}, \quad (1.52)$$

а контурный КПД определяется отношением выходной мощности к сумме выходной и контурной мощностей:



$$\eta_{\text{кон}} = \frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}} + P_{\text{кон}}}, \quad (1.53)$$

или согласно (1.50):

$$\eta_{\text{кон}} = \frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}} + P_{\text{э}}}, \quad (1.54)$$

В автоколебательном режиме  $P_{\text{ВХ}} = 0$  и полный КПД генератора равен произведению электронного и контурного КПД:

$$\eta = \eta_{\text{э}} \cdot \eta_{\text{кон}}. \quad (1.55)$$

Для генератора с резонансной колебательной системой и одним элементом связи с внешней нагрузкой (при двух полосной схеме включения в линию передачи):

$$\eta_{\text{кон}} = \frac{Q_{\text{Н}}}{Q_{\text{ВН}}} = \frac{Q_0}{Q_0 + Q_{\text{ВН}}}, \quad (1.56)$$

где  $Q_0$  - собственная добротность резонаторной системы;

$Q_{\text{ВН}}$  - внешняя добротность резонаторной системы;

$Q_{\text{Н}}$  - нагруженная добротность резонаторной системы.

### Положительные и отрицательные связи.

Рассматривая процесс усиления, необходимо иметь в виду, что между выходом и входом усилителя существуют многократные отражения сигнала, что обеспечивает его циркуляцию по замкнутому контуру внутренней обратной связи. Кроме внутренней обратной связи, которая в большинстве случаев является нежелательной, иногда создают внешнюю регулируемую обратную связь. Если колебания монохроматические, то обратная связь может быть положительной или отрицательной. В первом случае отражённый сигнал, поступающий на вход усилителя по линии обратной связи, совпадает по фазе с входным сигналом, что увеличивает коэффициент усиления. При отрицательной обратной связи отражённый сигнал поступает на вход усилителя в противофазе с входным сигналом, коэффициент усиления снижается, однако при этом повышается устойчивость работы усилителя, т.е. он становится менее склонным к самовозбуждению.

Возвращаясь к условиям самовозбуждения генератора, отметим, что фазовые условия самовозбуждения, определяемые положительной обратной связью можно представить в виде:

$$\Sigma\varphi = 2\pi n \text{ при } n = 0, 1, 2, 3, \dots, \quad (1.57)$$

где  $\Sigma\varphi$  - суммарный фазовый сдвиг, при обходе контура внешней обратной связи;  $n=0,1,2,3,\dots$  - целое число.

Выполнение фазового условия самовозбуждения не является достаточным для поддержания колебаний. При наличии положительной обратной

связи необходимо, чтобы удовлетворялись также амплитудные условия самовозбуждения, т.е. чтобы мощность, поступающая с выхода на вход по линии обратной связи, превышала минимальный сигнал, при котором возможно нарастание электромагнитных колебаний.

Амплитудные условия самовозбуждения следуют из баланса активных мощностей  $\sum P_a = 0$ , а фазовые условия связаны с балансом реактивных мощностей  $\sum P_p = 0$ . Рассмотрение того и другого балансов приводит к уравнениям:

$$-P_{\text{э}} = P_{\text{кон}} + P_{\text{вых}} = \frac{\omega_0 \cdot W}{Q_H}, \quad (1.58)$$

$$P_{\text{э,р}} = 2\Delta\omega W, \quad (1.59)$$

отличается от резонансной;

$P_{\text{э,р}}$  - реактивная мощность электронного потока;

$P_{\text{э}}$  - активная мощность электронного потока;

$W$  - энергия, запасаемая в резонаторной системе.

$$-\frac{P_{\text{э,р}}}{P_{\text{э}}} = \frac{2 \cdot Q_H \cdot \Delta f}{f_0},$$

здесь  $\Delta f = f_{\text{ген}} - f_0$ .

При возбуждении в резонаторе ВЧ поля  $E(t)$  электронными сгустками электронные мощности определяются интегралами по объёму  $V$  резонатора:

$$P_{\text{э}} = -\int_V \overline{J_{\text{конв}}(t) \bar{E}(t)} d\bar{V}, \quad \text{и} \quad (1.60)$$

$$P_{\text{э,р}} = -\int_V \overline{J_{\text{конв}}(t) \frac{d\bar{E}(t)}{d(\omega t)}} d\bar{V}, \quad (1.60,а)$$

где  $J_{\text{конв}} = f(t)$  - переменная составляющая плотности конвекционного тока. Горизонтальная черта над подынтегральным выражением означает усреднение во времени.

Возбуждение нерезонансных колебательных систем происходит при длительном взаимодействии электронных сгустков с полем бегущей волны в согласованной линии передачи. Электроны могут взаимодействовать как с «быстрыми» волнами, распространяющимися с фазовой скоростью, равной или превышающей скорость света, так и с «медленными» волнами, фазовая скорость которых значительно меньше скорости света в вакууме и близка к средней скорости электронов. «Быстрые» волны возбуждаются в приборах с линиями передачи волноводного типа, где используют криволинейные электронные потоки, взаимодействующие с поперечными компонентами высокочастотных электрических полей. «Медленные» волны распространяются в ЗС, представляющих собой периодические структуры, поля в которых оказываются суммой пространственных гармоник, состоящих из прямых и обратных волн, т.е. волн, у которых фазовые скорости или совпадают по направлению с групповыми скоростями. или направлены в

противоположные стороны. Электроны взаимодействуют с продольными компонентами электрического поля прямых или обратных волн.

Поток мощности, переносимый бегущей волной ЗС можно выразить в виде:

$$P = \pm \frac{E_z \cdot E_z^*}{(2 \cdot \beta^2 R_{CB})} = \pm \frac{E_{Zm}^2}{2 \cdot \beta^2 R_{CB}}, \quad (1.61)$$

где  $E_z = \dot{E}_{zm} \cdot e^{j\omega t}$ ,  $E_z^* = \dot{E}_{zm}^* \cdot e^{-j\omega t}$ ,  $\dot{E}_{zm}$  - комплексная амплитуда продольной напряженности электрического поля взаимодействующей волны.

Знак плюс относится к мощности, передаваемой в положительном направлении оси z, т.е. прямой волной, а минус характеризует передачу мощности в отрицательном направлении, т.е. обратной волной.

С параметрами, определяющими поток мощности, мы познакомимся более детально при рассмотрении приборов СВЧ.

## Глава 2.

### 2. Электровакуумные приборы СВЧ

К электровакуумным приборам СВЧ относятся различные по схеме приборы: клистроны, ЛБВ, ЛОВ, магнетроны и другие. Однако детальное рассмотрение показывает, что принцип взаимодействия электронного потока с электромагнитным полем у всех приборов во многом аналогичен – взаимодействие сводится к трем сравнительно простым процессам: модуляции по скорости; модуляции по плотности; переходу энергии электронных сгустков в СВЧ энергию. Только в одних приборах эти процессы совмещены в пространстве и во времени, а в других протекают раздельно. Наиболее типичным прибором с ярко выраженным разделением указанных процессов в пространстве, и во времени является. рдвухрезонаторный клистон.

#### 2.1. Пролетные усилительные клистроны

Клистроны – это приборы, в которых происходит отбор кинетической энергии у электронного потока тормозящим СВЧ – полем при кратковременном их взаимодействии. На рис. 2.1 показана схема устройства двухрезонаторного (усилительного) пролетного клистрона. На ней введены обозначения 1 – катод; 2 – ускоряющий электрод; 3 – входной резонатор; 4 – трубка дрейфа; 5 – выходной резонатор; 6 – коллектор;  $t_1$  – момент времени, определяющий пролет электронов через центр первого зазора;  $t_2$  – время пролета электронов через второй зазор.

При прохождении электронного потока через первый резонатор происходит модуляция его по скорости модулирующим ВЧ напряжением  $U_{m1}$ , подаваемым во входной резонатор. Будем полагать, что  $U_{m1} \ll U_0$ . Обозначим через  $t_1$  момент времени прохождения электрона через центр первого зазора. Мгновенное значение модулирующего напряжения на первом зазоре равно  $U = U_{m1} \sin \omega t_1$ . Скорость электронов на выходе первого зазора можно вычислить по формуле (1.40), где переменная составляющая скорости определена выражением:

$$V_1 = (\gamma V_0 U_{m1}) / 2U_0. \quad (2.1)$$

Глубина модуляции определяется отношением переменного и постоянного напряжений. Для возможности управления ею, вводится коэффициент модуляции электронного потока (М)

$$M = (\gamma U_{m1}) / 2U_0. \quad (2.2)$$

Как видно из формулы (2.2), коэффициент М пропорционален  $\gamma$ , амплитуде модулирующего напряжения  $U_{m1}$  и обратно пропорционален ускоряющему напряжению  $U_0$ .

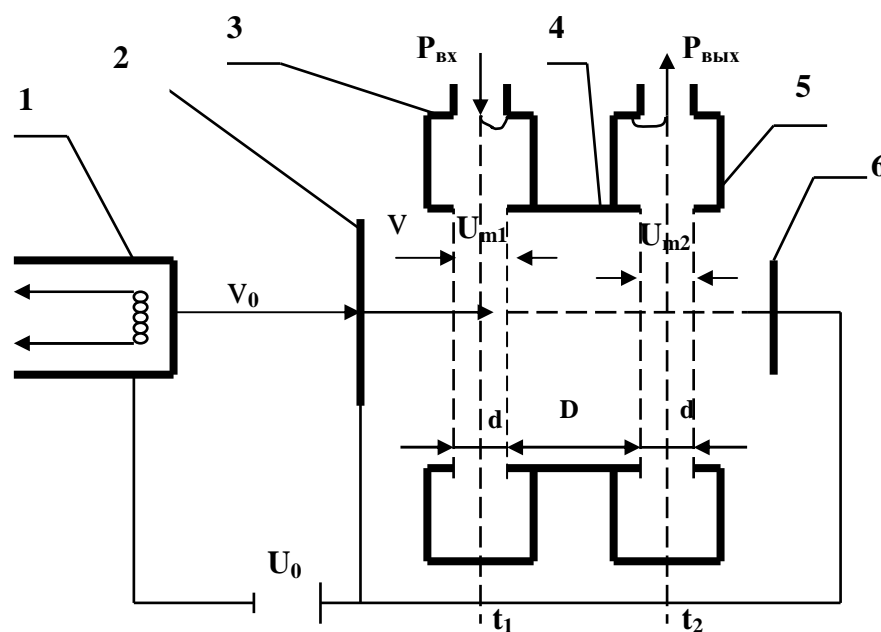


Рис. 2.1 Двухрезонаторный пролетный клистрон

Скорость электронов на выходе из входного резонатора выражается следующим соотношением:

$$V = V_0 \left( 1 + \gamma \frac{U_{m1}}{2U_0} \sin \omega t \right) = V_0 (1 + M \sin \omega t), \quad (2.3)$$

где  $\gamma$  – коэффициент эффективного взаимодействия электронного пучка с полем зазора (1.18), который принято выбирать в пределах 0,6-0,9.

Электроны, прошедшие первый резонатор оказываются **промодулированными по скорости**. Далее пролетая в пространстве до второго резонатора (пространство дрейфа), электроны модулируются по плотности.

Пренебрегая действием пространственного заряда, можно считать движение электронного потока в пространстве дрейфа чисто инерционным. Положим, что электрон пролетает через середину сеточного промежутка в момент времени  $t_1$ , а через середину сеточного промежутка второго резонатора в момент времени  $t_2$ . Тогда  $t_2$  будет определено следующим образом

$$t_2 = t_1 + \frac{D}{V_0 + V_1 \sin \omega t_1}, \quad (2.4)$$

где  $D$  – расстояние между резонаторами.

Преобразуем (2.4). Примем  $V_1 \ll V_0$ . Вынося во втором множителе  $V_1 / V_0$  и раскладывая в ряд множитель по малому параметру, получаем:

$$t_2 \approx t_1 + \frac{D}{V_0} - \frac{DV_1}{V_0^2} \sin \omega t_1. \quad (2.5)$$

Умножим обе части полученного уравнения на  $\omega$  тогда, опуская в дальнейшем знак приближенного равенства, запишем фазу прибытия сгруппированного потока во второй резонатор в виде следующего выражения:

$$\omega t_2 - \omega D / V_0 = \omega t_1 - (\omega D V_1 / V_0^2) \sin(\omega t_1). \quad (2.6)$$

Величина  $v_1$  здесь соответствует произведению  $V_0$  и  $M$ . Иначе формулу (2.6) можно записать в виде

$$\omega t_2 - \theta_0 = \omega t_1 - X \sin(\omega t_1), \quad (2.7)$$

где  $\theta_0 = \omega D / V_0$  - угол пролета в пространстве дрейфа электронов, не изменивших своей скорости, множитель  $X$  носит название параметра группировки. Преобразуя последнее выражение, придем к следующему виду параметра группировки  $X$ :

$$X = (\omega D / v_0) \cdot (\gamma U_{m1} / 2 U_0) = \theta_0 M. \quad (2.8)$$

Сгруппированный электронный поток, поступающий во второй резонатор, наводит в нем конвекционный ток.

Из закона сохранения заряда следует, что

$$dq_1 = dq_2, \quad (2.9)$$

где  $dq_1$  и  $dq_2$  - количество заряда, прошедшее через 1-ый и 2-ой зазоры резонаторов в моменты времени  $t_1$  и  $t_2$  соответственно.

Определим мгновенные значения конвекционных токов  $i_1$  и  $i_2$ . Используя равенство (2.9), запишем  $i_1 dt_1 = i_2 dt_2$ . Откуда определим ток во втором резонаторе:

$$i_2 = i_1 \frac{dt_1}{dt_2} = \frac{I_0}{dt_2 / dt_1}. \quad (2.10)$$

Модуляция пучка электронов по плотности в центре первого зазора практически отсутствует. Поэтому ток  $i_1$  равен постоянному току  $I_0$ , поступающему в зазор из электронной пушки.

Из (2.5) следует, что

$$\frac{dt_2}{dt_1} = 1 - \frac{\omega D \cdot V_1}{V_0^2} \cos \omega t_1 = 1 - X \cos \omega t_1. \quad (2.11)$$

Подставляя (2.11) в (2.9) запишем мгновенное значение конвекционного тока во втором резонаторе, равным:

$$i_2 = \frac{I_0}{1 - X \cos \omega t} \quad (2.12)$$

На рис. 2.2 показана зависимость конвекционного тока от времени, определенная по формуле (2.12) для трех значений параметра группирования  $X$ .  
 $X=0$   $i_2=I_0$ . Если  $X \ll 1$ , то

$$X \cdot \cos \omega t_1 \ll 1$$

и по формуле (2.12)

$$i_2(t_2) \approx I_0(1 + X \cos \omega t_1).$$

Так как связь  $t_2$  и  $t_1$  при  $X < 1$  незначительная, то зависимость должна быть приближенно синусоидальной с частотой  $\omega$ , той же частоте напряжения на входе резонатора. С увеличением  $X$  резко проявляется несинусоидальный характер кривой, но периодичность остается прежней ( $T=2\pi/\omega$ ). При  $X=1$  появляются бесконечно большие импульсы тока, соответствующие группированию части потока электронов, прошедших первый резонатор в момент времени  $t_1=0$ . Этому случаю соответствует значение момента прихода в сечение, соответствующее  $t_2$ , когда производная  $dt_2/dt_1$  равна нулю, в формуле (2.12)  $(1 - X \cos \omega t_1) = 0$ ,  $i_2$  становится бесконечно большим. При  $X > 1$  в пределах периода появляются два бесконечно больших импульса, так как производная  $dt_2/dt_1$  равна нулю в двух моментах времени.

Конвекционный ток в клистроне резко несинусоидальный, поэтому кроме первой гармоники (с частотой  $\omega$ , равной частоте входного сигнала) он должен содержать много других гармонических составляющих.

Функция (2.12) разложенная в ряд Фурье, имеет вид:

$$i_2 = I_0 + 2I_0 \sum_{n=1}^{\infty} J_n(nX) \cos n(\omega t_2 - \theta_0), \quad (2.13)$$

где  $n$  – номер гармонической составляющей, а  $J_n(nX)$  – функция Бесселя первого рода  $n$ -го порядка от аргумента  $nX$ .

Амплитудное значение гармоник с номером  $n$

$$I_{(n)} = 2J_n(nX)I_0 \quad (2.14)$$

Чаще всего интересуются только первой гармоникой тока. Максимум первой гармоники будет при максимуме функции Бесселя  $J_1(X)$ , а именно при  $X = 1,84$ . Сгруппированный электронный поток достигает наибольшего уплотнения в плоскости, где  $X=1$ , но в плоскости  $X=1,84$  достигает наибольшей величины первая гармоника электронного тока частоты  $\omega$ .

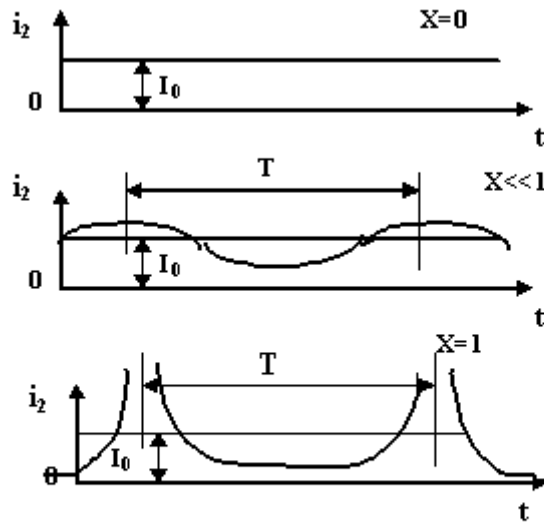


Рис. 2.2. Зависимость конвекционного тока от времени

ви-  
от  
фор-  
па-  
При

од-  
 $i_2(t_2)$   
си-  
рав-  
пер-  
ем  $X$   
со-  
ка,

Если выходной резонатор настроен на частоту модулирующего напряжения  $\omega$ , то заметную мощность в нем будут иметь только колебания этой частоты, и напряжение между его сетками будет практически синусоидальным. Следовательно, в формуле (2.13) можно учитывать только первую гармонику электронного тока, амплитуда которой  $I_1 = 2I_0 J_1(X)$ .

**Наведенный ток.** Рассмотрим, как происходит отбор энергии от модулированного по плотности электронного потока. Пусть сгруппированный электронный поток проходит в пространстве между сетками второго резонатора. Определим величину наведенного тока, появляющегося в этом резонаторе.

Вследствие группирования электронов конвекционный ток содержит гармонические составляющие, определяемые рядом (2.13). Поэтому и в наведенном токе должны быть те же гармоники.

При расчете наведенного тока пределы интегрирования в (1.17а) должны определяться координатами сеток выходного резонатора  $D-d_2/2$  и  $D+d_2/2$ . Запишем определение для наведенного тока первой гармоники

$$i_{\text{нав}(1)} = \frac{1}{d_2} \int_{D-d_2/2}^{D+d_2/2} I \cos(\omega t - \theta_0) dz \quad (2.15)$$

После интегрирования получим:

$$i_{\text{нав}(1)} = I_{\text{нав}(1)} \cos(\omega t - \theta_0), \quad (2.16)$$

где  $I_{\text{нав}(1)} = \gamma_2 I_{(1)}$ ,  $\gamma_2 = \sin \frac{\theta_2}{2} / \frac{\theta_2}{2}$ ,  $\theta_2 = \omega d / V_0$ ,  $\theta_0 = \omega D / V_0$ .

Здесь  $\theta_0$  – угол пролета невозмущенного электрона между серединами входного и выходного резонаторов, а угол  $\theta_2$  – угол пролета этого электрона между сетками выходного резонатора.

Величины  $\gamma_2$  по аналогии с (1.18) называют коэффициентом эффективного взаимодействия электронного потока с полем выходного резонатора. Все ранее сделанные замечания относительно зависимости коэффициента  $\gamma_1$  входного резонатора от угла пролета  $\theta$  справедливы и здесь, то есть зависимость  $\gamma_2$  от  $\theta_2$  такая же, как  $\gamma_1$  от  $\theta$  (см. рис. 1.7). Так как всегда  $\theta_2 \neq 0$ , то  $\gamma_2 < 1$  и амплитуда первой гармоники конвекционного тока  $I_{\text{нав}(1)} < I_{(1)}$ . Аналогично можно получить выражения для любой гармоники наведенного тока с номером  $n$ :

$$I_{\text{нав}(n)} = -\gamma_2 I_{(n)} = -2\gamma_2 J_n(nX) I_0 \quad (2.17)$$

Амплитуда наведенного тока  $I_1$  численно равна амплитуде конвекционного тока, умноженной на коэффициент взаимодействия пучка с зазором на частоте данной гармоники. Каждая гармоника наведенного тока, протекая через внешнюю цепь зазора, создает между сетками резонатора переменное напряжение  $U_2 \sin \omega t_2$ .

**Эквивалентные схемы резонаторов пролетного клистрона.** На рис. 2.3 приведены эквивалентные схемы выходного и входного резонаторов пролетного

клизотрона соответственно. На рис. 2.3,а изображен выходной резонатор, где  $Y_{p2}$  – собственная проводимость резонатора.

Проводимость резонатора  $Y_{p2}$

$$Y_{p2} = G_{p2} + jB_{p2} \quad (2.18)$$

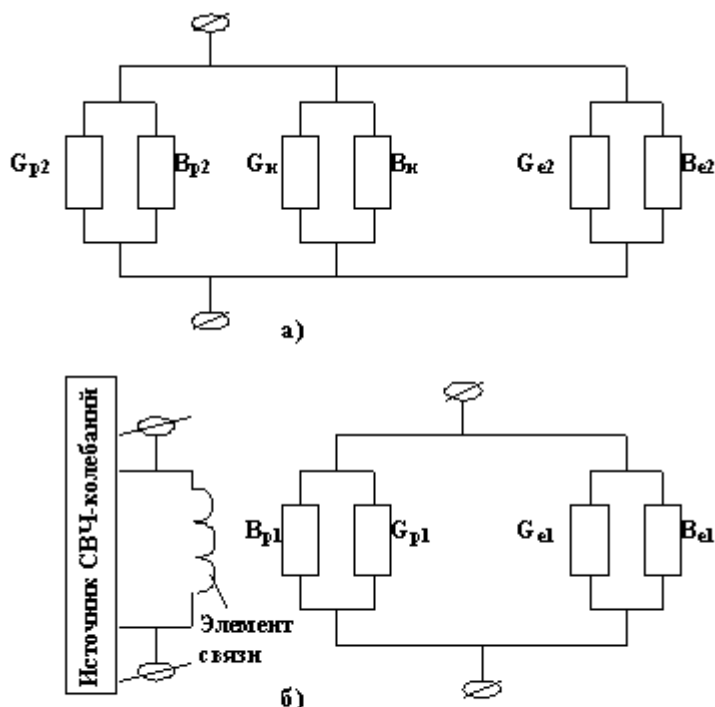


Рис. 2.3. Эквивалентные схемы а) выходного резонатора, б) входного резонатора

состоит из активной проводимости  $G_{p2}$ , характеризующей потери в резонаторе  $G_{p2}=G_{2п}$  и реактивной составляющей  $B_{p2}$ , определяемой емкостью и индуктивностью резонатора. Включена также проводимость нагрузки

$$Y_H = G_H + jB_H, \quad (2.19)$$

пересчитанная ко входу резонатора и электронная проводимость

$$Y_{e2} = G_{e2} + jB_{e2}, \quad (2.20)$$

обусловленная прохождением через зазор резонатора электронного потока.

Эквивалентная схема входного резонатора отличается тем, что у нее отсутствует проводимость нагрузки. Элемент связи первого резонатора с входной линией конструируется таким образом, чтобы обеспечить поступление максимальной мощности в резонатор от источника колебаний, то есть обеспечить согласование резонатора с входной линией.

**Параметры и характеристики двухрезонаторного пролетного клизотрона.** Наиболее важными параметрами двухрезонаторного клизотрона является коэффициент усиления по напряжению и по мощности, а также ширина полосы пропускания.

Коэффициент усиления по мощности определяется отношением:



$$K_p = 10 \lg \frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}}}, \quad (2.21)$$

где  $P_{\text{ВЫХ}}$  – выходная мощность, равная

$$P_{\text{ВЫХ}} = 2I_0^2 \gamma^2 \{J_1(X)\}^2 \frac{G_H}{(G_{23})^2}, \quad (2.22)$$

где  $I_0$  – постоянный ток пучка;

$G_{23} = G_{2n} + G_H + G_{2e}$  – эквивалентная проводимость выходного резонатора при резонансе, равная сумме трех проводимостей: проводимости потерь в резонаторе  $G_{2n}$ , эквивалентной проводимости нагрузки резонатора  $G_H$  и электронной проводимости  $G_{2e}$ , обусловленной прохождением электронного потока через зазор резонатора.

Мощность входного сигнала  $P_{\text{ВХ}}$ , поступающая на вход клистрона, расходуется на потери в самом резонаторе и на модуляцию электронного потока по скорости и при выполнении условия согласования определяется соотношением:

$$P_{\text{ВХ}} = \frac{1}{2} U_{m1}^2 (G_{1n} + G_{1e}) = \frac{1}{2} U_{m1}^2 G_{13}, \quad (2.23)$$

$G_{13} = G_{1n} + G_{1e}$  – эквивалентная проводимость входного резонатора при резонансе, равная сумме двух проводимостей: проводимости потерь в резонаторе  $G_{1n}$  и электронной проводимости.

Для удобства расчетов  $U_{m1}$  можно выразить через параметр группировки:

$$U_{m1} = \frac{2U_0}{\gamma \theta} X, \quad \text{тогда} \quad P_{\text{ВХ}} = \frac{2U_0^2 G_{13} X^2}{\gamma^2 \theta^2}.$$

Подставляя (2.22) и (2.23) в (2.21), можно выразить коэффициент усиления в децибелах. Максимум коэффициента усиления в зависимости от полной проводимости нагрузки достигается при условии сопряженного согласования нагрузки с выходным резонатором клистрона.

$$B_{2p} + B_H = 0; \quad G_H = G_{2p} + G_{2e}$$

Усилительные двухрезонаторные клистроны имеют мощность на выходе от нескольких десятков ватт до единиц киловатт в непрерывном режиме и примерно на порядок больше в импульсном. КПД - 25÷30%, полоса генерации составляет 0,1% от резонансной частоты, коэффициент усиления 30дБ.

### Частотно-умножительные клистроны

Представление конвекционного тока в клистроны в виде суммы большого числа гармоник, позволяет добиться эффективного умножения частоты. Так настраивая выходной резонатор на  $n$  – ую гармонику, на выходе получаем частоту, увеличенную в  $n$  раз по сравнению с частотой входного резонатора. Режим работы частотно - умножительного клистрона должен соответствовать оптимальному значению параметра группировки тоже для  $n$  – ой гармоники:

$$X_{onm}^{(1)} = 1,84; \quad X_{onm}^{(2)} = 1,53; \quad X_{onm}^{(3)} = 1,4; \quad X_{onm}^{(8)} = 1,22; \quad X_{onm}^{(16)} = 1,13.$$

Двухрезонаторные клистронные умножители обычно используются с коэффициентом умножения от 3 до 10.

Таким образом, частотно-умножительный клистрон отличается от усилительного соотношением размеров своих резонаторов и параметром группировки, требующим иных соотношений между напряжениями и углом пролета. Выходные параметры частотно-умножительного клистрона: КПД и выходная мощность, которые на порядок меньше, чем в усилительном. Однако количественно эти параметры уменьшаются с ростом номера гармоники сравнительно медленно.

### ***Многорезонаторные клистроны.***

В двухрезонаторном клистроне теоретически можно получить достаточно большой коэффициент усиления. Для этого достаточно увеличить расстояние между резонаторами, чтобы приблизить второй резонатор к оптимальному сечению, которое при малой входной мощности оказывается далеко за его пределами. Однако практически коэффициент усиления в двухрезонаторном клистроне не удается получить более 10-15 дБ. Это объясняется тем, что при больших расстояниях между резонаторами нельзя пренебрегать кулоновскими расталкивающими силами, возникающими между электронами сгустка, которые размывают сгусток и уменьшают его плотность. Последнее приводит к тому, что размеры сгустка возрастают и ему требуется значительно большее время для прохождения второго резонатора даже при бесконечно малой длине зазора. Поэтому лишь небольшая часть сгустка проходит зазор при максимальном значении тормозящего поля  $U_2$ , а остальная часть при уменьшенном значении  $U_2$ , когда торможение слабое, неэффективное. В результате сгусток отдает лишь небольшую часть своей кинетической энергии, и коэффициент усиления клистрона оказывается значительно меньше того значения, которое было рассчитано без учета кулоновских сил.

Указанное противоречие устраняется при использовании трехрезонаторного, или в общем случае многорезонаторного клистрона. Вводимые в пространство дрейфа один или несколько дополнительных ненагруженных резонаторов усиливают процесс модуляции по скорости и плотности, что позволяет при малом времени пролета и слабом входном сигнале получить достаточно плотные сгустки.

Трехрезонаторный клистрон работает следующим образом. Электронный поток, промодулированный по скорости, начинает модулироваться в пространстве дрейфа по плотности. Однако этот процесс ввиду малости входного сигнала протекает неинтенсивно и оптимальное сечение образуется справа от второго резонатора и поэтому в нем наводится очень слабый ток. Однако этот резонатор не нагружен, обладает большой добротностью и поэтому даже при слабом наведенном токе на нем возникает относительно большое напряжение, во много раз превышающее напряжение на зазоре входного резонатора. Указанное напряжение и используется для повторной модуляции электронного потока, причем коэффициент модуляции по скорости получается значительно большим.

Таким образом, промежуточный (один или несколько) резонатор используется для усиления процесса модуляции по скорости, что существенно ускоряет про-

цесс образования сгустков, а также повышается КУ и КПД клистрона. Ориентировочно максимальный коэффициент усиления (в дБ) для  $N$  резонаторного клистрона равен:

$$K_y = 15 + 20 (N-2) \text{ дБ.} \quad (2.24)$$

Получить коэффициент усиления в многорезонаторном клистроне более 60 дБ трудно из-за паразитных обратных связей и самовозбуждения колебаний.

Теоретически максимальный электронный КПД многорезонаторного клистрона достигает 73,8 % вместо 58,2 % для двухрезонаторного клистрона, а полный КПД – порядка 40 %. Аналогично и полоса усиливаемых частот многорезонаторного клистрона шире полосы усиления двухрезонаторного клистрона и достигает величины порядка 10%.

Основное применение многорезонаторных клистронов – в качестве мощных усилителей и умножителей частоты. Их мощность в непрерывном режиме составляет десятки и даже сотни киловатт, а в импульсном режиме – десятки мегаватт.

## 2.2 Отражательные клистроны

Клистрон, в котором электроны, пролетев сквозь зазор резонатора, возвращаются в него тормозящим полем отражателя, называется отражательным. Он представляет собой автогенератор СВЧ малой мощности, широко применяемый в качестве маломощных генераторов.

Отражательные клистроны используются, главным образом в тех случаях, когда требуется получение СВЧ-колебаний высокой стабильности.

На рис. 2.4 показано устройство и схема питания отражательного клистрона.

**Принцип работы.** Электроны, вылетающие из катода, ускоряются напряжением  $U_0$ , приложенным к резонатору, а после прохождения резонатора попадают в тормозящее поле участка резонатор – отражатель, где потенциал поля линейно уменьшается до  $-U_{отр}$ . В точке, где потенциал поля равен нулю, скорость электронов также падает до нуля. Здесь они начинают ускоренное движение в обратном направлении и снова пролетают через сетки резонатора по направлению к катоду. В конечном счете, электроны оседают на поверхности резонатора, обладающей положительным напряжением.

Модуляция по скорости и плотности. Процессы модуляции в отражательном клистроне, которые рассматривались в разделе 1.4, во многом аналогичны тем же

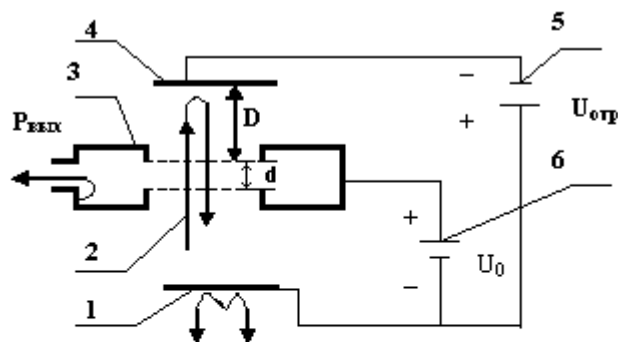


Рис. 2.4. Отражательный клистрон

- 1 – катод; 2 – электронный пучок; 3 – резонатор;  
4 – отражатель; 5 – источник напряжения отражателя;  
6 – источник ускоряющего напряжения клистрона.

процессам в двухрезонаторном клистроне. На рис. 1.35 иллюстрировался процесс образования сгустков и взаимодействие последних с СВЧ полем. Электрон 5 пролетает резонатор в тот момент, когда СВЧ напряжение равно нулю, поэтому закон движения этого электрона остается тем же, что и в статическом случае. Наоборот, электрон 3, пролетевший зазор на  $T/4$  раньше (в момент максимального ускоряющего напряжения), вылетает из резонатора с несколько большей кинетической скоростью, чем электрон 5, поэтому удаляется от резонатора несколько дальше, затрачивая на весь путь до возвращения в резонатор соответственно большее время. При определенных условиях моменты возвращения электронов 3 и 5 могут совпасть, как показано на рисунке 1.36.

Аналогично электрон 7, пролетевший через зазор на  $T/4$  позднее электрона 5 (в момент тормозящего поля), вылетает из резонатора с меньшей скоростью и поэтому затрачивает на прямое и обратное движение меньшее время. Момент возвращения электрона 7 также может совпасть с моментом возвращения электронов 3 и 5. При этом образуется сгусток, который пройдет через зазор от отражателя к катоду.

**Зоны генерации.** Если сгусток пройдет через резонатор в момент тормозящего для него поля (так как электрон - отрицательно заряженная частица, то при смене направления движения электрона на обратное ускоряющее поле становится для него тормозящим), он потеряет часть своей кинетической энергии и усилит имевшиеся в резонаторе СВЧ колебания. Если же сгусток пройдет через зазор в момент ускоряющего для него поля, он, наоборот, поглотит часть СВЧ энергии.

Для обеспечения эффективного торможения угол пролета должен составлять

$$\theta = \omega \tau_0 = 2\pi (n + 0,75), \quad n = 0, 1, 2, \dots,$$

где  $\tau_0$  – оптимальная величина времени пролета электронов центра сгустка в пространстве группировки по отношению к центру высокочастотного зазора резонатора.

$$\tau_0 = (n + 0,75)T. \quad (2.25)$$

С учетом геометрии прибора и подаваемых напряжений оптимальный угол пролета при условии  $d \ll D$  равен:

$$\theta = \omega \tau_{\text{пр}} = \frac{f 2\pi D \sqrt{U_0 \frac{8m}{e}}}{U_0 - U_{\text{отр}}}, \quad (2.26)$$

где  $f$  – частота генерируемых колебаний;  $d$  и  $D$  – размеры, показаны на рис. 2.4.

Следовательно, генерация в клистроне возможна не при любом времени пролета  $\tau$  электронов от зазора в сторону отражателя и обратно, а лишь при ряде дискретных значений  $\tau = \tau_0$ . Если время пролета не соответствует этому значению  $|\tau \neq \tau_0|$ , то центр сгустка будет проходить зазор не в момент максимального тормозящего для него поля, а раньше или позже. Поэтому сгусток будет либо отда-

вать СВЧ полю меньшую часть своей кинетической энергии, либо даже поглощать СВЧ энергию. В последнем случае самовозбуждение клистрона вообще не возможно.

Но время пролета электронов, как следует из (2.25) зависит от внутренних размеров клистрона и его напряжений. В частности, из физических соображений ясно, что чем меньше при заданном анодном напряжении напряженность тормозящего поля  $E = (|U_0| + U_{отр})/D$ , тем дальше будут удаляться электроны от резонатора в сторону отражателя и, следовательно, тем большее будет их время пролета  $\tau$ . Поэтому, меняя  $U_{отр}$  или  $D$ , можно в широких пределах изменять время пролета электронов и тем самым осуществлять в соответствии с формулой (2.25) генерацию в нужной зоне  $n$ .

Пример зависимости мощности радиации от напряжения  $U_{отр}$  приведен рис. 2.5.

Напряжения  $U_{отр}$ , соответствующие центрам зон, удовлетворяют ловию (2.25). И, наоборот, между значениями  $U_{отр}$  соответствуют случаю, когда сгустки проходят зазор, поглощая энергию  $\tau \approx \tau_0 + T/2$ , и поэтому самовозбуждение клистрона невозможно.

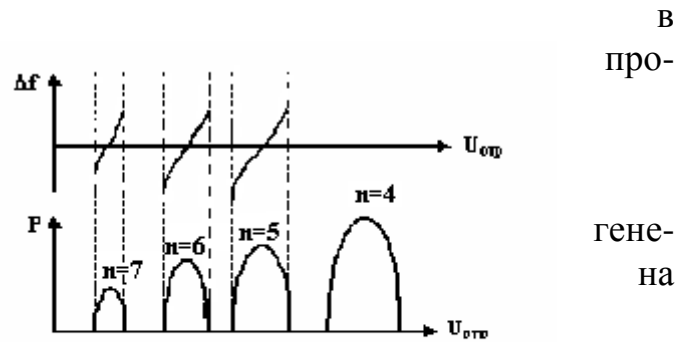


Рис. 2.5. Зависимости крутизны и мощности генерации от напряжения на отражателе

в про-  
гена  
на  
ус-  
нами  
ко-  
энер-

**Модуляция по плотности, конвекционный ток в отражательном клистроне.** Как и в двухрезонаторном клистроне, сгустки в отражательном клистроне можно представить в виде периодической последовательности импульсов конвекционного тока. Закон сохранения заряда позволяет записать мгновенный конвекционный ток  $i_2$ , поступающий в центр высокочастотного зазора после движения в пространстве группировки в виде (2.10)

$$i_2 = i_1 \frac{dt_1}{dt_2} = \frac{I_0}{dt_2/dt_1},$$

где  $t_1$  и  $t_2$  моменты первого и второго прохождений одного и того же электрона через центр зазора.

Суммарное время пролета электрона от центра зазора к отражателю и обратно:

$$\tau = t_2 - t_1 = \frac{d}{v} + \frac{2m}{eE} v. \quad (2.27)$$

Подставляя сюда значения  $v$  и  $E$ , и, умножая на круговую частоту колебаний  $\omega$ , имеем:

$$\omega t_2 - (\theta_2 + \theta_1) = \omega t_1 + X \sin \omega t_1, \quad (2.28)$$

где  $\theta_2$  – угол пролета в пространстве группировки и равен:

$$\theta_2 = \frac{2\pi}{e} \cdot \frac{\omega D V_0}{U_0 - U_{\text{отр}}}; \quad (2.29)$$

$\theta_1$  – угол пролета в зазоре резонатора, он определяется как:

$$\theta_1 = \frac{\omega d}{V_0}. \quad (2.30)$$

Параметр группировки в отражательном клистроне можно записать в виде соотношения:

$$X = V_1 \left( \frac{2m}{e} \cdot \frac{\omega D}{U_0 - U_{\text{отр}}} - \frac{\omega d}{V_0} \right) = \frac{V_1}{V_0} (\theta_2 - \theta_1) = \frac{\gamma U_1}{2U_0} (\theta_2 - \theta_1). \quad (2.31)$$

Отметим, что уравнение (2.31) совершенно аналогично по форме полученному выше уравнению (2.8). Однако вместо угла пролета в пространстве дрейфа здесь встречается разность углов пролета  $\theta_1$  и  $\theta_2$ . Имеется также различие в знаках в правых частях уравнений (2.7) и (2.28).

Но если положить в исходном соотношении (2.1) вместо  $\omega t_1$  величину  $\omega t_1 \pm \pi$ , то уравнения отражательного и двухрезонаторного клистронов становятся тождественными. С физической точки зрения это означает, что отсчет фазы следует производить относительно центра сгустка, как это делалось выше. С изменением начала отсчета фазы уравнение (2.28) дает:

$$\omega t_2 - (\theta_2 + \theta_1) = \omega t_1 - X \sin \omega t_1 \quad (2.32)$$

Дальнейший ход расчета конвекционного тока не отличается от анализа, проведенного для пролетных клистронов. Конвекционный ток  $i_2$ , поступающий в резонатор при возвращении электронов из пространства группировки равен в точности соотношению (2.12). Последний в свою очередь можно представить в виде суммы постоянной составляющей и бесконечного ряда гармоник, амплитуды которых определяются разложением в ряд Фурье. Следует учесть, однако, что высшие гармоники наведенного тока не могут быть использованы при обычной конструкции отражательного клистрона. Действительно, поскольку для модуляции и отбора энергии используется один и тот же резонатор, внешняя цепь зазора играет роль практически короткого замыкания для всех гармоник тока, кроме первой. Следовательно, достаточно ограничиться лишь рассмотрением первой гармоники конвекционного тока, мгновенное значение которой равно

$$(i_2)_1 = 2I_0 J_1(X_1) \cos[\omega t_2 - (\theta_2 + \theta_1)]. \quad (2.33)$$

Составляющая тока основной частоты, наводимого в резонаторе сгруппированным электронным потоком,

$$\dot{i}_H = -2\gamma I_0 J_1(X_1) \cos[\omega t_2 - (\theta_2 + \theta_1)], \quad (2.34)$$

$$u = -U_1 \sin \omega t_2. \quad (2.35)$$

Отрицательный знак в (2.34) учитывает изменение направления движения электронов после пребывания в пространстве группировки. Происхождение отрицательного знака в уравнении (2.35) обсуждалось выше в связи с изменением начала отсчета фаз.

Величина максимальной мощности равна:

$$P_{\max} = I_0 U_0 (2X_1 J_1(X_1)) / (\theta_2 - \theta_1) = I_0 U_0 X_1 J_1(X_1) / (\pi(n + 3/4) - \theta_1) \quad (2.36)$$

Хотя отражательный клистрон может генерировать колебания одной и той же частоты в различных зонах, однако его мощность, как показано на рис. 2.5, в этих зонах не одинакова. Она тем меньше, чем больше номер  $n$  зоны. Причем зоны, соответствующие малым значениям  $n$  могут вообще не возбуждаться. Так на рис. 2.5 невозбуждающаяся зона  $n=0$  показана пунктирной кривой. При высокой добротности резонатора (при малой нагрузке) клистрон, в принципе, может генерировать в любой зоне. Мощность генерации, близкую к  $0,582P_0$ , можно получить в любой зоне, если при работе в этой зоне нагрузку резонатора подобрать так, чтобы оптимальное сечение в установившемся режиме находилось близко от зазора, а  $U_{m1} \approx U_0$ . При выборе номера зоны, в которой предполагается получить наибольшую мощность от клистрона, приходится искать компромиссные решения.

Максимальный электронный КПД (коэффициент полезного действия) отражательного клистрона определяется из соотношения:

$$\eta_{\max} = P_{\max} / P_0 = X_1 J_1(X_1) / (\pi(n + 3/4) - \theta_1). \quad (2.37)$$

В отражательных клистронах возможна электронная перестройка частоты, за счет изменения напряжения на отражателе. При этом угол пролета  $\theta$  перестает быть равным  $2\pi(n+3/4)$  или  $2\pi(n-3/4)$ . Тогда наводимый в резонаторе ток и переменное напряжение на резонаторе не совпадают по фазе и в резонатор будет вноситься добавочная емкостная или индуктивная проводимость. Это приведет к изменению частоты клистронного генератора:

$$f = f_0 \left\{ 1 - \frac{1}{Q_H} \operatorname{tg} \left[ \frac{2\pi(n+0,75)}{U_0 + |U_{\text{отр}}|} \Delta U_{\text{отр}} \right] \right\}, \quad (2.38)$$

где  $U_{\text{отр}}$  - изменение напряжения на отражателе относительно напряжения, соответствующего центру зоны генерации.

Мощности, получаемые в отражательных клистронах, значительно меньше, чем в пролетных, но зато несколько расширяется рабочая полоса, с возможностью перестройки до 60 МГц.

### 2.3. Лампы бегущей и обратной волны О - типа

Для осуществления длительного взаимодействия электромагнитного поля с

потоком заряженных частиц в электро-нике СВЧ применяются медленные волны. Такое взаимодействие возникает, если заряженные частицы и электромагнитная волна движутся с одинаковыми скоростями. Так как скорость частиц не может превышать скорость света, для реализации длительного взаимодействия необходимо замедлять электромагнитные волны. Для создания медленных волн разработаны специальные металлические и металлодиэлектрические линии передачи, называемые *замедляющими системами* (см. 1.3.3). В лампах бегущей волны электронные сгустки взаимодействуют с продольной  $E_z$  компонентой электромагнитного поля.

Для существования незатухающих медленных E-волн необходимо, чтобы на поверхности ЗС продольное поверхностное сопротивление было отлично от нуля и имело индуктивный характер.

### 2.3.1. Элементы линейной теории ЛБВ

Обычно ЛБВ используются для усиления слабого сигнала. В этом случае приближенный анализ взаимодействия электронного потока с полем бегущей волны возможен на основе линейной теории ЛБВ. При этом предполагается, что все переменные составляющие величин, характеризующих электронный поток, много меньше их постоянных составляющих.

Задачу взаимодействия электронного потока с полем бегущей волны рассматривают в три этапа [4]:

- группирование электронов под действием электрического поля бегущей волны в замедляющей системе;
- возбуждение сгруппированным электронным потоком волн в замедляющей системе;
- нахождение дисперсионного уравнения ЛБВ (совместное решение результатов двух первых этапов).

Дисперсионное уравнение записывается в виде:

$$\frac{jR_{св} I_0 \beta_e \Gamma^2 \Gamma_0}{2U_0 (\Gamma_0^2 - \Gamma^2) (j\beta_e - \Gamma)^2} = 1, \quad (2.39)$$

где  $\Gamma_0$  – постоянная распространения волн в замедляющей системе без электронного потока (холодная система);  $\Gamma$  – постоянная распространения волн в замедляющей системе с электронным потоком (горячая система),

$$\beta_e = \omega / v_e.$$

Выражение (2.39) является уравнением четвертой степени относительно  $\Gamma$ . Его решение в рассматриваемой системе определяет постоянные распространения ( $\Gamma^{(i)}$ ) для четырех волн, которые равны следующим выражениям:

$$\left. \begin{aligned} \Gamma^{(i)} &= \alpha^{(i)} + j\beta^{(i)}, \quad i = 1, 2, 3, 4; \\ \Gamma^{(1)} &= \alpha^{(1)} + j\beta^{(1)}, \quad \alpha^{(1)} = \frac{\sqrt{3}}{2}\beta_0 C, \quad \beta^{(1)} = \beta_0(1 + C/2); \\ \Gamma^{(2)} &= \alpha^{(2)} + j\beta^{(2)}, \quad \alpha^{(2)} = -\frac{\sqrt{3}}{2}\beta_0 C, \quad \beta^{(2)} = \beta_0(1 + C/2); \\ \Gamma^{(3)} &= \alpha^{(3)} + j\beta^{(3)}, \quad \alpha^{(3)} = 0, \quad \beta^{(3)} = \beta_0(1 - C); \\ \Gamma^{(4)} &= \alpha^{(4)} + j\beta^{(4)}, \quad \alpha^{(4)} = 0, \quad \beta^{(4)} = -\beta_0(1 - C^3/4); \end{aligned} \right\} \quad (2.40)$$



где  $C = \sqrt{I_0 R_{св}} / 4U_0$  - параметр усиления; (2.41)

$R_{св}$  – сопротивление связи (1.36);

$I_0$  – ток пучка;  $U_0$  – ускоряющее напряжение;  $\beta_0$  – коэффициент фазы или фазовая постоянная волны нулевой пространственной гармоники в ЗС

( $\phi_0 = \beta_0 \cdot L$  – сдвиг фазы на один период  $L$  для нулевой пространственной гармоники).

Из анализа (2.40), следует, что три первых волны распространяются вдоль оси  $z$ , так как  $\beta^{(1)}$ ,  $\beta^{(2)}$  и  $\beta^{(3)}$  положительны, а четвертая – в обратном направлении ( $\beta^{(4)} < 0$ ), т.е. навстречу электронам. Ее постоянная распространения равна постоянной распространения в замедляющей системе без учета влияния электронного потока. При равенстве  $C=0,02 \div 0,2$  значение  $C^3$  очень мало, значит и  $\beta^{(4)}$  мало. Следовательно, четвертую волну можно исключить.

Фазовые скорости первых трех парциальных волн различны и определяются по формулам:

$$\left. \begin{aligned} V_{фр}^1 &= \omega / \beta^{(1)} = V_0 / (1 + C/2); \\ V_{фр}^2 &= \omega / \beta^{(2)} = V_0 / (1 + C/2); \\ V_{фр}^3 &= \omega / \beta^{(3)} = V_0 / (1 - C); \end{aligned} \right\} \quad (2.42)$$

Считая  $C \ll 1$  и подставляя значения  $\beta^{(i)}$ , в (2.42), сделаем вывод, что фазовые скорости первой и второй волн немного меньше скорости электронного потока ( $V_{ф} < V_0$ ), а скорость третьей волны немного больше.

Продольная составляющая поля любой парциальной волны с номером  $i$  с учетом (2.40) равна:

$$\begin{aligned} E_{Zm} \exp(j\omega t - \Gamma^{(i)}z) &= E_{Zm} \exp(-\alpha^{(i)} \cdot z) \exp(j\omega t - \beta^{(i)} \cdot z) = \\ &= E_{Zm}^{(i)} \exp(j\omega t - \beta^{(i)} \cdot z). \end{aligned} \quad (2.43)$$

При распространении этих волн вдоль оси  $z$  амплитуда первой волны ( $i=1$ ) убывает ( $\alpha^{(1)} > 0$ ), а второй ( $i=2$ ) возрастает ( $\alpha^{(2)} < 0$ ). Амплитуды третьей волны остаются неизменной ( $\alpha^{(3)} = 0$ ). Следовательно, только вторая парциальная волна характеризует процесс усиления в ЛБВ.

### 2.3.2 Лампы бегущей волны типа «О»

**Принцип работы и устройство лампы бегущей волны.** В приборах типа О происходит преобразование кинетической энергии электронов в энергию СВЧ – поля при длительном взаимодействии электронного потока с СВЧ – полем. Длительное взаимодействие позволяет получить необходимое группирование электронов при слабом входном сигнале. Обмен энергией между электронами и полем происходит в результате взаимодействия электронов с составляющей напряженности поля, совпадающей по направлению со скоростью электронов.

Рассмотрим устройство лампы бегущей волны на схеме, изображенной на рис. 1.37, введя элементную базу рис. 2.6. Электронная пушка образована катодом 1, управляющим электродом 2. Эта система электродов обеспечивает необходи-

мую начальную фокусировку пучка и регулировку его тока. Далее электронный поток попадает в замедляющую систему типа спирали 3, в конце которой находится выходное устройство, и коллектор 4, на который оседает электронный поток. На поверхность держателей спирали наносят слой поглотителя 5 для предотвращения самовозбуждения ЛБВО.

Взаимодействуя с СВЧ-полем, электронный поток, попадая либо в ускоряющую, либо в тормозящую фазу поля, меняет скорость. Происходит группировка электронов по скорости и по плотности одновременно. Вопрос о модуляции электронного пучка был рассмотрен в разделе 1.4. Как показано на рис. 1.38 электроны группируются вокруг электрона типа 1, начавшего движение в нулевом поле, соответствующем переходу от ускоряющей к тормозящей полуволне. Причем, если скорость электрона равна фазовой скорости волны, то электроны, группирующиеся слева относительно электрона 1, будут отбирать от СВЧ – поля столько же энергии, сколько будут отдавать энергии электроны, находящиеся в тормозящем поле. Поэтому нужно добиваться неравенства этих скоростей, а именно небольшого превышения скорости электронов по сравнению с фазовой скоростью волны. В этом случае группирование электронов происходит в тормозящем поле волны. Энергия, отдаваемая основной частью потока электронов полю волны превышает энергию, отбираемую от поля остальными электронами, и возможно усиление СВЧ – поля.

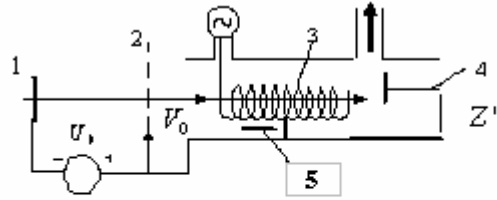


Рис. 2.6. Схема лампы бегущей волны

1-катод, 2- ускоряющий электрод  
3-замедляющая система, 4-коллектор,  
5-поглотитель.

Энергия усиливаемых СВЧ-колебаний, поступающих на вход лампы, распространяется по виткам замедляющей системы (ЗС) со скоростью света  $c = V_{гр}$  (1.33). Фазовая скорость  $V_{ф}$  (1.31) этих колебаний направлена вдоль оси лампы (спирали) и будет меньше групповой  $V_{гр}$  в  $\pi D/L$  раз, где  $D$  - диаметр спирали,  $L$  - шаг спирали. Эффективная передача энергии от электронов полю бегущей волны, как было показано выше, происходит при условии  $V_{ф} \leq V_0$ .

Обычно разность  $V_0$  и  $V_{ф}$  составляет не более 5-10 %. Для характеристики замедляющих систем используется коэффициент замедления, определяемый для спирали, из условия

$$K_3 = \frac{c}{V_{ф}} \approx \frac{\pi D}{L}. \quad (2.44)$$

Коэффициент замедления волны, распространяющейся по спирали, в первом приближении равен отношению длины витка к шагу спирали. Обычно  $K_3$  берут равным 10-20.

Фазовые скорости пространственных гармоник определяются отношением (1.31). Скорость нулевой гармоники ( $t=0$ ) запишем через замедленную длину волны и через напряжение источника питания ( $U_0$ ):

$$\lambda_{зам} = \lambda \frac{V_{ф}}{c} = \lambda \frac{\sqrt{U_0}}{505}.$$

Тогда фазовую скорость можно выразить следующим образом:

$$V_{\phi(0)} = \frac{c\sqrt{U_0}}{505}. \quad (2.45)$$

Коэффициент усиления один из важнейших параметров ЛБВ. Он может быть определен как отношение амплитуды выходного сигнала на конце замедляющей системы, длиной  $l$  ( $z = l$ ) к амплитуде входного сигнала в начале линии:

$$K_y(U) = \frac{E_{zm}(l)}{E_{zm}(0)}$$

Выражая величину  $K_y(U)$  в децибелах, коэффициент усиления записывают в виде:

$$K_y(U) = -9,54 + 47,3N \cdot C - L_{\pi}, \quad (2.46)$$

где  $L_{\pi}$  - потери, вносимые поглотителем от ( $L_{\pi} = 4 \div 15$  дБ);

$N$  - электрическая длина замедляющей системы, определяемая отношением:

$N = \frac{l}{\lambda_{зам}}$ .  $N$  принимается равной от 10 до 30, параметр усиления  $C$  выбирается в

пределах 0,02–0,2. Сопротивление связи  $R_{св}$  (1.36) связано с квадратом продольной составляющей напряженности поля  $E_{zm}$  и величиной потока мощности  $P$  в системе. С геометрией замедляющей системы  $R_{св}$  связано следующим образом:

$$R_{св} = \frac{c}{V_{\phi}} 13,3 \frac{\lambda_{зам}}{D\pi}. \quad (2.47)$$

В действительности СВЧ сигнал, поданный извне на вход ЛБВ, возбуждает не только нарастающую волну, но и две другие волны. Эти волны, не участвуя в усилении сигнала, создают потери на входе лампы. На большом расстоянии от входа сигнала поле представлено в основном нарастающей волной. Начальные потери сигнала составляют  $20 \lg 1/3 = -9,54$  дБ.

Реальная величина коэффициента усиления ЛБВ оказывается сравнительно высокой: порядка 20÷40 дБ. Особенно ценным свойством ламп прямой бегущей волны является их широкополосность. Коэффициент усиления ЛБВ при неизменном ускоряющем напряжении может оставаться почти неизменным в пределах 20÷50% и более от средней частоты.

Рабочая полоса частот ЛБВ частично определяется дисперсией замедляющей системы, то есть изменением фазовой скорости замедленной волны в зависимости от частоты. При фиксированном ускоряющем напряжении  $U_0$  скорость электронов остается неизменной. Следовательно, чем слабее дисперсия замедляющей системы, тем шире диапазон частот, в пределах которого может удовлетворяться условие синхронизма.

Максимальная величина электронного КПД ЛБВ определяется из уравнения:

$$\eta_{э} = (2 - 3)C \quad (2.48)$$

Коэффициент шума в относительных единицах

$$K_{ш} \approx 1 + \frac{1}{2} \left( 1 - \frac{\pi}{4} \right) \frac{T_k}{T_0} \cdot \frac{1}{C}, \quad (2.49)$$

где  $T_k$ - температура катода, °К;  
 $T_0$ - температура окружающей среды;  
 $C$ - параметр усиления.

В современных ЛБВ коэффициент шума в диапазоне 10 см составляет 3-4 дБ.

### 2.3.3 Лампы обратной волны типа «О»

Рассмотренная возможность самовозбуждения ЛБВ может быть использована для создания генератора СВЧ – лампы обратной волны типа О (ЛОВО).

В отличие от ЛБВ направление групповой скорости в ЛОВО противоположно направлению скорости электронов ( $V_0$ ). Направление фазовой скорости ( $V_{фм}$ ) и скорости электронов ( $V_0$ ) совпадает, поэтому при выполнении условия синхронизма для обратной гармоники ( $V_0 \geq V_{фм}$ ), можно обеспечить ее взаимодействие с электронным потоком. Наличие обратных пространственных гармоник является принципиально важным, так как приводит к появлению внутренней, положительной обратной связи. Обеспечить эффективную обратную связь удастся при использовании ЗС с отрицательной (аномальной) дисперсией (рис. 1.22,в).

Каждый элемент длины ЗС, на котором происходит взаимодействие поля замедляющей системы и электронного потока, можно рассматривать как распределенную обратную связь (ОС). Дисперсионные характеристики системы типа цепочка связанных резонаторов (см. рис. 1.28) приведены на рис. 1.29. Приборы, в которых реализуется принцип взаимодействия электронного пучка с волной, обладающих отрицательной дисперсией (обратной волной), называются лампами обратной волны. Такие лампы используются как генераторы колебаний СВЧ и обладают способностью к электронной перестройке в широком диапазоне частот. Широкодиапазонная перестройка частоты осуществляется путем изменения скорости пучка при выполнении амплитудного ( $V_0 \approx V_{фм}$ ) и фазового ( $\phi_{вх}, \phi_{вых} \ll 2\pi$ ) условий самовозбуждения.

#### Принцип действия ЛОВО

Схема генераторной ЛОВО показана на рис. 2.7. Электронный пучок, формируемый пушкой (1), движется со скоростью  $V_0$  вдоль ЗС (2), к коллектору (4). В любом электроном наблюдаются флуктуации токов и тока с любой той. Поэтому в замедляющей теме ЛОВ возбуждаются гармонические составляющие с быими частотами, которые пространяются в обоих на- лениях: в сторону поглоти- (3) выходного конца (Вы- Следствие периодичности структуры замедляющей

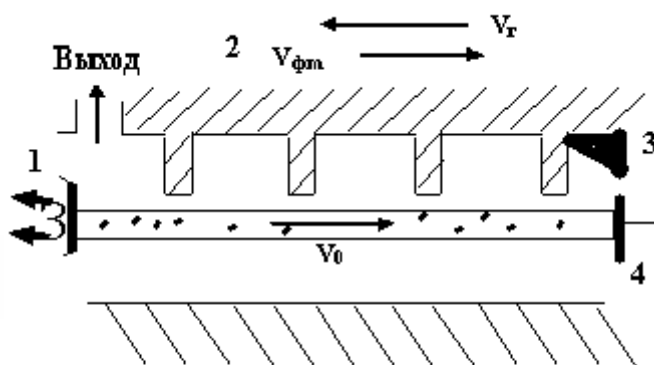


Рис. 2.7. Схема генераторной ЛОВО

1- электронная пушка, 2- замедляющая система,  
 3- поглотитель, 4 - коллектор.

токе  
 элек-  
 то-  
 сис-  
 гар-  
 лю-  
 рас-  
 прав-  
 теля  
 ход).

сис-

темы поле шумовых колебаний с любой частотой должно быть представлено бесконечной суммой пространственных прямых и обратных гармоник. Обратная пространственная гармоника, для которой выполняется условие синхронизма, вызывает модуляцию электронного потока по скорости и, следовательно, по плотности, что приводит к передаче энергии от электронного потока к полю, т.е. к росту всего поля в замедляющей системе. Если мощность, отдаваемая электронным потоком, превысит мощность потерь в системе, возникнет генерация колебаний.

Поглотитель в данной системе устраняет паразитную обратную связь, которая возникает при неполном согласовании. Введение поглотителя, хорошо согласованного с ЗС, в рабочем диапазоне частот, устраняет также возможность генерации колебаний в ЛОВ на прямых пространственных гармониках.

Различие в работе ЛОВ и ЛБВ наиболее отчетливо проявляется при сравнении продольного распределения высокочастотного потенциала волны в ЗС ( $V$ ) и переменной составляющей конвекционного тока пучка  $I$  в этих приборах (рис. 2.8). В ЛБВ обе величины возрастают вдоль пространства взаимодействия от пушки к коллектору приблизительно по экспоненциальному закону. Для ЛОВ величина  $M$  нарастает от коллектора к пушке, поскольку групповая скорость обратной волны переносит всю энергию, приобретенную от пучка справа налево.

Для ЛОВ величина  $M$  нарастает от коллектора к пушке, поскольку групповая скорость обратной волны переносит всю энергию, приобретенную от пучка справа налево.

Для самовозбуждения ЛОВ необходимо выполнение условий баланса амплитуд и баланса фаз:

$$\frac{\omega l_1}{V_{\phi(-1)}} - \frac{\omega l_2}{V_0} = (2m+1) \cdot \pi \quad m = 0, 1, 2, \dots, \quad (2.50)$$

где  $l_1$  и  $l_2$  - длины соответствующих участков петли обратной связи;  $m$ - номер гармоники ЛОВ или номер зоны колебаний (для  $m=0$  передается наибольшее значение энергии от электронного потока полю и наименьшее значение пускового тока;

$\frac{\omega l_1}{V_{1 \phi(-1)}}$  определяет сдвиг фазы, создаваемый волной обратной пространственной гармоники, а  $\frac{\omega l_2}{V_0}$  - электронным потоком.

Фазовая скорость обратной гармоники:

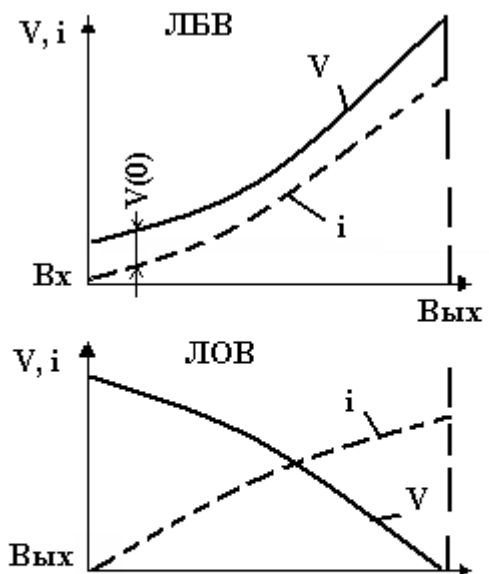


Рис. 2.8. Распределение амплитуды ВЧ напряжения и переменной составляющей тока пучка вдоль ЛБВ и ЛОВ

$$V_{\phi(-1)} = \frac{\omega}{\left(\beta_0 - \frac{2\pi}{L}\right)}. \quad (2.51)$$

Фазовые условия используют и для определения частоты генерируемых колебаний в различных зонах. Частота зависит от номера зоны  $m$ , а, в выбранной зоне, от  $V_0$ . Скорость электронов в приборах типа О связана с ускоряющим напряжением  $U_0$  обычным выражением (1.8):

Как видно, увеличение  $U_0$  сопровождается ростом частоты колебаний.

Ширина диапазона электронной перестройки частоты характеризуется коэффициентом перекрытия диапазона  $\delta_{\Pi} = f_{\max}/f_{\min}$ , где  $f_{\max}$  и  $f_{\min}$  - максимальная и минимальная граничные частоты диапазона ( $\delta_{\Pi} = 1,05 \div 1,5$ ).

Выходную мощность генераторной ЛОВ определяют по формуле:

$$P_{\text{вых}} = kU_0(I_0 - I_{\text{пуск}}),$$

где  $k$  – коэффициент, зависящий от параметра усиления (2.41) и электрической длины – число волн, укладываемых на длине ЗС, ( $N = l/\lambda_{\text{зам}}$ );  $I_0$  – ток пучка;  $I_{\text{пуск}}$  – пусковой ток, при котором начинается генерация.

В ЛОВ пусковой ток определяется из условия:

$$I_n \approx 0,124 \frac{U_0}{R_{\text{св}} N^3}, \quad N \approx 20.$$

При плохом согласовании концов ЗС на механизм распределенной обратной связи накладывается механизм дискретной обратной связи. Понятие дискретности можно связать с равенством скорости электронов пучка  $V_0$  и фазовой скорости волны  $V_{\phi}(\omega)$ , осуществляющейся в точности только на некоторой определенной частоте. На тех частотах, где обратная связь оказывается положительной, мощность генерации ЛОВО возрастает и пусковой ток понижается. Там же, где дискретная связь отрицательна, наблюдается снижение мощности и рост пускового тока. В результате этого зависимости  $P(U_0)$  и  $I_{\text{пуск}}(U_0)$  оказываются «изрезанными». Степень изрезанности всех этих характеристик возрастает с ухудшением согласования поглотителя и нагрузки, на которую работает ЛОВО.

Частота генерации в маломощных ЛОВО монотонно возрастает с увеличением  $U_0$ , что вполне естественно, так как обратные волны обладают аномальной дисперсией. Мощность генерации также возрастает с увеличением ускоряющего напряжения, поскольку с увеличением  $U_0$  возрастает мощность пучка. Наблюдаемая изрезанность этой кривой – результат неидеального согласования поглотителя с ЗС.

Выходные мощности ЛОВО находятся в пределах 100-1500 мВт при ускоряющих напряжениях  $U_0 = 180-1400$  В и токе в пучке электронов  $I_0 = 40-70$  мА. Электронный КПД  $\eta = 0,84 \cdot C$ .

Коэффициент полезного действия генераторных ЛОВ сравнительно невелик при значениях  $C = 0,02 - 0,2$  не более 20%.

Диапазон длин волн, в котором используются ЛОВ, составляет интервал от дециметров до долей миллиметров. Главным свойством ЛОВ, определяющим их практическую ценность, является возможность быстрой перестройки частоты за счет изменения ускоряющего напряжения.

## 2.4. Электронные приборы со скрещенными полями

### 2.4.1 Движение электронов в скрещенных полях

Рассмотрим задачу о движении электронов в скрещенных электрических и магнитных полях в отсутствии высокочастотных колебаний, подобно рассмотрению в [4]. Электрон движется между двумя параллельными плоскостями, где имеется электрическое поле  $E_0$  и магнитное поле  $B$ .

**Решение уравнения Лоренца.** Рассмотрим уравнение Лоренца (1.5), связывающее силы электрического и магнитного полей, действующих на заряд, с параметрами движущегося

ряда. Определим

ориентацию движения заряда. На рис. 2.9 показаны направления напряженно-постоянного электрического поля  $E$  и магнитной индукции  $B$  в выбранной системе координат.

Пусть в начальный момент времени  $t=0$ , электрон находится в начале координат  $x_0, y_0, z_0$  и обладает начальными скоростями  $v_{x0}, v_{y0}, v_{z0}$ .

Решим уравнение (1.5) для нашего случая, разложим векторы  $V, E, B$  по осям  $e_x, e_y, e_z$ :

$$[\nabla \mathbf{V}] = \begin{vmatrix} e_x & e_y & e_z \\ V_x & V_y & V_z \\ B_x & B_y & B_z \end{vmatrix} = e_x (V_y B_z + V_z B_y) + e_y (V_z B_x + V_x B_z) + e_z (V_x B_y + V_y B_x),$$

где  $v_x = \frac{dx}{dt}$ ,  $v_y = \frac{dy}{dt}$ ,  $v_z = \frac{dz}{dt}$ .

Таким образом, скалярное уравнение движения электрона в прямоугольной системе координат записывается в виде трех уравнений:

$$\frac{d^2 x}{dt^2} = -\frac{e}{m} \cdot (E_x + B_z \cdot \frac{dy}{dt} - B_y \cdot \frac{dz}{dt}), \quad (2.52)$$

$$\frac{d^2 y}{dt^2} = -\frac{e}{m} \cdot (E_y + B_x \cdot \frac{dz}{dt} - B_z \cdot \frac{dx}{dt}), \quad (2.53)$$

$$\frac{d^2 z}{dt^2} = -\frac{e}{m} \cdot (E_z + B_y \cdot \frac{dx}{dt} - B_x \cdot \frac{dy}{dt}). \quad (2.54)$$

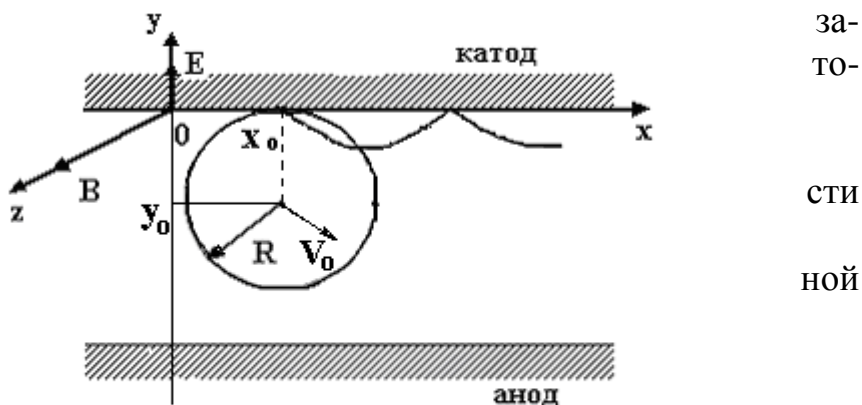


Рис. 2.9. Движение электрона в скрещенных электрическом и магнитном полях

Согласно рис. 2.9,  $E_x = E_z = 0$ ;  $E_y = +E$ ;  $B_x = B_y = 0$ ;  $B_z = +B$ .

Подставляя эти значения в (2.53), (2.54), (2.55) получим

$$\frac{d^2x}{dt^2} = \frac{-eB}{m} \cdot \frac{dy}{dt}, \quad (2.55)$$

$$\frac{d^2y}{dt^2} = -\frac{eE}{m} + \frac{eB}{m} \cdot \frac{dx}{dt}, \quad (2.56)$$

$$\frac{d^2z}{dt^2} = 0. \quad (2.57)$$

После интегрирования (2.55-2.57) уравнения движения электрона принимают вид:

$$x = x_0 + \frac{E}{B} \cdot t - \frac{m \cdot E}{e \cdot B^2} \cdot \left(1 - \frac{B}{E} \cdot V_{x0}\right) \cdot \sin \omega_{\text{ц}} \cdot t - \frac{V_{y0}}{\omega_{\text{ц}}} \cdot (1 - \cos \omega_{\text{ц}} \cdot t), \quad (2.58)$$

$$y - y_0 = -\frac{m \cdot E}{e \cdot B^2} \cdot (1 - \cos \omega_{\text{ц}} \cdot t) \left(1 - \frac{B}{E} V_{x0}\right) + \frac{V_{y0}}{\omega_{\text{ц}}} \sin \omega_{\text{ц}} t, \quad (2.59)$$

$$z = z_0 + V_{z0} t. \quad (2.60)$$

Здесь через  $\omega_{\text{ц}}$  – обозначена круговая циклотронная частота

$$\omega_{\text{ц}} = \frac{eB}{m}. \quad (2.61)$$

Уравнения (2.58), (2.59) и (2.60) дают полный ответ на вопрос о движении одиночного электрона, находившегося в момент  $t=0$  в точке с координатами  $(x_0, y_0, z_0)$  и имеющего составляющие начальной скорости  $v_{x0}, v_{y0}, v_{z0}$ , как показано на рис. 2.9.

В случае магнетронного генератора, у которого источником электронов является катод, лежащий в плоскости  $y=0$ , положим  $y_0=0$ , введем обозначения скорости переносного движения ( $V_{\text{ц}}$ ) и радиуса катящегося круга ( $R$ ):

$$V_{\text{ц}} = \frac{E}{B}, \quad R = \frac{E}{\omega_{\text{ц}} \cdot B} = \frac{mE}{eB^2}. \quad (2.62)$$

Тогда (2.58), (2.59) и (2.60) будут представлены в каноническом виде:

$$x = x_0 + v_{\text{ц}} \cdot t - R \sin(\omega_{\text{ц}} t), \quad y = -R(1 - \cos(\omega_{\text{ц}} t)), \quad z = z_0. \quad (2.63)$$

При  $x_0=0, V_{x0}=0, V_{y0}=0$  возведем выражения (2.63) в квадрат и сложим, то получим

$$(x - R)^2 + (y - V_{\text{ц}} \cdot t)^2 = R^2, \quad (2.64)$$

то есть (2.64) описывает окружность радиуса  $R$ . Центр окружности перемещается вдоль оси  $x$  со скоростью  $V_{\text{ц}}$ . Электрон, как бы помещенный на обод катящегося колеса ( $x_0=0$ ), движется по циклоидальной траектории.

Если начальная скорость электрона равна переносной скорости  $V_0=V_{\text{ц}}$ , и при  $1 - \frac{B}{E} V_{x0} = 0$ ,  $V_{x0}=V_{y0}=0$  выражения (2.58), (2.59) и (2.60) будут иметь вид:



$$y = y_0, \quad x = x_0 + \frac{E}{B}t, \quad z = z_0. \quad (2.65)$$

Траектория электрона вырождается в прямолинейную траекторию (случай ЛБВ М), параллельную поверхности электродов. Прямолинейное и равномерное движение электрона – это результат равенства электрической и магнитной сил. Происходит взаимная компенсация противоположно направленных сил, и электрон движется по инерции равномерно.

#### 2.4.2 Многорезонаторные магнетроны

Магнетрон является резонансным прибором М – типа, в котором замкнутый электронный поток взаимодействует с СВЧ полем замкнутой замедляющей системы. Типичный многорезонаторный магнетрон представляет собой устройство (рис. 2.10,а), в центре которого вдоль оси расположен цилиндрический катод с подогревателем, окруженный многорезонаторной системой, выполненной в медном анодном блоке. Магнитная индукция  $B$  направлена вдоль оси магнетрона. Анодное напряжение  $U_a$  между анодом и катодом создает электрическое поле, перпендикулярное магнитному.

Возьмем для простоты цилиндрические и плоские магнетроны со сплошным анодом (рис. 2.10,а и б). Пространственного заряда, создаваемого движущимися электронами учитывать не будем. Далее, электроны вылетают из катода с очень малыми начальными скоростями, поэтому величиной начальной скорости в статическом режиме магнетрона можно сразу пренебречь.

Уравнения (2.63), описывающие движение электрона под действием электрических и магнитных полей, применим к нахождению траектории электрона в магнетронах.

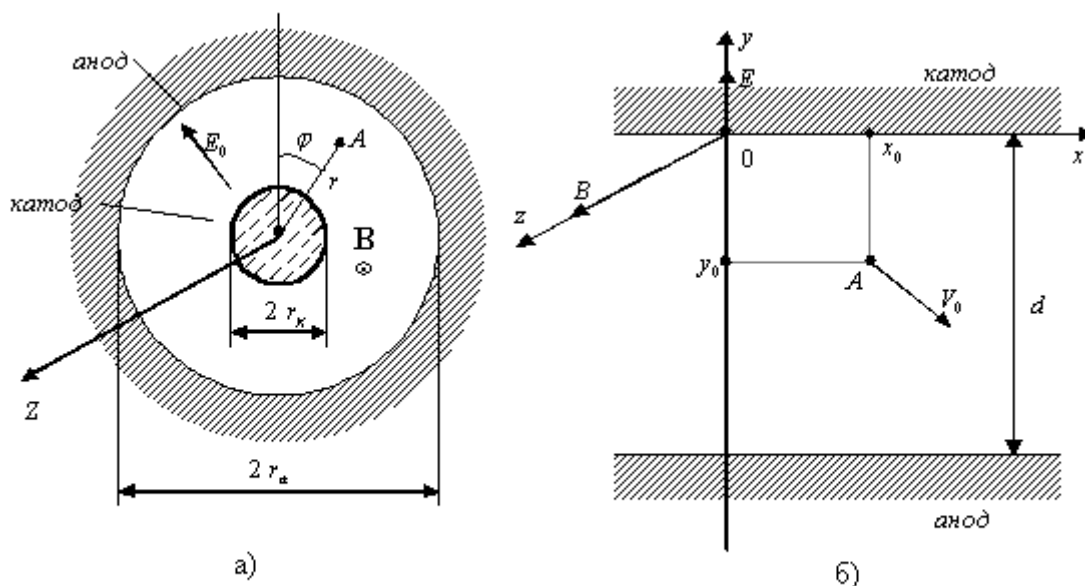


Рис. 2.10. Схема цилиндрического (а) и плоского магнетронов (б)

Как показали соотношения (2.63) в магнетроне электронный поток движется по циклоиде. На рис. 2.9 показана циклоидальная траектория движения электрона в плоском диоде в отсутствии ВЧ колебаний.

Если увеличивать индукцию магнитного поля  $B$ , радиус  $R$  согласно (2.62) будет постепенно уменьшаться. Электронные траектории, которые при  $B=0$  нормальны к поверхностям анода и катода, начинают искривляться. При  $d=2R$  электроны лишь касаются анода в вершине циклоиды. Наконец при  $d>2R$  электроны, не доходя до анода, совершают многократные колебания по циклоиде. Конвекционный ток на аноде становится равным нулю.

Критическое магнитное поле, при котором в статическом режиме прекращается анодный ток плоского диода определяется из условия:

$$B_{кр} = \frac{1}{d} \sqrt{\frac{2m}{e}} U_a \quad (2.66)$$

Можно рассуждать и так. При неизменной величине  $B$  существует критическое анодное напряжение  $U_{a,кр}$ , ниже которого ток через диод становится равным нулю. Используя соотношения для  $B_{кр}$ , построим параболу критического режима рис. 2.11.

**Устройство и принцип действия магнетрона.** Электроны эмитируются цилиндрическим катодом, коаксиальным анодному блоку. Статическое магнитное поле, создаваемое специальным магнитом, направлено вдоль оси прибора перпендикулярно радиальному статическому электрическому полю. Благодаря скрещенным электрическому и магнитному полям, эмитированные электроны образуют в пространстве взаимодействия облако, вращающееся с определенной скоростью вокруг катода.

В стенке анодного блока имеется ряд полостей – объемных резонаторов, которые образуют колебательную систему магнетрона. Вследствие флуктуаций электронов в колебательной системе возбуждаются слабые переменные электромагнитные поля. При выполнении условий синхронизма между электронным облаком и высокочастотным полем электроны группируются в сгустки, и их энергия передается электромагнитному полю резонаторов. Наличие сильной обратной связи, обусловленной свернутой в кольцо колебательной системой, приводит к установлению в магнетроне режима автоколебаний. Энергия этих колебаний выводится с помощью петли связи, расположенной в одном из резонаторов.

Многорезонаторные магнетроны используют в качестве мощных генераторов сантиметрового и миллиметрового диапазонов волн обычно в импульсном режиме.

**Свойства колебательной системы магнетрона.** Колебательные системы магнетронов образованы широкополосными периодическими замедляющимися системами, свернутыми в кольцо и замкнутыми на себя. В результате системы становятся резонансными и узкополосными, что резко увеличивает амплитуду ВЧ поля на резонансных частотах.

Рассмотрим замкнутую цепочку из  $N$  полых резонаторов как показано на рис. 2.12, расположенных на равных расстояниях, по внутренней поверхности анодного блока магнетрона.

Весь анодный блок представим как сложный кольцевой полый резонатор. Виды колебаний в этом резонаторе при полной идентичности отдельных резона-



Рис. 2.11. Парабола критического режима

торов достаточно характеризовать фазой колебаний. Обозначим через  $\varphi$  разность фаз колебаний в соседних резонаторах. Условие замкнутости цепочки резонаторов требует, чтобы при обходе вдоль всей окружности по внутренней поверхности анода полный сдвиг фазы был равен 0 или кратен  $2\pi$ :

$$\varphi \cdot N = 2\pi n, \quad n = 0, 1, 2, \dots$$

Отсюда вытекает, что разность фаз колебаний в резонаторах может принимать только дискретные значения, определяемые соотношением:

$$\varphi = \frac{2\pi n}{N}. \quad (2.67)$$

И можно сказать, что анодный блок магнетрона имеет столько видов колебаний сколько резонаторов  $N$ . При  $n=0$  колебания во всех резонаторах происходят синфазно. В случае  $n=N/2$  соседние резонаторы колеблются в противофазе, т.е. со сдвигом по фазе на  $\pi$ . Этот вид называется  $\pi$ -видом и является основным рабочим видом колебаний магнетронов. Кстати,  $\pi$ -вид возможен только при четном числе резонаторов  $N$ . Поэтому нечетное число резонаторов не находит применения в магнетронных генераторах.

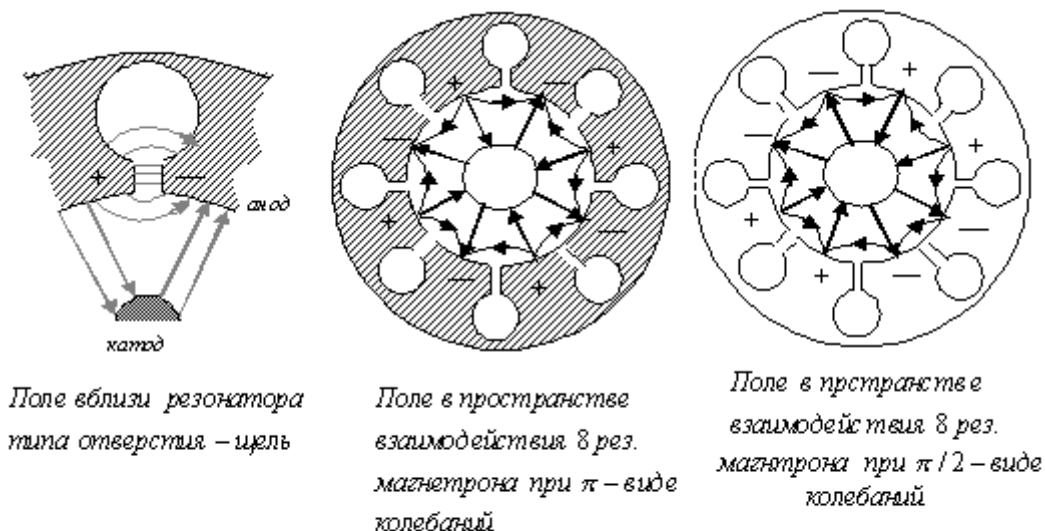


Рис. 2.12. Структура типов колебаний в анодном блоке магнетрона

Каждому из видов колебаний может соответствовать определенная структура поля (см. рис. 2.12), отличная от структуры других видов. Таким образом, следует сделать вывод, что существует  $N$  различных резонансных частот анодного блока, хотя все резонаторы в отдельности имеют одну и ту же резонансную частоту. Из  $N$  видов колебаний при полной симметрии анодного блока невырожденными являются колебания  $\pi$ -вида.

**Определение фазовых скоростей.** Предположим, что путь между двумя резонаторами волна проходит за время  $\Delta t$  ( $\Delta t = \frac{\varphi}{\omega}$ ), тогда фазовая скорость волн, бегущих в резонаторах в разных направлениях, будет определяется из условия

$$V_{\Phi} = \frac{2\pi r_a}{N} \frac{\omega}{\varphi} = \frac{2\pi r_a \omega N}{2\pi n N} = \frac{\omega r_a}{n}.$$

Так как колебательная система является периодической по азимуту, то поля бегущих волн любого типа колебаний могут быть разложены в ряд Фурье. Полученные пространственные гармоники имеют различную угловую скорость и одинаковую частоту. Если фазовый сдвиг  $\varphi_m$  на период структуры для  $m$ -ой пространственной гармоники составляет

$$\varphi_m = \varphi_0 + 2\pi m, \quad m = 0; \pm 1; \pm 2,$$

то фазовые скорости пространственных гармоник (ПГ) определяются по формуле

$$(V_{\Phi})_{nm} = \frac{\omega r_a}{n + mN}. \quad (2.68)$$

Чем выше номер  $n$  и  $m$ , тем меньше скорости прямых и обратных волн, вращающихся в пространстве взаимодействия. Волна, соответствующая случаю  $m=0$ , имеет при  $n=N/2$  наибольшую фазовую скорость и называется основной волной.

Величина фазовой скорости  $(V_{\Phi})_{nm}$  может быть сделана значительно меньше скорости света  $c$  в свободном пространстве.

$$K_3 = \frac{c}{(V_{\Phi})_{nm}} = \frac{\lambda f}{(V_{\Phi})_{nm}} = \frac{\lambda \omega (n + mN)}{2\pi \omega r_a} = \frac{\lambda (n + mN)}{2\pi r_a}, \quad (2.69)$$

$K_3$  - коэффициент замедления,  $\lambda$ -длина волны в свободном пространстве.

Среди основных волн при  $m=0$  наибольшее замедление имеют волны  $\pi$ -вида.

**Определение резонансных частот.** Следующим шагом в рассмотрении свойств “холодной” колебательной системы магнетрона является нахождение резонансных частот каждого из видов колебаний.

Рассмотрим эквивалентную схему (рис. 2.13) участка анода и катода равнорезонаторного магнетрона  $C_3$  – эквивалентная емкость одиночного резонатора;  $C'$  – сосредоточенная емкость между каждым сегментом анода и катодом.

Резонансная частота  $n$ -го вида колебаний  $\omega_n$  составляет

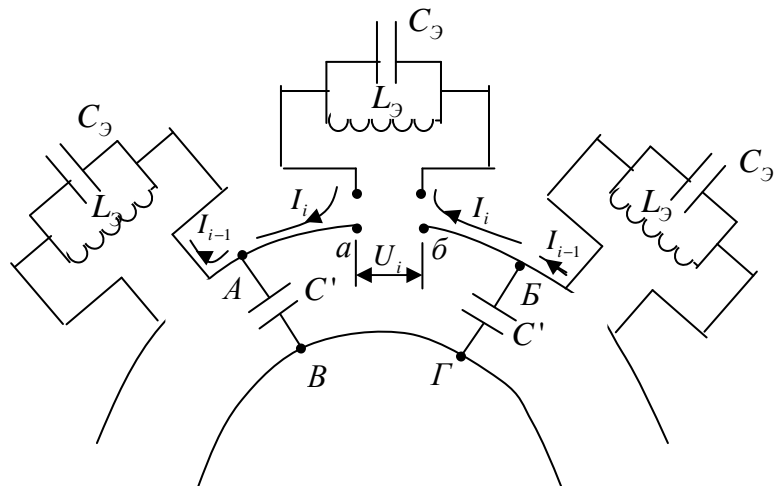


Рис. 2.13. Эквивалентная схема анодного блока магнетрона

$$\omega_n = \omega_0 / \sqrt{1 + \frac{C'}{C_3 \left(1 - \cos\left(\frac{2\pi n}{N}\right)\right)}}, \quad (2.70)$$

где  $\omega_0$  – резонансная частота одиночного резонатора.

Спектр видов колебаний таков, что резонансная частота  $\pi$ -вида  $\omega_\pi$  лишь незначительно отличается от частоты ближайшего вида, для которого  $n=(N/2)-1$ . Можно разность частот указанных видов колебаний выразить в процентах по отношению к резонансной частоте  $\pi$ -вида. Чем больше разделение частот, тем более стабильной будет работа магнетрона. В магнетроне при  $N=8$  при  $C'/C_3 = 0,5$  разделение частот небольшое – около 1%.

При увеличении числа резонаторов  $N$  разделение частот становится еще меньшим. Чем больше диаметр катода и поверхность сегмента анода, т.е. чем больше  $C'$ , тем сильнее разделены виды  $N/2$  и  $(N/2)-1$ . Таким образом, резонансная частота  $\pi$ -вида может лежать как выше, так и ниже частот других видов колебаний. В обоих случаях разность частот видов  $N/2$  и  $(N/2)-1$  оказывается очень небольшой и уменьшается при увеличении числа резонаторов.

Но для устойчивой работы магнетронного генератора желательно иметь разделение резонансных частот  $\pi$ -вида и ближайшего к нему вида  $(N/2)-1$  не менее 20%. С увеличением числа резонаторов разделение частот ухудшается. Существуют два метода увеличения разделения частот:

1. Применение связей в равнорезонаторном анодном блоке.
2. Использование разнорезонаторного анодного блока.

**Динамический режим магнетрона.** Рассмотрим теперь изменение движения электронов. Наиболее значительного взаимодействия электронов с полем можно ожидать лишь при условии, что электрон длительное время находится в одной и той же фазе по отношению к высокочастотному электрическому полю. При этом следует учесть, что в плоском магнетроне электрон совершает два движения: равномерное поступательное движение вдоль поверхности катода, характеризующее перемещением центра катящегося круга и вращательное движение, происходящее с циклотронной частотой. С другой стороны, высокочастотное поле в пространстве взаимодействия может быть представлено в виде бегущих волн. Поэтому условие постоянства фазы ВЧ поля относительно электрона можно считать как условие равенства фазовой скорости одной из волн и средней скорости  $V_{ц}$ , с которой перемещается центр катящегося круга - условие резонанса скоростей

$$V_{ц} = (V_{\phi})_{nm},$$

где  $(V_{\phi})_{nm}$  – фазовая скорость резонансной пространственной гармоники,  $n$ -го вида колебаний.

Действие других волн, если у них фазовые скорости сильно отличаются от средней скорости электронов, можно не учитывать.

Рассмотрим поведение типичных электронов, находящихся в разных участках пространства взаимодействия.

На рис. 2.14 изображено ВЧ электрическое поле резонаторов  $E$ , соответствующее фиксированному моменту времени и имеющее радиальную –  $E_r$  и

тангенциальную  $E_\tau$  составляющие, которые периодически изменяются по периметру анода. Кроме переменного  $E_\tau$  между анодом и катодом существует постоянное поле  $E_0$ . Тогда суммарное электрическое поле  $E_\Sigma$  в различных точках пространства взаимодействия имеет различную величину и направление. Действие же суммарного электрического поля на электрон сводится к изменению его скорости и направления движения. Рассмотрим результирующее электрическое поле  $E_\Sigma$ , действующее на электроны типов 1, 2, 3 и 4 (рис. 2.15).

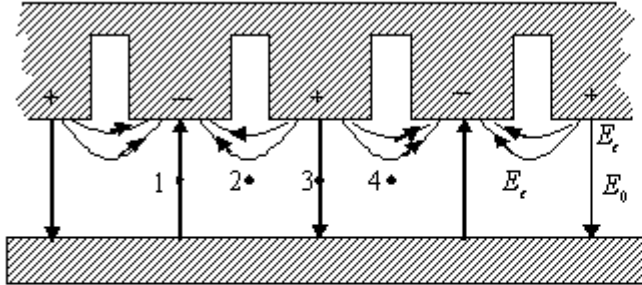


Рис. 2.14. Высокочастотное поле в резонаторном блоке магнетрона

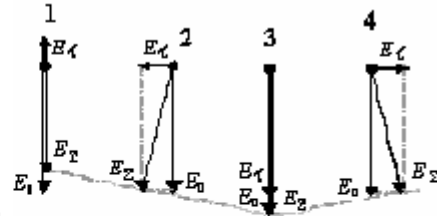


Рис. 2.15. Результирующее электрическое поле ( $E_\Sigma$ ), действующее на электроны.

Результирующее электрическое поле  $E_\Sigma$ , действующее на электрон типа 1, не изменяет своего направления, но уменьшается по абсолютной величине на величину  $E_r$ . Скорость центра катящегося круга, определяющего траекторию электрона относительно неподвижной системы координат равна теперь  $E_\Sigma/V$ . В результате электрон типа 1 начинает отставать от бегущей волны и постепенно смещается по направлению к электрону типа 4. Похожее изменение происходит и с электроном типа 3, для которого результирующее электрическое поле несколько увеличивается на величину  $E_r$  в сравнении со статическим режимом. Поэтому электрон типа 3 движется быстрее волны и постепенно приближается к электрону типа 4.  $V_4$  - поступательная скорость электрона типа 4 (см. рис. 2.16).

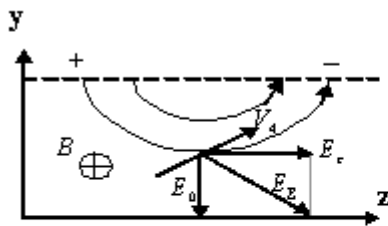


Рис. 2.16. Движение электрона к аноду

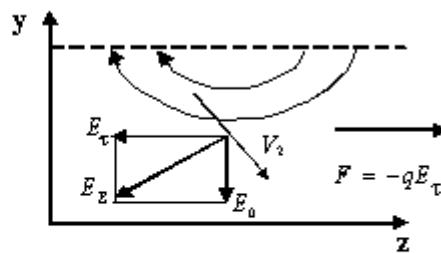


Рис. 2.17. Движение электрона к катоду

Электрон типа 4 движется в поле  $E_\Sigma$ , почти равным по абсолютной величине статическому электрическому полю, но имеющему некоторый наклон, направленный в сторону анода.

С физической точки зрения электрон типа 4 является “благоприятным” для поддержания автоколебаний в магнетроне. Электроны типа 1 и 3, попадая в тормозящее тангенциальное поле, также становятся “благоприятными” для генерации. Таким образом, электрон типа 4, а с ним и формирующейся в дальнейшем

сгусток, будет по мере взаимодействия с переменным полем, в процессе их совместного движения перемещаться по направлению к положительно заряженному электроду.

Рассмотрим теперь движение электрона типа 2 (см. рис. 2.17). Он после выхода из катода подвергается действию ускоряющего тангенциального поля.

Вектор поступательной скорости электронов отклонен к оси  $Z$ , т.е. к электроду с нулевым потенциалом. Следовательно, электроны, попавшие при вхождении в замедляющую систему в эту фазу поля; будут постепенно уходить на нижний электрод, т.е. в сторону катода. В конце первого циклоидального колебания электрон, поглотив часть энергии от поля, ударяется о катод и прекращает дальнейшее существование в пространстве взаимодействия. Таким образом, происходит пересортировка электронов на благоприятные и неблагоприятные, в результате которой образуются сгустки электронов относительно тех электронов, которые начали свое движение в максимуме тормозящего тангенциального поля. Модуляция электронного потока по плотности происходит автоматически, без участия какого-либо специально созданного внешнего управляющего устройства.

Основную роль в группировке (фазовой фокусировке) в магнетроне играет радиальная  $E_r$  составляющая. Под влиянием радиальной составляющей вектора напряженности переменного электрического поля электроны, движущиеся в скрещенных постоянных электрическом и магнитном полях, группируются в сгустки. Роль тангенциальной составляющей сводится к отбору энергии от электронов.

**Возникновение незатухающих колебаний в автогенераторе.** Механизм возникновения незатухающих колебаний в магнетроне такой же, как и в любом автогенераторе. Начальные колебания в резонаторах магнетрона возникают в результате флуктуаций электронного потока. Частота этих колебаний в общем случае несколько отличается от собственной резонансной частоты колебательных систем, так как анодный блок магнетрона образует систему сложно связанных контуров. Колебания поддерживаются за счет энергии источника постоянного напряжения анод-катод. Энергия электронного потока и взаимодействующего с переменным электрическим полем вблизи щелей резонаторов, передается полю волны. Для этого электронный поток должен быть сгруппирован в сгустки, время прохождения которых вблизи щели резонатора совпало бы со временем существования там поля в нужной фазе.

Потоки электронов к аноду создаются в некоторых областях пространства взаимодействия, образуя так называемые "электронные спицы" (рис. 2.18).

Число спиц зависит от характера высокочастотных колебаний. Причем каждому макси-

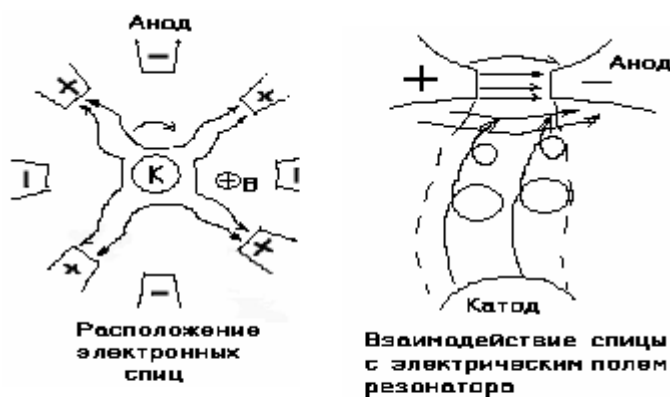


Рис. 2.18. Образование спиц в пространстве взаимодействия

муму тормозящего тангенциального поля в фиксированный момент времени соответствует одна спица. Следовательно, при виде колебаний  $n=1$  имеется всего одна спица, при  $n=2$  - две спицы и так далее. В случае  $\pi$ -вида колебаний в магнетроне существует наибольшее число спиц пространственного заряда.

Электроны в спицах перемещаются к аноду по сложным циклоидальным траекториям. Таким образом, характер их движения определяется суммарным воздействием постоянного и переменного магнитного поля. Спицы образуются вблизи участков катода, лежащих против тех сегментов анода, которые в данный момент сказываются благодаря наложению переменного электрического поля заряженными до положительного потенциала. Так как с изменением фазы колебаний меняются знаки заряда на сегментах анода, то изменяются и участки катода, вблизи которых формируются спицы. Спицы как бы вращаются в пространстве взаимодействия со скоростью

$$(V_{\phi})_{\text{сред}} = \frac{\omega_n (r_a + r_k)}{2(n + m \cdot N)}.$$

Направление фазовой скорости, которой совпадает с направлением вращения пространственного заряда. Эта волна и получает энергию электронного потока. Амплитуда колебаний в каждом из резонаторов должна быть одинаковой, но сдвиг фазы  $\Phi$  по-прежнему определяется уравнением.

Еще нужно отметить, что номер пространственной гармоники  $p$ , не оказывает влияния на число спиц. Таков в общих чертах принцип работы многорезонаторного магнетрона.

#### **Условия самовозбуждения колебаний типа бегущей волны.**

Условием отдачи электронами максимальной энергии ВЧ полю является совпадение фазовой скорости бегущей волны и средней скорости движения электронов в отсутствие колебаний, т.е.

$$V_{\text{ц}} = (V_{\phi})_{\text{пм.средн}}.$$

При этом всякое первоначальное колебание, возникшее в анодном блоке магнетрона должно нарастать до тех пор, пока не станут действовать ограничивающие нелинейные эффекты.

Учитывая, что  $E \approx \frac{U_a}{r_a - r_k}$ , можно записать условие для анодного напряжения ( $U_a$ ), при котором должно происходить самовозбуждение магнетрона для каждого вида колебаний при фиксированном номере гармоники  $m$

$$U_a = \frac{\omega_n (r_a^2 - r_k^2)}{2(n + mN)} \cdot B. \quad (2.71)$$

Прямые, определяемые уравнением (2.71) проходят через начало координат и пересекают параболу критического режима рис. 2.19.



При  $B = \text{const}$  и  $U_a > U_{\text{акр}}$  в заштрихованной области генерация колебаний типа бегущей волны невозможна: все электроны попадают на анод не позднее, чем через половину периода циклоидального колебания. Поэтому условиям самовозбуждения отвечают лишь участки прямых, выделенные жирными линиями, лежащих ниже параболы критического режима.

Таким образом, можно установить, что для каждого вида колебаний существуют оптимальные соотношения между постоянным анодным напряжением и индукцией магнитного поля.

### Пороговое анодное напряжение.

Одним из признаков возбуждения колебаний является протекание конвекционного электронного тока в анодной цепи магнетрона. Постоянное анодное напряжение ( $U_a$ ), при котором начинаются бесконечно малые СВЧ колебания, сопровождаемые попаданием электронов на анод, называется пороговым напряжением магнетрона ( $U_{\text{п}}$ ). Если проанализировать более подробно условия самовозбуждения магнетрона, то величина порогового напряжения оказывается связанной с магнитной индукцией

$$U_{\text{п}} = 1,01 \cdot 10^7 \left( \frac{r_a}{n\lambda} \right)^2 \left[ B \cdot 10^{-3} \left( 1 - \frac{r_k^2}{r_a^2} \right) \frac{n\lambda}{10,7} - 1 \right]. \quad (2.72)$$

График порогового напряжения можно объяснить так. При  $U_a < U_{\text{п}}$  самовозбуждение колебаний рассматриваемого вида невозможно, т.к. малое ВЧ напряжение, появившееся на сегментах анодного блока, например за счет шумов, не в состоянии сгруппировать и поднять электроны на анод магнетрона. Постоянное напряжение  $U_0$  (рис. 2.20), соответствующее касанию линии к параболе критического режима называется напряжением синхронизации. Между этой линией и параболой критического режима находится рабочая область СВЧ колебаний. Таким образом, рабочее анодное напряжение магнетрона всегда удовлетворяет условию  $U_a > U_{\text{п}}$ .

В области выше критической параболы анодный ток достигает величины, ограниченной лишь условиями пространственного заряда, СВЧ колебания в этой области отсутствуют. Генерация должна быть равна нулю также и под критической параболой при анодных напряжениях, лежащих ниже напряжения синхронизации  $U_0$ .

**Коэффициент полезного действия магнетрона.** Рассмотрим кинетическую энергию, которую имеет электрон в момент удара об анод, после того

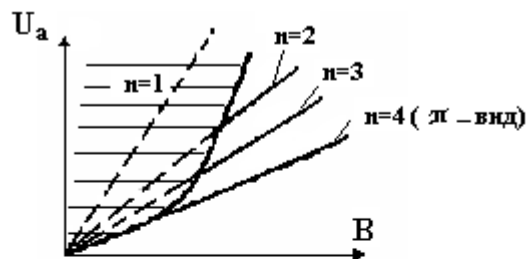


Рис. 2.19. Семейство прямых, приближенно определяющих условия самовозбуждения магнетрона по уравнению (2.71) при  $N=8$  и  $m=0$ .

Здесь же построена парабола критического режима

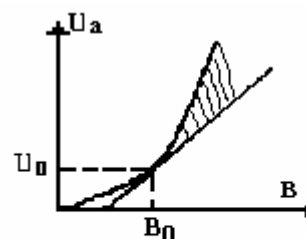


Рис. 2.20. Рабочая область заштрихована

как значительная часть потенциальной энергии электрона отдана ВЧ полю в пространстве взаимодействия.

Максимальная кинетическая энергия, рассеиваемая электроном на аноде, равна

$$W_{\text{расс.макс}} = \frac{m_0 V_{\text{макс}}^2}{2} = \frac{2m_0 a^2}{B^2 d^2},$$

где

$$V_{\text{макс}} = \omega_{\text{ц}} \cdot 2R = \frac{2U_a}{Bd},$$

здесь  $m_0$  – масса электрона.

Тот же электрон, находясь на катоде до начала движения в пространстве взаимодействия, обладает по отношению к аноду потенциальной энергией, равной  $W_n = eU_a$ , при нулевой кинетической энергии. Следовательно, энергия, отданная ВЧ полю, по закону сохранения энергии равна  $W_n - W_{\text{расс.макс}}$ . КПД рассматриваемого одиночного электрона имеет величину

$$\eta_{\text{э}} = \frac{W_n - W_{\text{расс.макс}}}{W_n} = 1 - \frac{U_a}{B^2} \cdot \frac{2m_0}{ed^2}. \quad (2.73)$$

Запишем это уравнение для цилиндрической системы, полагая  $d = r_a - r_k$ , тогда

$$\eta_{\text{э}} = 1 - \frac{U_a}{B^2} \cdot \frac{2m_0}{e(r_a - r_k)^2}. \quad (2.74)$$

Расчетные значения электронного КПД обычно превышают получаемую на практике величину  $\eta_{\text{э}}$ . Отчасти это и понятно, так как в рассмотренном расчете не были учтены потери энергии за счет "неблагоприятных" электронов, бомбардирующих катод, а так же потери электронов на боковые крышки магнетрона и некоторые другие факторы.

Величина электронного КПД существующих магнетронов сантиметрового диапазона составляет 50–60%, а в некоторых случаях и более. Столь высокая эффективность магнетронов делает их ценнейшими, мощными автогенераторами диапазона СВЧ.

Зависимость электронного КПД магнетрона от величины магнитной индукции можно сделать особенно наглядной, если учесть условие синхронизма при работе на любом фиксированном виде колебаний. Например, возьмем условие самовозбуждения (2.71) и подставим его в соотношение в (2.74), то при  $m=0$  имеем

$$\eta_{\text{э}} = 1 - \frac{m_0 \omega_n}{e} \cdot \frac{r_a + r_k}{r_a - r_k} \cdot \frac{1}{nB}.$$

Для данного магнетрона при неизменной частоте уравнение для  $\eta_3$  можно переписать в виде

$$\eta_3 = 1 - \frac{\text{const}}{nB}. \quad (2.75)$$

Полученное уравнение соответствует случаю, когда КПД отсчитывается при движении вдоль одной из прямых самовозбуждения на плоскости  $(U_a, B)$ , соответствующее графическое изображение зависимости  $\eta_3 = f(B)$  для

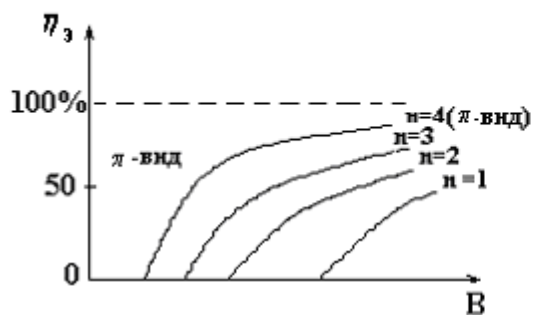


Рис. 2.21. Характеристика электронного КПД магнетрона для различных видов колебаний

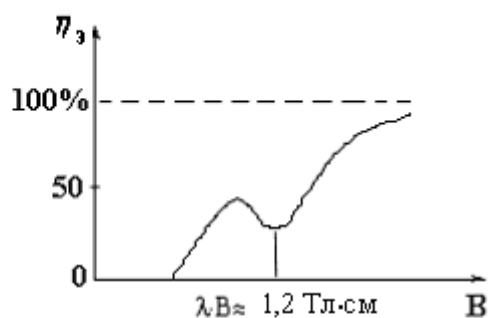


Рис. 2.22. Зависимость  $\eta_3(B)$  для разнорезонаторного магнетрона на  $\pi$ -виде колебаний

8-резонаторного магнетрона показано на рис. 2.21.

Кривые электронного КПД имеют вид отрезков гипербол. Чем ниже номер вида колебаний  $n$ , тем меньше электронный КПД при одной и той же величине магнитной индукции  $B$ . Говоря иначе, для достижения одного и того же электронного КПД наименьшее магнитное поле требуется при  $\pi$ -виде колебаний.

Неограниченный рост электронного КПД магнетрона при увеличении  $B$  легко понять с физической точки зрения, так как при этом происходит уменьшение радиуса катящегося круга и уменьшение вращательной энергии электрона, рассеиваемой на аноде.

В случае разнорезонаторных магнетронов имеются отклонения от зависимости  $\eta_3 = f(B)$ , показанной на рис. 2.22.

В некотором интервале значений магнитной индукции наблюдается "провал" на кривой зависимости  $\eta_3 = f(B)$ , как показано на рис. 2.22. Электронный КПД здесь резко падает наполовину, после чего вновь возрастает и далее изменяется по закону гиперболы. Исследования показали, что в центре "провала" произведение магнитной индукции, выраженной в теслах, на длину волны в сантиметрах имеет для всех магнетронов одинаковую величину, равную приблизительно 1,2 Тл·см.

"Провал" электронного КПД в разнорезонаторных магнетронах можно качественно объяснить с точки зрения циклотронного резонанса, возникающего при условии  $\omega = \omega_c$ . В пространстве взаимодействия разнорезонаторного

магнетрона, кроме поля  $\pi$ -вида возникают колебания  $n=0$ . При приближительном совпадении частоты генерируемых колебаний и циклотронной частоты характер движения электронов может измениться. Большую роль играет тот факт, что поле нулевой составляющей значительно медленнее убывает при удалении от анода, чем поле вида  $\pi$ -вида. Структура поля, форма спиц и взаимодействие их с полем  $\pi$ -вида ухудшаются.

Обычно разнорезонаторные магнетроны эксплуатируют при значениях магнитного поля, лежащего ниже центра "провала". Перейти в область большой величины магнитной индукции "за провалом" практически не удается из-за трудностей получения очень сильных магнитных полей.

Наибольшему электронному КПД соответствует режим, при котором электроны рассеивают на аноде наименьшую энергию. Поэтому можно было бы предположить, что линия порогового напряжения (рис. 2.19) является одновременно и линией максимального электронного КПД. Однако при очень малой амплитуде ВЧ напряжения электронный КПД ухудшается вследствие влияния полей других видов. Поэтому максимум электронного КПД, когда  $V=\text{const}$ , достигается при анодном напряжении, несколько превышающем пороговую величину. При дальнейшем увеличении  $U_a$  и приближении к критическому напряжению  $U_{кр}$  величина электронного КПД, естественно стремится к нулю. Если увеличивать магнитное поле, то характер зависимости  $\eta_3 = f(U_a)$  в основном сохраняется, но максимум ее становится более высоким, а вся кривая сдвигается в сторону более высоких анодных напряжений.

Приведем данные о зависимости электронного КПД от активной проводимости внешней нагрузки при неизменных величинах магнитного поля и анодного тока  $\eta_3 = f(G)$ . С уменьшением величины  $G'_H$ , то есть при ослаблении связи магнетрона с нагрузкой, происходит постепенное повышение амплитуды напряжения в пространстве взаимодействия. При соизмеримости амплитуды колебаний с постоянным анодным напряжением электроны начинают попадать на анод, не совершив большого числа циклоидальных колебаний. Электронный КПД должен снизиться. С другой стороны, при очень малых амплитудах  $\eta_3$  так же снижается за счет помех от других видов колебаний. Таким образом, должен существовать довольно широкий диапазон амплитуд и, следовательно, значений  $G'_H$ , при которых  $\eta_3$  магнетрона проходит через максимум.

Некоторое влияние на величину электронного КПД. Оказывает разделение видов колебаний. Электронный КПД магнетрона может так же снизиться за счет влияния поля связок около концов анодного блока. Это поле, не имеющее азимутальных вариаций, оказывает примерно такое же воздействие на пространственный заряд в магнетроне, какое оказывает поле нулевой составляющей в магнетронах разнорезонаторного типа. Для устранения подобных эффектов связки экранируются путем расположения их в кольцевых канавках, выточенных на торцах анодного блока.

Падение  $\eta_3$  наблюдается иногда из-за неточной центровки катода относительно анода. Размер катода так же влияет на величину  $\eta_3$ . С одной

стороны, при малой величине отношения  $r_k/r_a$  не выполняются условия синхронизма между электронами и высокочастотным полем, так как напряженность постоянного электрического поля имеет наибольшую величину у катода и уменьшается по направлению к аноду. С другой стороны, при  $r_k/r_a \rightarrow 1$ , возрастает энергия, рассеиваемая электронами на аноде магнетрона.

Полный КПД определяется с учетом КПД резонаторной системы  $\eta_{\text{полн}} = \eta_{\text{э}} \eta_{\text{рез}}$ , где  $\eta_{\text{рез}}$  связана с собственной, нагруженной и внешней добротностями колебательной системы (1.23). Величина внешней добротности ( $Q_{\text{внеш}}$ ) обычно не бывает ниже 100–200. Собственную добротность желательно иметь как можно выше. Типичная величина  $Q_0$  в сантиметровом диапазоне имеет порядок 1000. Практически КПД резонаторной системы может составлять от 90–95% до 50–60% на более коротких волнах. Типичные значения полного КПД составляют от 60–70% на дециметровых волнах до 20–30%, на сантиметровых волнах.

**Рабочие и нагрузочные характеристики магнетронов.** Анализ процессов, происходящих в магнетроне, проводят с помощью экспериментального исследования режимов работы магнетрона, на основании которого получают необходимые характеристики.

Характеристики магнетронов разделяют на две группы:

I) рабочие, связывающие выходные параметры генератора (мощность  $P$ , частота  $f$ , и коэффициент полезного действия (КПД)  $\eta$ ) с входными параметрами (анодное напряжение  $U_a$ , анодный ток  $I_a$  и магнитная индукция  $B$ ). Генератор работает на согласованную нагрузку ( $\Gamma_n=0$ ).

II) нагрузочные, связывающие выходные параметры генератора ( $P$ ,  $f$ ,  $\eta$ ) с импедансом нагрузки (коэффициентом отражения  $\Gamma$ ) при фиксированных параметрах питания, близких к номинальным ( $I_a = I_{a0}$ ;  $B = B_0$ ).

I - группа характеристик: семейство зависимостей тока от напряжения - вольт - амперные характеристики (ВАХ). Эти характеристики снимаются при постоянном магнитном поле и фиксированной нагрузке, соответствующей режиму согласования выходного устройства магнетрона. Параметрами при снятии  $U_a = f(I_a)$  являются  $B$ ,  $P_{\text{ген}}$ , частота генерируемых колебаний  $f$  и полный КПД -  $\eta_{\text{полн}}$ . Вольт - амперные характеристики магнетрона, снятые при условиях  $B = \text{const}$ ,  $P_{\text{ген}} = \text{const}$ ,  $f = \text{const}$ ,  $\eta_{\text{полн}} = \text{const}$  носят название рабочих характеристик. Эти характеристики принято строить в прямоугольной системе координат.

II-группа - нагрузочные характеристики, определяющие зависимость  $P_{\text{ген}}$  и  $f$  от полного сопротивления (полной проводимости) нагрузки. Нагрузочные характеристики  $P_{\text{ген}} = f(Z_n)$  и  $f = f(Z_{\text{ген}})$  при  $I_a = \text{const}$ , как и для других типов автогенераторов СВЧ, оказывается удобным строить на комплексной плоскости полного сопротивления нагрузки в полярной системе координат, так называемой диаграмме Вольперта - Смита.

**Рабочие характеристики магнетронов.** Построим рабочие характеристики магнетронов типа бегущей волны для семейства кривых постоянной ге-

нерируемой мощности  $P_{ген} = const$ . Генерируемая мощность магнетрона связана с анодным напряжением, постоянным анодным током и КПД соотношением  $P_{ген} = \eta_{полн} U_a I_a$ .

Если бы КПД магнетрона оставался неизменным и не зависел от  $I_a$ , то ВАХ магнетрона  $U_a = f(I_a)$  при  $P_{ген} = const$  имели бы вид равнобочных гипербол. Однако с ростом  $I_a$  электронный КПД уменьшается, за исключением узкой области малых токов. Поскольку КПД резонаторной системы, естественно, остается неизменным, с увеличением тока  $I_a$  линии постоянной мощности отклоняются от гипербол и несколько поднимаются, как показано на рис. 2.23, а, чем больше  $P_{ген}$ , тем выше и правее должны располагаться кривые  $P_{ген} = const$ .

Еще более просто можно рассуждать для нахождения хода кривых постоянной магнитной индукции  $B = const$ . Если исходить из условия синхронизации электронного облака и поля, например, в соотношении (2.71) можно предположить, что линии  $B = const$  являются прямыми параллельными оси абсцисс. В действительности, однако, увеличение постоянного  $I_a$  требует изменения формы пространственного заряда и некоторого повышения  $U_a$ . Таким образом, линии  $B = const$  на рабочей характеристике магнетронов должны иметь вид семейства параллельных прямых имеющих наклон под малым углом к оси абсцисс (см. рис. 2.23, б).

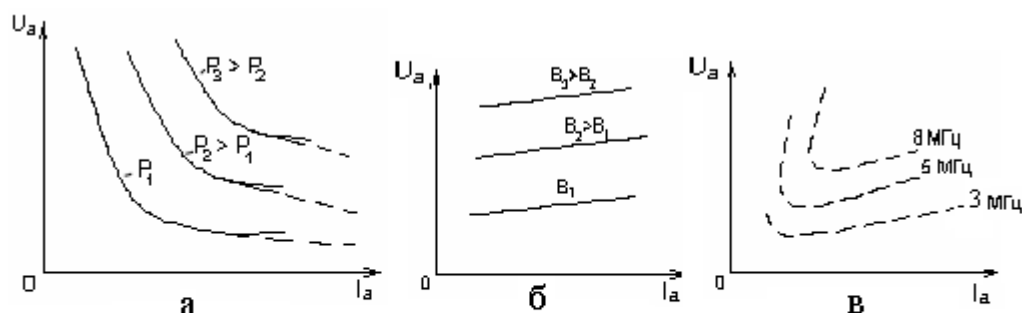


Рис. 2.23. Рабочие характеристики магнетрона при:

а – постоянной генерируемой мощности, б – постоянной магнитной индукции, в – постоянной частоте.

Ход кривых постоянной частоты генерируемых колебаний  $\nu_{ген} = const$  можно рассматривать с позиций электронного смещения частоты. Линии  $\nu_{ген} = const$  имеют крутой наклон в области малых токов рис. 2.23, в, но идут почти параллельно линиям  $B = const$  в той части графика, которая близка к нормальному рабочему режиму.

**Нагрузочные характеристики магнетронов.** В нагрузочных характеристиках магнетрона наблюдается затягивание частоты под действием внешней нагрузки значительно меньше, чем разделение видов колебаний. Магнетрон рассматриваем как, "одноконтурный" генератор, колебательная система которого в рассматриваемом диапазоне обладает свойствами параллельного резо-

нансного контура. Линии постоянной частоты на идеализированной форме нагрузочной характеристики (показанной на рис. 2.24)

$$v_{\text{ген}} = \frac{\omega_{\text{ген}}}{2\pi} = \text{const}, \quad (2.76)$$

совпадают с окружностями постоянной входной реактивной проводимости. Линии  $P_{\text{ген}} = \text{const}$  совпадают с окружностями постоянной активной проводимости.

Семейства  $P_{\text{ген}} = \text{const}$  и  $v_{\text{ген}} = \text{const}$  пересекаются под одним и тем же углом  $\alpha$ . На диаграмме заштрихована область срыва, где колебания в генераторе отсутствуют. Линии постоянной частоты в этой области диаграммы не имеют физического смысла.

Окружность с  $\rho = 1,5$  - это окружность постоянного КСВ, равного 1,5.

Вывод энергии нейтрона конструируется образом, чтобы обеспечить допустимую величину заваивания частоты. Поэтому, видно из рис. 2.24, максимум генерируемой мощности получается не в режиме гласования, а при тельно высоком КСВ грузки. В области высокогенерируемой мощности, сходятся линии постоянной частоты, работа магнетрона обычно бывает неустойчивой.

Рассмотрим работу магнетрона в импульс-режиме. Основные характеристики импульса подаваемого на магнетрон: а) длительность импульса  $\tau_{\text{имп}}$ , приложенного к магнетрону; б) крутизна фронта импульса. Длительность импульса не может выбираться произвольно. Нужно учитывать, что при коротких импульсах ( $\tau_{\text{имп}} = 0,1 - 0,3 \text{ мкс}$ ) начинает сказываться конечное время установления колебаний в магнетроне. При длинных импульсах ( $\tau_{\text{имп}} \approx 2 - 5 \text{ мкс}$ ) ограничивающими факторами становятся спад эмиссии катода и тепловой режим анода и катода. Спад эмиссионного тока так же делает возможным переход магнетрона на другой вид колебаний.

Оптимальная форма импульса, при которой могут быть обеспечены наиболее благоприятные условия возбуждения рабочего вида колебаний, должна быть в основном прямоугольной.

Для модуляции магнетронов импульсами малой длительности порядка (десятков наносекунд) возможно применение пьедестального импульса сравнительно большой продолжительности (порядка 1 мкс), на который накладыва-

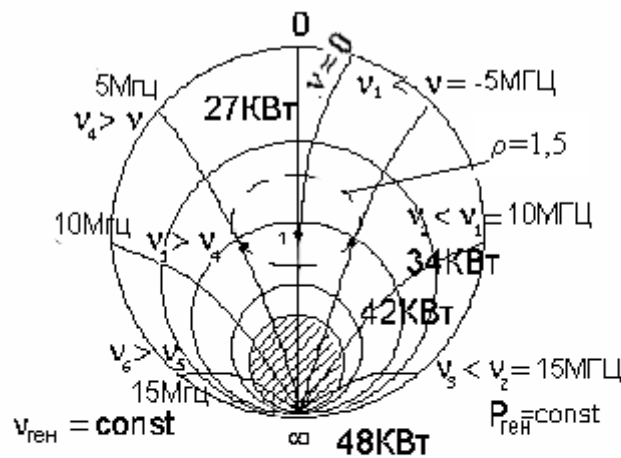


Рис. 2.24. Пример реальной нагрузочной характеристики импульсного магнетрона 3-сантиметрового диапазона при  $B=5,5 \text{ Тл}$  и  $I_a=10 \text{ см}$ . линии  $v=\text{const}=0$  соответствует номинальная частота, равная 9375 МГц.

маг-  
таким  
чить  
тяги-  
как  
си-  
ности  
со-  
ни-  
на-  
кой  
где  
ной  
трона  
той-

ту  
ном

вается с некоторым запаздыванием короткой рабочий импульс. Роль пьедестального импульса - предварительное возбуждение колебаний  $\pi$ -вида. Обычно опытным путем подбирают крутизну фронтов и форму плоской части импульса.

Другой трудной задачей является генерирование с помощью магнетрона длинных импульсов ( $\tau_{\text{имп}} = 10 \text{ мкс}$ ). Эта задача, однако, не может быть решена путем усовершенствования схемы модулятора. Опыт показывает, что необходимо переходить к новым типам катодов, способным работать без искрения и утомления и имеющим достаточный срок службы.

Импульсные магнетроны перекрывают миллиметровый, сантиметровой, дециметровый диапазоны длин волн. Импульсные выходные мощности составляют 100Вт–10КВт в коротковолновой части миллиметрового диапазона, до 1–2МВт в дециметровом диапазоне.

Они имеют относительно невысокие анодные напряжения, которые в небольших пределах изменяются от одного типа магнетрона к другому, например, при уменьшении уровня мощности в 1000 раз в диапазоне волн от 25 см до 0,4 м анодное напряжение уменьшается всего в 5 раз, а скважность увеличивается в 9 раз. Для этих же условий ток уменьшается в 40 раз.

Импульсные магнетроны являются распространенными приборами с фиксированной настройкой частоты. Возможности механической перестройки частоты невелики и составляют не более 5–10%. Электронная перестройка частоты возможна в очень узких пределах и часто сопровождается понижением стабильности работы и надежности магнетрона.

**Параметры и области применения магнетронов.** На данное время разработано и выпущено свыше 500 типов многорезонаторных магнетронов. Основные тенденции в развитии магнетронов связаны с наращиванием уровней мощности, укорочением  $\lambda$ , увеличением срока службы, повышением стабильности и надежности.

Высокая эффективность и сравнительная простота конструкции является основными достоинствами магнетронных генераторов.

Перспективами дальнейшего развития магнетронов является: а) повышение максимальной генерируемой мощности; б) освоение нижней части миллиметрового диапазона волн; в) повышение стабильности частоты; г) увеличение срока службы.

Происходящая сейчас специализация применения магнетронов предъявляет к ним ряд новых повышенных требований, например в отношении д) низкого уровня шумов на частотах близких к генерируемой частоте.

В некоторых случаях, в первую очередь, в установках промышленного нагрева, от магнетронов в основном требуется е) более высокий КПД при отсутствии специальных требований к стабильности частоты.



Особенностью магнетронных генераторов является наличие электронного смещения частоты. В среднем можно считать, что *крутизна электронной перестройки* составляет от десятых долей до единиц МГц на 1 А изменения тока

Практически электронная перестройка частоты используется только в маломощных магнетронах. В мощных приборах применяются *механические методы перестройки* такие, как введение стержней в полости резонаторов (индуктивная настройка), или перемещение колец, расположенных параллельно торцевым поверхностям анодного блока (емкостная настройка). Механические методы позволяют получить диапазон настройки до 10-15%.

На рис. 2.25 приведены зависимости мощности и КПД от частоты.

I и II кривая относится к предельным уровням мощности для магнетронов с фиксированной настройкой и механически перестраиваемых по частоте. Заштрихованная область - уровни мощности практических конструкций магнетронов. Нижняя граница совпадает с границей предельных уровней мощности для генераторов СВЧ на триодах.

III - усредненные и предельные значения КПД. Размеры вертикальных отрезков указывают на большие разбросы КПД от одного типа к другому, что в основном связано с различием уровней этих магнетронов.

КПД магнетронов на частоте выше 30000 МГц составляет 7-20% и уменьшается на частоте 80000 МГц до 3-5%, но является наиболее высоким из всех известных распространенных конструкций других мощных генераторов СВЧ.

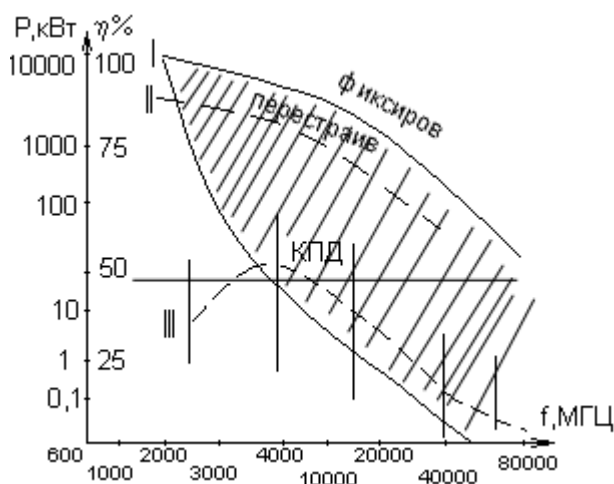


Рис. 2.25

Магнитные поля создаются постоянными магнитами или электромагнитами. Чем короче длина волны, тем большая требуется величина магнитной индукции. В сантиметровом диапазоне она составляет 0,1–0,8Тл, в дециметровом диапазоне 0,05-0,2Тл.

Для охлаждения анодного блока предусмотрен радиатор, обдуваемый воздухом. Водяное охлаждение применяется только в наиболее мощных приборах.

Магнетроны с длиной волны более 3см обычно имеют петлевую связь с передающей линией. В магнетронах более коротких волн размеры анодного блока которых малы, поэтому используют связь с помощью щели, возбуждающей выходной волновод через четвертьволновой трансформатор.

Магнетроны непрерывного режима работают в сантиметровом и дециметровом диапазонах, с выходными мощностями от 1-100 Вт и 100Вт - 100КВт соответственно.

Основные области применения магнетронов:

1. Передатчики импульсных р/л станций и р/л маяков;
2. Генераторы помех;

3. Передатчики в линиях им-пульсной СВЧ связи и радиотехнических системах;
4. Импульсные генераторы для питания линейных электронных ускорителей;
5. Генераторы непрерывного режима для промышленного и бытового нагрева;

### 2.4.3. Лампы бегущей и обратной волны типа М

Как было показано в разделе 2.4.1, при определенной взаимной ориентации векторов  $\vec{E}, \vec{B}, \vec{V}$  электрическая и магнитные силы имеют противоположные направления. В этих условиях их равнодействующая может быть равной нулю  $-e\{\vec{E}_0 + [\vec{V}\vec{B}]\} = 0$ , если  $v = v_{ц} = \frac{E_0}{B}$ . Следовательно, при такой скорости  $v = v_{ц}$  движения электрона в скрещенных полях будет прямолинейно и равномерно. Электрическое поле вдоль его пути неодинаково. Непосредственно над катодом действует поле  $E_{упр}$ , а далее между замедляющей системой и отрицательным электродом – поле  $E$ , которое вдвое сильнее управляющего.

$$E_0 = 2E_{упр}. \quad (2.77)$$

В условиях фазового синхронизма скорость электрона должна быть равна фазовой скорости прямой или обратной волны. Учитывая это условие, можно связать  $U_0$  со скоростью  $V_{\phi}$  и  $V_{ц}$  соотношением

$$U_0 = \frac{mv_{ц}^2}{2e} = \frac{mV_{\phi}^2}{2e} \approx 25,6 \cdot 10^4 \left( \frac{V_{\phi}}{c} \right)^2. \quad (2.78)$$

Подбираются такие условия, чтобы электроны в момент входа в пространство взаимодействия, образуемого верхним электродом замедляющей системы (анод) и нижним электродом (холодный катод), находились на вершине циклоиды. В этой точке имеется только горизонтальная составляющая скорости  $V_{0z}$  (начальная скорость), которая определяется по формуле  $V_{0z} = 2E_{упр}/B$ . Если начальная скорость электронов направлена параллельно электродам и равна скорости переноса, то траектория электронов прямолинейна. Скорость переноса в пространстве взаимодействия определяют по формуле  $V_{п} = E_0/B$ . Таким образом, при выполнении условия  $V_{0z} = V_{п}$  траектория оказывается прямолинейной. Электроны должны попасть при отсутствии высокочастотного поля в пространстве взаимодействия на коллектор.

Высокочастотный сигнал подводится через вход замедляющей системы. Если фазовая скорость пространственной гармоники  $V_{\phi m}$  равна скорости переноса  $V_{п}$ , то в пространстве взаимодействия происходит увеличение энергии СВЧ-поля, в результате уменьшения потенциальной энергии электронов. Для предотвращения самовозбуждения ставится поглотитель.

Дальнейшее рассмотрение работы ЛБВМ схоже с магнетроном.

Электронный коэффициент полезного действия может быть представлен в виде

$$\eta_{\text{Э}} = 1 - \frac{25,6 \cdot 10^4 \left( \frac{V_{\phi}}{c} \right)^2}{U_a} \quad (2.79)$$

Коэффициент усиления для ЛБВМ

$$K_y(U) = -6 + 54,6 CN, \quad (2.80)$$

где  $N$  - электрическая длина замедляющей системы,  
 $C$  - параметр усиления (2.41).

Коэффициент усиления ЛБВМ ниже, чем в ЛБВО, но достаточно велик 20-25 дБ, а КПД сравнительно высокий 40-60 %. Пусковой ток в генераторной ЛОВМ определяется формулой

$$I_0^n(\text{пуск}) = (2n - 1)^2 I_0^{(1)} \text{пуск}, \quad (2.81)$$

где  $n$  - номер генерируемой зоны.

*При большом токе пучка одновременно возможна генерация колебаний в нескольких зонах.*

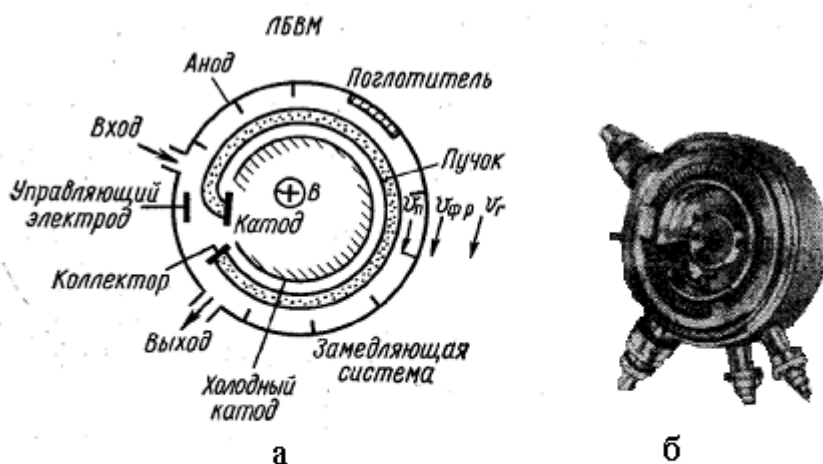


Рис. 2.26. ЛБВ цилиндрической конструкции:

а - схема, б- внешний вид.

В лампах обратной волны М, которые могут быть генераторными или усилительными устройствами взаимодействие электронного потока происходит с обратной пространственной гармоникой. На рис. 2.26,а и в показаны устройство и внешний вид ЛБВМ. Электроны, эмитируемые катодом, под действием напряжения электрического и магнитного поля двигаются по циклоидальной кривой и входят в пространство взаимодействия с вершины циклоиды. При выполнении таких же условий, как в ЛБВМ, электрон при отсутствии СВЧ поля двигается по окружности, если это цилиндрическая система и прямой в плоских электродах и попадает на коллектор. При выполнении условия синхронизма  $V_{\phi(-1)} = V_{\text{п}}$ , и токе пучка больше пускового, ЛОВМ возникают колебания. У коллекторного конца расположен поглотитель, который поглощает энергию, отраженную от выходного устройства, если согласование в нем недостаточно хорошее.

На рис. 2.27 показано устройство цилиндрической ЛОВМ. Как видно отличие ее в том, что имеется только один зазор, предназначенный для вывода энергии

из пространства взаимодействия, расположенный со стороны катода. Принцип действия такой системы аналогичен магнетрону.

Выходная мощность ЛОВМ и ЛБВМ практически линейно зависит от тока пучка как это показано на рис. 2.28.

$$P_{\text{вых}} = A(I_0^{(1)} - I_0 \text{ пуск}), \quad (2.82)$$

где  $A$  – некоторый коэффициент. Выходная мощность ЛОВМ в дециметровом диапазоне в непрерывном режиме достигает нескольких десятков киловатт, а в сантиметровом диапазоне – порядка нескольких сот ватт и а миллиметровом диапазоне – десятки ватт. Электронный КПД ЛОВМ также зависит от тока пучка. Сначала КПД растет с увеличением тока пучка, а затем после перехода ЛБВ в режим насыщения практически не изменяется. В этом режиме  $P_{\text{вых}}$  и затраченная мощность  $P_0 = I_0 U_0$  практически одинаково увеличиваются с ростом тока. Максимальное КПД можно определить по формуле:  $\eta_{\text{э}} = 1 - U_{\text{синх}} / U_0$ , где  $U_{\text{синх}}$  – потенциал синхронизации. Обычно КПД составляет 50-60%.

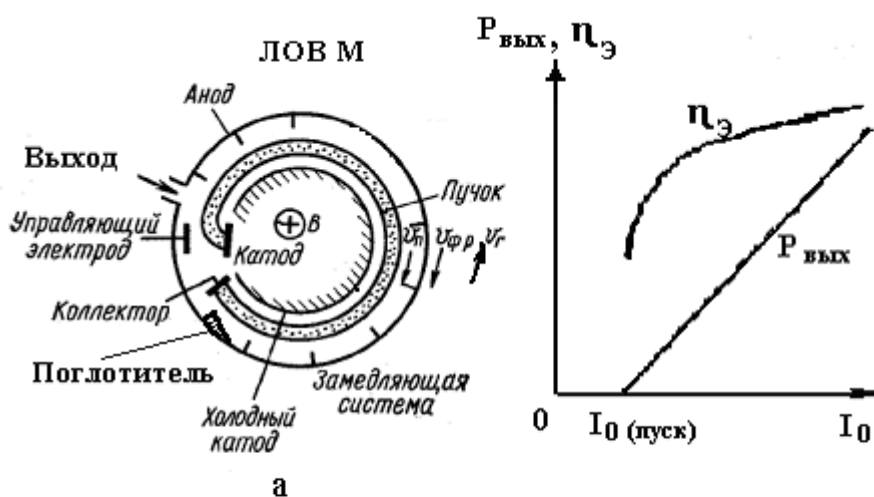


Рис. 2.27. Схема устройства цилиндрической генераторной ЛОВМ

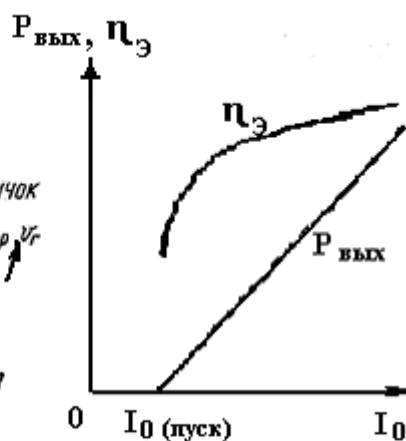


Рис. 2.28. Зависимости выходной мощности и электронного КПД от тока пучка

ЛОВМ применяется обычно в радиоаппаратуре радиопротиводействия, в которой наиболее полно реализуются преимущества этих приборов по обеспечению быстрой электронной перестройки частоты в широких пределах при достаточно большой выходной мощности.

## ГЛАВА 3.

### ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ СВЧ

Твердотельная электроника стала успешно развиваться с появлением биполярных и полевых транзисторов, которые вначале представляли собой типичные низкочастотные приборы. Когда был открыт эффект динамического отрицательного сопротивления, появилась возможность применять в полупроводниковых приборах принцип использования конечного времени пролета электронов, широко применяемый в вакуумной электронике СВЧ. Использование отрицательной дифференциальной проводимости полупроводникового материала, которая обусловлена переходом из одного энергетического минимума (долины) зоны проводимости в другой, где значительно ниже подвижность носителей, позволило расширить функциональные возможности полупроводниковой электроники СВЧ.

В настоящее время усилители на транзисторах успешно функционируют в сантиметровом диапазоне, конкурируя с усилителями на вакуумных ЛБВ. Также широко используются в технике СВЧ конструкции полупроводниковых диодов: диоды р-п, р-і-п, м-п (диоды с барьером Шоттки), диоды М-О-П (металл - оксид - проводник), диоды М-Д-П (металл - диэлектрик - проводник), диоды с накоплением заряда (ДНЗ), лавинно-пролетные диоды (ЛПД), диоды с объемной неустойчивостью.

#### 3.1. Элементы зонной теории твердого тела

Твердое тело, в том числе рассматриваемые полупроводниковые монокристаллы, представляет собой систему, состоящую из большого числа атомов, плотность которых составляет примерно  $10^{22}$  на  $1 \text{ см}^3$ . Каждый атом характеризуется дискретным спектром энергий  $E_1, E_2, E_3, \dots, E_n$ , разрешенных для электронов (рис. 3.1,а).

По представлениям квантовой механики состояние электрона в атоме характеризуется квантовыми числами. Электроны в атоме могут занимать только вполне определенные, разрешенные энергетические уровни. В нормальном (невозбужденном) состоянии атома электроны должны располагаться на самых низких энергетических уровнях.

Однако в соответствии с принципом Паули в одном и том же атоме (или в какой-либо квантовой системе) не может быть двух электронов, обладающих одинаковой совокупностью квантовых чисел. Поэтому электроны в соответствии с их состояниями распределяются по определенным оболочкам вокруг ядра. Атом кремния имеет 14 электронов, расположенных на трех оболочках по 2, 8 и 4 электрона. Атом германия имеет 32 электрона, расположенных на четырех оболочках по 2, 8, 18 и 4 электрона. Валентные электроны на последней (верхней) оболочке, имеющие наибольшую энергию, определяют электропроводность кремния и германия.

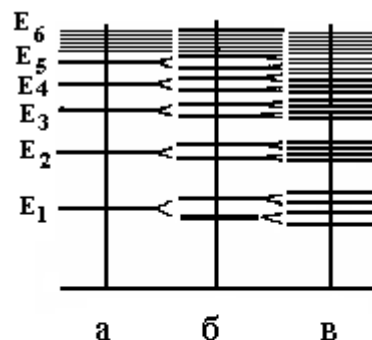


Рис. 3.1. Разрешенные энергетические уровни электронов в атоме

Практическое следствие из принципа Паули при рассмотрении германия и кремния заключается в том, что при объединении двух атомов происходит расщепление каждого уровня на два (рис. 3.1,б), а при объединении  $N$  атомов в кристалл происходит расщепление каждого энергетического уровня на  $N$  уровней (рис. 3.1,в).

Однако в соответствии с соотношением неопределенностей Гейзенберга частица не может одновременно иметь вполне точные значения, например, координаты  $x$  и соответствующей этой координате составляющей импульса  $p_x$ , причем неопределенности в значениях этих величин  $\Delta x$  и  $\Delta p_x$  должны удовлетворять условию:  $\Delta x \cdot \Delta p_x \geq \hbar$ , где  $\hbar = h/2\pi$  — постоянная Планка;  $h = 6,62 \times 10^{-34}$  Дж·с. Чем точнее определена одна из величин,  $x$  или  $p$ , тем больше становится неопределенность другой. В рассматриваемом случае находящиеся на очень близком энергетическом расстоянии отдельные уровни, образовавшиеся при объединении множества атомов Ge или Si в кристалл (рис. 3.1,в) становятся практически неразличимы. Совокупность близко расположенных энергетических  $N$  уровней называется энергетической зоной.

Разрешенные энергетические зоны отделены друг от друга запрещенными энергетическими участками, которые называются запрещенными зонами. Ширина запрещенных зон зависит от расстояния между атомами, т. е. определяется строением кристаллической решетки, а также строением и состоянием атомов, образующих монокристалл. Таким образом, рис. 3.1,в можно назвать зонной диаграммой кристалла, в которой разрешенные зоны чередуются с запрещенными.

Для рассмотрения физических процессов в полупроводниках и полупроводниковых приборах интерес представляют только три верхние энергетические зоны: валентная, запрещенная и свободная, или зона проводимости, так как именно эти зоны обуславливают электропроводность полупроводника.

Электропроводность возможна лишь тогда, когда возможен переход электрона на другой энергетический уровень. Это означает, что в проводимости могут участвовать электроны только тех зон, где есть свободные уровни, а при температуре абсолютного нуля они имеются лишь в самой верхней разрешенной зоне, которую называют, поэтому, зоной проводимости. Нижний энергетический уровень зоны проводимости обозначается  $E_c$ . В зоне проводимости находятся электроны, осуществляющие электропроводность кристалла.

Зона проводимости отделена от валентной запрещенной зоной, ширина которой обозначается  $\Delta E$ .

Валентная зона — это зона, в которой все энергетические уровни заняты при температуре абсолютного нуля, поэтому электроны этой зоны не могут участвовать в проводимости. Однако, с повышением температуры электроны переходят из валентной в зону проводимости, в валентной зоне образуются свободные уровни и появляется возможность перехода на них электронов с близлежащих уровней. Верхний энергетический уровень валентной зоны обозначается  $E_v$ .



Рис. 3.2. Энергетические диаграммы твердого тела

Электропроводность твердых тел зависит от взаимного расположения зоны проводимости и валентной зоны. Именно по характеру энергетических диаграмм твердые тела более четко разделяются на проводники (металлы), полупроводники и диэлектрики (изоляторы) (рис. 3.2).

В металлах зона проводимости и валентная зона перекрываются, и электроны валентной зоны могут легко переходить в зону проводимости независимо от температуры тела. В зоне проводимости электроны принадлежат всему твердому телу и свободно внутри него перемещаются. Проводящее состояние является обычным. В отличие от металлов при нулевой температуре у изоляторов и полупроводников зона проводимости пуста и электропроводность отсутствует.

Ширина запрещенной зоны  $\Delta E$  — один из важнейших параметров полупроводникового материала, определяющий его электрохимический потенциал. Чем больше число оболочек в атоме, тем меньше ширина запрещенной зоны и тем слабее связаны валентные электроны с ядром.

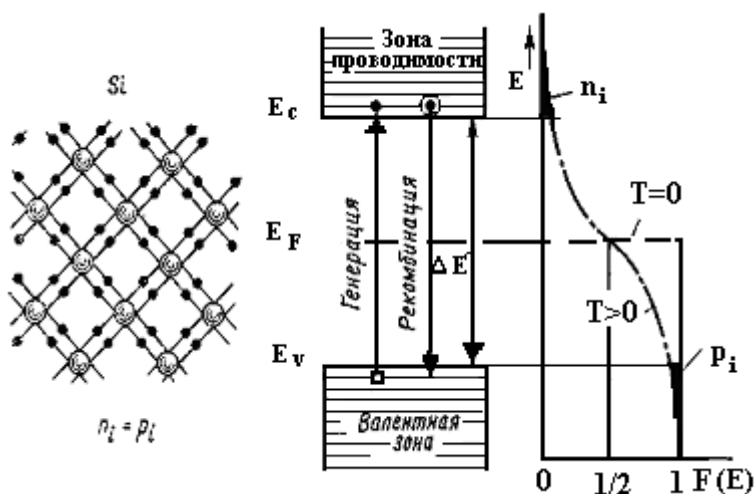


Рис. 3.3. Переход электронов из валентной зоны в зону проводимости-генерация носителей и обратный переход – рекомбинация носителей.

В отличие от металлов электропроводность полупроводника обусловлена электронами и дырками, т. е., кроме электронной, в полупроводнике существует и другой вид проводимости — дырочная. Переход электронов из валентной зоны в зону проводимости приводит к образованию дырок — вакантных уровней в валентной зоне. При температуре, отличной от нуля, в зоне проводимости полупроводника всегда имеется  $n$  электронов, а в валентной зоне —  $p$  дырок.

Переход электронов из валентной зоны в зону проводимости и образование дырки на освободившемся энергетическом уровне — процесс генерации носителей. Но наряду с генерацией происходит и рекомбинация носителей — переход электронов из зоны проводимости обратно в валентную зону на свободный уровень (рис. 3.3). При динамическом равновесии оба процесса идут непрерывно, их скорости равны.

Для производства полупроводниковых приборов необходимы материалы с определенной величиной ширины запрещенной зоны  $\Delta E$ , чтобы обеспечить электропроводность в заданном диапазоне рабочих температур прибора.

Из теории твердого тела известно, что если имеется набор энергетических уровней, то вероятность того, что электрон находится на энергетическом уровне с энергией  $E$ , описывается функцией Ферми — Дирака:

$$F(E) = \left[ \exp\left(\frac{E - E_F}{kT}\right) + 1 \right]^{-1}, \quad (3.1)$$

где  $E_F$  — параметр системы, называемый уровнем Ферми;  $kT$  — средняя тепловая энергия электрона;  $k$  — постоянная Больцмана, равная  $1,38 \cdot 10^{-23}$  Дж/К;  $T$  — температура в Кельвинах. При  $T = 300\text{K}$  (комнатная температура)  $kT = 0,25$  эВ.

Если  $E = E_F$ , то  $F(E) = 1/2$ , т. е. уровень Ферми можно определить как уровень, вероятность пребывания электрона на котором равна  $1/2$ .

Вероятность того, что энергетический уровень с энергией  $E$  занят электроном в зоне проводимости:

$$F_n(E) = \left[ \exp\left(\frac{E - E_F}{kT}\right) + 1 \right]^{-1}, \quad (3.2)$$

а вероятность того, что энергетический уровень с энергией  $E$  свободен или занят дыркой в валентной зоне:

$$F_p(E) = 1 - F_n(E) = \left[ \exp\left(\frac{E_F - E}{kT}\right) + 1 \right]^{-1}. \quad (3.3)$$

На рис. 3.3 показаны энергетическая диаграмма собственного полупроводника ( $n_i = p_i$ ), а также графики зависимости  $F(E)$  и распределение подвижных носителей зарядов по энергиям при различных температурах. При  $T = 0$  (пунктирная линия в виде ступеньки) валентная зона полностью заполнена, т. е.  $F_n(E) = 1$ , а зона проводимости пуста, т. е.  $F_p(E) = 0$ . Максимальное значение  $F(E)$  равно едини-



це. Уровень Ферми  $E_F$  совпадает с максимальной энергией, которую может иметь электрон при температуре абсолютного нуля.

С повышением температуры начинается термогенерация электронов и дырок, появляется вероятность нахождения электронов в зоне проводимости ( $F_n(E) > 1$ ) и дырок в валентной зоне ( $F_p(E) > 1$ ) (ступенька сглаживается — штрихпунктирная линия). Концентрации электронов в зоне проводимости и дырок в валентной зоне пропорциональны зачерненным площадкам на рис. 3.3.

В собственном полупроводнике уровень Ферми располагается в середине запрещенной зоны, и энергия уровня Ферми

$$E_F = (E_c - E_v) / 2, \quad (3.4)$$

если эффективные массы электрона и дырки равны ( $m_n^* = m_p^*$ ). Теперь можно определить при заданной температуре  $T$  количество свободных носителей заряда (электронов и дырок) в разрешенных зонах полупроводника, если известна ширина запрещенной зоны  $\Delta E$ , или определить их концентрации  $n$  и  $p$ .

Концентрация свободных электронов в зоне проводимости

$$n = \int_{E_c}^{\infty} D_n(E) F_n(E) dE, \quad (3.5)$$

где  $D_n(E)$  — плотность энергетических уровней в зоне проводимости. Под знаком интеграла — количество заполненных электронами в зоне проводимости уровней в элементарном интервале энергий  $dE$ .

В зоне проводимости  $E - E_F > 0$ , и экспонента в выражении для  $F_n(E)$  намного превышает единицу. В этом случае распределение Ферми — Дирака переходит в распределение Максвелла — Больцмана:

$$F_n(E) \approx \exp[-(E - E_F) / kT], \quad (3.6)$$

Если подставить значения  $D_n(E)$  и  $F_n(E)$ , то после некоторых преобразований можно получить выражение для  $n$ , которое будет содержать только числовые множители, универсальные константы, температуру  $T$ , а также параметры полупроводника.

Концентрация свободных электронов в зоне проводимости

$$n = AT^{3/2} \exp[-(E_c - E_F) / kT], \quad (3.7)$$

а свободных дырок в валентной зоне

$$p = AT^{3/2} \exp[-(E_F - E_v) / kT], \quad (3.8)$$

Коэффициент  $A$  содержит числовые множители и универсальные константы, он одинаков в выражениях для концентраций электронов  $n$  и дырок  $p$ , если допустить, что их эффективные массы равны ( $m_p^* = m_n^*$ ). Полученные соотношения для  $n$  и  $p$  справедливы для любых полупроводников. Они показывают, что с ростом температуры  $T$  концентрация подвижных носителей заряда растет по экспоненте.

Произведение концентраций подвижных носителей зарядов в полупроводнике

$$np = A^2 T^3 \exp(-\Delta E / kT)$$

не зависит от положения уровня Ферми и определяется только температурой и шириной запрещенной зоны  $\Delta E = E_C - E_F$ . Отношение концентраций носителей зарядов:

$$n/p = \exp[-2(E_i - E_F)/kT],$$

где  $E_i$  — энергия, соответствующая середине запрещенной зоны:

$$E_i = (E_C - E_V)/2.$$

Если разделить  $E_i$  на элементарный заряд  $q$ , то получим потенциал середины запрещенной зоны, который принято называть электрохимическим потенциалом:

$$\phi_i = E_i/q.$$

Если разделить величину средней тепловой энергии  $kT$  на элементарный заряд  $q$ , то получим величину, пропорциональную температуре и имеющую размерность потенциала:

$$\phi_T = kT/q \approx T/11600,$$

которую принято называть температурным потенциалом. При комнатной температуре  $\phi_T = 0,025$  В или 25 мВ.

В собственном полупроводнике концентрации свободных носителей зарядов одинаковы ( $n_i = p_i$ ), и их отношение

$$n/p = \exp[-2(E_i - E_F)/kT] = 1.$$

Отсюда следует, что  $E_F = E_i$ , т. е. что уровень Ферми в собственном полупроводнике расположен в середине запрещенной зоны. Концентрации свободных электронов и дырок в собственном полупроводнике равны

$$n_i = p_i = AT^{3/2} \exp(-\Delta E/2kT). \quad (3.9)$$

Из формулы видна сильная экспоненциальная зависимость концентраций  $n_i$  и  $p_i$  от температуры, а также от ширины запрещенной зоны при данной температуре. Концентрации электронов  $n$  и дырок  $p$  в примесных полупроводниках, выраженные через собственные концентрации  $n_i$  и  $p_i$ , можно получить, используя, вышеприведенные формулы:

$$n = n_i \exp[-(E_F - E_i)/kT] \quad (3.10)$$

и

$$p = p_i \exp[-(E_i - E_F)/kT]. \quad (3.11)$$

**Уровень Ферми в собственных и примесных полупроводниках.** В собственном полупроводнике, как уже было показано, уровень Ферми лежит в середине запрещенной зоны (рис. 3.3). При  $T > 0$  рождается «пара» (электрон и дырка). Концентрации электронов в зоне проводимости  $n$  и дырок в валентной зоне  $p$  в собственном полупроводнике равны:  $n_i = p_i$ .

Произведение концентраций электронов и дырок в собственном полупроводнике

$$n_i p_i = n^2_i = p^2_i = A^2 T^3 \exp(-\Delta E/kT) \quad (3.12)$$

не зависит от положения уровня Ферми и определяется только температурой  $T$  и шириной запрещенной зоны  $\Delta E$ .

В примесном полупроводнике  $n$ -типа (рис. 3.4) пятый валентный электрон донора слабо связан с ядром и становится свободным даже при слабых энергетических воздействиях. В результате введения доноров в запрещенной зоне появляется локальный уровень  $E_D$  вблизи зоны проводимости. При температуре  $T > 0$  в электронном полупроводнике происходят процессы генерации и рекомбинации носителей заряда, как и в собственном полупроводнике. Кроме того, большое число электронов с уровня донора  $E_D$  переходит в зону проводимости и намного увеличивает электропроводность вещества.

Энергия, необходимая для перевода электронов с уровней доноров в зону проводимости, называется энергией активации доноров  $E_{ак}$  электронного полупроводника. Например, для фосфора в кремнии  $E_{ак} \sim 0,044$  эВ.

Средняя тепловая энергия электронов при комнатной температуре  $kT = 0,025$  эВ. Поэтому можно утверждать, что при такой температуре все атомы доноров ионизированы, все электроны с донорных уровней переходят в зону проводимости.

В полупроводнике  $n$ -типа электронов в зоне проводимости оказывается больше, чем дырок в валентной зоне,  $n_n > p_n$  (зачерненные площадки на рис. 3.4), так как свободные носители заряда образуются не только в результате термогенерации как в собственном полупроводнике, где  $n_i = p_i$  (собственный механизм образования носителей), но и в основном за счет перехода электронов в зону проводимости с донорных уровней (примесный механизм образования носителей). Можно считать, что при нормальной температуре число образовавшихся дополнительных электронов примерно равно числу введенных атомов донора  $N_D$ .

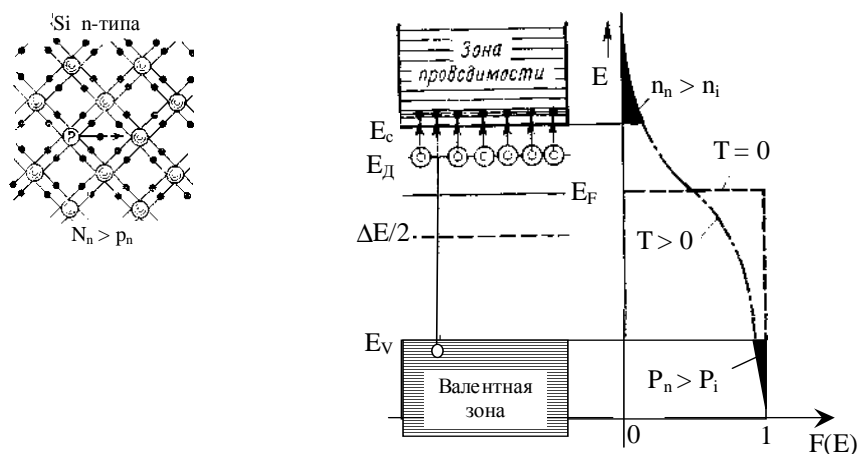


Рис. 3.4. Переход электронов с донорных уровней в зону проводимости

В полупроводнике  $n$ -типа электронов в зоне проводимости оказывается больше, чем дырок в валентной зоне,  $n_n > p_n$  (зачерненные площадки на рис. 3.4), так как свободные носители заряда образуются не только в результате термогенерации как в собственном полупроводнике, где  $n_i = p_i$  (собственный механизм образования носителей), но и в основном за счет перехода электронов в зону проводимости с донорных уровней (примесный механизм образования носителей). Можно считать,

что при нормальной температуре число образовавшихся дополнительных электронов примерно равно числу введенных атомов донора  $N_D$ .

Уровень Ферми смещается в электронном полупроводнике от середины запрещенной зоны вверх, к уровню доноров  $E_D$ , ближе к зоне проводимости. Чем больше концентрация доноров  $N_D$ , тем выше поднимается уровень Ферми. При высоких концентрациях  $N_D$  он может оказаться в запрещенной зоне выше уровня  $E_D$  или даже попасть в зону проводимости.

В примесном полупроводнике p-типа (рис. 3.5) в результате введения атомов акцепторной примеси в запрещенной зоне появляется локальный уровень акцепторов  $E_A$  вблизи валентной зоны. При нагревании монокристалла ( $T > 0$ ) происходит термогенерация носителей зарядов, сопровождаемая рекомбинацией. Кроме того, электрон валентной зоны, получив очень небольшую энергию  $E_{ак}$ , может быть захвачен атомом акцептора и перейти на уровень  $E_A$ , образовав в валентной зоне дырку.

Энергия  $E_{ак}$ , необходимая для перехода электрона верхнего уровня валентной зоны на уровень акцептора  $E_A$ , называется энергией активации дырочного полупроводника.

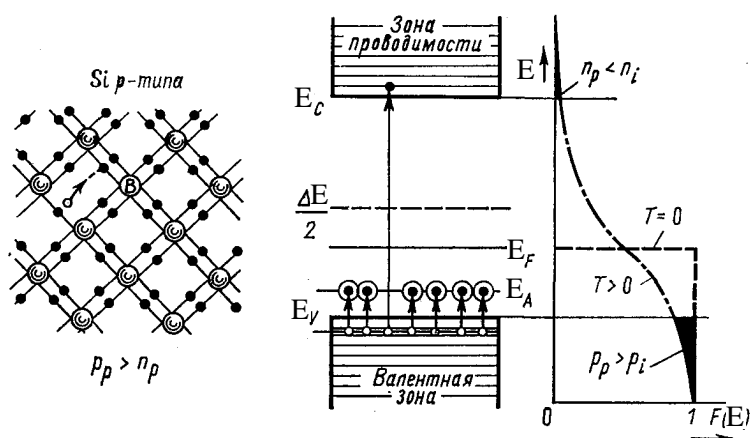


Рис. 3.5. В полупроводнике p-типа дырок в валентной зоне больше, чем электронов в зоне проводимости

Эта энергия невелика, например: для бора в кремнии  $E_{ак} = 0,046$  эВ. Значение  $E_{ак}$  сравнимо со средней тепловой энергией электронов при комнатной температуре ( $kT = 0,025$  эВ), и можно утверждать, что практически все атомы акцепторов ионизованы.

В полупроводнике p-типа дырок в валентной зоне оказывается больше, чем электронов в зоне проводимости,  $p_p > n_p$  (зачерченные площади на рис. 2.5), так как свободные носители заряда образуются не только в результате термогенерации как в собственном полупроводнике, где  $n_i = p_i$ , но и в основном за счет перехода электронов на акцепторные уровни. Можно считать, что при нормальной температуре число образовавшихся дополнительно дырок равно примерно числу введенных атомов акцептора.

Уровень Ферми смещается в дырочном полупроводнике вниз, к уровням акцептора  $E_A$ , ближе к валентной зоне. Чем больше концентрация атомов акцепто-

ров  $N_A$ , тем ниже опускается уровень Ферми. При высоких концентрациях  $N_A$  он может оказаться ниже уровня  $|E_A$  и даже сместиться в валентную зону.

Рассмотренные выше соотношения справедливы для концентраций примесей доноров  $N_D$  и акцепторов  $N_A$ , меньших  $10^{16} \text{ см}^{-3}$ . Однако существуют материалы, в которые вводится еще большее количество атомов примесей. Предельная концентрация примесей - порядка  $10^{19} \text{ см}^{-3}$ , т. е. один атом примеси на 1000 атомов основного вещества. Большого количества ввести невозможно без нарушения кристаллической решетки и резкого изменения свойств вещества (обычно в  $1 \text{ см}^3$  исходного материала, Ge или Si,  $10^{22}$  атомов).

Так как положение уровня Ферми зависит от концентрации примесей, то существует такое значение последней, когда уровень Ферми оказывается в разрешенной зоне. Это наступает при концентрации атомов примеси порядка  $\sim 5 \cdot 10^{18} - 10^{19} \text{ см}^{-3}$ . При такой высокой концентрации примеси уровень Ферми оказывается в зоне проводимости полупроводника n-типа или в валентной зоне полупроводника p-типа. Такие полупроводники называются вырожденными. Они используются в качестве исходного материала для изготовления некоторых типов полупроводниковых приборов. По своим свойствам вырожденные полупроводники приближаются к металлам и называются поэтому иногда полуметаллами.

Положение уровня Ферми зависит также от температуры монокристалла. С ростом температуры уровень Ферми примесного полупроводника приближается к середине запрещенной зоны. Любой примесный полупроводник при нагревании приближается к собственному, поскольку концентрации подвижных носителей обоих знаков сравниваются.

### Свойства p-n контактов

Электронно-дырочный переход схематически представлен на рис. 3.6. Поверхность, по которой контактируют слои p и n, называется металлургической границей, а прилегающая к ней область объемных зарядов – собственно p – n переходом. При контакте p- n-полупроводников электроны из полупроводника n-типа диффундируют в полупроводник p-типа, а дырки в обратную сторону до тех пор, пока уровни их химических потенциалов не выравняются. На контакте при этом появится разность потенциалов  $\phi_0$  (рис. 3.6). Высота потенциального барьера p – n перехода определяется контактной разностью потенциалов

$$|q| \phi_0 = E_{\phi n} - E_{\phi p},$$

где  $E_{\phi n}$  и  $E_{\phi p}$  – высота уровней Ферми. Для нахождения распределения потенциала и напряжения поля в области объемного заряда необходимо решать уравнение Пуассона.

Переход бывает ступенчатым, резким, с идеальной границей и постоянной концентрацией доноров  $N_D$  и акцепторов  $N_A$  по обе стороны

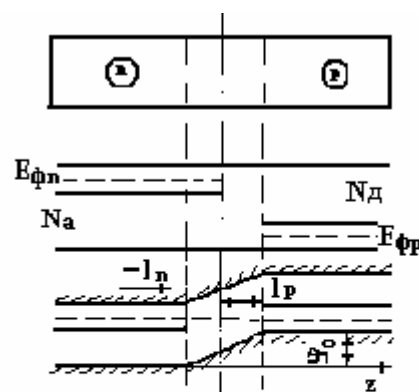


Рис. 3.6. Появление разности потенциалов  $\phi_0$  при контакте p-n полупроводников

металлургической границы и плавным, когда концентрация одного типа примеси постепенно уменьшается, другой растет.

Для резкого  $p - n$  перехода будем предполагать что справа от границы (рис. 3.6) на расстоянии  $z = l_p$  заряд равен  $qN_a$ , а слева на расстоянии  $z = -l_n$  заряд равен  $qN_d$ . За точку отсчета принимаем плоскость  $z = 0$ . При  $z = l_p$  потенциал равен контактной разности  $\phi_0$ , а электрическое поле отсутствует; при  $z = -l_n$  как потенциал, так и электрическое поле равны нулю. При этом толщина слоя объемного заряда равна:

$$l = l_n + l_p = \left( \frac{2\varepsilon_0\varepsilon \cdot \phi_0}{|q|} \frac{N_a + N_d}{N_a N_d} \right)^{1/2}. \quad (3.13)$$

Для плавного  $p - n$  перехода с линейным распределением примесей  $N(z)=az$ , где  $a$  – градиент концентрации примеси. Толщина слоя объемного заряда при этом равна:

$$l = l_n + l_p = \left( \frac{12\varepsilon_0\varepsilon \cdot \phi_0}{|q|a} \right)^{1/2}. \quad (3.14)$$

Итак, в области  $p - n$  перехода имеется объемный заряд доноров и акцепторов (положительный в  $n$ -области и отрицательный в  $p$ -области). Подвижные носители почти все ушли из перехода, поэтому его сопротивление очень велико.

Если к переходу приложить напряжение таким образом, чтобы объемный заряд увеличился (для этого надо к  $n$ -области приложить положительное напряжение), то сопротивление перехода еще более увеличится. Такое состояние перехода называется запертым, а соответствующее напряжение – обратным (обратно смещенный переход). При обратной полярности приложенного напряжения объемный заряд в переходе уменьшится и сопротивление перехода тоже. Такое состояние перехода будет называться открытым, а напряжение – прямым (прямо смещенный переход).

В отсутствие внешнего напряжения в условиях равновесия через переход в противоположных направлениях будут протекать равные токи носителей. Эти токи состоят из тока проводимости – дрейфовой компоненты, зависящей от электрического поля  $E$  и диффузионного тока, определяемого концентрацией носителей. Диффузионный ток – дырочный и электронный – обусловлены различием концентраций носителей заряда в  $p$ - и  $n$ - областях и уравниваются током проводимости. Диффузионный дырочный ток из  $p$ -области в  $n$ - область определяется концентрацией дырок в  $p$ -области, равной концентрации акцепторов  $N_a$ . Диффузионный электронный ток из  $n$ -области в  $p$ - область определяется концентрацией электронов ней, равной концентрации доноров  $N_d$ . В условии равновесия дырочный ток проводимости и электронный ток проводимости равны.

Равновесие этих токов нарушается, если на барьер подать напряжение. Токи проводимости от полярности напряжения не зависят. Диффузионные токи существенно зависят от напряжения, которое либо вычитается (при прямом на-

пряжении), либо складывается (при обратном напряжении) с величиной  $\varphi_0$  в экспоненте. Разница токов (диффузионных и проводимости) будет также изменяться экспоненциально.

Импульсные (высокочастотные свойства) р-п переходов в запертом состоянии определяются барьерной емкостью, которая образуется на слое пространственного заряда толщиной  $l$ : для резкого р-п перехода ( $N_a \gg N_d$ ) практически весь слой объемного заряда, находится в области:

$$C_b = \left[ \frac{\varepsilon_0^2 \varepsilon^2 |q| N_d}{2(\varphi_0 - U)} \right]^{1/2} \cdot S; \quad (3.15)$$

а для плавного перехода:

$$C_b = \left[ \frac{\varepsilon_0^2 \varepsilon^2 |q| N_d}{12(\varphi_0 - U)} \right]^{1/3} \cdot S. \quad (3.16)$$

Барьерная емкость зависит от разности потенциалов  $U$ . С увеличением напряжения, приложенного в обратном направлении ( $\varphi_0 + |U|$ ), барьерная емкость уменьшается. Это уменьшение емкости обусловлено тем, что с увеличением обратного напряжения возрастает среднее расстояние  $l$  слоя объемного заряда.

Импульсные свойства в открытом состоянии р-п контакта определяются диффузионной емкостью, обусловленной диффузией неосновных носителей в р-п переход:

$$C_D = \frac{dQ}{dU} = \frac{q \cdot l}{k_B \cdot T} \tau, \quad (3.17)$$

где  $Q$  – заряд инжектированных неосновных носителей, который определяется током через переход и временем жизни носителей  $\tau$ ,  $k_B$  – постоянная Больцмана. Диффузионная емкость, как и барьерная, ухудшают высокочастотные свойства р-п перехода, применяемого в конкретном приборе.

### Свойства контакта металл-полупроводник

В приборах с таким контактом осуществляется выпрямление слабых токов путем контакта металлической иглы с кристаллом полупроводника. Точечноконтактные диоды послужили позже основой для создания целого направления приборов, известных под названием приборов с контактом Шоттки. Структура и свойства контакта металл – полупроводник существенно зависят от соотношения термодинамических работ выхода металла -  $E_{\text{фм}}$  и полупроводника -  $E_{\text{фп}}$ . В случае  $E_{\text{фм}} > E_{\text{фп}}$  электроны переходят из полупроводника в металл до тех пор, пока возникающая контактная разность потенциалов не выровняет уровень Ферми металла и полупроводника.

Используя уравнение Пуассона, можно определить распределение потенциала и электрического поля вдоль контакта Шоттки. Толщина обедненного слоя полупроводника, получившего название запорного, равна:

$$l_n = \left( \frac{2\varepsilon_0 \varepsilon \cdot \varphi_0}{|e| n_0} \right)^{1/2}, \quad (3.18)$$

где  $n_0$  – концентрация электронов в полупроводнике.

Если к контакту приложить внешнюю разность потенциалов, повышающую высоту барьера (обратное напряжение – минус на металле), ширина обедненного слоя растет, сопротивление контакта увеличивается. Если внешнее напряжение приложено так, что барьер понижается (прямое напряжение – плюс на металле), то ширина обедненного слоя тоже уменьшается. В этом случае будет уменьшаться сопротивление обедненного слоя. На этом явлении основаны выпрямляющие свойства подобного контакта. Контакт Шоттки является основой многих устройств, используемых в технике СВЧ и в низкочастотной технике (радиотехнике, телевизионной и вычислительной технике).

### 3.2. Лавинно-пролетные диоды (ЛПД)

Механизм работы ЛПД основан на эффекте динамического отрицательного сопротивления, возникающего при лавинном пробое диода в результате ударной ионизации и последующего пролета носителей через запирающий слой. ЛПД-диоды (p-n, p-i-n, n-p-i-p - диод Рида) работают в диапазоне 0,5-500 ГГц. В сантиметровом диапазоне КПД ЛПД колеблется до 20-30 %, в дециметровом 60-70 %. ЛПД применяются в качестве генераторов, усилителей СВЧ, по мощности сравнимых с маломощными клистродами, и ЛБВ. Лавинный пробой существует при напряженностях поля порядка  $10^5 - 10^6$  В/см. Применение различных типов полупроводниковых диодов связано с нелинейным сопротивлением, нелинейной емкостью, дифференциальным сопротивлением и т. д.

Так на основе ЛПД, где проявляется эффект отрицательного дифференциального сопротивления, разработан ряд устройств (генераторов и усилителей) средней и малой мощности, перекрывающих весь диапазон СВЧ от 0,3 до 3000 ГГц. А на основе эффекта Ганна созданы сугубо твердотельные приборы СВЧ, работающие в различных режимах и являющиеся наряду с ЛПД перспективным классом приборов коротковолновой области диапазона СВЧ.

**Ударная ионизация.** При превышении носителями заряда определенной величины энергии, начнется ударная ионизация. Это означает образование пары новых носителей – электрона и дырки, при соударении носителя с нейтральным атомом.

Ударная ионизация характеризуется коэффициентами ударной ионизации -  $\alpha_p$  и  $\alpha_n$ , показывающими сколько электрон – дырочных пар создается на единице пути (1см) электроном и дыркой соответственно. На рис. 3.7 показано как меняются коэффициенты ионизации для разных материалов в зависимости от напряженности электрического поля, а выражение (3.19) дает представление о величине этих коэффициентов.

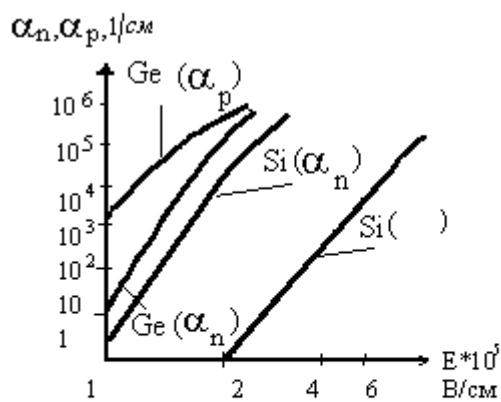


Рис. 3.7. Зависимости коэффициента ионизации от напряженности электрического поля



$$\alpha = A \cdot e^{\left(\frac{E}{G}\right)^m}, \text{ где} \quad \begin{array}{l} m(\text{Ge, Si}) = 1 \\ m(\text{GaAs}) = 2 \end{array} \quad (3.19)$$

Коэффициент ионизации  $\alpha$  достигает порядка  $10^5 \text{ см}^{-1}$ .

**Лавинное умножение носителей в p-n переходе.** Ударная ионизация при высокой напряженности поля вызывает рост обратного тока перехода, что вызывает лавинный пробой. Для определения увеличения токов в переходе по отношению к начальному состоянию вводят коэффициент умножения дырок или электронов. рис. 3.8 показано распределение дырочного и электронного токов в p-n переходе. Дырки, шедшие из n-области, двигаются по направлению к p-области, а электроны шедшие из p-области, перемещаются к n-области. На рисунке нанесены значения токов начальный и конечный момент движения. Вследствие ударной ионизации число движущихся дырок растет слева направо, а электронов наоборот.

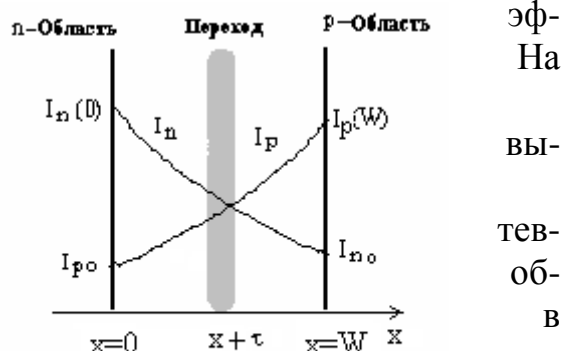


Рис. 3.8. p-n переход

Введем следующие обозначения: пусть -

$I_p/e$  - число дырок, проходящих через  $dx$  в 1 сек;

$I_n/e$  - число электронов, проходящих через  $dx$  в 1 сек

$\alpha_p dx$  - число электронно-дырочных пар, создаваемых 1 дыркой на пути  $dx$ .

$\alpha_n dx$  - число электронно-дырочных пар, создаваемых одним электроном на пути  $dx$ . Увеличение числа дырок на пути  $dx$ .

$$\frac{dI_p}{e} = \frac{I_p}{e} \alpha_p dx + \frac{I_n}{e} \alpha_n dx,$$

где  $dI_p = I_p \alpha_p dx + I_n \alpha_n dx$  - рост дырочного тока или

$$\frac{dI_p}{dx} = I_p \alpha_p + I_n \alpha_n \quad ; \quad \frac{dI_p}{dx} - I_p (\alpha_p - \alpha_n) = I \alpha_n. \quad (3.20)$$

$I = I_p + I_n$  - полный ток в переходе, тогда

$$-\frac{dI_n}{dx} - I_n (\alpha_n - \alpha_p) = I \alpha_p \quad (3.21)$$

Граничными условиями при решении дифференциальных уравнений (3.20) и (3.21) являются токи:

$$\begin{array}{ll} \text{при } x = 0 & I_{p0} \text{ и } I_n(0), \\ \text{при } x = W & I_{n0} \text{ и } I_p(W) \end{array}$$

где  $I_{p0}$  и  $I_{n0}$  – начальный дырочный и электронный токи перехода.

$$M_p = I_p(W)/I_p(0), \quad \text{при } I_{p0} \gg I_{n0}$$

$M_n = I_n(W)/I_n(0)$ , где  $M$  – коэффициент лавинного умножения.

Из (3.20) получаем уравнение, характеризующее лавинный пробой:

$$1 - \frac{1}{M_p} = \int_0^W \alpha_p \exp\left[-\int_0^x (\alpha_p - \alpha_n) dx'\right] dx. \quad (3.22)$$

**Лавинный пробой.** Лавинный пробой наступает при таком обратном напряжении на переходе, когда  $M$  стремится к  $\infty$ . Если  $I_{p0} \gg I_{n0}$  то из (3.22)

$$\int_0^W \alpha_p \exp\left[-\int_0^x (\alpha_p - \alpha_n) dx'\right] dx = 1 \quad \text{– условие лавинного пробоя,}$$

$$\text{если } I_{n0} > I_{p0}, \text{ то } \int_0^W \alpha_n \exp\left[-\int_0^x (\alpha_n - \alpha_p) dx'\right] dx = 1$$

$$\text{если } \alpha_n = \alpha_p, \text{ то } \int_0^x \alpha \cdot dx = 1. \quad (3.23)$$

Здесь  $\alpha$  является эффективным коэффициентом ионизации.

Каждый электрон или дырка создают в среднем по электронно-дырочной паре.

**Пролетный режим работы ЛПД.** В ЛПД чаще всего используется структура типа  $n^+ - p - i - p^+$  (диод Рида), изображенная на рис. 3.9, а.

Распределение концентрации примесей в областях структуры показано на рис. 3.9, б; напряженность поля на рис. 3.9, в;

коэффициент ионизации на рис. 3.9, г. Слой, где возможно лавинное умножение носителей из-за высокой напряженности поля, очень узкий ( $0 - x_L$ ). На участке  $0 - x_L$  в результате лавинного умножения должно появиться 95% носителей, возникающих в процессе лавинного умножения

$$\int_0^{x_L} \alpha(x) dx = 0,95.$$

Часть структуры от  $x = x_L$  до границы между  $i$ - областью и  $p^+$ -областью называют слоем дрейфа ( $x_L - W$ ).

**Принцип работы ЛПД.** Принцип работы рассмотрим с помощью пространственно-временной диаграммы, показанной на рис. 3.10

Там, где в слое умножения и дрейфа имеется большая напряженность поля, изменяющаяся по синусоидальному закону, образуется лавина свободных зарядов – электроны и дырки (сгусток носителей). Этот сгусток, запаздывающий на четверть периода от максимума поля (т.к. для развития лавины требуется время), создает короткий импульс лавинного тока  $i_{л}(t)$  на границе  $x=x_{л}$  между слоями умножения и дрейфа. Далее носители двигаются в электрическом поле слоя дрейфа, пока не достигнут его границы ( $x=W$ ) при угле пролета  $\theta_{др}$ . Если  $\theta_{др} < \pi$ , то электронный пакет в момент дрейфа находится в тормозящем полупериоде поля и от-

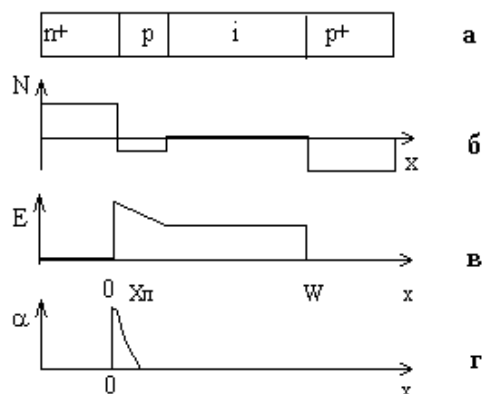


Рис. 3.9. Структура диода p-i-p-типа (а), распределение концентрации (б), напряженности поля (в), и коэффициента ионизации (г) вдоль оси x.

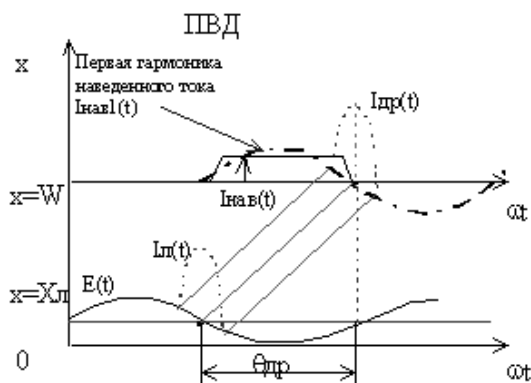


Рис. 3.10. Пространственно-временная диаграмма ЛПД

Дрейфовая скорость носителей в промежутке  $x_{л} - W$  постоянна, и поэтому обозначена прямыми линиями.  $\theta_{др} = \omega \tau = 2\pi f \tau$   
 $\theta_{др} < \pi$  – электроны всю энергию отдают полю  
 $\theta_{др} > \pi$  – электроны частично забирают энергию от поля  
 $\theta_{др} = 2\pi$  – эффект взаимодействия исчезает

дает свою энергию полю. В другие моменты носители либо отбирают энергию поля при  $\theta_{др} > \pi$ , либо вообще не взаимодействуют с полем при  $\theta_{др} = 2\pi$ . На рис.2.4 показано, что когда носители достигают границы  $x=W$ , появляется дрейфовый ток  $i_{др}(t)$  в виде узкого импульса, повторяющего  $i_{л}(t)$ . Одновременно изображена кривая наведенного тока  $i_{нав}(t)$ , созданного движением короткого сгустка на пути от  $x=x_{л}$  до  $x=W$ . На рис. 3.10 изображена также первая гармоника  $i_{нав1}(t)$  этого тока. Наведенный ток протекает во внешней цепи в отрицательный полупериод СВЧ-напряжения, что эквивалентно появлению отрицательной электронной проводимости в ЛПД ( $G_3^{-1}$ ). Максимальное значение  $G_3^{(-1)}$  имеет место при  $\theta_{др} \approx \pi$ , поэтому оптимальная длина области дрейфа связана с частотой СВЧ-

колебаний соотношением  $l_{др} \approx \frac{v_{нас}}{(2f)}$ , а скорость электронов становится равной

$v_{нас} \approx 10^5$  м/с. По величине амплитуды 1-ой гармоники  $I_{нав}$  и амплитуде поля можно найти мощность, передаваемую потоком носителей.

**Токи проводимости в слое умножения.** Для нахождения токов проводимости нужно определить, сколько электронно-дырочных пар сможет образоваться в данной структуре

$$S \frac{dp}{dt} = -\frac{1}{e} \frac{dI_p}{dx} + S\alpha V(n+p)$$

$$S \frac{dn}{dt} = -\frac{1}{e} \frac{dI_n}{dx} + S\alpha V(n+p) \quad (3.24)$$

I            II            III

$V$  – скорость электронов и дырок,  $S$  – площадь сечения структуры,  
 $I$  – изменение количества носителей в слое длиной 1 см за 1 сек,  
 $II$  – уменьшение числа носителей за 1 сек из-за прохождения тока,  
 $III$  – сколько пар носителей образуется за 1 сек электронами и дырками,  
 $I_p$  – дырочная составляющая и  $I_n$  – электронная составляющая тока.

Из уравнения непрерывности для слоя умножения (3.24) с учетом граничных условий, получаем конечное выражение для тока проводимости:

$$\tau_{л} \frac{dI}{dt} = -[I_p - I_n]_{0}^{x_{л}} + 2I \int_0^{x_{л}} \alpha dx, \quad (3.25)$$

где,  $I$  – суммарный ток. Он равен сумме электронного и дырочного токов ( $I = I_p = I_n$ ).

При большой  $E$ , характерной для лавинного пробоя, скорость носителей равна скорости насыщения – максимально возможной скорости носителей в полупроводниках ( $V_n = 10^7$  см/с – скорость насыщения). Поэтому время пролета носителей в слое умножения

Проанализируем значение тока проводимости для разных сечений полупроводниковой структуры.

В сечении  $x = 0$  при  $I(0) = I_p(0) = I_n(0)$  получаем:

$$I_p(0) - I_n(0) = 2I_p(0) - I_n(0) = 2I_{p0} - I(0) \quad (3.26)$$

$$\text{При } x = x_{л} \quad I_p(x_{л}) - I_n(x_{л}) = -2I_{n0} - I(x_{л}). \quad (3.27)$$

С учетом принятых условий получим дифференциальное уравнение для тока:

$$\frac{\tau_{л}}{2} \frac{dI}{dt} = I \left( \int_0^{x_{л}} \alpha dt - 1 \right) + I_0, \quad \text{где } I_0 = I_{p0} + I_{n0}, \quad t_{л} = \frac{l_{л}}{V_{нас}}, \quad (3.28)$$

где  $\tau_{л}$  – время пролета носителей заряд в слое лавинного умножения.

Условием пробоя является выражение, аналогичное (3.23).

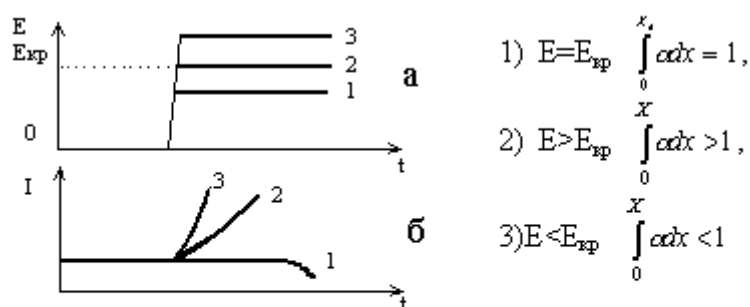


Рис. 3.11 Зависимости напряжения (а) и тока (б) от времени

$$\int_0^{x_1} \alpha dx = 1, \text{ при котором } I \rightarrow \infty.$$

Рассмотрим решения уравнения (3.28) при небольшом скачкообразном изменении поля. На рис. 3.11,а показаны изменения напряженности поля. Кривая 1 показывает асимптотическое убывание тока, кривая 2 соответствует значению тока равному  $I = I_0(1 + 2t/\tau_{л})$ , а кривая 3-его - рост по экспоненциальному закону. Во всех случаях наблюдается запаздывание тока относительно скачка напряженности поля, т.к. необходимо какое-то время для получения определенного числа носителей.

**Токи в слое умножения при малой амплитуде поля.** Здесь мы проанализируем конечное значение этих токов (рис. 3.11,б).

Режим малой напряженности электрического поля в слое умножения (режим слабого сигнала или малых амплитуд) обычно используется для анализа работы структур. Дается малое приращение к выбранным параметрам, как это демонстрируется ниже:

$$\left. \begin{aligned} E &= E_{др} + \dot{E}_{л} \exp(j\omega t) \\ \bar{\alpha} &= \bar{\alpha}_0 + \dot{\alpha}_{л} \exp(j\omega t) \\ I &= I_0 + \dot{I}_{л} \exp(j\omega t) \end{aligned} \right\} \begin{aligned} |\dot{E}_{л}| &\ll E_{др} \\ |\dot{\alpha}| &\ll \bar{\alpha}_0 \\ |\dot{I}_{л}| &\ll I_0 \end{aligned},$$

где  $-\dot{\alpha}_{л}$  и  $\dot{I}_{л}$  - комплексные амплитуды соответствующих величин. Коэффициент умножения линеаризуем по переменному напряжению  $\dot{\alpha}_{л} \approx \frac{d\alpha}{dE} \dot{E}_{л} = a' \dot{E}_{л}$ , где  $E_{л}$  - амплитуда переменной составляющей поля.

Если ширина слоя умножения  $x_l$  не- большая, то мал и переменный заряд в этом слое. Поэтому напряженность одинакова по всему слою  $\dot{U}_l = \dot{E}_l x_l$ , тогда полный ток слоя умножения представляется как

$$\dot{I}_{\text{полн}} = \dot{I}_l + \dot{I}_{\text{см}} = \left( \frac{1}{j\omega L_l} + j\omega C_l \right) \cdot \dot{U}_l,$$

где

$$\dot{I}_l = \frac{2I_0}{j\omega\tau_l} \alpha' \dot{U}_l = \frac{\dot{U}_l}{j\omega L_l}, \quad (3.29)$$

а лавинная индуктивность  $L_l = \frac{\tau_l}{2I_0\alpha'}$  – эквивалентна индуктивности, т.к. ток отстает от напряжения на  $90^\circ$ . При решении был учтен средний коэффициент ионизации  $\bar{\alpha} = \frac{1}{x_l} \int_0^{x_l} \alpha dx$  с условием пробоя  $\bar{\alpha} x_l = 1$ .

Полученный результат является следствием инерционности процесса образования носителей. Когда поле пройдет максимальное значение и начнет уменьшаться, концентрация носителей еще продолжает возрастать. Максимумы концентрации носителей и тока достигаются тогда, когда переменная составляющая поля проходит через нулевое значение (отставание по фазе на  $90^\circ$  рис. 3.12).

С увеличением напряжения на слое умножения предположение о малости сигнала перестает выполняться. Изменение тока по времени будет сильно отличаться от синусоидального закона и начнет приобретать импульсный характер. При этом максимумы импульсов тока практически не смещаются по отношению к максимальным значениям тока при слабом сигнале как показано на рис. 3.12. Поэтому слой умножения можно рассматривать как источник импульсов тока, запаздывающих по отношению к максимальному значению напряжения на четверть периода.

Токи смещения в слое умножения, иначе их представляют как емкостной ток, равен

$$I_{\text{см}} = \varepsilon S \frac{\partial E}{\partial t} = j\omega C_l \dot{U}_l, \quad (3.30)$$

где  $C_l = \varepsilon S / x_l$  – эквивалентная емкость слоя умножения, а  $S$  – площадь сечения, выбранной структуры.

Соответственно полная проводимость слоя умножения:

$$Y_l = \dot{I}_{\text{полн}} / \dot{U}_l = \frac{1}{j\omega L_l + j\omega C_l}. \quad (3.31)$$

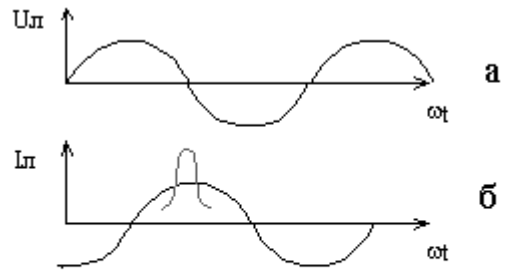


Рис. 3.12. Фазовое изменение напряжения (а) и тока (б).

Слой умножения может быть представлен эквивалентной схемой (рис. 3.13) параллельным колебательным контуром, содержащим индуктивность ( $L_{\pi}$ ) и емкость ( $C_{\pi}$ ).

Сопротивление слоя из выражения (3.31) будет равно

$$Z_{\pi} = \frac{1}{Y_{\pi}} = \frac{1}{j\omega C_{\pi}} \left(1 - \frac{\omega_{\pi}^2}{\omega^2}\right) \quad \omega_{\pi} = \frac{1}{\sqrt{L_{\pi} C_{\pi}}}; \quad f_{\pi} = \frac{\omega_{\pi}}{2\pi}$$

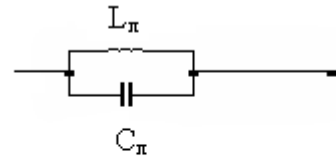


Рис. 3.13. Эквивалентная схема слоя умножения

где  $\omega_{\pi}$  – собственная резонансная частота контура, называемая круговой лавинной частотой, а  $f_{\pi}$  – линейная частота;  $V_H$  – скорость насыщения;  $j_0$  – плотность тока  $j_0 = I_0/S$ .

**Токи в слое дрейфа при малой амплитуде поля.** Рассмотрим теперь составляющие полного тока в слое дрейфа и его импеданс.

В случае малого сигнала плотность переменного заряда незначительна не только в слое умножения, но и в слое дрейфа. Поэтому амплитуда напряженности переменного поля остается практически неизменной и равной  $\dot{E}_{\pi}$  в слое умножения. Тогда падение напряжения на слое дрейфа  $\dot{U}_{др} = \dot{E}_{\pi} (W - x_{\pi})$ .

В слое дрейфа напряженность поля меньше критической величины  $E_{кр}$ , но достаточно велика, чтобы скорость носителей оставалась практически равной скорости насыщения, т.е.

$$\text{при } V \cong V_H, \quad E_{др} < E_{кр}$$

Полный ток тогда записывается как сумма наведенного и емкостного токов:

$$I_{\text{полн}} = I_{\text{нав}} + I_{\text{емк}}. \quad (3.32)$$

Время пролета в слое умножения определяется из выражения  $\tau_x = \frac{x}{V_H}$ .

Время пролета в слое дрейфа равно  $t_{др} = \frac{l_{др}}{V_H}$ .

Для нижней части сантиметрового диапазона длин волн  $l_{др} = 3-10$  мкм,  $C_{др} = \epsilon S (W - x_{\pi})$  – емкость слоя дрейфа.

Импеданс слоя дрейфа тогда будет равен:

$$Z_{др} = \frac{\dot{U}_{др}}{I_{\text{полн}}}. \quad (3.33)$$

Эквивалентная схема слоя дрейфа и слоя умножения изображена на рис. 3.14.

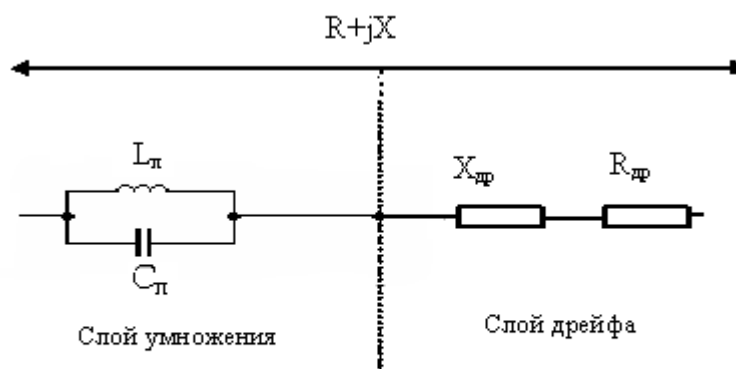


Рис. 3. 14. Эквивалентная схема слоя умножения и слоя дрейфа ЛПД.

Как видно из рис. 3.14 импеданс слоя дрейфа равен сумме  $X_{др}$  и  $R_{др}$   
 $Z_{др} = R_{др} + jX_{др}$ , где  $X_{др}$  и  $R_{др}$  записаны как

$$R_{др} = \frac{1}{\omega C_{др}} \cdot \gamma \left( \frac{1 - \cos \theta_{др}}{\theta_{др}} \right) \quad X_{др} = \frac{1}{\omega C_{др}} \left( -1 + \frac{\gamma \sin \theta_{др}}{\theta_{др}} \right). \quad (3.34)$$

Зависимости  $X_{др}$  и  $R_{др}$  от  $f$  и от  $\theta$ , вычисленные по (3.34) представлены на рис. 3.15. Активное сопротивление слоя дрейфа отрицательно на всех частотах  $\omega > \omega_{л}$  ( $\gamma < \theta$ ), кроме частот, на которых  $R_{др} = 0$  (при  $\theta_{др} = \omega \tau_{др} = 2\pi n$ ;  $n=1, 2$  и т.д.). Максимум отрицательного сопротивления наблюдается вблизи  $\theta_{др} = \pi$ . При дальнейшем увеличении  $\theta_{др}$  до  $3R_{др}/2\pi$  уменьшается до нуля. Таким образом, диапазон частот, в котором может быть обеспечено отрицательное сопротивление, велик. Однако ЛПД хорошо работает на частотах, соответствующих углу пролета  $\theta_{др} = \omega \tau_{др} = \pi$ . Из этого следует, что для узкого слоя умножения ( $x_{л} \ll 2W$ ) можно определить значение частоты, называемой **пролетной частотой**:

$$f_{пр} = \frac{\omega_{пр}}{2\pi} = \frac{1}{2\tau_{др}} = \frac{V_{н}}{2W}.$$

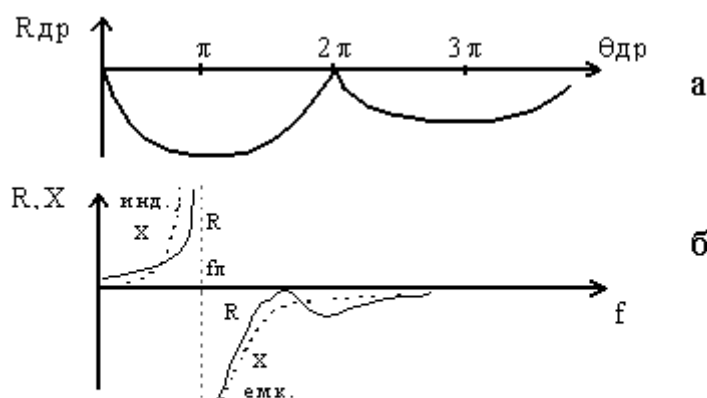


Рис. 3.15. Зависимости  $R_{др}$  от  $\theta_{др}$  (а),  
 $R$  и  $X$  от частоты  $f$  (б).



Итак, как показано на рис. 3.15, видно, что контур  $L_{л}, C_{л}$  характеризует процессы, происходящие в слое умножения, а  $X_{др}$  и  $R_{др}$  – процессы в слое дрейфа. Полное сопротивление ЛПД:  $Z=R+jX$ . Из рис. 3.15 видно, что на частотах выше  $f_{л}$  активное сопротивление отрицательно, а реактивное становится емкостным. Наличие отрицательного сопротивления и позволяет использовать ЛПД для создания генераторов и усилителей СВЧ.

Параметры и характеристики генераторов и усилителей на ЛПД в пролетном режиме.

На основе ЛПД можно создать СВЧ–генераторы (ГЛПД) рис. 3.16 и усилители (УЛПД) рис. 3.17. В обоих случаях ЛПД включен в колебательную систему. Генерацию или усиление объясняют наличием отрицательного сопротивления.

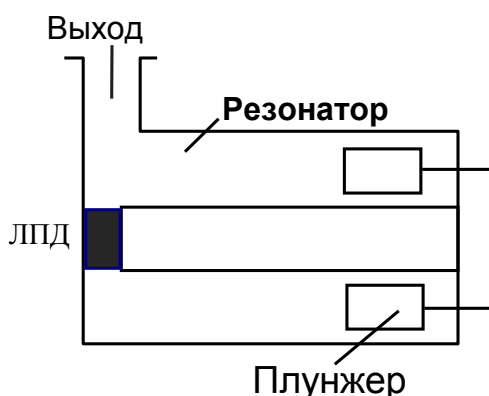


Рис. 3.16. ГЛПД

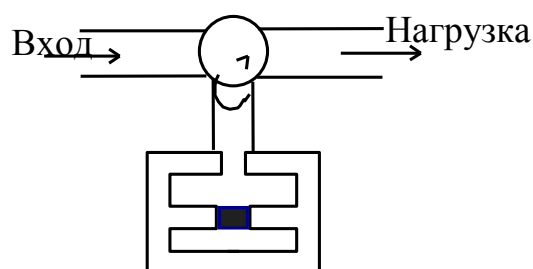


Рис. 3.17. УЛПД

Отрицательное сопротивление зависит от тока ЛПД, поэтому в ГЛПД с заданным сопротивлением нагрузки, резонатора и областей диода, существует минимальный пусковой ток, начиная с которого возможна генерация. При токе менее пускового, обеспечивается усилительный режим.

Схема ГЛПД содержит коаксиальный резонатор, перестраиваемый с помощью поршня и ЛПД, помещенного между центральным проводником резонатора и его торцом. ЛПД находится в пучности высокочастотного поля.

На рис. 3.17 показан простейший УЛПД. Источник сигнала и нагрузка включены в два плеча циркулятора. К третьему плечу присоединена колебательная система ЛПД. Усиление сигнала происходит в результате его взаимодействия с колебательной системой, имеющей отрицательную добротность, и может рассматриваться как поступление в нагрузку через циркулятор отраженного усиленного сигнала.

**Шумы.** Шумы имеют разные источники их возникновения: флуктуации лавинного тока ( $\sim 20$ дБ), флуктуации частиц в лавинах, образованных каждым носителем. Для уменьшения шума надо работать при частотах  $> f_{л}$ , но при малых токах, а это приводит к малому КПД.

Минимальные шумы у GaAs (20 дБ), максимальные - у Si. (Генераторы шума).

**Выходная мощность ГЛПД.** Мощность  $P_{0\text{ макс}}$  ГЛПД может быть оценена из соотношений, определяемых формулой

$$P_{0\text{ макс}} = \varepsilon S V_n E_{\text{кр}}^2 \quad (3.35)$$

Выходная мощность не может быть больше потребляемой. Поэтому для оценки максимальной выходной мощности рассматривают максимальную потребляемую диодом мощность.

$$P_{0\text{ макс}} = I_{0\text{ макс}} U_{0\text{ макс}},$$

где  $I_{0\text{ макс}}$   $U_{0\text{ макс}}$  - максимально допустимое и напряжение на диоде и соответствующий ему максимальный ток. Очевидно, что самое большое значение  $U_{0\text{ макс}}$  будет в том случае, когда на всей длине перехода  $W$  поле будет равным критическому  $U_{0\text{ макс}} = E_{\text{кр}}$ .

Конечное значение для максимальной мощности запишем как:

$$P_{0\text{ макс}} \cdot f^2 = V_n^2 E_{\text{кр}}^2 / 8\pi X_c \quad (3.36)$$

Условие (3.36) означает, что при постоянном сопротивлении прибора  $X_c$  потребляемая мощность обратно пропорциональна квадрату частоты:

$$P_{0\text{ макс}} \cdot f^2 = \text{const}. \quad (3.37)$$

$$\text{На более низких частотах} \quad P_{0\text{ макс}} \cdot f = \text{const}. \quad (3.38)$$

На рис. 3.18 показана расчетная зависимость мощности от частоты.

Плотность тока ЛПД не превышает нескольких сот ампер на  $\text{см}^2$ .

Возбуждение СВЧ-колебаний в генераторе происходит, когда отрицательное сопротивление ЛПД превысит суммарные потери в системе, включая потери в диоде, контуре и нагрузке  $-R_{\text{ном}}$ . Это имеет место при токах диода, превышающих значение пускового тока:

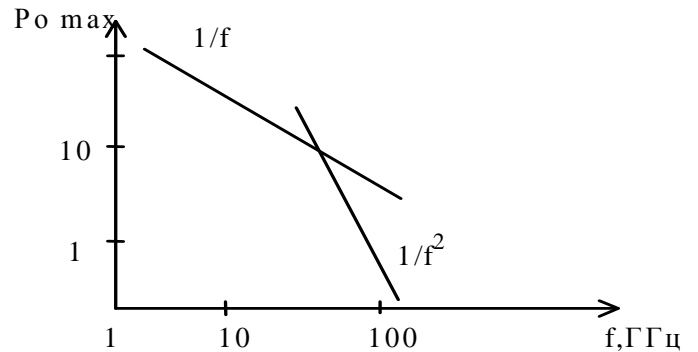


Рис. 3.18

$$I_{\text{пуск}} = \frac{\pi}{\alpha'} \frac{(\omega C)^2}{\omega C + \frac{\theta_{\text{др}} R_{\text{пот}}}{1 - \cos \theta_{\text{др}}}}, \quad (3.39)$$

где  $C = \frac{\varepsilon \varepsilon_0 s}{l}$ ,  $l$  и  $s$  - толщина и площадь полупроводниковой структуры;

$\theta_{\text{др}} = \omega \tau_{\text{др}}$  - угол пролета электронов в области дрейфа.

Выходная СВЧ-мощность  $P_{\text{вых}} = 0,5 U_m I_{10}$ , и  $\eta_3 \approx \frac{2U_m}{(\pi U_0)}$ .

Амплитуда переменной составляющей электрического поля:

$$U_m < U_{др}, U_m \leq \left(\frac{1}{4} - \frac{1}{2}\right) U_0,$$

где  $U_0 = U_l + U_{др}$ , состоящей из суммы  $U_{др}$  - падение напряжения на дрейфовом слое,  $U_l$  - падение напряжения в лавинном слое.

Мощность в пролетном режиме на частотах 4-12 ГГц  $\sim$  5 Вт – непрерывный режим. Однопролетный режим:

на частоте	50 ГГц	$P_{вых} = 0,5$ Вт,	$\eta = 10\%$ ,
на частоте	92 ГГц	$P_{вых} = 0,15$ Вт,	$\eta = 7\%$ .

Коэффициент усиления при каскадном включении  $\sim$  30 дБ (но шумы больше, поэтому во входных устройствах ЛПД усилители не применяют).

**Электронный КПД.** Теоретически максимальный электронный КПД генератора на ЛПД составляет  $\sim$  30 %. КПД уменьшается при  $U_l \sim U_{др}$ , т.к. должно быть  $U_l \ll U_{др}$

$$U_l \approx \left(\frac{1}{3} \div 1\right) U_{др}, \text{ поэтому КПД} - 10-15 \%$$

### 3.3. Диоды Ганна

Диоды Ганна относят к классу приборов с объемной неустойчивостью. Это приборы, в которых генерация или усиление колебаний вызваны неустойчивостью объемного заряда в полупроводниковом материале вследствие отрицательной дифференциальной подвижности.

**Дифференциальная проводимость полупроводников.** Для получения отрицательной дифференциальной проводимости полупроводников, необходимо, чтобы дрейфовая скорость в структуре уменьшалась с ростом поля. Но это произойдет только в том случае, если подвижность частиц будет уменьшаться с ростом поля, так как подвижность связана со скоростью.

Дифференциальная проводимость  $\sigma_d$  составляет

$$\sigma_d = \frac{dI}{dU} = \frac{ens}{1} \frac{dV}{dE} \quad (3.40)$$

Если подвижность не зависит от поля ( $\mu = \text{const}$ ), то  $dV/dE = \mu$ ,  $\sigma_d = \sigma$ . Дифференциальная проводимость станет отрицательной, если

$$\text{в формуле (3.40)} \quad (dV/dE) < 0. \quad (3.41)$$

Так как  $v = \mu E$ , то такое неравенство будет иметь смысл, когда подвижность зависит от поля. В конечном счете запишем  $\frac{dE}{E} < \frac{|d\mu|}{\mu}$ . Для получения отрицательной величины  $\sigma_d$  необходима зависимость подвижности от поля ( $\mu \neq \text{const}$ ), поэтому введем вместо  $\mu = v/E$  дифференциальную подвижность  $\mu_d = dV/dE$ . С уче-

том этого определения условие (2.23) означает, что для получения  $\sigma_d < 0$  необходима отрицательная дифференциальная подвижность  $\mu_d < 0$ .

Рассмотрим полупроводниковый материал (арсенид галлия), у которого наиболее ярко выражена возможность получения отрицательной дифференциальной проводимости. Осуществляется это благодаря тому, что в арсениде галлия, как и некоторых других полупроводниковых соединениях подвижность электронов в верхней долине меньше, чем в нижней.

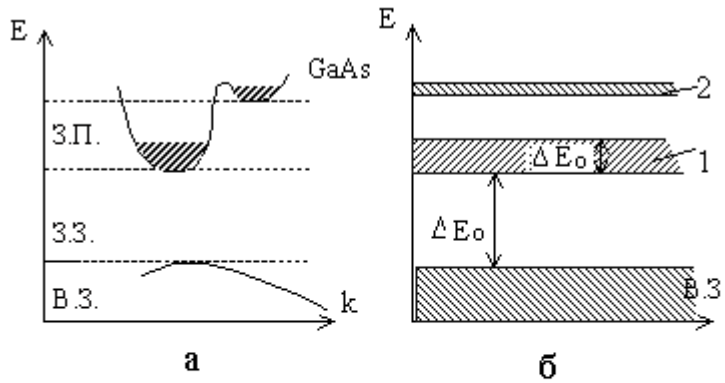


Рис. 3.19. Связь энергии электронов в валентной зоне и зоне проводимости GaAs с волновым числом  $k$  (а), появление двух подзон 1 и 2 в зоне проводимости (б).

У арсенида галлия имеется две группы носителей заряда (электронов) с разными дрейфовыми скоростями. Зона проводимости имеет сложную структуру. В частности имеется нижняя (центральная) и верхняя (боковая) долины (участки зоны проводимости), разделенные небольшой запрещенной зоной рис. 3.19.

При комнатной температуре  $kT \ll \delta T$ , верхняя долина не заселена. В состоянии термодинамического равновесия все свободные электроны совершают хаотическое (тепловое) движение. Это позволяет рассматривать коллектив электронов как электронный газ с некоторой электронной температурой  $T_э$ , определяемой из формулы  $mV_T^2 / 2 = \frac{3}{2} kT_э$ . В состоянии термодинамического равновесия средняя тепловая скорость электронов  $V_T$  определяется температурой кристаллической решетки (вещества), т.е.  $T_э = T$ .

В слабом поле  $E < 3$  кВ/см подвижность  $\mu_1$  электронов велика и составляет 6000-8500 см<sup>2</sup>/Вс. При  $E > 3,5$  кВ/см,  $|\mu_2|_{\text{див}} = \frac{\partial V}{\partial E} \approx 100$  см<sup>2</sup>/Вс. При  $E > 15 - 20$  кВ/см, средняя скорость электронов составляет  $V_{\text{нас}} = 10^7$  см/с. эффективная масса электронов в боковой долине  $m_{\text{эф2}} = 1,2m$ , а в центральной  $m_{\text{эф1}} = 0,07m$ , интервал энергии между минимумами энергии в долинах в GaAs,  $\Delta E_1 = 0,36$  эВ, а ширина запрещенной зоны  $\Delta E_0 = 1,4$  эВ. Соотношение концентрации электронов в долинах

$$\frac{n_2}{n_1} = \left( \frac{N_2}{N_1} \right) \exp\left( -\frac{\Delta E_1}{kT} \right), \quad (3.42)$$

где  $N_1$  и  $N_2$  - плотности энергетических состояний в долинах

$$\frac{N_2}{N_1} = \left( \frac{m_{\phi 2}}{m_{\phi 1}} \right)^{\frac{3}{2}} \approx 70.$$

При комнатной температуре ( $T \approx 300\text{K}$ )  $kT = 0,025$  эВ, следовательно,  $\Delta E_1 \gg kT$ , т.е.

$$\frac{n_2}{n_1} = 70 \exp\left(\frac{-0,36}{0,025}\right) \approx 7 \cdot 10^{-5}.$$

При этом все электроны находятся в нижней долине ( $n_1 \gg n_2$ ). Однако соотношение  $n_1$  и  $n_2$  резко изменится, если в полупроводнике создано сильное электрическое поле, т.е.  $n_2 > n_1$ .

При напряженности в полупроводнике более некоторой пороговой напряженности ( $E > E_{\text{п}}$ ), основная часть электронов переходит из нижней долины в верхнюю. Становится возможным получение соотношения  $n_2 > n_1$ . Для арсенида галлия  $E_{\text{п}} \approx 3$  кВ/см.

При увеличении электрического поля происходит увеличение скорости электронов на длине пробега между двумя соударениями при сохранении хаотичности движения из-за соударений.

Увеличение средней кинетической энергии в электрическом поле эквивалентно возрастанию электронной температуры по сравнению с температурой решетки ( $T_3 > T$ ). В этом случае электроны называются горячими электронами. Рост энергии электронов увеличивает вероятность перехода из долины 1 в долину 2. При этом нужно уже брать электронную температуру  $T_3$ .

В сильных полях  $E \approx 10^5$  В/м активные свойства диодов обусловлены переходом электронов из «центральной» энергетической долины в боковую и этот переход сопровождается, как показано выше, уменьшением подвижности и дрейфовой скорости, что эквивалентно появлению отрицательной дифференциальной проводимости (рис. 3.20).

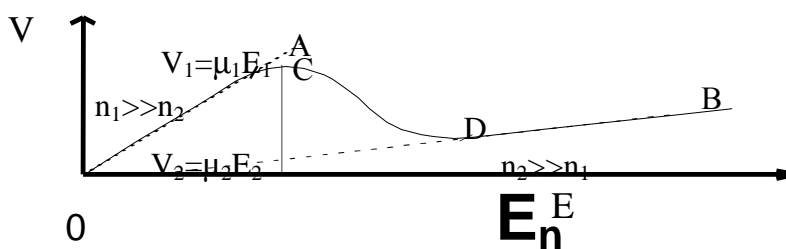


Рис. 3.20. Зависимость средней скорости электронов от электрического поля

На рис. 3.20 пунктирные прямые ОА и ОС соответствуют дрейфовым скоростям  $V_1 = \mu_1 E$  и  $V_2 = \mu_2 E$ , при этом  $\mu_1$  и  $\mu_2$  считаются постоянными, а  $\mu_1/\mu_2 \approx 50$ . Зависимость дрейфовой скорости электронов от  $E$  при  $E < E_n$  совпадает с отрезком ОА, так как все электроны находятся в нижней долине, а при слабых полях  $\mu = \text{const}$ . С ростом поля кривая отклоняется вниз от прямой ОА, так как обычно происходит уменьшение подвижности носителей. Однако при  $E = E_n$  начнется переход электронов в верхнюю долину и появятся электроны с подвижностью  $\mu_2 \approx \mu_1 / 50$ . После перехода всех электронов, зависимость  $V$  от  $E$  должна представляться участком ДВ. Переходный участок зависимости CD соответствует интервалу полей, когда имеются электроны в обеих долинах, но соотношение концентрации непрерывно изменяется: левее точки С ( $n_1 \gg n_2$ ), правее точки D  $n_2 > n_1$ . На участке CD можно говорить о некоторой эффективной (средней) скорости электронов обеих долин, определяемой по формуле:

$$V = \frac{(n_1 \mu_1 + n_2 \mu_2)}{n_0}, \quad (3.43)$$

где  $n_0 = n_1 + n_2$  - общее число электронов ( $n = 2 \cdot 10^{14} - 2 \cdot 10^{16} \text{ см}^{-3}$ ).

При  $n_2 \ll n_1$ ,  $V = V_1 = \mu_1 E$ , а при  $n_2 \gg n_1$ ,  $V = V_2 = \mu_2 E$ . Участку CD соответствует отрицательная дифференциальная подвижность:

$$\mu_0 = dv/dE < 0 \quad (3.44)$$

Время установления отрицательной дифференциальной проводимости (ОДП) определяется из времени разогрева электронного газа в «центральной» долине ( $\sim 10^{12}$  с для GaAs) и времени междолинного перехода ( $\sim 5 \cdot 10^{-4}$  с).

В реальных условиях к полупроводниковому образцу кроме постоянного напряжения приложено СВЧ – напряжение. Поэтому важно знать, как быстро устанавливается значение дрейфовой скорости при изменениях напряжения, т.е. какова частотная зависимость скорости. Инерционность определяется главным образом временем установления электронной температуры в нижней долине, которое близко к  $10^{-12}$  с. Влиянием самого времени перехода из нижней долины в верхнюю можно пренебречь, оно примерно равно  $5 \cdot 10^{-14}$  с. Инерционность процесса с повышением частоты приводит к повышению фазового сдвига между скоростью и полем. Выяснено, что на частоте 30 ГГц отставание по фазе составляет около  $45^\circ$ . Влияние фазового сдвига выражается в увеличении порогового поля и уменьшении абсолютного значения дифференциальной подвижности.

**Эффект Ганна.** Диоды Ганна (ДГ) основаны на явлениях, возникающих в объеме однородного полупроводника. Эффект Ганна заключается в возникновении СВЧ-колебаний тока в некоторых полупроводниковых кристаллах под действием сильного электрического поля. В арсениде галлия это примерно  $E_{n\text{GaAs}} \approx 3 \text{ кВ/см}$ .

Напряженность поля  $E$  в однородном образце пропорциональна напряжению, а ток – дрейфовой скорости. Поэтому можно считать, что кривая  $V(E)$  рис. 3.20 одновременно может изображать ВАХ. Но отрицательная дифференциальная

проводимость присуща лишь некоторой области образца (область объемной неустойчивости), напряженность поля в которой отличается от напряженности поля в остальном объеме. То есть пороговое значение напряженности  $E_{\pi}$ , при котором начинается междолинный переход электронов, достигается лишь в узкой области образца, где имеется неоднородность концентрации примеси или флуктуация электрического поля.

Возьмем образец длиной  $l$ , у которого имеется участок протяженностью  $d$ , где концентрация донорной примеси несколько меньше, чем в остальной части образца (рис. 3.21,а). Увеличение электрического сопротивления приводит к росту падения напряжения на участке  $d$ , т.е. к увеличению  $E$ . Тогда на участке  $d$  начинается переход электронов из нижней долины в верхнюю, сопровождающийся понижением дрейфовой скорости электронов (рис. 3.21,б).

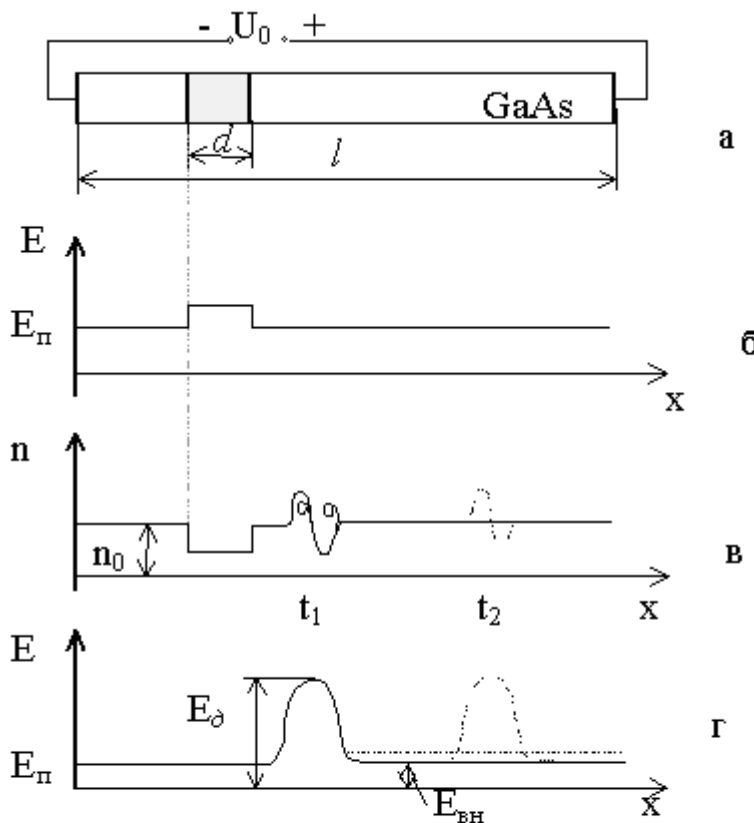


Рис. 3.21. Схема образца GaAs (а), изменение напряженности на неоднородности  $d$ , образца (б), образование домена в разные моменты времени  $t$  по длине образца (в), увеличение напряженности поля на домене (г).

Оказавшиеся в верхней долине электроны начинают отставать от электронов, оставшихся в нижней долине. В левой части участка наблюдается избыток электронов (отрицательный объемный заряд), а в правой – недостаток электронов, т.е. объемный положительный заряд донорных ионов, который теперь не компенсируется зарядом электронов. Образующийся двойной электрический слой объемного заряда (рис. 3.21,в, момент  $t_1$ ) называется электрическим доменом. Домен должен быть электрически нейтральным. Так как электроны двигаются, то и до-





Таким образом, при коротких об- разцах полупроводника их GaAs можно получить колебания СВЧ - диапазона. Период генерируемых колебаний определяется по времени прохождения домена через кристалл:

$$T_0 = \frac{l}{v_d}.$$

Критическая частота

$$f_{кр} = \frac{100}{l}, \quad (3.47)$$

где  $l$  - в микрометрах;  $f$  - в ГГц.

Другие виды неустойчивости здесь рассматриваться не будут.

**Режимы работы.** Пролетный режим генератора. Обычно так называют режим работы, в котором колебательная система, связанная с прибором имеет низкую добротность. В этом случае переменное напряжение на колебательной системе мало по сравнению с постоянным напряжением и не оказывает обратного влияния на процессы в образце из арсенида галлия. Если постоянное напряжение превышает пороговое значение, то в образце возникнут импульсы тока, частота следования которых определяется временем пролета. Этот режим уже рассмотрен как эффект Ганна.

Частота генерации в пролетном режиме определяется формулой:

$$f_{пр} = 1/T_0 = v_d/l = f_{пр},$$

где  $T_0$  – время пролета домена.

Условие пролетного режима определено неравенством:

$$n_0 l \geq 5 \cdot 10^{11} \text{ см}^{-2}$$

Коэффициент полезного действия в пролетном режиме максимален, когда домен занимает  $1/2l$  образца, а форма тока почти синусоидальная. Обычно  $\eta \leq 10\%$ . Мощность колебаний можно оценить по формуле:

$$P = \eta P_0, \quad (3.48)$$

где  $P_0$  – потребляемая от источника мощность  $P_0 = U_0 I_0$ .

Необходимое напряжение источника  $U_0 = E_n l$ , а ток

$$I_0 = en_0 v S,$$

где  $v$  - средняя дрейфовая скорость,  $S$  – площадь сечения образца,  $E_n$  - пороговое напряжение. Тогда мощность ( $P_0$ ) пропорциональна  $l$  и  $S$ . Однако площадь влияет на сопротивление образца в слабом поле  $R_0$ .

$R_0 = \frac{l}{en_0 \mu_1 S}$  - сопротивление образца в слабом поле.

Нельзя выбирать любую площадь образца, т.к. она влияет на сопротивление нагрузки  $R_n$ . Мощность связана с сопротивлением, частотой, длиной образца следующим образом:  $PR_n$  пропорционально  $l^2$  пропорционально  $1/f^2$ , следовательно, уменьшение как  $U_0$ , так и  $R_n$ , вызовет снижение полезной мощности. Длина образца  $l$  определяется выбранной рабочей частотой, а концентрация  $n_0$  - услови-

ем  $n_0 l \geq 5 \cdot 10^{11} \text{ см}^{-2}$ . Поэтому сопротивление образца  $R_0$  можно изменять только выбором площади  $S$ .

**Режим с задержкой образования домена.** Этот режим наблюдается, когда колебательная система обладает высокой добротностью, постоянное напряжение больше порогового значения ( $U_0 > U_{0n}$ ), а время пролета домена  $T_0$  меньше периода колебаний  $T$  ( $T_0 < T$ ). Домен здесь возникает при выполнении условия, когда  $U = U_0 + U_1 \sin \omega t$  достигнет порогового значения. Следующий домен возникает только с задержкой на время  $T - T_0$ . Период колебаний  $T$  определяется настройкой резонатора. При заданном времени пролета домена  $T_0$  возможна механическая перестройка частоты.

В энергетическом отношении режим с задержкой образования домена более выгоден, чем пролетный режим. Длительность импульса тока  $T_{и} = T - T_0$  может превышать сумму времени нарастания и рассасывания домена, которая определяет длительность импульса в пролетном режиме. КПД возрастает с увеличением длительности импульса  $T_{и}$  и достигает максимального значения примерно при  $T_{и} = T/2 = T_0$ , то есть на рабочей частоте, вдвое меньшей пролетной.

В режиме с задержкой образования домена КПД теоретически достигает 25%, а экспериментально получено 20% при отдаваемой мощности 100 Вт.

Представляют интерес также режимы:

подавление домена и ограничение накопления заряда (ОНОЗ). Но они выходят за рамки изучаемого курса.

**Особенности устройства и применения диодов Ганна.** Диоды изготавливают на основе монокристаллов или эпитаксиальных пленок арсенида галлия. В зависимости от выбираемого режима работы и параметров длина образцов составляет  $l = 5 \text{ мкм}$  до  $1 \text{ мм}$ , а площадь сечения -  $S = 2,5 \cdot 10^{-5} - 10^{-2} \text{ см}^2$ . Необходим хороший омический контакт, обладающий линейной ВАХ и малым сопротивлением по сравнению с сопротивлением образца.

Частотный диапазон, перекрываемый генераторами Ганна очень широк и составляет:  $100 \text{ МГц} - 150 \text{ ГГц}$ . На частотах от  $1$  до  $150 \text{ ГГц}$  диоды Ганна используют в основном для создания СВЧ генераторов.

Диоды Ганна достаточно эффективно работают при температурах  $200-250^\circ \text{C}$ . Перегрев диодов ограничивает максимальную мощность генераторов в непрерывном режиме и максимальную длительность импульсов в импульсном режиме.

Диоды Ганна включают в линии передачи и резонаторы, перестраиваемые по частоте. Коаксиально-волноводная секция с диодом Ганна, включаемая в волноводный тракт, показана на рис. 3.23. Короткозамыкающие поршни необходимы для перестройки генератора по частоте и согласования диода с нагрузкой. Увеличение рабочей частоты, требует уменьшения длины образца, следовательно, и напряжения питания.

Мощность колебаний пропорциональна квадрату напряжения питания. Для диодов Ганна, работающих в доменных режимах,  $P_f^2 = 2 \div 5 \cdot 10^3$

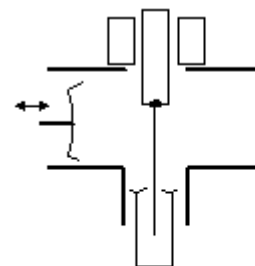


Рис. 3. 23. Настраиваемый узел в волноводе

Вт·ГГц<sup>2</sup>, а теоретический предел  $P_f^2 \approx 10^4$  Вт·ГГц<sup>2</sup>, что отражает преимущества этого режима. Коэффициент полезного действия зависит от режима работы и составляет от нескольких единиц до нескольких десятков процентов  $\eta = 10 - 20$  %. В отдельных генераторах он доходит и до 30%. Мощность в непрерывном режиме  $P_{\text{неп}} - 0,62$  Вт на частоте 12,8 ГГц при  $\eta = 3 - 4$ %. Ниже приведено еще несколько примеров работающих генераторов.  $P_{\text{имп}} - 2,1$  кВт на частоте 7,0 ГГц,  $\eta = 4$ %.

$P_{\text{неп}} - 100$  мВт на частоте 100 ГГц  $\eta = 5$ %.

**Перестройка частоты.** Перестройку частоты можно осуществлять, изменяя параметры резонатора (емкостную или индуктивную составляющие). Механическая перестройка колебательного контура происходит при плавном переходе из одного режима в другой. Электронная перестройка достигается изменением напряжения питания и примерно равна 5-20 МГц/В. Эта перестройка связана с изменением емкости домена.

Диоды Ганна не являются малошумящими приборами, так как эффективная температура электронов в области домена превышает температуру кристаллической решетки. Шум в диодах обусловлен также случайным изменением, неоднородностью момента зарождения домена.

Амплитудный шум генераторов примерно на 30 дБ меньше частотного, а последний близок к уровню шумов клистронов. Для лучших генераторов частотный шум составляет: - 110 дБ при смещении от основной частоты на 100 кГц; - 130 дБ при смещении на 1 МГц; - 160 дБ при смещении на 10 МГц.

В настоящее время генераторы на диодах Ганна находят применение в качестве СВЧ-гетеродинов и генераторов в маломощных передатчиках сантиметрового и миллиметрового диапазонов.

## Глава 4. Применение и регистрация СВЧ энергии

### 4.1. Нагрев СВЧ энергией

#### 4.1.1. Поведение диэлектриков в электромагнитных полях

Энергия СВЧ поля широко используется в промышленном производстве (сушка, сублимация, прогрев, дробление), в медицине (диатермия), в быту и торговле (СВЧ печи). Для этих применений выделены частотные диапазоны [6]:

433,2 МГц $\pm 2,0\%$	} - медицина
461,04 МГц $\pm 2,0\%$	
915,0 МГц $\pm 2\%$	- промышленные установки
2450 $\pm 50$ МГц	- СВЧ печи
5800 $\pm 75$ МГц	} - для новых разработок
22125 $\pm 125$ МГц	

Тепловую энергию можно передавать путем излучения, конвекции или теплопроводности. В процессе СВЧ нагрева передача тепла материалу производится путем облучения его энергией СВЧ. Энергия СВЧ - это очень удобный источник тепла, имеющий несомненные преимущества перед другими источниками. Нагрев энергией СВЧ является чистым методом, так как при его использовании отсутствуют какие-либо продукты сгорания, отличается большой гибкостью в применении и практически безынерционен в управлении. Легкость, с которой СВЧ энергия преобразуется в тепло, позволяет получить очень высокие скорости нагрева, возможность мгновенно изменять количество СВЧ энергии по всему объему диэлектрика, что обеспечивает хорошее качество продукции.

Не металлические материалы полупрозрачны для СВЧ - энергии. Поля СВЧ проникают в такие среды на значительную глубину, которая зависит от их свойств, параметров. В теории электромагнитного (ЭМ) поля каждая среда характеризуется параметрами: абсолютной диэлектрической проницаемостью  $\epsilon \left( \frac{\Phi}{\text{м}} \right)$ ,

абсолютной магнитной проницаемостью  $\mu \left( \frac{\Gamma_{\text{H}}}{\text{м}} \right)$  и проводимостью

$\sigma \left( \frac{1}{\text{Ом} \cdot \text{м}} \right)$ . В [12,13] приводятся относительные параметры  $\epsilon_r$  и  $\mu_r$ , которые свя-

заны с абсолютными величинами соотношениями:  $\epsilon = \epsilon_0 \cdot \epsilon_r$  и  $\mu = \mu_0 \cdot \mu_r$

Взаимодействуя со средой на атомном и молекулярном уровне, эти поля влияют на движение электронов. Электроны в таких материалах движутся более или менее свободно в зависимости от диэлектрической проницаемости и проводимости. Наличие проводимости делает диэлектрик неидеальным, с потерями. В диэлектрике с потерями, находящемся в ЭМ поле, текут токи: смещения, плот-

ность которого  $\vec{j}_{\text{см}} = i\omega \cdot \epsilon \vec{E} = \omega \epsilon \vec{E} \cdot e^{j\frac{\pi}{2}}$ , и проводимости с плотностью  $\vec{j}_{\text{пр}} = \sigma \vec{E}$ ,

сдвинутые на  $\frac{\pi}{2}$  при векторном представлении плотности полного тока  $\vec{j} = \vec{j}_{см} + \vec{j}_{пр}$  (рис. 4.1.). Когда число движущихся электронов становится настолько большим (проводимость велика), что индуцированные при их движении поля по своей интенсивности приближаются к возбуждающему полю, результирующее поле становится равным нулю. СВЧ- энергия не может проникнуть внутрь материала, так как возникает явление поверхностного эффекта.

Отношение тока проводимости к току смещения называется тангенсом угла потерь

$$\operatorname{tg} \delta = \frac{\sigma}{\omega \epsilon}. \quad (4.1)$$

Величина  $\operatorname{tg} \delta$  определяет коэффициент затухания  $\alpha$  ЭМ поля в этом диэлектрике:

$$\alpha = \omega \cdot \sqrt{\epsilon \mu \cdot \left( \frac{\sqrt{1 + \operatorname{tg}^2 \delta} - 1}{2} \right)}, \quad (4.2)$$

где  $\omega$  - круговая частота электромагнитного поля.

Величина коэффициента затухания  $\alpha$  определяет изменение амплитуды полей (электрического -  $E_m$  и магнитного -  $H_m$ ) в зависимости от расстояния от поверхности в образце  $x$ :

$$E_m(x) = E_m(0) \cdot e^{-\alpha x} \cdot e^{-j\beta x}, \quad (4.3)$$

здесь  $\beta$ -коэффициент фазы волны в диэлектрике.

В случае диэлектриков с малыми потерями ( $\operatorname{tg} \delta \ll 1$ ) выражение (4.2), после разложения в ряд по малому параметру, примет более простой вид:

$$\alpha = \frac{\omega \sqrt{\epsilon_r \cdot \mu_r}}{c} \cdot \frac{\operatorname{tg} \delta}{2} = \frac{\pi \sqrt{\epsilon_r \cdot \mu_r}}{\lambda} \cdot \operatorname{tg} \delta, \quad (4.4)$$

где  $\lambda = \frac{c}{f}$  длина волны в свободном пространстве.

Расстояние, на котором амплитуда поля электромагнитной волны уменьшается в  $e$  раз, называется глубиной проникновения поля в материал -  $\Delta$ . Для случая сред с малыми потерями величина  $\Delta$ , согласно (4.3) и (4.4), выражается как:

$$\Delta = \frac{1}{\alpha} = \frac{\lambda}{\pi \sqrt{\epsilon_r \cdot \mu_r} \cdot \operatorname{tg} \delta}. \quad (4.5)$$

Глубина проникновения поля в металл, когда  $\operatorname{tg} \delta \gg 1$ , из (4.5) с учетом (4.2) определяется формулой:

$$\Delta_{мет} = \sqrt{\frac{2}{\omega \mu \sigma}}. \quad (4.6)$$

В таблице 4.1. приведены электродинамические параметры [6,7] некоторых материалов на частотах, применяемых для СВЧ нагрева.

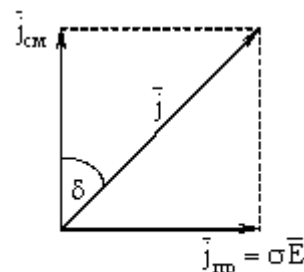


Рис. 4.1

ЭМ поле в неидеальном диэлектрике затухает, т.е. рассеивается. Потери энергии СВЧ поля в диэлектрике это тепловые (Джоулевы) потери, мощность  $P_{\text{пот}}$  которых определяется соотношением:

$$P_{\text{ном}} = \int_{V_{\varepsilon}} \sigma (\vec{E} \cdot \vec{E}^*) dV = \omega \varepsilon_0 \varepsilon_r \text{tg} \delta \int_{V_{\varepsilon}} |E|^2 dV, \quad (4.7)$$

где  $V_{\varepsilon}$  - объем нагреваемого диэлектрика;

$|E|$  - величина электрического поля в области, занятой диэлектриком, в каждом конкретном случае является функцией координат.

Эффективность преобразования энергии электрического поля в тепловую энергию, согласно (4.7), увеличивается пропорционально рабочей круговой частоте и квадрату напряженности электрического поля. Но увеличивать напряженность электрического поля нельзя, так как, начиная с некоторой величины (30 кВ/см), возникает пробой. Поэтому увеличение частоты - единственный путь увеличения удельной тепловой энергии потерь. Но с ростом частоты уменьшается глубина проникновения поля в вещество (4.5) и при большом объеме диэлектрика может оказаться, что ЭМ поле будет отсутствовать в его центральной части, которая останется без тепловой обработки.

Процесс термообработки сопровождается повышением температуры диэлектрика. Чтобы нагреть образец массой  $m$  на  $\Delta T$  градусов (от  $T_{\text{н}}$  - начальной до  $T_{\text{к}}$  - конечной температуры, т.е.  $\Delta T = T_{\text{к}} - T_{\text{н}}$ ) при его удельной теплоемкости  $\gamma \left[ \frac{\text{кал}}{\text{г} \cdot \text{град}} \right]$ , необходима энергия:

$$W = 4,1868 \cdot m \cdot \gamma_c \cdot \Delta T, \quad [\text{Дж}]. \quad (4.8)$$

Изменение энергии нагрева за время  $\Delta t$ , т.е.  $\frac{W}{\Delta t}$ , есть мощность  $P_{\text{пот}} = \frac{W}{\Delta t}$ , затрачиваемая ЭМ полем. Это равенство с учетом (4.7) и (4.8) переходит в соотношение:

$$\Delta t \omega \varepsilon_0 \varepsilon_r \cdot \text{tg} \delta \int_{V_{\varepsilon}} |E|^2 dV = 4,1868 \cdot m \gamma \Delta T. \quad (4.9)$$

Из (4.9) можно определить изменение температуры диэлектрика за время  $\Delta t$ , или наоборот.

$$\Delta T = \frac{\omega \varepsilon \cdot \text{tg} \delta \int_{V_{\varepsilon}} |E|^2 dV}{4,1868 \cdot m \gamma} \cdot \Delta t, \quad [\text{град}]. \quad (4.10)$$

Величина  $|E|^2$  определяется ЭМ полями в тех системах, которые применяются для нагрева, т.е. в волноводах, антеннах или резонаторной камере СВЧ печи.

Таблица 4.1.

Параметры $\varepsilon_r$ , $\text{tg} \delta$ и теплоемкости $\gamma$						
№	Материал	$f = 3 \cdot 10^8, \text{Гц}$		$f = 3 \cdot 10^9, \text{Гц}$		$\gamma, \frac{\text{кал}}{\text{г} \cdot \text{град}}$
		$\varepsilon_r$	$\text{tg} \delta$	$\varepsilon_r$	$\text{tg} \delta$	
						$T, \text{ } ^\circ\text{C}$

1	Мясо сырое постное	50	0,78	40	0,3	0,58	25
2	Мясо сырое жирное	—	—	33	0,32	0,65	4,4
3	Мясо сухое	2	0,02	2	0,02	0,3	25
4	Жир	2,5	0,12	2,5	0,07	0,43	25
5	Вода	77,5	0,016	76,7	0,157	1	25
6	Мышечная ткань	49 – 52	$1,6 \cdot 10^{-3}$	45 – 48	$2,96 \cdot 10^{-3}$	0,4	25
7	Печень	—	—	37	$3 \cdot 10^{-3}$	0,58	25
8	Стекло	5 – 7	0,01-0,02	5,2	0,01	0,2	20
9	Фарфор	5,8	0,0115	5,7	0,014	0,2	20
10	Снег, лед	—	—	—	—	0,43	0
11	NaCl	5,9	$2 \cdot 10^{-4}$	—	—	0,22	—
12	Тефлон	2,0	$2 \cdot 10^{-4}$	2,04	$2,8 \cdot 10^{-4}$	0,1	25
13	Бумага	2,75	0,066	2,7	0,056	—	23
14	Пихта зел.	15,8	0,19	15,0	0,18	—	23

Следует отметить, что при падении волны на диэлектрик определенной толщины некоторая часть поля отражается от поверхности. Если  $Z_{gn}$  – волновое сопротивление диэлектрика у поверхности, а  $Z_0$  волновое сопротивление свободного пространства, то величина отраженного сигнала [12] определяется коэффициентом отражения.

$$\dot{\Gamma} = \frac{Z_{gn} - Z_0}{Z_{gn} + Z_0} \quad (4.11)$$

$Z_{gn}$  представляет, с другой стороны, входное сопротивление диэлектрика и согласно [8] теории передающих линий, определяется по формуле:

$$\dot{z}_{gn} = z_g \frac{z_0 \cdot ch(l \cdot \gamma) + z_g \cdot sh(l \cdot \gamma)}{z_g \cdot ch(l \cdot \gamma) + z_0 \cdot sh(l \cdot \gamma)},$$

где  $\dot{Z}_q = \sqrt{\mu / \dot{\epsilon}}$  – волновое сопротивление диэлектрика;

$l$  – толщина диэлектрика.

$\gamma$  - комплексная постоянная распространения волны, определяется уравнением  $\gamma = \alpha + i\beta$ . Величина  $\alpha$  (4.2) в случае малых потерь определяется формулой (4.4), а фазовая постоянная при малых потерях равна

$$\beta = \left(1 + \frac{\text{tg}^2 \delta}{8}\right) \cdot \omega \sqrt{\mu \cdot \epsilon}$$

Волновое сопротивление свободного пространства  $Z_0 = 377 \text{ Ом}$ , волновое сопротивление диэлектрика при малых потерях [12] можно определять как

$$Z_{\text{qn}} \approx 377 \sqrt{\frac{\mu_r}{\epsilon_r}} \cdot \left[1 - \frac{3}{8} \cdot \text{tg}^2 \delta + i \cdot \frac{\text{tg} \delta}{2}\right] \quad (4.12)$$

Хотя уравнения для  $Z_{\text{qn}}$ ,  $\alpha$ ,  $\beta$  имеют приближенный характер и справедливы для случая малых потерь в среде, но они позволяют оценить сложные процессы взаимодействия электромагнитной волны с диэлектриком.

#### 4.1.2. Электродинамические системы нагрева диэлектриков полями бегущих волн

Электродинамические системы для нагрева в распространяющихся полях применяются для сушки бумаги, тканей, фанеры, пиломатериалов, для нагрева порошков, для раскалывания камней, гранита, для размораживания пищевых продуктов и т.п. Такие поля создаются с помощью волноводов (прямоугольных, цилиндрических, коаксиальных), антенн специальных конструкций, возбуждаемых волноводами, и работающими в дециметровом и сантиметровом диапазонах волн. Для создания полей бегущих ЭМ волн в таких линиях необходимо выполнить согласование сопротивлений всех узлов. Тогда, очевидно, нагрев диэлектрика происходит равномерно электрическим полем бегущей волны, а не использованная полностью энергия поступает в нагрузку.

**Прямоугольный волновод** для СВЧ нагрева полями бегущих волн работает на основной волне типа  $H_{10}$  (рис. 4.2) и его размеры сторон  $a/2$  и  $b$  задаются неравенствами

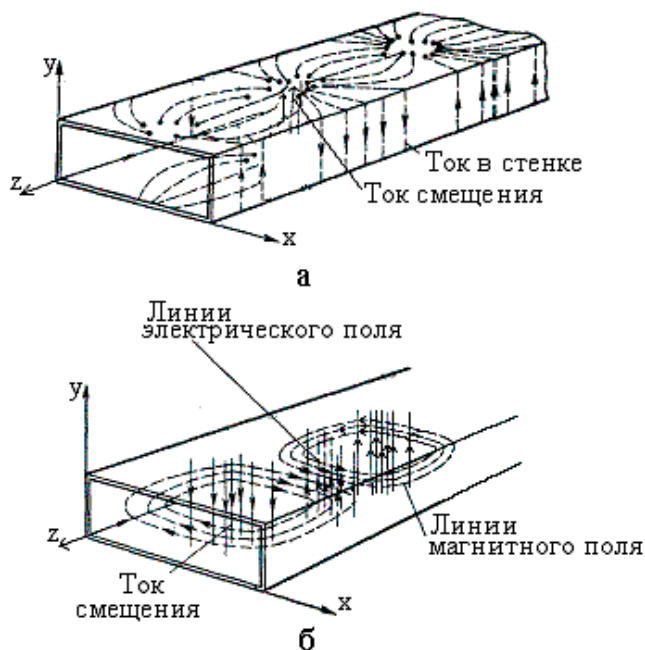


Рис. 4.2. Распределение поля волны  $H_{10}$  в волноводе.



$$0.5\lambda_0 < a < \lambda_0$$

$$0 < b < 0.5\lambda_0$$

или

$$1.2a < \lambda_0 < 1.8a$$

$$b < 0.5a$$

здесь  $\lambda_0 = \frac{3000}{f[\text{MHz}]}$  длина волны в свободном пространстве,  $f$  - рабочая частота.

Для данного волновода и типа волны необходимо рассчитать величину  $|E|^2$ ,

определяющую мощность потерь на нагрев. Для этого используем выражения компонент полей, которые получены из уравнений Максвелла с учетом граничных условий для данного волновода:

$$\left. \begin{aligned} E_y &= Z_c A \sin \frac{\pi \cdot x}{a} \\ H_x &= -A \sin \frac{\pi \cdot x}{a} \\ H_z &= -jA \frac{\lambda_b}{2a} \cos \frac{\pi \cdot x}{a} \\ H_y &= E_x = E_z = 0 \end{aligned} \right\} e^{-j\beta_b z} \quad (4.13)$$

где  $\beta_b$  - коэффициент фазы волны в волноводе;

$A$  - константа, определяется величиной передаваемой мощности  $P$  и размерами поперечного сечения волновода в виде

$$|A| = 0.103 \cdot \sqrt{\frac{P}{a \cdot b}} \sqrt{1 - \left(\frac{\lambda_0}{2a}\right)^2} \quad (4.14)$$

здесь  $a, b$  - в сантиметрах;

$Z_c$  - характеристическое сопротивление волновода для волны  $H_{10}$  определяется из соотношения

$$Z_c = \frac{120 \cdot \pi}{\sqrt{\varepsilon'_r - \left(\frac{\lambda_0}{2 \cdot a}\right)^2}}; \quad (4.15)$$

$$\lambda_g = \frac{\lambda_0}{\sqrt{\varepsilon'_r - \left(\frac{\lambda_0}{2a}\right)^2}}, \quad (4.16)$$

где  $\varepsilon'_r$  - действительная часть относительной диэлектрической проницаемости среды волновода.

Распределение ЭМ поля и токов волны типа  $H_{10}$  (структура) (рис. 4.2) построено в соответствии (4.13). Линии тока смещения направлены также как и линии электрического поля; их максимум в центре волновода, но он сдвинут вдоль оси  $z$  на  $\pi/2$ , относительно максимума поля  $E$ . По внутренним стенкам волновода течет ток проводимости, создаваемый тангенциальным магнитным полем  $H$ , проходящим вблизи стенок. Картина поля с совпадением направлений линий повторяется через расстояния, равные длине волны в волноводе  $\lambda_b$ . Картины токов и ЭМ поля позволяют выбрать место расположения диэлектрика для эффективного использования энергии электрического поля, а также форму технологиче-

ских отверстий, в которых следует помещать дополнительные ручки управления. Из рис. 4.2 следует, что диэлектрик должен быть конечной толщины, располагаться в плоскости  $uoz$ , проходящей через середину широкой стенки волновода, где линии электрического поля параллельны его границам; технологические отверстия в стенках не должны нарушать линий токов проводимости.

Амплитуда напряженности электрического поля из (4.13) с учетом (4.14) и (4.15) и при  $x=0,5a$  равна

$$E_{ym} = \frac{38,8}{\sqrt{\epsilon_r'}} \sqrt{\frac{P}{a \cdot b}} \sqrt{1 - \left(\frac{\lambda_0}{2a}\right)^2}. \quad (4.17)$$

Волна  $H_{10}$  распространяется вдоль оси  $z$  с фазовой скоростью

$$v_{\phi} = \frac{c}{\sqrt{\epsilon_r' - \left(\frac{\lambda_0}{2a}\right)^2}}. \quad (4.18)$$

При наличии потерь в стенках и диэлектрике поля ослабляются, их амплитуда уменьшается по закону:

$$k = k_0 \cdot e^{-\alpha_n z}, \quad (4.19)$$

где  $\alpha_n = \alpha_b + \alpha_g$  - постоянная затухания в системе, состоит из потерь в стенках волновода  $-\alpha_b$  и в диэлектрике  $\alpha_g$ , заполняющем волновод полностью.

Если можно пренебречь потерями в стенках волновода из-за их малости, то потери в диэлектриках следует учитывать, и тогда  $\alpha_n \sim \alpha_g$ . Величина  $\alpha_g$  определяется выражением:

$$\alpha_g = \frac{\omega \cdot \epsilon_r'' \sqrt{\epsilon_0 \cdot \mu_0}}{2 \sqrt{\epsilon_r' - \left(\frac{\lambda_0}{2a}\right)^2}}, \quad (4.20)$$

где  $\epsilon_r''$  - мнимая часть относительной диэлектрической проницаемости среды, связанная с проводимостью диэлектрика ( $\sigma$ ) соотношением:

$$\sigma = \omega \epsilon_0 \epsilon_r'' \left[ \frac{1}{\text{Ом}} \right] \quad (4.21)$$

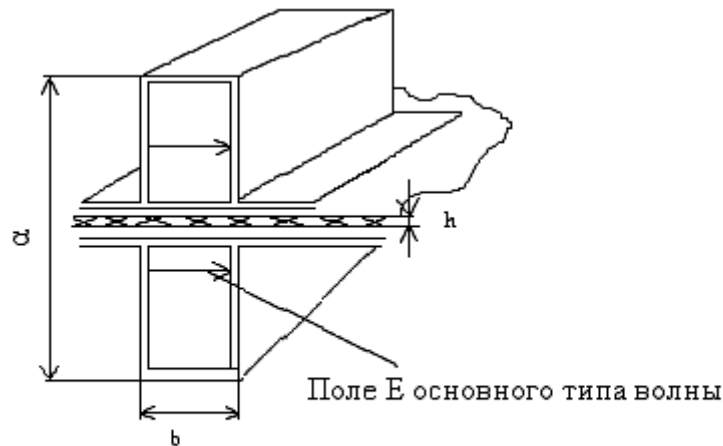


Рис. 4.3 Волновод, частично заполненный диэлектриком

Для случая волноводов, частично заполненных диэлектриком (рис. 4.3),

коэффициенты фазы  $\beta_\varepsilon$  и амплитуды  $\alpha_\varepsilon$  будут зависеть от геометрии поперечного сечения и места расположения диэлектрика в волноводе [6]:

$$\beta_\varepsilon = \frac{2 \cdot \pi}{\lambda_{b_\varepsilon}} = \beta_b + (\varepsilon' - 1) \frac{\Delta S}{S} \frac{\lambda_b}{\lambda_0^2} \sin^2 \frac{\pi x_1}{a},$$

$$\alpha_\varepsilon = \alpha_b + \frac{2\pi\lambda_b}{\lambda_0^2} \varepsilon'' \frac{\Delta S}{S} \sin^2 \frac{\pi x_1}{a},$$
(4.22)

где  $\alpha_b$  — коэффициент затухания в волноводе без диэлектрика;

$\Delta S$  и  $S$  — площади поперечного сечения диэлектрического образца и волновода;

$\lambda_b$  — длина волны в волноводе, в отсутствие диэлектрика;

$x_1$  — расстояние от стенки волновода до диэлектрика.

В тех случаях, когда необходимо знать зависимость коэффициента затухания диэлектрика от температуры или влажности, его можно измерить. Для этого измеряют мощность  $P_0$  на входе в систему, а затем — мощность  $P_1$  на выходе из секции волновода с диэлектриком длиной  $l$ . Выражение мощности при наличии потерь в диэлектрике может быть записано в виде:

$$P_e = P_0 \cdot e^{-2 \cdot \alpha_\varepsilon l}. \quad (4.23)$$

Преобразуемая в тепло мощность в согласованном режиме определяется соотношением:

$$P_o = P_0 \left(1 - \frac{P_e}{P_0}\right). \quad (4.24)$$

Из (4.23) и (4.24) определяется коэффициент затухания для нагреваемого образца по формуле:

$$\alpha_\varepsilon = \ln(P_e/P_0)/(2l). \quad (4.25)$$

Коэффициент затухания также зависит от температуры и влажности.

Чтобы достичь высокой эффективности использования энергии СВЧ генератора, следует стремиться согласовать волновое сопротивление пустого волновода с волновым сопротивлением волновода с диэлектриком.

Для увеличения интенсивности электрического поля в месте расположения тонких диэлектриков при фиксированной вводимой мощности можно использовать Н – образные (или П – образные) волноводы (рис.4.4), выполняя их из прямоугольного путем помещения вкладышей, регулирующих высоту и ширину выступающего гребня ( $\hat{a}$ -q). Критическая частота для Н – образного волновода, выполненного на основе прямоугольного волновода с размерами  $a \times b$ , будет определяться формулой [5]:

$$f_{крН} = \frac{1}{2\pi} \sqrt{\frac{q}{\mu \cdot \varepsilon \cdot l \cdot b \cdot d}}. \quad (4.26)$$

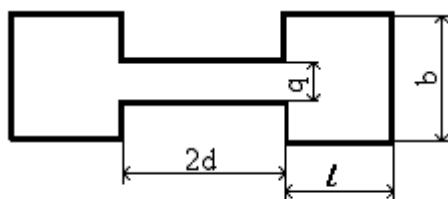


Рис. 4.4. Поперечное сечение Н – образного волновода.

Размеры прямоугольного волново- да  $a \times b$ , используемого для выполнения Н- образного волновода, следует уменьшить, чтобы не возбудилась волна  $H_{20}$ . Во-первых, она имеет минимальное поле  $E$  на выступах, во-вторых, при возбуждении одновременно воли  $H_{20}$  и  $H_{10}$  передаваемая мощность перераспределяется между ними и интенсивность электрического поля волны  $H_{10}$  снижается. При толстых слоях диэлектрика в волноводе или больших величинах  $\epsilon_r$  диэлектрика тоже могут возникнуть две волны.

Диэлектрики в волноводах можно перемещать (рис. 4.5) по направлению движения волны (согласно), против направления движения волны (встречно) и в поперечном направлении, т.е. когда энергия и диэлектрическая материя движутся во взаимно перпендикулярных направлениях. В этих типах перемещений различными будут величины потерь мощности нагрева диэлектрика до определенной температуры (или сушки), необходимые технологические длины линий при известных коэффициентах затухания и одинаковых величинах входных мощностей, так как непрерывно изменяется температура образца, влажность и коэффициент затухания при поглощении мощности.

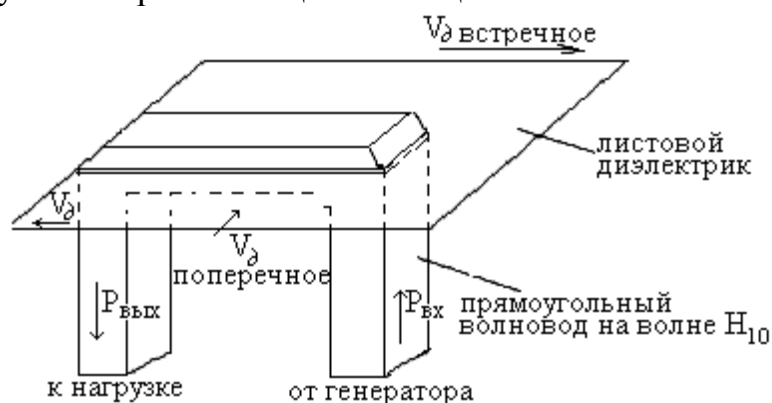


Рис 4.5. Распространение волны и направление движения ( $V_0$ ) диэлектрического листа в продольной щели широкой стенки волновода.

**Круглые волноводы** также можно использовать для нагрева (жидкостей, порошков и т.п.), если расположить по центру трубку из термостойкого стекла с малыми потерями или фторопласта, а нагреваемый образец помещать в нее.

В круглом волноводе для нагрева используется волна типа  $E_{01}$ , имеющая аксиальную симметрию электрического продольного поля и максимальную плотность энергии на оси (рис 4.6).

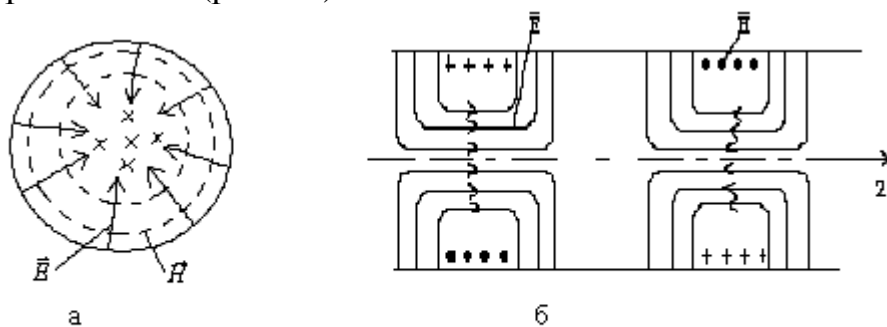


Рис. 4.6. Распространение поля волны  $E_{01}$  в круглом волноводе

Поля волны типа  $E_{01}$  в полярных координатах описываются следующими уравнениями:

$$\left. \begin{aligned} E_r &= AJ_1(4.81 \frac{r}{D}) \\ E_z &= -jA \frac{\lambda_0}{\lambda_{kp}} J_0(4.81 \frac{r}{D}) \\ H\varphi &= -\frac{A}{Z_E} J_1(4.81 \frac{r}{D}) \end{aligned} \right\} e^{-j\beta_b z} \quad (4.27)$$

$$E_\varphi = 0, \quad H_r = 0, \quad H_z = 0,$$

где  $J_0$  и  $J_1$  – функции Бесселя первого рода нулевого и первого порядков;

$D$  – внутренний диаметр волновода;

$\lambda_{kp} = 1,305 \cdot D$  - критическая длина волны типа  $E_{01}$ ;

$Z_E$  – характеристическое (волновое) сопротивление волновода с диэлектриком на волне  $E_{01}$ , имеющее следующий вид:

$$Z_E = 377 \sqrt{\epsilon'_r - \left( \frac{\lambda_0}{\lambda_{kp}} \right)^2}, [OM] \quad (4.28)$$

Амплитуда  $A$  определяется выражением:

$$A = |A| = 60 \frac{\sqrt{P}}{D} \sqrt{1 - \left( \frac{\lambda_0}{1,305 \cdot D} \right)^2}. \quad (4.29)$$

Энергия из прямоугольного волновода подается в круглый волновод с помощью трансформатора типа волны. Все изолирующие, крепежные элементы и вводные стеклянные трубки должны располагаться от этого трансформатора на расстоянии порядка длины волны в круглом волноводе на рабочем типе  $E_{01}$ . Ввод в круглый волновод диэлектрических материалов для нагрева производится через наружные металлические трубки, впаянные в основной волновод и работающие в запердельном режиме (режим не распространения всех типов волн). Входное сопротивление трубки в плоскости волновода в запердельном режиме будет чисто реактивным и энергия в нее не поступает.

В тех случаях, когда нагреваемые диэлектрики находятся вне волноводов, используются специальные излучатели. Для диатермии – спиральные антенны с отражателями в виде чаши (рис.4.7), для дробления камня - открытый конец волновода или открытый конец волновода со стержнями (рис.4.8). Спиральные антенны [10] позволяют получать осевое излучение с круговой поляризацией в направлении максимума диаграммы направленности (рис.4.7,б). Направление вращения плоскости поляризации совпадает с направлением намотки спирали. Волны с круговой поляризацией при отражении от поверхности диэлектрика изменяют на обратное направление вращения поляризации. Отраженная волна не может быть принята той же самой или соседней спиральной антенной с таким же направлением намотки спирали. Это позволяет устранять отраженные волны в тракте генератор – антенна. Спиральная антенна, кроме того, обладает широкополосными свойствами и ее можно присоединять к кабелю без сложного по конструкции дополнительного согласования.

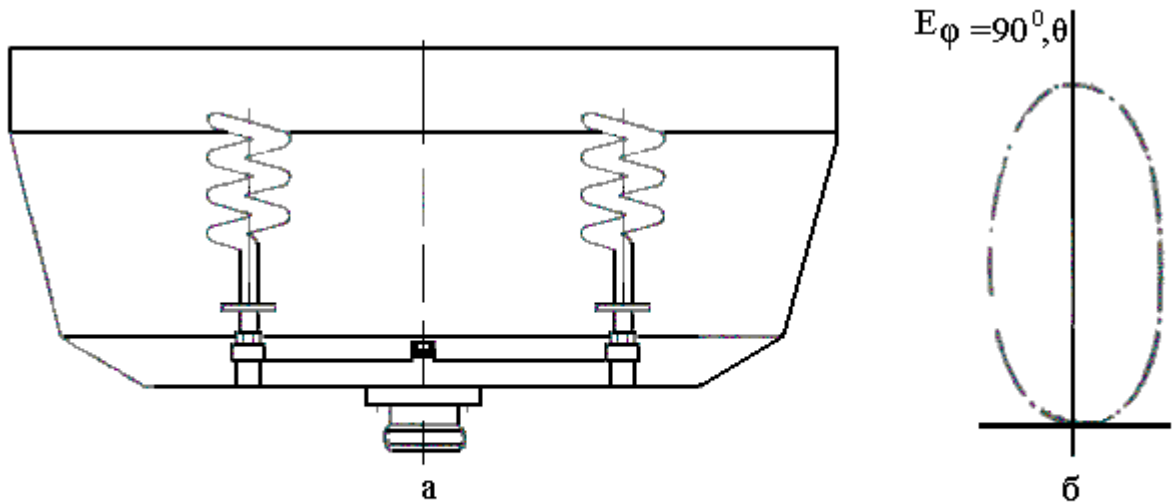


Рис.4.7. Спиральный излучатель в рефлекторной чаше с осевым излучением

Дробление камней (бетона) происходит из-за внутренних механических напряжений, возникающих при быстром нагреве и связанным с ним расширением небольших объемов каменной породы. При этом большое значение имеет локальность нагрева.

Использование в качестве облучателя открытого конца волновода [6], при мощности 2 киловатта, дробление поверхностного слоя начиналось через 20 секунд, но для каждого следующего действия требовалось время более 5-6 минут.

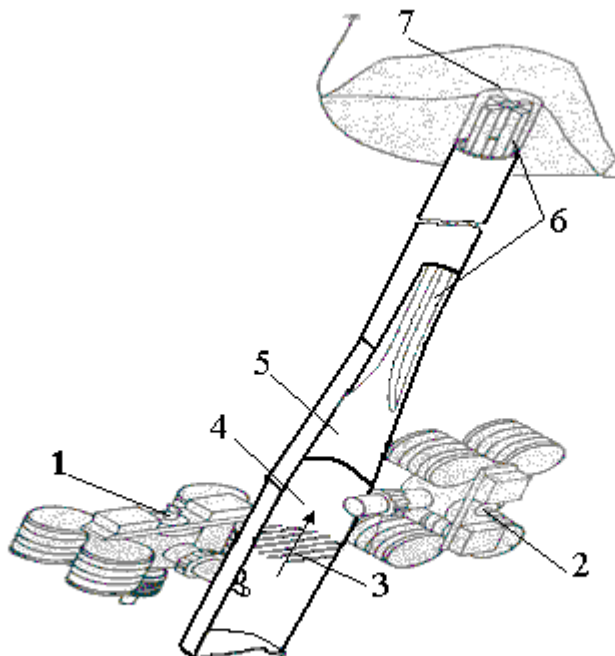


Рис. 4.8. Излучатель для дробления каменной породы с питанием от двух магнетронов

Лучшие показатели имел излучатель в виде зонда (рис.4.8), выполненный на основе круглого волновода (4), работающего на волне  $H_{11}$ . Поперечное сечение волновода плавно уменьшается (5) по направлению к излучателю. Из открытого конца круглого волновода выступают четыре стержня (6), электрически связан-

ные между собой, благодаря чему энергия системой излучается радиально. Для прочности стержни покрываются сверху колпачком из синтетического материала (7). Эти зондовые излучатели вводятся через отверстие внутрь камня. Подводимая к излучателю мощность от двух магнетронов (1 и 2), развязанных между собой с помощью поляризационной решетки (3), составляла более 6-7 киловатт. Время воздействия с последующим взрывом на глубине до 20 см., раскалыванием и дроблением составляет 2 минуты. Затем излучатель перемещается вглубь. Продвигаемый вперед излучатель позволяет получать лучшие результаты, чем неподвижный.

### 4.1.3. Нагрев в поле стоячей волны

Термообработка диэлектриков в ЭМ поле стоячей волны производится с помощью объемных резонаторов. Система такого типа называется СВЧ печью или микроволновой печью.

СВЧ печь - это конструктивное устройство, ограничивающее область взаимодействия между полями СВЧ и нагреваемым продуктом.

Печи СВЧ по уровню используемой мощности делятся на две группы:

-малой мощности от 0,6 кВт до 2 кВт для бытовых использований;

-мощные от 2 кВт до 100 кВт для промышленных применений.

СВЧ печь включает восемь основных узлов и элементов:

- 1) источник питания генератора и вспомогательных элементов;
- 2) СВЧ генератор - магнетрон, клистрон, ЛОВМ и т.п.;
- 3) секция передачи энергии от генератора к нагревательной камере;
- 4) устройство связи, обеспечивающее передачу энергии в камеру;
- 5) устройства распределяющие (выравнивающие) СВЧ поля, которые способствуют достижению равномерного нагрева образца в камере;
- 6) резонансная камера, работающая в полосе заданных частот;
- 7) герметизирующие уплотнения или ВЧ – ловушки, предотвращающие паразитные излучения;
- 8) устройства управления рабочим режимом и схемы безопасности.

Дополнительно СВЧ печь может иметь гриль, конфорки.

Особенности работы и характеристики генераторов СВЧ рассмотрены в главе 2.4. Вопросы пунктов 1) и 8) не составляют предмет рассмотрения для данного пособия. Некоторые особенности формирования ЭМ поля для нагрева в устройствах и конструктивные особенности элементов, относящиеся к пунктам 3, 4, 5, 6, 7, будут рассмотрены ниже. Остановимся на основных положениях, определяющих работу резонансной камеры.

**Нагревательная камера СВЧ печи** - призматический резонатор прямоугольной формы сечения, имеющий размеры  $a \times b \times h$  (рис. 4.9), выполненный из метал-

ла (нержавеющая сталь, алюминий и т.п., относительная магнитная прони-

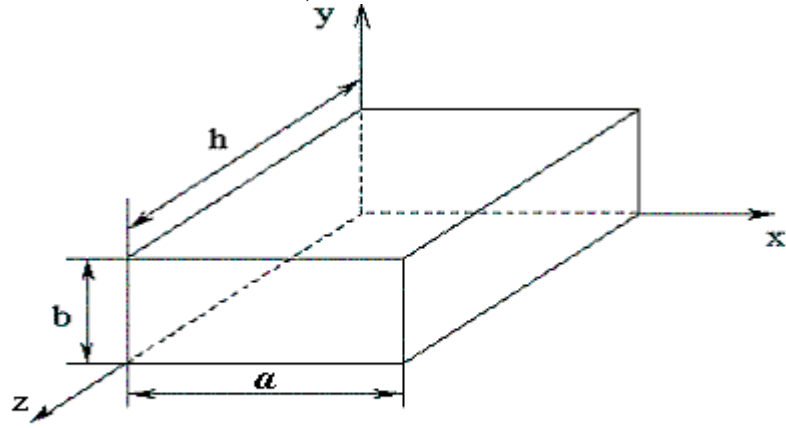


Рис.4.9

цаемость которых  $\mu_r = 1$ ).

Проводящие стенки необходимы для отражения электромагнитных волн, падающих на их поверхность. В результате этого между параллельными поверхностями камеры возникают стоячие волны электрических и магнитных полей, с разным числом полупериодов (вариаций) функций  $\cos$  или  $\sin$  вдоль соответствующих осей координат. Эти различия отмечаются в величинах индексов полей  $m$ ,  $n$ ,  $p$ . Индекс  $m$  означает число вариаций по оси  $x$  в пределах стороны  $\vec{a}$ , индекс  $n$  - по оси  $y$  на стороне  $b$ , а индекс  $p$  - число вариаций по оси  $z$  на длине резонатора  $h$ .

В прямоугольных резонаторах возможны колебания типов  $H_{mnp}$  и  $E_{mnp}$ , для которых выражения компонент поля [9] имеют вид:

колебания типа  $H_{mnp}$

$$\begin{aligned}
 H_x &= -D \cdot \frac{\pi n}{a} \cdot \frac{\pi p}{h} \cdot \sin\left(\frac{\pi m x}{a}\right) \cdot \cos\left(\frac{\pi n y}{b}\right) \cdot \cos\left(\frac{\pi p z}{h}\right); \\
 H_y &= D \cdot \frac{\pi n}{b} \cdot \frac{\pi p}{h} \cdot \cos\left(\frac{\pi m x}{a}\right) \cdot \sin\left(\frac{\pi n y}{b}\right) \cdot \cos\left(\frac{\pi p z}{h}\right); \\
 H_z &= D \cdot \left[ \left(\frac{\pi m}{a}\right)^2 + \left(\frac{\pi n}{b}\right)^2 \right] \cdot \cos\left(\frac{\pi m x}{a}\right) \cdot \cos\left(\frac{\pi n y}{b}\right) \cdot \sin\left(\frac{\pi p z}{h}\right); \\
 E_x &= j\omega_0 \mu D \cdot \frac{\pi n}{b} \cdot \cos\left(\frac{\pi m x}{a}\right) \cdot \sin\left(\frac{\pi n y}{b}\right) \cdot \sin\left(\frac{\pi p z}{h}\right); \\
 E_y &= -j\omega_0 \mu D \cdot \frac{\pi m}{a} \cdot \sin\left(\frac{\pi m x}{a}\right) \cdot \cos\left(\frac{\pi n y}{b}\right) \cdot \sin\left(\frac{\pi p z}{h}\right); \\
 E_z &= 0,
 \end{aligned} \tag{4.30}$$

где  $D$  - амплитудный коэффициент, он равен  $D = \sqrt{\frac{4P}{ab} \cdot \frac{g_{mn}^2}{\omega \mu \beta}}$ .

колебания типа  $E_{mnp}$

$$E_x = -D \cdot \frac{\pi n}{a} \cdot \frac{\pi p}{h} \cdot \cos\left(\frac{\pi m x}{a}\right) \cdot \sin\left(\frac{\pi n y}{b}\right) \cdot \sin\left(\frac{\pi p z}{h}\right);$$



$$\begin{aligned}
E_y &= -D \cdot \frac{\pi n}{b} \cdot \frac{\pi p}{h} \cdot \sin\left(\frac{\pi n x}{a}\right) \cdot \cos\left(\frac{\pi n y}{b}\right) \cdot \sin\left(\frac{\pi p z}{h}\right); \\
E_z &= D \cdot \left[ \left(\frac{\pi n}{a}\right)^2 + \left(\frac{\pi n}{b}\right)^2 \right] \cdot \sin\left(\frac{\pi n x}{a}\right) \cdot \sin\left(\frac{\pi n y}{b}\right) \cdot \cos\left(\frac{\pi p z}{h}\right); \\
H_x &= j\omega_0 \varepsilon D \cdot \frac{\pi n}{b} \cdot \sin\left(\frac{\pi n x}{a}\right) \cdot \cos\left(\frac{\pi n y}{b}\right) \cdot \cos\left(\frac{\pi p z}{h}\right); \\
H_y &= -j\omega_0 \varepsilon D \cdot \frac{\pi n}{a} \cdot \cos\left(\frac{\pi n x}{a}\right) \cdot \sin\left(\frac{\pi n y}{b}\right) \cdot \cos\left(\frac{\pi p z}{h}\right); \\
H_z &= 0,
\end{aligned} \tag{4.31}$$

где  $D = \sqrt{\frac{4P}{ab} \cdot \frac{g_{mn}^2}{\omega \varepsilon \beta}}$ ;  $g_{mn}^2 = \left(\frac{\pi m}{a}\right)^2 + \left(\frac{\pi n}{b}\right)^2$ ;  $\beta = \sqrt{\omega^2 \varepsilon \mu - g_{mn}^2}$ .

Чтобы представить особенности общего поля в СВЧ камере, рассмотрим поле одного из многих колебаний  $H_{mnp}$ , например  $H_{102}$ . Поле колебания  $H_{102}$  описывается следующими соотношениями (из системы (4.31), при значениях  $m=1$ ,  $n=0$ ,  $p=2$ ):

$$\left. \begin{aligned}
E_y &= E_0 \cdot \sin\left(\frac{\pi x}{a}\right) \cdot \sin\left(\frac{2\pi z}{h}\right) \\
H_x &= -j \cdot \frac{E_0}{Z_w} \cdot \frac{\lambda}{2h} \cdot \sin\left(\frac{\pi x}{a}\right) \cdot \cos\left(\frac{2\pi z}{h}\right) \\
H_z &= j \cdot \frac{E_0}{Z_w} \cdot \frac{\lambda}{2a} \cdot \cos\left(\frac{\pi x}{a}\right) \cdot \sin\left(\frac{2\pi z}{h}\right)
\end{aligned} \right\} \cdot e^{j\omega t} \tag{4.32}$$

и представлено на рис. 4.10. Из рис. 4.10,б (время фиксировано) видно, что линии электрического поля направлены между широкими стенками волновода, плотность их максимальна в плоскости, проходящей через центр широкой стенки и уменьшается до нуля у боковых и торцевых стенок ( $E_\tau = 0$  на границе), а вдоль оси  $z$  плотность линий  $\vec{E}$  дважды меняется. Линии магнитного поля (рис. 4.10,в) — замкнутые кривые, расположенные в параллельных плоскостях  $XOZ$ , окружают линии электрического поля, сдвинуты на  $90^\circ$  вдоль оси  $z$  относительно поля  $E$  (влияние множителя  $j$ ). Наличие у поля компоненты  $H_z$  подчеркивает, что это колебание типа  $H$ ; значения  $p=2$ ,  $m=1$  видны на всех рисунках 4.10., значение  $n=0$  — из рис. 4.10,б (поля  $H$ ,  $E$  не изменяются по оси  $y$ ).

Объемные резонаторы с одним видом колебания и даже с несколькими (два — три) практически не пригодны для нагрева диэлектрических материалов большого объема, так как в областях с большой напряженностью поля (рис. 4.10.б) будет

наблюдаться перегрев, а остальные части изделия будут нагреваться слабо. Чтобы получить равномерное распределение поля в резонаторе надо использовать много типов колебаний  $H_{mnp}$  и  $E_{mnp}$ , а для этого следует увеличить в необходимых пределах внутренние размеры (объем) резонатора. Поля этих колебаний, имеющие одну резонансную частоту или близкие, смешиваются друг с другом и образуют общее, почти равномерное в любом направлении и сечении электрическое и магнитное поля. Окончательно выбранные размеры резонатора должны быть такими, чтобы производственные допуски при изготовлении камеры не вызывали существенного нарушения работоспособности системы. Для заданных размеров резонатора существует определенный набор значений  $m, n, p$ , который наилучшим

образом удовлетворяет этому требованию [1].

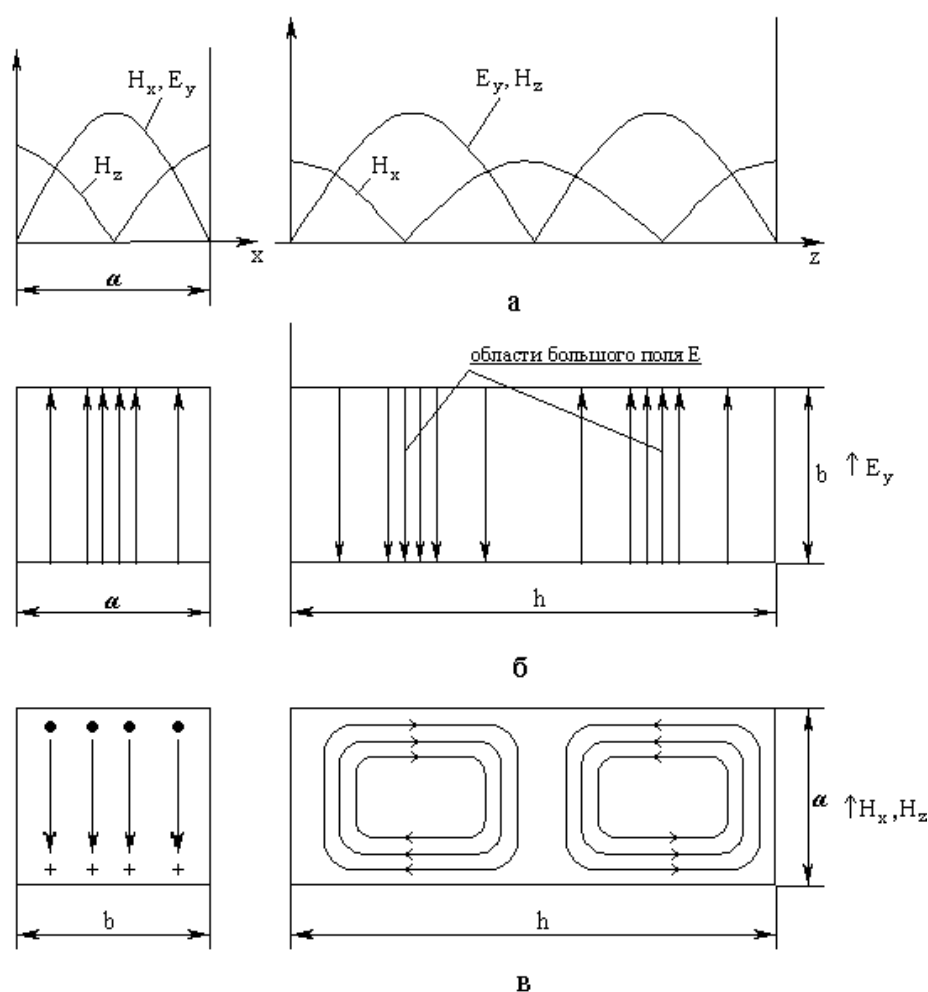


Рис. 4.10. Структура полей  $E$  и  $H$  колебания  $H_{102}$  в прямоугольном резонаторе.

Набор целых чисел  $m, n, p$ , удовлетворяющий (4.34) для заданной частоты  $f_{R0}$ , показывает, сколько видов колебаний может возникнуть в резонаторе. Число колебаний  $N$  в резонаторе объемом  $V=a \times b \times h$  можно определить [6,7] по формуле:

$$N = \frac{4}{3} \pi \cdot \left( \frac{f_{R0}}{c} \right)^3 \cdot V. \quad (4.35)$$

Для нахождения двух из трех индексов, например  $m, n$ , видов колебаний, которые могут возникать в заданном сечении камеры  $a \times b$  можно воспользоваться рис. 4.11 [10]. Таких колебаний может быть несколько десятков. Далее по (4.34) определяются индексы  $p$  для заданной частоты. Значение индексов  $p$  следует округлить до ближайших целых чисел, хотя при этом несколько изменяется и резонансная частота. Пользуясь графиками рис. 4.11, можно для сечения  $a^{1/2}h$  определить индексы  $n, p$  или индексы  $m, p$  для сечения  $a^{1/2}h$ , а затем найти значение третьего индекса.

Резонансная длина волны камеры - резонатора определяется выражением:

$$\lambda_R = \frac{2}{\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{h}\right)^2}}, \quad (4.33)$$

соответственно резонансная частота:

$$f_{R0} = \frac{c}{\lambda_R} = \frac{c}{2} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{h}\right)^2}, \quad (4.34)$$

здесь  $c$  - скорость света.

При разработке камеры СВЧ печи следует рассчитать резонансные частоты  $f_R$  всех возможных колебаний и построить спектр, чтобы убедиться насколько соответствует он заданному диапазону рабочих частот.

Энергия СВЧ поля  $W$ , вводимая в резонатор, определяется соотношением:

$$W = \frac{1}{2} \int_V \epsilon |E|^2 dV,$$

где  $|E|^2$  усредненная величина поля  $E$  всех колебаний, возбуждающихся в резонаторе. Для любого одного  $k$ -го колебания с учетом (4.30) или (4.31) величина  $|E|^2$  должна быть представлена в виде:

$$|E_k|^2 = E_{xk}^2 + E_{yk}^2 + E_{zk}^2.$$

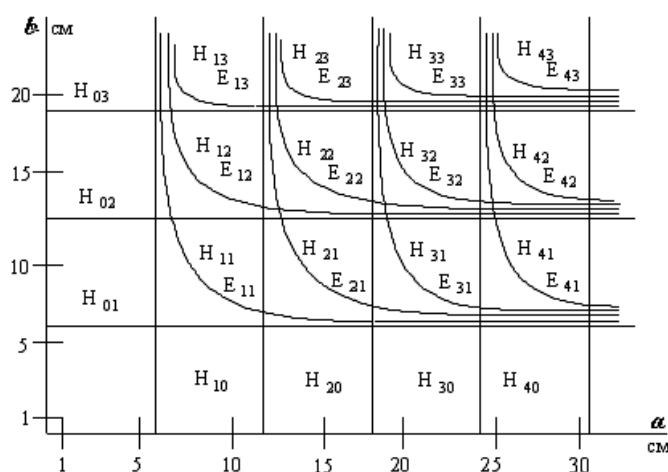


Рис. 4.11. Виды колебаний в прямоугольном волноводе в зависимости от размеров поперечного сечения  $a \times b$  на частоте 2400 МГц.

Результат вычисления интеграла по объему для любого одного из колебаний будет:

$$W = \frac{1}{16} \cdot \epsilon |E_k|^2 V.$$

За период колебания  $T$  в резонатор вводится от генератора мощность  $P_k$ , которая, делится (принимая) поровну, между  $N$  возбужденными одновременно колебаниями. Таких колебаний может быть минимум два, это определяется рассчитанным спектром частот колебаний и вырождением колебаний  $N$  и  $E$  при одинаковых индексах.

$$\frac{P_k T}{N} = \frac{1}{16} \cdot \epsilon |E_k|^2 V,$$

откуда

$$|E_k|^2 = \frac{16 \cdot P_k}{N \cdot \epsilon \cdot fV}. \quad (4.36)$$

**Внесение инородных тел в камеру СВЧ печи** вызывает изменение резонансной частоты  $f_{R0}$  до величины  $f$ . Если объем диэлектрика сравним с объемом камеры, или  $\epsilon_r \gg 1$ , но объем мал, то частота  $f$  может оказаться несовпадающей с рабочей частотой генератора. Этот момент при конструировании печи должен учитываться.

Как правило, объем диэлектрика  $V_\epsilon$  всегда много меньше объема камеры  $V$ . Это позволяет, используя метод малых амплитуд [1], определить изменение частоты камеры при внесении образца:

$$\frac{f - f_{R0}}{f_{R0}} = - \frac{(\epsilon_r - 1) \int |E|^2 dV}{\int_V |E|^2 dV}.$$

Предполагая, что поле в СВЧ камере при возбуждении нескольких колебаний усредняется и не зависит от координат, получим:

$$\frac{f - f_{R0}}{f_{R0}} = - \frac{(\epsilon_r - 1) V_\epsilon}{V}. \quad (4.37)$$

Используя полосу рабочих частот генератора, можно определить допуск на объем диэлектрика  $V_\epsilon$ , и, в соответствии с (4.5), можно оценить толщину разогреваемого образца.

Если в резонатор поместить идеально проводящее не магнитное тело, объем которого  $V_M$  ( $V_M \ll V$ ), то изменение частоты будет характеризоваться выражением:

$$\frac{f - f_{R0}}{f_{R0}} = \frac{-\epsilon_0 \int_{V_M} |E|^2 dV + \mu_0 \int_{V_M} |H|^2 dV}{\epsilon_0 \int_V |E|^2 dV + \mu_0 \int_V |H|^2 dV}. \quad (4.38)$$

Согласно (4.38), металлическое тело, вносимое в  $E$ -поле резонатора, как бы вытесняет электрическое поле. Концентрация электрического поля в новом объеме, чуть меньшем, возрастает, отчего эквивалентная емкость резонатора увеличивается. Частота резонатора  $f_m$  уменьшается по отношению к  $f_{R0}$ . А помещенное в магнитное поле тело, как бы увеличивает магнитное поле, появляется дополнительная параллельная индуктивность, а т.к. общая эквивалентная индуктивность

резонатора уменьшается, то частота резонатора увеличивается. Следовательно, если перемещать периодически проводящее тело из области электрического поля в область магнитного поля, то можно получать небольшое уменьшение или увеличение частоты камеры относительно резонансной. Эта возможность регулировки частоты используется в некоторых камерах СВЧ печи, где применяются так называемые "мешалки" полей. Их роль состоит в настройке резонансной частоты того или иного вида колебания на частоту работающего генератора  $f_R$ . Использование "мешалок" приводит к улучшению равномерности распределения электрического поля по объему камеры [6] разогреваемый объект при этом неподвижен. Если помещать мешалки вблизи ввода энергии в камеру, то они будут отражать некоторую часть энергии. Это также будет приводить к изменению частоты генератора СВЧ.

"Мешалки" выполняются из тонкого алюминия, имеют вид 4-х или 5-и лопастного вентилятора (лопасти посажены на ось свободно). Края каждой лопасти загнуты (высота загиба 2-3 см) и слегка развернуты (рис. 4.12,а).

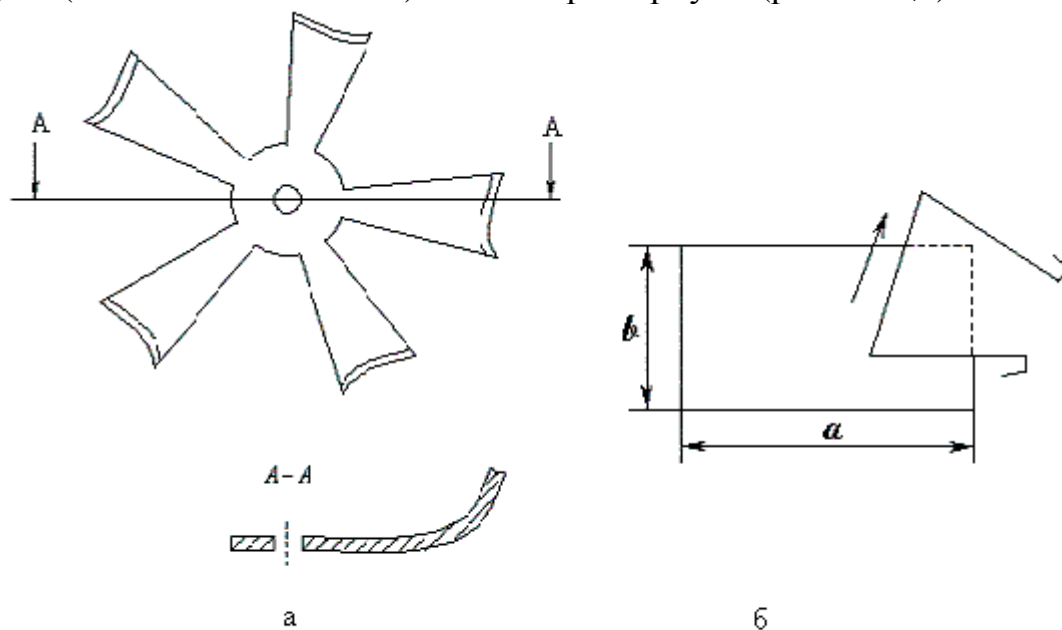


Рис. 4.12. Вид "мешалок"

Загнутыми краями лопастей "мешалка" обращена к стенке камеры, в которой расположен ввод энергии в виде открытого конца волновода, впаянного краями в стенку в камеру. "Мешалки" от рабочего объема камеры частично экранированы алюминиевой пластинкой с отверстием  $150 \times 150$  мм, симметрично расположенным против открытого волноводного ввода. Пластинка лежит свободно на диэлектрическом фторопластовом поддоне. Поток теплого воздуха после охлаждения магнетрона поступает в камеру на лопасти "мешалок" и приводит их в движение. Лопасти вращаются и при этом частично как бы перекрывают открытый конец волновода (рис. 4.12,б). Изменяется сопротивление излучения и возникает в волноводном тракте отраженная волна.

В последнее время в СВЧ печах вместо мешалок используются вращающиеся столики, это позволило уменьшить объем камеры. Вращающийся столик даже при малом числе колебаний позволяет равномерно разогревать объект и плавно изме-

нять сопротивление камеры в плоскости ввода энергии. Изменилась конструкция ввода энергии в резонатор. Камера внутри покрыта фторопластовой пленкой. Столик вращающийся, посуда выполняется из стекла.

Отметим, что при нагревании изменяются габариты, плотность вещества, испаряется вода, что приводит к уменьшению параметров  $\epsilon_r$  и  $\text{tg}\delta$  (рис. 4.13) и коэффициента затухания, все это изменяет сопротивление камеры  $Z_k$  на входе открытого конца волновода.

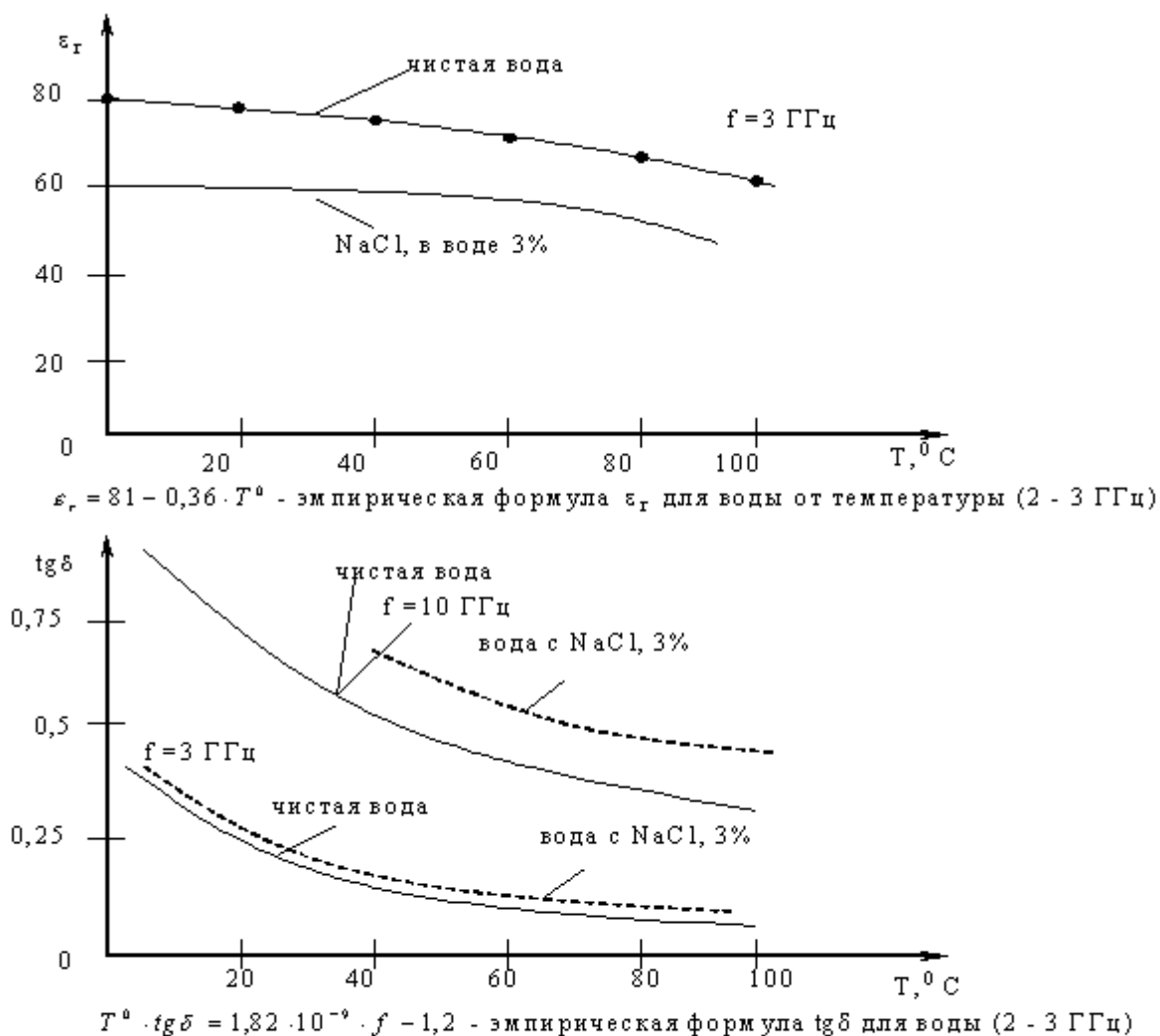


Рис. 4.13

**Оценка мощности, поступившей в камеру нагрева**, выполняется легко, если известны волновое сопротивление волновода и входное сопротивление СВЧ камеры. Стоячие волны в тракте между камерой и магнетроном характеризуются коэффициентом стоячей волны  $K_{ст}$ , равным:

$$K_{ст} = \frac{U_{\max}}{U_{\min}},$$

где  $U_{\max}$  - напряжение в пучности;

$U_{\min}$  - напряжение в узле стоячей волны.

Допустимая величина  $K_{ст}$  для работающих в СВЧ печах магнетронов  $1 \leq K_{ст} < 3$ , иначе сокращается срок службы генератора. В случае  $K_{ст} > 3$  отраженная волна возвращается в магнетрон, изменяет структуру поля в его пространстве взаимодействия и отдача энергии электронами нарушается. При неизменной величине подводимой мощности не израсходованная электронами энергия попадает на анод и частично на катод магнетрона, которые сильно перегреваются, что сокращает срок службы. К тому же, отраженная волна вызывает изменение частоты и уменьшение, в соответствии с нагрузочными характеристиками, мощности магнетрона на выходе. Изменение частоты приведет к нарушению числа и видов колебаний в СВЧ печи в полосе частот  $2450 \pm 50$  МГц. А это вызовет или не равномерный разогрев диэлектрика в камере, или его недостаточную температурную обработку.

Предположим, что сопротивление открытого конца волновода  $Z_w$  и входное сопротивление камеры  $Z_{kr}$  равны, в тракте волновода согласованный режим,  $K_{ст} = 1$ , тогда в камеру поступает мощность  $P_k$ , равная мощности генератора  $P_{ген}$ . В случае стоячих волн (будет при изменении сопротивления камеры  $Z_k$ ) между мощностями  $P_k$  и  $P_{ген}$  устанавливается соотношение:

$$P_k = P_{ген} (1 - |\Gamma|^2), \quad (4.39)$$

где  $|\Gamma| = \frac{K_{ст} - 1}{K_{ст} + 1}$  - величина коэффициента отражения.

Сопротивление камеры  $Z_{kr}$  тоже определяется [4] параметром  $|\Gamma|$ :

$$|Z_{kr}| = Z_w \cdot \frac{1 + |\Gamma|}{1 - |\Gamma|} = Z_w \cdot K_{ст}, \quad (4.40)$$

здесь величина волнового сопротивления волновода

$$Z_w = 2 \frac{b}{a} \cdot \frac{120\pi}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}}. \quad (4.41)$$

По нагрузочным характеристикам при требуемой мощности  $P_{ген}$  (4.39), которая бы в камере обеспечивала величину мощности  $P_k$  при заданной величине  $K_{ст} < 3$ , можно определить тип магнетрона. Из (4.40) определяется сопротивление камеры  $Z_{kr}$  в рабочем режиме.

Величина электрического поля  $E$  в камере может быть определена из соотношения:

$$P_k = \frac{1}{2} \cdot \frac{|E|^2}{Z_{kr}}. \quad (4.42)$$

**Дверца СВЧ камеры** должна удовлетворять следующим требованиям: хорошая видимость продукта и отсутствие излучения ЭМ поля.

Обзор продукта обеспечивается окнами, а защита от СВЧ излучения через окна и контакты дверцы с камерой - специальными их конструкциями.

Для предотвращения излучения непосредственно из окна последнее либо выполняется из диэлектрических материалов с металлическими присадками (свинцовые стекла), либо делается из стекла "пирекс" и покрывается перфорированной металлической пленкой - сеткой. Коэффициент прозрачности такой плен-

ки высок, поэтому обрабатываемый продукт хорошо виден. Коэффициент излучения определяется размерами отверстий, расстоянием между ними и может быть определен (в децибелах) по номограмме рис. 4.14 [10] или по формуле, приведенной на том же рисунке.

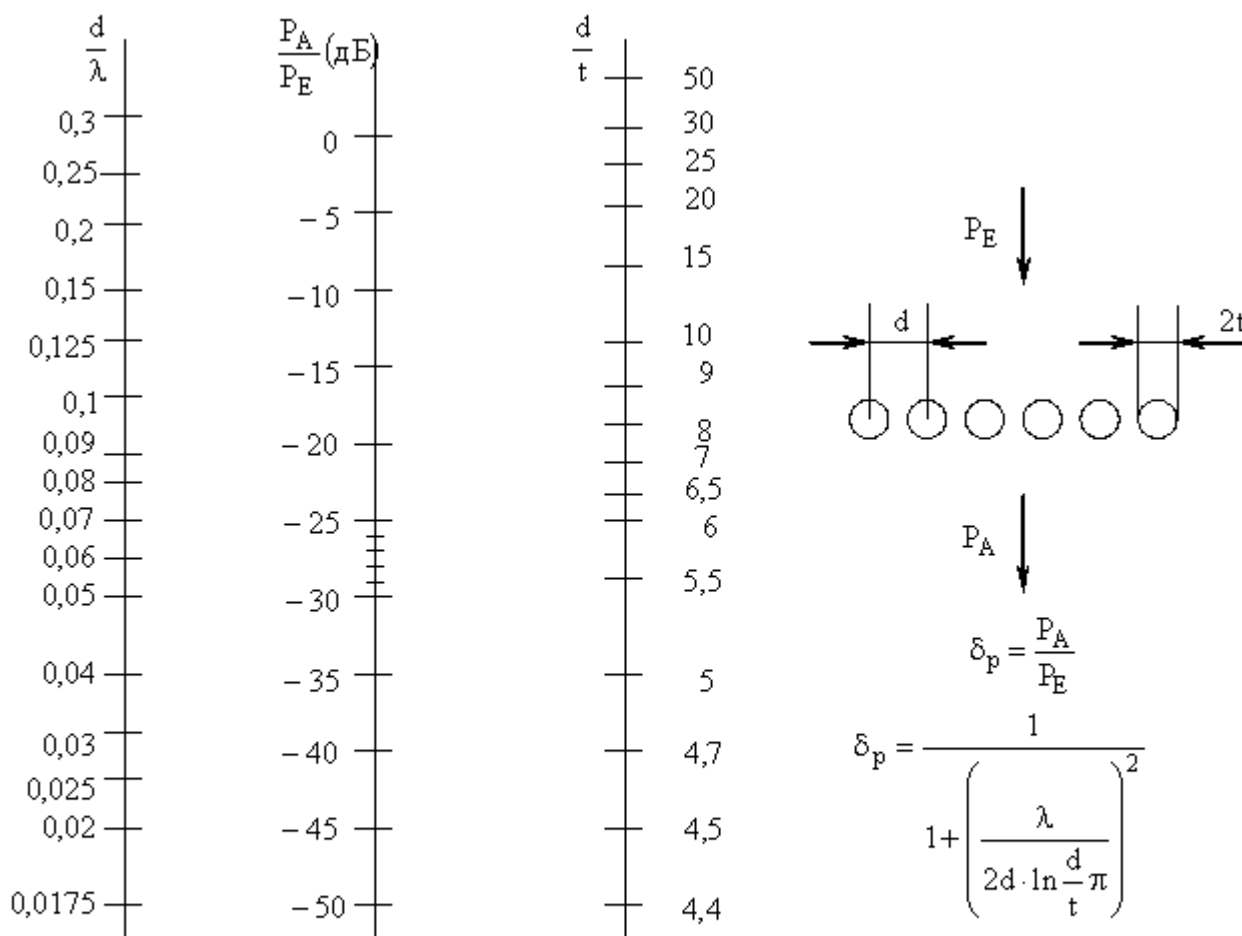


Рис. 4.14. Номограмма для определения паразитного электромагнитного излучения решетки. Формула  $\delta_p$  дает ослабление мощности в разгах

Чтобы дверки СВЧ печей были герметичны по высокой частоте, используются бесконтактные дверные ловушки (рис. 4.15).

В точках  $k-f$  (рис. 4.15,г) короткое замыкание, в точках  $c-e$  - дверной контакт, в точках  $a-b$  - вход в ловушку.

Входное сопротивление  $Z_{вх}$  однородного участка линии определяется [5] в виде:

$$Z_{вх} = Z_c \cdot \frac{Z_H + jZ_c \cdot \operatorname{tg}\beta l}{Z_c + jZ_H \cdot \operatorname{tg}\beta l}, \quad (4.43)$$

где  $l$  - длина однородного участка линии;

$Z_c$  - сопротивление нагрузки линии на конце участка;





Короткозамкнутая на конце линия длиной  $l = \frac{\lambda}{4}$ , в точках  $e - c$ , имеет входное сопротивление  $Z_{bxce} = jZ_c \cdot \operatorname{tg}\left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{4}\right)$ , равное бесконечности. В этом месте последовательно стоит контакт дверцы с сопротивлением  $R_k$ , которое не может изменить бесконечную величину. Последующий (рис. 4.15, г) участок линии ( $ec - ab$ ), длиной  $l = \frac{\lambda}{4}$  и нагруженный  $Z'_n = Z_{bxce} = \infty$ , имеет входное сопротивление в точках  $a-b$ ,

$$\text{согласно (4.43), равно } Z_{\text{вх}ab} = jZ_c \cdot \operatorname{ctg}\left(\frac{2\pi}{\lambda} \cdot \frac{\lambda}{4}\right) = 0.$$

Ловушка, таким образом, в плоскости прилегания дверцы к камере обеспечила нулевое сопротивление на рабочей частоте, что позволяет замкнуться СВЧ токам через прорезь камера – дверца накоротко. Можно в плоскости контакта (место примыкания дверцы к камере), при длине его  $l < \frac{\lambda}{10}$ , учитывать появляющуюся сосредоточенную емкость. Учет приведет к тому, что участок  $ab - ce$  станет линией, нагруженной на паразитную емкость и бесконечное сопротивление. Тогда длину этого участка можно укоротить.

Таблица 4.2.

Частота 2450 МГц			
	$\epsilon_r$	$\rho$ , Ом·см	$\operatorname{tg}\delta$
Кровь	25	2,5	11,76
Мышечная ткань	50	4,5	2,9
Жировая ткань	23	6 ÷ 10	6 ÷ 3

**Допустимые величины воздействующего СВЧ излучения на организм человека** установлены следующими: верхний предел во всем СВЧ диапазоне принят не более  $10 \text{ мВт/см}^2$  в течение одного часа; безопасный уровень характеризуется плотностью излучения  $1 \text{ мВт/см}^2$  в течение 8 часов, а уровень при котором не наблюдается каких-либо биологических нарушений равен  $10 \text{ мкВт/см}^2$  [6]. Разогрев мышечных и жировых тканей и кровеносных сосудов не одинаков при постоянной поступившей мощности, так как их параметры  $\epsilon_r$  и  $\operatorname{tg}\delta$  (табл. 4.2) различны. Облучение предельными дозами может привести к местному повышению температуры, особенно чувствительных к превышению тепла участков. В числе таких органов - хрусталик глаза, желчный пузырь, мочевого пузыря и яички.

## 4.2. Измерение СВЧ энергии

### 4.2.1. Калориметрические методы

Применение калориметрического метода для измерения СВЧ мощности является частным случаем его использования в области физических исследований [1, 2]. Широкое применение этого метода оправдано простотой превращения энергии электрического тока в теплоту. Устройство калориметрических ваттметров СВЧ весьма разнообразно и определяется в основном уровнем мощности, диапазоном частот и требуемой точностью.

Все калориметры можно разделить на две большие группы. Это калориметры с переменной температурой, в которых мощность определяется по изменению температуры рабочего тела калориметра, и калориметры с постоянной температурой. Тепловые процессы, происходящие в калориметре, в зависимости от оболочки калориметра могут носить адиабатический или изотермический характер. Для калориметрического тела с неидеальной изоляцией уравнение теплового равновесия имеет вид:

$$P = c_{t^{\circ}} d\theta/dt + \theta/R_{\text{тпл}}, \quad (4.44)$$

где  $P$  — мощность, рассеиваемая в рабочем теле;  $c_{t^{\circ}}$  — теплоемкость рабочего тела;  $d\theta/dt$  — скорость изменения температуры рабочего тела;  $\theta$  — разность температур рабочего тела и окружающей среды;  $R_{\text{тпл}}$  — тепловое сопротивление между рабочим телом и окружающей средой.

Общее решение уравнения (4.44) запишем как:

$$\theta = R_{\text{тпл}} P [1 - \exp(-t/R_{\text{тпл}} c_{t^{\circ}})], \quad (4.45)$$

где  $R_{\text{тпл}} c_{t^{\circ}}$  — тепловая постоянная времени калориметрической системы.

При идеальной теплоизоляции рабочего тела уравнение (4.44) упрощается:

$$P = c_{t^{\circ}} d\theta/dt \quad (4.46)$$

Решение уравнения (4.46) дает линейную зависимость температуры рабочего тела от времени (рис. 4.16), если  $P = \text{const}$ ,  $c_{t^{\circ}} = \text{const}$ .

При конечном, но достаточно большом значении теплового сопротивления скорость изменения температуры в начальный период практически такая же, как при  $R_{\text{тпл}} = \infty$  затем уменьшается и при  $t \gg R_{\text{тпл}} c_{t^{\circ}}$  стремится к нулю. При малых значениях  $R_{\text{тпл}}$  время установления температуры уменьшается, но одновременно уменьшается и степень разогрева рабочего тела.

На практике применяют устройства как с большим, так и с малым значением  $R_{\text{тпл}}$ . В первом случае величину измеряемой мощности определяют исходя из скорости изменения температуры рабочего тела  $d\theta/dt$  в момент подачи мощности, а во втором — по разности температур  $\theta$  в установившемся режиме.

### Калориметры с переменной температурой

К этой группе приборов относятся как статические, так и проточные калориметрические ваттметры. В статических приборах рабочее тело приемного

преобразователя, где энергия СВЧ колебаний превращается в тепловую, неподвижно и в процессе измерения не изменяет формы и физических свойств.

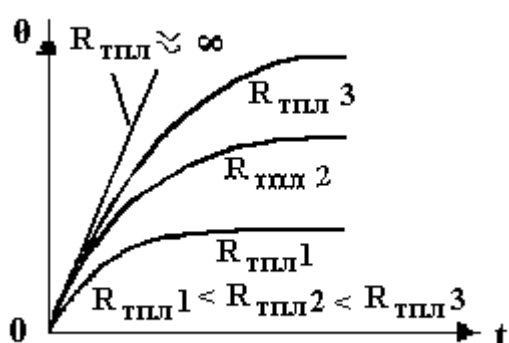


Рис. 4.16. Зависимость температуры рабочего тела от времени при различных значениях теплового сопротивления

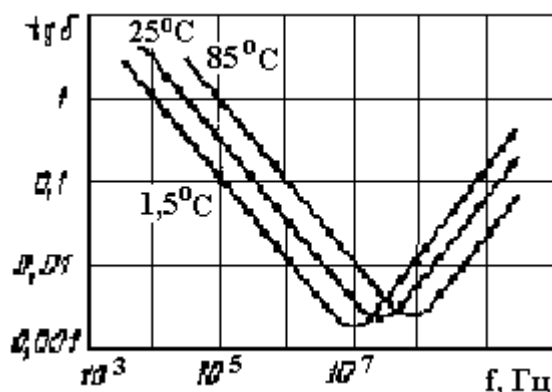


Рис. 4.17. Зависимость тангенса угла диэлектрических потерь дистиллированной воды от частоты.

В проточных калориметрах содержится жидкость, играющая роль переносчика тепла от рабочего тела во внешнюю среду. Если для этих целей используют дистиллированную воду или растворы на ее основе, то в силу ее большой поглощающей способности на частотах свыше 500 МГц ( $\text{tg} \delta > 0,01$ , рис. 4.17) она может одновременно выполнять роль объемного поглотителя мощности СВЧ в приемных преобразователях (калориметрических нагрузках) и является, таким образом, рабочим телом калориметра.

Рабочим телом для статических калориметров может служить как вода, так и твердые объемные и пленочные поглотители. В случае использования твердых материалов приборы называют «сухими» калориметрами. При высокой степени тепловой изоляции калориметрического тела величину измеряемой мощности определяют умножая известное значение теплоемкости рабочего тела на скорость изменения его температуры по формуле (4.46). В формулу подставляют значения теплоемкости в джоулях на градус Цельсия. Если теплоемкость выражена на градус Цельсия, то справедливо выражение

$$P_{\text{СВЧ}} = 4,28 c_{t^\circ} d\theta / dt \quad (4.47)$$

#### 4.2.2. БолOMETрические методы

В основу приборов, использующих болOMETрический метод, положено изменение сопротивления резистивного термочувствительного элемента под действием энергии СВЧ, превращенной им в теплоту. Изменение сопротивления термочувствительного элемента, пропорциональное подводимой СВЧ мощности, измеряют с помощью мостовых измерительных устройств, в одно из плеч которых он включен. На СВЧ применяют два вида термочувствительных элементов — болOMETры и термисторы (терморезисторы). В зависимости от вида применяемого термочувствительного элемента ваттметры называют болOMETрическими или термисторными.

Основными узлами болометрических и термисторных ваттметров являются приемные преобразователи, измерительное и отсчетное устройства.

**Приемные преобразователи.** Приемный преобразователь — это отрезок СВЧ тракта, на конце которого включен термистор или болометр. Эти элементы выполняют роль оконечной нагрузки и именно в них энергия СВЧ превращается в тепловую.

Термисторы изготовляют из полупроводниковой массы в виде бусинки диаметром 0,2 — 0,5 мм (рис. 4.18, а) или цилиндра диаметром 0,2—0,3 мм и длиной 1—1,5 мм (рис. 4.18, б). Полупроводниковая масса представляет собой порошкообразную смесь окислов меди, марганца, кобальта, титана и др., спекаемую в определенной среде. В электрическую цепь термистор включают с помощью вваренных в его тело выводов), 3 из платины, платино-иридиевого или платино-ролиевого сплава. Диаметр

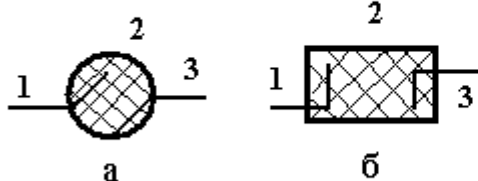


Рис. 4.18. Бусинковый (а) и цилиндрический (б) термисторы: 1,3 — выходы; 2 — рабочее тело.

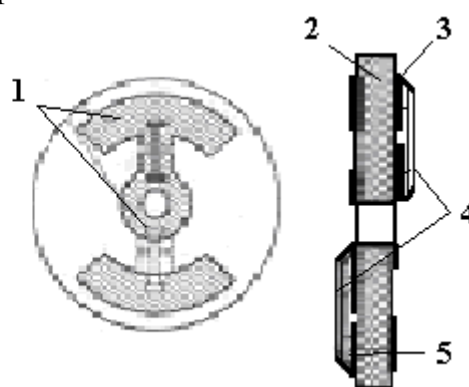


Рис. 4.19. Вставка с нитевидными пленочными болометрами для коаксиальных приемных преобразователей: 1 — электроды 2 — слюдяной диск; 3-контакты; 4- болометры; 5-опора.

выводов 25—50 мкм. Чтобы увеличить проводимость полупроводниковой массы, в нее добавляют порошок меди. Для жесткости конструкции термисторы типов Т8, Т9, ТШ, ТВ помещают в стеклянный баллон диаметром до 3 мм и длиной до 10 мм с проволочными выводами диаметром 0,8 мм. Широкое распространение получили также безбаллонные термисторы типа ТШ-1, терморезисторы СТЗ-18, "ТЗ-29, на базе которых созданы высокочастотные термисторные коаксиальные вставки.

Болометры представляют собой тонкую металлическую проволоку длиной 0,8—1,2 мм (проволочные болометры) или тонкую металлическую пленку (из платины, палладия), нанесенную в вакууме на основание (подложку) из стекла или слюды (пленочные болометры). Для изготовления проволочных болометров применяют платиновую проволоку диаметром 1 мкм. Пленочные болометры бывают нитевидными (рис. 4.19) и плоскими (рис. 4.20). В нитевидных тонкопленочных болометрах основанием служит нить из стекловолокна диаметром 3 мкм, а в плоских — слюда толщиной 30—50 мкм. Нитевидные болометры устанавливают на опорах в зазоре между серебряными электродами, которые наносят на слюдяной диск (основание), вжиганием (рис. 4.19). Механиче-

ское крепление и электрический контакт болометров с электродами осуществляют с помощью токопроводящей массы. Опора из нитей стекловолокна обеспечивает зазор между основанием и болометром, необходимый для согласования и создание определенного теплового режима.

Для включения в линию передачи болометры имеют серебряные или платиновые контакты. Серебряные контакты изготавливают вжиганием серебряной пасты или напылением в вакууме серебряной пленки на подслои из никеля или нихрома, платиновые – напылением в вакууме пленки платины.

Рассмотрим основные характеристики болометров и термисторов. К ним относятся:

- сопротивление болометра (термистора) в рабочей точке  $R_{t^0}$  (Ом) — сопротивление постоянному току, при котором болометр (термистор) согласуется с волновым сопротивлением линии передачи;
- температурный коэффициент сопротивления

$$\alpha_{t^0} = \frac{1}{R_{t^0}} \frac{dR_{t^0}}{dt^0} \quad (4.48)$$

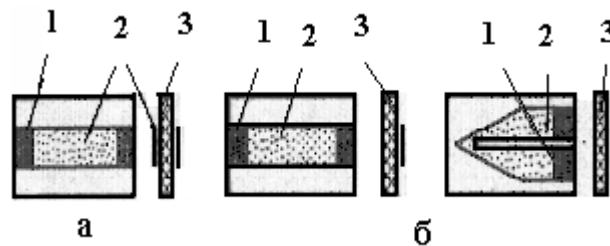


Рис. 4.20. Виды плоских болометров для коаксиальных (а) и волноводных (б) трактов: 1 — контакт; 2 — термочувствительная пленка; 3 — основание из слюды.

т. е. относительное изменение сопротивления в результате изменения температуры,  $1/^\circ\text{C}$ ;

— постоянная рассеяния

$$h_{t^0} = dP / d\theta \quad (4.49)$$

т. е. отношение приращения мощности, рассеиваемой в болометре (термисторе), к возникающему в результате этого повышению его температуры по сравнению с температурой окружающей среды,  $\text{Вт}/^\circ\text{C}$ ;

— чувствительность

$$S_{t^0} = dP / d\theta \quad (4.50)$$

т. е. отношение изменения сопротивления болометра (термистора) к изменению (приращению) мощности на нем,  $\text{Ом}/\text{Вт}$  или  $\% / \text{Вт}$ ;

— тепловая постоянная времени  $\tau$ , с. Характеризует скорость установления температуры болометра (термистора) при изменении его теплового режима и выражается временем, в течение которого предварительно нагретый болометр (термистор) остывает в  $e$  раз по сравнению с первоначальной разностью температур относительно окружающей температуры;

— максимально допустимая мощность рассеяния  $-P_{\text{срмакс}}$ , Вт.

Это мощность, которую термистор (болومتر) может рассеивать в течение длительного времени при условии, что необратимые изменения его характеристик останутся в пределах норм.

В общем случае, как болометрам, так и термисторам присуща нелинейная зависимость изменения сопротивления от изменения уровня рассеиваемой мощности и температуры окружающей среды. Для термисторов эту зависимость можно записать в виде:

$$R_{t^{\circ}} = R_{\infty} \exp B_{\tau} / t_K^{\circ} \quad (4.51)$$

Эта зависимость может быть записана как:

$$R_{t^{\circ}} = R_{t^{\circ}_{\text{окр}}} \exp \frac{B_{\tau} (t_{K \text{ окр}}^{\circ} - t_K^{\circ})}{t_{K \text{ окр}}^{\circ} - T_K^{\circ}} \quad (4.52)$$

Дифференцируя выражение (4.51) или (4.52), можно определить относительный температурный коэффициент сопротивления

$$\alpha_{t^{\circ}} = -B_{\tau} / (t_K^{\circ})^2 = -B_{\tau} / (t_{K \text{ окр}}^{\circ} + \theta)^2. \quad (4.53)$$

Из выражения (4.53) следует, что температурный коэффициент-сопротивления термистора является величиной отрицательной и существенно изменяется при изменении температуры термистора. Однако при изменении сопротивления термистора в пределах  $R_{t^{\circ}} \pm (5—10)\%$  температурный коэффициент можно принять постоянным, что позволит сравнивать характеристики термисторов.

Зависимость сопротивления проволочного болометра от температуры в общем виде может быть записана как:

$$R_{t^{\circ}} = R_{t^{\circ}_{\text{окр}}} (1 + \alpha_{t^{\circ}} \theta), \quad (4.54)$$

где  $\theta = t_K^{\circ} - t_{K \text{ окр}}^{\circ}$ .

Постоянная рассеяния  $h_{t^{\circ}}$  термочувствительного элемента определяет чувствительность и дрейфовые характеристики ваттметра. Дрейф показаний отсчетного прибора при изменении температуры окружающей среды прямо пропорционален постоянной рассеяния болометра (термистора):

$$\Delta P = h_{t^{\circ}} \Delta \theta. \quad (4.55)$$

Как термисторы, так и болометры используют для измерения малых уровней мощности непрерывных сигналов и среднего значения мощности импульсно-модулированных СВЧ сигналов. На практике предпочтение отдают термисторам из-за их более высокой чувствительности и большей энергии перегорания при одном и том же максимальном пределе измерений среднего значения мощности импульсно-модулированных сигналов. Благодаря малой постоянной времени, проволочные и нитевидные болометры на стекловолокне применяют

для измерения малых уровней импульсной мощности, На проволочных болометрах, имеющих высокую стабильность параметров, строят образцовую аппаратуру. Плоские пленочные болометры позволяют расширить динамический диапазон средних уровней мощности до 1000 мВт. Для уровней мощности ниже 10 мВт такие болометры малопригодны из-за большой постоянной рассеяния, обуславливающей значительный дрейф показаний при измерениях.

### 4.2.3. Термоэлектрические методы

Термоэлектрические методы основаны на преобразовании СВЧ энергии в тепловую с помощью высокочастотных дифференциальных термопар прямого или косвенного нагрева и измерении возникающей в результате нагрева термо-ЭДС пропорциональной рассеиваемой термопаре СВЧ мощности. Термопары выполняют одновременно согласованной нагрузки и дифференциального термометра. Зависимость между измеряемой мощностью и величиной термо-ЭДС. выражается соотношением

$$P_{\text{СВЧ}} \approx U_{\text{тэ.д.с.}} / K_{\text{прб СВЧ}}, \quad (4.56)$$

где ( $U_{\text{тэ.д.с.}}$  - напряжение термо-ЭДС на выходе термопары, мВ;  $K_{\text{прб СВЧ}}$  - коэффициент преобразования термопары, мВ/мВт. Термоэлектрический метод, как и болометрический, применяют в основном для измерения мощности малых уровней, однако он имеет ряд преимуществ перед болометрическим. Так как величина термо-ЭДС пропорциональна подводимой СВЧ мощности и практически не зависит от температуры окружающей среды, в термоэлектрических ваттметрах отсутствуют элементы и узлы для температурной компенсации. Приемные преобразователи термоэлектрических ваттметров не требуют длительного времени прогрева и при высокой чувствительности

термопар могут работать непосредственно со стрелочным индикатором. Поскольку термопара не требует начального подогрева, как термистор или болометр, потребление ваттметра от источников питания мало, что особенно важно в случае питания прибора от первичных источников (например, аккумуляторов).

**Приемные преобразователи термоэлектрических ваттметров.** Основным элементом приемного преобразователя является термопара (блок термопар). Материалы для ветвей дифференциальных термопар выбирают с таким расчетом, чтобы обеспечивались линейная температурная зависимость термо-ЭДС, малый температурный коэффициент сопротивления и высокая чувствительность. Наибольшее распространение получили термопары висмут-сурьма, копель-сурьма, хромель-копель. Чтобы расширить динамический диапазон измеряемых мощностей, две термопары и более соединяют последовательно по постоянному току. При этом по высокой частоте их соединяют параллельно и для оптимального согласования их общее активное сопротивление должно быть равным характеристическому сопротивлению линии передачи.



В диапазоне СВЧ в основном применяют пленочные термопары, представляющие собой тонкие металлические пленки, напыленные в вакууме на слюдяные, стеклянные или иные диэлектрические подложки.

К основным требованиям, предъявляемым к приемным термоэлектрическим преобразователям, относятся: хорошее согласование с СВЧ трактом; постоянство коэффициента эффективности в диапазоне частот и уровней сигнала при абсолютном его значении, близком к единице; высокий коэффициент преобразования; устойчивость к перегрузкам и воздействиям внешних факторов.

Коаксиальный преобразователь (рис. 4.22) состоит, из отрезка коаксиала с разделительной емкостью  $C_1$  в центральном проводнике, вставки с двумя термопарами  $R_T$  в кожуха с выходным разъемом.

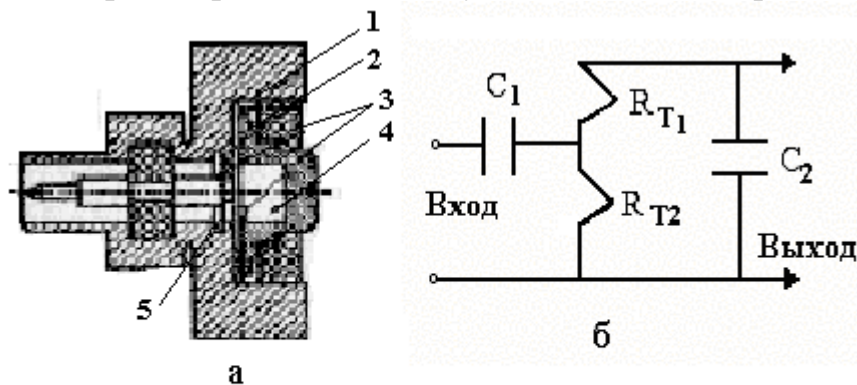


Рис. 4.22. Упрощенная конструктивная (а) и принципиальная электрическая (б) схемы коаксиального термоэлектрического преобразователя: 1 — вставка; 2 — отрезок коаксиала; 3 — термопары; 4 — заглушка; 5 — проточка.

Вставку 1 согласуют с трактом, подбирая размеры согласующей камеры в заглушке 4, и проточки 5 в корпусе отрезка линии передачи, а также рабочее сопротивление термопар.

### **Особенности согласования пленочных нитевидных термопар с СВЧ трактом. Коэффициент эффективности термоэлектрических преобразователей.**

С повышением частоты на согласование термоэлектрического преобразователя начинает заметно влиять реактивная составляющая его полного сопротивления, определяемая индуктивностью пленочных термопар, емкостью контактной системы и другими неоднородностями, возникающими при включении вставки в тракт. Индуктивность нитевидных пленочных термопар уменьшают их допустимым укорочением и соответствующим расположением над проводящими пленочными электродами. Согласуют термопары с помощью проточек в корпусе преобразователя в непосредственной близости к месту включения термопар. Образующей при этом резонансной системе свойственна низкая добротность из-за шунтирующего действия сопротивления термопары. Это определяет отсутствие значительных экстремумов КСВ преобразователей в широкой полосе частот (рис. 4.23).

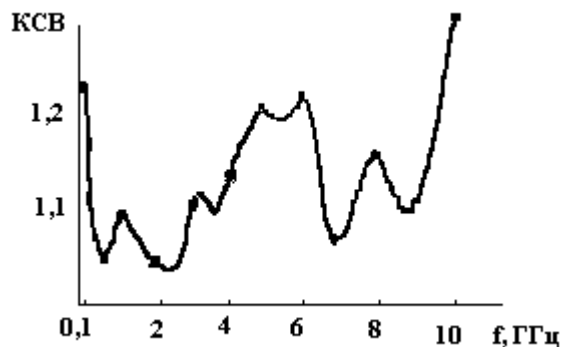


Рис.4.23. Частотная зависимость коаксиального термоэлектрического преобразователя М5-51.

В коаксиальных преобразователях наиболее часто применяют симметричное включение двух термопар. При этом сопротивление каждой термопары выбирают равным удвоенному волновому сопротивлению тракта. Так, для коаксиального тракта 50 Ом применяют термопары с рабочим сопротивлением 100 Ом каждая.

Оптимальное согласование короткой нитевидной пленочной термопары в волноводе обеспечивают выбирая рабочее сопротивление термопары приблизительно равным характеристическому сопротивлению тракта в месте ее включения. Рабочее сопротивление термопары для волноводных преобразователей выбирают таким, чтобы их общее сопротивление на СВЧ было около 100 Ом. Для согласования применяют плавные переходы с регулярного волновода на П-образный характеристическим сопротивлением 100 Ом на средней частоте рабочего диапазона. Волноводный преобразователь оканчивается короткозамкнутым отрезком прямоугольного или круглого (запредельного) волновода.

В термоэлектрический преобразователь входят элементы, приводящие к нежелательным потерям. Это разъем, держатели, диэлектрические опоры, диафрагмы для компенсации реактивностей, конструктивные конденсаторы и другие неоднородности. В связи с этим некоторая доля мощности не поступает непосредственно на термопару, что уменьшает коэффициент эффективности преобразователя. Для термоэлектрических преобразователей коэффициент эффективности представляет отношение коэффициента преобразования на СВЧ к коэффициенту преобразования на переменном токе низкой частоты (НЧ).

Поскольку термопары СВЧ относятся к термопарам прямого нагрева, их калибровка на постоянном токе является невозможной. В области нижних частот (до 3—4 ГГц) для коаксиальных преобразователей М5-51 отличие  $K_s$  от единицы незначительно (до 1—2%) и им зачастую пренебрегают. На более высоких частотах  $K_s$ , как для коаксиальных, так и для волноводных преобразователей может значительно отличаться от единицы, и при измерениях, его следует учитывать. Коэффициент термо-ЭДС. ( $dU_{ТЭ.д.с.}/dt^\circ$ ) материалов, применяемых для изготовления термопар, является нелинейной функцией температуры. Компенсируют нелинейность температурной характеристики термопары ( $U_{ТЭ.д.с.} = f(t^\circ)$  применяя для ветвей материалы с различной по знаку нелинейностью коэффициентов термо-э. д. с.

Однако для измерения мощности СВЧ практический интерес представляет не температурная, а амплитудная характеристика преобразователя ( $U_{ТЭ.д.с.} = \Phi(P)$ ). Линейность амплитудной характеристики в общем случае не обеспечивается при линейной температурной характеристике, так как при нагреве термопа-

ры преобразуемой мощностью изменяется коэффициент теплоотдачи и, следовательно, температура «горячего» спая нелинейно зависит от подводимой мощности. Поскольку качество термопреобразователя определяется степенью линейности его амплитудной характеристики, то это обуславливает соответствующие требования к материалам ветвей термопар и конструкции преобразователя, при выполнении которых рассмотренные нелинейности взаимно компенсируются в заданных интервалах мощностей и температур и при этом обеспечивается требуемая величина коэффициента преобразования.

**Возможности и достоинства метода.** Термоэлектрический метод используют в диапазонах метровых, дециметровых, сантиметровых и миллиметровых длин волн. С его помощью измеряют мощности от десятков нановатт до единиц ватт при использовании аттенуаторов и направленных ответвителей.

Погрешность измерения СВЧ мощности при использовании этого метода определяется следующими основными составляющими:

- погрешностью измерительного блока  $\delta_1$  включающей временной дрейф и погрешность отсчетного устройства (может быть сведена до  $\pm (1-2)\%$  и менее);

- погрешностью определения коэффициента преобразования на переменном токе низкой частоты  $\delta_2$ , зависящей от точности используемых методов измерения действующего значения напряжения (при использовании вольтметра класса 0,5 погрешность составит  $\pm (1-2)\%$ ; для схем со встроенным калибратором  $\delta_2$  — погрешность самокалибровки);

- погрешностью, обусловленной нелинейностью термопары в динамическом диапазоне,  $\delta_3$  (обычно составляет  $\pm (1-3)\%$ );

- погрешностью определения коэффициента эффективности преобразователя  $\delta K_3$  (при использовании калибраторов может принимать значение  $\pm (1,5-5)\%$ );

- погрешностью, обусловленной отражением мощности от входа преобразователя,  $\delta_4$ ;

- погрешностью из-за рассогласования  $\delta_{\text{расс}}$ . Расчетное значение погрешности данного метода измерения мощности составит  $\delta = \pm (2,5-10)\%$  без учета погрешности рассогласования.

К достоинствам метода следует отнести: применимость в широком диапазоне частот; малое время установления показаний, что позволяет использовать метод в автоматизированных системах; малую зависимость результата измерений от температуры окружающей среды; малое время подготовки к работе; технологическую возможность изготовления термопар на различные уровни мощности.

Недостатками метода являются: ограниченный сверху динамический диапазон; невысокая устойчивость к перегрузкам, ограничивающая применение термоэлектрических преобразователей при измерении средней мощности импульсно-модулированных колебаний. Пленочные нитевидные термопары, обладая малой тепловой инерционностью, уже при относительно небольшом значении энергии импульса перегорают. Так, например, термопары преобразователей М5-51 допускают подачу на них импульсной мощности: 10 Вт — при длительности импульсов 1 мкс, 4 Вт — при длительности импульсов 4 мкс, 2,5

Вт -при длительности импульсов  $10 \text{ мкс}$  и среднем значении мощности порядка  $10 \text{ мВт}$ .

Приведем краткие характеристики термоэлектрического ваттметра МЗ-21. Прибор предназначен для измерения мощности в коаксиальных и волноводных трактах. Он укомплектован одним коаксиальными и пятью волноводными преобразователями на диапазон частот  $0,1 - 37,5 \text{ ГГц}$ . Пределы измерений - от  $10 \text{ мкВт}$  до  $10 \text{ мВт}$ , погрешность  $\pm (10 + 100/P_B) \%$ , где  $P_B$  — показание прибора,  $\text{мкВт}$ .

## РАЗДЕЛ ВТОРОЙ

### Квантовые приборы СВЧ и оптического диапазона

#### ВВЕДЕНИЕ

Более 40 лет назад появилась новая область физики и техники – квантовая электроника. Квантовые генераторы и усилители из уникальных лабораторных установок превратились в серийно выпускаемые промышленностью приборы, которые находят все более широкое применение в самых различных областях науки и техники. Физика, химия, астрономия, медицина, технология производства, вычислительная техника, системы управления, связи, локации и навигации, специальная фотография, измерительная техника – таков далеко не полный перечень областей применения квантовых усилителей и генераторов. Перспективы их использования считаются настолько широкими, что трудно назвать область науки или техники, где квантовые устройства не будут применяться [2].

Кратко остановимся на основных достоинствах квантовых усилителей и генераторов, обусловивших их необычно широкую популярность. Для квантовых генераторов СВЧ диапазона волн характерна чрезвычайно высокая стабильность частоты ( $10^{-11}$  и выше). Высокая точность молекулярных и атомных стандартов частоты создала совершенно новые условия в метрологии времени. Астрономы получили в свое распоряжение часы, ход которых не связан с движением небесных тел.

Квантовые усилители СВЧ диапазона имеют исключительно низкий уровень собственных шумов. Поэтому использование квантовых (парамагнитных) усилителей позволяет получить в сантиметровом и дециметровом диапазонах длин волн выигрыш в чувствительности по сравнению с обычными приемными устройствами более чем в 100 раз. Такое значительное повышение чувствительности открыло большие возможности в радиоастрономии, космической связи, радиолокации, радионавигации и т.д.

Оптические квантовые генераторы в отличие от естественных источников света имеют излучение с высокой временной и пространственной когерентностью.

Высокая степень когерентности излучения лазера позволяет использовать хорошо изученные радиотехнические методы для решения ряда задач управления, связи, навигации и локации в оптическом диапазоне волн.

Чрезвычайно большая величина несущей частоты лазера ( $10^{14} - 10^{15}$  Гц) определяет высокую чувствительность доплеровских измерителей скоростей движения объектов, разрешающую способность по дальности локационных приборов, большую скорость передачи информации. Такая величина несущей частоты обеспечивает определение чрезвычайно малых перемещений, обнаружение дефектов обработки поверхностей и т. д.

Высокая направленность излучения, которая определяется пространственной когерентностью, обеспечивает огромную дальность действия систем оптического диапазона волн, возможность концентрации энергии в малом объеме пространства, разрешающую способность локационных систем по углу, помехозащищенность различных систем передачи информации.

Оптические квантовые генераторы являются непревзойденными источниками квазимонохроматического излучения. В оптическом диапазоне имеется реальная возможность создать эталон частоты с относительной нестабильностью  $10^{-13}$ .

Огромная интенсивность излучения лазера позволила впервые наблюдать генерацию гармоник, вынужденное комбинационное рассеяние и другие нелинейные эффекты.

Отмеченные особенности и ряд других ценных свойств излучения квантовых генераторов, особенно оптического диапазона волн, определяют тот интерес, который проявляется к ним в настоящее время. К разработке квантовых устройств и их использованию привлекается все более широкий круг работников самых различных специальностей.

Для четкого понимания принципов работы квантовых устройств и грамотной их эксплуатации требуется специальная подготовка. В настоящее время квантовая электроника становится обязательной дисциплиной во многих высших учебных заведениях.

Поэтому в данном учебном пособии излагаются основы теории, принципы построения и применения квантовых усилителей и генераторов, и которое, базируется на знаниях по математике, физике и химии в объеме соответствующих курсов технических высших учебных заведений.

Из раздела 1 следует, что в случае обычных электронных ламп мы имеем дело со свободными электронами вещества, эмиттированными в вакуумное пространство, где они получают энергию от специального источника питания (возбуждения). Затем под влиянием внешнего, обычно слабого, поля сигнала возбужденные электроны отдают свою энергию нагрузке-полю колебательной системы.

В квантовых приборах свободных электронов нет. В них участвуют связанные электроны, входящие в состав атомов вещества.

Описание процессов в таких системах существенно отличается от тех, которые были применены к свободным электронам. Механизм взаимодействия связанных зарядов с внешними полями также иной. Первичной колебательной системой, в которой реализуются усиленные колебания, в данном случае является само вещество-источник активных частиц. Однако самое общее ознакомление с принципами работы квантовых усилителей дает возможность сделать вывод о том, что между электронными усилителями обычного типа и квантовыми усилителями, несмотря на существенное различие, имеется и нечто общее. Общим для них является сам принцип электронного усиления.

Усиление осуществляется, как известно, за счет энергии местного источника (источника возбуждения или иначе питания). А это означает, что сигнал может лишь воздействовать на электронную схему, причем так, чтобы она вызвала "излучение" электронов, возбужденных местными источниками на той же частоте и в той же фазе, что и действующий на входе усилителя сигнал. Действительно, генератор или усилитель представляет собой устройство, преобразующее энергию источника питания (или иначе возбуждения) в энергию электромагнитного поля, обладающего тем или иным спектром. Это преобразование осуществляется с помощью носителей зарядов, находящихся либо в свободном, либо в связанном состояниях, что для конечного эффекта усиления совершенно безразлично.

## Глава 5

# ВЗАИМОДЕЙСТВИЕ ЭЛЕКТРОМАГНИТНОГО ИЗЛУЧЕНИЯ С ВЕЩЕСТВОМ

### 5.1. Принцип квантового усиления

В основе работы устройств квантовой электроники лежит принцип квантового усиления электромагнитных волн. Созданные на основе этого принципа квантовые приборы позволили сделать скачек по частоте на 4 порядка от  $10^{11}$  до  $10^{15}$  Гц (в длинах волн от 3 мм до 0,3 мкм).

Рассмотрим суть принципа квантового усиления электромагнитных волн [4]. Прежде всего, отметим, что квантовыми системами называют системы, состоящие из микрочастиц (электронов, ядер, атомов, молекул и т.д.) и подчиняющиеся квантовым законам, характерным для микромира.

В квантовой теории отдельный электрон, атом или молекулу (и более сложные системы) можно рассматривать только как объект, существующий в некотором разрешенном квантовом состоянии. Важнейшим свойством квантовой системы является то, что ее внутренняя энергия (энергия, не определяемая движением системы как целого) при определенных условиях может принимать лишь разрешенные дискретные значения  $E$ . Каждому из этих разрешенных значений энергии соответствует одно или несколько устойчивых состояний движения частиц в системе: энергия квантована. Квантование энергии является следствием волновых свойств микрочастиц.

Энергетические состояния, которыми может обладать квантовая система, для наглядности изображают графически в виде уровней энергии (рис.5.0) по аналогии с потенциальной энергией тела, поднятого на различные высоты (уровни). Самый нижний уровень, соответствующий возможной наименьшей энергии системы, называют основным и обычно обозначают  $E_0$ . Все остальные уровни  $E_1, E_2 \dots E_n$  называют возбужденными, так как для перехода на них с основного уровня надо возбудить систему, т.е. сообщить ей дополнительную энергию.

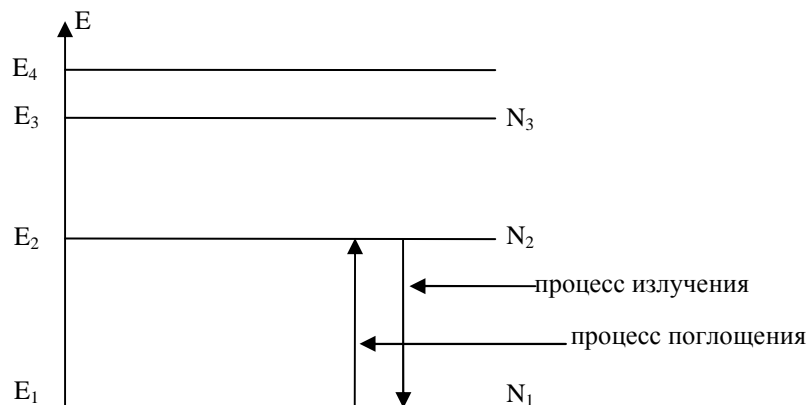


Рис. 5.0. Графическое представление квантовой системы

Число частиц на каждом уровне различно. Так как любая система стремится к устойчивому состоянию, то частицы по уровням распределены так, что  $E_1$  энергией обладает  $N_1$  число частиц, энергией  $E_2$  -  $N_2$  число частиц и так далее. Среднее число частиц, обладающее одинаковой энергией, приходящееся, как правило, на  $\text{см}^3$  называют населенностью данного уровня.

В условиях термодинамического равновесия распределение числа частиц по энергетическим уровням подчинено закону Больцмана

$$\frac{N_m}{N_n} = \frac{g_m}{g_n} e^{-\frac{E_m - E_n}{kT}}, \quad (5.1)$$

где  $g_m, g_n$  – кратность вырождения уровней (или число возможных состояний квантовой системы с определенным значением энергии),  $T$  – абсолютная температура,  $k$  – постоянная Больцмана ( $1,38 \cdot 10^{-23}$  Дж/град).

В определенных условиях взаимодействия, электрон может совершать скачкообразный переход с одного уровня на другой. При подобных переходах сверху вниз (т.е. с уровня с большей энергией на уровень с меньшей энергией) атом излучает, а при переходе снизу вверх поглощает порцию электромагнитной энергии, называемую квантом энергии или фотоном. Энергия испущенного (поглощенного) фотона определяется известным правилом частот Н. Бора:

$$\Delta E = E_m - E_n = \hbar \omega_{mn}, \quad (5.2)$$

где  $\hbar$  – постоянная Планка ( $1,05 \cdot 10^{-34}$  Дж/с),  $\omega_{mn}$  – частота кванта или частота перехода. Угловая частота связана с временной частотой  $\omega_{mn} = 2\pi \nu_{mn}$

$$\Delta E = \frac{h}{2\pi} 2\pi \nu_{mn} = h \nu_{mn}.$$

Как следует из выражения (5.2), частотный спектр (набор частот электромагнитных волн) тесно связан с энергетическим спектром, поэтому дискретностью энергетического спектра объясняется линейчатый характер спектров поглощения или излучения электромагнитных волн атомами.

Преобразуем распределение Больцмана, учитывая постулат Бора. Будем считать, что уровни не вырождены т.е.  $g_m = g_n = 1$

$$\frac{N_m}{N_n} = e^{-\frac{\hbar \omega_{mn}}{kT}}. \quad (5.3)$$

Из (5.3) следует, что если частота кванта  $\nu_{mn} = 10^{15}$  Гц. (оптический диапазон), то

$$\frac{N_m}{N_n} = e^{-\frac{h}{2\pi} \frac{2\pi \nu_{mn}}{kT}} = e^{-\frac{6,57 \cdot 10^{-34} \cdot 10^{15}}{1,38 \cdot 10^{-23} \cdot 300}} \approx e^{-2 \cdot 10^2} = \frac{1}{2,73^{200}}.$$

В СВЧ диапазоне  $\nu = 10^{11}$  Гц и следовательно

$$\frac{N_m}{N_n} = \frac{1}{2,73^{0,02}}.$$



Даже из таких простых оценок можно сделать следующий вывод: в оптическом диапазоне число частиц на уровнях, с ростом энергии уровня, очень резко падает, поэтому почти все частицы находятся на основном энергетическом уровне. В СВЧ диапазоне частицы по уровням распределены более равномерно.

Представьте себе, что квантовая система атомов находится в электромагнитном поле с частотой  $\omega$  и что  $\omega$  соответствует частоте перехода между первым и вторым энергетическими уровнями  $\omega = \omega_{12}$  (уровни не вырождены и  $E_2 - E_1 = \hbar\omega$ ). Тогда, как известно, между полем и системой атомов будет происходить эффективное взаимодействие. Поле будет индуцировать переход атомов, находящихся в состоянии с энергией  $E_1$  в состояние с энергией  $E_2$ . При этом переходе атом отбирает у поля квант энергии  $\hbar\omega$ , в результате происходит поглощение электромагнитной энергии веществом. Мощность, теряемую полем, можно подсчитать, зная вероятность индуцированного перехода в единицу времени. Допустим, что вероятность индуцированного перехода атома из состояния  $E_1$  в состояние  $E_2$  в единицу времени равно  $W_{12}$ . Число переходов совершенных в единицу времени будет  $N_1 W_{12}$ . Общая мощность поглощения поля веществом составит

$$P_{\text{погл}} = N_1 W_{12} \hbar \omega .$$

Одновременно с этим, электромагнитное поле будет взаимодействовать с атомами, находящимися в состоянии  $E_2$  и индуцировать их в состояние  $E_1$  (рис.5.0), при каждом переходе  $2 \rightarrow 1$  атом излучает квант энергии  $\hbar\omega$ . Главная особенность этого процесса в том, что излученное атомом поле является точной копией поля, индуцировавшего переход. Оно тождественно с ним по частоте, фазе, поляризации и направлению распространения. Благодаря этому исходное поле усиливается. Рассмотренный процесс излучения электромагнитных волн носит название индуцированного или вынужденного излучения и является основой работы квантовых устройств. Излучение может быть и самопроизвольным или спонтанным, при котором атом сам по себе совершает переход  $2 \rightarrow 1$ . При этом излучаемое поле не согласованно с исходным, ни по фазе, ни по направлению, ни по поляризации и представляет лишь некоторый шум. Усиления поля при спонтанном переходе не происходит.

Мощность, излучаемая атомами при индуцированных переходах сверху вниз ( $2 \rightarrow 1$ ) равна

$$P_{\text{изл}} = \hbar \omega W_{21} N_2 .$$

Суммарная мощность, поглощенная атомами очевидно равна

$$P = P_{\text{погл}} - P_{\text{изл}} = W \hbar \omega (N_1 - N_2) . \quad (5.4)$$

В условиях термодинамического равновесия  $N_1 > N_2$  и, следовательно,  $P > 0$ , что обозначает поглощение электромагнитных волн. Если же создать такое состояние вещества, при котором

$$N_2 > N_1, \quad (5.5),$$

то в этом случае  $P < 0$ . Это означало бы усиление электромагнитных волн. В создании ситуации, соответствующей неравенству (5.5) и состоит основная идея получения квантового усиления. Оказывается, что существует достаточно много

способов достижения такого состояния вещества, при котором энергетически более высоколежащие уровни имеют большую населенность, чем низележащие. Если это случается, то говорят, что в среде имеет место инверсия населенностей энергетических уровней. Среду, в которой осуществлена инверсия населенностей будем называть активной средой. Ниже мы рассмотрим некоторые способы создания инверсии населенности, реализуемые в различных приборах.

## 5.2. Типы квантовых переходов. Коэффициенты Эйнштейна

Квантовым переходом называют переход системы из одного квантового состояния в другое. В квантовых системах, обладающих дискретными уровнями энергии, существуют три типа переходов между энергетическими состояниями:

- 1) спонтанные переходы;
- 2) переходы, индуцированные электромагнитным полем;
- 3) безызлучательные, релаксационные переходы.

### 5.2.1. Спонтанные переходы и их вероятность

Спонтанные переходы – это самопроизвольные квантовые переходы частицы из верхнего энергетического состояния в нижнее, с меньшим значением энергии.

При спонтанных переходах излучается энергия в виде квантов. Квант энергии равен:

$$\hbar \cdot \omega_{i,j} = E_i - E_j,$$

где  $i, j$  – номера энергетических уровней.

Основной задачей теории квантовых переходов является вычисление вероятности перехода под воздействием внешних полей или в силу внутренних причин. Обозначим вероятность спонтанного перехода в единицу времени -  $A_{i,j}$ . Спонтанный переход возможен только сверху вниз, то есть:

$$A_{i,j} \neq 0, \text{ если } i > j$$

$$A_{i,j} = 0, \text{ если } i < j$$

Спонтанное излучение носит случайный характер. Различные частицы излучают несинхронно и независимо. Поэтому спонтанное излучение ненаправлено, некогерентно, неполяризовано и немонахроматично. Излучение обычных источников света есть результат спонтанных переходов. Поскольку спонтанное излучение некогерентно по отношению к внешнему полю оно играет роль собственных шумов в системе и служит затравочным толчком в процессе усиления и возбуждения колебаний.

Определим вероятность спонтанного перехода  $A_{i,j}$ . Рассмотрим классическую модель атома водорода (квантовомеханическое описание электромагнитного излучения является более полным, но во многих практически важных случаях проще и удобнее пользоваться классическим описанием, которое является частным случаем квантовомеханического). Положительно заряженное ядро и отрицательно заряженный электрон образуют электрический диполь. Момент электри-

ческого диполя численно равен:  $\bar{p} = e \cdot \bar{r}$  и направлен от отрицательного полюса к положительному (рис.5.1).

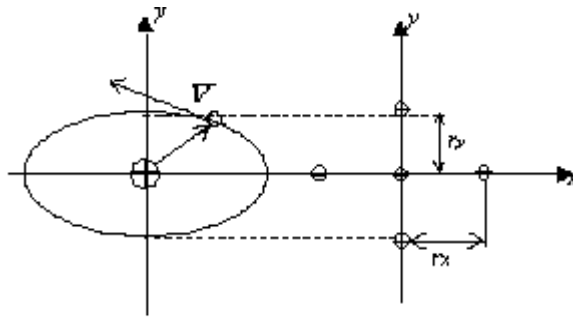


Рис. 5.1. Представление атома в виде двух осцилляторов

где  $e$  – заряд электрона,  $\bar{r}$  – радиус-вектор, соединяющий ядро с электроном.

Если дипольный момент  $\mathbf{p}$  гармонически изменяется с частотой  $\omega$ , то такой диполь называют осциллятором. Как видно (рис.5.1) движение электрона по эллиптической орбите можно заменить гармоническими колебаниями двух диполей. Дипольные моменты таких излучателей равны:

$$p_x(t) = -er_x \cos \omega t ; p_y(t) = -er_y \sin \omega t .$$

Естественно, можно ограничиться рассмотрением свойств одного из осцилляторов.

Из курса электродинамики известно, что любой диполь излучает, то есть теряет свою энергию. Потеря энергии связана с переходом частицы с одного уровня на другой.

Мощность, излучаемую таким диполем, можно рассчитать, используя следующую формулу:

$$\bar{P}_{изл} = \frac{2 \cdot \pi}{3} \cdot W_0 \cdot I_m^2 \cdot \left( \frac{l}{\lambda} \right)^2 , \quad (5.6)$$

где  $W_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$  – волновое сопротивление среды;  $I = j \cdot \omega \cdot q$  – действительное

значение тока;  $|I| = \omega \cdot q$  – модуль тока;  $l$  – длина диполя;  $\lambda = \frac{2 \cdot \pi \cdot c}{\omega}$  – длина

волны, излучаемая диполем;  $c = \frac{1}{\sqrt{\epsilon_0 \cdot \mu_0}}$  – скорость света за период;  $\mu_0$ ,  $\epsilon_0$  – маг-

нитная и электрическая постоянные вакуума.

Выразим квадрат тока через дипольный момент. Считая, что диполь эквивалентен проводнику с током длиной  $l$ , проекция дипольного момента на ось  $x$  равна

$$p_x = q \cdot l .$$

С учетом этого, выражение для мощности, излучаемой диполем в направлении оси  $x$  принимает вид:

$$P_{x_{изл}} = \frac{2}{3} \cdot \frac{\omega^4}{c^3} \cdot p_x^2.$$

Полагая, что направления  $x$  и  $y$  равновероятны, выражение для мощности излучаемой диполем в направлении оси  $y$ :

$$P_{y_{изл}} = \frac{2}{3} \cdot \frac{\omega^4}{c^3} \cdot p_y^2.$$

Суммарная мощность, излучаемая диполем

$$\sum P_{изл} = \frac{4}{3} \cdot \frac{\omega^4}{c^3} \cdot p^2, \quad (5.7),$$

где  $p = p_x = p_y$ .

Выражение (5.7) получено с точки зрения классической электродинамики. С точки зрения квантовой электродинамики (используя уравнение Шредингера), можно просто прийти к тому же выражению для мощности, но необходимо учесть, что с точки зрения квантовой электродинамики все квантуется. Тогда

$$\sum P_{ij_{изл}} = \frac{4}{3} \cdot \frac{\omega_{ij}^4}{c^3} \cdot p_{ij}^2. \quad (5.8)$$

Найдем связь между  $\sum P_{ij_{изл}}$  и вероятностью спонтанных переходов  $A_{i,j}$ .

Из предыдущего параграфа имеем:

$$P_{ij_{изл}} = \hbar \cdot \omega_{ij} \cdot W_{ij} \cdot (N_i - N_j). \quad (5.9)$$

Заменив в выражении (5.9)  $W_{ij}$  на  $A_{i,j}$ , получим выражение для мощности, излучаемой одним атомом:

$$P_{ij_{изл}} = \hbar \cdot \omega_{ij} \cdot A_{ij}.$$

Отсюда

$$A_{i,j} = \frac{P_{ij_{изл}}}{\hbar \cdot \omega_{ij}} = \frac{4}{3} \cdot \frac{\omega_{ij}^3}{\hbar \cdot c^3} \cdot p_{ij}^2. \quad (5.10)$$

Анализ выражения (5.10) показывает, что:

1) Вероятность спонтанных переходов  $A_{21}$  пропорциональна  $\omega_{ij}^3$ , это означает, что в области низких частот спонтанные переходы слабо выражены, в области высоких частот - велики и ими пренебрегать нельзя. Для оптического диапазона

$$A_{ij} = (10^3 \div 10^9) \cdot \frac{1}{c}.$$

2) Вероятность спонтанных переходов  $A_{ij} \sim p_{ij}^2$ . Дипольный момент  $p_{ij}$  сильно зависит от свойств уровней. Если  $p_{ij} = 0$  это значит, что такой переход между

уровнями  $i$  и  $j$  не может происходить в дипольном приближении, поэтому его называют запрещенным, если  $p_{ij} \neq 0$ , то переход между уровнями является разрешенным.

Величина, обратная вероятности спонтанного перехода  $A_{i,j}$  определяет среднее время жизни частиц на данном уровне  $\tau_{ij} = \frac{1}{A_{ij}}$ . Если ниже уровня  $i$  лежит несколько уровней  $j_1, j_2, j_3$  (рис. 5.2), то полное время жизни частиц определяется следующим образом:  $\tau_{ij} = 1 / \sum_j A_{ij}$

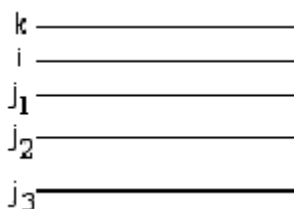


Рис. 5.2. Многоуровневая квантовая система

Если переход с  $i$ -го уровня на все нижележащие запрещен (см. рис. 5.2), то время жизни частиц на уровне  $i$  оказывается большим. Такой уровень называется метастабильным и играет существенную роль при построении квантовых устройств.

Для метастабильных уровней время жизни частиц составляет  $\tau = 10^{-3} \div 10^{-6} \text{ с}$ . Уровни, переход с которых разрешен, имеют время жизни  $\tau = 10^{-9} \text{ с}$  [4,5].

### 5.2.2. Индуцированные переходы и их вероятность

Как было определено ранее, индуцированные переходы между уровнями возникают при воздействии на систему высокочастотных внешних полей. Воздействующее электромагнитное поле может быть как монохроматическим, так и немонахроматическим. Рассмотрим кратко эти случаи:

**а) Случай немонахроматических полей.** Пусть дан атом с двумя уровнями энергии так, что  $E_j > E_i$  (рис. 5.3).

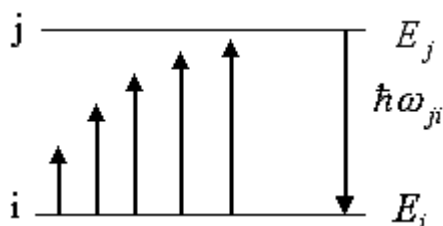


Рис. 5.3. Возбуждение двухуровневой квантовой системы немонахроматическим полем

Поместим двухуровневую квантовую систему во внешнее поле. Будем считать поле немонахроматическим — то есть имеющим набор частот. Возможны переходы как с  $j$  на  $i$  уровень, так и с  $i$  на  $j$  уровень. Вероятность перехода может зависеть от величины поля на частоте перехода  $\omega = \omega_{ji}$ . Величину поля на некоторой частоте характеризует спектральная плотность энергии поля  $\rho(\omega)$ .

По определению, вероятность индуцированных переходов должна быть пропорциональна  $\rho(\omega)$ , вычисленной на частоте перехода:

$$\omega_{ji} = \frac{E_j - E_i}{\hbar}.$$

Обозначим коэффициент пропорциональности между вероятностью индуцированного перехода  $W_{ij}$  и спектральной плотностью энергии поля  $\rho(\omega)$  через  $B_{ij}$ , тогда:

$$W_{ij} = B_{ij} \cdot \rho(\omega). \quad (5.11)$$

Коэффициент  $B_{ij}$  называется коэффициентом Эйнштейна. Он зависит от свойств уровней  $i$  и  $j$ .

Величина  $W_{ij}$  характеризует, (согласно рис. 5.3), процесс индуцированного поглощения, тогда как  $W_{ji}$  — процесс индуцированного излучения. Для определения  $W_{ij}$  в явном виде необходимо найти  $\rho(\omega)$  и  $B_{ij}$ .

Подсчитаем  $\rho(\omega)$  на частоте  $\omega = \omega_{ji}$ .

Предположим, что наше поле заключено в некотором объеме  $V$ . В качестве объема примем куб размером  $a \cdot a \cdot a$  (рис. 5.4).

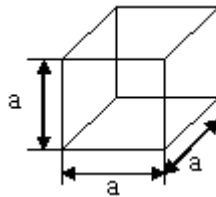


Рис. 5.4. Кубический резонатор

Такой куб представляет из себя резонатор с бесчисленным множеством собственных колебаний. Каждое из собственных колебаний характеризуется резонансной частотой  $\omega_{mnp}$ . Где числа  $m, n, p$  определяют число полуволн, укладываемых на ребрах резонатора в направлении  $x, y, z$  соответственно. Из всего множества колебаний выберем один  $S$ -ый тип колебаний ( $m, n, p = s$ ). Тогда:

$$\omega_s^2 = \frac{c^2 \cdot \pi^2}{a^2} \cdot (m^2 + n^2 + p^2). \quad (5.12)$$

Однако на частоте  $\omega_S$  может быть не один фотон, а множество. Обозначим через  $m_S$  число фотонов в S-ом типе колебаний на частоте  $\omega_S$ , а энергию кванта (фотона)  $\hbar \cdot \omega_S$ , тогда энергия поля S -го типа будет очевидно равна:

$$\hbar \cdot \omega_S \cdot m_S,$$

а плотность энергии S -го типа:

$$\rho_S = \frac{\hbar \cdot \omega_S}{V} \cdot m_S.$$

Чтобы получить вероятность  $W_{ij}$ , нужно очевидно просуммировать все  $\rho_S$ , находящиеся в единице спектрального интервала, вблизи частоты  $\omega = \omega_{ji}$ , то есть:  $\rho(\omega) = \sum_S \rho_S$ . Тогда выражение (5.11) примет вид:

$$W_{ij} = B_{ij} \cdot \sum_S \frac{\hbar \cdot \omega_S}{V} \cdot m_S.$$

До сих пор мы рассматривали вероятность индуцированного поглощения. Рассмотрим теперь обратные процессы, то есть переход с уровня j на i. Строгий расчет, основанный на законах квантовой электродинамики показывает, что:

$$W'_{ji} = B_{ij} \cdot \sum_S \frac{\hbar \cdot \omega_S}{V} \cdot (m_S + 1). \quad (5.14)$$

Причем если уровни не вырождены (то есть каждому уровню энергии отвечает одно состояние атома), то:

$$B_{ji} = B_{ij}.$$

Формула (5.14) принципиально отличается от формулы (5.13). Она показывает, что даже в отсутствии поля, ( $m_S = 0$ ), вероятность  $W'_{ji}$  не равна нулю. Иначе, переходы сверху вниз происходят и в отсутствие поля. Ясно, что такие переходы являются спонтанными. Вероятность их обозначена буквой  $A_{ji}$ . Следовательно, из (5.14) следует, что:

$$A_{ji} = B_{ji} \cdot \sum_S \frac{\hbar \cdot \omega_S}{V}. \quad (5.15)$$

Учитывая (5.15), имеем:

$$W'_{ji} = W_{ij} + A_{ji}. \quad (5.16)$$

Исключая из (5.16) вероятность спонтанных переходов можно полагать, что вероятность индуцированных переходов сверху вниз и снизу вверх одинакова, то есть:

$$W_{ji} = W_{ij}.$$

Выражение (5.15) позволяет вычислить коэффициент Эйнштейна  $B_{ji}$ . Для этого необходимо найти  $\sum_S \hbar \cdot \omega_S$ . Рассмотрим  $\sum_S \hbar \cdot \omega_S$  в единичной полосе равной 1 Гц, тогда  $\omega_S$  можно считать практически постоянной и равной  $\omega$  и под

знаком суммы останется 1. Обозначим  $\sum_s 1 = n(\omega)$ , где  $n(\omega)$  число типов колебаний, приходящихся на единичный частотный интервал (1 Гц). Тогда

$$A_{ji} = B_{ji} \cdot \frac{\hbar \cdot \omega_s}{V} n(\omega). \quad (5.17)$$

Для определения  $n(\omega)$  воспользуемся равенством (5.12), рассмотрим зависимость  $\omega_s$  от чисел  $m, n, p$ . Для этого выберем систему координат  $m, n, p$ .

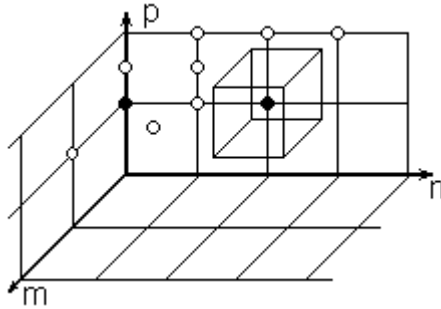


Рис. 5.5. К иллюстрации плотности генерируемых мод в резонаторе

В этой системе координат каждому типу колебаний, то есть тройке чисел  $m, n, p$  отвечает точка. Далее, все пространство  $m, n, p$  можно разбить на кубики (один из них показан на рис. 5.5), в каждом из которых содержится одна точка  $m, n, p$ . Длина ребра кубика равна 1 и, следовательно, объем кубика  $\Delta V$  равен 1. Иначе, в пространстве  $m, n, p$  объем, приходящийся на один тип колебаний равен 1. Тогда, в любой фигуре объема  $V$ , число типов колебаний  $N(\omega)$  будет численно равно  $V$ .

Из выражения (5.12) видно, что когда  $\omega_s$  изменяется от нуля до некоторого значения  $\omega$ , числа  $m, n, p$  изменяются от нуля до  $\left(\frac{a \cdot \omega}{c \cdot \pi}\right)^2$ , то есть:

$$m^2 + n^2 + p^2 = \left(\frac{a \cdot \omega}{c \cdot \pi}\right)^2.$$

Это выражение представляет собой уравнение сферы с радиусом, равным:

$$R = \frac{a \cdot \omega}{c \cdot \pi}.$$

Таким образом, полное число собственных частот полости, заключенных в интервале частот от 0 до некоторого значения  $\omega$ , численно равно  $\frac{1}{8}$  объема октанта сферы (так как  $m, n, p > 0$ ), то есть:

$$N(\omega) = V = \frac{1}{8} \cdot \frac{4}{3} \cdot \pi \cdot \left(\frac{a \cdot \omega}{c \cdot \pi}\right)^3 = \frac{1}{6} \cdot \frac{a^3 \cdot \omega^3}{c^3 \cdot \pi^2}.$$

Учитывая, что каждой определенной частоте отвечают два вида колебаний, отличающихся лишь поляризацией, полученное число нужно удвоить, получим:

$$N(\omega) = \frac{1}{3} \cdot \frac{a^3 \cdot \omega^3}{c^3 \cdot \pi^2}. \quad (5.18)$$



Для определения числа собственных колебаний, лежащих между  $\omega$  и  $\omega + d\omega$ , необходимо продифференцировать это выражение по частоте.

Дифференцируя (5.18) получим:

$$dN(\omega) = \frac{a^3 \cdot \omega^2}{c^3 \cdot \pi^2} \cdot d\omega.$$

Число типов колебаний в единичном интервале частот будет:

$$n(\omega) = \frac{dN(\omega)}{d\omega} = \frac{V \cdot \omega^2}{c^3 \cdot \pi^2},$$

где  $V = a^3$  – объем куба. Тогда выражение (5.17) примет вид:

$$A_{ji} = B_{ji} \cdot \frac{\hbar \cdot \omega_{ji}^3}{c^3 \cdot \pi^2}.$$

Это уравнение носит название соотношения Эйнштейна между коэффициентами  $A_{ji}$  и  $B_{ji}$ . Подставляя в место  $A_{ji}$  его выражение, полученное выше, найдем:

$$B_{ji} = \frac{4}{3} \cdot \frac{\pi^2}{\hbar^2} \cdot p_{ji}^2.$$

Таким образом, коэффициент Эйнштейна  $B_{ji}$  определяется только квадратом дипольного момента. Когда спонтанный переход запрещен, то запрещен и индуцированный, так как определяется одной и той же величиной.

Итак, используя полученные выражения для  $B_{ji}$ ,  $\rho(\omega)$ , можно из формулы (5.11) найти вероятность индуцированных переходов в явном виде.

В заключении следует отметить, что при определении вероятности индуцированных переходов в случае действия немонахроматических полей необходимо учитывать вероятность попадания энергии воздействующего поля  $F(\omega)$  в единичный частотный интервал (вблизи частоты перехода  $\omega_{ji}$ ). Тогда выражение (5.11) с учётом рассмотренного выше примет вид:

$$W_{ij} = \rho B_{ji} F(\omega), \quad (5.19)$$

где  $\rho = \frac{\hbar \omega_{ij}}{V} m$  – плотность энергии полного поля, воздействующего на квантовую систему;  $F(\omega) = \frac{\rho(\omega)}{\rho} du$  – вероятность того, что энергия поля оказывается в спектральном интервале  $d\omega$  вблизи частоты  $\omega_{ji}$ ;  $m$  – полное число фотонов в резонаторе.

При рассмотрении энергетических состояний квантовых систем в случае а) нас не интересовал вопрос о ширине дискретных энергетических уровней атомов или молекул, и эти уровни принимались бесконечно узкими. Излучательным переходам между такими изолированными уровнями энергии соответствовала бы бесконечно узкая спектральная линия излучения (или поглощения) на строго фиксированной частоте, т.е. идеальная монохроматическая волна. В действительности подобная ситуация не возможна хотя бы по той причине, что время жизни  $\tau$  в возбужденном состоянии конечно, а в силу соотношения неопределенностей  $\delta E \tau \approx \hbar$  это приводит к неопределенности в энергии состояния, т.е. к «размытию» энергетического уровня на величину порядка  $\delta E$ . Для изолированных атомов и

молекул это уширение очень мало ( $\delta E \ll E$ ), но принципиально энергетические уровни имеют конечную ширину. Более подробно на вопросах уширения спектральных линий мы остановимся ниже.

б) Случай монохроматических полей. Рассмотрим случай, когда внешнее поле имеет строго определенную частоту  $\omega = \omega_{ij}$ , а атом или любая другая частица обладает энергетическим спектром определенной ширины (рис.5.6).

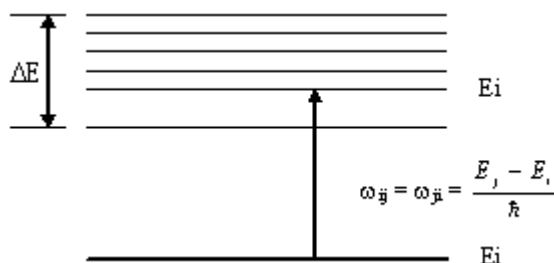


Рис. 5.6. Энергетический спектр квантовой системы при воздействии монохроматическим полем

При этом необходимо учитывать вероятность попадания атома в единичный частотный интервал (1Гц), тогда как попадание поля в этот интервал событие достоверное (вероятность его равна единице) (рис.5.7).

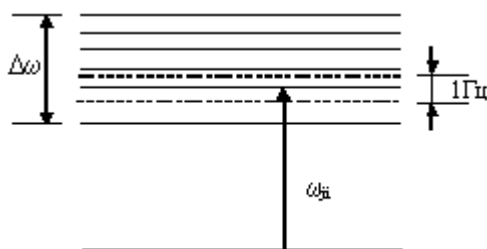


Рис. 5.7. К иллюстрации попадания атома в единичный спектральный интервал

С учетом сказанного, в выражении (5.19) необходимо заменить  $F(\omega)$  на вероятность того, что атом окажется в единичном спектральном интервале. Эту вероятность обозначим  $g(\omega)$ . Тогда:

$$W_{ji} = B_{ji} \cdot \rho \cdot g(\omega), \quad (5.20)$$

где  $\rho = \frac{\epsilon E^2}{2}$  – объемная плотность энергии воздействующего поля, приходящаяся на единицу объема ( в системе Си).

$g(\omega)$  - имеет различную форму в зависимости от ситуации.

Наибольшее количество частиц будет в середине участка, где  $\omega = \omega_{ij}$ . По мере того, как удаляемся от частоты  $\omega_{ij}$ , вероятность  $g(\omega)$  уменьшается, то есть график должен иметь вид (рис.5.8):

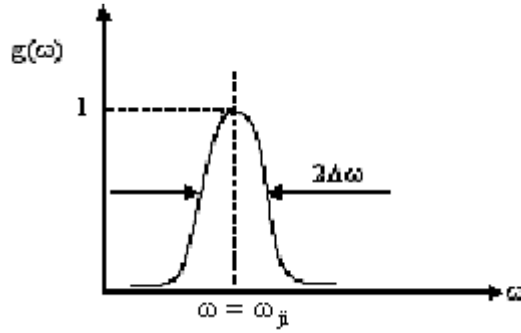


Рис. 5.8. Лоренцева линия

Примем, что эта зависимость такая же как у колебательного контура. Тогда (по аналогии)  $g(\omega)$  описывается следующим выражением:

$$g(\omega) = \frac{C}{1 + (\omega - \omega_{ji})^2 \cdot \tau^2}, \quad (5.21)$$

где  $\tau$  – время жизни частицы и носит название времени поперечной релаксации (определяется из эксперимента);  $C$  – константа.

Константу найдём из условия нормировки:

$$\int_{-\infty}^{\infty} g(\omega) \cdot d\omega = 1, \quad (5.22)$$

то есть вероятность того, что атом находится на какой либо частоте событие достоверное.

Подставив (5.21) в (5.22), заменяя  $(\omega - \omega_{ji}) \cdot \tau = x$  получим

$$\int_{-\infty}^{\infty} \frac{C \cdot d\omega}{1 + (\omega - \omega_{ji})^2 \cdot \tau^2} = \frac{C}{\tau} \cdot \int_{-\infty}^{\infty} \frac{dx}{1 + x^2} = \frac{C}{\tau} \cdot \arctg(x) \Big|_{-\infty}^{\infty} = \frac{C}{\tau} \cdot \left( \frac{\pi}{2} + \frac{\pi}{2} \right) = \frac{\pi \cdot C}{\tau} = 1;$$

отсюда  $C = \frac{\tau}{\pi}$ . С учетом этого

$$g(\omega) = \frac{1}{\pi \cdot \tau} \cdot \frac{1}{\frac{1}{\tau^2} + (\omega - \omega_{ji})^2}, \quad (5.23)$$

если  $\omega - \omega_{ji} \gg \frac{1}{\tau}$ , то есть выбранный интервал много больше  $\Delta\omega$ ,  $W_{ji} \rightarrow 0$ .

Выражение (5.20) можно представить в другом виде, выразив объемную плотность  $\rho$  энергии воздействующего поля через число фотонов в резонаторе:

$\rho = \frac{\hbar\omega}{V} m$  ( $m$ - число фотонов резонатора). Окончательно получим:

$$W_{ji} = B_{ji} \cdot \frac{\hbar \cdot \omega \cdot g(\omega)}{V} \cdot m, \quad (5.24)$$

где  $b_{ij} = B_{ji} \cdot \frac{\hbar \cdot \omega \cdot g(\omega)}{V}$  - носит название спектрального коэффициента

Эйнштейна и зависит от частоты. Окончательно имеем:  $W_{ji} = b_{ji} \cdot m$  [4,5,6,12].

**Сечение перехода, коэффициент поглощения и усиления.** В заключении данного подраздела следует отметить, что вероятность индуцированного излучения  $W_{ji}$ , в отличие от вероятности спонтанного излучения  $A_{ji}$ , зависит не только от конкретного перехода, но и от интенсивности падающей электромагнитной волны. В [12] показано, что для плоской волны выражение (5.24) можно представить в другом виде

$$W_{ji} = \sigma_{ji} I,$$

где  $I$  – плотность потока фотонов в падающей волне (количество фотонов, проходящих через  $1 \text{ см}^2$  в единицу времени);  $\sigma_{ji}$  – величина, имеющая размерность площади (она называется сечением вынужденного излучения) и зависящая только от характеристик данного перехода;  $\sigma_{ij}$  – называется сечением вынужденного поглощения, причем  $\sigma_{ji} = \sigma_{ij} = \sigma$

Учитывая связь между вероятностями индуцированного и спонтанного излучения и полагая, что  $I = c \cdot \rho/n$  [13] ( $c$  – скорость света,  $n$  – показатель преломления среды), запишем выражение для  $\sigma$

$$\sigma = \frac{\pi\omega}{3n\epsilon_0 c \hbar^2} |p_{ji}|^2 g(\omega). \quad (5.24a)$$

Отсюда видно, что сечение  $\sigma$  зависит от свойств среды ( $|p_{ji}|^2$ ,  $g(\omega)$ ) и частоты  $\omega$  падающей волны. Поэтому сечение перехода является очень важным и широко применяемым параметром перехода.

Другим важным параметром, описывающим взаимодействие излучения с веществом, является коэффициент  $\xi$ , определяемый следующим образом;

$$\xi = \sigma(N_i - N_j). \quad (5.24б)$$

Если  $N_i > N_j$ , то величина  $\xi$  характеризует поглощение излучения в веществе (аналогично в подразделе 5.1 -  $\Sigma P_{\text{погл}} > 0$ ) и называется коэффициентом поглощения. Достоинством этого параметра является то, что во многих случаях его можно непосредственно измерить. Если  $N_j > N_i$  коэффициент поглощения становится отрицательным и, следовательно, волна в среде будет усиливаться. В этом случае удобно ввести новую величину

$$\alpha = -\xi = \sigma(N_j - N_i), \quad (5.24в)$$

которая имеет положительное значение и называется коэффициентом усиления.

### 5.2.3. Безызлучательные переходы и их вероятность

Атомы и молекулы газа, в результате неупругих соударений друг с другом или электронами, теряют или приобретают энергию. При этом не происходит ни излучения, ни поглощения энергии электромагнитного поля. Такие энергетические переходы называют безызлучательными. Если рассмотрим некоторую систему атомов или ионов, внедрённых в узлы кристаллической решётки (например, ионы Cr в  $\text{Al}_2\text{O}_3$ ), то между ионами и решеткой будет происходить взаимодействие и обмен энергией. Таким образом, безызлучательные переходы в твёрдом теле происходят за счет колебательного движения кристаллической решётки. Если температура решётки поддерживается постоянной и равной  $T$ , то такая совокуп-

ность ионов образует термостатированный ансамбль. Роль термостата выполняет кристаллическая решётка.

Допустим, что такой ансамбль каким-либо способом выведен из состояния теплового равновесия, например, за счёт воздействия на него электромагнитным полем. Тогда, в период воздействия поля, частицы будут совершать переходы между уровнями, так что их распределение по уровням окажется отличным от больцмановского.

Если  $N^e$  – число частиц при больцмановском распределении, то можно утверждать, что в период действия поля  $N \neq N^e$ , (где  $N$  – распределение частиц в период действия поля). Однако опыт показывает, что после выключения поля система возвращается к прежнему состоянию теплового равновесия. Числа  $N_m$  начинают изменяться так, что вновь приходят к  $N_m^e$ . Этот процесс носит название процесса релаксации системы к состоянию теплового равновесия.

Опишем этот процесс[4]. Примем, что действие термостата состоит в индуцировании переходов между различными состояниями системы. Обозначим через  $\Gamma_{mn}$  – вероятность того, что система в единицу времени совершит переход из состояния  $m$  в состояние  $n$  под действием термостата. Тогда число систем совершающих в единицу времени переход из состояния  $m$  в состояние  $n$ , очевидно, будет  $\Gamma_{mn}N_m$ . На все уровни  $n$  из  $m$ -состояний совершит переход число частиц, равное  $\sum_{n \neq m} \Gamma_{mn}N_m$ , аналогично  $\sum_{n \neq m} \Gamma_{nm}N_n$  – число частиц переходящих из  $n$ -состояний в  $m$ , в единицу времени. Полное изменение числа частиц в  $m$ -состоянии в единицу времени равно, очевидно, разности этих величин, с другой стороны оно должно быть равно  $\frac{dN_m}{dt}$ . В результате находим

$$\frac{dN_m}{dt} = \sum_{n \neq m} (\Gamma_{nm}N_n - \Gamma_{mn}N_m). \quad (5.25)$$

Это уравнение учитывает взаимодействие системы с термостатом, но не учитывает другие типы взаимодействий. Числа  $\Gamma_{mn}$  и  $\Gamma_{nm}$  характеризуют степень этого взаимодействия: чем они больше, тем сильнее взаимодействие.

Вероятность безызлучательных переходов сверху вниз и снизу вверх не равны, т.е.

$$\Gamma_{nm} \neq \Gamma_{mn}.$$

Чтобы найти количественную связь между ними, рассмотрим систему атомов в состоянии термодинамического равновесия, когда  $N_m = N_m^e$  и не зависит от времени. Поэтому сумма в правой части равенства (5.25) равна нулю. Обычно постулируется, что равна нулю не только сумма, но её отдельные слагаемые. Тогда получим

$$\Gamma_{nm}N_n^e = \Gamma_{mn}N_m^e. \quad (5.26)$$

Равенство (5.26) выражает принцип детального равновесия и устанавливает

связь между  $\Gamma_{mn}$  и  $\Gamma_{nm}$ , так как  $N_n^e = Ce^{-\frac{E_n}{kT}}$ ;  $N_m^e = Ce^{-\frac{E_m}{kT}}$ , то

$$\Gamma_{nm} \cdot e^{-\frac{E_n}{kT}} = \Gamma_{mn} \cdot e^{-\frac{E_m}{kT}}, \quad (5.27)$$

если  $E_n > E_m$ , то из (5.27) следует, что  $\Gamma_{nm} > \Gamma_{mn}$ , т.е. вероятность безызлучательного перехода сверху вниз больше, чем снизу вверх. Связь между  $\Gamma_{nm}$  и  $\Gamma_{mn}$  можно записать также в виде

$$\Gamma_{nm} = \Gamma_{mn} \cdot e^{\frac{\hbar\omega_{nm}}{kT}}$$

Отсюда следует, что на низких частотах, когда  $\hbar\omega_{nm} \ll kT$ , величины  $\Gamma_{nm}$  и  $\Gamma_{mn}$  почти равны. На высоких частотах, когда  $\hbar\omega_{nm} \gg kT$ , имеет место неравенство

$$\Gamma_{nm} \gg \Gamma_{mn}.$$

Абсолютные значения  $\Gamma_{nm}$ , как правило определяются из экспериментальных данных. Расчёт их затруднителен.

В заключении следует отметить, что вероятности безызлучательных переходов зависят от следующих факторов:

- 1) от расстояния между энергетическими уровнями;
- 2) от природы энергетических уровней;
- 3) от температуры.

### 5.3. Уширение спектральных линий

Ширина спектральной линии излучения лазера зависит главным образом от двух факторов: от резонатора, с одной стороны, и усиливающей среды с другой. Косвенно она определяется и другими, практическими соображениями.

#### 5.3.1. Естественная ширина спектральной линии

Усиливающая среда характеризуется шириной спектральной линии  $\Delta\nu$ , связанной с двумя энергетическими уровнями  $E_1$  и  $E_2$ . Неопределённостью  $\Delta E$  измерения разности  $E_2 - E_1$ , определяется время жизни уровня 2, в соответствии с соотношением неопределённости Гейзенберга:

$$\Delta E \tau = h, \text{ т.е. } \Delta\nu = \frac{1}{\tau}, \quad (5.28)$$

где  $\tau$  – среднее время жизни частиц на уровне по спонтанным переходам.

Выгоднее выбирать переходы с большим временем жизни. Правда, ширина  $\Delta\nu$ , называемая естественной шириной линии, обычно очень мала и ею можно пренебречь по сравнению с другими.

Пример:  $\lambda = 0.5 \text{ мкм}$ ,  $\tau = 10^{-8} \text{ с}$ ,  $\Delta\lambda = 10^{-5} \text{ нм}$ .

#### 5.3.2. Ширина полосы пропускания резонатора Фабри-Перо

Ширина полосы пропускания резонатора Фабри-Перо зависит, главным образом, от его добротности  $Q$ . Теоретически добротность определяется выражением:

$$Q = \frac{\nu_p}{\Delta\nu_p} \quad \text{или} \quad \Delta\lambda = \lambda./Q. \quad (5.29)$$

Если  $Q=10^8$ ,  $\lambda=500 \text{ нм}$ , то  $\Delta\lambda=5*10^{-5} \text{ нм}$ .

В действительности ширина полосы пропускания резонатора ограничивается допусками, с которыми изготавливается резонатор, и обычно составляет  $\sim 10^{-4} \text{ нм}$ .

### 5.3.3. Причины уширения спектральной линии

В зависимости от типа лазера имеется много различных причин уширения лазерной линии. Все они приводят к уширению естественной линии, а также к увеличению ширины полосы пропускания объёмного резонатора.

**а) Доплеровское уширение.** Существенно в газовых лазерах. Оно состоит в том, что атомная линия частоты  $\nu$  смещается под действием броуновского движения атомов со скоростью  $v$ , образующей угол  $\theta$  с направлением излучения, в результате чего частота линии становится равной (в первом приближении)

$$\nu \left( 1 - \frac{v}{c} \sin(\theta) \right). \quad (5.30)$$

Доказывается, что контур спектральной линии в этом случае выражается кривой Гаусса

$$g(\omega) = \frac{\sqrt{\ln 2}}{\sqrt{\pi \Delta \nu_d}} = e^{-\ln 2 \left( \frac{\nu - \nu_0}{\Delta \nu_d} \right)^2}, \quad (5.31)$$

и имеет полуширину на уровне, равном половине от максимального,

$$\Delta \nu_d = \nu_0 \sqrt{\frac{2kT \ln 2}{mc^2}} = 7,18 \nu_0 \sqrt{\frac{T}{M}} 10^{-7}, \quad (5.32)$$

где  $m$  – масса атома или молекулы;  $M$  – атомный или молекулярный вес.

Ширина спектральной линии при доплеровском уширении может быть очень большой. Так, например, в газовых лазерах она может достигать многих сотен мегагерц и даже быть более 1000 МГц. В таких лазерах она превосходит уширения, обусловленные другими факторами [5,6,9].

**б) Уширение давлением.** Движение атома и молекул газа часто приводит к столкновениям между ними. Эти столкновения вызывают переход возбужденных атомов с уровня  $E_2$  на уровень  $E_1$ , снижая тем самым время жизни возбуждённого уровня. В результате увеличивается ширина спектральной линии излучения (поскольку  $\Delta \nu \tau = 1$ ).

**в) Штарковское уширение.** Под действием электрического поля уровни могут расщепиться на несколько близких энергетических уровней (поле снимает вырождение энергетических уровней). Эффект Штарка наблюдается, главным образом, в твёрдотельных лазерах большой мощности. Так в случае лазера на рубине штарковское уширение линии излучения имеет порядок 0,1 нм.

**г) Уширение за счет дефектов в твердом теле, создающих неоднородности в распределении потенциала.** Все виды уширения приводят к тому что линия излучения лазера становится более широкой по сравнению с той которая определяется добротностью резонатора. Таким образом спектр излучения лазера представляет собой ряд равноотстоящих друг от друга спектральных линий, приходящихся на моды резонатора Фабри-Перо. Огибающей этих линий служит кривая, обусловленная главной причиной уширения для рассматриваемого типа лазера. Можно показать, что в действительности ширина каждой из рассматриваемых линий меньше той ширины, которая обусловлена резонатором Фабри-Перо, и эта разница тем больше, чем больше отдаваемая мощность. Такое обострение спектральных линий обусловлено индуцированным излучением (рис. 5.9).

На практике важно различать два вида паразитного уширения спектральной линии:

- 1) Однородное уширение (Штарковское или обусловленное давлением), при котором все атомы играют одинаковую роль в статическом смысле;
- 2) Неоднородное уширение (доплеровское), когда волна частоты  $\nu$  может взаимодействовать только с атомами, частота которых равна

$$\nu = \nu_0 \left(1 - \frac{v}{c} \cos(\theta)\right). \quad (5.33)$$

Эти два вида уширения совершенно по-разному влияют на работу лазера [33].

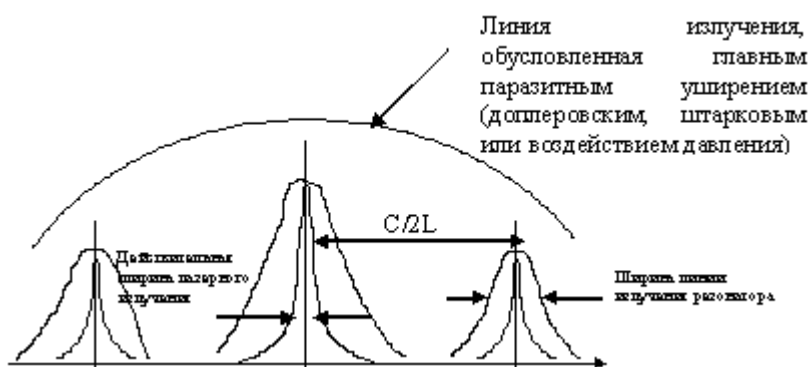


Рис. 5.9. Частотная избирательность различных явлений, используемых при лазерном излучении

#### 5.4. Общие уравнения баланса

Результатом взаимодействия электромагнитного поля с веществом является поглощение или усиление этого поля веществом.

$$P_{\text{погл}} = \hbar\omega_{12} W_{12}(N_1 - N_2) \quad (5.34)$$

Отсюда видно, что одной из главных задач при подсчёте мощности поглощения  $P_{\text{погл}}$  системой атомов, является определение числа частиц на различных энергетических уровнях. Если вещество находится в состоянии термодинамического равновесия, то распределение частиц подчиняется закону Больцмана. Если на вещество действует электромагнитное поле, то необходимо учитывать взаимодействие вещества не только с термостатом, но и с полем. Кроме того, необходимо учитывать и спонтанные переходы. Для определения этих чисел чаще всего используют уравнения баланса. Наша цель записать эти уравнения.

Выберем в системе энергетических уровней атомов некоторый уровень  $m$  (рис. 5.10)

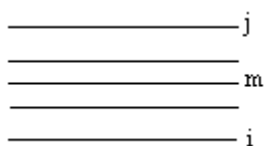


Рис. 5.10. Энергетические уровни атома



Пусть в момент времени  $t$  число частиц на этом уровне равно  $N_m$ . Подсчитаем скорость изменения  $N_m$ , равную  $\frac{d}{dt} \cdot N_m$ . В общем случае изменение числа частиц на  $m$ -уровне может происходить за счёт тепловых, индуцированных и спонтанных переходов, т.е.

$$\left(\frac{dN_m}{dt}\right) = \left(\frac{\partial N_m}{\partial t}\right)_{\text{менл}} + \left(\frac{\partial N_m}{\partial t}\right)_{\text{сп}} + \left(\frac{\partial N_m}{\partial t}\right)_{\text{инд}}, \quad (5.35)$$

где  $\left(\frac{\partial N_m}{\partial t}\right)_{\text{менл}}$  – скорость изменения числа частиц  $N_m$  за счёт взаимодействия с термостатом,

$\left(\frac{\partial N_m}{\partial t}\right)_{\text{сп}}$  – скорость изменения числа частиц  $N_m$  за счет самопроизвольных переходов,

$\left(\frac{\partial N_m}{\partial t}\right)_{\text{инд}}$  – скорость изменения числа частиц  $N_m$  под действием внешнего электромагнитного поля.

Каждое из этих слагаемых легко подсчитывается. Для тепловых переходов мы уже имели (см. уравнение 5.25)

$$\left(\frac{\partial N_m}{\partial t}\right)_{\text{менл}} = \sum_{n \neq m} (\Gamma_{nm} N_n - \Gamma_{mn} N_m). \quad (5.36)$$

Заменяя в (5.36) вероятности тепловых переходов  $\Gamma_{mn}$  на вероятности индуцированных переходов  $W_{mn}$  или на вероятности спонтанных переходов  $A_{mn}$ , получим вместо уравнения (5.35) следующее выражение:

$$\left(\frac{dN_m}{dt}\right) = \sum_{n \neq m} (\Gamma_{nm} N_n - \Gamma_{mn} N_m) + \sum_{n \neq m} W(N_n - N_m) + \sum_{n \neq m} (A_{nm} N_n - A_{mn} N_m), \quad (5.37)$$

причём, если  $E_n < E_m$ , то  $A_{nm} = 0$ . Выражение (5.37) – есть уравнение баланса частиц. Оно должно быть записано для каждого из рассмотренных уровней. Число этих уравнений бесконечно велико. Такие уравнения, называемые также кинетическими, не учитывают фазовых соотношений между полями мод и там, где фазовые соотношения становятся существенными, могут приводить к неверным результатам. Тем не менее, балансные уравнения позволяют качественно описать большинство основных явлений, имеющих место в лазерах, и в силу простоты этих уравнений в дальнейшем мы будем использовать только их. Записав балансное уравнение для каждого энергетического уровня квантовой системы и, просуммировав, все левые и правые части этих уравнений, получим следующее равенство  $\frac{d}{dt}(N_1 + N_2 + N_3 + \dots) = 0$ .

Отсюда следует, что полное число частиц есть величина постоянная, т.е.

$$N_1 + N_2 + N_3 + \dots = N. \quad (5.38)$$

Равенство (5.38) выражает закон сохранения числа частиц в системе. Таким образом, как бы частицы не были распределены между энергетическими уровнями, сумма их остаётся неизменной. Поэтому при расчётах одно из уравнений для

$N_m$  может быть заменено равенством (5.38). Кроме того, нужно учесть, что  $\Gamma_{nm} N_n^e = \Gamma_{mn} N_m^e$  накладывает ограничение на величину  $\Gamma$ .

При заданных электромагнитных полях, а, следовательно, и при заданных вероятностях  $W_{nm}$ , решение системы (5.37) позволяет определить все  $N_m$  [4].

#### 5.4.1 Взаимодействие монохроматического поля с двухуровневой квантовой системой

Применим теперь уравнение баланса к рассмотрению частного случая, когда система атомов, взаимодействующая с полем, является двухуровневой, причём будем полагать (рис. 5.11), что частота кванта близка к частоте поля. Конечной целью решения нашей задачи – расчёт мощности поглощения (5.4)

$$P_{\text{погл}} = \hbar\omega W(N_1 - N_2), \quad (5.39)$$

где  $W=W_{12}=W_{21}$ .

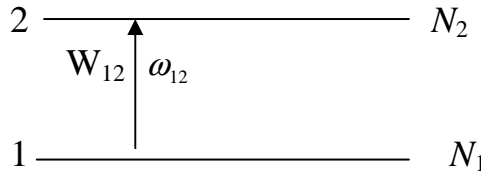


Рис. 5.11. Схема двухуровневого генератора для расчета  $P_{\text{погл}}$

Как видно (5.39), для расчёта  $P_{\text{погл}}$  нужно знать  $N_1$  и  $N_2$ . Займёмся их определением. При этом будем полагать, что уровни невырождены, то есть  $g_1=g_2=1$ , вероятности спонтанных переходов малы и ими можно пренебречь. Тогда уравнение баланса для  $N_1$  будет

$$\begin{cases} \frac{dN_1}{dt} = \Gamma_{21}N_2 - \Gamma_{12}N_1 + W(N_2 - N_1) \\ N_2 + N_1 = N \end{cases} \quad (5.40)$$

Для простоты рассмотрим случай стационарного взаимодействия, когда  $\frac{dN_1}{dt} = 0$ . При этом первое уравнение в (5.40) упрощается

$$\begin{cases} (\Gamma_{21} + W)N_2 - (\Gamma_{12} + W)N_1 = 0 \\ N_2 + N_1 = N \end{cases}.$$

Решая эту систему, получим:

$$N_1 = \frac{\Gamma_{21} + W}{\Gamma_{12} + \Gamma_{21} + 2W} N, \quad N_2 = \frac{\Gamma_{12} + W}{\Gamma_{12} + \Gamma_{21} + 2W} N. \quad (5.41)$$

Пользуясь принципом детального равновесия, исключим отсюда  $\Gamma_{ij}$

$$\Gamma_{21}N_2^e = \Gamma_{12}N_1^e, \quad (5.42)$$

поделив равенство (5.42) на  $\Gamma_{21} \Gamma_{12}$ , получим  $\frac{N_2^e}{\Gamma_{12}} = \frac{N_1^e}{\Gamma_{21}}$ . Обозначим  $\frac{N_m^e}{\Gamma_{ik}} = NT_1$ , где

$T_1$  – время жизни частицы на уровне, обусловленное взаимодействием частицы с термостатом. Время  $T_1$  – называется временем продольной релаксации. В результате имеем:

$$\Gamma_{12} = \frac{N_2^e}{NT_1}; \quad \Gamma_{21} = \frac{N_1^e}{NT_1}. \quad (5.43)$$

После подстановки (5.43) в (5.41) получим

$$N_1 = \frac{N_1^e + WT_1 N}{1 + 2WT_1} \quad (5.44)$$

$$N_2 = \frac{N_2^e + WT_1 N}{1 + 2WT_1}$$

Разность населённости

$$N_1 - N_2 = \frac{N_1^e - N_2^e}{1 + 2WT_1}. \quad (5.45)$$

Подставив (5.45) в (5.39), имеем

$$P_{\text{погл}} = \hbar\omega \frac{W}{1 + 2WT_1} (N_1^e - N_2^e). \quad (5.46)$$

Формула (5.46) даёт решение задачи о взаимодействии поля с двухуровневой квантовой системой. Анализ её проведём в следующем параграфе.

#### 5.4.2 Мощность поглощения, как функция частоты и амплитуды поля. Эффект насыщения

Проанализируем формулу (5.46) Для этого запишем её в явном виде, учтя что

$$W = B_{12}\rho g(\omega),$$

$$\text{где } B_{12} = \frac{4}{3} \cdot \frac{\pi^2}{\hbar^2} p_{12}^2, \quad \rho = \frac{\varepsilon \cdot E^2}{2}, \quad g(\omega) = \frac{\frac{1}{\tau\pi}}{\frac{1}{\tau^2} + (\omega - \omega_{12})^2}$$

Частотно-амплитудная зависимость получается сложной. Ограничимся рассмотрением двух крайних случаев:

а) случай слабых полей. Поле называется слабым, если имеет место неравенство  $WT_1 \ll 1$ . Тогда из (5.46)

$$P_{\text{погл}} = \hbar\omega B_{12} \frac{\varepsilon \cdot E^2}{2} g(\omega) (N_1^e - N_2^e). \quad (5.47)$$

Прежде всего, из этого соотношения видно, что мощность, поглощаемая от поля, линейно зависит от квадрата поля и, следовательно, прямо пропорциональна падающей мощности. Такая линейная связь между  $P_{\text{погл}}$  и  $P_{\text{пад}}$  характерна для линейной электродинамики, когда параметры среды не зависят от величины амплитуды поля. Поэтому, при  $WT_1 \ll 1$  мы работаем в пределах применимости линейной электродинамики.

Из (5.47) следует, что  $P_{\text{погл}}$  сложным образом зависит от частоты поля  $\omega$ .

Эта зависимость описывается функцией  $g(\omega)$ , совпадающей с резонансной кривой простого колебательного контура. Вид этой кривой представлен на рис. 5.12

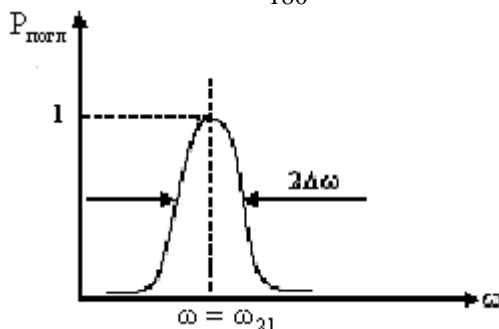


Рис. 5.12. Зависимость мощности поглощения от частоты поля

Максимум поглощения достигается при  $\omega = \omega_{21}$ . Полуширина кривой  $\Delta\omega$  по уровню  $0.5P_{max}$  пропорциональна  $\Delta\omega \approx 2/\tau$ .

Таким образом, сняв экспериментально зависимость  $P_{погл} = f(\omega)$ , можно определить время поперечной релаксации.

б) Случай сильных полей, когда  $WT_1 \gg 1$ . Тогда из выражения (5.47) имеем

$$P_{погл} = \hbar\omega \frac{1}{2T_1} (N_1^e - N_2^e). \quad (5.48)$$

1) Как видно из этого соотношения, в случае больших амплитуд полей, мощность поглощения не зависит от величины падающей мощности и, следовательно, среда нелинейная. Эффект нелинейной зависимости  $P_{погл}$  от падающей мощности носит название эффекта насыщения.

График зависимости  $P_{погл}$  от  $P_{пад}$  схематично изображен на рис. 5.13.



Рис. 5.13. Зависимость поглощаемой мощности от падающей

Существование эффекта насыщения можно пояснить, исходя из полученных ранее соотношений (5.44). Данные соотношения устанавливают связь между населённостями уровней в состоянии теплового равновесия и в состоянии инверсии. Из этих соотношений следует, что при слабых полях распределение частиц по уровням бoльцмановское (см. рис. 5.14).

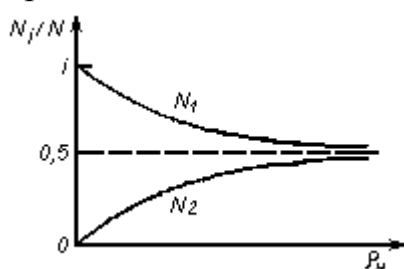


Рис. 5.14. Зависимость относительной населённости уровней от интенсивности (плотности) накачки

При увеличении амплитуды поля частицы с первого уровня будут переходить на второй. В результате населённость второго уровня будет увеличиваться, а первого падать. При увеличении поля до очень большой величины ( $W \gg 1/T_1$ ) населённости уровней становятся почти равными. Из рис. 5.14 следует, что эффект насыщения обусловлен зависимостью заселенностей уровней активных центров от плотности светового потока, проходящего через среду, что и подтверждает выражение (5.44) [7,8,9,12].

## Глава 6

### УСИЛЕНИЕ И ГЕНЕРАЦИЯ ЭЛЕКТРОМАГНИТНОГО ИЗЛУЧЕНИЯ

#### 6.1. Принцип работы квантовых усилителей и генераторов и пороговые условия генерации

В подразделах 5.1, 5.2 показано, что при некоторых условиях, суммарная мощность поглощения (а, следовательно, коэффициент поглощения  $\xi$  активного вещества) может оказаться отрицательным. В этом случае электромагнитная волна, проходя через квантовую систему, будет усиливаться. Действительно, для изменения интенсивности  $I(z)$  волны [2] имеем

$$I(z) = I(0)e^{-\alpha z}, \quad (6.1)$$

где  $\alpha = 2\sigma k = 2\sigma \frac{\omega}{c}$ .

Следовательно, при  $\xi < 0$  интенсивность электромагнитной волны увеличивается с удлинением пройденного через квантовую систему пути по экспоненциальному закону. Зависимость (6.1) обычно называют **законом Бугера**.

Поскольку  $\xi < 0$  только в квантовой системе с инверсной населенностью энергетических уровней, для обеспечения возможности **непрерывного** усиления электромагнитной волны необходимо поддерживать стационарную инверсную населенность. В импульсных квантовых устройствах такой необходимости нет, и процессы возбуждения квантовой системы и усиления могут быть разнесены во времени.

Выражение для коэффициента «усиления» (6.2), с учетом выражений (5.24а) и (5.24в) позволяет предъявить определенные требования к квантовой системе

$$\alpha = \frac{\pi\omega}{3nc\epsilon_0\hbar} |p_{ji}|^2 (N_j - N_i) \cdot g(\omega). \quad (6,2)$$

Прежде всего переход между соответствующими уровнями должен быть разрешенным  $|p_{ji}|^2 \neq 0$ , инверсия населенностей как можно больше, частота перехода  $\omega_{ji}$  должна быть как можно ближе к частоте электромагнитной волны (усиление носит резонансный характер).

Таким образом, квантовые системы при определенных условиях могут обеспечить усиление проходящих через них электромагнитных (световых) волн (среда будет действовать как усилитель). Для получения и поддержания инверсной насе-

ленности необходим источник энергии накачки, переводящий микрочастицы в возбужденное состояние.

Для того чтобы усилитель превратить в генератор, необходимо ввести подходящую положительную обратную связь. В лазере обратную связь обычно получают размещением активной среды между двумя зеркалами с высоким коэффициентом отражения. Усилитель становится регенеративным. При повышении коэффициента связи и коэффициента усиления такой усилитель может самовозбудиться, и будем иметь **квантовый генератор**. Частота генерации в нем определяется частотой настройки резонатора и частотой перехода  $\omega_{lm}$ .

Исходя из вышеизложенного, нетрудно установить структуру квантового генератора. Он должен включать три основных «элемента»: квантовую микросистему – активную среду, источник накачки (подсветки) и резонатор, обеспечивающий положительную обратную связь. Для **квантового усилителя** резонатор не является необходимым элементом, хотя обычно для повышения коэффициента усиления он используется.

Заметим здесь, что работа квантовых усилителей и генераторов, как отмечалось во вводной части курса, основана на индуцированных переходах, при которых излучение когерентно с «падающей» электромагнитной волной.

Принцип работы квантового генератора рассмотрим на примере газового лазера (рис. 6.1). В кварцевую трубку помещено газообразное рабочее вещество, инверсная населенность которого обеспечивается источником накачки. Открытый резонатор образован двумя зеркалами, плоскости которых перпендикулярны к оси трубки. При спонтанных переходах излучение фотонов обычно идет равновероятно во все стороны. Однако те фотоны, которые распространяются вдоль оси трубки, отражаются от зеркал и вновь возвращаются в активную среду. Выполняя роль индуцирующего поля, они вызывают вынужденные переходы

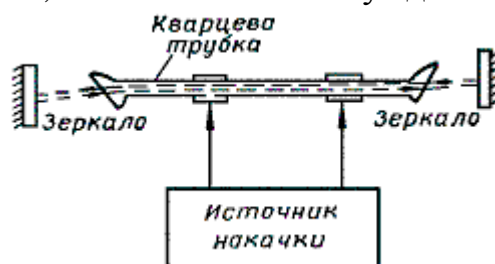


Рис. 6.1. Схема устройства лазера

возбужденных микрочастиц, затем повторно отражаются от зеркал и так далее. При каждом проходе квантовой микросистемы интенсивность поля нарастает и достигает, наконец, определенной конкретной величины. Часть световой энергии проходит через частично прозрачное зеркало и образует когерентный световой поток. Резонатор обеспечивает положительную обратную связь. Он «выбирает» из всех излученных фотонов только те, которые имеют вполне определенные частоты и направления распространения: частоты, при которых между зеркалами укладывается целое число полуволн, а направления распространения, для которых возможно многократное отражение фотонов от зеркал. При многократном отражении увеличивается время взаимодействия с рабочим веществом и повышается усиление соответствующей световой волны. В лазерах именно резонатор обуславливает временную и пространственную когерентность светового излучения. Однако следует отметить, что в лазерах генерация возможна лишь при выполнении некоторого порогового условия. Например, в Лазаре генерация начнется то-

гда, когда усиление активной среды компенсирует потери в нем (скажем, потери, обусловленные частичным выходом излучения из резонатора через зеркало). Усиление излучения за один проход в активной среде (т.е. отношение выходной и входной плотностей потока фотонов) равно  $\exp[\sigma(N_2 - N_1)\ell]$ , где  $\ell$  - длина активной среды. Если потери в резонаторе определяются только пропусканием зеркал, то порог генерации будет достигнут при выполнении условия  $R_1 R_2 \exp[2\sigma(N_2 - N_1)\ell] = 1$ , ( $R_1$  и  $R_2$  - коэффициенты отражения зеркал по интенсивности). Это условие показывает, что порог достигается тогда, когда инверсия населенностей приближается к некоторому критическому значению  $(N_2 - N_1)_{\text{пор}}$ , называемому *пороговой инверсией* и определяемому соотношением

$$(N_2 - N_1)_{\text{пор}} = -\ln(R_1 R_2) / 2\sigma\ell.$$

Как только достигнута пороговая инверсия, генерация разовьется из спонтанного излучения. Действительно, фотоны, которые спонтанно испускаются вдоль оси резонатора, будут усиливаться.

Достаточно полную информацию о работе квантовых усилителей и генераторов можно получить, рассматривая лишь оптически активные уровни и те, между которыми осуществляется накачка. Обычно ограничиваются анализом либо трехуровневых, либо четырехуровневых квантовых систем. Весьма часто, однако, при рассмотрении стационарных режимов, когда суммарное число частиц на оптически активных уровнях не изменяется во времени, можно ограничиться описанием этих двух уровней (двухуровневая идеализация) с соответствующими константами релаксации, а накачку учесть феноменологически. При этом анализ значительно упрощается [2].

## 6.2. Методы получения инверсии населенностей в квантовых генераторах и усилителях

При конструировании оптических генераторов и усилителей принципиальными этапами являются выбор активного вещества и способа его возбуждения, а также выбор электродинамической системы (резонатора), обеспечивающей эффективное взаимодействие электромагнитного поля с активным веществом. Как следует из п.б.1, для выполнения условий самовозбуждения лазера необходимо создать инверсную населенность энергетических уровней. Микрочастицы в возбужденное состояние могут быть переведены различными методами. Кратко остановимся на них.

В обычных условиях одновременное наличие трех процессов взаимодействия внешнего поля с квантовой системой: поглощения, излучения под действием индуцирующего поля и спонтанного излучения - поддерживает равновесное состояние системы. В условиях термодинамического равновесия среднее число частиц активного вещества, находящихся в двух энергетических состояниях, подчиняется распределению Больцмана (5.1) Поэтому в обычных условиях вещества поглощают электромагнитную энергию. Увеличение интенсивности внешнего поля приводит к увеличению поглощения до тех пор, пока число частиц на обоих уровнях, между которыми происходит обмен энергией поля и вещества, не выровняется. Такое состояние системы называют состоянием насыщения (п.5.4.2). В этом состоянии энергия поглощаемых квантов будет равна энергии, отводимой от системы релаксационными процессами. Однако можно создать условия, когда

число частиц на верхнем энергетическом уровне  $N_2$  превысит число частиц на нижнем энергетическом уровне  $N_1$ . Такая система с инверсией населенности неустойчива. Релаксационные процессы («спин-спиновые» и «спин-решеточные» взаимодействия), спонтанное излучение и взаимодействие с внешним полем будут возвращать систему в невозбужденное состояние. Для поддержания инверсной населенности необходимо, чтобы время жизни частиц на верхнем энергетическом уровне и постоянная времени релаксации превышали время перевода частиц из основного состояния на верхний энергетический уровень и время опустошения верхнего уровня за счет индуцированного излучения.

В оптических квантовых генераторах используются следующие методы создания инверсии населенности (отрицательной температуры):

- возбуждение световым потоком (оптическая накачка);
- возбуждение электронным ударом;
- возбуждение за счет неупругих столкновений атомов;
- возбуждение при диссоциации молекул (при столкновениях молекул);
- газодинамическая накачка
- инжекция носителей заряда через  $p$ - $n$  переход и другие.

Возбуждение активного вещества оптической накачкой (некогерентным светом) применимо при любом агрегатном состоянии вещества. Остальные методы возбуждения (кроме последнего) пригодны лишь в газообразных средах. Последний метод используется для инвертирования населенности в полупроводниках [2].

### 6.2.1. Двухуровневая схема

Выше (п.5.4.2) было показано что при  $g_1=g_2=1$

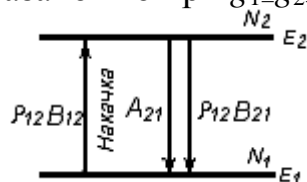


Рис. 6.2. Двухуровневая схема накачки

в двухуровневой схеме (из-за наличия эффекта насыщения) нельзя достичь инверсии с помощью электромагнитного поля, а следовательно, и усиления.

Если взять два уровня с разными статистическими весами  $g_1$  и  $g_2$ , то аналогично получим, что в пределе  $\rho_n \rightarrow \infty$  ( $\rho_n = \rho_{12}$  – плотность энергии поля накачки)

число частиц на верхнем уровне  $N_2 \rightarrow \frac{g_2}{g_1 + g_2} N$ , а на нижнем уровне

$N_1 \rightarrow \frac{g_1}{g_1 + g_2} N$ . Хотя при  $g_2 > g_1$  число частиц на уровне  $E_2$  может быть больше, чем

на уровне  $E_1$  населенность верхнего уровня  $N_2/g_2$  всегда меньше населенности нижнего уровня  $N_1/g_1$ , т. е. инверсии населенностей и усиления достигнуто не будет.

Полученные результаты однозначно указывают, что по двухуровневой схеме квантовые усилители и генераторы с оптической накачкой работать не могут. Аналогичные результаты получатся, если в системе уравнений (5.40) учесть безызлучательные процессы, а также применить накачку с помощью газового разряда, возбуждением частицами высоких энергий или газодинамическую накачку.



Инверсию населенностей в двух-уровневой системе можно получить, используя метод сортировки частиц. Именно этот метод привел к созданию первого квантового генератора – мазера на пучке молекул аммиака.

### 6.2.2. Трехуровневая схема

Результаты, полученные выше, показывают, что в двухуровневой квантовой системе создание инверсии населённости за счёт действия электромагнитного поля невозможно. В лучшем случае  $N_2 \approx N_1$ , но никогда не будет  $N_2 > N_1$ . Это однако не означает, что электромагнитное излучение вообще не может привести к возникновению инверсии населённости. Можно показать, (впервые это было замечено Н.Г.Басовым и А.М.Прохоровым в 1955 году) что в многоуровневых системах воздействие мощного электромагнитного излучения может привести к возникновению инверсии.

Механизм создания инверсии населенностей в трехуровневых схемах поясняет рис. 6.4. В зависимости от того, между какими уровнями достигается инверсия, различают трехуровневые схемы первого и второго типов. В схемах первого типа рабочий переход заканчивается в основном состоянии (рис. 6.3а), а в схемах второго типа – в возбужденном (рис. 6.3б). Накачка осуществляется по возможности селективно на уровень  $E_3$ . По трехуровневой схеме первого типа работает рубиновый лазер, а по схеме второго типа – гелий-неоновый газовый лазер.

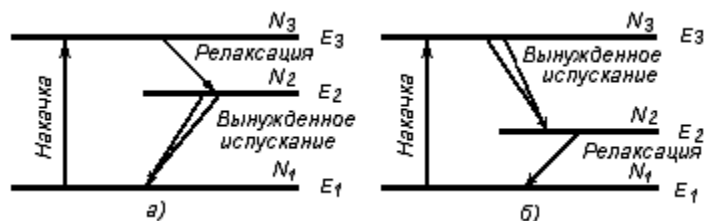


Рис. 6.3 Трехуровневые схемы работы квантовых усилителей и генераторов первого (а) и второго (б) типов

В трехуровневых схемах канал накачки, и канал усиления частично разделены. Это позволяет использовать для достижения инверсии наиболее универсальный метод оптической накачки, а также накачку с помощью газового разряда. Возможность получения инверсии населенностей с помощью оптической накачки в трехуровневой схеме довольно очевидна. Например, если в схеме второго типа осуществить селективный переход  $E_1 \rightarrow E_2$ , то уровень  $E_3$  окажется инверсно заселенным относительно уровня  $E_2$  (при  $kT \ll E_2 - E_1$ ). Из рис. 6.3 можно заключить, что накопление частиц на верхнем лазерном уровне ( $E_2$  в схеме первого типа и  $E_3$  — в схеме второго типа) будет в том случае, если релаксационные процессы  $E_3 \rightarrow E_2$  в схеме, изображенной на рис.6.3а и  $E_2 \rightarrow E_1$  в схеме (рис.6.3б) идут достаточно быстро, а верхний рабочий уровень является метастабильным.

Определим зависимость населенностей уровней от плотности (интенсивности) возбуждающего излучения накачки  $\rho_n = \rho_{13}$ . В качестве примера рассмотрим трехуровневую схему первого типа. Предположим, что возбуждение системы (накачка) осуществляется чисто оптическим путем в канале  $1 \rightarrow 3$ , а внешнее возбуждение в каналах  $2 \rightarrow 3$  и  $1 \rightarrow 2$  отсутствует (или пренебрежимо мало). Скорость релаксации  $3 \rightarrow 2$  обозначим  $\omega_{32}$ . Она может осуществляться за счет излучательных ( $A_{32}$ ) и безызлучательных ( $\Gamma_{32}$ ) переходов, так что  $u_{32} = A_{32} + \Gamma_{32}$ . Рас-

смотрим сначала режим усиления, когда активное вещество не находится в резонаторе. Соответствующие переходы изображены на рис. 6.4а.

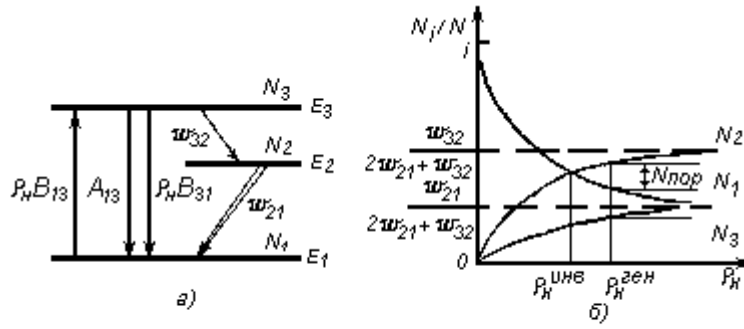


Рис. 6.4. Трехуровневая схема (а) и зависимость относительной населенности уровней от интенсивности накачки (б) в отсутствие (сплошные линии) и при наличии (пунктирные линии) резонатора

Кинетические уравнения в этом случае для стационарного режима будут иметь вид

$$\left. \begin{aligned} \frac{dN_3}{dt} &= c_n B_{13} N_1 - [c_n B_{31} + (w_{32} + A_{31})] N_3 = 0, \\ \frac{dN_2}{dt} &= w_{32} N_3 - w_{21} N_2 = 0, \\ N_1 + N_2 + N_3 &= N. \end{aligned} \right\} \quad (6.3)$$

Приняв для простоты кратности вырождения уровней  $g_1=g_2=g_3=1$  и решая систему уравнений (6.3), находим населенности уровней:

$$\begin{aligned} N_1 &= \left[ \frac{w_{21}(A_{31} + w_{32} + c_n B_{31})}{w_{21}(w_{32} + A_{31}) + (2w_{21} + w_{32})c_n B_{31}} \right] N, \\ N_2 &= \left[ \frac{w_{32} c_n B_{31}}{w_{21}(w_{32} + A_{31}) + (2w_{21} + w_{32})c_n B_{31}} \right] N, \\ N_3 &= \left[ \frac{w_{21} c_n B_{31}}{w_{21}(w_{32} + A_{31}) + (2w_{21} + w_{32})c_n B_{31}} \right] N. \end{aligned}$$

Зависимость относительной населенности уровней  $N_i/N$  ( $i=1, 2, 3$ ) от плотности накачки согласно выражениям (6.4) представлена на рис. 6.4б. При больших плотностях накачки населенности основного и верхнего состояний в пределе  $\rho_n \rightarrow \infty$  стремятся к

$$\lim_{c_n \rightarrow \infty} N_1 = \lim_{c_n \rightarrow \infty} N_3 = \frac{w_{21}}{2w_{21} + w_{32}} N,$$

а населенность уровня  $E_2$  при  $\rho_n \rightarrow \infty$  стремится к

$$\lim_{c_n \rightarrow \infty} N_2 = \frac{w_{32}}{2w_{21} + w_{32}} N.$$

При  $u_{32} > u_{21}$ , как видно из рис. 6.4б и соотношений (6.4), начиная с некоторого значения плотности накачки  $\rho_n$  между уровнями  $E_2$  и  $E_1$  будет наблюдаться

инверсия населенностей ( $N_2 > N_1$ ). Величина  $c_n^{\text{инв}}$  называется *пороговой плотностью накачки по инверсии*. С увеличением  $c_n > c_n^{\text{инв}}$  инверсия увеличивается.

Отметим, что пороговая накачка для генерации будет превышать пороговую накачку по инверсии, поскольку для возникновения генерации необходимо выполнить еще условия самовозбуждения. Приравнявая выражения (6.4) и (6.4а), найдем  $c_n^{\text{инв}}$ :

$$c_n^{\text{инв}} = \frac{u_{21}(u_{32} + A_{31})}{B_{31}(u_{32} - u_{21})}. \quad (6.5)$$

Из проведенного рассмотрения вытекает, что для накопления частиц на уровне  $E_2$  и создания максимальной инверсии населенностей наиболее выгодны системы с большим значением  $\omega_{32}$  (переход  $3 \rightarrow 2$  должен быть быстрым), малым значением  $\omega_{21}$  (уровень  $E_2$  должен быть метастабильным) и большим коэффициентом Эйнштейна  $B_{13}$  (оптический переход  $1 \rightarrow 3$  должен быть разрешен).

В соотношениях (6.3), а следовательно, и (6.4) вероятность перехода  $\omega_{21}$  считалась постоянной, не зависящей от скорости накачки. Это справедливо в отсутствие генерации, когда опустошением уровня  $E_2$  за счет вынужденных переходов можно пренебречь. Если же активное вещество помещено в резонатор, то после превышения инверсии над некоторым пороговым значением начинает развиваться процесс генерации. Наличие интенсивного излучения на частоте  $\omega_{21}$  и связанных с ним вынужденных переходов  $2 \rightarrow 1$  и  $1 \rightarrow 2$  вызывает изменение населенности уровней. Возрастание интенсивности накачки  $\rho_n$  приводит к увеличению инверсии  $\Delta N = N_2 - N_1$  и согласно (6.2) – к увеличению усиления. Это увеличивает  $\rho_{21}$ , что в свою очередь обуславливает выравнивание населенностей  $N_2$  и  $N_1$  т. е. происходит процесс насыщения инверсии населенностей и коэффициента усиления. Увеличение поступления частиц на уровень  $E_2$ , вызываемое ростом накачки, компенсируется возрастанием числа активных переходов  $2 \rightarrow 1$ . Поэтому в режиме генерации инверсия  $\Delta N = N_2 - N_1$  остается приблизительно постоянной, как изображено пунктирными линиями на рис. 6.4б. Ее значение примерно равно пороговой перенаселенности, при которой усиление превышает потери в генераторе и которая достигается при пороговой накачке, равной  $c_n^{\text{ген}}$  [11].

### 6.2.3. Четырехуровневая схема

В этой схеме каналы генерации и накачки полностью разделены, что позволяет получать инверсию населенностей при минимальных уровнях накачки. Очевидно, трехуровневые схемы обоих типов можно рассматривать как частный случай четырехуровневой схемы.

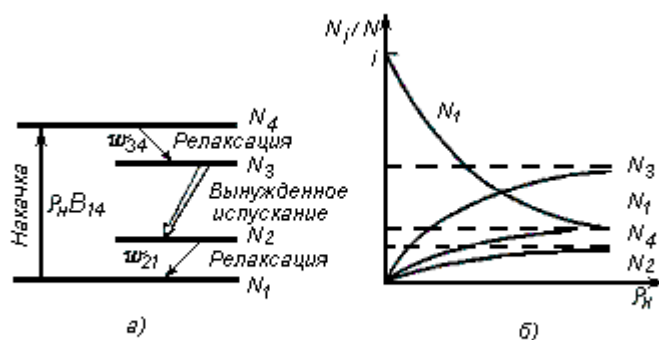


Рис. 6.5. Четырехуровневая схема накачки (а) и зависимость относительной населенности уровней от интенсивности накачки (б) в отсутствие (сплошные линии) и при наличии (пунктирные линии) резонатора

Механизм создания инверсии между рабочими лазерными уровнями  $E_3$  и  $E_2$  в четырехуровневой схеме и переходы, определяющие сущность процесса, изображены на рис. 6.5 а. Так же, как и в трехуровневой системе, накачка переводит атомы из основного состояния  $E_1$  в верхнее состояние  $E_4$ . Инверсия достигается между уровнями  $E_3$  и  $E_2$ . Для этого необходимо, чтобы уровень  $E_3$  быстро заполнялся, а уровень  $E_2$  опустошался за счет переходов  $E_4 \rightarrow E_3$  и  $E_2 \rightarrow E_1$  соответственно. Наличие реально существующих переходов  $E_4 \rightarrow E_1$ ,  $E_4 \rightarrow E_2$ ,  $E_3 \rightarrow E_1$ ,  $E_1 \rightarrow E_2$  и  $E_3 \rightarrow E_4$  приводит к уменьшению инверсии, т. е. снижает качество усилителя и генератора.

Анализ кинетики заселения уровней в четырехуровневой схеме может быть проведен с помощью кинетических уравнений, аналогично тому, как это было сделано для трехуровневой схемы. В данном случае эти уравнения будут лишь более громоздкими. Зависимость населенности уровней от интенсивности накачки в четырехуровневой схеме представлена на рис. 6.5б. Инверсия населенностей между уровнями  $E_3$  и  $E_2$  будет достигаться при соблюдении неравенства

$$w_{21}w_{43} > [w_{42}(w_{31} + w_{32}) + w_{32}w_{43}] \frac{g_3}{g_2}, \quad (6.6)$$

где  $w_{mn}$  – вероятности переходов  $m \rightarrow n$ .

Из рис. 6.5б видно, что в четырехуровневой схеме (при  $kT \ll E_2 - E_1$ ) ничтожное накопление частиц на третьем уровне приводит к инверсии населенностей, т. е. пороговая плотность излучения накачки по инверсии  $c_{и}^{инв}$  в четырехуровневой схеме равна нулю. Поскольку здесь не учитываются потери в активном веществе, реальная пороговая накачка в активных материалах, работающих по четырехуровневой схеме, несколько больше нуля, но может быть очень мала.

Так же, как и в трехуровневой схеме, зависимость населенности уровней в режиме генерации будет отличаться от изображенной на рис. 6.5б. За счет интенсивных вынужденных переходов с испусканием в канале генерации  $E_3 \leftrightarrow E_2$  значение  $N_3$  будет уменьшаться, а  $N_2$  – возрастать, приводя к насыщению коэффициента усиления и населенностей уровней.

В заключение отметим следующие моменты, относящиеся как к трехуровневым, так и к четырехуровневым схемам.

а) Для исключения термического заселения необходимо, чтобы энергетические расстояния между уровнями  $E_4—E_3$  и  $E_2—E_1$  были больше  $kT$ . Однако они не должны быть слишком большими, поскольку в противном случае большая часть энергии накачки будет расходоваться бесполезно. Это приведет к уменьшению к. п. д. [в пределе  $\eta < (E_3 - E_2)/(E_4 - E_1)$ ] и разогреву активного вещества, если избыточная энергия при релаксационных процессах в конечном итоге выделится в виде тепла.

б) При оптической накачке, когда источник накачки излучает в широкой области спектра, необходимо, чтобы верхний уровень  $E_4$  (или система уровней) был достаточно широким. Это нужно для более полного использования излучения накачки

в) Для исключения самопоглощения, приводящего к переходам  $E_1 \rightarrow E_2$  и  $E_3 \rightarrow E_4$ , желательно, чтобы релаксационные процессы осуществлялись за счет неоптических (безызлучательных) переходов.

г) Время жизни на верхнем лазерном уровне  $E_3$  должно определяться излучательными процессами, а вероятность безызлучательных переходов с этого уровня должна быть минимальной [2,4,11,13].

## Глава 7

### Двухуровневые квантовые генераторы и усилители

В квантовых приборах СВЧ используются переходы, дающие излучение в СВЧ диапазоне. Здесь рассмотрены только два типа этих приборов. Квантовые стандарты частоты (генератор на пучке молекул аммиака) и квантовые парамагнитные усилители.

#### 7.1. Квантовые стандарты частоты

Квантовые стандарты частоты разделяются на пассивные и активные.

В пассивных КСЧ используются линия поглощения атомов или молекул. В активных КСЧ производится сравнение частоты вспомогательного генератора с частотой излучения квантового генератора.

В КСЧ используются вещества с узкими спектральными линиями и очень слабой зависимостью частоты квантовых переходов от внешних воздействий.

На ширину линии влияют соударения между частицами и частиц со стенками, электрические и магнитные поля, эффект Доплера. Естественная ширина спектральной линии обратно пропорциональна времени жизни. Для используемых в КСЧ уровне аммиака  $\Delta\nu \approx 10^{-8}$  Гц, а для уровней сверхтонкой структуры щелочных элементов  $\Delta\nu \approx 10^{-13}$  Гц.

Рассмотрим факторы, вызывающие уширение линии. При низком давлении газа и большой длине свободного пробега соударений между молекулами практически нет, а так как пучок сфокусирован, можно пренебречь и соударениями молекул со стенками резонатора. Поэтому уширения спектральных линий, вследствие соударений, в молекулярных генераторах нет. Для уменьшения доплеровского уширения следует увеличивать фазовую скорость поля вдоль оси резонатора. Для этого в резонаторе используют тип колебаний с однородным распределением поля вдоль оси. Практически ширина спектральной линии определяется временем пролета молекул сквозь резонатор и составляет несколько килогерц. Такой полосой пропускания прибор может обладать в режиме усиления, когда поток молекул, поступающих в резонатор, ниже порогового. После установления режима автоколебаний спектральная линия генерации становится уже по сравнению с линией усиления.

Задачи, которые требуют применения КСЧ, решаются, как правило, в низкочастотном диапазоне. Поэтому главная задача при построении КСЧ - переход от очень стабильного сигнала СВЧ молекулярного генератора к сигналу той же относительной стабильности, но лежащей в низкочастотном диапазоне. Первым, примененным в КСЧ, был молекулярный генератор на аммиаке, поэтому принцип работы конструктивные особенности рассмотрим на этом веществе.

##### 7.1.1. Принцип действия молекулярного квантового генератора

Молекулярный генератор – квантовый генератор, в котором активной средой является молекулярный газ или молекулярный пучок. В молекулярном генераторе на пучке молекул аммиака электромагнитные колебания генерируются вследствие вынужденного излучения микрочастиц. Он построен на двухуровне-

вой схеме, причем, рабочими уровнями является пара инверсионных колебательных уровней с частотой перехода  $\nu_{\text{пер}}=29870$  МГц ( $\lambda=12,5$  мм).

Покажем устройство генератора на схеме, изображенной на рис. 7.1.

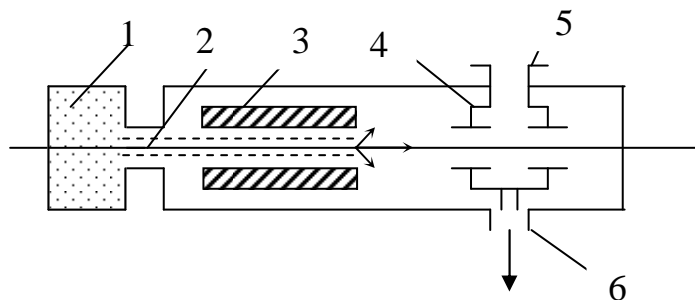


Рис. 7.1. Блок-схема генератора.

1-источник молекул, 2-пучок молекул, 3-сортирующая система, 4-объемный резонатор, 5-выходной волновод, 6-откачной патрубком.

Из источника 1 молекулы аммиака  $\text{NH}_3$  выходят через капилляры и образуют молекулярный пучок 2. Пучок проходит сквозь сортирующую систему 3, в которой из пучка удаляются молекулы, находящиеся на нижнем энергетическом уровне пары рабочих инверсионных состояний. Далее пучок, в котором число молекул на верхнем уровне преобладает, пропускается через объемный резонатор 4, настроенный на частоту используемого перехода. Электромагнитное поле резонатора воздействует на пролетающие молекулы и вызывает их вынужденное излучение. Таким образом осуществляется обратная связь. Если мощность излучения превысит сумму потребляемых мощностей, генератор самовозбуждается.

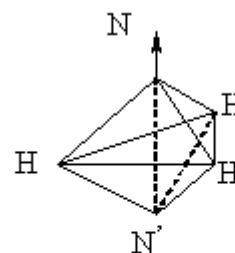


Рис. 7.2 Молекула аммиака

**Энергетические уровни молекул аммиака.** Молекула аммиака имеет вид правильной пирамиды как показано на рис. 7.2 (тетраэдр), в вершине которого помещается атом азота, а в вершинах треугольника, образующего основание, находятся атомы водорода. Данная система имеет много степеней свободы. Она может совершать поступательное, вращательное, колебательные движения.

В молекулярном генераторе используются уровни, которые получаются вследствие их расщепления за счет так называемого инверсионного колебания атома азота относительно плоскости Н-Н-Н (рис.7.2). Потенциальная энергия атома азота  $E(z)$  в зависимости, от расстояния его на оси  $Z$  от плоскости, в которой лежат атомы водорода, приведены на рис. 7.3.

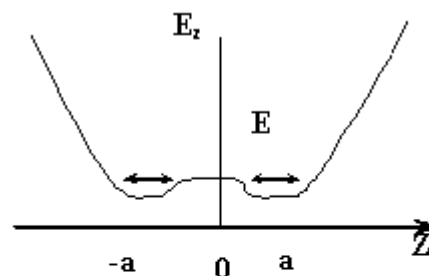


Рис. 7.3. Распределение потенциальной энергии атома азота относительно плоскости атомов водорода.

Минимум потенциальной энергии атома азота достигается при смещении атома от плоскости Н на  $\pm a$ . В точках  $z=\pm a$  силы отталкивания и притяжения молекул уравниваются, и при таких положениях атома азота система обладает наименьшей потенциальной энергией. Атом азота как бы колеблется в потенциальной яме. Он может перейти из одной потенциальной ямы в другую. Это движение оказывается очень важным и происходит за счет туннельного эффекта.

В спектре молекул аммиака можно делить пару уровней. Уровням с большей энергией атома азота соответствуют антифазные (антисимметричные) колебания  $E_a$ . Уровням с меньшей энергией – колебания в потенциальных ямах (симметричные)  $E_s$ . При установлено, что энергия симметричного состояния меньше энергии асимметричного состояния (рис. 7.4).

$$\nu = \frac{E_a - E_s}{h}, \quad (7.1)$$

где  $\nu$ - частота линии поглощения или

Сортировка частиц в молекулярном генераторе осуществляется с помощью неоднородного электрического поля. Действие такого поля на молекулы зависит от ее энергии. Так, попадая в постоянное электрическое поле  $E_0$ , молекула аммиака приобретает дополнительную потенциальную энергию, так что уровни ее энергии смещаются (эффект Штарка).

Величина и знак этого смещения различны для разных уровней. На рис. 7.4 изображена зависимость положения двух уровней молекулы аммиака  $E_a$  и  $E_s$  от величины поля  $E_0$ . При увеличении абсолютной величины  $E_0$  энергия молекул в состоянии  $E_a$  уменьшается, а в состоянии  $E_s$  увеличивается. Так как система всегда стремится занять состояние с наименьшей энергией, то молекулы в антисимметричном состоянии будут группироваться в области, где  $E_0=0$ , а в симметричной области, где  $E_0$  стремится к  $E_{\max}$ .

Сортирующее устройство рис. 7.5, называемое квадрупольным конденсатором, создает неоднородное электрическое поле  $E_0$ . В центральной части пучка поле равно нулю, а по мере удаления от оси пучка плавно нарастает. Поэтому рис. 7.4 отражает также зависимость энергии молекул в состоянии  $E_a$  и  $E_s$  от ее расстояния  $r$  от оси симметрии конденсатора. Т.к. по законам механики любая система испытывает силу, направленную в сторону уменьшения ее потенциальной энергии, то молекулы пучка, которые в конденсаторе находятся в верхнем состоянии  $E_a$ , приближаются к центру и фокусируются. Молекулы же, находящиеся в нижнем состоянии  $E_s$ , удаляются от центра пучка и не попадают в резонатор.

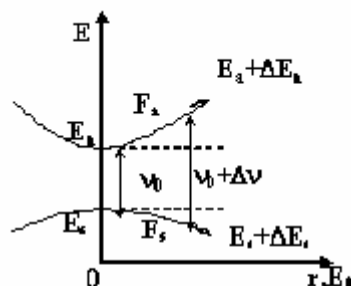


Рис. 7.4. Энергетические уровни молекулы аммиака

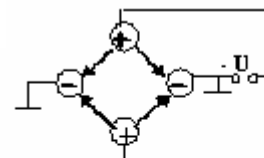


Рис. 7.5. Квадрупольный конденсатор

### 7.1.2. Условие самовозбуждения генератора



После селектора молекулы попадают в резонатор, где поле усиливается. Кроме усиления имеют место и потери, в частности

1. потери на излучение ( $P_{\text{пот.изл}}$ ),
2. потери на тепловые преобразования ( $P_{\text{тепл}}$ ).

Общие потери равны  $P_{\text{пот}} = P_{\text{пот.изл}} + P_{\text{тепл}}$ .

Мощность, отдаваемая молекулами  $\text{NH}_3$  должна превышать мощность потерь. В стационарном режиме величины  $P_{\text{пот}}$  и  $P_{\text{изл}}$  равны.

$P_{\text{изл}}$  – мощность, отдаваемая молекулами  $\text{NH}_3$  за счет индуцированного излучения. Как  $P_{\text{изл}}$ , так и  $P_{\text{пот}}$  зависят от амплитуды поля

$$P_{\text{изл}} \sim h\nu_{21} W(N_2 - N_1), \text{ где } W \sim E_0^2, \quad (7.2)$$

$$P_{\text{пот}} \sim E_0^2. \quad (7.3)$$

Из (7.3) следует, что потери пропорциональны энергии поля. На графике (рис. 7.6) также видно, что  $P_{\text{изл}}$  также растет с ростом  $E_0^2$ . При значительном увеличении  $E_0^2$ , вследствие эффекта насыщения, рост  $P_{\text{изл}}$  замедляется и стремится к некоторому пределу. Если в рассматриваемом типе молекулярного генератора в стационарном режиме  $P_{\text{пот}} = P_{\text{изл}}$ , то выходная мощность определяется как точка ( $P_0$ ) пересечения двух кривых  $P_{\text{изл}}$  и  $P_{\text{пот}}$ . Здесь все потери связаны с выводом энергии из молекулярного генератора.

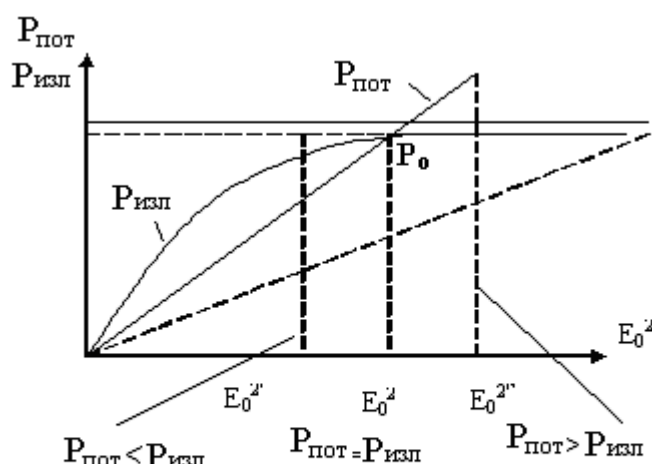


Рис. 7.6. Графики зависимости  $P_{\text{пот}}$  и  $P_{\text{изл}}$  от величины электрического поля  $E_0$ .

Покажем, что среднее значение мощности, излучаемой молекулярным генератором в точке  $P_0$ , стабильно. Пусть увеличилась величина  $E_0^2$  до  $E_0^{2''}$ . В этом случае  $P_{\text{пот}} > P_{\text{изл}}$ , что ведет к уменьшению генерируемой мощности, пока она не станет равной  $P_0$ . Если запасенная энергия  $E_0^2$  уменьшилась до  $E_0^{2'}$ , где потери  $P_{\text{пот}} < P_{\text{изл}}$  и мощность начнет возрастать вплоть до  $P_0$ . Таким образом, при любых отклонениях средней мощности от равновесного значения молекулярный генератор стремится вернуться к этому стационарному состоянию.

Условием генерации генераторов является:

1. Баланс амплитуд.
2. Баланс фаз.

Баланс амплитуд состоит в том, что происходит нарастание поля из-за инверсной населенности и индуцированных переходов.

Потери в резонаторе состоят из:

1. Тепловых потерь в стенках резонатора (потери мощности в стенках резонатора  $P_a$ ).
2. Потерь, связанных с наличием устройств связи ( $P_{\text{св}}$ ).

$$P_{\text{пот}} = P_a + P_{\text{св}} \quad - \text{общие потери в системе} \quad (7.4)$$

$$P_{\text{изл}} > P_{\text{пот}} \quad - \text{баланс амплитуд} \quad (7.5)$$

$$P_{\text{изл}} = \frac{h\nu W(N_2^e - N_1^e)}{1 + 2WT_1} \quad (7.6)$$

$$\text{или } P_{\text{пот}} = kE^2, \quad (7.7)$$

где  $k$  - коэффициент пропорциональности,  $E$  - напряженность поля,  $N^e$  – населенность в отсутствие поля,  $N$  – населенность при наличии поля.

Определим условия возникновения порога самовозбуждения. Пусть  $P_{\text{пот}} = P_{\text{изл}}$ . Разделим обе части на  $W\nu$ , где  $W$  - энергия запасенная в резонаторе,  $\nu$  - частота генерации, тогда получим

$$\frac{P_{\text{пот}}}{W\nu} = \frac{P_{\text{изл}}}{W\nu} = \frac{1}{Q}. \quad (7.8)$$

Добротность выражается следующими соотношениями

$$\frac{W\nu}{P_{\text{пот}}} = Q \quad \text{или} \quad Q = \frac{\nu_0 L}{r}, \quad (7.9)$$

где  $L$  - длина резонатора,  $r$  - сопротивление потерь.

Рассмотрим случай слабых полей  $W \ll 1$ ,

$$\text{т.е. } P_{\text{изл}} = h\nu W(N_2^e - N_1^e), \quad (7.10)$$

где  $W \approx B_{12}\rho(\nu) \cdot g(\nu)$ ,  $\rho\nu = 1/2 \epsilon E^2$  и  $g(\nu) = \frac{1}{\sqrt{1 - \left(\frac{\nu - \nu_0}{\Delta\nu}\right)^2}}$ ,

тогда, энергия запасенная в резонаторе, определится как:

$$W = \frac{1}{2} B_{12} \epsilon E^2 \frac{1}{\sqrt{1 - \left(\frac{\nu - \nu_0}{\Delta\nu}\right)^2}}. \quad (7.11)$$

Используя, приведенные соотношения, путем несложных преобразований, получим разность населенностей:

$$(N_2^e - N_1^e)_{\text{пор}} = \frac{\nu \sqrt{1 - \left(\frac{\nu - \nu_0}{\Delta\nu}\right)^2}}{B_{12} h Q} = \frac{r \nu \sqrt{1 - \left(\frac{\nu - \nu_0}{\Delta\nu}\right)^2}}{B_{12} h \nu L}. \quad (7.12)$$

### Параметры молекулярного генератора

Мощность генерации определена соотношением (7.6):

Общие потери определяются суммой тепловых потерь, потерь, выделяемых на нагрузке и потерь на излучение.

$$P_{\text{вых}} = P_{\text{теп}} + P_{\text{изл}} + P_{\text{нагр}}. \quad (7.13)$$

При хорошо выполненном резонаторе тепловые потери и потери на излучение много меньше мощности, выделяемой на нагрузку.

Пусть  $P_{\text{вых}} \approx P_{\text{н}}$ . Для самовозбуждения генератора необходимо чтобы селектор поставлял не менее  $10^{14}$  мол/сек. Значит  $2WT_1 \gg 1$  и

$$P_{\text{ген}} = hv \frac{N_2^e}{2T_1} \approx P_{\text{н}} . \quad (7.14)$$

Проведем количественную оценку  $P_{\text{ген}}$ .

$N_1^e$  - можно не учитывать, т.к. оно очень мало,  $h=1,05 \cdot 10^{-34}$ ,  $\nu_{21}=24 \cdot 10^9$ ,  $\frac{N_2^e}{T_1} \approx 10^{14} \frac{\text{мол}}{\text{сек}}$ , тогда  $P \approx 1,05 \cdot 10^{-34} \cdot 2 \cdot 3,14 \cdot 24,5 \cdot 10^9 \cdot 10^{14} \approx 10^{-9} [\text{Вт}]$ .

Из этого примера видно, что все молекулярные генераторы СВЧ маломощны.

**Стабильность частоты аммиачного генератора.** При попадании активных молекул в резонатор происходят следующие процессы. Шумовое поле аммиака (представляющее собой спонтанное излучение) вызывает в резонаторе свободные колебания. Это поле, частота которого близка к частоте перехода, вызывает поляризацию молекул аммиака. При этом диполи молекул ориентируются по направлению поля. Среда готова к индуцированному излучению. В результате поляризации осуществляются синхронные переходы, при которых возникает индуцированное излучение. В резонаторе возникают вынужденные колебания, которые усиливаются индуцированным излучением. При определенном числе активных молекул и соответственно при определенной величине поля и поляризации, фаза вынужденного колебания совпадает с фазой исходного поля. Возникает баланс фаз, при котором происходит генерация. В этом случае, если дипольные моменты параллельны, дипольный момент равен числу активных молекул.

Если  $P \sim E$ , то мощность  $P$  пропорциональна напряженности излучения  $E_{\text{изл}}$ .  $E_{\text{изл}} \sim P/Z_k$ , где  $Z_k$  - сопротивление контура.

Значит  $Z_k = R + j(\omega L - 1/\omega C) = R + jL(\omega - \omega_0^2/\omega) = R + j\omega_0 L(\omega/\omega_0 - \omega_0/\omega)$ ,

где  $\omega$  - частота вынужденных колебаний,  $\omega_0$  - резонансная частота или частота генерации

$$(\omega/\omega_0 - \omega_0/\omega) = (\omega^2 - \omega_0^2)/\omega\omega_0 = (\omega - \omega_0)(\omega + \omega_0)/\omega\omega_0 = 2(\omega - \omega_0)/\omega_0 \quad (7.15)$$

$$Z_k = R \left[ 1 + j2 \frac{\omega_0 L}{R} \frac{\omega - \omega_0}{\omega_0} \right]. \quad (7.16)$$

Так как частота резонатора не зависит от  $R$ , то в (7.16)  $R$  опускаем и, следовательно:

$$E_{\text{изл}} \sim P / (1 + j2Q(\omega - \omega_0)/\omega_0). \quad (7.17)$$

По аналогии запишем

$$P \sim E / (1 + j2Q_L(\omega - \omega_0)/\omega_0),$$

где  $Q_L$  - добротность молекулярной линии  $Q_L = 1/2\omega_0\tau$ .

Из этих двух выражений следует:

$$\bar{E}_{\text{изл}} \sim \frac{\bar{E}}{\left(1 + 2jQ \frac{\omega - \omega_0}{\omega_0}\right) \left(1 + 2jQ \frac{\omega - \omega_{21}}{\omega_{21}}\right)}. \quad (7.18)$$

Чтобы  $E_{\text{изл}}$  и  $E$  были в фазе необходимо, чтобы знаменатель был чисто активным, т.е.  $Q \frac{\omega - \omega_0}{\omega_0} + Q \frac{\omega - \omega_{21}}{\omega_{21}} = 0$  - условие баланса фаз.

Оценим количественно стабильность такого генератора. Возьмем характерные величины добротностей  $Q \sim 10^4$ ,  $Q_{\text{л}} \sim 10^7$ ,  $\omega \sim \omega_{21}$ . В первом приближении, определим частоту  $\omega$  и нестабильность частоты генератора ( $\Delta\omega$ ) при  $\omega_{21} \approx \omega_0$

$$\omega = \omega_{21} - \frac{Q}{Q_{\text{л}}} (\omega_{21} - \omega_0), \quad \Delta\omega = \frac{Q}{Q_{\text{л}}} \Delta\omega_0^{\text{р}}.$$

Зададим  $\Delta\omega_0 = 10^3$  Гц, а  $\omega_{\text{л}} = 24 \cdot 10^9$  Гц, тогда получим

$$\frac{\Delta\omega}{\omega_{\text{л}}} = \frac{Q}{Q_{\text{л}}} \frac{\Delta\omega_0^{\text{р}}}{\omega_{\text{л}}} = \frac{10^{-3} 10^3}{24 \cdot 10^{10}} \approx 5 \cdot 10^{-11}, \quad \frac{Q}{Q_{\text{л}}} \sim 10^{-5},$$

$\Delta\omega_0^{\text{р}}$  - изменение частоты в резонаторе ( $10^3$  Гц) за счет несогласованности с нагрузкой, за счет не точной настройки резонатора, за счет изменения напряжения на селекторе (например,  $\Delta\omega \cdot \omega_{\text{л}} = 5 \cdot 10^{-11} \cdot 24 \cdot 10^9 = 1$  Гц).

Стабильность частоты, достигаемая в квантовых стандартах частоты (КСЧ) в первую очередь определяется характеристиками использованного молекулярного генератора

Вслед за генератором на аммиаке были построены генераторы на других веществах и все они практически работают в СВЧ диапазоне.

Характерными являются:

1. Цезиевый стандарт стал основой эталона времени и частоты.
2. Генератор на молекулах формальдегида  $\text{CH}_2\text{O}$   $f=72838$  МГц ( $\lambda=4$  мм).
3. На молекулах синильной кислоты  $\text{HCN}$   $f=88630$  МГц ( $\lambda=3$  мм).
4. На молекулах тяжелого аммиака  $\text{ND}_3$   $f=1656$  МГц.
5. На атомарном водороде. Нестабильность  $10^{-3}$ ,  $f=1420,405$  МГц  $R_{\text{вых}}=10^{12}-10^{11}$  Вт,  $\lambda=21$  см, используется как первичный стандарт частоты.

Устройство этих генераторов аналогично генератору на  $\text{NH}_3$ . Исключение составляет генератор на водороде, у которого используется не электрический дипольный переход, а магнитный дипольный переход.

## 7.2. Квантовые парамагнитные усилители СВЧ (КПУ)

КПУ – это СВЧ приборы, в которых используются энергетические уровни парамагнитных веществ.

### 7.2.1. Парамагнетизм в кристаллах

В кристаллах каждый  $i$ -ый электрон, входящий в состав атома или молекулы, обладает определенным орбитальным  $\mu_{ei}$  и спиновым  $\mu_{si}$  магнитными моментами. Полный магнитный момент  $M$  атома или молекулы определяется суммой спиновых и орбитальных магнитных моментов отдельных электронов:

$$M = \sum_i (\mu_{si} + \mu_{Li}), \quad (7.19)$$

где  $\mu_{si} = g \cdot \beta \cdot S$ ;  $\mu_{Li} = g \cdot \beta \cdot L$ ;  $g$  - фактор спектроскопического расщепления или  $g$ -фактор, для спиновых моментов  $g=2$ ;  $\beta$  - магнетон бора;  $S$  – спиновый момент;  $L$  – орбитальный механический момент.

Взаимодействие магнитных моментов атомов и молекул с внешним магнитным полем вызывает намагничивание веществ. В зависимости от характера этого взаимодействия различают парамагнетики и диамагнетики.

**Парамагнетики** – это вещества, у которых под действием внешнего магнитного поля, возникает магнитный момент, направленный вдоль поля. Атомы и молекулы парамагнетиков обладают постоянным отличным от нуля магнитным моментом.

**Диамагнетики** – это вещества атомы, ионы, молекулы которых имеют целиком заполненные электронные оболочки, спины скомпенсированы, момент  $M=0$  (например, инертные газы).

Устойчивый парамагнетизм наблюдается в веществах, обладающих частично заполненными внутренними электронными оболочками с нескомпенсированными спинами. В кристаллах орбитальные моменты электронов в значительной степени скомпенсированы действием электрического поля решетки и их вклад в полный момент  $M$  атома или молекулы невелик, на спин же поле кристаллической решетки влияет незначительно. Поэтому полный магнитный момент оценивается суммой спиновых магнитных моментов отдельных электронов.

Постоянным магнитным дипольным моментом, вступающим во взаимодействие с внешним магнитным полем, обладают электроны атомов группы железа (никель, кобальт, хром и др.) и группы редкоземельных элементов (лантан, неодим, самарий, европий, гадолиний и др.), а также лантаниды (соединения Al и Cu).

**Магнитный момент во внешнем поле.** Рассмотрим взаимодействие магнитного момента электрона с внешним полем. Электрон, обладающий одновременно механическим и магнитным моментами, представляет собой миниатюрный намагниченный волчок (рис. 7.7). Это в равной мере относится и к спиновому и орбитальному моментам.

1. В момент включения  $H_0$  ось волчка образовала с направлением  $H_0$  угол  $\varphi$ .

2. После приложения поля на волчок начинает действовать момент сил, стремящихся

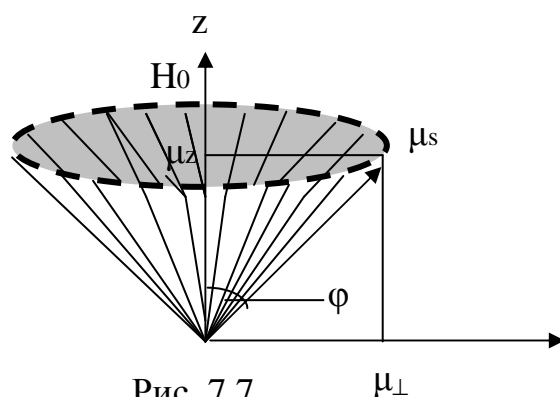


Рис. 7.7

совместить его ось с направлением на  $H_0$ .

Однако, за счет гироскопических свойств ось волчка не совмещается с направлением поля, а прецессирует относительно этого поля, описывая некоторый конус прецессии и сохраняя угол  $\varphi$ . Угловая скорость прецессии,  $w_{пр} = \gamma H_0$ , где  $\gamma = \mu_s/S$ .

Из рис. 7.7 видно, что с полем  $H_0$  взаимодействует только  $M_z$  составляющая магнитного момента, т.е.  $M_{\perp}$  в среднем за период равна 0, ввиду вращательного движения  $M_{\perp}$ .

При решении статических задач, можно магнитный волчок заменить магнитным диполем, магнитный момент которого равен его  $Z$  – ой составляющей (т.е. его среднему значению). Причем, магнитный момент может иметь разные знаки.

Магнитный диполь во внешнем поле может иметь только два положения равновесия: устойчивое и неустойчивое. При малейшем отклонении от устойчивого состояния момент сил стремится вернуть диполь в устойчивое состояние.

Рассмотрим влияние внешнего поля на парамагнитное вещество.

### 7.2.2. Энергетические уровни парамагнитных ионов

Энергетические уровни парамагнитных ионов. Если парамагнитный ион поместить в магнитное поле, то он будет взаимодействовать с этим полем. Энергия взаимодействия определяется следующим образом:  $\epsilon_{вз} = \mu H$ .

Для повышения плотности активных частиц в качестве рабочего вещества КПУ используют не свободные парамагнитные ионы, а парамагнитные кристаллы. Обычно парамагнитные ионы вводятся в диамагнитную решетку в качестве примесей, замещающих небольшую часть диамагнитных ионов. Кристаллическое поле оказывает существенное влияние на магнитные свойства парамагнитных ионов. Так как кристаллическое поле различно для разных направлений, оно ориентирует орбиты электронов таким образом, чтобы энергия иона в целом была минимальна. Стабилизация орбит оказывается настолько жесткой, “орбиты замораживаются”, что при переходах, соответствующих взаимодействию иона с магнитным полем, орбитальное движение практически не меняется. Следовательно, основной вклад в формирование энергетических уровней вносит спиновое взаимодействие. Орбитальное движение просто сдвигает все уровни энергии на одну величину. Энергия спинового взаимодействия атома ( $\epsilon$ ), если магнитное поле направлено вдоль оси  $z$ , выражается как

$$\epsilon = g\beta S_z H_z, \quad (7.20)$$

где  $S_z$  и  $H_z$  – проекции энергии на ось  $z$ .

Проекция спинового момента изменяется в пределах  $-S < S_z < S$   $S = |S|$ . Ориентация спинового момента квантована, т.е.  $S$  принимает только определенные значения, отличающиеся на единицу:

$$-S, -S+1, -S+2, \dots, 0, 1, 2, \dots, S-1, S;$$

Каждому значению спинового момента соответствует один энергетический уровень. Всего таких уровней будет  $2S+1$ . Например:  $S=3/2$ , получим ряд:  $-3/2, -1/2, +1/2, +3/2$ , - 4 уровня; если  $S=5/2$ , то будет ряд:  $-5/2, -3/2, -1/2, 1/2, 3/2, 5/2$  – 6 уровней.

Если внешнего поля нет ( $2S+1$ ), уровни сливаются в один. При введении атома в магнитное поле напряженности  $H$  уровни расщепляются.

Явление расщепления энергетических уровней в магнитном поле носит название эффекта Зеемана.

Допустим, мы имеем три электрона с суммарным спином  $S=3/2$ . Значения энергии атома в магнитном поле будут следующие, при  $S_z = -3/2, -1/2, +1/2, +3/2$  (рис. 7.8). Введенные на рисунке обозначения соответствуют следующим выражениям:

1)  $\epsilon_{3/2} = 3/2 g \beta H_z$ ; 2)  $\epsilon_{1/2} = 1/2 g \beta H_z$ ; 3)  $\epsilon_{-1/2} = -1/2 g \beta H_z$ ; 4)  $\epsilon_{-3/2} = -3/2 g \beta H_z$ ;  $\Delta \epsilon = g \beta H_z$ .

Кроме того, необходимо учесть следующее. Даже в отсутствие магнитного поля энергия различных спиновых состояний может оказаться различной. На парамагнитный ион, находящийся в кристаллической решетке, дополнительно действует внутрикристаллическое поле (электрическое – эффект Штарка), которое изменяет систему уровней. Уровень энергии с квантовым числом  $n$  расщепляется на несколько уровней. Положение их орбит несколько изменяется, что и приводит к изменению энергии иона в целом (рис. 7.9). Этот эффект не зависит от величины магнитного поля. Тогда выражение для энергии парамагнитного иона в кристаллической решетке примет вид для рубина:

$$\epsilon = g \beta S_z H \cos \theta + D[S_z^2 - 1/3 S(S+1)], \quad (7.21)$$

где  $D$  – коэффициент, зависящий как от величины кристаллического поля, так и от его симметрии,  $\theta$  – угол между направлением поля и кристаллографической осью.

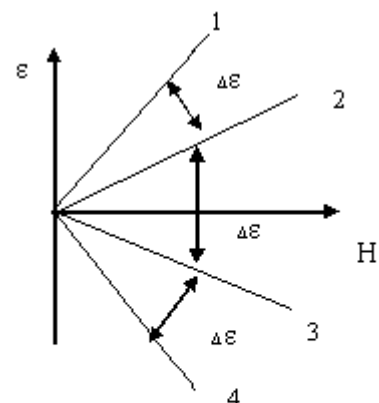


Рис. 7.8. Энергетические уровни иона хрома в магнитном поле

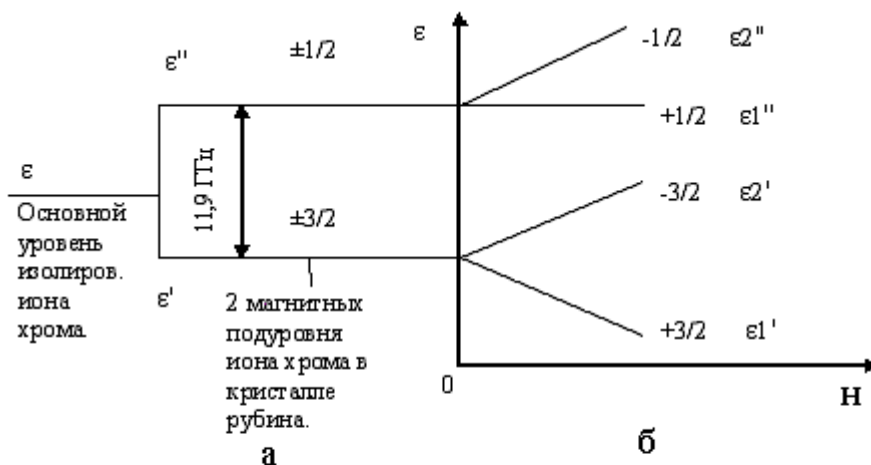


Рис. 7.9. Расщепление уровня иона хрома под действием внутрикристаллического электрического поля (а) и под действием внешнего магнитного поля (б).

При  $\theta = 0$ , значения энергетических уровней становятся равными

1)  $\epsilon_{-3/2} = -3/2 g \beta H + D[9/4 - 1/2(3/2+1)] = -3/2 g \beta H + D$ , 2)  $\epsilon_{-1/2} = -1/2 g \beta H - D$ , 3)  $\epsilon_{+1/2} = -1/2 g \beta H - D$ , 4)  $\epsilon_{+3/2} = +3/2 g \beta H + D$ .

Частота перехода между уровнями, полученными за счет расщепления электрическим полем, равна 11,9 ГГц. Во внешнем поле  $H$  каждый такой уровень расщепляется еще на два, причем величина этого расщепления висит как от величины поля, так и от угла  $\theta$  между направлением поля и кристаллографической осью кристалла. На рис. 7.10 показано распределение энергетических уровней для а)  $\theta = 0$ , б) для  $\theta = 54,44^\circ$ .

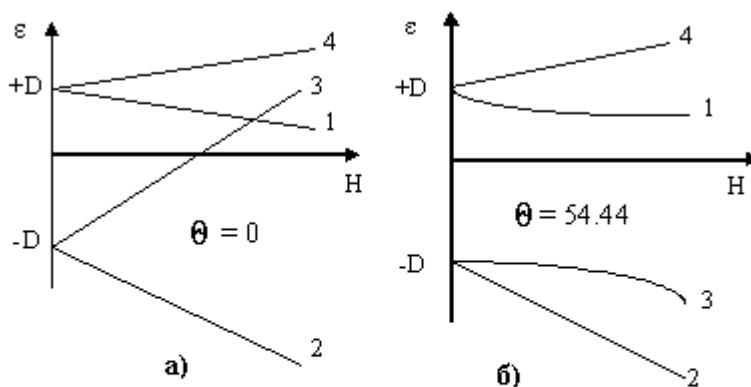


Рис. 7.10. Распределение энергетических уровней иона хрома для  $\theta = 0$  (а),  $\theta = 54,44^\circ$  (б).

за-  
ду

В квантовых усилителях наиболее часто используются режимы, соответствующие углам  $\theta = 54,44^\circ$  и  $\theta = 90^\circ$ .

В сантиметровом диапазоне в КПУ используется рубин.

В миллиметровом – рутил (окись  $\text{TiO}_2: \text{Cr}^{+3}$ , или  $\text{TiO}_2: \text{Fe}^{+3}$ ).

### 7.2.3. Электронный парамагнитный резонанс (ЭПР)

Явление ЭПР наблюдается при наложении на парамагнитное вещество двух взаимно перпендикулярных магнитных полей: слабого переменного  $H_1$  и сильного постоянного  $H_0$ . Переменное поле такой частоты, при которой наступает наиболее интенсивное поглощение энергии этого поля.

Резонансная частота, при которой происходит резкое селективное поглощение оказывается равной частоте Лармора:  $\nu_L = e/m_s H_0$  для сравнения  $\omega_0 = \gamma H_0$ .

Как было выяснено раньше, влияние постоянного поля  $H_0$  на парамагнетик вызывает его прецессию. Наложение перпендикулярной переменного поля  $H_1$  приводит к изменению угла этой прецессии, т.е. происходит изменение энергии взаимодействия.

1. Рассмотрим случай, когда вращение вектора поля  $H_1$  совпадает по направлению с прецессионным вращением вектора  $\mu_i$ , являющегося суммой орбитального и спинового магнитных моментов. При этом частота переменного поля отличается от частоты прецессии, будет происходить то ускорение, то замедление прецессии электрона, т.к. результирующее поле в каждый момент времени определяется мгновенным значением геометрической суммы двух полей  $H_0$  и  $H_1$ . Значение дополнительного вращательного момента будет равно нулю.

2. При совпадении частоты переменного поля с частотой прецессии спинов  $\omega = \omega_{\text{пр}}$  появится постоянно действующий на спины опрокидывающий момент. Действие этого момента выражается в возникновении дополнительных колебаний, накладываемых на основную прецессию спинов. При этом создаются усло-



вия, при которых резко возрастает внутренняя энергия атома за счет энергии взаимодействия между его спинами и внешним переменным полем.

3. Если вектор  $H_1$  вращается в обратную сторону по отношению к вращению магнитного момента  $\mu_i$ , вызванного прецессией, то результирующий эффект в среднем, независимо от соотношений частот поля и прецессии, тоже будет равен 0, как и при  $\omega$  не равное  $\omega_{пр}$  в первом случае.

Поле СВЧ может взаимодействовать со спинами или орбитальным магнитным моментом, если:

1.  $\nu = \nu_{пр}$

2. существует  $H_0 + H_1$

3. направление вращения вектора  $H_1$  должно совпадать с вращением  $\mu_i$ .

С квантовой точки зрения магнитный момент электрона равен одному магнитону Бора:  $\mu = h/2m_e$ .

Энергия взаимодействия с внешним полем  $H_0$  равна  $\epsilon_{вз} = -\mu H_0$

( $-\mu, +\mu$ ) (т.е. момент либо совпадает с полем  $H_0$ , либо ему противоположен), основной уровень расщепится на два (рис. 7.9):

$$\epsilon_{\max} = E + \mu H_0, \quad \epsilon_{\min} = E - \mu H_0. \quad (7.22)$$

Переходу между этими уровнями соответствует частота:

$$\omega_{21} = [E + \mu H_0 - E - \mu H_0] / h = 2 \mu H_0 / h \quad (7.23)$$

Таким образом, условию ЭПР отвечает равенство частоты внешнего поля  $H_1$  частоте перехода  $\omega_{21}$ , т.е.

$$\omega = \omega_{21} = 2 \mu H_0 / h = e H_0 / m_e = \gamma H_0. \quad (7.24)$$

Здесь наблюдается мгновенный переход частиц под воздействием энергии  $2H_0$  с одного уровня на другой с поглощением фотона при переходе “вверх” и излучением фотона при переходе “вниз”.

Из изложенного вытекает, почему в обычной среде ЭПР проявляется в виде поглощения энергии, а в активной среде с инверсной населенностью наблюдается усиление. Действительно, в обычной среде параллельных волчков больше, чем антипараллельных, и поэтому переходов вверх больше, чем число переходов вниз – среда в целом поглощает энергию. В состоянии инверсии – наоборот.

Такое распределение частиц по уровням обладает свойствами устойчивого равновесия; если каким-либо способом изменить это распределение, то через некоторое время оно снова вернется в равновесное состояние из-за спин – решеточной и спин – спиновой релаксации. При равновесном распределении, когда большинство частиц обладает минимальной энергией, поглощение электромагнитной энергии всегда должно преобладать над ее излучением, что и наблюдается в ЭПР. Наблюдение ЭПР осуществляется с помощью радиоспектроскопов или радиоспектрометров.

Изменяя величину постоянного поля  $H_0$ , мы изменяем положение энергетических уровней, когда расстояние между уровнями равно энергии кванта,  $\Delta\epsilon = h\nu$ , то происходит резонансное поглощение СВЧ поля.

Если квантовая система является многоуровневой рис. 7.11, то резонансное поглощение ВЧ может иметь место при различных значениях постоянного магнитного поля  $H_0$ .

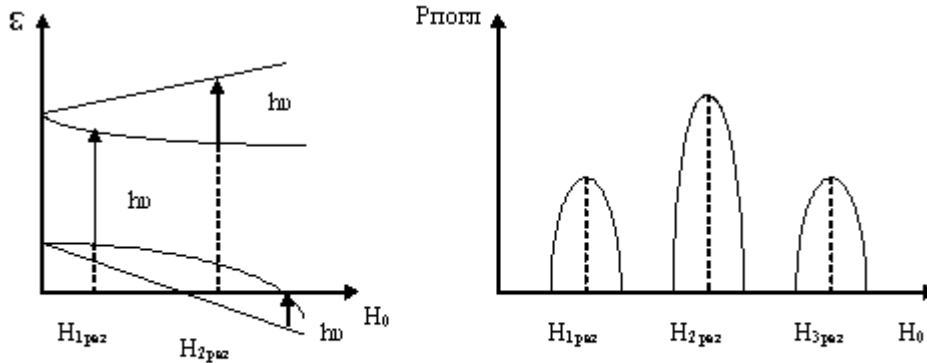


Рис. 7.11. Резонансное поглощение в многоуровневых квантовых системах.

В КПУ широко используются трех- и четырехуровневые схемы.

Рассмотрим 3-х уровневые схемы. На рис. 7.12 показан один из вариантов трехуровневой схемы, где частота накачки определена переходами с первого на третий уровень ( $\nu_H = (\epsilon_3 - \epsilon_1)/h$ ), частота сигнала со второго на первый ( $\nu_{сиг} = \nu_{21} = (\epsilon_2 - \epsilon_1)/h$ ).

Получение коэффициента усиления достигается путем инвертированности среды. Коэффициент инверсии ( $I_{21}$ ) определяется из соотношения разности населенностей между рабочими уровнями в возбужденном состоянии по отношению к разности населенностей этих же уровней, но находящихся в невозбужденном состоянии

$$I_{21} = (N_2 - N_1) / (N_2^e - N_1^e). \tag{7.25}$$

Из (7.25) видно, что необходимо знать величины разности населенной между уровнями. При ее расчете используют уравнения баланса. Спонтанными переходами можно пренебречь, оставляются кинетические уравнения на основе переходов, показанных на рис. 7.12: достаточно взять изменения населенности во времени на втором уровне, т.к. для определения второго неизвестного, возьмем сумму всех частиц системы, распределенных по уровням.

$$\begin{aligned} dN_2/dt &= \Gamma_{12} \cdot N_1 + \Gamma_{32} \cdot N_3 - (\Gamma_{21} + \Gamma_{23}) \cdot N_2 = 0, \\ N_1 + N_2 + N_3 &= N, \end{aligned}$$

где  $dN_2/dt$  – стационарный (установившийся) режим, остальные параметры показаны на рисунке. Решая эти уравнения, легко находим значения  $N_1$  и  $N_2$ .

Недостатками трехуровневых схем является то, что в создании инверсии участвуют три (рис. 7.12) уровня. Если генерация осуществляется между 3 и 2 уровнем ( $\nu_{32}$ ), то в этом случае уровень, на котором частицы

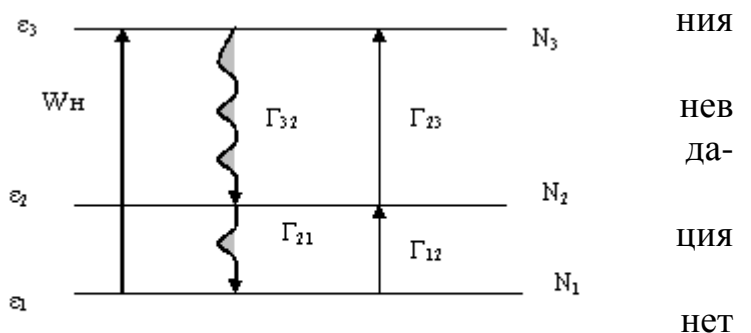


Рис. 7.12. Трехуровневая квантовая система

должны накапливаться. Если генерация соответствует частоте  $\nu_{21}$ , то здесь нижний рабочий уровень является основным. Инверсия в таком случае достигается с трудом.

Позже будут отмечены и другие недостатки.

Четырехуровневые схемы представляют больший интерес. Но и по системам накачки они тоже имеют разновидности рис. 7.13, а, б.

а) Система уровней симметрична, одновременно происходит накачка переходов  $1 \rightarrow 3$ ,  $2 \rightarrow 4$ . Здесь на втором уровне уменьшается число частиц, на третьем увеличивается и рабочими уровнями являются 2 и 3.

б) На этой схеме инверсия достигается за счет первоначально перехода частиц с первого на четвертый, а затем уже с четвертого на третий.

Коэффициент усиления в 4-х уровневых схемах  $>$ , чем в 3-

Недостаток: не всегда можно получить условие симметричности уровней.

**Коэффициент инверсии.**

Основным параметром, определяющим мощность излучения и соответственно генерации, является коэффициент инверсии. Запишем формулу, определяющую мощность излучения:

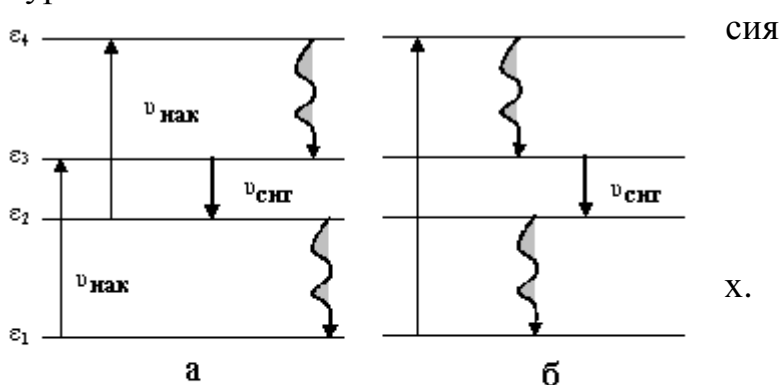


Рис. 7.13. Четырехуровневые схемы а- система уровней симметрична с накачкой между переходами 1-3 и 2-4, б- система уровней симметрична с накачкой между переходом 1-3.

$$P_{\text{изл}} = h\nu W(N_2 - N_1); \quad (7.26)$$

Чтобы добиться большей мощности излучения необходимо добиться большей инверсии населенностей. На практике величинами  $N$  неудобно пользоваться, поэтому вводят понятие коэффициента инверсии:

$$I_{21} = (N_2 - N_1) / (N_1^e - N_2^e); \quad (7.27)$$

При вычислении мощности излучения для термостатированного состояния, учитывается коэффициент инверсии:

$$P_{\text{изл}} = h\nu W(N_2^e - N_1^e) I_{21}$$

1. В состоянии термодинамического равновесия коэффициент инверсии равен  $-1$ , что легко доказать:

$$I_{21} = (N_2^e - N_1^e) / (N_1^e - N_2^e) = -(N_1^e - N_2^e) / (N_1^e - N_2^e) = -1;$$

2. В состоянии с инверсией:

$$I_{21} > 0;$$

Чем лучше качество активного элемента, тем больше  $I_{21}$ . В настоящее время используют активные элементы с коэффициентом инверсии порядка нескольких единиц.

В состоянии термодинамического равновесия по закону Больцмана:

$$N_2^e / N_1^e = \exp(-h\nu_{21}/kT).$$

По аналогии с законом Больцмана формально можно записать:

$$N_2 / N_1 = \exp(-h\nu_{21}/kT_{\text{э}}), \quad (7.28)$$

где  $T_{\text{э}}$  – эффективная спиновая температура.

Если  $N_2 > N_1$ , то  $T_{\text{э}} < 0$  (состояние инверсной населенности), т.е.  $T_{\text{э}} = -|T_{\text{э}}|$ , тогда выражение (7.28) можно переписать следующим образом:

$$N_2 / N_1 = \exp(h\nu_{21}/k|T_{\text{э}}|).$$

В диапазоне СВЧ  $h\nu$  всегда  $< kT$  и следовательно  $< k|T_{\text{э}}|$ . Тогда, используя разложение экспоненты в ряд и ограничиваясь двумя членами разложения:

$$n_2^e/n_1^e = 1 - h\nu/kT; \quad n_2/n_1 = 1 + h\nu/k|T_{\text{э}}|; \quad (7.29)$$

Для определения коэффициента инверсии, подставим (7.29) в (7.27):

$$I_{21} \approx -T/T_{\text{э}}. \quad (7.30)$$

Качество усилителя будет тем выше, чем ближе  $|T_{\text{э}}|$  к нулю.

Таким образом, коэффициент инверсии и эффективная спиновая температура  $T_{\text{э}}$  тесно связаны и любая может быть использована для характеристики усилителя.

Инверсия населенности для трехуровневых схем достигается на переходах, на которых выполняется условие:

$$\nu_{\text{н}}/2 > \nu_{\text{с}}; \quad (7.31)$$

Аналогично проводится расчет для 4-х уровней схемы. В результате получим выражение для коэффициента инверсии:

$$I = (\nu_{\text{н}}/\nu_{\text{с}}) - 1; \quad (7.32)$$

**Пример:** Рассмотрим случай  $\nu_{\text{н}} = 3\nu_{\text{с}}$ : Для 3-х уровней  $I_{21} = 0.5$ ;

Для 4-х уровней  $I = 2$ ;

Для 4-х уровневой схемы коэффициент усиления в 4 раза больше, чем для 3-х уровневой схемы.

Изучение ЭПР осуществляется с помощью спектрометров ЭПР, которые по устройству близки к квантовым парамагнитным усилителям (КПУ), поэтому более подробно остановимся на рассмотрении КПУ.

#### 7.2.4. Устройство квантового парамагнитного усилителя

На рис. 7.14 показана схема квантового парамагнитного усилителя с отражательным резонатором. Сигналы, распространяющиеся в противоположных направлениях, разделяют при помощи циркулятора – устройства, пропускающего электромагнитные волны между двумя соседними плечами только в одном направлении. Входной сигнал попадает в циркулятор, а из него в резонатор с парамагнитным кристаллом.

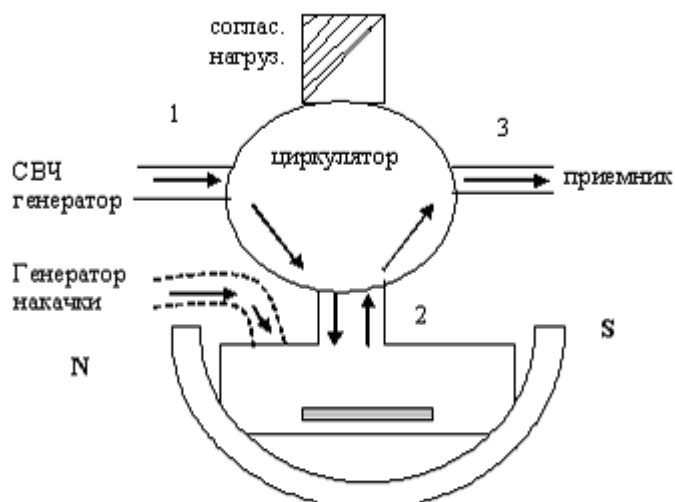


Рис. 7. 14. Схема квантового усилителя

В резонаторе происходит усиление, и усиленный отраженный сигнал через циркулятор поступает на приемник сигнала. Шумы, излучаемые входом приемника, а также отраженная от него мощность при плохом согласовании приемника и источника сигнала могут попасть через циркулятор на вход резонатора. Для устранения этого нежелательного явления вслед за приемником к циркулятору подключают согласованную нагрузку. Нагрузку следует охлаждать, поскольку ее шумы поступают на вход усилителя.

Специфическим является требование 2-х частотности резонатора, т.е. способности его резонировать на частоте сигнала  $\nu_c$  и накачки  $\nu_n$ . Парамагнитный кристалл помещают в пучности магнитных полей сигнала и накачки с учетом требуемой ориентации оси кристалла и постоянного магнитного поля.

Генератор накачки насытит какой-либо переход (1-3). Тогда на переходе (3-2) или (2-1) будет инверсия населенности.

Подавая в резонатор слабый сигнал с  $h\nu_{32}$  или  $h\nu_{21}$ , будем иметь его усиление.

Можно использовать схемы и с проходными резонаторами. Особенность такого усилителя в наличии отдельных передающих линий для входного и выходного сигналов. Для прохождения усиливаемого сигнала входного и выходного сигналов. Для прохождения усиливаемого сигнала только в одном направлении применяются ферритовые вентили.

Применение многорезонаторных схем улучшает некоторые из параметров парамагнитных усилителей. Пассивные резонаторы изменяют форму частотной характеристики усилителя так, чтобы сделать ее максимально плоской, т.е. наиболее близкой к оптимальной прямоугольной форме. Так, например, при усилении 10дБ добавление 1-го пассивного резонатора позволяет более чем вдвое расширить полосу усилителя. Так, в диапазоне см. волн 2-х резонаторного усилителя на рубине при  $K_0 = 100$  имеет полосу пропускания  $\Delta\nu \sim 40\text{МГц}$ , при ширине спектральной линии  $\Delta\nu_{\text{макс}} \sim 60\text{МГц}$ .

Парамагнитные усилители бегущей волны применяются, когда необходимо увеличение усиления. В них усиление происходит за счет применения замедляющих систем. Замедляющие системы замедляют групповую скорость  $V_{\text{гр}}$ , обеспечивая при этом длительное взаимодействие и наращивание коэффициента усиле-

ния. Замедляющую систему помещают в волновод накачки или объемный резонатор с резонансной частотой равной частоте накачки.

### 7.2.5. Особенности квантового парамагнитного усилителя и области применения

КПУ СВЧ реализованы в большом диапазоне частот 300 МГц ( $\lambda=1\text{м}$ ) до 9600 МГц ( $\lambda=3,1\text{м}$ ). С их помощью получено значительное увеличение чувствительности радиолокаторов, систем глобальной и космической связи.

Основное характерное достоинство КПУ – низкий уровень собственных шумов. В этом смысле у КПУ единственный конкурент – п/п параметрические усилители, имеющие при охлаждении  $T_{\text{ш}}=20\text{-}30^\circ\text{ К}$ , но являющиеся менее стабильными. В КПУ  $T_{\text{ш}}=5\text{-}10^\circ\text{ К}$ .

Для сравнения у ЛБВ  $T_{\text{ш}}=1000^\circ\text{ К}$ .

КПУ применяются: наземная радиолокация; лабораторные измерения; космическая связь; радиолокация планет; радиоастрономия.

Особенно важно применение КПУ в радиоастрономии. КПУ без увеличения антенн повышают чувствительность приемного устройства во много раз.

## ГЛАВА 8.

### ОБЩИЕ ВОПРОСЫ ПОСТРОЕНИЯ ОПТИЧЕСКИХ КВАНТОВЫХ ГЕНЕРАТОРОВ И УСИЛИТЕЛЕЙ

Освоение новых диапазонов более высоких частот электромагнитного спектра является постоянно действующим закономерным процессом в резонансных, вычислительных, информационных, радиоэлектронных системах, таких как системы связи, локации, навигации, бытовая радиоэлектроника и др. Уже десятилетия ведутся активные научные и опытно-конструкторские работы в области создания новых систем в оптическом диапазоне. Имеются немалые достижения в этой области, однако значительных результатов предстоит достичь в ближайшем будущем. В настоящее время оптический диапазон широко используется на практике для такого рода систем. Отсюда следует, что современные радиоинженеры должны знать основы радиоэлектроники оптического диапазона.

#### 8.1. Диапазон оптических волн и его особенности

**Диапазон оптических волн.** Оптическими волнами называют электромагнитные волны, длина которых лежит в пределах от 0,01 до 1000 мкм. Этим волнам соответствуют колебания с частотами от  $3 \cdot 10^{16}$  до  $3 \cdot 10^{11}$  Гц.

$$\nu[\text{Гц}] = \frac{3 \cdot 10^{14}}{\lambda} [\text{мкм}]^{-1}$$

По физическим свойствам и техническим особенностям полученных электромагнитных волн весь этот спектр частот не является однородным. Принято делить весь оптический спектр на поддиапазоны, в которых указанные свойства более однородны (см. таблицу 8.1).

Таблица 8.1

<i>Наименование поддиапазонов</i>	$\lambda$ , мкм	$\nu$ , Гц	Принятое волновое число для спектроскопа $K$ , $\text{см}^{-1}$
<i>Инфракрасные, дальние волны</i>	10 – 1000	$3 \cdot 10^{13}$ – $3 \cdot 10^{11}$	$10^3$ – 10
<i>Инфракрасные, ближние волны</i>	0,75 – 10	$4 \cdot 10^{14}$ – $3 \cdot 10^{13}$	$1,33 \cdot 10^4$ – $10^3$
<i>Видимые волны</i>	0,4 – 0,75	$7,5 \cdot 10^{14}$ – $4 \cdot 10^{14}$	$2,5 \cdot 10^4$ – $1,33 \cdot 10^4$
<i>Ультрафиолет, ближний</i>	0,1 – 0,4	$3 \cdot 10^{15}$ – $7,5 \cdot 10^{14}$	$10^5$ – $2,5 \cdot 10^4$
<i>Ультрафиолет, дальний</i>	0,01 – 0,1	$3,16 \cdot 10^{16}$ – $3 \cdot 10^{15}$	$10^6$ – $10^5$

Освоение оптического диапазона обусловлено, в первую очередь, ростом экономических и культурных потребностей общества, а так же требованиями обороны.

Следует отметить, что одним из замечательных явлений в истории наук XX века было возникновение квантовой электроники – нового раздела физики и тех-

ники. Наука, изучающая усиление, генерацию и преобразование электромагнитных волн.

В настоящее время на базе квантовой электроники формируется оптическая радиоэлектроника как область техники. Оптическая радиоэлектроника, аналогично технике радиодиапазона, СВЧ включает в себя описание закономерностей распространения оптических волн, описание законов, устройств излучения и приема этих волн, анализ процессов генерации, модуляции, усиления, детектирования и т.д.

Всякая система радио-, СВЧ диапазона включает в себя электронные приборы (лампы, полупроводниковые приборы и др.) и пассивные электрические линейные элементы цепи (резисторы, конденсаторы, колебательные контуры разного вида и др.).

В оптическом диапазоне волн радиосистема также имеет квантовые электронные приборы (лазеры, мазеры, ФЭУ и др.) и линейные устройства в виде оптических линз, зеркал, телескопических систем, оптических фильтров, модуляторов и др.

**Особенности оптического диапазона волн.** До появления лазеров или оптических квантовых генераторов в оптическом диапазоне волн отсутствовали источники когерентного монохроматического излучения. Принципиально нет никакой разницы между квантовыми приборами СВЧ и оптического диапазона. Для получения инверсии необходимо, чтобы число частиц на верхнем уровне было больше, чем на нижнем. Однако в оптическом диапазоне имеется ряд особенностей, которые существенно влияют на конструкции лазеров и их выходные характеристики.

Первой особенностью квантовых генераторов оптического диапазона является то, что у веществ используются возбуждённые уровни, разделённые друг от друга и от уровня основного состояния оптическими переходами.

По закону Больцмана практически населённым является только уровень основного состояния. Населённости верхних возбуждённых уровней будут практически равны нулю.

Второе отличие состоит в способе накачки. Несмотря на то, что в настоящее время появились достаточно мощные монохроматические источники накачки (полупроводниковые лазерные диоды), накачка производится либо немонахроматическими источниками, либо применяются другие методы возбуждения.

Третье отличие состоит в том, что резонаторы оптического диапазона отличаются от обычных объёмных резонаторов СВЧ. В оптическом диапазоне используют открытые резонаторы.

Наконец, в оптическом диапазоне нельзя не учитывать спонтанные переходы, т.к. вероятность спонтанных переходов пропорциональна кубу частоты.

Тепловые (безызлучательные) переходы, как правило, для рабочих уровней можно не учитывать, но для холостых переходов их вероятности могут быть очень велики [3,5].



## 8.2. Классификация используемых в оптико-электронных квантовых приборах лазеров

Оптические квантовые генераторы по различным признакам могут быть разделены на определенные группы. По агрегатному состоянию вещества различают:

- газовые лазеры;
- лазеры на твердом теле;
- полупроводниковые оптические квантовые генераторы (п/п лазеры);
- жидкостные лазеры.

Для накачки и получения индуцированного излучения в газовых лазерах могут использоваться энергетические уровни атомов, молекул и ионов. В соответствии с этим газовые лазеры подразделяют на атомарные, молекулярные и ионные.

Атомарные лазеры следует рассматривать как источники главным образом инфракрасного (ИК) излучения небольшой мощности. Их спектр содержит небольшое число линий излучения в видимой и ближней инфракрасной областях. Основное число линий сосредоточено в интервале 1 – 25 мкм.

Ионные оптические квантовые генераторы излучают в основном видимые и ультрафиолетовые электромагнитные колебания, а молекулярные лазеры являются источниками мощного ИК излучения с длинами волн, превышающими 5 мкм.

Если в атомарных и молекулярных лазерах «рабочими» излучателями являются электрически нейтральные частицы (атомы и молекулы), то в ионных – заряженные частицы (ионы).

По типу резонансной системы лазеры подразделяются на линейные и кольцевые. Кольцевые оптические квантовые генераторы называют также оптическими квантовыми генераторами бегущей волны. В кольцевых лазерах луч света распространяется по замкнутому контуру В зависимости от режима работы различаются лазеры непрерывного излучения и импульсные.

Исторически сложилось так, что первые образцы лазеров на твердом теле и полупроводниковые оптические квантовые генераторы работали только в импульсном режиме, а первые газовые лазеры – лишь в непрерывном. В настоящее время от генераторов всех классов получен как импульсный, так и непрерывный режим работы.

В зависимости от частоты излучения различают лазеры инфракрасного, видимого и ультрафиолетового диапазонов волн.

По положению двух энергетических уровней, между которыми совершается рабочий переход атомов активного вещества, лазер можно подразделять на трехуровневые и четырехуровневые (см. рис.6.4, 6.5).

Деление оптических квантовых генераторов на трех- и четырехуровневые весьма условно, так как конечный уровень «2» может быть незначительно удален от основного. Принято считать лазер четырехуровневым, если  $E_2 \geq 8kT$  [2].

## 8.3. Структурная схема оптико-электронного квантового прибора

На рис.8.1. представлена обобщенная структурная схема, являющаяся основной практически любого оптико –электронного квантового прибора (ОЭКП) [9]. Краткое рассмотрение ее во введении позволяет показать назначение основных и вспомогательных элементов ОЭКП и систематизировать их по

функциональному назначению. Назначение системы накачки 2, получающей энергию от источника питания 1, – обеспечить возбуждение рабочей среды (лазера) 3, создать в ней условия для поддержания неравновесного состояния с инверсией населенностей. Оптический резонатор, представляющий собой систему минимум из двух параллельных друг другу отражателей 4 и 5, является частотно и пространственно-селективным устройством положительной обратной связи, а также в значительной степени определяет и энергетические характеристики лазера. Предельно достижимые частотные, временные и пространственные характеристики излучения получаются в ОЭКП с помощью помещенных внутрь резонатора селектора мод 6 и модулятора 7, а также с помощью внешней по отношению к резонатору оптической системы формирования излучения, состоящей из внешнего селектора мод 8, внешнего модулятора (временных характеристик излучения) 9, устройства управления излучением в пространстве 10 (устройства отклонения луча). Для стабилизации параметров излучения (мощности излучения  $P_{\text{вых}}$ , его частоты  $\nu_{\text{ген}}$ , положения оси лазерного луча в пространстве, а также систем автоматической юстировки отражателей резонатора и т.п.) служит устройство 11, функционально связанное с элементами резонатора и источника накачки 2. Для ОЭКП связного, локационного назначения, для технологических ОЭКП необходима взаимная связь между объектом исследований, передающей частью ОЭКП и его информационной частью. Эту связь, чаще всего непосредственно через излучение (прошедшее через объект, отраженное от него и принятое приемной частью ОЭКП 13), осуществляет устройство обработки информации 14 и решающее устройство 15, в котором на основе полученной информации принимается решение об изменении режима работы лазера, об осуществлении требуемого закона сканирования или его изменения, а также о дозировании во времени энергии (мощности), падающей на объект. Этот блок непосредственно связан с элементами передающей части ОЭКП и управляет их работой. При необходимости устройства 14 и 15 снабжаются блоком памяти 16, где хранится информация о законах управления всем устройством и возможном ее изменении в зависимости от поступления в

процессе функционирования новой информации.

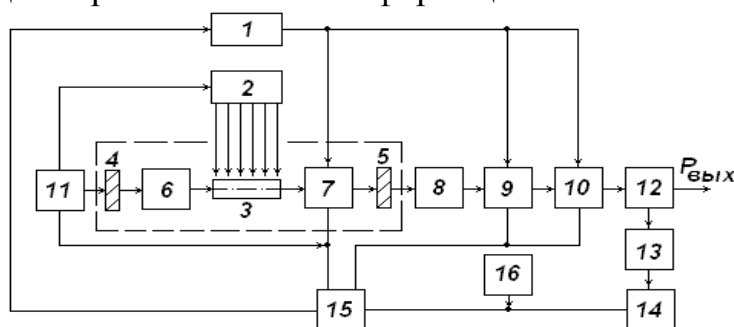


Рис. 8.1.Обобщенная структурная схема ОЭКП

Представленная схема не претендует на исчерпывающую полноту. Возможны ее дальнейшая детализация и дополнение. Однако рассмотренные элементы структурной схемы являются в значительной степени обязательными для большинства ОЭКП независимо от тех задач, которые они призваны решать [9].

Как следует из рис. 8.1 важнейшим элементом лазера как и квантового генератора радиодиапазона является резонатор.

Прежде чем перейти к рассмотрению отдельных типов лазеров и их характеристик, выделим резонаторы в отдельный раздел.

#### 8.4. Оптические резонаторы

**Особенности оптических резонаторов.** Как и в диапазоне СВЧ, в оптическом диапазоне наиболее эффективное взаимодействие электромагнитного поля с веществом осуществляется при помещении его внутрь резонансной системы.

Однако между резонаторами СВЧ и оптическими существует много принципиальных различий. Так как добротность любого резонатора пропорциональна размерам и корню квадратному из длины волны, то на оптических частотах в принципе нельзя создать резонансные системы, у которых размеры сравнимы с длиной волны. Кроме того, такую систему трудно изготовить, да и количества активного вещества в ней будет недостаточно для усиления или генерации.

Следовательно, оптический резонатор должен иметь размеры много больше длины волны излучения, то есть в резонаторе возможно существование многих типов колебаний (мод). Пусть, резонатор представляет собой куб с ребром  $L$  и объемом  $V = L^3$  с идеально отражающими стенками. Если эта полость заполнена электромагнитным излучением, то в ней может существовать бесконечно большое число стоячих электромагнитных волн, характеризующихся волновыми векторами  $\vec{k}$  с проекциями

$$k_x = \frac{2\pi l}{L}, \quad k_y = \frac{2\pi m}{L}, \quad k_z = \frac{2\pi n}{L},$$

где  $l, m, n$  – целые числа, указывающие число полуволн, укладываемых вдоль соответствующей оси резонатора.

В пространстве этих волновых векторов на каждый тип колебаний приходится объем

$$\frac{dk_x dk_y dk_z}{dldmdn} = \left( \frac{2\pi}{L} \right)^3.$$

Общее число типов колебаний  $N_k$ , приходящихся на объем, в котором вектор  $\vec{k}$  может изменяться от 0 до  $\vec{k}_0$ , равно общему объему сферы радиусом  $R$ , деленному на объем, приходящийся на один тип колебаний

$$N_k = 2 \frac{\left( \frac{4\pi}{3} \right) k^3}{\left( \frac{2\pi}{L} \right)^3} = \frac{k^3 L^3}{3\pi^2} = \frac{8\pi\nu^3}{3c^3} L^3. \quad (8.1)$$

Здесь множитель 2 говорит о том, что каждому значению  $\vec{k}$  соответствует два возможных значения поляризации.  $N_k$  – число типов колебаний, приходящихся на интервал частот  $\Delta\nu$ . Плотность типов колебаний на единичный интервал частот и на единицу объема  $V = L^3$

$$N(\omega) = \frac{1}{V} \frac{dN_k}{d\nu} = \frac{8\pi\nu^2}{c^3}. \quad (8.2)$$

Число колебаний в объеме  $V$  и в интервале частот  $\Delta\nu$

$$\Delta N = \frac{8\pi\nu^2}{c^3} V \Delta\nu. \quad (8.3)$$

Видно, что с ростом  $\nu$  в заданном частотном интервале  $\Delta\nu$ , расстояние между модами уменьшается. При  $V = 1\text{см}^2$ ,  $\lambda = 1\text{мкм}$ ,  $\Delta\nu = 1\text{см}^{-1} = 30\text{Гц}$ ,

$$\Delta N = 10^9, \frac{\Delta\nu}{\Delta N} = 30\text{Гц}.$$

Если еще учесть, что с ростом  $\nu$  падает добротность резонатора для каждой моды, то ясно, что резонансные кривые мод начнут перекрываться и резонатор потеряет свои свойства. Проведенный анализ убедительно показывает, что в оптическом диапазоне в качестве резонансных систем нельзя использовать замкнутые полости с отражающими стенками, подобные СВЧ резонаторам. Для успешного применения резонансных систем в оптическом диапазоне необходимо найти пути разрежения их спектра [4,15].

**Открытые резонаторы и их элементарная теория.** В настоящее время наибольшее распространение в качестве резонансных систем в оптическом диапазоне получили открытые резонаторы, впервые предложенные нашим ученым А.М. Прохоровым и американским ученым Диком. Открытый резонатор представляет собой два зеркала, размещенных на некотором расстоянии  $L$  друг от друга (см. рис. 8.2.).

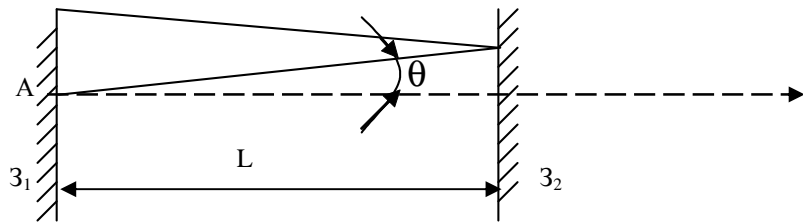


Рис. 8.2. Открытый резонатор Фабри - Перо

Между зеркалами ( $Z_1$  и  $Z_2$ ) помещается активное вещество. Зеркала могут быть либо плоскими, и в этом случае резонатор носит название Фабри - Перо, либо сферическими.

Разрежение спектра в таком резонаторе возможно потому, что типы колебаний с большими значениями  $k_x$  и  $k_y$  будут представлять собой суперпозицию волн, распространяющихся под значительными углами  $\theta$  к оси резонатора. Такие типы колебаний будут затухать значительно сильнее тех, для которых  $\theta$  близко к нулю, так как для них возрастают потери на излучение через боковую поверхность резонатора, если зеркала имеют конечные размеры. Поэтому в данном резонаторе будет существовать бесконечный набор типов колебаний со всеми возможными значениями  $k_x$  и  $k_y$ .

Хотя строгий расчет характеристик таких резонаторов проводится с помощью теории электромагнитного поля, представление о некоторых основных характеристиках открытых резонаторов можно получить, применяя законы геометрической оптики. Рассмотрим для примера элементарную теорию резонатора Фабри - Перо.

В приближении геометрической оптики зеркала будем считать безграничными в поперечных размерах, а распространяющиеся в них волны – плоскими. Тогда, каждое собственное колебание резонатора будет образовываться в результате сложения плоских волн, движущихся в противоположных направлениях между отражающими зеркалами. Сложения волн, распространяющихся вдоль оси  $z$  резонатора, с учетом граничных условий на зеркалах приводит к стоячим волнам, которые обычно называют продольными модами резонатора. Резонансные частоты продольных мод определяются граничными условиями на зеркалах и тем условием, что за время возвращения луча в точку  $A$  (рис.8.2), после двукратного отражения от зеркал резонатора, фаза поля может измениться на число, кратное  $2\pi$

$$\omega_q t_{2np} = 2\pi q, \quad (8.4)$$

где  $q=1,2,3\dots$

Из (8.4) можно найти выражение для резонансных частот резонатора  $\omega_q$  и расстояние между ними  $\Delta\omega$ , учитывая, что время двойного прохода луча в резонаторе  $t_{2np} = \frac{2L}{c}$  ( $c$  – скорость света)  $\omega_q = q \frac{\pi}{L} c$ ,  $\Delta\omega = \omega_q - \omega_{q-1} = \pi \frac{c}{L}$ . Видно, что индекс  $q$  характеризует число полуволн, укладываемых вдоль оси резонатора. Так как длина резонатора имеет обычно порядок  $5 \div 100$  см, а  $\omega$  в оптическом диапазоне имеет порядок  $10^{15}$  рад/с., то число  $q$  будет очень большим ( $10^5 - 10^6$ ).

Кроме продольных типов колебаний, существуют также поперечные колебания, образованные плоскими волнами, распространяющимися под некоторым углом  $\theta$  к оси резонатора.

Собственные частоты этих колебаний определяются формулой:  $\omega_q = \frac{q \cdot \pi \cdot c}{L \cdot \cos \theta}$

причем,  $\theta$  может принимать любые непрерывные значения. Видно, что колебания в данном случае будут сильно вырождены, то есть одной и той же частоте колебаний будет соответствовать множество типов колебаний, различающимся между собой углом  $\theta$  [13].

**Добротность резонаторов.** Одной из основных характеристик резонатора является его добротность. Здесь, как и в обычных резонаторах, потери энергии в резонаторе делятся на потери за счет связи с нагрузкой и потери, обусловленные неидеальностью резонатора.

Для вывода световой энергии из резонатора одно или оба зеркала делаются частично пропускающими. Если коэффициент отражения зеркал обозначить через  $r$ , то уменьшение энергии в резонаторе за один проход будет равно:

$dW = -W \cdot (1-r)$ , а за время  $dt$

$$dW = -W \cdot (1-r) \cdot \frac{c}{n \cdot L} \cdot dt, \quad (8.5)$$

где  $n$  – положительное целое число.

Поскольку затухание энергии в резонаторе определяется через добротность

$$W = W_0 \cdot e^{-\frac{\omega}{Q} t}, \quad (8.6)$$

то

$$dW = -W \cdot \frac{\omega}{Q} \cdot dt. \quad (8.7)$$

Сравнивая (8.5) и (8.7), получим

$$Q = \frac{L \cdot n \cdot \omega}{c \cdot (1-r)} = \frac{2 \cdot \pi \cdot L \cdot n}{\lambda \cdot (1-r)}. \quad (8.8)$$

Например, для  $r = 0,5$ ,  $\lambda = 1$  мкм,  $L = 100$  см,  $Q = 10^7$ .

Среди внутренних потерь в резонаторе отметим: дифракционные явления, обусловленные конечными размерами отражателей, непараллельность зеркал, их шероховатость и т.д.

Оценим сначала влияние дифракции. Луч света, выходящий из резонатора, будет отклоняться за счет дифракции на угол  $\alpha_d \sim \lambda/D$ . Вероятность того, что луч света выйдет из резонатора, огибая зеркало, будет равна отношению линейного отклонения к поперечному размеру зеркала  $\frac{\lambda \cdot L}{2 \cdot D^2}$ .

Эту величину следует сравнить с коэффициентом пропускания зеркала  $(1-r)$ , тогда выражение для добротности примет вид:

$$Q = \frac{L \cdot n \cdot \omega}{c \cdot \left(1-r + \frac{c \cdot \pi \cdot L}{n \cdot \omega \cdot D^2}\right)} = \frac{2 \cdot \pi \cdot L}{\lambda \cdot \left(1-r + \frac{\lambda \cdot L}{2 \cdot D^2}\right)}. \quad (8.9)$$

При  $\lambda = 1$  мкм,  $D = 1$  см,  $L = 100$  см,  $\frac{\lambda \cdot L}{2 \cdot D^2} = 0,5 \cdot 10^{-2}$ , то есть в большинстве реальных случаев дифракционные потери вносят в добротность ничтожный вклад и ими можно пренебречь.

Оценим теперь влияние на добротность непараллельности зеркал (рис. 8.3.)

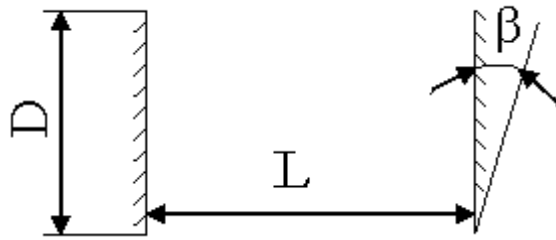


Рис. 8.3. Резонатор Фабри – Перо с квадратными зеркалами

а  $2\beta$ . Число отражений за время  $t$  есть  $\frac{c \cdot t}{L \cdot n}$ . Поэтому, время  $t$ , за которое луч переместится на Пусть правое зеркало повернуто на угол  $\beta$ . При отражении от непараллельных зеркал квант отклонится на угол равный  $2\beta$ . При каждом отражении угол отклонения увеличивается расстояние  $D$ , то есть уйдёт из резонатора, определится из условия:  $D = \beta \cdot L \cdot \left(\frac{c \cdot t_1}{2 \cdot L \cdot n}\right)^2 = \beta \cdot \frac{c^2 \cdot t_1^2}{2 \cdot L \cdot n^2}$  и будет равно

$$t_1 = \sqrt{\frac{2 \cdot D \cdot L \cdot n^2}{\beta \cdot c^2}}.$$

Следовательно,

$$Q = \omega \cdot t_1 = \frac{\omega \cdot n}{c} \cdot \sqrt{\frac{2 \cdot L \cdot D}{\beta}}. \quad (8.10)$$

Сравнивая (8.10) и (8.8), можно получить условие, определяющее точность, с которой необходимо выдерживать параллельность зеркал

$$\frac{L \cdot n \cdot \omega}{c \cdot (1-r)} < \frac{\omega \cdot n}{c} \cdot \sqrt{\frac{2 \cdot L \cdot D}{\beta}}, \quad (8.11)$$

Откуда 
$$\beta < \frac{2 \cdot D \cdot (1-r)^2}{L}.$$

Для реальных резонаторов лазера допустимый перекося составляет единицы и десятки угловых секунд.

**Волновая теория открытых резонаторов.** Строгое исследование электромагнитного поля в открытом резонаторе основывается на решении уравнений Максвелла с определенными граничными условиями. Однако для большинства практических случаев все размеры резонатора: расстояние между зеркалами, радиус кривизны, апертура – много больше длины волны излучения и поле очень близко к поперечному электромагнитному полю. Тогда для описания резонатора можно воспользоваться принципом Гюйгенса.

Метод заключается в том, что берется какое-то произвольное распределение поля на поверхности одного зеркала, а затем, используя принцип Гюйгенса, вычисляется распределение поля на поверхности другого зеркала. Полученное распределение принимают за исходное и вычисление повторяется. Оказалось, что после определенного числа отражений, распределение амплитуд и фаз поля на поверхности зеркал больше не изменяется, становится стационарным. Такое распределение соответствует нормальным типам колебаний резонатора. Предположим, что у одного из зеркал резонатора возбуждена начальная волна с распределением  $U_p$  и что она может распространяться в резонаторе в прямом направлении и обратно, попеременно отражаясь от его зеркал. После  $q$  проходов поле у одного зеркала  $U_q$ , связанное с полем, отраженным у другого зеркала, можно записать согласно принципу Гюйгенса:

$$U_{q+1} = \frac{j \cdot k}{4\pi} \cdot \oint U_q \cdot \frac{e^{-ikr}}{r} \cdot (1 + \cos \theta) ds, \quad (8.12)$$

где  $k$  – постоянная распространения в среде,

$r$  – расстояние от произвольной точки первого зеркала до произвольной точки второго зеркала,

$\theta$  – угол, который  $r$  образует с нормалью к плоскости зеркал (осью резонатора),

$s$  – поверхность зеркала.

Далее предполагается, что после многочисленных проходов распределение поля у зеркал будет подвергаться незначительным изменениям от отражения к отражению и со временем поля около зеркал становятся одинаковыми с точностью до постоянной величины.

$$U_q = \left(\frac{1}{\gamma}\right)^q \cdot U(x, y), \quad (8.13)$$

где  $U(x, y)$  – функция, описывающая распределение поля по поверхности зеркал, не меняющаяся от отражения к отражению,  $\gamma$  – комплексная постоянная, независимая от координат. Подставляя (8.13) в (8.12), получаем интегральное уравнение

$$v = \gamma \cdot \oint_s k(r, \theta) \cdot v \cdot ds, \quad (8.14)$$

с ядром интегрального уравнения

$$k(r, \theta) = \frac{L \cdot K}{4 \cdot \pi \cdot r} \cdot (1 + \cos \theta) \cdot e^{-i \cdot k \cdot r} . \quad (8.15)$$

Функцию распределения  $U(x, y)$ , которая удовлетворяет уравнению (8.14), можно рассматривать как нормальный тип колебаний интерферометра. Действительная часть определяет потери за один проход, а мнимая – фазовый сдвиг за один проход, добавляемого к геометрическому фазовому сдвигу.

Таким образом, в резонаторе, в зависимости от условий, могут устанавливаться отдельные типы колебаний, причем для каждого типа колебаний характерны своё распределение поля на поверхности зеркал, свои потери и фазовый сдвиг за один проход. Условие резонанса в этом случае можно записать

$$\frac{2 \cdot \pi \cdot L}{\lambda} + \varphi = \pi \cdot q, \quad (8.16)$$

где  $\varphi$  – фазовый сдвиг, соответствующий определенному типу колебаний. Из (8.16) видно, что в резонаторе каждому типу колебаний соответствует в общем случае и своя частота

$$\omega_q = \frac{q \cdot \pi \cdot c}{L} - \frac{c \cdot \varphi}{L} . \quad (8.17)$$

Из (8.17) видно, что при постоянных  $q$  и  $L$  частота зависит от фазового сдвига и будет различной для различных типов колебаний. Разность частот колебаний одного и того же типа, но отличающихся по индексу  $q$  на единицу, равна

$$\Delta \omega = \omega_{q+1} - \omega_q = \frac{\pi \cdot c}{L}, \quad (8.18)$$

и совпадает с разностью частот продольных мод резонаторов Фабри - Перо.

Сокращенно нормальные типы колебаний называют модами и обозначают как  $TEM_{mnq}$ , где  $m$  и  $n$  – целые числа равные  $0, 1, 2, \dots$ , обозначающие число изменений знака поля на поверхности зеркал в двух ортогональных направлениях, а  $q$  продольным или аксиальным индексом. В соответствии с этим моды, характеризующиеся одними и теми же индексами  $m$  и  $n$ , но разными  $q$ , объединяются под общим названием поперечной моды. Если нужно выделить колебание с определенным индексом  $q$ , то говорят о продольной моде.

Рассмотрим некоторые конкретные примеры мод резонаторов (рис.. 8.4.).

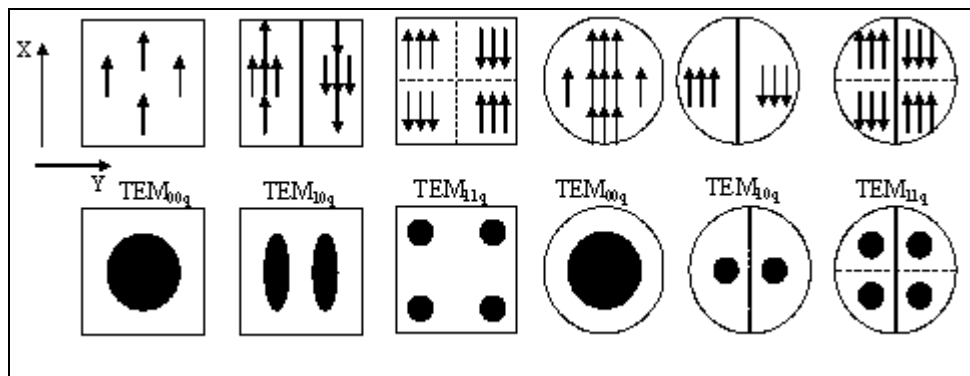


Рис. 8.4. Структура поля для некоторых мод низшего порядка



Для резонаторов с прямоугольными зеркалами индекс  $m$  означает число изменений направления поля вдоль большей стороны (оси  $x$ ), а  $n$  – вдоль меньшей стороны (оси  $y$ ). Для круглых зеркал буква  $n$  означает число изменений поля по окружности, а  $m$  – вдоль радиуса. Из рис. 8.4., например, видно, что для плоских зеркал поле моды  $TEM_{10}$  два пятна, симметрично расположенных относительно оси  $x$ , поле моды  $TEM_{11}$  – четыре пятна и т.д.

Чем больше индексы  $m$  и  $n$ , тем более высоким считается тип колебаний, тем сложнее распределение поля на зеркалах. Обычно число  $q=10^5-10^6$  много больше чисел  $m$  и  $n$ , которые обычно характеризуются числами от 1 до 10. Поэтому разность частот между различными поперечными модами, относящимися к одному и тому же индексу  $q$  (порядка нескольких десятков и сотен кГц), много меньше расстояния между продольными модами (сотни МГц).

Рассмотрим теперь резонатор, в котором помещено вещество, способное усиливать электромагнитное излучение. Поскольку активное вещество обладает определенной спектральной характеристикой, то оно будет усиливать лишь те спектральные составляющие пассивного резонатора, которые попадают в пределы спектральной линии вещества. На рис. 8.5. изображены спектральная характеристика вещества и моды пассивного резонатора. Из рисунка следует, что генерировать будут лишь те моды, для которых усиление больше потерь. Если учесть, что центральная частота линии  $\nu_0$  обычно составляет  $10^{14} \div 10^{15}$  Гц, а разность частоты соседних продольных мод 100-1000 МГц, то становится понятным большое значение индекса  $q$  для генерирующих мод лазера. Зная полуширину линий и длину резонатора, можно оценить число генерирующих продольных мод как:

$$j = \frac{\Delta\omega_0}{\frac{c \cdot \pi}{L}}$$

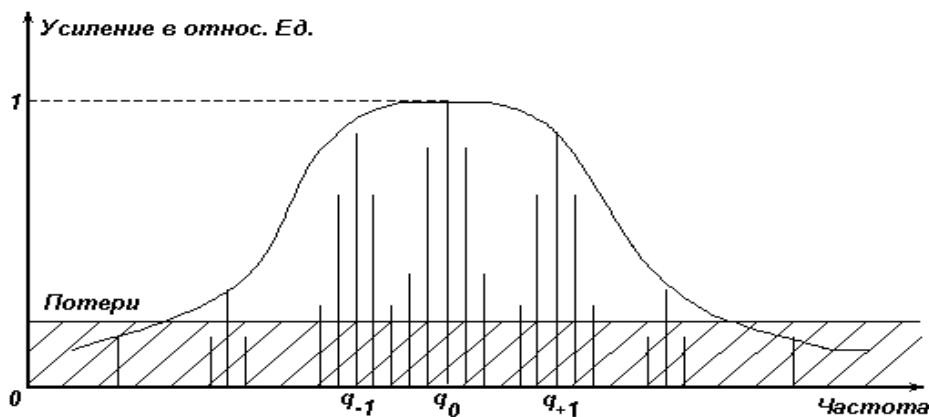


Рис.8.5. Спектральная характеристика вещества и моды пассивного резонатора

Так, например, для гелий-неонового лазера с  $\Delta\nu_0 = 800$  МГц,  $L=1$  м,  $j=6$

Из рис. 8.5. следует и другой важный вывод. Поскольку каждому поперечному типу колебаний соответствуют определенные потери тем больше, чем больше

$m$  и  $n$ , различные для разных типов, то видно, что условия возбуждения различных мод неодинаковы. В первую очередь должны возбуждаться моды с наименьшими потерями. Поскольку наименьшими потерями обладает мода  $TEM_{00}$ , то с увеличением накачки первой будет выходить в генерацию она. По мере возрастания мощности накачки будут возбуждаться и другие моды [4,13,14].

**Классификация оптических резонаторов.**[11]. Кроме открытого резонатора с плоскими зеркалами в оптических квантовых генераторах используются системы других конфигураций. Широкое распространение получили оптические резонаторы, образованные двумя сферическими зеркалами или одним сферическим и одним плоским зеркалами. Различные типы резонаторов со сферическими зеркалами представлены на рис. 8.6, а—г. Для специальных целей применяются более сложные системы, например составные и кольцевые резонаторы.

Расчет поля и определение собственных колебаний в резонаторе произвольной конфигурации могут быть выполнены аналогично тому, как в резонаторе с плоскими зеркалами. Рассмотрим кратко свойства различных типов оптических резонаторов.

**1. Конфокальный резонатор.** Конфокальным называют открытый резонатор, образованный одинаковыми сферическими зеркалами, оси и фокусы которых совпадают (рис.8.6а). Так как фокус сферического зеркала радиусом  $R$  расположен на расстоянии  $R/2$ , то это означает, что радиусы кривизны зеркал равны длине резонатора. Конфокальный резонатор имеет следующие особенности:

а) Распределение фазы по поверхности зеркал однородно, т. е. отражающие поверхности являются поверхностями равных фаз.

б) По сравнению с плоским резонатором поле в конфокальном резонаторе более плотно сконцентрировано у его оси и спадает на краях значительно быстрее. Это приводит к тому, что дифракционные потери в конфокальном резонаторе значительно меньше, чем для аналогичных типов колебаний в плоском резонаторе. Распределение поля для разных  $N$  приведено на рис.8.7.

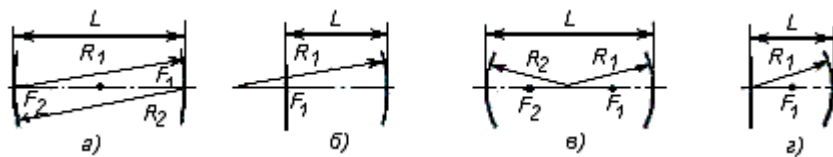


Рис. 8.6. Схемы различных типов резонаторов со сферическими зеркалами: а – конфокальный ( $R_1=R_2=L$ ); б – полуконфокальный ( $R_1=2L$ ,  $R_2=\infty$ ); в – концентрический ( $R_1=R_2=L/2$ ); г – полуконцентрический ( $R_1=L$ ,  $R_2=\infty$ )

в) В распределении поля не наблюдается характерных для плоского резонатора перегибов (рис. 8.7). Основное  $TEM_{00}$ -колебание имеет гауссово распределение как в направлении  $x$ , так и в направлении  $y$ . Радиус светового пятна  $r_{1/e}$  соответствующий уменьшению поля в  $TEM_{00}$ -моды в  $e$  раз,

$$r_{1/e} = \sqrt{\lambda L / \pi}. \quad (8.19)$$

Пучок, распределение поля в поперечном направлении которого характеризуется функцией Гаусса, называется *гауссовым пучком*.

г) Собственные частоты определяются по формуле

$$v_{mng} = \frac{\omega_{mng}}{2\pi} = \frac{c}{4L} (2g + 1 + m + n). \quad (8.20)$$

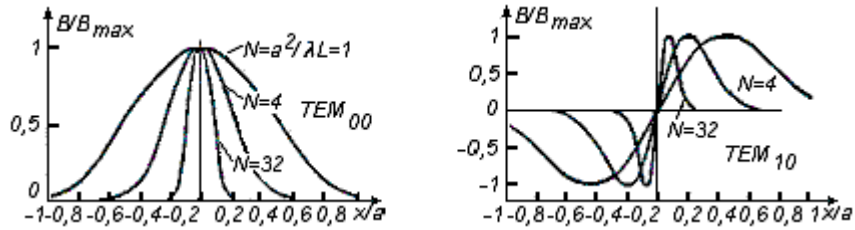


Рис. 8.7. Распределение относительной амплитуды  $E/E_{max}$  для  $TEM_{00}$  и  $TEM_{10}$ -колебаний в конфокальном резонаторе

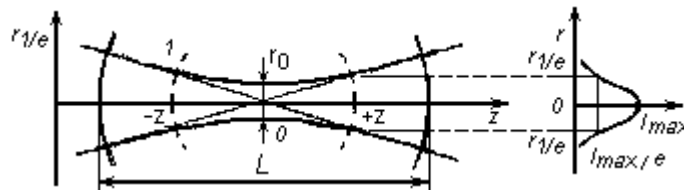


Рис.8.8. Радиус светового пятна синфазные поверхности распределение интенсивности в конфокальном резонаторе

Отсюда видно, что спектр собственных частот в конфокальном резонаторе сильно вырожден.

д) Поле внутри резонатора в наибольшей степени сконцентрировано вблизи общего фокуса зеркал, т. е. в центре резонатора. Поверхности равных фаз представляют собой сферические поверхности. Радиус кривизны синфазной поверхности, пересекающей оптическую ось резонатора в точке  $z$  (начало отсчета  $z=0$  совпадает с общим фокусом зеркал), определяется соотношением

$$R(z) = z(1 + L^2 / 4z^2). \quad (8.21)$$

След этих поверхностей изображен на рис. 8.8. пунктирными линиями. С приближением к центру резонатора радиусы кривизны синфазных поверхностей увеличиваются. При  $z=0$   $R \rightarrow \infty$ , т. е. синфазная поверхность является плоской.

е) Распределение поля внутри резонатора в поперечном направлении для  $TEM_{00}$ -моды приближенно описывается функцией Гаусса, причем радиус светового пятна  $r_{1/e}$ , изменяется с координатой  $z$  по закону

$$r_{1/e} = \left(\frac{\lambda L}{2\pi}\right)^{1/2} \sqrt{1 + \left(\frac{2z}{L}\right)^2}. \quad (8.22)$$

Эта зависимость представлена на рис. 8.8. сплошной линией. Радиус пятна минимален в центре резонатора. Он в  $\sqrt{2}$  раз меньше, чем на поверхности зеркал.

Перечисленные особенности распределения поля в конфокальном резонаторе имеют как преимущества, так и недостатки. То, что благодаря фокусирующему

действию зеркал поле стремится сконцентрироваться вдоль оси резонатора что, обуславливает минимальные дифракционные потери. конфокальный резонатор менее чувствителен к разъюстировке зеркал. Однако, это иногда мешает полностью использовать объем активного вещества в резонаторе.

**2. Резонаторы с произвольными сферическими зеркалами.** Это резонаторы, состоящие из двух соосных сферических зеркал радиусами  $R_1$  и  $R_2$ , расположенных на расстоянии  $L$  друг от друга. Свойства таких резонаторов легко определить, если найдена соответствующая конфокальная система, в которой две синфазные поверхности совпадают с поверхностями зеркал резонатора. Если известны радиусы  $R_1$  и  $R_2$  зеркал и расстояние между ними, то из выражения (8.21) можно найти длину, а также радиусы кривизны и координаты зеркал эквивалентного конфокального резонатора. Если известно  $L_{\text{эк}}$ , то распределение поля внутри и вне резонатора будет таким же, как в эквивалентном конфокальном резонаторе. В частности, радиус светового пятна будет определяться соотношением (8.22), где вместо  $L$  будет стоять  $L_{\text{эк}}$ . Если резонатор образован двумя одинаковыми сферическими зеркалами радиусами  $R_1=R_2=R$ , расположенными на расстоянии  $L$  друг от друга, то из (8.21) получаем, что длина эквивалентного конфокального резонатора

$$L_{\text{эк}} = \sqrt{L(2R - L)}. \quad (8.23)$$

Эквивалентный конфокальный резонатор определяет только собственные функции произвольного сферического резонатора. Нахождение собственных значений не может быть сведено к задаче о конфокальном резонаторе. В частности, собственные частоты  $\nu_{mq}$  в общем случае будут определяться не соотношением (8.20), а более сложными выражениями. В обобщенном сферическом резонаторе частотное вырождение мод обычно исчезает.

Сферический резонатор может быть либо устойчивым, либо неустойчивым (в последнем случае резонатор по существу теряет свои резонансные свойства). Например, если резонатор образован двумя одинаковыми сферическими зеркалами радиусами  $R$ , причем  $L > 2R$ , то, как следует из (8.23), ему невозможно подобрать эквивалентный конфокальный резонатор. Это означает, что в таких резонаторах невозможно образование устойчивого стационарного распределения электромагнитного поля с малыми потерями.

Пусть  $R_1$  и  $R_2$  — радиусы зеркал, а  $L$  — расстояние между ними. Можно показать, что резонатор будет устойчивым и ему всегда можно подобрать эквивалентный конфокальный резонатор, если выполняется неравенство

$$0 \leq \left(\frac{L}{R_1} - 1\right)\left(\frac{L}{R_2} - 1\right) \leq 1. \quad (8.24)$$

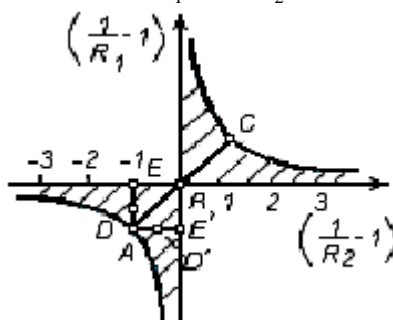


Рис. 8.9. Диаграмма устойчивости оптических резонаторов

На рис. 8.9. приведена диаграмма, иллюстрирующая условие устойчивости оптических резонаторов. Устойчивые области соответствуют заштрихованным участкам. Отметим характерные точки на этой диаграмме.

Точке  $A$  с координатами  $(-1; -1)$  соответствует резонатор с плоскими зеркалами, расположенный на границе устойчивости. Точке  $B$  с координатами  $(0; 0)$  соответствует конфокальный резонатор.

Точке  $C$  с координатами  $(1; 1)$  соответствует резонатор, образованный двумя одинаковыми сферическими зеркалами, оси и центры кривизны которых совпадают, т. е.  $R_1=R_2=L/2$ . Такой *концентрический резонатор* (см. рис. 8.6,в) также лежит на границе, разделяющей устойчивую и неустойчивую области. Дифракционные потери в нем очень быстро возрастают с увеличением типа колебаний, т. е. с увеличением индексов  $m$  и  $n$ . Поэтому такой резонатор применяют для селекции (отбора) неаксиальных колебаний: за счет большой разницы в дифракционных потерях между аксиальными  $TEM_{00}$  и неаксиальными модами последние в концентрическом резонаторе обладают малой добротностью-и, следовательно, гасятся.

Точкам  $D, D'$  с координатами  $(-1; -0,5)$  и  $(-0,5; -1)$  соответствует *полуконфокальный резонатор*, т. е. резонатор, образованный одним плоским и одним сферическим зеркалом, радиус кривизны которого равен удвоенной длине резонатора (см. рис. "8.6 б). Он находится в середине устойчивой области и его свойства аналогичны конфокальному резонатору с удвоенной длиной.

Точкам  $E$  и  $E'$  с координатами  $(-1; 0)$  и  $(0; -1)$  соответствует *полуконцентрический резонатор* (см. рис. 8.6., г). Он обладает особенностями, характерными для концентрического резонатора, а также применяется для селекции неаксиальных колебаний.

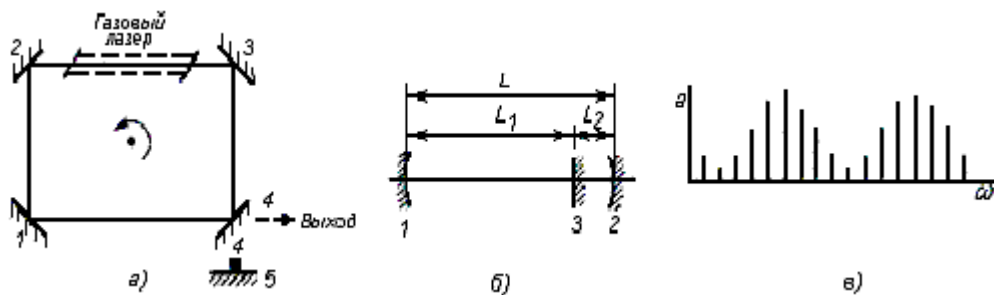


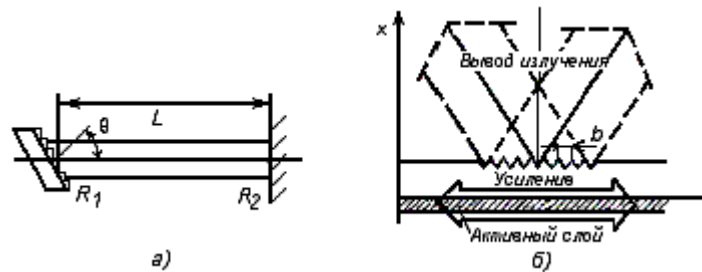
Рис. 8.10. Схемы кольцевого (а) и составного (б) резонаторов; зависимость добротности аксиальных колебаний от частоты в составном резонаторе (в)

**3. Кольцевой резонатор.** Кольцевым называют открытый резонатор, зеркала которого обеспечивают распространение электромагнитных волн по замкнутому контуру (рис. 8.10а). Система из четырех зеркал  $1, 2, 3$  и  $4$  образует замкнутый контур. В одно из плеч резонатора помещают активный элемент, например, газоразрядную трубку газового лазера. В кольцевом резонаторе, как и в рассмотренных оптических резонаторах, может существовать стоячая волна, образованная интерференцией двух волн, бегущих в противоположных направ-

лениях. Если каким-либо образом устранить одну из бегущих волн, например, сделав зеркало 4 полупрозрачным и поставив дополнительное зеркало 5 (как показано на рис. 8.10а), то в таком резонаторе можно осуществить режим бегущей волны. В бегущей волне поле в среднем более однородно, что позволяет полнее использовать активное вещество.

Кроме того, если вращать кольцевой резонатор вокруг оси, то длина пути для волн, распространяющихся по направлению и против направления вращения, будет различной. Это может быть использовано для измерения скорости вращения и построения лазерных гироскопов.

**4. Составной резонатор.** Такой резонатор представляет собой два (или более) связанных между собой резонатора. Схема составного резонатора приведена на рис. 8.10б. Зеркало 3 является полупрозрачным. Свойства такого резонатора аналогичны свойствам двух связанных контуров. Как видно из рис. 8.10в, составной резонатор можно использовать для селекции аксиальных колебаний.



**Рис. 8.11.** Схема резонатора с брэгговским зеркалом (а) и с распределенной обратной связью(б)

Дополнительная спектральная селекция колебаний в составном резонаторе определяется тем, что наибольшей добротностью будут обладать те типы колебаний, для которых условие резонанса наилучшим образом удовлетворяется в каждой из составных частей резонатора.

**5. Резонатор с брэгговским зеркалом.** Иногда возникает необходимость плавно перестраивать собственную частоту резонатора и тем самым плавно изменять частоту генерации лазера в пределах контура спектральной линии активного вещества. В более широком диапазоне перестройка частоты может быть осуществлена, если одно из зеркал сделать селективным, т. е. обладающим большим коэффициентом отражения в узком спектральном диапазоне. При этом необходимо, чтобы была возможность каким-либо образом изменять спектральный диапазон. В качестве селективного элемента может быть использована, например, дифракционная решетка, которую -устанавливают вместо одного из зеркал резонатора. Подобный резонатор называется *резонатором с брэгговским зеркалом* (рис. 8.11 а). Дифракционная решетка располагается под углом к оптической оси резонатора. Если на такое «зеркало» падает плоская электромагнитная волна, распространяющаяся вдоль оптической оси резонатора, то за счет дифракции на решетке она

будет отражаться назад точно в противоположном направлении при выполнении условия Вульфа Брэгга:

$$b \sin \theta = m\lambda / 2,$$

где  $b$  — период дифракционной решетки;  $\theta$  — угол между нормалью к плоскости решетки и оптической осью резонатора;  $m=1; 2; 3; \dots$  — порядок дифракции.

Изменяя угол  $\theta$ , можно изменять длину волны  $\lambda$  соответствующую максимуму отражательной способности зеркала в направлении оптической оси. Такой резонатор, применяется, в частности, для перестройки частоты генерации жидкостных лазеров на органических красителях.

**6. Резонатор с распределенной обратной связью.** Это особый вид резонатора, в котором торцевые зеркала вообще могут отсутствовать, а положительная обратная связь обеспечивается рассеянием на периодических неоднородностях, образующих строго периодическую решетку. Важно, чтобы эта решетка образовывалась пространственными периодическими изменениями одного из параметров, от которого зависят условия распространения света. Это может быть коэффициент преломления, коэффициент усиления, толщина пленки или кристалла. Расстояние  $b$  между неоднородностями должно удовлетворять условию Вульфа — Брэгга. В оптических резонаторах, реализованных таким образом, решетка является одновременно фильтром. и отражателем, выполняя роль зеркал «обычных» резонаторов.

Распределенная обратная связь используется, в частности, в некоторых типах полупроводниковых лазеров [11,22].

**Многомодовость** [6]. Наличие большого числа резонансных частот резонатора ведет к тому, что лазер может одновременно генерировать колебания на нескольких частотах. Понять это легко, если учесть, что фактическая ширина спектральных линий излучения активного вещества (см. волновая теория резонаторов) может быть в ряде случаев намного

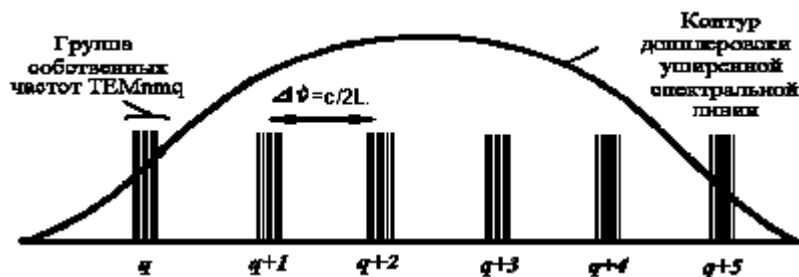


Рис. 8.12. К иллюстрации многомодовости оптических резонаторов

больше, чем резонансные полосы резонатора и расстояния между ними. Это может привести к тому, что условия самовозбуждения будут соблюдаться сразу для нескольких частот. Так, найденная [6] ширина линии активного вещества при доплеровском уширении составляет около 1 000 МГц, а ширина резонансных линий равна примерно 1 МГц, при расстоянии между частотами продольных мод около 150 МГц. Таким образом, может существовать картина, подобная изображенной на рис. 8.12. Резонатор, как видно, из широкой полосы спектра активного веществ-

ва будет вырезать узкие линии, соответствующие его собственным частотам. Поскольку затухание составляющих спектра при больших  $m$  и  $n$  сильно возрастает, то число поперечных мод, частоты которых располагаются вблизи частот продольных мод, подобно боковым частотам обычно невелико.

Следует иметь в виду, что частоты колебаний, генерируемых лазером, не точно совпадают с собственными частотами резонатора вследствие взаимодействия колебательных систем вещества как излучающих диполей с резонатором.

Наличие многих частот в излучении создает сложную картину распределения интенсивности в поперечном сечении пучка света. С увеличением мощности накачки число генерируемых частот обычно увеличивается. Многочастотность является одной из причин пикового характера излучения (см. п. 9.3.2)

Действительно, интерференция волн приводит к тому, что в некоторые моменты они могут гасить друг друга, а в другие моменты усиливать. В результате лазер вместо непрерывного колебания или длинного импульса генерирует серию очень коротких импульсов. Эти импульсы тем короче и мощнее, чем большее число волн в излучении. Наблюдаемые интерференционные импульсы имеют иногда продолжительность до  $10^{-12}$  с.

Пиковая картина излучения лазера обусловлена, однако, не только этой причиной. К другим причинам относятся явления релаксации в веществе и неодновременное возбуждение отдельных участков активного вещества.

Из-за всех указанных причин излучение лазера не монохроматично и имеет конечную ширину линии излучения.

**Селекция мод резонаторов.** Многомодовость вредна при использовании лазеров во многих областях применения, так как при этом ограничивается полоса передаваемого сообщения и снижается помехоустойчивость. Она ухудшает когерентность излучения и его направленность. Поэтому постоянно ведутся активные работы по разработке одномодовых и одночастотных лазеров. Так, например, уменьшая длину лазера  $L$ , можно сделать расстояние между частотами поперечных мод больше полосы усиления активного вещества. Однако во всех случаях это связано с уменьшением их к. п. д. или мощности и усиления. Многие методы селекции мод резонаторов лазера основаны на создании условий, при которых увеличиваются потери для нежелательных мод. Так, например, используя рост дифракционных потерь с ростом поперечных индексов  $m$  и  $n$ , можно добиться генерации только на нескольких поперечных модах или даже одной или нескольких продольных модах, изменяя кривизну зеркал или расстояние между ними.

Один из самых простых методов получения генерации на одной частоте заключается в соответствующем выборе уровня мощности накачки. Именно: если уровень мощности накачки превосходит пороговое значение лишь настолько, что усиление превышает потери только на центральной частоте контура усиления, то лазер будет генерировать лишь на одной частоте, совпадающей с центральной частотой контура. В этом случае мощность генерации невелика и необходимо поддерживать уровень накачки в определенных пределах.

Кроме изменения параметров резонатора и мощности накачки, для селекции мод используются дополнительные устройства, располагаемые внутри или вне резонатора. Одним из таких устройств, с помощью которого подавляются поперечные моды высокого порядка, является диафрагма, расположенная внутри резонатора. Распределение поля в плоскости поперечного сечения резонатора таково,



что основной тип  $TEM_{00q}$  концентрируется вблизи оси резонатора, а поля других типов  $TEM_{11q}$ ,  $TEM_{20q}$  и т. д. простираются от оси тем дальше, чем более высоким порядком они соответствуют. Поэтому введение в резонатор диафрагмы с отверстием, расположенным вблизи оси резонатора, будет увеличивать потери типов колебаний высокого порядка.

Существенно отметить, что одномодовый режим генерации, при котором возбуждается одна низшая поперечная мода, позволяет получить в дальней зоне предельно узкую диаграмму излучения, ширина которой соответствует дифракционному пределу (см. п. 8.6).

Один из наиболее эффективных методов селекции продольных типов колебаний основан на использовании системы связанных резонаторов. С этим и другими методами подробнее можно ознакомиться в книгах, посвященных лазеру [8,13].

## 8.5. Свойства лазерного излучения. Характеристики излучения лазеров

**Когерентность.** Излучение лазера обладает рядом замечательных свойств – высокая когерентность и монохроматичность.

Рассмотрим понятие когерентности в применении к лазеру. Когерентностью называют согласованное протекание во времени нескольких волновых процессов, или свойство, отражающее стабильность характеристик одной или нескольких электромагнитных волн.

Когерентность характеризует, например, способность электромагнитных волн интерферировать друг с другом. Рассмотрим интерференцию двух волн, осуществляемых каким-либо способом. В интерференционной картине будут наблюдаться максимумы (при  $\Delta r k = 2\pi n$ ) и минимумы (при  $k\Delta r = (2n+1)\pi$ ,  $k = 2\pi/\lambda = 2\pi\nu/c$ ). Если  $I_{max}$  – максимальная интенсивность света, а  $I_{min}$  – минимальная интенсивность света, то можно ввести понятие контрастности интерференционной картины:  $\gamma = (I_{max} - I_{min})/I_{max}$ , которая определяет ее видность (рис. 8.13.)

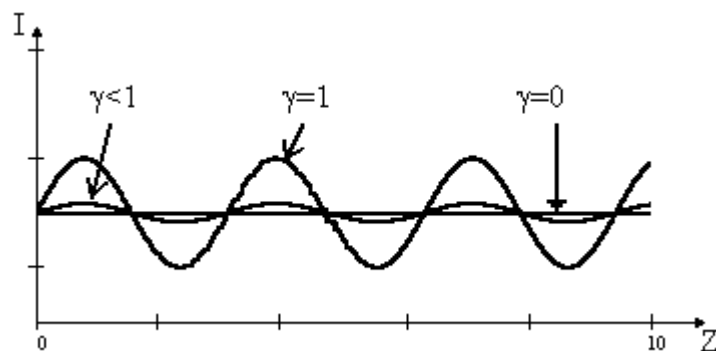


Рис. 8.13. Распределение интенсивности света на экране

При  $\gamma=1$  излучение полностью когерентно, при  $\gamma=0$  полностью не когерентно. Для всех реальных источников света  $\gamma < 1$ , что связано с двумя причинами:

1. излучение не строго монохроматично;
2. источник имеет конечные размеры, причем фаза лучей, испускаемых из разных точек источника, отличается друг от друга.

Рассмотрим первую причину – условие  $k\Delta r = 2\pi n$  выполняется в нескольких точках, так как  $\Delta\nu \neq 0$  и картина размывается. Оценим количественно [4]. Как известно атомы излучают электромагнитные волны не непрерывно, а в течении некоторых интервалов времени, промежутки между которыми случайны (рис. 8.14)

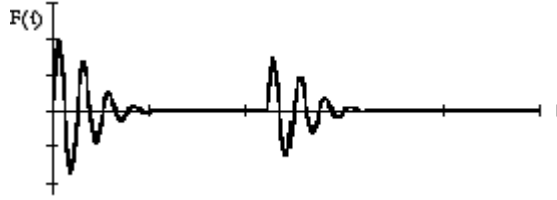


Рис. 8.14. Цуги, излучаемые атомом

Часть сигнала, в течение, которого его фаза меняется непрерывно, называется цугом. За время цуга фаза излучения сохраняется (и когерентность сохраняется). Длительность цуга  $\tau$  связана с шириной спектра следующим образом  $\Delta\omega = 2\pi/\tau$ , (то есть с монохроматичностью). Время  $\tau$  – называется временем когерентности, а длину цуга в пространстве  $\ell = c \cdot \tau = 2\pi c / \Delta\omega = \lambda_0^2 / \Delta\lambda$  – длиной когерентности.

Очевидно, что контрастная интерференционная картина получается в том случае, если смещение максимума, связанное с  $\Delta\lambda \neq 0$ , будет много меньше расстояния между соседними порядками интерференции:

$$d(k\Delta r) = 2\pi\Delta r\Delta\lambda / \lambda_0^2 \ll 2\pi$$

Откуда получаем:

$$\Delta r \ll \lambda_0^2 / \Delta\lambda; \quad \Delta r \ll \ell$$

То есть для получения интерференционной картины с хорошей видностью разность хода  $\Delta r$  должна быть много меньше длины когерентности  $\ell$ .

Пример: для белого света  $\tau \approx 10^{-14}$  с;  $\ell \approx 1$  мкм для гелий – неоновом лазере (в многомодовом режиме)  $\lambda_0 = 0.63$  мкм,  $\Delta\nu = 0.5$  ГГц,  $\tau = 2 \cdot 10^{-10}$  с,  $\ell = 60$  см, а для одностотного режима работы  $\Delta\nu = 0.5$  Гц,  $\tau = 2 \cdot 10^{-2}$  с,  $\ell = 2 \cdot 10^7$  м.

Вторая причина, ухудшающая когерентность – демонстрируется с помощью известного опыта Юнга. Не трудно показать, что видность интерференционной картины (и когерентность) зависит от стабильности фазы излучения в плоскости поперечного сечения пучка. Это свойство называют пространственной когерентностью ( $k\Delta r + \Delta\phi(r) = 2\pi n$ ).

В случае тепловых источников, когерентность излучения ухудшается с увеличением их размеров, т.к. разные элементы нагретого тела излучают независимо друг от друга по всем направлениям. Но по мере удаления от источника это случайное расходящееся поле все больше приближается к плоской волне, и поле будет пространственно когерентно в пределах круга диаметром:  $d = 0.16\lambda L / R$  (по теореме Ван Ситтерта – Цернике), где  $R$  – поперечный размер источника,  $L$  – расстояние от источника до точки наблюдения ( $L \gg R$ ). Интенсивность излучения пропорциональна  $1/L^2$  и эффективное когерентное излучение нельзя получить от тепловых источников.

Излучение лазера образуется в результате согласованного индуцированного излучения атомов во всем объеме активной среды, и в пределах всего поперечно-

го сечения пучка есть жесткая фазовая корреляция, так что волна имеет плоский или сферический фронт. Поэтому излучение лазера обладает пространственной когерентностью, и лазеры позволяют получить когерентный свет высокой интенсивности.

Причины, ухудшающие пространственную когерентность реальных лазеров – это дефекты и неоднородности рабочего вещества, генерация многих поперечных мод, и так далее. Как правило, где нужна высокая когерентность, работают на одной поперечной моде.

**Монохроматичность.** Монохроматичность характеризует степень «размазанности» излучения по спектру. Монохроматическим является гармоническое колебание, например, такого вида  $E(t) = E_0 \sin(\omega_0 t + \varphi_0)$ , где  $E_0$  – амплитуда частоты  $\omega_0$ .  $\varphi_0$  – начальная фаза, эти величины постоянны, а время  $t$  меняется от  $-\infty$  до  $+\infty$ . Количественной характеристикой степени монохроматичности является ширина спектральной линии на уровне 0.5 от ее максимума или спектральный диапазон  $\Delta\omega(\Delta\lambda)$  занимаемого группой линий. Более объективной характеристикой является не абсолютная, а относительная ширина спектра  $\Delta\omega/\omega_0 = \Delta\lambda/\lambda_0$ , где  $\omega_0, \lambda_0$  – соответствуют максимуму спектра. Величина  $\Delta\omega/\omega_0$  называется спектральной чистотой излучения.

Ширина спектральной моды, выделяемой резонатором, определяется его добротностью  $Q$

$$\Delta\omega = \omega_0 / Q = 1 / \tau, \quad (8.49)$$

где  $\tau$  – время жизни фотона в резонаторе.

В свою очередь  $Q$ , определяется потерями в резонаторе, но резонатор лазера выполняет не только функции пассивного фильтра. Ширина спектральной линии генерируемой лазером в одномодовом режиме, будет существенно меньше величины, определяемой (8.49) для пассивного резонатора. Если внутри резонатора находится активный элемент, компенсирующий потери на частоте генерации, то для такой моды в режиме генерации,  $Q \rightarrow \infty$  и  $\Delta\omega \rightarrow 0$ . Теоретический предел ширины спектральной линии лазерного излучения определяется двумя факторами:

- 1) шумами, обусловленными тепловым излучением в резонаторе;
- 2) шумами, связанными со спонтанным излучением активного вещества.

В оптическом диапазоне шумы за счет спонтанного излучения преобладает над тепловыми шумами. Если учитывать только шумы, вызванные спонтанными переходами, то окажется, что спектральная линия, выходного лазерного излучения имеет лоренцову форму с полушириной:

$$\Delta\omega = 4 \hbar \omega_0 (\Delta\omega_c)^2 / P, \quad (8.50)$$

где  $P$  – выходная мощность лазерного излучения. Предполагается, что лазер генерирует в стационарном, непрерывном режиме на одной моде и его параметры не изменяются во времени. Величину  $\Delta\omega$  можно рассматривать как теоретический предел ширины спектральной линии генерации.

Для лазера с  $P_{\text{вых}} = 1 \text{ мВт}$ ,  $\lambda = 0.63 \text{ мкм}$ ,  $\omega_0 / 2\pi = 4.76 \cdot 10^{14} \text{ Гц}$ ,  $\hbar \omega_0 = 3.15 \cdot 10^{-19} \text{ Дж}$ . Для  $Q \approx 10^8$  из (8.50) получаем  $\omega_0 / 2\pi = 0.2 \text{ Гц}$ , что соответствует спектральной чистоте  $\Delta\omega / \omega_0 = 5 \cdot 10^{-16}$ .

Для получения такой спектральной чистоты накладывают определенные требования на стабильность длины резонатора. Так как  $\Delta\omega / \omega_0 = \Delta L / L$ , то есть при  $L = 1 \text{ м}$ , допустимое отклонение резонатора от заданной величины

$\Delta L = 5 \cdot 10^{-16} = 5 \cdot 10^{-7} \text{ нм}$ . Если учесть, что типичные размеры атомов составляют величину  $\sim 0.1 \text{ нм} = 1 \text{ \AA}$ , то очевидно, что стабилизировать длину резонатора в таких изделиях трудно. Это означает, что в реальных условиях монохроматичность лазерного излучения определяется изменениями длины резонатора, естественными тепловыми эффектами, вибрациями. В очень жестких условиях было достигнуто  $\Delta\omega/\omega_0 = 10^{-14}$ . При обычных условиях, но с использованием методов стабилизации частоты, можно получить спектральную чистоту  $\sim 10^{-12} - 10^{-13}$ . В видимой области этому соответствуют  $\Delta\omega/2\pi = 50 \text{ Гц}$ . Ни один из «классических» источников излучения не способен обеспечить такие параметры. Например, при выделении одной спектральной линии у ртутных или натриевых ламп в лучшем случае достигают  $\Delta\omega/\omega_0 = 10^{-6}$ .

В большинстве случаев, если не принимать специальные меры по селекции мод, лазер будет генерировать в многомодовом режиме. При генерации в импульсном режиме спектр излучения для каждой моды будет уширен до  $\nu_{\text{ген}} \approx 1/\tau_{\text{и}}$ , где  $\tau_{\text{и}}$  – длительность импульса. Поэтому, в режиме генерации гигантских импульсов при  $\tau_{\text{и}} = 1 \text{ нс}$ , даже при одномодовой генерации  $\Delta\nu_{\text{ген}} = 1000 \text{ МГц}$ . В пиковом режиме генерации значение  $\Delta\nu_{\text{ген}}$  для каждой моды будет определяться длительностью отдельных пиков.

Спектр лазерного излучения также может быть уширен вследствие неоднородности активного материала (например, за счет термических деформаций), что особенно характерно для твердотельных лазеров

**Направленность:** определяет расходимость светового пучка в пространстве и характеризуется плоским и телесным углом, в котором распространяется большая часть излучения. Расходимость пучка служит мерой его отклонения от параллельности.

«Классические» источники света обладают очень низкой направленностью излучения и обычно излучают в пределах телесного угла от  $2\pi$  до  $4\pi$ .

Источники, излучающие по закону абсолютно черного тела, дают ненаправленное излучение: поверхностная плотность излучаемой электромагнитной энергии не зависит от направления в пределах всего телесного угла  $4\pi$  (источник Ламберта). От них можно получить направленное излучение с помощью диафрагм, линз, зеркал, но при этом большая часть излучаемой энергии теряется.

Лазерный излучатель по своей природе обладает высокой степенью направленности, которая определяется свойствами резонатора. Расходимость лазерного пучка при генерации на одной аксиальной моде близка к предельно достижимой расходимости, которая ограничивается явлением дифракции.

Дифракционная расходимость является следствием естественного расширения лазерного пучка по мере движения световой волны в пространстве. Она может быть пояснена с помощью принципа Гюйгенса – Френеля (рис. 8 15) [11].

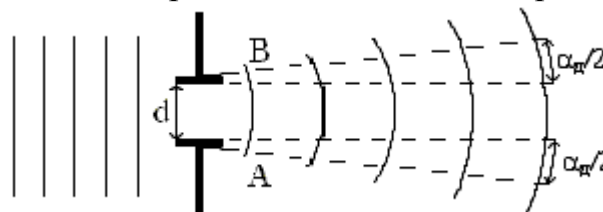


Рис. 8.15. Дифракционная расходимость плоской электромагнитной волны

На диафрагму диаметра  $d$  падает идеальная плоская электромагнитная волна с полной пространственной когерентностью. После диафрагмы вследствие дифракции волновой фронт искажается, и излучение приобретает дифракционную расходимость.

Мерой дифракции и мерой расходимости служит угол  $\alpha_0$ , под которым из отверстия виден диаметр первого темного дифракционного кольца. Вычисления дают: при  $d \gg \lambda$  угол

$$\alpha_d = 2,44 \cdot \frac{\lambda}{d}. \quad (8.51)$$

Для гауссова пучка, который получается в резонаторе со сферическим зеркалом, распределение интенсивности в поперечном сечении представляется гладкой экспоненциальной кривой. С увеличением расстояния  $r$  от оси пучка интенсивность убывает по закону  $\exp(-\frac{r}{r_1/l})$ ,

$$\text{где } r_1/l \approx \frac{\lambda \cdot z}{\pi \cdot r_0}, \quad (r_0 = (\frac{\lambda \cdot L}{2\pi})^2 - \text{радиус шейки пучка внутри резонатора (рис.8.8)).}$$

Поэтому угол дифракционной расходимости гауссова пучка определяется

$$\alpha_d = \frac{2\lambda}{\pi \cdot r_0} \approx 1,27 \frac{\lambda}{d_0}, \quad (8.52)$$

где  $d_0$  – диаметр пучка в наиболее узкой его части.

Уравнения (8.51) и (8.52) показывают, что при одинаковых  $d$  - расходимость гауссова пучка примерно в два раза меньше расходимости плоского пучка.

В полупроводниковых лазерах (инжекционных) генерация происходит неравномерно по всему объему активного элемента, а в областях, образующих отдельные «нити». Диаметр этих «нитей»  $a$  меньше общего диаметра светового пучка. Если «нити» взаимно некоррелированы, то  $\alpha_0 \approx 2,44 \cdot \frac{\lambda}{a}$ , если коррелированы – то  $\alpha_0 \sim 2,44 \cdot \frac{\lambda}{d}$ , то есть эквивалентны множеству синхронно излучающих антенн.

При  $d = 1\text{см}$  и  $\lambda = 1\text{мкм}$   $\theta_0 = 0.12$  мрад. Расходимость луча современных газовых лазеров составляет десятки угловых секунд – десятые доли градусов, полупроводниковых – единицы градусов. Высокая направленность излучения лазера обеспечивает большую дальность и скрытность передачи информации. Значительно улучшает помехозащищенность оптических систем связи при использовании приемных устройств с малым углом.

**Яркость.** Яркость какого-либо источника электромагнитной волны характеризует мощность излучения, испускаемого с единицы поверхности в единичном телесном угле в направлении, перпендикулярном излучательной поверхности. Эта величина имеет размерность  $\frac{Вт}{м^2 \cdot ср}$  и называется энергетической яркостью или излучательной способностью.

$$B_{E_\lambda} = \frac{P}{S \cdot \Omega},$$

где  $P$  – мощность излучения;  $S$  – площадь излучаемой поверхности;  $\Omega$  – телесный угол (в стерадианах).

$B_{E_\lambda}$  является объективной физической характеристикой электромагнитного излучения.

Кроме энергетической яркости вводится понятие фотометрической яркости  $B_{v_\lambda}$  или просто яркости. Фотометрическая яркость служит для оценки эффективного воздействия света на глаз человека. Она определяется многими факторами и может быть различной у разных людей, зависит от уровня освещенности (дневное и ночное зрение), психологического состояния и т.д. Но наиболее сильной является ее зависимость от длины волны. Эта зависимость, полученная путем усреднения многочисленных данных, была утверждена международной комиссией по освещению (МКО) в качестве стандартной кривой видности (рис. 8.16) [11].

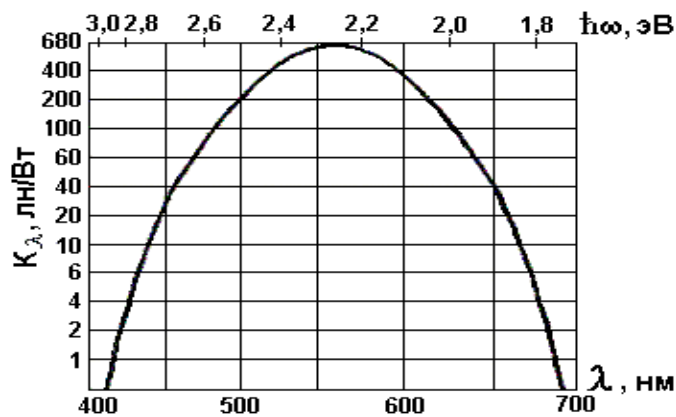


Рис. 8.16. Кривая видности для дневного зрения

Эта кривая выражает чувствительность глаза к свету с различными длинами волн. Таким образом, переход от энергетических величин к фотометрическим осуществляется через коэффициент  $K_\lambda$ , зависящий от длины волны. Этот коэффициент является световым (фотометрическим) эквивалентом потока излучения (энергетического) и называется спектральной световой эффективностью монохроматического излучения или видностью.

Для данной длины волны эта связь определяется соотношением

$$B_{v_\lambda} = K_\lambda \cdot B_{E_\lambda},$$

для немонахроматического света

$$B_{v_\lambda} = \int B_{v_\lambda} \cdot d\lambda = \int K_\lambda \cdot B_{E_\lambda} \cdot d\lambda.$$

Полный световой (фотометрический) поток источника света связан со спектральной плотностью потока излучения соотношением

$$\Phi_\lambda = \int K_\lambda \cdot \Phi_\lambda \cdot d\lambda.$$

Для нормального дневного зрения максимум функции видности приходится на  $\lambda = 555\text{ нм}$  (зеленый свет). На этой длине волны (рис.16) мощность излучения 1 Вт эквивалентна световому потоку 680 Лм, то есть  $K_{\lambda=555\text{ нм}} = K_{max} = 680 \text{ Лм/Вт}$ . При

$\lambda_{min} = 380 \text{ нм}$  и  $\lambda_{max} = 780 \text{ нм} \Rightarrow K_{\lambda} \rightarrow 0$  (это фиолетовая и красная границы видимой области света).

Наряду с яркостью для описания энергетических характеристик лазерного излучения часто используют такие понятия как мощность излучения (непрерывная, импульсная) и энергии излучения, которая характеризует энергию, переносимую световым потоком, и применима обычно к импульсным лазерам. Энергия излучения определяет энергию электромагнитного поля в одном лазерном импульсе. Но энергетическая яркость лазерного излучения часто является более важной характеристикой, чем мощность, энергия и интенсивность излучения. Это связано с тем, что во многих практических случаях интерес представляет интенсивность, которую можно получить при фокусировке лазерного излучения линзой. А эта величина пропорциональна яркости пучка.

Существенным является то, что хотя интенсивность пучка можно увеличить (за счет фокусировки), его яркость при этом остается неизменной. Если, например, уменьшить диаметр лазерного пучка от  $d_1$  до  $d_2$  с помощью линз, фокусы которых совпадают (при этом  $\frac{d_2}{d_1} = \frac{f_2}{f_1}$ ), то интенсивность пучка увеличится, но расходимость выходного пучка ( $\sim \frac{\lambda}{d_2}$ ) будет больше расходимости входного ( $\sim \frac{\lambda}{d_1}$ ), так что яркость останется неизменной (если пренебречь потерями) [5,11,12,13,16].

**Характеристики излучения лазера,** Основными техническими характеристиками лазера, определяющими возможности использования их в конкретных системах, являются: [2]

1) рывном режиме.

Мощность излучения твердотельных и жидкостных лазеров в непрерывном режиме невелика и составляет примерно 10 милливатт, а в импульсном режиме больше 100 ГВт при  $\tau_n$  равным не **Длина волны** или частота излучаемых электромагнитных колебаний. Современные лазеры генерируют в диапазоне частот  $1,3 \cdot 10^{15} \div 5,57 \cdot 10^{11}$  Гц. Использование колебаний высокой частоты позволяет применять антенны сравнительно малых размеров и передавать информацию в широкой полосе частот (до  $10^{12}$  Гц). Именно поэтому информационная емкость оптических систем связи может быть, на несколько порядков выше емкости радиотехнических систем.

2) **Мощность излучения** в непрерывном или импульсном режиме. Импульсный режим обычно характеризуется длительностью импульса  $\tau_n$ , частотой следования импульсов  $F_n$ , средней мощностью  $P_{cp} = P_u \tau_n F_n$  и энергией излучения в импульсе  $W_u = P_u \tau_n$ . Мощность излучения большинства газовых лазеров в непрерывном режиме составляет единицы - десятки милливатт, а в импульсном режиме - сотни ватт - единицы киловатт. В последнее время мощность излучения газовых лазеров достигает десятков киловатт и в не несколько наносекунд.

Мощность, излучаемая полупроводниковыми лазерами колеблется в пределах сотни милливатт - единицы ватт в непрерывном режиме и от 10 Вт до 100 кВт в импульсном режиме. Достигнутая энергия излучения в импульсе у твердотельных и жидкостных лазеров больше 1000 Дж. Столь значительная энергия концентрируется обычно в небольшом объеме, что позволяет использовать лазер в сис-

темах связи с чрезвычайно удаленными объектами, в установках для сварки самых тугоплавких материалов и для других целей. Так, если сфокусировать излучение мощностью  $10^9$  Вт в пятно диаметром  $10^{-2}$  мм, то плотность мощности составит  $10^{15}$  Вт/см<sup>2</sup>. При воздействии мощности такой плотности любой материал немедленно испаряется.

**3) Коэффициент полезного действия.** Коэффициент полезного действия лазера, как и любого другого генератора определяется отношением излучаемой мощности к мощности накачки  $\eta = P_{\text{изл}} / P_{\text{накачк}}$  (выражается в процентах).

$\eta$  – большинства газовых, твердотельных, жидкостных лазеров составляет доли процента. Однако, лазер на CO<sub>2</sub> обладает достаточно большим  $\eta \approx 30-40\%$ , полупроводниковые лазеры могут обеспечить  $\eta > 50\%$  [2].



## ГЛАВА 9.

### ОПТИЧЕСКИЕ ГЕНЕРАТОРЫ НА ТВЕРДОМ ТЕЛЕ

#### 9.1. Устройство твердотельных генераторов

Современный лазер, в котором в качестве активного вещества используется твердое тело состоит из активного элемента 1 (см. рис. 9.1.), системы накачки 2, резонатора 3, системы охлаждения 4.

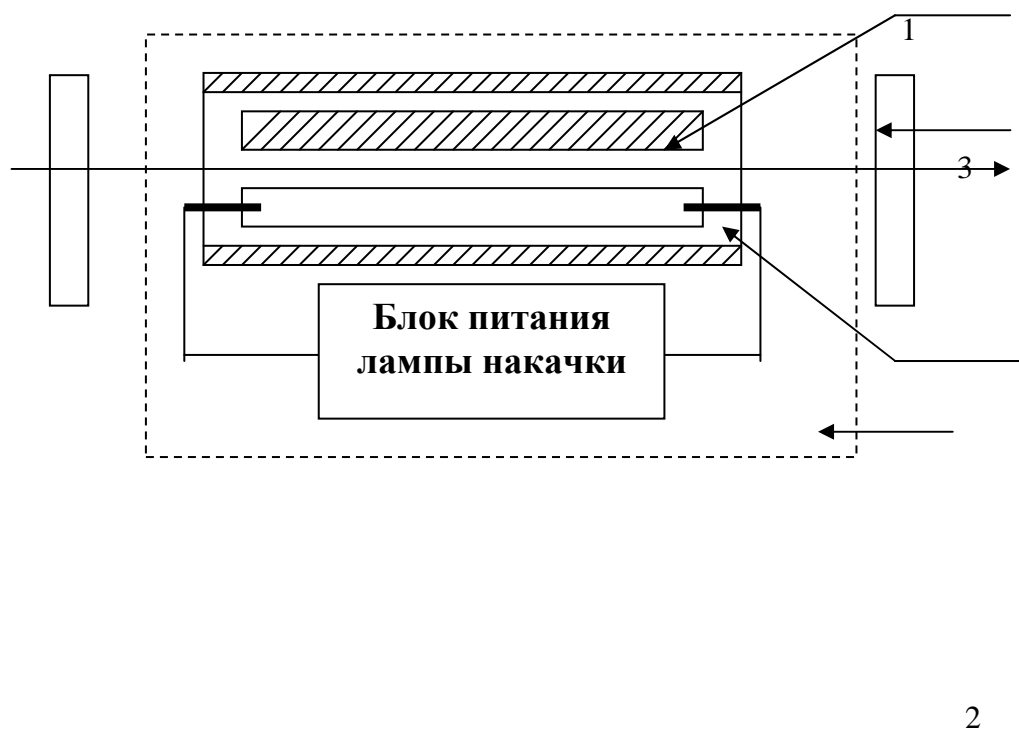


Рис. 9.1. Схема оптического квантового генератора на твердом теле

Рассмотрим наиболее характерные свойства основных элементов твердотельного лазера.

**Активный элемент.** В первых твердотельных лазерах, в качестве активного элемента, применялись кристаллы синтетического рубина.  $Al_2O_3$  с примесью трехвалентных ионов  $Cr^{3+}$ . Введение ионов  $Cr^{3+}$  в рубин придает ему красный цвет. Содержание ионов хрома в рубине обычно не велико (0,05%).

В кристалле на атомы  $Cr$  действует сильное электрическое поле, создаваемые атомами  $O$  и  $Al$ . Вклад самих атомов  $Cr$  в это поле мал, так как мала их концентрация. Поэтому взаимодействие атомов  $Cr$  друг с другом слабо. Их энергетический спектр соответствует спектру свободного атома  $Cr$ , помещенного в сильное электрическое поле кристалла, расщепляющее уровни атома (эффект Штарка) на ряд уровней  ${}^4A_2$ ,  ${}^4F_1$ ,  ${}^4F_2$ . (рис. 9.4).

В настоящее время рубиновые лазеры вытесняются неодимовыми.

Генерационные возможности материалов, используемых в качестве активных сред, в большой степени зависят от способности этих материалов поглощать подводимую энергию накачки. Чем больше поглощается энергии накачки в нужной

полосе спектра, тем большую энергию излучения можно получить, тем более высоким будет к.п.д. активной среды.

Важнейшим свойством, характеризующим активные элементы лазеров, является способность их к длительной работе. Наиболее долговечным является гранат, который может обеспечить до миллиона и более вспышек, в то время как рубин в режиме модулированной добротности выдерживает всего несколько десятков тысяч вспышек.

**Элементы системы накачки.** В системе накачки твердотельного лазера можно выделить три основных элемента:

- а) Лампа накачки, световое излучение которой обеспечивает создание инверсии населенности рабочих уровней активного вещества;
- б) Осветительная система, с помощью которой излучение лампы с минимальными потерями передается в лазерный кристалл;
- в) Электрическая схема блока питания лампы накачки.

Рассмотрим кратко характеристики практически используемых элементов лампы накачки и осветительной системы.

. К лампам накачки предъявляются следующие требования:

- 1) Они должны обладать высокой эффективностью преобразования электрической энергии в световую;
- 2) Хорошим согласованием спектра излучения со спектром поглощения активного элемента;
- 3) Важнейшими параметрами лампы накачки являются: предельная энергия вспышки, которую лампа может выдержать без разрушения оболочки, и постоянная величина, определяемая эмпирическим соотношением [17].

$$(C I_0^4 / l^3)_{\text{пред}} = k,$$

где  $C$  – емкость конденсатора, разряжающегося через лампу;  $U_0$  – рабочее напряжение лампы накачки;  $l$  – длина разрядной трубки лампы накачки;  $k$  – постоянная, зависящая от конструктивных особенностей лампы накачки.

Основными источниками накачки твердотельных лазеров являются газоразрядные лампы, которые имеют максимум излучения в видимой области спектра.

Из газоразрядных ламп более широкое распространение получили импульсные ксеноновые газоразрядные лампы. Эффективность преобразования электрической энергии в световую для таких ламп может достигать 50%. Однако на спектральную область, где расположены линии поглощения активного элемента, например, рубина, приходится менее 30% излучаемой энергии. Остальная энергия является бесполезной и даже вредной, так как она затрачивается на нагревание активного элемента, оболочки лампы накачки и осветителя. Отсюда вытекает неполное соответствие спектра излучения источника накачки спектру поглощения активного тела.

Для накачки рубиновых лазеров, работающих в непрерывном режиме, применяют ртутные капиллярные лампы сверхвысокого давления, спектр излучения которых хорошо согласуется со спектром поглощения рубина. Лампы такого типа характеризуются очень высокой яркостью излучения и в номинальном режиме способны отдавать до 600 Вт/см с единицы длины разрядного промежутка.

Как указывалось выше, использование в качестве источника накачки специальных ламп, имеющих широкий спектр излучения, приводит к тому, что лишь часть этого излучения приходится на полосы поглощения активных элементов.

Остальная часть энергии расходуется на нагрев оболочки лампы накачки, осветителя, активного элемента, что приводит к целому ряду нежелательных явлений (увеличению порога генерации, расходимости пучка излучения и т. д.). Что вызывает необходимость применения охлаждения системы. Более оптимальным является использование узкополосных источников накачки, спектр излучения которых совпадает с полосой поглощения активного элемента.

В настоящее время в качестве таких источников стали применять светодиоды некогерентного излучения на основе тройных соединений типа  $GaAs_{1-x}P_x$ .

**Осветительная система.** Сопоставление спектров поглощения активных элементов и спектров излучения ламп накачки показывает, что коэффициент использования излучения импульсных ламп в лазерах мал. Для обеспечения эффективной передачи энергии вспышки лампы накачки активному элементу используют специальные осветители или отражатели. Внутренняя поверхность отражателя полируется и покрывается отражающим материалом (серебром, золотом). Для исключения ненужного спектра излучения лампы накачки, используются также диэлектрические интерференционные покрытия. Наиболее широкое применение находят осветители на основе эллиптического цилиндра (см. рис. 9.2а).

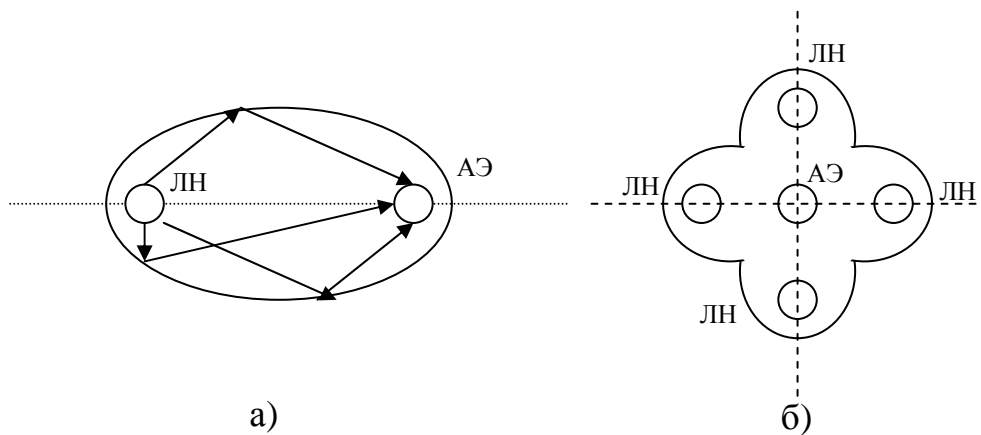


Рис. 9.2. Отражатели эллиптические (а), полиэллиптические (б)

В основе работы этого осветителя лежит свойство эллипса, состоящее в том, что всякий луч, проходящий через одну из фокальных осей эллипса, после отражения от его поверхности, обязательно пройдет через вторую фокальную ось. Таким образом, если активный элемент расположить вдоль одной фокальной оси, а лампу накачки вдоль другой, то большая часть светового потока лампы, отражаясь от поверхности осветителя, собирается в другой фокальной оси, где расположен активный элемент. Еще большая концентрация энергии на элементе получается в полиэллиптическом отражателе (см. рис. 9.2б). Такие отражатели используются, если в качестве источника применяют стержневые лампы. Более полную информацию об осветителях можно найти в [13, 17].

## 9.2. Генераторы на рубине

Рубиновый лазер был исторически первым твердотельным квантовым генератором оптического диапазона волн. Возможность получения высокой энергии генерации с хорошими оптическими, термическими и механическими свойствами кристаллов рубина позволяет и в настоящее время применять лазеры на рубине в некоторых научных исследованиях.

Упрощенная схема энергетических уровней иона  $Cr$  в рубине показана на рис. 9.3.

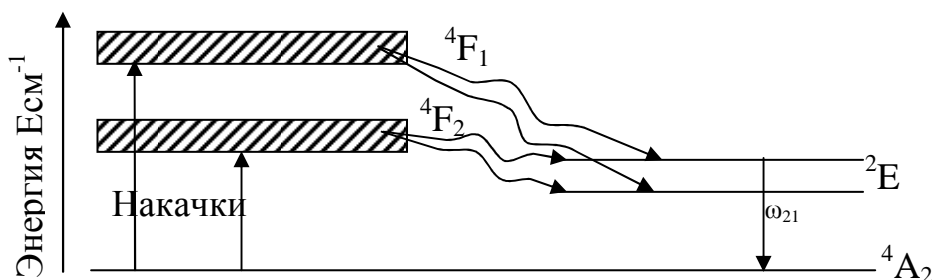


Рис. 9.3. Схема энергетических уровней иона хрома в рубине

Из рисунка видно, что имеется основное состояние  ${}^4A_2$  и возбужденное состояние  ${}^2E$ . Переход между этими уровнями и используется для генерации света. Уровни  ${}^4A_2$  и  ${}^2E$  считаются узкими. Наряду с ними имеются сравнительно широкие полосы энергии  ${}^4F_2$  и  ${}^4F_1$ . Переходы между уровнями полосы  ${}^4F_2$  и уровнем  ${}^4A_2$  лежат в зеленом интервале спектра, а переход  ${}^4F_1 - {}^4A_2$  – в голубом.

Если атом  $Cr$  возбудить, переводя его из основного состояния в полосу  ${}^4F_2$  или  ${}^4F_1$ , то за очень короткое время ( $\sim 10^{-8}$  с) он перейдет из этих полос на уровень  ${}^2E$ . При переходе на уровень  ${}^2E$  атом  $Cr$  не излучает. Его энергия тратится на возбуждение колебаний кристаллической решетки рубина. Возможность возвращения атома из полос  ${}^4F_2$  и  ${}^4F_1$  снова на уровень  ${}^4A_2$  хотя и существует, но скорость этого процесса пренебрежимо мала по сравнению со скоростью перехода атомов на уровень  ${}^2E$ . На уровне  ${}^2E$  атом «живет»  $10^{-3}$  с. Такое большое время жизни позволяет накапливать атомы на уровне  ${}^2E$ . В таком случае населенность уровня  ${}^2E$  окажется больше, чем  ${}^4A_2$  и между этими уровнями возникает инверсия населенности.

Теперь при наличии положительной обратной связи в резонаторе лазера достаточно появления первичных спонтанных переходов иона хрома с одного из подуровней в основное состояние, например с  ${}^2E$  на  ${}^4A_2$ , чтобы возникло индуцированное излучение с длиной волны  $\lambda = 0,6943$  мкм (что соответствует красному свету).

Рубиновый твердотельный генератор относится к трехуровневым генераторам. Наличие широких полос в спектре ионов  $Cr^{3+}$  в рубине является «Третьим уровнем» – благоприятным обстоятельством для создания инверсии.

При использовании в качестве активных частиц ионов редкоземельных элементов (неодима, диспрозия, эрбия) инверсия населенностей создается по четырехуровневой схеме [4,5,13,15,17].

### 9.3. Неодимовый лазер

Неодимовые лазеры являются самыми популярными из твердотельных лазеров. В этих лазерах активной средой обычно является кристалл  $Y_3Al_5O_{12}$  [сокращенно называемый *YAG* (yttrium aluminum garnet), иттрий-алюминиевый гранат], в котором часть ионов  $Y^{3+}$  замещена ионами  $Nd^{3+}$ , либо просто стекло активированное ионами неодима  $Nd^{3+}$ . Неодимовые лазеры могут генерировать на нескольких линиях; самая сильная из них, а следовательно, и наиболее широко используемая – это линия с длиной волны равной 1,06 мкм.

На рис.9.4 представлена упрощенная схема энергетических уровней *Nd:YAG*.

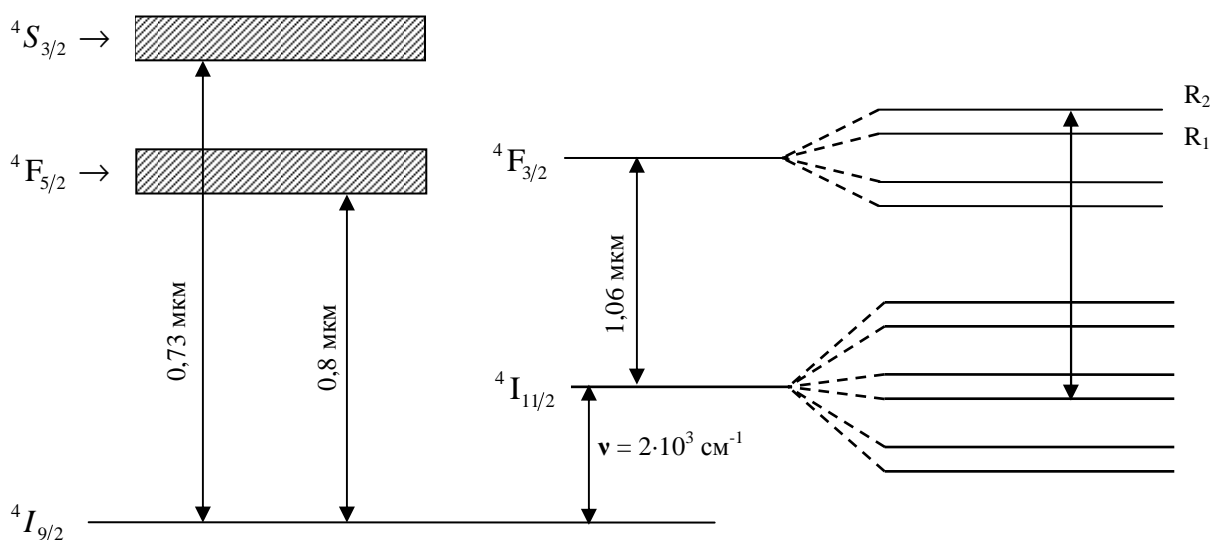


Рис. 9.4. Упрощенная схема энергетических уровней *Nd:YAG*

Эта схема во многом похожа на схему уровней ионов неодима в стекле, так как кристаллическое поле оказывает незначительное влияние на рассматриваемые энергетические уровни. Лазерный переход с  $\lambda=1,06$  мкм является наиболее сильным из переходов  ${}^4F_{3/2} \rightarrow {}^4I_{11/2}$ . Две основные полосы накачки расположены на длинах волн 0,73 и 0,8 мкм соответственно. Эти полосы связаны посредством быстрой безызлучательной релаксации с уровнем  ${}^4F_{3/2}$ , а нижний уровень  ${}^4I_{11/2}$  связан также быстрой безызлучательной релаксацией с основным состоянием  ${}^4I_{9/2}$ . Кроме того, разница между энергиями уровней  ${}^4I_{11/2}$  и  ${}^4I_{9/2}$  почти на порядок величины больше, чем  $kT$ . Из этого следует, что неодимовый лазер работает по четырехуровневой схеме. Как и в случае рубина, лазерный переход в неодимовом лазере (преимущественно) является однородно уширенным, и соответствующая ширина составляет  $\Delta\nu_0 = 6,5 \text{ см}^{-1} = 195 \text{ ГГц}$  при температуре  $T = 300^0 \text{ К}$ . В последнем случае время жизни верхнего лазерного уровня также велико ( $\tau = 0,23 \text{ мс}$ ) вследствие того, что этот переход электродипольно запрещен.

*Nd:YAG*-лазер может работать как в непрерывном, так и в импульсном режиме. В обоих случаях обычно используются линейные лампы в схемах с одноэллипсным осветителем. Для работы в импульсном и непрерывном режимах применяются ксеноновые лампы среднего давления (4-6 атм.) соответственно. Размеры стержней обычно такие же, как и в случае рубина. Выходные параметры *Nd:YAG*-лазера могут быть сведены к следующим:

1) в непрерывном режиме выходная мощность без усилителя до 150 Вт и после одного каскада усиления до 700 Вт; 2) в режиме модулированной добротности выходная мощность до 50 МВт; 3) в режиме синхронизации мод длительность импульса до ~20 пс. Как в импульсном, так и в непрерывном режиме тангенциальная эффективность составляет около 1÷3%. *Nd:YAG*-лазеры широко используются в различных областях, таких, например, как обработка материалов, измерение расстояний и лазерная хирургия. Размеры стержня из стекла с  $Nd^{3+}$  могут значительно превышать размеры стержня из *Nd:YAG* (и могут достигать 1 м в длину и нескольких десятков см в диаметре). Благодаря значительно более низкой температуре плавления стекло получить гораздо легче, чем кристаллы *YAG*. Лазер на стекле с неодимом можно применять вместо *Nd:YAG*-лазера во всех экспериментах, где требуется достаточно низкая частота следования импульсов, при которой не возникало бы проблем, связанных с низкой теплопроводностью стержня. Очень важное применение стекол с неодимом – это применение их в качестве усилителей в лазерных системах для получения очень высокой энергии в экспериментах по лазерному термоядерному синтезу. В настоящее время уже создана лазерная система на основе стекла с неодимом, дающая импульсы с пиковой мощностью более 20 ТВт и полной энергией ~15 кДж и более [12,15,17].

#### 9.4. Режимы работы твердотельных лазеров. Основные уравнения

Для работы твердотельных лазеров характерны следующие режимы работы:

- 1) Непрерывный режим генерации.
- 2) Режим свободной генерации (в простом резонаторе при накачке от лампы - вспышки). При этом генерируется световой импульс длительностью несколько сотен микросекунд.
- 3) Режим с модуляцией добротности. Здесь лазер генерирует импульсы длительностью от единиц до десятков наносекунд. Мощность импульса достигает до сотен МВт.
- 4) Режим синхронизации мод. В этом случае генерируется целая последовательность импульсов. Длительность каждого импульса достигает 10 пикосекунд. Частота повторения этих импульсов порядка 100 МГц. Мощность этих импульсов (пиков) велика.

Используя кинетические уравнения (6.3), приведем их к виду удобному для описания режимов работы твердотельных лазеров на примере генератора, работающего по трехуровневой схеме накачки (рис. 6.3).

Уровень 1 есть основное состояние, уровень 2 является метастабильным и выше расположен уровень 3, через который осуществляется накачка метастабильного уровня. В соответствии с реальной ситуацией примем, что уровень 3-короткоживущий, так что  $N_3 = 0$ . Кроме того, будем считать, что квантовый выход

равен 1, так что любая частица, заброшенная на уровень 3, обязательно совершит переход на метастабильный уровень и далее излучательный переход в состояние 1. Считаем, что безызлучательные переходы 2-1 малы. Примем для простоты, что  $g_1=g_2=1$  в генерации участвует только один тип колебаний. Тогда число фотонов в нем равно  $m$  (число фотонов в поле резонатора лазера).  $N_1, N_2, N_3$  – населенности уровней лазера. Но нас будет интересовать разность населенностей, обозначим ее через  $n = N_2 - N_1$ . В результате балансные уравнения примут вид [4]

$$\begin{aligned} \frac{dN_1}{dt} &= bmn + \frac{N_2}{\tau} - WN_1, \\ \frac{dN_2}{dt} &= bmn - \frac{N_2}{\tau} + WN_1. \end{aligned} \quad (9.1)$$

В рассматриваемом приближении сумма  $N_2+N_1=N$  – есть постоянная величина,  $-WN_1$  – характеризует скорость опустошения уровня 1 за счет действия поля накачки (с такой же скоростью будет заполняться уровень 2),  $b$  – спектральный коэффициент Эйнштейна (см. подраздел 5.2.2). Вычитая из второго уравнения системы (9.1) первое и выражая  $N_2, N_1$  через  $n$  и  $N$ , получим

$$\frac{dn}{dt} = -2bmn - (W + \frac{1}{\tau})n + (W - \frac{1}{\tau})N. \quad (9.2)$$

К этому уравнению следует добавить соотношение, определяющее число фотонов  $m$ . Вычислим скорость изменения  $m$  равную  $dm/dt$ .

Если бы процессы индуцированного испускания и поглощения фотонов отсутствовали, то со временем поле в резонаторе уменьшалось бы за счет потерь резонатора. Обозначим через  $\tau_p$  – время затухания поля в резонаторе, тогда  $\tau_p = Q/\omega$  ( $Q$ -добротность резонатора).

Следовательно:

$$\frac{dm}{dt} = -\frac{m}{\tau_p}.$$

Учет индуцированных переходов может быть проведен очень просто. Изменение разности населенностей в единицу времени за счет индуцированных переходов равно, согласно (9.2),  $2bmn$ . При каждом переходе число фотонов в резонаторе изменяется на  $\pm 1$ , а разность населенностей на  $\pm 2$ . Поэтому изменение числа фотонов за счет индуцированных переходов просто в два раза меньше изменения разности населенностей и противоположно ему по знаку, т.е.

$$\left(\frac{dm}{dt}\right)_{\text{инд}} = bmn.$$

В результате полная скорость изменения числа фотонов в резонаторе будет:

$$\frac{dm}{dt} = bmn - \frac{m}{\tau_p}. \quad (9.3)$$

Уравнения (9.2), (9.3) образуют замкнутую систему, описывающую (при сделанных допущениях) работу лазера.

Аналогичным образом могут быть записаны уравнения для 4-х уровневой схемы работы лазера. Наиболее простой вид они будут иметь в том случае, когда нижний рабочий уровень короткоживущий и его можно принять пустым. В этом

случае разность населенностей  $n$  равна населенности метастабильного уровня  $N_3$  и для  $n$  легко получить

$$\frac{dn}{dt} = -bmn - \frac{n}{\tau} + W(N - n). \quad (9.4)$$

Здесь под  $(N - n)$  понимается населенность основного состояния. Уравнение для числа фотонов остается тем же что и в 3-х уровневой схеме.

### 9.4.1. Режим непрерывной генерации твердотельного лазера

Рассмотрим теперь стационарные решения полученных уравнений, когда ни число фотонов, ни разность населенностей не зависят от времени. Этот случай, строго говоря, может быть реализован в непрерывном лазере или же в импульсном – длительность импульса накачки в котором достаточно велика. Для определенности остановимся на 3-х уровневом лазере. Уравнения (9.2) и (9.3) примут вид:

$$\begin{aligned} m(bn - \frac{1}{\tau_p}) &= 0, \\ -2bmn - (W + \frac{1}{\tau})n + (W - \frac{1}{\tau})N &= 0 \end{aligned} \quad (9.5)$$

Из первого уравнения системы (9.5) следует, что возможно два решения этой системы:

1) при  $m=0$ ; из второго уравнения (9.5) получаем

$$n = \frac{W - \frac{1}{\tau}}{W + \frac{1}{\tau}} N. \quad (9.6)$$

В этом режиме число фотонов  $m$  в резонаторе лазера равно 0. При этом, с ростом уровня накачки, пропорционального  $W$  величина разности населенностей  $n$  монотонно возрастает. При  $W \rightarrow \infty$  величина  $n = N$ , что означает полную перекачку частиц на уровень 2. Этот режим, очевидно, соответствует состоянию отсутствия генерации.

2) при  $m \neq 0$ ;

$$\begin{aligned} n &= \frac{1}{b\tau_p}, \\ m &= \frac{N_{2nop}}{\tau} \tau_p (\alpha - 1), \end{aligned} \quad (9.7)$$

где приняты обозначения:

$$\begin{aligned} N_{2nop} &= \frac{1}{2} (N + \frac{1}{b} \tau_p), \\ \alpha &= \frac{W}{W_{nop}}; \\ W_{nop} &= \frac{1}{\tau} \cdot \frac{N + \frac{1}{b} \tau_p}{N - \frac{1}{b} \tau_p}. \end{aligned} \quad (9.8)$$

Отсюда следует, что когда  $\alpha > 1$ , то есть  $W > W_{nop}$ , становится возможным существование режима 2. (при  $\alpha < 1$ ,  $m < 0$ , что не имеет физического смысла).



При  $m > 0$  лазер генерирует. Разность населенностей  $n$  в этом режиме не зависит от уровня накачки, тогда как выходная мощность лазера (пропорциональная  $m$ ) линейно возрастает с ростом  $W$

В 3-х уровневом лазере порог уровня накачки определяется двумя факторами: 1) временем жизни метастабильного уровня; 2) потерями резонатора.

Анализ выражений (9.3), (9.4) для четырехуровневой системы проводится аналогично. В 4-х уровневом лазере  $W_{nop}$  ограничена в основном потерями в резонаторе. При достижении  $W_{nop}$ ,  $n$  прекращает расти. Оценим  $P_{вых}$  генератора непрерывного режима.

Полная мощность рассеиваемая, в резонаторе, и излучаемая через зеркало есть  $\frac{\hbar\omega \cdot m}{\tau_p}$ . Только часть этой мощности используется, и именно та, которая теряется на излучение через зеркало. Если через  $\tau_c$  обозначить часть времени жизни фотонов в резонаторе, обусловленную потерями излучения за счет прохождения его через зеркало, а  $\tau_0$  – время жизни, обусловленное всеми прочими видами потерь, то  $1/\tau_p = 1/\tau_c + 1/\tau_0$ .

Полезная мощность, развиваемая лазером, равна  $\frac{\hbar\omega \cdot m}{\tau_c}$ . В результате, для 3-х уровневого генератора:

$$P_{вых} = \frac{N_{2nop}}{\tau} (\alpha - 1) \frac{\tau_p}{\tau_c} \hbar\omega.$$

Оценки показывают, что для рубинового лазера длиной активного элемента  $\ell = 2 \div 5$  см и диаметром  $d = 2 \div 3$  мм в генерации участвует только часть этого объема. Примем диаметр генерирующей области  $d = 1$  мм, а  $l = 4$  см. Тогда, используя полученное уравнение, получим  $P_{вых} \approx 10$  Вт [4,10,19].

#### 9.4.2. Пички и режим свободной генерации

Первые же экспериментальные исследования работы твердотельных лазеров выявили существенную нестационарность их излучений. Даже при непрерывном питании лампы накачки излучение имеет вид импульсов, амплитуда которых и расстояние между которыми могут изменяться во времени. Эти импульсы называют пичками. Для рубиновых лазеров длительность пичков  $\sim 1$  мкс, период  $\sim 10$  мкс. Характер пичковой генерации зависит от свойств резонатора и уровня накачки. В резонаторах с плоскими зеркалами пички, как правило, носят нерегулярный характер. В резонаторах со сферическими зеркалами последовательность пичков может быть регулярной, незатухающей или затухающей к постоянному уровню.

Анализ пичкового режима является сложной математической задачей [4]. Это объясняется тем, что в генерации могут принимать участие много типов колебаний, сильно взаимодействующих друг с другом, однако, физические причины образования пичков просты и могут быть поняты на основе следующих соображений.

Пусть в начальный момент времени  $t = 0$  в лазере включается сильное поле накачки, уровень которой будем предполагать постоянным. Под действием этого поля частицы перекачиваются на верхний метастабильный уровень. В некоторый момент времени  $t = t_1$ , разность населенностей  $n$  становится равной пороговой

$n_{\text{пор}}$ . При  $t > t_1$ , разность населенностей становится больше пороговой  $n > n_{\text{пор}}$ , и в результате возникает индуцированное поле, амплитуда которого быстро нарастает (рис 9.5.).

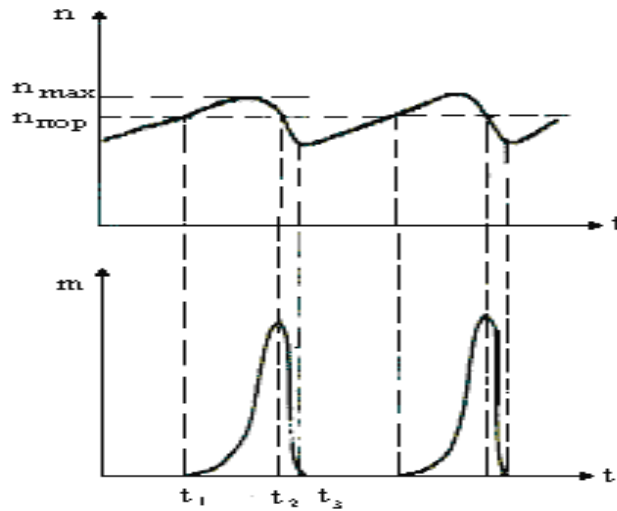


Рис. 9.5. Зависимость разности населенностей  $n$  и числа фотонов  $m$  в резонаторе лазера от времени

Возникновение поля и появление индуцированных переходов приводит к опустошению верхнего уровня. Рост  $n$  замедляется, затем прекращается и далее  $n$  начинает падать, достигая при  $t = t_2$  порогового значения. При  $t > t_2$ ,  $n < n_{\text{пор}}$  усиление в резонаторе меньше потерь и поле начинает падать. К моменту времени  $t = t_3$  поле исчезает, и  $n$  оказывается меньше порогового значения. Под действием накачки  $n$  снова начинает увеличиваться и процесс повторяется.

В тех случаях, когда пички не являются столь глубокими, как показано на рис.9.5, а представляют собой сравнительно небольшие колебания  $m$  относительно стационарного значения числа фотонов, их можно описать аналитически. Рассмотрим случай одномодовой генерации трехуровневого лазера. Примем, что

$$\begin{aligned} m &= m_{cm} + x(t) \\ n &= n_{cm} + y(t) \end{aligned} \quad (9.8a)$$

где  $m_{cm}$  и  $n_{cm}$  – стационарные значения  $m$  и  $n$ , а  $x$  и  $y$  малые отклонения  $m$  и  $n$  относительно стационарных значений. Используя уравнения (9.2), (9.3) и (9.7) из предыдущего параграфа для  $x$  и  $y$  получим линейные дифференциальные уравнения

$$\begin{aligned} \frac{dx}{dt} &= b \cdot m_{cm} \cdot y, \\ \frac{dy}{dt} &= -2bn_{cm} \cdot x - \left(W - \frac{1}{\tau}\right) \frac{N}{n_{cm}} \cdot y. \end{aligned} \quad (9.9)$$

При этом произведения малых величин  $x$  и  $y$  в (9.9) опущены.

Уравнения (9.9) легко решаются. Исключим из (9.9)  $x$ . Тогда для  $y$  будет

$$\frac{d^2 y}{dt^2} + \frac{N}{n_{cm}} \left(W - \frac{1}{\tau}\right) \frac{dy}{dt} + 2b^2 n_{cm} m_{cm} y = 0. \quad (9.10)$$

Это уравнение колебаний маятника в среде с трением, причем коэффициент трения пропорционален величине  $\frac{N}{n_{cm}}(W - \frac{1}{\tau})$ . Если он мал, то частота колебаний  $\omega$ , очевидно, будет

$$\Omega_0^2 = 2bn_{cm}m_{cm}, \quad (9.11)$$

а затухание колебаний  $1/T_0$  ( $T_0$  – время, в течение которого амплитуда колебаний уменьшается в  $e$  раз) равно

$$\frac{1}{T_0} = \frac{1}{2} \frac{N}{n_{cm}} (W - \frac{1}{\tau}). \quad (9.12)$$

Более точно  $1/T_0$  частота колебаний определяется соотношением

$$\Omega = \sqrt{\Omega_0^2 - \frac{1}{T_0^2}}.$$

Если  $T_0$  очень мало, так, что  $\frac{1}{T_0} > \Omega_0$ , то закон изменения  $y$  во времени становится аperiодическим. Обычно для лазера  $\frac{1}{T_0} \ll \Omega_0$ , и  $y(t)$  как и  $x(t)$  есть затухающие колебания (пички) излучения.

Подставляя в (9.8а) явные выражения для  $n_{cm}$  и  $m_{cm}$ , получим

$$\Omega^2 = \frac{2bN_{2пор}}{\tau} (\alpha - 1). \quad (9.13)$$

Выразим входящую сюда величину « $b$ » через параметры активного элемента и резонатора. Для этого учтем, что  $bn$  есть, как следует из уравнения (9.3) предыдущего параграфа, коэффициент усиления поля в резонаторе в единицу времени. Пусть  $n=N$ , то есть все частицы перекачены на верхний рабочий уровень. Величина  $bN$  дает предельный коэффициент усиления, который можно достичь в данном лазере. Зная это, можно подсчитать предельный коэффициент усиления на единицу длины резонатора  $\alpha_{пред}$ . Если обозначить через  $v$  среднюю скорость распространения волны в резонаторе лазера, то, очевидно,

$$\alpha_{пред} = \frac{bN}{v}. \quad (9.14)$$

Если активный элемент заполняет резонатор полностью, то  $\alpha_{пред}$  определяется концентрацией и спектроскопическими характеристиками ионов, участвующих в усилении поля. Обозначим  $\alpha_{пред}$  для этого случая через  $\chi$ . Для рубина, например,  $\chi = 0.4 \text{ см}^{-1}$  для 0.05% концентрации  $\text{Cr}^{3+}$  и при комнатной температуре. Если длина активного элемента  $l$  меньше длины резонатора  $L$ , то, очевидно  $\alpha_{пред}$  будет во столько раз меньше  $\chi$ , во сколько раз  $l$  меньше  $L$ . То есть

$$\alpha_{пред} = \chi \frac{l}{L}. \quad (9.15)$$

Из (9.14) и (9.15) находим

$$b = \frac{l}{L} \cdot \frac{v}{N}, \quad (9.16)$$

подставляя это значение в (9.13), получим следующее выражение для частоты пичков:

$$\Omega^2 = \frac{2\chi\vartheta}{\tau} \frac{N_{2nop}}{N} (\alpha - 1) \frac{l}{L}. \quad (9.17)$$

В случае малых потерь резонатора  $N_{2nop} \approx \frac{1}{2}N$  и (9.17) упрощается

$$\Omega^2 = \frac{\chi\vartheta}{\tau} \frac{l}{L} (\alpha - 1). \quad (9.18)$$

Период следования пичков  $t_0 = \frac{2\pi}{\Omega}$  равен

$$t_0 = 2\pi \sqrt{\frac{\tau}{\chi\vartheta} \frac{L}{l} \cdot \frac{1}{\alpha - 1}}. \quad (9.19)$$

Из (9.19) следует, что для рубинового лазера (при следующих параметрах:  $\alpha = 2$ ,  $L = 1\text{ м}$ ,  $l = 10\text{ см}$ ,  $\tau = 10^{-3}\text{ с}$ ,  $\chi = 0.4\text{ см}^{-1}$ ,  $\vartheta = 3 \cdot 10^{10}\text{ см./с}$ ) период следования пичков  $t_0 = 10\text{ мкс}$ , что по порядку величины соответствует экспериментальным данным.

Аналогичным образом можно оценить время затухания пичков. Оно имеет порядок  $10^{-3}\text{ с}$  и уменьшается с ростом накачки.

Пичковая генерация типична для подавляющего большинства лазеров, работающих в режиме свободной генерации. Этот режим реализуется при разряде батарей конденсаторов блока питания через лампу накачки. Время, в течение которого происходит разряд, составляет по порядку величины  $10^{-3}\text{ с}$ . Время затухания пичков в лазере имеет тот же порядок величины. Поэтому выход на стационарный режим в таких лазерах (т.е. регулярные пички), как правило, не происходит. Более того, в ряде случаев такой выход невозможен при любом увеличении длительности импульса накачки. Наличие пичков, а также хаотичность их во времени, является недостатком лазеров и препятствует ряду их применений.

Приведём некоторые цифры, характеризующие режим свободной генерации. Энергия излучения в импульсе колеблется от сотых долей джоуля до тысяч джоулей. Длительность импульса составляет  $10^{-3}\text{ с}$ , и, соответственно импульсная мощность достигает десятков и сотен кВт, к.п.д. – может достигнуть нескольких процентов.

### 9.4.3. Лазер с модуляцией добротности

В режиме свободной генерации излучение твердотельных лазеров представляет собой импульс общей длительностью порядка 1 мс. Этот импульс состоит из множества пичков. Для ряда приложений лазеров таких как дальнометрия, оптическая локация, желательнее иметь короткий и мощный моноимпульс. Получение моноимпульсов излучения длительностью 10 – 100 нс. и энергией в импульсе порядка 100 МВт возможно в лазере с модуляцией добротности. Принцип действия такого лазера состоит в следующем.

В режиме свободной генерации излучение света начинается в тот момент, когда разность населенностей рабочих уровней равна пороговому значению. Максимум излучения приходится на момент времени, когда эта разность вновь равна пороговой. В промежутке между этими моментами времени разность населенностей достигает максимального значения (см. рис. 9.7.) Очевидно, что энергия, излученная в «пичке», определяется тем, насколько разность населенностей «успела» превысить свое пороговое значение. Энергия, запасенная избыточным чис-

лом атомов верхнего уровня равна  $(N_2 - N_{2\text{пор}})_{\text{max}} \hbar \omega$ . Обычно величина  $(N_2 - N_{2\text{пор}})_{\text{max}}$  мала и энергия, излученная в одном «пичке», также невелика. Для увеличения ее уменьшают добротность резонатора до малой величины каким-либо способом. В результате  $N_{2\text{пор}}$  становится очень большим и генерация не возникает. В некоторый момент времени  $t = t_0$  потери, внесенные ранее в резонатор «выключаются» так, что добротность резонатора резко возрастает. Разность населенностей достигает большой величины. В результате генерируется импульс света малой длительности и достаточно высокой энергии.

Проведем оценку импульса излучаемого лазером с модуляцией добротности. Используем уравнения для  $n$  и  $m$ , полученные выше: допустим, что добротность модулятора включается мгновенно в момент времени  $t = 0$ . Время изменения  $m$  и  $n$  и время длительности импульса порядка  $10^{-8}$  с. Будем полагать, что во время генерации импульса вероятности индуцированных переходов намного превышают вероятности спонтанных переходов и переходов за счет накачки. Тогда изменением  $n$  за счет  $1/\tau$  можно пренебречь и можно пренебречь изменением  $n$  за счет накачки, так что уравнение (9.3) не изменится, а (9.2) примет вид

$$\frac{dn}{dt} = -2bmn \quad (9.20)$$

Поделив уравнение (9.3) на (9.20), получим

$$\frac{dm}{dn} = -\frac{1}{2} + \frac{1}{2b\tau_p n}. \quad (9.21)$$

Очевидно, переменные  $m$  и  $n$  в (9.21) легко разделяются, и интегрирование (9.21) дает следующее соотношение

$$m = m_{\text{нач}} + \frac{1}{2} \left[ (n_{\text{нач}} - n) + n_{\text{пор}} \ln \frac{n}{n_{\text{нач}}} \right], \quad (9.22)$$

где введено обозначение  $n_{\text{пор}} = \frac{1}{b\tau_p}$ ,  $m_{\text{нач}}$  и  $n_{\text{нач}}$  – начальные значения числа фотонов в резонаторе и разности населенностей соответственно.

Найдем максимальное значение числа фотонов в резонаторе. При этом  $\frac{dm}{dt} = 0$  и, как видно из (9.3),  $n = n_{\text{пор}}$ . Поэтому

$$m_{\text{max}} = m_{\text{нач}} + \frac{1}{2} \left[ (n_{\text{нач}} - n_{\text{пор}}) + n_{\text{пор}} \ln \frac{n_{\text{пор}}}{n_{\text{нач}}} \right]. \quad (9.23)$$

Для оценки ограничимся случаем, когда  $n_{\text{нач}} \gg n_{\text{пор}}$ . Тогда из (9.23) приближенно имеем

$$m_{\text{нач}} \approx \frac{n_{\text{нач}}}{2}. \quad (9.24)$$

Пусть потери в резонаторе связаны только с выходом излучения через второе зеркало, коэффициент отражения которого  $R$ . Тогда число фотонов, вышедших из резонатора за время двойного прохода резонатора, можно оценить как

$$\Delta m = m(1 - R), \quad (9.25)$$

а мощность, излучаемая при этом, будет равна  $\hbar \omega \frac{\Delta m}{\Delta t}$ , где  $\Delta t = \frac{2L}{c}$ .

В результате, используя (9.24), найдем максимальную выходную мощность генератора:

$$P_{\max} = \hbar\omega \frac{c}{4L} n_{\text{нач}} (1 - R). \quad (9.26)$$

Для определения выходной энергии генератора и длительности импульса следует подставить (9.22) во второе уравнение (9.20) и провести его интегрирование. Ограничимся лишь оценками. Энергия ( $\bar{\omega}_{\text{нач}}$ ), запасенная в образце на рабочем переходе перед началом генерации, очевидно, равна

$$\bar{\omega}_{\text{нач}} = N_2 \hbar\omega,$$

где  $N_2$  – число атомов на уровне 2.

Энергия, излученная в импульсе, меньше  $\bar{\omega}_{\text{нач}}$ , так как импульс заканчивается в момент, когда часть частиц остается на втором уровне и полного высвечивания не происходит. При значительном превышении порога можно принять, что в конце генерации разность населенностей равна нулю, так что  $N_{2\text{кон}} = \frac{1}{2} N$ . В результате высвободившаяся энергия, очевидно, равна

$$\bar{\omega}_{\text{изл}} = (N_{2\text{нач}} - \frac{1}{2} N) \hbar\omega. \quad (9.27)$$

Выражение (9.27) дает оценку выходной энергии генератора в рассмотренном приближении.

Длительность импульса можно оценить, пользуясь соотношением

$$\bar{\omega}_{\text{изл}} \approx P_{\max} \tau. \quad (9.28)$$

Проведем численные оценки для рубинового лазера. Пусть кристалл имеет объем  $5 \text{ см}^3$  и плотность активных частиц в нем составляет  $1.6 \cdot 10^{19} \text{ см}^{-3}$ . Примем длину резонатора  $L=1 \text{ м}$ , коэффициент отражения зеркала  $R=0.5$ ,  $N_{2\text{нач}} = 0.9N$  и соответственно  $n_{2\text{нач}} = 0.8N$ . Величина  $\hbar\omega$  на длине волны рубинового лазера равна  $2.85 \cdot 10^{-19} \text{ Дж}$ . Подставляя эти значения в формулы (9.26), (9.27), (9.28), найдем:

$$P_{\max} = 0.7 \cdot 10^9 \text{ Вт},$$

$$\bar{\omega} = 9.1 \text{ Дж},$$

$$\tau = 13 \text{ нс}.$$

Таким образом, в режиме модуляции добротности, лазер излучает наносекундные импульсы с высокой пиковой мощностью, достигающей сотен мегаватт [4,13,14].

#### 9.4.4. Синхронизация мод в твердотельном лазере и

##### генерация ультракоротких импульсов (УКИ)

Обычно, твердотельные лазеры генерируют на достаточно большом числе типов колебаний (мод). При этом, однако, фазовая связь между полями различных мод отсутствует. Усложнением схемы генератора можно добиться жесткой связи между фазами колебаний мод. В таком случае говорят, что генератор работает в режиме синхронизации мод. Если связь между фазами мод устанавливается самопроизвольно, то говорят что, имеет место самосинхронизация мод.

При появлении синхронизации мод, временные характеристики лазеров резко меняются. Чтобы понять происходящие изменения, рассмотрим схему лазера бегущей волны (рис. 9.6).

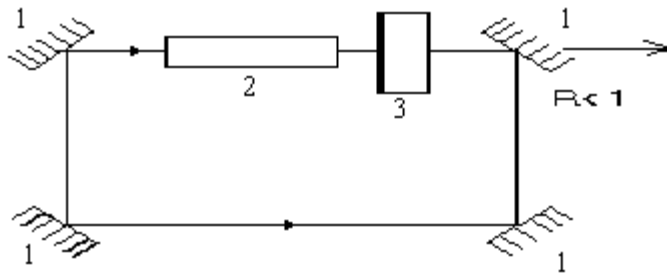


Рис. 9.6. Схема лазера бегущей волны с синхронизацией мод:

1 –зеркала; 2 –активный элемент; 3 –модулятор

В резонаторе такого генератора поле может быть представлено в виде суммы волн, бегущих со скоростью света в каком-то направлении, например, по часовой стрелке. Для простоты будем полагать, что этих типов колебаний всего два

$$E_1 = E_m \cdot \sin(\omega_0 \cdot t - k_0 \cdot x)$$

$$E_2 = E_m \cdot \sin(\omega_1 \cdot t - k_1 \cdot x) ,$$

где  $k_{0,1} = \omega_{0,1} / c$  — волновое число;  $\omega_{0,1}$  — частота типов колебаний.

Расстояние между ними по частоте  $\omega_1 = 2\pi c / L$  совпадает с расстоянием между продольными модами. Соответственно волновые числа волн отличаются на  $k_1 = 2\pi / L$ , где  $L$  — полная длина резонатора.

Примем также, что амплитуды волн равны, а фазы колебаний в момент времени  $t = 0$  в точке  $z = 0$  совпадают и далее жестко между собой связаны. Тогда поле в точке  $z$  в момент времени  $t$  будет

$$E_1 + E_2 = E_m \cdot [ \sin(\omega_0 t - k_0 z) + \sin(\omega_1 t - k_1 z) ] =$$

$$= 2 E_m \cdot \cos[(\omega_0 - \omega_1) \cdot t / 2 - (k_0 - k_1) \cdot z / 2] \cdot \sin[(\omega_0 + \omega_1) \cdot t / 2 - (k_0 + k_1) \cdot z / 2] =$$

$$2 E_m \cos(\omega_1 t - k_1 z) / 2 \cdot \sin[(\omega_0 + \omega_1) / 2] t - (k_0 + k_1 / 2) z]. \quad (9.29)$$

Отсюда видно, что произведение сомножителей перед последним синусом может быть взято за амплитуду поля, медленно меняющуюся во времени и пространстве. Интенсивность света  $I$  пропорциональна квадрату амплитуды поля, так что из (9.29) получим

$$I(z, t) = E_m^2 [1 + \cos(\omega_1 t - k_1 z)]. \quad (9.30)$$

График зависимости  $I(z, t)$  представлен на рис.9.7. Как видно, распределение поля неравномерно, причем поле в интервале  $L/2$  превышает среднее значение и в таком же интервале оно меньше среднего. Вся картина поля не остается неизменной в пространстве, а перемещается вдоль оси резонатора  $z$  со скоростью света  $c = \omega_1 / k_1$ . В некоторый момент времени максимум интенсивности света набегаем на выходное зеркало. В этот момент через зеркало проходит максимальная мощность излучения. Через промежуток времени  $2L / c$  у зеркала окажется минимум интенсивности. При этом мощность излучения минимальна. Вывод: при синхронизации двух типов колебаний интенсивность выходного излучения является периодической функцией времени с периодом  $2L / c$ .

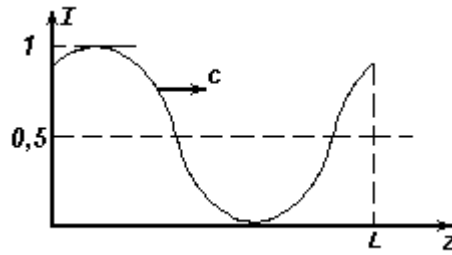


Рис. 9.7. Распределение интенсивности поля в резонаторе при синхронизации двух типов колебаний

Качественно эта картина остается верной и при синхронизации большого числа мод. Однако распределение интенсивности света в резонаторе гораздо более неравномерно, чем при двух синхронизованных модах. Амплитуда результирующего поля при синхронизации типов колебаний отлична от нуля в интервале  $2L/N$  ( $N$  – число синхронизованных мод). На остальной части резонатора поле близко к нулю (рис.9.8).

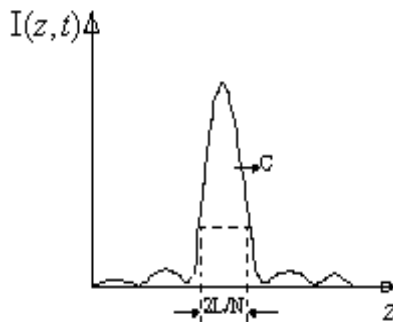


Рис. 9.8. Распределение поля в резонаторе лазера при синхронизации  $N$  мод

Максимум поля по-прежнему перемещается в направлении оси резонатора со скоростью света. Каждый раз, когда максимум интенсивности оказывается у выходного зеркала, генератор излучает импульс света.

Таким образом, в режиме синхронизации мод лазер выдает импульс света с периодом повторения  $T = 2L/c$  и длительностью импульсов  $\tau = T/N$ . Если  $2L = 1$  м, то  $F_{\text{повт}} = 1/T = 300$  МГц, а  $\tau$  при синхронизации 100 мод будет  $3 \cdot 10^{-11}$  с. Пиковая мощность превышает среднюю в  $(T/\tau) = 100$  раз.

Переход к генерации с синхронизацией мод может быть осуществлен с помощью модулятора потерь, помещенного в резонатор лазера. Если коэффициент пропускания этого модулятора изменяется с частотой  $F_{\text{ном}} = c/2L$ , то при отсутствии синхронизации мод, потери энергии в модуляторе будут пропорциональны среднему за период коэффициенту пропускания модулятора.

При синхронизации типов колебаний формируется импульс, который может проходить через модулятор как раз в те моменты времени, когда потери в модуля-



торе малы. В результате режим с синхронизацией мод оказывается энергетически более выгодным и, следовательно, в схеме, представленной на рис.9.8 он разовьется в первую очередь. Могут быть и другие схемы получения режима синхронизации мод.

В настоящее время в режиме синхронизации типов колебаний могут быть получены импульсы света с длительностью, измеряемой долями пс и пиковой мощностью  $10^{11}$  Вт (число синхронизированных мод ограничено, так как это зависит от многих факторов [4]).

#### 9.4.5. Спектральные и угловые характеристики твердотельных лазеров

Хотя излучение лазера обладает на несколько порядков меньшей шириной спектра (чем тепловые источники) полностью монохроматичным назвать нельзя. Например, в рубиновом лазере, если не принять специальные меры  $\Delta \nu = 30$  ГГц, т.е.  $5 \cdot 10^{-5}$  от  $\nu_0$ . Для многих приложений это недопустимо. Кроме того, наряду с нерегулярным во времени излучением (“пичками”) наблюдается также нерегулярное изменение спектра – в генерацию входят в разных “пичках” различные продольные моды. Разберем причины нерегулярности спектра излучения лазера во времени.

Вначале генерация лазера развивается на том аксиальном типе колебаний, который расположен наиболее близко к центру линии люминесценции и, следовательно, обладает большим усилением. По мере возрастания амплитуды поля этого колебания, разность населений рабочих уровней будет падать, а с ней и коэффициент усиления. Однако, поскольку поле аксиальной моды представляет собой стоячие волны, разность населенностей будет падать в основном вблизи пучностей стоячей волны (см. рис.9.9).

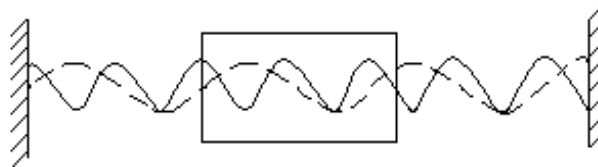


Рис. 9.9. Распределение поля аксиальной моды в резонаторе

Высвечивание активных ионов в минимумах очевидно не происходит. Это благоприятствует возбуждению тех мод, пучности которых расположены в области узлов основной моды. Таким образом, происходит срыв генерации одной моды (или группы мод) и в генерацию входят другие моды.

Расчеты показывают, что за счет неоднородности распределения поля мод, генерация происходит на многих модах в частотном интервале

$$\Delta \nu = 0.7 \cdot \Delta \nu_{\text{л}} \cdot (\delta \nu / \Delta \nu_{\text{л}})^{1/3} \cdot (1 - 1/\alpha) \quad (9.31)$$

где  $\alpha = W / W_{\text{пор}}$ ,

$\Delta \nu$  – ширина спектра генерации,  $\Delta \nu_{\text{л}}$  – ширина линии люминесценции,

$\delta \nu$  – расстояние между соседними модами,

$\alpha$  – превышение накачки ее порогового значения.

Как видно из (9.31), ширина спектра генерации  $\Delta\nu$  зависит от уровня накачки. На пороге генерации  $\alpha \rightarrow 1$  и  $\Delta\nu \rightarrow 0$ . При повышении уровня накачки  $\Delta\nu$  растет, очень быстро достигая насыщения  $\Delta\nu_{\text{пред}}$  при  $\alpha \rightarrow 0$ . Это предельное значение, очевидно, равно

$$\Delta\nu_{\text{пред}} = 0.7 \cdot \Delta\nu_{\text{л}} \cdot (\delta\nu / \Delta\nu_{\text{л}})^{1/3}. \quad (9.32)$$

Оценим  $\Delta\nu_{\text{пред}}$ . При длине резонатора  $L = 0,5$  м -  $\delta\nu = c / 2L = 300$  МГц, пусть далее  $\Delta\nu_{\text{л}} = 300$  ГГц, тогда  $\Delta\nu_{\text{пред}} \approx 20$  ГГц, что согласуется с экспериментальными результатами.

Из изложенного следует, что устранение продольной неоднородности поля мод, должно приводить к сужению спектра излучения лазера.

Устранение неоднородности взаимодействия вещества и поля можно достичь различными способами, например, двигая активный элемент вдоль оси резонатора в процессе генерации, или двигая поле в активном элементе с помощью электрооптических кристаллов, помещенных в оптический резонатор (рис. 9.10).

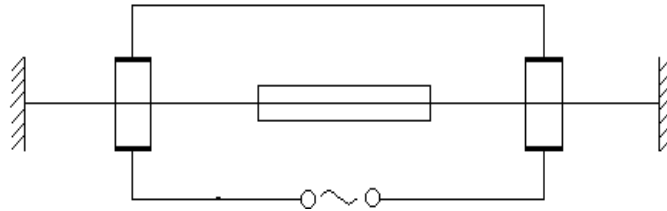


Рис. 9.10. К иллюстрации устранения неоднородностей с помощью двух электрооптических кристаллов

В последнем случае, электрическое поле изменяет скорость распространения света в кристаллах, поэтому меняется  $\lambda = v / \nu = c / \nu$ , и при изменении напряжения стоячая волна двигается в активном элементе. Важным является также то, что в таких лазерах происходит не только сужение спектра генерации, но и улучшение ее кинетики. Последовательность пичков приобретает хорошую регулярность.

Существенно зависят спектральные и временные характеристики лазера от сферичности зеркал резонатора. При достаточно большой сферичности в генерацию вовлекается большое число поперечных мод. Когда имеет место сильное перекрытие спектров мод с различными аксиальными индексами, последовательность пичков регуляризуется, спектр сужается. Но угловая расходимость излучения при этом возрастает. Использование различных методов сужения спектра излучения лазера можно добиться его ширины до нескольких десятков МГц.

В заключении раздела отметим, что в твердотельных лазерах (при генерации многих поперечных мод) угловая расходимость излучения значительно превышает дифракционную (она составляет  $(1 \div 5) \cdot 10^{-3}$  рад.). Предельное значение угловой расходимости при выходе в генерацию многих поперечных мод и при высоком уровне накачки, можно оценить по формуле [4].

$$\alpha_{\text{пред}} \cong 0.7 \cdot [\ln(1/r)]^{1/3} \cdot [\lambda/d]^{1/3}, \quad (9.33)$$

где  $r$  – коэффициент отражения выходного зеркала,  $d$  – диаметр активного элемента.

## ГЛАВА 10

### Газовые лазеры

#### 10.1. Особенности газов как активного вещества для лазеров

Активная среда газовых лазеров имеет ряд отличий от активного вещества твердотельных лазеров.

Так как расстояние между молекулами в газовой среде больше, то взаимодействие частиц друг с другом ослаблено. Это ведет к сужению спектра. Ширина линии определяется в основном разными скоростями излучающих атомов и молекул во время их движения (доплеровское уширение линии).

Важной особенностью является то, что газовая среда обладает очень высокой однородностью. Потери на рассеяние почти отсутствуют. Это позволяет использовать в газовых лазерах большие расстояния между зеркалами. Поэтому с помощью газовых лазеров легко получить высокую направленность и монохроматичность излучения.

#### 10.2. Механизмы возбуждения газовых лазеров

В газовых лазерах возбуждение атомов (молекул) обычно осуществляется с помощью электрического разряда, т.е. накачка достигается достаточно большим электрическим током (постоянным или переменным), который течет через газовую среду.

Возбуждение молекулы газа может происходить при ее соударении с электроном (например, при электрическом разряде в газе). При этом в газоразрядной плазме наряду с электронами присутствуют и ионы. Однако, средняя энергия последних мала, и процессом возбуждения молекул (атомов) газа ионами обычно пренебрегают. Можно считать также, что электроны имеют максвелловское распределение энергий по скоростям (строго говоря, это приближение справедливо не во всех случаях), и их температура  $T_e$  и плотность (концентрация)  $N_e$ , зависящая от величины тока разряда, считаются известными.

В газовом разряде наиболее важными процессами, приводящими к обмену энергией между частицами, являются:

- электронные удары (соударения) первого рода, при которых атом получает энергию от электрона;
- электронные удары второго рода, при которых атом передает свою энергию электрону;
- спонтанное излучение возбужденных атомов;
- поглощение атомами излучения внешнего поля;
- индуцированное излучение атомов.

Скорость, с которой протекают эти процессы, определяется числом атомов в данном состоянии и вероятностями переходов в единицу времени для каждого атома.

Вероятность перехода в единицу времени  $\left(\sum_{ij} W_{ij}\right)$  равна обратной величине времени жизни атома в данном состоянии. Обозначив через  $\sigma_{ij}$  сечение соударений, вызывающих переход между уровнями  $i$  и  $j$ , и через  $v$  относительную скорость соударяющихся частиц, можно выразить вероятность перехода в виде

$$W_{ij} = T_{ij}^{-1} = N \overline{y_{ij} x}. \quad (10.1)$$

Черта в формуле (10.1) означает усреднение по скоростям.

Из принципа детального баланса ( $W_{ji}g_j = W_{ij}g_i$ ) следует, что вероятность  $W_{ij}$  процесса, в котором атом переходит с нижнего уровня  $E_i$  на более высокий уровень  $E_j$ , связана с вероятностью обратного процесса  $W_{ji}$  соотношением

$$W_{ij} = b \cdot \frac{g_j}{g_i} W_{ji}, \quad (10.2)$$

где  $b$  — интеграл по энергетическому распределению частиц.

В случае принятого нами допущения о максвелловском распределении энергий электронов по скоростям при температуре  $T_e$  величина  $b$  представляет собой фактор Больцмана и выражение (10.2) можно переписать в виде

$$\frac{W_{ji}}{W_{ij}} = \frac{T_{ij}}{T_{ji}} = \frac{g_i}{g_j} e^{-(E_j - E_i)/kT_e}. \quad (10.3)$$

Стационарное распределение может быть получено, если учесть все возможные процессы в кинетических уравнениях для изменения числа частиц на уровнях. В простейшем случае (двухуровневая система) это уравнение имеет вид

$$\frac{dN_2}{dt} = \frac{N_1}{T_{12}} - \frac{N_2}{T_{21}} - \frac{N_2}{\tau_{21}}, \quad (10.4)$$

где  $\tau_{21}$  — время спонтанного перехода  $2 \rightarrow 1$ .

Скорость накопления частиц на уровне «2» зависит от числа частиц на нижнем уровне «1— $N_1$ » и вероятности перехода с уровня «1» на уровень «2» (1-й член в уравнении). Переходы с уровня «1» на уровень «2» происходят при этом за счет электронных ударов первого рода.

За счет электронных ударов второго рода (при соударении возбужденного атома с электронами) часть частиц с уровня «2» возвращается на уровень «1», уменьшая скорость накопления частиц на верхнем уровне (член  $N_2/T_{21}$  в уравнении (10.4)). Спонтанное излучение со временем перехода с уровня «2» на уровень «1» —  $\tau_{21}$  также обедняет уровень «2» со скоростью  $N_2/\tau_{21}$ .

В стационарном состоянии, когда  $\frac{dN_2}{dt} = 0$ , из уравнения (10.4) получаем

$$\frac{N_2}{N_1} = \frac{T_{21} \tau_{21}}{T_{12} (\tau_{21} + T_{21})}. \quad (10.5)$$

Если спонтанное излучение играет основную роль в обеднении возбужденного уровня (при условии, что процесс спонтанного излучения происходит значительно быстрее переходов с возбужденного уровня за счет электронных ударов второго рода, т.е.  $\tau_{21} \ll T_{21}$ ), то

$$\frac{N_2}{N_1} \approx \frac{\tau_{21}}{T_{12}}, \quad (10.6)$$

или с учетом (10.3)

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \cdot \frac{\phi_{21}}{T_{21}} e^{-(E_2-E_1)/kT_e}. \quad (10.7)$$

Поскольку  $E_2 > E_1$ , экспоненциальный множитель, входящий в формулу (10.7), меньше единицы (при высоких электронных температурах экспоненциальный множитель незначительно отличается от единицы). Отношение  $\tau_{21}/T_{21}$  представляет собой величину отклонения распределения от бoльцмановского (характерного для равновесных систем). Вероятность перехода  $W_{12}$  и, следовательно, отношение  $N_2/N_1$  могут увеличиваться по желанию экспериментатора путем увеличения электронной концентрации, но не беспредельно, поскольку применимость формулы (10.7) ограничена предположением  $\frac{\phi_{21}}{T_{21}} \ll 1$ .

Таким образом, в двухуровневых системах рассмотренный механизм возбуждения электронными ударами не приводит к инверсии населенности, так как  $N_2$  всегда меньше  $N_1$ .

Для трехуровневой системы по вышеизложенной методике при тех же ограничениях ( $\tau_{32} \ll T_{31}$ ) можно определить отношение

$$\frac{N_3}{N_2} \approx \frac{g_3}{g_2} \cdot \frac{\phi_{32}}{\phi_{21}} \cdot \frac{T_{21}}{T_{31}} e^{-(E_3-E_2)/kT_e}. \quad (10.8)$$

При благоприятных соотношениях между вероятностями переходов и временами спонтанных переходов экспоненциальный множитель может значительно превышать единицу, и при  $\frac{\phi_{32}}{T_{31}} \ll 1$  (в области применимости формулы (10.8)) можно получить состояние с инверсией населенности  $\frac{N_3}{N_2} > 1$ .

### 10.3. Гелий – неоновый лазер

Гелий неоновый (*He-Ne*) лазер относится к типу атомарных газовых лазеров. Этот тип лазеров был создан в начале 1961 г в США. Гелий-неоновые лазеры нашли наибольшее распространение по сравнению с другими типами газовых лазеров.

Схематично конструкция *He-Ne* лазера может быть представлена следующим образом (см. рис. 10.1).

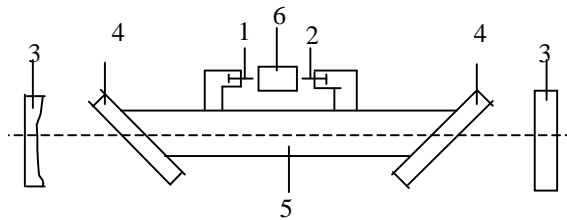


Рис. 10.1. Конструкция *He-Ne* лазера: 1. анод. 2. катод. 3. зеркала резонатора. 4. выходные окна. 5. трубка. 6. источник питания

Лазер состоит из газоразрядной трубки 5 и резонатора, образованного зеркалами 3, источника питания -6. Трубка заполнена смесью газов гелия и неона. С торцов она закрывается плоскими окнами из оптического стекла, наклоненными под углом Брюстера к продольной оси трубки, при этом обеспечиваются наименьшие потери на торцах и высокая степень поляризации.

В гелий-неоновых лазерах для возбуждения активной среды используется тлеющий разряд, который осуществляется в трубке либо постоянным током, либо полем высокой частоты.

Для пояснения принципа действия *He-Ne* лазера рассмотрим энергетические уровни атомов гелия и неона (см. рис.10.2),

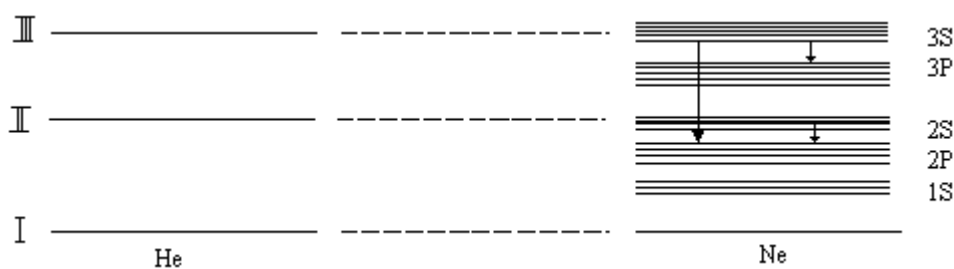


Рис. 10.2. Энергетические уровни атомов гелия и неона

Для получения инверсии населенности используется рассмотренный выше механизм столкновения второго рода. В качестве рабочего газа служат атомы неона, вспомогательного – гелия.

При включении накала трубки, в ней возникает тлеющий разряд, в котором появляется большое число свободных электронов и ионов.

Атомы гелия из основного состояния 1 вследствие неупругих столкновений со свободными электронами могут переходить в возбужденные состояния 2, 3. Верхние уровни неона также заселяются за счет соударений с электронами. Уровни 2 и 3 гелия энергетически очень близки к уровням 2S и 3S неона, поэтому при столкновениях возбужденных атомов гелия с невозбужденными атомами неона происходит резонансная передача возбуждений: атомы гелия переходят в основное состояние 1, а атомы неона – в возбужденные состояния 2S или 3S. Для наиболее эффективного заселения уровней 2S и 3S неона за счет столкновений второго рода создают избыточную концентрацию гелия (обычно отношение парциальных давлений гелия и неона выбирают примерно 10:1).

Вследствие резонансной передачи возбуждений, населенности уровней 2S и 3S значительно возрастают и, при определенных условиях разряда, возникает инверсная населенность на переходах  $3S \rightarrow 3P$  ( $\lambda = 3,39 \text{ мкм}$ );  $2S \rightarrow 2P$  ( $\lambda = 1,15 \text{ мкм}$ );  $3S \rightarrow 2P$  ( $\lambda = 0,63 \text{ мкм}$ ). На практике переход  $3S \rightarrow 2P$  используют особенно часто.

Уровень 2P опустошается за счет спонтанных переходов на метастабильный уровень 1S. Перевод атомов неона из метастабильного состояния 1S в основное, осуществляется за счет диффузии на стенки газоразрядной трубки. Для облегчения диффузии возбужденных частиц к стенкам, диаметр трубки должен быть около 7 мм.

Атомы неона, находящиеся на более высоких энергетических состояниях, имеют меньшее время жизни и не успевают добраться до стенок трубки. Уменьшение населенности уровня 1S вызывает дополнительные переходы атомов Ne с уровня  $2P \rightarrow 1S$ . Эти переходы, обедняющие нижний лазерный уровень, приводят

к увеличению инверсии, следовательно, к более интенсивной лазерной генерации. Выходная мощность лазера при этом возрастает. В лазере с возбуждением на постоянном токе изменение выходной мощности от тока разряда имеет вид (рис.10.3):

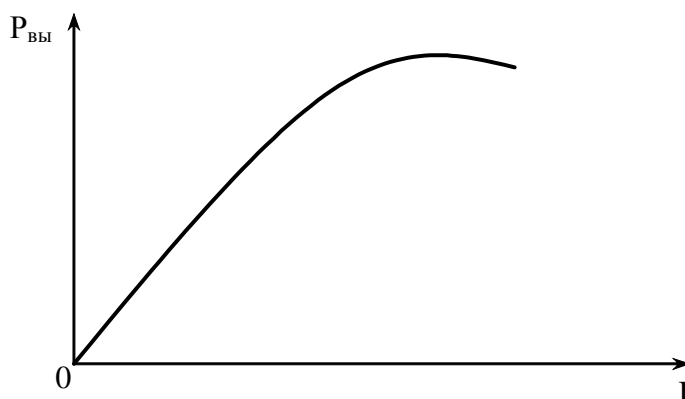


Рис. 10.3. Зависимость выходной мощности от тока разряда

Уменьшение выходной мощности при больших, токах объясняется тем, что большая плотность электронов увеличивает вероятность столкновения атомов *He*, находящихся в метастабильном состоянии, с электронами. За счет этих столкновений метастабильные состояния *He* разрушаются и, вследствие этого, уменьшается число атомов *He*, возбуждающих атомы *Ne*.

Длина *He-Ne* лазера – от 10 см до 2 м. Чем длиннее лазер, тем больше направленность лазерного луча, которая зависит также от модового состава излучения и кривизны зеркал резонатора (расходимость составляет несколько угловых минут).

*He – Ne* лазер характеризуется высокой стабильностью излучения. Относительный уход частоты за 1 час составляет  $10^{-9}$ . За счет принятия специальных мер, уменьшающих шумы лазера, удалось достигнуть высокой монохроматичности излучения (ширина спектра генерации – единицы герц). Мощность, излучаемая *He-Ne* лазером в непрерывном режиме, составляет от долей милливатта до 100 мВт. К.п.д. мал и составляет сотые доли процента.

Наряду с лазерами, активной средой которых являются нейтральные атомы, существенное значение приобретают атомарные импульсные газовые лазеры на парах твердых веществ: свинца, меди, марганца, таллия и др. Эти генераторы имеют наиболее высокие мощности излучения в сине-зеленой области спектра. Они обладают наивысшим к.п.д. по сравнению со всеми атомарными лазерами [17], так как в лазерах этого типа (в качестве рабочих уровней) используются сравнительно низко расположенные энергетические уровни атомов металлов. Верхним рабочим уровнем является первый резонансный, нижним - метастабильный уровень. На возбуждение первого резонансного уровня расходуется большая доля энергии разряда. Например, при применении ртути она составляет около 60%. ( на возбуждение уровня 17эВ, соответствующего верхнему рабочему уровню неона, идет только около 1% энергии). Одна из уникальных особенностей лазеров на переходах с резонансного на метастабильный уровень, и в том числе лазера на парах меди,- чрезвычайно высокий коэффициент усиления, на проход. Эти

лазеры легко могут работать в режиме сверхсветимости, генерация в них может возникать за счет усиления спонтанного излучения, т. е. без резонатора, а только с одним зеркалом или даже без зеркал. Однако однонаправленное излучение и более низкий порог получаются при использовании одного зеркала со 100%-ным отражением на одном из торцов трубки при выводе излучения с другого ее торца. Использование схемы генерации на лазерных переходах, оканчивающихся на метастабильных уровнях, вызывает необходимость иметь короткие импульсы возбуждения порядка длительности существования инверсии. Проведенные эксперименты свидетельствуют о способности этих приборов работать при высоких частотах повторения импульсов порядка единиц-десятков килогерц при мощности излучения около 1 кВт/м.

В настоящее время наибольшее применение находят генераторы на парах меди ( $\lambda=0,5106$  и  $0,5782$  мкм). Для обеспечения требуемой величины паров меди температура внутри разрядной трубки поддерживается равной  $1500-1600^{\circ}\text{C}$  (что соответствует давлению паров меди  $53,2$  Па). Для создания такой температуры применяются специальные подогреватели, которые несколько усложняют конструкцию лазера и требуют дополнительного расхода энергии. В лазерах на парах меди достигнута средняя мощность  $\sim 40$  Вт и частоты повторения импульсов  $\sim 15$  кГц. Эти лазеры представляют интерес для подводной связи и дистанционного поиска погруженных объектов, а также для некоторых лазерных фотохимических применений.

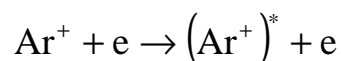
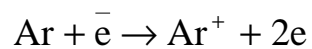
Широкое применение лазеров в различных областях связано с возможностью обеспечения устойчивой его работы при воздействии механических и климатических факторов [4,9,12,21].

#### 10.4. Ионный аргоновый лазер

Из ионных лазеров наибольшее распространение получили аргоновые лазеры (непрерывного действия на длине волны  $\lambda = 0,48$  мкм).

Инверсия населенности в аргоновом лазере создается следующим образом:

Сначала атомы  $Ar$  сталкиваясь с электроном, образуемым в плазме газового разряда, ионизируются (то есть переходят в состояние  $Ar^+(3p^5)$  рис. 10.4). Затем, при втором соударении, происходит возбуждение  $Ar$  (переход одного электрона из  $3p$  в  $4p$  оболочку).



На самом деле уровни  $3p^44P$  и  $3p^44S$  состоят из множества подуровней, поэтому лазер может генерировать на многих переходах между подуровнями этих состояний в широком диапазоне длин волн ( $0,46 - 0,52$ ) мкм (т.е. в сине-зеленой области спектра).



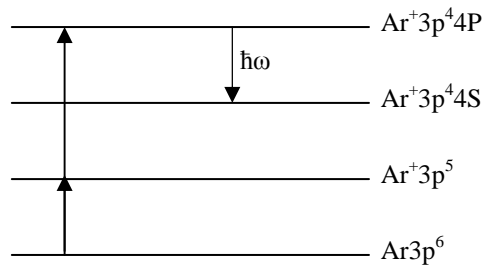


Рис. 10.4. Энергетические уровни аргона

Для ионизации газа данный лазер требует протекания через газоразрядную трубку токов очень большой плотности – до нескольких тысяч А/см<sup>2</sup>. При использовании токов такой плотности, необходимо применять специальные меры для сохранения газоразрядной трубки от перегрева. Поэтому все аргоновые лазеры работают, как правило, с водяным охлаждением. Для уменьшения бомбардировки стенок трубки ионами и электронами используют продольное магнитное поле, которое концентрирует заряженные частицы вдоль оси трубки (см. рис. 10.5). Газовый разряд создается в капилляре диаметром 6 мм. По конструкции аргоновые лазеры несколько отличаются от гелий – неоновых.

Катод эмитирует электроны, которые под действием электрического напряжения движутся к аноду ионизируя атомы  $Ar^+$ . Ионы  $Ar^+$  сосредотачиваются у катода, что приводит к уменьшению давления газа у анода. Это может сорвать заряд. Для этого в конструкции применен отводный канал, который выравнивает давление газа в основном (рабочем) капилляре. Вследствие однонаправленного движения ионов аргона в разрядном промежутке, контур спектральной линии имеет асимметричную

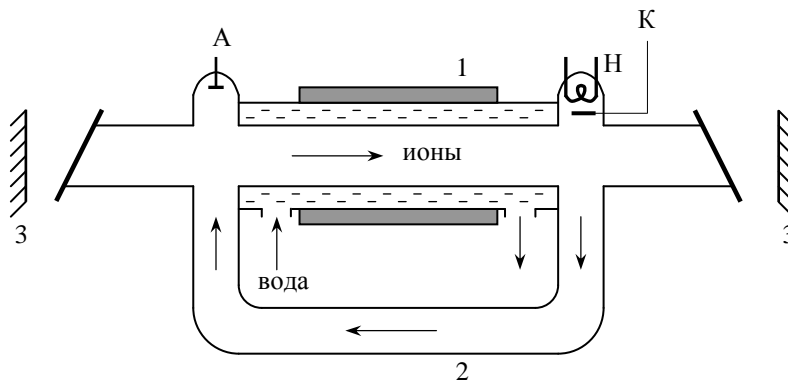


Рис. 10.5. Конструкция лазера на Ar  
1-соленоид. 2-отводной канал. 3-зеркала

форму (такой же вид имеет кривая усиления). Блок питания представляет собой мощный выпрямитель (около 100 кВт), выходное напряжение которого составляет около 400 В. В качестве источника накачки можно использовать высокочастотное возбуждение разряда. Это позволяет уменьшить нагрев стенок газоразрядной трубки за счет уменьшения бомбардировки ионами.

Аргоновые лазеры являются самыми мощными лазерами в видимом диапазоне. Мощность их в непрерывном режиме достигает сотни Вт, в импульсном дос-

тигает сотен кВт. Недостаток – малый к.п.д. – 0,1% и малые сроки службы (~50 часов). При больших мощностях срок службы ограничен десятками часов, из-за разрушения стенок трубки и загрязнения выходных окон распыленными частицами электродов. Большая мощность генерации приводит также к быстрому разрушению диэлектрических зеркал.

В заключении следует отметить, что существуют лазеры на парах металлов, работающих на ионных переходах. Они так же как и лазеры на атомарных переходах многочисленны и разнообразны. Наиболее распространенный в настоящее время лазер этого типа – гелий кадмиевый ( $He - Cd$ ). В непрерывном режиме они позволяют получить с приборов средних размеров мощность генерации  $P = 10 \div 100$  мВт на  $\lambda = 441.6$  нм при к.п.д. около 0.1 %.

Основными областями применения аргоновых лазеров являются фотохимия, лазерная технология и медицина [5,9,13].

### 10.5. Молекулярный лазер на углекислом газе

Низкий к.п.д. атомарных и ионных лазеров, работающих в непрерывном режиме, обусловлен следующим: 1). Возбуждение газовой смеси происходит на высокорасположенные электронные уровни, так что  $\frac{\hbar\omega_{изл}}{E_{возб}}$  не превышает 0,1. 2).

Кроме того, при определении к.п.д. необходимо учесть, что не все электроны имеют достаточную энергию в газовом разряде для возбуждения верхнего рабочего уровня. В существующих атомарных и ионных лазерах доля таких электронов, имеющих энергию порядка  $E_{возб}$ , относительно мала. Основная же часть энергии в типичных условиях газоразрядной плазмы тратится на возбуждение самых низких уровней. Поэтому КПД этих лазеров, определяемый как  $\eta = \eta_n \frac{\hbar\omega_{изл}}{E_{возб}}$ , составляет доли процента.

Для повышения КПД газовых лазеров выгодно в качестве рабочих использовать низко расположенные уровни частиц. Наиболее подходящими с этой точки зрения являются возбужденные колебательные уровни молекул. Как известно, любая молекула совершает колебательное и вращательное движение. Энергия этих движений квантована таким образом, что разрешены лишь состояния с определенными значениями энергии. Обычно энергия вращательного движения меньше энергии колебательного движения. Поэтому, одному колебательному энергетическому состоянию с квантовым числом  $\nu$  соответствует много вращательных, характеризующимися квантовыми числами  $j$ , отличающимися на единицу. Подобно электронным состояниям атома, молекула имеет основное невозбужденное колебательное состояние с минимальной энергией и возбужденные состояния (рис.10.6).

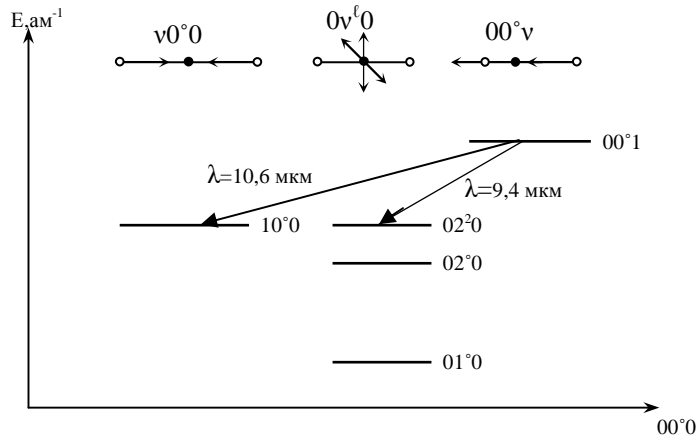


Рис. 10.6. Схема низших колебательных состояний молекулы  $CO_2$ . Вверху схематически показаны типы колебаний атомов в молекуле

Рассмотрим молекулу  $CO_2$ , которая представляет из себя линейную цепочку, в центре которой располагается атом углерода  $C$ . Атомы в молекуле могут колебаться вокруг положения равновесия. Молекула  $CO_2$  имеет три следующих типа колебаний:

1. Симметричный тип колебаний. Атомы кислорода колеблются вдоль оси молекулы симметрично относительно атома углерода.

2. Деформационные колебания. В этом случае атом углерода колеблется в направлении, перпендикулярном продольной оси молекулы. Данный тип колебаний является двукратно – вырожденным, поскольку при этом имеется два взаимно – перпендикулярных направления колебаний.

3. Асимметричные колебания. При этом атом углерода колеблется вдоль продольной оси молекулы (см. рис.10.6).

Все вышеперечисленные типы колебаний обозначаются соответственно тремя цифрами  $\nu 0^\circ 0$ ,  $0 \nu^l 0$ ,  $0 0^\circ \nu$ . В этих обозначениях  $\nu$  – число квантов, запасенных в данном виде колебаний (номер колебательного уровня), а  $l$  – поляризация деформационного колебания. Если число  $\nu$  стоит первым в тройке цифр, то этот уровень соответствует симметричному типу колебаний. Когда не равна нулю средняя цифра, то это означает, что уровень относится к деформационному типу колебаний, а когда последняя – к асимметричному. Расстояния между колебательными уровнями одного вида колебаний всегда эквидистантны.

Основное состояние молекулы обозначается  $0 0^\circ 0$ . Самым низшим возбужденным уровнем молекулы  $CO_2$  является уровень  $0 1^\circ 0$ . Далее идут  $0 2^\circ 0$ ,  $0 2^2 0$ ,  $1 0^\circ 0$ ,  $0 0^\circ 1$  и т.д. Обычно частоты переходов между колебательными уровнями много меньше частот между различными электронными состояниями молекул и лежат в миллиметровой и инфракрасной области спектра. В лазере на  $CO_2$  в качестве лазерных переходов используются колебательно – вращательные переходы  $0 0^\circ 1 \leftrightarrow 1 0^\circ 0$  и  $1 0^\circ 0 \leftrightarrow 0 2^\circ 0$  ( $\lambda = 10,6$  и  $9,4$  мкм соответственно).

Инверсия населенности в стационарном режиме на колебательно – вращательных переходах молекулы  $CO_2$  образуется за счет того, что скорость разрушения колебательных состояний  $1 0^\circ 0$  и  $0 2^\circ 0$  много больше скорости разрушения состояния  $0 0^\circ 1$ . Вероятность перехода  $CO_2$  из состояния  $1 0^\circ 0$  вниз возрастает при введении в газовую смесь атомов гелия, которые при столкновениях с молекула-

ми  $CO_2$  разрушают это состояние. Вероятность же разрушения состояния  $00^0 1$  при столкновениях с атомами гелия мала.

Возбуждение лазера на  $CO_2$  осуществляется (в основном) за счет столкновений второго рода. В качестве вспомогательного газа используются молекулы азота –  $N_2$ , находящиеся на возбужденном колебательном уровне (см. рис. 10.7а). Конструкция лазера на  $CO_2$  с непрерывной прокачкой рабочей смеси показана на рис. 10.7б.

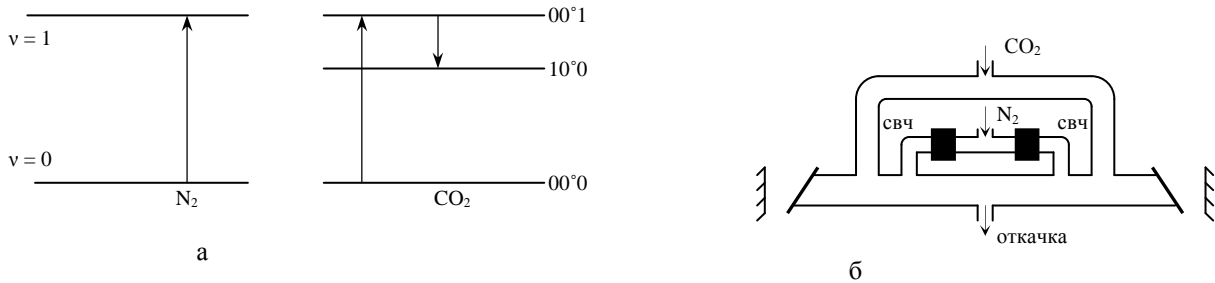


Рис. 10.7.

а) -схема колебательных уровней  $N_2$  и  $CO_2$ , б) -конструкция лазера на  $CO_2$

Через разрядную трубку в области высокочастотного разряда – прокачивается чистый азот. Возбужденные молекулы  $N_2$ , находящиеся в метастабильных колебательных состояниях попадают затем в область взаимодействия. В этой области возбужденный азот смешивается с углекислым газом, и при столкновениях осуществляется резонансная передача энергии возбуждения от молекул  $N_2$  к молекулам  $CO_2$ . Смесь газов участвовавшая в процессе генерации откачивается. Прокачка газовой смеси позволяет удалить из рабочей части лазера вещества, образующиеся в результате химических реакций.

В лазере на углекислом газе используются трубки большого диаметра (до 10 см) и большой длины (до 20 м). При этом в непрерывном режиме может быть получена большая мощность (~ до 10 кВт). Основное достоинство лазера на углекислом газе – высокий к.п.д. (до 30%) [4,17]. Это самый высокий к.п.д. из всех лазеров, исключая полупроводниковые лазеры. Высокий к.п.д. молекулярного лазера объясняется во первых тем, что отношение излучаемой энергии к энергии накачки

верхнего лазерного уровня (квантовый выход) –  $\frac{h\nu_{нзл}}{E_{00^0 1}} = 0,4$ . Во вторых, почти

вся энергия электронов газового разряда идет на возбуждение как молекул  $N_2$ , так и самих молекул  $CO_2$ . Кроме того соседние колебательные уровни молекулы  $N_2$  обладают хорошей эквидистантностью. Практически любое возбужденное состояние молекулы  $N_2$  способно передать колебательный квант энергии молекуле,  $CO_2$ , переходя при этом в состояние с колебательным числом на единицу меньшим. Таким образом одна молекула азота  $N_2$ , переведенная в самое верхнее колебательное состояние, может, при столкновениях, возбудить несколько молекул  $CO_2$ .

Особенностью молекулы углекислого газа является то, что ее верхний рабочий уровень обладает очень большим временем жизни (его разрушение происходит, главным образом, только за счет столкновений). Благодаря этому верхний

рабочий уровень, в отсутствие генерации, может служить накопителем возбуждённых молекул. При мгновенном включении резонатора с модулированной добротностью молекулы, переходя на нижний уровень излучают мощный импульс. От лазера, который в непрерывном режиме давал мощность –  $P = 3$  Вт, при модуляции добротности с помощью вращающегося зеркала, была получена мощность более 100 кВт в импульсе длительностью 100 нс.

Наиболее интенсивное излучение на длине волны 10,6 мкм. подавляет излучение на длине волны 9,4 мкм. Наша атмосфера прозрачна для излучения с длиной волны 10,6 мкм. Это позволяет использовать эти лазеры в оптической локации в диапазоне 9,2 – 11,4 мкм., с излучаемой мощностью  $P = 10 – 20$  Вт. Использование специальных секционированных трубок, позволяет увеличить излучаемую мощность до 100 Вт в непрерывном режиме [32].

Лазеры на  $CO_2$  используются в нелинейной оптике, локации, технологии, медицине.

Отечественные серийные лазеры имеют следующие марки: ЛГ – 15  $P = 10$  Вт; ЛГ – 22  $P = 40$ Вт; ЛГК – 21  $P = 10$  кВт.

**Газодинамический  $CO_2$ -лазер** [12,17,22,23]. Значительным достижением в развитии мощных газовых лазеров является создание газодинамических лазеров (ГДЛ). Идея их создания принадлежит ученым В.К. Конюхову и А.М. Прохорову. Газодинамический  $CO_2$ -лазер заслуживает особого внимания, поскольку инверсия населенностей в нем создается не электрическим разрядом, а за счет быстрого расширения газовой смеси (содержащей  $CO_2$ ), предварительно нагретой до высокой температуры. Инверсия населенностей возникает в потоке в области расширения. В литературе имеются сообщения, что газодинамические  $CO_2$ -лазеры дают самые высокие мощности.

Принцип работы газодинамического лазера можно кратко описать следующим образом (см. рис. 10.8а). Предположим, что вначале газовая

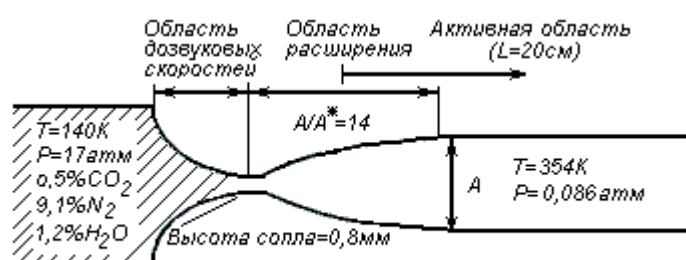


Рис. 10.8а. Схема, иллюстрирующая принцип действия газодинамического  $CO_2$ -лазера

смесь находится при высокой температуре (например,  $T = 1400$  К) и высоком давлении (например,  $p = 17$  атм) в соответствующем резервуаре. Поскольку газ первоначально находится в термодинамическом равновесии, молекулы  $CO_2$  будут иметь значительную населенность уровня  $00^0_1$  (~10% населенности основного состояния; см. рис.10.8а). Разумеется, населенность нижнего уровня является более высокой (~25%), и, следовательно, инверсия населенностей отсутствует. Предположим теперь, что газовая смесь расширяется через какие-то сопла (рис.10.8б).



Рис. 10.8б. Схема, иллюстрирующая пространственное распределение численности  $N^*$  верхнего и нижнего лазерных уровней

Поскольку расширение является адиабатическим, трансляционная температура смеси становится очень низкой. За счет V—T-релаксации численности как верхнего, так и нижнего лазерных уровней будут стремиться к новым равновесным значениям. Однако, поскольку время жизни верхнего уровня больше времени жизни нижнего, релаксация нижнего уровня произойдет на более ранней стадии процесса расширения (рис. 10.8 б). Таким образом, в направлении по потоку от зоны расширения будет существовать достаточно широкая область с инверсией численностей. Протяженность этой области  $L$  приблизительно определяется временем, необходимым для передачи возбуждения от молекулы  $N_2$  молекуле  $CO_2$ . При этом оба лазерных зеркала выбираются прямоугольной формы и располагаются так, как показано на рис. 10.8 в. Такой способ создания инверсии численностей будет эффективным лишь в том случае, если в процессе расширения температура и давление смеси понижаются за время, которое мало по сравнению со временем жизни верхнего лазерного уровня и велико по сравнению с временем жизни нижнего лазерного уровня. Чтобы удовлетворить этим условиям, расширение должно проходить со сверхзвуковыми скоростями (число Маха =4). В заключение следует заметить, что начальную высокую температуру газовой смеси получают за счет горения специально подобранного топлива (например, горение  $CO$  и  $H_2O$  или бензола  $C_6H_6$ , а также окиси азота  $N_2O$ ; при этом автоматически образуется смесь  $CO_2/H_2O$  в отношении 2:1).

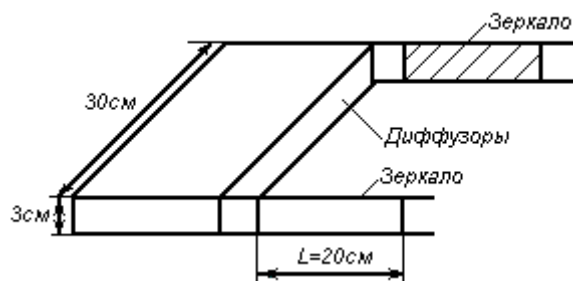


Рис. 10.8в. Схема, иллюстрирующая геометрию резонатора

Согласно опубликованным данным, газодинамические  $CO_2$ -лазеры излучают мощность до 80 кВт, причем химический к.п.д. (определяется как отношение выходной энергии лазера к полной химической энергии, которую можно получить при сгорании топлива) составляет 1 %. Лазеры этого типа могут работать в непрерывном режиме только в течение короткого времени, что обусловлено нагревом

отдельных элементов (например, зеркала) лазерным пучком.

Очевидно, что газовые лазеры, работающие на колебательно-вращательных переходах, не ограничиваются лишь  $CO_2$ -лазером. Другими примерами такого типа лазеров, о которых следует упомянуть, являются  $CO$ - ( $\lambda \approx 5$  мкм) и  $HCN$ -лазеры. Последний может генерировать на таких больших длинах волн, как  $\lambda = 773$  мкм = 0,773 мм, достигая, таким образом, миллиметровых длин волн. Благодаря высоким значениям мощности и к.п.д.  $CO$ -лазер представляет значительный интерес. Экспериментально достигнутые [22] выходные мощности таких лазеров превышают 100 кВт, а к.п.д. — 60%. Однако для реализации таких параметров газовую смесь необходимо охладить до низкой температуры (77—100 К<sup>0</sup>). В генерацию лазера при  $\lambda = 5$  мкм дают вклад несколько вращательно-колебательных переходов.

Накачка колебательных уровней молекулы  $CO$  осуществляется при возбуждении электронным ударом.

## Глава 11

### ЖИДКОСТНЫЕ ЛАЗЕРЫ

#### 11.1. Характеристика и особенности жидкостных лазеров

Жидкостные оптические квантовые генераторы (ЖКГ) характеризуются использованием в качестве активного лазерного вещества жидких сред – растворов неорганических соединений редкоземельных элементов (преимущественно неодима) или растворов органических красителей. В связи с этим различают ЖКГ на основе жидких неорганических сред и ЖКГ на основе органических красителей. Эти лазеры имеют ряд достоинств, делающих их весьма перспективными для некоторых областей применения.

Во-первых, в жидкостных лазерах можно обеспечить примерно такую же концентрацию активных частиц, как и в активных элементах твердотельных лазеров, что позволяет получить высокий энергоъем с единицы объема активной среды. Во-вторых, в ЖКГ в силу высокой однородности активной среды отсутствуют потери излучения, имеющие место в твердотельных лазерах из-за неоднородности и дефектов структуры активных элементов. В-третьих, охлаждение активной среды достаточно просто осуществляется за счет прокачки ее через кювету, что позволяет при импульсном режиме работы полностью обновить рабочий объем жидкости, участвующий в генерации, в интервале между последующими импульсами накачки, а при непрерывном режиме работы обеспечить ее эффективное охлаждение.

Достоинством ЖКГ является также их повышенная устойчивость к термическим напряжениям по сравнению с твердотельными лазерами. Долговечность жидкостных генераторов определяется в основном фотохимической стойкостью жидкостей к оптическому излучению накачки и стойкостью торцов кюветы к поверхностным напряжениям, вызванным высокой плотностью мощности или энергии излучения.

К недостаткам таких лазеров можно отнести небольшой срок сохраняемости рабочих растворов (через 1-2 месяца в растворах наблюдаются значительные изменения, которые могут привести к исчезновению генерации). Кроме того, в

жидких генерирующих средах появляются термические линзы с фокусным расстоянием, меняющимся в течение импульса излучения. Вследствие этого расходимость выходного излучения, особенно в мощных лазерах, существенно возрастает и достигает величин порядка единиц, а иногда и десятков градусов.

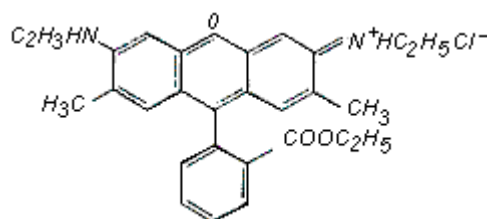
Следует отметить, что каждый из типов ЖКГ имеет и свои отличительные свойства.

Так, в ЖКГ на основе неорганических соединений редкоземельных элементов используются весьма агрессивные, в химическом отношении, жидкости. Активные же среды ЖКГ на основе растворов органических красителей имеют большую ширину спектральной линии излучения, достигающую 100нм, что позволяет создавать на их основе квантовые генераторы, перестраиваемые в широком диапазоне длин волн. Грубую перестройку можно осуществлять сменой красителя, поскольку диапазоны длин волн излучения различных красителей отличаются друг от друга, а плавную - введением в резонатор прибора регулируемых селективных элементов, позволяющих обеспечить генерацию на любой длине волны в пределах спектральной линии излучения данного раствора. Известный в настоящее время набор красителей позволяет создать ЖКГ с излучением, перестраиваемым в диапазоне от ультрафиолетовой до ближней ИК области спектра. Учитывая выше сказанное, остановимся на рассмотрении ЖКГ на органических красителях [4,12,22,24,25].

## 11.2. Жидкостные лазеры на органических красителях

В жидкостных лазерах активной средой является раствор некоторого органического красителя в воде или этиловом и метиловом спиртах. Эти красители обычно подразделяются на следующие типы: 1) полиметиновые ( $\lambda = 0,7—1$  мкм), 2) ксантеновые ( $\lambda = 0,5—0,7$  мкм), 3) кумариновые ( $\lambda = 0,4—0,5$  мкм) и 4) сцинтилляторные ( $\lambda < 0,4$  мкм) красители. Благодаря перестройке длины волны в широком спектральном диапазоне, а также их простоте лазеры на органических красителях играют все более важную роль в различных применениях (от спектроскопии до фотохимии).

**Фотофизические свойства органических красителей.** Органические красители представляют собой большие и сложные молекулярные системы с сопряженными двойными связями. В качестве примера приведем структурную формулу широко используемого красителя родамина 6G (ксантеновый краситель).



Эти красители, как правило, имеют сильную полосу поглощения в УФ- или видимой области спектра и при возбуждении светом с соответствующей длиной волны излучают интенсивные широкополосные спектры флуоресценции, как показано на рис. 11.1 для родамина 6G в этаноле.

Структуру энергетических уровней молекул красителя нетрудно понять, если



воспользоваться моделью свободных электронов. Продемонстрируем это на примере цианинового красителя, структурная формула которого приведена на рис. 11.2 а.  $\pi$ -электроны атомов углерода образуют два плоских облака, одно из которых расположено выше, а другое ниже плоскости молекулы (рис.11.2 б). Этим распределением  $\pi$ -электронов и определяются

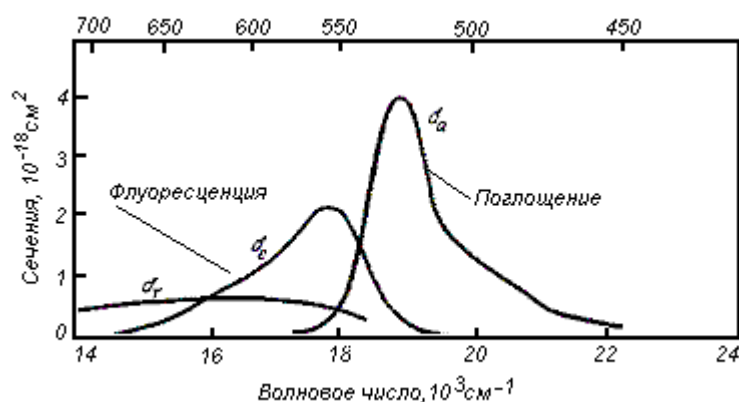


Рис. 11.1. Спектры флуоресценции для родомина 6G в этаноле

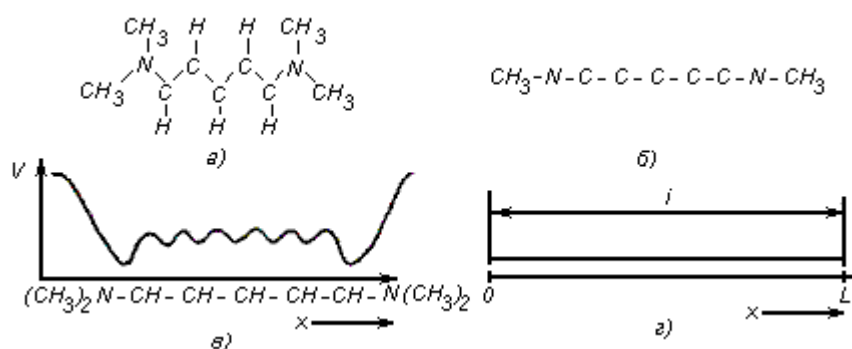


Рис. 11.2. Иллюстрация модели свободных электронов, объясняющей структуру электронных энергетических уровней молекулы красителя

электронные состояния молекул. В модели свободных электронов предполагается, что  $\pi$ -электроны движутся свободно в пределах их плоских распределений и их движение ограничено только отталкивательным потенциалом группы на каждом конце красителя. Поэтому энергетические уровни электронов представляют собой уровни свободного электрона, находящегося в потенциальной яме, форма которой показана на рис.11.2в. Если эту яму приближенно представить в виде прямоугольной (рис.11.2г), то энергетические уровни электрона в такой яме хорошо известны и определяются выражением  $E_n = h^2 n^2 / 8m_e L^2$ , где  $n$  — целое число,  $m_e$  — масса электрона, а  $L$  — ширина потенциальной ямы. Здесь важно отметить, что в  $\pi$ -электронном облаке (молекулярные системы с неспаренными электронами называются радикалами. Такие системы легко вступают в реакцию друг с другом, образуя при этом систему со спаренными электронами) молекулы красителя находится четное число электронов. Если предположить, что это число равно  $2N$ , то основное состояние молекулы будет соответствовать ситуации, при которой все электроны находятся на  $N$  нижних энергетических уровнях (см. рис.11.3). На самом деле каждый уровень может быть занят двумя электронами с

противоположными спинами. Таким образом, это молекулярное состояние имеет нулевой результирующий спин (синглетное состояние); на рис.11.3 оно обозначается через  $S_0$ . На этом же рисунке самый верхний занятый уровень и следующий над ним (пустой) уровень обозначены двумя квадратами, расположенными друг над другом. Первое возбужденное синглетное состояние (обозначенное на рисунке через  $S_1$ ) получается при переходе одного из двух высоколежащих электронов на следующий, более высокий уровень без переворота спина. Если спин при этом переворачивается, то образуется триплетное состояние (на рисунке  $T_1$ ; полный спин  $S=1$ ). Возбужденные синглетное ( $S_2$ ) и триплетное ( $T_2$ ) состояния получаются, когда электрон поднимается на следующий уровень, и т. д. Заметим, что на рис. 11.3 каждое электронное состояние в действительности состоит из ряда колебательных (жирные горизонтальные линии на рисунке) и вращательных (тонкие линии) уровней. Обычно расстояние между колебательными уровнями составляет  $1400\text{—}1700\text{см}^{-1}$ , а между вращательными уровнями в 100 раз меньше. Поскольку в жидкостях механизмы, вызывающие уширение линий, оказываются гораздо более сильными, чем в твердых телах, вращательные уровни в жидкостях не разрешаются, что приводит к сплошному спектру между колебательными уровнями.

Посмотрим теперь, что происходит с молекулой под действием электромагнитного излучения. Прежде всего напомним, что правила отбора требуют, чтобы  $\Delta S = 0$ . Следовательно, синглет-синглетные переходы являются разрешенными, а синглет-триплетные — запрещенными. Таким

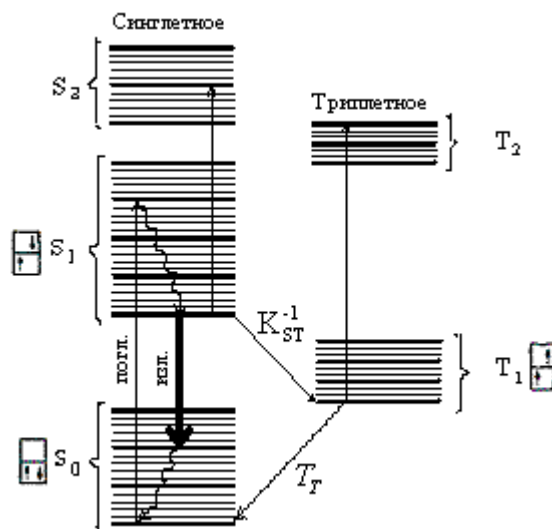


Рис. 11.3. Типичная схема энергетических уровней красителя в растворе. Синглетные и триплетные уровни приведены отдельно слева и справа

образом, в результате взаимодействия с электромагнитным излучением молекула может перейти из основного состояния  $S_0$  на один из колебательных уровней состояния  $S_1$ . При этом спектр поглощения представляет собой широкий бесструктурный переход (см., например, рис. 11.1), поскольку вращательные и колебательные уровни являются неразрешенными. Важная особенность красителей со-

стоит в том, что они имеют чрезвычай- но большую величину дипольного матричного элемента  $\mu$ . Это объясняется тем, что л-электроны свободно движутся на расстояниях, сравнимых с размером молекулы  $a$ , а поскольку  $a$  — достаточно большая величина, отсюда следует, что  $\mu$ , также велико ( $\mu \approx e a$ ). Таким образом, сечение поглощения, которое пропорционально  $a^2$ , также велико ( $\sim 10^{-16}$  см<sup>2</sup>). Молекула в возбужденном состоянии релаксирует за очень короткое время (безызлучательная релаксация,  $\tau_{\text{безызл}} \approx 10^{-12}$  с) на самый нижний колебательный уровень состояния  $S_1$  (точнее говоря, в этом случае будет происходить термализация колебательно-вращательных уровней состояния  $S_1$ ). С этого уровня она совершает излучательный переход на некоторый колебательный уровень состояния  $S_0$  (флуоресценция). Вероятность перехода определяется соответствующим фактором Франка—Кондона. Из сказанного ясно, почему флуоресценция имеет форму широкой и бесструктурной полосы, сдвинутой в длинноволновую область спектра по отношению к полосе поглощения (см. рис. 11.1). Перейдя на возбужденный колебательно-вращательный уровень основного состояния  $S_0$ , молекула возвращается на самый низкий колебательный уровень за счет другой очень быстрой (порядка пикосекунд) безызлучательной релаксации. Заметим, что из самого нижнего уровня состояния  $S_1$  молекула может также перейти в состояние  $T_1$ . Этот процесс обусловлен столкновениями и называется синглет-триплетной конверсией. Аналогичным образом преимущественно за счет столкновений осуществляется переход  $T_1 \rightarrow S_0$ . Этот переход частично происходит также и за счет излучательных процессов (излучательный переход  $T_1 \rightarrow S_0$ , как отмечалось выше, является запрещенным). Излучение, возникающее в результате таких переходов, называется фосфоресценцией. Будем характеризовать эти процессы релаксации следующими тремя постоянными: 1) временем жизни  $\tau_{\text{сп}}$  состояния  $S_1$ , определяемым спонтанным излучением; 2) скоростью (с<sup>-1</sup>) синглет-триплетной конверсии  $k_{ST}$  и 3) временем жизни  $\tau_T$  состояния  $T_1$ . Если время жизни состояния  $S_1$  обозначить через  $\tau$ , то можно записать следующее соотношение [см. 12]:

$$1/\tau = 1/\tau_{\text{сп}} + k_{ST}$$

Излучательное время жизни  $\tau_{\text{сп}}$ , очень невелико (всего несколько наносекунд), что обусловлено большой величиной матричного элемента  $\mu$ , дипольного момента. Поскольку  $k_{ST}^{-1}$  обычно гораздо больше ( $\sim 100$  нс), то, следовательно, наибольшее число молекул из состояния  $S_1$  будет релаксировать за счет флуоресценции. Вследствие этого квантовый выход флуоресценции (число испущенных за счет флуоресценции фотонов, деленное на число атомов, переведенных в состояние  $S_1$ ) близок к единице. Действительно, для квантового выхода имеем  $\eta = \tau/\tau_{\text{сп}}$ . Время жизни триплетного состояния  $\tau_T$  зависит от экспериментальных условий и, в частности, от содержания кислорода в растворе. Оно может колебаться от  $10^{-7}$  с в растворах, насыщенных кислородом, до  $10^{-3}$  с и более в бескислородных растворах.

### 11.2.1. Способы накачки лазеров на основе красителей

**Накачка излучением импульсных лазеров.** Для возбуждения растворов красителей в импульсном режиме чаще всего используются основные частоты излучения и гармоники твердотельных лазеров [вторая гармоника рубинового (347

нм), вторая (530 нм) третья (353 нм) и четвертая (285 нм) гармоники неодимовых лазеров], а также излучение импульсных газовых лазеров: азотного (337 нм) и ксенонового (172,5 нм). Мощность излучения накачки разных импульсных источников варьируется обычно в пределах от нескольких киловатт до десятков мегаватт.

При лазерном возбуждении наиболее распространены две основные геометрии накачки: поперечная, когда направление возбуждающего потока перпендикулярно направлению генерации красителя, и продольная, когда оба эти направления совпадают (рис. 11.4 а и б).

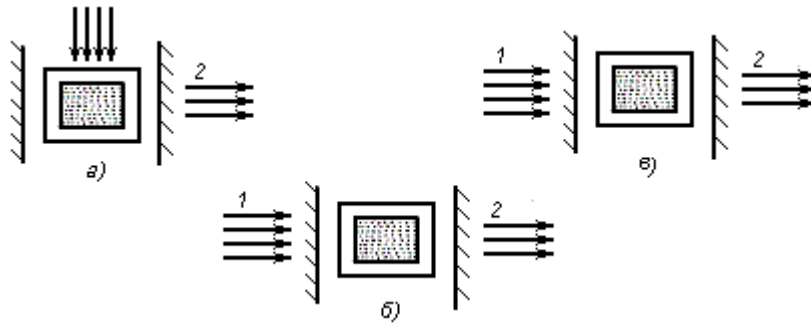


Рис. 11.4. Схемы накачки лазера на основе красителей: а — поперечная; б — продольная в — продольная с призмой полного внутреннего отражения. 1 — накачка, 2 — генерация

Генерация в продольной схеме особенно эффективна, если возбуждение проводится через селективное зеркало, хорошо пропускающее излучение накачки и обладающее высоким коэффициентом отражения в спектральной области генерации красителя. Этому требованию хорошо удовлетворяет схема с призмой полного внутреннего отражения (рис. 11.4, в), в которой излучение накачки с незначительными потерями на отражение проходит через призму со стороны ее ребра и слегка сфокусированным пучком попадает в объем активного вещества. Генерация красителя формируется в резонаторе, образованном призмой полного внутреннего отражения и выходным полупрозрачным зеркалом. Такая схема эффективно работает при любых длинах волн излучения накачки и генерации.

При оптимально подобранных параметрах раствора и резонатора энергетическая эффективность поперечной и продольной схем возбуждения генерации одинакова.

При возбуждении, в продольной схеме линейно поляризованным светом, излучение лазера на красителе всегда полностью поляризовано. В поперечной схеме степень поляризации генерируемого излучения зависит от угла между электрическим вектором возбуждающего излучения и осью резонатора лазера на красителе. Стопроцентная поляризация обычно наблюдается лишь при ортогональном направлении вектора поляризации накачки к оси лазера.

Лазерная накачка красителей может осуществляться с большой частотой повторения импульсов. При циркуляции раствора, обеспечивающей достаточно быструю смену активной среды в рабочей зоне лазера, частота следования импульсов генерации красителя ограничивается лишь рабочей частотой возбуждающего лазера. Для таких целей в настоящее время наиболее удобными являются лазеры

периодического действия (азотный, ксеноновый, на парах меди), а также на основе иттрий-алюминиевого граната с неодимом.

При возбуждении излучением твердотельных или газовых лазеров с синхронизацией мод в красителях может быть осуществлен режим генерации ультракоротких импульсов. Для этого достаточно, чтобы оптические длины резонаторов задающего генератора и лазера на красителе были равны или кратны. Длительность импульсов генерации при таком способе возбуждения может составлять около 5 пс [12].

К. п. д. лазеров на красителях с возбуждением импульсными лазерными источниками достигает десятков процентов. Максимальный к. п. д. (75%) был реализован для этанольного раствора родамина 6G при накачке излучением 2-й гармоники лазера на стекле с неодимом.

При возбуждении импульсными лазерами генерация на красителях может быть осуществлена практически на любых длинах волн в пределах спектрального диапазона от 340 до 1100 нм (рис. 11.5).

**Непрерывное возбуждение.** Источником непрерывной накачки красителя в большинстве случаев служит аргоновый ионный лазер, имеющий на нескольких сильных линиях в синей и зеленой частях спектра выходную мощность несколько ватт. Для превышения порога генерации красителя излучение аргонового лазера такой мощности должно фокусироваться в раствор в малую область диаметром 10 — 20 мкм. Эффективное использование энергии накачки обеспечивается лишь при согласовании с высокой точностью профилей зоны накачки и генерации красителя. Схема непрерывного лазера на красителе приведена на рис. 11.6 [22].

Одной из проблем, связанных с непрерывным оптическим возбуждением жидкостных лазеров, является устранение термооптических искажений, возникающих в активной жидкости и элементах резонатора при прохождении через них мощного пучка накачки. Обычно термооптические искажения раствора развиваются через несколько микросекунд после начала возбуждения. Поэтому, чтобы устранить их влияние, необходимо быстро прокачивать раствор через зону генерации, так чтобы за несколько микросекунд произвести полную смену жидкости в активной области. Аналогичные требования к прокачке раствора вызваны необходимостью вывода из зоны генерации продуктов фотораспада, образующихся под действием накачки. При скоростях тока, обеспечивающих смену раствора в генерирующем объеме за времена короче длительности метастабильного

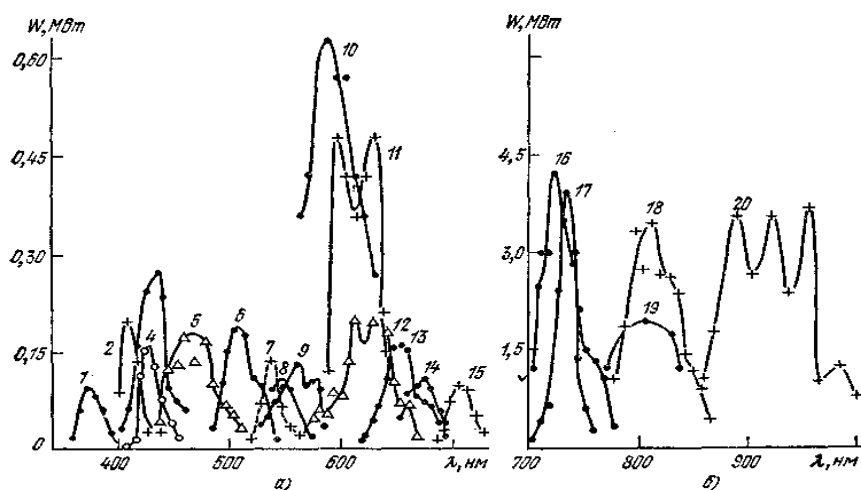


Рис. 11.5. Мощность генерации различных красителей в области 360—1000 нм при возбуждении излучением рубинового лазера: а — второй гармоники; б — основной частоты. 1) 2,5-дифенилоксазол; 2) 1,4-дистирилбензол; 3) 1-стирил-4 [ω-винил (п-бифенил)] бензол; 4) РОРОР; 5) эскулин; 6) тетрафенилбутадиеи; 7) амидин; 8) уранин; 9) флуоресцеин; 10) родамин 6G; //) родамин G; 12) полиметиновый № 4546; 13) полиметиновый № 4545; 14) 3,3'-диэтилоксадикарбоцианиинйодид; 15) полиметиновый №3115; 16) полиметиновый №4567; 17) 1.11-Дибензил-4,4'хинокарбоцианиинйодид; 18) 1.11, 3, 3,3'-гексаметилениндокарбоцианиинйодид; 19) полиметиновый № 29; 20) полиметиновый № 118; 21) 1,1'-диэтилхино- (2,2/)-трикарбоцианиинйодид

триплетного уровня ( $\tau_2 = 10^{-6}$  с), уменьшается вредное влияние Т-Т поглощения. В этом случае, несмотря на непрерывный режим накачки и генерации лазера на красителе, процесс возбуждения молекул красителя по существу носит кратковременный импульсный характер. Простая и надежная конструкция непрерывного лазера на красителе изображена на рис. 11.7 [12]. В этой схеме кювета с красителем отсутствует, а генерация происходит в тонкой плоскопараллельной струе раствора, формируемой с помощью специального сопла. Плоскость струи ориентирована под углом Брюстера к оси резонатора, чтобы свести к минимуму отражение от ее поверхности и скомпенсировать астигматизм сферических зеркал для наклонных пучков резонатора. Излучение Аг-лазера вводится в резонатор через призму.

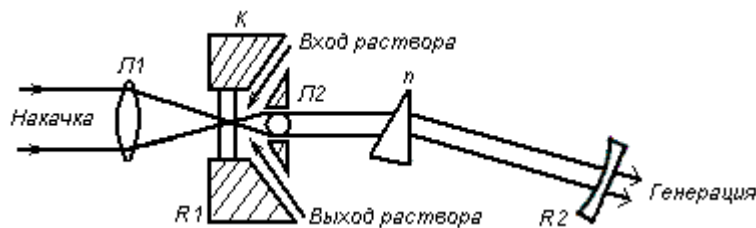


Рис. 11.6. Схема непрерывного лазера на красителе: Л1, Л2 — линзы; R1, R2 — зеркала; К — кювета; П — призма

Непрерывные лазеры на красителях генерируют во всей области видимого спектра и позволяют получать достаточно узкие перестраиваемые по частоте линии генерации с минимальной угловой расходимостью излучения (рис. 11.8). Мощности излучения современных непрерывных лазеров на основе красителей достигают нескольких ватт. Они обладают хорошей стабильностью (около 2% за время  $\sim 10$  мс), при этом флуктуации амплитуды их излучения определяются в основном флуктуациями мощности возбуждающего Аг-лазера.

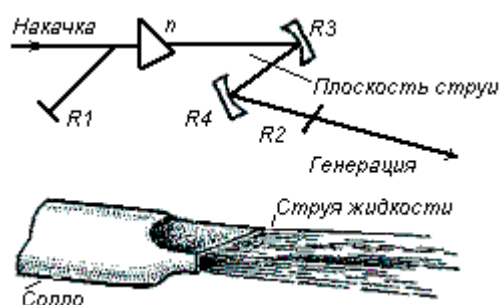


Рис. 11.7. Схема струйного лазера на красителе R1—R4 — зеркала резонатора; П — призма. Внизу расположены конфигурация сопла и вытекающей струи активного вещества

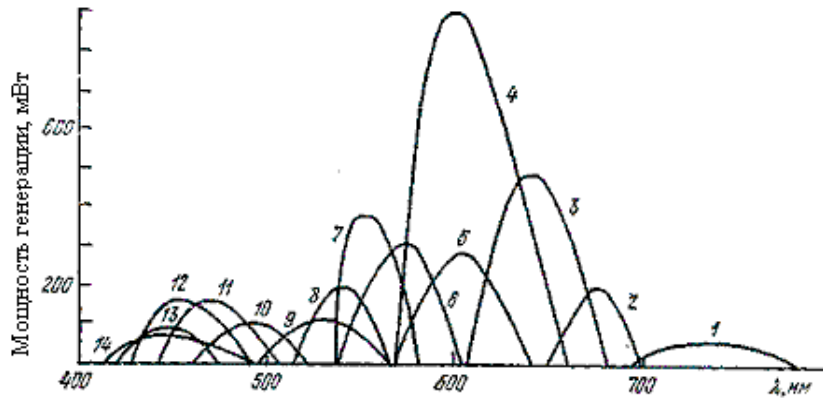


Рис. 11.8. Спектр мощности непрерывного струйного лазера при использовании различных органических красителей: 1 — нильский голубой; 2 — крезилфиолет перхлорат; 3 — родамин В; 4 — родамин 6G (мощность возбуждения 4 Вт); 5 — родамин 6G (мощность возбуждения 2 Вт); 6 — родамин 11O; 7 — флуоресцеин; 8 — кумарин В; 9 — кумарин 7; 10 — кумарин 102; 11 — 7-диэтиламино-4-метилкумарин; 12 — кумарин 2; 13 — кумарин 120; 14 — карбостирол (78)

Лазер на красителе с непрерывной накачкой может работать в режиме активной или пассивной синхронизации мод. При пассивной синхронизации необходима сильная фокусировка излучения лазера на красителе в кювету с насыщающимся поглотителем. Одна из возможных схем непрерывного лазера на красителе с пассивной синхронизацией мод приведена на рис. 11.9 [13].

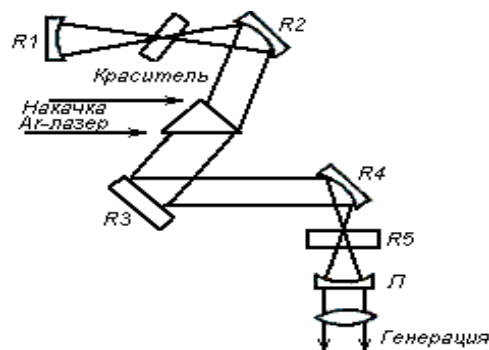


Рис. 11.9. Схема резонатора непрерывного лазера на основе красителя с пассивной синхронизацией мод насыщающимся поглотителем: R1—R4 — зеркала; Л — линза; З — затвор

Излучение накачки вводится в резонатор через кварцевую призму и фокусируется зеркалом на кювету с генерирующим раствором. На другом конце резонатора находится подобная кювета с раствором насыщающего поглотителя. Обе кюветы ориентированы под углом Брюстера к оси резонатора и предусматривают непрерывную прокачку растворов. Такой лазер способен генерировать импульсы длительностью 1—2 пс. Устойчивая синхронизация мод обычно наблюдается при не-

большом превышении уровня накачки над порогом генерации.

**Накачка импульсными лампами.** Генерация на красителях может быть возбуждена с помощью разнообразных источников некогерентного света: трубчатых и коаксиальных импульсных ламп, «пинчевых» источников, излучением разряда в воздухе и плазменным фокусом магнитоплазменного компрессора.

При использовании трубчатых ламп накачка красителя обычно производится при режимах разряда, обеспечивающих длительность импульсов излучения от 1 до 500 мкс при вводимой энергии от десятков джоулей до килоджоуля (в зависимости от типа ламп). Трубчатые лампы позволяют создавать сравнительно простые конструкции лазерных головок на основе красителей, излучающие в спектральной области 400 — 700 нм с энергией генерации в импульсе 5 — 10 Дж [17]. К. п. д. генерации лазера с накачкой трубчатыми лампами при использовании наиболее эффективных красителей; (родамин 6 G) превышает 1%..

Для возбуждения генерации красителей разработано несколько вариантов специальных коаксиальных импульсных ламп, обеспечивающих импульсы излучения накачки длительностью в несколько десятых долей микросекунды [18—20]. Отличительной особенностью конструкции таких ламп является весьма малая толщина кольцевого разрядного промежутка (десятые доли миллиметра). Осуществление разряда с такой малой длительностью предъявляет жесткие требования к индуктивности питающего конденсатора и другим элементам разрядной цепи.

Благодаря высокой интенсивности и малой длительности накачки коаксиальные лампы позволяют возбуждать генерацию большего числа красителей и в более широкой спектральной области (до 340 нм), чем трубчатые лампы. Однако выходная энергия генерации, достигаемая при использовании таких ламп, существенно ниже, чем при накачке лампами трубчатого типа.

С помощью коаксиальной лампы с высоким энерговыделением удастся осуществить генерацию излучения на красителе (родамине 6 G) с выходной энергией до 400 Дж при достаточно высоком к. п. д. (0,8%) [22].

Как правило, при ламповом возбуждении импульс генерации красителя короче импульса накачки и имеет беспичковый характер. При использовании активных или пассивных модуляторов добротности в генераторах на красителях с ламповым возбуждением легко осуществляется синхронизация мод. Лазеры на красителях с ламповой накачкой могут работать и в режиме периодического действия с частотой повторения импульсов  $\sim 100$  Гц. В последнем случае необходимым условием является прокачка рабочего раствора через кювету.

### **11.3. Спектральные характеристики и управление спектром лазера на основе красителей**

Спектр генерации органических красителей в резонаторе без селективных элементов при любом способе возбуждения обычно составляет 5 — 20 нм.[22]. В большинстве случаев спектр имеет четкую полосчатую структуру, обусловленную интерференцией генерируемого излучения при отражении от плоскопараллельных элементов резонатора (поверхностей зеркал, стенок кюветы и др.). У красителей, полосы, люминесценции которых заметно перекрываются с полосами поглощения, спектр генерации не совпадает с максимумом полосы люминесценции и сдвинут относительно него в длинноволновую часть спектра. Сред-



ная частота генерации таких красителей зависит от концентрации молекул в растворе, коэффициентов отражения зеркал, длины активного слоя и базы резонатора. Положение спектра генерации некоторых красителей существенно меняется также при изменении температуры раствора.

Для перестройки и сужения спектра генерации красителей в резонатор лазера вводятся спектрально-селективные элементы (селекторы).

В качестве селекторов применяются спектральные светофильтры, клиновые и плоско-параллельные интерферометры Фабри — Перо, дифракционные решетки, дисперсионные призмы и некоторые другие устройства. Типичные схемы таких генераторов с использованием селекторов приведены на рис. 11.10. Введение одного селективного элемента сужает линию генерации до 0,05 — 1,0 нм без существенного уменьшения выходной мощности и позволяет перестраивать спектр генерации в пределах нескольких сотен ангстрем на одном и том же растворе красителя. Для получения более узких линий генерации следует использовать комбинации различных селекторов. Получение линий шириной порядка  $10^{-3}$  —  $10^{-4}$  нм

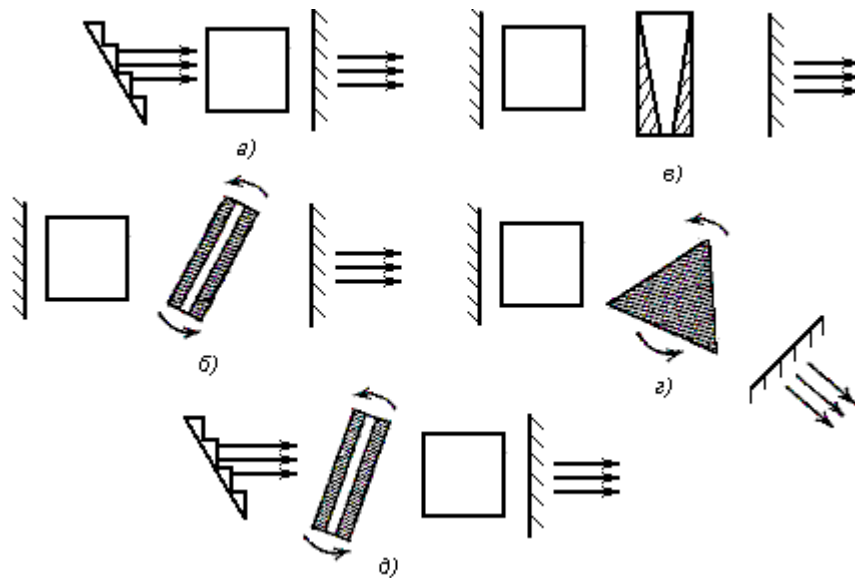


Рис. 11.10. Схема лазера на основе красителей с селекторами: а — с дифракционной решеткой; б—в — с плоскопараллельным и клиновым интерферометром Фабри — Перо, г — с дисперсионной призмой; д — с дифракционной решеткой и интерферометром Фабри — Перо

неизбежно связано со значительными потерями энергии генерируемого излучения, однако является весьма эффективным способом повышения спектральной плотности энергии генерации.

Генерация узкой перестраиваемой линии на красителе может быть получена и без внесения в резонатор селективных элементов, на принципе "распределенной обратной связи" (РОС). В лазерах с РОС обратная связь осуществляется за счет брэгговского отражения от периодической структуры, возникающей в активной среде в результате модуляции ее показателя преломления [22]. Вследствие резонансного характера брэгговского отражения лазерные колебания возбуждаются лишь в узкой спектральной области. Изменяя период модуляции, можно осуществить плавную перестройку частоты. Необходимая модуляция показателя преломления среды может быть получена, например, при интерференции двух когерент-

ных лучей в активной среде лазера на красителе. Генерация в лазере с РОС возбуждается на длине волны  $\lambda_r = n_p \lambda_n / n \sin \theta$ , где  $n_p$  и  $n$  — показатели преломления раствора и граничащей с ним среды;  $\lambda_n$  — длина волны накачки;  $\theta$  — угол интерференции пучков накачки. Перестройка длины волны генерации может осуществляться изменением угла пересечения интерферирующих лучей, изменением показателей преломления раствора или граничащей с ним среды, а также перестройкой длины волны излучения накачки  $\lambda_n$ . Спектральная ширина линии генерации в лазерах с РОС обычно составляет 0,01 — 0,001 нм.

## Глава 12

### ПОЛУПРОВОДНИКОВЫЕ ИЗЛУЧАТЕЛИ СВЕТА

В настоящее время широко используются оптические источники излучения двух типов- светоизлучающие диоды и лазерные диоды. Излучение светоизлучающих диодов некогерентно и имеет большую расходимость. Для светоизлучающего диода характерен широкий спектральный диапазон  $|\Delta\lambda/\lambda| \approx 0.03$ . Можно считать, что выходная мощность изменяется пропорционально току через диод, хотя имеется тенденция к насыщению при высоких уровнях мощности, что обусловлено ростом температуры полупроводника. Частота модуляции ограничивается 100 МГц, коэффициент полезного действия сравнительно невелик.

Работа лазерного диода более сложна. При низких токах накачки механизм излучения остается тем же, что и в светодиодах (спонтанное излучение), а после превышения порога генерация преобладает над спонтанным излучением. В результате изменяется и характер излучения – оно становится более направленным и когерентным, а в спектре появляется линейчатая структура. Как пороговый ток, так и спектр чувствительны к температуре и поэтому при изменении окружающих условий могут меняться. Одной из основных областей применения, в настоящее время, полупроводниковых излучателей – использование их в качестве оптического источника излучения в оптоволоконных системах связи. Поэтому, рассмотрим эти излучатели более подробно.

#### 12.1. Полупроводниковые лазеры

Квантовая электроника является, в основном, электроникой связанных состояний, которым соответствуют дискретные уровни энергии и относительно узкие линии резонансных переходов. Чем более изолирован от внешних воздействий электрон активного центра, тем в большей степени его время жизни в связанном состоянии определяется временем спонтанного распада и тем ближе ширина линии резонансного перехода к предельно узкой. И наоборот, чем больше подвержен внешним воздействиям тот электрон, переходы между связанными состояниями которого - рассматриваются как возможные лазерные переходы, тем, вообще говоря, шире соответствующие линии люминесценции и усиления. Примеры хорошо известны. Это уже рассмотренные выше лазеры.

До сих пор мы обсуждали лазеры, активные центры которых характеризуются наличием относительно узких дискретных уровней энергии. Однако получение инверсии возможно и в тех случаях, когда в энергетическом спектре активных центров выделяются широкие энергетические зоны разрешенных состояний, разделенные отчетливо проявляющейся запрещенной зоной. Примером являются полупроводниковые лазеры.

Отличительной особенностью полупроводниковых лазеров является инверсия на переходах между состояниями в электронных энергетических зонах полупроводникового кристалла. Эти энергетические зоны возникают вследствие расщепления уровней энергии валентных электронов атомов, составляющих кри-

сталлическую решетку в сильном периодическом, внутрикристаллическом поле собственных атомов кристалла (см. рис. 12.1а) В этом поле индивидуальные атомы теряют свои валентные электроны, локализованность которых, исчезает. Другими словами, в сильном периодическом поле валентные электроны атомов полупроводникового кристалла обобществляются. Коллективные движения обобществленных электронов в разрешенных энергетических зонах полупроводникового кристалла обеспечивают электропроводность кристалла.

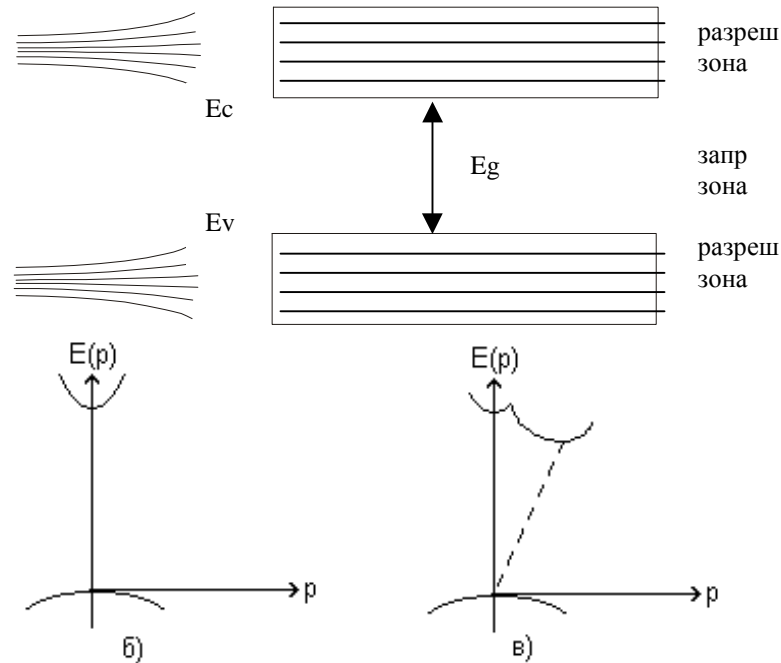


Рис. 12.1. а) Энергетические зоны в полупроводниковом кристалле, б) и в) Межзонные оптические переходы в полупроводниках

Как известно, носителями тока в полупроводниках являются электроны в зоне проводимости и дырки в валентной зоне. Пусть, с помощью какого-либо внешнего воздействия (накачки), созданы избыточные, по отношению к равновесным, электроны в зоне проводимости и дырки в валентной зоне. Возвращение к равновесию, т. е. рекомбинация избыточных электронов и дырок, может происходить излучательно на оптическом переходе зона — зона (рис. 12.1а)

Вероятность излучательных переходов зона — зона велика, если переходы являются прямыми, т. е. если они могут быть изображены вертикальной прямой на диаграмме, показывающей зависимость энергии  $E$  носителей тока от их квазиимпульса  $p$  (рис. 12.1б). Дело в том, что оптические переходы зона — зона завершаются созданием свободных носителей тока. Следовательно, при переходе должен выполняться закон сохранения импульса. Ввиду малости импульса фотона это означает, что при переходе квазиимпульс электрона должен остаться неизменным. В случае, так называемых прямозонных полупроводников, для которых максимум потенциальной кривой, ограничивающей валентную зону сверху, и минимум потенциальной кривой, ограничивающей зону проводимости снизу, имеют

место при одном и том же значении квазиимпульса  $p$  (рис. 12.1в), вертикальные переходы соединяют потолок валентной зоны и дно зоны проводимости, т. е. те части обеих разрешенных зон, которые в первую очередь заселяются избыточными носителями тока. В этом случае переход разрешен, и создание неравновесной заселенности максимально облегчено. Прямозонные полупроводники благоприятны созданию лазеров. В случае непрямозонного полупроводника (рис. 12.1в) носители тока накапливаются в состояниях с разными значениями квазиимпульса, вертикальные переходы невозможны из-за отсутствия рекомбинирующих частиц, а невертикальные (непрямые) переходы запрещены.

Большая вероятность излучательных переходов в прямозонных полупроводниках и большая плотность состояний в зонах позволяет получать исключительно высокие значения коэффициентов усиления, превышающие в некоторых особо благоприятных случаях  $10^4 \text{ см}^{-1}$ . В лазерах на основе твердых диэлектриков с примесными активными центрами, рассмотренных в предыдущих разделах и называемых обычно твердотельными лазерами, в отличие от полупроводниковых лазеров используются переходы между дискретными уровнями энергии изолированных ионов. Плотность активных частиц и вероятность переходов в них существенно ниже, поэтому для них характерны значительно меньшие коэффициенты усиления и, следовательно, заметно большие линейные размеры активных элементов.

Важным отличием полупроводниковых лазеров от остальных лазеров на конденсированных средах является электропроводность полупроводников, позволяющая осуществлять их накачку электрическим током и тем самым непосредственно преобразовывать электрическую энергию в лазерное излучение.

Таким образом, полупроводниковые лазеры занимают в квантовой электронике особое место, отличаясь от всех прочих лазеров рядом важных-особенностей.

С практической точки зрения наиболее существенны следующие достоинства полупроводниковых лазеров:

1. Компактность, обусловленная гигантским коэффициентом усиления в полупроводниках.
2. Большой к. п. д., обусловленный высокой эффективностью преобразования подводимой энергии в лазерное излучение, при накачке достаточно совершенных полупроводниковых монокристаллов электрическим током.
3. Широкий диапазон длин волн генерации, обусловленный возможностью выбора полупроводникового материала с шириной запрещенной зоны, соответствующей излучению на переходах зона — зона практически в любой точке спектрального интервала от 0,3 до 30 мкм.
4. Плавная перестройка длины волны излучения, обусловленная зависимостью спектрально-оптических свойств полупроводников и прежде всего ширины запрещенной зоны от температуры, давления, магнитного поля и т. п.
5. Малоинерционность, обусловленная малостью времен релаксации и практически безынерционностью создания неравновесных электронов и дырок при накачке электрическим током, приводящая к возможности модуляции излучения изменением тока накачки с частотами, достигающими 10 ГГц.
6. Простота конструкции, обусловленная возможностью накачки постоянным током и приводящая к совместимости полупроводниковых лазеров с интеграль-

ными схемами полупроводниковой электроники, устройствами интегральной оптики и волоконных оптических линий связи.

Недостатки полупроводниковых лазеров являются, как это часто бывает, продолжением их достоинств. Малые размеры приводят к низким значениям выходной мощности или энергии. Кроме того, полупроводниковые лазеры, как и все приборы полупроводниковой электроники, чувствительны к перегрузкам (разрушаются при потоках оптического излучения в несколько мегаватт на квадратный сантиметр) и к перегреву, приводящему к резкому повышению порога самовозбуждения и даже к необратимому разрушению при нагреве свыше некоторой характерной для каждого типа лазера температуры.

Лазерная генерация получена при использовании многих различных полупроводниковых материалов, общим числом в несколько десятков. Эти материалы, кроме обычного требования чистоты и монокристалличности, должны обладать высокой оптической однородностью и малой вероятностью безызлучательной рекомбинации электронов и дырок [26,27,28.29].

### 12.1.1.Рекомбинация электронов и дырок

Рекомбинацией электронов и дырок в полупроводниках называется процесс, приводящий к переходу электрона из зоны проводимости в валентную зону, в результате которого происходит исчезновение пары электрон проводимости — дырка валентной зоны. Рекомбинация всегда означает переход носителя тока на наиболее низкие уровни энергии — либо в валентную зону, либо на примесный уровень в запрещенной зоне.

При термодинамическом равновесии рекомбинация уравнивает процесс тепловой генерации носителей, а скорость этих взаимно обратных процессов такова, что их совместное действие приводит к установлению распределения Ферми для электронов и дырок по энергиям.

Известно много механизмов рекомбинации, которые отличаются друг от друга направлением передачи энергии, выделяемой при рекомбинации. Если избыток энергии выделяется в виде кванта излучения, то происходит излучательная рекомбинация. Этот элементарный акт генерации света в полупроводниках подобен излучательному распаду возбужденного состояния в системах с дискретным спектром. Возможна также безызлучательная рекомбинация, при которой высвобождающаяся энергия расходуется на возбуждение колебаний кристаллической решетки, т. е. в конечном счете, идет на нагрев кристалла. Очевидно, безызлучательная рекомбинация полностью подобна безызлучательной релаксации энергии возбуждения в системах с дискретным спектром.

При излучательной рекомбинации полное число актов излучения пропорционально произведению  $np$  концентраций электронов ( $n$ ) и дырок ( $p$ ). При небольших концентрациях носителей этот канал рекомбинации мало эффективен. При больших концентрациях, превышающих  $10^{16}$ - $10^{17}$   $см^{-3}$ , полупроводники становятся эффективными источниками света рекомбинационного излучения в относительно узком диапазоне длин волн вблизи края собственного поглощения полупроводника.

Рекомбинационное излучение происходит на переходах зона — зона (см. рис. 12.1а). Межзонный характер носит рекомбинационная люминесценция не только

в собственных, но и в сильно легированных полупроводниках. В последнем случае, энергетический спектр полупроводника вблизи краев запрещенной зоны сильно искажен ввиду того, что при сильном легировании примесные уровни уширены. При этом энергия квантов рекомбинационного излучения может несколько отличаться от номинальной ширины запрещенной зоны легируемого полупроводника. Однако практически для всех процессов излучательной рекомбинации, используемых в полупроводниковых лазерах, характерна близость энергии перехода  $\hbar\omega$  к энергии запрещенной зоны  $E_g$ .

Наличие спонтанной излучательной рекомбинации свидетельствует о возможности создания лазера. Для того чтобы в спектре излучения спонтанной рекомбинации возникло усиление, вынужденное испускание фотонов должно преобладать над их поглощением. Необходимым условием для этого является наличие инверсии населенностей. В полупроводниковых лазерах, следовательно, должна существовать инверсия на переходах излучательной рекомбинации.

### 12.1.2. Условия и методы получения инверсии населенностей в полупроводниках

**Условия получения инверсии.** Анализ условий достижения инверсии требует знания уровней энергии, сечений их возбуждения, времен релаксации и т. п. При общем рассмотрении такого широкого класса лазерных активных сред, какими являются полупроводники, этот путь нерационален, даже если и возможен. К счастью, достаточно общие термодинамические соображения, учитывающие вместе с тем специфику статистики электронов в полупроводниках, могут дать общие условия инверсии в этих материалах.

Независимо от конкретного механизма излучательной рекомбинации, возникающие при этом фотоны, подчиняются общим законам теории излучения. Скорость заполнения фотонами частоты  $\omega$  некоторой радиационной моды объема  $V$  составляет

$$dm_\omega / dt = (A + Bm_\omega) / V, \quad (12.1)$$

где  $m_\omega$  — число фотонов, имеющих в моде;  $B$  — коэффициент вынужденного испускания. Первый член в (12.1) обусловлен спонтанным испусканием, второй соответствует разности скоростей вынужденного испускания и поглощения фотонов. В рассматриваемом случае излучательной рекомбинации в каждом акте испускания фотона одна электронно-дырочная пара исчезает, а в каждом акте поглощения одна такая пара возникает.

Связь между коэффициентами спонтанного  $A$  и вынужденного  $B$  излучений можно найти из термодинамических соображений. При температуре  $T$  в соответствии со статистикой Бозе-Эйнштейна равновесное число заполнения моды (для двух поляризаций) составляет

$$N = 2 / (\exp(\hbar\omega/kT) - 1). \quad (12.2)$$

При равновесии  $dm_\omega / dt = 0$ , следовательно, в силу (12.1) имеем

$$A / B = -m_\omega. \quad (12.3)$$

Дальнейший анализ требует учета специфики полупроводников. Выделим в спектре электронных состояний два уровня с энергиями  $E_2 > E_1$ . Скорость излучательной рекомбинации на переходе  $E_2 \rightarrow E_1$  пропорциональна произведению концентраций электронов на уровне  $E_2$  и дырок на уровне  $E_1$ .

Электроны, как известно, подчиняются статистике Ферми — Дирака. Вероятность того, что электрон находится в состоянии с энергией  $E$ , задается распределением Ферми (3.1).

Вероятность найти на уровне энергии  $E$  дырку равна вероятности того, что этот уровень не занят электроном, и составляет, следовательно,

$$1-f(E) = \exp((F-E) / kT + 1)^{-1}. \quad (12.4)$$

Тогда скорость спонтанной рекомбинации, пропорциональная числу электронов на уровне  $E_2$  и числу дырок на уровне  $E_1$ , может быть представлена в виде

$$A = A_0 f(E_2) (1-f(E_1)), \quad (12.5)$$

где  $A_0$  — коэффициент пропорциональности.

Аналогично, величина  $B$ , определяющая в (12.1) разность скоростей вынужденного излучения и поглощения, составляет

$$B = B_2 f(E_2) (1-f(E_1)) - B_1 f(E_1) (1-f(E_2)), \quad (12.6)$$

где  $B_1$  и  $B_2$  — коэффициенты пропорциональности. Подставляя эти выражения для  $A$  и  $B$  в (12.3) и учитывая равновесные распределения (12.2) и (3.1), мы получаем уравнение [26]

$$e^{\frac{h\omega}{kT}} - 1 = 2 \frac{B_2}{A_0} \left( \frac{B_2}{B_1} e^{\frac{E_2-E_1}{kT}} - 1 \right). \quad (12.7)$$

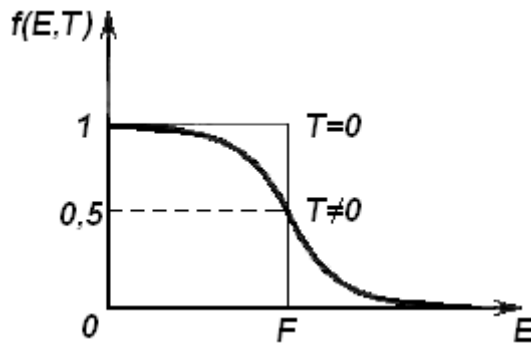


Рис. 12.2. Распределение Ферми ( $F$  — уровень Ферми).

По предположению, фотоны с энергией  $h\omega$ , возникают в результате прямозонной излучательной рекомбинации, происходящей между уровнями  $E_2$  и  $E_1$ ; значит,  $h(\omega) = E_2 - E_1$ . Тогда уравнение (12.7) всегда удовлетворяется, если  $B_2 - B_1 = A_0/2$ . Это означает, что

$$B = A_0/2 [(f(E_2) - f(E_1))]. \quad (12.8)$$

Следовательно, разность между скоростями вынужденного испускания фотонов на переходе зона — зона при излучательной рекомбинации и поглощения фотонов на том же переходе положительна при условии

$$f(E_2) > f(E_1). \quad (12.9)$$

Если  $f(E)$  — распределение Ферми (12.4), представленное на рис. 12.2, то для пары уровней  $E_2 > E_1$  при термодинамическом равновесии это условие не выполняется.

Вспомним теперь, что уровни  $E_2$  и  $E_1$  разделены запрещенной зоной и находятся соответственно в зоне проводимости и в валентной зоне. Неравновесные носители, т. е. неравновесные электроны в зоне проводимости и дырки в валентной зоне, создаваемые каким-то источником накачки, обладают конечными временами жизни в зонах. На протяжении этого времени термодинамическое равновесие не соблюдается, и, следовательно, единого уровня Ферми для всей системы



в целом не существует. Однако, если в электронном и дырочном газах в отдельности за время, меньшее времени жизни носителей в зонах, устанавливается свое квазиравновесное состояние фермиевского типа, но с одной и той же температурой для всей системы, а равновесие между этими газами отсутствует, то можно ввести так называемые квазиуровни Ферми отдельно для электронов в зоне проводимости  $F_n$  и дырок в валентной зоне  $F_p$ .

Допустимость введения отдельных квазиравновесных функций распределения для дырок в зоне проводимости и электронов в валентной зоне оправдывается тем, что, по крайней мере, в нескольких полупроводниках время термализации электронов в пределах зоны (0,1 пс) на три-четыре порядка меньше характерного времени межзонной термализации (1—10 нс).

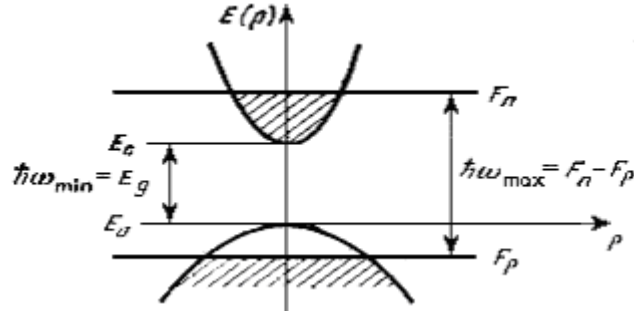


Рис. 12.3. Инверсное распределение электронов в однодолинном полупроводнике.

Вернемся к условию инверсии (12.9). Представим, в соответствии с проведенным выше обсуждением,  $f(E_2)$  и  $f(E_1)$  в виде

$$f(E_2) = (1 + \exp(E_2 - F_n) / kT)^{-1}; \quad f(E_1) = (1 + \exp(E_1 - F_p) / kT)^{-1}, \quad (12.10)$$

где  $F_n$  — квазиуровень Ферми для электронов, а  $F_p$  — для дырок, из неравенства (12.9) легко получить эквивалентное ему, но более наглядное условие:

$$F_n - F_p > E_2 - E_1. \quad (12.11)$$

Так как минимальное значение  $E_2 - E_1$  равно ширине запрещенной зоны  $E_g$ , условие инверсии на межзонных переходах приобретает простой вид:

$$F_n - F_p > E_g. \quad (12.12)$$

Следовательно, накачка, создающая неравновесность, должна быть достаточно сильной для того, чтобы квазиуровни Ферми оказались внутри соответствующих разрешенных зон энергии. Это означает, что электронный и дырочный газы вырождены, при этом все уровни в валентной зоне с энергией  $E > F_p$  практически полностью незаселены, а все уровни в зоне проводимости с энергией  $E < F_n$  практически полностью заселены электронами (рис. 12.3). Тогда фотоны, энергия которых лежит в интервале

$$E_g < \hbar \omega < F_n - F_p, \quad (12.13)$$

не могут вызвать переходы валентная зона — зона проводимости и поэтому не поглощаются. Обратные переходы зона проводимости — валентная зона возможны. Вынужденная излучательная рекомбинация на этих переходах как раз и создает лазерное излучение. Неравенства (12.13) определяют ширину полосы соответствующего усиления.

Интенсивность излучательной рекомбинации определяется конкретными особенностями зонной структуры полупроводника, квадратом матричного элемента соответствующего перехода, плотностью рекомбинирующих пар. Скорость излучательной рекомбинации, т. е. число переходов с излучением в единичном

объеме в единицу времени, зависит, очевидно, от тех же факторов. Так как излучательный и безызлучательный каналы рекомбинаций параллельны, то результирующая скорость рекомбинации  $1/\tau$  равна сумме скоростей рекомбинаций излучательной ( $1/\tau_{изл}$ ) и безызлучательной ( $1/\tau_{безызл}$ ):

$$1/\tau = 1/\tau_{изл} + 1/\tau_{безызл}. \quad (12.14)$$

Очевидно, относительная доля актов излучательной рекомбинации во всем процессе рекомбинации в целом равна отношению соответствующих скоростей

$\frac{1/\tau_{изл}}{1/\tau}$ . Это же отношение указывает долю всей энергии, заключенной в неравно-  
весных электронах и дырках, которая высвобождается по излучательному каналу рекомбинации. Тем самым мы определили, так называемый внутренний квантовый выход излучательной рекомбинации:

$$\eta_{внутр} = (1/\tau_{изл}) / (1/\tau_{изл} + 1/\tau_{безызл}). \quad (12.15)$$

Эта величина характеризует качество полупроводникового материала. Правильный выбор легирования и изготовление совершенных кристаллов позволяет получать для многих полупроводниковых материалов значение  $\eta_{внутр}$ , приближающееся к 100%.

Для грубой оценки порядка величин скоростей безызлучательной рекомбинации можно пользоваться экспериментально определенными значениями сечений соответствующих процессов. При концентрации носителей  $10^{19} \text{ см}^{-3}$  эффективная скорость рекомбинации составляет  $10^6 - 10^{13} \text{ с}^{-1}$ . Приведенные оценки показывают, что в реальных материалах действительно возможно достижение внутреннего квантового выхода излучательной рекомбинации, приближающегося к 100%..

Итак, при выполнении условия инверсии и неравенства  $\tau_{изл} \ll \tau_{безызл}$  в полупроводниковых кристаллах возможно эффективное лазерное излучение.

### **Методы получения инверсной населенности в полупроводниках.**

Для создания в полупроводнике состояния инверсии населенностей (состояния с отрицательной температурой) необходимо, чтобы число электронов, обладающих энергиями зоны проводимости, превышало число электронов, находящихся в валентной зоне. Это условие может быть выполнено следующим образом. Через полупроводник, состоящий из  $p$ -,  $n$ -частей в прямом направлении пропускается большой ток. При этом в  $p$ -области возникает большое число электронов проводимости, а в  $n$ -области – большой избыток дырок. Такой способ получения неравновесного распределения называется инжекцией носителей через  $p - n$  переход.

Наличие запрещенных зон с большой энергетической шириной требует для создания инверсии больших токов внешнего источника, что вызывает большие тепловые потери в малых объемах  $p - n$  перехода. Поэтому иногда приходится охлаждать полупроводники до температур жидкого азота или жидкого гелия.

Другим методом возбуждения полупроводников является оптическая накачка. Недостаток этого метода возбуждения (с помощью обычных световых источников) – малый к. п. д. вследствие широкого энергетического спектра излучения источников накачки. Кроме того, затруднено просветление поверхности возбуждаемого образца для широкого спектра, что приводит к значительному отражению

падающего света. Использование монокроматического излучения (например, излучения лазера на рубине) позволяет устранить часть трудностей, но при этом из-за преобразования более коротковолнового излучения в длинноволновое неизбежны энергетические потери.

Наилучшие результаты получены Н. Г. Басовым и его сотрудниками, применившими для накачки полупроводниковых лазеров рубиновый оптический квантовый генератор, энергия фотонов которого близка к ширине запрещенной зоны арсенида галлия, вследствие этого потери на поглощение относительно малы, Полученная выходная импульсная мощность достигала 30 кВт.

Для оптической накачки применимо также двухфотонное возбуждение, при котором ширина запрещенной зоны равна удвоенной энергии фотона падающего излучения. Такая накачка применима для полупроводникового (п/п) лазера на арсениде галлия при облучении светом неодимового лазера ( $\lambda = 1,06 \text{ мкм}$ ) и для п/п на сернистом кадмии (накачка рубиновым лазером).

Еще одним способом накачки полупроводниковых лазеров является накачка электронным пучком. Пучок электронов, обладающих энергией более  $20 \text{ кэВ}$ , формируемый электронной пушкой, направляется на полупроводниковый кристалл. Электроны, проникая в глубь кристалла на несколько микрон (микрометров), теряют около 30% своей энергии на создание электронно-дырочных пар, которые затем при рекомбинации испускают фотоны света. Один падающий электрон создает примерно  $10^4$  электронно-дырочных пар.

Высокие плотность тока и энергия электронов в пучке приводят к сильному разогреву полупроводника и требуют интенсивного охлаждения.

Электроннолучевая накачка позволяет создавать излучатели и на тех полупроводниках, в которых не могут быть созданы удовлетворительные  $p - n$  переходы, что обеспечивает работу п/п излучателей этого типа практически во всем диапазоне – от ультрафиолетовой до дальней инфракрасной области спектра. Коэффициент полезного действия п/п лазера с электроннолучевой накачкой (не считая к. п. д. электронной пушки) теоретически не может превышать 12%.

Наиболее совершенным методом возбуждения полупроводника (из рассмотренных) является инжекция носителей через  $p - n$  переход. В этом случае отношение энергии когерентного излучения ко всей затраченной энергии при определенных условиях может быть близким к единице, так как электрическая энергия непосредственно преобразуется в световую.

### 12.1.3. Инжекционные полупроводниковые лазеры

Как выше отмечено, наиболее распространенным является возбуждение полупроводников постоянным током, осуществляющим инжекцию электронов и дырок в область  $p-n$  - перехода полупроводникового диода. Создаваемые таким образом, так называемые инжекционные (или диодные) полупроводниковые лазеры, получили наибольшее распространение в силу своей простоты, надежности и высокого к.п.д.

При контакте двух полупроводников, из которых один имеет проводимость  $p$ -типа, а другой  $n$ -типа, в пограничном слое возникают потенциальные барьеры, которые заметно изменяют концентрации носителей внутри слоя. Свойства пограничного слоя, т. е. области перехода от полупроводника  $p$ -типа к полупровод-

нику  $n$ -типа или, иначе говоря, области электронно-дырочного ( $p$ — $n$ ) перехода, зависят от приложенных напряжений.. Такая зависимость во многих случаях приводит к нелинейной вольтамперной характеристике полупроводникового контакта, приобретающего свойства полупроводникового диода.

Хорошо известно, что  $p$ - $n$  - переходы широко применяются в современной полупроводниковой электронике для выпрямления электрического тока, преобразования, усиления и генерации электрических колебаний. Мы кратко рассмотрим их лазерное применение, основанное на создании неравновесных носителей в области перехода.

$p$ - $n$ -переходы создаются путем образования внутри единого монокристалла нужного распределения донорной и акцепторной примесей (рис. 12.4). Если эти примеси полностью ионизованы, то в левой части кристалла, где велика концентрация акцепторов  $N_a$ , доминирует дырочная проводимость с концентрацией основных носителей  $p = N_a - N_d$ . Соответственно, в правой части кристалла преобладает электронная проводимость с концентрацией основных носителей  $n = N_d - N_a$  ( $N_d$ —концентрация доноров). Между  $p$ - и  $n$ -областями располагается переходный слой технологического  $p$  —  $n$ -перехода, в котором концентрации примесей резко изменяются по толщине слоя. Для некоторой тонкой, области переходного слоя, где доноры и акцепторы компенсируют друг друга ( $N_d = N_a$ ), характерна собственная проводимость ( $i$ ). Строго говоря,  $p$ - $n$ -переход, по существу, является  $p$ - $i$ - $n$ -переходом.

Требуемое распределение доноров и акцепторов создается различными технологическими приемами. Для кремния и германия (IV группа периодической таблицы Менделеева) акцепторами являются элементы III группы, а донорами — элементы V группы, например бор и фосфор соответственно. Для арсенида галлия  $GaAs$  — полупроводникового соединения типа  $A^3B^5$  — акцепторной примесью являются цинк и кадмий (II группа), а донорной — селен и теллур (VI группа). В лазерных диодах целесообразно применять сильно легированные полупроводники, в которых электронный и дырочный газы в  $n$ - и  $p$ -областях сильно вырождены.

Критерием вырождения фермионов является существенное отличие их распределения по энергиям от больцмановского. Из формулы (3.1) и рис.12.2 видно, что это отличие, малое при  $\exp[(E-F)/kT] \gg 1$ , заметно уже при  $\exp [(E-F)/kT] < 1$  (слабое вырождение) и становится крайне существенным при  $\exp [(E-F)/kT] \ll 1$  (сильное вырождение). Так как в запрещенной зоне электронов нет, нижней границей их энергий является положение дна зоны проводимости  $E_c$ . Следовательно, условием сильного вырождения электронов является выполнение неравенства [4,26,29]

$$\exp((E_c - F_n) / kT) < 1. \quad (12.16)$$

Это означает, что уровень Ферми электронов  $F_n$  должен лежать внутри зоны проводимости. Аналогично, при сильном вырождении дырок их уровень Ферми  $F_p$  должен лежать внутри валентной зоны. Тогда даже без инжекции уровень Ферми лежит в  $p$ -области внутри валентной зоны и в  $n$ -области—внутри зоны проводимости (рис. 12.5а).

В равновесном полупроводнике с  $p$  —  $n$ -переходом в отсутствие тока через переход, концентрация основных носителей — дырок в дырочной области  $p_p$  велика и постоянна. В области перехода концентрация дырок уменьшается и в элект-

тронной области, где дырки являются неосновными носителями, принимает малое значение  $p_n$ .

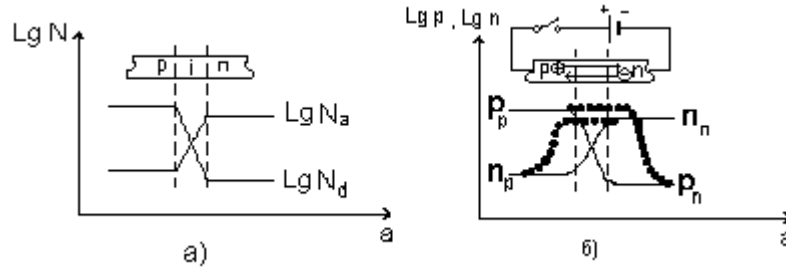


Рис. 12.4. Схема  $p$ - $n$ -перехода(а) и инжекции носителей в  $p$ - $n$ -переходе(б)

Аналогично, концентрация электронов изменяется от большого значения  $n_n$  в  $n$ -области (основные носители) до малого значения  $n_p$  в  $p$ -области (неосновные носители).

**Принцип действия полупроводникового лазера.** Пусть мы имеем два вырожденных полупроводника  $n$  и  $p$ -типов. Если теперь соединить эти два полупроводника, то в момент соединения в области соединения, называемой  $p$ - $n$ -переходом, возникает небольшой градиент электронов и дырок. В силу этого электроны, находящиеся в зоне проводимости  $n$ -области, будут переходить на более низкие энергетические уровни в  $p$ -области и в ней, вблизи  $p$ - $n$ -перехода, будет образовываться отрицательный объемный заряд. Такой же положительный заряд образуется в  $n$ -области. Образование этих зарядов приводит к появлению на  $p$ - $n$ -переходе контактной разности потенциалов  $U_k$ , которая препятствует дальнейшему движению электронов и дырок из области в область. В полупроводниках устанавливается равновесие, обязательным условием которого является то, что в обеих областях уровень Ферми должен быть одинаковым  $F_n = F_p$ . Границы энергетических зон при этом смещаются (рис. 12.5а) на величину, пропорциональную контактной разности потенциалов:

$$U_k = \frac{E_c^p - E_c^n}{e} = \frac{E_v^p - E_v^n}{e}.$$

Величина  $U_k$  определяется концентрацией электронов и дырок в  $n$  и  $p$ -областях

$$\frac{n_p}{n_n} = \frac{p_n}{p_p} = e^{(E_c^n - E_c^p)/kT} = e^{-eU_k/kT}$$

Откуда

$$U_k = \frac{kT}{e} \ln \frac{n_n}{n_p} = \frac{kT}{e} \ln \frac{p_p}{p_n}.$$

В этих выражениях  $n_n$  и  $n_p$  обозначают концентрацию электронов в  $n$  и  $p$ -областях соответственно, а  $p_n$  и  $p_p$  - обозначения для концентрации дырок,  $e$ -заряд электрона.

Очевидно, что в первый момент времени, в месте соединения, может выполниться условие инверсии населенностей (12.12). При переходе из области в область электроны и дырки рекомбинируют, излучая свет, который будет усиливаться до тех пор, пока выполнено условие (12.12). По мере установления равновесия  $F_n$  стремится к  $F_p$  и условие (12.12) нарушается. Поэтому при тепловом равновесии усиление излучения в  $p$ - $n$ -переходе получить не возможно.

Для того, чтобы снова создать инверсию населенностей в  $p-n$  – переходе, нужно добиться, чтобы уровни  $F_n$  и  $F_p$  снова разошлись на расстояние большее ширины запрещенной зоны  $E_g$ . Это возможно, если к переходу приложено внешнее напряжение так, что «плюс» источника напряжения соединен с  $p$  - областью, а «минус» с  $n$ -областью, то через переход течет положительный ток (через диод ток протекает в прямом направлении). При этом дырки из  $p$  - области устремляются в  $n$ -область, а электроны из  $n$ -области в  $p$  - область. Дырки, пришедшие в  $n$ -область, и электроны, пришедшие в  $p$  - область, становятся в этих областях неосновными носителями. Они должны рекомбинировать с соответствующими им основными носителями, имеющимися в  $p$ - и  $n$ -областях, прилегающих к  $p-n$ -переходу. При этом уровни Ферми смещаются на величину  $eU$ :

$$F_n - F_p = eU. \quad (12.17)$$

Если напряжение  $U$  включено встречно  $U_k$  (прямое смещение), так что суммарное напряжение на переходе равно  $\Delta U = U_k - U$ , то разность между границами зон в  $n$  и  $p$ - областях уменьшается (рис.12.5б).

Отношение населенностей теперь определяется не  $U_k$ , а  $\Delta U$ . Таким образом, в области  $p-n$  –перехода одновременно возникает вырождение по электронам и по дыркам, и при достаточной концентрации неравновесных электронов и дырок возникает состояние инверсии населенностей (рис.12.5б). Неравновесные носители будут рекомбинировать, испуская кванты  $\hbar\omega$ , а через  $p-n$  –переход будет протекать электрический ток  $j$ , который тем больше, чем больше приложенное напряжение. В общем случае этот ток выражается через сумму двух токов  $j = j_n + j_p$  ( $j_n$  – электронный ток,  $j_p$  –дырочный). Минимальная плотность тока  $j$ , при которой индуцированное излучение превышает индуцированное поглощение, называется пороговой плотностью тока  $j_{пор}$ . Если  $j > j_{пор}$ , то  $p-n$  переход является усиливающей средой для света, распространяющегося в плоскости  $p-n$  –перехода.

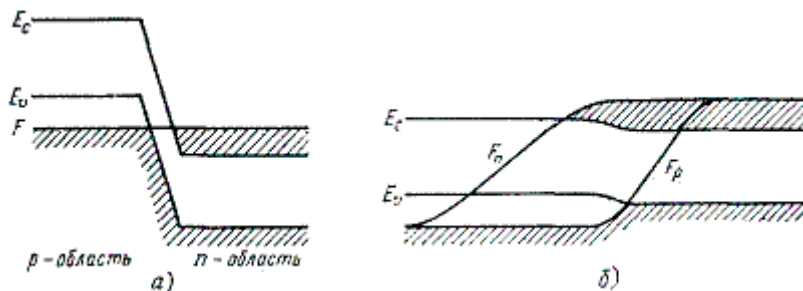


Рис. 12.5. Квазиуровни Ферми в вырожденном  $p-n$ -переходе: а) без инжекции, б) при инжекции. Косая штриховка показывает области, заполненные электронами

Искажение зонной структуры вблизи  $p-n$ -перехода при инжекции носителей в положительном направлении приводит к выполнению условия инверсии (12.13), как это показано схематически на рис. 12.5б. Ширина активной области, в которой  $F_n - F_p > E_c - E_v = E_g$ , может быть значительно больше технологической ширины перехода. В этой области в спектральном интервале созданы условия для усиления на вынужденных рекомбинационных переходах зона — зона.

Лазерный эффект при инверсии методом инжекции носителей в  $p-n$ -переход реализован во многих однодолинных прямозонных полупроводниках. Одним из лучших является лазер на арсениде галлия. Диод представляет собой тонкий срез монокристалла GaAs с поперечными и продольными размерами порядка 0,1— 1 мм. Образец вырезан из сильно легированного материала  $n$ -типа

(донорная примесь Te, Se). После диффузии или имплантации  $p$ -типа (акцепторы Zn, Cd) верхняя часть образца приобретает проводимость  $p$ -типа, и недалеко от поверхности (на расстоянии 10—100 мкм) образуется планарный слой  $p$ — $n$ -перехода толщиной 1—10 мкм. Контактные поверхности  $p$ - и  $n$ -областей покрываются золотом. Кристалл крепится (пайка, клей) к теплопроводящей подложке обычно  $n$ -областью. Наилучшей подложкой является алмаз, часто используется сапфир. На рис. 12.6 показана схема устройства такого лазера. При пропускании большого тока в прямом направлении (с плотностью тока  $1000 \text{ А/см}^2$ ) инверсия возникает в тонком слое  $p$  —  $n$ -перехода. Коэффициент усиления велик. Поэтому боковые грани кристалла ( $n = 3,6$ ;  $R = 30\%$ ) могут играть роль френелевских зеркал, и генерация возникает при малой длине кристалла.

Одним из удачных технологических приемов образования совершенных плоскопараллельных зеркал резонатора полупроводникового лазера является скалывание боковых граней кристалла.

Излучение из лазера выходит через узкие полоски, образованные пересечением активного слоя с частично отражающими гранями кристалла. Характерные угловые размеры диаграммы излучения, соответствующей этому сечению, составляют  $5^\circ \times 50^\circ$ . Более высокая направленность может быть достигнута при помощи внешнего резонатора. Тогда грани полупроводникового диода должны быть либо просветлены, либо ориентированы под брюстеровским углом к оси резонатора. Как правило, техническое осуществление таких конструкций встречает серьезные трудности.

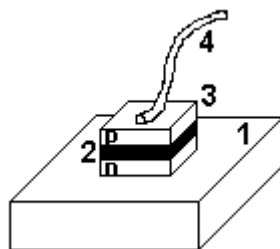


Рис. 12.6. Схема инжекционного лазера: 1 — подложка, 2 — плоскость  $p$ — $n$ -перехода, 3 — кристалл полупроводника, 4 — контактный провод

До сих пор мы рассматривали  $p$  —  $n$ -переходы, образованные путем распределения  $p$ - и  $n$ -примесей в одном и том же монокристалле. В таких переходах, называемых также гомопереходами, по обе стороны границы раздела свойства кристалла одинаковы. Существенное улучшение характеристик полупроводниковых лазеров и прежде всего резкое снижение пороговой плотности тока и связанная с этим возможность работы в непрерывном режиме при комнатных температурах были достигнуты применением анизотропных гетеропереходов, которые образуются наращиванием монокристаллического слоя одного полупроводника на монокристаллическую подложку другого полупроводника (рис. 12.7). Главное различие между гетеропереходом и гомопереходом заключается в том, что запрещенные зоны изготавливаются из двух различных материалов. Поэтому несоответствие коэффициентов кристаллических решеток этих материалов может привести к возникновению большого количества неизлучающих рекомбинационных центров на переходе, а это уменьшает вероятность излучения света.

Примерами являются пары GaAs — AlGaAs, CdTe— CdSe и т. д. Вводя должным образом акцепторные и донорные примеси, можно создать полупроводниковые диоды с гетеропереходом. Схема устройства диода полупроводникового лазера с двойным гетеропереходом приведена на рис. 12.7.

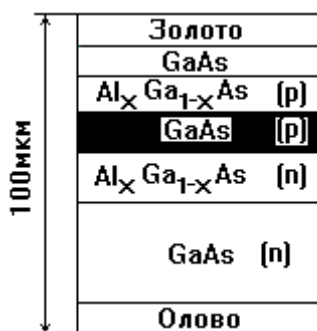


Рис. 12.7. Схема диода полупроводникового лазера с двойным гетеропереходом (ДГ-лазер). Активная область заштрихована

В гетероструктуре  $p$ — $n$ -переход носит наиболее резкий характер, отсутствует диффузионное растекание инжектируемых носителей, концентрирующихся при инжекции в хорошо определенной узкой области перехода. Поэтому уменьшается величина пороговой плотности тока инжекции. Кроме того, показатель преломления трехкомпонентного полупроводника AlGaAs существенно меньше, чем бинарного полупроводника GaAs. Следовательно, возникает эффект оптического волновода, излучение концентрируется в активной области и не проникает в поглощающую область, не подверженную накачке. Все это в целом, а также монтаж диода на подложке высокой теплопроводности, позволило снизить пороговую плотность тока при комнатной температуре более чем в 100 раз, доведя ее до нескольких сотен ампер на квадратный сантиметр. В результате получены устойчивые режимы генерации полупроводниковых лазеров непрерывного действия с выходной мощностью до 100 мВт при комнатной температуре, что резко увеличило практическую значимость этих приборов. Основной областью их применения является лазерная связь, прежде всего по волоконно-оптическим линиям.

Гетероструктуры были предложены для создания полупроводниковых инжекционных лазеров Ж. И. Алферовым в 1963 г. Это предложение было реализовано под его руководством в 1968 г.

Диапазон длин, волн излучения полупроводниковых лазеров в соответствии с (12.12) определяется шириной запрещенной зоны. Арсенид галлия ( $E_g = 1,5$  эВ) излучает в области 0,84 мкм. Полупроводниковые растворы переменного состава позволяют перекрывать широкий спектральный диапазон. Так, система AlGaAs дает излучение в зависимости от состава в области 0,63—0,90 мкм, система AlGaSb—в области 1,20—1,80 мкм, система GaInAs — в области 0,9—3,4 мкм и т. д. В некоторых твердых растворах, например в  $Pb_x Sn_{1-x} Se$ ,  $Cd_x Hg_{1-x} Te$ , при изменении процентного содержания компонент ширина запрещенной зоны проходит через нуль. Это позволяет создавать длинноволновые лазеры с длиной волны до 30—40 мкм. В последнем случае для предотвращения термического заселения зоны проводимости необходимо глубокое охлаждение накачиваемого кристалла [4,26,30,31].



**Длина волны излучения диодных полупроводниковых лазеров может меняться, в широких пределах.** В отличие от лазеров других типов, она определяется переходами не между дискретными уровнями энергии атомов или молекул, а между зонами разрешенных состояний в полупроводнике и зависит от многих факторов, влияющих на его зонную структуру, таких как давление, температура, магнитное поле. Выделение одной моды излучения при этом облегчено большим спектральным расстоянием между модами, обусловленным малыми размерами активной среды лазера. Как правило, расстояние между двумя соседними модами составляет  $0,5\text{—}3\text{ см}^{-1}$ . Спектральная ширина одной моды несколько превышает  $10^{-4}\text{—}10^{-3}\text{ см}^{-1}$ , что обусловлено нестабильностью температуры и тока инжекции.

Наибольшее распространение получила перестройка лазеров путем изменения давления и температуры.

Для лазеров на основе тройных соединений свинца, таких как  $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ ,  $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ , всестороннее сжатие, осуществляемое обычно гидростатически, изменяет параметры кристаллической решетки и тем самым уменьшает ширину запрещенной зоны. В результате этого длина волны генерируемого излучения увеличивается. Изменение это значительно. В доступной области давлений 0 - 15 кбар возможна перестройка на  $1000\text{см}^{-1}$ . В качестве примера укажем диодный лазер на  $\text{PbSe}$ , перестраиваемый давлением в 14 кбар с 8,5 мкм на 22 мкм. Преимущества перестройки лазеров, давлением заключены в отсутствии влияния давления на пороговый ток и условия самовозбуждения лазера. Недостатки сводятся к заметной технической сложности устройства в целом и к существенной инерционности перестройки.

Другим способом перестройки является понижение температуры лазера от максимально возможной до гелиевой. Для этого широко распространенного метода характерна величина перестройки в  $100\text{—}200\text{ см}^{-1}$ . Недостатком метода является существенная нелинейность зависимости длины волны излучения от температуры. Кроме того, изменение температуры, как это следует из предыдущего рассмотрения, сильно меняет порог самовозбуждения лазера. Технически, при наличии автономных криогенных устройств замкнутого цикла, температурный способ перестройки длины волны излучения диодных лазеров достаточно удобен. Разумеется, возможно сочетание температурной перестройки и перестройки давлением одного и того же лазера, что существенно расширяет рабочий диапазон.

Тонкая и относительно безынерционная перестройка осуществляется во время импульса тока инжекции через  $p-n$ -переход за счет его разогрева джоулевым теплом. Токовая перестройка носит квазинепрерывный характер, так как смещение полосы усиления сопровождается сдвигом частот спектральных мод лазера вследствие температурного изменения показателя преломления активной среды. Эти процессы идут с разными скоростями, вследствие чего диапазон перестройки не является непрерывным, а состоит из отдельных отрезков кривых плавной перестройки мод лазера, вызванной изменением показателя преломления. Область плавной перестройки током составляет обычно  $0,5\text{—}3\text{ см}^{-1}$ .

Удобным методом сканирования частоты лазера является добавление к постоянному току инжекции, меньшему порогового, переменной компоненты заданной формы и частоты. Максимальная скорость токовой перестройки определяется тепловой инерцией лазера и может достигать  $10^4\text{ см}^{-1}\cdot\text{с}^{-1}$ . Подстройка лазера

на некоторую заданную длину волны излучения и стационарное удержание этой длины волны требует высокой стабильности температуры лазера и тока накачки.

Основной областью применения перестраиваемых диодных лазеров является спектроскопия высокого разрешения и обнаружение различного рода примесей в газах и жидкостях.

**Пороговый ток.** Оценим величину порогового тока накачки при температурах, близких к абсолютному нулю. Для этого воспользуемся выражением для пороговой разности населенностей [4].

$$(N_2 - N_1)_{\text{пор}} = [\hbar Q B_{12} g(\omega)]^{-1}. \quad (12.18)$$

Будем считать, что в  $p$ - $n$  – переходе при  $T \rightarrow 0^\circ \text{K}$  населенность уровней, прилегающих к вершине валентной зоны, стремиться к нулю. ( $N_1 \rightarrow 0$ ), а нижние уровни дна зоны проводимости населены за счет инжекции ( $N_2 \neq 0$ ). В этом случае

$$(N_2 - N_1)_{\text{пор}} = (\Delta n_p)_{\text{пор}}$$

Для того, чтобы найти связь между пороговым током и изменением концентрации электронов воспользуемся уравнением непрерывности. Будем считать, что  $j_n \gg j_p$ , и все процессы протекают за счет инжекции электронов в дырочную область.

Тогда

$$\frac{\partial n_p}{\partial t} = \frac{1}{e} \text{div} \overline{j_n} - \frac{n_p - n_{p0}}{\tau} - g, \quad (12.19)$$

где  $g$  – число электронно-дырочных пар, рекомбинирующих с испусканием вынужденного излучения в единице объема за единицу времени,  $\tau$  – время жизни электронов в зоне проводимости.

В стационарном режиме уравнение (12.19) приобретает вид

$$\text{div} \overline{j_n} = \frac{e}{\tau} \Delta n_p - \frac{e}{g} \quad (12.20)$$

и выражает условие сохранения числа электронов: поток электронов, втекающих в элементарный объем за счет инжекции, равен количеству электронов, теряющихся за счет тепловой и вынужденной рекомбинации.

На пороге генерации индуцированным излучением можно пренебречь ( $g \rightarrow 0$ ), тогда в одномерном случае

$$\left(\frac{\partial j_n}{\partial z}\right)_{\text{пор}} = \frac{e}{\tau} (\Delta n_p)_{\text{пор}}.$$

Если вдоль перехода толщиной  $d$  изменение плотности тока постоянно, то

$$\frac{\partial j_n}{\partial z} = \frac{j_n}{d} \text{ и}$$

$$(j_n) = j_{\text{пор}} = \frac{ed}{\tau} (\Delta n_p)_{\text{пор}} \quad (12.21)$$

Подставляя в (12.21) (12.18) и выражение для

$$B_{12} = \frac{(c/n)^3 \pi}{2\omega^2} \cdot \frac{1}{\hbar \omega \tau},$$

получим

$$j_{\text{пор}} = \frac{2ed\hbar\omega^2}{\hbar Q g(\omega)(c/n)^3 \pi}. \quad (12.22)$$

Подставляя в (12.22) выражения для (8.8) и (5.23), окончательно, для плотности порогового тока при  $T \rightarrow 0^0 \text{K}$ , можно записать

$$j_{\text{пор}} = \frac{8\pi^2 e d (1-r) n^3}{\lambda^2 L} \quad (12.23)$$

Здесь  $n$  – показатель преломления активного вещества,  $L$  – длина резонатора,  $r$  – коэффициент отражения зеркал резонатора. Подставляя в (12.23)  $e=4,8 \cdot 10^{-10}$  ед. СГСЕ=1,44 СИ,  $d=10$  мкм,  $r=0.3$ ,  $\tau = 10^{-10}$  с,  $\lambda = 0.84$  мкм,  $L = 100$  мкм, получим  $j_{\text{пор}} = 10^3$  А/см<sup>2</sup>.

#### 12.1.4. Оптические и электрические свойства полупроводниковых лазеров

**Моды излучения.** Модовая структура полупроводникового лазера определяется в первую очередь геометрией резонатора. Резонатор Фабри-Перо образован двумя гранями полупроводникового кристалла, параллельными плоскости  $zOy$ , причем толщина кристалла  $L$  всегда мала. Очевидно, что в объемном резонаторе существуют моды ТЕ и ТМ. Анализируя трехмерный случай, можно выделить следующие типы мод (рис. 12.8) [33].

а) Продольные моды (распространяющиеся вдоль оси  $Oz$ ). Это моды резонатора Фабри-Перо. Расстояние между такими модами составляет несколько ангстрем.

б) Боковые моды (распространяющиеся вдоль оси  $Oy$ ). Они зависят от способа получения поверхностей кристалла, от ширины резонатора, ширины полоски. Обычно наблюдается несколько мод низкого порядка, отстоящих друг от друга на несколько десятков ангстрем.

в) Поперечные моды (вдоль оси  $Ox$ ). Эти моды перпендикулярны плоскости  $p$ - $n$ -перехода и обусловлены разностью показателей преломления. Очевидно, что обычно стремятся к тому, чтобы возбуждалась только основная мода. При этом толщина активного слоя в лазерах должна быть менее одного микрометра. В лазерах с шириной полоски (толщиной активного слоя) более 20–30 мкм обычно наблюдается резкий переход через порог, что ведет к возбуждению мод высокого порядка.

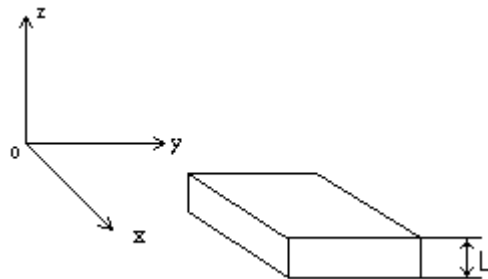


Рис. 12.8. К определению типов мод в резонаторе полупроводникового лазера

**Пороговый ток [22].** Пороговый ток сильно зависит от температуры, и эта зависимость имеет вид

$$I_s = I_0 \exp \frac{T}{T_0}$$

где  $T_0 = 120^0 K$ .

В полупроводниковых лазерах с простой гетероструктурой пороговый ток уменьшается с увеличением расстояния между р-п-переходом и гетероструктурой. В лазерах с двойной гетероструктурой он пропорционален объему резонатора. Температурные зависимости порогового тока в инжекционных лазерах разного типа с арсенидом галлия в качестве рабочего вещества представлены на рис. 12.9 [28]

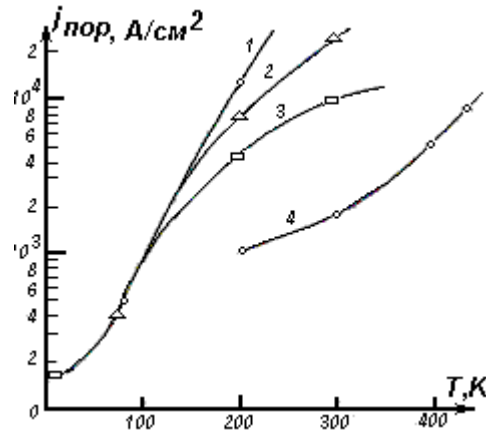


Рис. 12.9. График температурной зависимости пороговой плотности тока в инжекционных лазерах различных типов на основе арсенида галлия. 1- диффузионные диоды, 2-эпитаксильные сильно легированные диоды, 3- гетероструктуры, 4-двусторонние гетероструктуры

**Модуляция излучения.** Самый простой метод модуляции излучения полупроводниковых лазеров состоит в модуляции тока инжекции  $j$ . Если инжекционный ток превышает пороговый ток, то наблюдается некоторая задержка генерации излучения, описываемая приближенным выражением

$$t_e = t_0 \ln \left( \frac{j}{j - j_{пор}} \right).$$

Поскольку электрические и оптические явления, наблюдающиеся в лазере, взаимосвязаны, при высоких частотах модуляции наблюдается резонанс между инжектируемыми носителями заряда и электрическим полем. Резонансная частота  $\nu_r$  тем выше, чем больше отношение  $j/j_{пор}$  (рис.12.10). Она дается формулой

$$\nu_r = \frac{m}{2\pi} \sqrt{\frac{1}{\tau_n \tau_p} \left( \frac{j}{j_{пор}} - 1 \right)},$$

где  $\tau_n$  и  $\tau_p$  - времена жизни электронов и дырок по отношению к спонтанной рекомбинации, а  $m$  - логарифм коэффициента усиления по току (типичные значения  $m$  лежат между 1 и 3). Рассматриваемые явления подобны тем, которые наблюдаются в газовых лазерах. На рис. 12.10 представлены типичные кривые зависимости мощности излучения от частоты модуляции тока накачки.

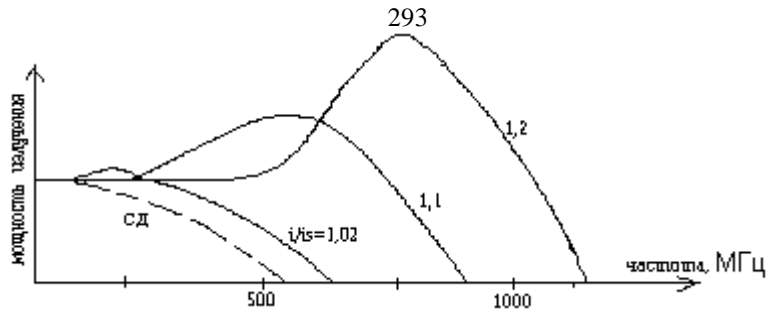


Рис.12.10. Зависимость излучаемой мощности от частоты модуляции при различных значениях отношения тока модуляции к пороговому току. Ниже порога показана аналогичная зависимость для светодиода (штриховая)

**Основные характеристики полупроводниковых лазеров.** К числу основных характеристик полупроводниковых лазеров, определяющих возможность их использования в системах связи и передачи информации, относятся: мощность излучения и ее зависимость от тока модуляции, диаграмма направленности излучения, спектр излучения, срок службы

**1. Выходная мощность, эффективность и к.п.д. лазера [22].** Если ток накачки превышает пороговое значение, то избыток мощности накачки над пороговым значением может быть полностью преобразован в мощность когерентного излучения (даже если внутренний квантовый выход излучательной рекомбинации заметно меньше единицы), поскольку эффективное излучательное время жизни избыточных носителей тока убывает с ростом интенсивности электромагнитного поля.

Во внешнюю среду выходит некоторая часть  $f$  (функция выхода) когерентного излучения, генерируемого в активной среде, которая зависит от характеристик резонатора и внутренних потерь  $\alpha$ . Для продольных типов колебаний в плоском резонаторе при незначительном превышении порога генерации в стационарном режиме величина  $f$  может быть представлена в виде [22]

$$f(\alpha, L, R) = \left( 1 + \alpha L / \ln \frac{1}{\sqrt{R_1 R_2}} \right)^{-1},$$

причем суммируется излучение, испускаемое в обе стороны; если же одно зеркало глухое, то

$$f(\alpha, L, R) = \left( 1 + 2\alpha L / \ln \frac{1}{R} \right)^{-1}.$$

Выходная мощность инжекционного лазера

$$P = (\hbar \omega / e) \eta_p f(\alpha, L, R) (j - j_{пор})$$

где  $j$  — ток накачки;  $j_{пор}$  — пороговый ток;  $\eta_p$  — эффективность накачки, т. е. отношение числа электронно-дырочных пар, образующихся в активной области лазера, к числу электронов, расходуемых на создание электронно-дырочных пар за то же время. Величина  $\eta_p$  может отличаться от единицы вследствие утечек мощности накачки и неполного совпадения рабочего объема активной области лазера с объемом, в котором действует накачка. Для лазера с электронным возбуждением

$$P = \frac{\hbar \omega}{e} \frac{\varpi(1-r)}{E_I} \eta_p f(\alpha, L, R) (j - j_{пор})$$

где  $\hbar\omega$  — энергия быстрых электронов;  $r$  — коэффициент их отражения от поверхности;  $E_1$  — средняя энергия, расходуемая на образование одной электронно-дырочной пары. Согласно [22], по эмпирическим данным эта энергия составляет

$$E_1 \approx 2.8E_g + \delta E$$

где  $\delta E = 0.5 - 1.0$  эВ. С теоретической точки зрения  $E_1$  определяется на основе законов сохранения энергии и импульса при ионизационных потерях энергии быстрыми электронами. Предельная эффективность лазера с электронным возбуждением ограничивается отношением  $\hbar\omega/E_1$  и составляет около 0,3 (поскольку  $\hbar\omega \approx E_g$ ).

Для лазера с оптическим возбуждением фотонами с энергией  $\hbar\omega_p$

$$P = \frac{\hbar\omega}{\hbar\omega_p} (1 - R_p) \eta_p f(\alpha, L, R) (J - J_{\text{пор}}),$$

где  $J$  и  $J_{\text{пор}}$  — поток фотонов накачки и его пороговое значение соответственно;  $R_p$  — коэффициент отражения фотонов накачки. Отношение  $\hbar\omega/\hbar\omega_p$  называют энергетическим выходом. Коэффициент полезного действия определяется отношением мощности  $P$  к полной затрачиваемой мощности, включая мощность, теряемую в схемах формирования потока накачки (питание электронной пушки и фокусирующей системы, питание лазера или лампы накачки и т. д.), и поэтому зависит от конкретных устройств. В инжекционном лазере следует учесть джоулевы потери в объеме полупроводника и в контактах, определяемые током накачки и остаточным сопротивлением  $R_s$ . Тогда к. п. д.

$$\eta = P / (IU + I^2 R_s)$$

где  $U$  — напряжение, приложенное к  $p$ - $n$ -переходу и близкое по величине к  $\hbar\omega/e$ .

Дифференциальной эффективностью инжекционного лазера называют величину

$$\eta_{\text{диф}} = dP / d(IU) = (\hbar\omega / eU) \eta_p f(\alpha, L, R) \quad (12.24)$$

максимальные значения которой достигают 0,7 — 0,8. Согласно формуле (12.24) дифференциальная эффективность, описывающая наклон выходной характеристики [например, ватт-амперной  $P(j)$  рис.12.11], представляется постоянной величиной. На практике встречаются некоторые отклонения от идеализированной модели.

Уменьшение  $\eta_{\text{диф}}$  и даже падение выходной мощности при большой накачке обычно обусловлено тепловой перегрузкой излучающего элемента. Тепловое сопротивление и предельная рассеиваемая мощность диода сильно зависят от размеров диода и рабочей температуры. Так, в инжекционных лазерах совершенной конструкции при гелиевых температурах удается отвести 30 — 40 Вт тепла, прежде чем выходная мощность излучения перестанет расти; при охлаждении жидким азотом эта величина составляет не менее 10 Вт; при комнатной температуре — около 1 Вт.

«Смягченная» вблизи порога ватт-амперная характеристика нередко наблюдается вследствие изменения объема возбужденной области лазера

(т. е. за счет роста величины  $\eta_p$ ). На сверхлинейном участке характеристики наблюдается развитие картины ближней зоны от одного яркого пятна до образования яркой полосы (в достаточно однородных лазерах).

«Жесткая» характеристика возникает в тех случаях, когда имеется неоднородность в распределении накачки вдоль оси резонатора. Резкий рост выходной мощности на начальном участке возникает вследствие оптического режима (переходные процессы и автомодуляция). Обычно уже при небольшом превышении порога генерации возникают более или менее регулярные пульсации излучения с характерной частотой повторения  $10^8$  —  $10^9$  Гц. Нередко наблюдаются процессы типа перезарядки ловушек, приводящие к задержкам генерации по отношению к фронту импульса накачки на 10 — 30 нс.

Расчеты на основе скоростных уравнений показывают, что установление режима генерации носит периодический характер [22], причем период пульсаций  $\theta$  интенсивности составляет (при малой глубине)

$$\theta^2 \approx \frac{4\pi^2 \tau_e \tau_p}{(I/I_t) - 1}$$

где  $\tau_e$  — время жизни электронов;  $\tau_p$  — время жизни фотонов. Если длительность импульсов сравнима или меньше  $\theta$ , то выходная характеристика существенно отличается от стационарной.

В непрерывном режиме ограничения мощности обуславливаются перегревом излучателя, в импульсном — процессами саморазрушения кристалла (главным образом торцов диода). Критические значения плотности оптического потока зависят от условий и режима работы. Для лазеров на основе сильно легированных  $p-n$  - переходов в GaAs при  $77^0$  К и длительности импульсов накачки 0,5 мкс критическая мощность, на 1 мм ширины диода, составляет около 50 Вт.

По экспериментально измеренной зависимости мощности излучения полупроводникового лазера от тока накачки можно определить пороговый ток и энергоотдачу. На кривых такой зависимости (рис.12.11) обнаруживаются нелинейности (перегибы и изломы), выраженные в той или иной степени. Кроме того, мощность, излучаемая лицевой гранью ограничивается величиной, выше которой происходит разрушение граней кристалла, что делает несовместимыми большую мощность излучения и непрерывный режим работы.

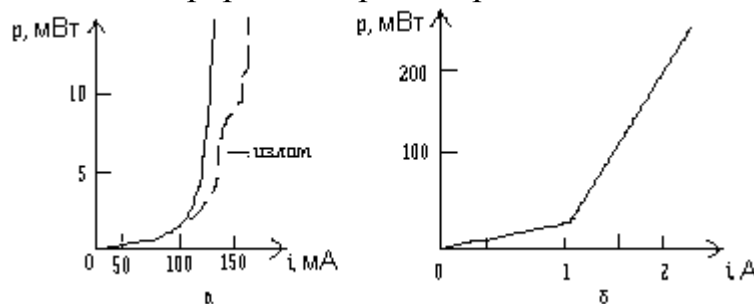


Рис. 12.11. Примерная зависимость излучаемой мощности от тока накачки  $P(i)$  для полупроводникового лазера:

*a* - в непрерывном режиме; *б*- в импульсном режиме с коэффициентом заполнения менее 10%. К.п.д лазера измеряется крутизной  $P/i$

## 2. Диаграмма направленности излучения

Направленность излучения полупроводниковых лазеров значительно хуже, чем газовых лазеров (рис.12.12), по следующим причинам:

а) из-за большого показателя преломления полупроводникового кристалла оказываются большими углы выхода излучения из кристалла;

б) малая толщина активной зоны лазера оказывает сильное влияние на дифракционные эффекты.

Диаграмма направленности излучения полупроводникового лазера несимметрична: ее ширина менее  $20^\circ$  в плоскости, параллельной переходу, и более  $40^\circ$  в перпендикулярной плоскости. В случае, когда существует несколько поперечных мод, наблюдается несколько боковых лепестков диаграммы направленности излучения в дальней зоне.

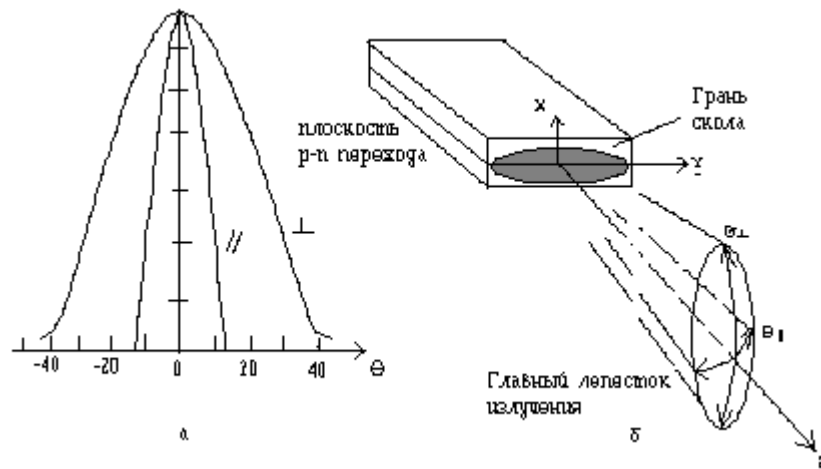


Рис. 12.12. Диаграмма направленности излучения полупроводникового лазера.

*a* - диаграмма направленности; приведена только активная зона, ширина диаграммы направленности измеряется на уровне половинной мощности; *б*- зависимость излучаемой мощности от угла ( во взаимно ортогональных направлениях  $Ox$  и  $Oy$ )

**3.Спектр излучения.** В спектральной области, где можно определить коэффициент усиления резонатора, существуют только такие продольные моды, для которых длина пробега света в резонаторе туда и обратно равна целому числу длин волн. Если  $p$ -порядок моды, а  $\lambda$ - длина волны излучения, то для этой моды имеем

$$p\lambda = 2n(\lambda)L,$$

а для следующей соседней моды порядка  $p+1$  с длиной волны  $\lambda^1$  аналогично можно написать

$$(p+1)\lambda^1 = 2n(\lambda^1)L.$$

Отсюда получаем следующую формулу для интервала между соседними модами:

$$\Delta\lambda = \lambda^2 / 2nL(1 - \lambda dn / nd\lambda)$$

Таким образом, кажущийся показатель преломления резонатора есть групповой показатель преломления.

Простой числовой пример ( $\lambda = 0,9$  мкм,  $n = 3,6$  и  $L = 300$  мкм) показывает, что интервал между модами равен  $\Delta\lambda = 0,4$  нм. Этот элементарный расчет объясняет и регулярную линейчатую структуру спектра излучения полупроводникового лазера (рис. 12.13) [33].



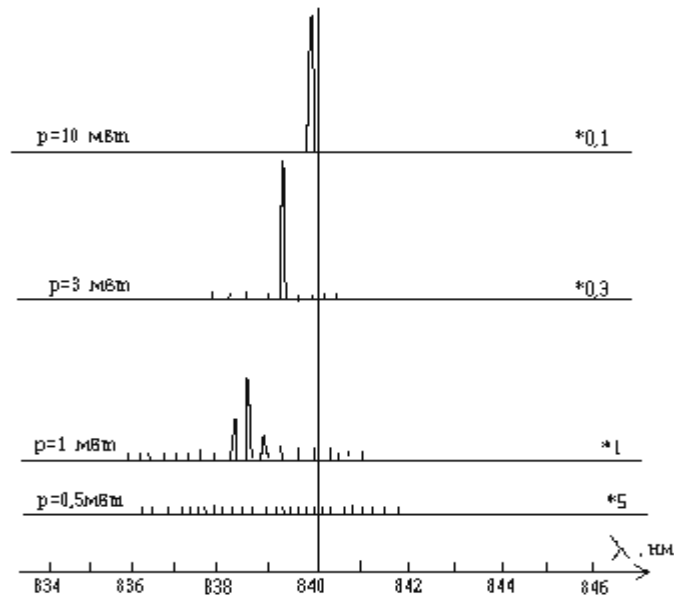


Рис. 12.13. Зависимость спектра излучения полупроводникового лазера от излучаемой оптической мощности. Справа на горизонтальных осях указан масштабный множитель для вертикальной шкалы

## 12.2. Светоизлучающие диоды

Как выше было отмечено, наряду с лазерным диодом, широко используются светоизлучающие диоды (СИД). Особенностью работы СИД является использование для генерации спонтанного излучения [16,27,29,34].

Структура простейшего светоизлучающего диода представлена на рис. 12.14:

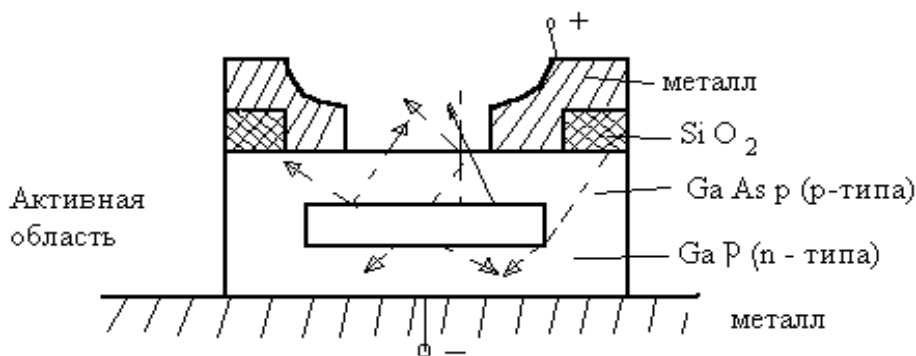


Рис. 12.14. Структура светоизлучающего диода

СИД строятся на основе полупроводниковых соединений типа  $A^{III} B^V$  (GaAs, GaP, AlAs, InAs, GaSi). Диоды на соединениях  $In_x Ga_{1-x} As_y P_{1-y}$ , в зависимости от содержания  $x$  и  $y$  имеют рабочую длину волны в диапазоне от 1 мкм до 1.6 мкм (изменяется ширина запрещенной зоны, меняется и длина волны). Энергия выделяющихся фотонов определяется выражением (12.13).

Подставляя в (12.13) значение  $hc$ , а также зная, что  $1\text{эВ} = 1.6 \cdot 10^{-19}$  Дж, можно определить  $\Delta E_g$  в эВ, необходимую для получения излучения с  $\lambda$  (в мкм). Из вы-

ражения  $\lambda = 1.24 / (E_g - \Delta E)$ , мкм, следует, что для получения видимого излучения с  $\lambda = 0.38—0.78$  мкм, полупроводник должен иметь  $\Delta E_g > 1.7$  эВ.

Внесение в полупроводник различных примесей позволяет получить свечение различного цвета. Цвет свечения зависит от соотношения токов через переходы. Для арсенид - галлиевого СИД длина генерируемого света составляет около 900 нм. При добавке к арсениду галлия примеси Al длина волны генерируемого света уменьшается и может составлять, например 780 нм.

Для получения СИД с ещё более короткой длиной волны, лежащей в видимой области спектра, необходимо применять в них арсенид-фосфид галлия или фосфид галлия. Наоборот, СИД, излучающие свет с  $\lambda = 1.3$  мкм, изготавливают из полупроводниковых материалов, содержащих арсенид-фосфид галлия и индия.

В отличие от лазерного диода (ЛД)-в СИД отсутствует положительная обратная связь (зеркала), кроме того длина усиливающей области обычно больше чем в ЛД и достигает 1.5 мм.

Полупроводниковые светоизлучающие диоды в силу физики процессов, происходящих в них, существенно уступают по параметрам излучения полупроводниковым лазерам. Спонтанное время жизни значительно больше, чем индуцированное, поэтому СИД является менее быстродействующим прибором. Однако относительная простота технологии изготовления, низкая стоимость и более высокий срок службы вследствие менее напряженного режима работы делают целесообразным применение СИД в системах, где широкая линия излучения не является существенным недостатком: в локальных сетях, объектовых волоконно-оптических системах передач (ВОСП).

### 12.2.1. Устройство СИД на гомо- и гетеропереходах

В светодиодах, как и в лазерных диодах, используются как гомопереходы так и гетеропереходы. Полупроводниковые структуры на основе гетеропереходов целесообразно использовать для конструирования эффективных светодиодов с высокой излучательной способностью. В таких структурах можно удачно сочетать соответствующее легирование области рекомбинации с малым коэффициентом поглощения на длине волны излучения для окружающих областей, выполняемых для этой цели из материала с большей шириной запрещенной зоны. Локализация носителей тока в светодиоде приводит к увеличению внешней эффективности эмиссии света, а использование волноводной геометрии существенно улучшает спектральные характеристики излучателя.

Время жизни неосновных носителей в области рекомбинации зависит от концентрации в ней носителей. Если плотность инжектированных носителей мала по сравнению с равновесной, то время жизни для излучательной рекомбинации ( $\tau_i$ ) оказывается обратно пропорциональным концентрации дырок или электронов соответственно для материала *p*- или *n*-типа. Эффективность излучения зависит от соотношения времен жизни для излучательной и безызлучательной ( $\tau_{би}$ ) рекомбинации:

$$\eta_i = \frac{1}{1 + (\tau_i / \tau_{би})}.$$

Чтобы при той же эффективности излучения  $\eta_i$  увеличить быстродействие диода, необходимо уменьшить  $\tau_n$ , за счет дополнительного легирования, не допуская при этом чрезмерного снижения  $\tau_{\text{би}}$ . Следует учитывать, что для  $GaAs$  и сходных с ним соединений центры безызлучательной рекомбинации образуются при уровнях легирования, близких к пределу растворимости этой примеси при данных условиях выращивания.

На практике наиболее подходящей акцепторной примесью для соединения  $GaAs$  —  $AlGaAs$ , выращиваемого методом жидкостной эпитаксии, является германий. Уровень легирования германием может быть поднят вплоть до  $10^{19} \text{ см}^{-3}$  без заметного снижения квантовой эффективности. При этом время жизни неосновных носителей снижается по крайней мере до 5 нс. Диоды с двумя гетеропереходами при высоких уровнях легирования германием имеют времена спада и подъема свечения менее 5 нс.

Следствием столь малых значений времени затухания излучения является способность таких диодов работать при очень высоких частотах модуляции. Для диода с двумя гетеропереходами предельная частота модуляции на уровне половины выходной мощности равна 180 МГц.

Основным источником шума в структурах с гетеропереходами обычно являются плохие омические контакты. Для приборов с хорошими контактами уровень шумов собственно диодов значительно ниже уровня дробовых шумов любых реальных систем.

Важным параметром светодиода с точки зрения его согласования со световодом является яркость излучения. Не менее существенна и конфигурация диода, оптимальные размеры которой зависят от размеров волокна и его числовой апертуры, а также от того, используется ли одиночный световод или пучок волокон. Основная задача применения светодиодов в волоконнооптических системах сводится к получению необходимого уровня мощности, вводимой в световод или пучок световодов.

С точки зрения эффективности согласования с пучком световодов представляется полезной параболическая конфигурация (рис. 12.15). Внешний диаметр параболы выбирается в соответствии с диаметром пучка. С помощью такого устройства при той же плотности тока удастся получать самые высокие выходные мощности, собирая свет не только с поверхности диода, но также и со всех четырех его торцов.

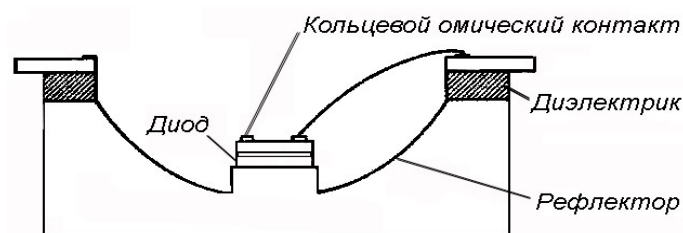


Рис. 12.15. Параболическая структура со светодиодом, позволяющая использовать как поверхностную, так и торцевую эмиссию света

Использование эмиссии света из торца привлекательно, как своей простотой, так и тем, что такое излучение имеет определенную направленность. Свет можно сфокусировать, подать на торец световода с помощью линз или просто приставить диод к световоду. В режиме некогерентного излучения можно получить световую мощность порядка десятка микроватт, тогда как выше порога, где эффективность диода и направленность его излучения значительно выше, вполне достижим уровень мощности порядка милливатта.

### 12.2.2. Типы конструкций СИД и их особенности

Можно выделить два типа конструкций СИД: поверхностные — с выводом излучения в направлении, перпендикулярном плоскости  $p$ - $n$ -перехода, и торцевые СИД, излучающие в направлении, параллельном этой плоскости.

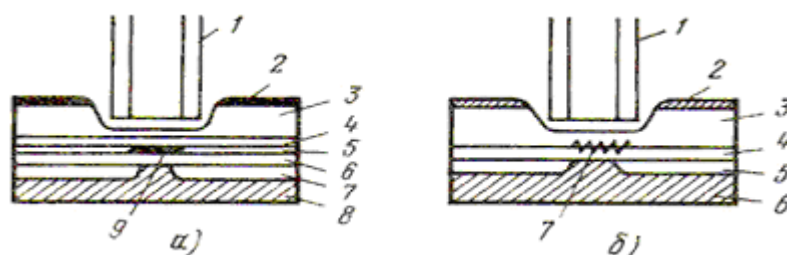


Рис. 12.16. Поверхностные СИД:

*a* — на двойной гетероструктуре (ДГС): 1 — волоконный световод; 2 — контакты; 3 — подложка  $n$ -JnP; 4 — широкозонный слой  $n$ -JnP; 5 — узкозонный слой  $p$ -GaAs AsP; 6 — широкозонный слой  $p$ -JnP; 7 — диэлектрик ( $\text{SiO}_2$ ); 8 — контакт и радиатор; 9 — активная область; *б* — на гомопереходе: 1 — волокно; 2 — контакт; 3 —  $n$ -GaAs; 4 —  $p$ -GaAs; 5 —  $\text{SiO}_2$ ; 6 — контакт и радиатор; 7 — активная область

Обычно в поверхностных СИД спонтанное излучение в активном слое происходит равновероятно во всех направлениях. Преобразование энергии накачки в свет происходит с высоким внутренним квантовым выходом, близким к 100 %. Однако полное внутреннее отражение на границе раздела «полупроводник—воздух» с малым критическим углом  $\theta_c = \arcsin(1/n) \approx 15^\circ$  ( $n$  — показатель преломления полупроводника;  $n=3,6$  для GaAs;  $n=3,5$  для JnP) приводит к тому, что большая часть света остается в кристалле и внешняя квантовая эффективность составляет, в лучшем случае, несколько процентов. Излучение с поверхности происходит с диаграммой направленности ламбертовского типа, что усугубляет проблему ввода излучения поверхностного СИД в волоконный световод.

Наиболее рациональной является конструкция поверхностного СИД (рис.12.16), в которой волокно непосредственно стыкуется с излучающей площадкой, диаметр ее близок к диаметру сердцевины волокна. Эпоксидная смола с  $n=1,5$  играет роль иммерсии. Необходимо помнить, что ни одна оптическая система не может повысить эффективность ввода излучения поверхностного СИД в волоконный световод, если его излучающая площадка больше сердцевины световода. Уменьшение излучающей площади, с сохранением мощности излучения, приводит к возрастанию плотности тока накачки, что сокращает срок службы

прибора. Вместе с тем известны конструкции поверхностных СИД с малой площадкой и согласующими линзовыми системами.

Поверхностный СИД изготавливается как на основе ДГС (рис. 12.16, а), так и на основе структур с диффузионными  $p-n$  - гомопереходами (рис. 12.16, б). В первом случае свет выводится через полупроводник с широкой запрещенной зоной, что снижает внутренние потери на поглощение. На гомопереходах СИД имеют низкую стоимость и высокую надежность вследствие простоты технологии.

Спектр излучения поверхностных СИД соответствует спектру спонтанного излучения, ширина его по уровню половинной мощности равна 30-50 нм и не зависит от тока накачки. Центральная длина волны излучения  $\lambda_0$ , как и в лазерах, определяется шириной запрещенной зоны, т. е. материалом. Ватт-амперная характеристика поверхностных СИД линейна (рис. 12.17) вплоть до значений тока накачки  $I=I_{max}$  при которых нагрев активного  $p-n$ -перехода снижает выходную мощность. Мощность, излучаемая СИД при  $I=I_{max}$ , составляет величину 1-10 мВт, мощность, вводимая при этом в волоконный многомодовый световод,—50-500 мкВт. Высокая линейность ватт-амперной характеристики делает поверхностный СИД одним из основных источников для аналоговых систем передачи.

Ширина полосы частот модуляции поверхностных СИД определяется в общем случае временем жизни носителей  $\tau$ . Переменная составляющая мощности излучения  $P_\Omega$  и мощность излучения в отсутствие модуляции  $P_0$  связаны соотношением

$$P_\Omega / P_0 = 1 / [1 + (\Omega \tau)^2]^{1/2}.$$

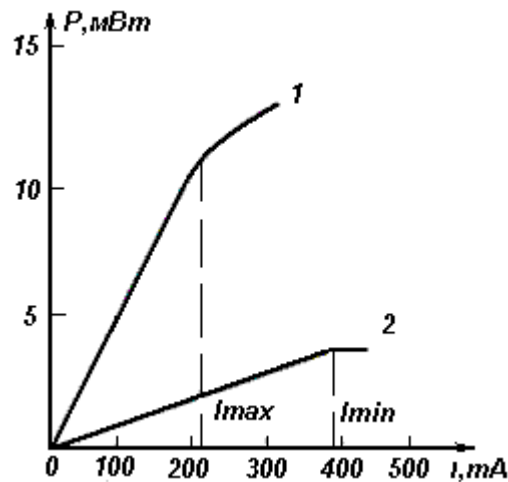


Рис. 12.17. Ватт-амперные характеристики поверхностных СИД на двойных гетероструктурах. 1 — GaAlAs; 2 — InGaAs/InP

Это означает, что верхняя граничная частота модуляции по уровню  $0,707F_g=1/(2\pi\tau)$ . Величина  $\tau$  зависит от тока инжекции, концентрации электронов и дырок, толщины перехода и скорости рекомбинации. Для выпускаемых промышленностью поверхностных СИД типовое значение  $F_g$  не превышает 100 МГц.

В торцевых светоизлучающих диодах, генерируемое излучение выходит из прибора наружу параллельно поверхности  $p-n$  перехода, то есть через боковую грань прибора. Это позволяет уменьшить размеры сечения пучка света, эмити-

руемого СИД, и повысить эффективность ввода света от диода в волоконный световод, имеющий диаметр сечения сердечника менее 200 мкм. Диаметр излучающей площадки для СИД, с выходом излучения через боковую грань, достигает 50-60 мкм.

Прогресс в развитии торцевых СИД связан с появлением в последние годы конструкций, в которых осуществляется усиление спонтанного излучения без обратной связи — *суперлюминесцентных СИД*. Суперлюминесцентные СИД по параметрам излучения занимают промежуточное положение между лазерами и поверхностными СИД со спонтанным излучением. Конструкция люминесцентных СИД представляет собой двойную гетероструктуру с полосковым контактом, который с одной только стороны доходит до торца кристалла (рис.12.14). Таким образом, основное отличие от лазера с полосковым контактом состоит в отсутствии резонатора Фабри — Перо, т. е. в отсутствии положительной обратной связи. Спонтанное излучение происходит равновероятно во все стороны, однако часть его удерживается, направляется планарным световодом и усиливается за счет вынужденного излучения. При этом, в процессе распространения происходит сужение спектра излучения, так как спектральные компоненты, расположенные у максимума линии спонтанного излучения, усиливаются сильнее. Спектр излучения суперлюминесцентных СИД сплошной, так же как и у поверхностных, однако значительно уже (3-5 нм). Диаграмма направленности более узкая, чем у поверхностных СИД, и несимметричная, как у лазеров, с угловыми размерами около 120...40°. Эффективность ввода излучения суперлюминесцентных СИД в многомодовые волокна выше, чем у поверхностных. Мощность излучения лежит в пределах 1-10 мВт, мощность, вводимая в многомодовый световод, 0,1-1 мВт. Светоизлучающие диоды являются наиболее подходящими источниками для низкоскоростных систем передачи информации с использованием многомодовых волоконных световодов.

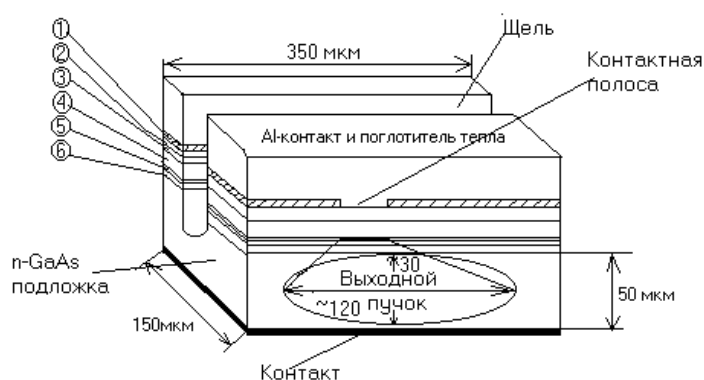


Рис. 12.18. -Схематическое изображение светодиода на основе двойной гетероструктуры с торцевым излучением

Типичная для СИД эффективность излучения, характеризующая зависимость отношений мощности на выходе, составляет 550мкВт/мА. Генерируемый СИД свет не поляризован. Спектр излучения непрерывный (рис.12.19) [8,17,27,29].

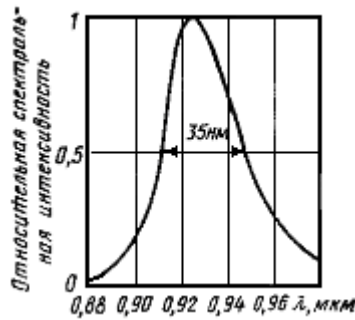


Рис. 12.19. Характерный спектр излучения СИД на основе GaAs

## Глава 13.

### ПРИБОРЫ УПРАВЛЕНИЯ ЛАЗЕРНЫМ ИЗЛУЧЕНИЕМ

Современная оптоэлектроника, как и “обычная” электроника, является твердотельной. Это означает, что работа, как отдельных элементов, так и оптоэлектронных систем в целом построена на основе различных оптических явлений, протекающих в твердых телах. Анализ этих явлений базируется на рассмотрении процессов взаимодействия электромагнитного излучения с атомными системами, проведенном ранее, вначале нашего курса.

Лазерное излучение можно использовать как высокоэффективный переносчик информации [9]. Чтобы осуществить передачу информации о событии, представляющем собой изменение во времени параметра  $A$  некоторого реального физического процесса (например, изменение температуры контролируемого объема), требуется в общем случае реализовать преобразование функции  $A(t)$  в соответствующее изменение во времени одного из параметров лазерного излучения. К определяющим параметрам лазерного излучения относится интегральная величина потока излучения  $P_{\text{л}}$ , сосредоточенного в эффективной зоне лазерного пучка с расходимостью  $\theta$ , значения частот спектральных линий  $\nu_1, \nu_2 \dots \nu_n$  тонкой структуры и соответствующие им значения начальных фаз колебаний  $\varphi_1, \varphi_2, \varphi_n$  направление оси лазерного пучка в пространстве, определяемое единичным вектором  $\bar{k}$ . В тех случаях, когда тонкая структура спектра не играет существенной роли, в расчетных формулах используется центральная частота рабочего перехода лазера  $\nu_0$ , значение начальной фазы  $\varphi_0$ , ширина спектра  $\Delta\nu$ .

Указанное выше преобразование двухэтапное: сначала (в датчике) происходит преобразование параметра анализируемого процесса  $A$  в параметр промежуточного носителя информации  $Q$  (параметр  $Q$  – напряжение электрического поля), а затем изменение  $Q$  переводится в изменение одного из параметров излучения, которые далее обозначим символом  $\chi$ . Преобразование  $Q(t) \rightarrow \chi(t)$  осуществляется специальными устройствами управления излучением.

Приборы управления по назначению делятся на два класса:

- 1) Модуляторы
- 2) Дефлекторы

Модулятор предназначен для воспроизведения произвольного, строго говоря, случайного закона изменения рабочего параметра излучения. Модуляция – это

изменение параметров света (его фазы, амплитуды, поляризации, частоты и т.д.) в зависимости от управляющего сигнала. Дефлектор служит для изменения направления распространения луча (сканирования). Причем, закон изменения направления луча после дефлектора (закон сканирования), в большинстве случаев, является детерминированным, рассчитанным заранее. Приборы управления излучением могут быть сконструированы в виде внешних или внутри резонаторных элементов. Внешние устройства изменяют параметры излучения после выхода луча. Внутренние же элементы управления излучением обеспечивают преобразование параметров луча в процессе его формирования непосредственно в генераторе.

### 13.1. Методы модуляции, классификация

В настоящее время известно большое число методов и физических явлений, на основе которых созданы оптические модуляторы [35,36]. Выделим и кратко охарактеризуем те из них, которые нашли применение в современном приборостроении:

1) Оптико-механические модуляторы интенсивности излучения. Действуют по принципу “механической шторки”. Область их применения ограничена главным образом из-за высокой инерционности и малой надежности. К достоинствам этих модуляторов следует отнести малые искажения пучка, малые потери света и широкий рабочий спектральный диапазон.

2) Интерференционные модуляторы. Модуляция интенсивности и фазы достигается в них за счет малых, не превышающих полдлины волны, механических перемещений отдельных элементов интерференционной оптической схемы. Для чего используются обычно пьезоэлектрические, магнитострикционные элементы и электромагнитные устройства.

3) Модуляторы на управляемом оптическом контакте. При малых перемещениях одной из двух соприкасающихся оптических поверхностей происходит перераспределение энергии в проходящем и отраженном пучках. Модуляторы групп 2 и 3 чувствительны к изменениям температуры.

4) Акустооптические модуляторы. Модуляция излучения в них происходит за счет дифракции на неоднородном по величине показателе преломления объемной структуры, возникающей за счет изменения давления в возбуждаемой ультразвуковой волне с управляемыми параметрами. Модуляторы этого типа достаточно светосильны, полоса модуляции достигает  $10^6$  Гц. Недостаток – чувствительны к вибрациям и температуре.

5) Полупроводниковые модуляторы. В них осуществляется управление коэффициентом пропускания за счет изменения концентрации носителей тока. Модуляторы контактные, работоспособны в инфракрасном диапазоне. Недостаток – требуют больших плотностей тока управления, и, следовательно, сложных систем принудительного охлаждения.

б) Модуляторы, основанные на явлении наведенной анизотропии. Их можно подразделить на три подгруппы:

- а) Модуляторы, основанные на магнитооптических эффектах;
- б) На эффекте фотоупругости;
- в) Электрооптические модуляторы.



На основе указанных эффектов можно получить любой вид модуляции. Наиболее быстродействующими и перспективными на сегодня являются электрооптические и акустооптические модуляторы, на которых остановимся подробно.

### 13.2. Электрооптические модуляторы света

Работа электрооптических модуляторов света основана на использовании электрооптического эффекта в веществах, сущность которого заключается в изменении показателя преломления вещества под действием электрического поля. Он проявляется в анизотропных средах, и средах которые становятся анизотропными под действием электрического поля.

Для анизотропных сред связь между компонентами векторов электрической индукции  $\vec{D}$  и напряженностью электрического поля  $\vec{E}$  задается в общем случае более сложными соотношениями. Для изотропных  $-\vec{D} = \epsilon \vec{E}$ , для анизотропных  $\vec{D} = \hat{\epsilon} \vec{E}$  Проектируя на оси координат, получим:

$$\begin{aligned} D_x &= \epsilon_{xx}E_x + \epsilon_{xy}E_y + \epsilon_{xz}E_z \\ D_y &= \epsilon_{yx}E_x + \epsilon_{yy}E_y + \epsilon_{yz}E_z \\ D_z &= \epsilon_{zx}E_x + \epsilon_{zy}E_y + \epsilon_{zz}E_z \end{aligned}$$

Девять величин  $\epsilon_{ij}$  образуют тензор диэлектрической проницаемости. Плотность электрической энергии поля кристалла, в общем случае, равна  $\rho_E = (\vec{E} \cdot \vec{D}) / 2$ , подставляя  $\vec{E}$  и  $\vec{D}$ , получим  $\rho_E = 0,5 \{ \epsilon_{xx}E_x^2 + \epsilon_{yy}E_y^2 + \epsilon_{zz}E_z^2 + 2\epsilon_{xy}E_xE_y + 2\epsilon_{xz}E_xE_z + 2\epsilon_{yz}E_yE_z \}$ . Такой результат получается при произвольном выборе координатных осей. Вид тензора  $\hat{\epsilon}$  зависит от выбора системы координат так, что все  $\epsilon_{ik} = 0$ , где  $i \neq k$  т.е. матрица  $\hat{\epsilon}$  может быть диагонализирована и приведена к виду

$$\epsilon = \begin{pmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix}$$

Система координат, в которой тензор  $\epsilon$  диагонализирован, называется главной системой координат, а ее оси-главными диэлектрическими осями. В этом случае плотность энергии

$$\rho_E = 0,5 \{ \epsilon_{xx}E_x^2 + \epsilon_{yy}E_y^2 + \epsilon_{zz}E_z^2 \}. \quad (13.1)$$

Величины  $\epsilon_x, \epsilon_y, \epsilon_z$  называют главными значениями тензора.

Введем вместо  $\epsilon_i$  показатели преломления. Для изотропных сред

$n = \sqrt{\epsilon_{отн}} = \sqrt{\frac{\epsilon}{\epsilon_0}}$ , где  $\epsilon_0$ -диэлектрическая проницаемость вакуума. Для анизотропных сред

$$n_x = \sqrt{\frac{\epsilon_x}{\epsilon_0}}, \dots n_y = \sqrt{\frac{\epsilon_y}{\epsilon_0}}, \dots n_z = \sqrt{\frac{\epsilon_z}{\epsilon_0}}.$$

Числа  $\varepsilon_i$ -можно трактовать как показатели преломления для анизотропных сред. Введем координатные оси  $(\xi, \eta, \zeta)$ , где

$$\xi = \frac{D_x}{\sqrt{2\rho_E\varepsilon_0}}, \dots, \eta = \frac{D_y}{\sqrt{2\rho_E\varepsilon_0}}, \dots, \zeta = \frac{D_z}{\sqrt{2\rho_E\varepsilon_0}}.$$

Направление этих осей совпадает с направлением осей  $(x, y, z)$  Тогда получим :

$$\frac{\xi^2}{n_x^2} = \frac{\eta^2}{n_y^2} = \frac{\zeta^2}{n_z^2} = 1, \quad (13.2)$$

где  $\xi, \eta, \zeta$ -безразмерные величины, имеют смысл некоторых показателей преломления. Рассмотрим уравнение (13.2). Это уравнение эллипсоида с главными осями вдоль направлений  $x, y, z$ . Длины полуосей эллипса равны, соответственно -  $n_x, n_y, n_z$ . Этот эллипсоид определяет поверхность постоянной плотности энергии и называется эллипсоидом преломления или оптической индикатрисой кристалла. Его использование удобно при анализе оптических свойств анизотропных кристаллов.

Эллипсоид образуется следующим образом. Из произвольной точки  $O$  кристалла отложим по всем направлениям радиусы-векторы равные, в некотором масштабе, показателю преломления кристалла в этом направлении. Концы векторов, в общем случае образуют 3х-осный эллипсоид (см. рис. 13.1)

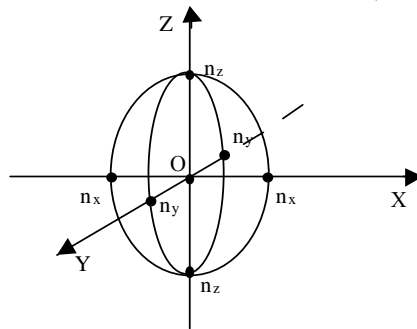


Рис. 13.1. Эллипсоид показателей преломления

Оптическая индикатриса полностью характеризует оптические свойства кристаллов. Для изотропных кристаллов  $n_x = n_y = n_z = n_0$ . Любое ее сечение есть окружность, поэтому оптические свойства таких кристаллов не зависят от направления, т.е. являются оптически изотропными. Если кристалл анизотропен, возможны два случая

1) При  $n_x = n_y \neq n_z$  сечение окружности можно провести только одно. Оптическая индикатриса – эллипсоид вращения. Такие кристаллы называются одноосными.

2) В кристаллах с  $n_x \neq n_y \neq n_z$  таких сечений может быть только два. Направления, перпендикулярные этим сечениям, называются оптическими осями кристалла.

Рассмотрим анизотропную прозрачную не проводящую ( $\sigma = 0$ ), немагнитную ( $\mu = 1$ ) среду. Пусть в этой среде распространяется плоская монохроматическая волна, вектор электромагнитного смещения которого равен  $\bar{D} = D_0 e^{-jkr}$ , где  $\bar{r}$ -радиус вектор,  $\bar{k}$  – волновой вектор ( $\bar{k} = 2\pi / \lambda = \omega/v$ ). Распространение свето-

вых волн в такой среде описывается уравнениями Максвелла. В случае отсутствия объемных зарядов имеем

$$\operatorname{div} \bar{D} = 0 \text{ или } \operatorname{div} \bar{D} = (\nabla \bar{D}) = 0, \text{ где } \nabla = (d/dx)\mathbf{i} + (d/dy)\mathbf{j} + (d/dz)\mathbf{k},$$

$\mathbf{D} = D_0 e^{-i(K_x x + K_y y + K_z z)}$ . Тогда  $(\nabla \cdot \mathbf{D}) = (\mathbf{K} \cdot \mathbf{D}) = 0$ . Отсюда следует, что  $\mathbf{K} \perp \mathbf{D}$

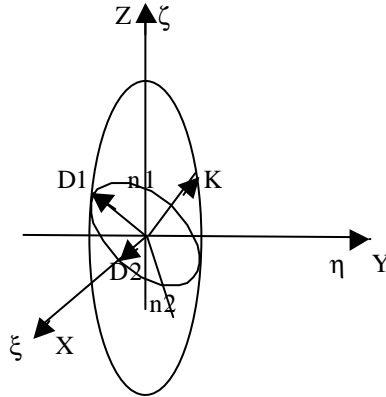


Рис. 13.2. Эллипсоид показателей преломления и нейтральные оси, соответствующие данному направлению распространения световой волны

В качестве анизотропной среды выберем одноосный кристалл. Оптическая индикатриса – эллипсоид вращения (рис. 13.2). Вектор  $\mathbf{K}$  оканчивается на поверхности индикатрисы. Решая основные уравнения Максвелла можно найти ориентацию вектора  $\mathbf{D}$  в сечении оптической индикатрисы. Рассечем индикатрису плоскостью перпендикулярной вектору  $\mathbf{K}$ . В этой плоскости лежит вектор  $\mathbf{D}$ . Сечение в виде эллипса показывает, что вектор  $\mathbf{D}$  может быть ориентирован в сечении двояко: либо вдоль малой полуоси, либо вдоль большой полуоси малого эллипса. Отсюда следует, что световая волна в анизотропном кристалле распадается на две линейно

-поляризованные монохроматические волны: (обыкновенную с коэффициентом преломления  $n_0$  и необыкновенную с коэффициентом преломления  $n_n$ ). Волны имеют одинаковую частоту, одинаковое направление волнового вектора, но различаются показателями преломления, а, следовательно, и разными скоростями распространения  $v_1 = c/n_1$ ;  $v_2 = c/n_2$ , поляризацией. Этот эффект носит название двулучепреломления. Некоторые кристаллы являются изотропными только до приложения внешнего поля. У большинства электрооптических кристаллов анизотропия появляется после приложения электрического поля. Двойное лучепреломление не испытывают волны, распространяющиеся вдоль оптической оси [7,9.11,37].

### 13.2.1. Электрооптический эффект Поккельса

Появление оптической анизотропии следует рассматривать как результат изменения диэлектрической проницаемости вещества под действием электрического поля  $\epsilon_{ik}(E) = f(E)$ , где  $E$  – напряженность электрического поля, а не поля световой волны. При  $E = 0$   $\epsilon_{ik}(0) = \epsilon^0_{ik}$ -тензор диэлектрической проницаемости кристалла в отсутствии поля. При наличии малых полей  $\epsilon_{ik}$  можно разложить в ряд

$$\epsilon_{ik}(E) = \epsilon^0_{ik} + \sum \beta_{iks} E_s. \quad (13.3)$$

Где числа  $\beta_{iks}$  образуют тензор третьего ранга, симметричного по индексам  $i$  и  $k$  ( $\beta_{iks} = \beta_{kis}$ ), таких чисел 27. Если  $\beta_{iks} \neq 0$  -возникает линейный электрооптиче-

ский эффект, его называют эффектом Поккельса. Он появляется в кристаллах под влиянием внешнего электрического поля. Одноосный кристалл обретает свойства двуосного. Если  $\beta_{iks} = 0$ , то зависимость  $\varepsilon_{ik} = f(E)$  становится квадратичной. Такой эффект носит название квадратичного электрооптического эффекта или эффекта Керра. Эффектом Керра обладают некоторые жидкости, такие как нитробензол, сероуглерод, а также кристаллы группы перовскитов  $ABO_3$ .

Используя линейный или квадратичный электрооптический эффекты можно легко построить фазовый или амплитудный электрооптический модулятор. Рассмотрим кратко эффект Поккельса [4,7,9]. На практике эффект Поккельса широко применяется в одноосных кристаллах типа  $KH_2PO_4$  (KDP).

В отсутствие внешнего поля оптическая индикатриса кристалла KDP представляет собой эллипсоид вращения, главные полуоси которого имеют длины  $n_H = n_z$  по оси  $z$  (оптической оси),  $n_x = n_y = n_0$  – по оси  $x$  и  $y$ .

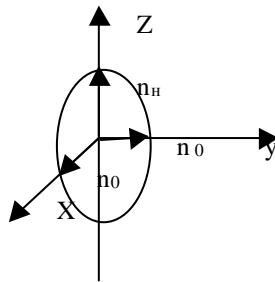


Рис. 13.3. Оптическая индикатриса кристалла KDP в отсутствие поля

Электрическое поле прикладывают вдоль оптической оси кристалла  $-z$ . В результате эллипсоид вращения оптической индикатрисы деформируется (превращается из эллипсоида вращения в трехосный эллипсоид, при этом ось  $z$  перестает быть оптической осью кристалла) и поворачивается. Главные диэлектрические оси  $x$  и  $y$  поворачиваются на  $45^\circ$ . Графически это можно пояснить следующим образом (см. рис. 13.4). Если приложить поле вдоль оси  $z$ , то круг в плоскости сечения перпендикулярной оси  $z$  превращается в эллипс. В отсутствие поля кристалл характеризовался двумя главными показателями преломления ( $n_H$  и  $n_0$ ). При наличии поля число главных показателей преломления возрастает до трех,  $n_1, n_0, n_2$

$$n_1 = n_0 + 0.5 n_0^3 r E_z, \quad n_2 = n_0 - 0.5 n_0^3 r E_z \quad (13.4)$$

где  $r$  – электрооптическая константа.

Теперь рассмотрим, что происходит со световой волной проходящей через электрооптический кристалл. На практике используют два варианта: свет распространяется вдоль направления внешнего поля, т.е. вдоль оси  $z$  (продольный электрооптический эффект.); свет распространяется перпендикулярно направлению распространения поля (поперечный электрооптический эффект).

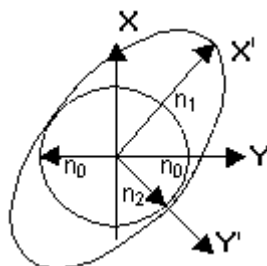


Рис. 13.4. Сечение индикатрисы плоскостью  $z=0$ 

В случае продольного электрооптического эффекта внутри кристалла распространяются две линейно – поляризованные волны, поляризации которых направлены вдоль оси  $x'$  (скорость волны  $v_1=c/n_1$ ), вдоль  $y'$  (скорость волны  $v_2=c/n_2$ ), с учетом (13.4) имеем:

$$\begin{aligned} v_1 &= c / n_0(1 + n_0^2 r E_z / 2)^{-1} \\ x_2 &= \frac{c}{n_0} \left( 1 - n_0^2 r \frac{E_z}{2} \right) \end{aligned} \quad (13.5)$$

$v_{1,2}$  связаны с фазой волны. Фаза волны  $\varphi_{1,2}$  равна:

$$\varphi_{1,2} = k \cdot L$$

где  $k = \frac{\pi}{x_{1,2}}$ ;  $L$  – длина кристалла.

После прохождения волной пути  $L$  в кристалле разность фаз, указанных волн, составит:

$$\Delta\varphi = \pi L \left( \frac{1}{x_1} - \frac{1}{x_2} \right) \approx \frac{2p \cdot n_0^3 r E_z L}{L}$$

Отметим, что разность фаз не зависит от длины кристалла. Если фиксировано приложенное к кристаллу напряжение  $U = E_z L$ , то

$$\Delta\varphi = 2p \cdot n_0^3 r \frac{U}{L} \quad (13.6)$$

То есть набег фаз зависит от величины приложенного напряжения.

Если компоненты линейно-поляризованной световой волны, после прохождения через кристалл, приобретают разность фаз, равную  $\frac{\pi}{2}$ , выходная волна остается линейно-поляризованной, но ее плоскость поляризации поворачивается на  $90^\circ$ .

Для определения напряжения  $U_{\pi/2}$ , требуемого для получения такой фазовой задержки, следует в (4.6) положить  $\Delta\varphi = \pi$ , откуда:

$$U_{\pi/2} = \frac{L}{2n_0^3 r} \quad (13.7)$$

где  $U_{\pi/2}$  – полуволновое напряжение, необходимое для поворота плоскости поляризации входного линейно-поляризованного светового пучка на  $90^\circ$ . Эта величина при  $L = 0,547$  мкм для кристаллов KDP составляет 7,5 кВ.

Рассмотрим поперечный эффект [37]. Предположим, что свет распространяется вдоль оси  $y$ . В этом случае внутри кристалла распространяются две линейно-поляризованные волны, поляризации которых направлены по осям  $x$  (скорость  $v_1$ ) и  $z$  (скорость  $v_3 = c/n_n$ ). После прохождения пути  $L$  в кристалле разность фаз указанных волн составит

$$\Delta\varphi = \omega L \left( \frac{1}{v_3} - \frac{1}{v_1} \right) = \frac{2\pi L}{\lambda} n_n - \frac{2\pi L}{\lambda} n_0 \cdot \left( 1 + \frac{n_0^2 r E}{2} \right) \approx \frac{2\pi L}{\lambda} (n_n - n_0) - \pi n_0^3 r E L / \lambda.$$

Здесь первый член обусловлен естественной анизотропией кристалла, а второй – искусственной анизотропией, наведенной внешним полем.

Таким образом, из рассмотренного выше следует, что используя продольный электрооптический эффект можно построить фазовый модулятор. Фазовый модулятор можно легко преобразовать в амплитудный или в модулятор интенсивности световой волны. Рассмотрим некоторые способы модуляции, применяемые на практике.

**Амплитудная модуляция света.** Электрооптический модулятор света впервые был создан Биллингсом [37], который применил для этой цели кристалл ADP. Схема экспериментальной установки Биллингса показана на рис. 13.5. Кристалл ADP с приложенным вдоль оси  $z$  напряжением помещался между скрещенными поляризаторами. Плоскость входного поляризатора параллельна оси  $x$ , выходного — оси  $y$ . При наличии поля  $E_z$  главные диэлектрические оси занимают положение  $x'$  и  $y'$ . Показатели преломления для волн, поляризованных вдоль этих направлений, определяются по тем же формулам, что и в случае KDP. Падающий на кристалл свет имеет равные компоненты поля  $E$  по осям  $x'$  и  $y'$ . Комплексные амплитуды  $E_{x'}$  и  $E_{y'}$  на выходе кристалла можно записать в виде

$$E_{x'} = Ae^{i\varphi_{x'}} = Ae^{i(\omega/c)[n_0 - (n_0^3/2)r_{63}E_z]L},$$

$$E_{y'} = Ae^{i\varphi_{y'}} = Ae^{i(\omega/c)[n_0 + (n_0^3/2)r_{63}E_z]L}. \quad (13.7)$$

Дифференциальный фазовый сдвиг, называемый *з а п а з д ы в а н и е м*, определяется как

$$\Gamma = \varphi_{y'} - \varphi_{x'} = \frac{\omega n_0^3 r_{63} E_z L}{c} = \pi \frac{U}{U_{1/2}}, \quad (13.8)$$

где  $U = E_z L$  — приложенное напряжение;  $U_{1/2}$  — напряжение, при котором  $\Gamma = \pi$ , т. е.

$$U_{1/2} = \frac{\lambda_0}{2n_0^3 r_{63}}, \quad (13.9)$$

где  $\lambda_0 = 2\pi c / \omega$  — длина волны в свободном пространстве. Используя данные для ADP [37]:  $n_0^3 r_{63} = 95 \cdot 10^{-12} \text{ м/В}$ , получаем  $U_{1/2} = 10 \text{ кВ}$ , при  $\lambda_0 = 0,5 \text{ мкм}$ .

Принцип действия модулятора состоит в следующем. При  $U = 0$   $\Gamma = 0$  и на выходном конце кристалла ADP  $E_{y'}$  и  $E_{x'}$  — находятся в фазе, так что поляризация выходящего луча остается неизменной. Поэтому скрещенный выходной поляризатор не пропускает этот луч. При  $U = U_{1/2}$  запаздывание равняется  $\Gamma = \pi$ . В результате этого выходящий из кристалла луч поляризован вдоль оси  $y$ , и выходной поляризатор полностью пропускает световой луч.

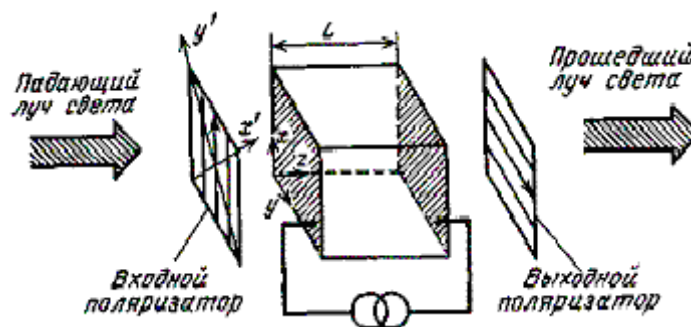


Рис. 13.5. Продольный амплитудный модулятор света, состоящий из кристалла ADP и двух скрещенных поляризаторов. Оси  $x, y, z$ -кристаллографические оси кристалла ADP. Оси  $x^1, y^1$ -главные диэлектрические оси, наведенные электрическим полем  $E$ , приложенным параллельно оси  $z$

Для того чтобы найти ту часть интенсивности падающего света, которая „проходит через устройство при произвольном напряжении  $U$ , положим интенсивность на входе, равной единице, т. е.  $E_{x^1} = E_{y^1} = 1$ . Выходная интенсивность равняется  $0.5 \cdot (E_y E_y^*)$  и с учетом (13.7)

$$\frac{I_{\text{прох}}}{I_{\text{пад}}} = \frac{1}{2} E_y E_y^* = \frac{1}{2} \left( \frac{1}{\sqrt{2}} e^{i\varphi_{x^1}} - \frac{1}{\sqrt{2}} e^{i\varphi_{y^1}} \right) \left( \frac{1}{\sqrt{2}} e^{-i\varphi_{x^1}} - \frac{1}{\sqrt{2}} e^{-i\varphi_{y^1}} \right).$$

При помощи (13.8) последнее равенство может быть записано в виде

$$\frac{I_{\text{прох}}}{I_{\text{пад}}} = \sin^2 \frac{\Gamma}{2} \quad (13.10)$$

Если запаздывание изменяется по закону

$$\Gamma = \Gamma_0 + \Gamma_m \cos \omega_m t, \quad (13.11)$$

то интенсивность проходящего света будет иметь составляющую, которая изменяется как  $\cos(\omega_m t)$ . Эта составляющая максимальна при  $\Gamma_0 = \pi/2$ . Постоянная составляющая запаздывания  $\Gamma_0 = \pi/2$  может быть обеспечена либо постоянным напряжением  $U = 1/2 U_{1/2}$ , либо «четвертьволновой» пластинкой дwoякопреломляющего кристалла.

О модуляции света СВЧ колебаниями (9 ГГц) в кристалле KDP по схеме рис. 13.5 впервые было сообщено в [38].

**Поперечная схема электрооптической модуляции.** В схеме модуляции, показанной на рис. 13.5, направления модулирующего поля и распространения света совпадают с осью  $z$ . По ряду причин, которые далее станут понятными, желательно прикладывать модулирующее поле перпендикулярно направлению распространения света. Это так называемая поперечная схема модуляции. В случае KDP мы можем создать такой модулятор в виде прямоугольной призмы с ребрами, параллельными осям  $x', y', z$ , прикладывая вдоль оси  $z$  модулирующее электрическое поле, а вдоль направления  $y'$  пропуская свет, прошедший предварительно через входной поляризатор с плоскостью поляризации, составляющей угол  $45^\circ$  с осью  $z$ , найдем, что в этом случае, запаздывание равняется

$$\Gamma = \varphi_{x^1} - \varphi_z = \frac{\omega}{c} \left( n_0 - n_n - \frac{n_0^3 r_{63} E_z}{2} \right) L_{y^1}, \quad (13.12)$$

и включает член  $(\omega/c)(n_0 - n_n)/L_{y^1}$ , вызванный естественным двойным лучепреломлением. Из вида эллипсоида показателей следует, что этот член зависит от направления распространения луча. В случае непараллельного пучка световых лучей запаздывание  $\Gamma$  зависит от угла, под которым распространяется рассматриваемый луч. Если результирующий разброс становится заметным ( $\geq \pi/4$ ), то глубина модуляции снижается. Это ограничивает допустимый угол расхождения входного пучка, что в большинстве практических случаев является весьма серьез-

езным ограничением. От подобного ограничения, обусловленного наличием естественного двойного лучепреломления, свободны кубические кристаллы, например, класса 43m, для которых  $n_n = n_0$ . В простейшей схеме, с подобного рода кристаллом, модулирующее поле параллельно оси  $z$ ), а световой луч, поляризованный под углом  $45^\circ$  к оси  $z$ , распространяется нормально к плоскости (1 1 0).

Для случая  $E_x = E_y = 0$  уравнение индикатрисы

$$\frac{x^2 + y^2 + z^2}{n_0^2} + 2r_{41}E_z xy = 1, \quad (13.13)$$

Главными диэлектрическими осями являются оси  $x'$ ,  $y'$ ,  $z$ , для которых

$$n_{x'} = n_0 - \frac{n_0^3 r_{41} E_z}{2}, \quad n_{y'} = n_0 + \frac{n_0^3 r_{41} E_z}{2}, \quad n_z = n_0,$$

а запаздывание для описанного выше устройства равняется

$$\Gamma = \varphi_{x'} - \varphi_z = \frac{\omega}{2c} n_0^3 r_{41} E_z L_{y'} = \frac{\pi}{\lambda_0} n_0^3 r_{41} E_z L_{y'}, \quad (13.14)$$

где  $\lambda_0$  — длина волны в свободном пространстве,  $r_{41}$  — электрооптическая постоянная. Сравнение с (13.8) показывает, что здесь  $\Gamma \sim E_z L_{y'}$ , в то время как в продольном случае,  $\Gamma \sim U$  и не зависит от длины. Наиболее эффективно кубические кристаллы используются, когда поле приложено нормально к плоскости (110), а луч распространяется под прямым углом к плоскости (1 1 0). В этом случае запаздывание равняется

$$\Gamma = \frac{\omega}{c} n_0^3 r_{41} E L = \frac{2\pi}{\lambda_0} n_0^3 r_{41} E L, \quad (13.15)$$

т. е. удваивается по сравнению с (13.14). Доказательство формулы (13.15) предоставляется читателю в качестве задачи.

**Фазовая модуляция.** Рассмотрим случай, когда свет падает на электрооптический кристалл так, что вектор  $\mathbf{D}$  параллелен одной из наведенных приложенным полем диэлектрических осей. Это достигается, например, в устройстве, показанном на рис. 13.5, если поляризация луча на входе параллельна оси  $x'$  (или  $y'$ ), а выходной поляризатор отсутствует. Из (13.7) видно, что электрическое поле просто изменяет фазу на величину  $\Delta\varphi_{x'} = -(\pi n_0^3 / \lambda_0) r_{63} E_z L_z$ . Если, например,  $E_z = E_m^{(z)} \cos \omega_m t$ , то электрическое поле световой волны на выходе из кристалла равняется

$$E_{x'} = A e^{i(2\pi/\lambda_0)n_0 L - i(\pi/\lambda_0)n_0^3 r_{63} L_z E_m^{(z)} \cos \omega_m t}, \quad (13.16)$$

что соответствует гармонической фазовой модуляции с индексом модуляции  $\delta = (\pi/\lambda_0)n_0^3 r_{63} L_z E_m^{(z)}$ .

Прежде чем закончить параграф, отметим, что у всех кристаллов, исследованных до настоящего времени, наблюдается малое изменение величины коэффициентов  $r_{ij}$  в диапазоне от нулевых частот до СВЧ. Это связано с тем, что как ионная, так и электронная компоненты поляризации имеют резонансы на частотах свыше  $10^{12}$  Гц, и в области низких частот, частотной зависимости не наблюдается.



Из предыдущего рассмотрения ясно, что критерием качества электрооптического материала, если не касаться преимуществ, связанных с симметрией кристалла, является фактор  $n_0^3 r_{ij}$ .

Во многих опытах по модуляции критерий качества основывается на минимальной величине мощности, которую необходимо затратить для получения данного запаздывания.

**Модуляция света посредством квадратичного электрооптического эффекта.** Центросимметричные кристаллы, в которых отсутствует линейный электрооптический эффект, могут быть применены для модуляции света с использованием квадратичного электрооптического эффекта.

Модулятор, в котором применяется кристалл класса  $m3m$  показан на рис. 13.6. Запаздывание выражается формулой

$$\Gamma = \frac{\pi(\epsilon - \epsilon_0)^2 n_0^3}{\lambda_0} (g_{11} - g_{12}) L_x E_z^2.$$

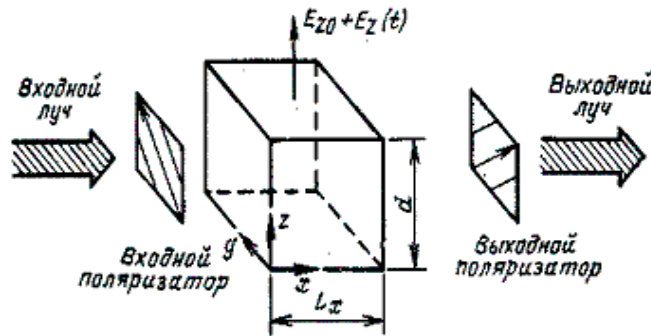


Рис. 13.6. Амплитудный модулятор света, использующий квадратичный электрооптический эффект в кристалле КТН (класс  $m3m$ )

Чтобы эффективное запаздывание было линейным, наложим переменное электрическое поле  $E_z(t)$ , содержащее «информацию», на постоянное поле смещения  $E_{z0} \gg E_z(t)$ . Тогда зависящее от времени запаздывание равняется

$$\Gamma(t) \approx \frac{2\pi(\epsilon - \epsilon_0)^2 n_0^3}{\lambda_0} (g_{11} - g_{12}) L_x E_{z0} E_z(t). \quad (13.17)$$

Сравнивая (13.17) с выражением (13.8) для запаздывания в случае линейного электрооптического эффекта ( $\Gamma = \frac{2\pi n_0^3}{\lambda_0} L_z E_z r_{63}$ ) получаем, что эффективный линейный электрооптический коэффициент, в квадратичном случае, равняется

$$r_{эфф} = (\epsilon - \epsilon_0)^2 (g_{11} - g_{12}) E_{z0}. \quad (13.18)$$

Следовательно, приращение напряжения, вызывающее  $\Delta\Gamma = \pi$ , составляет

$$(\Delta U)_{1/2} = \frac{\lambda_0 d}{2n_0^3 (\epsilon - \epsilon_0)^2 (g_{11} - g_{12}) E_{z0} L_x}, \quad (13.19)$$

где  $U(t) = E_z(t)d$ ;  $d$  — толщина кристалла в направлении оси  $z$ . Таким образом, величина управляющего напряжения, которому отвечает разность хода в  $\lambda/2$ , обратно пропорциональна приложенному полю смещения.

**Пример.** В качестве примера вычислим величину  $(\Delta U)_{1/2}$  для кристалла  $\text{KTa}_{0,65}\text{Nb}_{0,35}\text{O}_3$  (КТН), приняв  $E_{z0} = 2 \text{ кВ/см}$ ,  $L_x = d = 1 \text{ см}$ ,  $\lambda = 0,5 \text{ мкм}$ . При комнатной температуре  $(g_{11} - g_{12}) = 0,174 \text{ ж}^4/\text{с}^2$  и  $\varepsilon/\varepsilon_0 \sim 10^4$ . Результат вычисления -  $(\Delta U)_{1/2} = 75 \text{ В [37]}$ .

### 13.3.2. Параметры и характеристики модуляторов света

1. Полоса пропускания (полоса модулирующих частот) модулятора света определяется разностью верхней и нижней граничных частот.
2. Полуволновое напряжение – минимальное напряжение, необходимое для изменения коэффициента пропускания модулятора от минимального до максимального или изменения фазовой задержки на  $\pi$  радиан. Это напряжение зависит от свойств применяемых в конструкции модулятора электрооптических кристаллов.
3. Полоса прозрачности модуляторов света, определяющая диапазон длин волн лазерного излучения, пропускаемых прибором без заметных потерь (до 3 дБ), зависит от типа используемого в нем кристалла.
4. Входная апертура модуляторов света определяется толщиной кристалла и обычно не превышает нескольких мм. Увеличение апертуры требует повышения управляющего напряжения.
5. Коэффициент контрастности модулятора света, определяется отношением коэффициента пропускания прибора (при полном просветлении) к коэффициенту пропускания (при полном затемнении).
6. Емкость модулятора света. От ее величины зависит полоса пропускания.

### 13.3. Акустооптические методы управления параметрами лазерного излучения

Для управления параметрами излучения лазера, наряду с электрооптическим эффектом, возможно использование рефракционных и дифракционных явлений, возникающих при распространении оптического излучения в среде, пространственная модуляция коэффициента преломления которой создается ультразвуковыми (акустическими) волнами.

Пространственное изменение показателя преломления обусловлено упруго-оптической связью, которая существует между деформациями, создаваемыми в среде акустической волной, и показателем преломления.

При распространении акустических волн в среде образуется «фазовая решетка», представляющая собой периодически следующие друг за другом места сжатия и разрежения среды. При сжатии и разрежении изменяется плотность среды и, как следствие, показатель преломления, который можно представить в виде

$$n_r = n_0 \Delta n \cos 2\pi f_{\text{зв}}(t - x/v_{\text{зв}}) = n_0 + \Delta n \cos(\Omega_{\text{зв}} t - k_{\text{зв}} x), \quad (13.20)$$

где:  $n_0$ —показатель преломления среды в отсутствие акустической волны;

$\Delta n$ —амплитуда изменения показателя преломления;

$f_{\text{зв}}$ —частота акустических колебаний;

$v_{\text{зв}} = \lambda f_{\text{зв}}$ —скорость звука в среде;

$x$ —текущая координата;

$\lambda$ —длина звуковой волны;

$$\Omega_{3B} = 2\pi f_{3B}; \kappa_{3B} = \Omega_{3B}/v_{3B}.$$

Амплитуда изменения коэффициента преломления  $\Delta n$  определяет количественную сторону, так называемого, явления фотоупругости (упругооптический эффект), заключающегося в изменении оптической индикатрисы показателя преломления среды при воздействии статических и переменных механических напряжений. Подробно это явление рассмотрено в [9].

При создании одномерных напряжений (деформаций) в изотропной среде её свойства описываются одноосным эллипсоидом показателей преломления с оптической осью, совпадающей с направлением деформации. В целом же, характер деформации оптической индикатрисы зависит от свойств среды, от числа отличных от нуля упругооптических коэффициентов ( $p_{kl}$ ), которые по аналогии с электрооптическими константами связывают изменения поляризационных констант  $\Delta a_k$  с амплитудами деформаций  $\varepsilon_l$ . Эта зависимость имеет вид

$$\Delta a_k = \sum_{l=1}^n p_{kl} \cdot \varepsilon_l. \quad (13.21)$$

Максимально возможное число упругооптических коэффициентов  $p_{kl} = 36$ , число ненулевых коэффициентов зависит от свойств среды (кристаллов). Выражение для  $\Delta n$  вдоль направления деформации в случае изотропной среды представляется в виде:

$$(\Delta n)_l = 0,5n^3 p_{kl} \varepsilon_l. \quad (13.22)$$

Амплитуда деформации  $\varepsilon_l$  выражается через мощность звуковой волны  $P_{3B}$  и параметры активной среды :

$$\varepsilon_{kl} = [2P_{3B}/(aL\rho v_{3B})]^{1/2}, \quad (13.23)$$

где  $\rho$ —плотность среды;  $aL$ —площадь сечения акустического столба в направлении распространения звуковой волны.

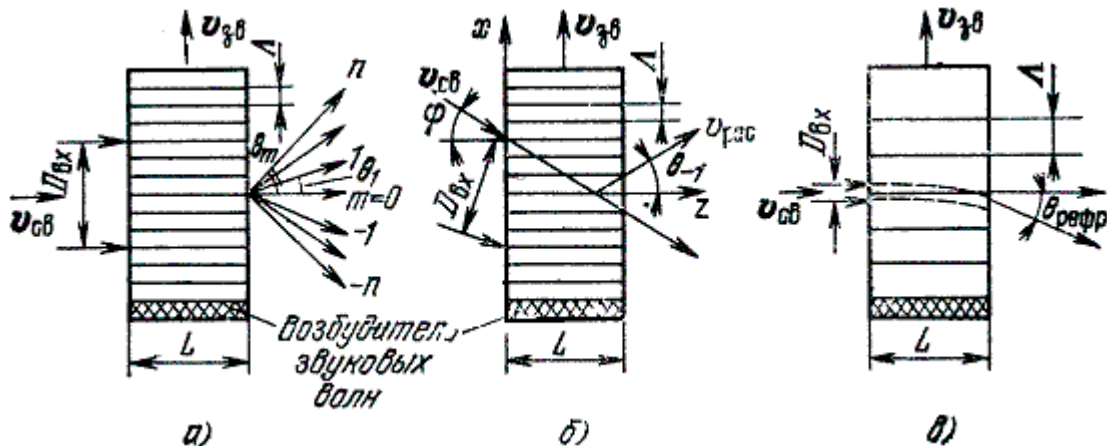


Рис. 13.7. Взаимодействие световой волны с акустически возмущенной средой

При использовании бегущих акустических волн фазовая решетка в среде движется со скоростью звука. При использовании стоячей звуковой волны период фазовой решетки равен  $L/2$ ; с учетом неравенства  $c/v_{3B} \gg 1$ , где  $c$ —скорость света, «фазовую решетку» можно считать практически неподвижной для распространяющейся световой волны.

При прохождении оптического излучения через акустически возмущенную среду различают три случая взаимодействия:

1) диаметр светового пучка  $D_{\text{ex}}$  (рис.13.7,а) значительно больше длины акустической волны, угол между направлением распространения света и фронтом акустической волны равен  $90^\circ$ . В этом случае имеет место «нормальная» дифракция (дифракция Рамана—Ната), характеризующаяся появлением многих дифракционных лепестков, угловое положение которых определяется из условия  $\sin\theta_m = m\lambda / \Lambda$ , где  $m$ —целое число;

2)если пучок света падает на бегущую акустическую волну не нормально к направлению распространения звука, а под небольшим углом  $\varphi$  (рис. 13.7,б), то при условии  $D_{\text{вх}} \gg \Lambda$  наблюдается дифракция с ярко выраженной дискриминацией дифракционных максимумов высшего порядка. Если для длины взаимодействия  $L$  и угла  $\varphi$  выполняются условия

$$\pi L \lambda / \Lambda^2 \gg 1, \quad (13.24)$$

$$\varphi = \arcsin(\lambda / 2\Lambda), \quad (13.25)$$

то, практически, все падающее излучение дифрагирует в один дифракционный максимум с направлением

$$\theta_{-1} = -\arcsin(\lambda / \Lambda). \quad (13.26)$$

Соотношения (13.24) и (13.25) определяют условия дифракции Брегга ;

3)при диаметре светового пучка, меньшем длины звуковой волны, излучение по мере прохождения через среду, испытывает рефракцию (рис. 13.7,в). Угол рефракции определяется градиентом показателя преломления, который равен

$$\text{grad } n = dn / dr = -(2\pi \Delta n / \Lambda) \sin(2\pi r / \Lambda). \quad (13.27)$$

Все три взаимодействия светового пучка с акустическим столбом могут быть использованы для создания акустических модуляторов и сканирующих устройств. Такая возможность следует из соотношений (13.22) и (13.23), которые указывают и на способ управления дифракционными явлениями путем изменения интенсивности звуковой волны и ее длины волны. Модуляция параметров акустической волны приведет к модуляции параметров фазовой решетки и, следовательно, к изменению параметров дифрагированного оптического излучения, которое обнаруживается при прохождении его через пространственные или поляризационные фильтры. Изменяются не только амплитудные и фазовые характеристики оптического пучка, направление его распространения, но и поляризационное состояние, так как дифракция в анизотропных средах сопровождается поворотом плоскости поляризации рассеянной (дифрагированной) волны.

Рассмотрим подробнее дифракцию света на ультразвуковых волнах с целью выяснения количественной связи между интенсивностью и углами отклонения световых пучков с углами дифракции и параметрами наведенной в среде фазовой решетки [9].

### 13.3.1. Дифракция Рамана – Ната

Если выполняется условие  $D_{\text{вх}} \gg \Lambda$ , а ширина акустического столба  $L$  мала, то для взаимно ортогонального распространения световой и акустической плоских монохроматических волн оптическое излучение, прошедшее пространственную фазовую решетку, изменение коэффициента преломления которой описывается (13.20), можно представить в виде

$$E(x, L) = E_0(x, 0) \exp\{-i k l [n_0 + \Delta n \cos(kx - \Omega_{\text{зв}} t)]\}. \quad (13.28)$$

Соотношение (13.28) справедливо для бесконечно тонкой фазовой решетки, режима бегущей волны и не учитывает дифракционных явлений на границах ре-

шетки. Используя принцип Гюйгенса— Кирхгофа, по известной структуре поля, на выходе фазовой решетки можно найти распределение оптического излучения в окружающем пространстве. При  $kr \gg 1$ , где  $r$ —расстояние от точки на фазовой решетке до точки наблюдения (условие дальней зоны), выражение для напряженности электрического поля световой волны с учетом (13.28) примет вид

$$E(L, \Theta) = E_0 \frac{a^2}{\lambda} \exp\left(\frac{i \cdot 2\pi L n}{\lambda}\right) \sum_{m=-\infty}^{\infty} J_m\left(\frac{2\pi \Delta n L}{\lambda}\right) \times \frac{\sin\left[\left(\frac{\pi a}{\lambda}\right)\left(\sin \Theta - \frac{m\lambda}{\Lambda}\right)\right]}{\left(\frac{\pi a}{\lambda}\right)\left(\sin \Theta - \frac{m\lambda}{\Lambda}\right)} \cdot \exp\left[-im(\Omega_{3B}t) - \frac{\pi}{2}\right], \quad (13.29)$$

где  $J_m$ —функция Бесселя  $m$ -го порядка;  $a$ —поперечный размер светового пучка;  $\theta$ —угол между нормалью к поверхности фазовой решетки и направлением распространения световой волны. Если фазовая решетка индуцирована в среде стоячей ультразвуковой волной, то выражение для распределения поля на ее выходе имеет вид

$$E(L, t) = E_0 \exp[ikL[n + \Delta n \cos(k_{3B}x) \cos(\Omega_{3B}t)]], \quad (13.30)$$

а поле в дальней зоне, соответственно,

$$E(L, \Theta) = \frac{E_0 a^2}{\lambda} \exp(-i \cdot k \ell_n) \sum_{m=-\infty}^{\infty} (-i) J_m(kL \Delta n \cos(\Omega_{3B}t)) \times \frac{\sin\left[\left(\frac{a\pi}{\lambda}\right)\left(\sin \Theta - \frac{m\lambda}{\Lambda}\right)\right]}{\left(\frac{a\pi}{\lambda}\right)\left(\sin \Theta - \frac{m\lambda}{\Lambda}\right)} \quad (13.31)$$

Из (13.29) и (13.31) следует, что при

$$\Theta_m = \arcsin(m\lambda / \Lambda) \quad (13.32)$$

в направлении  $\theta_m$  поле принимает максимальные значения. Условие (13.32) определяет направление дифракционных максимумов, интенсивность поля в которых зависит от максимального значения разности фаз ( $2\pi L \Delta n / \lambda$ ), приобретаемой волной при ее прохождении через фазовую решетку. В случае бегущей ультразвуковой волны интенсивность поля в  $m$ -м дифракционном максимуме пропорциональна квадрату функции Бесселя  $m$ -го порядка и не зависит от времени, если  $\Delta n$  постоянно:

$$E_m^2 = E_0^2 J_m^2(2\pi L \Delta n / \lambda).$$

В дифракционных максимумах частота излучения не совпадает с частотой падающего излучения, и этот частотный сдвиг растет с ростом номера дифракционного максимума:

$$\omega_m = \omega_0 + m\Omega_{3B}. \quad (13.33)$$

Соотношение (13.33) отвечает требованию соблюдения закона сохранения энергии при дифракции света на ультразвуковых волнах.

Для стоячей акустической волны выражение для интенсивности света в  $m$ -м дифракционном максимуме имеет вид

$$E_m^2 = E_0^2 J_m^2(2\pi L \Delta n \cos \Omega_{3B}t / \lambda) \quad (13.34)$$

Согласно (13.22), (13.32) и (13.34) соотношение между интенсивностями неотклоненного и дифрагированного излучения можно изменять, управляя мощностью акустической волны. В силу распределения излучения по всем дифракционным максимумам к.п.д. модуляторов, использующих рассматриваемый вид дифракции, невысок, а максимальная частота модуляции ограничена временем прохождения акустической волны через поперечное сечение светового пучка.

При использовании стоячих акустических волн модуляция может быть осуществлена в соответствии с (13.34) на частоте ультразвуковой волны  $\Omega_{зв}$ , на частоте модуляции мощности ультразвуковой волны, а также на частоте, удвоенной по сравнению с частотой ультразвуковой волны за счет наведенного двулучепреломления, которое дважды за период акустической волны испытывает проходящее через среду оптическое излучение. Частота ультразвука в режиме стоячих волн выбирается с учетом размеров акустооптической среды ( $a$ ) или акустооптического резонатора ( $L_p$ ) и должна отвечать условию

$$\Omega_{зв}^m = 2\pi m v_{зв} / (2L_p).$$

Относительная расстройка частоты модулятора на стоячих волнах составляет единицы процентов при еще приемлемом изменении  $\Delta n$ , но эта узкополосность модулятора в какой-то мере компенсируется наличием достаточно широкого интервала, определяемого частотными характеристиками возбудителя, внутри которого размещается набор собственных частот резонатора  $\Omega_m$  [9].

### 13.3.2. Бреговская дифракция

Рассмотрим распространение плоской монохроматической волны света в изотропной среде, коэффициент преломления которой меняется в соответствии с (13.20), а направление распространения составляет угол  $\varphi$  с нормалью к поверхности фазовой решетки (рис.13.7, б). При условии  $\lambda / \Lambda \ll 1$  волновое уравнение для вектора напряженности электрического поля волны с хорошим приближением принимает вид

$$\Delta E(x, z, t) = \mu_0 (\partial^2 / \partial t^2) [\varepsilon(x, t) E(x, z, t)]. \quad (13.35)$$

Решение (13.35) в силу условия (13.33) удобно представить в виде суперпозиции плоских волн:

$$E(x, y, z) = \sum_{-\infty}^{\infty} E_m \cdot \exp\{+i(\omega + m\Omega_{зв})t - (k \sin \varphi + m k_{зв})x - k \cos \varphi z\}, \quad (13.36)$$

где  $E_m$ —амплитуда плоской волны, отклоненной в направлении  $m$ -го дифракционного максимума.

Подставляя (13.20), (13.36) в (13.35) и учитывая, что  $\Delta n/n \ll 1$  и  $m\Omega_{зв}/\omega \ll 1$  для любых  $m$ , для граничных условий

$$\begin{cases} E_m(z=0) = E_{m0} \text{ при } m=0, \\ E_m(z=0) = 0 \text{ при } m \neq 0, \end{cases} \quad (13.37)$$

получаем решение (13.35) для  $m$ -й составляющей поля в следующем виде

$$E_m = E_{m0} \frac{k \Delta \varepsilon}{2 \varepsilon \cos \varphi} \exp(-i \beta_m Z) \int_{-1/2}^{+1/2} E_{m=-1}(Z') \cdot \exp(i \beta_m Z') dz', \quad (13.38)$$

где

$$\beta_m = \frac{2mk_{зв}k \left( \frac{v_{зв}^2}{c^2} - \sin^2 \varphi \right) - m^2 k_{зв}^2 \left( 1 - \frac{v_{зв}^2}{c^2} \right)}{2k \cos \varphi}, \quad (13.39)$$

а  $\Delta \varepsilon$ — амплитуда изменения диэлектрической проницаемости, соответствующая величине  $\Delta n$ .

Согласно (13.38) максимальное значение амплитуды волны, дифрагированной в направлении дифракционного максимума с  $m=\pm 1$ , имеет место при выполнении условия  $\beta_{\pm 1}=0$  или с учетом (13.39):

$$\sin\varphi \approx \pm\lambda / 2L = \sin\theta_{\pm 1}.$$

Последнее равенство совпадает с (13.25) и определяет угол Брегга. Но, для того чтобы из всех  $m$  направлений сохранился лишь  $m = -1$  дифракционный максимум, необходимо наложить дополнительное ограничение на «толщину» фазовой решетки  $L$ , при которой за счет интерференции плоских волн амплитуды  $E_m$  для всех других направлений, кроме исходного ( $m = 0$ ) и брегговского ( $m = -1$ ), станут пренебрежимо малы. Таким условием является условие (13.24).

Согласно (13.38) отношение интенсивности волны, дифрагированной в направлении максимума с  $m = \pm 1$ , к интенсивности исходной представляется в виде

$$\frac{E_{-1}^2}{E_0^2} = \frac{k^2(\Delta\varepsilon)^2}{4\varepsilon^2 \cos^2 \varphi} \frac{\sin^2\left(\frac{\beta_1 L}{2}\right)}{\left(\frac{\beta_1 L}{2}\right)^2} \quad (13.40)$$

При выполнении условий (13.24) и (13.25) интенсивность поля, дифрагированного в  $m=-1$  порядок, определится из условия

$$E_{-1}^2 = E_0^2 \sin^2 2\pi L \Delta n / \lambda. \quad (13.41)$$

Из соотношения (13.41) следует, что при определенной мощности ультразвуковой волны падающее излучение может быть полностью отклонено на угол Брегга. Согласно (13.22) и (13.23) это возможно, если

$$P_{зв} = a\rho v_{зв}^3 / (Lk^2 n_0^6 p_{kl}^2), \quad (13.42)$$

то есть чем выше показатель преломления среды и ее упругооптические коэффициенты, чем меньше плотность ее и скорость распространения звука в ней, тем выше эффективность модулятора или дефлектора, построенного на основе дифракции Брегга.

Полученные результаты относятся к идеальной модели плоских волн – как акустических, так и световых. На практике, световой и акустический пучки имеют конечную ширину и расходимость. Интересны два предельных случая, когда расходимость акустического пучка  $\theta_{зв}$  значительно больше расходимости светового пучка  $\theta_{св}$  ( $\theta_{зв} \gg \theta_{св}$ ) и когда расходимость акустического пучка значительно меньше расходимости светового пучка ( $\theta_{зв} \ll \theta_{св}$ ).

В первом случае при изменении частоты акустической волны условие дифракции Брегга (13.25) при правильно выбранном угле падения  $\varphi$  может быть выполнено для светового пучка малой расходимости во всем диапазоне углов дифракции, равном  $\Delta\theta_{-1} = 2\theta_{зв}$ , причем число разрешаемых элементов  $N_{ак}$  в этом угле сканирования будет равно  $N_{ак} \approx 2\theta_{зв} / \theta_{св}$ .

В случае расходимости пучков, вызванной дифракционными явлениями (т. е.  $\theta_{св} \approx \lambda / a$ ,  $\theta_{зв} \approx \lambda / L$ ), разрешение определяется геометрией акустического столба и соотношением длин волн света и звука:

$$N_{ак} \approx 2aL / (L\lambda). \quad (13.43)$$

Если мощность акустического поля достаточно велика, то за счет малой расходимости светового пучка удастся отклонить практически всю падающую мощность в направлении угла Брегга, но при этом на отклонение будет затрачена только часть мощности акустического пучка.

При условии  $\theta_{св} \gg \theta_{зв}$  на отклонение затрачивается вся мощность акустического пучка, расходимость светового пучка возрастает на  $\theta_{зв}$ , а под углом Брегга отклоняется только часть мощности лазерного излучения, причем разрешающая способность модулятора уменьшается по сравнению с (13.34) почти вдвое.

Диапазон изменения частоты акустической волны при брегговской дифракции для заданной расходимости звукового столба  $\theta_{зв}$  и угле падения  $\varphi$  определяется через векторное равенство постоянных распространения, следующее из закона сохранения импульса:

$$k_с = k + k_{зв}, \quad (13.44)$$

где  $k_с$  — вектор постоянной распространения отклоненного светового пучка.

Расходимость акустической волны  $\theta_{зв}$  соответствует диапазону изменения постоянной распространения  $\Delta k_{зв}$ , и из (13.44) он равен

$$\Delta k_{зв} = 2k\theta_{зв}\cos\varphi = 2\pi\Delta f_{зв} / v_{зв}.$$

Отсюда допустимый диапазон изменения частоты ультразвука

$$\Delta f_{зв} = 2\theta_{зв}v_{зв}\cos\varphi / \lambda. \quad (13.45)$$

Дифракция Брегга в таком диапазоне частот в изотропных средах требует достаточно большой расходимости ультразвукового пучка, что приводит как к неэффективному использованию мощности акустического пучка, так и к отклонению не всей мощности оптического излучения. Отмеченные недостатки в значительной степени устраняются при использовании анизотропных сред, где требуемое разрешение достигается при меньшей расходимости звуковой волны.

К акустооптическим средам предъявляется ряд требований, основные из которых следующие:

- высокий коэффициент акустооптического качества, определяемый величинами  $\rho_{kl}, \rho, v_{зв}$ ;
- малые потери в области рабочих длин волн оптического излучения;
- малое акустическое поглощение на рабочих звуковых волнах;
- высокая оптическая однородность;
- высокая химическая стойкость, механическая прочность, стабильность параметров, наличие отработанной технологии выращивания кристаллов высокого качества, а для ряда применений и стойкость к лазерному излучению.

Из кристаллических сред в настоящее время наиболее перспективными для видимого и ближнего инфракрасного диапазона следует признать парателлуриит ( $\text{TeO}_2$ ), молибдат свинца ( $\text{PbMo}_4$ ), соли альфа-йодистой кислоты ( $\alpha\text{-HfO}_3$ ). В инфракрасной области предпочтительны теллур, арсенид галлия, германий. Твердые материалы, как правило, обладают меньшим, чем жидкости, звукопоглощением, и поэтому могут применяться на более высоких частотах, вплоть до 500-700 МГц.

Электрооптические модуляторы и дефлекторы являются широкополосными приборами, способными работать на высоких частотах. Нижний предел постоянной времени, который определяется измерительной аппаратурой, составляет 3 нс, поэтому ожидается хорошая работа данных приборов в диапазоне гигагерц. Верхний частотный предел определяется допустимой мощностью рассеяния в цепи управления, которая равняется  $P = \frac{1}{2}CV^2\Delta f$ , где  $C$  — емкость прибора и  $\Delta f$  — полоса частот. Для наиболее перспективного материала (танталониобат лития) требуемая управляющая мощность менее 0,2 мВт/МГц. Электрооптические полупроводники



обладают большими потерями и, следовательно, имеют верхний частотный предел, более низкий, чем электрооптические диэлектрики, подобные танталониобату лития. Потери, связанные с наличием свободных носителей в полупроводниках, можно использовать для модуляции интенсивности отклоняемого луча.

## РАЗДЕЛ ТРЕТИЙ

### ПРИЕМ И ПЕРЕДАЧА ИЗЛУЧЕНИЯ В ОПТОЭЛЕКТРОНИКЕ

#### Глава 14. Приемники оптического излучения

Фотоприемник – это прибор, который поглощает энергию оптического излучения и преобразует её в электрическую с определенным КПД.

Классификация фотоприемников достаточно условна и может осуществляться по многим признакам.

1. По среде, в которой происходит движение фотоэлектронов или оптически генерированных пар носителей заряда. Различают вакуумные и твердотельные (полупроводниковые) фотоприемники.

2. По структуре полупроводникового материала фотоприемника. Приемники с однородной структурой полупроводника. Приемники с р-п-переходами. Например: Фоторезисторы, фотодиоды, фототранзисторы и т.д.

3. Приборы с внутренним усилением фототока, лавинные фотодиоды, фототранзисторы, канальные фотоумножители т.д.

4. По времени отклика на оптический сигнал и по частотным свойствам.

5. По области применения (фотометрия, солнечные элементы, оптоэлектронные схемы, волоконно-оптическая связь и т.д.

#### 14.1. Физические принципы приема оптического излучения

##### 14.1.1. Поглощение света в твердых телах

В основу изложения этой главы взяты материалы из [41]. При прохождении света через вещество, его интенсивность уменьшается. Часть энергии излучения поглощается и идет на увеличение энергии электронов или теплового движения атомов. Рассмотрим возможные переходы электронов в кристаллах под действием света.

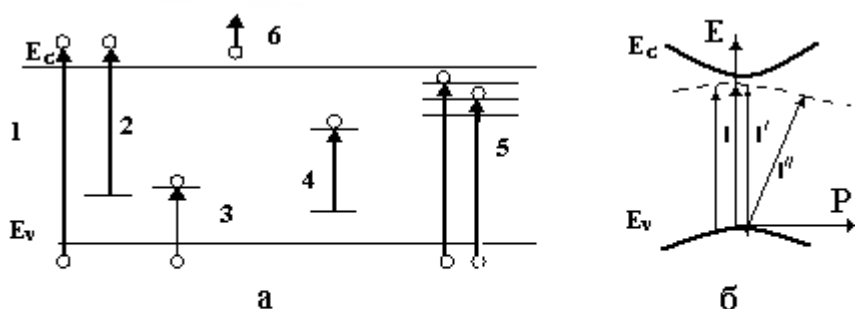


Рис. 14.1. Переходы электронов в кристаллах под действием света (а); прямые и непрямые межзонные переходы (б).

На рис. 14.1,а введены следующие обозначения переходов:

1-зона – зона  $h\nu > \Delta E = E_c - E_v$ ; 2- локальный уровень примеси - зона или дефект кристалла – зона; 3 - зона – локальный уровень примеси или зона – дефект

кристалла; 4- внутри центровые пере- ходы; 5- зона – экситонный уровень; 6- свободные носители.

Переходы 1,3 изменяют электропроводность твердых тел (внутренний фотоэффект); при переходах 4, 5, 6 – электрон не освобождается и процесс поглощения света не приводит к изменению электропроводности кристалла.

**Межзонные переходы.** При поглощении электроном фотона должны выполняться законы сохранения энергии и импульса, поэтому более полно процесс поглощения света описывается с помощью диаграммы, учитывающей изменение энергии  $E$  и импульса  $P$ . На рис. 14.1,б показаны прямые и непрямые межзонные переходы.

Сплошная линия соответствует полупроводнику, у которого минимумы энергии электрона и энергии дырки (она увеличивается вниз на рисунке) приходятся на одно и то же значение импульса (так называемые прямые зоны). В наиболее простом случае кинетическая энергия электрона связана с импульсом соотношением  $E = P^2 / 2m$ .

Стрелки 1 (рис. 14.1,б) изображают прямые переходы. Они соответствуют минимумам энергии электрона и энергии дырки. Эти переходы происходят без изменения импульса. Импульс фотона  $h\nu / c$  очень мал и соответствующим приращением импульса электрона, поглотившего фотон, можно пренебречь.

В случае материалов с прямыми зонами преобладают прямые межзонные переходы, не требующие участия третьей частицы. GaAs ( $\Delta E = 1,4$  эВ); CdSe (1,8 эВ); CdS (2,5 эВ); ZnS (3,7 эВ).

На рис. 14.1,б - 1' изображены непрямые межзонные переходы. В этом случае, происходит изменение импульса электрона, поглотившего фотон. В процессе поглощения участвует третья частица – фонон (квант колебательной энергии кристалла).

Может осуществляться случай, когда экстремумы энергий электронов и дырок на диаграмме  $E (P)$  приходятся на различные  $P$  (штриховая линия на рис. 14.1,б). Здесь наблюдаются только непрямые (1'') переходы (рис. 14.1,б). К материалам с непрямыми переходами относятся, например, Ge ( $\Delta E = 0,7$  эВ); Si (1,1 эВ); AlAs (2,2 эВ); GaP (2,3 эВ); SiC (2,4 – 3,1 эВ).

Если под действием света осуществляются переходы с участием примесных уровней, переходы 2,3 (рис. 14.1,а), то третьей частицей, обеспечивающей выполнение закона сохранения импульса, может быть примесный центр.

Количественно поглощение света описывается следующим образом. Пусть на вещество падает поток излучения  $\Phi_0$ , а на плоскость на глубине  $x$  под поверхностью вещества - поток  $\Phi(x)$ . В слое толщиной  $dx$  поглощается энергия  $d\Phi(x)$ , которая пропорциональна падающему потоку и толщине  $dx$ , то есть  $-d\Phi(x) = \alpha\Phi(x)dx$ , где  $\alpha$  – коэффициент пропорциональности,

$$\frac{d\Phi(x)}{\Phi(x)} = -\alpha dx. \text{ Интегрируя, получаем: } \Phi(x) = \Phi_0(x)e^{-\alpha x}. \quad (14.1)$$

При  $x = x^* = 1/\alpha$  имеем  $\Phi(x) = \Phi_0/e$ .

$x^*$  - называется длиной поглощения света. Это расстояние, на котором поток уменьшается в  $e$  раз ( $e=2,72$ ).

Пример: если при собственном поглощении  $\alpha = 10^5 \text{ см}^{-1}$ , то  $x^* = 0,1$  мкм. В случае примесного поглощения света  $\alpha$  гораздо меньше, так как концентрация примесных электронов или акцепторных уровней сравнительно мала (обычно порядка  $10^{16} - 10^{18} \text{ см}^{-3}$ ). При концентрации примеси около  $10^{17} \text{ см}^{-3}$  коэффициент примесного поглощения  $\alpha = 10 \text{ см}^{-1}$ .

В кристаллах с большой долей ионной связи возможно возбуждение колебаний решетки световыми волнами. Поглощение такого типа становится

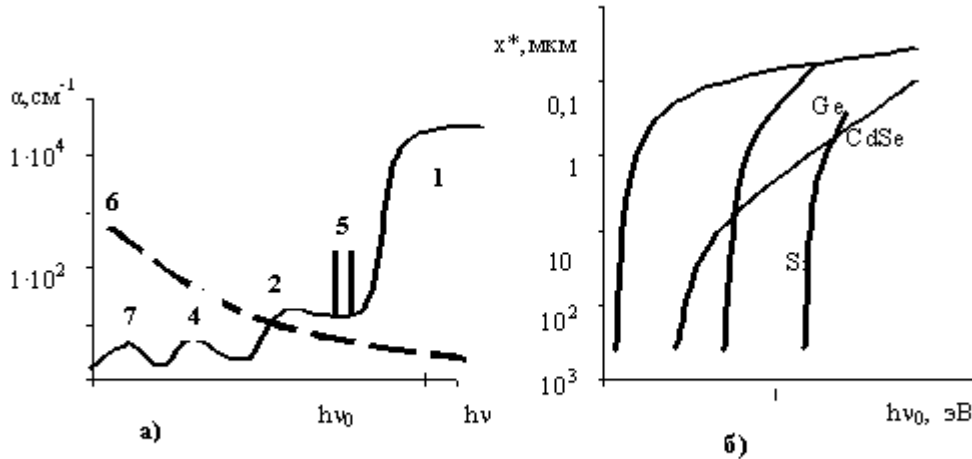


Рис. 14.2. Схематический спектр поглощения света твердым телом (а) и спектры собственного поглощения нескольких веществ (б)

особенно сильным в области резонанса ( $\sim 10^{13}$  Гц). Поглощение света свободными носителями заряда пропорционально их концентрации

Различные типы поглощения проявляются при различных энергиях квантов  $h\nu$ . Обобщенная зависимость  $\alpha$  от  $h\nu$  выглядит примерно так, как показано на рис. 14.2. Собственное поглощение начинается на частоте  $\nu_0$ , соответствующей

$h\nu = \Delta E$ . Примесное поглощение создает полосы 2 и 4, (нумерация полос та же, что и для переходов, изображенных на рис. 14.1). В широкой полосе частот присутствует слабое поглощение света свободными электронами или дырками (6), в инфракрасной области проявляется пик (7), связанный с поглощением излучения ионами решетки ( $h\nu \approx 0,1 \text{ эВ}$ ).

У края собственного поглощения располагаются узкие линии экситонного поглощения (5). Экситон представляет собой слабо связанные электрон и дырку, образующие систему, которая имеет серию энергетических уровней. Экситон является нейтральным образованием и его появление не приводит к изменению электрических характеристик образца. Если температура достаточно высока, чтобы под действием тепловой энергии электрон перешел в зону проводимости, то при экситонном поглощении получится тот же результат, что и при собственном поглощении света.

В фотоприемниках обычно используют собственное поглощение. Примесное поглощение используют для расширения спектральной характеристики в длинноволновую область. Примеры зависимости  $x^*$  от энергии фотонов  $h\nu$  в области собственного поглощения кристаллов приведены на рис. 14.2,б.

### 14.1.2 Явление фотоэффекта

Под фотоэффектом будем понимать изменение электрических и физических характеристик материалов под действием светового потока. В данном пособии будут рассмотрены, в основном, явления внутреннего фотоэффекта и приборы, основанные на этом явлении.

Введем понятие фотопроводимости и фототока. Фотопроводимость – это изменение сопротивления однородного полупроводника под действием света. Разность потенциалов или направленное движение носителей заряда, возникающее в полупроводнике при условии разделения пар носителей носит название фото-ЭДС (фототока). Разделение носителей может происходить за счет разного коэффициента диффузии ( $D_{\text{диф}}$ ) – диффузионная фото-ЭДС, за счет электрического поля – фотогальванический эффект, за счет магнитного поля – фотомагнитный эффект, за счет комбинации этих способов разделения зарядов.

Приборы, в которых происходит преобразование светового потока в фототок или фото-ЭДС, называют фотоэлементами.

**Условия существования фотоэффекта.** Если оптически генерированные неравновесные носители (n-электроны и p-дырки) претерпевают разделение в объеме освещенного полупроводника, происходит не только изменение его фотопроводимости но и образование разности потенциалов между участками освещенного образца и возникновение фототока.

Рассмотрим некоторые принципы разделения зарядов и условие возникновения фототока или фото-ЭДС:

1. *Фотогальванический эффект.* Неравномерно освещенный полупроводник, помещенный в магнитное поле дает сравнительно большую величину фото-ЭДС. И это явление называется фотомагнитоэлектрическим эффектом (ФМЭ).

2. *Эффект Дембера (диффузионный фотоэффект).*

При неравномерном освещении однородного полупроводника, характеризующегося неодинаковыми коэффициентами диффузии электронов и дырок  $D_n \neq D_p$  образуется градиент концентрации  $\Delta n$  и  $\Delta p$  вдоль оси, совпадающей с направлением освещения, который создает диффузионные потоки.

3. *Градиентный фотоэффект.* Образование фото - ЭДС при, освещении полупроводника с неравномерно распределенной примесью в объеме, имеющего градиент проводимости  $dy/dx$  – называется объемно *градиентным* фотоэффектом.

4. *Фотопьезоэлектрический эффект.* При освещении односторонне сжатого полупроводника возникает фото - ЭДС за счет *фотопьезоэлектрического* эффекта. Фото - ЭДС возникает на гранях, перпендикулярных направлению сжатия, величина ее пропорциональна интенсивности света и давлению, вызывающего деформацию кристалла.

5. *Фото-термо - ЭДС* возникает при освещении неравномерно нагретого полупроводника. В этом случае изменяется величина термо-ЭДС, определяемая разностью температур между холодными и нагретыми участками образца.

Рассмотрим подробно каждый из типов, создающий фото - ЭДС.

**Эффект Дембера.** Если свет сильно поглощается полупроводниковым элементом, то образуется градиент концентрации неравновесных носителей заряда вдоль оси, совпадающей с направлением освещения. При этом создаются диффузионные потоки электронов и дырок от освещенной поверхности к неосвещенной

$$I_n = eD_n \text{grad } n, \quad (14.2)$$

$$I_p = -eD_p \text{grad } p. \quad (14.3)$$

Если коэффициенты диффузии электронов  $D_n$  и дырок  $D_p$  сделать одинаковыми, то диффузионный ток дырок станет уравновешен диффузионным током электронов. При этом положительные и отрицательные заряды распределялись бы по объему полупроводника равномерно.

Неодинаковость диффузионных токов электронов и дырок приводит к образованию пространственного заряда и созданию статического электрического поля с напряженностью  $E$ , тормозящего более быстрые и ускоряющего более медленные носители.

В состоянии равновесия диффузионные токи уравновешиваются дрейфовыми.

$$I = E_x e[\mu_n n(x) + \mu_p p(x)] - e(D_n - D_p)(\text{grad } n)_x, \quad (14.4)$$

где  $\mu_n$  и  $\mu_p$  - подвижности электронов и дырок,  $E_x$  - напряженность электрического поля в направлении  $x$ .

откуда

$$E_x = \frac{D_n - D_p}{\mu_n n(x) + \mu_p p(x)} e(\text{grad } n)_x = \frac{D_n - D_p}{\sigma(x)} e(\text{grad } n)_x.$$

Полная разность потенциалов между освещенной и не освещенной поверхностью полупроводника может быть вычислена путем интегрирования  $E_x$  по толщине образца (с учетом граничных условий на поверхности и с использованием  $eD = \mu kT$ ).

$$U = \int_0^d E_x dx \approx \frac{kT}{\sigma} (\mu_n - \mu_p) \Delta \bar{n}, \quad (14.5)$$

где  $\Delta \bar{n}$  - среднее приращение концентрации электронов в объеме. Образование фото-ЭДС Дембера показано схематически на рис. 14.3.

Величина фото-ЭДС определяется разностью коэффициентов диффузии основных и неосновных носителей заряда и изменяется пропорционально интенсивности освещения образца.

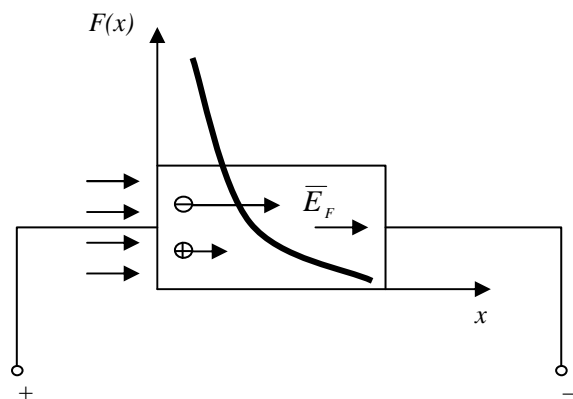


Рис. 14.3. Образование фото-ЭДС Дембера

### 14.1.3. Фотомагнитоэлектрический эффект (ФМЭ) и фотогальванический эффекты

При фотомагнитоэлектрическом эффекте (ФМЭ) потоки электронов и дырок, движущихся от освещенной поверхности образца к неосвещенной, отклоняются магнитным полем в разные стороны.

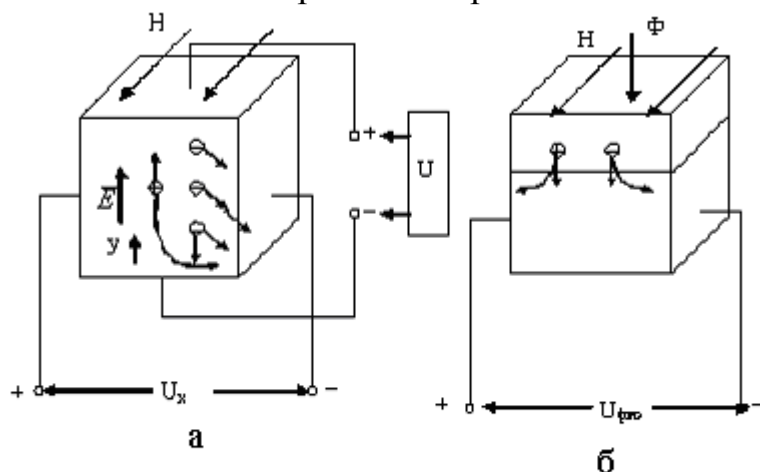


Рис. 14.4. Схема эффекта Холла (а), ФМЭ-ЭДС (б).

ФМЭ эффект возникает в результате действия магнитного поля на диффузионный ток электронов и дырок, потоки электронов и дырок, движущихся от освещенной поверхности образца к неосвещенной, отклоняются магнитным полем в разные стороны. Электрическое поле (фото ЭДС) возникает в направлении, перпендикулярном направлению магнитного поля и направлению диффузионного тока. Возникновение ФМЭ аналогично механизму эффекта Холла (рис. 14.4).

Эффект Холла (рис. 14.4,а) заключается в образовании поперечной разности потенциалов вследствие отклонения в магнитном поле равновесных носителей заряда, движущихся в продольном направлении под действием внешнего электрического поля.

ФМЭ эффект (рис. 14.4,б) есть результат отклонения в противоположных направлениях одинакового количества генерированных светом электронов и дырок, движущихся перпендикулярно магнитному полю за счет градиента концентрации. Величина ФМ ЭДС может достигать больших значений, до нескольких десятков вольт.

ФМЭ эффект используется на практике для создания фотомагнитных приемников (ФМП) излучения на основе узкозонных полупроводников.

#### Фотогальванический эффект

Приборы на основе фотогальванического эффекта (ФГЭ) наиболее широко применяются и поэтому рассмотрим его более подробно.

Световой поток создает градиент концентрации пар неравновесных носителей заряда. Возникает диффузионное перемещение их от освещенной поверхности вглубь системы – к р-п- переходу (см. рис. 14.5). Часть пар носителей рекомбинирует в объеме полупроводника, но большая часть носителей, образовавшаяся

на расстоянии  $L$  с обеих сторон перехода, диффундирует к барьерному слою. Здесь, в области потенциального барьера происходит разделение пар: основные носители, для которых контактное поле – тормозящее, остаются в своем объеме, неосновные носители свободно переходят через p-n-переход, втягиваясь ускоряющим их контактным полем, в зону перехода и образуют фототок ( $I_{\Phi}$ ).

Если  $\Phi$  – падающий световой поток спектрального состава, соответствующего области собственного поглощения  $h\nu \geq \Delta E$ , тогда большая часть светового потока поглощается в наружном слое (из-за большого коэффициента поглощения) p – полупроводника, но некоторая часть проникает глубже во внутреннюю часть системы – n-область. Равновесие токов, текущих через переход, нарушается.

Если система разомкнута, то по мере перехода неосновных носителей через барьер на границах p-n-перехода будут накапливаться объемные заряды, препятствующие дальнейшему переходу носителей: возникает фото-ЭДС. Полярность фото-ЭДС обратна контактной разности потенциалов. Поток неосновных носителей через освещенный переход снижает потенциальный барьер до тех пор, пока возросший ток диффузии основных носителей не скомпенсирует фототок.

Предельная величина фото-ЭДС ( $U_{\Phi}$ ) при больших освещенностях равна контактной разности потенциалов ( $U_K$ ) системы, которая при высокой степени легирования p и n - области близка к  $\Delta E$

$$U_{\Phi} \rightarrow U_K \approx E_g / e, \quad (14.6)$$

здесь  $E_g = \Delta E$  – ширина запрещенной зоны,  $e$  – заряд электрона.

Если p-n-переход замкнут на нагрузку, то при освещении перехода в цепи течет ток, направление которого соответствует потоку неосновных носителей (т.е. совпадает с направлением обратного тока, текущего через переход при подаче внешнего напряжения в запиорном направлении).

На рис. 14.6 показана энергетическая схема освещенного перехода.

Считаем, что число фотонов, поглощаемых в объеме наружного полупроводника в единицу времени, равно  $\Phi/h\nu$ , а количество генерируемых пар носителей определяется величиной  $\eta \frac{\Phi}{h\nu}$ .

В этом выражении  $\eta$  – квантовый выход внутреннего ФЭ:

$$\eta = \frac{\text{количество возбужденных пар носителей}}{\text{кол-во поглощенных квантов излучения}} \leq 1. \quad (14.7)$$

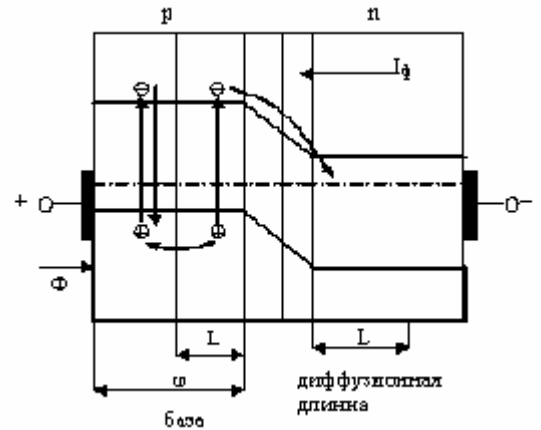


Рис.14.5. Диффузные перемещения неравномерных носителей заряда на p-n-переходе.



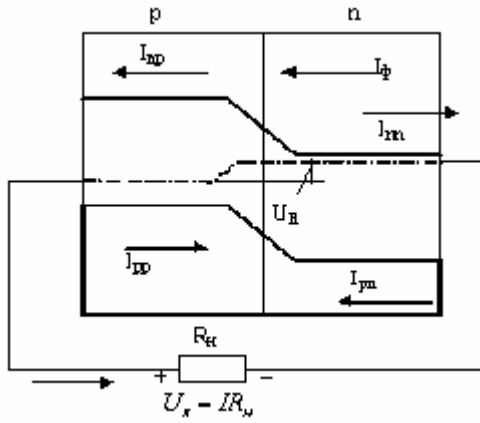


Рис. 14.6. Энергетическая схема освещенного перехода.

Эффективность разделения носителей  $\beta$  определяется как:

$$\beta = \frac{\text{число пар носителей, дошедших до границы } p-n\text{-перехода}}{\text{полное число генерированных пар}}. \quad (14.8)$$

Эффективность разделения носителей зависит от толщины  $p$  – области:  $v = f(\omega)$  и  $v \approx 1$ , если  $\omega < L$ .

Фототок соответствует количеству пар, подошедших к переходу:

$$I_{\Phi} = \frac{1}{h\nu} \eta \beta \Phi = S_{\lambda} \Phi, \quad (14.9)$$

где  $S_{\lambda}$  – коэффициент пропорциональности, характеризует спектральную чувствительность фотозлемента.

Полный ток, протекающий во внешней цепи будет определен следующим выражением:

$$I = I_{\Phi} + I_{pn} + I_{np} - I_{pp} - I_{nn} = \frac{U_R}{R_H}, \quad (14.10)$$

где  $U_R$  – напряжение ( $U_R = U_{\Phi} - U_K$ ), возникающее на сопротивлении нагрузки при освещении ФЭ,  $I_{\Phi}, I_{pn}, I_{np}, I_{pp}, I_{nn}$  – обозначение токов приведено на рис. 14.6.

Полярность этого напряжения соответствует прямому напряжению на переходе, вследствие чего оно вызывает увеличение диффузионных токов основных носителей через переход.

$$I_{nn} = I_{np} \exp\left(\frac{eU_R}{kT}\right), \quad (14.11)$$

$$I_{pp} = I_{pn} \exp\left(\frac{eU_R}{kT}\right). \quad (14.12)$$

Подставив (14.11) и (14.12) в полное значение тока  $I$ , получим

$$I = I_{\Phi} - I_S \left[ \exp\left(\frac{eU_R}{kT}\right) - 1 \right], \quad (14.13)$$

где  $I_S$  – обратный ток через переход.

В режиме *короткого замыкания* (КЗ)  $R_H = 0$ ,  $U_R = 0$  и

$$\exp\left(-\frac{eU_R}{kT}\right) = 1.$$

Ток по внешней цепи (ток КЗ ФЭ) определяются  $I_{КЗ} = I_\Phi = S_\lambda \Phi$ .

Зависимость тока от светового потока – линейна в определенных пределах. Фототок в режиме КЗ не зависит от температуры.

Напряжение на фотогальваническом (ФГ) элементе получается:

$$U_R = \frac{kT}{e} \ln\left(1 + \frac{I_R - I}{I_S}\right) = \frac{kT}{e} \ln\left(1 + \frac{S_\lambda \Phi - I}{I_S}\right). \quad (14.14)$$

В режиме *холостого хода* (ХХ) ток равен нулю ( $I = 0$ ), а напряжение определяется следующим выражением:

$$U_{XX} = \frac{kT}{e} \ln\left(1 - \frac{I_\Phi}{I_S}\right).$$

Напряжение ХХ (фото-ЭДС) фотоэлемента изменяется при увеличении светового потока по логарифмическому закону. Только в области малых световых потоков

$$\frac{I_\Phi}{I_S} \ll 1 \quad \text{и} \quad U_{XX} = \frac{kT}{e} S_\lambda \Phi. \quad (14.15)$$

При больших световых потоках рост фото-ЭДС замедляется, т.к. напряжение холостого хода не может быть больше контактной разности потенциалов, которая не превышает величины, соответствующей ширине запрещенной зоны полупроводника.

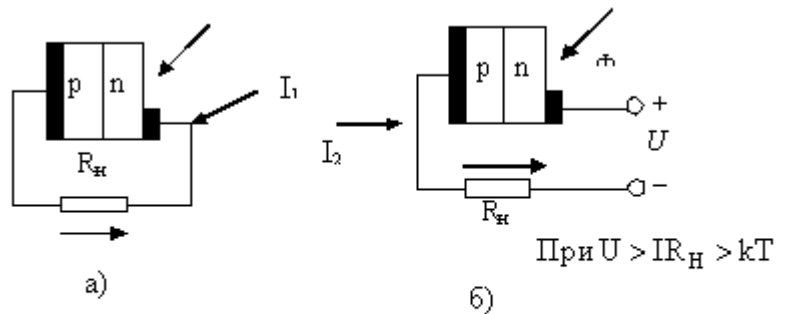


Рис. 14.7 -Схема включения ФЭ с электронно-дырочным переходом

На рис. 14.7 показаны схемы включения ФЭ с электронно-дырочным переходом в фотогальваническом (рис. 14.7,а) и фотодиодном (рис. 14.7,б) режимах.

На рис. 14.7 введены обозначения токов:

$$I_1 = I_\Phi - I_S \exp\left(\frac{e(IR_H)}{kT} - 1\right) \quad \text{и} \quad I_2 = I_\Phi + I_S \quad (14.27) \quad + I_S$$

(14.15)

На рис. 14.13 приведены вольт – амперные характеристики освещенного и неосвещенного p-n-перехода.

Отрезок О-а, отмеченный на рис. 14.8 показывает величину тока КЗ, освещенного р-п-перехода (по направлению этот ток совпадает с обратным током диода). Отрезок О-б – соответствует  $U_{XX}$ , т.е. величине генерируемой им фото-ЭДС. Участок между точками а и б (IV квадрант) представляет собой В-А характеристику диода в ФГ- режиме.

Фотодиодный режим соответствует III квадранту (рис. 14.8). Семейство В-А характеристик ФД эквидистантно.

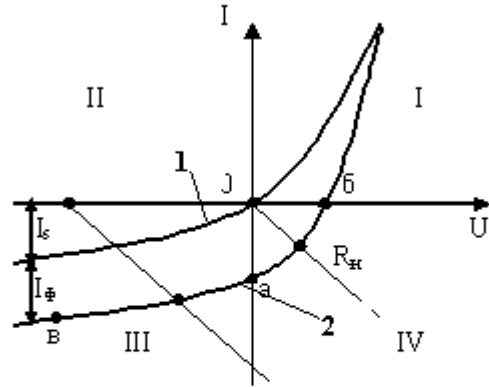


Рис. 14.8 - В-А характеристики неосвещенного и освещенного р-п перехода

Первая характеристика – цифра 1 на рис.

14.8 соответствует темновому току ФД, т.е. обратному току через не освещенный р-п-переход  $I_T = I_s$ . Вторая характеристика (2) - освещенному диоду. Эквидистантность заключается в том, что одному и тому же  $\Delta\Phi$  соответствует одинаковое  $\Delta I_\Phi$ .

## 14.2. Параметры фотоприемников

Важнейшим параметром фотоприемника является *чувствительность*. Она отражает изменения электрического состояния на выходе фотоприемника при подаче на его вход единичного оптического сигнала.

В соответствии с характеризующим параметром различают чувствительности фотоприемника к потоку излучения  $S_{\Phi_e}$ , к световому потоку  $S_{\Phi_v}$ , к облученности  $S_{E_e}$  и к освещенности  $S_{E_v}$ .

В зависимости от измеряемого параметра на выходе фотоприемника различают токовую  $S_I$  и вольтовую чувствительности  $S_U$  фотоприемника. Примеры определения статической чувствительности фотоприемника приведены в выражениях:

$$\begin{aligned} S_{I\Phi_v} &= I_\Phi / \Phi_v; & S_{IE_v} &= I_\Phi / E_v; \\ S_{U\Phi_e} &= U_\Phi / \Phi_e; & S_{UE_e} &= U_\Phi / E_e; \end{aligned} \quad (14.16)$$

где  $S_{I\Phi_v}$  – токовая чувствительность к световому потоку;

$S_{IE_v}$  – токовая чувствительность к освещенности;

$S_{U\Phi_e}$  – вольтовая чувствительность к монохроматическому потоку излучения;

$S_{UE_v}$  – вольтовая чувствительность к потоку облученности.

Дифференциальная чувствительность определяется отношением малых приращений измеряемых величин, например, дифференциальная токовая чувствительность фотоприемника к освещенности:

$$S_{IE_v} = \Delta I_\Phi / \Delta E_v \quad (14.17)$$

Чувствительность зависит от длины волны падающего излучения. Поэтому различают интегральную и монохроматическую (спектральную) чувствительность фотоприемника к немонахроматическому излучению заданного спектрального состава.

**Квантовый выход** внутреннего фотоэффекта  $\eta_1$  определяется числом неравновесных носителей (пар), созданных каждым поглощенным фотоном.

**Скорость генерации** неравновесных носителей  $G$  – определяется числом неравновесных носителей, возникающих в  $1\text{см}^3$  материала за 1 секунду при его облучении и определяется:

$$G_{(x)} = \eta_1 \frac{\alpha \cdot \Phi_1(x)}{h\nu}. \quad (14.18)$$

В области собственного поглощения ( $\eta_1=1$ ) при  $\Phi_1=\text{const}$  скорость генерации уменьшается с ростом частоты. Фототок и чувствительность пропорциональны  $G$  и имеют примерно ту же зависимость от частоты- рис. 14.9,а.

**Инерционность фотоприемников.** Инерционность зависит о различных причин и может быть охарактеризована *постоянными времени нарастания и спада фототока*. Если фототок растет по закону

$$I_{\Phi} = I_{\text{max}} \cdot [1 - \exp(-t/\tau_1)],$$

а спадает по закону  $I_{\Phi} = I_{\text{max}} \cdot \exp(-t/\tau_2)$ , то  $\tau_1$  – постоянная времени нарастания, а  $\tau_2$  – постоянная времени спада (рис. 14.9,б).

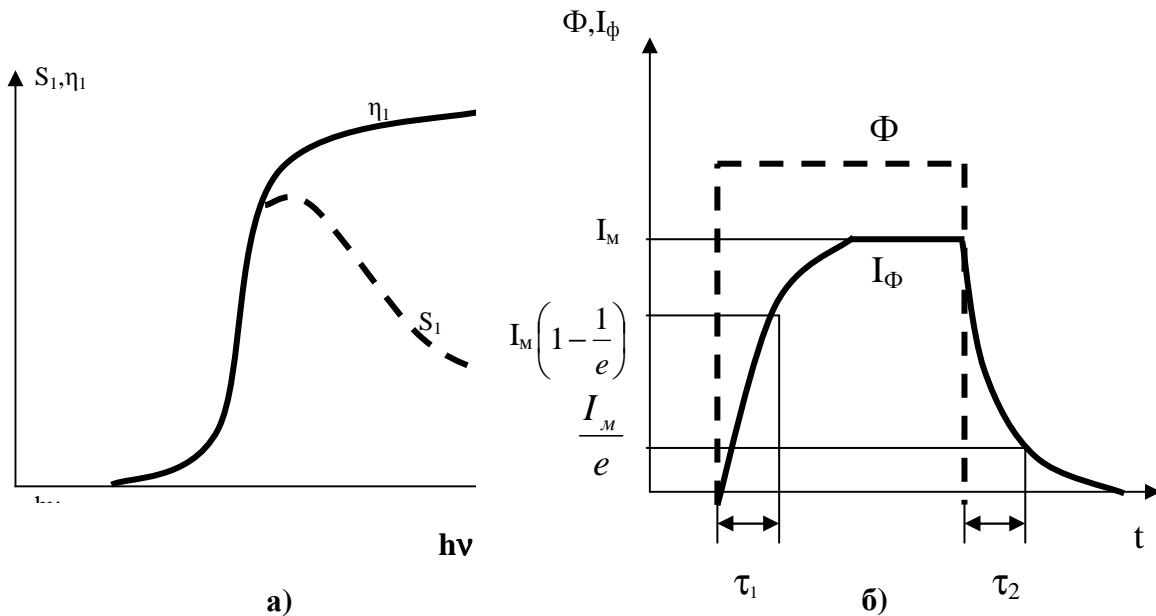


Рис. 14.9. а) зависимость  $S_1$  и  $\eta_1$  от энергии фотонов; б) – изменение фототока  $I_{\Phi}$  со временем.

У одних приемников фототок полностью определяется величиной  $G(x)$  и объемом материала, в котором происходит поглощение света, у других – величиной  $G(x) \cdot Z(E)$ , где  $Z(E)$  – коэффициент усиления, зависящий от напряженности поля  $E$ .

Токи, идущие через фотоприемник в темноте и при освещении, испытывают **шумовые** (случайные) отклонения  $\Delta I$  от среднего значения тока  $\bar{I}$ . Так как отклонения имеют разные знаки и их среднее значение равно  $\overline{\Delta I} = 0$ , то мерой случайных отклонений является среднее значение квадрата отклонения  $\overline{\Delta I^2}$ .

Вследствие хаотичности теплового движения свободных носителей (электронов) их концентрация в различных участках полупроводника изменяется со временем. Связанные с этим колебания возникающей ЭДС и тока пропорциональны температуре  $T$  и определяют среднеквадратичную составляющую тока, обусловленную **тепловыми шумами**:

$$\overline{\Delta I^2} = \frac{1}{R} kT \Delta f, \quad (14.19)$$

где  $R$  – сопротивление образца,  $\Delta f$  – полоса частот.

**Дробовой шум** вызван колебанием числа электронов, проходящих через прибор при постоянном внешнем напряжении (Среднеквадратичная составляющая тока, обусловленная током сигнала):

$$\overline{\Delta I^2} = 2eI \Delta f. \quad (14.20)$$

*Шум такого типа определяется случайным характером процессов генерации, рекомбинации и диффузии электронов и дырок.*

**Пороговая чувствительность.** Это уровень светового потока  $\Phi_n$ , когда сигнал равен шуму, т. е.  $\overline{I_\phi^2} = \overline{\Delta I^2}$ . Т. к.  $\sqrt{\overline{\Delta I^2}}$  и  $\Phi_n$  могут зависеть от площади  $S$  приемника и полосы  $\Delta f$ , то  $\Phi_n^* = \frac{\Phi_n}{\sqrt{S \cdot \Delta f}}$ , (14.21)

где  $\Phi^*$  – приведенный пороговый поток.

### 14.3. Типы фотоприемников

#### 14.3.1. Полупроводниковые фотоприборы

**Фоторезисторы.** При освещении однородного полупроводника его электропроводность увеличивается. Это явление называется *фотопроводимостью*, а соответствующий прибор – фоторезистором. Схема наблюдения фотопроводимости показана на рис. 14.10,а.

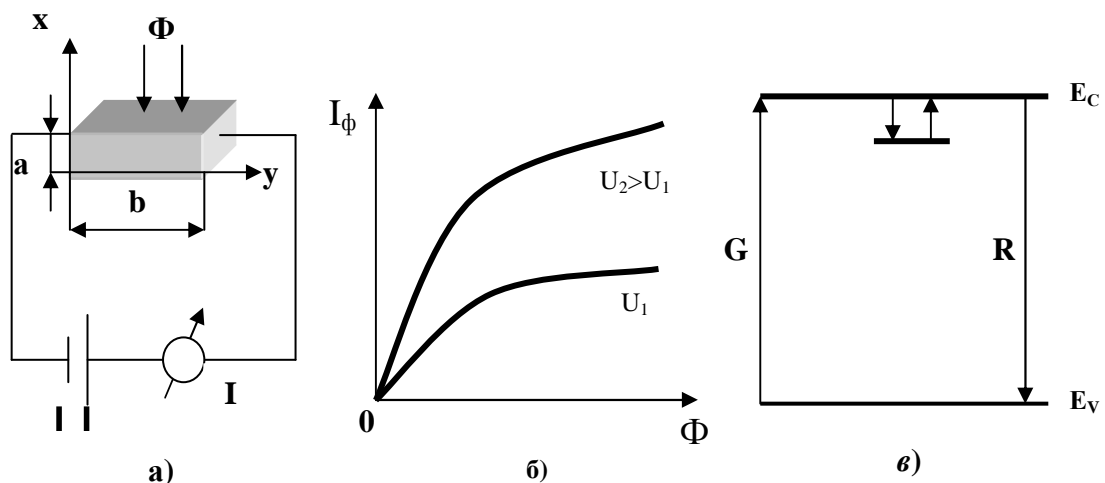


Рис. 14.10.-Фотопроводимость: а-схема наблюдения; б-фототок при раз-

Если под действием света возникают только электроны в зоне проводимости, то добавочный ток (фототок):

$$I_{\phi} = e n_{\phi} V_d \cdot S, \quad (14.22)$$

$$V_d = \mu E, \quad (14.23)$$

где  $V_d$  – дрейфовая скорость,  $e$  – заряд электрона,  $n_{\phi}$  – концентрация неравновесных (избыточных) фотоэлектронов,  $\mu$  – подвижность электронов,  $E$  – напряженность поля,  $S$  – сечение образца.

Если свет создает электронно-дырочные пары, то возникает и дырочная составляющая фототока. В большинстве случаев достаточно рассматривать фототок, связанный с носителями одного типа.

Концентрация фотоэлектронов определяется скоростью генерации и временем жизни в зоне проводимости:

$$n_{\phi} = G\tau. \quad (14.24)$$

При этом у освещенной поверхности ( $x = 0$ ) скорость генерации будет определяться как

$$G(0) = \eta_1 \frac{\alpha \Phi_1(0)}{h\nu}, \quad (14.25)$$

$\Phi_1(0)$  – плотность падающего потока. Полагая, что генерация идет равномерно по объему в слое полупроводника толщиной  $x^*$  (в этом случае поглощается  $2/3$  фотонов), получим:

$$I_{\phi} = q\eta_1 \alpha \frac{\Phi_1(0)}{h\nu} \tau V_d \alpha \cdot x^*. \quad (14.26)$$

Так как  $\alpha = \frac{1}{x^*}$ , произведение  $\tau V_d$  имеет смысл расстояния  $y$ , которое проходит электрон за время  $\tau$ , фототок

$$I_{\phi} = q\eta_1 \frac{\Phi y}{h\nu b}, \quad (14.27)$$

где  $y = \tau V_d = \tau \mu E$ . Фототок растет с увеличением  $\Phi$  и напряженности поля  $E$  (или напряжения  $U = bE$ ).

Экспериментальные зависимости  $I_{\phi}(\Phi)$  показывают ослабление зависимости между  $I_{\phi}$  и  $\Phi$  при больших световых потоках (см. рис. 14.10,б). Это связано с тем, что при больших  $\Phi$  и соответствующих  $n_{\phi}$  время жизни  $\tau$  электронов уменьшается, например, из-за роста при освещении концентрации носителей противоположного знака (в данном случае дырок).

В области собственного поглощения с увеличением частоты может наблюдаться уменьшение  $I_{\phi}$ , связанное с падением  $\tau$  в тонких слоях полупроводника у поверхности вследствие более быстрой рекомбинации носителей через локальные энергетические состояния.

Инерционность фоторезисторов определяется часто не столько временем жизни  $\tau$ , сколько участием ловушек (рис. 14.10,в). После попадания электронов в полосу проводимости они могут захватываться пустыми уровнями примесей или дефектов решетки, причем время нахождения электронов на этих уровнях растет с ростом их глубины и с понижением температуры. Только после теплового освобождения электронов они могут рекомбинировать с дырками (переход R). Про-

цессы захвата носителей ловушками («прилипание» электронов и дырок) затягивает рост фототока после включения света и спад после его выключения.

В результате фоторезисторы обладают большей инерционностью, чем другие ФП. Как это обстоятельство, так и температурная зависимость параметров резисторов ограничивает их использование.

Фоторезисторы могут иметь в качестве чувствительного элемента монокристалл полупроводника, пленку полупроводника на диэлектрике или таблетку прессованного порошкообразного материала. Последний способ применяют, например, при изготовлении промышленных фоторезисторов из сульфида или селенида кадмия. Фоторезисторы CdSe имеют темновое сопротивление  $\sim 10^6$  Ом, максимум спектральной характеристики в области  $\lambda_{\text{макс}} = 0,7$  мкм, работают при  $U = 20$  В и имеют постоянные времена роста и спада фототока  $\tau_1 = 40$  мс и  $\tau_2 = 20$  мс. Их интегральная чувствительность при использовании общего излучения теплового источника света с цветовой температурой 2840 К составляет  $S_1 = 600 \frac{\text{mA}}{\text{лм}}$  (при освещенности 200 лк).

Фоторезисторы на основе сульфида свинца являются менее инерционными ( $\tau = 0,1$  мс) и обладают высокой обнаружительной способностью  $D^* (2,4 \text{ мкм}, 780 \text{ Гц}, 1 \text{ Гц}) = 1,5 \cdot 10^{11} \text{ см} \cdot \text{Гц}^{1/2} \cdot \text{Вт}^{-1}$  при 295 К. Понижение температуры приводит к росту  $D^*$ .

**Фотодиоды (ФД) с p-n-переходами.** В ФД с p-n-переходом разделение генерированных оптическим излучением неравновесных пар носителей осуществляется на p-n-переходе. При этом возможны два режима работы ФД.

1. Гальванический режим (рис. 14.11,а) – на фотодиод не подается внешнее питание, а нагрузкой является высокоомный индикатор – регистрируется фото-ЭДС; нагрузкой служит низкоомный индикатор – регистрируется фототок. В этом случае фотоэлектроны не накапливаются в n-области, а так как напряжение  $U_{\text{ф}} = 0$ , то фототок, проходящий через переход ( $I_{\text{ф}}$ ) равен:

$$I_{\text{ф}} = -I, \quad I_{\text{ф}} \sim \Phi_0, \quad (14.28)$$

В таких условиях диод работает как фотоэлемент. Например, солнечные батареи на кремнии Si ( $\Delta E = 1,1$  эВ). Здесь все фотоны солнечного излучения способны создавать электронно-дырочные пары. Фото-ЭДС таких ФЭ составляет несколько десятых долей вольта.

Поэтому их часто соединяют последовательно для получения  $U$  в несколько вольт.

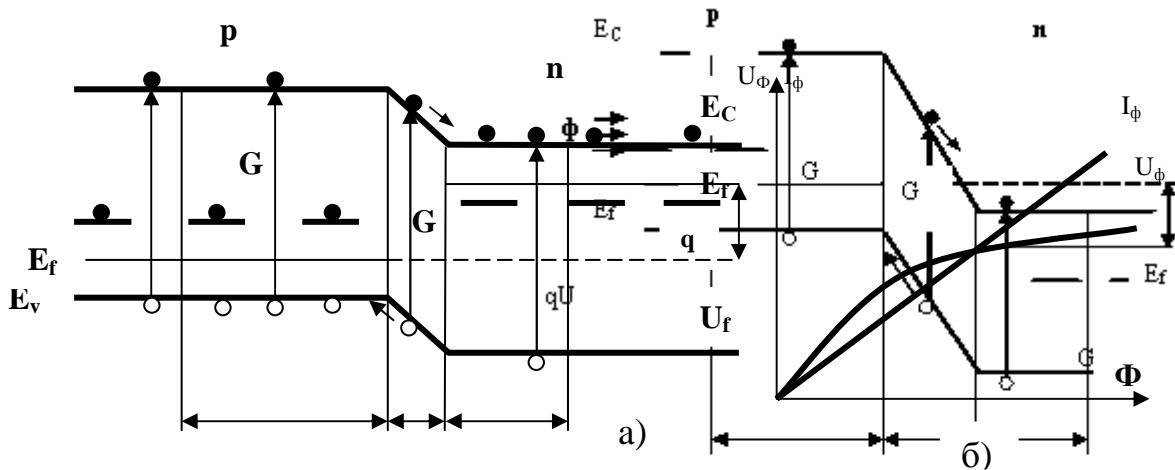


Рис. 14.11. а)- р-п-переход при освещении, но в отсутствии внешнего напряжения ( $U$ ); б) зависимость  $I_\Phi$  и  $U_\Phi$  от  $\Phi$ .

На рис. 14.11,б изображены световые характеристики для фотогальванического и фотодиодного режимов.

2. Фотодиодный режим работы – на ФД подается обратное напряжение, при этом обратный ток диода в широких пределах линейно зависит от падающего светового потока, нагрузка диода высокоомная. Этот режим применяется при использовании ФД в качестве входного элемента электронных устройств (операционные усилители и усилители постоянного тока).

Таким образом, если р-п-переход включен в запирающем направлении, то практически все напряжение будет падать на обедненной области, шириной  $d$  и схема энергетических зон приобретет вид, показанный на рис. 14.12.

Теперь разделение электронов и дырок, возникающих при освещении, производится более сильным полем, а ширина области высокого сопротивления увеличивается ( $d \sim \sqrt{U}$ ). Это увеличивает ту часть кристалла, из которой электроны и дырки быстро выводятся полем.

Пример: Если дрейфовая скорость электронов  $V_d = 1 \cdot 10^6 \frac{\text{см}}{\text{с}}$ ,  $d = 10^{-4} \text{ см}$ ,

то время пролета электронами области поля  $\tau_i = 1 \cdot 10^{-10} \text{ с}$ . С такой же быстротой эта составляющая фототока будет следовать за изменением интенсивности света. Однако через переход проходят и электроны (дырки), рожденные светом в слоях  $l_1$  и  $l_2$  и достигшие перехода путем диффузии. Соответствующие времена гораздо больше ( $\tau_d \approx 1 \cdot 10^{-7} \text{ с}$ ) и именно они могут определять инерционность ФД. Для ослабления этой составляющей фототока р-п-переход следует формировать у самой освещенной поверхности (уменьшая толщину  $x_1$  слоя р-типа).

Так как большая часть света поглощается на глубине  $x^* = 1/\alpha$  ( $\alpha$  – коэффициент поглощения) то при  $x_1 \approx 0$  и  $d = x^*$  роль света, поглощенного в слое  $l_2$ , и приходящих отсюда дырок также ослаблена (рис. 14.12). Графики вольт-амперных зависимостей р-п-перехода при различных световых потоках приведены на рис. 14.13.

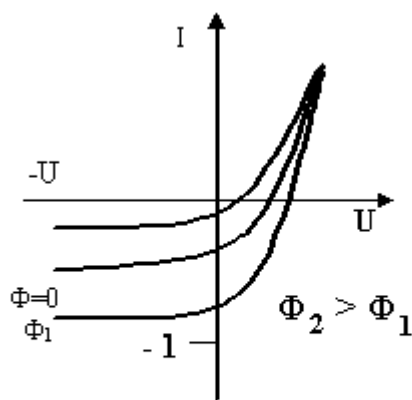


Рис. 14.13. ВА характеристики ФД при различных световых потоках



При включении перехода в запирающем направлении ( $U < 0$ ) и при  $qU \gg kT$  ток будет определяться выражением

$$I = -(I_n + I_\phi). \quad (14.29)$$

При увеличении  $\Phi$ , фототок  $I_\phi$  увеличивается и может значительно превысить темновой ток ( $I_\phi > I_n$ ).

Если весь световой поток  $\Phi$ , падающий на ФД, поглощается в слое  $x_1 + d + l_2$  (т. е. эта толщина больше  $2x^*$ ), то все возникшие электроны и дырки примут участие в создании фототока и

$$I_\phi = e\eta_1\Phi/h\nu, \quad (14.30)$$

где  $\eta_1$  - квантовый выход процесса генерации электронно-дырочных пар.

Поскольку ФД работают при сравнительно высоких обратных напряжениях, то с целью повышения напряжения пробоя их изготавливают из монокристаллических материалов. Толщина базы фотодиода (т.е. глубина залегания р-п-перехода) делается обычно значительно меньше диффузионной длины неосновных носителей заряда, поэтому почти все генерируемые светом носители разделяются полем перехода, не успев рекомбинировать. Постоянная времени ФД определяется временем диффузии электронов через базу. И при толщине базы несколько мкм может составлять  $\approx 10^{-8}$  с.

Пути улучшения частотных свойств ФД:

1. Уменьшение толщины базы до 1-2 мкм.
2. Так как коэффициент диффузии электронов  $D_n > D_p$ , выгодно в качестве базы использовать р-область (в которой электроны – неосновные носители).
3. Уменьшение емкости р-п-перехода и конструктивной емкости фотодиода.

Кроме германия (Ge) и кремния (Si) для изготовления ФД используются и другие материалы (GaAs; CdS, CdTe, AlSb, InP, InSb, InAs и др.)

Физические явления и ВАХ диода с р-п-переходом подробно были рассмотрены в разделе 14.1.1. Чтобы полностью исключить поглощение света правее перехода, можно увеличить  $d$  до  $2x^*$ . Например, в полупроводниках типа р-і-п.

**Р-і-п-фотодиод.** Уменьшение толщины базы ограничено трудностями. Улучшение частотных характеристик ФД может быть получено при использовании гетеропереходов или барьеров Шоттки.

В гетеропереходах, освещаемых со стороны широкозонного полупроводника, поглощение излучения происходит не в базе, а в области, лежащей ниже перехода, и в самом переходе. База служит лишь своеобразным светофильтром и защитным слоем, толщина её может быть сделана достаточно большой без ухудшения частотных характеристик ФД. Конструкции р-і-п-фотодиодов изображены на рис. 14.14, а и б.

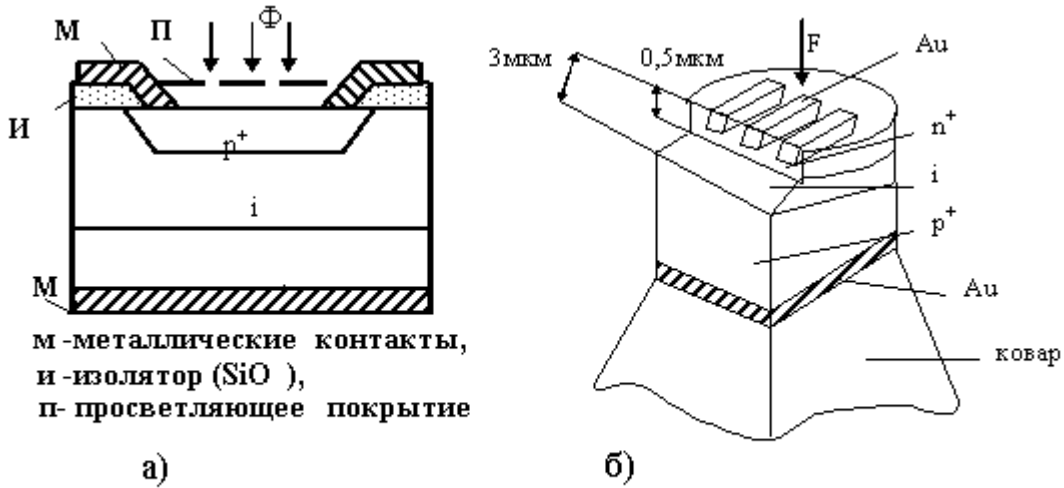


Рис. 14.14. Конструкции p-i-n- диода

В кремниевом (Si) p-i-n-фотодиоде, который является наиболее распространенным ФД толщина i-области - 50 мкм,  $p \pm 3$  мкм. При освещении такого диода светом  $\lambda = 0,9$  мкм (от GaAs излучателя)  $x^* - 30$  мкм и около 80% света поглощается в i-слое. Следовательно, и быстродействие диода определяется временем  $\tau_i$  - их пролета через i-слой. Дрейфовая скорость электронов в кремнии сначала растет с увеличением напряжения поля, а затем испытывает насыщение при  $U_d \approx 5 \cdot 10^6$  см/с. В этих условиях время пролета равно:

$$\tau_i = \frac{50 \cdot 10^{-4} \text{ см}}{5 \cdot 10^6 \text{ см/с}} = 10^{-9} \text{ с},$$

а время  $\tau_d$ , определяемое диффузией электронов из p-области или дырок из n-области ( $\tau_d > \tau_i$ ), не играет существенной роли.

В общем случае следует учитывать еще одну составляющую  $\tau_{rc}$  постоянной времени, связанной с сопротивлением R и емкостью C цепи. При малом сопротивлении нейтральных областей диода, а также внешней цепи, при широком переходе (зарядовая емкость p-n-перехода  $C \sim d^{-1}$ , а  $d \sim \sqrt{U}$ ) имеем  $\tau_{rc} < \tau_i$ . Обнаружительная способность кремниевых фотодиодов достигает значения  $1 \cdot 10^{13} \text{ см} \cdot \text{Гц}^{1/2} \cdot \text{Вт}^{-1}$  при  $\lambda = 1$  мкм,  $T = 300\text{К}$ .

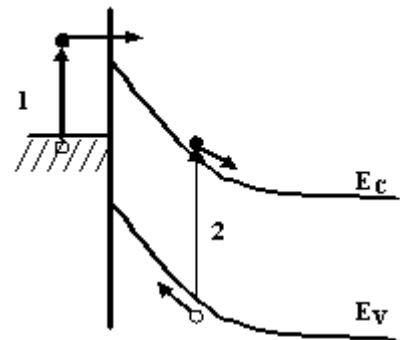


Рис. 14.15 Фотозффект в поверхностном барьере включенном в запирающем направлении

**ФД с поверхностными барьерами.** ФП с поверхностным барьером Шоттки (рис. 14.15) также обладает высоким быстродействием и эффективностью. Подобные барьеры, образующиеся на контакте металл – полупроводник, могут быть получены на материалах, на которых невозможно создать p-n-переход. Если электронный полупроводник контактирует с металлом, у которого работа выхода меньше работы выхода для полупроводника, то определенное число электронов переходит из полупроводника в металл. Ионизированная донорная

примесь в полупроводнике образует слой положительного пространственного заряда, обладающий высоким сопротивлением. При включении диода в запирающем направлении (минус на металле) ширина барьера увеличивается в соответствии с формулой

$$d = \left[ \frac{2\varepsilon\varepsilon_0(U_k + U)}{eN_d} \right]^{1/2}, \quad (14.31)$$

где  $\varepsilon$  – диэлектрическая проницаемость;  $\varepsilon_0$  – диэлектрическая постоянная;  $U_k$  – контактная разность потенциалов;  $U$  – внешнее напряжение;  $N_d$  – концентрация доноров.

Если  $N_d = 1 \cdot 10^{17} \text{ см}^{-3}$ , то при напряжении в несколько вольт  $d \approx 1 \cdot 10^{-5} \text{ см}$ . Тонкий слой металла толщиной  $1 \cdot 10^{-6} \text{ см}$  может быть нанесен на полупроводник методом вакуумного распыления. Свет направляют на кристалл сквозь эту почти прозрачную пленку (в случае диода Au-Si слой Au пропускает 95% излучения с  $\lambda = 0,63 \text{ мкм}$ ). Если  $d > x^*$ , то основная часть света с энергией фотонов  $h\nu > \Delta E$  поглощается в области, в которой присутствует сильное электрическое поле (переход 2 рис. 14.15). Возникшие пары быстро разделяются полем, и время пролета  $\tau_i$  может быть сделано очень малым ( $\tau_i = 10^{-10} - 10^{-11} \text{ с}$ ), особенно в том случае, когда  $x^*$  и толщина  $d < 1 \text{ мкм}$ .

По мере увеличения  $f x^*$  уменьшается, и область поглощения света сдвигается к металлическому слою. Тем не менее, вследствие близости барьера к поверхности это поглощение остается в пределах поля барьера и диффузионная составляющая тока отсутствует. Чувствительность диодов Шоттки, как и p-i-n-переходов является высокой и достигает  $S_i = 0,5 \text{ А/Вт}$ . Особенностью барьеров Шоттки является возможность регистрации фотонов с  $h\nu > \Delta E$ . При энергии фотонов, большей чем высота барьера, со стороны Me, электроны могут вводиться в полупроводник из металла (переход 1 на рис. 14.15). Соответственно расширяется область спектральной чувствительности ФД.

Чтобы предупредить сильное отражение света от поверхности диода, на нее наносят просветляющее покрытие, например, пленку сульфида цинка, с  $n = 2,3$  для  $\lambda = 0,63 \text{ мкм}$ . Толщину пленки устанавливают такой, чтобы при интерференции лучей, отраженных от границ пленки, получился минимум, т. е. чтобы отражение отсутствовало для лучей данной  $\lambda$ .

**Лавинные фотодиоды (ЛФД).** Если к ФД приложить высокое обратное смещение, то  $E$  в барьере оказывается настолько высоким ( $5 \cdot 10^5 - 1 \cdot 10^6 \text{ В/см}$ ), что развиваются процессы ударной ионизации атомов решетки ускоренными электронами. Это приводит к росту обратного тока при данном световом потоке  $\Phi$  (рис. 14.16). Как темновые, так и фотоэлектроны, приобретают в высоком поле энергию, большую, чем они теряют при столкновениях с атомами решетки. Если полученная электроном энергия превышает энергию ионизации  $E_i$  (обычно  $\Delta E < E_i < 1,5\Delta E$ ), то электрон может создать новую электронно-дырочную пару. При достаточно протяженной области поля возникшие электрон и дырка тоже могут ускориться до энергии  $E_i$  и совершить новые ионизации, т. е. будет наблюдаться лавинное нарастание числа носителей заряда.

Увеличение тока, вследствие ионизации при больших  $E$ , характеризуется коэффициентом умножения носителей  $M = I/I_0$ , где  $I$  – ток при больших напря-

жениях  $U$ , а  $I_0$  – при малых. Величина  $M$  отражает общее число ионизаций, совершенных как первичными, так и вторичными электронами и дырками. Она определяется отношением числа электронов, вошедших в область поля ( $n_0$ ), к числу электронов, вышедших из него ( $n$ ), т.е.

$$M = n/n_0. \quad (14.32)$$

Квантовый выход ионизации зависит от коэффициента ударной ионизации (число пар, созданных электроном на 1 см пути) для электронов ( $\alpha_i$ ) для дырок ( $\beta_i$ ), а также ширины  $d$  области поля. При  $\alpha_i = \beta_i$  и однородном поле шириной  $d$  имеем  $N = \alpha_i d$ , в случае неоднородного поля  $N = \int_0^d \alpha_i dx$ . Так как  $\alpha_i \sim \exp(-c_1 / E^2)$ , а для

барьера Шоттки  $E \sim \sqrt{U}$  ( $U$  – напряжение на барьере), то  $N$  может быть представлено в виде  $N = a_1 \exp(-v_1/U)$ , где  $a_1$ ,  $v_1$ ,  $c_1$  – постоянные, значение  $N$  изменяется от 0 при малых  $U$  до 1 при напряжении пробоя  $U_B$ , когда ток через диод резко возрастает ( $M \rightarrow \infty$ ). На практике часто используется степенное представление зависимости  $N(U)$ , а именно  $N = (U/U_0)^m$ , где  $m = 2 \div 4$  для разных материалов.

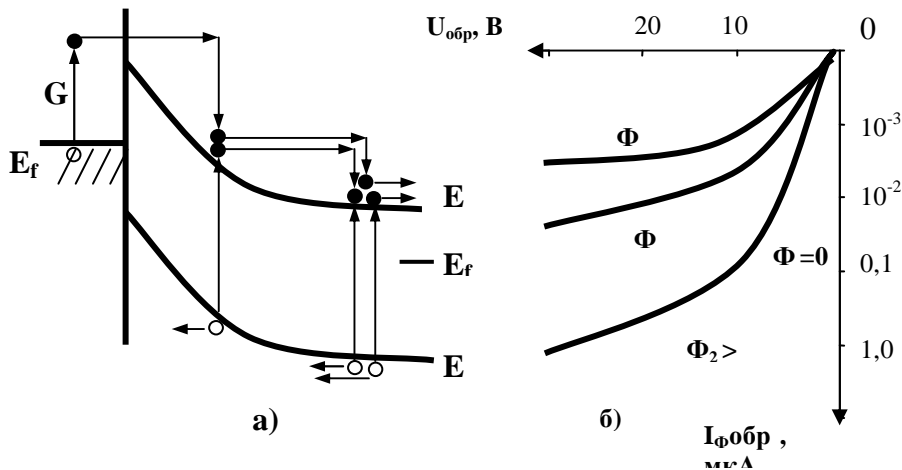


Рис. 14.16. Процесс лавинного размножения носителей заряда в диоде Шоттки (а) и В – А характеристика лавинного фотодиода (б).

Лавинное умножение фотоносителей получено как в поверхностных барьерах (рис. 14.21), так и p-n- переходах. При больших  $U$  сильные токи приводят к разогреву полупроводника, что увеличивает темновой ток и уменьшает фототок. Так как коэффициент умножения

$M \sim 1 \cdot 10^3$ , ЛФД пригоден для регистрации слабых световых сигналов. В то же время из-за сильной зависимости  $M(U)$  использование ЛФД затруднено необходимостью применения высокостабильного напряжения. Инерционность ЛФД составляет около  $1 \cdot 10^{-9}$  с.

**Специальные типы фотодиодов. Фото диоды с продольным фотоэффектом.** На характеристики и параметры ФД кроме материала p-n-перехода существенное влияние оказывает место падения светового потока относительно области p-n-перехода. Рассмотрим некоторые типы фотодиодов. Если на одной из

областей р-п-перехода сделать два изолированных электрода А и В, то при локальном освещении поверхности оптического окна ФД узким световым лучом, смещенным относительно центра окна, между этими электродами генерируется продольная фото-ЭДС, в отличие от обычной поперечной фото-ЭДС, возникающей между р-п-областями при несмещенном потоке (рис. 14.17).

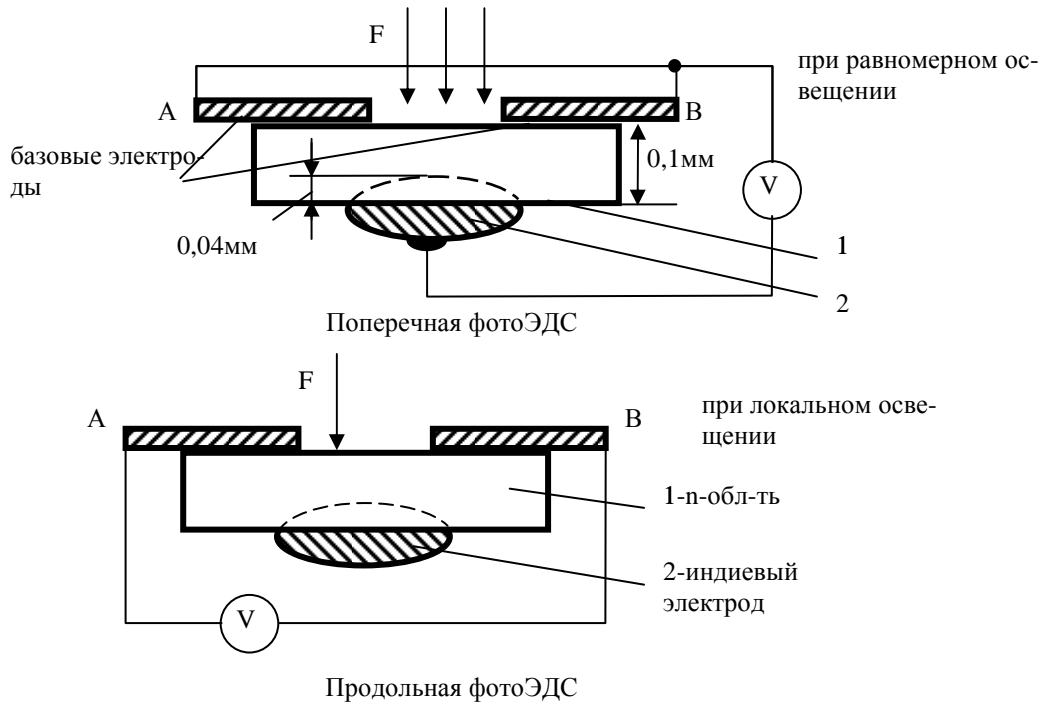


Рис. 14.17. Конструкция германиевого координатно-чувствительного ФЭ

Продольный ФЭ используется для создания фотоприемников чувствительных к месту падения светового луча, т.е. координатно-чувствительных ФЭ, применяющихся для регистрации координат источника излучения (в качестве следящих датчиков).

Материалом для них является германий, антимонид и арсенид индия.

Механизм образования продольной фото-ЭДС (рис. 14.18). Р<sup>+</sup>-область в Ge

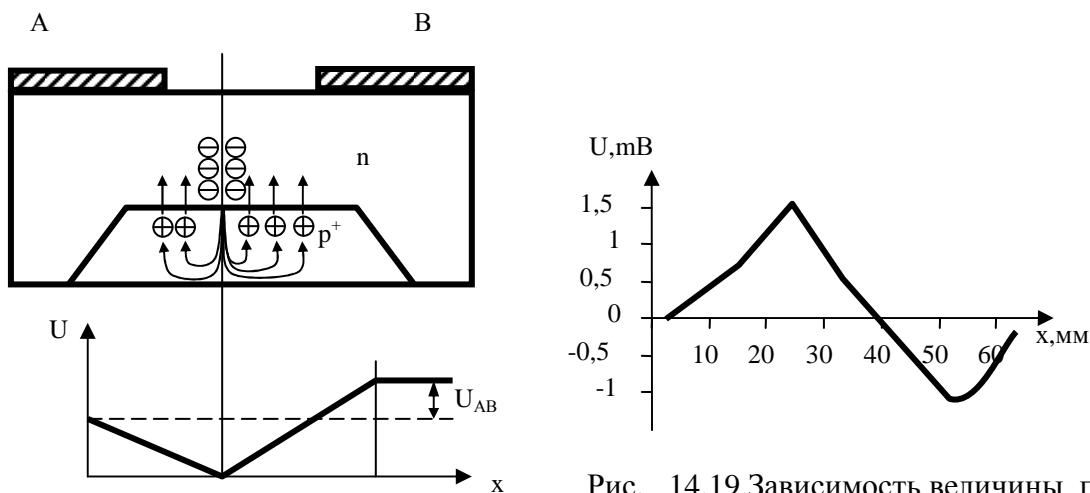


Рис. 14.18. Механизм образования продольной фото-ЭДС.

Рис. 14.19. Зависимость величины продольной фото-ЭДС от положения светового пятна

с высокой концентрацией акцепторной примеси обладает значительно большей электропроводностью, чем базовая область n германия.  $P^+$  - область можно считать эквипотенциальной. Дырки, втягиваемые в  $p^+$ -область через освещенный участок p-n-перехода сразу же растекаются по  $p^+$ -слою, не образуя в нем объемного заряда. Поперечная фото-ЭДС, эквивалентная прямому смещению на p-n-переходе, обуславливает обратный переход избыточных дырок в n-область.

При протекании обратного тока через переход в n-слое Ge, обладающим высоким продольным сопротивлением, возникает продольное поле, которое вызывает движение электронов от освещенного участка к местам обратной инжекции дырок для нейтрализации создаваемого ими объемного заряда.

Непременным условием образования продольной фото-ЭДС является несимметричность p-n-перехода. От положения светового пятна зависит величина продольной фото-ЭДС (рис. 14.19). Одним из видов несимметричного электронно-дырочного перехода является барьер Шоттки. Известны координатно-чувствительные ФЭ с барьером Шоттки, например, со структурой кремний-золото, где гетеропереходы принципиально несимметричны. Примером продольного ФЭ с гетеропереходом может служить тонкопленочный фотоэлемент n-CdS – p-Cu<sub>2</sub>O, в котором слой Cu<sub>2</sub>O – эквипотенциален. Чувствительность координатных ФЭ достигает  $\approx 300 \frac{\text{мВ}}{\text{мм} \cdot \text{мВт}}$  (для Si – Au). Применение их обеспечивает точность определения направления светового луча до долей угловой секунды.

**Фотодиоды с магнитным эффектом (ФМЭ).** Фотомагнитный эффект используется на практике для создания фотомагнитных приемников (ФМП) излучения. ФМП на основе узкозонных полупроводников используются, главным образом, в качестве приемников инфракрасного излучения. Рассмотрим ФМП излучения на антимониде индия InSb (Рис. 14.20). Рабочая длина волны этого ФМП находится в пределах от  $\lambda = 6,2 \text{ мкм}$  до  $7,5 \text{ мкм}$ . На рис. 14.20 введены обозначения:

1 – антимонид индия p-типа,  $p = 5 \cdot 10^{16} \text{ см}^{-3}$  удельное сопротивление материала  $\gamma \approx 0,03 \text{ Ом} \cdot \text{см}$ , 2 – магниты, t – толщина пластинки, длина образца между электродами l – 50 мкм,  $S \approx 1 \text{ мм}^2$ .

Поверхности пластинки обрабатывают так, чтобы рекомбинация со стороны падения света была меньше, чем на противоположной стороне.

Если электроды ФМП замкнуты на небольшое внешнее сопротивление, то при освещении наружной поверхности течет ток КЗ

$$j_{\text{фмэ}} = \frac{eN_q \varepsilon L}{1 + \frac{\phi S}{L}} (m_n + m_p) \frac{H}{C} \quad (14.33)$$

$N_q$  – поток квантов излучения,  $\varepsilon$  – квантовый выход внутреннего фотоэффекта, L – диффузионная длина неосновных носителей заряда,

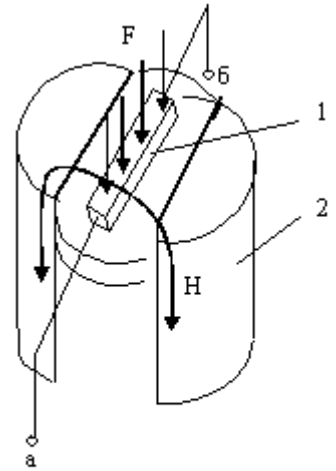


Рис.14.20. Фотомагнитный приемник с фотомагнитным эффектом.

$S$  – скорость поверхностной рекомбинации носителей на освещаемой поверхности,  $\tau$  – время жизни носителей,  $\mu_n$  и  $\mu_p$  – подвижности электронов и дырок,  $H$  – напряженность магнитного поля.

При разомкнутой внешней цепи величина фотоЭДС (напряжение между

точками а и б) (рис.14.20) равно:

$$U_{xx} = \frac{J_{фмэ}}{\sigma l},$$

где  $\sigma$  – удельная электрическая проводимость полупроводника,  $l$  – длина образца между электродами.

При  $m_n = 80000 \frac{\text{см}^2}{\text{В} \cdot \text{с}}$  и  $\tau < 0,2$  мкс обес-

печивается высокое быстродействие. Зависимость фототока КЗ фотомангнитного ФМП от величины индукции магнитного поля показана на рис. 14.21.

Чувствительность, определяемая как отношение падение напряжения на наружном сопротивлении к единице мощности излучения  $\left(\frac{U}{P}\right)$ ,

составляет величину порядка 1 – 3 В/Вт. Обнаружительная способность:

$$D^* (300^\circ \text{К}, 4000 \text{Гц}, 1 \text{Гц}) = 3 \cdot 10^8 \text{ см} \cdot \text{Гц}^{1/2} \text{Вт}^{-1}.$$

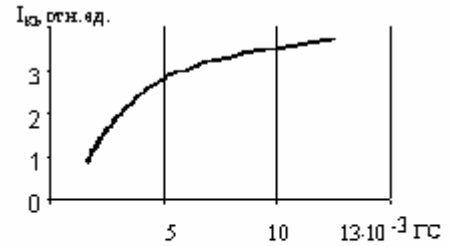


Рис.14.21. Зависимость фототока от напряженности магнитного поля.

**Фототранзисторы и фототиристоры.** Фототранзистор (рис. 14.22) -

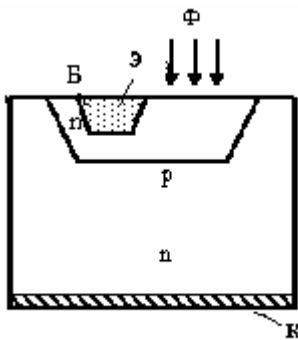


Рис.14.22. Фототранзистор

представляет собой структуру фототранзистора из чередующихся слоев р-п-р или п-р-п, как правило, в планарном (плоском) исполнении. Освещение ведется со стороны базы и эмиттера через оптическое окно. Внешнее напряжение смещает коллекторный переход в обратном направлении, а переход эмиттер - база может иметь либо обратное, либо небольшое прямое смещение в зависимости от выбранного режима работы.

Если оба перехода смещены в обратном направлении, то в отсутствие освещения ток коллектора  $I_K$  будет равен обратному току перехода коллектор – база, т.е.  $I_{КТ} = I_{КБО}$ . При этом учтем, что обратное смещение перехода эмиттер – база (Э-Б) невелико, равно 0 или не превышает  $U_T$ . Направим на переход эмиттер – база световой поток  $\Phi$ , в обеих областях перехода будет происходить генерация пар носителей заряда. Через переход эмиттер – база увеличится ток экстракции неосновных носителей. Это приведет к снижению потенциального барьера и к росту тока базы. В свою очередь ток коллектора увеличится  $I_K = \beta \cdot I_B + (\beta + 1)I_{КБО}$ , где  $\beta$  - коэффициент передачи тока базы. При дальнейшем увеличении светового потока переход Э-Б окажется смещенным в прямом направлении. Токи инжекции неосновных носителей превысят токи экстракции неосновных носителей, и ток коллектора возрастет еще больше. Эти рассуждения справедливы и для случая, когда в темновом режиме переход Э – Б имеет небольшое прямое смещение.

В темноте практически все внешнее напряжение падает на коллекторном переходе. Освещение средней части (базы) ведется через тонкий слой эмиттера. Возникающие в базе и в обоих переходах фотоэлектроны попадают в области эмиттера и коллектора, а дырки собираются в средней области. В результате к левому р-п-переходу оказывается приложенным дополнительное напряжение в прямом направлении и возникает инжекция темновых электронов через сниженный барьер в базу и далее – в коллекторный переход (рис. 14.23,а).

Таким способом первоначальный ток фотоносителей может быть усилен  $\sim 1 \cdot 10^2$  раз. Соответственно чувствительность фототранзистора значительно выше, чем у обыкновенного ФД. С другой стороны, участие процессов диффузии носителей заряда увеличивает инерционность и постоянная времени возрастает  $\tau = 10^{-5} - 10^{-6}$  с.

Сужение базы, необходимое для уменьшения  $\tau$  приводит к уменьшению чувствительности фототранзистора вследствие уменьшения числа поглощенных фотонов.

Рис. 14.22 и 14.23,а и б иллюстрируют работу фототранзистора. Вольт – амперные характеристики фототранзистора приведены на рис. 14.23,б.

**Фототиристор** имеет чередующиеся слои р, п, р, п – типов проводимости и соответственно 3 р-п-перехода, из которых средний называется коллекторным, а два крайних эмиттерными. Структура включается так, чтобы коллекторный переход был включен в обратном направлении, а 2 эмиттерных – в прямом (+ на внешней р-области, а – на п-области).

Если напряжение на всем тиристоре увеличить до  $U_n$ , при котором эмиттерные переходы заметно понизятся (или при  $U < U_n$ , но с помощью управляющего электрода одной из баз эмиттерный переход включается в прямом направлении), то через тиристор начинает течь значительный ток, который приводит к накоплению в р-базе положительного заряда, а в п-базе отрицательного. Это снижает высоту р-п-переходов и вызывает резкий рост тока. При этом общее падение  $U$  на тиристоре снижается, т. к. токи сами теперь поддерживают нужную степень накопления зарядов. Т. о. тиристор может находиться в состоянии, соответствующем большому или малому токам, т. е. тиристор может работать как ключ в электрической цепи.

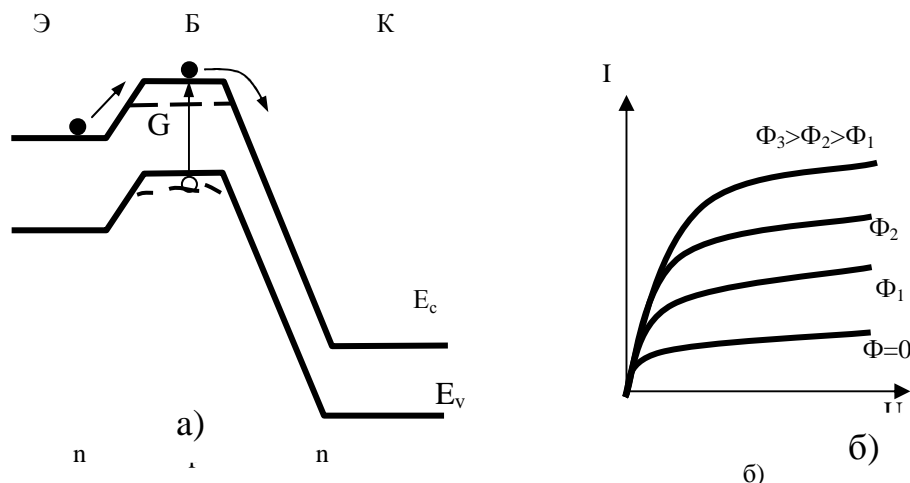


Рис. 14.23. Фототранзистор а) – энергетическая диаграмма фото-транзистора; б) – В А характеристики.





На рис. 14.25 изображена цепочка МДП-структуры, являющейся частью многоэлементного фоточувствительного экрана.

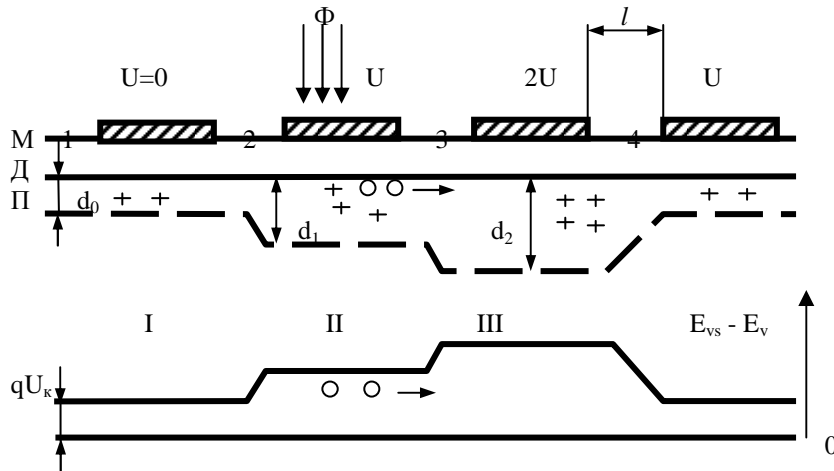


Рис. 14.25. Серия взаимодействующих элементов электрической схемы МДП-структуры, к которым приложены различные напряжения. Штриховая линия – граница области пространственного заряда.  $E_{vs} - E_v$  – глубины потенциальных ям и дырок.

В случае I на структуру не подано внешнее напряжение. Наличие обедненного слоя толщиной  $d_0$  связано с контактной разностью потенциалов  $U_k$ .

Случай II соответствует случаю, когда присутствует внешнее напряжение и структура освещается через полупрозрачный металлический электрод (затвор). Накопленный за время освещения заряд дырок может быть сдвинут вдоль цепочки структур, если на соседний элемент подано более высокое напряжение.

Случай III. В этом случае глубина потенциальной ямы для дырок больше, и они вследствие диффузии перейдут от второго элемента к третьему. В промежутке между затворами 2 и 3 присутствует ускоряющее дырки поле, которое проникает и в область скопления дырок, поэтому переход дырок к электроду 3 может быть быстрым.

Электрический сигнал от элемента 2, пропорциональный интенсивности  $\Phi$  и времени действия света, падавшего на этот элемент, может быть проведен по цепочке элементов и выведен во внешнюю цепь. Аналогичным образом могут быть получены видеосигналы от других элементов строки и элементов других строк. Таким образом, микроскопические ФЭ расположены в строке настолько близко ( $l \approx 3 \text{ мкм}$ ), что они могут взаимодействовать, передовая друг другу заряды. Отсюда *приборы с зарядовой* (или объемной) связью. Эта связь между МДП-структурами осуществляется через общую для всех структур толщу полупроводников.

Регистрирующие изображение устройства рассмотренного типа (из МДП-элементов) потребляют малую энергию во время считывания, требуют небольших  $U$  (10 – 20 В) и при числе элементов  $\sim 1 \cdot 10^5 \text{ см}^{-2}$  могут иметь хорошую разрешающую способность (140 линий на 1 мм).

Высокая чувствительность экрана определяется тем, что он работает в условиях накопления светового действия. Так как в чувствительных элементах используются поверхностные барьеры, то облегчается подбор области спектральной чувствительности прибора, путем выбора материала с  $\Delta E \leq h\nu$ . Кремниевые ПЗС имеют срок службы  $1 \cdot 10^4$  ч и регистрируют изображения телескопов.

### 14.3.2. Фотоэлектронные приборы СВЧ

Среди приборов классической электроники СВЧ особую группу составляют приборы, в которых термоэмиссионный катод заменен на фотокатод. Такие приборы служат для извлечения информации из модулированного светового потока и используются в качестве детекторов и смесителей в оптических системах связи и локации, а также применяются для исследования спектра излучения оптических квантовых генераторов.

В фотодетекторе, состоящем из фоточувствительного катода (нелинейный элемент) и СВЧ колебательной системы, первоначально происходит преобразование светового потока в фототок, т. е. модуляция потока фотонов переходит в модуляцию электронного потока, а затем переменная составляющая тока преобразуется в выходной сигнал. В ряде фотоприемников применяют дополнительное умножение фототока.

Выражение для выходной мощности фотоэмиссионных приборов СВЧ удобно записать в виде

$$P = \frac{1}{2} |m \cdot M \cdot I_0|^2 |f(\omega)|^2 R_{\text{ЭКВ}}, \quad (14.34)$$

где  $I_0$  — средняя составляющая тока на фотокатод;  $m = i/I_0$  — параметр, характеризующий глубину модуляции;  $M$  — коэффициент умножения тока;  $f(\omega) \leq 1$  — функция, определяющая частотную полосу пропускания фотодетектора;  $R_{\text{ЭКВ}}$  — эквивалентное сопротивление, величина которого характеризует связь переменной составляющей фототока с внешней нагрузкой. Эквивалентное сопротивление  $R_{\text{ЭКВ}}$  является основным параметром любого фотодетектора.

Важной характеристикой фотоэмиссионных детекторов является также квантовая эффективность  $\eta_0$  катода, т. е. число свободных электронов, образуемых одним фотоном на фотокатод. Однако останавливаться здесь на рассмотрении квантовой эффективности оптических детекторов не будем, так как эти приборы интересуют нас в первую очередь как СВЧ электронные приборы. Отметим лишь, что разработанные в настоящее время фоточувствительные слои имеют максимальную квантовую эффективность в синей, фиолетовой и ультрафиолетовой областях спектра, в то время как большинство существующих оптических квантовых генераторов работают на частотах красной и инфракрасной областей спектра. Следовательно, в настоящее время использование оптических квантовых генераторов, работающих в диапазоне  $50 — 5\,000 \text{ А}^\circ$ , и создание фоточувствительных слоев с более высокой «красной границей» являются одним из основных направлений повышения эффективности фотодетекторов, работа которых основана на внешнем фотоэффекте.

Большинство фотоэмиссионных приборов СВЧ могут быть использованы как в режиме демодуляции, так и в режиме оптического гетеродинамирования. Если на вход фотодетектора поступает световой поток, промодулированный по ин-

тенсивности на частоте  $\omega$ , то на фотокатоде образуется фототок, средняя величина которого  $I_0 = \phi \eta_0 e$  пропорциональна квантовой эффективности  $\eta_0$  ( $\phi$  — световой поток в фотонах в секунду,  $e$  — заряд электрона).

В режиме гетеродинирования, когда на фотокатод поступают оптический сигнал с несущей  $\omega_c$  и сигнал от местного генератора на частоте  $\omega_r$ , на поверхности фотокатода возникают биения двух когерентных сигналов, приводя, в частности, к появлению разностной частоты  $\omega_c - \omega_r$ . Амплитуда переменной составляющей фототока в этом случае на частоте  $\omega_c - \omega_r$  равна  $2\sqrt{I_c I_r}$ , где  $I_c$  и  $I_r$  — средние значения фототоков, возникших соответственно под действием внешнего сигнала и светового потока от местного гетеродина.

При оптическом гетеродинировании возможно получение более высоких значений постоянной составляющей фототока и тем самым достижение более высокого уровня выходной мощности фотодетектора. Оптическое гетеродинирование позволяет увеличивать отношение сигнал/шум, делая фотоприемник более чувствительным, и повышает его разрешающую способность.

Весь большой класс существующих в настоящее время фотоэмиссионных приемников по аналогии с обычными СВЧ приборами можно разбить на две основные группы: приборы резонансного типа (узкополосные) и нерезонансного типа (широкополосные). Приборы типа фотоэлектронных умножителей принято выделять в отдельную группу.

### Фотоэмиссионные приборы СВЧ резонансного типа

Типичным представителем группы приборов резонансного типа является фотоклистрон, схема подобного прибора представлена на рис. 14.26. К основным элементам фотоклистрона относятся фотокатод, который при воздействии на него модулированного по интенсивности светового сигнала создает модулированный по плотности электронный поток, и проходной резонатор, настроенный на частоту модуляции. Фотоклистрон явился исторически первым фотоэлектронным СВЧ прибором, с помощью которого было исследовано фотосмещение двух оптических сигналов.



Рис. 14.26. Фотоклистрон.

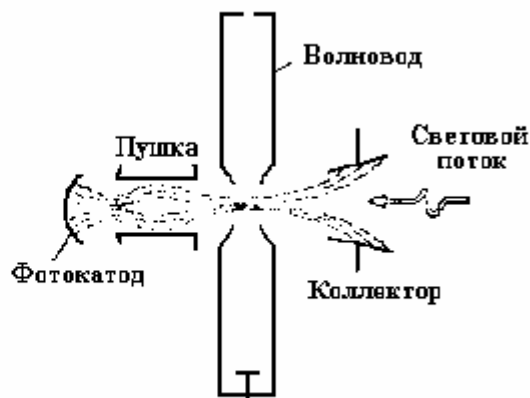


Рис. 14.27. СВЧ-фотодиод.

Расчёт высокочастотных процессов в фотоклистроне сводится к ранее рассмотренному анализу возбуждения резонатора заданным током. При достаточно высокой добротности резонатора эквивалентное сопротивление  $R_{экв}$  фотоклистро-

на может достигать порядка нескольких сотен Ом. Это позволяет создать на базе фотоклистрона сравнительно высокочувствительные приемники оптических сигналов. Основной недостаток таких фотоприемников, ограничивающий их применение в оптических системах связи — узкополосность. Рабочая полоса частот может быть расширена только за счет уменьшения  $R_{\text{экв}}$ , что, в конечном счете, ведет к резкому ухудшению чувствительности приемника.

Схема прибора резонансного типа, названного СВЧ фотодиодом, показана на рис. 14.27. Катод, представляющий собой участок сферической поверхности, покрытой фоточувствительным слоем, является частью электронной пушки, фокусирующей электронный поток. Область взаимодействия потока электронов с электромагнитным полем образована срезанными вершинами двух полых конусов, которые являются элементами волновода. Эквивалентное сопротивление такого прибора достигает порядка 200 Ом. СВЧ фотодиод был использован при изучении процесса фотосмещения аксиальных типов колебаний импульсных квантовых генераторов. В результате квадратичного детектирования биений между аксиальными типами колебаний ток фотоэмиссии был промодулирован по плотности СВЧ сигналом. В волноводе, возбуждаемом этим током, наблюдалось СВЧ излучение на частотах 8,2 и 10,28 ГГц.

В качестве узкополосного фотоприемника может быть использована ЛОВ, в которой термоэлектронный катод заменен на фотокатод. Этот прибор привлекателен тем, что в нем может быть достигнут сравнительно высокий уровень выходной мощности.

### Фотоэлектронные умножители

Хорошо известными приборами для демодуляции амплитудно-модулированных световых сигналов являются электростатические фотоэлектронные умножители (ФЭУ), которые состоят из фотокатода и нескольких каскадов электронного умножения — динодов с вторичной электронной эмиссией. Диноды умножительной секции ФЭУ могут быть как отражательного типа, так и проходного. Фотоэлектронные умножители характеризуются большим умножением фототока ( $10^5$ — $10^7$  раз), что обуславливает их высокую чувствительность. К сожалению, применение электростатических ФЭУ ограничено частотами модуляции порядка нескольких сотен мегагерц. Это связано с тем, что на высоких частотах эквивалентное сопротивление  $R_{\text{ЭКВ}}$  электростатического ФЭУ сильно уменьшается из-за влияния времени пролета электронов между динодами. В фотоэлектронных умножителях с проходными динодами, выполненными в виде тонких пленок, эффекты времени пролета сказываются слабее, чем в ФЭУ с отражательными диодами.

К детектирующим устройствам типа фотоумножителей, способным работать в широкой полосе частот (вплоть до СВЧ диапазона), относятся динамические фотоэлектронные умножители со скрещенными полями. Устройство динамического ФЭУ со скрещенными полями показано на

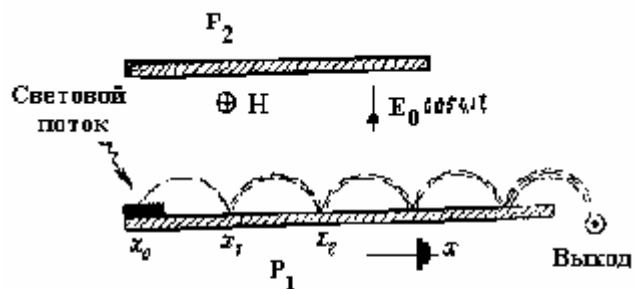


Рис.14.28. Устройство динамического ФЭУ со скрещенными полями.

Устройство динамического ФЭУ со скрещенными полями показано на

рис. 14.28. Между электродами  $P_1$  и  $P_2$  действуют два пространственно-однородных поля: высокочастотное электрическое поле  $E_0 \cos(\omega t)$  и постоянное магнитное поле  $H$ . Электрод  $P_1$  имеет покрытие с большим коэффициентом вторичной эмиссии.

Чтобы пояснить принцип работы этого прибора, рассмотрим движение отдельного электрона, который в начальный момент времени находится на плоскости  $P_1$  в точке  $x_0$  (см. рис. 14.28.). Пусть полярность электрического поля, приложенного между электродами  $P_1$  и  $P_2$ , в начальный момент времени такова, что электрон начинает ускоренно двигаться в направлении плоскости  $P_2$ .

Под действием постоянного магнитного поля  $H$  траектория электрона будет искривляться. Если магнитное поле  $H$  направлено так, как показано на рис. 14.28, то электрон закручивается по направлению движения часовой стрелки. Величину напряженности магнитного поля можно подобрать такой, что поворот электрона под действием магнитного поля к плоскости  $P_1$  совпадет с изменением направления высокочастотного электрического поля. Тогда электрон будет увеличивать свою кинетическую энергию за счет энергии поля  $E_0$ . Падение электрона на плоскость  $P_1$  в точке  $x_1$  приводит к появлению вторичных электронов. К этому моменту времени электрическое поле вновь меняет свою полярность и процесс повторится, но уже с большим количеством электронов.

Основная особенность рассмотренной умножительной системы заключается в наличии фазовой группировки электронов, если частоты СВЧ сигнала и электрического поля  $E_0 \cos(\omega t)$  лежат в одном диапазоне. Из анализа уравнений движения электронов в такой системе следует, что электроны, покидающие электрод  $P_1$  в самых различных фазах, группируются до возвращения на этот электрод. С каждым циклом умножения фазовая группировка электронов улучшается.

Известны эксперименты, в которых динамический ФЭУ со окрещенными полями использовались для детектирования светового потока, промодулированного СВЧ сигналом в диапазоне 3 ГГц. Исследования показали, что можно получить коэффициент усиления по току порядка  $10^5$  и  $R_{экв} \sim 100$  Ом. Метод динамического умножения обеспечивает чувствительность и отношение сигнал/шум как и у обычных электростатических ФЭУ, но на более высоких частотах.

### *Широкополосные фотоэмиссионные приборы СВЧ*

Так как световой когерентный луч может нести огромный объем информации, то вполне естественно, что особое внимание уделяется широкополосным фотодетекторам. Наиболее простой по конструкции широкополосный фотоприемник — вакуумный фотоэлемент

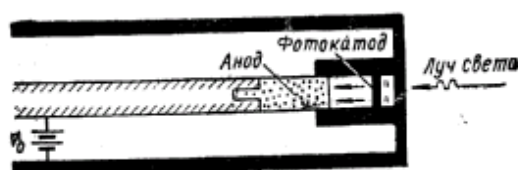


Рис. 14.29. Плоский фотодиод, помещенный в волновод.

Он представляет собой миниатюрный плоский фотодиод, помещенный в волновод или выполненный в виде центрального проводника коаксиальной линии (рис. 14.29). Вакуумный фотоэлемент имеет очень низкую чувствительность. Это объясняется отсутствием умножения фототока и уменьшением эквивалентного сопротивления на высоких частотах из-за влияния времени пролета. За счет повышения напряжения на диоде можно уменьшить время пролета электронов. Однако даже при повышении напряжения втрое против рекомендуемого  $R_{\text{экв}}$  вакуумного фотоэлемента не превышает 6 Ом для 10-см диапазона и 0,5 Ом для 3-см диапазона.

Для широкополосной демодуляции оптических сигналов с частотами модуляции в миллиметровом диапазоне предложен фотодиод с распределенной эмиссией. Принципиальная схема этого прибора приведена на рис. 14.30. На фотокатод под действием светового луча возбуждается «бегущая волна» фототока. Параметры прибора подбираются таким образом, чтобы скорость бегущей волны фототока совпала со скоростью волны в линии передачи; в результате этого волна тока отдает энергию в линию передачи. Фотодиод с распределенной эмиссией предполагается использовать вплоть до частот 100 ГГц; при этом на верхних частотах можно ожидать значений  $R_{\text{экв}} \sim 100 \text{ Ом}$ .

Наиболее совершенным широкополосным фотоэмиссионным детектором в настоящее время является лампа бегущей волны с фотокатодом (фото-ЛБВ и ФЭУ-ЛБВ). Эквивалентное сопротивление этого прибора может достигать величины  $10^6 \text{ Ом}$  при ширине диапазона рабочих частот порядка октавы. Фото-ЛБВ имеет низкий уровень собственных шумов, высокую чувствительность и в случае специальной конструкции может быть применена для демодуляции ЧМ светового потока (взаимодействие с поперечными волнами электронного пучка). Схема фото-ЛБВ приведена на рис. 14.31. Если на фотокатод падает световой луч, модулированный по интенсивности на частоте  $\omega$ , то, благодаря квадратичной характеристике



Рис.14.30. Принципиальная схема фотодиода с распределенной эмиссией.

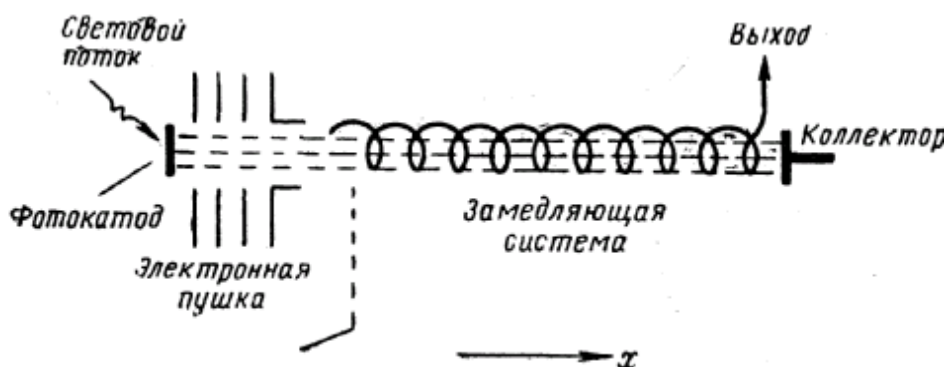


Рис. 14.31. Схема фото-ЛБВ.

фотокатода, в пучке электронов, покидающих его, содержится гармоническая составляющая тока на частоте  $\omega$ . В замедляющей системе в результате взаимодействия модулированного потока электронов с полем возбуждаемой им бегущей электромагнитной волны происходит усиление сигнала.

## ГЛАВА 15.

### СРЕДСТВА ПЕРЕДАЧИ ОПТИЧЕСКОГО ИЗЛУЧЕНИЯ

#### 15.1. Диэлектрические световоды

В интегральной оптике основой любого устройства являются планарный диэлектрический волновод и различные типы трехмерных канальных и полосковых волноводов, рассмотренных в [42-43]. Трехмерные волноводы в отличие от планарных обеспечивают дополнительное ограничение оптического излучения в поперечном направлении. Ограничение светового потока в поперечном направлении позволяет снизить управляющие напряжения волноводных модуляторов, рабочие мощности инжекционных лазеров, входящих в структуру интегрально-оптических схем. Рассмотрим диэлектрический световод (параметры и геометрические размеры приведены на рис. 15.1). Для облегчения задачи будем рассматривать случай симметричного световода (обе внешние среды одинаковы).

##### 15.1.1. Распространение волн в световодах

Для того, чтобы в световоде распространялись волны, необходимо выполнение условия  $\epsilon_2 < \epsilon_1$ , т.е.  $n_1 < n_2$ . Волновой вектор плоской волны, введенной внутрь световода будет определяться соотношением:

$$k_1 = \gamma \text{ (по оси Oz), и } k_1 = \beta \text{ (по оси Ox) } \quad |k_1| = kn_1,$$

где  $\gamma$  - продольная постоянная распространения,  $\alpha$  и  $\beta$  - поперечные постоянные распространения в разных средах.

Волна преломляется на границе сред, и в среде 2 её волновой вектор будет равным  $k_2 = \gamma$  (по оси Oz), и  $k_2 = -j\alpha$  (по оси Ox)  $|k_2| = kn_2$ .

Проекция этих векторов на ось Oz одинаковы, поскольку вдоль всей плоскости раздела сред сохраняется равенство фаз.

Но нам необходимо полное внутреннее отражение, так как только в этом случае волна будет отражаться от границ раздела и оставаться внутри световода.

При отражении волны от поверхности оптически менее плотной среды наблюдается предельный угол падения  $i_c$ , который соответствует полному внутреннему отражению. Причем  $\sin i_c = n_2/n_1 = \eta$ , что выполняется при условии  $\theta < \theta_c$ . Составляющая вектора  $k_2$ , параллельная оси Ox запишется в виде  $kn_2 \cos r =$

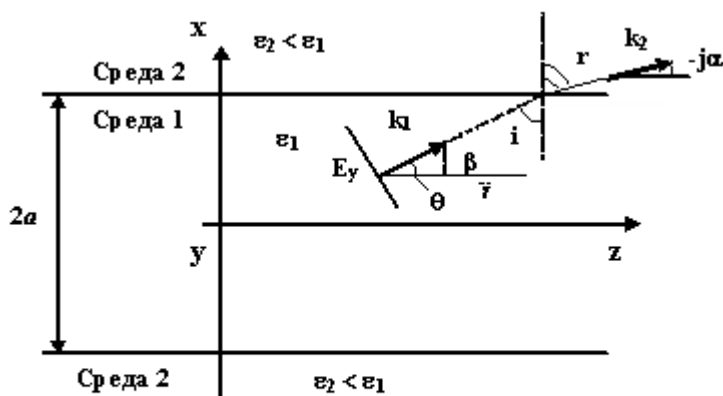


Рис. 15.1. Параметры и геометрические размеры диэлектрического световода



-  $j\alpha$  (углы  $\theta$ ,  $r$ ,  $i$  показаны на рис. 15.1) где поперечная постоянная распространения  $\alpha$  определяется следующим выражением:

$$\alpha = kn_2 \sqrt{\frac{\cos^2 \theta}{\eta^2} - 1}, \quad (15.1)$$

а напряженность электрического поля во второй среде будет равна:

$$\begin{aligned} E_y^2 &= E_0 e^{j(\omega t - \gamma z)} e^{-\alpha x}, \\ \gamma &= kn_2 \sin r = kn_1 \cos \theta, \end{aligned} \quad (15.2)$$

Амплитуда неоднородной волны во второй среде, как сверху, так и снизу от первой среды, экспоненциально уменьшается при удалении в поперечном направлении от границ световода. Внутри световода происходит переотражение также, как в металлическом волноводе. Суперпозиция падающей и отраженной волн даст нам бегущую волну вдоль оси Oz и стоячую волну вдоль оси Ox. При этом:

$$E_y^{(1)} = E_0 e^{j(\omega t - \gamma z)} \begin{cases} \cos \beta x - \text{«четная» волна} \\ \sin \beta x - \text{«нечетная» волна.} \end{cases}$$

Полная напряженность электрического поля в диэлектрическом световоде на критической частоте при  $r=1$  будет равна:

$$E_{\text{полн}}(n\omega_c) \approx e^{j(\omega t - \gamma z)} 2j \sin \beta X \approx \sin \beta(x + a). \quad (15.3)$$

Так как  $a\beta = n(\pi/2)$ , имеем:

$$\text{Если } n=2p, \text{ то } E(n\omega_c) \approx \sin\left(p \frac{\pi}{a} x\right).$$

$$\text{Если } n=2p+1, \text{ то } E(n\omega_c) \approx \cos\left(\frac{\pi}{2} + pn\right) \frac{x}{a}.$$

Эти выражения полностью идентичны формулам, относящимся к металлическому волноводу.

**Уравнения дисперсии и условие согласования фаз.** Известно, что в общем случае процесс распространения электромагнитных волн в среде является дисперсным, т.е. скорость распространения зависит от частоты. Дисперсия складывается из двух различных явлений: 1) дисперсия, обусловленная средой (показатель преломления зависит от частоты); 2) в оптическом волноводе имеется дисперсия, обусловленная неоднородностью среды. Существует также дисперсия мод. Рассмотрим это явление подробнее.

На основании определения векторов  $k_1$  и  $k_2$  можно сразу написать:

$$\begin{aligned} k_1^2 &= k^2 n_1^2 = \gamma^2 + \beta^2, \\ k_2^2 &= k^2 n_2^2 = \gamma^2 - \alpha^2. \end{aligned} \quad (15.4)$$

Отсюда, исключив  $\gamma$ , получаем

$$\alpha^2 + \beta^2 = k^2 (n_1^2 - n_2^2). \quad (15.5)$$

Напишем теперь условие согласования фаз. Отраженная волна приобретает фазовый сдвиг  $\Phi$ , возникающий при отражении от каждой границы раздела сред как снизу, так и сверху и зависящий от угла падения

$$\Phi = \pm 2 \operatorname{arctg} \left[ \frac{\sqrt{\sin^2 i - \eta}}{\cos i} \right]. \quad (15.6)$$

Для нахождения дисперсионного уравнения, приравняем полный сдвиг фазы, при прохождении волны от нижней стенки до верхней и обратно, числу кратному  $2\pi$ :

$$4kn_1a \sin \theta + 2\Phi = 2\pi n. \quad (15.7)$$

Определим чему равен фазовый сдвиг. Для этого воспользуемся (15.6) и (15.1) и формулой для постоянной распространения  $\beta$ :

$$\beta = kn_1 \sin \theta. \quad (15.8)$$

Найдем

$$\Phi = -2 \operatorname{arctg} \left( \frac{\alpha}{\beta} \right). \quad (15.9)$$

В действительности отражение происходит не от границы волновода, а несколько глубже, как бы от поверхности, расположенной на некотором расстоянии  $x_0$  далее границы раздела двух сред. Таким образом, световые пучки смещаются при отражении от границы раздела двух сред.

**Моды диэлектрического волновода.** Нахождение зависимости  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\theta$  от

частоты приводит к трансцендентным уравнениям. Из уравнения дисперсии (15.7) и уравнений (15.8) и (15.9), получаем два варианта:

$$\begin{aligned} \text{если } n = 2p, \text{ то } \beta \operatorname{tg} \beta a &= \alpha, \\ \text{если } n = 2p+1, \text{ то } -\beta \operatorname{ctg} \beta a &= \alpha. \end{aligned} \quad (15.10)$$

Рассмотрим один вариант  $n = 2p$ . На рис. 15.2 показано графическое построение зависимостей  $\alpha$  и  $\beta$  от частоты. Так как  $\alpha$  и  $\beta$  должны одновременно удовлетворять уравнениям (15.5) и (15.10), то при каждом новом

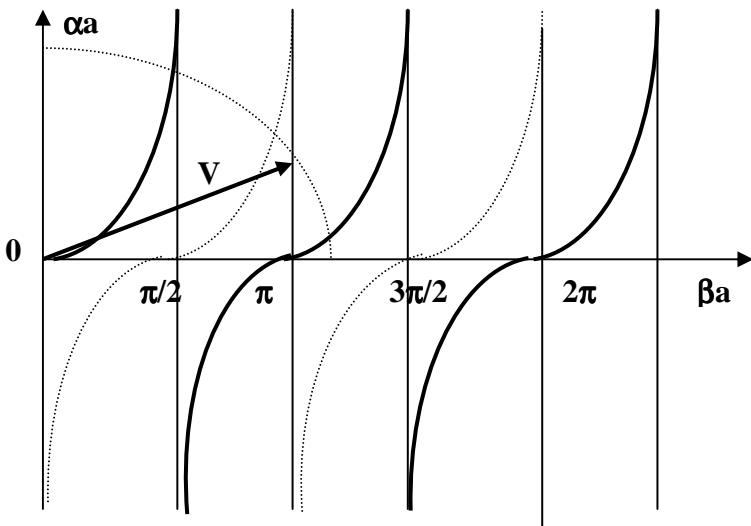


Рис. 15.2. Изменение мод в зависимости от приведенной частоты

значении частоты они определяются точкой пересечения построенной кривой (15.10), с дугой окружности, которая описывается уравнением (15.5). При увеличении частоты радиус такой окружности увеличивается и значения  $\alpha$  и  $\beta$  также увеличиваются. Диаграмма, представленная на рис. 15.2 показывает, как появляются разные моды распространения. Значения  $n = 2p$  соответствуют нечетным

модам, а значения  $n = 2p+1$  – четным модам. Для удобства часто вводят приведенную частоту

$$V = ka\sqrt{(n_1^2 - n_2^2)}. \quad (15.11)$$

Окончательно характеристические уравнения мод записываются в виде

$$a\beta \operatorname{tg} \beta a = \alpha a, \quad -a\beta \operatorname{ctg} \beta a = \alpha a, \quad (\alpha^2 + \beta^2)a^2 = V^2.$$

В зависимости от того, составляющая какого поля – электрического или магнитного – оказывается перпендикулярной направлению распространения волны, различают поперечные электрическую ТЕ- и магнитную ТМ-моды. Если моды распространяются вдоль оси  $z$  (рис. 15.1) с координатной зависимостью амплитуды, пропорциональной  $\exp(-j\gamma z)$ , то ТЕ-мода будет содержать только  $E_y$ ,  $H_x$ ,  $H_z$ , а ТМ мода  $H_y$ ,  $E_x$ ,  $E_z$  компоненты электромагнитного поля. В первом случае в поперечной плоскости волновода оказывается ограниченным электрическое поле волны, во втором – магнитное. Для волноводных структур это определяет различные режимы их работы. Волноводные моды являются основными рабочими модами устройств интегральной оптики. Число волноводных мод зависит от толщины волновода  $2a$ , частоты поля  $\omega$  и показателей преломления сред, образующих волновод. При данной  $\omega$  число волноводных мод растет с увеличением  $a$ .

**Критические частоты.** Поведение каждой моды при изменении частоты можно рассмотреть следующим образом. Критические частоты удовлетворяют следующим условиям:  $\beta a = \pi/2$ ,  $\alpha = 0$ , откуда получаем

$$\omega_c = \frac{\pi c}{2a\sqrt{n_1^2 - n_2^2}}. \quad (15.12)$$

Новые моды появляются при всех частотах  $n\omega_c$  ( $n = 0, 1, 2, \dots$ ). В диэлектрическом волноводе существует нулевая мода с  $n = 0$ , т.е. для диэлектрического световода нет частотного порога. В диэлектрическом волноводе при распространении волны от нижней стенки до верхней сдвиг фазы может быть скомпенсирован изменением фазы  $\Phi$  при отражении (обратном проходе от верхней стенки до нижней).

Когда возникает мода распространения, мы имеем  $\alpha = 0$ , откуда в силу формулы (15.1)  $\cos \theta_c = \eta$ , то есть угол – предельный. В этом случае распространение волны возможно, так как мы находимся на пределе преломления. Но условие  $\alpha = 0$  означает, что волна распространяется во второй среде. При этом фазовая скорость равна фазовой скорости во второй среде. При увеличении частоты  $\alpha$  и  $\beta$  увеличиваются, а угол  $\theta$  убывает. Поле не так глубоко проникает во внешнюю среду, и волна концентрируется внутри световода. В пределе, когда частота  $\omega$  стремится к бесконечности,  $\theta = 0$  и волна полностью удерживается в первой среде.

В частном случае нулевой моды  $n = 0$  легко показывается, что

$\beta a \sim V$  при  $\omega \rightarrow 0$ , и мы снова получаем  $\theta = \theta_c$  при  $\omega = 0$ . Но согласно (15.8), на критической частоте выполняются соотношения

$$\sin \theta_c = \frac{\beta}{kn_1} = \frac{n\pi \lambda}{2a 2\pi} = \frac{n\lambda}{4a} = n \frac{\omega_c}{\omega},$$

где  $\omega_c = \pi c/2an_1$ , т.е.  $\omega_c$  – это критическая частота металлического волновода, заполненного диэлектриком с показателем преломления  $n_1$ .

На критической частоте поле внутри диэлектрического световода точно такое же, как в соответствующем металлическом волноводе.

С точки зрения дисперсии желательно иметь одномодовый режим. Для этого необходимо выполнить условие  $\omega < \omega_c$ . Максимальный размер такого одномодового световода определяется выражением

$$(4a)_{\text{одномод}} \leq \frac{\lambda}{\sqrt{n_1^2 - n_2^2}}. \quad (15.13)$$

Таким образом, можно получить одномодовый световод с поперечными размерами во много длин волн, если достаточно мала разность показателей преломления  $\Delta n = n_1 - n_2$ .

**Расширение волнового пакета.** Волновой пакет, распространяющийся в дисперсной среде, постепенно расширяется. При перемещении на расстояние  $l$  временное уширение, связанное с вкладом полосы частот  $\omega \pm \Delta\omega$ , приблизительно дается выражением

$$\frac{\Delta t}{\Delta\omega} = l \frac{|\partial\vartheta_{\text{гр}}/\partial\omega|}{\vartheta_{\text{гр}}^2}.$$

Можно также ввести закон дисперсии в виде зависимостей  $n(\omega)$  или  $n(\lambda)$ . Например:

$$\Delta\omega \approx \sqrt{\frac{c}{\left| 2 \frac{dn}{d\omega} + \omega \frac{d^2n}{d\omega^2} \right|}}. \quad (15.14)$$

Уширение волнового пакета, обусловленное средой распространения, пропорционально корню квадратному из пройденного расстояния. Полоса передаваемых частот  $\Delta\omega$  выражается в единицах ГГц·км<sup>1/2</sup>.

Для моды TE<sub>0</sub> в пределе низких частот ( $\omega \rightarrow 0$ ) ширина импульса  $T$  увеличивается в  $\sqrt{2}$  раз на характеристическом расстоянии

$$L = \frac{c^3 T^2}{6\omega a^2 (n_1^2 - n_2^2)^2}. \quad (15.15)$$

Можно сделать заключение, что:

а) длительность импульса входит в это выражение в квадрате;

- б) частота входит в первой степени;
- в) выгодно использовать среды с близкими показателями преломления.

### 15.1.2. Типы световодов и их характеристики

Оптический волновод состоит из двух слоев с низкими значениями показателя преломления и расположенной между ними волноводной областью с высоким значением показателя преломления. По характеру изменения показателя преломления в поперечной плоскости оптические волноводы делятся на волноводы со ступенчатым (рис. 15.3,а) и плавным, или градиентным (рис. 15.3,б) изменениями показателя преломления [42]. На рис. 15.3 изображены профили показателя преломления и распределение электромагнитного поля для мод  $TE_0$  и  $TE_1$  в планарных волноводах со ступенчатым (а) и градиентным (б) изменениями показателя преломления, где  $n_3$  –показатель преломления подложки волновода;  $n_2$  – показатель преломления волноводного слоя;  $n_1$  - показатель преломления среды над волноводом.

В ступенчатых волноводах показатель преломления волноводного слоя постоянен. Для реализации волноводного режима в нем необходимо выполнение условия  $n_2 > n_{эф} > n_1$ , где  $n_{эф} = \gamma/k_0$ . Нормированная диаграмма  $\omega - \beta$  для планарного волновода со ступенчатым изменением показателя преломления при различных значениях степени его асимметрии  $\alpha$ . изображена на рис. 15.4 в виде графических решений дисперсионного уравнения волновода (15.7). Это уравнение определяет постоянную распространения как функцию частоты  $\omega$  и толщины волновода  $2a$  для основных мод различного порядка  $m$ . По осям рис. 15.4 отложены нормированная частота  $V^* = k_0 2a \sqrt{n_2^2 - n_3^2}$  и нормированный показатель преломления  $b = (n_{эф}^2 - n_3^2)(n_2^2 - n_3^2)^{-1}$ . Параметром кривых служит степень асимметрии волноводной структуры  $a^* = (n_3^2 - n_1^2)(n_2^2 - n_3^2)^{-1}$ . Для симметричного волновода со ступенчатым изменением показателя преломления  $n_3 = n_1$  и  $a^* = 0$ . У волноводов с ярко выраженной асимметрией  $n_3 \neq n_1$ ,  $n_3 \rightarrow n_2$ ,  $a^* \rightarrow \infty$ . Максимальное число распространяющихся по волноводу мод ограничено соотношением  $M = \frac{1}{2} + \frac{2d}{\lambda_0} (2n_3 A_1)^{1/2}$ , где

$A_1 = n_2 - n_1$ . При этом основная мода симметричного волновода со ступенчатым законом изменения показателя преломления не имеет отсечки. Для такого волновода электромагнитное поле в пределах волноводного слоя изменяется по синусоидальному закону, а вне его затухает по экспоненте.

В градиентном волноводе граница раздела подложки и волноводной области не ярко выражена и определяется профилем показателя преломления, задаваемого соотношениями:

$$n(x) = n_3 + \Delta n(x), \quad x > 0; \quad n(x) = n_1, \quad x < 0,$$

где  $x$  - координата, вдоль которой задается градиент  $n$ .

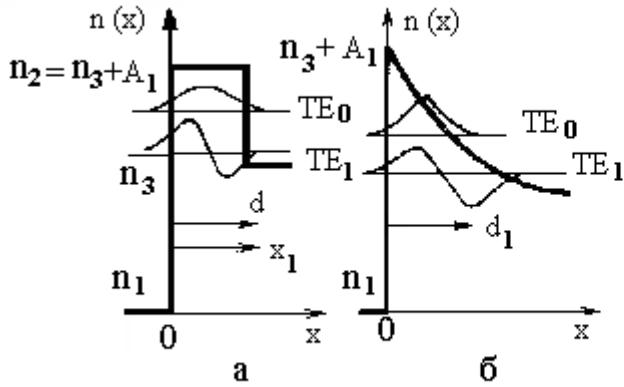


Рис. 15.3. Графики показателя преломления и распределения ЭМП для мод  $TE_0$  и  $TE_1$  в планарных волноводах со ступенчатым (а) и градиентным (б) изменениями  $n$ .

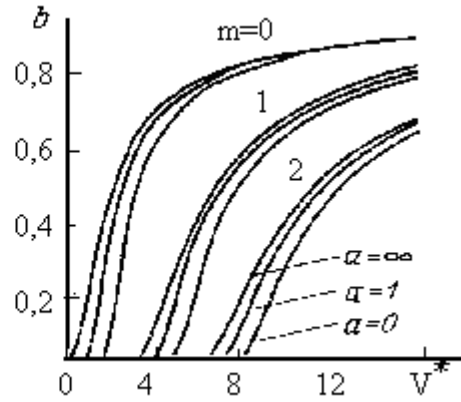


Рис. 15.4. Дисперсионные уравнения волновода.

Градиент  $n$  с достаточной точностью может быть аппроксимирован зависимостью вида  $\Delta n(x) = A_1 \exp(1 - x/d_x)$ , где  $a_x$  - характеристическая толщина волноводной области. Число мод, распространяющихся по такому волноводу, ограничивается неравенством  $M \leq 1/4 + (4a_1/\lambda_0)(2n_3A_1)^{1/2}$ . Распространение электрического поля в пределах  $0 < x < x_j$ , где  $x_j$  - точка поворота моды - соответствует периодической кривой. За пределами этих границ оно затухает по экспоненте.

В планарных волноводах ограничение светового канала имеет место только в одной плоскости. В интегральной оптике это ухудшает согласование микроволноводов с другими ее компонентами и волоконными световодами, осложняет локализацию оптических связей, снижает эффективность волноводных устройств. Значительно большими степенями свободы характеризуются трехмерные волноводы, которые и получили основное распространение в интегральной оптике и интегральных оптоэлектронных устройствах.

**Трехмерные волноводные структуры.** Основные типы трехмерных плосковых и канальных волноводов схематически представлены на рис. 15.5. Деление трехмерных волноводов на плосковые и канальные в некоторой степени условно, тем не менее, оно отражает конструктивные особенности структур, влияющие на возможное число волноводных мод, специфику распределения полей и, следовательно, на основные их характеристики. В канальных волноводах гребенчатого типа (рис. 15.5, а б) значительный скачок  $\Delta n$  в поперечном направлении приводит к многомодовому их режиму. Погруженные канальные структуры (рис. 15.5, в-д) характеризуются промежуточными свойствами по сравнению с

гребенчатыми канальными и полосковыми волноводами. Число распространяющихся мод в них определяется размером канала и разностью показателей преломления материалов, образующих волновод. Полосковый волновод, или волновод с наложенной полоской (рис. 15.5,е-к), представляет собой планарный тонкопленочный или диффузионный волновод с узкой диэлектрической областью, нанесенной на его поверхность. В таких волноводах практически вся световая мощность локализована под полоской диэлектрика и почти не проникает в нее. В результате неровности границ раздела между полоской и волноводом не приводят к значительным потерям излучения. В полосковых волноводах легко реализуются одномодовый или маломодовый режим работы.

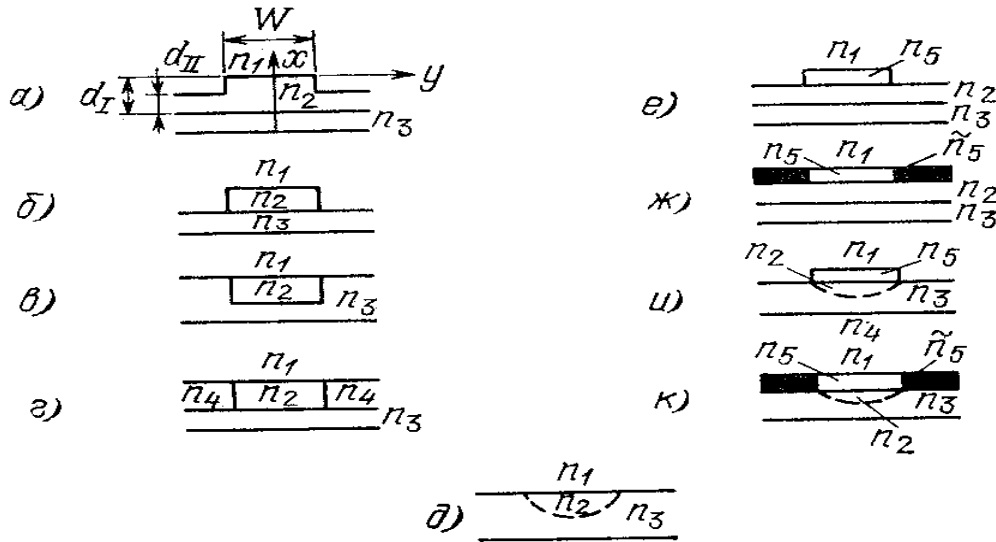


Рис. 15.5. Основные типы трехмерных волноводов:  
а-д – канальные, е-к – полосковые.

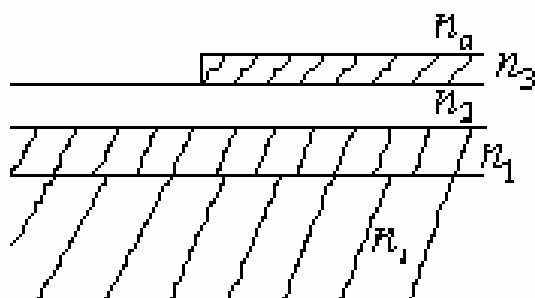
Наиболее перспективными для интегральной оптики являются различные виды полосковых и погруженных канальных волноводов. Основные их достоинства связаны с широкими технологическими и конструктивными возможностями создания эффективных волноводных структур на их основе. К достоинствам полосковых структур следует отнести слабую чувствительность к нарушениям границ диэлектрической полоски и увеличение допуска на ее формат в одномодовом режиме. В интегрально-оптических схемах ее ширина может достигать десяти и более микрон, что в значительной степени упрощает операцию стыковки волноводов.

Для канальных и полосковых структур все волноводные характеристики, в том числе и дисперсионные, зависят не только от формы и относительных размеров волновода, но и от электродинамических параметров материалов, образующих волновод.

В настоящее время разработаны основы теории и технологии тонкопленочных микроволноводов, которые позволяют рассчитывать и изготавливать волноводные структуры с заданными свойствами. Перечень материалов, пригодных для создания на их основе волноводных структур, включает различные стекла, в том числе халькогенидные, применяемые для среднего и длинноволнового инфракрасных диапазонов, различные полимеры, ферриты, активные диэлектрики, электрооптические кристаллы, керамику и полупроводники.

### 15.1.3. Элементы связи

Волноводный элемент связи между планарными структурами рис.



$$n_1, n_3 \text{ и } n_2 > n_1, n_3$$

Рис.15.6. Волноводный элемент связи между планарными структурами.



Рис.15.7. Связь волноводов с  $n_1$  и  $n_3$  через третий слой с  $n_2$ .

Связь между двумя волноводными планками осуществляется за счет затухающих полей, которые проникают в промежуточный слой с  $n_2$ . Обмен энергией можно рассчитать с помощью метода связанных мод.

Два изолированных тонкопленочных волновода связаны через третий слой, обладающий суживающимися соединениями в местах соприкосновения с волноводами 1 и 2 (рис. 15.7).

Этот метод использует гибридную технологию. Если поверхностная волна распространяется слева в плёнке с  $n_1$ , то тогда она достигает суживающегося края данной плёнки, энергия излучается в пространство. Однако, поскольку  $n_2 > n_3$ , энергия излучается в промежуточную плёнку, а не в подложку. Затем она распространяется в промежуточной плёнке в виде поверхностной волны. Аналогичный процесс происходит, если поверхностная волна распространяется справа в плёнке с  $n_3$ . Согласно теореме взаимности, энергия может быть выведена из промежуточной плёнки в зависимости от направления распространения волны. Таким образом, энергия может быть передана справа налево (или наоборот) через промежуточную плёнку. С помощью плавного сужения можно достичь эффективности связи, равной почти 100%.

Недостаток: общая подложка.

Если плёнки расположены на разных подложках, то передачу из одного волновода в другой можно осуществить с помощью решетки (см. рис. 15.8). Если необходима полная передача энергии, то длина решётки и длина перекрытия между волноводами должны быть равны длине связи. Здесь эффективность связи может быть равна 65%.

На рис. 15.9 показан наиболее простой ввод – непосредственная фокусировка на торец световедущего слоя при помощи линзы (трудность – требует субмик-



ронной юстировки). На рис. 15.10 показана торцевая безлинзовая стыковка волноводной структуры и полупроводникового лазера.

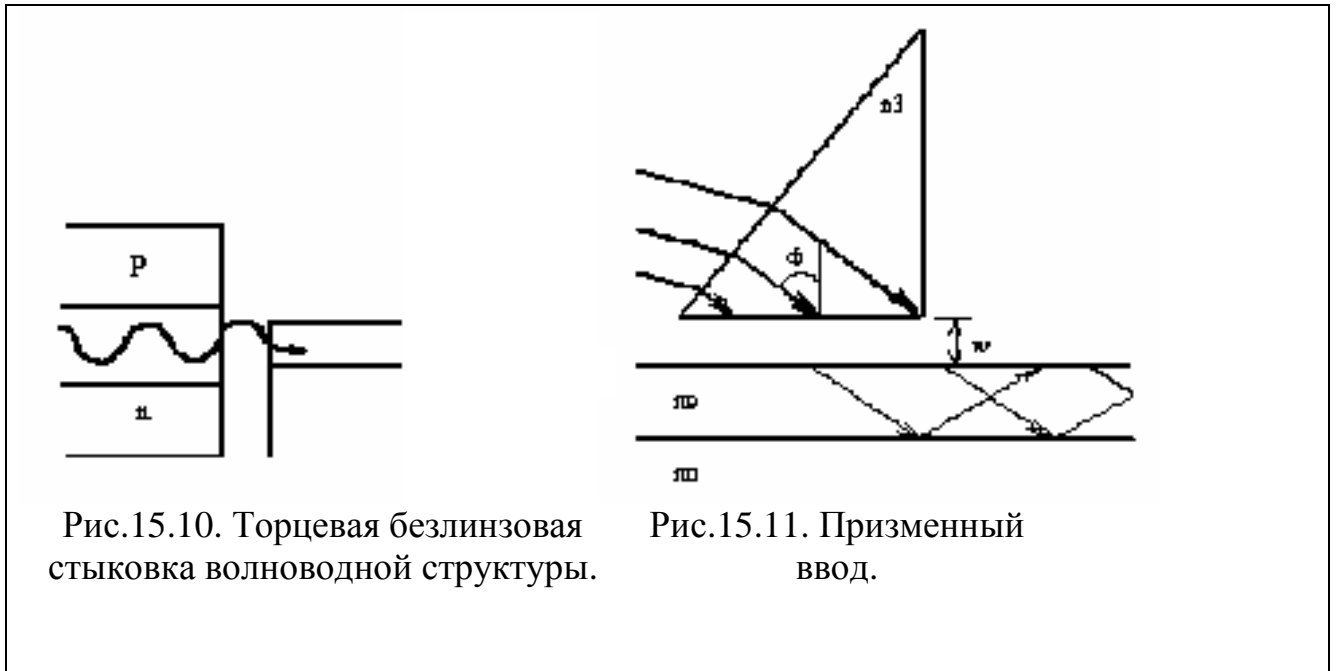


Рис.15.10. Торцевая безлинзовая стыковка волноводной структуры.

Рис.15.11. Призмный ввод.

Эффективность ввода при помощи этих методов довольно низкая. На рис. 15.11 изображен призмный ввод. За счет эффекта оптического туннелирования энергия проникает в волновод.

Изменяя угол  $\Phi$ , можно возбуждать различные моды в планарном волноводе. В случае вывода различные моды будут характеризоваться различными углами вывода световых волн. Эффективность ввода – 88 % .

Самый оптимальный способ связи - это решетчатый элемент связи (рис. 15.12), когда штрихи нанесены на волноводный слой с периодом  $d$ .

Если  $d \approx \lambda$ , и разность фаз для двухсоседних штрихов равна  $K \cdot \lambda \cdot \sin \theta_0$ .

Волноводная мода возбуждается, если

$$\Delta\phi := 2\pi \cdot n_0 \cdot \frac{\pi}{\lambda} \cdot \sin(\theta_0) + \frac{2\pi m}{d} := \beta_m$$

Варьируя угол падения пучка, можно выбирать моду, для которой будет осуществлена передача энергии. Эффективность ввода в этом случае достигает 71 %.

**Вывод излучения.** Вывод излучения можно осуществить, используя клиновидный конец волновода (рис. 15.13).



Рис.15.12. Решетчатый элемент связи.

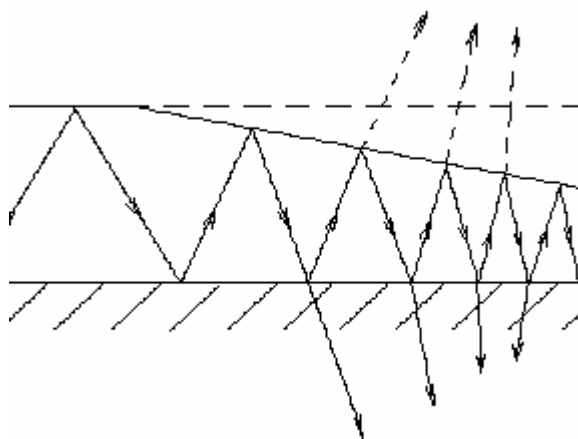


Рис.15.13. Вывод излучения через клиновидный конец волновода.

Мода, подойдя к клину, испытывает многократное отражение. Распространяясь вдоль клина, каждый раз излучает всё меньшую и меньшую энергию. На каком-то этапе угол падения может стать меньше угла внутреннего отражения, энергия излучается в воздух, что приводит к потерям. Чтобы этого не происходило, сужение клина должно быть плавным. Эффективность вывода – 50-70% .

## 15.2. Волоконно – оптические световоды

Оптические волокна послужили основой для разработки и создания оптических систем связи высокой эффективности, обеспечивающие возможность передачи большого потока информации на любые расстояния.

### 15.2.1. Свойства оптического волокна

Подобно плоским диэлектрическим волноводам в интегрально – оптических схемах, волоконные световоды различной длины применяются для направленной передачи оптического излучения, которое вводят через торцевой, обычно плоский конец световолокна. В простейшем случае оптическое волокно представляет собой тонкую нить круглого сечения из прозрачного в нужной спектральной области материала (стекло, кварц, различные полимеры).

По волоконно-оптическим линиям связи (ВОЛС) со скоростью передачи – 2,5 Гбит/с можно одновременно организовывать от 7680 до 300000 телефонных каналов тональной частоты (ТЧ) со скоростью передачи 64 кБит/с. Важнейшим фактором в развитии оптических систем явилось появление оптического квантового генератора – лазера. Лазерные системы работают в оптическом диапазоне длин волн, что составляет частоту порядка  $10^{14}$  Гц.

**Оптическое волокно.** Рассмотрим теперь цилиндрическое стеклянное волокно, состоящее из внутренней сердцевины с показателем преломления  $n_1$  (или в некоторых случаях  $n_B$ ) и окружающей её оболочки с показателем преломления  $n_2$ , причем здесь также выполняется условие  $n_1 > n_2$  . Расчет проводится на основе

[43]. Торец волокна срезан под прямым углом к его оптической оси. На рис. 15.14 изображен луч, входящий в волокно с торца. Этот луч

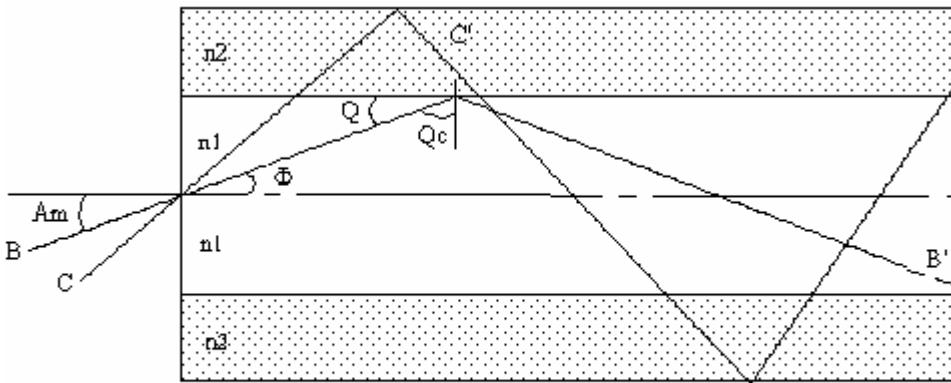


Рис.15.14. Траектория лучей в волокне

распространяется вдоль волокна путем многократных отражений от границы сердцевина – оболочка. Распространение луча будет осуществляться при условии  $Q < Q_c$ ,

для этого необходимо, чтобы угол наклона луча к оптической оси волокна ( $\Phi$ ) был меньше угла ( $\Phi_m$ ), где  $\Phi_m = \pi/2 - Q_c$ .  $A < A_m$ . Для определения величины углов  $\Phi_m$  и  $A_m$  воспользуемся законом Снеллиуса, приняв  $n_a = 1$ ,

Чем больше  $A_m$ , тем большая часть падающего на торец света может быть введена в волокно и будет в нем распространяться за счет полного внутреннего отражения. Числовая апертура (NA) волокна определяет способность микробио-ективов собирать свет и определяется как

$$(NA) = \sin A_m = \sqrt{2n \cdot \Delta n} = \sqrt{\frac{\Phi}{\Phi_0}}, \quad (15.16)$$

где  $\Phi_0$  – мощность, излучаемая источником,  $\Phi$  – мощность, введенная в волокно.

Из (15.16) видно, что для того, чтобы ввести в волокно как можно больше энергии, необходимо обеспечить большие значения величин  $n$  и  $\Delta n$ . Очевидно, что лучшее, что может быть сделано – это использовать для изготовления волокна стекло с большим показателем преломления и не покрывать его оболочкой. В этом случае полное внутреннее отражение будет происходить на поверхности стекло-воздух.

**Типы оптических волноводов (ОВ).** Ступенчатые волоконные световоды. Различают световоды со ступенчатым профилем, у которых показатель преломления сердцевин  $n_B$  одинаков по всему поперечному сечению (рис. 15.15,а), и градиентные - с плавным профилем, у которых  $n_B$  уменьшается от центра к периферии (рис. 15.15).

Канализация света в волокне основывается на явлении полного отражения. Как в прямолинейном, так и в изогнутом волоконном световоде направляемые волны можно наглядно представить лучами, распространяющимися зигзагообразно, под углами к его оси и каждый раз испытывающими полное внутреннее отражение на поверхности раздела. Для этого необходимо, чтобы коэффициент преломления материала световода был обязательно больше коэффициента преломления окружающей его среды (стекло, кварц, различные полимеры). Пучок света, испытывая полное внутренне отражение и распространяясь в однородной

прозрачной нити, в идеальном случае не должен терять энергию. Однако практически потери все же происходят. Учитывая большую длину волокна (в отличие от полоскового волновода в интегрально-оптических схемах), они могут оказаться значительными. Заметная часть потерь приходится, прежде всего, на рассеяние света на поверхности волокна.

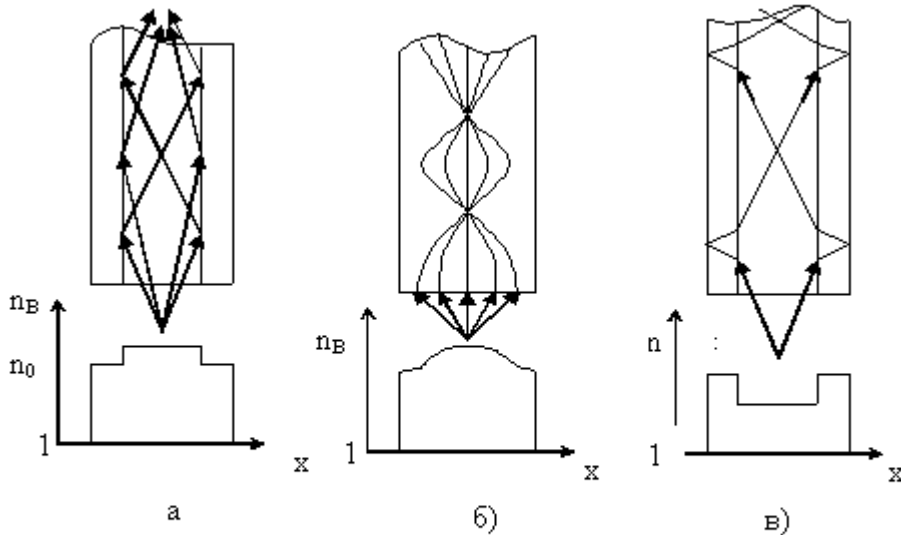


Рис. 15.15. Устройство и ход лучей в оптическом (а), градиентном (б) и газо-диэлектрическом (в) световодах. Внизу распределение показателя преломления по радиальному сечению.

Низкая механическая прочность делают тонкую прозрачную нить не пригодной для непосредственного применения в качестве световода. Положение радикально изменяется при переходе к двухслойному световоду, представляющему собой световедущую сердцевину (жилу), покрытую оболочкой из материала, показатель  $n_o$  которого меньше, чем жилы  $n_B$ . Как и в интегрально-оптическом волноводе, при полном внутреннем отражении световая волна проникает в оболочку, поэтому, для того, чтобы волна не затухала, в соответствующей области спектра оболочка должна быть прозрачной. Необходимо также обеспечить высокое совершенство границы раздела между сердцевиной и оболочкой. Если толщина оболочки превышает несколько длин волн света, интенсивность световой волны, достигающей внешней поверхности оболочки оказывается пренебрежимо малой. Поэтому, на процесс распространения канализируемых в сердцевине волн, состояние поверхности оболочки практически не влияет. Световод со ступенчатым профилем коэффициента отражения можно покрыть еще одной защитной, упрочняющей, например, полимерной оболочкой, снабдить армирующими элементами, обеспечивающими необходимую механическую прочность и стойкость к воздействию окружающей среды, и превратить волоконный световод в оптический кабель, пригодный для практического использования. Оптическая линия связи может содержать как один, так и множество световодов.

## 15.2.2 Параметры и характеристики оптического волокна

Покажем, что только часть света (пропорциональная  $(NA)^2$ ), излучаемая малоразмерным диффузным источником, помещенным на оптической оси волокна вблизи его торца, может быть введена в волокно и, следовательно, будет в нём

распространяться. Предположим, что мощность, излучаемая в единицу телесного угла в направлении под углом  $\theta$  к нормали его поверхности, определяется выражением

$$I(\theta) = I_0 \cos \theta. \quad (15.17)$$

Полная мощность  $P_0$ , излучаемая таким источником, находится интегрированием  $I(\theta)$  по всем направлениям:

$$P_0 = \int_0^{\pi/2} (I_0 \cos \theta)(2\pi)(\sin \theta)d\theta = -2\pi I_0 [\cos^2 \theta / 2]_{\theta=0}^{\pi/2} = \pi I_0. \quad (15.18)$$

Мощность  $P$ , введенная в волокно, диаметр сердцевины которого больше диаметра источника, определяется следующим интегралом:

$$\begin{aligned} P &= \int_0^{\alpha_m} (I_0 \cos \theta)(2\pi)(\sin \theta)d\theta = -2\pi I_0 [\cos^2 \theta / 2]_{\theta=0}^{\alpha_m} = \\ &= \pi I_0 \sin^2 \alpha_m = P_0 (NA)^2, \quad P/P_0 = (NA)^2 = 2n\Delta n. \end{aligned} \quad (15.19)$$

Отсюда ясно, что для того, чтобы ввести в волокно как можно больше света, необходимо обеспечить большие значения величин  $n$  и  $\Delta n$ . Очевидно, что лучшее, что может быть сделано – это использовать для изготовления волокна стекло с большим показателем преломления и не покрывать его оболочкой. Но при использовании таких волокон возникают две проблемы.

Первая состоит в том, что при полном внутреннем отражении часть волны света проникает сквозь отражающую поверхность. Она является затухающей волной. Её амплитуда уменьшается экспоненциально с увеличением расстояния от поверхности волокна и поэтому обычно не может распространяться в среде с более низким показателем преломления. Однако, наличие всяких неровностей и неоднородностей на отражающей поверхности может привести к преобразованию затухающей волны в волну распространения. В жгуте из волокон без оболочек условия на отражающей поверхности неизбежно и неконтролируемо изменяются, поскольку отдельные волокна входят в контакт друг с другом и окружающей средой. В результате значительная доля распространяющейся мощности выводится из них, что приводит к большим потерям.

Вторая проблема связана с тем, что любой короткий световой импульс, введенный в волокно, состоит из ряда лучей, которые распространяются вдоль оси волокна, наклоненным к ней под некоторым углом. Показатель преломления среды можно рассматривать как меру скорости распространения света  $V$  в этой среде, то есть

$$V = c/n. \quad (15.20)$$

Следовательно, осевой луч будет проходить расстояние вдоль волокна за время  $n_1 l/c$ , в то время как наиболее наклоненный, который ещё может распространяться в волокне, то же самое расстояние пройдет за время, определяемое соотношением

$$\frac{n_1 l}{c \cos \phi_m} = \frac{n_1 l}{c \cos \theta_A} = \frac{n_1^2 l}{n_0 c}, \quad (15.21)$$

где  $c$  – скорость света в вакууме.

Таким образом, если оба эти луча введены в волокно одновременно, то на выходе волокна они окажутся разделенными во времени на интервал  $\Delta t$ , определяемый формулой

$$\Delta t = (n_1/n_2)(l/c) \Delta n. \quad (15.22)$$

В результате световой импульс, содержащий лучи под всеми возможными углами, окажется размытым во времени в процессе своего распространения по волокну на величину, определяемую выражением

$$\Delta t/l = (n_1/n_2)(\Delta n/c). \quad (15.23)$$

Можно определить разницу времен распределения световых импульсов вдоль осевого и наиболее наклонного лучей по формуле, учитывающей групповой показатель преломления сердцевины  $n_{\Gamma}$

$$\Delta t/l = (n_{\Gamma}/n_2)(\Delta n/c). \quad (15.24)$$

Время  $\Delta t$  для  $l = 1$  км и значения  $\Delta n = 1 \cdot 10^{-2}$  составляет несколько десятков наносекунд, что соответствует предельной частоте модуляции несколько десятков МГц (верхняя граничная частота, выраженная в герцах, по величине приближенно равна скорости передачи информации в битах в секунду). Принято говорить, что такие искажения сигнала (не связанные с немонахроматичностью света) обусловлены межмодовой (многолучевой) дисперсией и называется временной дисперсией волокна.

Как было показано, дисперсионные свойства оптических материалов характеризуются зависимостью показателя преломления от частоты (от длины волны в свободном пространстве). Поэтому, необходимо выразить величины  $V_{\Gamma}$  и  $n_{\Gamma}$  через  $n$  и  $\lambda$ . Отметим, что

$$n_{\Gamma} = c / V_{\Gamma} = c(d\beta / d\omega) = c(d / d\omega)(\omega n / c) = n + (\omega dn / d\omega). \quad (15.25)$$

$$a \quad dn / d\omega = (dn / d\lambda)(d\lambda / d\omega) \quad \text{или} \quad d\omega / d\lambda = -(2\pi c / \lambda^2).$$

Подставляя полученные выражения в (15.25), находим

$$n_{\Gamma} = n + \frac{2\pi c}{\lambda} \frac{dn}{d\lambda} \left( -\frac{\lambda^2}{2\pi c} \right) = n - \lambda \frac{dn}{d\lambda}. \quad (15.26)$$

Таким образом,

$$V_{\Gamma} = c / n_{\Gamma} = c / [n - \lambda dn / d\lambda]. \quad (15.27)$$

Тогда время прохождения  $t$  световым импульсом расстояния  $l$  будет равно

$$t = \frac{l}{V_{\Gamma}} = \frac{n_{\Gamma} l}{c} = \left[ n - \lambda \frac{dn}{d\lambda} \right] \frac{l}{c}. \quad (15.28)$$

Если свет имеет ширину спектра  $\Delta \lambda$  относительно  $\lambda$  и если среда дисперсионная, то световой импульс в процессе распространения, как было уже показано, искажается, но интервал времени  $\Delta t$  теперь равен:

$$\Delta t = \frac{dt}{d\lambda} \Delta \lambda = \frac{l}{c} \frac{dn_{\Gamma}}{d\lambda} \Delta \lambda = \frac{l}{c} \left[ \frac{dn}{d\lambda} - \frac{dn}{d\lambda} - \lambda \frac{d^2 n}{d\lambda^2} \right] - \frac{1}{c} \lambda \frac{d^2 n}{d\lambda^2} \Delta \lambda. \quad (15.29)$$

Покрытие сердцевины волокна стеклянной оболочкой, имеющей немного меньший показатель преломления, приводит к возникновению трёх эффектов: 1) если покрытие имеет высокое качество и толщину, достаточную для удержания затухающей волны, то оно существенно уменьшает потери; 2) уменьшению временной дисперсии; 3) уменьшению вводимой в волокно мощности света.

Если  $\Delta n \ll n$ , то выражение (15.23) для временной дисперсии волокна можно преобразовать к виду

$$\Delta t/l \approx \Delta n/c. \quad (15.30)$$

Взаимосвязь между временной дисперсией волокна  $\Delta t$ , и шириной полосы частот, занимаемой сигналом, и максимальной скоростью передачи информации  $B$ , можно записать в следующем виде:

$$B \approx 2\Delta f \approx 1/\Delta t, \quad (15.31)$$

откуда

$$(\Delta f)l \approx c/2\Delta n. \quad (15.32)$$

Лучи, падающие на торец световода под углом, выходящим за пределы  $\theta_A$  не испытывают полого внутреннего отражения от поверхности сердцевинки, преломляются в оболочке и могут приводить к нежелательным паразитным связям, например, между световодами в многожильных кабелях. Для их устранения желательно, чтобы вторая защитная оболочка была светопоглощающей.

Диаметр сердцевинки волоконного световода может, в зависимости от его назначения, составлять от нескольких микрометров до нескольких сотен микрометров, а толщина оболочки – от нескольких десятков микрометров вплоть до миллиметра (один из стандартных диаметров оболочки – 125 мкм.). Когда диаметр сердцевинки соизмерим с длиной волны света, по световолокну, как и по волноводу, может распространяться только дискретная совокупность волн. Им соответствуют различные углы падения лучей на поверхность раздела сердцевинка – оболочка, а это, в свою очередь приводит к различной длине пути, которые проходят лучи (рис.15.15,а). В результате, передаваемый по световоду, сигнал искажается. Если на вход световода подать короткий световой импульс, то при достаточно большой длине может случиться, что на выходе импульсы, переносимые разными модами, окажутся разделенными по времени, т. е. на выходе вместо одного импульса появится серия импульсов. Фазовая и групповая скорости каждой моды в световоде зависят от частоты, то есть световод является дисперсной системой. Вызванная этим волноводная дисперсия является одной из причин искажения передаваемого сигнала. Различие групповых скоростей различных мод в многомодовом режиме называется модовой дисперсией. Она является весьма существенной причиной искажения сигнала, поскольку он переносится по частям многими модами. Если входной импульс удлинить, то серия импульсов на выходе, в конце концов, сольется в один импульс увеличенной длительности. По этой причине нельзя беспредельно увеличивать частоту модуляции распространяющегося по волноводу сигнала: период модуляции должен превышать разницу времени прохождения световода отдельными модами.

### Одномодовые волокна

Волоконный оптический волновод (ВОВ) может работать в одномодовом режиме. Для этого нужно, чтобы выполнялось условие, аналогичное для плоского волновода

$$h_{кр} \approx \frac{\lambda}{4} \sqrt{2n_1 \cdot \Delta n} \quad (15.33)$$

под  $h_{кр}$  – подразумевается диаметр сердцевины волокна. Таким образом, для одномодового режима диаметр световолокна и различие показателей преломления  $\Delta n = n_1 - n_2$  должны быть достаточно малыми.

В одномодовых световодах дисперсия как таковая не проявляется. Качественная передача сигнала в этом случае все же ограничена тем, что используемое излучение не может быть строго монохроматическим, а занимает конечный спектральный интервал. Из-за этого возникает материальная дисперсия, обусловленная зависимостью показателя преломления сердцевины  $n_1$ , а значит и скорости распространения от длины волны света. Если принять, что ширина спектральной полосы одномодового полупроводникового лазера составляет несколько нанометров, а световода – несколько десятков нанометров, то различие во времени прохождения пути длиной 1 км для граничных длин волн составит по порядку величины  $1 \cdot 10^{-10}$  с, т. е. предельные частоты модуляции сигнала равны  $1 \cdot 10^{10}$  и  $1 \cdot 10^9$  Гц.

Возможны также другие виды дисперсии, проявляющиеся в тех случаях, когда межмодовую и материальную дисперсии удается свести к минимуму. Рассмотрение на основе волновой теории показывает, что даже в одномодовом режиме скорость распространения света по волокну зависит от частоты независимо от материальной дисперсии. Этот вид дисперсии называют волноводной (внутримодовой). С практической точки зрения важно, что в определенном спектральном интервале материальная и волноводная дисперсии могут компенсировать друг друга. При этом может проявиться еще один вид дисперсии, обусловленный тем, что, например, из-за механических напряжений в волокне, отклонения формы сердцевины от цилиндрической, волны с взаимно ортогональной поляризацией распространяются с различными скоростями. Такая дисперсия, называемая поляризационной, приводит к уширению светового сигнала на величину порядка  $1 \cdot 10^{-11}$  с при длине волокна 1 км. Как следует из изложенного, для всех видов дисперсии уширение импульса пропорционально длине световолокна. Поэтому пропускную информационную способность приводят к единице длины световолокна, а полосу пропускания выражают в мегагерцах-километрах или гигагерцах-километрах.

Следует иметь в виду, что широкая полоса пропускания достигается ценой усложнения технологии изготовления и его эксплуатации. Эффективное введение энергии в световод требует того, чтобы и источник излучения был одномодовым лазером. К этому нужно также добавить, что числовая апертура одномодового волокна заметно меньше, чем многомодового: 0,1 и 0,2 – 0,3 соответственно.

Достоинства одномодового световода: малы потери на пропускание, которые количественно оцениваются затуханием светового сигнала (в дБ на километр). Высокая пропускная способность; невосприимчивость к электромагнитным помехам; не подвергаются коррозии; малые размеры, малая масса; передача недоступна для прослушивания и перехвата.

В одномодовом волокне отсутствует модовая дисперсия, и сигнал искажается значительно меньше, чем в многомодовом.

## Потери в оптических волокнах



В основном, потери света в ОВ обусловлены двумя причинами: первая – поглощение, вторая – рассеяние. Имеется множество других потерь, сравнительно малых по отношению к поглощению и рассеянию и мы их рассмотрим в методическом пособии, при разборе конкретных задач.

**Поглощение в оптическом волокне.** Поглощение определяется свойствами материала и рабочей длиной волны. Оно имеет место при возбуждении в материале электронных переходов и резонансов с последующими неизлучательными релаксационными процессами. В результате увеличивается тепловая энергия, накапливаемая в материале; Потери в волокне зависят не только от качества материалов сердцевины. Значительную роль играет также и материал оболочки. При полном внутреннем отражении электромагнитные волны проникают через раздел сердцевина – оболочка и распространяются в оболочке. И если оболочка имеет плохое качество или большое поглощение, то она будет вносить заметный вклад в общие потери в волокне.

**Рассеяние в оптическом волокне** частично может обуславливаться свойствами материала, но в основном определяется нарушениями геометрической формы ОВ. Оно происходит тогда, когда мода распространения света изменяется таким образом, что часть оптической энергии покидает волокно.

По своей природе стекло является неупорядоченной структурой, в которой имеются микроскопические отклонения от средней плотности материала, а также локальные микроскопические изменения в составе. Каждое из указанных изменений приводит к флуктуации показателя преломления. Это справедливо для любого стеклообразного материала, однако и при качественном изготовлении в нём наблюдается рассеяние света, известное как рэлеевское. Характерная особенность данного явления состоит в том, что рассеиваемая мощность, а, следовательно, и потери обратно пропорциональны длине волны в четвёртой степени. Именно рэлеевское рассеяние, а не край полосы УФ поглощения является основной причиной потерь в кварцевых ОВ на длинах волн короче 1,5 мкм. Основные нарушения геометрии поверхности и большие дефекты в сердцевине волокна приводят к значительным локальным потерям. Резкие изгибы ОВ приводят к тому, что часть света не будет отражаться от оболочки, а будет в ней распространяться и таким образом теряться. Теоретически рассеиваемая мощность экспоненциально зависит от радиуса изгиба. На практике, однако, минимально допустимый радиус изгиба определяется исходя из механических свойств волокна.

Оптические волокна имеют очень малое (по сравнению с другими средами) затухание светового сигнала в волокне. Лучшие образцы российского волокна имеют затухание 0,22 дБ/км на длине волны 1,55 мкм, что позволяет строить линии связи длиной до 100 км без регенерации сигналов. Для сравнения, лучшее волокно Sumitomo на длине волны 1,55 мкм имеет затухание 0,154 дБ/км. В оптических лабораториях США разрабатываются еще более "прозрачные", так называемые фтороцирконатные волокна с теоретическим пределом порядка 0,02 дБ/км на длине волны 2,5 мкм.

### 15.2.3. Ввод и вывод излучения

Трудность связи между волноводами и волокнами состоит преимущественно в том, что сердцевина волокна имеет относительно малые размеры и её трудно связать с волноводом так, чтобы соединение было достаточно жестким, а эффективность связи высокой.

Один из методов соединения волокна с волноводом состоит в следующем. На плёночный волновод наносится жидкая капля с высоким показателем преломления (рис. 15.16). В этом случае, получаем утечку энергии под углом  $\theta$ . Следовательно, вставив конец стеклянного волокна в эту жидкость и ориентируя волокно под углом  $\theta$  к нормали, извлекаем энергию из волокна. Коэффициент эффективности связи может достигать достаточно большой величины ( $\eta=50\%$ ). Но такая связь не слишком подходит для стационарных установок. Кроме того эффективность связи при передаче энергии из волокна в волновод обычно хуже, т.к. поле в волокне не имеет соответствующего экспоненциального распределения, которое необходимо для возбуждения волны утечки, что приводит к сильному



Рис. 15.16 Элемент связи между волноводом и волокном



Рис. 15.17. Элемент связи через суживающийся край пленки.

рассогласованию полей.

Другой метод – связь через суживающийся край пленки рис. 15.17. Эффективность такого метода достаточно высока, но применяется в основном к многомодовым волокнам.

Следующий метод – через туннелирующий слой (Рис. 15.18). Связь, с применением суживающегося края края плёнки, и наложенного снаружи волокна не круглой, а прямоугольной формы. Для лучшего согласования конец полоскового волновода суживается и между полосковыми волноводами и внешним волокном вводится слой с низким показателем преломления. Он действует как туннелирующая область, в пределах которой происходит передача энергии ( $\eta=50\%$ ).

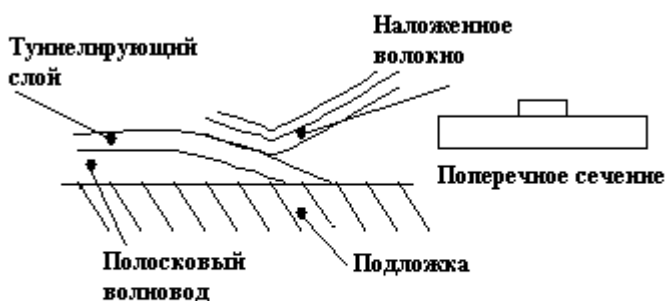


Рис. 15.18. Элемент связи через туннелирующий слой

оптических кабельных систем передачи (генераторов, фотоприёмников, разъёмных и неразъёмных соединителей, ответвителей и других элементов).

## ЗДЕЛ ЧЕТВЕРТЫЙ

### ОПТОЭЛЕКТРОННЫЕ СИСТЕМЫ И ИХ ПРИМЕНЕНИЯ

#### Глава 16. Системы обработки и хранения информации

Эффективная и своевременная обработка больших массивов информации является одной из важнейших научно-технических проблем. Развитие науки и техники идет по пути необычайно быстрого и постоянно увеличивающегося роста потока информации, подлежащего переработке и использованию. Появление большого количества задач, решение которых целиком зависит от возможности переработки больших массивов информации, выдвигают высокие требования к своевременной и эффективной ее обработке.

В данной главе мы рассмотрим возможности применения оптико – электронных устройств в области обработки и хранения информации. В основу изложения этой главы взяты материалы из [44-48].

Любой физический процесс, несущий в себе ту или иную информацию, называется сигналом. В оптике носителем информации является световая волна. Следовательно, световая волна может рассматриваться как оптический сигнал. Преобразование таких сигналов производит оптический процессор.

Чем больший объем информации может быть закодирован в параметрах носителя информации, чем больше скорость его прохождения через систему обработки, тем больше быстродействие такого устройства. Использование когерентного светового поля (лазерного излучения) в качестве переносчика информации позволяет создать достаточно гибкую систему обработки информации, способную обрабатывать большой объем данных в очень короткое время.

Оптические системы в большинстве случаев представляют собой специализированные процессоры для выполнения основных математических операций, одно- и двумерного преобразований Фурье, вычисления корреляционных функций. Такие системы высокоэффективны и обладают большим быстродействием. На базе таких систем могут быть построены гибкие оптические процессоры, способные решать широкий класс радиотехнических задач.

#### 16.1. Структура системы оптической обработки сигналов

Сейчас имеется много оптических и цифровых процессоров, которые при высокой скорости обработки данных (до миллиарда операций умножения и сложения в секунду) обеспечивают высокую точность измерения параметров выходных сигналов. Одним из перспективных направлений применения оптических методов является обработка радиолокационных сигналов. Быстродействие радиолокационных станций (РЛС) с системой оптической обработки информации ( $10^{12} \dots 10^{14}$  бит/с) значительно превышает быстродействие РЛС с традиционной электронной обрабатывающей системой ( $10^6 \dots 10^8$  бит/с).

Структурная схема системы оптической обработки сигналов приведена на рис. 16.1. Для выполнения заданного алгоритма обработки на аналоговое оптическое вычислительное устройство поступают оптические сигналы, в которых закодирована информация, выдаваемая источником информации. Преобразование электрических сигналов в оптические выполняется источником света и преобразователем входных сигналов в оптический аналог. После выполнения заданного алгоритма обработки аналоговое оптическое вычислительное устройство формирует на выходе системы распределения амплитуд и фаз света, в которых содержатся результаты обработки сигналов, поступающих от источника инфор-



Рис. 16.1. Структурная схема системы оптической обработки сигналов.

мации. Преобразователь выходных сигналов измеряет параметры световых сигналов, поступающих на его выход, и выдает получателю информации конечный результат обработки в необходимом виде (как правило, в виде электрических сигналов). Преобразователь входных сигналов является устройством ввода информации в оптическую обрабатывающую систему, поэтому требования к нему во многом зависят от тех требований, которые предъявляются к системе обработки информации. Так например, при обработке оптической системы, предназначенной для обработки сигналов, поступающих на антенную решетку, требования к устройству ввода будут определяться характеристиками РЛС: дальностью обнаружения, разрешающей способностью системы по направлению, полосой пропускания, динамическим диапазоном обрабатываемых сигналов, быстродействием. Таким образом, устройство ввода должно быть многоканальным (50 - 100 каналов), иметь высокую чувствительность, широкую полосу пропускания (десятки МГц), идентичные характеристики отдельных каналов, низкий уровень собственных шумов, большой динамический диапазон обрабатываемых сигналов (не менее 40 дБ). В качестве устройств ввода применяются различные типы пространственно-временных модуляторов света, принцип работы которых основан на взаимодействии света с динамическими неоднородностями, реализуемыми различными физическими эффектами в твердых и жидких веществах.

**Оптические методы обработки информации.** В решении проблемы эффективной обработки больших массивов информации особый интерес представляют оптические методы. Это обусловлено исключительно высокой информационной емкостью светового поля как переносчика информации, высокой скоростью распространения оптических сигналов и легкостью осуществления целого ряда интегральных операций над двумерными массивами информации. Потенциальные возможности оптических методов обработки информации позволяют надеяться в будущем на существенный прогресс в развитии вычислительной техники.

Наряду с применением оптических и оптоэлектронных элементов в средствах дискретной счетно-решающей техники в последние годы интенсивно развиваются аналоговые оптические методы и средства обработки информации, в основе которых лежат явления переноса информации пространственно-модулированным волновым полем и дифракции света.

В отличие от оптоэлектронных устройств ЦВМ, где параллельная обработка информации может осуществляться за счет пространственной дискретизации, в аналоговых оптических вычислительных устройствах она реализуется естественным образом при эквивалентном числе каналов, определяемом разрешающей способностью оптической системы и используемого носителя информации. Так, например, при размере обрабатываемого транспаранта  $24 \times 24$  мм, записанного на носителе с разрешающей способностью  $100 \text{ мм}^{-1}$  (легко реализуемый случай), эквивалентное число параллельных каналов составит  $5,76 \cdot 10^6$ .

Эффективность обработки информации с помощью оптических методов можно проиллюстрировать на примере осуществления преобразования Фурье. Как известно, операция двумерного Фурье-преобразования над когерентным оптическим сигналом осуществляется обыкновенной сферической линзой за время, равное времени распространения света в оптической системе, т. е. за  $10^{-9}$  сек. при оптической длине анализатора в 30 см. Если учесть, что в  $1 \text{ см}^2$  носителя может содержаться до  $10^9$  бит информации (при разрешающей способности носителя  $3000 \text{ мм}^{-1}$ ), то потенциальная производительность такого аналогового оптического анализатора составит  $10^{18}$  бит/сек. $\cdot\text{см}^2$ . Это на много порядков превышает производительность самых совершенных ЭЦВМ.

На практике скорость обработки информации ограничивается скоростью ввода и вывода информации, а не временем обработки оптических сигналов. Однако и с учетом ограничений быстродействия, связанных с вводом и выводом информации, аналоговые оптические вычислительные устройства все же оказываются намного более производительными, чем ЭЦВМ (приблизительно на 4 порядка).

Аналоговая оптическая обработка информации может быть реализована с помощью как когерентных, так и некогерентных световых полей. Аналоговые оптические вычислители, использующие некогерентный свет, непосредственно оперируют с интенсивностью светового поля, а фазовая модуляция в таких системах может быть осуществлена только косвенным образом. Они конструктивно просты, но решают сравнительно узкий круг задач, связанных с обработкой главным образом одномерных сигналов (спектральный и корреляционный анализ и некоторые другие операции). Когерентные оптические системы обработки информации оперируют с амплитудами и фазами светового поля. Эти системы более гибкие и позволяют решать широкий круг задач по обработке двумерных комплексных функций, задаваемых в виде некоторого распределения амплитуд и фаз светового поля. Однако они конструктивно сложнее, чувствительны к механическим воздействиям и предъявляют более жесткие требования к качеству оптических элементов, пространственно-временным модуляторам (управляемым транспарантами) и источникам света. Следует отметить, что почти все математические операции над оптическими сигналами, реализуемые в когерентном свете, могут быть выполнены и в некогерентном свете, однако степень сложности практической

реализации будет зависеть от вида операции, т. е. решаемой задачи. Так, например, операция двумерного Фурье-преобразования в когерентном свете реализуется исключительно просто, а в некогерентном свете необходимо использовать серию масок, число которых зависит от требуемой точности выполнения операции.

### 16.1.1. Принцип аналоговой оптической обработки информации с помощью когерентных световых полей

Интерес к применению оптических методов в аналоговой вычислительной технике обусловлен тем фактом, что любая оптическая система по своей природе является аналоговым устройством обработки информации, поступающей на ее вход в виде светового поля, модулированного в пространстве и во времени.

Обобщенная структурная схема аналоговой когерентной оптической системы обработки информации приведена на рис. 16.2 и включает в себя оптический квантовый генератор ОКГ, расширитель пучка 7, устройство ввода информации 2, аналоговое оптическое вычислительное устройство 3, устройство вывода информации 4, запоминающее устройство 5.

Устройство ввода осуществляет пространственную модуляцию сформированной расширителем пучка плоской монохроматической волны по амплитуде, фазе или поляризации в соответствии с обрабатываемыми сигналами, поступающими на его вход. Основной задачей устройства ввода является преобразование подлежащих обработке сигналов в когерентные оптические сигналы (поля). Обрабатываемые сигналы могут иметь различную физическую природу, однако обычно они преобразуются либо в электрические, либо в оптические. В первом случае устройство ввода должно иметь электронный вход и оптический выход, а во втором — оптический вход и оптический выход.

Аналоговое оптическое вычислительное устройство осуществляет над сформированным когерентным оптическим сигналом требуемую математическую операцию. Результат обработки в общем случае представляет собой некоторую световую картину. Устройство вывода преобразует результат вычисления в электрический сигнал для трансляции в ЭЦВМ (или иное электронное устройство) или регистрирует на светочувствительный носитель (например, фотопленку).

Когерентные аналоговые оптические вычислительные устройства позволяют относительно просто выполнять следующие математические операции над функциями двух переменных: умножение, интегрирование, сложение и вычитание, преобразование Фурье, свертку, взаимную корреляцию, дифференцирование, преобразования Гильберта, Френеля и др.

Перечисленные операции могут быть реализованы различными способами. Обычно используют разного рода маски, с помощью которых осуществляют пространственную модуляцию поля по требуемому закону, или голограммы. Запоминающее устройство (оптическое или голографическое) служит для хранения требуемого набора эталонных масок или голограмм и является обязательным элементом оптико-электронных информационно-поисковых систем и многофункциональных аналоговых оптических вычислительных устройств.

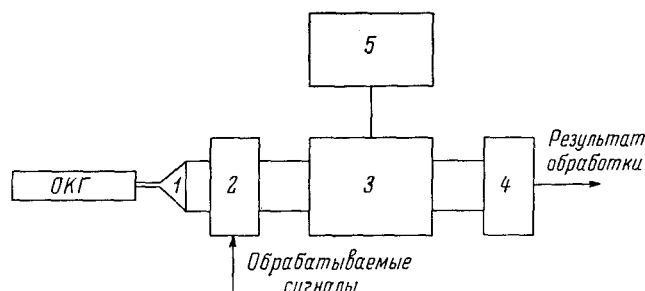


Рис. 16.2. Структурная схема аналоговой оптической системы обработки информации

Можно показать, что даже на основе двух базовых операций — умножения и преобразования Фурье, которые проще всего реализуются в оптике, возможно выполнить широкий класс операций по обработке информации: операцию свертки, дифференцирование, интегрирование с весом, преобразования Френеля, изменение масштаба аргумента функции, восстановление функции из ее спектральной плотности, сложение и вычитание и другие.

Когерентные аналоговые оптические вычислители обладают следующими основными достоинствами.

1. Обеспечивают возможность параллельной обработки больших массивов информации благодаря своей двумерности и чрезвычайно высокой информационной емкости светового поля и используемого носителя.

2. Обеспечивают высокую скорость выполнения операций, не зависящую от степени их сложности, которая практически ограничивается скоростью ввода—вывода информации, а не скоростью собственно обработки, определяемой временем прохождения света от входа до выхода системы.

Хотя в целом аналоговые оптические вычислители уступают ЦВМ по точности вычисления, однако, благодаря скорости вычислений, простоте аппаратуры и ее обслуживания, отсутствию необходимости программирования они оказываются более эффективными при решении ряда практических задач. Чаще всего эти задачи связаны с выполнением операций спектрального и корреляционного анализа над двумерными комплексными функциями. Решение их на ЭЦВМ сопряжено со значительными затратами машинного времени и потому не всегда экономически оправдано. В некоторых частных случаях оптические методы обработки не уступают электронным по точности и даже превосходят их.

Следует отметить, что в значительной части задач обработки информации эвристического типа не требуется высокой абсолютной точности промежуточных вычислений. В частности, в задачах опознавания изображений, где оптическими методами осуществляется формирование корреляционного поля анализируемых изображений, важна не абсолютная точность вычислений кросскорреляционных функций, а обеспечение требуемого относительного превышения максимума сигнала опознавания над фоном.

Оптические методы обработки информации, несмотря на их еще короткую историю развития, находят уже достаточно широкое применение в вычислительной технике, кибернетике, радиотехнике, радиолокации и связи, экспериментальной физике, астрономии, метеорологии и геофизике, машиностроении и приборо-

строении, криминалистике, микроэлектронике, медицине, биологии, биофизике и др.

В настоящее время развитие методов и средств оптической обработки информации идет по пути создания специализированных аналоговых оптических вычислительных устройств для решения широкого круга прикладных задач, связанных в большинстве случаев с необходимостью осуществления спектрального или корреляционного анализа над двумерными массивами информации, а также пространственной фильтрации.

Оптические устройства обработки информации можно классифицировать по ряду различных признаков: области применений, назначению решаемых задач, виду обрабатываемых сигналов и т.д. Можно условно объединить, существующие устройства оптической обработки информации, в следующие четыре группы осуществляющие: 1) спектральный анализ; 2) пространственную фильтрацию; 3) корреляционный анализ; 4) интегральные преобразования (Гильберта, Френеля и др.).

## 16.2. Способы выполнения математических операций оптическими системами

**Оптическое умножение.** В оптической системе обработки информации сигналы обычно перемножаются путем модуляции светового потока объектом, устанавливаемом в плоскости изображений (транспарантом). Функция пропускания транспаранта  $T(x,y)$  известна и умножается на входной сигнал, заданный в виде распределения амплитуд и фаз светового потока. Предположим, что в фокальной плоскости  $P_1$  (рис. 16.3,а) расположен транспарант с коэффициентом прозрачности

$$T(x_1, y_1) = S_o(x_1, y_1) \cdot e^{i\theta(x_1, y_1)} \quad (16.1)$$

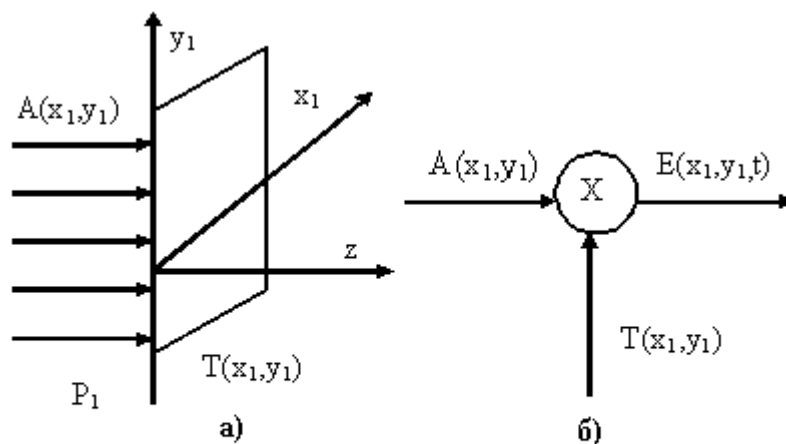


Рис. 16.3. Оптический модулятор (а) и его эквивалентная схема (б).



Обозначим когерентную монохроматическую световую волну, распространяющуюся в направлении оси  $z$  и освещающую оптическую систему

$$A(x_1, y_1) = E_o \cdot e^{i(\phi_o + \omega \cdot t)}$$

Заданное пространственное распределение коэффициента прозрачности транспаранта является модулирующей функцией уже световой волны  $A(x_1, y_1) = E_o \cdot e^{i(\phi_o + \omega \cdot t)}$ . На выходе такого "модулятора" комплексное распределение амплитуды светового поля можно записать в следующем виде (см. рис. 16.3, б):

$$E_1(x_1, y_1, t) = A(x_1, y_1) \cdot T(x_1, y_1) = E_o \cdot S_o(x_1, y_1) \cdot e^{i(\omega t + \phi_o + \theta(x_1, y_1))} \quad (16.2)$$

Как видно, выражение (5.2) описывает операцию умножения постоянной величины  $A(x_1, y_1) = const$  функцию  $T_1(x_1, y_1)$ . Если вплотную к транспаранту поставить другой транспарант с функцией пропускания  $T_2(x_1, y_1)$ , то на выходе получим результат перемножения в виде

$$E_1(x_1, y_1, t) = A(x_1, y_1) \cdot T_1(x_1, y_1) \cdot T_2(x_1, y_1) \quad (16.3)$$

Следовательно, суммарная модуляционная характеристика транспаранта, полученного при наложении нескольких транспарантов, равна произведению модуляционных характеристик отдельных транспарантов (экранов), а носителем информации являются световая волна. Таким образом, в оптическом решающем устройстве сигналы перемножаются путем модуляции светового потока с помощью оптических элементов (экранов, транспарантов, линз, решеток и т.п.), амплитудное пропускание которых соответствует перемножаемым сигналам, Оптическое умножение и более сложных функций можно осуществить с помощью оптической системы, показанной на рис. 16.4.

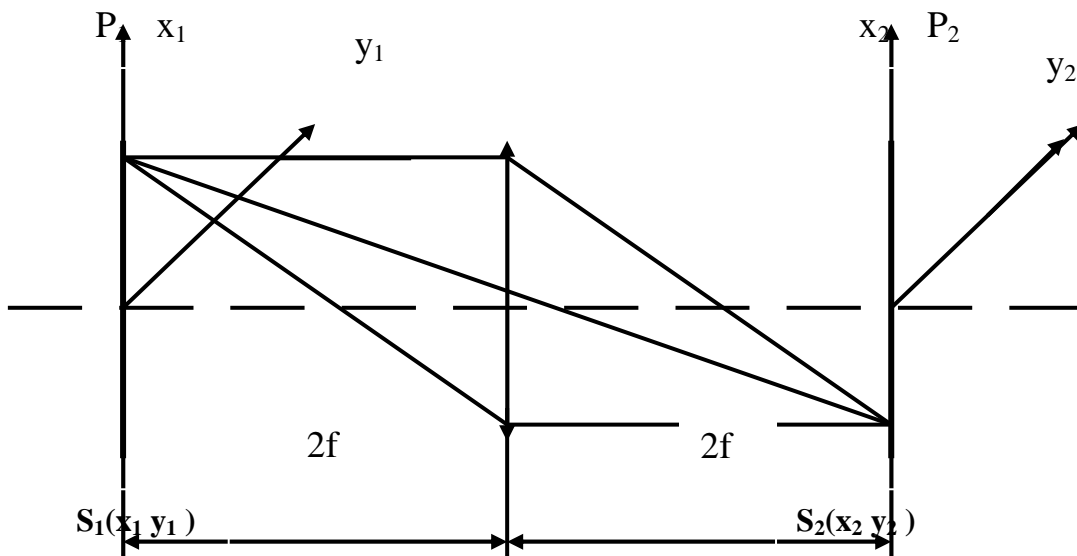


Рис. 16.4 Оптическое умножение комплексных функций

**Сложение сигналов.** В оптическом решающем устройстве световое поле в любой точке пространства есть результат суммирования световых полей, приходящих в эту точку. При этом вычисляется сумма, т.е. учитываются амплитуды и фазы поля. Если требуется строго определенное соотношение фаз отдельных сигналов при их суммировании, то возникает сложная задача контроля фазовых сдвигов сигналов. В свою очередь величина фазового сдвига сигнала зависит от длины оптического пути, пройденного сигналом. Так как длина волны света, например,  $\lambda = 0,63$  мкм, то изменение пути всего на несколько долей микрометра приводит к существенному изменению фазовых сдвигов сигналов. В качестве примера на рис. 16.5 показано, как складываются спектры двух сигналов. В передней фокальной плоскости линзы расположен экран с записью сигналов  $S_1(x,y)$  и  $S_2(x,y)$ , соответственно на верхней и нижней половинах экрана. В задней фокальной плоскости линзы световой поток  $E_o$ , прошедший экран, собирается, образуя сумму спектров входных сигналов. На практике устанавливаются весьма высокие требования к точности взаимного расположения сигналов.

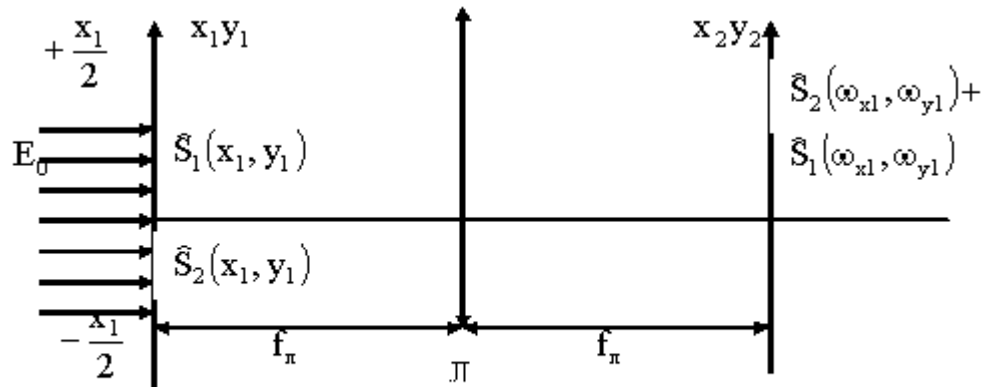


Рис. 16.5. Схема суммирования спектров двух сигналов

По условию  $\lambda/T_x \ll l$ , точность ( $T_x = l/f_x$ ;  $f_x$  - пространственная частота сигнала) установки сигнала в передней фокальной плоскости должна быть порядка минимального периода пространственной частоте сигнала. На практике сложение сигналов находит достаточно широкое применение и является одной из основных операций при обработке информации.

**Интегрирование.** Операция интегрирование выполняется линзой, в фокусе которой амплитуда и фаза соответствуют интегралу от функции распределения света в плоскости изображений линзы:

$$E = \frac{E_o}{\lambda \cdot f_\Lambda} \cdot \int \int_x S(x, y) dx dy \quad (16.4)$$

В общем случае можно утверждать, что если линза выполняет преобразование Фурье над сигналом  $S(x,y)$ , то в точке  $(\omega_x = 0, \omega_y = 0)$  выходного сигнала выполняется операция интегрирования сигнала. Интегрирование такого рода есть вычисление постоянной составляющей сигнала. Операцию интегрирования легко

реализовать с помощью элементарной оптической системы, если сделать равными единице экспоненциальный множитель в преобразований Фурье, выполняемом той системой. Это означает, что надо приравнять нулю переменные  $x_1$  и  $y_1$ . Практически это можно сделать с помощью точечной диафрагмы или оптического приемника малого размера, который должен располагаться точно на оптической оси (рис. 16.6а). В одномерных многоканальных

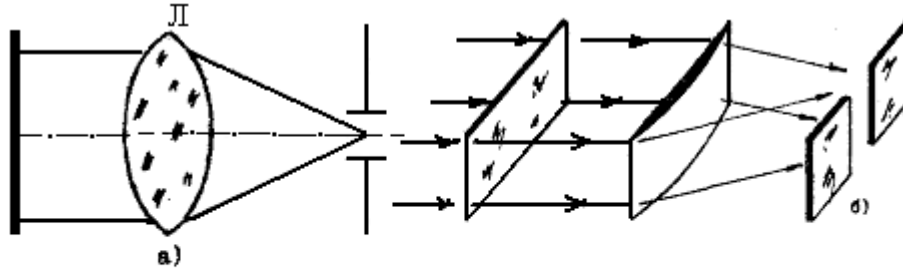


Рис. 16.6. Схема интегрирования одномерных и двумерных сигналов

системах используется щелевая диафрагма, причем направление щели совпадает с линией фокусов цилиндрической линзы (рис. 16.6, б).

**Дифференцирование функций.** Дифференцирование функций осуществляется в частотной области, т.е. в области преобразования Фурье. Применяя теорему о спектре производной временного сигнала, для  $n$ -ой производной можно получить выражение

$$S_n(\omega) = \int_{-\infty}^{\infty} \frac{d^n}{dt^n} x \cdot e^{-i \cdot \omega \cdot t} dt = (i \cdot \omega)^n \cdot S(\omega), \quad (16.5)$$

из которого следует, что первая производная функции определяется произведением спектра  $S(\omega)$  на угловую частоту ( $\omega$ ).

Для пространственных сигналов это означает, что в плоскости преобразования Фурье необходимо поставить фильтр, состоящий из маски с нулевой прозрачностью в центре и постепенно увеличивающейся прозрачностью к краям, (т.е. по мере возрастания пространственной частоты), а так же полуволновой пластинки, осуществляющей сдвиг фазы (перемену знака) для отрицательных пространственных частот.

### 16.2.1. Вычисление корреляционных функций

Рассмотрим оптическую систему, изображенную на рис. 5.7. Цилиндрическая линза  $z_1$  выполняет преобразование Фурье по оси  $x$  в плоскости  $(\omega_x, y)$ , а линза  $z_2$  - проектирует изображение сигнала  $S(x, y)$  по оси  $y$  на ту же плоскость. По оси  $y$  ( $\omega_x=0$ ) распределение светового поля соответствует одномерному многоканальному интегрированию. Если же в плоскости  $x_1 y_1$  два сигнала, т.е.

$$S_{x_1}(0, y_1, x_o) = \int_{x'(y)}^{x''(y)} S(x_1, y_1) \cdot q(x_1 - x_o, y_1) dx_1, \quad (16.6)$$

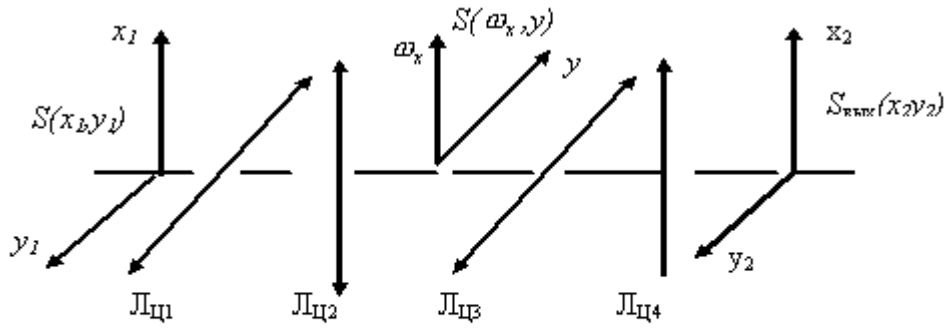


Рис. 16.7. Одномерный спектральный анализатор на цилиндрических линзах

где  $q(x_1-x_0, y)$  – функция пропускания дополнительного экрана в плоскости  $(x, y)$ , наложенного на экран с пропусканием  $S(x_1, y_1)$ . Выбирая различные функции  $q(x_1, y_1)$  и пределы интегрирования  $x'(y)$  и  $x''(y)$  (размер и форму отверстия входной диафрагмы), можно обеспечивать различные интегральные операции. Линия взаимное положение экранов (переменная  $x_0$ ), на оси  $y$  выходной плоскости можно получить взаимно корреляционную функцию сигналов  $S(x)$ , и  $q(x)$  в различных каналах по оси  $y$ .

Если тот же экран с функцией пропускания ( $q(x_1-x_0, y_1)$ ) повернуть, наложив его на экран с функцией пропускания, будет равно зеркальному отображению  $q(x_1-x_0)$ , т.е.  $q(x_0-x_1)$ . Такая система выполняет операцию свертки сигналов  $S(x)$  и  $q(x)$ . Но такие оптические системы обладают следующими недостатками; необходимость относительного перемещения сигналов, вычисление интеграла последовательно во времени в процессе перемещения сигналов.

**16.2.2. Преобразование Фурье в оптической системе**

Разберем, как элементарная оптическая система, схема которой дана на рис. 16.8, может быть описана математически [48]. Если представим источник  $U$ , расположенный на плоскости  $P_1$ , в виде бесконечно малой  $\delta$  величины и запишем его величину освещенности в виде соотношения  $E_s(0, 0) = \delta(x, y)$  по определению двумерная  $\delta$ -функция

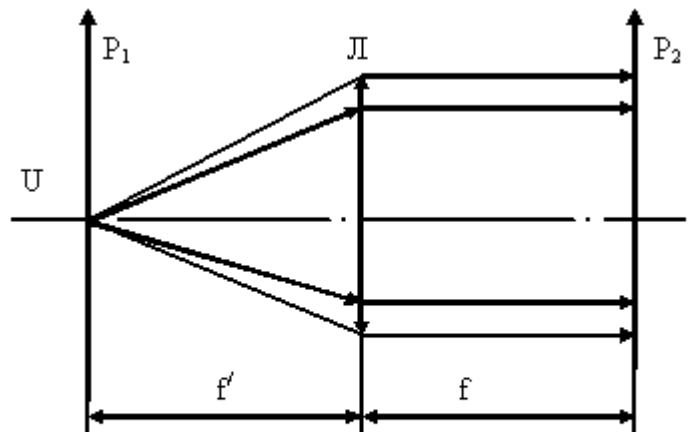


Рис. 16.8. Элементарная оптическая система

$$\lim_{x \rightarrow 0} I \rightarrow \begin{cases} 0 & \text{при } x < 0 \\ \infty & \text{при } 0 < x < x_0 \\ 0 & \text{при } x > 0 \end{cases} \quad (16.7)$$

обладающая следующими двумя свойствами:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (16.8,а)$$

$$\int_{-\infty}^{\infty} f(x) \cdot \delta(x - x_0) dx = f(x_0) \quad (16.8,б)$$

Последнее свойство часто называется фильтрующим свойством и  $\delta$ - функции. Найдем

$$S(\omega_x) = \int_{-\infty}^{\infty} \delta(x) \cdot e^{-j \cdot \omega_x \cdot x} dx \quad (16.9)$$

формально преобразование Фурье для  $\delta$ - функции:

Применив формулу Эйлера, получим:

$$S(\omega_x) = \int_{-\infty}^{\infty} \delta(x) \cdot \cos(\omega_x \cdot x) dx - \int_{-\infty}^{\infty} \delta(x) \cdot \sin(\omega_x \cdot x) dx \quad (16.10)$$

Поскольку по определению  $\delta(x)$  равна нулю на всей оси, кроме точки  $x = 0$ , где она стремится к бесконечности в выражении (16.10) Из первого интеграла можно вынести постоянный множитель

$$\cos(\omega_x \cdot x) = 1 \quad x = 0,$$

а из второго интеграла множитель

$$\sin(\omega_x \cdot x) = 0 \quad \text{при } x = 0.$$

Учитывая свойство (16.8,б), получим

$$S(\omega_x) = e^{j \cdot \omega_x \cdot x} = 1 \quad x = 0. \quad (16.11)$$

Можно  $\delta$  функцию представить в виде обратного преобразования Фурье от равномерной спектральной плотности

$$\delta(x) = \frac{1}{2 \cdot \pi} \cdot \int_{-\infty}^{\infty} S(\omega_x) \cdot e^{j \cdot \omega_x \cdot x} d\omega_x = \frac{1}{2 \cdot \pi} \cdot \int_{-\infty}^{\infty} e^{j \cdot \omega_x \cdot x} d\omega_x \quad (16.12)$$

Аналогичным образом можно написать соотношения для двумерной  $\delta$  - функции;

$$\delta(x - x_0, y - y_0) = \begin{cases} \infty & x = x_0 \\ 0 & y = y_0 \end{cases}$$

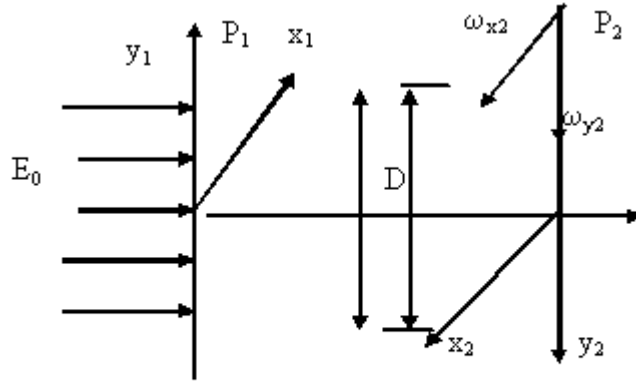
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - x_0, y - y_0) dx dy = 1$$

в остальной области

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cdot \delta(x - x_0, y - y_0) dx dy = f(x_0, y_0).$$

Согласно законам геометрической оптики сферическая линза преобразует распределение освещенности в виде  $\delta$ -функции в равномерную освещенность плоскости 1. Если сравнить этот хорошо известный экспериментальный факт с выражением для преобразования Фурье  $\delta$ -функции, то можно сделать важный вывод; идеальная сферическая линза обладает свойством выполнять преобразование, аналогичное преобразованию Фурье.

Рассмотрим, как осуществляется это преобразование, на примере оптической системы, изображенной на рис. 16.9.



**Рис. 16.9. Преобразование Фурье в оптической системе**

Примем следующие допущения: в системе не происходит поглощения и отражения электромагнитной энергии; система не имеет аберраций.

Предположим, что в передней фокальной плоскости  $P_1$  линзы  $L$  расположен транспарант с коэффициентом прозрачности

$$S(x_1, y_1) = S_o(x_1, y_1) \cdot e^{j\theta(x_1, y_1)}.$$

Обозначим когерентную монохроматическую световую волну, распространяющуюся в направлении  $z$  и освещающую оптическую систему  $E_o \exp(i \cdot \omega \cdot t)$ . Используя положение геометрической оптики, мы можем считать, что заданное пространственное распределение коэффициента прозрачности является модулирующей функцией для световой волны  $E_o \exp(i \cdot \omega \cdot t)$ . На выходе такого модулятора, комплексное распределение амплитуды светового поля можно записать в следующем виде

$$E_1(x_1, y_1, t) = E_o \cdot S_o(x_1, y_1) \cdot e^{j(\theta(x_1, y_1) + \omega \cdot t)} \quad (16.13)$$

Найдем распределение амплитуды светового поля на выходе оптической системы - задней фокальной плоскости линзы  $L$ . Эта линза с апертурой  $D$  будет осуществлять суммирование дифрагированной энергии в фокальной плоскости. Для оценки результата такого суммирования для некоторого направления  $V$  необходимо найти выражение для фаз дифрагированной энергии во всех точках плоскости, перпендикулярной направлению  $V$  и выразить его через координату плоскости  $x_1, y_1$ . При этом будем считать, что плоская волна, падающая на линзу  $L$ , образует с её осью очень малый угол. В общем виде, уравнение плоскости, перпендикулярной направлению  $V$  имеет вид

$$a_1 \cdot x_1 + b_1 \cdot y_1 + c_1 \cdot z_1 = 0, \quad (16.14,a)$$

где  $a_1, b_1, c_1$ , - направляющие косинусы плоскости, перпендикулярной направлению  $V$ . Тогда расстояние до точки в плоскости  $P_1$  до этой плоскости равно

$$D_2 = a_1 \cdot x_1 + b_1 \cdot y_1 \quad (16.14,b)$$

следовательно, для фаз дифрагированной энергии можно написать

$$\theta = \theta(x_1, y_1) + \frac{2 \cdot \pi}{\lambda} \cdot D_2 = \theta(x_1, y_1) + \frac{2 \cdot \pi}{\lambda} \cdot (a_1 \cdot x_1 + b_1 \cdot y_1), \quad (15.15)$$

где  $\lambda$  - длина волны света.

Далее определяется положение точки суммирования дифрагированной энергии, определяемое значениями  $A_i, B_i, C_i$ , пользуясь тем, что свет в направлении  $V$ , проходящий через центр оптической системы, не отклоняется от своего направления. Если расстояние от центра линзы до этой точки равно  $R$ , то

$$a_1 = \frac{x_2}{R} \quad b_1 = \frac{y_2}{R} \quad c_1 = \frac{F}{R}, \quad (16.16)$$

где  $x_2, y_2$  - координаты выходной плоскости.

Обозначим через  $L$  оптическую длину пути от рассматриваемой точки до входной плоскости  $P_1$  с координатами  $x_1$  и  $y_1$  до точки суммирования в выходной плоскости  $P_2$  с координатами  $x_2$  и  $y_2$ . Подставив соотношение (16.16) в выражение (16.15), получим

$$\theta = \theta(x_1, y_1) + \frac{2 \cdot \pi}{\lambda} \cdot L + \frac{2 \cdot \pi \cdot c_1}{F} \cdot (x_1 \cdot x_2 + y_1 \cdot y_2) \quad (16.17)$$

Теперь можно написать выражение для распределения освещенности  $E_2(x_2, y_2, t)$  в плоскости  $P_2$  в виде двойного интеграла

$$E(x_2, y_2, t) = E_0 e^{j\omega \cdot t} \iint_{P_2} S_0(x_1, y_1) e^{j\left\{\theta(x_1, y_1) + \frac{2 \cdot \pi}{\lambda} L + \frac{2 \pi c_1}{\lambda F} (x_1 x_2 + y_1 y_2)\right\}} dx_1 dy_1 \quad (16.18)$$

Введем вспомогательные обозначения

$$\omega_{x_2} = \frac{2\pi}{\lambda \cdot F} \cdot x_2, \quad \omega_{y_2} = \frac{2\pi}{\lambda \cdot F} \cdot y_2,$$

которые носят название пространственных частот и имеют размерность  $\mu^{-1}$ . Они являются пространственными аналогами частоты изменения сигналов во времени, имеющей размерность  $c^{-1}$ . Если дифрагированные углы малы, то  $c_1 \approx 1$ . Можно считать, что для всех значений  $x_2$  и  $y_2$ , величина  $\alpha$  будет постоянной. Поэтому член, содержащий  $\alpha$  в выражении (16.18), можно опустить, так как он характеризует постоянный сдвиг по всей плоскости P2. Так как обработки сигналов осуществляются за много периодов светового излучения, то  $E_o \exp(i \cdot \omega \cdot t)$  можно заменить его среднеквадратичным значением  $E_o$  и принять  $E_o = 1$ .

Тогда на основе принятых допущений окончательно получим:

$$E(x_2, y_2) = \iint_{P_2} S_o(x_1, y_1) \cdot e^{j \cdot (\omega_{x_2} \cdot x_1 + \omega_{y_2} \cdot y_1 + \theta(x_1, x_2))} dx_1 dy_1 \quad (16.19)$$

Это выражение является двумерным преобразованием Фурье функции  $S(x_1, y_1)$ . Если осуществлять только пространственную амплитудную модуляцию (например, в плоскости P1 задать плоское изображение), то  $\theta(x_1, x_2) = const$  и выражение (16.19) упростится

$$E(\omega_x, \omega_y) = S(\omega_x, \omega_y) = \iint S_o(x, y) \cdot e^{j \cdot (\omega_x \cdot x + \omega_y \cdot y)} dx dy \quad (16.20)$$

Итак, в когерентной идеальной оптической системе комплексное распределение амплитуды освещенности светового поля в передней фокальной плоскости линзы связано с комплексным распределением светового поля в задней фокальной плоскости линзы двумерным преобразованием Фурье. Функция  $S(\omega_x, \omega_y)$  называется пространственным спектром объекта  $S(x, y)$ . Плоскость, в которой дифракционная картина объекта  $S(x, y)$  описывается его преобразованием Фурье называется Фурье-плоскостью.

### 16.2.3. Пространственная фильтрация

Простота осуществления операций прямого и обратного фурье-преобразований над двумерными функциями (в общем случае комплексными) в сочетании с голографическим способом регистрации световых полей открывает широкие возможности для синтеза схем пространственной фильтрации с произвольным видом передаточной функции [47-48].

Оптические системы способны выполнять общие линейные преобразования поступающих на вход данных. Объем данных бывает настолько большим, что существенно превышает возможности наблюдателя. Тогда для уменьшения количества данных решающую роль будет играть линейное преобразование, выделяющие нужные группы, таким образом, чтобы они могли восприниматься наблюдателем.



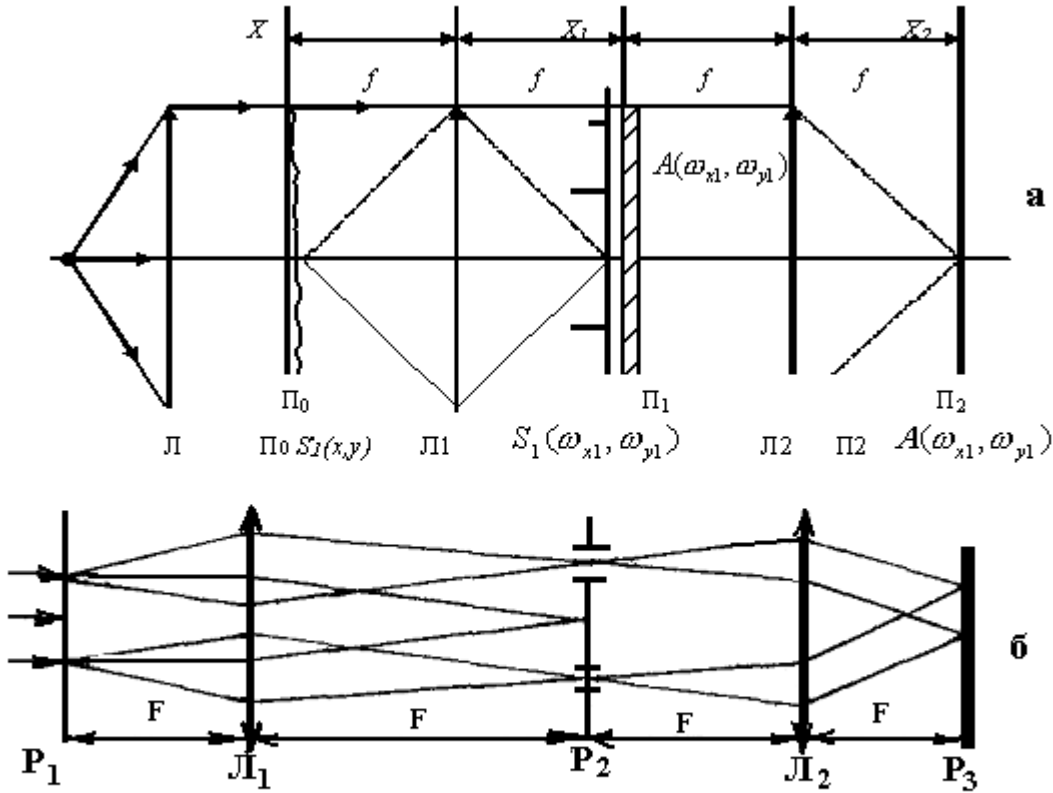


Рис. 16.10. Принцип двойной дифракции в когерентной оптической системе

Поскольку сферическая линза может выполнять преобразование Фурье, то помещая в Фурье-плоскости различные маски или транспаранты с заданным распределением прозрачности, можно определенным образом влиять на полученный пространственный спектр сигнала, т.е. осуществлять фильтрацию сигнала. Рассмотрим схему оптического пространственного фильтра, изображенного на рис. 16.10. Если в данной системе уберем фильтр

$A(\omega_{x1}, \omega_{y1})$ , то тогда когерентный световой поток, пройдя через транспарант с функцией пропускания  $S_1(x_1, y_1)$  создаст, благодаря дифракции в фокальной плоскости линзы  $L_1$ , пространственный спектр объекта  $S(\omega_{x1}, \omega_{y1})$ . Свет, пройдя через линзу  $L_2$ , образует изображение  $S(x_2, y_2)$  в плоскости  $\Pi_2$  с координатами  $x_2$  и  $y_2$  так что изображение  $S_2(x_2, y_2)$  можно рассматривать как результат операции интегрирования, происходящей над полем  $S(\omega_{x1}, \omega_{y1})$  в зрачке линзы  $L_2$ , т.е. изображение  $S_2(x_2, y_2)$  соответствует обратному преобразованию Фурье. Теперь введем фильтр с пропусканием  $A(\omega_{x1}, \omega_{y1})$  на который проектируется спектр сигнала  $S(\omega_{x1}, \omega_{y1})$ . На выходе плоскости  $\Pi_2$  имеем напряженность электрического поля

$$S'_2(x_1, y_1) = A(\omega_{x1}, \omega_{y1}) \cdot S(\omega_{x1}, \omega_{y1}), \quad (16.21)$$

изменение которого по пространственным частотам изменено в соответствии с характеристикой фильтра.

В качестве иллюстрации общих положений рассмотрим несколько примеров оптических пространственных фильтров. Если в плоскости спектра сигнала П1 стоит диафрагма с отверстием, как показано на рис. 16.11, то такая оптическая система соответствует полосовому фильтру в области верхних частот. Ширина полосы фильтра определяется шириной отверстия диафрагмы. Это явление особенно хорошо можно продемонстрировать на фильтрации пространственного спектра объекта, имеющего вид сетки. Так в спектральной плоскости П линзы Л (см. рис. 16.10,а) приведено изображение не изменённого спектра сетки. Из-за периодичности предмета появляется набор отдельных спектральных компонент; ширина каждой из них определяется апертурой падающей световой волны.

При помещении в спектральную плоскость узкой щели, так чтобы через неё проходил только один ряд спектральных компонент, в плоскости изображения П3 получим спектр сетки отфильтрованной этой щелью (рис. 16.12,б). Если в спектральную плоскость поместить ирисовую диафрагму и установить её так, чтобы через неё проходила только осевая Фурье компонента, то при последовательном расширении диафрагмы можно шаг за шагом проследить Фурье-синтез сетки. Если же убрать диафрагму, а вместо неё на оси линзы в фокальной

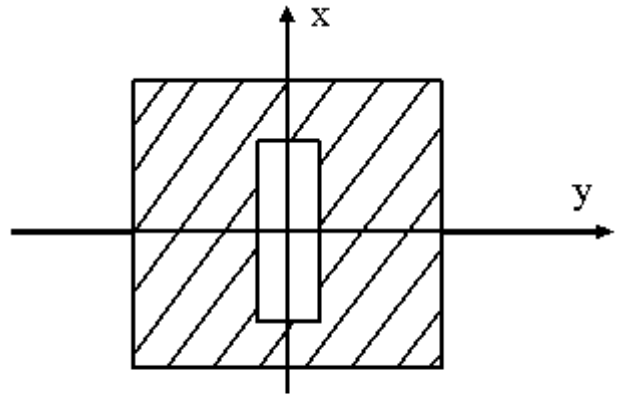


Рис. 16.11-Полосовой пространственный фильтр

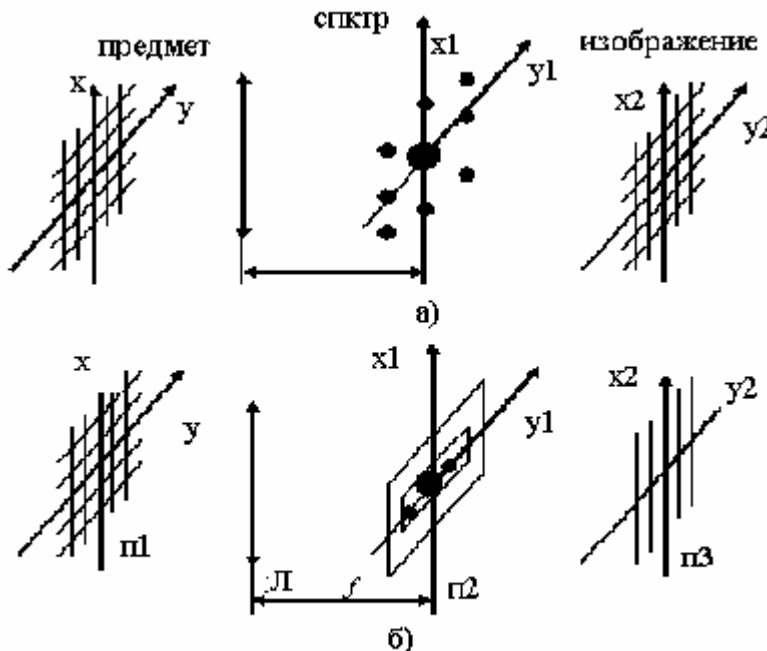


Рис. 16.12. Пространственная фильтрация

плоскости поместить маленький экран, который закрывал бы только центральный порядок, или компоненту “нулевой частоты”, то мы получим изображение сетки с обращенным контрастом (рис. 16.12,б).

### 16.2.4. Согласованная фильтрация

Для решения различных задач обнаружения сигнала в шумах наибольшее распространение получил метод оптической согласованной фильтрации [48]. Достоинством оптической согласованной фильтрации (ОСФ) является высокая информационная емкость, банк эталонных сигналов, многоканальность, обработка больших массивов данных, за время, равное времени распространения света в оптической системе от модулятора до фотоприемника. Наряду с достоинствами этот метод имеет и недостатки: необходимости изготовления голографических согласованных фильтров (ГСФ). Юстировка этих фильтров в частной плоскости оптической системы должна проводиться с очень высокой точностью, которая предъявляется также и всей оптической системе.

#### Создание согласованного голографического фильтра (СГФ)

Рассмотрим вопрос о записи согласованных голографических фильтров. Если полезный известной формы обозначить через  $S(x)$ , а помеху через  $n(x)$ , то оптический сигнал на входном транспаранте когерентной оптической системы представляется в виде суммы

$$f(x) = S(x) + n(x) \quad (16.22)$$

спектральной плоскости спектр  $f(x)$  будет состоять из спектров сигнала и шума

$$F(\omega_{x2}) = S(\omega_{x2}) + N(\omega_{x2})$$

Если в спектральной плоскости поставить пространственный фильтр с характеристикой пропускания

$$H(\omega_{x2}) = S^*(\omega_{x2}), \quad (16.23)$$

то суммированный сигнал на выходе фильтра будет равен

$$S(\omega_{x2}) \cdot S^*(\omega_{x2}) = N_2(\omega_{x2}) \cdot S^*(\omega_{x2}) \quad (16.24)$$

Если теперь произвести над этим сигналом обратное преобразование Фурье, то в фокальной плоскости второй линзы образуется автокорреляционная функция сигнала и корреляционная функция сигнала и шума

$$E_3(x_3, y_3) = \int_{-\infty}^{\infty} S(x_1) \cdot S^*(x_1 - x_3) dx + \int_{-\infty}^{\infty} n(x_1) \cdot S^*(x_1 - x_3) dx_1 \quad (16.25)$$

Поскольку сигнал и шум не коррелированы, то второе слагаемое будет мало по сравнению с первым и сигнал будет выделен из шума.

Рассмотрим, как создается комплексный оптимальный фильтр оптическими методами. В частности на принципах голографии (см. рис. 16.13). Эталонный  $S(x)$  сигнал устанавливается в виде транспаранта в предметной плоскости  $P_1$ . Линза 5 осуществляет прямое преобразование Фурье этого сигнала. С помощью призмы 3 и поворотного полупрозрачного зеркала формируются опорный и сигнальный лучи, которые пересекаются в светочувствительной среде под углом  $\theta$ . Запись голограммы Фурье-образа с помощью лазера (1). Равномерная засветка транспаранта осуществляется и помощью коллиматора 4. Светочувствительная среда регистрирует интенсивность интерференционной картины

$$I(\omega) = [S(\omega) + e^{j\omega x_0}]^2 = 1 + S^2 + S^* \cdot e^{j\omega x_0} + S \cdot e^{-j\omega x_0} \quad (16.26)$$

Далее, устанавливая голограмму фильтра в оптическую систему, осуществляют процесс распознавания образов.

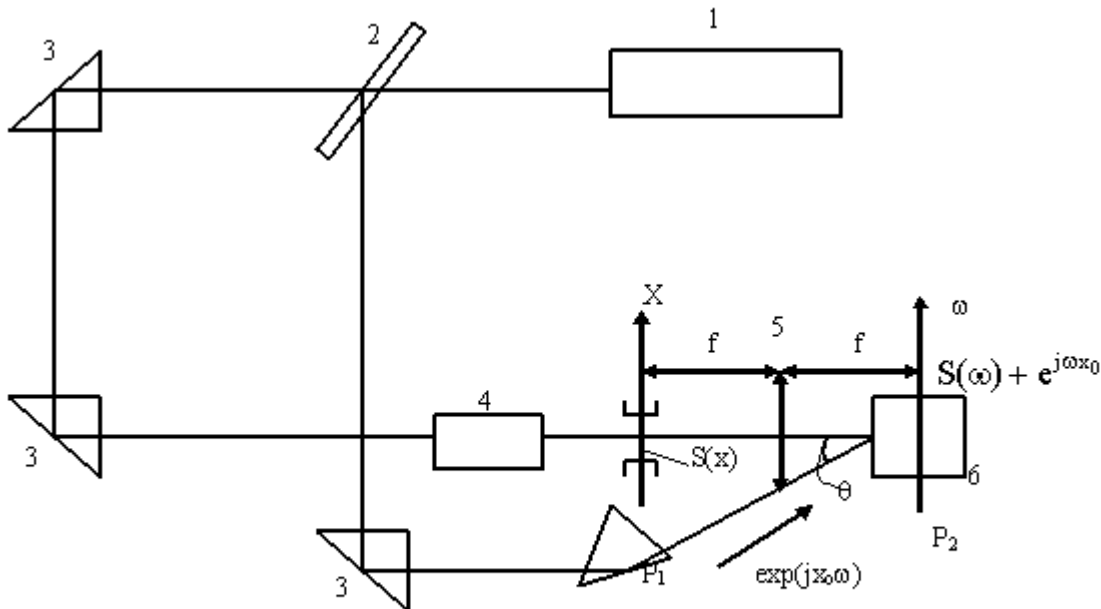


Рис. 16.13 - Схема записи голографического согласованного фильтра

**Фильтрация в оптическом корреляторе.** Когерентные оптические корреляторы можно синтезировать как в предметной, так и в частотной областях. При обработке двумерных функций удобнее использовать корреляторы, синтезируемые в частотной области, поскольку они не требуют перемещений анализируемых функций относительно друг друга. Рассмотрим именно эти корреляторы, реализуемые на базе схемы пространственной помощи фильтрации с голографических согласованных фильтров.

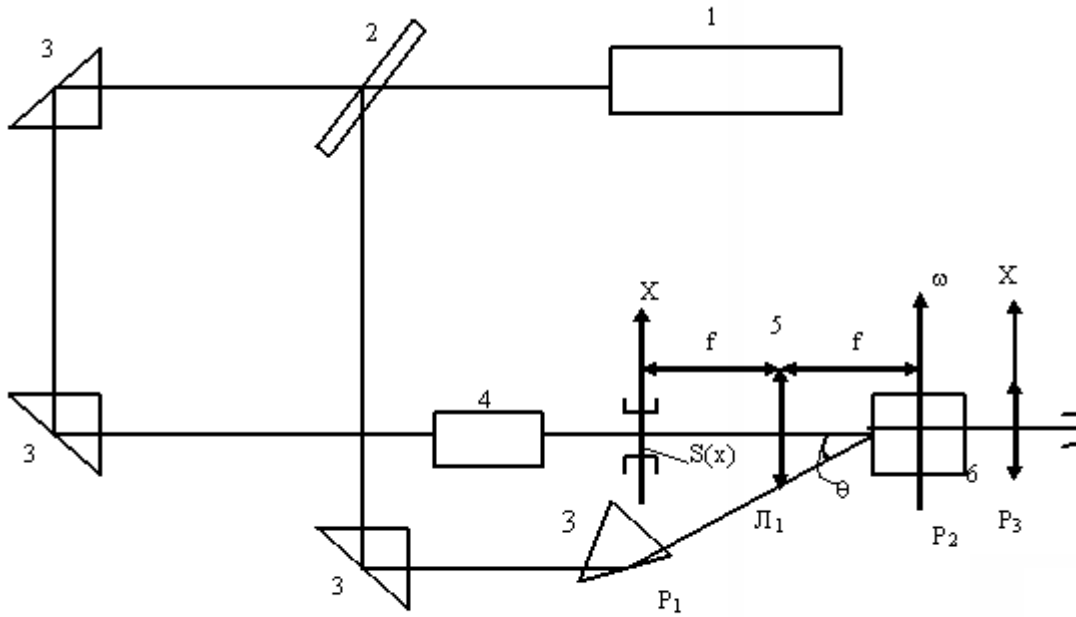


Рис. 16. 14. Схема фильтрации с помощью голографического фильтра

опустим теперь через согласованный фильтр (б) световой сигнал, имеющий образ входного сигнала  $f(x)$ , т.е.  $F(\omega)$  рис. 16.14. В плоскости  $P_1$  устанавливается транспарант анализируемого сигнала. В плоскости  $P_2$  - СГФ. Линза  $\Pi_2$  осуществляет обратное преобразование Фурье. На выходе фильтра, возникает световой сигнал с амплитудой

$$H(\omega) = F(\omega) \cdot I(\omega) = F[1 + S^2] + F \cdot S^* \cdot e^{j\omega x_0} + F \cdot S \cdot e^{-j\omega x_0} \quad (16.27)$$

Этот сигнал имеет три составляющих. Основной пучок, отображаемый первым членом в выражении (16.27), распространяется в направлении оптической оси. Второй пучок, описываемый членом  $F \cdot S^* \cdot \exp(j \cdot \omega \cdot x_0)$  идет в направлении опорного пучка. Третий член  $F \cdot S \cdot \exp(-j \cdot \omega \cdot x_0)$  описывает пучок, отклоняющийся относительно оси в противоположном направлении. Сигнал, передаваемый опорным пучком

$$F \cdot S^* = (S + N) \cdot S^* = S^2 + N \cdot S^* \quad (16.28)$$

состоит из двух слагаемых. В результате преобразования Фурье, функция  $S^2$  превратится в сигнал  $S \cdot S^*$ , т.е. в сигнал автокорреляции функции  $S(x)$ . Этот сигнал появится в центре верхнего бокового изображения, в виде яркой точки. Первое слагаемое нижнего бокового пучка

$$F \cdot S = (S^* + N) \cdot S = S \cdot S^* + N \cdot S \quad (16.29)$$

превратится в сигнал автосвертки  $S \otimes S$ . Сигналы  $N \cdot S^*$  и  $N \cdot S$  вследствие случайной природы шума и определенного рассогласования в фильтре, изготовленном для обнаружения только полезного сигнала  $S(x)$  значения свертки  $n \otimes S$  кросс корреляции  $N \cdot S^*$  будут малы.

Метод согласованной пространственной фильтрации при решении задач контроля безусловно уступает интерферометрии и по чувствительности и по точности, однако он не требует сложной расшифровки интерферограмм — является интегральным методом и легко поддается автоматизации.

Достоинством оптических методов измерений является простота, экспрессность измерений и легкость автоматизации, что весьма существенно при внедрении этих методов в промышленность. Обеспечиваемая при этом точность вполне достаточна для практики. Оптические методы измерений и контроля особенно эффективны, когда необходимо дать интегральную оценку качества или нельзя использовать контактные методы измерений.

Согласованный пространственный фильтр осуществляет параллельный корреляционный анализ входного изображения. Результат анализа может быть представлен в виде световой картины, т. е. поля функций корреляции входного и эталонного изображений. Схема согласованной пространственной фильтрации не чувствительна к положению опознаваемого объекта в пределах анализируемого изображения, однако допустимые изменения его масштаба и угловой ориентации не должны превышать соответственно  $\pm(5\div 10)\%$  и  $2\div 5^\circ$  по отношению к эталону в зависимости от структурной сложности объекта. Отношение сигнал/шум на выходе фильтра колеблется от нескольких десятков до нескольких сотен в зависимости от структуры объекта и окружающего его фона. Неинвариантность схемы фильтрации к масштабу и ориентации объекта может быть устранена путем использования многоканального фильтра, либо осуществлением поиска по масштабу и ориентации. В последнем случае увеличивается время анализа. Применения согласованной пространственной фильтрации связаны как с формированием меры близости в виде функции взаимной корреляции при решении задач, связанных с необходимостью выполнения операций обнаружения или идентификации, изображений, так и с необходимостью измерения интенсивности функции взаимной корреляции при решении задач измерения или контроля. С помощью согласованной пространственной фильтрации могут быть решены следующие задачи:

- обнаружение объектов на некотором сложном фоне;
- навигационная привязка летательных аппаратов по наземным ориентирам;
- исследование скорости движения облаков по фотоснимкам, получаемым с метеорологических спутников;
- автоматический дактилоскопический поиск (опознавание полных отпечатков пальцев);
- автоматическая идентификация ИК спектров веществ;
- ввод буквенно-цифровой информации в ЦВМ;
- интегральный контроль качества печатных плат и интегральных схем;
- интегральный контроль предела усталости лопаток турбин;
- идентификация и распознавание электрических сигналов;
- оптимальная фильтрация электрических сигналов;

В оптике могут быть реализованы не только корреляционные методы распознавания. Весьма вероятно, что распознающий автомат будущего будет гибридной

оптико-электронной системой, в которой операции по анализу изображений будут осуществляться оптическими методами, а алгоритмическая обработка — цифровыми. Сущность метода контроля качества печатных плат состоит в сравнении рисунка печатного монтажа до и после температурных воздействий. С этой целью на рисунок печатного монтажа платы в нормальных условиях изготавливают ГСФ и измеряют интенсивность корреляционного пятна на выходе согласованного фильтра.

Затем плату последовательно нагревают и охлаждают до температур, оговариваемых в технических условиях, и при нормальной температуре опять устанавливают в схему согласованной фильтрации. Из-за неодинаковых температурных коэффициентов расширения печатных

проводников и материала платы возникают остаточные деформации,

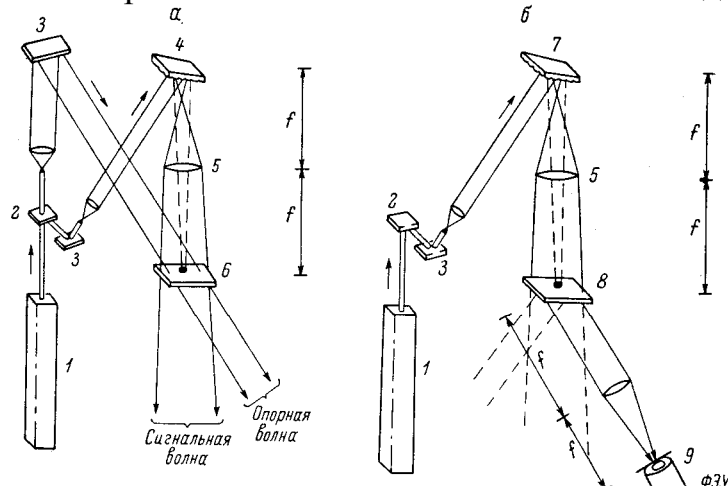


Рис.5.15. Схема получения ГСФ (а) и схема измерений (б).

1 – лазер, 2 – полупрозрачные зеркала, 4 – испытуемая плата, 5 – линза, 6 – фотопластинка, 7 – плата после температурного воздействия, 8 – фурье-голограмма, 9 – ФЭУ.

которые изменяют положение печатных проводников на плате. В результате смещения проводников интенсивность корреляционного пятна будет изменяться в зависимости от величины смещения (деформации). Измеряя интенсивность корреляционного пятна до и после температурных воздействий, можно оценить величину возникших при этом остаточных деформаций и осуществить технологический контроль качества печатных плат. На рис. 16.15 представлена схема получения ГСФ непосредственно с поверхности платы, т. е. в отраженном свете, а не в проходящем, как это обычно делается, и схема обработки. Зависимость интенсивности корреляционного пятна от величины деформации иллюстрирует рис. 15.16. Аналогичным образом можно обнаруживать начинающиеся усталостные разрушения в механических деталях.

Метод согласованной пространственной фильтрации при решении задач контроля безусловно уступает интерферометрии и по чувствительности и по точности, однако он не требует сложной расшифров-

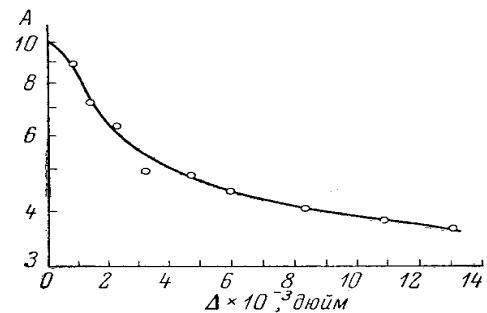


Рис. 16.16. Зависимость амплитуды корреляции A от величины смещения печатных проводников

ки интерферограмм — является интегральным методом и легко поддается автоматизации.

Достоинством оптических методов измерений является простота, легкость измерений и автоматизации, что весьма существенно при внедрении этих методов в промышленность. Обеспечиваемая при этом точность вполне достаточна для практики. Оптические методы измерений и контроля особенно эффективны, когда необходимо дать интегральную оценку качества или нельзя использовать контактные методы измерений.

### 16.2 5. Распознавание образов

С помощью СГФ можно осуществлять распознавание образов для различных конкретных применений [45-48]. Так, например, если изготовить транспарант с буквой Т и записать Его Фурье-образ в виде голограммы по оптической схеме, изображённой на рис. 16.14. После этого поместить в плоскость R транспарант с набором разных фигур, включая и букву Т. То при освещении когерентным пучком лазера в боковых изображениях должны появиться точки, в местах обнаружения буквы Т.

Рассмотрим этот случай, т.е. что произойдет, если анализируемый сигнал сместить в другую точку в плоскости изображений. Такое смещение можно описать сигналом вида

$$S(x_1) = S_1(x + x_1)$$

Его Фурье -образ равен

$$S_1(\omega) = \int S(x_1 + x) \cdot e^{-j\omega \cdot x_0} dx = S(\omega) \cdot e^{j\omega \cdot x_1}$$

и отличается от Фурье-образа  $S(a)$  исходного, несмещенного сигнала множителем  $\exp(j\omega x_1)$  который, приводит только к дополнительному наклону каждого из боковых пучков.

В результате фильтрации на выходе оптической системы получается сигнал

$$S_1(\omega) \cdot S^*(\omega) \cdot e^{j\omega \cdot x_0} = S^* \cdot S \cdot e^{j\omega \cdot (x_0 + x_1)}. \quad (16.30)$$

Поворот оси распространения сигнала в плоскости Фурье-координат вызывает смещение точки, вокруг которой располагается отфильтрованный сигнал. Таким образом, поступательное перемещение сигнала в плоскости  $P_1$  приводит к такому же поступательному перемещению сигнала свертки  $S \otimes S$  или корреляции во второй предметной плоскости  $P_2$ .

При повороте объекта в плоскости изображений на угол  $\varphi_1$  сигнал автокорреляции, например, прямоугольника» монотонно уменьшается и в минимуме равен (где  $a$  и  $b$  - ширина и высота прямоугольника), т.е. уменьшается в  $a/b$  раза по сравнению с исходным сигналом.

Если изменить масштаб сигнала  $S(x)$ , т.е. взять увеличений) в  $\beta$  - раз сигнал при одной и той же полной интенсивности:



$$S_2(x) = \left[ \frac{1}{\sqrt{\beta}} \right] \cdot S\left(\frac{x}{\beta}\right),$$

то спектр нового сигнала растянется в  $1/\beta$  раз вдоль координаты  $\omega_x$  и  $\omega_y$ , что приводит также к ослаблению сигнала.

Устройства, в которых используются пространственные фильтры, называются оптическими корреляторами. Такие корреляторы позволяют производить распознавание образов, т.е. идентифицировать объекты.

Рассмотренные системы оптической обработки информации являются основой сверхбыстродействующих когерентно – оптических вычислительных машин. Такие машины удобно использовать для параллельной обработки информации, поскольку световой поток, переносящий информацию, имеет двумерную структуру.

## **ГЛАВА 17. РАЗЛИЧНЫЕ ПРИМЕНЕНИЯ ОПТИКО-ЭЛЕКТРОННЫХ КВАНТОВЫХ ПРИБОРОВ И СИСТЕМ**

В настоящее время оптико - электронные квантовые приборы (ОЭКП) широко применяются в различных областях науки и техники, таких как физика, химия, биология, электроника, медицина и бытовая радиоаппаратура. Такое широкое распространение ОЭКП получили благодаря особым свойствам лазерного излучения: высокой монохроматичности, малой расходимости, возможности формирования сверхкоротких импульсов. Прежде всего следует отметить широкие перспективы применения ОЭКП в различных информационных системах: системы локации, связи, навигации, дальнометрии, высотомеры и т.п. Особенностью оптических информационных систем являются их широкополосность и, следовательно, возможность передачи больших объемов информации в короткий промежуток времени. Рабочая полоса канала передачи информации может быть увеличена по сравнению с самой широкополосной системой радиодиапазона в  $10^5 - 10^6$  раз, что делает принципиально возможным создание многоканальных систем, использующих излучение одного лазера, без опасности возникновения при этом взаимных помех.

Благодаря возможности получения направленности излучения, близкой к дифракционному пределу  $\theta_d$  ( $\theta_d = 1,22\lambda/D$ , где  $D$  – диаметр сечения лазерного пучка), мощность излучения, необходимая для передачи одинаковых объемов информации, в оптическом диапазоне во много раз меньше, чем в радиодиапазоне. Малая расходимость излучения лазера позволяет направлять всю энергию излучения в строго заданных направлениях, обеспечивает пространственную скрытность и высокую помехозащищенность систем передачи информации, делает принципиально возможным значительное снижение массо-геометрических параметров прямо передающих устройств. Использование ОЭКП для локации и навигации, за счет отмеченных свойств, значительно повышает точность определения координат объектов по сравнению с локационными системами радиодиапазона.

## 17.1. Применение твердотельных лазеров

**Технология.** Одним из достоинств лазера является его высокая интенсивность излучения, как отмечалось выше. Лазеры, излучающие короткие импульсы (в режиме синхронизации мод или модуляции добротности), позволяют достигать за время длительностью  $10^{-8} \div 10^{-11}$  с, мощности, составляющие  $10^8 \div 10^{10}$  Вт. Такой мощности вполне достаточно для мгновенного нагрева или испарения многих веществ. Благодаря этому лазеры широко используются для резки, сварки, сверления и другой обработки материалов. Для этих целей из твердотельных лазеров используются рубиновые лазеры и лазеры на стекле с неодимом. Плотность мощности лазерного излучения увеличивают с помощью оптических систем, за счет фокусировки света. Лазерный луч не является строго направленным, а обладает конечной угловой расходимостью  $\theta$ , которая обычно составляет  $n(1 \div 10)$  угловых минут. Поэтому в фокальной плоскости линзы получаем не точку, а пятно с радиусом  $r$  (см. рис. 3.40) равным

$$r = f \frac{\theta}{2},$$

где  $f$  – фокусное расстояние линзы.

Минимальная площадь поперечного сечения лазерного луча равна при этом

$$S = \frac{\pi}{4} f^2 \theta^2$$

Тогда, зная энергию  $E_r$ , излучаемую лазером за время действия импульса генерации  $\tau$ , плотность мощности в пределах фокального пятна можно определить по следующей формуле:

$$P = \frac{4E_r}{\pi f^2 \theta^2 \tau}$$

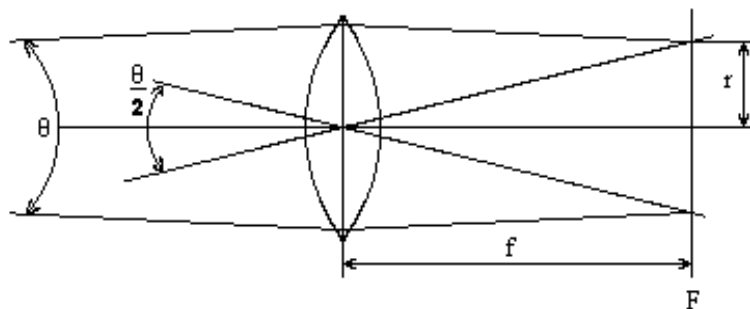


Рис.17.1. Фокусировка лазерного излучения конечной расходимости

При фокусировке излучения плотность мощности повышается в тысячи раз по сравнению с не фокусированным светом. При этом  $P$  может достигать величин

$10^{15} \div 10^{18}$  Вт/см<sup>2</sup>. Следует учитывать, что только ~ 1% этой мощности поглощается материалом и идет на нагрев. Остальная часть мощности отражается от обрабатываемой поверхности.

Возможность получения столь высоких плотностей мощности излучения послужила основанием для создания ряда технических установок для термообработки, сварки, обработки тонких пленок, получения отверстий в различных материалах, в том числе и сверхтвердых – алмазе, керамике и т. д.

Промышленные лазерные установки могут пробивать отверстия и производить их обработку в различных металлах, сплавах и неметаллических материалах корунде, ситаллах, а также в полупроводниках. При этом глубина обрабатываемых отверстий может достигать 4 мм, точность обработки не хуже  $\pm 0,01$  мм. Так для пробивки отверстия диаметром 1,44 мм требуется 550 импульсов излучения рубинового лазера с энергией 2-3 Дж и длительностью импульса 0,6 мс.

При использовании лазеров с энергией импульсов излучения порядка нескольких джоулей может производиться сварка таких материалов, как молибден, никель, тантал, кобальт, медь, золото при толщине свариваемых полос порядка 0,5 – 0,8 мм, с помощью лазеров можно сваривать стекло, кварц, полупроводниковые материалы.

Импульсные лазерные установки используются для сварки элементов миниатюрных устройств, соединений электродов транзисторов, монтажа электронных блоков, корпусов интегральных схем.

**Системы подводного наблюдения.** Весьма важной задачей, стоящей на пути освоения Мирового океана с его бесчисленными ресурсами, является создание систем наблюдения и связей под водой.

Появление лазеров, работающих в зеленой области спектра (вторая гармоника излучения лазера на аллюминиево иттриевом гранате активированном неодимом), излучение которых проходит через морскую воду с минимальными поглощениями, привело к разработке и созданию лазерно-телевизионных систем подводной локации и обнаружения (обзора дна, поиска затонувших кораблей, обнаружения мин и т.д.).

**Оптические локаторы.** Вследствие малой расходимости луча лазера поиск цели затруднен, однако оптические дальномеры можно успешно использовать в космических и авиационных системах совместно с приборами сантиметрового и миллиметрового диапазонов.

Дальность действия локатора можно определить как

$$R = \sqrt{\frac{A_0 P_{пер} D_{пр}^2 N_{пр} K}{4\pi\theta_{пер}^2 \frac{S}{N} P_{пр}}}$$

где  $A_0$  – поперечное сечение цели;  
 $P_{пер}$  – мощность лазерного передатчика;  
 $D_{пр}$  – диаметр приемной оптики;  
 $N_{пр}$  – эффективность приемной оптики;  
 $K$  – отражающая поверхность цели;

$\theta_{\text{пер}}$  – расходимость луча локатора;

$S/N$  – отношение сигнал/шум;

$P_{\text{пр}}$  – предельная чувствительность приемника.

Для обеспечения точного определения дальности, излучение лазера модулируется по синусоидальному закону с помощью электрооптического модулятора.

Чувствительность приемной системы локатора в дневное время примерно  $3 \cdot 10^{-10}$  Вт/м<sup>2</sup> и  $3 \cdot 10^{-14}$  Вт/м<sup>2</sup> в ночное время. При этом отношении сигнал/шум в полосе сервосистемы равно 40:1.

Основным преимуществом локаторов оптического диапазона перед радиолокаторами является возможность за счет выходной направленности излучения резко увеличить разрешающую способность по углам и дальности и обеспечить высокую точность измерения. С помощью оптического локатора можно очень точно измерить скорость движущегося объекта  $v$ . Согласно эффекту Доплера, относительное изменение частоты сигнала, отраженного от такого объекта, равно

$$\frac{\Delta \nu}{\nu} = \frac{v}{c}.$$

Чем больше значение  $v$ , тем больше абсолютный частотный сдвиг  $\Delta \nu$  и тем точнее можно определить скорость.

## 17.2. Применение газовых лазеров

Различные виды газовых лазеров имеют существенные отличия между собой по техническим и эксплуатационным параметрам и характеристикам. Это обуславливает различия в областях их применения.

Высокая степень монохроматичности, стабильности пространственно-временной когерентности позволяет успешно использовать гелий-неоновые лазеры в системах связи, интерферометрии, метрологии, спектроскопии.

Например, создание приборов со стабильностью  $10^{-12}$  и выше открывает возможность создания на их основе эталонов длины. Высокая направленность излучения обеспечивает эффективность их использования в строительстве при нивелировании и разбивочных работах, при прокладке железных дорог, тоннелей шахт.

Ионные лазеры, имеющие достаточно высокую мощность излучения (единицы – десятки Вт) с длиной волны 0,5 мкм находят применение в подводной локации, телевидении и связи, в системах аэрофоторазведки, в приборах ночного видения, для подсветки объектов и целей, в цветном телевидении, медико-биологических исследованиях, используются также в голографических приборах, в системах записи и считывания информации, в приборах спектрального анализа.

Молекулярные лазеры характеризуются большим разнообразием разработанных и созданных типов лазеров. Они излучают в широком спектральном диапазоне длин волн от 0,15 до 774 мкм, имеют большие мощности излучения.

Высокие мощности непрерывной генерации, полученные в газоразрядных молекулярных лазерах на длине волны 10,6 мкм, позволяет с успехом применять их в ряде технологических установок для резки, сверления стекла, металлов, в медицинских и биологических устройствах.

Повышение частотной стабильности  $\text{CO}_2$ -лазеров делает их пригодными для систем дальней связи, особенно космической.

Области применения сверхмощных молекулярных лазеров (газодинамических и других) определяются их термическим действием. Они могут быть использованы в системах передачи энергии на расстояния, например, обеспечить питание ИСЗ (искусственный спутник земли) с земли, в системах связи в атмосфере и космосе.

Мощность, расходимость и ширина спектральной линии излучения существующих лазеров обеспечивают возможность передачи информации на расстояние 160 млн. км (до планеты Марс) со скоростью 1 млн. бит/с, 40 тысяч бит/с на расстояние 2700 млн. км (до планеты Юпитер) и 1000 бит/с на расстояние 6400 млн. км (до Плутона). Для сравнения отметим, что при использовании обычной радиолинии скорость передачи информации с корабля “Маринер” составляла 8 бит/с.

Газовые лазеры можно использовать в качестве маяков для визуальной ориентировки космонавтов.

Широкие возможности открывает использование газовых лазеров в системах локации, авиационной разведки, в лазерных телевизионных системах.

Основное преимущество телевизионных систем с использованием лазеров заключается в возможности получения изображения высокого качества за счет повышения разрешающей способности и улучшения отношения сигнал/шум.

Широкие перспективы открываются перед газовыми лазерами в медицине, главным образом в хирургии. Однако развиваются также и некоторые диагностические методы с помощью лазера.

В хирургии сфокусированный лазерный пучок (чаще  $\text{CO}_2$ -лазер) используется вместо обычного скальпеля. Инфракрасное излучение  $\text{CO}_2$ -лазера сильно поглощается молекулами воды в ткани, что приводит к интенсивному испарению этих молекул с последующим рассечением тканей. Принципиальные преимущества “светового скальпеля” состоят в следующем.

1. Разрез может быть произведен с высокой точностью.
2. Возможность проводить операции на недоступных участках.
3. Резкое уменьшение потери крови вследствие прижигающего действия лазерного света на кровеносные сосуды (вплоть до сосудов диаметром  $\sim 0,5$  мм).
4. Ограниченное проникновение в соседние ткани (в пределах десятков микрометров).

В офтальмологии лазеры уже давно применяются для лечения отслоения сетчатки (обычно  $\text{Ar}^+$ -лазер). Зеленое излучение сильно поглощается красными кровяными тельцами, и последующее тепловое воздействие может привести к “привариванию” сетчатки или коагуляции её сосудов. Все большее применение находят лазеры в отоларингии, потому что она имеет дело с такими органами, как трахея, глотка, среднее ухо, недоступность которых затрудняет операцию. Лазер также продемонстрировал свою полезность при лечении сильно кровоточащих поражений в желудочно-кишечном тракте. При этом лазерный пучок (от Nd: YAG или  $\text{Ar}^+$ -лазера) направляется на залечиваемый участок при помощи специального световода. Перспективным является применение лазеров в гинекологии, дерматологии (для удаления татуировки и для лечения некоторых сосудистых заболеваний) и т. д.

Развитие лазерной техники и голографии открыло возможность создания нового класса запоминающих устройств, но своей эффективности значительно превосходящих существующие. Предложенные в 1966 году голографические системы памяти позволяют довести плотность записи информации до  $10^8 - 10^9$  бит/см<sup>2</sup>, что на несколько порядков превышает значения, достигнутые в запоминающих устройствах на магнитных лентах, дисках и сердечниках. Одновременно с резким увеличением плотности записи информации обеспечивается и малое время поиска и выборки информации, не превышающие долей или единиц микросекунд.

### 17.3. Области применения полупроводниковых источников излучения

Основные достоинства инжекционных полупроводниковых источников света, такие как: малогабаритность, простота методов накачки и модуляции – позволяют использовать их в устройствах и приборах различного назначения: в дальномерах, световодных и атмосферных линиях связи, в системах подсвета приборов ночного видения и видения сквозь дождь и туман, в бытовой радиоаппаратуре и др.

Дальнометрия. Одной из первых областей применения полупроводниковых квантовых генераторов (ПКГ) является дальнометрия. Структурная схема одного из дальнометров на ПКГ приведена на рис. 17.2.

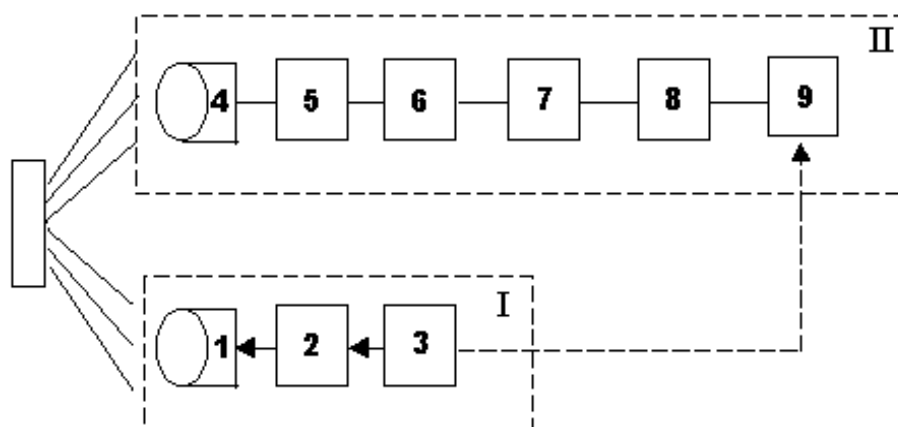


Рис. 17.2. Структурная схема тактического дальномера на ПКГ: I – передающее устройство; II – приемное устройство; 1- коллимирующая система передатчика; 2 – ПКГ; 3 – система накачки; 4 – приемная оптическая антенна; 5 – интерференционный входной фильтр; 6 – фотоприемник; 7 – предусилитель; 8 – пороговое устройство; 9 – счетчик дальности

Отсчет дальности производится подсчетом количества периодов синусоидального напряжения известной частоты (15МГц) за период времени между зондирующим и отраженным сигналами.

В качестве излучателя используется ПКГ с импульсной мощностью 100 Вт. Его масса с батареями питания составляет 4 кг. Расходимость пучка излучения, при использовании коллимирующей оптики, составляет 5 мрад.. При увеличении импульсной мощности до 1кВт, дальность действия может быть более 1 км.

**Световодные линии связи.** В последнее время во многих крупных городах ощущается достаточно заметный дефицит пропускной способности действующих линий связи. Дальнейшее их расширение связано со значительными трудностями из-за перегруженности городской территории подземными коммуникациями. Аналогичная картина перегруженности наблюдается и в системах межгородской и межобластной связи.

Поэтому в последнее время усилился интерес к световодным линиям связи, надежность работы которых не зависит от погодных условий и которые отличаются пропускной способностью, на несколько порядков превышающей пропускную способность кабельных линий связи.

Для линий световодной связи протяженностью до 1000 м, работающих со средними скоростями передачи информации (линии 50 Мбит/с) в качестве источников излучения применяются световоды (СИД), работающие в диапазоне длин волн до 0,9 мкм.

В линиях световодной связи средней протяженности (1 – 10км) применяют многомодовые градиентные волокна с шириной полосы до 200÷1000 МГц/км, потерями 2÷5 дБ/км при  $\lambda=0,85$  мкм и 1÷2 дБ/км при  $\lambda=1,3$  и 1,55 мкм. Значения скорости передачи в таких системах находятся в пределах 30÷100 Мбит/с, но могут достигать 300 Мбит/с и более. Источниками излучения служат ПКГ.

В линиях большой протяженности (10 – 100 км и более) для наземной связи, где скорости передачи информации 200 – 400 Мбит/с, источниками излучения служат только одномодовые ПКГ ( $\lambda=1,3$  мкм;  $\lambda=1,55$  мкм). Информация в таких системах осуществляется только в дискретной (цифровой) форме.

**Переговорные устройства.** Под таким названием известны системы одноканальной связи на основе ПКГ, предназначенные для ведения переговоров в условиях оптической видимости. Передающее устройство построено на основе лазерного диода, мощность излучения которого равна 5 Вт при частоте следования импульсов 8 – 10 кГц. Полоса модулирующих частот 300 – 3400 Гц. Для коллимации пучка излучения до расходимости 10 – 60' используется объектив диаметром 40мм. Мощность, потребляемая приемопередающим блоком, не превышает 2 Вт.

Самолетные системы предупреждения столкновений в воздухе, строятся на основе ПКГ с мощностью излучения 100 Вт, работающих с частотой 1000Гц. Дальность действия таких систем в обычную погоду 3 км.

#### **17. 4. Акусто – оптические системы обработки радиосигналов**

Спектральный анализ является одним из основных методов исследования радиосигналов. Важной характеристикой систем, осуществляющих спектральный анализ, является возможность работы в реальном масштабе времени. При этом

последовательность спектров может вычисляться одновременно по перекрывающимся во времени сигналам. Такой режим работы спектроанализатора отличается от избирательного, когда для всей полосы анализируемых частот применяется один перестраиваемый узкополосный фильтр. При исследовании сигналов в реальном масштабе времени предполагается, что анализ спектра выполняется за время, меньшее, чем необходимо для заполнения всей апертуры устройства ввода входными данными.

Сейчас разработаны эффективные цифровые алгоритмы, значительно снижающие трудоемкость спектрального анализа. Спектральный анализ в реальном масштабе времени для сигналов с шириной спектра менее 2МГц выполняется, как правило, с помощью специализированных цифровых устройств. При обработке сигналов с шириной спектра более 10МГц значительно возрастает объем и стоимость цифровой аппаратуры, поэтому перспективным является применение акустооптических процессоров.

Разработаны акустооптические анализаторы спектра на объемных акустических волнах со следующими параметрами: ширина полосы анализа 1ГГц, разрешающая способность 1МГц, объем аппаратуры с источником питания менее 0,003 м<sup>3</sup>. Существует акустооптический приемник радиосигналов со средней частотой 2,8ГГц и шириной полосы анализа 884МГц. Приемник имеет время когерентного накопления 0,5..50 мс, точность измерения частоты  $\pm 250$ кГц.

Акустооптические спектроанализаторы в зависимости от переменной интегрирования при выполнении преобразования Фурье могут быть разделены на две группы устройств: с пространственным интегрированием, когда преобразование осуществляется по пространственным переменным, и с временным интегрированием, когда преобразование выполняется по временной переменной. В каждой из этих групп можно организовать как одномерное, так и двумерное преобразование Фурье [47-48]. Рассмотрим схемы спектроанализаторов с пространственным интегрированием.

#### 17.4.1. Классический акустооптический спектроанализатор

Рассмотрим классический акустооптический спектроанализатор [46], показанный на рис. 17.1. Во входной плоскости  $P_1$  установлен пространственный модулятор 2, освещаемый сферической волной от точечного источника света 1 с радиусом кривизны  $d_1$ . Преобразующая линза 3 с фокусным расстоянием  $F_1$  размещается в плоскости  $P_2$  на расстоянии  $d_2$  от входной плоскости  $P_1$ . Выходная плоскость 4, в которой помещается фотоприемник, расположена на расстоянии  $d_3$  от линзы.

Выражение для распределения амплитуд поля в плоскости наблюдения  $P_3$  можно записать в виде:

$$E(x_3, y_3) = (iD_2D_3/(\lambda_{св}(D_2 + F_1^{-1} + D_3))) \cdot \Phi(x_3, y_3, D_3 - (D_3^2/(D_2 - F_1^{-1} + D_3))) \cdot S(x_3, y_3), \quad (17.1)$$

где



$$S(x_3, y_3) = \int_p T(x_1, y_1) \cdot \Phi \left( x_1, y_1, D_1 + D_2 - \frac{D_2^2}{D_2 - F_1^{-1} + D_3} \right) \exp \left[ -jv \frac{D_2 D_3}{D_2 - F_1^{-1} + D_3} (x_3 x_1 + y_3 y_1) \right] dx_1 dy_1;$$

$$\Phi(x_i, y_i; D_i) = \exp \left[ j \frac{2\pi}{\lambda_{ce}} \frac{D_i}{2} (x_i^2 + y_i^2) \right];$$

$T(x_1, y_1)$  – функция пропускания АОМ,  $D_i = 1/d_i$ ;  $i = 1, 2, 3$ .

Как видно из (17.1), модуль распределения амплитуд поля на выходе системы пропорционален модулю Фурье-спектра  $S(\omega_x, \omega_y)$  функции пропускания  $T(x_1, y_1)$ . Если линзой выполняется условие преобразования Фурье  $D_1 + D_2 - (D_2^2 / (D_2 - F_1^{-1} + D_3))$  (17.2)

то в этом случае получаем

$$\dot{S}(x_3, y_3) = \int_{p_1} \dot{T}(x_1, y_1) \exp \left[ -i \frac{2\pi}{\lambda_{ce}} \cdot \frac{D_2 D_3}{D_2 - F_1^{-1} + D_3} (x_3 x_1 + y_3 y_1) \right] dx_1 dy_1 = S(\omega_x, \omega_y)$$

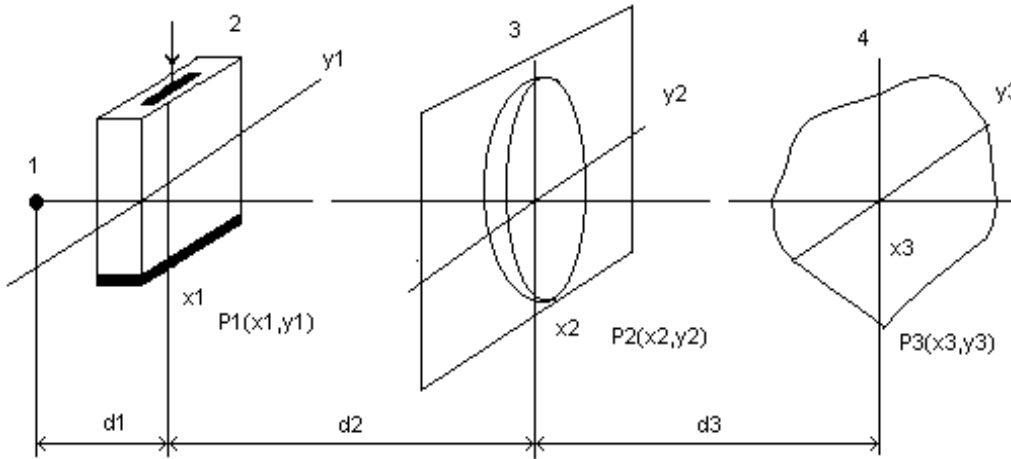


Рис. 17.3 Классический акустооптический спектроанализатор.

Пространственные угловые частоты связаны с координатами плоскости  $P_3$  соотношениями:

$$\omega_x = \frac{2\pi}{\lambda_{ce}} \cdot \frac{D_2 D_3}{D_2 - F_1^{-1} + D_3} x_3; \omega_y = \frac{2\pi}{\lambda_{ce}} \cdot \frac{D_2 D_3}{D_2 - F_1^{-1} + D_3} y_3$$

Из (17.2) следует, что при освещении АОМ плоской волной условием формирования Фурье-образа является условие  $D_3 = F_1^{-1}$ , т.е. плоскость наблюдения  $P_3$  должна находиться в задней фокальной плоскости линзы. Если АОМ освещается неплоским пучком, то это приводит к сдвигу плоскости, в которой формируется спектр Фурье.

При освещении модулятора плоской световой волной величина  $D_2$  оказывает влияние только на фазовый множитель. При  $D_2 = F_1^{-1}$  фазовый множитель равен



На рис. 17.5 показан автоколлимационный спектроанализатор [46], в котором, как и в предыдущей схеме, при освещении системы плоской световой волной ( $D_1 = 0$ ) спектр Фурье будет формироваться в задней фокальной плоскости интегрирующей линзы. При этом пространственные частоты  $\omega_x = 2\pi x_3 / (\lambda_{св} d_3)$ ,  $\omega_y = 2\pi y_3 / (\lambda_{св} d_3)$ .

Рассматриваемая схема позволяет осуществлять преобразование Фурье с переменным масштабом, зависящим от отношения  $d_2 / F_1$ . При больших  $d_2 / F_1$  апертура преобразующей линзы должна значительно превосходить апертуру АОМ, поэтому на практике расстояние  $d_2$  должно быть близким к нулю, и основным преимуществом схемы является минимальное влияние конечных размеров апертуры линзы.

### Основные характеристики акустооптического спектроанализатора

Важнейшими характеристиками спектроанализатора является полоса пропускания, разрешающая способность по частоте, ДД обрабатываемых сигналов, быстродействие. Полоса пропускания анализатора определяется полосой пропускания АОМ, которая зависит от геометрии взаимодействия световых и акустических пучков, центральной частоты преобразователя, физических свойств материалов звукопровода пьезопреобразователя и связующих слоев. Существенное влияние на полосу АОМ оказывает также электрические свойства пьезопреобразователя, входное электрическое сопротивление которого в общем случае является комплексной величиной, что требует осуществления электрического согласования с источником сигнала. Результирующая полоса АОМ является комбинацией трех частотных характеристик: акустооптической, акустической и электродинамической.

Разрешающая способность анализатора определяется обычно как наименьший интервал между двумя спектральными линиями равной интенсивности, при котором они еще различаются. Согласно критерию Рэлея две спектральные линии считаются разрешенными, если пик дифракционного максимума от одной составляющей. При этом результирующая освещенность в месте перекрытия составляет около 80% от максимальной. Разрешающая способность акустооптического анализатора света по Рэлею при идеальной оптической системе и отсутствии поглощения акустических волн в материале звукопровода определяется длительностью выборки входного сигнала  $T \Delta f_{\min} = 1/T = v_{зв}/a$ , где  $a$  – апертура АОМ в направлении распространения акустооптической волны.

Для повышения разрешающей способности звукопровод модулятора нужно изготавливать из материала с малой скоростью распространения звуковой волны при максимально достижимой апертуре. Такое требование является противоречивым из-за поглощения акустической энергии в материале звукопровода. Скорость звука и коэффициент поглощения  $\lambda$  связаны соотношением  $\lambda p_{зв}/f^2 \approx v_{зв}^{-5}$ , где  $f$  – частота акустических колебаний.

Таким образом, материал с малой скоростью распространения акустических волн обладает большим поглощением звука, что существенно влияет на характеристики анализатора спектра.

### 17.4.2. Акустооптические процессоры

Частным случаем спектрального анализа является определение взаимной функции неопределенности сигнала, которую можно записать в виде:

$$Q(\omega, t) = \int_{-\infty}^{+\infty} f'(t) \dot{g}^*(t-\tau) \exp(j\omega\tau) d\tau \quad (17.4)$$

Соотношение (17.4) можно рассматривать как спектр произведения двух сигналов, который зависит от их относительной задержки. Функции неопределенности находят широкое применение в радиолокации. Эти функции представляют собой отклик оптимальной линейной системы, осуществляющей обработку радиолокационных сигналов, на входной сигнал на фоне белого гауссовского шума.

Два АОМ позволяют получить произведение функций, стоящих под интегралом в (6.10).

На рис. 17.6 [46] приведена схема оптического процессора с двумя ортогонально расположенными АОМ, осуществляющего операцию временного интегрирования произведения трех функций.

В качестве источника света 1 используется светоизлучающий диод, интенсивность света которого модулируется во времени сигналом  $f(t)$ . Первый АОМ 4, расположенный в горизонтальном направлении, модулируется сигналом  $g(t)$  и освещается светом с помощью линз 2 и 3. В результате интенсивности света, дифрагированного модулятором, оказывается промодулированной произведением сигналов  $f(t)g(t - x / v_{зв})$ . Световой пучок, выходящий из модулятора 4, расширяется в вертикальной плоскости и фокусируется в горизонтальной плоскости с помощью сферической линзы 5 для освещения вертикально расположенного АОМ 6, на вход которого подается сигнал  $h(t)$ . В результате свет, дифрагированный модулятором 6, будет промодулирован произведением сигналов  $f(t)g(t - x / v_{зв}) h(t - y / v_{зв})$ .

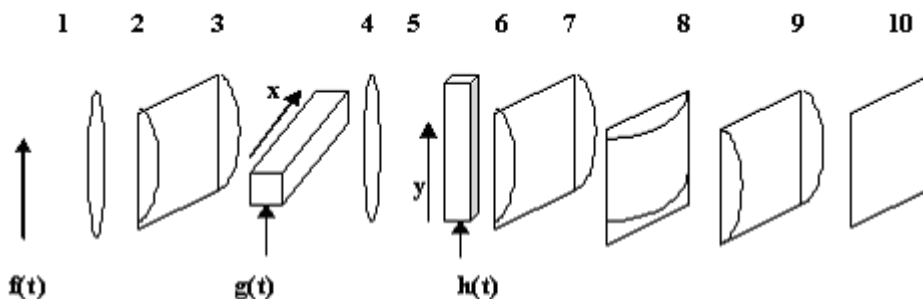


Рис. 17.6 Акустооптический процессор тройного произведения с временным интегрированием - архитектура оптического процессора

Цилиндрические линзы 7 и 9 формируют изображение модулятора 6 в выходной плоскости 10 в вертикальном направлении, а сферическая 5 и цилиндрическая 8 линзы модулятора 4 – в горизонтальном направлении. Как и в когерентном корреляторе, к полученному в выходной плоскости распределению света необходимо добавить опорную световую волну.

Результирующее распределение света детектируется матрицей фотодетекторов, на выходе которой формируется пространственное распределение зарядов, пропорциональное интенсивности света. Величина заряда, накапливаемого в течение времени  $T$  в каждом фотодетекторе с координатами  $x$ , пропорционально функции:

$$Q(x,y) = \int_T f(t) g(t - x / v_{3B}) h(t - y / v_{3B}) dt. \quad (17.5)$$

Так как оптическая система (рис. 17.4) состоит из двух одномерных корреляторов с временным интегрированием, работающих одновременно в двух ортогональных направлениях, то результат обработки получается не простым повторением одномерных преобразований, а при совместной двумерной обработке. Оптическая система с такой архитектурой, называемая процессором тройного произведения, всегда выполняет операцию общего вида (17.5). При этом обеспечивает значительную гибкость, так все три входных сигнала – электрические и путем их изменения можно перестроить процессор для выполнения новой математической операции, используя те же оптические компоненты. Если входные сигналы

$$f(t) = S(t) \exp(jkt^2), \quad g(t) = S(t) \quad \text{и} \quad h(t) = \exp(-jkt^2),$$

где  $k$  – постоянный коэффициент, то:

$$Q(x,y) = \exp[-j k (y/v_{3B})^2] \int S(t) S(t - x / v_{3B}) \exp[-j2 (ky / v_{3B}) t] dt. \quad (17.6)$$

Соотношение (17.6) представляет собой функцию неопределенности сигнала  $S(t)$ , умноженную на квадратичный фазовый множитель, который при необходимости может быть устранен при последующей обработке. В этой же системе может быть выполнен иной вид обработки, если  $f(t) = S(t) \exp(jk_1 t^2 + jk_2 t^2)$ ,  $g(t) = (-jk_1 t^2)$  и  $h(t) = \exp(-jk_2 t^2)$ . Тогда:

$$Q(x,y) = \exp[-j k_1 (x/v_{3B})^2 - j k_2 (y/v_{3B})^2] \int S(t) \exp[-j(2 k_1 (x/v_{3B})) t] \exp[-j(2 k_2 y / v_{3B}) t] dt. \quad (17.7)$$

Выражение (17.7) представляет собой преобразование Фурье сигнала  $S(t)$ , вычисленное по алгоритму ЛЧМ – преобразования по обеим пространственным координатам [46-48], если не учитывать квадратичный фазовый множитель. Спектр сигнала  $S(t)$  отображается по двум координатам в виде функций  $k_1 x / (2\pi v_{3B})$  и  $k_2 y / (2\pi v_{3B})$ , которые являются частотными переменными по осям  $x$  и  $y$ . Максимальные частоты сигнала  $S(t)$  равны соответственно  $k_1 a / (2\pi v_{3B})$  и  $k_2 a / (2\pi v_{3B})$ , где  $a$  – максимальный размер апертуры по осям  $x$  и  $y$  обоих АОМ. Разрешение по частоте равно максимальному значению частоты, деленному на число элементов разрешения по каждой координате двумерной матрицы фотодетекторов. Если  $k_1$  выбрать из условия  $k_1 a / (2\pi v_{3B}) = \Delta f$ , где  $\Delta f$  – полоса пропускания АОМ, то спектр сигнала  $S(t)$  отобразится по оси  $x$  с низким разрешением, Если при этом

$k_2 = k_1/N$ , где  $N$  – число элементов разрешения в фотоприемнике (например,  $N=10^3$ ), то только тысячная часть спектра отображается по оси  $x$  при каждом положении  $x$  и разрешение по  $y$  равно  $\Delta f/10^6$ . Предполагается, что время интегрирования  $T \geq 10^6 / \Delta f$ . Такое сочетание преобразования по оси  $x$  с низким разрешением в широкой полосе частот и преобразование по оси  $y$  с высоким разрешением в узкой полосе представляет собой двумерный «свернутый» спектр, или, что то же самое, двумерный растр записи  $10^6$  отсчетов одномерных Фурье – образов сигнала  $S(t)$ . Формирование одномерного спектра по двум пространственным координатам в описанной архитектуре позволяет отображать до  $10^6$  отсчетов спектральных составляющих, так как двумерная матрица фотодетекторов имеет такой размер.

Процессор, выполняющий операцию произведения функций, можно использовать также для вычисления «свернутой» корреляционной функции одномерных сигналов длительностью, большей времени задержки акустического сигнала в звукопроводе модулятора. Достоинство устройства определения свернутой корреляционной функции заключается в возможности вычислять корреляционную функцию двух сигналов с очень большой длительностью и в то же время вводить большие относительные задержки между сигналами. Одно из этих свойств можно реализовать в одномерном варианте процессора, но одновременно их можно обеспечить только в двумерных оптических системах.

### 17. 4.3. Перспективы развития акустооптических систем

К настоящему времени проектирование акустооптических систем обработки информации прошло законченный цикл. Первые акустооптические системы в основном представляли двумерные процессоры. При двумерной обработке сигналов одновременно выполняется множество операций, что обеспечивает оптическим процессорам очень высокую потенциальную мощность вычислений. Вследствие способности оптических систем производить двумерную обработку стало возможным создание процессоров, приспособленных для обработки изображений, требующей большого числа вычислений.

Одномерные архитектуры появились благодаря разработке эффективных одноканальных акустооптических устройств ввода. В одномерных оптических процессорах акустооптические устройства ввода используются как широкополосные высококачественные преобразователи электрических сигналов в оптический аналог. Широко распространенным процессором такого типа является акустооптический спектроанализатор, построенный на основе акустооптического устройства ввода, линзы, выполняющей одномерное преобразование Фурье, и линейки фотодетекторов.

Отличием нового поколения архитектур двумерных оптических систем является применение в качестве входных преобразователей одномерных акустооптических систем ввода. Классические двумерные процессоры, построенные на двумерных пространственных модуляторах света, осуществляют непосредственную обработку данных с помощью сферических линз, предназначенных для выполнения преобразований Фурье или двумерных линейных операций, инвариантных к сдвигу. Двумерные акустооптические системы с одномерными входными устройствами дают возможность осуществлять более широкий класс линейных операций и обеспечивают большой выбор способов их выполнения. Динамиче-

ские процессы обработки, такие, как временное интегрирование и систолическая обработка, реализуемые с одномерными входными устройствами, в сочетании с высоким быстродействием, достигаемым в двумерных оптических системах, позволяют создавать наиболее производительные и многофункциональные оптические процессоры. Если несколько снизить излишнее быстродействие, присущее таким системам, то можно существенно повысить точность, используя методы двоичного умножения с помощью аналоговых сверток. Это может привести к реализации оптических процессоров, точность которых сравнима с точностью ЭВМ, и тем самым расширить возможности применения оптических вычислений для решения общих проблем, включающих численные решения различных линейных математических задач.

Убедительным аргументом в пользу одномерных устройств является их высокий уровень развития. В нашей стране и за рубежом разработаны многочисленные оптические устройства, предназначенные, как правило, не для оптической обработки информации. Тем не менее, многие из них эффективно используются в качестве входных или выходных преобразователей в одномерных оптических процессорах. Наиболее широко применяются такие элементы, как акустооптические устройства ввода, полупроводниковые источники света (лазерные диоды и светоизлучающие диоды), полупроводниковые фотодетекторы (фотодиоды и приборы с зарядовой связью). Малая потребляемая мощность и низкая стоимость данных устройств делает возможной реализацию оптических процессоров, имеющих преимущество не только по быстродействию, но и по потребляемой мощности, габаритным размерам и стоимости.

Развитие новых архитектур оптических систем зависит в первую очередь от состояния и уровня разработки оптических элементов и устройств. Если будут созданы надежные и эффективные двумерные пространственные модуляторы света, работающие в реальном масштабе времени, то их можно будет использовать в двумерных архитектурных оптических процессорах, в которых одномерные устройства применяются как широкополосные входные преобразователи, а двумерные – в качестве программируемой памяти большой емкости для оптической системы. Разработка двумерных входных устройств, приспособленных к динамической обработке, например многоканальных АОМ, приведет к новым архитектурам эффективных оптических процессоров.

Нелинейные оптические устройства также могут дать новый импульс развитию оптических вычислительных систем. Если будут созданы линейки и матрицы оптических бистабильных устройств, то окажется возможной реализация оптических процессоров, выполняющих не только линейные преобразования, но и операции более широкого класса. Быстродействие таких оптических процессоров должно превысить быстродействие цифровых вычислительных машин, но не потому, что оптические логические элементы обладают большей скоростью, чем электронные, а потому, что оптическому процессору легко придать свойства трехмерной обработки.

## СПИСОК СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ ПАРАМЕТРОВ

БТР	- биполярные транзисторы
ВАХ	- вольт- амперная характеристика
ВЧ	- высокочастотное поле
ГЛПД	- генератор на лавинно-пролетном диоде
ЖКГ	-жидкостный квантовый генератор
ЗС	- замедляющие системы
КПД	— коэффициент полезного действия
КСЧ	- квантовые стандарты частоты
ЛД	-лазерный диод
ЛБВ и ЛОВ	- лампы бегущей и обратной волны
ОС	- обратная связь
п/п	-полупроводник
ПСВЧ	— приборы сверхвысоких частот
ПВД	- пространственно-временные диаграммы
ПТР	- полевые транзисторы
СЗС	— спиральные замедляющие системы
СВЧ	— сверхвысокая частота
СИД	-светодиод
УЛПД	- усилитель на лавинно-пролетном диоде
ЭМВ, ЭМП	- электромагнитные волны и поля
ЭМ	- электромагнитные
ЭДС	- электродвижущая сила
р/л	- радиолокационные

### Латинские обозначения

$A_{ij}, A_1$	- вероятность спонтанного перехода между уровнями $i \rightarrow j$ в единицу времени, вектор-потенциал
$B_{ji}$	-коэффициент Эйнштейна
$B$	- реактивная проводимость
$b_{ij}$	-спектральный коэффициент Эйнштейна
$c$	— скорость распространения света в вакууме ( $c = 3 \cdot 10^8$ м/сек)
$\mathbf{D}$ и $\mathbf{B}$ ( $\bar{D}$ и $\bar{B}$ )	– векторы электрического смещения и магнитной индукции
$D_{пр}$	-диаметр приемной оптики
$d$	-поперечный размер источника излучения
$\Delta E, E_g$	–ширина запрещенной зоны в полупроводнике
$E$ и $H$	- напряженности электрического и магнитного полей
$e$	- заряд электрона
$g(\omega), g_i$	-форма линии, кратность вырождения $i$ -го уровня
$G_0$	- активная проводимость резонатора



- $h$  - постоянная Планка ( $1,05 \cdot 10^{-34}$  Дж/с)  
 $f, \nu, f_{зв}$  - частота, частота акустических колебаний  
 $f(E)$  - распределение Ферми (вероятность нахождения электрона или дырки на уровне энергией  $E$ )  
 $F, F_n, F_p$  - энергия Ферми, квазиуровень Ферми для электронов, для дырок  
 $g$  - фактор спектроскопического расщепления  
 $i_{конв}, i_{навед}, i_{полн}$  - конвекционный и наведенный токи и ток, регистрируемый во внешней цепи схемы  
 $j_{пор}$  - пороговый ток накачки  
 $i, j$  - номера энергетических уровней  
 $I_{max}, I_{min}$  - максимальная и минимальная интенсивности света  
 $I_0, I$  - интенсивность светового потока на входе и выходе системы  
 $\bar{J}_{полн}, \bar{J}_{конв}, \bar{J}_{см}$  - плотности полного, конвекционного токов и тока смещения  
 $k$  - волновое число  
 $K_3, K, K_y$  - коэффициент замедления, постоянная Больцмана ( $1,38 \cdot 10^{-23}$  Дж/град), коэффициент усиления  
 $L, L_n$  - период замедляющей структуры, орбитальный механический момент, длина оптического резонатора, длина активного элемента в оптических квантовых генераторах, длина когерентности; потери, вносимые поглотителем  
 $m$  - полное число фотонов в поле излучения  
 $M$  - коэффициент модуляции электронного потока, полный магнитный момент атома или молекулы  
 $n_{1,2}$  - показатель преломления среды  
 $N, N_a, N_d$  - концентрация электронов в пучке, концентрация акцепторов и доноров в материале,  
 $N_i^e$  - населенность  $i$ -го уровня в условиях термодинамического равновесия,  
 $N$  - полное число частиц в  $1 \text{ см}^3$  вещества  
 $P_{(i)}$  - мощность (индексом обозначена принадлежность мощности к объекту: погл - мощность поглощения; изл- мощность излучения; прох- мощность, проходящая через вещество; с - потери мощности; ср - средняя мощность, зв- мощность звуковой волны)  
 $q$  - электрический заряд, число ступеней в ЗС  
 $Q_0, Q_{вн}$  и  $Q_n$  - собственная, внешняя и нагруженная добротности резонатора  
 $R_{св}$  - сопротивление связи  
 $r$  - сопротивление потерь  
 $r_{ij}$  - электрооптическая константа в анизотропных средах  
 $S$  - спиновый момент  
 $t_{2пр}$  - время двойного прохождения луча в резонаторе  
 $T_1, T_2$  - времена продольной и поперечной релаксаций  
 $T$  - абсолютная температура  
 $\Delta t$  - длительность импульса  
 $U_{\lambda/2}$  - полуволновое напряжение, необходимое для поворота плоскости поляризации входного линейно-поляризованного пучка на  $90^\circ$   
 $V_\phi$  и  $V_g, V_{зм}$  - фазовая и групповая скорости волны, фазовая скорость в ЗС  
 $V$  - объем резонатора

$W_k, W_p$  - кинетическая и потенциальная энергии,  
 $W_{ij}$  - вероятность индуцированного перехода атомов из состояния  $i$  в состояние  $j$  в единицу времени  
 $X$  - параметр группировки  
 $Z_c, Z_H$  - волновое сопротивление, сопротивление нагрузки

### Греческие обозначения

$\alpha_d$  - дифракционный предел  
 $\delta$  - глубина проникновения поля в стенки резонатора  
 $\beta$  - фазовая постоянная волны в линии, магнетон бора  
 $\gamma$  - коэффициент (взаимодействия) эффективности электронного потока  
 $\Gamma_{ik}$  - вероятности тепловых переходов в единицу времени  
 $\varphi_{1,2}; \varphi_m$  - фаза волн, фазовый набег волны на периоде ЗС  
 $\Delta\varphi$  - разность фаз  
 $\Psi$  - угол намотки спирали ЗС  
 $\lambda, \lambda_{zm}$  - длина волны излучения, длина замедленной волны в ЗС  
 $\Delta\lambda$  - интервал длин волн между соседними модами  
 $\Lambda$  - длина звуковой волны  
 $\theta, \theta_{zv}$  - угол пролета электронов, расхожимость акустической волны,  
 $\rho, \rho(\omega)$  - плотность объемного заряда, спектральная плотность энергии поля  
 $\eta, \eta_{вых}$  - коэффициент полезного действия, внутренний квантовый выход излучательной рекомбинации  
 $\eta_{э}, \eta_{кон}$  - электронный КПД, контурный КПД  
 $\mu_{cm}, \mu$  - абсолютная магнитная проницаемость стенок резонатора, подвижность электронов  
 $\mu_{ei}, \mu_{si}$  - орбитальные и спиновые магнитные моменты  
 $\sigma_{cm}$  - удельная проводимость материала стенок резонатора  
 $\tau, \tau_n$  - время пролета электрона и длительность импульса  
 $\tau_{би}, \tau_n$  и  $\tau_p$  - время безызлучательной рекомбинации, время жизни электронов и дырок  
 $v_{zv}$  - скорость звука в среде  
 $\omega_{mn}$  - частота кванта или частота перехода  
 $\Delta\omega$  - расстояние между резонансными частотами резонатора, ширина спектра

### Русские обозначения

$\Gamma$  - постоянная распространения

Остальные обозначения введены по тексту.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

### К РАЗДЕЛУ 1

1. Никольский В.В., Никольская Т. И. Электродинамика и распространение радиоволн: Учеб. пособие для вузов.: – М., Наука, 1989. – 554 с.
2. Федоров Н.Д. Электронные приборы СВЧ и квантовые приборы.- М.: Радио и связь, 1981. – 207с.
3. Березин В.М., Буряк В.С., Гутцайт Э.М., Маринин В.П. Электронные приборы СВЧ.-М.: Высшая школа, 1985. – 296с.
4. Лебедев И. В. Техника и приборы СВЧ. Том 1. Техника СВЧ. Том 2. Приборы СВЧ. Учебник для студ. вузов. – М.: Высшая школа, 1970. – 440с.
5. Григорьев А.Д. Электродинамика и техника СВЧ: Учеб. Для Вузов. М.: - Высшая школа, 1990.- 335с.
6. Пюшнер Г. Нагрев энергией сверхвысоких частот. Пер. с англ./ Под ред. Э.Я. Пастрона. – М.: Энергия, 1968. – 312 с.
7. СВЧ энергетика./ Под ред. Э. Окресса. Пер. с англ. / Под ред. В. Г. Алыбина, Э. Я. Пастрона. т.т 1, 2, 3. – М.: изд. Мир, 1971.
8. Конструкции СВЧ устройств и экранов. /Под ред. А. М. Чернушенко и др. – М.: Радио и связь, 1983. – 400с.
9. Никольский В. В. Теория электромагнитного поля. – М.: Высшая школа, 1978. – 370с.
10. Радионов В. М. Линии передачи и антенны СВЧ. Сб. номограмм. – М.: Сов. радио, 1965. – 200с.
11. Фельдштейн А. Л., Явич Л. Р., Смирнов В. П. Справочник по элементам волноводной техники. – М.: Госэнерго, 1963. –296с.
12. Петров Б.М. Электродинамика и распространение радиоволн: Учебник для вузов. – 2-е изд. испр.- М.: Горячая линия – Телеком, 2003.-558 с.; ил.
13. Хиппель А.Р. Диэлектрики и их применение. –М.–Л.: Сов.радио, 1959.
14. Кацман Ю.А. Приборы СВЧ.-М.: Высшая школа, 1983.- 346с.
15. Тугов Н.М., Глебов Б.А., Чарыков Н.А. Полупроводниковые приборы. М.: Энергоатомиздат. 1992 – 575с.
16. Бонч-Бруевич В.Л., Калашников С.Г. Физика полупроводников. М.: Наука, 1990-685с.
17. Левитский С.М., Кошечкина С. В. Вакуумная и твердотельная электроника СВЧ: Учебное пособие. Высшая школа, 1988.- 272с.
18. Микро-электронные устройства СВЧ под ред. Г.И. Веселова. М.: Высшая школа, 1988.
19. Куц Г.Г., Соколова Ж.М., Шангина Л.И. Приборы и устройства оптического и СВЧ диапазонов Учебное пособие. – Томск: Изд-во НТЛ, 2003. - 560 с.

### К РАЗДЕЛАМ 2-4

1. Басов Н.Г., Прохоров А.М. ЖЭТФ, 1954. —с.27, 431с.
2. Пестов Э.Г., Лапшин Г. М. Квантовая электроника. — М.: Воениздат, 1972.—334 с.
3. Басов Н.Г., Вул Б.М., Попов Ю.М. ЖЭТФ, 1959. —с.37, 587с..

4. Коваленко Е. С. и др. Введение в квантовую электронику. — Томск, ТГУ, 1974. — 445с.
5. Квантовая электроника. Маленькая энциклопедия. / Под ред. М. Е. Жаботинского. — М.: Сов. энциклопедия, 1969.— 432с.
6. Смирнов В.А. Введение в оптическую радиоэлектронику. — М. Сов. радио, 1973. — 203 с.
7. Тарасов Л. В. Физические основы квантовой электроники (оптический диапазон). — М.: Сов. радио, 1976.—368 с.
8. Бирнбаум Дж. Оптические квантовые генераторы: Пер. с англ./ Под ред. Ф. С. Файзулова. — М.: Сов. радио, 1967.—359 с.
9. Пахомов И.И., Рожков О.В. и др. Оптико-электронные квантовые приборы: Учебное пособие для вузов / Под редакцией И. И. Пахомова. — М.: Радио и связь, 1982. — 456с.
10. Ханин Я. И. Динамика квантовых генераторов. — М.: Сов. радио, 1975.— 495 с.
11. Пихтин А.Н. Физические основы квантовой электроники и оптоэлектроники: Учеб. пособие для вузов по специальности “Полупроводники и диэлектрики”. — М.: Высш. шк., 1983.—304 с.
12. Звелто О. Принципы лазеров: Пер. с англ. — М.: Энергия, 1969. — 400с.
13. Микаэлян А.Л. и др. Оптические генераторы на твердом теле. — М.:Сов. радио, 1967. — 384 с.
14. Климов Ю. М. Расчет оптико-электронных приборов с лазерами.— М.: Сов. радио, 1979. — 260 с.
15. Ищенко Е.Ф., Климов Ю. М. Оптические квантовые генераторы. — М.: Сов. радио, 1968.— 468 с.
16. Красюк Б.А., Корнеев Г.И. Оптические системы связи и световодные датчики. Вопросы технологии.-М.: Радио и связь, 1985.-192с.
17. Рябов С. Г., Торопкин Г. Н., Усольцев И. Ф. Приборы квантовой электроники. — М.: Сов. радио, 1976.— 308 с.
18. Рябцев Н.Г. Материалы квантовой электроники. — М.: Сов. радио, 1972.— 380 с.
19. Методы расчета оптических квантовых генераторов. т.1,т.2 / Под ред. Б.И.Степанова, т. 2. — Минск: Наука и техника, 1968, — 465 с.
20. Дьяков В.А. Введение в квантовую электронику. — М.: Энергия, 1969.— 261с.
21. Аллен Л., Джонс Д. Основы физики газовых лазеров: Пер. с англ./Под ред. Е. А. Верного. — М.: Наука, 1970.—207с.
22. Справочник по лазерам: Пер. с англ./ Под ред. Прохорова т.1, т.2.— М.: Сов. радио, 1978.— 500 с.,400 с.
23. Конюхов В.К., Матросов И.В., Прохоров А.М. и др. Газодинамический квантовый генератор непрерывного действия на смеси углекислого газа, азота и воды. — Письма в ЖЭТФ, т.12, №10, 1970. — с.461-463.
24. Степанов Б.И, Рубинов А.Н., Мостовников В.А. Оптическая генерация в растворах сложных молекул. — Письма в ЖЭТФ, т.5,1967.— с.144-148.
25. Степанов Б.И., Рубинов А.Н. Оптические квантовые генераторы на растворах органических красителей.— УФН, т.5, №1, 1968. — с.45-74.

26. Карлов Н. В. Лекции по квантовой электронике: Учеб. пособие. — М.: Наука, 1983. — 320с.
27. Чео П. К. Волоконная оптика: Приборы и системы: Пер. с англ. — М.: Энергоатомиздат, 1988. — 280с.
28. Лескович В.И., Пак Г.Т. и др. О температурной зависимости порогового тока полупроводниковых лазеров на арсениде галлия. — ФТП, т.1, №9, 1967. — с.1440-1442.
29. Гауэр Дж. Оптические системы связи: Пер с англ. — М.: Радио и связь, 1989. — 504с.
30. Елисеев П.Т., Страхов В.П.. Полупроводниковый квантовый генератор направленного действия с выходной мощностью в несколько ватт.—ЖТФ, т.40, №7,1970. — с.1564-1565.
31. Елисеев П.Т. Лазеры на гетеропереходах в сб.Квантовая электроника/ Под ред. Н.Г.Басова. — М.: Сов.радио, №6(12),1972. — с.3-30.
32. Ораевский А. Н. Молекулярные генераторы.— М.: Наука, 1964.— 294с.
33. Козанне А.,Флере Ж. и др. Оптическая передача и обработка информации: Пер. с фр. — М.:Мир, 1984. — 504с.
34. Фотоника./ Под ред. М. Балкански. — М.: Мир, 1978. —415с.
35. Мустель Е. Р., Парыгин В. Н. Методы модуляции и сканирования света. — М.: Наука, 1970.—270 с.
36. Ребрин Ю. К. Управление оптическим лучом в пространстве.— М.: Сов. радио, 1977. — 336 с.
37. Ярив А. Квантовая электроника и нелинейная оптика. Пер. с англ./Под ред. О.Г. Вендика, Я.И. Ханина. — М.: Сов.радио, 1973. — 456 с.
38. Kaminow L. Phys. Rev. Letters, 6, 1961. — 528 с.
39. Применение лазеров: Пер. с англ./ Под ред. В. П. Тычинского — М.: Мир, 1974.—445 с.
40. Хирд Г. Измерение лазерных параметров. Пер. с англ. /Под ред. Ф.С.Файзулаева. — М.: Мир, 1970. — 235 с.
41. Кульчин Ю.Н.. Распределенные волоконно – оптические измерительные системы. – М.: Физматлит, 2001 г., 272 с.
42. Бусурин В.И., Носов Ю.Р.. Волоконно - оптические датчики: физические основы, вопросы расчета и применения. - М.: Энергоатомиздат, 1990 г.
43. Шевцов Э.Л., Белкин М.Е. «Фотоприемные устройства», М., «Радио и связь».
44. Акаев А., Майоров А. Оптические методы обработки информации. М.: Высшая школа. - 1988.-132с.
- 45.Свет В. Д. Оптические методы обработки сигналов. - М.; Энергия, 1971.-104с.
46. Кондратенков Г. С. Обработка информации когерентными оптическими системами. - М.: Советское радио, 1972.- 206 с.
47. Гудмен Дж. Введение в Фурье оптику. - М.: Мир, 1970. - 364 с.
48. Сороко Л. М. Основы топографии и когерентной оптики. -М.: Наука, 1971.-616с.
49. Куш Г.Г., Соколова Ж.М., Шангина Л.И. Приборы и устройства оптического и СВЧ диапазонов Учебное пособие. – Томск: Изд-во НТЛ, 2003. - 560 с.

Учебное издание

Г.Г. Кущ, Ж.М. Соколова, Л.И. Шангина

**ПРИБОРЫ И УСТРОЙСТВА  
ОПТИЧЕСКОГО И СВЧ ДИАПАЗОНОВ**

**Учебное пособие**

для студентов, обучающихся по направлению подготовки 654200 (210300) «Радиотехника» и 654400 (210400) «Телекоммуникации» по дисциплинам «Оптоэлектронные и квантовые приборы и устройства», «Устройства СВЧ и антенны».

Формат 60x84 1/16. Усл. печ. л..

Тираж экз. Заказ

Томский государственный университет  
систем управления и радиоэлектроники.

634050, Томск, пр. Ленина, 40.

Тел. (3822) 533018.